

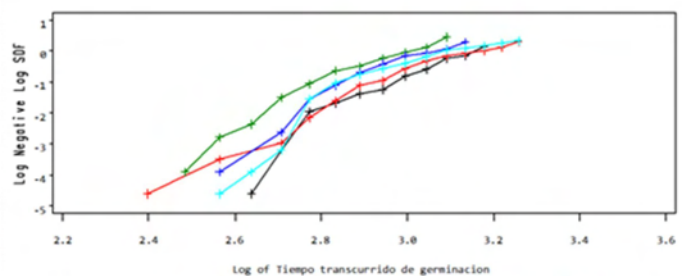
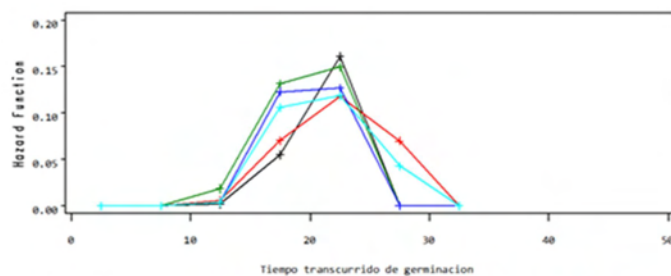
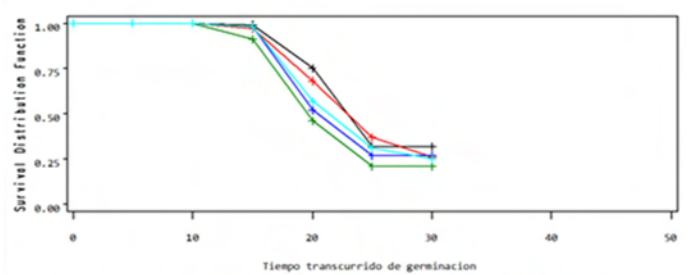
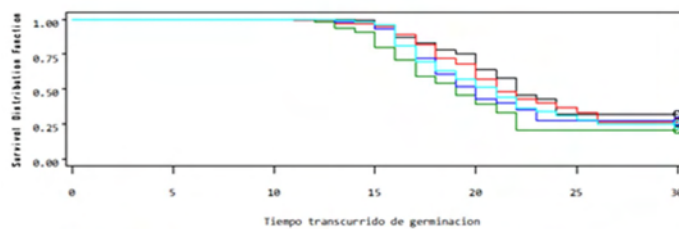
## Manual de procedimientos para la elaboración de tablas de vida en plantas, insectos y experimentos de laboratorio

Aplicaciones en SAS (Statistical Analysis System)



Cesar Coromoto Vale Montilla

Manual de procedimientos para la elaboración de tablas de vida en plantas, insectos y experimentos de laboratorio



UNIVERSIDAD DE LOS ANDES



PUBLICACIONES VICERRECTORADO ACADÉMICO



## UNIVERSIDAD DE LOS ANDES

### Autoridades universitarias

- **Rector**  
Mario Bonucci Rossini
- **Vicerrectora Académica**  
Patricia Rosenzweig Levy
- **Vicerrector Administrativo**  
Manuel Aranguren Rincón
- **Secretario (E)**  
Manuel Joaquín Morocoima

### SELLO EDITORIAL PUBLICACIONES DEL VICERRECTORADO ACADÉMICO

- **Presidenta**  
Patricia Rosenzweig Levy
- **Coordinadora**  
Marysela Coromoto Morillo Moreno
- **Consejo editorial**  
Patricia Rosenzweig Levy  
Marysela Coromoto Morillo Moreno  
Marlene Bauste  
María Teresa Celis  
Jonás Arturo Montilva  
Joan Fernando Chipia L.  
María Luisa Lazzaro  
Alix Madrid  
Francisco Grioso

### COLECCIÓN TEXTOS UNIVERSITARIOS: CIENCIAS NATURALES

Sello Editorial Publicaciones del  
Vicerrectorado Académico.  
de la Universidad de Los Andes.

Los trabajos publicados en esta  
colección han sido rigurosamente  
seleccionados y arbitrados por  
especialistas en las diferentes  
disciplinas.

### Sello Editorial Publicaciones Vicerrectorado Académico de la Universidad de Los Andes

Manual de procedimientos para la  
elaboración de tablas de vida en plantas,  
insectos y experimentos de laboratorio.  
Aplicaciones en SAS (Statistical  
Analysis System).

Primera edición digital, 2024

© Universidad de Los Andes  
Sello Editorial Publicaciones del  
Vicerrectorado Académico  
© Cesar Coromoto Vale Montilla

Hecho el depósito de ley  
Depósito Legal: 2024000242



#### *Corrección de estilo:*

Carlos Gregorio Perdomo Ramírez

#### *Diagramación:*

Cesar Coromoto Vale Montilla  
Marysela C. Morillo Moreno

#### *Diseño de Portada:*

La fotografía de la portada fue tomada por el  
autor y corresponde a un árbol de la especie  
*Hura crepitans L.*, del cual provienen las  
semillas para el ensayo de germinación tratado  
en la obra. Las gráficas que acompañan la  
imagen muestran la distribución de  
probabilidades de germinación de sus semillas.

### SELLO EDITORIAL PUBLICACIONES DEL VICERRECTORADO ACADÉMICO DE LA UNIVERSIDAD DE LOS ANDES

Av. 3 Independencia,  
Edificio Central del Rectorado,  
Mérida, Venezuela.  
[publicacionesva@ula.ve](mailto:publicacionesva@ula.ve)  
[publicacionesva@gmail.com](mailto:publicacionesva@gmail.com)  
<http://www2.ula.ve/publicaciones>  
[academico](http://bdigital2.ula.ve/bdigital/)  
<http://bdigital2.ula.ve/bdigital/>

**Prohibida la reproducción total o  
parcial de esta obra sin la  
autorización escrita de los autores y  
editores.**

Editado en la República Bolivariana de  
Venezuela

### COLECCIÓN DE TEXTOS UNIVERSITARIOS

Esta colección contempla la edición de textos  
académicos que sirven de apoyo docente en las  
áreas del conocimiento existentes en la  
Universidad: Ciencias Humanísticas y Sociales, las  
Ciencias Naturales, la Ingeniería y la Tecnología, la  
Medicina y las ciencias de la salud y las ciencias  
agrícolas.

Entre los objetivos específicos de esta colección  
resaltan:

- Estimular la edición de libros al servicio  
de la docencia.
- Editar la obra científica de los profesores  
de nuestra Casa de Estudios.
- Publicar las investigaciones generadas en  
los centros e institutos de investigación.

Hasta ahora, un número considerable de textos  
universitarios ha sido publicado por miembros de  
nuestra planta profesoral, obras de las que se han  
beneficiado por igual estudiantes y docentes, en la  
búsqueda del mejoramiento de la calidad de  
nuestra educación de pre y posgrado.



UNIVERSIDAD  
DE LOS ANDES



PUBLICACIONES  
VICERRECTORADO ACADÉMICO

# **Manual de procedimientos para la elaboración de tablas de vida en plantas, insectos y experimentos de laboratorio.**

**Aplicaciones en SAS  
(Statistical Analysis System)**



**Universidad de Los Andes  
Mérida 2024 - Venezuela**





# **Manual de procedimientos para la elaboración de tablas de vida en plantas, insectos y experimentos de laboratorio.**

**Aplicaciones en SAS  
(Statistical Analysis System)**

**Cesar Coromoto Vale Montilla**

Centro de Investigaciones Agrícolas, Biológicas, Educativas y Sociales (CIABES)  
Universidad de Los Andes  
Núcleo Universitario Rafael Rangel -NURR-  
Trujillo, Venezuela.

**COLECCIÓN TEXTOS UNIVERSITARIOS CIENCIAS  
NATURALES**

Sello Editorial Publicaciones del Vicerrectorado Académico  
Universidad de Los Andes



# AGRADECIMIENTOS

La consolidación de este proyecto no hubiera sido posible sin la valiosa contribución de la Universidad de Los Andes, específicamente de algunas de sus instituciones. En lo particular, agradezco al Núcleo Universitario “Rafael Rangel”, especialmente al Departamento de Biología y Química, en el cual he ejercido la docencia e investigación desde mis inicios como Profesor y al Centro de Investigaciones Agrícolas, Biológicas, Educativas y Sociales (CIABES) del NURR – ULA, por el uso de sus equipos de laboratorio. Especial agradecimiento a los revisores anónimos, por su interés en mejorar la calidad del manuscrito.

Finalmente, mi agradecimiento al Vicerrectorado Académico, con especial referencia al Sello Editorial Publicaciones del Vicerrectorado Académico de la Universidad de Los Andes, Coordinadora y Consejo Editorial por su apoyo, lo cual hizo posible llevar a feliz término la publicación de esta obra.





# ÍNDICE

<b>INTRODUCCIÓN</b> .....	13
<b>CAPÍTULO 1 Consideraciones Generales sobre el Análisis de Sobrevivencia</b> .....	15
Introducción.....	15
1.1. Definiciones básicas de Análisis de Supervivencia.....	17
1.1.1. Análisis de Supervivencia.....	17
1.1.2. Variables tiempo – evento.....	18
1.1.3. Censura.....	19
1.1.4. Naturaleza de los tiempos de falla en estudios ecologicos.....	22
1.1.5. Función de supervivencia.....	26
1.1.6. Función de riesgo.....	28
1.1.7. Razón de riesgo.....	28
<b>CAPÍTULO 2 Características y Procedimientos de Análisis</b> .....	31
2.1. Estructura de datos.....	31
2.2. Métodos de análisis.....	32
2.2.1. Método 1. Análisis de tiempo hasta el evento (Tablas de vida, Kaplan-Meier) usando PROC LIFETEST.....	32
2.2.1.1. Especificación de los procedimientos generales para PROC LIFETEST.....	32
2.2.1.2. Especificación de las opciones para PROC LIFETEST.....	40
2.2.2. Método 2. Análisis de supervivencia con modelos de regresión Cox usando PROC PHREG.....	42
2.2.2.1. Sintaxis de los procedimientos.....	44
2.2.2.2. Uso básico de PROC PHREG.....	45
2.2.2.3. Especificación de las opciones para PROC PHREG.....	45
2.2.2.4. Posibles causas de violación de las suposiciones de riesgo proporcional.....	57
2.2.2.5. Pruebas para los supuestos de riesgo no proporcional.....	

2.2.2.6. Procedimiento con Proc Format para transformar una variable continua en variable categorica. ....	59
2.2.3. Método 3. Análisis de supervivencia con modelos de regresión pa- ramétrica utilizando PROC LIFEREG.....	68
2.2.3.1. Sintaxis de los procedimientos.....	69
2.2.3.2. Especificación de las opciones para PROC LIFEREG.....	70
2.2.3.3. Uso básico de PROC LIFEREG.....	71

<b>CAPÍTULO 3 Metodologías y Procedimientos Aplicados al Crecimiento de Plantas, Entomología y Control Biológico.....</b>	<b>85</b>
3.1. Aplicación básica del Análisis de Supervivencia en estudios eco- lógicos.....	85
3.1.1. Principios de análisis de supervivencia en estudios ecologicos.....	86
3.1.2. Modelos de datos de supervivencia en estudios ecologicos.....	88
3.1.3. Ejemplos introductorios para la interpretación de datos ecológicos.	90
3.1.3.1. Tiempos de emergencia y floración en especies invasoras (Fox 2001) .....	90
3.1.3.2. Factores físicos y biológicos involucrados en la supervivencia y riesgos de mortalidad de plántulas de de <i>A. religiosa</i> bajo regeneración na- tural (Hernández, López, Cruz y Luna, 2022) .....	103
3.1.3.3. Supervivencia de plántulas de <i>A. religiosa</i> bajo condición de do- sel.....	108
3.2. Algunos ejemplos metodológicos típicos extraídos de la literatura científica aplicados al crecimiento de plantas, entomología y control bio- lógico.....	125
3.2.1. Germinación de semilla.....	125
3.2.1.1. Métodos estadísticos de análisis de supervivencia aplicados en germinación de semilla.....	126
a. Ejemplo 1: Germination Data Analysis by Time-to-Event Ap- proaches (Romano & Stevanato, 2020) .....	128
b. Ejemplo 2: Germinación de tres especies de <i>Cephalocereus</i> (Cac- taceae) endémicas del Istmo de Tehuantepec, México (Barcenas et al., 2013) .....	130
c. Ejemplo 3: Germinación y emergencia diaria de cariósides y diásporas de pastos nativos e introducidos (Quero et al., 2017) .....	131
d. Ejemplo 4: Factores que influyen en la viabilidad y germinación de <i>Carpinus caroliniana</i> Walt. (Rodriguez, 2019) .....	134



3.2.2. Análisis de regeneración natural y crecimiento de plantas.....	139
a. Ejemplo 1: Factores que influyen en la emergencia y supervivencia de plántulas de <i>Carpinus caroliniana</i> in situ (Rodríguez, 2019) .....	139
b. Ejemplo 2: Supervivencia en plantaciones de <i>Pinus pseudostrobus</i> Lindl. en función del sistema de producción y preacondicionamiento en vivero (Sigala, Gonzalez & Prieto, 2015) .....	141
3.2.3. Entomología y control biológico.....	142
a. Ejemplo 1: Impacto de diferentes temperaturas nas características biológicas de <i>Helicoverpa armigera</i> (Hübner, 1808) (Lepidoptera: Noctuidae) durante três gerações (Alvez, 2017) .....	142
b. Ejemplo 2: Survival and behavior of the insecticide-exposed predators <i>Podisus nigrispinus</i> and <i>Supputius cincticeps</i> (Heteroptera: Pentatomidae). (Castro et al., 2013) .....	146
c. Ejemplo 3: Lethal and sub-lethal effects of <i>Beauveria bassiana</i> (Cordycipitaceae) strain NI8 on <i>Chrysoperla rufilabris</i> (Neuroptera: Chrysopidae) (Portilla, Snodgrass, & Luttrell, 2017) .....	149
d. Ejemplo 4: Bioactivity of essential oils from Artemisia against <i>Diaphania hyalinata</i> and its selectivity to beneficial insects (Lopez et al., 2018) .....	152
e. Ejemlo 5: Survival and feeding avoidance of the eucalyptus defoliator <i>Thyrinteina arnobia</i> exposed to the proteinase inhibitor berenil (Marinho et al., 2012) .....	155
3.2.4. Impacto de enfermedades fungosas y bacterianas en el crecimiento y sobrevivencia de las plantas.....	157
a. Ejemplo 1: Applications of survival analysis in botanical epidemiology (Scherer & Ojiambo, 2004) .....	158
b. Ejemplo 2: Survival analysis to determine the length of latent period of <i>Mycosphaerella pinodes</i> on peas ( <i>Pisum sativum</i> L.) (Setti, Bencheikh, Henni & Claire, 2010) .....	163
3.3. Aplicación práctica de las metodologías utilizadas para el desarrollo de tablas de vida.....	167
3.3.1. Germinación de semilla.....	167
3.3.1.1. Análisis de tiempo y riesgo de germinación en semillas de <i>Hura crepitans</i> L imbibidas en agua (Vale 2023) .....	168
3.3.1.2. Análisis de supervivencia en germinación de <i>Cassia moschata</i> Kunth. con aplicación de hormonas vegetales (Vale 2023) .....	227
3.3.1.3. Análisis de supervivencia en germinación de <i>Swietenia macrophylla</i> King con aplicación de hormonas vegetales (Vale-Montilla, 2022) ...	260
3.3.2. Entomología y otros experimentos de laboratorio.....	291
3.3.2.1. Efecto de dos dietas alternantes vegetales en el desarrollo larval de dos especies de mariposas diurnas; <i>Morpho menelaus</i> y <i>Morpho he-</i>	

<i>lenor</i> bajo condiciones de cautiverio (Peña, 2017) .....	291
3.3.2.2. Experimentos de laboratorio.....	347
3.3.2.2.1. Relacion entre dieta alimenticia y desarrollo de tumores en ratas de laboratorio (King et al., 1979) .....	349
<b>EPÍLOGO</b> .....	395
<b>BIBLIOGRAFÍA</b> .....	403
<b>EL AUTOR</b> .....	415

# INTRODUCCIÓN

Este manual, presenta algunos de los procedimientos utilizados para el cálculo y desarrollo de análisis de sobrevivencia en algunas especialidades biológicas. Así, el usuario podrá entender las relaciones funcionales entre los componentes ambientales y biológicos que, con herramientas cuantitativas y cualitativas, permiten formalizar modelos e hipótesis en el crecimiento de los organismos. Con esta base, se podrán formular nuevas interrogantes e hipótesis que permitan la incorporación de una serie de consideraciones tanto conceptuales como matemáticas para generar otras investigaciones relacionadas con el crecimiento de los organismos.

En vista de la utilidad de los análisis de sobrevivencia para casos particulares de la investigación biológica, el objetivo del presente manual es contribuir con su uso a través de paquetes estadísticos para analizar y agilizar estructuras de datos que involucran diferentes variables del crecimiento o de otro tipo, evaluadas en tiempo, sitio o tratamientos. Con el manual se pretende ilustrar al usuario una metodología que le permita el procesamiento de los datos para verificar cuales métodos se adaptan mejor, determinar qué criterios tienen más y mejor relación con el método aplicado. Es decir, presenta una serie de procedimientos para el análisis de datos experimentales utilizando el paquete estadístico SAS® (SAS, 2020).

El propósito general es ofrecer a profesionales y estudiantes dedicados a la investigación biológica, una herramienta para la aplicación de análisis de sobrevivencia. Su objetivo consiste en manejar los detalles de exploración y descripción de datos, y servir de guía práctica para ilustrar las posibilidades de las técnicas de análisis de datos en diferentes actividades relacionadas con las plantas, insectos u otras investigaciones biológicas. Discute el uso de procedimientos con PROC LIFETEST, PROC PHREG y PROC LIFEREG, como herramientas para el ajuste de curvas de sobrevivencia.

En el texto del manual, se incluye la aplicación práctica de las metodologías utilizadas para el desarrollo de tablas de vida. Se detallan algunos procedimientos para el análisis de germinación de semillas de las especies *Hura crepitans* L, *Cassia moschata* Kunth y *Swietenia macrophylla* King bajo tratamientos pregerminativos a base de tiempos de imbibición en agua y en hormonas vegetales. Otro de



los estudios analizados, se basa en el uso de dietas alimenticias y desarrollo de fases e instares larvales en insectos. Finalmente, un estudio para investigar la relacion entre dieta alimenticia y desarrollo de tumores en ratas de laboratorio.

El texto del Manual se ha organizado de la siguiente manera:

**Capítulo 1:** Consideraciones generales sobre el Análisis de Supervivencia.

**Capítulo 2:** Características y procedimientos del análisis.

**Capítulo 3:** Metodologías y Procedimientos Aplicados al Crecimiento de Plantas, la Entomología y el Control Biológico. Hace una reseña bibliográfica con algunos ejemplos sobre el uso de metodologías y procedimientos aplicados en germinación de semillas y crecimiento de plantas, Entomología, control biológico y experimentos de laboratorio. En el Capítulo se incluye la aplicación práctica de las metodologías utilizadas para el desarrollo de tablas de vida en estudios sobre germinación de semillas, efecto de dietas vegetales en el desarrollo larval de insectos y relación entre dieta alimenticia y desarrollo de tumores en ratas de laboratorio.

Para la aplicación del Manual no se requiere de conocimientos avanzados en estadística o cálculo matemático. Aunque para las discusiones y ejemplos se utiliza el paquete estadístico SAS®, se requiere algún conocimiento de su programación, de interés general en el ajuste de las curvas.

*Cesar C. Vale Montilla*

# CAPÍTULO 1

## Consideraciones Generales sobre el Análisis de Supervivencia

### Introducción

El Análisis de Supervivencia utiliza métodos estadísticos para el estudio de la ocurrencia y tiempos en los cuales ocurren los eventos. Son métodos que fueron diseñados para estudios de mortalidad. Sin embargo, actualmente se utilizan para el estudio de diferente clase de eventos en Ciencias Sociales y Naturales, aunque no implican diferencias en las técnicas (Allison, 2010).

El Análisis de Supervivencia fue diseñado para datos longitudinales sobre la ocurrencia de eventos, definiendo un evento como un cambio cualitativo ubicado en el tiempo, es decir, una transición de un estado discreto a otro. Por ejemplo, cambios en altura y diámetro de planta, germinación de semillas, cambios del estado de huevo a larva, etc. Para ello, se requiere conocer cuando ocurrieron los cambios y ubicar estos eventos en el tiempo. Las transiciones entre eventos pueden ocurrir rápido y conocer el tiempo exacto en el que ocurren, mientras que otros también pueden ocurrir en tiempo corto, pero se desconoce el tiempo exacto o el tiempo resulta confuso. Si el evento de interés es una plantación de especies perennes, se puede conocer el año de plantación, mientras que en ella ocurren eventos relativamente cortos que abarcan la duración total de las observaciones.

La mejor manera de abordar un análisis de supervivencia es de manera prospectiva. Monitorear un grupo de individuos durante un lapso de tiempo definido, registrando los eventos de interés. En este sentido, existen diferentes tipos de eventos que pueden afectar las tasas de supervivencia. Si el evento es mortalidad, se puede registrar la causa de mortalidad. Algunos eventos pueden ocurrir dos o más veces sobre el mismo individuo (Allison, 2010).

Se puede realizar un análisis de supervivencia cuando los datos consisten solo de tiempo de los eventos, pero una finalidad común es estimar modelos causales o predictivos en los cuales el riesgo de un evento es dependiente de covariables, para los cuales se debe disponer de mediciones de las covariables. Algunas de ellas, por ejemplo, sexo y raza, pueden ser constantes en el tiempo. Otras son variables en el tiempo y los datos deben incluir en detalle su variación temporal, por ejemplo, variables ambientales.

Los datos de supervivencia tienen dos rasgos comunes que son difíciles de manejar con métodos estadísticos convencionales, ellos son: censura y covariables dependientes del tiempo. Para el análisis de ambos problemas se dispone del siguiente ejemplo: una muestra de 250 plantas de una misma especie se les hizo seguimiento durante 5 años para observar floración. El evento de interés fue la primera floración. El propósito fue determinar cómo la ocurrencia y tiempo transcurrido para la floración dependió de varias covariables (variables predictoras). Algunas de estas covariables (procedencia, edad a la floración y altura en msnm) permanecieron constantes durante el intervalo de medición. Otras (altura y diámetro de la planta) debieron cambiar en algún momento durante la continuación del periodo. ¿Cómo analizar esta información utilizando métodos convencionales? Una posibilidad es con regresión logística, con variable dependiente dicotómica: floración y no floración. Pero este análisis ignora información relacionada con el tiempo transcurrido para la floración. Otra solución es hacer la variable dependiente como el lapso del tiempo transcurrido entre el inicio del experimento y la primera floración y luego, utilizar un modelo convencional de regresión lineal. Pero, ¿qué hacer con las plantas que no florecieron durante el mismo periodo? Tales casos son referidos como censurados.

Todos los métodos de análisis de supervivencia permiten censurar y algunos también permiten las covariables dependientes del tiempo. Para censurar, el truco es proyectar un procedimiento que combine la información en los casos censurado y no censurado en una vía que produzca estimados consistentes de los parámetros de interés. Se puede utilizar el método de máxima verosimilitud o verosimilitud parcial. Las covariables dependientes del tiempo también pueden utilizar estos métodos de verosimilitud (Allison, 2010).

El SAS® dispone de seis procedimientos para los análisis de supervivencia, cuyas características generales, son las siguientes:



**LIFETEST:** Principalmente diseñado para análisis univariado de tiempo de los eventos. Produce tablas de vida y gráficas de curvas de supervivencia, también llamadas funciones de sobreviviente. El procedimiento prueba si las curvas de supervivencia son semejantes en dos o más grupos. También prueba asociaciones entre tiempos de eventos y covariables constantes en el tiempo, pero no produce estimados de los parámetros.

**LIFEREG:** Estima modelos de regresión con censura, datos continuos en el tiempo, bajo varias alternativas distribucionales. Admite variedades de censurar, pero no para covariables dependientes del tiempo.

**PHREG:** Usa métodos de Cox con verosimilitud parcial para modelos de regresión con datos censurados. Es menos restrictivo que LIFEREG, además, los métodos de estimación admiten covariables tiempo – dependientes. Manipula datos continuos y discretos en el tiempo.

**LOGISTIC:** Designado para problemas generales en datos categóricos, pero es efectivo y flexible en modelos de supervivencia para datos discretos en el tiempo con covariables tiempo – dependientes.

**GENMOD:** Estima los mismos modelos de supervivencia discretos que PROC LOGISTIC, pero puede admitir eventos repetidos utilizando métodos de ecuaciones generalizadas (GEE).

**NLMIXED:** Estima modelos paramétricos de efectos mixtos al azar para eventos repetidos.

## **1.1. Definiciones básicas del Análisis de Supervivencia**

### **1.1.1. Análisis de Supervivencia**

El análisis de supervivencia es un conjunto de métodos para analizar datos en los que la variable dependiente está representada por el tiempo hasta que ocurre un evento de interés. El evento puede ser cualquier investigación de interés, por ejemplo, germinación de semilla, muerte por efecto de tratamientos, ocurrencia de una enfermedad, efecto de tratamientos con insecticidas, etc. El tiempo para que ocurra el evento o el tiempo de supervivencia se puede medir en días, semanas, meses, años, etc. Por ejemplo, si el caso de interés es una enfermedad, entonces el tiempo de supervivencia puede ser el tiempo en horas, días, semanas, etc., hasta que el organismo desarrolle la enfermedad.

Los datos de sobrevivencia consisten en una variable de respuesta (tiempo hasta el evento, tiempo de falla o tiempo de supervivencia) que mide la duración del tiempo hasta que ocurre un evento específico y posiblemente un conjunto de variables independientes que se cree que están asociadas con la variable de tiempo de ocurrencia del evento. Estas variables independientes (variables concomitantes, covariables o factores de pronóstico) pueden ser cualitativas o categóricas como el sexo o la especie, o cuantitativas, como la edad o la temperatura. El propósito del análisis de sobrevivencia es modelar la distribución subyacente de la variable de tiempo de falla y evaluar su dependencia de las variables independientes.

En resumen, el análisis de supervivencia se centra en dos datos importantes:

- Si un organismo sufre o no el evento de interés durante el período de estudio (es decir, una variable dicotómica o indicadora, a menudo codificada como 1 = ocurrió el evento o 0 = el evento no ocurrió durante el período de observación del estudio).
- El tiempo de seguimiento para cada individuo que se está evaluando

### 1.1.2. Variables tiempo – evento

El tiempo cero, es el momento en el que los participantes se consideran en riesgo para el resultado de interés. El tiempo de seguimiento se mide desde el tiempo cero (el inicio del estudio o desde el punto en el que se considera que el participante está en riesgo) hasta que ocurre el evento, finaliza el estudio o se pierde el participante, lo que ocurra primero. Los individuos al analizar, a menudo ingresan o son seleccionados en estudios de cohortes y ensayos de laboratorio durante un período de varios días, meses o años. Por lo tanto, es importante registrar el tiempo de entrada para que el tiempo de seguimiento se mida con precisión. El interés radica en el tiempo hasta el evento, pero por varias razones no siempre se puede medir el tiempo hasta el evento (por ejemplo, el participante abandona el estudio por pérdida o el período de observación del estudio finaliza).

Existen características únicas de las variables de tiempo – evento. Primero, los tiempos para el evento son siempre positivos y sus distribuciones a menudo están sesgadas. Por ejemplo, en un estudio que evalúa el tiempo hasta la recaída en individuos de alto riesgo, la mayoría de los eventos (recaídas) pue-

den ocurrir al principio del seguimiento y muy pocos ocurren posteriormente. Por otro lado, en un estudio del tiempo hasta la muerte en una muestra comunitaria, la mayoría de los eventos (muertes) pueden ocurrir más adelante en el seguimiento. No se aplican los procedimientos estadísticos estándar que asumen la normalidad de las distribuciones. Se pueden invocar procedimientos no paramétricos, excepto por el hecho de que existen problemas adicionales. Específicamente, en un estudio, los datos completos (tiempo real para los datos del evento) no siempre están disponibles para cada individuo. En muchos estudios, los individuos se evalúan durante un período de tiempo (días o meses) y el estudio finaliza en una fecha calendario específica. Por lo tanto, los individuos cuya evaluación se inicia más tarde (luego de iniciada la investigación), son seguidos durante un período más corto que los individuos que se les hace seguimiento desde el inicio de la evaluación. Para algunos individuos, por alguna circunstancia, abandonan el estudio antes del final del período de seguimiento y otros pueden morir durante el período de seguimiento (asumiendo que el resultado de interés no es la muerte).

En cada uno de estos casos, se dispone de información de seguimiento incompleta. El verdadero tiempo de supervivencia (a veces llamado tiempo de falla) no se conoce porque el estudio finaliza o porque para un participante que abandona el estudio antes de experimentar el evento. Lo que se conoce es que el tiempo de supervivencia de los individuos es mayor que su último tiempo de seguimiento observado. Estos tiempos se llaman tiempos censurados.

### 1.1.3. Censura

Según Emmert y Dehmer (2019), para realizar un análisis de supervivencia, es necesario registrar el tiempo hasta el evento  $t_i$  para los sujetos  $i \in \{1, \dots, N\}$  de un grupo. Sin embargo, esto no siempre es posible y solo tenemos información parcial del tiempo del evento. En tal caso, se habla de censura. Específicamente, un sujeto tiene una censura del tiempo de supervivencia si el evento aún no ha ocurrido o no ocurrió para este organismo.

La censura es de dos tipos, derecha e izquierda. Generalmente se encuentran datos censurados por la derecha. Los datos censurados por la izquierda pueden ocurrir cuando el tiempo de supervivencia de un individuo es incompleto en el lado izquierdo del período de seguimiento del individuo. Por ejemplo, se puede hacer un seguimiento de un individuo para detectar cualquier trastorno infeccioso o síntoma,

desde el momento en que se detecta el trastorno. Es posible que no se conozca el momento exacto de exposición al agente infeccioso (Prinja, Gupta y Varma, 2010).

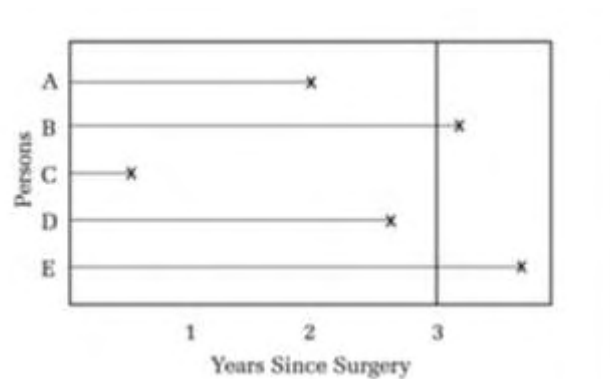
Por esta razón, se habla de censurar a la izquierda si se observa el evento, pero no el inicio del proceso, ya que, por lo general, la infección se diagnostica en algún momento, pero la infección en sí ocurrió con anterioridad, en un punto de partida desconocido. También existen los sujetos censurados por la derecha. La principal característica de los datos de sobrevivencia es la presencia de censura que ocurre hasta el final del estudio; en todas las unidades es importante considerar esos tiempos para el análisis. Según Bolfarine *et al.* (1991) se pueden considerar tres tipos de censura:

(a) Censura Tipo 1: ocurre cuando el estudio se cierra después de un periodo de tiempo preestablecido (esto es, bajo control del investigador). Por ejemplo, se fija un tiempo para el estudio, al final del cual, se cierra el estudio, independientemente de las unidades que fallaron, de manera que se pierde información sobre determinadas unidades. Todos los sujetos comienzan y terminan el estudio al mismo tiempo (duración fija del estudio). Los ejemplos son experimentos de laboratorio: seguimiento de germinación durante 30 días.

(b) Censura Tipo 2: ocurre cuando el estudio se cierra luego de haber ocurrido un evento de interés en un número preestablecido de individuos. Por lo tanto, se cierra el estudio con las unidades que fallaron (presentaron el evento) y nada se puede afirmar respecto de las otras unidades. Todos los sujetos comienzan el estudio al mismo tiempo, pero el estudio termina cuando un número fijo predeterminado de sujetos han experimentado el evento (duración flexible del estudio). Los ejemplos son experimentos de laboratorio. Por ejemplo, un investigador siguiendo un experimento con 100 ratas de laboratorio puede decidir que el experimento finalizará cuando haya muerto el 50 % de ellas; en germinación, cuando haya germinado el 50% de las semillas (tiempo medio de germinación,  $T_{50}$ ).

(c) Censura Aleatoria: se presenta cuando la unidad es retirada sin haber ocurrido la falla o evento. Por ejemplo, cuando la unidad falla por una razón diferente a la estudiada. En ensayos clínicos, un factor que causa censura aleatoria es la pérdida de un segmento (loss-to-follow-up), o sea, por alguna razón el individuo sale del estudio. Los sujetos ingresan al estudio en diferentes momentos, pero la duración del estudio es fija. Algunos ejemplos son los ensayos clínicos.

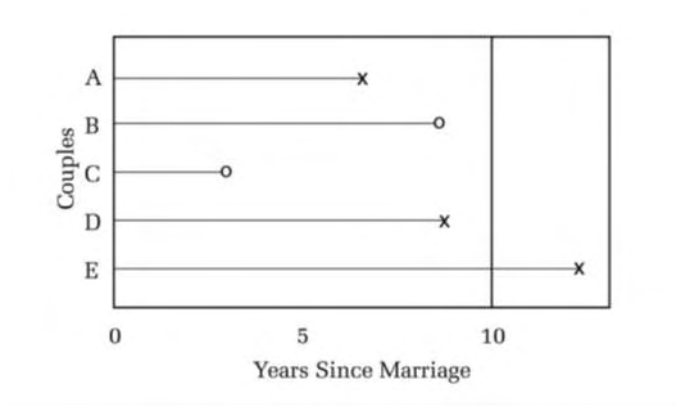
Para Allison (2010), la situación de censura más simple y común se refleja en la Figura 01. Esta figura describe datos de un estudio en el cual todas las personas (individuos) fueron operadas del corazón en el tiempo 0 y se les hace seguimiento durante 3 años. El eje horizontal representa el tiempo y cada línea horizontal (A – E) representa un individuo. Una X indica una muerte ocurrida en este punto en el tiempo. La línea vertical en 3 es el punto en el cual se detiene el seguimiento a los individuos. Se observan algunas muertes ocurridas en el tiempo 3 o antes, por tanto, estos tiempos de muerte no son censurados. Algunas muertes ocurridas después del tiempo 3 no son observadas y estos tiempos de muerte se consideran censurados en el tiempo 3. Por ejemplo, los individuos A, C y D tienen tiempos de muerte no censurados, mientras que B y E tienen tiempos de muerte censurados a la derecha. Las observaciones que son censuradas en esta vía son referidas como censuradas Tipo I.



**Figura 01. Datos censurados a la derecha.** Fuente. Allison (2010).

La censura al azar ocurre cuando las observaciones culminan por razones que no están bajo el control del investigador. Existen algunas razones posibles para que esto suceda. Por ejemplo, hacer el seguimiento de una muestra de matrimonios durante 10 años, comenzando desde el casamiento, registrando el tiempo hasta los divorcios. Claramente, los matrimonios que permanecen casados después de 10 años son censurados por un mecanismo Tipo I. Pero en algunas parejas puede morir el marido o la esposa antes de los 10 años. Algunas parejas pueden mudarse y hacer imposible su ubicación. Otras parejas pueden negarse a continuar en el estudio luego de 5 años. Estas clases de censuras son descritas en la Figura 02, donde la “O” para las parejas B y C indican que las observaciones están censuradas en este punto en el tiempo.





**Figura 02. Datos censurados al azar.** Fuente. Allison (2010).

La censura al azar puede también producirse cuando existe un tiempo de terminación, pero los tiempos de entrada en el estudio varían aleatoriamente entre individuos. Al considerar el ejemplo del seguimiento de individuos hasta la muerte, luego de ser operadas, el escenario más probable es uno en el cual los individuos son operados en tiempos diferentes, pero el estudio termina sobre un dato individual. Todos los individuos que permanecen vivos sobre este dato se consideran censurados, pero sus tiempos de sobrevivencia de la operación son variables. Esta censura se considera al azar ya que los tiempos de entrada no están bajo control del investigador.

#### 1.1.4. Naturaleza de los tiempos de falla en estudios ecológicos

Los ecólogos obtienen datos sobre el tiempo de los eventos al observar repetidamente a individuos identificados de forma individual. Al igual que con el análisis de varianza de medidas repetidas, esta estructura de observaciones repetidas conduce a métodos estadísticos especiales. Estas observaciones pueden tener lugar de forma continua (por ejemplo, al estudiar los tiempos de abandono de abejas recolectoras) o pueden estar en la forma de censos a intervalos (para estudiar la supervivencia de las plantas marcadas o muestras). En cada tiempo de observación, el investigador determina si el evento de interés ha ocurrido para cada individuo. En estudios simples, son posibles tres resultados generales:

1. El evento aún no ha ocurrido y el individuo todavía se encuentra en su estado original. En este caso se dice que el individuo ha sobrevivido; "supervivencia" en este sentido puede referirse a permanecer vivo, así como a permanecer en un sitio, semilla sin germinar, etc.
2. El evento ha ocurrido de manera observable. En este caso, los estadísticos hablan de falla y del tiempo del evento de cada individuo. En un entorno ecológico, una falla se refiere a que se ha observado que el individuo abandona la flor que poliniza, estar visiblemente muerto, la planta ha comenzado a florecer o la semilla germina de manera observable.
3. Por último, es posible que el individuo se haya perdido del estudio sin que se haya observado una falla o evento de interés. Este tipo de punto de datos se denomina punto de datos censurado por la derecha, porque el investigador solo puede asegurar de que el tiempo de falla real es mayor que el último tiempo de supervivencia registrado.

Por ejemplo, un tiempo de falla (en este caso, un "tiempo de abandono") no se puede asignar a una abeja recolectora perteneciente a un estudio, si se observa que fue comida por un depredador, pero el investigador sabe que el momento de abandonar la búsqueda de alimento (es decir, el tiempo de falla) habría sido al menos tan grande como el momento de la depredación (es decir, el tiempo de censura). Del mismo modo, no se puede asignar un tiempo de floración o tiempo de muerte a una planta que ha perdido su marca, porque el investigador no está seguro de su destino en cualquier tiempo posterior al anterior o último censo, aunque está claro que el tiempo de falla fue posterior al censo anterior.

Los investigadores a menudo descartan los puntos de datos censurados por la derecha, ya sea porque asumen que estos datos no son útiles, o porque creen que los métodos para manejar datos censurados son complejos. Ambas creencias están equivocadas. La primera creencia lleva a los investigadores a descartar datos potencialmente importantes y pueden contribuir a resultados muy sesgados. También es posible que los puntos de datos estén censurados por la izquierda. Este es el caso cuando las fallas ya han ocurrido de forma observable para algunos individuos cuando comienza el estudio (Kalbfleisch y Prentice, 1980; Lawless, 2003).

Un ejemplo sería un estudio de los tiempos de eclosión de pupas de una población de lepidóptera, en la que hay crisálidas abiertas presentes al inicio del estudio. Censurar a la izquierda puede evitarse con frecuencia mediante una planificación cuidadosa (por ejemplo, comenzando el estudio con anterioridad a la eclosión). En muchos experimentos controlados, la censura por la izquierda no es posible, porque las fallas o eventos no pueden ocurrir antes de comenzar el experimento (por ejemplo, las semillas no pueden germinar antes de que el investigador siembre las semillas). La censura por la izquierda es a veces inevitable en los estudios de campo.

Finalmente, la censura de intervalo ocurre cuando las mediciones se toman a intervalos lo suficientemente grandes como para que solo se conozca que un evento ocurrió dentro de un intervalo particular. Obviamente, dado que siempre se registran datos a intervalos discretos, se podría considerar que la mayoría de los datos censurados por la derecha están censurados por intervalos.

Existen cuestiones tanto biológicas como estadísticas a considerar aquí. Biológicamente, si se registran datos diariamente, se conoce que el evento ocurrió entre ayer y hoy. ¿Son los datos por hora significativos e importantes? No es probable para la mortalidad de las plantas, pero quizás sí lo es en los estudios de comportamiento animal. Se deben diseñar estudios de modo que el muestreo se realice en una escala de tiempo relevante para las cuestiones de interés. Estadísticamente, como las mediciones ocurren a intervalos más cortos, los resultados de los dos tipos de análisis convergen, (llamando a los datos censurados por intervalo y censurados por la derecha) (Collett, 2004). El mejor consejo podría ser: si el intervalo de muestra es largo (en un sentido biológico), se consideran los datos como censurados por intervalo, de lo contrario se consideran como censurados a la derecha. En caso de duda, se analizan los datos en ambos sentidos.

Los diseños experimentales a menudo incluyen censura: los investigadores a menudo planifican la finalización de los estudios antes que todos los individuos hayan fallado o presentado el evento. Obviamente, esto es necesario cuando los individuos tienen vidas muy largas. Dos tipos de diseños se pueden distinguir. En la censura de Tipo I, se puede planear terminar un estudio en un tiempo determinado, en cuyo caso el número de fallas es una variable aleatoria. Con censura Tipo II, el estudio se completa después de un número particular de fallas, en cuyo caso el tiempo de finalización del estudio es una variable aleatoria (Lawless, 2003). En los estudios ecológicos, la censura de Tipo I es más co-

mún porque los estudios a menudo terminan cuando finalizan las temporadas de campo. En la mayoría de estudios ecológicos, parte de la censura no es planificada y aleatoria, por ejemplo, algunos individuos experimentales mueren o de lo contrario, se pierden en el estudio en tiempos aleatorios durante el curso del estudio. En consecuencia, la discusión asume que la censura es de Tipo I y / o aleatoria. Este tipo de estructura significa que hay tres elementos importantes que requieren consideración en estudios de tiempo de eventos. Primero, las mediciones se repiten a lo largo del tiempo. En segundo lugar, la mayoría de los estudios incluyen puntos de datos censurados. Finalmente, incluso si no hay datos censurados, los tiempos de falla no están distribuidos normalmente bajo cualquier transformación estándar.

La no normalidad de los datos de tiempo de falla puede ser en parte una consecuencia de la estructura experimental. Las distribuciones normales son simétricas, con colas infinitamente largas en ambos lados. Pero al comenzar un experimento controlado en un tiempo particular, el investigador establece una línea clara de demarcación: no puede ocurrir ninguna falla antes del comienzo de un estudio controlado. En muchos ajustes ecológicos, esta línea de corte es en realidad más avanzada en el tiempo. Por ejemplo, en estudios de tiempo de floración que comienza cuando el investigador siembra las semillas, generalmente hay un tiempo mínimo que debe transcurrir antes de que cualquier planta comience a florecer. En este sentido, los datos de tiempo de falla a menudo están íntimamente relacionados al tiempo particular en el que se inicia el experimento, a diferencia de la mayoría de los otros tipos de datos (Fox, 1990a).

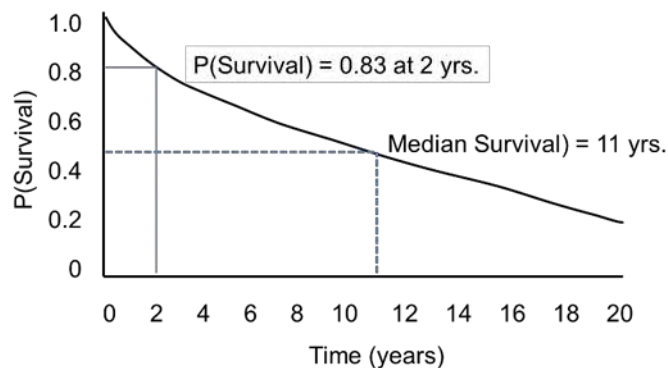
También hay razones ecológicas y biológicas por las que los datos de tiempo de falla pueden no estar distribuidos normalmente. Por ejemplo, en los estudios de supervivencia de plantas marcadas, no hay razón para esperar una distribución normal del tiempo hasta la muerte. En la mayoría de las poblaciones, es probable que las muertes se concentren por el tiempo de los eventos ambientales importantes, como heladas, sequías o migraciones de herbívoros. Fallas que ocurren cuando se alcanza algún umbral biológico, por ejemplo, el inicio de la reproducción o las muertes por senescencia, también es probable que conduzcan a conjuntos de datos que no se distribuyen normalmente.

A menudo es biológicamente importante que un conjunto de datos sea, por ejemplo, Gamma en lugar de distribución Weibull. Por ejemplo, la diferencia entre los tipos clásicos I, II y III de funciones de supervivencia, es precisamente que un Tipo II implica un riesgo constante de mortalidad (distribución exponencial), mientras que los Tipos I y III implican riesgos crecientes y decrecientes, respectivamente.

### 1.1.5. Función de supervivencia

La función de supervivencia,  $S(t)$  da la probabilidad de que un individuo sobreviva más de un tiempo especificado  $t$ . Da la probabilidad de que la variable aleatoria  $T$  exceda el tiempo especificado  $t$ . La función de supervivencia es fundamental para un análisis de supervivencia. La función de superviviente a menudo se expresa como una curva de Kaplan-Meier. La caída vertical en una curva de Kaplan-Meier indica un evento.

En el análisis de supervivencia, se utiliza información sobre el estado del evento y el tiempo de seguimiento para estimar una función de supervivencia. Por ejemplo, un estudio realizado durante 20 años para evaluar la supervivencia después de una enfermedad, en el cual el resultado es la mortalidad por todas las causas, la función de supervivencia (o curva de supervivencia) podría ser la que se muestra en la Figura 03.



**Figura 03. Curva de supervivencia y probabilidad de sobrevivir.** Fuente. Sullivan (2016).

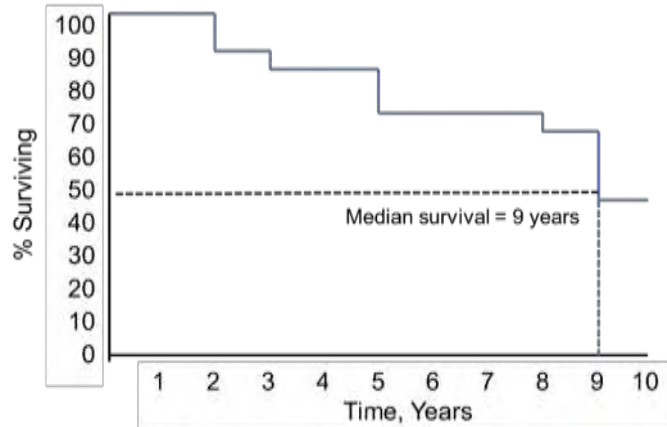
El eje horizontal representa el tiempo en años y el eje vertical muestra la probabilidad de sobrevivir o la proporción de individuos que sobreviven. A continuación, las características:

- En el momento cero, la probabilidad de supervivencia es 1,0 (o el 100% de los individuos están vivos).
- A los 2 años, la probabilidad de supervivencia es de aproximadamente 0,83 o 83 % (sobrevive el 83%).
- A los 10 años, la probabilidad de supervivencia es de aproximadamente 0,55 o 55 %.
- La mediana de supervivencia es de aproximadamente 11 años (el 50% de los individuos está vivo).

Una curva de supervivencia plana (es decir, una que se mantiene cerca de 1,0) sugiere una supervivencia muy buena, mientras que una curva de supervivencia que desciende bruscamente hacia 0 sugiere una supervivencia deficiente. La figura anterior muestra la función de supervivencia como una curva suave. En la mayoría de las aplicaciones, la función de supervivencia se muestra como una función escalonada en lugar de una curva suave.

Existen varios métodos paramétricos que se utilizan para modelar los datos de supervivencia y difieren en términos de las suposiciones que se hacen sobre la distribución de los tiempos de supervivencia en la población. Entre las distribuciones están: exponencial, Weibull, Gompertz, log-logística y log-normal (Cox y Oakes, 1984). Quizás la más popular es la distribución exponencial, que asume que la probabilidad de que un participante sufra el evento de interés es independiente del tiempo que el individuo ha estado libre de eventos. Otras distribuciones hacen diferentes suposiciones sobre la probabilidad de que un individuo desarrolle un evento (es decir, puede aumentar, disminuir o cambiar con el tiempo).

Los métodos no paramétricos no hacen suposiciones sobre cómo cambia la probabilidad de que un individuo desarrolle el evento con el tiempo. Utilizando métodos no paramétricos, se estima y gráfica la distribución de supervivencia. Las curvas de supervivencia a menudo se trazan como funciones escalonadas, como se muestra en la Figura 04. El tiempo se muestra en el eje X y la supervivencia (proporción de individuos en riesgo) se muestra en el eje Y. Se observa en la Figura 04 que la probabilidad de supervivencia es del 100% durante 2 años y luego desciende al 90%. La mediana de supervivencia es de 9 años (es decir, el 50% de la población sobrevive 9 años).



**Figura 04. Funcion de sobrevivencia indicando el valor de la mediana.** Fuente: Sullivan (2016).

### 1.1.6. Función de riesgo

La función de riesgo  $h(t)$  da el potencial instantáneo por unidad de tiempo para que ocurra el evento, dado que el individuo ha sobrevivido hasta el tiempo  $t$ . Es la probabilidad de falla en un período de tiempo infinitesimalmente pequeño entre  $Y$  y  $Y + \Delta y$ , dado que el sujeto ha sobrevivido hasta el momento  $y$ . En este sentido, el riesgo es una medida de tiempo: cuanto mayor es el riesgo entre los tiempos  $y_1$  y  $y_2$ , mayor es el riesgo de falla en este intervalo de tiempo. La función de riesgo tiene su propia importancia, ya que proporciona una idea de las tasas de fallas condicionales. Puede utilizarse para identificar una forma de modelo específica y es el vehículo mediante el cual se lleva a cabo el modelado matemático de los datos de supervivencia.

### 1.1.7. Razón de riesgo

La razón de riesgo (HR) es similar al riesgo relativo. Se ha utilizado para describir el resultado de ensayos terapéuticos en los que la pregunta es, ¿hasta qué punto el tratamiento puede acortar la duración de una enfermedad? La razón de riesgo es una estimación de la razón de la tasa de riesgo en el grupo tratado frente al grupo de control. Por ejemplo, si hay dos grupos, el grupo 1 y el grupo 2,  $HR = 4,5$  para el tratamiento significa que el riesgo (de recaída) para el grupo 2 es 4,5 veces mayor que el del grupo 1. Si  $HR = 1$ , entonces Grupo 1  $h(t) =$  Grupo 2  $h(t)$ . Por lo tanto, el efecto del tratamiento y el control sobre el tiempo, viene dado por un número de valor real. Aquí, es importante enfatizar que la



razón de los dos riesgos es para cualquier punto de tiempo  $t$  dado por HR (Tratamiento vs. Control) y, por lo tanto, no es la razón integrada en el tiempo. Específicamente, da el riesgo relativo instantáneo (en contraste, RR (riesgo relativo) cuantifica el riesgo acumulado integrado en el tiempo).

En contraste con la comparación de las curvas de supervivencia para el tratamiento y el control, por ejemplo, mediante una prueba de rango logarítmico, la cual da solo una distinción binaria, un HR dice algo sobre la magnitud y la dirección de esta diferencia. En pocas palabras, el HR tiene la siguiente interpretación:

- $HR (T \text{ vs. } C) > 1$ : el grupo de tratamiento experimenta un riesgo mayor que el grupo de control; se favorece al grupo de control,
- $HR (T \text{ vs. } C) = 1$ : No hay diferencia entre el grupo de tratamiento y el de control,
- $HR (T \text{ vs. } C) < 1$ : el grupo de control experimenta un riesgo mayor que el grupo de tratamiento); se favorece el grupo de tratamiento.

Por ejemplo, para  $HR (T \text{ vs } C) = 1,3$ , el riesgo del grupo de tratamiento aumenta en un 30% en comparación con el grupo de control y para  $HR (T \text{ vs } C) = 0,7$  el riesgo del grupo de tratamiento es 30 % menor en comparación con el grupo de control (Barraclough, Simms y Govindan, 2011; Spruance *et al.* 2004).



# CAPÍTULO 2

## Características y Procedimientos de Análisis

### 2.1. Estructura de datos

Los procedimientos LIFETEST, LIFEREG y PHREG utilizan la misma estructura básica de datos. Es una estructura estándar. Para cada caso en la muestra, debe existir una variable (por ejemplo, DUR = duración) que contiene el tiempo en que ocurre un evento o, para casos censurados, el último tiempo en el cual este caso fue observado, ambas medidas desde el origen seleccionado. Una segunda variable (llamada STATUS), necesaria si algunos de los casos son censurados o si se quieren distinguir diferentes clases de eventos. A la variable STATUS se asignan valores arbitrarios que indican el estatus del individuo en el tiempo registrado en la variable DUR. En caso de existir un único evento, es común disponer STATUS=1 para casos no censurados y STATUS=0 para casos censurados. Para PROC LIFEREG Y PROC PHREG, los cuales estiman modelos de regresión, el registro también debe contener valores de covariables, en caso de que las covariables sean constantes en el tiempo. Otras estructuras más complejas se requieren para covariables tiempos – dependientes (Allison, 2010).

**Tabla 01. Estructura de datos estándar para el manejo de la información**

Observación	DUR	STATUS	TRATAM	EFECTO
1	XX	1	1	1
2	XXX	0	2	0

DUR: Tiempo en días de la aplicación de tratamientos; STATUS= 1 para individuos vivos o no censurados; STATUS = 0 individuos muertos o censurados; TRATAM: Tratamiento aplicado; EFECTO: Efecto del tratamiento

## 2.2. Métodos de análisis

### 2.2.1. MÉTODO 1. Análisis de tiempo hasta el evento (Tablas de vida, Kaplan-Meier) usando PROC LIFETEST

La metodología para la elaboración de Tablas de vida (PROC LIFETEST), calcula estimaciones no paramétricas de la distribución de supervivencia de datos que podrían censurarse por la derecha debido al abandono o finalización del estudio. De estos datos, se conoce que el tiempo de vida excede el valor aportado. Generalmente, se quiere comparar curvas de supervivencia para determinar si los datos de dos organismos proceden de la misma distribución de supervivencia. Esta metodología calcula test de clasificación y un test de razón de verosimilitud para probar la homogeneidad de las funciones de supervivencia en los datos. También genera diagramas que permiten examinar las curvas de supervivencia.

Esta metodología se puede utilizar para estudiar la distribución de supervivencia de una población específica. Por ejemplo, estudiar la esperanza de vida de 40 ratas expuestas a un agente cancerígeno en una prueba pequeña y aleatoria y asignarlas a uno de dos grupos de tratamiento. La esperanza de vida es el tiempo desde que se realiza la prueba aleatoria hasta la muerte. El hecho de interés es la muerte a causa del cáncer por la introducción del agente cancerígeno. Se puede utilizar la metodología de Tablas de vida para determinar si las distribuciones de sobrevivencia entre los dos grupos de tratamiento, difieren entre sí. Con el procedimiento LIFETEST se puede solicitar el límite de producto (Kaplan y Meier) o la estimación de la tabla de vida (actuarial) de la distribución, calculando pruebas no paramétricas para comparar las curvas de sobrevivencia de dos o más grupos.

Los datos de sobrevivencia consisten en una variable de respuesta (tiempo hasta el evento, tiempo de falla o tiempo de sobrevivencia) que mide la duración del tiempo hasta que ocurre un evento específico y posiblemente un conjunto de variables independientes que se cree que están asociadas con la variable de tiempo de falla. Estas variables independientes (variables concomitantes, covariables o factores pronósticos) pueden ser cualitativas o categóricas, como el sexo o la raza, o cuantitativas, como la edad o la temperatura. El propósito del análisis de sobrevivencia es modelar la distribución subyacente de la variable de tiempo de falla y evaluar la dependencia de la variable de tiempo de falla de las variables independientes.

En los estudios longitudinales, es interesante estimar una curva de "sobrevivencia" para la población. ¿Qué proporción de la población sobrevive más allá de un intervalo de tiempo especificado sin que ocurra un evento en particular? La forma más sencilla de describir la supervivencia en una muestra, es calcular la Tabla de vida (LT). La distribución de los tiempos de supervivencia se divide en un cierto número de intervalos. Para cada intervalo, podemos calcular el número y la proporción de casos u objetos que ingresaron al intervalo respectivo 'vivos', el número y la proporción de casos que fallaron en el intervalo respectivo (número de eventos terminales o número de casos que murieron o número de semillas que germinaron), y el número de casos que se perdieron o censuraron en el intervalo respectivo. Con base en esos números y proporciones, se pueden calcular varias estadísticas adicionales, como el número de casos en riesgo, la proporción de fallas, la proporción de sobrevivientes, la función de supervivencia, la razón de riesgo y el tiempo medio de supervivencia. Este procedimiento se utiliza para muestras más grandes, en que los intervalos de tiempo son lo suficientemente grandes como para dividirse en unidades más pequeñas (Booth y Hirschl, 2005; Ives, Funk y Dennis, 2007).

#### **2.2.1.1. Especificación de los procedimientos generales para PROC LIFETEST**

Con la instrucción PROC LIFETEST se solicita el procedimiento. Las siguientes opciones pueden aparecer en la instrucción PROC LIFETEST. Si no se solicitan opciones, PROC LIFETEST calcula y muestra estimaciones de límite de producto de la distribución de sobrevivencia dentro de cada estrato y prueba la igualdad de las funciones de sobrevivencia en todos los estratos. Los siguientes procedimientos están disponibles en PROC LIFETEST:

```
PROC LIFETEST < options >;  
TIME variable < *censor (list) >;  
BY variables;  
FREQ variable;  
ID variables;  
STRATA variable < (list) > < ... variable < (list) > >;  
SURVIVAL options;  
TEST variables;
```

Para el uso más simple solo se requieren la instrucción PROC LIFETEST y la instrucción TIME. Puede utilizar la declaración STRATA para dividir los datos en varios estratos. A continuación, se estima una función de sobreviviente separada para cada estrato y se realizan pruebas de homogeneidad de los estratos. Sin embargo, si la opción GROUP = también se especifica en la instrucción STRATA, se realizan pruebas estratificadas para probar las k muestras definidas por la variable GROUP = mientras se controla el efecto de las variables STRATA. Puede usar la instrucción SURVIVAL para generar las estimaciones de la función de sobreviviente en un conjunto de datos. Puede especificar covariables en la instrucción TEST.

Todas las sentencias, excepto la sentencia TIME, son opcionales y no hay un orden obligatorio para las sentencias que siguen a la sentencia PROC LIFETEST. La sentencia TIME se utiliza para especificar las variables que definen el tiempo de supervivencia y el indicador de censura. La declaración STRATA especifica una variable o un conjunto de variables que definen los estratos para el análisis. La instrucción SURVIVAL le permite especificar una transformación que se utilizará en el cálculo de los intervalos de confianza; también le permite generar intervalos de confianza simultáneos. La instrucción TEST especifica una lista de covariables numéricas que se probarán para determinar su asociación con

el tiempo de supervivencia de la respuesta. Cada variable se prueba individualmente y también se calcula una estadística de prueba conjunta. La declaración ID proporciona una lista de variables cuyos valores se utilizan para identificar observaciones en las estimaciones de límite de producto de la función de supervivencia.

#### TIME Statement

TIME variable < \*censor(list) >; se utiliza para indicar la variable de tiempo de falla, que opcionalmente puede ir seguida de un asterisco, el nombre de la variable de censura y una lista entre paréntesis de valores que corresponden a la censura por la derecha. Los valores de censura deben ser numéricos, valores no perdidos. Por ejemplo, la declaración

Time T \* flag (1,2); identifica la variable T que contiene los valores del evento o tiempo censurado. Si la variable flag tiene valor 1 o 2, el valor correspondiente de T es un valor censurado por la derecha.

#### BY Statement

BY variables; puede especificar una instrucción BY para obtener análisis separados de las observaciones en grupos definidos por las variables BY. La declaración BY es más eficiente que la declaración STRATA para definir estratos en grandes conjuntos de datos. Sin embargo, si se usa la instrucción BY para definir estratos, PROC LIFETEST no agrupa los estratos para probar la asociación del tiempo de supervivencia con covariables, tampoco prueba la homogeneidad entre los grupos definidos por la opción BY. El tamaño del intervalo se calcula por separado para cada grupo BY. Cuando los intervalos se determinan de forma predeterminada, pueden ser diferentes para cada grupo BY. Para hacer que los intervalos sean los mismos para cada grupo BY, se utiliza la opción INTERVALS = en la instrucción PROC LIFETEST.

Cuando aparece una instrucción BY, el procedimiento espera que el conjunto de datos de entrada sea sorteado en el orden de las variables BY. Si el conjunto de datos de entrada no está ordenado en orden ascendente, se utiliza una de las siguientes alternativas:

- Se ordenan los datos utilizando el procedimiento SORT con una declaración BY.
- Se especifican las opciones NOTSORTED o DESCENDING en la instrucción BY para el pro-



cedimiento. La opción NOTSORTED no significa que los datos no estén ordenados, sino que los datos están organizados en grupos (de acuerdo con los valores de las variables BY) y que estos grupos no están necesariamente en orden alfabético o numérico creciente.

- Se crea un índice sobre las variables BY utilizando el procedimiento DATASETS.

#### FREQ Statement

FREQ variable; la declaración FREQ identifica una variable que contiene la frecuencia de ocurrencia de cada observación. PROC LIFETEST trata cada observación como si apareciera n veces, donde n es el valor de la variable FREQ para la observación. La declaración FREQ es útil para producir tablas de vida cuando los datos ya están en forma de conjunto de datos de resumen. Si no es un número entero, el valor de frecuencia se trunca a un número entero. Si el valor de frecuencia es menor que uno, no se utiliza la observación.

#### STRATA Statement

STRATA variable < (list) > < ... variable < (list) > > < /options >; La declaración STRATA indica las variables que determinan los niveles de estratos para los cálculos. Los estratos se forman de acuerdo con los valores no perdidos de las variables designadas. La opción MISSING se puede usar para permitir los valores perdidos como un nivel de estrato válido. Otras opciones le permiten especificar varias pruebas de k-muestras, pruebas de tendencia y pruebas estratificadas.

En la sintaxis anterior, list es una lista de puntos finales para una variable numérica. Los valores de la variable pueden tener formato o no. Si la variable es una variable de carácter, o numérica y no aparece ninguna lista, entonces los estratos están definidos por los valores únicos de la variable de estratos. Se puede especificar más de una variable en la instrucción STRATA, y cada variable numérica puede ir seguida de una lista. Cada intervalo contiene su punto final inferior pero no contiene su punto final superior. Los estratos correspondientes están formados por la combinación de niveles. Si una variable es numérica y va seguida de una lista, los niveles de esa variable corresponden a los intervalos definidos por la lista. La especificación de variables STRATA puede tener cualquiera de las siguientes formas (Cuadro 01):

**Cuadro 01. Formas de especificación para variables Strata.**

Lista separada por espacios en blanco	strata edad (5 10 20 30)
Lista separada por comas	strata edad (5,10,20,30)
x to y	strata edad (5 to 10)
x to y by z	strata edad (5 to 30 by 10)
Combinacion de los anteriores	strata edad (5,10 to 50 by 10)

Fuente. SAS (2020).

Por ejemplo, la especificación: STRATA edad (5,10 to 50 by 10 años) sexo; indica los siguientes niveles para la variable Edad. Es lo mismo que escribir: STRATA edad (5,10 to 50 by 10) sexo; indica los siguientes niveles para la variable edad:

$$\{(-\infty, 5), [5, 20), [20, 30), [30, 40), [40, 50), [50, \infty)\}$$

Esta declaración también especifica que los estratos de edad se subdividen aún más por los valores de la variable Sexo. En este ejemplo, hay 6 grupos de edad por 2 grupos de sexo, formando un total de 12 estratos.

La especificación de varias variables (por ejemplo, A B C) es equivalente a la sintaxis A \* B \* C ... de la instrucción TABLES en el procedimiento FREQ. El número de niveles de estratos generalmente crece muy rápidamente con el número de variables de STRATA, por lo que debe tener cuidado al especificar la lista de variables de STRATA. Las siguientes opciones pueden aparecer en la instrucción STRATA después de una barra ("/"). Aparte de la opción MISSING, estas opciones están dedicadas a las pruebas de dos o más muestras de datos de supervivencia.

GROUP=variable, especifica la variable cuyos valores identifican las diversas muestras cuyas curvas de supervivencia subyacentes se van a comparar. Las pruebas se estratifican según los niveles de las variables STRATA. Por ejemplo, en un ensayo en el que se van a comparar dos formas de terapia en dos centros de investigación, se especifica la variable que identifica las terapias como la variable GROUP = y la variable que identifica los centros como la variable STRATA, para realizar una prueba estratificada de 2 muestras para comparar las terapias controlando el efecto de los centros.

MISSING; permite que valores perdidos sean un nivel de estrato o valores válidos de GROUP= variable.

**NODETAIL** No exhibe rangos estadísticos y las correspondientes matrices de covarianza para varios estratos. Si se especifica la opción **TREND**, se suprime la exhibición de registros para calcular las pruebas de tendencia.

**NOTEST** Suprime las pruebas de k muestras, pruebas estratificadas y pruebas de tendencia.

**TREND** Calcula las pruebas de tendencia para probar la hipótesis nula de que la razón de riesgo de la población k es la misma frente a alternativas ordenadas. Si solo hay una variable **STRATA** y la variable es numérica, los valores sin formato de la variable se utilizan como puntajes; de lo contrario, las puntuaciones son 1, 2, ..., en el orden dado de los estratos.

**TEST=(list)** Permite seleccionar funciones de peso para k-muestras de prueba, pruebas estratificadas o también, pruebas de tendencia. Se puede especificar **list** que contiene una o más de las siguientes palabras clave:

**LOGRANK:** especifica la prueba de rango log (log-rank test).

**WILCOXON:** especifica la prueba de Wilcoxon. También es referida como prueba Gehan o prueba Breslow.

**TARONE:** especifica la prueba Tarone-Ware.

**PETO:** especifica la prueba Peto-Peto. También es referida como prueba Peto-Peto-Prentice.

**MODPETO:** especifica una prueba modificada de Peto-Peto.

**LR:** especifica la prueba razón de probabilidad o verosimilitud (likelihood ratio test), basada en el modelo exponencial.

**ALL:** especifica todas las pruebas no paramétricas con  $\rho_1=1$  y  $\rho_2=0$  para la prueba de FLEMING (1, 0).

Por defecto, **TEST= (LOGRANK WILCOXON LR)** para k-sample tests, y **TEST= (LOGRANK WILCOXON)** para pruebas estratificadas y de tendencia.

**SURVIVAL Statement**

**SURVIVAL options;** crea un conjunto de datos de salida con los resultados de la estimación de la fun-

ción de sobreviviente. Aunque puede utilizar la opción OUTSURV = para producir el conjunto de datos de salida, la instrucción SURVIVAL le permite generar bandas de confianza y especificar una transformación del tiempo de supervivencia en el cálculo de los intervalos de confianza puntuales y las bandas de confianza. Las opciones de las sentencias (ALPHA =, INTERVALS =, REDUCEOUT y TIME-  
LIST =) que se aplican a los datos OUTSURV = también se pueden especificar en las sentencias SURVIVAL (Cuadro 02).

**Cuadro 02. Resumen de opciones de la instrucción SURVIVAL**

Tarea	Opciones	Descripción
Especificación de Datos	OUT=	Especifica la salida de datos
Especificación de Transformación	CONFTYPE=	Especifica la transformación para el cálculo de intervalos de confianza puntuales y simultáneos para la función de sobreviviente: $\text{LogLog } g(x) = \log(-\log(x))$ $\text{Lineal } g(x) = x$ $\text{Log } g(x) = \log(x)$ $\text{Logit } g(x) = \log((x/(1-x)))$
Especificación de bandas de Confidencia	CONFBAND=	Especifica las bandas de confidencia para la salida
	BANDMAX=	Especifica el tiempo máximo para las bandas de confidencia
	BANDMIN=	Especifica el tiempo mínimo para las bandas de confidencia
Standard Errors	STDERR	Salida de errores estándar

Fuente. SAS (2020).

#### TEST Statement

TEST variables; especifica una lista de covariables numéricas (continuas) cuya asociación se desea probar con el tiempo de falla. Se calculan dos conjuntos de estadísticas de clasificación. Estas estadísticas de clasificación y sus varianzas se agrupan en todos los estratos y se muestran pruebas estadísticas univariadas (marginales) para cada una de las covariables. Además, se muestra una secuencia de pruebas estadísticas para los efectos conjuntos de las covariables. El primer elemento de la secuencia es el estadístico univariado de mayor valor. Luego, se agregan otras variables sobre la base del mayor aumento en la prueba estadística conjunta. El proceso continúa hasta que se hayan agregado todas las variables o hasta que las variables restantes sean linealmente dependientes de las variables agregadas previamente.

### 2.2.1.2. Especificación de las opciones para PROC LIFETEST

PROC LIFETEST < options >;

**ALPHA=value** Especifica un número entre 0,0001 y 0,9999 que establece el nivel de confianza de los intervalos de confianza para la función de superviviente. Por ejemplo, la opción ALPHA = 0,05 es un valor predeterminado que solicita un intervalo de confianza del 95% para el SDF en cada tiempo.

**INTERVALS=values** Especifica una lista de puntos finales de intervalo para el método de tabla de vida. Se supone que el intervalo inicial comienza en cero, se especifique o no cero en la lista. Cada intervalo contiene su punto final inferior pero no contiene su punto final superior. La opción INTERVALS = se puede especificar de cualquiera de las siguientes formas (Cuadro 03):

**Cuadro 03. Formas de especificación para la opción INTERVALS=**

Lista separada por espacios en blanco	intervals=1 3 5 7
Lista separada por comas	intervals=1,3,5,7
x to y	intervals=1 to 7
x to y by z	intervals=1 to 7 by 1
Combinación de los anteriores	intervals=1,3 to 5,7

Fuente. SAS (2020).

Si se selecciona Límite de producto, como método de estimación, se puede reducir el número de estimaciones de supervivencia que se indican, definiendo puntos finales para cada intervalo. Por el contrario, al seleccionar Tabla de vida, como método de estimación, puede especificar información sobre la construcción de intervalos que se utilizan en el cálculo de la función de supervivencia. Por defecto, se usan 10 intervalos. También puede especificar un valor diferente, seleccionando número e introduciendo un valor. También puede construir intervalos especificando ancho del intervalo, seleccionando WIDTH= y se introduce un valor.

**NINTERVAL=value** Especifica el número de intervalos utilizados para calcular las estimaciones de la tabla de vida de la función de superviviente. Este parámetro es anulado por la opción WIDTH = o la opción INTERVALS = Cuando se especifica la opción NINTERVAL =, PROC LIFETEST intenta encontrar un intervalo para los puntos finales. Se utiliza la opción INTERVALS = para controlar los puntos finales del intervalo. El valor predeterminado es NINTERVAL = 10.

METHOD=type Especifica el método utilizado para calcular las estimaciones de la función de supervivencia. Puede especificar producto limite (Kaplan-Meier = KM) o Tabla de vida (LT). Los dos métodos generan resultados similares. Sin embargo, el método de límite de producto se basa en los tiempos de supervivencia actuales y el método de tabla de vida agrupa los tiempos de supervivencia en intervalos. Cuando el tamaño muestral es muy grande, puede resultar útil agrupar los tiempos de supervivencia en intervalos y realizar el análisis de tabla de vida. El método predeterminado es el límite de producto. Los métodos son los siguientes:

PL | KM: Especifica si se utiliza el estimado producto – limite (PL) o de Kaplan-Meier (KM).

ACT | LIFE | LT: Especifica si se utilizan los estimados tabla de vida o actuariales.

Por defecto, se utiliza METHOD=PL.

PLOTS= (type <(NAME=name)> <, ..., type <(NAME=name)> >) Crea gráficos de estimaciones de supervivencia u observaciones censuradas, donde type, corresponde al tipo de gráfico y NAME, es un nombre de entrada de catálogo de hasta ocho caracteres. Los valores válidos de t son los siguientes:

CENSORED | C: Especifica una gráfica de observaciones censuradas por estrato (solo para método producto-limite).

SURVIVAL | S: Especifica una gráfica de estimados SDF vs tiempo.

LOGSURV | LS: Especifica una gráfica de –LOG (SDF estimada) vs tiempo

LOGLOGS | LLS: Especifica una gráfica del LOG (-LOG (SDF estimada)) vs LOG (tiempo).

HAZARD | H: Especifica una gráfica de la función de riesgo estimada vs tiempo (solo para método de tabla de vida - LT).

PDF | P: Especifica una gráfica de la función de densidad de probabilidad vs tiempo (solo para método de tabla de vida). Se requiere el uso de parenthesis en la especificación del tipo de gráfica. Por ejemplo,

PLOTS = (S): Solicita una gráfica de la función de supervivencia estimada vs tiempo.

PLOTS = (S(NAME=SURV2), H(NAME=HAZ2)): Solicita un gráfico de la función de supervivencia

estimada frente al tiempo y un gráfico de la función de riesgo estimada frente al tiempo, con `Surv2` y `Haz2` como sus nombres de catálogo, respectivamente. Esta opción genera un diagrama de la función de distribución de supervivencia que se usa para describir el tiempo de vida de la población de interés.

Para incluir los límites de confianza, se selecciona incluir intervalos de confianza puntuales. Los límites de confianza para los diagramas están controlados en la configuración de Métodos para los intervalos de confianza. El valor predeterminado es un intervalo de confianza del 95%. Para que todos los diagramas de diferentes estratos aparezcan en un solo diagrama con la leyenda, se selecciona la casilla Superponer estratos en un solo diagrama. Si ha asignado una variable al rol Variables de estrato y la casilla Mostrar diagrama de función de supervivencia está seleccionada, todos los diagramas para los diferentes estratos están en un único diagrama de manera predeterminada.

### **2.2.2. METODO 2. Análisis de supervivencia con modelos de regresión Cox usando PROC PHREG**

Para análisis de supervivencia se utilizan dos tipos de modelos de regresión: el modelo de tiempo de falla acelerada (AFT), disponible en `LIFEREG` y el modelo de Cox de riesgo proporcional, disponible en `PHREG`. El modelo AFT asume una forma paramétrica para los efectos de las variables explicatorias en la función de supervivencia. De igual manera, el modelo de Cox también asume una forma paramétrica para los efectos de las variables explicatorias, pero sigue una forma no especificada para la función de supervivencia. El procedimiento `PHREG` proporciona métodos de selección de modelos escalonados que incluyen modelos hacia atrás (`backward`), hacia adelante (`forward`) y escalonamiento (`stepwise`). `PROC PHREG` permite especificar niveles de significancia para la entrada y eliminación de efectos, agregar efectos en un orden secuencial, especificar el número de variables en el modelo para la selección hacia adelante o hacia atrás y seleccionar los mejores subconjuntos (Gordon, 2020).

La regresión de Cox se considera un procedimiento "semiparamétrico" porque no es necesario especificar la función de riesgo de referencia,  $h_0(t)$ . Hay dos suposiciones sobre el modelo de riesgo proporcional de Cox: las razones de riesgo de dos individuos son independientes del tiempo y son válidas solo para covariables independientes del tiempo. Esto significa que las funciones de riesgo para dos individuos, en cualquier momento, son proporcionales. En otras palabras, si un individuo tiene un riesgo

de muerte en algún momento inicial que es dos veces más alto que el de otro individuo, entonces en todos los momentos posteriores el riesgo de muerte sigue siendo el doble.

El modelo de análisis de regresión de riesgos proporcionales de Cox tiene en cuenta el efecto de varias variables a la vez y examina la relación de la distribución de supervivencia con estas variables. Es similar al análisis de regresión múltiple, pero la diferencia es que la variable dependiente es la función de riesgo en un tiempo  $t$ . Se basa en intervalos de tiempo muy pequeños, denominados “gatillos” de tiempo, que contienen como máximo un evento de interés. Es un enfoque semiparamétrico para la estimación de pesos en un modelo de riesgo proporcional.

El coeficiente  $\beta_0(t)$  es la función de riesgo de la línea de base y se define como la probabilidad de experimentar el evento de interés cuando todas las demás covariables son iguales a cero. Y es el único componente del modelo dependiente del tiempo. El modelo no hace suposiciones sobre la función de riesgo de línea de base y asume una forma paramétrica para el efecto de las covariables sobre el riesgo.

El signo de los coeficientes de regresión,  $\beta_i$ , juega un papel en el riesgo de un sujeto. Un cambio en estos coeficientes de regresión o covariables aumentará o disminuirá el riesgo de referencia. Un signo positivo para  $\beta_i$  significa que el riesgo de un evento es mayor y, por lo tanto, el pronóstico del evento de interés para ese sujeto en particular es mayor. Del mismo modo, un signo negativo significa que el riesgo del evento es menor. Además, hay que tener en cuenta que el valor en sí, también juega un papel importante. Por ejemplo, el valor de una variable igual a uno significaría que no tendrá ningún efecto sobre el riesgo. Para un valor menor que uno, reducirá el riesgo y para un valor mayor que uno, aumentará el riesgo.

El modelo de Cox puede manejar datos censurados por la derecha, pero no puede manejar datos censurados por la izquierda o censurados por intervalos directamente. Hay algunas covariables que no obedecen al supuesto de riesgo proporcional. Se les permite seguir siendo parte del modelo, pero sin estimar su efecto. A esto se le llama estratificación. El conjunto de datos se divide en  $N$  conjuntos de datos más pequeños basados en valores únicos de las covariables de estratificación. Cada conjunto de datos más pequeño tiene su propio riesgo de línea de base, que constituye la parte no paramétrica del modelo, y todos tienen parámetros de regresión comunes, que conforman la parte paramétrica del mo-



delo. No hay un parámetro de regresión para las covariables estratificadas.

### 2.2.2.1. Sintaxis de los procedimientos

Los siguientes procedimientos están disponibles en PROC PHREG.

PROC PHREG < options >;

ASSESS keyword </ options >;

CLASS variables;

MODEL response < \*censor (list) > = variables </options >;

STRATA variable < (list) > < ...variable < (list) > > </option >;

TEST equation1 <... equationk > </option >;

FREQ variable;

WEIGHT variable </option >;

ID variables;

OUTPUT <OUT=SAS-data-set>;

BASELINE <OUT=SAS-data-set> <COVARIATES=SAS-data-set>;

Todas las declaraciones, excepto MODEL son opcionales. Los elementos dentro de < > son opcionales y no hay un orden obligatorio para las declaraciones que siguen a la instrucción PROC PHREG. La declaración CLASS determina las variables explicativas categóricas. La declaración MODEL especifica las variables que definen el tiempo de supervivencia, la variable de censura y las variables explicativas. La declaración STRATA especifica una variable o un conjunto de variables que definen los estratos para el análisis. La declaración TEST contiene ecuaciones que definen hipótesis lineales sobre los parámetros del modelo. La declaración de ID especifica las variables con valores que se utilizan para etiquetar las observaciones. Las declaraciones OUTPUT y BASELINE crean conjuntos de datos que contienen las estimaciones de supervivencia. Se pueden incluir declaraciones de programación de pasos de DATOS para crear variables explicativas dependientes del tiempo.

### 2.2.2.2. Uso básico de PROC PHREG

Para utilizar PROC PHREG, se debe disponer para cada sujeto, una variable de tiempo y una variable de censura (o evento) para indicar el estado de supervivencia. Además, tener valores para una o más covariables que se quieran analizar para determinar el posible efecto sobre la supervivencia. Para solicitar el procedimiento, se necesita la instrucción PROC, seguida de una instrucción MODEL que indique el tiempo y (opcionalmente) la variable de censura y las covariables que se utilizarán en el análisis (Cantor, 2003). La sintaxis básica es:

```
PROC PHREG;
```

```
MODEL timevar * censvar (censval (s)) = covariable (s);
```

Aquí timevar, es la variable de tiempo y censvar es la variable de censura. Si no se censuran observaciones, no se necesita censvar. En la sintaxis, censval (s) representa los valores de censvar que indican que un tiempo está censurado y covariable (s) representa la lista de variables a considerar como las covariables de la regresión. La instrucción MODEL también puede incluir una opción para especificar cómo se manejan los vínculos en la variable de tiempo. El valor predeterminado es el método de Breslow. PROC PHREG usa solo aquellas observaciones para las cuales no estén ausentes todas estas variables. Por lo tanto, no se deben incluir covariables en las que muchos sujetos tienen valores perdidos.

### 2.2.2.3. Especificación de las opciones para PROC PHREG

```
PROC PHREG < options >;
```

Las siguientes opciones se pueden utilizar en PROC PHREG:

**COVOUT**     Agrega la matriz de covarianza estimada de los parámetros al conjunto de datos **OUTEST =**. La opción **COVOUT** no tiene ningún efecto a menos que se especifique la opción **OUTEST =**.

**COVM**        Solicita que la matriz de covarianza basada en el modelo (que es la inversa de la matriz de información observada) se presente y se utilice en el análisis si también se especifica la opción **COVS**. La opción **COVM** no tiene ningún efecto si no se especifica la opción **COVS**.

```
COVSANDWICH < (AGGREGATE) >
```

**COVS < (AGGREGATE) >** solicita la estimación para la matriz de covarianza. Cuando se especifica esta opción, se utiliza en las pruebas de Wald para probar la hipótesis nula global, las hipótesis nulas de parámetros individuales y las hipótesis en las declaraciones TEST. Además, se calcula una prueba de puntuación modificada en la prueba de la hipótesis nula global, y la tabla de estimaciones de parámetros tiene una columna StdErrRatio adicional, que contiene las razones de la estimación robusta del error estándar en relación con la estimación basada en el modelo correspondiente. Opcionalmente, puede especificar la palabra clave AGGREGATE entre paréntesis después de la opción COVSANDWICH (o COVS), que solicita un resumen de los residuos de puntuación para cada patrón de ID distinto en el cálculo de la estimación de covarianza de sándwich robusta. Esta opción AGGREGATE no tiene ningún efecto si no se especifica la instrucción ID.

**DATA=SAS-data-set** Nombra el conjunto de datos SAS que contiene los datos a analizar. Si se omite la opción DATA =, el procedimiento utiliza el conjunto de datos SAS creado más recientemente.

**MULTIPASS** Solicita que, para cada iteración de Newton-Raphson, PROC PHREG recompile los conjuntos de riesgos correspondientes a los tiempos del evento para el estilo de respuesta (inicio, parada) y vuelva a calcular los valores de las variables dependientes del tiempo definidas por las declaraciones de programación para cada observación en el conjunto de riesgo. Si no se especifica la opción MULTIPASS, PROC PHREG calcula todos los conjuntos de riesgos y todos los valores de las variables y los guarda en un archivo de utilidad. La opción MULTIPASS tiene efecto solo cuando se usa el estilo de respuesta (iniciar, detener) o cuando hay variables explicativas dependientes del tiempo.

**NOPRINT** Suprime toda la salida. Esta opción desactiva temporalmente el sistema de entrega de salida (ODS).

**NOSUMMARY** Suprime la visualización del evento y las frecuencias de observación censuradas.

**OUTEST=SAS-data-set** Crea un conjunto de datos de salida que contiene estimaciones de los coeficientes de regresión. Si se usa la opción COVOUT, el conjunto de datos también contiene la matriz de covarianza estimada de los parámetros.

**SIMPLE** Muestra estadísticas descriptivas simples (media, desviación estándar, mínimo y máximo) para cada variable explicativa en la declaración MODEL.

### 2.2.2.3.1. Especificación de los procedimientos generales para PROC PHREG

ASSESS Statement.

ASSESS < var=(list) > < PH > < /options >; La declaración ASSESS realiza los métodos gráficos y numéricos para verificar la idoneidad (cumplimiento de supuestos) del modelo de regresión de Cox. Los métodos se derivan de sumas acumuladas de residuos de martingale durante los tiempos de seguimiento o valores de covariables. Debe especificar al menos uno de los siguientes para crear un análisis.

VAR=(list) Especifica la lista de variables explicativas para las que se evalúan sus formas funcionales. Para cada variable de la lista, los residuos acumulativos de martingala se grafican contra los valores de la variable explicativa junto con 20 patrones residuales simulados (o n si se especifica NPATHS = n).

PROPORTIONAL HAZARDS

PH Solicita la comprobación del supuesto de riesgos proporcionales. Para cada variable explicativa del modelo, los registros observados del componente se grafican frente al tiempo de seguimiento junto con 20 patrones simulados (o n si se especifica NPATHS = n). Las siguientes opciones se pueden especificar después de una barra inclinada (/).

NPATHS=n Especifica el número de patrones residuales simulados que se mostrarán en un gráfico de residuos de martingala acumulativo o un gráfico de registros del proceso. El valor predeterminado es n = 20.

CRPANEL Solicita una gráfica con cuatro paneles, cada uno con los residuos acumulativos de martingala observados y dos patrones residuales simulados.

RESAMPLE < =n > Solicita que la prueba suprema de tipo Kolmogorov se calcule en 1000 patrones simulados o en n patrones simulados si se especifica n.

SEED=n Especifica una semilla entera para el generador de números aleatorios utilizado en la creación de realizaciones simuladas para gráficos y para las pruebas supremas de tipo Kolmogorov. La

especificación de una semilla le permite reproducir gráficos y valores  $p$  idénticos para las evaluaciones del modelo de la misma especificación PHREG. Si no se especifica la opción SEED =, o si especifica una semilla no positiva, se deriva una semilla aleatoria del tiempo del día.

#### CLASS Statement

CLASS variables. Las variables de clasificación en lugar de variables numéricas cuantitativas deben enumerarse en la declaración CLASS. Para cada variable explicativa listada en la declaración CLASS, se generan variables indicadoras para los niveles asumidos por la variable CLASS. Si se utiliza la declaración CLASS, debe aparecer antes de la declaración MODEL.

#### MODEL Statement

MODEL response < \*censor (list) > = variables < /options >;

MODEL (t1, t2) < \*censor(list) > = variables < /options >; La declaración MODEL identifica las variables que se utilizarán como variables de tiempo de falla o evento, la variable de censura opcional y las variables explicativas. Se pueden especificar dos formas de sintaxis MODEL; la primera forma permite una variable de respuesta, mientras que la segunda forma permite dos variables para el estilo de entrada del proceso de conteo.

En la primera declaración MODEL, que precede al signo igual, está el nombre de la variable de tiempo de falla. Opcionalmente, esto puede ir seguido de un asterisco, el nombre de la variable de censura y una lista de valores de censura (separados por espacios en blanco o comas si hay más de uno) entre paréntesis. Si la variable de censura toma uno de estos valores, se considera censurado el tiempo de falla correspondiente. Las variables que siguen al signo igual son las variables explicativas (a veces llamadas variables independientes o covariables) del modelo.

En lugar de una única variable de tiempo de falla, la segunda declaración MODEL identifica un par de variables de tiempo de falla. Sus nombres están entre paréntesis y significan los puntos finales de un intervalo semicerrado (t1, t2) durante el cual el sujeto está en riesgo. Si la variable de censura toma uno de los valores de censura, el tiempo t2 se considera censurado.

La variable censura y las explicativas deben ser numéricas. Las variables de tiempo de falla deben contener valores no negativos. Cualquier observación con un tiempo de falla negativo se excluye del análisis, al igual que cualquier observación con un valor faltante para cualquiera de las variables enumeradas en la declaración MODEL. Puede especificar las siguientes opciones en la declaración MODEL.

#### Ties-Handling Option

TIES=method especifica cómo manejar los empates en el tiempo de falla. Esta opción puede tomar los siguientes valores:

BRESLOW utiliza la probabilidad aproximada de Breslow. Valor predeterminado.

DISCRETE reemplaza el modelo de riesgos proporcionales por el modelo logístico discreto

EFRON utiliza la probabilidad aproximada de Efron.

EXACT calcula la probabilidad condicional exacta bajo el supuesto de riesgos proporcionales de que todos los tiempos de eventos empatados ocurren antes de los tiempos censurados del mismo valor o antes de valores mayores. Esto es equivalente a sumar todos los términos de la probabilidad marginal que son consistentes con los datos observados.

#### Model-Specification Options

ENTRYTIME=variable

ENTRY=variable especifica el nombre de la variable que representa el tiempo de truncamiento izquierdo. Esta opción no tiene ningún efecto cuando se especifica el estilo de entrada del proceso de recuento.

NOFIT realiza la prueba de puntuación global, que prueba la significancia conjunta de todas las variables explicativas en la declaración MODEL. No se estiman parámetros. Si se especifica la opción NOFIT junto con otras opciones de la instrucción MODEL, NOFIT tiene prioridad y todas las demás opciones se ignoran, excepto la opción TIES =.

SELECTION=method especifica el método utilizado para seleccionar el modelo. Los métodos disponibles son

BACKWARD | B solicita la eliminación hacia atrás.

FORWARD | F solicita la selección hacia adelante.

NONE | N se ajusta al modelo completo especificado en la declaración MODEL. Este es el valor predeterminado.

SCORE solicita la mejor selección de subconjunto. Identifica un número específico de modelos con la estadística de chi-cuadrado más alta para todos los tamaños de modelos posibles, que van desde una variable explicativa hasta el total de variables explicativas enumeradas en la declaración MODEL.

STEPWISE | S solicita una selección escalonada o paso a paso.

**Model-Building Options** Las siguientes opciones permiten especificaciones adicionales para los métodos de selección de modelo hacia atrás (BACKWARD), hacia adelante (FORWARD), registros (SCORE) y paso a paso (STEPWISE). No tienen ningún efecto cuando SELECTION = NONE. Solo las opciones INCLUDE =, START =, STOP = y BEST = funcionan con el método SCORE.

**BEST=n** Se utiliza exclusivamente con el método de selección de modelo SCORE. Esta opción, especifica que se mostrarán n modelos con las estadísticas de chi-cuadrado más alta para cada tamaño de modelo. Si se omite la opción y existen menos de 10 variables explicativas, se enumeran todos los modelos posibles para cada tamaño de modelo. Si se omite la opción y hay más de 10 variables explicativas, entonces el número de modelos seleccionados para cada tamaño de modelo es, como máximo, igual al número de variables explicativas enumeradas en la declaración MODEL.

**DETAILS** Produce una visualización detallada en cada paso del proceso de construcción del modelo. Produce una tabla de "Análisis de variables que no están en el modelo" antes de mostrar la variable seleccionada para la selección FORWARD o STEPWISE. Para cada modelo ajustado, produce la tabla "Análisis de estimaciones de máxima verosimilitud".

INCLUDE=*n* Incluye las primeras *n* variables explicativas enumeradas en la declaración MODEL en cada modelo. El valor de *n* varía de 1 a *s*, donde *s* es el número de variables explicativas en la declaración MODEL. El valor predeterminado de *n* es 0

MAXSTEP=*n* Especifica el número máximo de veces que las variables explicativas pueden entrar y salir del modelo antes de que finalice el proceso de construcción del modelo STEPWISE. El valor predeterminado para *n* es el doble del número de variables explicativas en la declaración MODEL. La opción no tiene ningún efecto para otros métodos de selección de modelos.

SEQUENTIAL Ordena que las variables se agreguen al modelo en el orden especificado en la instrucción MODEL o que se eliminen del modelo en el orden inverso especificado en la instrucción MODEL.

SLENTRY=*value*

SLE=*value* Especifica el nivel de significancia (un valor entre 0 y 1) para ingresar una variable explicativa en el modelo en el método FORWARD o STEPWISE. Para todas las variables que no están en el modelo, se ingresa la que tiene el valor *p* más pequeño si el valor *p* es menor o igual que el nivel de significancia especificado. El valor predeterminado es 0,05.

SLSTAY=*value*

SLS=*value* Especifica el nivel de significancia (un valor entre 0 y 1) para eliminar una variable explicativa del modelo en el método BACKWARD o STEPWISE. Para todas las variables del modelo, la que tiene el mayor valor *p* se elimina si el valor *p* excede el nivel de significancia especificado. El valor predeterminado es 0,05.

START=*n* Comienza el proceso de selección del modelo FORWARD, BACKWARD o STEPWISE con las primeras *n* variables explicativas enumeradas en la declaración MODEL. El valor de *n* varía de 0 a *s*, donde *s* es el número total de variables explicativas en la declaración MODEL. El valor predeterminado de *n* es *s* para el método BACKWARD y 0 para los métodos FORWARD y STEPWISE. La opción START = *n* especifica solo que las primeras *n* variables explicativas aparecen en el primer modelo, mientras que INCLUDE = *n* especifica que las primeras *n* variables explicativas se incluirán en



cada modelo. Para el método SCORE,  $START = n$  especifica que los modelos más pequeños contienen  $n$  variables explicativas, donde  $n$  varía de 1 a  $s$ . El valor predeterminado de  $n$  es 1.

$STOP=n$  Especifica el número máximo (método FORWARD) o mínimo (método BACKWARD) de variables explicativas que se incluirán en el modelo final. El valor de  $n$  varía de 0 a  $s$ , donde  $s$  es el número de variables explicativas en la declaración MODEL. El valor predeterminado de  $n$  es 0 para el método BACKWARD y  $s$  para el método FORWARD. Para el método SCORE,  $STOP = n$  especifica que los modelos más grandes contienen  $n$  variables explicativas, donde  $n$  varía de 1 a  $s$ . El valor predeterminado de  $n$  es  $s$ . La opción  $STOP =$  no tiene ningún efecto para el método STEPWISE.

### STOPRES

SR Especifica que la adición y eliminación de variables se basará en el resultado de la prueba de puntuación de verosimilitud para probar la significancia conjunta de variables que no están en el modelo. Esta estadística se denomina chi-cuadrado residual. En el método FORWARD, la opción STOPRES ingresa las variables explicativas en el modelo de una en una hasta que el chi-cuadrado residual se vuelve insignificante (es decir, hasta que el valor  $p$  del chi-cuadrado residual excede el valor  $SLENTRY =$ ). En el método BACKWARD, la opción STOPRES elimina las variables del modelo de una en una hasta que el chi-cuadrado residual se vuelve significativo (es decir, hasta que el valor  $p$  del chi-cuadrado residual se vuelve menor que el valor  $SLSTAY =$ ). La opción STOPRES no tiene ningún efecto para el método STEPWISE.

Optimization Options Se permiten cuatro criterios de convergencia:  $ABSFCNV =$ ,  $FCONV =$ ,  $GCONV =$  y  $XCONV =$ . Si se especifica más de un criterio de convergencia, la optimización finaliza en cuanto se cumple uno de los criterios. Si no se especifica ninguno de los criterios, el valor predeterminado es  $GCONV = 1E-8$ .

$ABSFCNV=value$ . Especifica el criterio de convergencia de la función absoluta. La terminación requiere un pequeño cambio en la función objetivo (función logarítmica de probabilidad parcial) en iteraciones posteriores, donde  $lk$  es el valor de la función objetivo en la iteración  $k$ .

CONVERGELIKE=value Es la misma como se especifica ABSFCNV=.

CONVERGEPARM=value Es la misma como se especifica XCONV= option.

FCNV=value. Especifica el criterio de convergencia de funciones relativas. La terminación requiere un pequeño cambio relativo en la función objetivo (función logarítmica de verosimilitud parcial) en iteraciones posteriores, donde  $l_k$  es el valor de la función objetivo en la iteración  $k$ .

GCONV=value. Especifica el criterio de convergencia de gradiente relativo. La terminación requiere que la reducción de la función de predicción normalizada sea pequeña, donde  $l_k$  es la probabilidad logarítmica parcial,  $g_k$  es el vector de gradiente (primeras derivadas parciales de la probabilidad logarítmica parcial) y  $H_k$  es la matriz hessiana negativa (segundas derivadas parciales de la probabilidad logarítmica parcial probabilidad), todo en la iteración  $k$ .

MAXITER=n Especifica el número máximo de iteraciones permitidas. El valor predeterminado para  $n$  es 25. Si no se logra la convergencia en  $n$  iteraciones, la salida mostrada y todos los conjuntos de datos creados por PROC PHREG contienen resultados que se basan en la última iteración de máxima verosimilitud.

RIDGING=ABSOLUTE | RELATIVE | NONE Especifica la técnica para mejorar la probabilidad logarítmica cuando su valor es peor que el del paso anterior. Para RIDGING = ABSOLUTE, los elementos diagonales del Hessian negativo (esperado) se inflan sumando el valor de la cresta. Para RIDGING = RELATIVE, los elementos diagonales se inflan con un factor igual a 1 más el valor de la cresta. Para RIDGING = NONE, se utiliza el método de búsqueda de línea burdo de dar medio paso en lugar de surcar.

SINGULAR=value Especifica el criterio de singularidad para determinar las dependencias lineales en el conjunto de variables explicativas. El valor predeterminado es 10-12.

XCONV=value Especifica el criterio de convergencia de parámetros relativos. La terminación requiere un pequeño cambio de parámetro relativo en iteraciones posteriores, donde  $\hat{\theta}_i$  es la estimación del  $i$ -ésimo parámetro en la iteración  $k$ .

## TEST Statement

< label: > TEST equation1 < , ... , equationk >< /option > ;

La declaración TEST prueba las hipótesis lineales sobre los coeficientes de regresión. PROC PHREG realiza una prueba de Wald para la hipótesis conjunta especificada en una sola declaración TEST. Cada ecuación especifica una hipótesis lineal; las ecuaciones múltiples (filas de la hipótesis conjunta) están separadas por comas. La etiqueta (un nombre SAS válido), se utiliza para identificar la salida resultante y siempre debe incluirse. Puede enviar varias declaraciones de TEST. La forma de una ecuación es la siguiente:

term < term ... > < = < term < term ... >>>

Aquí term, es una variable o una constante o una constante multiplicada por una variable. Es cualquier variable explicativa en la declaración MODEL. Cuando no aparece ningún signo igual, la expresión se establece en 0. El siguiente código ilustra los posibles usos de la instrucción TEST:

```
Proc phreg;
```

```
Model time= a1 a2 a3 a4;
```

```
Test1: Test a1, a2;
```

```
Test2: Test a1=0, a2=0;
```

```
Test3: Test a1=a2=a3;
```

```
Test4: Test a1=a2, a2=a3;
```

```
Run;
```

La primera y segunda declaración TEST son equivalentes, al igual que las declaraciones tercera y cuarta de TEST. Las siguientes opciones se pueden especificar en la instrucción TEST después de una barra (/).

**AVERAGE** Permite evaluar el efecto promedio de las variables en la declaración TEST. Una estimación general del efecto del tratamiento se calcula como un promedio ponderado de los coeficientes del tratamiento, como se ilustra en el siguiente código:

```
TREATMENT: test trt1, trt2, trt3, trt4 / average;
```

E Especifica que se impriman los coeficientes lineales y las constantes. Cuando se especifica la opción AVERAGE junto con la opción E, los pesos óptimos del efecto promedio también se imprimen en las mismas tablas que los coeficientes.

**PRINT** Muestra cálculos intermedios.

**WEIGHT Statement**

**WEIGHT** variable < /option >; en la instrucción **WEIGHT** se identifica la variable en el conjunto de datos de entrada que contiene los pesos de los casos. Cuando aparece la declaración de **WEIGHT**, cada observación en el conjunto de datos de entrada se pondera por el valor de la variable **WEIGHT**. Los valores de **WEIGHT** pueden ser no integrales y no están truncados. Las observaciones con valores negativos, cero o faltantes para la variable **WEIGHT** no se utilizan en el ajuste del modelo. Cuando no se especifica la instrucción **WEIGHT**, a cada observación se le asigna un peso de 1. La instrucción **WEIGHT** está disponible solo para **TIES = BRESLOW** y **TIES = EFRON**.

La siguiente opción se puede especificar en la instrucción **WEIGHT** después de una barra (/):

**NORMALIZE**

**NORM** Hace que los pesos especificados por la variable **WEIGHT** se normalicen para que sumen el tamaño real de la muestra. Con esta opción, la matriz de covarianza estimada de los estimadores de parámetros es invariante a la escala de la variable **WEIGHT**.

**BASELINE Statement**

```
BASELINE < OUT= SAS-data-set >< COVARIATES= SAS-data-set > < keyword=name ... keyword=name > < /options > ;
```

La declaración BASELINE crea un nuevo conjunto de datos SAS que contiene las estimaciones de la función de superviviente en los momentos del evento de cada estrato para cada patrón de valores de variables explicativas (x), creados en el conjunto de datos COVARIATES =. De forma predeterminada, el conjunto de datos también contiene las estimaciones de la función de superviviente correspondientes a las medias de las variables explicativas para cada estrato. Si solo desea estas estimaciones, puede omitir la opción COVARIATES =. No se crea ningún conjunto de datos BASELINE si el modelo contiene una variable dependiente del tiempo, definida por medio de una declaración de programación.

La siguiente lista explica las especificaciones en la declaración BASELINE.

OUT=SAS-data-set nombra el conjunto de datos de salida de BASELINE. Si se omite la opción OUT =, el conjunto de datos se crea y se le da un nombre predeterminado usando la convención DATAn.

COVARIATES=SAS-data-set nombra el conjunto de datos que contiene el conjunto de valores de variables explicativas para las que se estiman las funciones de superviviente. Debe haber una variable correspondiente en el conjunto de datos COVARIATES = para cada variable explicativa en el modelo final.

Las siguientes opciones aparecen en BASELINE después de un slash (/).

ALPHA=value especifica el nivel de significancia del intervalo de confianza de la función de supervivencia. El valor puede estar entre 0 y 1. El defecto es 0,05, que resulta en un intervalo de confianza de 95%.

CLTYPE=method Especifica el método utilizado para calcular los límites de confianza para  $S(t, z)$ , la función de superviviente para un sujeto con un vector covariable fijo  $z$  en el tiempo del evento  $t$ . La opción CLTYPE = puede tomar los siguientes valores:

LOG especifica que los límites de confianza para  $\log(S(t, z))$  deben calcularse utilizando la aproximación de la teoría normal. Los límites de confianza para  $S(t, z)$  se obtienen transformando hacia atrás los límites de confianza para  $\log(S(t, z))$ . El valor predeterminado es CLTYPE = LOG.

LOGLOG especifica que los límites de confianza para el log (-log (S (t, z))) deben calcularse utilizando una aproximación teórica normal. Los límites de confianza para S (t, z) se obtienen transformando de nuevo los límites de confianza para log (-log (S (t, z))).

NORMAL especifica que los límites de confianza para S (t, z) deben calcularse directamente utilizando la aproximación de la teoría normal.

METHOD=method especifica el método utilizado para calcular las estimaciones de la función de superviviente. Los dos métodos disponibles son:

CH | EMP | NELSON calcula la estimación de la función de riesgo acumulativo de Nelson (empírica) de la función de superviviente; es decir, la función de supervivencia se estima exponencializando la función de riesgo acumulativo empírico negativo.

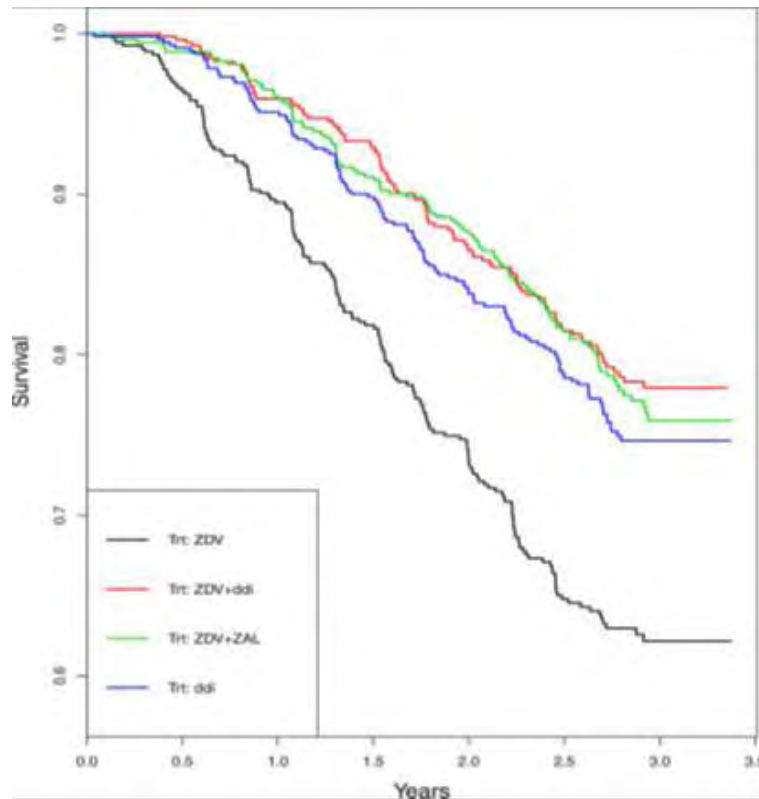
PL calcula la estimación del límite de producto de la función de superviviente. El valor predeterminado es METHOD = PL.

NOMEAN excluye las estimaciones de la función de superviviente correspondientes a las medias muestrales de las variables explicativas.

#### 2.2.2.4. Posibles causas de violación de las suposiciones de riesgo proporcional

Existen dos escenarios prácticos que pueden conducir a violaciones de la suposición de riesgo proporcional (PH) (Sheng y Ghosh, 2020):

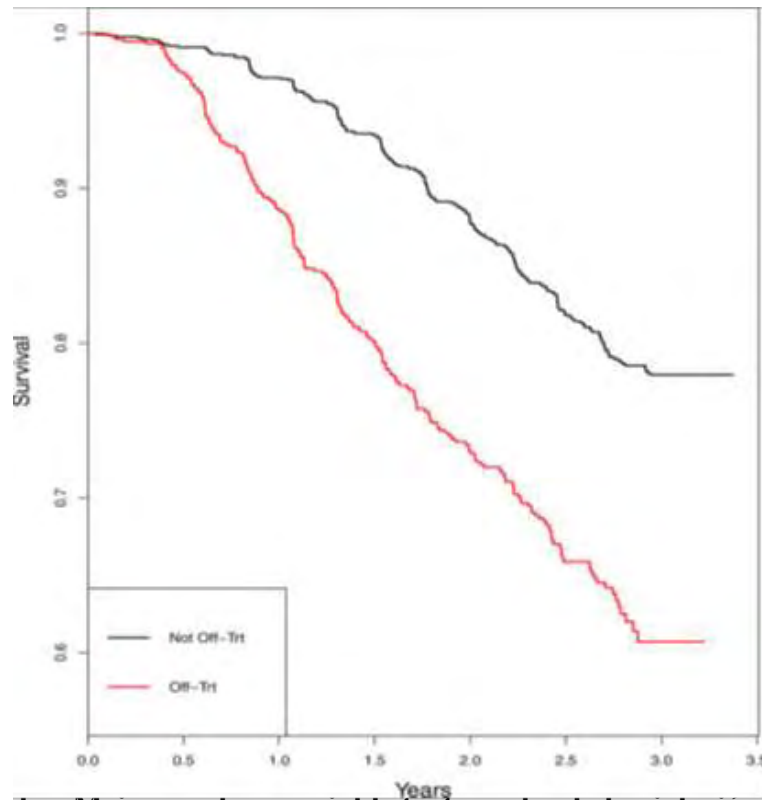
1. Cruce de curvas de supervivencia: cuando las covariables representan grupos de tratamiento (por ejemplo, valores de  $x = 0, 1$ ), se pueden calcular las estimaciones no paramétricas de las funciones de supervivencia de Kaplan-Meier (KM) y comparar los dos grupos utilizando las gráficas (Figura 05). Si los KM se cruzan, entonces es probable que se viole la suposición de PH, ya que, si la suposición de PH es cierta, entonces  $S_1(t) = S_0(t) \gamma_0$ , para algunos  $\gamma_0 > 0$ . Muchos otros modelos, por ejemplo, el modelo de tiempo de falla acelerado (AFT) o el modelo de probabilidades proporcionales (PO), tampoco permiten el cruce de funciones de supervivencia. En la Figura 05, al comparar cuatro tratamientos diferentes, se observa que la curva roja cruza a la curva verde dos veces y, por lo tanto, probablemente conduce a una violación de la suposición de PH.



**Figura 05. Curvas Kaplan-Meier para cuatro tratamientos donde el cruce de curvas de supervivencia viola los supuestos de PH.** Fuente. Sheng y Ghosh (2020).

2. Efectos retardados del tratamiento: cuando dos grupos de tratamiento no muestran diferencias significativas en las tasas de supervivencia en la fase inicial del ensayo, pero luego se observa un efecto retardado entre las curvas. En otras palabras, para algún intervalo de tiempo inicial  $t \in (0, t_0]$ , podemos observar  $S_1(t) \approx S_0(t)$ , pero luego podemos observar  $S_1(t) > S_0(t)$  para  $t > t_0$ , violando así la suposición de PH de que  $\log H_1(t) - \log H_0(t) = \log \gamma_0$ .

En tales casos, aunque las curvas de supervivencia no se cruzan, la suposición de PH aún puede violarse, ya que la separación del logaritmo del riesgo acumulativo puede no ser constante en todos los puntos de tiempo. Nuevamente en la Figura 05, vemos que este es el caso cuando se comparan las curvas azul y verde; una característica más prominente de este efecto retardado se puede ver en la Figura 06.



**Figura 06. Curvas Kaplan-Meier por la covariable tratam, donde la violación de supuestos de PH es representada por retardos en el efecto de los tratamientos.** Fuente. Sheng y Ghosh (2020).

#### 2.2.2.5. Pruebas para los supuestos de riesgo no proporcional

En modelos de regresión de riesgo proporcional, es importante evaluar si los datos en diferentes grupos de tratamiento o con ciertas covariables, cumplen los criterios de suposición de riesgo proporcional (PH). Básicamente, el modelo de regresión de riesgos proporcionales (PH) tiene dos supuestos, que cuando se satisfacen permiten su basamento en las inferencias estadísticas y las predicciones que arroja el modelo. El primer supuesto es que la independencia temporal de las covariables en la función de riesgo, o el cociente de la función de riesgo para dos individuos con diferentes covariables de regresión, no varía con el tiempo, lo que también se conoce como el supuesto de PH, en datos de análisis de supervivencia continuos o discretos. El segundo supuesto es que la relación entre el riesgo acumulado logarítmico y una covariable, es lineal.



Existen 3 formas para evaluar supuestos de riesgo proporcional. La primera, es un análisis gráfico que incluye curvas de Kaplan Meier con y sin una transformada  $-\log-\log$ . La segunda, es el uso de covariables dependientes del tiempo, y la tercera, es la bondad de ajuste, una especie de prueba estadística.

### 2.2.2.5.1. Análisis Gráfico

En este manual, se utilizarán las curvas de sobrevivencia de  $\log(-\log(S(t)))$  vs  $\log t$  para evaluar el supuesto de riesgos proporcionales, indicando el cumplimiento de la suposición de PH si las dos curvas muestran la misma tendencia sin cruce entre ellas. Además, se puede aplicar en variables continuas, binarias y categóricas. Para las covariables binarias y categóricas, se trazan las curvas dentro de cada subgrupo y se verifican sus tendencias, originando una imagen aproximada de la relación PH entre cada grupo. Para las covariables continuas, las prácticas similares también son útiles después de la discretización de las covariables continuas, teniendo cuidado con los umbrales. Son curvas que se definen con Proc Lifetest y la opción de graficación Plot= (LLS). Para las covariables binarias, con datos evaluados con valores 0 – 1 (datos de dos muestras), cuando los datos satisfacen la suposición PH, las siguientes expresiones se consideran verdaderas.

Las gráficas de las curvas de Kaplan-Meier demuestran la diferencia entre dos tratamientos. Cuando existe una buena separación entre las curvas indica una diferencia significativa en el tiempo de recurrencia para los diferentes niveles de la covariable. Cuando existe poca separación o solapamiento entre las curvas indica que la covariable no está asociada con el tiempo de recurrencia del evento. En las siguientes figuras no se observa solapamiento y existe distanciamiento entre la línea roja y la línea azul (Figura 07), lo cual indica que ellas son significativamente diferentes. En la Figura 08 se observa lo contrario, ya que existe solapamiento entre las curvas roja y azul, no están distanciadadas, lo cual indica que no son significativamente diferentes. Además, las gráficas de  $\text{Log}(-\text{Log}(S(t)))$  de los dos grupos tienen líneas consideradas paralelas y no existe un problema significativo con la suposición para los modelos. La Figura 07 y Figura 08, definen dos tipos de curvas, suposición de PH no violada y violada, respectivamente:

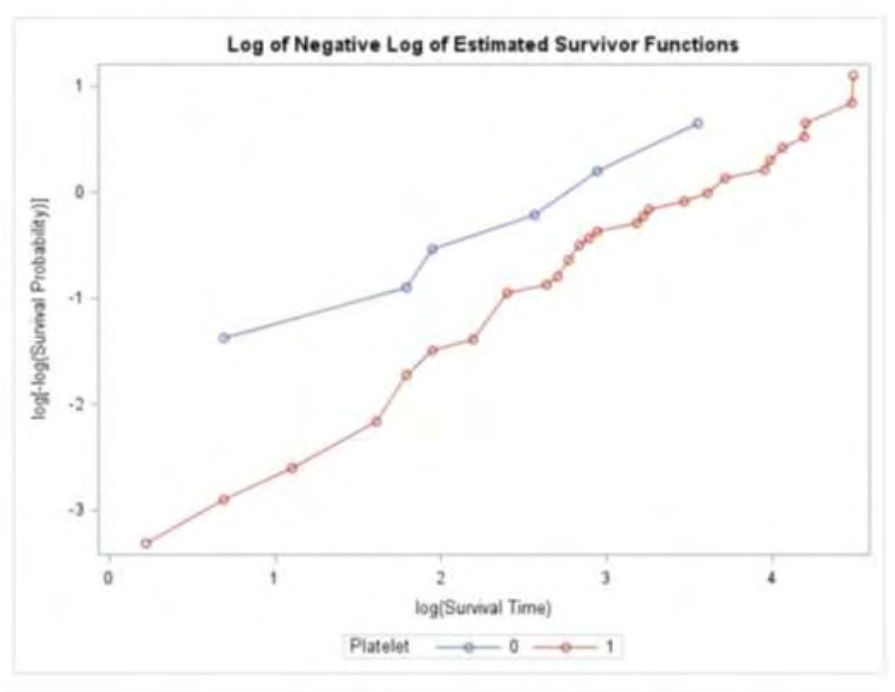


Figura 07. Gráfica de Log (-Log (S(t))) con suposición PH no violada. Fuente. Ying (2018).

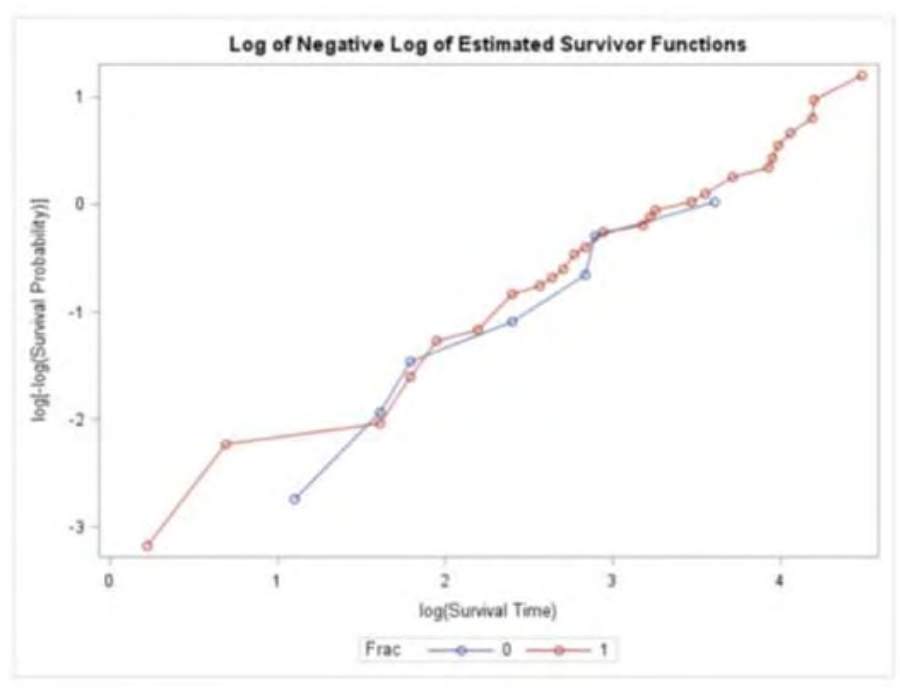


Figura 08. Gráfica de Log (-Log (S(t))) con suposición PH violada. Fuente. Ying. (2018).

### 2.2.2.5.2. Uso de modelos con variables explicatorias tiempo dependientes

El segundo método es establecer en el modelo de regresión de Cox, variables tiempo dependiente, para tasar o gravar la suposición PH. Agregando dentro del modelo una interacción con el tiempo, es decir,  $X \cdot \log(t)$ , se prueba la significancia. La hipótesis nula es que los datos de sobrevivencia satisfacen la suposición PH. Si el valor de  $p$  es menor que el nivel de significación (usualmente  $\alpha=0,05$ ) la suposición de PH fracasa. Una forma común de representar la interacción entre dos variables en un modelo de regresión lineal es incluir una nueva variable que es el producto de las dos variables en cuestión. Para representar la interacción entre una covariable  $x$  y tiempo en un modelo de Cox, podemos escribir

$$\text{Log } h(t) = \alpha(t) + \beta_1 x + \beta_2 X t.$$

Factorizando la  $X$ , podemos reescribir esto como:

$$\text{Log } h(t) = \alpha(t) + (\beta_1 + \beta_2 t) X$$

En esta ecuación, el efecto de  $X$  es  $\beta_1 + \beta_2 t$ . Si  $\beta_2$  es positivo, entonces el efecto de  $X$  aumenta linealmente con el tiempo; si es negativo, el efecto disminuye linealmente con el tiempo.  $\beta_1$  se puede interpretar como el efecto de  $X$  en el tiempo 0, el origen del proceso.

Al modelar una regresión de Cox, para agregar una variable tiempo dependiente, se hace de la manera siguiente:

```
Proc PHREG Data= Ratas;
```

```
Model Tiempo*Status (0) = Edad EdadT;
```

```
EdadT = Edad*Tiempo;
```

```
Run;
```

### 2.2.2.5.3. Prueba del supuesto de proporcionalidad con la declaración ASSESS

Se han definido diferentes tipos de residuos para modelos de regresión de Cox. Uno de ellos, residuos de martingala, se puede utilizar para probar la no proporcionalidad. Este método ha sido incorporado en la declaración ASSESS en PROC PHREG. Desafortunadamente, esta declaración no funcionará si el modelo contiene covariables dependientes del tiempo. A continuación, se explica cómo utilizar

la declaración ASSESS:

Ods Graphics On;

Proc Phreg Data = Flor;

Model Tiempo \* Floración (0) = Edad Variedad Suelo / Ties = Efron;

Assess Ph / Resample;

Run;

Ods Graphics Off;

En la declaración ASSESS, la opción PH solicita una evaluación del supuesto de riesgos proporcionales. Para cada covariable, esta declaración produce una representación gráfica del proceso de puntuación empírica, basado en los residuos de martingala. Se muestra un gráfico para cada variable. En cada gráfico, la línea continua es el proceso de puntuación empírico observado. Las líneas punteadas son procesos de puntuación empíricos basados en 20 simulaciones aleatorias que incorporan el supuesto de riesgos proporcionales. Si el proceso observado se desvía notablemente de los procesos simulados, es una prueba en contra del supuesto de riesgos proporcionales.

En la esquina inferior derecha del gráfico, aparece una información de la evaluación cuantitativa en forma de valor p, produciendo un tipo de prueba Kolmogorov-Smirnov. El valor p es producido por la opción RESAMPLE, que genera los 1,000 caminos simulados. Esa opción también produce una tabla resumen con los resultados para cada covariable. Sin embargo, esta prueba no dice nada sobre la naturaleza de la violación. Si solo se requiere la tabla, pero no los gráficos, simplemente no se ejecutan las declaraciones ODS.

#### **2.2.2.5.4. Prueba del supuesto de proporcionalidad con residuales de Schoenfeld**

También se realizan pruebas de riesgos no proporcionales con los residuales de Schoenfeld. A diferencia de la sentencia ASSESS, este método funciona incluso, si tienen covariables cambiantes en el tiempo, definidas por el método del proceso de conteo (pero no con el método de declaraciones de programación). Los residuos de Schoenfeld tienen una propiedad inusual: en lugar de una única residual

para cada individuo, hay un residual separado para cada covariable para cada individuo. No se definen para observaciones censuradas (faltan en el conjunto de datos de salida).

El uso principal de estos residuos es detectar posibles desviaciones del supuesto de riesgos proporcionales. Si se satisface la suposición, los residuos de Schoenfeld deben ser independientes del tiempo, no estar correlacionados. Probando la suposición de riesgos proporcionales es equivalente, por lo tanto, a probar si los residuos de Schoenfeld están correlacionados con el tiempo o con alguna función del tiempo. Existen 3 pasos para implementar la prueba de hipótesis. El primero, es correr PROC PHREG para calcular los residuales de Schoenfeld. El segundo paso, es correr PROC RANK para resortear el tiempo de sobrevivencia de la variable y obtener los rangos por tiempo de sobrevivencia. El tercer paso, es correr un PROC CORR para probar la correlacion entre los tiempos de sobrevivencia rankeados y los residuales de Schoenfeld.

Además, los residuales Schoenfeld están basados en el principio que, para una covariable regresora, la suposición restrictiva  $\beta(t_k) = \beta$  para todo valor de tiempo  $t_k$ . Esto implica que una gráfica de  $\beta(t_k)$  versus tiempo tendrá inclinación cero. La hipótesis nula es teniendo inclinación cero, la cual es equivalente a probar que el log de la función de razón de riesgo (Hr) es constante en el tiempo.

Se puede utilizar las siguientes declaraciones:

```
/*Residuales Schoenfeld*/
```

```
Proc Phreg Data = Flor;
```

```
Model Tiempo * Floración (0) = Edad Variedad Suelo / Ties = Efron RL;
```

```
Output Out= Resrank Resch=Schedad Schvariedad Schsuelo;
```

```
Run;
```

```
Data C;
```

```
Set Resrank;
```

```
Tiempo1=Log (Tiempo);
```

```
Tiempo2=Tiempo**2;
```

```
Proc Corr;

Var Tiempo Tiempo1 Tiempo2 Schedad Schvariedad Schsuelo;

Run;

/*Generación de gráfica de residuales Schoenfeld para variedad como función del tiempo*/

Proc Gplot Data = Resrank;

Symbol1 V=Dot C=Black Width=1 I=Sm80s;

Plot Schvariedad *Días/Haxis=Axis1 Vaxis=Axis2;

Axis1 Label=('Días');

Axis2 Label = (A=90 'Residuales Schoenfeld Para Variedad);

Run;
```

Hay que tener en cuenta que la palabra clave RESSCH va seguida arbitrariamente de los nombres de las tres variables elegidas correspondientes a las tres covariables del modelo. Estas variables contienen los residuos de Schoenfeld en el conjunto de datos de salida b.

#### **2.2.2.5.4.1. Jerarquización de los residuales de Schoenfeld por el procedimiento RANK.**

El procedimiento RANK origina los rangos para una o más variables numéricas de un juego de datos. Esta salida muestra los resultados de jerarquía de los valores de una variable, es decir, se indica ubicación del orden de finalización del resultado de cada variable, por ejemplo, en un ensayo de germinación el tratamiento con el más bajo número de días de germinación finaliza en primer lugar. Un procedimiento simple puede ser el siguiente:

```
Proc Rank Data=B Out=Resrank Ties=Means;

Var Tiempo;

Ranks Tiemporank;

Run;
```

```
Proc Print Data= Resrank;
```

```
Run;
```

Sin embargo, si se requiere de procedimientos más completos, con el procedimiento se indica como PROC RANK puede:

- a. Revertir el orden de jerarquía donde el valor más alto pueda recibir el primer rango, el siguiente valor más alto recibir el rango 2 y así sucesivamente;
- b. Por defecto, calcula el rango en orden ascendente;
- c. Rango de observaciones separadamente por valores de múltiples variables;
- d. Rango de observaciones por grupos, utilizando el comando BY;
- e. Manejar valores enlazados o encadenados (TIES).

```
Proc Rank Data=B Out= Resrank Group= Normal= Descending Ties=low;
```

```
By <Descending> Variable-1 <...><Descending> Variable-N <Notsorted>;
```

```
Var Data-Set-Variables(S);
```

```
Ranks New-Variables(S);
```

```
Run;
```

```
Proc Print Data= Resrank;
```

```
Run;
```

Con el programa anterior, con DATA=B se especifica el nombre del juego inicial de datos y con OUT=Resrank se crea una salida de resultados; GROUP= significa la partición de las observaciones en grupos definidos; NORMAL= computa los registros con calidad normal de los rangos, la variable resultante aparece normalmente distribuida y se pueden utilizar las pruebas BLOM | TUKEY | VW, donde VW es la prueba de van der Waerden y se puede utilizar para una prueba no paramétrica. La opción DESCENDING invierte la dirección de los rangos, el valor más alto recibe el primer rango y el siguiente más largo, recibe el segundo rango y así sucesivamente. De otra manera, los valores son ran-

quedados del más pequeño al mas alto valor. La opción NOTSORTED especifica que las observaciones no son necesariamente sorteadas en orden alfabetico o numérico, sino pudiera ser en orden cronológico y el proceimiento no utiliza índices. La opción TIES= especifica los rangos para valores enlazados y se pueden utilizar las opciones HIGH | LOW | MEAN, donde HIGH asigna los mas altos de los rangos correspondientes, LOW los rangos mas pequeños y MEAN la media de los rangos correspondientes.

El procedimiento RANK crea un juego de datos que contiene el rango o rangos de los registros, pero, no crea salidas impresas, para ello se requiere usar PROC PRINT, PROC REPORT, o cualquier otro procedimiento para imprimir la salida de la información.

El siguiente paso es modificar el conjunto de datos para incorporar dos funciones del tiempo, el LOG y el cuadrado, seguido de PROC CORR para calcular las correlaciones de los residuos con el tiempo y sus dos funciones. También podrían considerarse otras funciones del tiempo.

```
Data C;
```

```
Set Resrank;
```

```
Tiempo1=Log (Tiempo);
```

```
Tiempo2=Tiempo**2;
```

```
Proc Corr;
```

```
Var Tiempo Tiempo1 Tiempo2 Schedad Schvariedad Schsuelo;
```

```
Run;
```

También de esta forma:

```
Data C;
```

```
Set Resrank;
```

```
Tiempo1=Log (Tiempo);
```

```
Tiempo2=Tiempo**2;
```

```
Proc Corr;
```



```
Var Tiempo Tiempo1 Tiempo2 Schedad Schvariedad Schsuelo;
```

```
With Tiemporank;
```

```
Run;
```

### 2.2.2.6. Procedimiento con Proc Format para transformar una variable continua en variable categórica.

Con el siguiente ejemplo se evalúa el número de nódulos en las raíces de una especie leguminosa en vivero y su relación con la sobrevivencia y el crecimiento de las plantas.

```
PROC FORMAT;
```

```
VALUE nodulos
```

```
1-3 = 'N1'
```

```
4-9 = 'N2'
```

```
10 - alto = 'N3';
```

```
RUN;
```

```
Proc Phreg Data = nodul;
```

```
Class nodulos (REF = 'N1') / PARAM = REF;
```

```
Model días*status (0) = nodulos;
```

```
ASSESS PH / RESAMPLE SEED = 12345;
```

```
FORMAT nodulos nodulos.;
```

```
RUN;
```

Al usar PROC FORMAT, se divide la cantidad de nódulos en tres categorías: N1, N2 y N3. Luego en PROC PHREG, se agrega una declaración CLASS para categorizar la cantidad de nódulos. Se puede indicar el nivel de referencia entre paréntesis (REF='N1') y la opción PARAM=REF para especificar la razón de riesgo utilizando la referencia codificada, es decir, el riesgo de N2 y N3 se comparará a N1 en las razones de riesgo. Se remueve VAR (nódulos) de la declaración ASSESS ya que no se chequea

la suposición de linealidad si los nodulos es una covariable categórica en lugar de una covariable continua. Sin embargo, se puede chequear la suposición proporcional para covariables categóricas.

### **2.2.3. MÉTODO 3. Análisis de sobrevivencia con modelos de regresión paramétrica utilizando PROC LIFEREG**

El procedimiento LIFEREG ajusta los modelos de tiempo de falla acelerado a los datos de tiempo de falla o evento, incluidos los datos censurados por la izquierda, censurados por la derecha o censurados por intervalos. Los modelos para la variable de respuesta consisten en un efecto lineal que se compone de las covariables y un término de perturbación aleatorio. Las opciones de distribución para el término de perturbación aleatoria incluyen las distribuciones de valor extremo, normal, logística, exponencial, Weibull, lognormal, log-logístico y gamma de tres parámetros (Gordon, 2020).

#### **2.2.3.1. Sintaxis de los procedimientos**

Las siguientes opciones están disponibles para PROC LIFEREG.

PROC LIFEREG < options >;

BY variables;

CLASS variables;

INSET <keyword-list> < / options >;

MODEL response=<effects> < / options >;

OUTPUT < OUT=SAS-data-set > Keyword=name <. keyword=name > < options >;

PROBPLOT < / options >;

WEIGHT variable;

La declaración MODEL es necesaria y declara las variables utilizadas en la regresión del modelo, así como la distribución utilizada para el componente de error del modelo. Sólo se puede utilizar una única sentencia MODEL. En caso de existir varias sentencias MODEL, solo se utiliza la última. Los efectos principales y los términos de interacción se pueden especificar en esta declaración, de forma similar al procedimiento GLM. Los valores iniciales se pueden especificar en la instrucción MODEL o en un

conjunto de datos `INEST =`. Si no se especifican valores iniciales, las estimaciones iniciales se obtienen mediante mínimos cuadrados ordinarios. La declaración `CLASS` determina las variables explicativas categóricas. La instrucción `WEIGHT` identifica una variable con valores que se utilizan para ponderar las observaciones. Las observaciones con ponderaciones cero o negativas no se utilizan para ajustar el modelo, aunque se pueden calcular los valores predichos para ellas. La declaración `OUTPUT` crea un conjunto de datos de salida que contiene valores y residuales predichos.

### 2.2.3.2. Especificación de las opciones para PROC LIFEREG

`PROC LIFEREG < options >`; Puede especificar las siguientes opciones en la sentencia `PROC LIFEREG`:

`COVOUT` Origina la matriz de covarianza estimada en el conjunto de datos `OUTEST =` si se logra la convergencia.

`DATA=SAS-data-set` Especifica el conjunto de datos de entrada

`GOUT=graphics-catalog` Especifica un catalogo de gráficas para guardar los gráficos de salida.

`INEST= SAS-data-set` Especifica un conjunto de datos de entrada que contiene estimaciones iniciales para todos los parámetros del modelo.

`NAMELEN=n` Especifica la longitud de los nombres de efectos en las tablas y los conjuntos de datos de salida en `n` caracteres, donde `n` es un valor entre 20 y 200. La longitud predeterminada es de 20 caracteres.

`NOPRINT` Suprime la visualización de la salida. Esta opción desactiva temporalmente el sistema de entrega de salida (ODS).

`ORDER=DATA | FORMATTED | FREQ | INTERNAL` Especifica el orden de clasificación para los niveles de las variables de clasificación, especificadas en la declaración `CLASS`. Este orden determina qué parámetros del modelo corresponden a cada nivel de los datos. A continuación, se ilustra cómo `PROC LIFEREG` interpreta los valores de la opción `ORDER =`.

ORDER= niveles sorteados en By

DATA orden de aparición en input

FORMATTED valores formateados

FREQ cuentas decrecientes; niveles con la mayoría de observaciones vienen primero en el orden

INTERNAL valor no formateado

Por defecto, ORDER = FORMATTED. Para FORMATTED e INTERNAL, el orden de clasificación depende de la máquina.

OUTEST=SAS-data-set declara un conjunto de datos de salida con las estimaciones de los parámetros, la probabilidad de registro maximizada y, si se especifica la opción COVOUT, la matriz de covarianza estimada.

XDATA= SAS-data-set detalla un conjunto de datos de entrada que contiene valores para todas las variables independientes en la declaración MODEL y variables en la declaración CLASS para el trazado de probabilidad. Si hay covariables especificadas en una instrucción MODEL y se solicita una gráfica de probabilidad, con una instrucción PROBPLOT, se declaran valores fijos para los efectos en la instrucción MODEL con el conjunto de datos XDATA =.

### 2.2.3.3. Uso básico de PROC LIFEREG.

PROC LIFEREG utiliza una sintaxis muy parecida a la de PROC PHREG. Después de la instrucción PROC LIFEREG, requiere una declaración MODEL. Esta declaración puede tomar una de estas tres formas. La primera, que se puede utilizar si no hay censura a la izquierda o por intervalo, es el mismo que el de PROC PHREG, excepto por la especificación opcional de una función de distribución (D).

MODEL TIMEVAR\*CENSVAR(CENSVALS) = covariates / D = distribution;

Aquí Timevar es la variable de tiempo, Censvar es la variable que declara si el tiempo está censurado a la derecha o no, Censvals es una lista de valores de Censvar que indican que el tiempo está censurado a la derecha y Covariates es la lista de covariables a considerar en el modelo. El tipo de distribu-

ción de la variable de tiempo se especifica después de la barra. Las opciones permitidas son: exponencial, Weibull, gamma, logarítmico normal y logístico. Weibull es el predeterminado. Si desea permitir la censura a la izquierda y a intervalos, cambie la declaración MODEL de la siguiente manera:

MODEL (TIMELO, TIMEHI) = covariates / D = distribution;

Aquí TIMELO y TIMEHI son variables de tiempo. Si falta TIMELO, entonces TIMEHI se considera un tiempo censurado por la izquierda. Si TIMEHI falta, entonces TIMELO se considera un tiempo censurado por la derecha. Si no falta ninguno y  $TIMELO < TIMEHI$ , entonces el tiempo se censura por intervalo en el intervalo (TIMELO, TIMEHI). Si ambos son iguales y no faltan, entonces el tiempo no se censura en su valor común. En estas dos declaraciones, las covariables se pueden omitir. En ese caso, solo se estimará la intersección, el parámetro de escala y, si se elige gamma como distribución, para un parámetro de forma.

PROC LIFEREG no tiene ninguna disposición de especificar un método para ingresar o eliminar covariables. Sin embargo, puede probar varios modelos diferentes en la misma sesión SAS, utilizando varias declaraciones MODEL. Estos modelos pueden diferir tanto en la distribución como en las covariables utilizadas. Esto permite realizar varias pruebas de significancia "a mano".

#### BY Statement

BY variables; puede especificar una instrucción BY para obtener análisis separados de las observaciones en grupos definidos por las variables BY. Cuando aparece una instrucción BY, el procedimiento espera que el conjunto de datos de entrada se clasifique en el orden de estas variables. Si el conjunto de datos de entrada no está ordenado en orden ascendente, se usa una de las siguientes alternativas:

- a. Ordenar los datos utilizando el procedimiento SORT con una declaración BY.
- b. Especificar la opción de sentencia BY NOTSORTED o DESCENDING en la sentencia BY. La opción NOTSORTED no significa que los datos no están ordenados, sino que los datos están organizados en grupos (de acuerdo con los valores de las variables definidas por BY) y que estos grupos no están necesariamente en orden alfabético o numérico creciente.
- c. Crear un índice en las variables BY utilizando el procedimiento DATA SET.

### CLASS Statement

CLASS variables; Las variables de clasificación en lugar de variables numéricas cuantitativas deben enumerarse en la declaración CLASS. Para cada variable explicativa listada en la declaración CLASS, se generan variables indicadoras para los niveles asumidos por la variable CLASS. Si se utiliza la declaración CLASS, debe aparecer antes de la declaración MODEL.

### INSET Statement

INSET < keyword-list > < / options >; El cuadro o tabla de información resumida producida en los gráficos realizados con la declaración PROBPLOT se llama recuadro. Se utiliza la declaración INSET para personalizar la información y apariencia que se muestra en el recuadro de inserción. Para ello, se especifican las palabras clave correspondientes a la información que se desea mostrar. Por ejemplo, las siguientes declaraciones producen una gráfica de probabilidad con el número total de observaciones, el número de observaciones censuradas por la derecha, el nombre de la distribución y el parámetro de forma de Weibull estimado en el recuadro:

```
Proc lifereg data=epidemic;
```

```
Model life = dose / dist = Weibull;
```

```
Probplot;
```

```
Inset nobs right dist shape;
```

```
run;
```

De forma predeterminada, las entradas insertadas se identifican con etiquetas. Sin embargo, puede proporcionar una etiqueta personalizada especificando la palabra clave para esa entrada seguida del signo igual (=) y la etiqueta entre comillas. Por ejemplo, la siguiente instrucción INSET produce un recuadro que contiene el número de observaciones y el nombre de la distribución, etiquetado como "Tamaño de muestra" y "Distribución" en el recuadro.

```
Inset nobs= 'Tamaño de muestra ' dist='Distribution';
```

Si se especifica una palabra clave que no se aplica al gráfico que está creando, la palabra clave se ignora. Si se especifica más de una instrucción INSET, solo se usa la primera. A continuación, se enumeran las palabras clave disponibles en la declaración INSET para mostrar estadísticas resumidas, parámetros de distribución e información de ajuste de distribución.

CONFIDENCE: coeficiente de confianza para los intervalos de confianza

DIST: nombre de la distribución

INTERVAL: número de observaciones censuradas por intervalo

LEFT: número de observaciones censuradas por la izquierda

NOBS: número de observaciones

NMISS: número de observaciones con valores perdidos

RIGHT: número de observaciones censuradas por la derecha

SCALE: valor del parámetro de escala

SHAPE: valor del parámetro de forma

UNCENSORED: número de observaciones sin censura.

Las siguientes opciones controlan la apariencia del cuadro. Todas las opciones se especifican después de la barra inclinada (/) en la instrucción INSET.

CFILL=color, especifica el color de la caja de llenado.

CFILLH=color, especifica el color del encabezado del cuadro de relleno.

CFRAME=color, especifica el color del marco.

CHEADER=color, especifica el color del texto en el encabezado.

CTEXT=color, especifica el color del texto.

FONT=font, especifica el tipo de *Fuente* para el texto.

HEIGHT=value, especifica la altura del texto.

HEADER='quoted string' especifica texto para el encabezado o el título del cuadro.

NOFRAME omite el marco alrededor de la caja.

POS= value <DATA | PERCENT> Determina la posición del recuadro (Inset). El valor puede ser un punto de la brújula (N, NE, E, SE, S, SW, W, NW) o un par de coordenadas (x, y) entre paréntesis. Las coordenadas se pueden especificar en unidades de porcentaje de pantalla o unidades de datos de eje. El valor predeterminado son las unidades de porcentaje de pantalla.

REFPOINT= name Declara el punto de referencia para un recuadro (Inset) que está posicionado por un par de coordenadas con la opción POS =. Utiliza la opción REFPOINT = junto con las coordenadas POS =. La opción REFPOINT = especifica qué esquina del marco de inserción ha especificado con coordenadas (x, y), y puede tomar el valor de BR (abajo a la derecha), BL (abajo a la izquierda), TR (arriba a la derecha) o TL (arriba a la izquierda). El valor predeterminado es REFPOINT = BL. Si la posición insertada se especifica como un punto de la brújula, la opción REFPOINT = se ignora.

MODEL Statement.

MODEL response<\*censor (list)>=effects < / options >;

MODEL (lower, upper) =effects < / options >;

MODEL events/trials=effects < / options >; Sólo se puede utilizar una sola sentencia MODEL con LIFEREG. Si hay varias sentencias MODEL, solo se utiliza la última. La etiqueta opcional se utiliza para etiquetar las estimaciones del modelo en el conjunto de datos SAS de salida y el conjunto de datos OUTEST =.

La primera sintaxis de MODEL es apropiada para censurar por la derecha. La variable respuesta posiblemente esté censurada por la derecha. Si la variable de respuesta es censurada por la derecha, entonces una segunda variable, denotada como censura, debe aparecer después de la variable de respuesta con una lista de valores entre paréntesis, separados por comas o espacios en blanco, para indicar censura. Es decir, si la variable de censura toma un valor dado en la lista, la respuesta es un valor censurado por la derecha; de lo contrario, es un valor observado.



La segunda sintaxis de MODEL especifica dos variables, inferior y superior, que contienen valores de los puntos finales del intervalo de censura. Si los dos valores son iguales (y no faltan), se supone que no hay censura y se observa el valor de respuesta real. Si falta el valor inferior, el valor superior se utiliza como valor censurado por la izquierda. Si falta el valor superior, el valor inferior se toma como un valor censurado por la derecha. Si ambos valores están presentes y el valor inferior es menor que el valor superior, se supone que los valores especifican un intervalo de censura. Si el valor inferior es mayor que el valor superior o faltan ambos valores, entonces la observación no se utiliza en el análisis, aunque los valores predichos todavía se pueden obtener si no falta ninguna de las covariables. El Cuadro 04 resume las formas de especificar la censura:

**Cuadro 04. Resumen de las formas específicas para la censura.**

Bajo	Superior	Comparación	Interpretación
No ausente	No ausente	Igual	No censurado
No ausente	No ausente	Bajo < superior	Intervalo censurado
Ausente	No ausente		Valor superior usado como censurado a la izquierda
No ausente	Ausente		Valor bajo usado como censurado a la derecha
No ausente	No ausente	Bajo > superior	Observación no utilizada
Ausente	Ausente		Observación no utilizada

Fuente. SAS (2020).

La tercera sintaxis de MODEL especifica dos variables que contienen datos de recuento para una respuesta binaria. El valor de la primera variable, eventos, es el número de éxitos. El valor de la segunda variable, ensayos, es el número de intentos. Los valores de ambos eventos y (pruebas-eventos) deben ser no negativos y las pruebas deben ser positivas para que la respuesta sea válida. Los valores de las dos variables no necesitan ser enteros y no se modifican para que sean enteros.

Los efectos que siguen al signo igual son las covariables del modelo. Los efectos de orden superior, como interacciones y términos anidados, están permitidos en la lista, de forma similar al procedimiento GLM. Los nombres de variables y las combinaciones de nombres de variables que representan términos de orden superior pueden aparecer en esta lista. Las variables de clase se pueden utilizar como efectos y las variables indicadoras se generan para los niveles de clase. Si no especifica ninguna covariable después del signo igual, se ajusta un modelo de solo intercepción.

Ejemplos de tres declaraciones MODEL válidas son

a: model time\*flag (1,3) =temp;

b: model (start, finish) =;

c: model r/n=dose;

La declaración a de MODEL, indica que la respuesta está contenida en una variable llamada tiempo y que, si la bandera de la variable toma los valores 1 o 3, la observación está censurada por la derecha. La variable explicativa es temp, que podría ser una variable de clase. La declaración b de MODEL, revela que la respuesta se encuentra en el intervalo entre los valores de las variables inicio y finalización y que no existen covariables, excepto por un término de intersección predeterminado. La declaración c de MODEL, revela una respuesta binaria, con la variable r que contiene el número de respuestas y la variable n que contiene el número de intentos. Las siguientes opciones pueden aparecer en la declaración MODEL.

ALPHA=value establece el nivel de significancia de los intervalos de confianza para los parámetros de regresión y las probabilidades de supervivencia estimadas. El valor debe estar entre 0 y 1. Por defecto, ALPHA = 0,05.

CONVERGE=value establece el criterio de convergencia. La convergencia se declara cuando el cambio máximo en las estimaciones de los parámetros entre los pasos de Newton-Raphson es menor que el valor especificado. El cambio es relativo si el parámetro es mayor que 0,01 en valor absoluto; de lo contrario, es un cambio absoluto. Por defecto, CONVERGE = 1E-8.

CONVG=value establece el criterio de convergencia relativa de Hessian. El valor debe estar entre 0 y 1. Después de determinar la convergencia con el cambio en el criterio de parámetro especificado con la opción CONVERGE =, la cantidad se calcula y se compara con el valor, donde g es el vector de gradiente, H es la matriz de Hesse para los parámetros del modelo, y f es la función logarítmica de verosimilitud. Si tc es mayor que el valor, se muestra una advertencia de que se ha excedido el criterio relativo de convergencia de Hesse. Este criterio detecta el caso ocasional en el que se satisface el cambio en el criterio de convergencia de parámetros, pero no se ha alcanzado un máximo en la función loga-

rítmica de verosimilitud. Por defecto, CONVG = 1E-4.

CORRB Produce la matriz de correlacion estimada de los parámetros.

COVB Produce la matriz de covarianza estimada de los parámetros estimados.

DISTRIBUTION=distribution-type

DIST=distribution-type

D=distribution-type especifica el tipo de distribución asumido para el tiempo de falla. Por defecto, PROC LIFEREG ajusta una distribución de valor extremo de tipo 1 al registro (log) de la respuesta. Esto es equivalente a ajustar la distribución de Weibull, ya que el parámetro de escala para la distribución de valor extremo está relacionado con un parámetro de forma de Weibull y la intersección está relacionada con el parámetro de escala de Weibull en este caso. Cuando se especifica la opción NOLOG, PROC LIFEREG modela la respuesta no transformada con una distribución de valor extremo de tipo 1 como predeterminada. Los siguientes son valores válidos para el tipo de distribución:

EXPONENTIAL La distribucion exponencial, la cual es tratada como una distribucion Weibull restringida.

GAMMA Una distribucion Gamma generalizada. La distribucion standard Gamma de dos parámetros no esta disponible en PROC LIFEREG.

LLOGISTIC Una distribucion loglogística.

LNORMAL Una distribucion lognormal

LOGISTIC Una distribucion logística (equivalente a LLOGISTIC cuando se especifica la opcion NOLOG)

NORMAL Distribucion normal (equivalente a LNORMAL cuando se especifica la opcion NOLOG).

WEIBULL Distribucion Weibull. Si se especifica NOLOG, ajusta una distribucion de valor extremo tipo 1 a datos brutos no transformados.

De manera predeterminada, PROC LIFEREG transforma la respuesta con el logaritmo natural antes de ajustar el modelo especificado con las opciones Gamma, Llogistic, Lnormal O Weibull. Puede suprimir la transformación de registro con la opción NOLOG. La siguiente tabla resume las distribuciones resultantes cuando las opciones de distribución se utilizan en combinación con la opción NOLOG (Cuadro 05).

**Cuadro 05. Distribuciones resultantes cuando las opciones de distribución anteriores se utilizan en combinación con la opción NOLOG.**

Distribution=	Nolog Especificada?	Distribución Resultante
Exponential	No	Exponencial
Exponential	Si	Un Parámetro De Extremo Valor
Gamma	No	Gamma Generalizada
Gamma	Si	Gamma Generalizada Con Respuesta No Transformada Logistic No
Logistic		Logística
Logistic	Si	Logística (Nolog Sin Efecto)
Llogistic	No	Log-Logística
Llogistic	Si	Logística
Lnormal	No	Log-Normal
Lnormal	Si	Normal
Normal	No	Normal
Normal	Si	Normal (Nolog Sin Efecto)
Weibull	No	Weibull
Weibull	Si	Valor Extremo

Fuente. SAS. (2020).

INITIAL=values establece valores iniciales para los parámetros de regresión. Esta opción puede resultar útil en caso de dificultad de convergencia. Los valores especificados se utilizan para inicializar los coeficientes de regresión para las covariables especificadas en la declaración MODEL. El parámetro de intercepción se inicializa con la opción INTERCEPT = y no se incluye aquí. Los valores se asignan a las variables en la instrucción MODEL en el mismo orden en que se enumeran en la instrucción MODEL. Hay que tener en cuenta que una variable de clase requiere valores k-1 cuando la variable de clase toma k niveles diferentes. El orden de los niveles de clase está determinado por la opción ORDER

=. Si no hay un término de intersección, la primera variable de clase requiere k valores iniciales. Si se usa una instrucción BY, todas las variables de clase deben tener el mismo número de niveles en cada grupo BY o no se pueden especificar valores iniciales significativos. La opción INITIAL = se puede especificar de la siguiente manera (Cuadro 06).

**Cuadro 06. Valores iniciales para los parámetros de regresión.**

Tipo de lista	Especificación
Lista separada por espacio en blanco	Initial= 3 4 5
Lista separada por comas	Initial= 3, 4, 5
X to Y	Initial= 3 to 5
X to Y by Z	Initial= 3 to 5 by 1
Combinación de métodos	Initial= 1, 3 to 5, 9

Fuente. SAS (2020).

Por defecto, PROC LIFEREG calcula estimaciones iniciales con mínimos cuadrados ordinarios. La opción INITIAL = se sobrescribe con la opción INEST =.

INTERCEPT=value inicializa el término de intersección a valor. De forma predeterminada, la intersección se inicializa mediante una estimación de mínimos cuadrados ordinarios.

ITPRINT muestra el historial de iteraciones, la evaluación final del gradiente y la evaluación final de la segunda derivada de la matriz negativa, es decir, el negativo del Hessiana.

MAXITER=n establece el número máximo permitido de iteraciones durante la estimación del modelo. De forma predeterminada, MAXITER = 50.

NOINT mantiene fijo el término de intersección. Debido a la transformación logarítmica habitual de la respuesta, el parámetro de intercepción suele ser un parámetro de escala para la respuesta no transformada o un parámetro de ubicación para una respuesta transformada.

NOLOG solicita que no se realice ninguna transformación logarítmica de la variable de respuesta. De manera predeterminada, PROC LIFEREG modela el log de la variable de respuesta para las opciones de distribución Gamma, Llogistic, Lognormal Y Weibull.

NOSCALE mantiene fijo el parámetro de escala. Hay que tener en cuenta que, si la transformación lo-

garítmica se ha aplicado a la respuesta, el efecto del parámetro de escala es una transformación de potencia de la respuesta original. Si no se especifica ningún valor `SCALA =`, el parámetro de escala se fija en el valor 1.

`NOSHAPE1` mantiene fijo el primer parámetro de forma, `SHAPE1`. Si no se especifica ningún valor `SHAPE1 =`, `SHAPE1` se fija en un valor que depende del tipo de Distribución.

`SCALE=value` inicializa el parámetro de escala. Si se especifica la distribución de Weibull, este parámetro de escala es el parámetro de escala de la distribución de valor extremo de tipo 1, no el parámetro de escala de Weibull. Con una transformación logarítmica, el modelo exponencial es el mismo que un modelo Weibull con el parámetro de escala fijo en el valor 1.

`SHAPE1=value` inicializa el primer parámetro de forma. Si la distribución especificada no depende de este parámetro, esta opción no tiene ningún efecto. La única distribución que depende de este parámetro de forma es la distribución gamma generalizada.

`SINGULAR=value` establece la tolerancia para probar la singularidad de la matriz de información y la matriz de productos cruzados para las estimaciones iniciales de mínimos cuadrados. Aproximadamente, la prueba requiere que un pivote sea al menos este valor multiplicado por el valor de la diagonal original. Por defecto, `SINGULAR = 1E-12`.

**OUTPUT Statement.**

`OUTPUT <OUT=SAS-data-set> keyword=name <...keyword=name>;` crea un nuevo conjunto de datos SAS que contiene estadísticas calculadas después de ajustar el modelo. Se requiere al menos una especificación del formulario palabra clave = nombre.

Todas las variables del conjunto de datos original se incluyen en el nuevo conjunto de datos, junto con las variables creadas como opciones para la declaración `OUTPUT`. Estas nuevas variables contienen valores ajustados y cuantiles estimados. Si desea crear un conjunto de datos SAS permanente, debe especificar un nombre de dos niveles. Cada instrucción `OUTPUT` se aplica a la instrucción `MODEL` anterior. Las siguientes especificaciones pueden aparecer en `OUTPUT statement`:

`OUT =` conjunto de datos SAS declara el nuevo conjunto de datos. De forma predeterminada, el proce-

dimiento utiliza la convención DATAn para nombrar el nuevo conjunto de datos.

keyword=name define las estadísticas que se incluirán en el conjunto de datos de salida y da nombres a las nuevas variables. Especifique una palabra clave para cada estadística deseada, un signo igual y la variable que contiene la estadística.

Las palabras claves permitidas y las estadísticas que representan son las siguientes:

**CENSORED** especifica una variable indicadora para señalar la censura. La variable toma el valor 1 si la observación está censurada; de lo contrario, es 0.

**CDF** especifica una variable que contiene las estimaciones de la función de distribución acumulativa evaluada en la respuesta observada.

**CONTROL** especifica una variable en el conjunto de datos de entrada para controlar la estimación de cuantiles. Si la variable especificada tiene el valor 1, se calculan las estimaciones para todos los valores enumerados en la lista **QUANTILE =** para esa observación en el conjunto de datos de entrada; de lo contrario, no se calculan estimaciones. Si no se especifica ninguna variable **CONTROL =**, se estiman todos los cuantiles para todas las observaciones. Si la variable de respuesta en la instrucción **MODEL** es binomial, esta opción no tiene ningún efecto.

**CRESIDUAL | CRES** especifica una variable que contiene los residuos de Cox-Snell  $-\log(S(u_i))$ , donde  $S$  es la función de supervivencia estándar. Si la variable de respuesta en la declaración del modelo correspondiente es binomial, los residuos no se calculan y esta variable contiene valores perdidos.

**SRESIDUAL | SRES** especifica una variable que contiene los residuos estandarizados. Si la variable de respuesta en la declaración del modelo correspondiente es binomial, los residuos no se calculan y esta variable contiene valores perdidos.

**PREDICTED | P** especifica una variable con las estimaciones de cuantiles. Si la variable de respuesta en el enunciado del modelo correspondiente es binomial, entonces esta variable contiene las probabilidades estimadas,  $1-F(-x'b)$ .

QUANTILES | QUANTILE | Q da una lista de valores para los cuales se calculan los cuantiles. Los valores deben estar entre 0 y 1, no inclusivos. Para cada valor, se estima un cuantil correspondiente. Esta opción no se usa si la variable de respuesta en la declaración MODEL correspondiente es binomial. La opción QUANTILES se puede especificar de la siguiente manera (Cuadro 07):

**Cuadro 07. Valores para calcular los Cuantiles.**

Tipo de lista	Especificación
Lista separada por espacio en blanco	.2 .4 .6 .8
Lista separada por comas	.2, .4, .6, .8
X to Y	.2 to .8
X to Y by Z	.2 to .8 by .1
Combinación de métodos	.1, .2 to .8 by .2

*Fuente.* SAS (2020).

Por defecto, QUANTILES = 0,5. Cuando la respuesta no es binomial, se agrega una variable numérica, \_PROB\_, al conjunto de datos de SALIDA (OUTPUT) siempre que se especifica la opción QUANTILES =. La variable \_PROB\_ da el valor de probabilidad para las estimaciones de cuantiles. Estos son los valores tomados de la lista QUANTILES = y se dan como valores entre 0 y 1, no como valores entre 0 y 100.

STD\_ERR | STD especifica una variable con las estimaciones de los errores estándar de los cuantiles estimados o  $x^b$ . Si la respuesta usada en la declaración MODEL es una respuesta binomial, entonces estos son los errores estándar de  $x^b$ . De lo contrario, son los errores estándar de las estimaciones de cuantiles. Estas estimaciones se pueden utilizar para calcular intervalos de confianza para los cuantiles. Sin embargo, si el modelo se ajusta al logaritmo del tiempo del evento, normalmente se pueden calcular mejores intervalos de confianza transformando los intervalos de confianza para la respuesta logarítmica.

XBETA especifica una variable que contiene el valor calculado de  $x^b$ , donde  $x$  es el vector covariable y  $b$  es el vector de estimaciones de parámetros.



### PROBPLOT Statement.

PROBPLOT | PLOT < / options >; Puede usar la declaración PROBPLOT para crear una gráfica de probabilidad a partir de datos de la tabla de vida. Los datos pueden estar sin censura, censurados por la derecha o censurados arbitrariamente. Puede especificar cualquier número de sentencias de PROBPLOT después de una sentencia MODEL. La sintaxis utilizada para la respuesta en la declaración MODEL determina el tipo de censura asumido al crear la gráfica de probabilidad. El modelo que se ajusta a la declaración MODEL se traza junto con los datos. Si hay covariables en el modelo, se establecen en valores constantes especificados en el conjunto de datos XDATA = al crear la gráfica de probabilidad. Si no se especifica ningún conjunto de datos XDATA =, las variables continuas se establecen en sus valores medios generales y las variables categóricas especificadas en la declaración CLASS se establecen en sus niveles más altos. Puede especificar las siguientes opciones para controlar el contenido, el diseño y la apariencia de una gráfica de probabilidad.

ANNOTATE=SAS-data-set

ANNO=SAS-data-set, especifica un conjunto de datos ANNOTATE, que permite agregar características a la gráfica de probabilidad. El conjunto de datos ANNOTATE = que especifica en la declaración PROBPLOT se utiliza para todos los gráficos creados por la declaración.

NOTA: El valor predeterminado para el color, es el primer color en la lista de colores del dispositivo.

### WEIGHT Statement

WEIGHT variable; si se desea utilizar pesos para cada observación en el conjunto de datos de entrada, coloque los pesos en una variable en el conjunto de datos y especifique el nombre en una declaración de WEIGHT. Los valores de la variable WEIGHT pueden ser no integrales y no están truncados. Las observaciones con valores no positivos o faltantes para la variable de ponderación no contribuyen al ajuste del modelo. La variable WEIGHT multiplica la contribución a la probabilidad logarítmica de cada observación.

# CAPÍTULO 3

## Metodologías y procedimientos aplicados al crecimiento de plantas, entomología y control biológico

### 3.1. Aplicación básica del Análisis de Supervivencia en estudios ecológicos

Los datos sobre la ocurrencia y el momento de eventos discretos, por ejemplo, la germinación de esporas, la aparición de enfermedades o la muerte de propágulos, se registran comúnmente en estudios epidemiológicos. Al analizar estos datos de "tiempo transcurrido hasta el evento", el análisis de supervivencia es superior a las técnicas estadísticas convencionales porque se pueden utilizar observaciones censuradas, es decir, casos en los que el evento no ha ocurrido ni ocurre hasta el final del estudio. Para el análisis de supervivencia son fundamentales dos funciones matemáticas, la *función de superviviente*, que describe la probabilidad de que un individuo "sobreviva" a un evento epidemiológico (es decir, que el evento no ocurra) o también que experimente un evento de interés, por ejemplo, germinación, hasta un momento dado, y la *función de riesgo*, que origina el valor instantáneo del riesgo de que el evento ocurra en ese momento, ya que no ha ocurrido previamente. Estas funciones se pueden comparar entre dos o más grupos utilizando estadísticas de chi-cuadrado. Los efectos de las covariables discretas o continuas sobre los tiempos de supervivencia se pueden cuantificar con dos tipos de modelos, el *modelo de tiempo de falla acelerado, ATF* y el *modelo de riesgos proporcionales, PH* (Scherer y Ojiambo, 2004).

Un “*evento*” se define como un cambio cualitativo que ocurre en un momento particular en el tiempo. Además, el término de tiempo hasta el evento también se usa comúnmente en la literatura ecológica. Independientemente de la naturaleza del evento de interés, se utiliza el término análisis de sobrevivencia a lo largo de este manual; en consecuencia, la variable dependiente se denominará *tiempo de sobrevivencia* (Fox, 2001; Muenchow, 1986).

Los datos sobre la ocurrencia y el momento de los eventos se recopilan de manera rutinaria durante los estudios de campo y de laboratorio. Los ejemplos incluyen el tiempo hasta la primera aparición de síntomas en una población de plantas bajo estudio; el tiempo de germinación en una población de esporas de hongos o de semillas; el tiempo hasta la muerte en una población de huevos de nematodos; o el tiempo de defoliación en una población de hojas enfermas. A pesar de la ocurrencia común de datos de sobrevivencia en fitopatología, el análisis de sobrevivencia es poco utilizado para examinar tales conjuntos de datos (Scherin y Ojiambo, 2004).

### **3.1.1. Principios de análisis de sobrevivencia en estudios ecológicos**

Hay varias preguntas que un investigador se haría en relación con los datos de sobrevivencia (Cuadro 08). ¿Primero, es de interés estimar la distribución del tiempo de sobrevivencia para un grupo de individuos? Entre otras cosas, esto permite calcular el riesgo de que el evento ocurra dentro de un intervalo dado y calcular las cantidades derivadas, tales como la vida media residual, es decir, el tiempo hasta la ocurrencia del evento para un individuo que ha sobrevivido (es decir, no experimentado el evento) hasta el comienzo del intervalo. Un segundo objetivo podría ser comparar las distribuciones del tiempo de sobrevivencia entre dos o más grupos, por ejemplo, organismos sometidos a diferentes tratamientos tras el diagnóstico de una enfermedad, ataque de insectos, etc. Finalmente, se desea cuantificar los efectos de una o más variables independientes (covariables) sobre los tiempos de sobrevivencia, como un esfuerzo por desarrollar un modelo para describir o predecir los tiempos de sobrevivencia en una población.

**Cuadro 08. Aplicaciones generales de análisis de sobrevivencia y su implementación con procedimientos de Statistical Analysis System (SAS).**

Aplicación	Salida principal	implementación con SAS
Estimación de las distribuciones del tiempo de supervivencia.	Función de superviviente, $S(t)$ ; función de riesgo, $h(t)$ ; y cantidades derivadas (p. ej., vida útil residual mediana)	PROC LIFETEST
Comparación de las distribuciones del tiempo de supervivencia entre grupos	Prueba las estadísticas para la hipótesis de igualdad entre los grupos: pruebas de rango logarítmico, Wilcoxon y de razón de verosimilitud	PROC LIFETEST
Modelado del efecto de covariables discretas y / o continuas sobre los tiempos de supervivencia	Estimaciones de parámetros y niveles de significancia asociados; cociente de riesgo	PROC LIFEREG PROC PHREG
Una comparación detallada de las características del análisis de supervivencia de los tres paquetes SAS, SPSS y BMDP viene dada por Pelz y Klein (1996).		

Fuente. Pelz y Klein (1996).

Los investigadores a menudo se ven tentados a utilizar la metodología estadística convencional para el análisis de los datos. Por ejemplo, el análisis de regresión logística podría aplicarse para cuantificar la importancia de ciertas covariables al clasificar a los individuos en dos grupos, aquellos que experimentaron o no el evento durante el período de observación. Sin embargo, este enfoque puede resultar en una pérdida considerable de información porque no se consideran las diferencias en el momento en que ocurren los eventos. Alternativamente, se podría utilizar un análisis de regresión de mínimos cuadrados para identificar las covariables que influyen en los tiempos de supervivencia. El mayor inconveniente radica en el hecho de que los datos de supervivencia a menudo se censuran, es decir, contienen observaciones de las que no se sabe cuándo ocurrió el evento, ya sea porque el individuo correspondiente se perdió del conjunto de datos durante el período de estudio, o porque el estudio terminó antes de que todos los individuos experimentaran el evento. Ambas causas de censura ocurren comúnmente durante las investigaciones ecológicas de campo o laboratorio.

Con la metodología estadística convencional, las observaciones censuradas tendrían que eliminarse o se tendrían que hacer suposiciones *ad hoc*, por ejemplo, estableciendo sus tiempos de supervivencia en el período de tiempo total, desde el inicio hasta la terminación del estudio. Por el contrario, los mé-

todos de estimación de parámetros basados en verosimilitud que se utilizan en el análisis de sobrevivencia pueden extraer información relevante, tanto de observaciones censuradas como no censuradas, produciendo así estimaciones confiables de parámetros. Además, el análisis de sobrevivencia es el único método que puede acomodar fácilmente las covariables dependientes del tiempo, es decir, variables independientes cuyos valores cambian durante el curso del estudio. Como ejemplo, una investigación epidemiológica en la que el evento de interés es el tiempo de defoliación en relación con la gravedad de la enfermedad foliar. La gravedad de la enfermedad es una covariable dependiente del tiempo, ya que los valores de gravedad cambian con el tiempo y lo hacen de manera diferente para cada hoja (Schermer y Ojiambo, 2004; Lee y Wenyu, 2003; Allison, 2010).

### 3.1.2. Modelos de datos de sobrevivencia en estudios ecológicos

El enfoque más poderoso para el análisis de sobrevivencia se basa en el uso de modelos para describir y predecir los efectos de covariables discretas y / o continuas en los tiempos de sobrevivencia en una población. En la mayoría de los paquetes estadísticos se encuentran disponibles dos enfoques generales: el modelo de tiempo de falla acelerado (AFT) y el modelo de riesgos proporcionales (PH) (Allison, 2010; Pelz y Klein, 1996). Las principales diferencias entre los dos métodos se resumen en el Cuadro 09. El modelo PH, también conocido como modelo de regresión de Cox (Cox, 1972), es más flexible y más fácil de usar porque no requiere que el investigador asuma una distribución de probabilidad particular que represente los tiempos de sobrevivencia. Esto puede ser ventajoso en situaciones en las que no existen razones biológicas *a priori* en las cuales elegir una distribución sobre otra. El enfoque de PH se basa en la premisa de que, para una covariable determinada, las funciones de riesgo en diferentes niveles de esa covariable, son proporcionales. En otras palabras, las funciones de riesgo son similares en forma para diferentes niveles de la covariable y solo se ve afectado el riesgo relativo de ocurrencia de eventos en un momento dado. El modelo AFT es más apropiado en situaciones donde estos supuestos no se satisfacen, por ejemplo, si la covariable modifica principalmente el momento de los períodos de alto riesgo (Fox, 2001). En la práctica, es relativamente sencillo probar si el supuesto de PH se cumple mediante el desarrollo de un modelo de PH que incluye un término de interacción entre el tiempo y la covariable de interés, además de la covariable en sí; *si el término de interacción es estadísticamente significativo, el modelo AFT es más apropiado* (Lee y Wenyu, 2003).

**Cuadro 09. Enfoques generales de modelos para datos de sobrevivencia y su implementación usando procedimientos del Sistema de Análisis Estadístico (SAS)<sup>a</sup>**

Propiedad	Modelo tiempo de falla acelerado (AFT)	Modelo de riesgo Proporcional (HR)
Formulación del Modelo <sup>b</sup>	$\log_e T = \beta_0 + \beta_1 X_1 + \beta_2 X_2$	$\log_e h(t) = \log_e \lambda_0(t) + \beta_1 X_1 + \beta_2 X_2$
Método de Estimación	Maxima probabilidad (paramétrica)	Probabilidad Parcial (semi-paramétrica)
Especificar Distribución de sobrevivencia	Debe especificarse (ejemplo: exponencial, Gompertz, Weibull, gamma)	Forma no especificada
Efecto de covariables	Efectos multiplicativos sobre tiempo de sobrevivencia	Efectos multiplicativos sobre la función de riesgo
Covariables Tiempo-dependientes	No disponible	Incluidas
Implementación con SAS	PROC LIFEREG	PROC PHREG
<sup>a</sup> Procedimientos similares disponibles en otros paquetes (SPSS – BMDP). Una comparación de los tres paquetes aparece en Pelz & Klein (1996). <sup>b</sup> Ejemplos para modelos con 2 covariables, $X_1$ y $X_2$ . T = tiempo sobrevivencia, $h(t)$ = función de riesgo, $\lambda_0(t)$ = línea base de función de riesgo, y $\beta_i$ = parámetros a estimar. La línea base de la función de riesgo, que representa la función de riesgo para un individuo cuyos valores de $X_i$ son iguales a cero, no es estimada.		

Fuente. Pelz y Klein (1996).

### 3.1.3. Ejemplos introductorios para la interpretación de datos ecológicos

**3.1.3.1. Tiempos de emergencia y floración en especies invasoras (Fox 2001).** Esta investigación fue desarrollada por Fox (2001), como parte de un estudio más amplio sobre la ecología y la evolución en el rábano silvestre *Raphanus sativus* L. (Brassicaceae). Se estudió el tiempo de emergencia y floración en tres poblaciones de rábano silvestre y en dos cultivares de rábano doméstico. El rábano silvestre es una maleza anual (a veces bienal) en muchas partes del hemisferio norte. En América del Norte, *R. sativus* es especialmente común en las zonas costeras de California y en el Valle de Sacramento. Existen dos fuentes originales del viejo mundo: introducciones inadvertidas por colonos europeos y escapes de variedades de rábano cultivadas. Estos últimos son de la misma especie que el rábano silvestre. También existen híbridos con una especie introducida estrechamente relacionada, *R. raphanistrum*.

Se discuten dos preguntas muy relacionadas de esta investigación: 1) ¿Cuáles son las distribuciones de los tiempos de emergencia y floración en estas poblaciones, y 2) ¿estas difieren entre poblaciones? Para solucionar estas preguntas se utilizaron dos métodos estadísticos diferentes: análisis de las tablas de vida y modelos de tiempo de falla acelerado. Estos métodos son complementarios, cada uno con ventajas específicas: las tablas de vida y las pruebas estadísticas asociadas son bastante útiles para los análisis de datos exploratorios, pero carecen del poder estadístico de los modelos de tiempo de falla acelerado. Por otro lado, los modelos de tiempo de falla acelerado requieren más suposiciones que los métodos de tablas de vida.

#### 3.1.3.1.1. Métodos experimentales

Se recolectaron semillas de individuos seleccionados al azar en poblaciones alrededor de Santa Bárbara, California. El sitio de Campus Point está en un acantilado con vista al Océano Pacífico, y el sitio de Coal Oil Point a varios cientos de metros del océano (ambientes húmedos). El tercer sitio, Storke Road, se encuentra a varios kilómetros tierra adentro (más seco y cálido). Los cultivares domésticos utilizados fueron Black Spanish (variedad de floración tardía) y Rapid Cycling, un producto de selección artificial para un tiempo de generación corto. Al diseñar experimentos de tiempo de falla, ya sea en el campo o laboratorio, es importante considerar las formas de las distribuciones de tiempo de falla y

los patrones y magnitudes de la censura de datos. Por ejemplo, si hay mucha censura de datos, se necesitan muestras grandes, que a menudo involucran a cientos de individuos, para comparar tratamientos. La razón de esto radica en que, para comparar grupos, se necesita un número mínimo de tiempos de eventos reales, en contraposición a los tiempos de censura. *Si todos los puntos de datos son censurados, no es posible realizar una comparación entre grupos.* También se necesitan muestras grandes si las colas de la distribución son de particular interés ecológico (Fox, 1990b). Ya que el estudio se consideró un experimento piloto, no existía una estimación a priori del número de puntos de datos censurados que se esperaban, ni de las formas de las distribuciones del tiempo de falla. En consecuencia, se utilizaron muestras relativamente grandes, plantando 180 semillas de cada uno de los dos cultivares domésticos, 140 de los sitios Storke Road y Campus Point, y 130 del sitio Coal Oil Point, para un total de 770 semillas. Se plantó una semilla en cada maceta de 10 cm. Debido a que las hipótesis de interés se refieren a las diferencias entre poblaciones en tiempo de emergencia y antesis, las semillas se consideraron como unidades experimentales y las semillas individuales de la misma población como réplicas.

Las limitaciones de tiempo establecían que el experimento se terminara en 150 días (días después de la siembra). Para cada semilla individual, se registró un tiempo de emergencia TEMERG y un tiempo de antesis TANTH. Cada TEMERG y TANTH podría ser un tiempo real de emergencia o antesis, respectivamente, o un tiempo de censura. También se registraron las variables para censura EMRGCENS y ANTHCENS, asignando valores de 0 para puntos de datos no censurados y 1 para puntos de datos censurados.

### **Métodos estadísticos**

Los datos de antesis requieren un manejo especial, porque muchos de los datos están censurados. Pero en los datos de emergencia, la pequeña cantidad de puntos de datos censurados, el ANOVA podría considerarse una forma útil para comparar las poblaciones. El proceso de germinación comenzó tiempo después del primer riego. Ya que no existe evidencia de latencia de las semillas en esta especie, se podría esperar que los eventos de emergencia se agrupen en algún momento después del inicio del experimento, y desaparecer después de ese punto. Ninguna transformación estándar consigue normalizar estos datos. Por lo tanto, aunque hay muy poca censura en los datos de emergencia, la comparación de



medias usando ANOVA resultaría inapropiada. Afortunadamente, existen varios métodos para analizar datos de tiempo de falla que pueden manejar correctamente los datos censurados y que no requieren datos distribuidos normalmente. Estos incluyen análisis de tabla de vida y dos tipos diferentes de modelos de regresión: modelos de tiempo de falla acelerado y modelos de riesgos proporcionales.

### **Métodos y aplicación de tabla de vida**

Las hipótesis se pueden probar con estimaciones de las tablas de vida y son útiles para el análisis de datos exploratorios. A partir de la medición de la tasa de fallas o germinación, es simple estimar la proporción de quienes germinan en cada intervalo. Existen otras cuatro estadísticas que pueden derivarse de esta información. La función estimada de supervivencia acumulada,  $S_i$ , es la fracción de la cohorte que aún no ha germinado. La función de densidad de probabilidad estimada,  $P_i$ , da la probabilidad de que una semilla germine en el tiempo 0 en el  $i$ -ésimo intervalo. Por eso  $P_i$  también se denomina tasa de sobrevivencia incondicional. La función de riesgo estimada,  $h_i$ , a veces es llamada tasa de sobrevivencia condicional: estima la probabilidad de falla en el  $i$ -ésimo intervalo dado que un individuo ha sobrevivido hasta el comienzo de ese intervalo. Finalmente, la expectativa de vida mediana estimada,  $e_i$ , da la mediana del tiempo hasta el evento, de un individuo que ha sobrevivido hasta el comienzo del  $i$ -ésimo intervalo;  $e_0$  da la esperanza de vida media de un recién emergido.

La importancia ecológica de la función de supervivencia acumulativa  $S_i$ , es obvia, ya que refleja la velocidad a la que se acumulan las fallas o germinaciones, mientras que la función de riesgo refleja el valor per cápita de riesgo entre los que quedan sin germinar. La mediana, en lugar de la media, calcula la esperanza de vida; la mediana es a menudo una estimación más útil y robusta de la tendencia central cuando los datos no están distribuidos simétricamente.

Hay varias formas de comparar poblaciones o tratamientos estadísticamente. Se pueden utilizar errores estándar de las estimaciones de la tabla de vida para comparaciones por pares. Si los datos no están censurados, las pruebas no paramétricas, como las pruebas G o las pruebas  $\chi^2$  (chi cuadrado), se pueden utilizar para probar la independencia de los grupos. Con datos censurados, las pruebas de Wilcoxon o log-rank pueden utilizarse en los datos del tiempo de germinación para probar la heterogeneidad entre los grupos. Ambas pruebas comparan lo observado con el número esperado de germinación en cada in-

tervalo. La prueba de rango logarítmico es más poderosa que la de Wilcoxon cuando las funciones de riesgo de las diferentes muestras son proporcionales entre sí, y cuando no hay censura o solo censura aleatoria.

Estas pruebas, así como los análisis de la tabla de vida, se pueden realizar en SAS con el procedimiento LIFETEST (que utiliza la aproximación  $\chi^2$  a las pruebas de rango logarítmico y Wilcoxon descritas). Si estas pruebas revelan heterogeneidad y hay más de dos grupos, se puede utilizar Wilcoxon o rango logarítmico para construir las estadísticas para las comparaciones múltiples entre grupos. Por otra parte, el método Kaplan-Meier (KM) (predeterminado en el procedimiento LIFETEST) difiere de la tabla de vida en que la función de supervivencia acumulada cambia sólo cuando un evento es realmente observado.

La separación de las curvas de supervivencia implica que las poblaciones pueden diferir en el tiempo de emergencia o de floración. Esto es soportado con las pruebas de Wilcoxon y log-rank calculadas por el procedimiento LIFETEST: Wilcoxon  $\chi^2 = 439,1$ ,  $gl = 4$ ,  $p = 0,0001$ , rango logarítmico  $\chi^2 = 335,8$ ,  $gl = 4$ ,  $p = 0,0001$ . Estas pruebas indican que las cinco poblaciones son heterogéneas, pero no indican cuales poblaciones difieren entre sí. Para contestar esa pregunta, se puede utilizar el procedimiento STRATA mediante una prueba de Tukey con el procedimiento ADJUST=Tukey, que produce todas las comparaciones.

Los datos de antesis también se calculan mediante el procedimiento LIFETEST. En este caso se puede presentar una censura considerable, porque muchas plantas no han florecido cuando el estudio termina, además de que pueden ocurrir algunas muertes. Sin embargo, las curvas de supervivencia pueden implicar que las poblaciones difieren. Esta conclusión también puede respaldarse por las pruebas de Wilcoxon y de rango logarítmico calculadas por el procedimiento LIFETEST. Supóngase que la prueba de Wilcoxon  $\chi^2 = 582,2$ ,  $gl = 4$ ,  $p = 0,0001$  y la prueba de log-rank  $\chi^2 = 580,6$ ,  $gl = 4$ ,  $p = 0,0001$ . Ya que estas cinco poblaciones son heterogéneas, ¿cuáles difieren entre sí? Para examinar esta pregunta, se utiliza nuevamente el procedimiento STRATA mediante una prueba de Tukey con el procedimiento ADJUST=Tukey, que produce todas las comparaciones. Estas comparaciones múltiples, realizadas en el nivel de significancia de 0,05, pueden sugerir las diferencias entre poblaciones. Además, pueden sugerir que las poblaciones de tal planta difieren en el tiempo de emergencia y floración.

## Aplicación de modelos de regresión

Los modelos *de tiempo de falla acelerado* y *modelo de riesgos proporcionales* hacen diferentes suposiciones sobre el efecto de las covariables como, por ejemplo, tratamiento o tamaño de semilla. *Los modelos de tiempo de falla acelerado* asumen que los tratamientos y las covariables afectan el tiempo de falla o germinación multiplicativamente (es decir, la germinación de las semillas se "aceleran"). Una interpretación alternativa es que, bajo modelos de tiempo de falla, las covariables hacen correr más rápido o más lento el reloj, por lo que cualquier período de alto riesgo cambiará en el tiempo cuando las covariables cambian. *El modelo de riesgos proporcionales* asume que las covariables afectan las funciones de riesgo de diferentes grupos multiplicativamente. En modelos de riesgos proporcionales, los períodos de alto riesgo siguen siendo los mismos, pero la posibilidad de que un individuo caiga en uno de esos períodos variará con las covariables individuales (Kalbfleisch y Prentice, 1980).

¿Cómo elegir entre ambos modelos? Probablemente el enfoque más útil es considerar las hipótesis ecológicas y preguntarse si los tratamientos cambian el momento de los períodos de alto riesgo (lo que sugiere el uso de modelos de tiempo de falla acelerado) o si los tratamientos simplemente cambian la posibilidad de falla (lo que sugiere el uso de modelos de riesgos proporcionales). *Una comprobación adicional de la idoneidad del modelo de riesgos proporcionales es graficar  $\log(-\log[S(t)])$  contra el tiempo. Si el modelo de riesgos proporcionales es apropiado, las curvas de los diferentes grupos deben ser aproximadamente paralelas entre sí, para un determinado nivel de una covariable (Kalbfleisch y Prentice, 1980).* Una limitación de este enfoque es que deben realizarse comparaciones entre grupos de tratamiento dentro de los niveles de una covariable. Si todas las variables predictoras son covariables, no se pueden hacer comparaciones, independientemente de si el modelo de riesgos proporcionales es apropiado. En consecuencia, es probablemente mejor confiar en el examen de las hipótesis ecológicas.

## Descripción y aplicación de modelos de tiempo de falla acelerado

En un modelo de tiempo de falla acelerado, los tiempos de falla  $T$  se modelan como

$$\ln(T) = X'\beta + \sigma\varepsilon;$$

donde  $X$  es una matriz de valores de covariables y  $X'$  es su transpuesta,  $\beta$  un vector de parámetros de regresión,  $\varepsilon$  un vector de errores de una distribución de sobrevivencia especificada, y  $\sigma$  un parámetro. Las covariables en  $X$  pueden ser variables ficticias (dummy) correspondientes a efectos categóricos, o variables continuas, o ambas, o sus interacciones. Por ejemplo,  $X$  para los datos de emergencia del rábano es una matriz de 5 x 5 de variables ficticias para las cinco poblaciones de origen y 5 coeficientes de regresión en  $\beta$ . Para los datos de floración,  $X$  incluye una variable adicional, fecha de emergencia y hay un coeficiente de regresión adicional para la fecha de emergencia en  $\beta$ . En consecuencia, existen dos pasos para utilizar estos modelos: 1) elegir una distribución de supervivencia y, 2) estimar los parámetros para la distribución de supervivencia y los coeficientes de regresión  $\beta$ .

Existen tres formas diferentes de elegir una distribución de sobrevivencia. Primero, se puede elegir una distribución basada en fundamentos ecológicos o biológicos a priori. En segundo lugar, se puede tomar un enfoque empírico: después de que se hayan recopilado los datos, comparar las estimaciones de la tabla de vida de sus funciones de riesgo o supervivencia con las funciones de riesgo o supervivencia de varias distribuciones (Exponencial, Weibull, Gamma, Lognormal, Log-logista). A menudo es biológicamente importante que un conjunto de datos sea, Gamma en lugar de Weibull. Por ejemplo, la diferencia entre las funciones de supervivencia clásicas de Deeevey Tipos I, II y III, es precisamente que un Tipo II implica un riesgo constante de mortalidad (una distribución exponencial), mientras que los Tipos I y III implican riesgos crecientes y decrecientes, respectivamente.

Estos métodos conducen naturalmente a un conjunto de pruebas  $\chi^2$  para coeficientes de regresión. Estos prueban la importancia de la contribución adicional de la covariable a la función de probabilidad maximizada, más allá de las contribuciones de las otras covariables que ya están en el modelo. Por tanto, un valor no significativo de  $\chi^2$  significa que el modelo es adecuado sin esa covariable en particular, pero no implica que se puedan eliminar todas las covariables con valores  $\chi^2$  no significativos. De hecho, es posible que un conjunto de covariables sea significativo conjuntamente, aunque ninguna de ellas sea significativa individualmente. Para examinar la posible importancia de las covariables con valores  $\chi^2$  no significativos, entonces, se eliminan los términos no significativos uno a la vez, se reevalúa el modelo cada vez y examina los cambios en los otros términos.

Cuando existen varios niveles de una covariable categórica, por ejemplo, tratamientos o la población, se elige un nivel arbitrario como nivel de referencia con un coeficiente de regresión de 0 y los coeficientes de regresión de los otros niveles proporcionan comparaciones con el nivel de referencia. Por lo tanto, *un valor significativo de  $\chi^2$  para un nivel en particular significa solo que el nivel es significativamente diferente del nivel de referencia y puede o no ser significativamente diferente de niveles distintos al nivel de referencia*; es necesario realizar múltiples comparaciones para examinar esta hipótesis. Por defecto, el SAS elige como nivel de referencia el último nivel. Por ejemplo, si existen 5 niveles o tratamientos, elige el tratamiento 5. Sin embargo, el investigador puede elegir el tratamiento de referencia de acuerdo a su elección, utilizando como ejemplo, el siguiente procedimiento:

```
Proc Phreg Data=Jabillo;
Class Tratam (REF='0');
Model Días*Status (0) = Tratam Tratdias / Ties=Efron R1;
Test Tratam6=Tratam12; Test Tratam6=Tratam18; Test Tratam6=Tratam24;
Test Tratam12=Tratam18; Test Tratam12=Tratam24; Test Tratam18=Tratam24;
Tratdias = Tratam*(Días < 15);
Run;
```

En este caso, se seleccionó el tratamiento sin inmersión (0 horas) para comparar con 6 – 12 – 18 y 24 horas de inmersión de semillas en agua.

Para los datos de emergencia del rábano, parecía razonable elegir entre la distribución Log-logística o Lognormal, ya que la especie tiene poca o ninguna latencia de semillas, se puede esperar que la tasa de germinación y la posterior emergencia aumenten rápidamente después de la aplicación de agua, alcanzando una tasa máxima y luego declinar.

Para los datos de anthesis de rábano, se pensó que era probable que los datos se ajustaran mejor a una distribución gamma. El riesgo gamma se acerca monótonamente a un valor constante y, por lo tanto, los modelos gamma a menudo pueden proporcionar buenos ajustes para los procesos de desarrollo inevitables. Además, una distribución gamma es una suma de variables distribuidas exponencialmente;

si se piensa en los intervalos entre las divisiones celulares como aproximados por una distribución exponencial, nuevamente se esperaría que una etapa de desarrollo de nivel superior, como el tiempo de floración, se distribuyera en forma gamma. Tal intuición a menudo es incorrecta, por lo que se compara el modelo gamma con modelos de otras distribuciones (Exponencial, Weibull, Lognormal o Log-logística).

Hay varias formas de realizar esta comparación, incluida una prueba del multiplicador de Lagrange y una prueba de razón de probabilidad (likelihood ratio test). En principio son equivalentes, pero la prueba de razón de verosimilitud puede ser más robusta para muestras pequeñas.

La elección entre las distribuciones Gamma y Weibull basadas en el conjunto de datos propuesto, requiere una justificación biológica, más que estadística. La distribución Gamma parece una descripción biológicamente más razonable del tiempo hasta la antesis de una planta anual, porque el riesgo Gamma tiende hacia una constante a medida que el tiempo aumenta, mientras que el riesgo de Weibull tiende a cero o al infinito. Los análisis de los modelos de tiempo de falla del rábano se muestran en los Cuadros 10 y 11.

**Cuadro 10. Análisis del modelo de tiempo de falla acelerada para tiempo de emergencia de *Raphanus sativus*, utilizando 764 valores no censurados, 6 valores censurados a la derecha y una distribución Log-logística.**

Variable	G.L	Estimado	SE	X <sup>2</sup>	P
Intercepto	1	1,868	0,014	17442,78	0,0001
Población	4			395,67	0,0001
Black Spanish	1	-0,254	0,018	194,03	0,0001
Coal Oil Point	1	-0,072	0,020	13,09	0,0003
Campus Point	1	-0,084	0,019	19,63	0,0001
Rapid Cycling	1	-0,298	0,018	250,84	0,0001
Storke Road	0	0	0		
Parámetro de escala	1	0,100	0,003		
<sup>a</sup> Loglikelihood= 116,3					

Fuente. Fox (2001).

En ambos ejemplos, los coeficientes de regresión para la población de Storke Road es cero. Como se mencionó anteriormente, *cuando hay múltiples niveles de una variable de clase (en este caso, población), SAS toma un nivel (el último) como nivel de referencia, y todos los demás se comparan con este nivel. Por tanto, la población de Storke Road se considera la población de referencia en virtud de su orden.* Significa que, los coeficientes de otras poblaciones se comparan con Storke Road, y las pruebas de significancia para cada población prueban si difiere o no de Storke Road.

El análisis de los datos de emergencia (Cuadro 10) muestra que la población de origen contribuye significativamente al modelo. El hecho de que el coeficiente de regresión de cada población sea significativo, indica que cada uno difiere de la población de referencia (Storke Road). ¿Cuáles otras poblaciones difieren entre sí? Para examinar esta pregunta, se ejecutan múltiples comparaciones, análogas a las utilizadas en los análisis de tablas de vida, excepto que en este caso se utilizan parámetros de regresión como las estadísticas de comparación, en lugar de las puntuaciones de clasificación de Wilcoxon.

**Cuadro 11. Análisis del modelo de tiempo de falla acelerada para tiempo de antesis de *Rap-hanus sativus*, utilizando 210 valores no censurados, 554 valores censurados a la derecha y una distribución Gamma.**

Variable		G.L	Estimado	SE	X <sup>2</sup>	P
Intercepto		1	4,974	0,090	3047,02	0,0001
Tiempo de emergencia		1	0,019	0,010	3,68	0,06
Población		4			495,43	0,0001
Black Spanish		1	0,449	0,069	42,90	0,0001
Coal Oil Point		1	0,124	0,058	4,58	0,03
Campus Point		1	0,324	0,063	26,70	0,0001
Rapid Cycling		1	-0,682	0,054	154,35	0,0001
Storke Road		0	0	0		
Parámetro de escala		1	0,394	0,021		
Parámetro de forma		1	-0,822	0,179		
<sup>a</sup> Loglikelihood= -232,48						

Fuente. Fox (2001).

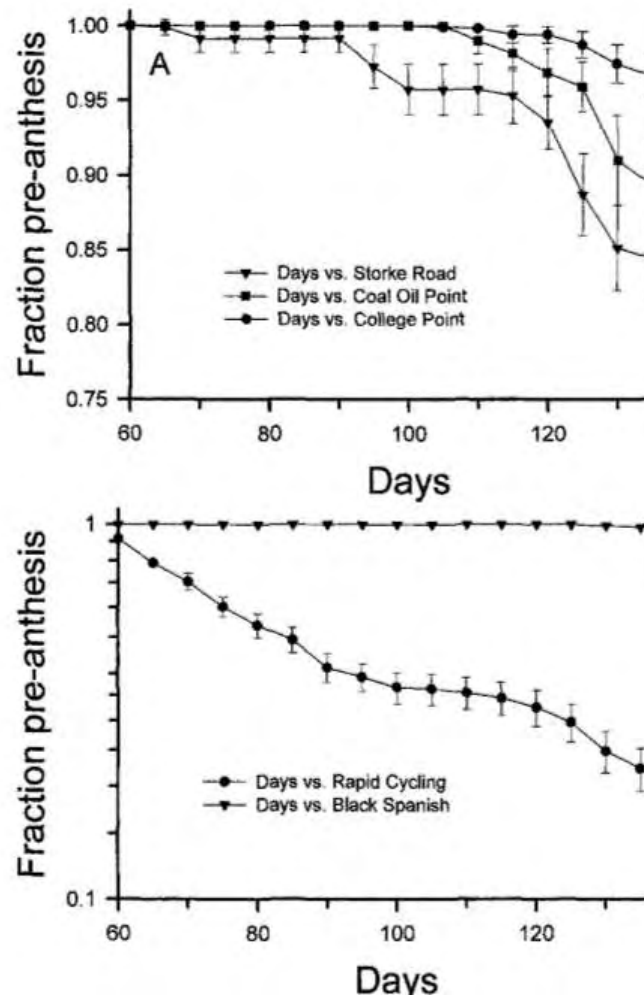
Los análisis de las tablas de vida mostraron una heterogeneidad significativa entre las poblaciones, y las comparaciones por pares sugirieron que todos los pares difieren, excepto Coal Oil Point - Campus Point y posiblemente Black Spanish - Rapid Cycling. Los coeficientes de regresión del Cuadro 10 y su matriz de covarianza estimada llevó a una conclusión algo diferente: la comparación Coal Oil Point - Campus Point resulta en una alta probabilidad de ser de la misma población, pero la comparación Black Spanish - Rapid Cycling ahora muestra diferencias significativas entre estas poblaciones. Todas las demás comparaciones se mantuvieron cualitativamente sin cambios. El modelo de tiempo de falla acelerado ha confirmado los patrones generales sugeridos por el análisis de la tabla de vida, pero su mayor poder estadístico también hizo posible encontrar diferencias significativas entre poblaciones que no fueron reveladas por el análisis de la tabla de vida.

Para interpretar los parámetros de regresión en el Cuadro 10, hay que recordar que corresponden a los efectos multiplicativos de las covariables sobre la probabilidad de emergencia de las plántulas. El tiempo medio de emergencia para una planta de Rapid Cycling es  $e^{(1,8680 - 0,298)} p_0(t) = 4,807 p_0(t)$ , mientras que para una planta de Storke Road es  $e^{1,868} p_0(t) = 6,475 p_0(t)$ . *Cuanto más negativo sea el coeficiente, más temprano será el tiempo medio de emergencia.* La distribución de referencia  $p_0(t)$  es la distribución log-logística con el parámetro de escala que se muestra en el Cuadro 10 (0,10).

El análisis de los datos de antesis (Cuadro 10) también muestra que cada población contribuye significativamente al modelo. La emergencia tardía probablemente tiende a retrasar la antesis. Esto es de esperar, porque la variable dependiente es el tiempo desde la siembra hasta la antesis, que debe ser mayor que el tiempo desde la siembra hasta la emergencia. Por tanto, incluir el tiempo de emergencia en el modelo es análogo a incluir un efecto de bloque en ANOVA.

Los coeficientes de regresión se interpretan de la misma manera que para los datos de tiempo de emergencia. Sólo difieren las distribuciones de probabilidad de referencia (la emergencia se distribuye como log-logística, mientras que se considera que la antesis tiene distribución gamma). Sin embargo, esto no afecta la interpretación cualitativa de los parámetros de regresión. Por ejemplo, el análisis de los datos de antesis muestra que las plantas Rapid Cycling alcanzan la antesis mucho antes y las plantas Black Spanish mucho más tarde que todas las demás. Esto se corresponde bien con los resultados que se muestran en la Figura 09.





**Figura 09. Curvas de sobrevivencia para antesis en rabano. Las barras son del error estándar de la tabla de vida. Todas las curvas finalizan en dato censurado. (A) Poblaciones Santa Barbara; (B) Cultivares. Fuente. Fox (2001).**

El coeficiente significativo para la población de Coal Oil Point indica que difiere significativamente de la población de Storke Road. Por el contrario, las comparaciones por pares utilizadas siguiendo el análisis de la tabla de vida sugirió que estas poblaciones no difirieron. Como con los datos de la emergencia, las comparaciones múltiples realizadas originan más diferencias entre poblaciones, que no fueron identificados por el análisis de la tabla de vida. Estas comparaciones, realizadas al nivel de significancia de 0,05, sugieren que todas las po-

blaciones difieren entre sí, excepto posiblemente el par Black Spanish - Campus Point, cuyas diferencias fueron marginalmente significativas. ¿Cual análisis es correcto? El modelo de tiempo de falla acelerado tiene más poder estadístico y, por lo tanto, es razonable tener más confianza en estos resultados que en el análisis de la tabla de vida.

### Descripción y aplicación al modelo de riesgo proporcional

Como se mencionó anteriormente, en un modelo de riesgos proporcionales, el efecto de las covariables es cambiar la posibilidad de caer en un período de alto riesgo. Las covariables actúan multiplicativamente sobre la función de riesgo, en lugar del tiempo de falla (como en los modelos de tiempo de falla acelerado). La función de riesgo para el  $i$ -ésimo grupo es así:

$$h_i(t) = h_0(t) \exp(\sum \beta_j X_{ij})$$

donde  $h_0(t)$  es una función de riesgo de referencia que cambia por las covariables  $X$  y los coeficientes de regresión  $\beta$ .

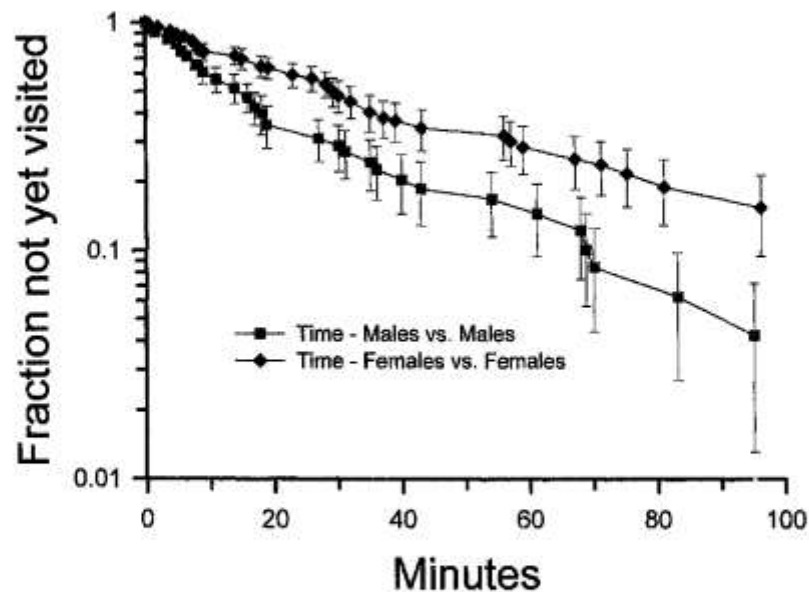
El modelo de riesgos proporcionales de Cox, que se utiliza mucho en epidemiología, estima la función de riesgo de referencia de forma no paramétrica. Los coeficientes de regresión  $\beta$  se estiman luego numéricamente con un procedimiento de máxima probabilidad. La distribución de Weibull también se puede utilizar para modelos de riesgos proporcionales, comparando dos grupos, tratando el parámetro  $p$  como la razón de los dos coeficientes de regresión (Kalbfleisch y Prentice, 1980).

Debido a que las covariables actúan sobre el tiempo de falla en los modelos de tiempo de falla acelerados y sobre la función de riesgo en los modelos de riesgos proporcionales, los coeficientes de regresión pueden tener significados opuestos. Un coeficiente positivo en un modelo de riesgos proporcionales significa que la covariable aumenta el riesgo, lo que reduce el tiempo de falla. Un coeficiente positivo en un modelo de tiempo de falla acelerado significa que la covariable aumenta el tiempo de falla.

En otro ejemplo, Muenchow (1986), estuvo interesado en probar si las flores masculinas y femeninas de la planta dioica *Clematis linguisticifolia* eran igualmente atractivas para los insectos polinizadores, en contra de la hipótesis alternativa de que las flores masculinas son más atractivas. Reconoció que esto podría tratarse como un problema de tiempo de falla ya que las diferencias en el atractivo deberían

conducir a diferencias en el tiempo hasta la primera visita del insecto. Al tratar esto como un problema de tiempo de falla, se examinaron estas hipótesis con un diseño experimental creativo: se observó pares de flores en la misma planta y se registraron los tiempos de espera hasta la primera visita del insecto. Para cada observación, también se registró la hora del día, la temperatura del aire y una categorización de la densidad de flores alrededor de 1 m de la planta objetivo. Los tiempos de espera se registraron al minuto más cercano.

Este estudio es uno en el que los supuestos del modelo de riesgos proporcionales de Cox parecen ser razonables a priori. Si hay diferencias en el atractivo, es probable que actúen para aumentar el "riesgo" (es decir, la posibilidad de una visita) del género atractivo, en relación con el riesgo del género menos atractivo. Otra forma de ver esto, es que ser miembro del género atractivo no debería tener ningún efecto sobre la cantidad de insectos en el área, pero sí afecta las posibilidades de una visita una vez que un insecto se encuentra a cierta distancia de la planta. La estimación de Muenchow de la función de supervivencia se muestra en la Figura 10. Señaló que estos datos parecen haber sido extraídos de una distribución exponencial, pero prefirió usar un modelo de Cox para sus pruebas estadísticas.



**Figura 10.** Estimado producto – límite de la función de supervivencia de los tiempos de espera de los insectos polinizadores de flores de *Clematis linguistifolia*. Fuente. Fox (2001).

Una parte importante del estudio de Muenchow es la forma en que analizó las cuatro covariables: género, categoría de densidad de flores, temperatura y hora del día. Informó que el tiempo de inicio y los coeficientes de temperatura no fueron significativamente diferentes de cero, es decir, estos factores no parecen influir en el tiempo de espera para las visitas. Debido a que las plantas masculinas tenían más flores, el género y la densidad de flores estaban correlacionados. En consecuencia, Muenchow analizó los datos estratificando: examinó el efecto de la densidad de flores dentro de cada género y el efecto del género dentro de cada categoría de densidad de flores.

Tanto el género como la densidad de flores tuvieron efectos significativos de forma independiente: dentro de un género, los insectos visitaron grupos densos de flores más rápido que otros grupos, y dentro de una categoría de densidad, los insectos visitaron a las flores machos más rápido que a las femeninas. Muenchow concluyó que las flores masculinas eran más atractivas, ya que tenían flores más densas y porque tenían algún carácter atractivo desconocido.

### **3.1.3.2. Factores físicos y biológicos involucrados en la supervivencia y riesgos de mortalidad de plántulas de *Abies religiosa* (Kunth) Schltl. & Cham. bajo regeneración natural (Hernández, López, Cruz y Luna,2022)**

Los factores físicos y biológicos involucrados en la supervivencia y riesgos de mortalidad de plántulas, son importantes para comprender los procesos naturales asociados con su establecimiento en un contexto espacio-temporal. Basados en estos antecedentes, se aprovecha el estudio desarrollado por Hernández, López, Cruz y Luna (2022), cuyo objetivo consistió en identificar los factores ambientales que definen el nicho de regeneración de *Abies religiosa* con base en la supervivencia y riesgos de mortalidad de sus plántulas.

### 3.1.3.2.1. Materiales y métodos

#### Zona de estudio

El clima en la zona de estudio corresponde a un templado-húmedo con precipitación media anual entre 900 y 1,200 mm. El régimen de lluvias es en verano y la temperatura media va de los 10 a los 14°C. Los suelos son incipientes, de textura gruesa en las proximidades del cono cinerítico del Tláloc y en el resto del área negros, profundos, muy ricos en materia orgánica y de textura media (migajones o francos) (Sánchez y López, 2003), cuyos valores de pH varían de 5.5 a 7.1. Esta porción del Parque Nacional corresponde a los bosques de la zona templada subhúmeda de las regiones montañosas de México (Rzedowski 2006).

#### Especie de estudio

Se consideró como plántula a los individuos de *A. religiosa* recién emergidos y que mantenían sus hojas embrionarias. Dado que *A. religiosa* no forma banco de semillas (Román, 2002; Leck *et al.*, 2008), se tuvo la certeza de que los individuos evaluados pertenecen a la cohorte de 2019. Algunos de ellos iniciaron la producción de hojas verdaderas sobre el primer conjunto de ramillas insertas en el primer nudo del tallo constituyendo su módulo fundamental de desarrollo. Los puntos de crecimiento forman un nudo a partir del cual se extienden las ramas plagiotrópicas y verticiladas (Hallé *et al.*, 1978) que al morir dejan cicatrices visibles sobre el tallo.

En México, los bosques de *A. religiosa* están compuestos de fragmentos de tamaños variables, que difieren en composición, estructura, fases de desarrollo y sucesión secundaria inducidas por perturbaciones naturales y antropógenas. Estas perturbaciones alteran su biomasa en pie, las funciones de sus ciclos biogeoquímicos e inducen una mayor vulnerabilidad a plagas, enfermedades, invasiones biológicas y contaminación del aire (Fenn *et al.*, 2002; Alvarado y Hernández, 2002; Montoya *et al.*, 2020). Las perturbaciones más importantes en estos bosques son la deforestación, la apertura de tierras para la agricultura, el pastoreo, los incendios y el aprovechamiento selectivo de madera (Madrigal, 1967; Challengier, 1998, Montoya *et al.*, 2020).

## Localización de parcelas de observación y supervivencia de plántulas

Se seleccionó un bosque fisionómicamente dominado por *A. religiosa* en el que se ubicaron 17 parcelas de 1 m<sup>2</sup> (1 m × 1 m), dentro de las cuales se marcaron y numeraron todas las plántulas de la cohorte 2019. A cada plántula se le midió altura (cm) y número de hojas embrionarias, presencia y número de nudos en su tallo y su emergencia sobre musgo o suelo directo. Las plántulas fueron recensadas y evaluadas a intervalos periódicos de tres semanas durante 12 meses, de agosto de 2019 a septiembre de 2020. El seguimiento de plántulas consistió en el registro de su estado: viva o muerta y los daños ocasionados por enemigos naturales (herbívoros, patógenos), así como daños debidos a caídas de ramas.

Los daños y causas potenciales de muerte se asociaron con: a) daño físico en los tallos, ramas, acículas y hojas embrionarias; b) ramoneo por consumo total o parcial de tallos, ramas, acículas y hojas embrionarias; c) pudrición por patógenos en tallos, acículas, ramas y hojas embrionarias; d) desecación por clorosis o amarillamiento total o parcial de tallos, ramas, acículas y hojas embrionarias, pérdida de turgencia en tallos o desecación total de la plántula y e) remoción de plántulas por extracción de musgo y hongos que realizan los habitantes de la zona.

Se determinaron las propiedades del suelo circundante a las parcelas. En cada parcela, se tomaron fotografías hemisféricas y con base en ellas, el dosel se estratificó en dos categorías: parcialmente cerrado y abierto. Los análisis se realizaron con los procedimientos LIFETEST y LIFEREG de SAS.

## Ambiente de luz

El ambiente lumínico en el que se desarrollan las plántulas se infirió a partir de fotografías hemisféricas del dosel, tomadas en el centro de cada parcela y a un metro de altura del suelo. Las fotografías se tomaron con una cámara Nikon Coolpix E-5400 provista con un lente hemisférico (Fisheye FC-E9, 0.2x), con un ángulo de visión de 180°, orientada hacia el norte magnético y nivelado horizontalmente. Las fotografías se analizaron con el software HemiView® (Rich *et al.*, 1999), configurado con los datos de latitud, longitud, altitud de las parcelas y fecha de cada fotograma. Las variables seleccionadas del ambiente de luz en cada parcela fueron: la proporción de aperturas en el dosel, proporción de luz directa (DIR), proporción de luz difusa (DIF) y proporción global del sitio (GLO).

Los porcentajes de apertura del dosel por parcela se agruparon en dos categorías: doseles parcialmente cerrados y doseles abiertos, cuyos valores de la media  $\pm$  error estándar y el rango (mínimo-máximo) fueron:  $0,211 \pm 0,003$  y  $0,191 - 0,221$ ; y  $0,228 \pm 0,002$  y  $0,222 - 0,236$ , respectivamente. Una prueba de Kruskal-Wallis (R Core Team 2020) mostró diferencias estadísticas significativas entre ambas condiciones de dosel.

### **Ambiente edáfico**

El ambiente edáfico en las parcelas se infirió a partir de las propiedades físicas y químicas de los suelos. Se recolectaron seis muestras compuestas de suelo de un kilogramo, tomadas a 20 cm de profundidad de cada esquina de las parcelas. Las propiedades de los suelos determinadas fueron: contenido de Nitrógeno (N; %), Fosforo (P; mg Kg-1), Potasio (K; Cmol Kg-1), Sodio (Na; Cmol Kg-1), sulfatos (SO<sub>4</sub>: Cmol Kg-1), pH, porcentaje de materia orgánica, densidad aparente (g cm-3) y porcentajes de arcilla, arena y limo. Las propiedades físicas y químicas se determinaron siguiendo la norma oficial mexicana, NOM-021-SEMARNAT-2000 (SEMARTAT 2000), que establece las especificaciones de fertilidad, salinidad y clasificación de suelos, estudio, muestreo y análisis.

Por otro lado, se colocaron seis registradores de datos (Extech Instruments, Modelo RHT10®) en el suelo a 10 cm de profundidad en las proximidades de las parcelas, configurados para registrar cada 10 minutos la humedad, temperatura mínima, máxima y punto de saturación. La información registrada se descargó en cada re-censo y de sus estadísticas se seleccionó valor mínimo, máximo, media y coeficientes de variación.

### **Análisis de sobrevivencia**

En virtud de que es prácticamente imposible conocer el tiempo exacto a la muerte (T) de las plántulas, los análisis de sobrevivencia y riesgos de mortalidad se hicieron conforme al diseño de censos por intervalos. La función de sobrevivencia empleada fue  $S(t) = \Pr \{T > t\} = 1 - F(t)$ , donde S(t) es la probabilidad de sobrevivencia de una plántula después de un tiempo t y F(t) es una función de distribución acumulada. El riesgo de mortalidad quedó definido por el número de plántulas muertas por intervalo de tiempo, dividido por el número de sobrevivientes en el mismo intervalo. La función de riesgo de mortalidad estima la probabilidad de que una plántula muera en un intervalo de tiempo dado (Allison, 2010).

Ambas funciones se obtuvieron con el método de tablas de vida de datos agrupados, mediante el procedimiento LIFETEST (SAS, 2012). LIFETEST estima las probabilidades de sobrevivencia y riesgo de mortalidad para datos que proporcionan: a) los límites de los intervalos, b) el número de muertes en cada intervalo y c) el número de casos censados en cada intervalo. Este procedimiento permite comparar pares de curvas de sobrevivencia y probar la hipótesis nula, de que las funciones de sobrevivencia son idénticas entre pares de curvas. Las curvas de sobrevivencia estimadas se compararon con la prueba de Mantel y Haenszel, conocida como Log-Rank (SAS, 2012).

El efecto de las variables intrínsecas y ambientales sobre los tiempos de sobrevivencia se estimaron con el modelo de máxima verosimilitud mediante el procedimiento LIFEREG (SAS, 2012). El procedimiento LIFEREG ajusta modelos paramétricos a datos de tiempos de muerte, que se registran bajo el esquema de censos en intervalos y permite analizar variables categóricas con valores múltiples. Los coeficientes de regresión y sus errores estándar ( $\beta_i \pm ee$ ) se mejoraron con base en una prueba de bondad de ajuste, para establecer el tipo de distribución de los tiempos en los que ocurrieron las muertes. La distribución observada se comparó con las distribuciones teóricas Weibull, Exponencial, Gamma, Log-logística y Log-normal. Los coeficientes de regresión de las variables ( $\beta_i$ ) se transformaron con la función  $100(e^{\beta_i} - 1)$ , que estima el porcentaje esperado de aumento en los tiempos de sobrevivencia por cada unidad de incremento en la covariable si el resto de las variables permanecen sin cambios (Allison, 2010).

## Resultados

Se registraron 1,565 plántulas distribuidas en las 17 parcelas. La densidad inicial promedio  $\pm$  una desviación estándar por parcela, fue de  $92,1 \pm 29,8$  plántulas  $m^{-2}$ , con variación de 41 a 126 plántulas  $m^{-2}$ . Las curvas de sobrevivencia de las plántulas fueron distintas entre las categorías de dosel (Log-Rank:  $\chi^2 = 8,35$ ,  $gl = 1$ ,  $p < 0,05$ ). La probabilidad de sobrevivencia bajo dosel abierto fue mayor ( $0,29 \pm 0,01$ ) que bajo dosel parcialmente cerrado ( $0,20 \pm 0,01$ ). Las variables con efectos positivos en los tiempos de sobrevivencia fueron: altura y número de nudos; proporción de luz directa, densidad aparente del suelo, temperatura media, humedad media y su coeficiente de variación, y la variación de la saturación de humedad del suelo.



En el último censo las plántulas vivas remanentes fueron 336 con densidad promedio por parcela de  $19,8 \pm 12,9$  plántulas  $m^2$ , con un mínimo de 1, un máximo de 54. El porcentaje final de plántulas muertas fue de 75,8 al 12<sup>avo</sup> mes de seguimiento. Los riesgos de mortalidad de plántulas bajo ambos doseles a lo largo del año fueron heterogéneos y diferentes entre ellos. En ambas condiciones de dosel, los riesgos máximos se observaron durante la estación seca del año, pero los riesgos fueron distintos dependiendo de la apertura del dosel.

### 3.1.3.3. Supervivencia de plántulas de *A. religiosa* bajo condición de dosel

Presentan como antecedentes la afirmación de que los factores físicos y biológicos involucrados en la sobrevivencia y riesgos de mortalidad de plántulas, son importantes para comprender los procesos naturales asociados con su establecimiento en un contexto espacio-temporal. Su objetivo consistió en identificar qué factores ambientales definen el nicho de regeneración de *Abies religiosa* con base en la supervivencia y riesgos de mortalidad de sus plántulas.

Métodos: Se establecieron 17 parcelas de 1  $m^2$  y en ellas, todas las plántulas de la cohorte 2019 fueron localizadas, numeradas y medidas en altura, número de hojas embrionarias, número de nudos y daños visibles y seguidas durante un año. Se determinaron las propiedades del suelo circundante a las parcelas. En cada parcela, se tomaron fotografías hemisféricas y con base en ellas, el dosel se estratificó en dos categorías: parcialmente cerrado y abierto. Los análisis se realizaron con los procedimientos LIFE-TEST y LIFEREG de SAS.

Las categorías de dosel seleccionadas a partir de los valores de la proporción de aperturas en el dosel fueron: dosel parcialmente cerrado ( $0,211 \pm 0,003$ ) y dosel abierto o con claros de luz en el dosel ( $0,228 \pm 0,002$ ), las cuales resultaron significativamente distintas (Kruskal-Wallis;  $\chi^2 = 14,075$ ; g.l. = 1;  $p \leq 0,05$ ). De las 1565 plántulas, 1017 se localizaron bajo dosel parcialmente cerrado y las restantes 548 bajo claros de luz.

El porcentaje de plántulas muertas bajo dosel parcialmente cerrado fue de 79 % ( $n = 803$ ) mientras que bajo claros de luz fue de 70,1 % ( $n = 384$ ). En cambio, el porcentaje de plántulas vivas a la semana 53 bajo dosel parcialmente cerrado fue de 17,9 % ( $n = 182$ ) mientras que bajo claros de luz fue de 28,1 % ( $n = 154$ ). Las causas posibles de mortalidad de plántulas de *A. religiosa* se muestran en el Cuadro

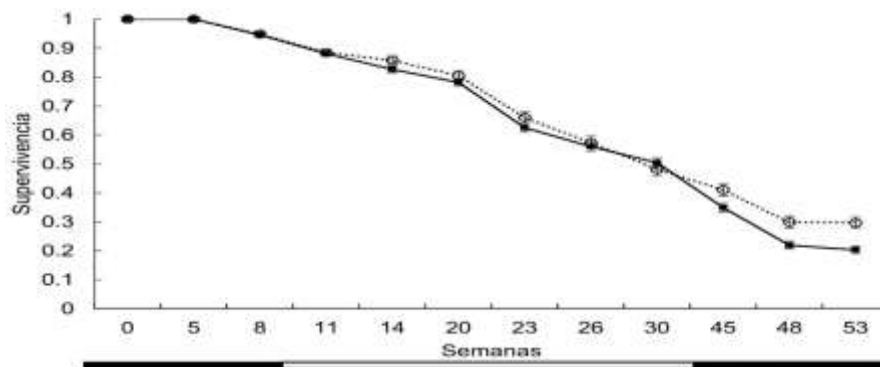
12. El daño ocasionado por desecación o clorosis fue la causa más importante asociada a la muerte de plántulas, seguido de los daños físicos y por pudrición o presencia de patógenos. El daño por remoción de musgo u hongos fue la causa de muerte de menor impacto durante las 53 semanas de seguimiento.

**Cuadro 12. Fluctuaciones en las tasas y posibles fuentes de mortalidad de plántulas de *Abies religiosa* durante 53 semanas en el Monte Tláloc.**

Censos (Sem)	Muertes (Acum)	Vivas (N)	Desapar (%)	Fuentes de mortalidad potencial				
				Daño fís. (%)	Herbiv (%)	hongos (%)	Desec. (%)	Extracc. musgo u hongos (%)
1	83	1481	1	1	17	58	7	0
(5)	(83)		(0,01)	(0,1)	(1,1)	(3,7)	(0,4)	(0)
2	101	1378	3	2	20	76	2	1
(8)	(184)		(0,02)	(0,1)	(1,3)	(4,9)	(0,1)	(0,1)
3	71	1308	3	14	24	19	14	0
(11)	(255)		(0,02)	(0,9)	(1,5)	(1,2)	(0,9)	(0)
4	74	1232	4	14	14	25	15	6
(14)	(329)		(0,03)	(0,9)	(0,9)	(1,6)	(1,0)	(0,4)
5	236	943	57	43	61	17	91	25
(20)	(565)		(3,6)	(2,7)	(3,9)	(1,1)	(5,8)	(1,6)
6	113	868	19	68	6	12	16	11
(23)	(678)		(1,2)	(4,3)	(0,4)	(0,8)	(1,0)	(0,7)
7	106	760	21	23	43	13	26	0
(26)	(784)		(1,3)	(1,5)	(2,7)	(0,8)	(1,7)	(0)
8	193	567	21	88	37	25	42	1
(30)	(977)		(1,3)	(5,6)	(2,4)	(1,6)	(2,7)	(0,1)
9	187	367	34	2	0	3	182	0
(45)	(1164)		(2,2)	(0,1)	(0)	(0,2)	(11,6)	(0)
10	16	347	38	4	0	0	12	0
(48)	(1180)		(2,4)	(0,3)	(0)	(0)	(0,8)	(0)
11	7	336	42	0	0	3	4	0
(53)	(1187)		(2,7)	(0)	(0)	(0,2)	(0,3)	(0)
Totales	1187	336	42	259	222	251	411	44
%	75,8%	21,5%	2,7	16,5%	14,2%	16%	26,3%	2,8%

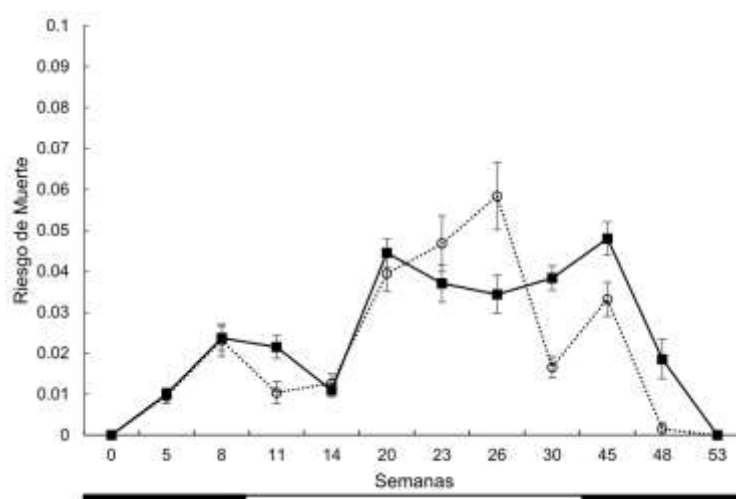
Fuente. Hernández et al. (2022).

Las curvas de supervivencia de plántulas bajo las dos condiciones de dosel se muestran en la Figura 11. Al final de las 53 semanas de seguimiento, la probabilidad de supervivencia de las plántulas bajo claros de luz fue de  $0,29 \pm 0,019$ , mientras que bajo dosel parcialmente cerrado fue de  $0,20 \pm 0,012$ . Las curvas de supervivencia bajo las dos condiciones de dosel fueron distintas (Log-Rank:  $\chi^2 = 8,35$ ,  $gl = 1$ ,  $p = 0,0038$ ).



**Figura 11. Cambios en la supervivencia de plántulas ( $\pm$  error estándar) de *Abies religiosa* bajo dosel parcialmente cerrado (cuadros) y claros de luz (círculos). Barra negra representa la estación lluviosa y barra blanca la estación seca del año. Fuente. Hernández et al. (2022).**

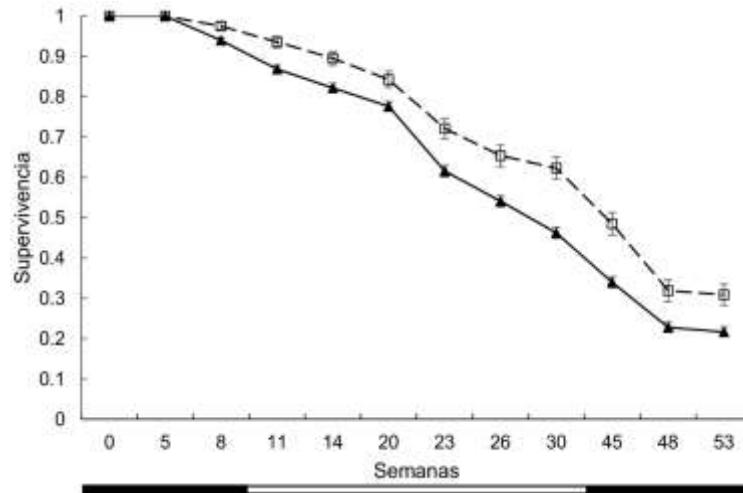
Bajo condiciones de dosel parcialmente cerrado se tuvieron las más altas probabilidades de riesgo de muerte (Figura 12). Sin embargo, durante los intervalos comprendidos en las primeras 13 y las últimas ocho semanas de seguimiento (45 a 53), los riesgos de mortalidad fueron los más bajos ( $< 0,03$ ). En cambio, hubo 30 semanas (de la 14 a la 45) en las que los riesgos de mortalidad se incrementaron, llegando a sus máximos durante las semanas 26 bajo claros de luz y en la 20 en el dosel parcialmente cerrado (Figura 12).



**Figura 12. Riesgos de mortalidad ( $\pm$  error estándar) de plántulas bajo condiciones de claros de luz (círculos) y dosel parcialmente cerrado (cuadros). Barra negra representa la estación lluviosa y barra blanca la estación seca del año. Fuente. Hernández et al. (2022).**

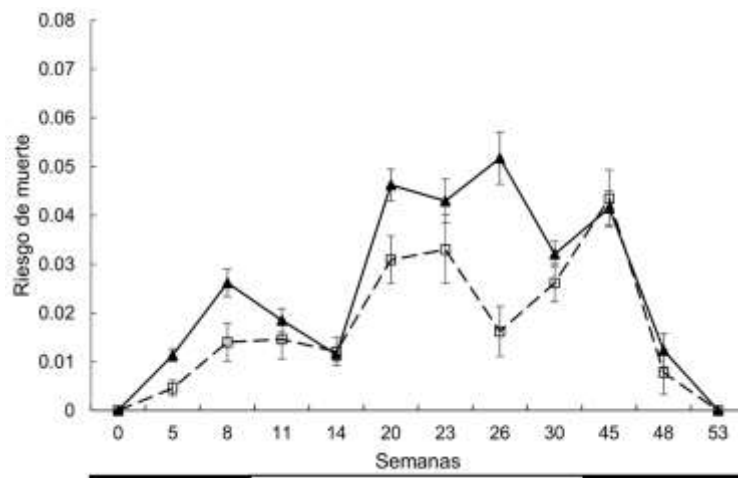
### 3.1.3.3.1. Supervivencia de plántulas sobre musgo y suelo directo.

La fracción de plántulas que emergieron sobre suelo directo y musgo fue de 79,3 % ( $n = 1241$ ) y 20,7% ( $n = 324$ ), respectivamente. A la semana 53, la probabilidad de supervivencia de plántulas sobre suelo directo fue de  $0,309 \pm 0,026$ , mientras que sobre musgo fue de  $0,217 \pm 0,011$ . Las curvas de supervivencia de plántulas emergidas en ambas condiciones (Figura 13) fueron distintas (Log-Rank:  $\chi^2 = 20,63$ ,  $gl = 1$ ,  $p < 0,0001$ ), por lo que la supervivencia de plántulas sobre musgo es una condición de mayor riesgo para su establecimiento.



**Figura 13. Tasas de supervivencia de plántulas ( $\pm$  error estándar) que emergieron sobre suelo directo (cuadros) y sobre musgo (triángulos). Barra negra representa la estación lluviosa y barra blanca la estación seca del año. Fuente. Hernández et al. (2022).**

En la Figura 14 se evidencia que los riesgos de muerte de las plántulas fueron consistentemente mayores en aquellas que emergieron sobre musgo que sobre suelo directo. Los riesgos más elevados de mortalidad se presentaron entre las semanas 20 a 45, con un máximo en la semana 26 en plántulas sobre musgo y en la semana 45 en suelo directo. Los riesgos más bajos de mortalidad se presentaron después de la semana 45 y durante las primeras 14 semanas de censos.



**Figura 14. Fluctuación en los riesgos de mortalidad de las plántulas ( $\pm$  error estándar) de *Abies religiosa* sobre musgo (triángulos) y sobre suelo directo (cuadros). Barra negra representa la estación lluviosa y barra blanca la estación seca del año. Fuente. Hernández et al. (2022).**

### 3.1.3.3.2. Influencia de covariables ambientales en la supervivencia de plántulas.

El modelo exponencial fue seleccionado con base en la magnitud del valor de log-likelihood (-1371,2). De acuerdo con Allison (2010), las menores magnitudes de este valor para cada modelo corresponden a mejores ajustes de los datos a las distribuciones comparadas. El Cuadro 13 muestra el análisis de regresión y las variables con efectos significativos en los tiempos de supervivencia de las plántulas de *A. religiosa*. Las variables con efectos positivos fueron: altura y número de nudos, mientras aquella con efectos negativos fue el número de hojas embrionarias. En cuanto al ambiente edáfico, la densidad aparente y las variables derivadas de temperatura y humedad del suelo presentaron efectos en los tiempos de supervivencia. Para el ambiente lumínico, la variable con efecto significativo positivo fue la proporción de luz directa (DIR), mientras que la proporción de luz global (GLO) estimó efectos negativos.

La transformación  $100(e^{\beta_i} - 1)$  predice que el aumento de un centímetro en la altura o adición de un nudo en el tallo, incrementa en 2,7 y 17,2 % los tiempos de supervivencia respectivamente. Por el contrario, el incremento en el número de hojas embrionarias disminuye los tiempos de supervivencia de las plántulas hasta en un 4,3 %.

Por otra parte, las covariables del ambiente edáfico incrementan considerablemente los tiempos de supervivencia de las plántulas de *A. religiosa*. Un incremento en una unidad de la densidad aparente del suelo incrementa 199 % los tiempos de supervivencia de las plántulas. Por su parte, un aumento de 1 °C en la temperatura media del suelo ( $T_{med}$ ) extiende un 68,5 % los tiempos de supervivencia, mientras que la humedad media y su variabilidad amplían los tiempos de supervivencia de las plántulas en un 12,6 y 11,1 %, respectivamente. En cambio, un incremento porcentual de uno en la humedad máxima del suelo ( $H_{max}$ ) reduce un 4,7 % la supervivencia de plántulas.

Finalmente, la saturación promedio de humedad en los suelos ( $D_{med}$ ) y la máxima ( $D_{max}$ ) reducen significativamente los tiempos de supervivencia hasta en un 41,6 %, mientras que la variabilidad de este ( $D_{cv}$ ) lo incrementa en un 5,6 y 1,8 %, respectivamente.

**Cuadro 13. Coeficientes de regresión ( $\beta_i \pm$  error estándar) de covariables cualitativas y cuantitativas, para los tiempos de supervivencia de plántulas de *Abies religiosa***

Variables	G.L.	Coeficientes ( $\beta \pm ee$ )	$\chi^2$	p	
Variables intrínsecas					
Altura	1	0.0271 $\pm$ 0.0131	4.29	0.0384	
Hojas embrionarias	1	-0.0442 $\pm$ 0.0066	44.43	<.0001	
Número de nudos	1	0.1589 $\pm$ 0.0214	55.37	<.0001	
Daños visibles					
Tipo de daño	5				
Físico	1	-3.7404 $\pm$ 1703.699	0	0.9982	
Herbivoría	1	-3.8799 $\pm$ 1703.699	0	0.9982	
Pudrición	1	-4.241 $\pm$ 1703.699	0	0.998	
Desecación	1	-3.5567 $\pm$ 1703.699	0	0.9983	
Remoción	1	-3.7747 $\pm$ 1703.699	0	0.9982	
Sin daño aparente	1	0	0	0.9982	
Ambiente de luz					
VisSky	1	-3.4933 $\pm$ 2.6488	1.74	0.1872	
DIF	1	12.1786 $\pm$ 6.446	3.57	0.0588	
DIR	1	99.6181 $\pm$ 49.8326	4	0.0456	
GLO	1	-109.309 $\pm$ 55.3059	3.91	0.0481	
Ambiente edáfico					
N	1	0	.	.	
P	1	0	.	.	
K	1	0	.	.	
Arcilla	1	-0.0001 $\pm$ 0.0155	0	0.9962	
Arena	1	0.0084 $\pm$ 0.0155	0.3	0.5869	
Limo	1	0	-	-	
pH	1	-0.1144 $\pm$ 0.221	0.27	0.6046	
Dap	1	1.0968 $\pm$ 0.3124	12.33	0.0004	
MO	1	0.0354 $\pm$ 0.0422	0.7	0.4024	
Na	1	0	-	-	
SO <sub>4</sub>	1	0	-	-	
Temperatura del suelo	T <sub>med</sub>	1	0.5221 $\pm$ 0.1371	14.5	0.0001
	T <sub>min</sub>	1	0.004 $\pm$ 0.0099	0.16	0.6884
	T <sub>max</sub>	1	-0.004 $\pm$ 0.0141	0.08	0.7783
	T <sub>CV</sub>	1	-0.0124 $\pm$ 0.0082	2.27	0.1322
Humedad del suelo	H <sub>med</sub>	1	0.1186 $\pm$ 0.0325	13.29	0.0003
	H <sub>min</sub>	1	0.0028 $\pm$ 0.0029	0.9	0.3424
	H <sub>max</sub>	1	-0.0491 $\pm$ 0.0163	9.04	0.0026
	H <sub>CV</sub>	1	0.1059 $\pm$ 0.0203	27.31	<.0001
Punto de Saturación de humedad	D <sub>med</sub>	1	-0.5372 $\pm$ 0.1306	16.91	<.0001
	D <sub>min</sub>	1	-0.0062 $\pm$ 0.008	0.61	0.4347
	D <sub>max</sub>	1	-0.019 $\pm$ 0.007	7.3	0.0069
	D <sub>CV</sub>	1	0.0549 $\pm$ 0.0097	32.31	<.0001

Fuente. Hernández et al. (2022).

### **3.1.3.3.3. Discusión**

#### **3.1.3.3.3.1. Variables intrínsecas de las plántulas**

La altura de las plántulas y el número de nudos incidieron positivamente en los tiempos de supervivencia. El incremento en altura en la fase inicial de emergencia de las plántulas está directamente asociado con la movilización de recursos desde la semilla a la producción de raíz, tallo, hojas embrionarias, hasta el desarrollo suficiente de tejido fotosintético y raíces. El crecimiento, la producción del primer nudo en el tallo y la generación de hojas verdaderas son buenos indicadores del desempeño de las plántulas y les facultan para enfrentar las fuertes presiones selectivas ejercidas sobre ellas. La asignación de energía al crecimiento inicial es crítica para su supervivencia; los recursos de la semilla se agotan con el crecimiento en un tiempo relativamente corto. La producción de hojas embrionarias tiende a favorecer la ganancia de carbono, pero no son suficientes a medida que la plántula continúa su crecimiento.

La transición de plántulas con hojas embrionarias a plántulas con hojas verdaderas producidas en las ramillas del primer nudo, es crucial para su supervivencia. Estos aspectos morfológicos influyen las estrategias de asignación de recursos y favorecen el nicho de regeneración de la especie. Kitajima y Myers (2008), mencionan que estas estrategias ecofisiológicas favorecen la supervivencia mediada por un balance positivo de carbono en plántulas. La producción de hojas verdaderas favorece el balance positivo de carbono incrementando los tiempos de supervivencia de la plántula a la vez que se independiza de los recursos de la semilla. En cambio, la asignación de energía a la producción de un mayor número de hojas embrionarias reduce las probabilidades de supervivencia de las plántulas. Es factible que, si los recursos finitos de la semilla se asignan a la producción de un mayor número de hojas embrionarias, se hará a costa del crecimiento del tallo, y de la producción de ramillas y hojas verdaderas.

La diferencia entre asignar energía a la producción inicial de hojas embrionarias sobre el tallo o al crecimiento del tallo y producción de ramillas con hojas verdaderas hace la diferencia entre la supervivencia y muerte de las plántulas de oyamel. Moles y Leishman (2008), encontraron que un tallo de mayor tamaño y hojas verdaderas debe proporcionar un mejor acceso al agua, a los minerales y a la luz, que una plántula con solo hojas embrionarias.



Las plántulas tienen probabilidades de supervivencia más altas bajo claros de luz que bajo dosel parcialmente cerrado. La supervivencia de plántulas está directamente relacionada con la proporción de radiación solar directa (DIR) que penetra a través de los claros de luz e incide sobre ellas. Las plántulas bajo claros de luz sobreviven mejor que bajo dosel parcialmente cerrado ya que están expuestas a mayores niveles de radiación solar directa. La condición de claros de luz integradas a las covariables edáficas y su variabilidad impactan positivamente los tiempos de supervivencia de las plántulas de *A. religiosa*.

La contribución relativa de ellos en orden descendente es: proporción de luz directa, densidad aparente del suelo, temperatura media, humedad media y su coeficiente de variación, y la variación del punto de saturación. Nuestros resultados sugieren que las aperturas de claros de luz y la consecuente proporción de luz directa que incide sobre las plántulas de *A. religiosa*, activan su fotosíntesis, coopera a su mantenimiento y mejoran su supervivencia.

Johnson y Smith (2005), estudiaron la supervivencia y mortalidad de plántulas de *Abies fraseri* (Pursh) Poiret en función de la fotosíntesis y sus relaciones hídricas bajo condiciones de aperturas de dosel o dosel parchado y dosel abierto o sin dosel. Ellos concluyen que la ganancia de carbono fue mayor en plántulas bajo condiciones de dosel abierto que bajo dosel parchado. Las plántulas de *A. fraseri* en dosel abierto asimilaron de 3,3 a 4,5 veces más carbón que bajo doseles parchados. Las plántulas de *A. fraseri* redujeron su ganancia de carbono debido, principalmente, a limitaciones de incidencia de luz solar, lo que ocasionó una mortalidad de hasta 98 % bajo dosel cerrado versus 79 % bajo doseles abiertos (Johnson y Smith, 2005).

Por otra parte, los trabajos de Marañón *et al.* (2004) sobre el nicho de regeneración en bosques mediterráneos evidenciaron que la fotosíntesis de las plántulas se activa más efectivamente bajo condiciones de aperturas en el dosel que bajo condiciones de dosel cerrado. Además, los resultados en diversos trabajos en ecosistemas de templados revelan que tanto las aperturas en el dosel como la incidencia de radiación solar directa incrementan la heterogeneidad tanto de la temperatura superficial como de humedad relativa y propiedades físicas y químicas de los suelos (Scharenbroch y Bockheim, 2007, He *et al.*, 2015, Matiz *et al.*, 2019).

La densidad aparente del suelo es susceptible de generar condiciones adecuadas para el desarrollo de las primeras raíces de las plántulas de *A. religiosa*. Sus bajos valores (media = 1,05 g cm<sup>-3</sup>, rango 0,89 – 1,16) revelan que son suelos porosos, bien drenados y no compactados, lo que beneficia el crecimiento y configuración del sistema radicular de las plántulas y el suministro de agua y nutrientes a ellas. La densidad aparente de estos suelos, son muy similares a la de la Reserva de la Biosfera Volcane ( $< 1$  g cm<sup>-3</sup>) encontrados por Cruz *et al.* (2020). He *et al.* (2015) encontraron que las propiedades físicas del suelo, incluidos sus espacios porosos, determinantes de la densidad aparente, se mejoran bajo la influencia de pequeños claros de luz en el dosel del bosque. Estos autores, encontraron que la humedad del suelo, la capacidad de retención de agua y su absorción por las plantas, mejoró con la composición relativa de la porosidad del suelo. En nuestro trabajo, la baja densidad aparente de los suelos y su asociación con las aperturas en el dosel del bosque, promueven el nicho de regeneración de *A. religiosa*. Los resultados de He *et al.* (2015), respaldan congruentemente a los nuestros, y apoyan la idea de que la densidad aparente del suelo, aunque es un componente hasta ahora poco estudiado, contribuye a la generación de micrositios idóneos y adecuados para la regeneración natural de *A. religiosa*.

Es evidente que tanto la temperatura superficial como la humedad del suelo son modificadas por la radiación solar que penetra a través de las aperturas del dosel, y que estas prolongan también los tiempos de supervivencia de las plántulas. Sin embargo, incrementos abruptos en ambas variables pueden tener resultados desfavorables en la supervivencia de las plántulas, y en tal caso ambas resultarán en filtros ecológicos para su reclutamiento.

La sobrevivencia de plántulas emergidas sobre suelo directo fue considerablemente mayor que las emergidas sobre musgo. La desecación de plántulas es la *Fuente* principal de mortalidad, lo que sugiere un suministro deficiente o una ineficaz captura de agua por las plántulas. En contraste, la saturación máxima (100 %) de humedad en el suelo tuvo efectos negativos sobre las plántulas, las cuales murieron al ser atacadas y debilitadas por hongos patógenos, la segunda fuente de mortalidad de plántulas, lo que sugiere que estos enemigos naturales circunscriben la regeneración natural de *A. religiosa* a las condiciones propiciadas por los claros de luz en el dosel. Los resultados sugieren que la disponibilidad de agua en el suelo para las plántulas es crítica, ya que su exceso favorece el ataque y proliferación de hongos, mientras que su deficiencia promueve su muerte por desecación. Los riesgos de muerte tienden

a ser mayores bajo condiciones de dosel parcialmente cerrado. Bajo estas condiciones, las plántulas están más expuestas a saturación prolongada en la humedad del suelo y al daño físico por caída de ramas, así como a un ambiente de luz reducido.

La sobrevivencia de plántulas de la cohorte de *A. religiosa* fue mayor bajo condiciones de claros en el dosel que bajo dosel parcialmente cerrado. Mientras que los riesgos de mortalidad durante la estación seca del año y bajo condiciones de dosel parcialmente cerrado fueron altos y sostenidos; en cambio, bajo dosel abierto los riesgos de mortalidad fueron heterogéneos e inconsistentemente sostenidos.

Los cambios en las covariables ambientales, mediados por las aperturas en el dosel del bosque favorecen el crecimiento en altura de las plántulas y la producción de ramillas y hojas verdaderas, y estos a su vez los tiempos de supervivencia. La importancia individual de covariables que contribuyen proporcional y favorablemente a la supervivencia de las plántulas fue: proporción de luz directa, densidad aparente del suelo, temperatura media, humedad media y su coeficiente de variación. En conjunto estas covariables están íntimamente relacionadas entre sí y definen el nicho de regeneración de *A. religiosa* desde sus fases iniciales de desarrollo y durante el primer año de su ciclo de vida.

#### 3.1.3.3.4. Conclusiones

Estas covariables definen el nicho de regeneración de *A. religiosa*. Los análisis de la sobrevivencia y riesgos de mortalidad de plántulas son de importancia prioritaria para comprender su destino y el curso potencial de regeneración natural de los bosques. La supervivencia de plántulas durante los primeros días y semanas después de su emergencia es un cuello de botella de extraordinaria importancia en el ciclo de vida de las plantas (Leck *et al.* 2008). Durante la fase de plántula, ocurren los más altos riesgos de muerte y elevadas tasas de mortalidad, por lo que es un estadio determinante en los procesos de regeneración natural de los bosques (Kitajima y Fenner, 2000).

La noción de nicho de regeneración fue propuesta por Grubb (1977), para referirse a todos los requerimientos necesarios del ambiente físico y biológico asociados con la germinación de semillas, la supervivencia y establecimiento de las plántulas. Esta noción es una expresión íntimamente relacionada con aquellos procesos ecológicos idóneos para la germinación, el establecimiento y reclutamiento de individuos juveniles. Una vez germinada la semilla, el crecimiento, mantenimiento y defensa de las

plántulas dependen de los recursos finitos que contiene la semilla, hasta el agotamiento completo de ellos. La independencia de la plántula de los recursos de la semilla significa que deberá enfrentar con éxito los múltiples factores de riesgo o morir. Por esta razón es pertinente valorar la importancia de los factores ambientales y biológicos que ponen en riesgo la supervivencia de las plántulas.

En particular, las perturbaciones ocurridas en el dosel del bosque ocasionan aperturas de distintos tamaños que impactan diferencialmente, en espacio y tiempo, el ambiente físico y la regeneración natural. Las perturbaciones del dosel de tamaño pequeño (< 100 m<sup>2</sup>), impactan las condiciones ambientales del sotobosque, particularmente a nivel del piso forestal, así como las propiedades físicas y químicas de los suelos (He *et al.*, 2015). En el estrato rasante del suelo forestal, la radiación solar (directa, difusa y global), la temperatura y contenido del agua del suelo son críticos para el establecimiento de las plántulas de *Abies religiosa* (Kunth) Schlttdl. & Cham., y la regeneración natural del bosque.

Las variaciones de las propiedades del suelo bajo la influencia de los claros de luz tienen un peso relativo crítico en la germinación de semillas, en el establecimiento de plántulas, en el reclutamiento de individuos y de la estructura y dinámica de los bosques (Canham y Marks, 1985; Platt y Streng, 1989; Yamamoto, 2000). Por lo tanto, un estudio de las propiedades físicas y químicas de los suelos en los claros del bosque puede proporcionar una mejor comprensión de la capacidad del suelo en el suministro de agua y nutrientes a las plantas.

En los análisis que intentan explicar las bajas tasas de supervivencia de plántulas se han empleado predictores tanto biológicos como ambientales con los cuales presumiblemente están correlacionados. Entre los predictores más sobresalientes están: la cobertura vegetal (Beckage *et al.*, 2005), disturbios por incendios inducidos por descargas eléctricas (Ángeles y López, 2009), distancia a los árboles progenitores (Packer y Clay, 2003), el microclima (Blennow y Lindkvist, 2000), la humedad en el suelo (Rey y Alcántara, 2000; Caspersen y Kobe, 2001; Arrieta y Suárez, 2005), las aperturas de claros de luz en dosel (Lara *et al.* 2009), la disponibilidad de la luz y las afectaciones de las propiedades físicas y químicas de los suelos (Baraloto *et al.*, 2005; Curt *et al.*, 2005; Kunstler *et al.*, 2005, He *et al.*, 2015), la herbivoría (Osunkjaya *et al.*, 1992), la capacidad fotosintética de las plántulas (Johnson y Smith, 2005), el tamaño y crecimiento de las plántulas (Cruz y López, 2004) y la masa de las semillas (Leishman y Westoby, 1994).

En esta investigación, se plantea un examen detallado de las causas probables de mortalidad y los factores de riesgo que se traducen en probabilidades variables de supervivencia de las plántulas de *A. religiosa*. El análisis de las causas de mortalidad de las plántulas es un componente valioso para una mejor comprensión del nicho de regeneración de *A. religiosa*. En este contexto, la hipótesis subyacente es que, en el sotobosque, el ambiente de luz y ciertas propiedades fisicoquímicas de los suelos definen el nicho de regeneración de *A. religiosa*. El ambiente de luz ocasionado por aperturas de tamaño variable en el dosel y los suelos bajo la influencia de ellos, mejoran las probabilidades de supervivencia de plántulas reduciendo sus riesgos de mortalidad.

Se derivan las siguientes interrogantes: ¿qué parámetros del ambiente de luz son de mayor importancia relativa en la supervivencia de las plántulas de *A. religiosa*?, ¿qué factores físicos y químicos de los suelos son los más importantes en la supervivencia de las plántulas? y ¿qué variables intrínsecas de las plántulas de *A. religiosa* son más relevantes para su supervivencia durante las primeras semanas de vida, y para su establecimiento futuro? Para responder a estas preguntas, se plantearon los siguientes objetivos: 1) analizar la supervivencia y riesgos de mortalidad de una cohorte de plántulas de *A. religiosa* y, 2) evaluar la importancia relativa de las propiedades edáficas, el ambiente de luz y los daños ocasionados por entes biológicos y físicos, sobre la mortalidad y supervivencia de plántulas. Para lograrlo, se optó por un análisis demográfico y seguimiento periódico del destino de una cohorte de las plántulas de *A. religiosa*, tal que aporte a una mejor comprensión de su nicho de regeneración (Grubb, 1977).

#### **Referencias bibliográficas del artículo "Factores físicos y biológicos involucrados en la supervivencia y riesgos de mortalidad de plántulas de *A. religiosa* bajo regeneración natural.**

Alvarado, D. y Hernández, T. (2002). Decline of sacred fir in the Desierto de los Leones National Park. In: Fenn M, de Bauer LI, Hernández-Tejeda T, eds. Urban air pollution and forests: resources at risk in the Mexico City Air Basin. New York: Springer-Verlag. pp. 243-260. DOI: [https://doi.org/10.1007/978-0-387-22520-3\\_10](https://doi.org/10.1007/978-0-387-22520-3_10)

Allison P. 2010. Survival analysis using the SAS® System: A practical guide. USA, Cary, North Carolina: SAS Publishing. ISBN 978-1-59994-640-5

- Ángeles, E. y López, L. (2009). Supervivencia de una cohorte de plántulas de *Abies religiosa* bajo diferentes condiciones postincendio. *Boletín de la Sociedad Botánica de México* 84: 25-33. DOI: <https://doi.org/10.17129/botsci.2289>
- Arrieta, S. y Suárez, F. (2005). Spatial patterns of seedling emergence and survival as a critical phase in holly (*Ilex aquifolium* L.) woodland recruitment in Central Spain. *Forest Ecology and Management* 205: 267-282. DOI: <https://doi.org/10.1016/j.foreco.2004.10.009>
- Baraloto, C., Goldberg, D. y Bonal, D. (2005). Performance trade-offs among tropical tree seedlings in contrasting microhabitats. *Ecology* 86: 2461-2472. DOI: <https://doi.org/10.1890/04-1956>
- Beckage, B., Lavine, M. y Clark, J. (2005). Survival of tree seedlings across space and time: estimates from long-term count data. *Journal of Ecology* 93: 1177-1184. DOI: <https://doi.org/10.1111/j.1365-2745.2005.01053.x>
- Blennow, K. y Lindkvist, L. (2000). Models of low temperature and high irradiance and their application to explaining the risk of seedling mortality. *Forest Ecology and Management* 135: 289-301. DOI: [https://doi.org/10.1016/S0378-1127\(00\)00287-5](https://doi.org/10.1016/S0378-1127(00)00287-5)
- Canham, C. y Marks, P. (1985). The response of woody plants to disturbance patterns of establishment and growth. In: Pickett TA, White PS. eds. *The Ecology of Natural Disturbance and Patch Dynamics*. Londres: Academic Press, Inc. pp 197-216. DOI: <https://doi.org/10.1016/B978-0-12-554520-4.50016-2>
- Caspersen, J. y Kobe, R. (2001). Interspecific variation in sapling mortality in relation to growth and soil moisture. *Oikos* 92: 160-168. DOI: <https://doi.org/10.1034/j.1600-0706.2001.920119.x>
- Challenger, A. (1998). *Utilización y conservación de los ecosistemas terrestres de México. Pasado, presente y futuro*. México, DF: Comisión Nacional para el Conocimiento y Uso de la Biodiversidad. ISBN 970-9000-02-0.
- Cruz, G., Guerra, E., Valderrábano, J. y Campo, J. (2020). Indicadores de calidad de suelos en bosques templados de la Reserva de la Biosfera los Volcanes, México. *Revista Terra Latinoamericana* 38: 781-793. DOI: <https://doi.org/10.28940/terra.v38i4.421>
- Cruz, J. y López, L. (2004). Demography of the seedling bank of *Manilkara zapota* (L.) Royen, in a subtropical rain forest of Mexico. *Plant Ecology* 172: 227-235. DOI: <https://doi.org/10.1023/B:VEGE.0000026341.46440.f5>
- Curt, T., Coll, L., Prévosto, B., Balandier, P. y Kunstler, G. (2005). Plasticity in growth, biomass allocation and root morphology in beech seedlings as induced by irradiance and herbaceous competition. *Annals of Forest Science* 62: 51-60. DOI: <https://doi.org/10.1051/forest:2004092>
- Fenn, de Bauer M., Hernández, T. eds. 2002. *Urban air pollution and forests: resources at risk in the Mexico City Air Basin*. New York: Springer-Verlag. <https://doi.org/10.1007/978-0-387-22520-3>

- Grubb, P. (1977). The maintenance of species richness in plant communities: the importance of the regeneration niche. *Bio-logical Reviews* 52: 107-145. DOI: <https://doi.org/10.1111/j.1469-185X.1977.tb01347.x>
- Hallé, F., Oldeman, R. y Tomlinson P. (1978). *Tropical Trees and Forests. An architectural analysis.* Berlin-Heidelberg-New York: Springer-Verlag. pp. 441. ISBN: 978-3-642-81190-6.
- He, Z., Liu, J., Su, S., Zheng, S., Xu, D., Wu, Z., Hong, W. y Wang, J. (2015). Effects of Forest Gaps on Soil Properties in *Castanopsis kawakamii* Nature Forest. *Plos one* 10: e0141203. DOI: <https://doi.org/10.1371/journal.pone.0141203>
- Johnson, D. y Smith, W. (2005). Refugial forests of the southern Appalachians: photosynthesis and survival in current-year *Abies fraseri* seedlings. *Tree Physiology* 25: 1379-1387. DOI: <https://doi.org/10.1093/treephys/25.11.1379>
- Kitajima, K. y Fenner M. (2000). Ecology of Seedling Regeneration. In: Fenner M, ed. *Seeds: The Ecology of Regeneration in Plant Communities.* London: CAB Publishing, pp 331-359. ISBN: 9780851994321. <http://dx.doi.org/10.1079/9780851994321.0331>
- Kitajima, K. y Myers, J. (2008). Seedling ecophysiology: strategies toward achievement of positive net carbon balance. In: Leck MA, Parker VT, Simpson RL. eds. *Seedling Ecology and Evolution.* Cambridge: Cambridge University Press. Pp. 172-188. ISBN 978-0-521-87305-5
- Kunstler, G., Curt, T., Bouchaud, M. y Lepart, J. (2005). Growth, mortality, and morphological response of European beech and downy oak along a light gradient in sub-Mediterranean forest. *Canadian Journal of Forest Research* 35: 1657-1668. DOI: <https://doi.org/10.1139/x05-097>
- Lara, R., Sánchez, L. y Corral, J. (2009). Regeneración de *Abies religiosa* en claros del dosel versus sotobosque, Parque Nacional Cofre de Perote, México. *Agrociencia* 43: 739-747.
- Leck, M., Parker, V., Simpson, R. (2008). *Seedling Ecology and Evolution.* Cambridge: Cambridge University Press. ISBN: 978-0-521-87305-5
- Leishman, M. y Westoby, M. (1994). The role of large seed size in shaded conditions: experimental evidence. *Functional Eco-logy* 8: 205-214. DOI: <https://doi.org/10.2307/2389903>
- Madrigal, S. (1967). Contribución al conocimiento de la ecología de los bosques de oyamel *Abies religiosa* (HBK) Schl. et Cham en el Valle de México. *Boletín técnico* 18, México, DF: Instituto Nacional de Investigaciones Forestales.
- Marañón, T., Camarero, J., Castro, J., Díaz, M., Espelta, J., Hampe, A., Jordano, P., Valladares, F., Verdú, M. y Zamora, R. (2004). Heterogeneidad ambiental y nicho de regeneración. In: Valladares F, ed. *Ecología del Bosque Mediterráneo en un Mundo Cambiante.* Madrid: Ministerio de Medio Ambiente, pp. 69-49. ISBN: 84-8014-552-8



- Matiz, J., Rodríguez, G. y Alfaro, C. (2019). Modelos de temperatura del suelo a partir de sondeos superficiales y sensores remotos para el área geotérmica de Paipa, Boyacá-Colombia. *Boletín de Geología* 41: 71-88. DOI: <https://doi.org/10.18273/revbol.v41n2-2019004>
- Moles, A. y Leishman, M. (2008). The seedling as part of a plant's life history strategy. In: Leck MA, Parker VT, Simpson RL. eds. *Seedling Ecology and Evolution*. Cambridge: Cambridge University Press. pp. 189-1238. ISBN 978-0-521-87305-5
- Montoya, E., Guzmán, R. y López, L. (2020). Fragmentation dynamics in an *Abies religiosa* forest of central Mexico. *Canadian Journal of Forest Research* 50: 680-688. DOI: <https://doi.org/10.1139/cjfr-2019-0235>
- Osunkjaya, O., Ash, J., Hopkins, M. y Graham, A. (1992). Factors affecting survival of tree seedlings in North Queensland rainforests. *Oecologia* 91: 569-578. DOI: <https://doi.org/10.1007/BF00650333>
- Packer, A. y Clay, K. (2003). Soil pathogens and *Prunus serotina* seedling and sapling growth near conspecific trees. *Ecology* 84: 108-119. DOI: [https://doi.org/10.1890/0012-9658\(2003\)084\[0108:SPAPSS\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2003)084[0108:SPAPSS]2.0.CO;2)
- Platt, W. y Streng, D. (1989). Gaps in forest ecology. Special Feature: Treefall Gaps and Forest Dynamics. *Ecology* 70: 535-576. DOI: <https://doi.org/10.2307/1940194>
- R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. <http://www.R-project.org/>
- Rey, P. y Alcántara, J. (2000). Recruitment dynamics of a fleshy-fruited plant (*Olea europaea*): connecting patterns of seed dispersal to seedling establishment. *Journal of Ecology* 88: 622-633. DOI: <https://doi.org/10.1046/j.1365-2745.2000.00472.x>
- Rich, P., Wood, J., Vieglais, D., Burek, K. y Webb, N. (1999). Hemiview Manual, version 2.1. Delta-T Devices, Ltd. [https://delta-t.co.uk/wp-content/uploads/2020/02/HemiView-User-Manual-\\_2.1.pdf](https://delta-t.co.uk/wp-content/uploads/2020/02/HemiView-User-Manual-_2.1.pdf) (accessed August 20, 2020)
- Román R. 2002. Ecología de semillas y plántulas de *Abies religiosa* (HBK) Schl. et Cham. en el parque nacional "Cumbres del Ajusco". BSc. Thesis. Universidad Nacional Autónoma de México.
- Rzedowski, J. (2006). Bosque de Coníferas. In: Rzedowski J, ed. *Vegetación de México*. Ciudad de México, México: 1ra. Edición digital, Comisión Nacional para el Conocimiento y Uso de la Biodiversidad. Págs. 295-327. ISBN: 9681800028.
- Sánchez, A. y López, L. (2003). Clasificación y ordenación de la vegetación del norte de la Sierra Nevada, a lo largo de un gradiente altitudinal. *Anales del Instituto de Biología, Universidad Nacional Autónoma de México, Serie Botánica* 74: 47-71.
- SAS. 2012. *Statistical Analysis System*. V.9.0 ed. SAS. Institute. Inc. N.C. USA: Cary.



Scharenbroch, B. y Bockheim, J. (2007). Impacts of forest gaps on soil properties and processes in old growth northern hardwood-hemlock forests. *Plant Soil* 294: 219-233. DOI: <https://doi.org/10.1007/s11104-007-9248-y>

SEMARNAT [Secretaría de Medio Ambiente y Recursos Naturales]. 2020. Anuario Estadístico de la Producción Forestal 2017. <http://dsiappsdev.semarnat.gob.mx/datos/portal/publicaciones/2020/2017.pdf> (accessed November 16, 2020).

SEMARNAT. (2000). Norma Oficial Mexicana NOM-021-RECNAT-2000, que establece las especificaciones de fertilidad, salinidad y clasificación de suelos. Estudios, muestreo y análisis. *Diario Oficial*, 31 de diciembre de 2002.

Yamamoto, S. (2000). Forest gap dynamics and tree regeneration. *Journal of Forest Research* 5: 223-229. DOI: <https://doi.org/10.1007/BF02767114>

## **3.2. Algunos ejemplos metodológicos típicos extraídos de la literatura científica aplicados al crecimiento de plantas, entomología y control biológico**

### **3.2.1. Germinación de semilla**

Los datos de emergencia y germinación presentan un problema estadístico especial: en general, no se sabe si luego de finalizado el experimento, las semillas que no germinaron son capaces de germinar, están latentes, han germinado y luego murieron, o siempre estuvieron muertas. Las semillas inviables, obviamente, no deben considerarse como parte de un estudio de población. Cuando se da el caso de que muy pocas semillas no emergen, se pueden analizar los datos asumiendo primero que estas eran viables. Esto significa que son tratadas como puntos de datos censurados, con el final del estudio como fecha de censura. Un segundo caso es asumir que las semillas eran inviables, por lo que las semillas que no emergen pueden ser excluidas del análisis. Luego ambos resultados se comparan y se observa si son o no cualitativamente iguales. Un enfoque alternativo será examinar las semillas no emergentes para determinar su viabilidad mediante una prueba de tetrazolio (Scott y Jones 1990), y así clasificar correctamente cada semilla. Esto sería necesario si los dos análisis estadísticos difieren cualitativamente.

Según McNair, Sunkara y Frobish (2012) y Scott, Jones y Williams (1984), en los experimentos de germinación de semillas se han propuesto numerosos métodos para analizar los datos resultantes, la mayoría de los cuales se dividen en tres clases: índices de germinación, tiempo medio de respuesta, análisis probit, ajuste de curvas de germinación acumulada, análisis de regresión no lineal, regresión logística y análisis de tiempo hasta el evento (análisis de sobrevivencia, análisis de tiempo de falla y análisis de confiabilidad). El análisis del tiempo transcurrido hasta el evento, tiene ventajas importantes sobre los otros métodos. Los tipos de análisis de tiempo hasta el evento para analizar los datos de germinación de semillas incluyen métodos no paramétricos (tablas de vida y estimadores de Kaplan-Meier, además de otros métodos para comparar dos o más grupos de semillas) y métodos semiparamétricos (modelo de riesgos proporcionales de Cox, que permite la inclusión de covariables categóricas y cuantitativas, así como efectos aleatorios). Los métodos de análisis del tiempo transcurrido hasta el

evento se pueden aplicar a muchos otros tipos de datos biológicos, como por ejemplo los tiempos de emergencia de las plántulas, tiempos de floración, tiempos de desarrollo de huevos y etapas larvales, ninfales y adultos en insectos o en etapas de vida de cualquier organismo biológico.

Scott, Jones y Williams (1984), discuten las comparaciones entre estos métodos e ilustran con datos de germinación en tomate (*Lycopersicon esculentum* Mill.). Para ellos, la germinación de semillas implica no solo respuestas cualitativas de los individuos, sino también respuestas poblacionales que se distribuyen a lo largo del tiempo. Según ellos, el análisis de varianza o los métodos de regresión son apropiados para un análisis de datos donde se observa la germinación de todas las semillas viables, pero son inapropiados cuando algunas semillas viables no germinan. Estos datos faltantes (censurados) complican el análisis estadístico y la interpretación posterior. Las pruebas de germinación deben diseñarse para determinar la naturaleza de las respuestas censuradas, que posteriormente pueden adaptarse mediante el análisis de supervivencia.

### **3.2.1.1. Métodos estadísticos de análisis de supervivencia aplicados en germinación de semilla.**

#### **Análisis tiempo – evento**

El método describe una variable de respuesta binaria, donde los datos se recopilan como recuentos a lo largo del tiempo. Consiste en modelar el tiempo de respuesta para cada individuo (semilla) en la muestra. Con respecto a los estudios de germinación, se pueden recolectar datos como el tiempo de germinación de cada semilla, el número de semillas que germinan en cada intervalo de tiempo o el número acumulado de semillas que germina durante el tiempo de observación. El análisis del tiempo transcurrido hasta el evento lo abordan mediante tres formas diferentes: *no paramétrica*, *semiparamétrica* y *totalmente paramétrica*.

#### **Métodos no paramétricos**

Los métodos no paramétricos no hacen suposiciones sobre una distribución de probabilidad subyacente. Así es como el evento de germinación cambia con el tiempo, en función de la probabilidad de desarrollo de la semilla. Utilizando métodos no paramétricos, se estima y gráfica la distribución de supervivencia o la curva de supervivencia. Es posible comparar dos o más distribuciones de supervivencia. Al respecto, los métodos más utilizados son la prueba de rango logarítmico, también conocida co-

no prueba de rango logarítmico de Mantel, la prueba de rango logarítmico de Cox Mantel y la prueba de Mantel Haenszel y sus versiones (Peto *et al.*, 1977).

Los métodos de análisis de supervivencia también se pueden ampliar para evaluar factores de riesgo o exposiciones simultáneamente, de forma similar al análisis de regresión logística lineal y múltiple. El modelo de riesgo proporcional de Cox es el más general de los modelos de regresión porque no se basa en suposiciones sobre la naturaleza de la forma de la distribución de supervivencia subyacente. En este modelo, la variable de respuesta es el "riesgo". El modelo asume que la tasa de riesgo subyacente, en lugar del tiempo de supervivencia, es una función de las variables independientes o covariables (Lewicki y Hill, 2006).

### **Métodos paramétricos**

En modelos completamente paramétricos, se asume una distribución de probabilidad específica de la *línea de base de riesgo / sobrevivencia*, de acuerdo con una distribución de probabilidad definida. Los modelos paramétricos se pueden expresar en: (1) modelo de riesgo proporcional, donde un cambio de una unidad en una variable explicativa causa cambios proporcionales en el riesgo; y (2) modelo de tiempo de falla acelerado (AFT), donde un cambio de una unidad en una variable explicativa provoca un cambio proporcional en el tiempo de sobrevivencia. El modelo AFT se calcula en función del número de semillas que germina en cada intervalo de tiempo antes de ese momento y del número de semillas que podrían haber germinado al final del intervalo anterior. El enfoque AFT permite considerar una variedad de funciones de distribución estadística para los tiempos de germinación en un lote de semillas y compara las respuestas para diferentes lotes de semillas mediante la estimación de un '*factor de aceleración*' que resume las tasas relativas de germinación de lotes de semilla, pero asume una forma común de respuesta. Enfatizan que otros autores (McNair *et al.*, 2012; Scott y Jones, 1982; Barak *et al.*, 2018), revisaron modelos no paramétricos y semiparamétricos aplicados a datos de germinación, mientras que Onofri *et al.* (2010); Winkler, Hülber y Hietz (2005); Hirsch *et al.* (2012), consideraron una variedad de modelos AFT para describir las respuestas de germinación para diferentes especies.

A continuación, se detallan algunos trabajos científicos tipo, extraídos de la literatura, en los cuales se explican las metodologías aplicadas:

**a. Ejemplo 1: Germination Data Analysis by Time-to-Event Approaches (Romano y Stevanato, 2020).**

## **Materiales y métodos**

### **Material vegetal**

El estudio tiene como objetivo dar una contribución al conocimiento sobre la confiabilidad de estos métodos mediante la evaluación de enfoques principales a datos de germinación en cohortes de semillas de remolacha azucarera (*Beta vulgaris* L.). Los resultados obtenidos confirmaron que, si bien los diferentes enfoques presentan ventajas y desventajas, en general podrían representar una valiosa herramienta para analizar los datos de germinación aportando parámetros cuya utilidad depende del propósito de la investigación.

### **Métodos aplicados**

El análisis de los datos de germinación consistió en aplicar métodos no paramétricos, semiparamétricos y paramétricos a los mismos datos de germinación de semilla de remolacha azucarera, basado en que ningún trabajo de investigación en germinación ha examinado la aplicación de modelos no paramétricos, semiparamétricos y totalmente paramétricos a los mismos datos, comparando el significado biológico de los parámetros estimados, por lo que el objetivo del trabajo fue una contribución al conocimiento de la efectividad de los métodos de tiempo hasta el evento, mediante el análisis de los datos de germinación de semillas de remolacha azucarera, utilizando dos covariables categóricas: el estrés osmótico y el genotipo, y comparar los tres enfoques para verificar la confiabilidad de los resultados.

## **Conclusiones**

El análisis del tiempo transcurrido hasta el evento demostró ser una herramienta confiable para analizar los datos de germinación. Sin embargo, la elección del método más apropiado dependerá del propósito para el cual debe ejecutarse el análisis. Consideran que las curvas de supervivencia de Kaplan-Meier (KM) proporcionan una primera idea útil de la forma de la función de supervivencia para cada tratamiento / genotipo, centrándose en que no se produzca un evento de germinación. También hacen una reflexión que sus resultados son gráficamente intuitivos y podría ser un método útil para comparar

más grupos de semillas a través de las pruebas de rango logarítmico y en términos de mediana o cuartil de tiempos de supervivencia. Al mismo tiempo, la función de riesgo no paramétrica, que proporciona una idea de las tasas de fallas condicionales, podría ser una herramienta gráfica para analizar de forma automática datos de germinación, ya que se enfoca en el evento que ocurre. Además, aplicando este método, no son necesarios supuestos paramétricos sobre las distribuciones del tiempo de germinación.

Consideran que el modelo de Cox proporciona información útil cuando se necesita considerar un conjunto de covariables que influyen en la germinación. Sin embargo, deben tenerse en cuenta algunas limitaciones sobre la interpretación de los resultados, tales como el retraso en el inicio de la germinación, para evaluar la validez del modelo cuando se aplica a los datos de germinación (McNair, Sunkara y Frobish, 2012), y las suposiciones de riesgo proporcional deben verificarse antes de aplicar este método. Si se cumplen los supuestos, el modelo podría considerarse una herramienta útil para estudios sobre el potencial de germinación en semillas sembradas en ambientes peculiares o para probar diferentes condiciones de almacenamiento de semillas, ajustando parámetros como la temperatura y la humedad relativa para estimar qué factor es más importante para determinar la descomposición de la germinación con el tiempo.

Finalmente, la ventaja del enfoque del tiempo de falla acelerado es que el efecto de las covariables en la supervivencia se puede describir en términos absolutos más que en términos relativos como una razón de riesgo. Además, ya que la forma de la línea base del riesgo de germinación generalmente se conoce, existe una buena posibilidad de utilizar la forma de distribución adecuada. El modelo AFT se puede interpretar en términos de la velocidad de aquellos procesos fisiológicos que terminan con el evento de germinación, y por esta razón, podría ser efectivo en aquellos experimentos cuyo objetivo sea para evaluar la velocidad de germinación en respuesta a factores ambientales, tales como probar los efectos de un herbicida sobre la germinación de ciertas malezas o para evaluar la precocidad germinativa de algunas variedades en comparación con otras, en condiciones estándar o después de tratamientos específicos.

**b. Ejemplo 2: Germinación de tres especies de *Cephalocereus* (Cactaceae) endémicas del Istmo de Tehuantepec, México (Barcenás *et al.*, 2013).**

Se estudió el proceso germinativo de tres especies de *Cephalocereus* (*C. apicicephalium*, *C. nizanensis*, *C. totolapensis*), bajo las hipótesis de que las semillas pierden la capacidad germinativa al transcurrir el tiempo (días) después de la recolección y que la germinación es diferente entre especies, pero no entre poblaciones de la misma especie. Para cada una de las especies, se recolectaron semillas de poblaciones diferentes y se sembraron en cuatro fechas distintas, colocando 30 semillas en dos o tres cajas de Petri. Las cajas se colocaron en una cámara con ambiente controlado y la germinación se registró diariamente durante 30 días, considerando una semilla germinada, al emerger la radícula.

### **Material y métodos**

#### **Material biológico**

De las poblaciones conocidas de las tres especies, se seleccionaron dos o tres por especie y en cada una de ellas se recolectaron tres frutos maduros de tres individuos. Las semillas se sacaron del fruto por disección; en cada fruto se encontraron una o más semillas con la radícula expuesta, por lo que se considero que todas estaban maduras. De acuerdo al criterio de apariencia propuesto por Ayala *et al.* (2004) las semillas se clasificaron en viables, abortadas y germinadas. Las semillas no germinadas se lavaron con agua corriente, se colocaron sobre papel absorbente y se dejaron secar a la sombra a temperatura ambiente (17- 21°C); una vez secas se guardaron en frasco color ambar por 100 días hasta el inicio de los experimentos.

#### **Ensayos de germinación**

Las semillas se sembraron en cuatro fechas distintas, 100 días después de la recolección de los frutos (ddc), 130 ddc, 190 ddc y 220 ddc. Además, se partió del supuesto de que la presencia de semillas germinadas es un indicador de la inexistencia de latencia para las tres especies y que éstas germinarían después de estar almacenadas 100 días, una vez que las condiciones de humedad fueran favorables. Las semillas fueron desinfectadas superficialmente por inmersión en una solución de hipoclorito de sodio al 10% durante 5 minutos, luego se enjuagaron con agua destilada y se sembraron en cajas de Petri de 9 cm de diámetro con papel filtro estéril saturado con agua destilada. Se colocaron 30 semillas en cada

caja y de dos a tres cajas por población. Las cajas de Petri se colocaron en una cámara con ambiente controlado a una temperatura diurna de 24°C y nocturna de 20°C, bajo un fotoperiodo de 14 horas de luz y 10 horas de oscuridad. La germinación se registró diariamente durante un periodo de 30 días (Baskin y Baskin, 2001), considerando una semilla germinada cuando emergía su radícula (Dubrovsky, 1998).

### **Análisis estadístico**

Se utilizó el análisis de varianza no balanceado mediante el modelo general lineal y la comparación de medias con la prueba de Tukey ( $P < 0.05$ ), para evaluar diferencias en los porcentajes de germinación entre las especies y entre las poblaciones de cada especie, a partir de datos transformados por arcoseno. Las diferencias observadas en las curvas de germinación para especies, poblaciones y fechas de siembra fueron analizadas con el procedimiento LIFETEST y comparadas entre pares de curvas por el mismo método; se empleó la prueba de Wilcoxon si las curvas de germinación se cruzaron y, en caso contrario, se utilizó la prueba de Log-Rank (Fox, 2001).

### **c. Ejemplo 3: Germinación y emergencia diaria de cariósides y diásporas de pastos nativos e introducidos (Quero et al., 2017).**

El objetivo del trabajo consistió en evaluar la tasa de germinación diaria (TGD) en cariósides clasificados por tamaño (CCT) y la tasa de emergencia diaria (TED), tanto en CCT como en diásporas, en los pastos nativos Banderita (*Bouteloua curtipendula* (Michx.) Torr.) y Navajita (*B. gracilis* (Willd. ex Kunth) Lag. ex Griffiths), y los introducidos Buffel (*Cenchrus ciliaris* L.) y Rhodes (*Chloris gayana* Kunth).

### **Materiales y métodos**

#### **Material genético**

El material experimental consistió de propágulos (semilla) de dos especies de pasto nativas del Desierto Chihuahuense: Navajita y Banderita variedad Reno, así como dos especies introducidas de África: Rhodes, variedad Bell y Buffel, variedad T4464. La semilla utilizada fue comercial, la cual se caracterizó inicialmente por contenido de humedad y prueba de viabilidad con tetrazolio (0,1%), con el



fin de uniformizar la tasa de germinación y emergencia por especie y tamaño de carióspside. Los experimentos tanto en laboratorio como en invernadero se realizaron en tres ocasiones, en laboratorio con intervalo de 30 d y en invernadero de 64 d. En cada ocasión que se realizó el experimento, se obtuvieron 200 g de carióspsides de manera manual mediante fricción con tapete y almohadilla de caucho corrugado a partir de diásporas comerciales. Para clasificar carióspsides por tamaño (CCT) se midió el grosor de 200 semillas botánicas y después se separó con ayuda de tamices, y se conservaron en sobres de papel en condiciones de laboratorio ( $20 \pm 3$  °C).

### **Fase de laboratorio**

Para envejecimiento acelerado (PEA), los propágulos (carióspsides y diásporas) permanecieron en estufa (42°C, 100 % HR) por 12, 24 y 36 h y una vez concluido el tiempo de estrés (TE), se lavó el material biológico y el testigo (sin estrés) con hipoclorito de sodio (0,6 %) por 3 min e inmediatamente se sumergió a la semilla en Captan (1 g L<sup>-1</sup> agua) por 3 min y se enjuagó con agua destilada. Posteriormente, se sembró por especie, CCT o diásporas y el testigo (CCT y diásporas, sin estrés), en cuatro repeticiones de 100 unidades de dispersión con base en semilla pura viable (SPV). Los sustratos incluyeron: a) Papel filtro, para sembrar por CCT en caja plástica transparente (20×15×8 cm) con flujo ligero de aire; b) Peat Moss, para sembrar diásporas a 1 cm de profundidad en charolas de 40×30×5 cm sin tapa. Las siembras se colocaron en cámara de ambiente controlado a 22°C con luz fluorescente durante las 24 h. Para el riego por CCT, se aplicaron 30 mL de agua destilada en el papel filtro a cada una, cada 48 h, y para diásporas 100 mL por aspersion sobre el Peat Moss, cada 72 h. La germinación de plántulas normales para CCT y diásporas se consideró efectiva cuando fue evidente la primera hoja, y se contabilizó cada 24 h; en CCT, hasta 15 d después de siembra (dds), y en diásporas hasta 25 dds.

### **Fase de invernadero**

Los CCT, diásporas y charolas (con 70 orificios individuales; 4×2×10 cm), se desinfectaron por inmersión en solución de hipoclorito de sodio al 0,6 % durante 3 min; posteriormente, se enjuagaron con agua destilada y se secaron a la sombra. Los sustratos utilizados fueron suelos tipo Calcisol de Salinas Hidalgo, San Luis Potosí y Vertisol de Atotonilco El Grande, Hidalgo. Los suelos se sometieron a esterilización en autoclave durante 6 h para evitar germinación de semilla de maleza. La siembra en ambos

suelos se realizó a capacidad de campo a profundidades de 0,5 – 1,0 – 2,0 – 3,0 y 5,0 cm en charolas de unícel y cada cavidad (unidad experimental) midió 10 × 3.5 cm, después se sembró con SPV, para obtener cinco plántulas por unidad experimental (seis cariósides para tamaño mediano y grande en las cuatro especies; siete cariósides en tamaño chico en Banderita y ocho en Navajita, Buffel y Rhodes). Se regó a saturación cada 48 h con agua potable. La cuantificación de plántulas se realizó cada 24 h por 64 d, para CCT y diásporas, y fue efectiva cuando la primera hoja fue evidente.

### **Tratamientos evaluados y variables respuesta**

En laboratorio, para la PEA los tratamientos incluyeron la combinación de tres CCT en dos especies nativas o introducidas por cada tiempo de estrés (TE) (0, 12, 24 o 36 h), i.e., 24 tratamientos para las especies nativas y 24 para las introducidas; por otra parte, en diásporas, la combinación de dos pastos nativos o introducidos en cada TE, lo cual resultó en ocho tratamientos para cada grupo de especies. Las variables respuesta fueron la tasa de germinación diaria (TGD) y la tasa de emergencia diaria (TED). En invernadero los tratamientos resultaron de la combinación de cinco diferentes profundidades de siembra (PS), dos especies nativas o introducidas, tres tipos de CCT sembradas en dos tipos suelo, i.e., 60 tratamientos por grupo; mientras que, para diásporas, los tratamientos resultaron de la combinación de dos especies nativas o dos introducidas sembradas en cinco PS y en dos tipos de suelo, lo que resultó en 20 tratamientos para especies nativas y 20 para introducidas. La variable respuesta evaluada fue la tasa de emergencia diaria tanto para cariósides como para diásporas.

### **Análisis estadístico**

Los datos se analizaron por regresión logística (Hosmer y Lemeshow, 2001) mediante procedimientos para análisis de tiempos de vida (LIFEREG). La comparación de la germinación en el tiempo se realizó por razón de verosimilitudes generalizada ( $\alpha = 0,05$ ), con base en los resultados de la rutina LIFEREG. Con la finalidad de comparar las curvas de germinación en el tiempo se utilizaron modelos estadísticos donde se presentan datos censurados. El modelo utilizado es:  $\log(T_i) = x'_i \beta + \sigma e_i$ , donde  $T_i$  es el tiempo a germinación,  $x'_i$ , representa un vector de covariables asociadas al tiempo de germinación (por ejemplo, tamaños de semilla, tratamiento, etc),  $\beta$  representa un vector de parámetros de regresión asociados a las covariables,  $\sigma$  es un parámetro de escala positivo y  $e_i$  es un error aleatorio.

**d. Ejemplo 4: Factores que influyen en la viabilidad y germinación de *Carpinus caroliniana* Walt. (Rodríguez, 2019)**

El objetivo consistió en evaluar los factores que influyen en la viabilidad, germinación y supervivencia de la especie arbórea *Carpinus caroliniana* del Bosque Mesófilo de Montaña. Bajo condiciones controladas, evaluaron la densidad de semillas en muestras de dos distancias diferentes de la base del árbol. Además, para conocer el potencial de germinación de la especie, se cuantificó la viabilidad de la semilla mediante la prueba de tinción con tetrazolio, en diferentes años de colecta (2015, 2016 y 2017). Se aplicaron diferentes tratamientos pregerminativos para incrementar el porcentaje de germinación y, como parte complementaria, las semillas fueron pesadas individualmente para evaluar la influencia del peso en la germinación. Adicionalmente, realizaron estudios en campo, donde se evaluó la densidad de plántulas de *C. caroliniana*, su establecimiento y supervivencia, bajo diferentes tratamientos de apertura de dosel y remoción de hojarasca.

**Determinación de viabilidad y germinación en semillas de la especie arbórea *C. caroliniana*.**

**Objetivo general**

Determinar los factores que influyen en la viabilidad y germinación de semillas de la especie arbórea *C. caroliniana* del bosque mesófilo de montaña.

**Objetivos específicos**

Evaluar la densidad de semillas a diferentes distancias del árbol mediante la tamización y germinación de tapetes de suelo.

Evaluar la viabilidad de semillas de *C. caroliniana* mediante la prueba de azul de tetrazolio.

Evaluar la germinación y capacidad de almacenamiento en semillas de *C. caroliniana* mediante tratamientos pregerminativos.

## Hipótesis

En estudios previos se ha demostrado un escaso potencial de germinación, supervivencia y establecimiento de plántulas de *C. caroliniana*, por lo que se sugiere que la distancia de dispersión, el tiempo de almacenamiento y el uso de tratamientos pregerminativos tendrá un efecto positivo en la germinación de las semillas de esta especie.

1. Los tapetes de suelo colectados cerca de la base del árbol progenitor presentarán un banco de semillas con mayor densidad, y la emergencia de plántulas incrementará en comparación con tapetes de suelo localizados al borde de la copa del árbol.
2. Las semillas de colectas recientes (año de colecta) presentarán mayor viabilidad, en comparación con lotes de semillas con mayor tiempo (uno y dos años) de almacenamiento.
3. La aplicación de tratamientos pregerminativos incrementará la velocidad y porcentaje de germinación de las semillas de *C. caroliniana*.

## Materiales y métodos

La densidad de semillas se evaluó mediante la extracción de tapetes de suelo; la viabilidad se determinó mediante una prueba con tetrazolio en semillas con diferente año de colecta (2015, 2016 y 2017), y la germinación se evaluó mediante la aplicación de diferentes tratamientos pregerminativos. Adicionalmente, se evaluó el tiempo de viabilidad con la aplicación del mejor tratamiento pregerminativo a las colectas de los diferentes años.

## Colecta de material biológico

Las semillas se obtuvieron de colectas directas de los árboles de un remanente de Bosque Mesófilo de Montaña en el Municipio de Uruapan, Michoacán. Previo a la colecta de semillas, se realizó el monitoreo fenológico de la especie durante los meses de febrero-julio del año 2015, 2016 y 2017, con la finalidad de coleccionar frutos y semillas maduros de 8 individuos. Los frutos y semillas se colectaron con una garrocha extensible y mantas para la captura de la semilla. Las semillas se limpiaron removiendo la bráctea alar y se almacenaron en bolsas herméticas de plástico a una temperatura de 4°C.

### **Densidad de semillas a diferente distancia en *Carpinus caroliniana***

Se extrajeron tapetes de suelo de 20 x 20 x 8 cm en el mes de agosto del año 2017, eliminando la capa de hojarasca, alrededor de cinco árboles seleccionados aleatoriamente. De cada árbol se colectaron 12 tapetes de suelo a dos distancias, seis bajo la copa del árbol (2m de la base del árbol) y seis más al borde de la copa del árbol (4m desde la base del árbol). Los tapetes fueron colocados y etiquetados en bolsas de plástico herméticas tipo ziploc para su traslado al laboratorio.

Para la densidad de semillas de *C. caroliniana* se utilizaron dos métodos:

**Tamizado:** Los tapetes de suelo fueron secados y tamizados para separar la materia orgánica, extraer y contabilizar el número de semillas presentes en cada bloque.

**Germinación:** los tapetes de suelo se colocaron en charolas de plástico (60 x 40 x 7 cm), y se mantuvieron en condiciones de invernadero con riego cada tercer día y/o a capacidad de campo. Se llevó a cabo el registro de la emergencia de plántulas semanalmente, tomando como referencia la emergencia de la plántula sobre el sustrato.

### **Prueba de viabilidad**

La prueba de viabilidad se realizó en lotes de semillas de tres diferentes años de colecta (en el mes de junio de 2015, 2016 y 2017). Las semillas se hidrataron en agua destilada por 12 horas, posterior al remojo, se realizó un corte longitudinal a las semillas dejando los embriones expuestos. Las semillas se colocaron en vasos de precipitados en una solución de Cloruro de Tetrazolio (0,12g/30ml de agua) por 72 horas y cubiertas con papel aluminio. Al finalizar el tiempo de inmersión, se procedió a retirar el excedente y contabilizar las semillas viables (teñidas de rojo). Cada lote contó con tres réplicas de 100 semillas.

### **Germinación de semillas de *Carpinus caroliniana* mediante tratamientos pregerminativos**

La germinación en semillas colectadas en junio del año 2018 se evaluó mediante la aplicación de tres tratamientos pregerminativos y un control, en cada uno de los tratamientos se utilizaron 400 semillas, divididas en cuatro réplicas de 100 semillas. Los tratamientos aplicados fueron: 1) Control (sin tratamiento pregerminativo y sin hormona), 2) Remojo en agua por una semana, 3) Adición directa en

sustrato de 0,143 g de Ácido Indol-tres-butirico (AIB, raizon®) y 4) Inmersión en Ácido giberelico (AG3, SIGMA®) a 1000 ppm por 24 horas. Previo a la aplicación de los tratamientos, las semillas fueron pesadas y colocadas individualmente en viales Eppendorf de 1,5 ml. Después de aplicados los tratamientos se realizó la siembra en charolas de plástico (60 x 40 x 7 cm), con sustrato hecho a base de Agrolita y Peat moss (turba) en proporciones de 1:1. Las charolas se etiquetaron con el tratamiento correspondiente y se dispersaron de forma aleatoria dentro de la casa sombra. Se mantuvieron bajo condiciones de invernadero, con riego cada tercer día y/o a capacidad de campo. La germinación se registró cada tercer día durante 10 semanas.

### **Ácido giberélico (AG3) en la germinación de semillas de *Carpinus caroliniana***

Adicionalmente, se probó si la aplicación de ácido giberélico (AG3) incrementa la germinación de las semillas con diferente año de colecta (2015, 2016 y 2017). Para ello, las semillas de *Carpinus caroliniana* fueron estratificadas por cuatro semanas a 4°C, en sustrato humedecido con agua destilada a capacidad de campo fueron sometidas a un tratamiento pregerminativo con AG3 (SIGMA®), en dos concentraciones: 1000 ppm (T1) y 2000 ppm (T2) y un control (T3). Se utilizaron tres lotes de semillas (2015, 2016 y 2017), que fueron pesadas individualmente y colocadas en viales Eppendorf de 1,5 ml para la aplicación del tratamiento, generando un total de nueve tratamientos. Las semillas se sumergieron en una solución de AG3 por 24 horas, al finalizar el tiempo de inmersión, las semillas se enjuagaron y fueron sembradas en charolas tipo Flex de 72 cavidades individuales, en sustrato hecho a base de agrolita y Peat moss (turba) en proporciones de 1:1, a una profundidad de 1 cm. Se utilizaron 100 semillas por tratamiento con un total de 900 semillas. Las semillas fueron colocadas en invernadero. El riego se realizó cada tercer día buscando mantener una humedad constante y/o a capacidad de campo. Los registros de germinación se realizaron cada tercer día durante 12 semanas, tomando como semilla germinada la emergencia de la plántula sobre el sustrato.

## Análisis estadísticos

Se cumplió con los supuestos de normalidad de los datos (prueba de Shapiro-Wilk W). En los tapetes de suelo se realizó la comparación de la densidad de semillas y la emergencia de plántulas mediante un análisis de varianza. En la prueba de viabilidad, los porcentajes de semillas viables se transformaron al arcoseno para cumplir con los supuestos de andeva y se realizó un análisis de varianza ( $p < 0,05$ ). Se utilizó un modelo lineal generalizado (GLM) con distribución binomial, para evaluar diferencias en la germinación entre los tratamientos, y el peso individual por semilla como covariable. Adicionalmente, se calculó la velocidad de germinación (VG) y se realizó un análisis de supervivencia (Kaplan-Meier) para evaluar si existen diferencias entre las curvas de germinación de los tratamientos. Los análisis se realizaron con el software JMP Versión 8.0 (SAS Institute).

El modelo de riesgos proporcionales (Modelo de Cox) permite evaluar el efecto de covariables en el tiempo de vida hasta la ocurrencia del evento de interés (Bonfil y Soberón, 1999; Corpas y Lara, 2009). En este caso, este modelo fue adaptado para evaluar el efecto del peso y el tratamiento germinativo en la germinación de semillas de *Carpinus caroliniana*. Cuando el valor del exponente del coeficiente (coef) es mayor a 1, se considera que la covariable tiene un mayor riesgo asociado con la ocurrencia del evento. En este estudio, la condición anterior significa un efecto positivo en la germinación. El análisis se realizó con el paquete “survival” del programa R Versión 3.5.1.

Para evaluar el efecto del peso de la semilla en la germinación, esta variable se dividió en cuatro categorías: bajo (6 a 8 mg), medio (8 a 10 mg), alto (10 a 12 mg) y muy alto (12-17 mg). En cada uno de los tratamientos pregerminativos se contabilizó el número de semillas germinadas por categoría y se obtuvo el porcentaje de germinación. La categoría de peso alto fue la que presentó el mayor porcentaje de germinación en todos los tratamientos y por tal motivo se utilizó como grupo de referencia. Adicionalmente, este modelo también fue aplicado para evaluar la probabilidad de germinación de los tratamientos pregerminativos de las semillas respecto al tratamiento control. Finalmente, se realizó la comparación del porcentaje de germinación de semillas en diferentes concentraciones de AG3 y el tiempo de almacenamiento.

### 3.2.2. Análisis de regeneración natural y crecimiento de plantas

#### a. Ejemplo 1: Factores que influyen en la emergencia y supervivencia de plántulas de *Carpinus caroliniana* in situ (Rodríguez, 2019).

##### Objetivo general

Explorar los factores que influyen en el establecimiento y supervivencia de plántulas de la especie arbórea *Carpinus caroliniana* del bosque mesófilo de montaña en condiciones de campo.

##### Objetivos específicos

Evaluar la densidad y supervivencia de plántulas in situ bajo tres condiciones de apertura de dosel (sombra, intermedio y expuesto).

Evaluar el efecto de remoción de hojarasca en el establecimiento y supervivencia de *Carpinus caroliniana* in situ.

##### Hipótesis

Dado que el tipo de vegetación donde se desarrolla la especie bajo estudio, produce una alta acumulación de hojarasca que le confiere un sustrato propicio para la germinación de la semilla, esté podría al mismo tiempo estar afectando el reclutamiento de plántulas en los siguientes estadíos.

1. La densidad de plántulas puede variar dependiendo de las condiciones de apertura de dosel. En condiciones intermedias de luz in situ, se encontrará una mayor densidad de plántulas que en condiciones de ausencia o extrema insolación.
2. En cuanto a la remoción de hojarasca en parcelas in situ, se espera encontrar mayor reclutamiento de plántulas en tratamientos con remoción de hojarasca a diferencia de las que no presenten remoción.



## **Materiales y métodos**

### **Evaluación de la emergencia y supervivencia de plántulas in situ de *C. caroliniana* Walt. bajo distintas aperturas de dosel**

Se evaluó la supervivencia in situ de plántulas de *C. caroliniana* en tres tratamientos de aperturas de dosel (sombra, intermedio y expuesto), en un rodal de Bosque Mesófilo de Montaña. En cada tratamiento se delimitaron seis parcelas de 3 x 3 m, y se cuantificó la apertura de dosel con imágenes hemisféricas analizadas con WinScanopy TM Versión 2014. En cada parcela se marcaron mensualmente las plántulas emergidas y se registró la supervivencia de cada cohorte de julio a diciembre de 2017 (época de lluvias) y enero a junio de 2018 (época seca). En cada registro de emergencia se midió la humedad y temperatura del suelo con el equipo Soil moisture kit 150  $\Delta$ T.

### **Efecto de la remoción de hojarasca en el establecimiento y sobrevivencia in situ de *C. caroliniana*.**

Con la finalidad de evaluar el efecto de la cantidad de hojarasca en el establecimiento y supervivencia de plántulas de *C. caroliniana*, previo a la época lluviosa de 2018 se establecieron 63 cuadrantes de 1 m<sup>2</sup> dentro del área del Bosque Mesófilo de forma aleatoria bajo árboles semilleros. Los cuadrantes se dividieron en dos tratamientos de remoción de hojarasca: remoción parcial de hojarasca (dejando un 50%), remoción total y control (Sin remoción de hojarasca). El establecimiento de las parcelas se realizó en el mes de mayo previo a la dispersión de semilla. La remoción de hojarasca se realizó manualmente teniendo especial cuidado para no remover la semilla. Las evaluaciones se realizaron de forma mensual de junio a diciembre de 2018, en cada visita se marcó y registró la emergencia y supervivencia de las plántulas. Las plántulas registradas por mes fueron tomadas como eventos independientes de germinación denominadas cohortes. Se midió la humedad y temperatura del suelo, así como el grosor de la hojarasca en cada uno de los cuadrantes.

### **Análisis estadístico**

Se probó la normalidad de los datos (Prueba de Shapiro-Wilk W) y se aplicó un análisis de varianza (ANOVA) de una vía para evaluar si existen diferencias en el número de plántulas emergidas entre los tratamientos de apertura de dosel (sombra, intermedio y expuesto) y remoción de hojarasca (remoción parcial, remoción total y control sin remoción). Se realizó un análisis de supervivencia (Estimador

Kaplan Mier) para conocer si el tiempo de supervivencia de las plántulas fue diferente entre los tratamientos de apertura de dosel y remoción de hojarasca. Los análisis estadísticos se realizaron con el Software JMP Versión 8.0 (SAS Institute).

El efecto de la cobertura de dosel, humedad y temperatura del suelo en la supervivencia de las plántulas se evaluó con un análisis de riegos proporcionales (modelo de Cox). Esta prueba semiparamétrica mide el efecto de un conjunto de variables en el tiempo que ocurre un evento (Bonfil y Soberon, 1999), en este caso la muerte de una plántula. Su interpretación se basa utilizando la relación de riesgo (HR), en donde un valor igual a 1 indica que la variable no ejerce ningún efecto y un valor mayor a 1 indica que la variable modifica la incidencia de mortalidad de las plántulas. El análisis se realizó con el paquete “survival” del programa R Versión 3.5.1.

**b. Ejemplo 2: Supervivencia en plantaciones de *Pinus pseudostrobus* Lindl. en función del sistema de producción y preacondicionamiento en vivero (Sigala, Gonzalez y Prieto, 2015).**

**Material vegetal**

Se seleccionaron ejemplares de *Pinus pseudostrobus* cultivados en tres sistemas de producción (Charola de poliestireno 60 mL (0+1), bolsa de polietileno 630 mL (0+2) y charola de poliestireno + bolsa de polietileno (60-560 mL) (1+1). Durante el periodo de producción se aplicaron riegos cada tres días, de acuerdo a las rutinas convencionales del vivero. Para inducir estrés hídrico y favorecer el preacondicionamiento, previo al trasplante en campo, las plantas se regaron durante 45 días bajo el siguiente régimen 1) nivel bajo, un riego diario durante la mañana; 2) nivel moderado, un riego a saturación cada 9 días; y 3) nivel alto, un riego a saturación cada 15 días.

**Plantacion**

Se eligieron dos sitios de condiciones contrastantes (sitio 1 y sitio 2), separados a una distancia de 200 m. En cada sitio se plantaron 24 individuos por tratamiento, divididos en tres repeticiones y situados a 1 m de separación; en total se emplearon 432 plantas. El diseño experimental fue en parcelas subdivididas, las grandes correspondieron a los sitios, las medianas el sistema de producción y las parcelas chicas al nivel de preacondicionamiento.

### Análisis estadístico

Una vez establecida la plantación, se evaluó mensualmente la supervivencia durante nueve meses, con un registro final a los 14 meses. En cada medición se asignaron valores de 0 o 1 para las plantas muertas y vivas, respectivamente. Asimismo, se consideraron las variables morfológicas iniciales: diámetro al cuello de la raíz (mm), altura del tallo (cm), peso seco de la parte aérea (PSA) y peso seco de la raíz (PSR) (g). Las dos últimas se obtuvieron de un muestreo destructivo realizado en cada tratamiento, antes de plantar. Las diferencias de supervivencia entre tratamientos, se analizaron mediante la prueba Log-Rank a partir de curvas de supervivencia construidas por el método Kaplan-Meier (Kaplan y Meier, 1958) para ello, se tomó en cuenta el estatus de cada planta (viva o muerta) al final del periodo de evaluación, así como el tiempo de vida de la misma en meses. Este análisis se hizo con el procedimiento LIFETEST. Para estimar el efecto de los factores estudiados, en función de las variables morfológicas como covariables, se aplicó una regresión de riesgos proporcionales de Cox. Este modelo estima un coeficiente  $\beta$  para cada factor o covariable y prueba la hipótesis nula que  $\beta = 0$  con el estadístico  $\text{Chi}^2$  (chi cuadrada). Dicho coeficiente explica el efecto de un factor o una covariable en la función de riesgo; es decir, si el coeficiente  $\beta$  es negativo, significa que el riesgo de muerte se reduce con el incremento de la covariable, mientras que un coeficiente  $\beta$  positivo, indica lo contrario (Williams, 2008). El análisis se efectuó mediante el procedimiento PHREG. El diámetro fue la variable morfológica que más se relacionó con el riesgo de mortalidad en los sitios de plantación.

### 3.2.3. Entomología y control biológico

#### a. Ejemplo 1: Impacto de diferentes temperaturas nas características biológicas de *Helicoverpa armigera* (Hübner, 1808) (Lepidoptera: Noctuidae) durante três gerações (Alvez, 2017)

Para que el control de insectos plaga sea efectivo, es necesario conocer las características biológicas del insecto, especialmente las relacionadas con factores abióticos. La temperatura es el principal factor que actúa sobre la dinámica poblacional de los insectos, interfiriendo en la supervivencia, el desarrollo, la reproducción y el comportamiento. Entre las plagas más dañinas, *Helicoverpa armigera* tiene una amplia distribución geográfica con registros de presencia en Europa, Asia, África, Oceanía y América. Además de encontrarse en casi todos los continentes, esta especie tiene un hábito migratorio. Cuando

las condiciones de temperatura son adversas y hay escasez de alimentos, *H. armigera* puede migrar hasta cientos de kilómetros a lugares favorables para la supervivencia. Por tanto, el objetivo de este trabajo fue evaluar la influencia de la temperatura en el desarrollo y adaptación en tres progenies de *H. armigera*. Se concluyó que *H. armigera* puede adaptarse cerca de las temperaturas óptimas, alrededor de los 25°C y a bajas temperaturas, cuando oscilan como lo hacen en el campo. Además, probablemente puede desarrollarse fuera de temporada si permanece en diapausa durante un período prolongado.

### **Materiales y métodos**

Para realizar los bioensayos, la primera generación de *H. armigera* partió de adultos criados en laboratorio, los cuales fueron recolectados en cultivos de cítricos en la ciudad de Botucatu, SP. Los insectos se mantuvieron a una temperatura de 25°C, humedad relativa (HR) al 70% y fotofase 14 horas, en jaulas fabricadas en tubo de PVC, forradas con papel, que las polillas utilizan como sustrato para la oviposición. Para la dieta de los adultos se ofreció una dieta a base de miel, diluida al 30%. Periódicamente, se retiraron los huevos y se transfirieron a vasos desechables (500 ml). Después de la eclosión, las orugas se transfirieron individualmente a cápsulas de propileno de 6 cm de diámetro, y se ofreció como alimento una dieta artificial y una metodología de preparación (Greene, Leppla y Dickerson, 1976; Parra, 2001). Los tratamientos utilizados para evaluar el impacto de la temperatura en las características biológicas de *H. armigera* fueron siete temperaturas constantes (15 - 20 - 22,5 - 25 - 27,5 - 30 y 35°C) y dos temperaturas variables (13,5 - 25° C y 18 - 32°C), alternando 12 horas de luz y 12 horas de oscuridad, durante el período de tres generaciones. Las temperaturas variables fueron elegidas debido a la ocurrencia de *H. armigera* en dos estados brasileños: Paraná, con un registro de temperaturas más bajas (13.5 - 25°C) y Mato Grosso, con un registro de temperaturas más altas (18 - 32°C). En la primera generación de insectos se utilizaron 100 repeticiones por temperatura y para la segunda y tercera generación se evaluó el máximo número posible de repeticiones, según el desarrollo del insecto a las respectivas temperaturas.

### **Influencia de la temperatura sobre aspectos biológicos de *H. armígera***

La fase larvaria se evaluó diariamente, hasta que los insectos alcanzaron la fase prepupal. Para cada temperatura, la cantidad y duración de cada estadio fue evaluada por una muestra de 15 individuos. Además, se evaluó diariamente la mortalidad larvaria y la duración del período larvario de todos los insectos hasta que se alcanzó la fase prepupa.

En la fase prepupa, se evaluó diariamente la mortalidad y la duración del período prepupa de todos los insectos hasta que se alcanzó la fase de pupa. En la fase de pupa se evaluó la duración (días) del período, el cual se determinó desde la prepupa hasta la emergencia del adulto. Para calcular la viabilidad se cuantificó el número de adultos que emergieron en relación al número de pupas. Para la evaluación de la masa (g) de la pupa se utilizó una balanza analítica de precisión (Ohaus - EX124), considerando tres decimales y se consideró un intervalo de 24 horas después de la pupa, ya que las pupas pierden agua con el paso del tiempo. Después de pesar, las pupas se sexaron (macho / hembra) para determinar la proporción de sexos a cada temperatura.

En la fase adulta se evaluó la duración (días) del adulto, la cual se determinó desde la emergencia hasta la muerte y si existía deformidad en el adulto. Así, se calculó desde la fase huevo hasta la muerte del adulto, el número de días que duró cada generación de *H. armigera* a cada temperatura, iniciando así una nueva generación hasta llegar a la tercera generación. Los adultos fueron separados por sexos (machos y hembras) y transferidos a jaulas de PVC de 21 cm de alto por 16 cm de diámetro. Se colocó una tela transparente en la parte superior de cada jaula, que se cambiaba a diario. La dieta ofrecida para adultos fue una solución de miel al 30% empapada en algodón que se colocó en el fondo de la jaula sobre una cápsula de propileno de 6 cm de diámetro. Para facilitar la recolección de huevos, se colocó un papel a los lados de cada jaula, donde ocurrió la oviposición. La recolección de huevos se realizó a diario. Las parejas permanecieron en las jaulas hasta la muerte, tras lo cual fueron descartadas.

El número promedio de huevos se calculó considerando el número total de huevos durante la vida de las polillas dividido por el número total de hembras. Para calcular el número de orugas, se cuantifica el número de orugas que eclosionaron en relación con el número de huevos. Por lo tanto, se determinó la duración en días del período de oviposición hasta la eclosión para cada temperatura.

## **Análisis estadístico**

Inicialmente, los modelos de regresión de Cox (Lee, 1992) se ajustaron para modelar el tiempo hasta completar cada una de las cuatro fases por separado, considerando generaciones y temperaturas e interacciones como factores. Asimismo, los modelos de regresión de Cox para la modelización del tiempo se ajustaron para completar cada uno de los siete estadios por separado dentro de las primeras fases, considerando como factores las generaciones y las temperaturas e interacciones. Para las dos variables descritas anteriormente, se utilizaron estimadores de producto límite de Kaplan-Meier para estimar las curvas de supervivencia y la prueba Log-Rank ajustada por Sidak (Westfall *et al.*, 1999) para las comparaciones entre las curvas de supervivencia. Se utilizaron procedimientos Phreg para ajustar los modelos de regresión de Cox y el procedimiento de prueba de vida para ajustar los estimadores de producto límite de Kaplan-Meier. Se ajustaron modelos de regresión de Cox condicional de PWP multivariante (Prentice, Williams y Peterson, 1981) para modelar el tiempo hasta completar las cuatro fases consecutivas, considerando la temperatura y la masa larvaria como covariables continuas, según generaciones. En este modelado es posible obtener las funciones de riesgo de completar la generación. Para ajustar estos modelos, se utilizaron procedimientos phreg.

En el análisis de la deformación adulta y la proporción de sexos, se utilizó un modelo lineal generalizado con la distribución binomial y la función de enlace logit con generación y temperatura e interacciones como factores. Para el análisis de las variables número de huevos por hembra, número de orugas por hembra y masa de pupa, se utilizaron modelos lineales generalizados con la distribución gamma y función de enlace logarítmico (Nelder y Wedderburn, 1972; Diggle *et al.*, 2002), incluyendo como factores: generación, temperatura e interacciones. Para los modelos lineales generalizados, la calidad de los ajustes se realizó mediante el análisis residual y el análisis de desviación. Para las comparaciones entre tratamientos se utilizó la prueba de Tukey-Kramer del procedimiento Genmod del programa estadístico SAS.

**b. Ejemplo 2: Survival and behavior of the insecticide-exposed predators *Podisus nigrispinus* and *Supputius cincticeps* (Heteroptera: Pentatomidae). (Castro *et al.*, 2013)**

Las chinches pentatómidas son depredadores importantes de las orugas defoliantes en los sistemas agrícolas y forestales, y el conocimiento del impacto de los insecticidas en los enemigos naturales es información importante para los programas de manejo integrado de plagas (MIP). De esta manera, se evaluó la toxicidad y la respuesta subletal conductual de los depredadores *Podisus nigrispinus* y *Supputius cincticeps* expuestos a deltametrina, metamidofos, espinosad y clorantraniliprol, insecticidas comúnmente utilizados para controlar la oruga del frijol terciopelo (*Anticarsia gemmatalis*) en cultivos de soja. Con la excepción de la deltametrina para *S. cincticeps*, todos los insecticidas mostraron una mayor toxicidad aguda para la presa que para estos enemigos naturales, proporcionando un control eficaz de *A. gemmatalis*.

La concentración de campo recomendada para deltametrina, metamidofos y espinosad para controlar *A. gemmatalis* causó una mortalidad del 100% de las ninfas de *P. nigrispinus* y *S. cincticeps*. El clorantraniliprol fue el insecticida menos tóxico y más selectivo para estos depredadores, lo que resultó en una mortalidad de menos del 10% cuando se expuso a 10 veces la concentración de campo recomendada durante un período de 72 h.

Se encontraron cambios en el patrón de comportamiento de los depredadores para todos los insecticidas, especialmente metamidofos y espinosad, que mostraron irritabilidad (es decir, evitación después del contacto) con ambas especies de depredadores. Sin embargo, no se observó repelencia de insecticidas (es decir, evitación sin contacto) en ninguno de los insectos probados. Los efectos letales y subletales de los plaguicidas sobre los enemigos naturales son de gran importancia para el MIP, y los resultados indican que la sustitución de insecticidas piretroides y organofosforados en sus dosis de campo por clorantraniliprol puede ser un factor clave para el éxito de los programas MIP de *A. gemmatalis* en soja.

## **Metodología**

### **Insectos**

Estos enemigos naturales se crían con pupas del gusano de la harina amarillo *Tenebrio molitor* L. (Coleoptera: Tenebrionidae) en condiciones ambientales controladas ( $25 \pm 2^\circ\text{C}$ ,  $70 \pm 5\%$  de humedad relativa y 12:12 fotoperiodo luz: oscuridad). Los adultos y las larvas del gusano de la harina amarillo se

crían en una bandeja de plástico que contiene harina de trigo mezclada con levadura ( $\approx 5\%$ ) y verduras como zanahoria, camote y mandioca, como alimento y humectación una vez a la semana. Las orugas de *A. gemmatalis* se crían con dieta artificial y sus adultos en jaulas de madera (30 x 30 x 30 cm) con los costados reticulados, cubiertas de vidrio y se alimentan con algodón empapado en solución nutritiva en el fondo de las jaulas. Se observaron diariamente ninfas de *P. nigrispinus* y *S. cincticeps* y larvas de *A. gemmatalis* para obtener insectos del tercer estadio para su uso en los bioensayos.

### **Insecticidas**

Todos los insecticidas utilizados están registrados para el control de *A. gemmatalis* en campos de soja brasileños. Los insecticidas utilizados y sus respectivas formulaciones comerciales fueron: el piretroide deltametrina (Decis 25 EC; 25 g ia L<sup>-1</sup>), el organofosforado metamidofos (Tamaron BR SC; 600 g ia L<sup>-1</sup>), la diamida clorantropilprol (Premio CS; 200 g ia L<sup>-1</sup>) y el spinosyn spinosad (Tracer 480 CS; 480 g ia L<sup>-1</sup>).

### **Bioensayos concentración-mortalidad**

Los bioensayos de concentración-mortalidad se realizaron utilizando placas de Petri (9,0 cm de diámetro x 2,0 cm de alto) con el fondo completamente cubierto con hojas de soja del cultivar “BRSMT pintado”, tratadas con soluciones insecticidas. Para cada tratamiento, las hojas de soja se sumergieron durante cinco segundos a diferentes concentraciones de cada solución de insecticida (diluidas en agua) y luego se dejaron secar a la sombra durante una hora antes de colocarlas en las placas de Petri. Cada placa de Petri recibió diez larvas de tercer estadio de *A. gemmatalis* o diez ninfas de tercer estadio de *P. nigrispinus* o *S. cincticeps*. Los bioensayos se establecieron siguiendo un diseño completamente al azar con cinco a ocho concentraciones y seis repeticiones. Las concentraciones utilizadas se establecieron mediante bioensayos preliminares con un rango de diluciones de 10 veces para cada insecticida y especie para permitir el reconocimiento del rango de concentración que conduce a una variación de la mortalidad entre 0% y 100%. La mortalidad se evaluó después de 72 h de exposición y los insectos se consideraron muertos si no se movían cuando se les pinchaba con un cepillo de pelo fino. Los depredadores no fueron alimentados durante la exposición al insecticida en este bioensayo, ya que pueden sobrevivir hasta más de 14 días sin presas como fuente de alimento.



### Bioensayo tiempo – mortalidad con insecticidas bajo dosis de campo

La toxicidad aguda (letal) hacia las chinches depredadoras de las concentraciones máximas recomendadas de insecticida para el control de *A. gemmatalis* (clorantraniliprol-13,3  $\mu\text{g ai mL}^{-1}$ , deltametrina-50  $\mu\text{g ai mL}^{-1}$ , espinosad-240  $\mu\text{g ai mL}^{-1}$ , y metamidofos-1500  $\mu\text{g ia mL}^{-1}$ ) se estimó utilizando ninfas de tercer estadio de *P. nigrispinus* y *S. cincticeps*. Se colocaron diez ninfas de cada especie sobre el papel de filtro impregnado con insecticida pegado (con resina cola sintética blanca al agua) al fondo de una placa Petri (9 cm de diámetro x 2 cm de alto), cuyas paredes internas se cubrieron con teflón PTFE, para evitar el escape de insectos. El disco de papel de filtro se consideró tratado cuando se removió durante 5 s con 1 mL de solución correspondiente a cada concentración de campo recomendada para el insecticida. Se utilizaron cinco réplicas para cada combinación de insecticida y especies depredadoras, además de un tratamiento de control en el que solo se aplicó agua (destilada y desionizada) a los papeles de filtro. La mortalidad de insectos se observó cada 30 minutos durante la exposición inicial de 24 horas y luego a intervalos de 5 horas hasta la muerte de todos los insectos o hasta que alcanzaron la etapa adulta. Las pupas de *Tenebrio molitor* se proporcionaron *ad libitum* a las ninfas depredadoras de la chinche apestosa a lo largo de los bioensayos. Las ninfas se registraron como muertas si no podían moverse cuando se les pinchaba dorsalmente con un cepillo fino. Todos los bioensayos se realizaron simultáneamente en las mismas condiciones de la cría del insecto siguiendo un diseño completamente aleatorizado.

### Análisis estadístico

Los resultados de los bioensayos tiempo-mortalidad se sometieron a análisis Probit mediante PROC PROBIT, generando curvas de concentración-mortalidad y se calcularon las tasas de selectividad y toxicidad. Para medir la selectividad de insecticidas sobre especies depredadoras, se calculó la selectividad diferencial con intervalos de confianza del 95% en base a los valores de  $LC_{50}$  de insecticidas para plaga (*A. gemmatalis*) y depredadores (*P. nigrispinus* y *S. cincticeps*). Los datos de tiempo-mortalidad se sometieron a análisis de supervivencia mediante el procedimiento no paramétrico LIFE-TEST. Este procedimiento permite estimar las curvas de supervivencia obtenidas mediante estimadores de Kaplan-Meier generados a partir de la proporción de ninfas de tercer estadio que sobreviven desde el principio hasta el final del experimento. Los resultados generales de los bioensayos se sometieron a

análisis de varianza multivariante (PROC GLM utilizando la declaración MANOVA). Posteriormente, cada parámetro se sometió a un análisis univariado de varianza y a la prueba HSD de Tukey ( $p < 0,05$ ), cuando fue apropiado (PROC UNIVARIATE). Se probaron las diferencias por pares en el tiempo empleado en cada mitad de las arenas medio tratadas (es decir, evitación de insecticidas) utilizando la prueba t de Student para pares ( $p < 0,05$ ) para cada insecticida y especie. Se verificó la homogeneidad de la varianza y la normalidad de los errores y los datos se transformaron cuando fue necesario (PROC UNIVARIATE; GPLOT PROC).

**c. Ejemplo 3: Lethal and sub-lethal effects of *Beauveria bassiana* (Cordycipitaceae) strain NI8 on *Chrysoperla rufilabris* (Neuroptera: Chrysopidae) (Portilla, Snodgrass y Luttrell, 2017)**

Se hizo un bioensayo utilizando una cepa nativa del Delta del Mississippi (NI8 ARSEF8889) de *Beauveria bassiana* (Bals. -Criv.) Vuill. (Cordycipitaceae), aislada de *Lygus lineolaris* (Palisot de Beauvois) (Hemiptera: Miridae), en una crisopa, *Chrysoperla rufilabris* (Burmeister) (Neuroptera: Chrysopidae) a 4 concentraciones de pulverización ( $7,02 \times 10^4$ ,  $10^5$ ,  $10^6$ , y  $10^7$  esporas por mL) para evaluar su efecto sobre la tasa reproductiva y duración de la vida de los adultos de este depredador de insectos.

El método de aplicación simuló pulverización atomizada y las concentraciones probadas fueron similares a las utilizadas para medir el impacto del hongo en *L. lineolaris*. Se encontraron efectos significativos de *B. bassiana* en adultos de *C. rufilabris* y la gravedad del impacto dependió de la concentración probada. *B. bassiana* impactó todas las medidas demográficas de la reproducción y sobrevivencia de *C. rufilabris*. La tasa de aumento intrínseco y finito y las tasas de reproducción bruta y neta de adultos tratados con las concentraciones de prueba más altas disminuyeron significativamente, mientras que el tiempo de duplicación aumentó para adultos tratados con las concentraciones de pruebas más bajas.

Basándose en estas observaciones, *C. rufilabris* si será afectado por aerosoles de *B. bassiana* dirigidos a *L. lineolaris* si los adultos están presentes en el momento y lugar del tratamiento. La concentración letal medida,  $CL_{50}$ , de 2,11 esporas viables por  $mm^2$  se compara con una  $CL_{50}$  de 2,75 esporas por  $mm^2$  determinada previamente para *L. lineolaris*. Se requirieron mayores concentraciones de esporas por  $mm^2$  para la respuesta de esporulación ( $RS_{50}$ ) del hongo entomopatógeno sobre *C. rufilabris* (13,60

esporas viables por  $\text{mm}^2$ ) de las concentraciones que se necesitaron para la respuesta de mortalidad ( $\text{LC}_{50}$ ).

## Metodología

### Colonias de *Chrysoperla rufilabris*

Los adultos de *Chrysoperla rufilabris* fueron aproximadamente, 400 adultos (2-3 días de edad). Para asegurar la copulación, los insectos se mantuvieron colectivamente en el recipiente original obtenido del proveedor comercial (caja de cartón cilíndrica de 3 L cubierta con tela de algodón de organdí). Se colocó una esponja con solución de azúcar-agua (10%) en una placa Petri de 11 cm de diámetro dentro de la jaula. Los insectos se mantuvieron en una cámara de crecimiento a  $25^{\circ}\text{C}$ , 55% de humedad relativa (HR) y un fotoperiodo de 12:12 h L: D hasta que se observó la primera oviposición.

### Cultivo de *Beauveria bassiana* cepa N18

La cepa N18 de *B. bassiana* se obtuvo de fuentes almacenadas de polvo de esporas, producida con regularidad para el programa de investigación de *L. lineolaris* (Portilla et al., 2016). La concentración de inóculo ( $1,20 \times 10^{11}$  esporas por g) se suspendió en 50 ml de Tween-80 al 0,04% y se diluyó para obtener concentraciones finales de  $7 \times 10^7$  esporas por ml. La viabilidad del inóculo se midió según la metodología de Portilla et al. (2014) y Portilla et al. (2016). Las concentraciones de prueba más bajas ( $7 \times 10^4$ ,  $10^5$  y  $10^6$ ) para este estudio se extrapolaron en base a la dilución de la concentración más alta ( $7 \times 10^7$ ). Los datos resultantes se analizaron mediante análisis de varianza. Alícuotas (6 ml) de la suspensión de mayor concentración ( $7 \times 10^7$ ) proporcionaron 395 esporas viables por  $\text{mm}^2$  en el área de pulverización objetivo cuando se aplicaron utilizando una torre de pulverización de laboratorio Potter Precision, siguiendo los procedimientos de Portilla et al. (2014).

### Procedimientos para bioensayos

Se prepararon diluciones en serie de 4 concentraciones de prueba de la cepa N18 ( $1,2 \times 10^4$ ,  $10^5$ ,  $10^6$  y  $10^7$  esporas por ml) para tratar las hembras de *C. rufilabris* y evaluar el efecto de la cepa N18 en sus tasas de reproducción. Para evitar la infección cruzada, solo se utilizaron hembras adultas de 4 a 5 días de edad recibidas del proveedor comercial. Las hembras seleccionadas se rociaron con N18 usando el

método de inoculación directa (administración por rociado atomizado) descrito por Portilla *et al.* (2014) y Portilla *et al.* (2016). Los insectos tratados se mantuvieron en una cámara de crecimiento a 25°C, 55% de HR y un fotoperíodo de 12:12 h L: D. Cada tratamiento de ensayo (concentración individual) se replicó 4 veces con 10 hembras adultas por réplica (200 hembras en total). Los insectos de control se rociaron de manera similar (12,5 kPa por pulgada<sup>2</sup>) con 6 ml de agua (control de agua). Los tratamientos de concentraciones de NI8 ( $1,20 \times 10^4$ ,  $10^5$ ,  $10^6$  y  $10^7$  esporas por ml) se administraron de manera similar en un volumen de pulverización de 6 ml de solución de *B. bassiana*. Después de la aplicación, las hembras de *C. rufilabris* se colocaron individualmente en vasos de 29,7 ml con una dieta sólida desarrollada para bioensayos de *L. lineolaris* (Portilla *et al.*, 2014). No se proporcionó ninguna fuente de alimento adicional. Las hembras se examinaron diariamente para determinar la mortalidad y la oviposición. Se contó el número de huevos depositados cada día por cada hembra. Las hembras con huevos se retiraron y se colocaron en una nueva taza de dieta hasta que murió la última hembra. Los insectos muertos se retuvieron en tazas de dieta individuales durante 10 días y se observaron diariamente para detectar la esporulación.

### **Análisis estadístico**

Para comparar los parámetros de la tabla de fertilidad en *C. rufilabris* rociados con diferentes concentraciones de *B. bassiana* se utilizó un ANOVA de una vía seguido de la prueba de diferencia significativa de Tukey. Las estimaciones no paramétricas de la función de supervivencia de las hembras de *C. rufilabris* se compararon entre tratamientos mediante el procedimiento PROC LIFETEST. Las diferencias estadísticas en la supervivencia de las hembras de *C. rufilabris* se declararon con base en la estadística de rango logarítmico y utilizando el procedimiento PROC GLM para detectar diferencias entre las concentraciones a los 3, 5 y 10 días después de la aplicación. Los datos de mortalidad y esporulación para cada grupo de hembras de *C. rufilabris* y cada concentración fueron analizados por PROBIT, utilizando el logaritmo común (log a la base 10) del valor de concentración.

#### **d. Ejemplo 4: Bioactivity of essential oils from *Artemisia* against *Diaphania hyalinata* and its selectivity to beneficial insects (Lopez *et al.*, 2018)**

La demanda de insecticidas eficaces en el control de plagas con baja toxicidad para los enemigos naturales y los polinizadores, está aumentando constantemente. Una buena alternativa para los insecticidas sintéticos son los compuestos naturales, incluidos los aceites esenciales (OE).

Este trabajo evaluó la toxicidad de los aceites esenciales extraídos de *Artemisia annua*, *A. absinthium*, *A. camphorata*, *A. dracunculus* y *A. vulgaris* contra el gusano del melón *Diaphania hyalinata* (Lepidoptera: Crambidae), una plaga de Cucurbitaceae, y su selectividad para la hormiga de fuego *Solenopsis saevissima* (Smith) (Hymenoptera: Formicidae) y la abeja jataí *Tetragonisca angustula* (Latreille) (Meliponinae).

Las plantas para la extracción de OE se cultivaron en un invernadero con fertilización mineral. Los insectos del bioensayo pertenecían al segundo estadio de *D. hyalinata* y formas adultas de *S. saevissima* y *T. angustula*. El aceite esencial de *A. annua* indujo una alta tasa de mortalidad en *D. hyalinata* (96%) durante un período de 48 h. El mismo aceite esencial fue selectivo para el depredador *S. saevissima* (42% de mortalidad) y el polinizador *T. angustula* (74% de mortalidad), mientras que causó una alta mortalidad en *D. hyalinata*. La actividad insecticida del aceite de *A. annua* se atribuyó al sinergismo de sus constituyentes, a saber, alcanfor y 1,8-cineol. Por lo tanto, este aceite esencial contiene componentes que son prometedores para un uso eficaz como insecticida debido a su alta toxicidad y rápida acción contra *D. hyalinata*, así como a su baja toxicidad para depredadores y polinizadores.

#### **Metodología.**

##### **Cultivos de *Artemisia* spp. y extracción de aceites esenciales**

Las plantas fueron cultivadas en invernadero durante 8 meses. Las especies seleccionadas fueron *Artemisia annua*, *Artemisia absinthium*, *Artemisia camphorata*, *Artemisia dracunculus* y *Artemisia vulgaris*. La siembra se realizó utilizando vasos de plástico de 300 mL que contenían un sustrato comercial para acelerar el enraizamiento y su posterior trasplante a las macetas.

Las plantas de cada maceta se fertilizaron con 6,5 g de sulfato de amonio, 40 g de fósforo y 3,0 g de cloruro de potasio. Se hidrodestilaron hojas frescas trituradas (100 g) de cada especie vegetal en un aparato Clevenger durante 2 horas y se extrajo el hidrolato con pentano (3 x 40 mL). Las fases orgánicas combinadas se secaron sobre sulfato de magnesio anhidro y se filtraron. El filtrado se concentró a presión reducida en un rotavapor y el residuo se almacenó en viales de vidrio a -4°C.

### **Análisis cualitativo y cuantitativo de aceites esenciales**

El análisis cuantitativo se realizó en un aparato Shimadzu GCMS5050 equipado con una columna capilar DB-5 (30 mx 0,25 mm, con un espesor de película de 0,25  $\mu\text{m}$ ) acoplada con el detector de ionización de llama. El helio fue el gas portador usado a un caudal de 1.8 mL min<sup>-1</sup> y la temperatura del inyector se fijó en 220°C. La temperatura inicial de la columna fue de 40°C. La isoterma se mantuvo durante 2 min, seguido de calentamiento desde 3°C min<sup>-1</sup> hasta 240°C. La isoterma se mantuvo durante 15 min antes de inyectar 1.0  $\mu\text{L}$  de la muestra [1% de la muestra en CH<sub>2</sub>Cl<sub>2</sub> (m / v)], en una relación de división de 1:20 y una presión de columna de 100 kPa. El análisis cualitativo se realizó en un aparato Shimadzu GC17A equipado con una columna capilar DB-5 (30 mx 0,25 mm, con un espesor de película de 0,25  $\mu\text{m}$ ) utilizando ionización por impacto de electrones (70 eV). La temperatura del inyector, la isoterma, la presión de la columna y la velocidad de calentamiento fueron similares a los del análisis GC / FID. El análisis se realizó por triplicado y la concentración de cada constituyente se calculó como porcentaje del área del pico correspondiente al área total de todos los picos. Los componentes se identificaron comparando los tiempos de retención, en relación con series de alcanos (C<sub>9</sub> - C<sub>27</sub>) y comparando los espectros de masas con la base de datos de bibliotecas de Wiley y Nist (05, 08 y 11).

### **Bioensayos con insecticidas**

Se obtuvieron larvas de *D. hyalinata* de segundo estadio de una población mantenida en el laboratorio. Se recolectaron adultos de *S. saevissima* y *T. angustula* de nidos ubicados en terrenos exteriores. El peso promedio de cada insecto se estimó midiendo la masa de diez insectos en una balanza analítica. Los bioensayos se realizaron mediante aplicación tópica. Cada solución se preparó en acetona y se aplicó (0,5  $\mu\text{L}$ ) en el tergum abdominal de cada insecto individual utilizando una microjeringa de Hamilton (10  $\mu\text{L}$ ). Para el control negativo, los insectos se trataron con un volumen igual de acetona. Co-

mo control positivo, se utilizó el insecticida comercial aceite de neem ( $12 \text{ g L}^{-1}$  de azadiractina). El aceite de neem se utilizó como estándar de eficacia porque la azadiractina es el componente natural que se utiliza con mayor frecuencia para el control de las lombrices. Además, el aceite de neem es un insecticida de origen natural, similar a los productos evaluados en este trabajo.

En este trabajo se realizaron cinco bioensayos. En el primero se seleccionó el aceite esencial con mayor actividad insecticida frente a *D. hyalinata*. En el segundo y tercer bioensayos se determinaron las curvas dosis-mortalidad y la velocidad de acción del aceite seleccionado sobre *D. hyalinata*. En el cuarto bioensayo se determinaron los compuestos responsables de la actividad insecticida del aceite esencial seleccionado. En el quinto bioensayo, se determinó la selectividad del aceite esencial para depredador y abeja.

### **Selección de aceites esenciales con actividad insecticida contra *D. hyalinata***

El diseño experimental fue completamente al azar con seis repeticiones. Cada réplica consistió en una placa de Petri (9 cm de diámetro) que contenía orugas de segundo instar de *D. hyalinata*. La dosis utilizada en este bioensayo fue de  $20 \mu\text{g}$  del compuesto por mg de larva. Los tratamientos comprendieron aceite de neem (control positivo), control negativo (acetona) y aceites esenciales de *A. annua*, *A. absinthium*, *A. camphorata*, *A. dracuncululus*. Después de la aplicación, se colocó un disco de hoja de chayote dentro de la placa de Petri para alimentar a las orugas. Las placas de Petri se colocaron en una incubadora a  $25 \pm 0,5^\circ\text{C}$  y  $75 \pm 5\%$  de humedad relativa con un fotoperiodo de 12 h. Se evaluó la mortalidad de insectos 48 h después de la aplicación de los tratamientos. Los datos de mortalidad de larvas se sometieron a análisis de varianza y las medias se compararon mediante la prueba de Tukey a  $p < 0,05$  (PROC GLM). Los aceites esenciales que indujeron más del 80% de mortalidad en *D. hyalinata* fueron seleccionados para el siguiente bioensayo. Estos criterios se utilizan en Brasil para evaluar la eficacia de un producto en el control de plagas (Bacci *et al.* 2007).

### **Curva dosis – mortalidad para aceites esenciales contra *D. hyalinata***

El diseño experimental, las condiciones y las evaluaciones fueron similares a los del bioensayo anterior. Los tratamientos incluyeron dosis de aceites esenciales seleccionados en el bioensayo y controles previos. Se utilizaron seis dosis (2,5 - 5,0 - 12,5 - 15,0 - 17,5 - 20 mg g<sup>-1</sup> de insecto). La mortalidad



de los insectos se evaluó a las 48 h de la aplicación del tratamiento, ya que este tiempo permite determinar la dosis para matar el 50% de los insectos (LD<sub>50</sub>) mediante el análisis probit. Los compuestos más activos fueron sometidos a bioensayos de toxicidad contra *D. hyalinata* siguiendo el diseño experimental aleatorizado, con seis repeticiones. Cada unidad experimental incluyó diez insectos en una placa de Petri (9 cm de diámetro) cubierta de organza. Las curvas de dosis-mortalidad para *A. annua* y aceite de neem se construyeron usando cuatro y seis dosis, respectivamente. Estas dosis se establecieron mediante bioensayos preliminares con cuatro concentraciones para cada compuesto para identificar el rango de concentración que induce una mortalidad mayor que cero y menor que 100%. Los datos de dosis-mortalidad se sometieron a análisis probit (PROC PROBIT) para estimar las curvas de dosis-mortalidad. Se aceptaron las curvas que presentaban probabilidades mayores a 0,05 por la prueba  $\chi^2$ . También se estimaron las dosis letales que indujeron 50 y 95% de mortalidad (LD<sub>50</sub> y LD<sub>95</sub>).

#### **Velocidad de acción del aceite esencial seleccionado contra *D. hyalinata***

Los insectos de *D. hyalinata* se sometieron a bioensayos de tiempo-mortalidad cuando se expuso al LD<sub>90</sub> del aceite más activo (aceite esencial de *A. annua*) y control. Se sometieron cien insectos a cada tratamiento. Cada unidad experimental involucró diez insectos en una placa de Petri (9 cm de diámetro) cubierta de organza. La mortalidad de las larvas se registró después de 48 h. Los intervalos entre las evaluaciones para cada tratamiento (que van de 30 min a 3 h) se determinaron antes. Las curvas de supervivencia se estimaron mediante el método de límite de producto de Kaplan-Meier y se compararon mediante la prueba de rango logarítmico en  $p < 0,05$  (PROC LIFETEST). También se estimó el tiempo medio letal (LT<sub>50</sub>) de los insectos.

#### **e. Ejemplo 5: Survival and feeding avoidance of the eucalyptus defoliator *Thyriniteina arnobia* exposed to the proteinase inhibitor berenil (Marinho *et al.*, 2012)**

Los inhibidores de proteinasas son reconocidos como potenciales agentes fitosanitarios contra las plagas de insectos y su uso es una alternativa para el manejo integrado de plagas. Berenil es un ejemplo de inhibidor de tripsina sintético y se evaluó su potencial de uso como insecticida contra *Thyriniteina arnobia* (Stoll) (Lepidoptera: Geometridae), uno de los principales defoliadores de eucalipto en Brasil.

Se evaluaron la supervivencia de los insectos y los rasgos de su ciclo de vida, incluido el tiempo de



desarrollo, el peso y la viabilidad de la pupa, y los parámetros de la tabla de vida de *T. arnobia* en larvas criadas en hojas de eucalipto que contienen 0,00%, 0,06%, 0,12%, 0,25%, 0,50% y 0,75%. (p / v) del inhibidor de tripsina sintético berenil. Además, también se evaluó la preferencia alimentaria y el consumo de hojas de las orugas de *T. arnobia*. Berenil retrasó el desarrollo larvario. La supervivencia de las larvas se vio gravemente afectada por el berenil, que también retrasó el desarrollo larvario. Las concentraciones subletales de berenil comprometieron los parámetros de la tabla de vida de *T. arnobia* reduciendo su tasa neta de reproducción y la tasa de crecimiento de la población, al tiempo que extendieron el tiempo de generación. Berenil también disuadió a *T. arnobia* y no provocó la alimentación compensatoria de las larvas. La actividad del berenil como insecticida, repelente y antialimentario contra *T. arnobia* sugiere su uso potencial contra las orugas defoliantes del eucalipto.

## **Materiales y métodos**

### **Insectos, plantas y productos**

La colonia de insectos utilizada se estableció a partir de huevos de *T. arnobia* obtenidos de una colonia de laboratorio, mantenida en condiciones controladas ( $26 \pm 3$  ° C; 12 h / 12 h L / D;  $60 \pm 15\%$  h.r.). En los experimentos se utilizaron plantas de *Eucalyptus grandis* libres de residuos del pesticida Berenil (acetato de diminazeno) y el dispersante agrícola Gotafix ( $C_{35}H_{64}N_{11}$ ), que se utilizó para mejorar la adherencia y penetración del PI en la superficie de la hoja.

### **Experimento de tabla de vida**

El experimento se estableció en tres bloques, cada uno conteniendo 13 réplicas de cada concentración del inhibidor de tripsina berenil (0%, 0,06%, 0,12%, 0,25%, 0,5% y 0,75% p / v) usando agua como solvente y conteniendo 0,06 % (v / v) Gotafix. Cada réplica incluía una oruga recién emergida, individualizada en un plato de plástico (9 cm de diámetro). Se utilizaron treinta y nueve larvas por concentración de berenil. Se utilizaron plantas de *E. grandis* (50 cm de altura); las hojas de la planta se sumergieron en solución de berenil y se colocaron dentro de la placa de Petri, donde posteriormente se soltaron las orugas. El pecíolo de la hoja estuvo involucrado en algodón húmedo cuando se cosechó para retener la humedad. Cuando las larvas alcanzaron el quinto estadio, se transfirieron a vasos de plástico (500 ml). La mortalidad de las larvas se evaluó diariamente; También se registraron el peso de la

pupa y la viabilidad de la pupa. Durante la etapa adulta, las polillas se emparejaron y colocaron en una jaula cilíndrica (8,5 cm de diámetro y 19 cm de altura), cerrada con una placa de Petri (9 cm de diámetro) y con sus paredes internas cubiertas con papel de filtro. Cada pareja se consideró como una réplica. Para obtener suficientes larvas para estudios de pupa y estadios adultos (por posible mortalidad), se mantuvieron 120 insectos en las mismas condiciones de cada tratamiento para su eventual reposición, en caso de ser necesario.

### **Análisis estadístico**

Las curvas de supervivencia se obtuvieron mediante estimadores de Kaplan-Meier generados a partir de la proporción de orugas que sobreviven cada día desde el inicio hasta la conclusión del experimento utilizando el procedimiento LIFETEST de SAS. Los insectos que sobrevivieron hasta la emergencia de los adultos se trataron como datos censurados (Allison, 1995). Las curvas de supervivencia de cada concentración de berenil se compararon mediante el método de regresión de Cox (procedimiento PHREG).

#### **3.2.4. Impacto de enfermedades fungosas y bacterianas en el crecimiento y supervivencia de las plantas**

La aplicación del análisis de supervivencia en la patología de plantas ha sido limitada. Un ejemplo clásico es el de Westra et al. (1994), quienes para optimizar el tiempo de las inspecciones previas a la cosecha por pudrición anular bacteriana, causada por *Clavibacter michiganensis* subsp. *sepedonicus* en campos de producción de semilla de papa, desarrollaron modelos de análisis de supervivencia para predecir el tiempo desde la siembra hasta (i) el inicio de los síntomas foliares de la enfermedad y (ii) la ocurrencia de la máxima incidencia de síntomas foliares en el campo. El análisis se basó en datos recopilados de tres cultivares en diversas regiones productoras de papa de los Estados Unidos durante un período de 3 años.

Dependiendo de la ubicación y el cultivo, los modelos predijeron 50 a 72 y 83 a 115 días después de la siembra para las probabilidades de supervivencia,  $p = 0,95$  y  $0,05$ , para el inicio de los síntomas y la incidencia máxima de la enfermedad, respectivamente. En otras palabras, a los 50 a 72 días después de la siembra, el 95% de las parcelas habían “sobrevivido” sin mostrar síntomas foliares de la enfermedad,

mientras que a los 83 a 115 días después de la siembra, solo el 5% de las parcelas aún no habían alcanzado la máxima incidencia de la enfermedad. Debido a que los modelos de supervivencia proporcionan estimaciones de la probabilidad de aparición de síntomas y la incidencia máxima de la enfermedad en momentos determinados durante la temporada de crecimiento, permiten la predicción de estos dos eventos y la programación óptima de las inspecciones de los campos de producción de semillas de papa para detectar la pudrición anular bacteriana.

#### **a. Ejemplo 1: Applications of survival analysis in botanical epidemiology (Scherm y Ojiambo, 2004)**

Este trabajo, utiliza la mancha foliar en plantas de arándano, causada por *Septoria* del arándano. La mancha foliar, causada por el hongo *Septoria albopunctata*, puede ser perjudicial para ciertos cultivares de arándanos cultivados en Georgia y otros estados del sureste de Estados Unidos (Cline, 2002; Scherm *et al.*, 2003). La enfermedad es más severa después de la cosecha a finales del verano y otoño, cuando no hay frutos. La infección de la hoja durante este período reduce la fotosíntesis, originando una defoliación prematura, los cuales pueden tener un impacto negativo en la iniciación de los botones florales en el otoño y rendimiento de frutos en la primavera siguiente (Roloff *et al.*, 2004; Brannen *et al.*, 2003; Cline, 2002; Ojiambo *et al.*, 2002). Basado sobre estas consideraciones, se iniciaron experimentos de campo para cuantificar las relaciones entre la gravedad de la mancha foliar, la defoliación y el rendimiento en arándanos del sur. Se utilizan los datos recopilados durante este estudio para demostrar la aplicación de varios procedimientos de análisis de supervivencia, centrándose en cuantificar los efectos de las covariables relacionadas con el huésped y la enfermedad a tiempo para defoliación de hojas individuales.

#### **Recolección de datos**

El estudio se realizó en una plantación experimental de arándano rabbiteye maduro "Premier" (*Vaccinium ashei*). Las plantas quedaron sin tratar con fungicidas durante toda la temporada. Se seleccionaron y marcaron 50 sesiones con diferentes niveles de severidad de la mancha foliar, y a cada hoja presente en los 20 cm distales de estos brotes se le asignó un número para realizar un seguimiento a lo largo del tiempo. Las poblaciones totales de las hojas monitoreadas consistieron en 567 individuos. Se

hizo una distinción entre las hojas de la mitad superior del segmento del brote y las de la parte inferior del brote, con valores de 0 y 1 asignados a hojas inferiores y superiores, respectivamente. Con intervalos entre 6 a 12 días durante el resto del otoño, se evaluaron las hojas individualmente para la gravedad de la enfermedad (expresada como número de manchas por hoja) y tiempo de defoliación (expresado en días posteriores al inicio de la marcación). La última evaluación de la enfermedad se realizó al completar 35 días del inicio del experimento, después de lo cual el conteo de manchas foliares se volvió imposible debido a la senectud. La evaluación por defoliación continuó durante otras 2 semanas. La defoliación avanzó lentamente durante el primer mes del estudio seguido de una tasa más rápida de pérdida de hojas. Al final del estudio, el 33,9% de las hojas aún estaban presentes (observaciones censuradas). La gravedad promedio de la enfermedad se duplicó desde la primera a la tercera evaluación seguida de una disminución hacia la última fecha de evaluación por defoliación de las hojas afectadas más severamente.

### **Análisis estadístico**

Ya que el conjunto de datos contenía múltiples observaciones censuradas (es decir, hojas que no se habían defoliado al final del estudio) y debido a que el tiempo hasta la defoliación, la variable dependiente, puede interpretarse como un “tiempo de sobrevivencia”, se utilizaron técnicas de análisis de sobrevivencia para describir y modelar los datos (Dungan *et al.*, 2003).

Un supuesto fundamental del análisis de sobrevivencia es que las observaciones de diferentes individuos sean estadísticamente independientes del tiempo. Esto fue una preocupación en el estudio ya que los tiempos de defoliación de las hojas ubicadas en el mismo brote podrían resultar correlacionadas. Para probar la validez del supuesto de independencia, utilizaron el procedimiento ARIMA para determinar si los tiempos de defoliación de las hojas dentro de los brotes estaban autocorrelacionados. Solo 3 de los 50 brotes mostraron significancia ( $p < 0,05$ ) en patrones de autocorrelación, lo que sugiere que las observaciones de diferentes hojas en el mismo brote fueron independientes.

### **Funciones de sobrevivencia y riesgo**

Como primer paso para análisis de sobrevivencia, es útil estimar las distribuciones del tiempo de sobrevivencia del evento de interés. Esto permite calcular cantidades derivadas (como la vida media re-

sidual) y comparar distribuciones del tiempo de supervivencia entre los grupos de tratamiento. En el análisis de supervivencia son fundamentales dos funciones matemáticas, la función de superviviente,  $S(t)$ , y la función de riesgo,  $h(t)$ .  $S(t)$  es un tipo de función de distribución acumulativa que describe la probabilidad de que un individuo presente en el tiempo 0 sobrevivirá hasta el tiempo  $t$ . Es por lo tanto definido como  $S(t) = \Pr(T \geq t)$ , donde  $T$  es el momento del evento (por ejemplo, muerte) para el individuo de interés. Por el contrario,  $h(t)$  es la función de densidad de probabilidad condicional que describe el instante riesgo de que el evento ocurra en el momento  $t$ , dado que no ha ocurrido previamente.

El riesgo es un número adimensional que puede interpretarse como número de eventos por intervalo de tiempo (1). Por lo tanto, tomando la inversa del riesgo da el tiempo esperado hasta que ocurra el evento. Es importante señalar que, a diferencia de  $S(t)$ ,  $h(t)$  no está delimitado por un límite de 1.

Debido a que se realizaron evaluaciones a intervalos que contienen múltiples eventos de defoliación, se utilizó el método de tabla de vida (PROC LIFETEST) para obtener estimaciones de las funciones de supervivencia y riesgo. Como se esperaba, la función de superviviente estimada se parece mucho a la curva de defoliación. La función de riesgo mostró que el riesgo instantáneo de defoliación se mantuvo muy bajo durante aproximadamente 1 mes después del inicio del estudio, tras lo cual aumentó exponencialmente. La vida media residual de las hojas en el inicio del estudio fue de 46,4 días; este valor disminuyó a 6,5 días para las hojas aún presentes al final del estudio.

### **Comparación entre grupos**

Se comparan las funciones de superviviente entre dos o más grupos de tratamiento utilizando la prueba de chi-cuadrado, de las cuales las pruebas de rango logarítmico, Wilcoxon y razón de verosimilitud son las más comunes. Brevemente, la prueba de razón de verosimilitud se basa en el supuesto de que la función de riesgo es constante en cada grupo, lo que a menudo es inverosímil con datos biológicos. La prueba de rango logarítmico es más apropiada cuando se cumplen los supuestos de riesgos proporcionales y cuando hay censura limitada. La prueba de Wilcoxon es generalmente superior en todas las demás situaciones.

Se aplicó el procedimiento LIFETEST para comparar funciones de superviviente de las hojas colocadas en la mitad superior de los brotes (es decir, hojas más jóvenes;  $n = 286$ ) y los de las secciones inferiores de los brotes ( $n = 281$ ). Las funciones de superviviente y de riesgo para los dos grupos de hojas resultaron muy similares durante el primer mes del estudio, después de lo cual aumentó el riesgo de defoliación de las hojas inferiores en relación con el de las hojas superiores. La diferencia en la supervivencia entre los dos grupos fue relativamente pequeña pero estadísticamente significativa ( $p < 0,0001$ ) según la prueba de rango logarítmico y la prueba de Wilcoxon. La vida media residual de las hojas inferiores y hojas superiores al final del estudio fue de 6,1 y 6,8 días, respectivamente.

En una segunda comparación de grupos, se comparó el riesgo de defoliación para hojas con baja gravedad de la enfermedad al comienzo del estudio (con  $\leq 10$  manchas foliares de *Septoria* en la primera fecha de evaluación;  $n = 256$ ) y los que estaban gravemente infectados en ese momento ( $> 10$  puntos en la primera evaluación;  $n = 311$ ). La diferencia en supervivencia entre los dos grupos fue pronunciada y altamente significativa ( $p < 0,0001$ ), lo que sugiere que la infección grave por *S. albopunctata* es uno de los principales contribuyentes a la defoliación prematura. De hecho, al final del estudio, el riesgo de defoliación de las hojas iniciales superiores, la severidad de la mancha fue más de cinco veces mayor que la de las hojas con baja gravedad inicial de la enfermedad.

### **Modelado de datos de supervivencia**

El enfoque más poderoso para el análisis de supervivencia se basa en el uso de modelos para describir y predecir los efectos de covariables discretas y / o continuas sobre los tiempos de supervivencia en una población. El modelo de regresión de Cox, es más flexible y fácil de usar porque no requiere que el investigador asuma una distribución de probabilidad particular representando tiempos de supervivencia. Este modelo se basa en la premisa de que, para una covariable determinada, las funciones de riesgo en diferentes niveles de esa covariable son proporcionales o similares en forma para diferentes niveles de la covariable y solo resulta afectado el riesgo relativo de la ocurrencia del evento en cualquier momento dado. El modelo AFT es más apropiado en situaciones donde estos supuestos no son satisfechos, por ejemplo, si la covariable modifica principalmente el tiempo de períodos de alto riesgo (Fox 2001). En la práctica, es relativamente sencillo probar si el supuesto de riesgo proporcional (PH) se cumple desarrollando un Modelo de PH que incluye un término de interacción entre el tiempo y la covariable

de interés, además de la propia covariable. Si el término de interacción es estadísticamente significativo, el modelo AFT es más apropiado.

Para el estudio, se consideraron dos covariables al desarrollar un modelo para el tiempo de defoliación de las hojas de arándano, posición de la hoja en el brote (una variable binaria) y gravedad de la enfermedad en la última fecha de evaluación antes de la defoliación (una variable continua). Cuando se ajustó un modelo de PH a los datos utilizando el procedimiento PHREG en SAS, los términos de interacción entre ambas covariables y el tiempo resultaron estadísticamente significativos ( $p < 0,05$ ). Por lo tanto, se rechazó el enfoque PH y en su lugar se siguió un modelo AFT. El modelo fue implementado con PROC LIFEREG usando el modelo Weibull para describir la distribución del tiempo de supervivencia subyacente. Esta elección de la distribución de probabilidad se basó en la observación que  $h(t)$  aumentó a un ritmo creciente con el tiempo, una propiedad consistente con una distribución de Weibull de tiempos de supervivencia (Allison, 2010). Eventos de defoliación dentro de dos fechas de evaluación sucesivas fueron tratados como censurados por intervalo en el análisis porque se desconocía la fecha de defoliación dentro del intervalo de evaluación.

El modelo AFT resultante describió bien los datos; los parámetros y pruebas estadísticas asociadas revelaron que el efecto de la enfermedad sobre el riesgo de defoliación fue muy significativo ( $p < 0,0001$ ), mientras que la posición de la hoja no lo fue ( $p = 0,1946$ ). Este último resultado entra en conflicto con el obtenido por comparación directa de las funciones de supervivencia de las hojas colocadas en la parte inferior y superior de las secciones de los brotes. La explicación más probable es que la gravedad de la enfermedad y la posición de la hoja se confundieron, es decir, la enfermedad progresó más rápidamente y / o alcanzó mayores niveles en las hojas (más viejas). Por lo tanto, la posición de la hoja ya no era significativa considerado conjuntamente con la gravedad de la enfermedad.

La estimación del parámetro AFT para la gravedad de la enfermedad fue  $\beta = -0,0020$ . Recordando que la variable dependiente en el modelo AFT es  $\log_e T$ , este valor se puede usar para calcular el cambio porcentual en el tiempo hasta la defoliación por cada hoja manchada adicional basada en la fórmula  $100\% \times (e^\beta - 1)$  (Allison, 2010); basado en este cálculo, cada lugar adicional durante el período de evaluación disminuyó el tiempo de defoliación en un 0,2%. Del mismo modo, el porcentaje de cambio en el riesgo de defoliación por cada aumento de una unidad en la severidad de la hoja se puede calcular

como  $100\% \times (e^{-\beta/\sigma} - 1)$ , donde  $\sigma$  es el parámetro de escala de la función de Weibull; esta conversión indica que cada mancha foliar adicional aumenta el riesgo de defoliación en un 3,3%. Debido a que las manchas foliares causadas por *S. albopunctata* pueden alcanzar cantidades sustanciales en hojas de cultivares de arándanos rabbiteye (el máximo en el presente estudio fue de 317 manchas por hoja, y > 25% de las hojas tenían 100 o más manchas), el análisis reveló un considerable y muy significativo efecto cuantitativo de la enfermedad sobre el riesgo de defoliación.

**b. Ejemplo 2: Survival analysis to determine the length of latent period of *Mycosphaerella pinodes* on peas (*Pisum sativum* L.) (Setti, Bencheikh, Henni y Claire, 2010)**

El tizón causado por *Mycosphaerella pinodes* es una de las enfermedades más devastadoras del guisante que causa pérdidas de rendimiento de más del 50% y puede causar la pérdida total del cultivo en condiciones epidémicas. Durante este estudio, se desarrolló una herramienta estadística y de evaluación de la enfermedad para el tizón de *Mycosphaerella* en los guisantes, el período de latencia se utilizó para discriminar entre diferentes tratamientos. Se registró el tiempo hasta la aparición de los primeros picnidios (duración del período de latencia). Las plántulas que no mostraron ninguna producción de estructura de picnidios durante el período de observación de 20 días se registraron como observaciones censuradas por la derecha. Utilizando análisis de supervivencia no paramétrico (estimaciones de Kaplan-Meier) y semiparamétrico (riesgos proporcionales de Cox) bajo diferentes hipótesis relacionadas con factores que podrían influir en el período de latencia.

Durante estas investigaciones, se hizo una regresión del período de latencia contra la duración de la humedad de la hoja, el cultivar de guisante, la concentración del inóculo, la edad de la planta y la agresividad del aislado. Tanto la regresión de Cox como la prueba de Kaplan-Meier mostraron la importancia de la duración de la humedad de la hoja, la concentración del inóculo y la agresividad del aislamiento en los tiempos de supervivencia, por lo que la mediana de la duración del período latente fue respectivamente 15 y 16 para tn0203 y md0202. La mediana para la humedad de 3 hojas fue 14, 16 y 17 respectivamente para 06, 48 y 72 h LWD. Tanto el cultivar como la edad de la planta no tuvieron riesgo significativo para la aparición de la estructura de los picnidios. Del mismo modo, utilizando la regresión semiparamétrica de riesgo proporcional de Cox, las 2 covariables (mayor humedad de la hoja y mayor dosis de inóculo con un aislamiento agresivo, se asociaron significativamente con el tiempo de



sobrevivencia. Por lo tanto, la razón de riesgo fue respectivamente 1,205 y 1,423 para LWD y la concentración de inóculo, respectivamente.

## **Materiales y métodos**

### **Material vegetal**

Pea cv Onward y cv Merveille de Kelvedon "MK", las variedades más cultivadas conocidas por diferir en su nivel de resistencia cuantitativa de altamente susceptible a resistencia moderada, respectivamente, se usaron en todos los experimentos. Las semillas de cada cultivar se sembraron en macetas de 20 cm de diámetro que contenían una mezcla de tierra/compost sin esterilizar. Se plantaron 10 semillas por maceta y las plántulas se diluyeron a 5. Las plantas se mantuvieron en un invernadero.

### **Producción de material fúngico e inóculo**

En el estudio se utilizaron dos aislados de *M. pinodes*, md0203 y tn0203, obtenidos de 2 localidades que presentaron bajo y alto puntaje de agresividad en el cv Onward. Las cepas se cultivaron en medio PDA durante 10 días a 21°C. Se recogieron conidios de cultivos de 10 días de antigüedad añadiendo 10 ml de agua destilada estéril para eliminar las esporas. La suspensión de esporas se filtró a través de 2 capas de cheesecloth para eliminar el micelio y los fragmentos de agar. La concentración de esporas se determinó utilizando un hemocitómetro. La suspensión de conidios se diluyó con agua destilada estéril para obtener la concentración final requerida para cada experimento.

### **Estudio del período de humedad de la hoja**

Plantas de dos semanas de edad de cv Onward y cv Merveille de Kelvedon se pulverizaron hasta escurrimiento con una suspensión de conidios de  $4 \times 10^6$  esporas/ml. Las plántulas de guisantes se sometieron luego a una duración de la humedad de la hoja (LWD) de 24, 48 y 72 h. Durante el período húmedo, las plantas se cubrieron con bolsas de polietileno transparente rociadas por dentro con agua destilada para permitir la infección. Las plantas sin cobertura o sin embolsar, se consideraron no expuestas a un período húmedo. Al final de este período, las plántulas se destaparon y se mantuvieron en un invernadero no controlado donde la temperatura osciló entre 15 y 25 °C.

## **Estudio de concentración de inóculo y edad de la planta**

El efecto de la concentración del inóculo se investigó en plantas de quince días (etapa de tres y cinco hojas) de cv Onward. Las plantas se inocularon por aspersión hasta el escurrimiento con una suspensión de esporas que contenía  $2,5 \times 10^3$ ;  $3,5 \times 10^5$  y  $5,2 \times 10^7$  esporas/ml. Las suspensiones se aplicaron con un atomizador de aspersión con boquilla ajustable para formar una alta densidad de gotas finas sobre las partes aéreas de las plantas. Posteriormente, inmediatamente después de la inoculación las plantas se cubrieron durante 48 h con bolsas de polietileno transparente y se rociaron en su interior con agua destilada para permitir la infección. Después del período de incubación, las plantas se destaparon y se mantuvieron en un invernadero no controlado donde la temperatura osciló entre 15 y 25 °C.

## **Análisis estadístico y modelado**

### **Estimación de Kaplan-Meier de la función de supervivencia**

El período de latencia, definido como el período desde la inoculación del huésped hasta la aparición de los primeros picnidios en las hojas, se denominó datos de supervivencia. Para estimar el período de latencia (LP), las plantas se observaron diariamente desde el momento de la inoculación hasta 20 días. La evaluación de la enfermedad de la infección por *M. pinodes* en las hojas también se registró utilizando una escala de gravedad de la enfermedad (DS) de 0 a 5 (Tivoli, 1999), donde 0, sin lesión; 1, algunas manchas dispersas; 2, numerosas manchas; 3, 10 - 15% área foliar necrótica y presencia de motas; 4, 50% del área foliar cubierta por lesiones; 5, 75 - 100% del área foliar deshidratada o necrótica.

Los picnidios en las hojas se registraron con la ayuda de una lupa (10 X). En los datos de supervivencia ocurrieron datos censurados porque muchos individuos no presentaron el evento al final de la observación, y el intervalo censurado ya que solo se sabe que el tiempo de supervivencia está entre dos tiempos de observación.

En este estudio, el conjunto de datos contenía múltiples observaciones censuradas (es decir, plantas que no presentaron síntomas al final del período de evaluación y aquellas con síntomas pero que no formaron una estructura de picnidios), por lo tanto, la variable dependiente se considera como "tiempo de supervivencia" (Esker *et al.*, 2006; Scherm y Ojiambo, 2004; Garrett *et al.*, 2004; Padovan y Gibb, 2001). De hecho, la función de supervivencia  $S(t)$  mide la probabilidad de que un individuo sobreviva

más allá del tiempo  $t$ :  $S(t) = P [T > t]$ . Sea  $T$  el tiempo de supervivencia. Consideraron a  $T$  como una variable aleatoria con función de distribución acumulativa  $P(t) = \Pr (T \leq t)$  y función de densidad de probabilidad  $p(t) = dP(t)/dt$ .

La función de supervivencia más optimista  $S(t)$  es el complemento de la función de distribución,  $S(t) = \Pr (T > t) = 1 - P(t)$ . Otra representación de la distribución de los tiempos de supervivencia es la función de riesgo, que evalúa el riesgo instantáneo de fallecimiento en el tiempo  $t$ , condicionado a la supervivencia a ese tiempo:

Además, se utilizó el método de Kaplan-Meier para estimar las tasas de supervivencia en una tasa entre 2 grupos y se probó la significación estadística mediante la prueba de rango logarítmico. Se consideró estadísticamente significativo un valor de  $p \leq 0,05$ . Las curvas de supervivencia son funciones escalonadas monótonas no crecientes iguales a 1 en el tiempo cero y 0 cuando el tiempo se acerca al infinito. Las diferencias estadísticas entre las curvas de supervivencia se calcularon mediante la prueba de rango logarítmico. Debido a que esta prueba no se puede calcular para datos censurados por intervalos, en este caso se utilizó el punto medio de cada intervalo.

### **Modelado del tiempo de supervivencia utilizando riesgos proporcionales de Cox**

Los modelos de regresión de Cox utilizan la función de riesgo para estimar el riesgo relativo de falla o de ocurrencia del evento. La función de riesgo,  $h(t)$  es una estimación de la muerte potencial por unidad de tiempo en un instante dado, ya que el caso ha sobrevivido hasta ese instante (Kleinbaum, 1996). El modelo de riesgos proporcionales de Cox examina la influencia de posibles covariables en el riesgo de muerte de un individuo (Collett, 2003; Ojiambo *et al.*, 2002; Kleinbaum, 1996; Dungan *et al.*, 2003). El riesgo en el tiempo  $t$ , es la probabilidad de que un individuo que ha sobrevivido hasta el tiempo  $t$  muera en el siguiente período de tiempo (Ojiambo *et al.*, 2002, Scherm y Ojiambo, 2004, Muenchow, 1986).

### 3.3. Aplicación práctica de las metodologías utilizadas para el desarrollo de tablas de vida

#### 3.3.1. Germinación de semilla

En modelos completamente paramétricos, se asume una distribución de probabilidad específica de la línea de base de riesgo / supervivencia, de acuerdo con una distribución de probabilidad definida. Los modelos paramétricos se pueden expresar en: (1) modelo de riesgo proporcional, donde un cambio de una unidad en una variable explicativa causa cambios proporcionales en el riesgo; y (2) modelo de tiempo de falla acelerado (AFT), donde un cambio de una unidad en una variable explicativa provoca un cambio proporcional en el tiempo de supervivencia. El modelo AFT se calcula en función del número de semillas que han germinado en cada intervalo de tiempo anterior y del número de semillas que podrían haber germinado al final del intervalo anterior. El modelo de tiempo de falla acelerado asume una forma paramétrica para los efectos de las variables explicativas y generalmente asume una forma paramétrica para la función de superviviente subyacente de que el efecto de las covariables en la distribución del tiempo de un evento es multiplicativo en el tiempo del evento. El enfoque AFT permite considerar una variedad de funciones de distribución estadística para los tiempos de germinación de semillas en diferentes lotes de semillas y compara las respuestas para los diferentes lotes de semillas mediante la estimación de un '*factor de aceleración*' que resume las tasas relativas de germinación de la semilla. Romano y Stevanato (2020) y McNair, Sunkara y Frobish (2012) revisaron modelos no paramétricos y semiparamétricos aplicados a datos de germinación, mientras que Onofri, Gresta y Tei (2010) consideraron una variedad de modelos AFT para describir las respuestas de germinación para diferentes especies de malezas. Otros autores aplicaron modelos no paramétricos, semiparamétricos y modelo AFT (Scott y Jones, 1982; Barak *et al.*, 2018; Winkler *et al.*, 2005; Hirsch *et al.*, 2012).

Sin embargo, pocas investigaciones han examinado en conjunto la aplicación de modelos no paramétricos, semiparamétricos y totalmente paramétricos a los mismos datos comparando el significado biológico de los parámetros estimados.

### **3.3.1.1. Análisis de tiempo y riesgo de germinación en semillas de *Hura crepitans* L imbibidas en agua (Vale, 2023)**

#### **Materiales y métodos**

##### **Sistema de recolección y procesamiento de frutos**

Se utilizaron semillas extraídas de frutos colectados de árboles plantados en los jardines de la Villa Universitaria del Núcleo Rafael Rangel, Universidad de Los Andes, municipio Pampanito, estado Trujillo, Venezuela, a 392,49 msnm. La recolección de los frutos de jabillo se hizo durante el mes de marzo de 2020, con una podadora de extensión. Para extraer las semillas y culminar el secado, las cápsulas se colocaron dentro de una caja con fondo de malla metálica para permitir la aireación y puestas al sol hasta el secado y dehiscencia. Luego de la dehiscencia, las semillas se mantuvieron a temperatura ambiente en el interior de bolsas de papel, hasta la siembra realizada en el mes de abril de 2020.

##### **Diseño experimental**

Antes de la siembra, las semillas se sometieron a diferentes tratamientos de tiempos de inmersión en agua, los cuales fueron:

T1: Sin inmersión de semillas en agua (testigo, 0 tiempo);

T2: Inmersión de semillas en agua durante 6 horas;

T3: Inmersión de semillas en agua durante 12 horas;

T4: Inmersión de semillas en agua durante 18 horas;

T5: Inmersión de semillas en agua durante 24 horas;

El experimento fue realizado mediante un diseño de bloques al azar, cinco tratamientos y cinco repeticiones de diez envases de polietileno cada una.

##### **Siembra de semillas y evaluación de la germinación**

Las semillas se colocaron en un envase plástico limpio con la cantidad de agua suficiente para cubrir las 100 semillas de cada tratamiento. Posterior al remojo por tiempo de inmersión de cada tratamiento, las semillas se extendieron en un lugar sombreado y se dejaron secar durante 30 minutos. La siembra de las 2 semillas se hizo por siembra directa en cada envase (6 cm de diámetro x 14 cm de altura), a una profundidad de 0,5 cm. Los envases se mantuvieron a campo abierto (intemperie) durante la eva-

luación del experimento. Mientras las semillas estuvieron en germinación, los riegos con agua se continuaron 2 veces por día. Para efectos prácticos, se consideró que una semilla había germinado al emerger el talluelo de la plántula, llevando un registro diario de semillas germinadas por tratamiento durante 30 días continuos después de la siembra (DDS). Cada semilla individual seleccionada para la germinación por tratamiento se consideró una unidad experimental. Las semillas que no germinaron al final de este período se trataron como censuradas a la derecha.

### **Hipótesis y nivel de significación**

Debido a que las hipótesis de interés se refieren a las diferencias entre tratamientos de imbibición en diferentes tiempos de inmersión en agua, las semillas se consideraron como unidades experimentales. Afortunadamente, los métodos utilizados para analizar los datos de tiempo de falla pueden manejar correctamente los datos censurados y no requieren datos distribuidos normalmente (Fox, 2001; Allison, 2010).

El procedimiento LIFETEST permite comparar pares de curvas de supervivencia y detectar si existen diferencias estadísticamente significativas entre ellas y probar la hipótesis nula de que las funciones de supervivencia son iguales entre curvas. La hipótesis nula ( $H_0$ ) es una hipótesis que la investigación trata de refutar, rechazar o anular, mientras que la hipótesis alternativa ( $H_1$ ) es lo que la investigación plantea que es la causa de un fenómeno, indicando que las funciones de supervivencia no son iguales. El nivel de significancia del intervalo de confianza de la función de supervivencia por defecto es 0,05, que resulta en un intervalo de confianza de 95%.

### **Análisis de datos**

Se utilizaron tres métodos estadísticos diferentes: análisis de las tablas de vida (LIFETEST o métodos de Kaplan-Meier), modelo de riesgo proporcional de Cox y el modelo de tiempo de falla acelerada (ATF). Se adoptó un esquema de “observación continua” durante la duración del ensayo, asumiendo que los tiempos de germinación tuvieran un “dato exacto” para cada semilla (McNair *et al.*, 2012).

Las variables tratamientos de inmersión en agua y días (tiempo para germinar), son variables cuantitativas, cada una con cinco y 100 valores, respectivamente. Para el análisis, en cada medición se asignaron valores de 1 o 0 para las semillas germinadas y no germinadas, respectivamente. Luego de cada

conteo diario, cada semilla germinada fue removida del experimento. Se utilizaron códigos para discriminar entre los tratamientos ( $T_1=0$ ,  $T_2=6$ ,  $T_3=12$ ,  $T_4=18$  y  $T_5=24$ ). Los datos se analizaron con el paquete SAS (SAS, 2020; Fox, 2001), de la manera siguiente:

### **Análisis no paramétrico**

Los datos de germinación fueron descritos y modelados bajo dos probabilidades relacionadas: supervivencia y función de riesgo, ambas dependientes del tiempo. Este análisis se hizo con el procedimiento PROC LIFETEST, que imprime por defecto, los estadísticos Log-rank y Wilcoxon, así como valores de chi-cuadrado para las comparaciones entre tratamientos. Las diferencias de supervivencia entre covariables, se analizaron a partir de curvas de supervivencia construidas por el método Kaplan – Meier (Kaplan y Meier, 1958; Klein *et al.*, 2001; Gómez *et al.*, 2009); para ello, se tomó en cuenta el estatus de cada semilla (status 1 = germinación, status 0 = censura debido a pérdida durante el experimento o semilla no germinada al final del experimento), así como el tiempo de germinación de la misma en días. Según Onofri *et al.* (2010), los tiempos de germinación se usan para calcular la probabilidad de “germinación/emergencia”, es decir, la probabilidad que una semilla individual germine después de un tiempo específico  $t$ , considerando el evento inicial (imbibición) en tiempo  $t = 0$ .

Este análisis compara la respuesta de las curvas de germinación por el método de máxima verosimilitud o probabilidad, utilizando la prueba de Wilcoxon en caso de presentarse cruzamientos entre curvas de germinación y, en caso contrario, la prueba de Log-Rank (Bárcenas *et al.*, 2013; SAS, 2020; Fox, 2001). La prueba Log-Rank otorga igual peso al principio y final de la germinación y asume que las funciones de riesgo para los tratamientos son paralelas, sin cruzamiento entre curvas, indicando que las razones de riesgo son constantes o proporcionales en todos los períodos de tiempo. Solo en esta situación es una prueba eficaz (Fleming *et al.*, 1987; Lee, 1996; Buyske *et al.*, 2000; Lee y Wenyu, 2003). Cuando la razón de riesgo no es constante, las pruebas de Gehan – Wilcoxon y la prueba Tarone – Ware pueden ser más eficaces (Tarone y Ware, 1977; Pepe y Fleming, 1989). La prueba Peto-Peto también es eficiente cuando se viola la suposición de riesgo proporcional (Kleinbaum y Klein, 2005).

### **Análisis semiparamétrico**

La razón de riesgo debe ser constante con respecto al tiempo, razón por la cual el modelo de Cox es llamado modelo de riesgo proporcional. El modelo examina la influencia de potenciales covariables sobre el riesgo de sobrevivencia de un individuo (semilla). El riesgo en el tiempo  $t$ , es la probabilidad que un individuo, quien ha sobrevivido al tiempo  $t$ , germine en el siguiente menor período de tiempo. El modelo asume que el riesgo para dos grupos es proporcional (Dixon y Newman, 1991; Collett, 2003).

El modelo de regresión de Cox tiene en cuenta la información proporcionada por todas las semillas de los diferentes tratamientos, incluidas aquellas que no germinaron durante el seguimiento (datos censurados). El período de seguimiento se divide en intervalos temporales, calculándose la tasa de riesgo como la probabilidad, entre no haber presentado el suceso hasta ese momento, y presentarlo en el periodo temporal subsiguiente. Esta tasa de riesgo (HR) varía a lo largo de la duración del estudio, por lo que el modelo calcula una hipotética tasa instantánea media cuando el periodo de tiempo tiende a cero. La HR representa el cociente de las tasas instantáneas entre el tratamiento control o de referencia (Molina 2015).

Para ajustar el modelo a los datos, se utilizó el procedimiento PHREG, con los términos de interacción entre tratamientos y el tiempo, además de la covariable por separado (Cox, 1972). En caso de resultar el término de interacción estadísticamente significativo, el modelo no es apropiado (Lee y Wen-yu, 2003). Además, se utilizaron otros métodos para verificar y avalar las suposiciones de riesgo proporcional. El primero, un análisis gráfico de Kaplan – Meier, utilizando las curvas de  $\log(-\log(S(t)))$  versus  $\log(t)$ . El segundo, mediante el análisis de residuales de Schoenfeld. También se corrió una prueba para analizar la multicolinealidad entre las variables (diagnóstico VIF). Un valor de  $VIF > 5$  se consideró evidencia de multicolinealidad.

En ambos análisis (semiparamétrico y paramétrico), se puede elegir arbitrariamente uno de los niveles como nivel o tratamiento de referencia, con coeficiente de regresión con valor de cero (0) y todos los demás tratamientos se comparan con este nivel y las pruebas de significancia de chi cuadrada para cada tratamiento prueban si difiere o no del tratamiento de referencia y puede o no, ser significativa-



mente diferente de otros niveles, requiriendo múltiples comparaciones (Allison, 2010). En la investigación, el tratamiento inmersión de semillas en agua durante 24 horas, se consideró de referencia, al considerar la opción por defecto que presenta el SAS de utilizar el último tratamiento como nivel de referencia (en este caso el 5 y último tratamiento) y los coeficientes de otros tratamientos, se compararon con él, y las pruebas de significancia para cada tratamiento se probaron para evaluar sus diferencias. Se utilizó la prueba de razón de probabilidad para las comparaciones, con una prueba de chi cuadrada ( $\chi^2$ ) para determinar si los modelos con covariables adicionales significativamente mejoran el modelo nulo. Sin embargo, puede elegir cualquier tratamiento como de referencia, utilizando el siguiente ejemplo al seleccionar el tratamiento de 0 horas de inmersión:

```
Proc Phreg Data=Jabillo;
Class Tratam (REF='0');
Model Días*Status (0) = Tratam Tratdias / Ties=Efron R1;
Test Tratam6=Tratam12; Test Tratam6=Tratam18; Test Tratam6=Tratam24;
Test Tratam12=Tratam18; Test Tratam12=Tratam24; Test Tratam18=Tratam24;
Tratdias = Tratam*(Días < 15);
Run;
```

El valor nulo para la razón de riesgo (HR) es 1, que indica igualdad de probabilidad de producción del suceso en los dos grupos a comparar en el siguiente intervalo de tiempo. Una  $HR > 1$  indica más riesgo de producción en el tratamiento de intervención que en el de referencia y una  $HR < 1$ , menor riesgo en el tratamiento de intervención que en el tratamiento de referencia.

### **Análisis paramétrico**

En este caso, para la distribución de supervivencia se asumió el modelo de tiempo de falla acelerado (AFT, por sus siglas en inglés), que postula una relación directa entre predictores y tiempo de supervivencia (George *et al.*, 2014). Para determinar la distribución más apropiada y mejorar el potencial del modelo AFT, se ajustaron y compararon con PROC LIFEREG, cuatro distribuciones paramétricas (Weibull, Lognormal, Log-Logística y Exponencial). Para establecer el mejor ajuste entre las distribu-

ciones, se utilizó el Criterio de Máxima Verosimilitud (ML).

### Programas SAS para análisis de datos

Data Jabillo;

Input Sem Tratam Días Status;

Label Sem='Semillas por tratamiento'

    Tratam='Tiempos de Inmersión en horas'

    Días='Tiempo transcurrido de Germinación'

    Status='Status de Germinación';

SemDias=Sem\*Días;

TratamDias=Tratam\*Días;

Datalines;

1    0    14    1

2    0    15    1

.....

.....

.....

99   0    30    0

100  0    30    0

.....

.....

1    6    11    1

2    6    13    1

```

.....
.....
99    6    30    0
100   6    30    0
1     12   12    1
.....
4     12   13    1
5     12   13    1
.....
.....
1     24   13    1
2     24   14    1
.....
.....
99    24   30    0
100   24   30    0
;

```

### **Métodos no paramétricos y tablas de vida con el procedimiento LIFETEST**

Con el siguiente programa, se utiliza el procedimiento LIFETEST para estimar la tabla de vida de las funciones de sobrevivencia por intervalos de tiempo para los tratamientos de inmersión de semilla en agua, utilizando METHOD=LT. Las gráficas de las funciones de sobrevivencia se solicitan con la opción PLOT (PLOT=S). La variable tiempo de falla es referida como DÍAS (días de germinación). La variable STATUS se refiere a la censura por los eventos de la germinación (pérdida o semilla no ger-

minada y semilla germinada) y la variable STRATA para estratificar la germinación en función de los tratamientos (TRATAM). La función STRATA se utiliza para comparar las funciones de sobrevivencia que calcula de manera predeterminada el programa (Log-Rank, Wilcoxon y -2Log (LR). Además de estas pruebas, con la opción TEST=ALL se calculan otras pruebas (Tarone, Peto, Peto modificada y Fleming). El SAS realiza las comparaciones utilizando métodos que asumen datos exactos. Con la opción PLOT= (S, H) se solicita las gráficas de sobrevivencia y razón de riesgo.

```
Proc Print Data=Jabillo;
```

```
Run;
```

```
Proc Lifetest Data=Jabillo Method=LT Plots= (S, H);
```

```
Time Días*Status (0);
```

```
Strata Tratam;
```

```
Run;
```

El programa siguiente utiliza los procedimientos LIFETEST para estimar la función de sobrevivencia por Kaplan – Meier, también conocido como estimador producto – limite para describir el proceso de germinación para cada día, arreglados en orden ascendente para la totalidad de tratamientos de germinación, utilizando METHOD=KM. El programa también estratifica para las cinco combinaciones entre tratamientos, en que la opción ADJUST origina los valores de probabilidad para las múltiples comparaciones entre tratamientos, utilizando la prueba de Tukey. Describe los datos en resultados de 6 columnas

```
Proc Lifetest Data=Jabillo Method=Km Plots= (S, LS);
```

```
Time Días*Status (0);
```

```
Strata Tratam / ADJUST = Tukey;
```

```
Run;
```

Según Onofri et al. (2010), los tiempos de germinación se usan para calcular la probabilidad de “germinación/emergencia”, es decir, la probabilidad que una semilla individual germine después de un tiempo específico  $t$ , considerando el evento inicial (imbibición) en tiempo  $t = 0$ .

### **Prueba de supuestos de proporcionalidad mediante PROC LIFETEST.**

El siguiente programa comprueba los supuestos de PH para la variable tratamientos en el modelo Cox, utilizando PROC LIFETEST, y luego gráfica  $-\log(-\log(S(\text{DÍAS})))$  versus  $\log(\text{DÍAS})$  para las cuatro funciones de sobrevivencia de los tratamientos. Con la opción PLOT= (LLS) se solicita la gráfica de sobrevivencia Log – Log (Log-Log survivor functions):

```
Proc Lifetest Data=Jabillo Method=Km Plots= (LLS);
```

```
Time Días*Status (0);
```

```
Strata Tratam;
```

```
Run;
```

### **Métodos semiparamétricos**

Para ajustar un modelo de PH (riesgo proporcional) a los datos, se utilizó el procedimiento PHREG en SAS, con los términos de interacción entre las covariables y el tiempo.

### **Comprobación de supuestos PH**

¿Cómo saber si los datos satisfacen la suposición de PH y qué sucede si se viola el supuesto? En este caso, se debe comprender que las violaciones de la suposición de PH son equivalentes a interacciones entre una o más covariables y tiempo. Si el efecto de una variable varía con el tiempo, la suposición de PH se viola para esa variable. Por tal motivo, se requiere el uso de interacciones entre variables y tiempo. Para esos casos, se requiere examinar métodos para probar y modelar riesgos no proporcionales. Entre ellos: el uso de residuales, incorporar las interacciones en el modelo y la estratificación, que subsume las interacciones en la función arbitraria del tiempo (Allison, 2010).

### **Prueba del supuesto de proporcionalidad con la declaración ASSESS**

Se han definido una variedad de diferentes tipos de residuos para modelos de regresión de Cox (Collett, 2003). Entre ellos, los residuos de martingala para probar la no proporcionalidad, utilizando la declaración ASSESS en PROC PHREG. Desafortunadamente, esta declaración no funcionará si el modelo contiene covariables dependientes del tiempo (Allison, 2010):

Ods Graphics On;

Proc Phreg Data = Jabillo;

Model Días\*Status (0) =Tratam / Ties = Efron;

Assess Ph / Resample;

Run;

Ods Graphics Off;

También pueden realizarse pruebas de riesgos no proporcionales con los residuales de Schoenfeld. A diferencia de la sentencia ASSESS, este método funciona incluso si se tienen covariables cambiantes en el tiempo. Los residuos de Schoenfeld tienen una propiedad inusual: en lugar de una única residual para cada individuo, hay un residual separado para cada covariable y para cada individuo. No se definen para observaciones censuradas (faltan en el conjunto de datos de salida).

### **Prueba se residuales de Schoenfeld**

El uso principal de estos residuos es detectar posibles desviaciones del supuesto de riesgos proporcionales. Si se satisface la suposición, los residuos de Schoenfeld deben ser independientes del tiempo. Probar la suposición de riesgos proporcionales es equivalente, por lo tanto, a probar si los residuos de Schoenfeld están correlacionados con el tiempo o con alguna función del tiempo (Allison, 2010).

Proc Phreg Data = Jabillo;

Model Días\*Status (0) =Tratam / Ties = Efron;

Output Out=B Rersch=Schtratam;

Run;

La palabra clave RESSCH va seguida de los nombres de las variables elegidas, correspondientes a las covariables del modelo. Estas variables contienen los residuos de Schoenfeld en el conjunto de datos de salida.

**Correlaciones entre los residuales de Schoenfeld para tratamientos y variables de tiempo de germinación (días, días1, días2), con la jerarquización de los tiempos de germinación por PROC RANK:**

El primero paso, es correr PROC PHREG para calcular los residuales de Schoenfeld. El segundo paso, es correr PROC RANK para resortejar el tiempo de sobrevivencia de la variable y obtener los rangos por tiempo de sobrevivencia. El tercer paso, es correr un PROC CORR para probar la correlacion entre los tiempos de sobrevivencia rankeados y los residuales de Schoenfeld.

El procedimiento RANK origina los rangos para una o más variables numéricas de un juego de datos. Esta salida muestra los resultados de jerarquía de los valores de una variable, es decir, se indica ubicación del orden de finalización del resultado de cada variable, por ejemplo, en un ensayo de germinación el tratamiento con el más bajo número de días de germinación finaliza en primer lugar. La variable a ser clasificada son los días de germinación (Días), utilizando la opción RANKS para clasificarla, eligiendo por nombre de la variable jerarquizada DIASRANK. El procedimiento es el siguiente:

```
Proc Print Data=B;
```

```
Run;
```

```
Data Event;
```

```
Set B;
```

```
Run;
```

```
Proc Rank Data=Event Out=Resrank Ties=Mean;
```

```
Var Días;
```

```
Ranks Diasrank;
```

Run;

Proc Print Data =Resrank;

Run;

Luego se utiliza PROC CORR para obtener las correlaciones entre la variable jerarquizada (DIAS-RANK) y las variables que contienen los residuales de Schoenfeld, para este caso los tratamientos probados (SCHTRATAM). Si se alcanza la suposición de PH (suposición de riesgos proporcionales) para esta covariable, entonces las correlaciones deben ser cercanas a cero. El valor de probabilidad (p) obtenido del PROC CORR prueba si esta correlacion es cero es el mismo valor de p que se utilizó para probar la suposición de PH. Es decir, si el valor de p obtenido es menor que 0,05 ( $p < 0,05$ ), sugiere que la suposición de PH se viola para la variable jeraquizada. El siguiente paso es modificar el conjunto de datos para incorporar dos funciones del tiempo, el LOG y el cuadrado, seguido de PROC CORR para calcular las correlaciones de los residuos con el tiempo y sus dos funciones.

Data C;

Set Resrank;

Días1=Log (Días);

Días2=Días\*\*2;

Proc Corr;

Var Diasrank Días1 Días2 Schtratam;

Run;

Se puede obtener el mismo valor de p al correr una regresión lineal con cada variable predictora (una a la vez), utilizando PROC REG y examinando los valores de p para los coeficientes de regresión.

Por ejemplo:

Proc Reg Data=Resrank;

Model Diasrank=Schtratam;



Run;

En el siguiente método, veremos cómo estimar un modelo de Cox que permita que el efecto de la covariable de interés cambie con el tiempo. Esto también dará una prueba definitiva del supuesto de proporcionalidad para esta variable. Usando tanto la declaración ASSESS como los residuales de Schoenfeld, acabamos de ver cómo ejecutar un programa para determinar la evidencia o no de que la variable de interés viola o no la suposición de riesgos proporcionales, lo que implicaría que el efecto de la variable puede variar o no con el tiempo desde su lanzamiento.

### **Prueba de supuestos con interacciones con el tiempo como covariables dependientes del tiempo**

Una forma común de representar la interacción entre dos variables en un modelo de regresión lineal es incluir una nueva variable que es el producto de las dos variables en cuestión. Para representar la interacción entre un covariable  $x$  y tiempo en un modelo de Cox, se puede definir:

$$\text{Log } h(t) = \alpha(t) + \beta_1 x + \beta_2 xt.$$

Factorizando la  $x$ , se puede reescribir esto como:

$$\text{Log } h(t) = \alpha(t) + (\beta_1 + \beta_2 t) x$$

En esta ecuación, el efecto de  $x$  es  $\beta_1 + \beta_2 t$ . Si  $\beta_2$  es positivo, entonces el efecto de  $x$  aumenta linealmente con el tiempo; si es negativo, el efecto disminuye linealmente con el tiempo.  $\beta_1$  se puede interpretar como el efecto de  $x$  en el tiempo 0, el origen del proceso (Allison, 2010). Se puede estimar este modelo definiendo una covariable tiempo dependiente  $z = xt$ .

El modelo se puede adaptar para diferentes etapas del proceso germinativo. A continuación, se detallan tres modelos: a) para los primeros 15 días del proceso (0 – 15 días); b) para los siguientes 15 días (16 – 30 días); un tercer modelo que incluye todo el proceso de germinación (para el proceso total de 30 días):

#### **a. /\*Primer Modelo Cox de estimación de Interacciones con Variables de Tiempo Primeros 15 Días de la germinación\*/**

Proc Phreg Data=Jabillo;

Class Tratam;

Model Días\*Status (0) = Tratam Tratdias / Ties=Efron Rl;

Test Tratam0=Tratam6; Test Tratam0=Tratam12; Test Tratam0=Tratam18;

Test Tratam6=Tratam12; Test Tratam6=Tratam18; Test Tratam12=Tratam18;

Tratdias = Tratam\*(Días < 15);

Run;

**b. /\*Segundo Modelo Cox de estimación de Interacciones con Variables de Tiempo 15-30 Días de la germinación\*/**

Proc Phreg Data=Jabillo;

Class Tratam;

Model Días\*Status (0) = Tratam Tratdias / Ties=Efron Rl;

Test Tratam0=Tratam6; Test Tratam0=Tratam12; Test Tratam0=Tratam18;

Test Tratam6=Tratam12; Test Tratam6=Tratam18; Test Tratam12=Tratam18;

Tratdias = Tratam\*(Días > 15);

Run;

**c. /\*Tercer Modelo Cox de estimación de Interacciones con Variables de Tiempo Total de la germinación (30 Días)\*/**

Proc Phreg Data=Jabillo;

Class Tratam;

Model Días\*Status (0) = Tratam Tratdias / Ties=Efron Rl;

Test Tratam0=Tratam6; Test Tratam0=Tratam12; Test Tratam0=Tratam18;

Test Tratam6=Tratam12; Test Tratam6=Tratam18; Test Tratam12=Tratam18;

Tratdias=Tratam\*Días;

Run;

En resumen, ahora tenemos una forma definitiva de probar las violaciones de supuesto de PH: para cualquier covariable, simplemente se agrega al modelo una covariable dependiente del tiempo que representa la interacción de la covariable original y tiempo. Si la covariable de interacción no tiene un coeficiente significativo ( $p > 0,05$ ), entonces se puede concluir que la suposición de PH no se viola para esa variable. Por el contrario, si la variable de interacción tiene un coeficiente significativo ( $p < 0,05$ ), existe evidencia de no proporcionalidad.

### **Método gráfico para evaluar el ajuste del modelo paramétrico**

Existe otra vía para discriminar entre diferentes distribuciones de probabilidad, utilizando el diagnóstico gráfico. Anteriormente con PROC LIFETEST se obtuvieron gráficas de los estimados de la función de supervivencia, los cuales se pueden utilizar para evaluar dos de los modelos de distribución. Específicamente, la opción PLOTS=LS, produce una gráfica de  $-\log \hat{S}(t)$  versus  $t$ . Si la verdadera distribución es Exponencial, la gráfica debe producir una línea recta, con un origen en cero (0). La opción PLOTS=LLS produce una gráfica de  $\log [-\log \hat{S}(t)]$  versus  $\log t$ , la cual debe ser una línea recta si la verdadera distribución es Weibull.

Existen dos limitaciones de este método. Primero, PROC LIFETEST no produce gráficos apropiados para evaluar las distribuciones Gamma, Lognormal o Loglogística. Segundo, estos gráficos no se ajustan para efectos de covariables. Ambas limitaciones se satisfacen con la declaración PROBLOT en PROC LIFEREG.

Después de la declaración MODEL, se coloca PROBLOT, el cual produce estimados no paramétricos de la función de supervivencia usando un método modificado Kaplan – Meier que ajusta las covariables. Si luego se aplica una transformación a los estimados de supervivencia, que, al graficar contra el logaritmo del tiempo, debe aparecer como una línea recta si el modelo especificado es correcto (Allison, 2010).

## Métodos paramétricos con el procedimiento LIFEREG

Los modelos de tiempo de falla acelerado asumen que los tratamientos y las covariables afectan el tiempo de falla multiplicativamente (es decir, las etapas del proceso de germinación se "aceleran"). Una interpretación alternativa es que, bajo modelos acelerados de tiempo de falla, las covariables hacen correr más rápido o más lento el reloj, por lo que cualquier período de alto riesgo cambiará en el tiempo cuando las covariables cambian (Kalbfleisch y Prentice, 1980). En modelos de riesgos proporcionales, los períodos de alto riesgo siguen siendo los mismo, pero la posibilidad de que un individuo caiga en uno de esos períodos variará con las covariables individuales.

¿Cómo se elige entre modelos de tiempo de falla acelerado y de riesgo proporcional? Probablemente el enfoque más útil es considerar las hipótesis ecológicas y preguntarse si los tratamientos se espera que realmente cambien el momento de los períodos de alto riesgo (lo que sugiere el uso de modelos de tiempo de falla acelerado) o si los tratamientos simplemente cambian la posibilidad de falla (lo que sugiere el uso de modelos de riesgos proporcionales). Una comprobación adicional de la idoneidad del modelo de riesgos proporcionales es graficar  $\log(-\log[S(t)])$  contra el tiempo. Si el modelo de riesgos proporcionales es apropiado, las curvas de los diferentes grupos deben ser aproximadamente paralelas entre sí, para un determinado nivel de una covariable (Kalbfleisch y Prentice, 1980). Una limitación de este enfoque es que deben realizarse comparaciones entre grupos de tratamiento dentro de los niveles de una covariable. Si todas las variables predictoras son covariables, no se pueden hacer comparaciones, independientemente de si el modelo de riesgos proporcionales es apropiado. En consecuencia, es probablemente mejor confiar en el examen de las hipótesis ecológicas.

Para Fox (2001), existen dos pasos para utilizar estos modelos: 1) elegir una distribución de supervivencia y, 2) estimar los parámetros para la distribución de supervivencia y los coeficientes de regresión  $\beta$ . Según este autor, existen tres formas diferentes de elegir una distribución de supervivencia. Primero, se puede elegir una distribución basada en fundamentos ecológicos o biológicos a priori. En segundo lugar, se puede tomar un enfoque empírico: después de que se hayan recopilado los datos, comparar las estimaciones de la tabla de vida de sus funciones de riesgo o supervivencia con las funciones de riesgo o supervivencia de varias distribuciones (Exponencial, Weibull, Gamma, Lognormal, Log-logist).

Para los datos de emergencia del jabillo, es razonable elegir entre la distribución Log-logística o Lognormal: dado que la especie tiene poca o ninguna latencia de semillas, esperando que la tasa de germinación y la posterior emergencia aumenten rápidamente después de la aplicación de agua, alcanzando una tasa máxima y luego declinar. Se comparan con la distribución Gamma.

La emergencia de las plántulas se evaluó con un modelo de regresión de máxima verosimilitud, utilizando el procedimiento LIFEREG que proporciona modelos de regresión para datos tomados a intervalos periódicos de tiempo y el análisis de variables cuantitativas, por lo que es posible probar la asociación de la emergencia con estas variables. Para mejorar la estimación de los coeficientes de regresión se determinó el tipo de distribución de la función de la supervivencia bajo cada tratamiento y semilla, por comparación con las distribuciones teóricas Exponencial, Weibull, Log-Normal y Log-Logística, utilizando la opción DIST=.

Para graficar los datos de estas distribuciones, con LIFETEST se pueden graficar o evaluar dos de las distribuciones. Con PLOT=LS se produce una gráfica de  $(-\log S(t))$  vs  $t$ . Si la distribución es Exponencial, esta gráfica debe producir una línea fuerte con origen en 0. Con PLOT=LL, se produce una gráfica de  $\log(-\log S(t))$  vs  $\log t$ , que debe originar una línea fuerte si la verdadera distribución es Weibull. Para las distribuciones Gamma, Lognormal o Loglogistic existen limitaciones ya que LIFETEST no produce gráficos apropiados y los gráficos no se ajustan para el efecto de las covariables. Ambas limitaciones se resuelven con PROBLOT en LIFEREG (Allison, 2010).

El primer paso es observar las gráficas definidas con LIFETEST (LS y LLS) y luego terminar de hacer la selección con el menor valor establecido por AIC en LIFEREG.

```
Proc Sort Data=Jabillo;
```

```
By Descending Tratam;
```

```
Run;
```

```
Proc Lifereg Data=Jabillo;
```

```
Class Tratam;
```

```
Model Días* Status (0) = Tratam Tratdías /Dist= Lognormal;
```

Probplot;

Run;

## Resultados

En ocasiones el remojo en agua a temperatura ambiente incrementa la velocidad de germinación en semillas que no experimentan latencia o con ligeros niveles de esta, también se utiliza conjuntamente con un tratamiento más fuerte o seguido de este (CATIE, 2000). Por su parte Schmidt (2000), ha planteado que donde la latencia física es relativamente débil, como en las semillas frescas de leguminosas, el remojo en agua a temperatura ambiente es muchas veces suficiente para permitir la permeabilidad de la corteza seminal, mientras que el efecto de este mismo tratamiento en semillas duras varía con la especie; esto indica que en algunas especies las semillas se convierten en gradualmente permeables y en otras existe un pobre efecto del remojo continuo. De lo anterior se deriva la importancia de comparar los tratamientos de inmersión en agua.

Además, los datos de germinación presentan un especial problema estadístico: no se conoce si las semillas que no germinaron son capaces de germinar, están latentes, germinaron y luego murieron sin emerger el talluelo de la plántula o siempre fueron inviables. En el presente caso, las muestras presentaron mediana a alta germinación, sin saber en cual condición de las anteriores se encontraban, es decir, sin tener conocimiento de la condición germinativa que tenían las semillas sin germinar. En el análisis de sobrevivencia son datos conocidos como censurados a la derecha y para estimar válidamente la media o la varianza del tiempo de germinación se requiere de los métodos utilizados a continuación.

### a. Métodos no paramétricos y tablas de vida con el procedimiento LIFETEST:

Según Scott y Jones (1982), la dispersión de las respuestas de germinación sobre el intervalo de tiempo  $i$  es descrito por: 1) Función de sobrevivencia ( $S_i$ ); 2) la función de densidad de probabilidad ( $f_i$ ) y, 3) la función de riesgo ( $h_i$ ). La función de sobrevivencia estima la proporción acumulada de semillas no germinadas, así la curva de germinación acumulada viene dada por  $1 - S_i$ . La función de densidad de probabilidad estima la probabilidad que una semilla germinará en un intervalo de tiempo dado. Se calcula como:

$$F_i = P_{i-1} * q_i / h_i$$

Donde  $p$  es la probabilidad de sobrevivir al final del intervalo de tiempo  $i - 1$ ,  $q$  es la probabilidad de germinación en el intervalo  $i$  y  $h_i$  es la amplitud del intervalo. La densidad de probabilidad puede interpretarse como la distribución de las tasas de germinación en el tiempo. La función de riesgo representa la probabilidad que una semilla germinará en un intervalo de tiempo particular, dado que no ha germinado. Esto representa la tasa de germinación condicional, puesto que el valor es condicional sobre el número de semillas que potencialmente pueden germinar en cada intervalo de tiempo (Scott, Jones y Willans, 1984).

### Comparación entre tratamientos de inmersión

Uno de los principales propósitos de la germinación para describir el proceso biológico, consiste en estimar algunos indicadores del lote de semillas, que permiten comparar tratamientos o cuantificar los efectos de factores experimentales sobre el proceso de germinación/emergencia. Entre ellos, la media y la mediana del tiempo de germinación (tiempo medio de germinación ( $T_{50}$ ): número de días transcurridos después de la incubación, para alcanzar el 50% del porcentaje de germinación final). En algunos casos también se requieren otros tiempos para caracterizar mejor la primera y última fase del proceso, los tiempos para el 25 % o 75 % de germinación ( $T_{25} - T_{75}$ ), (Onofri *et al.*, 2010).

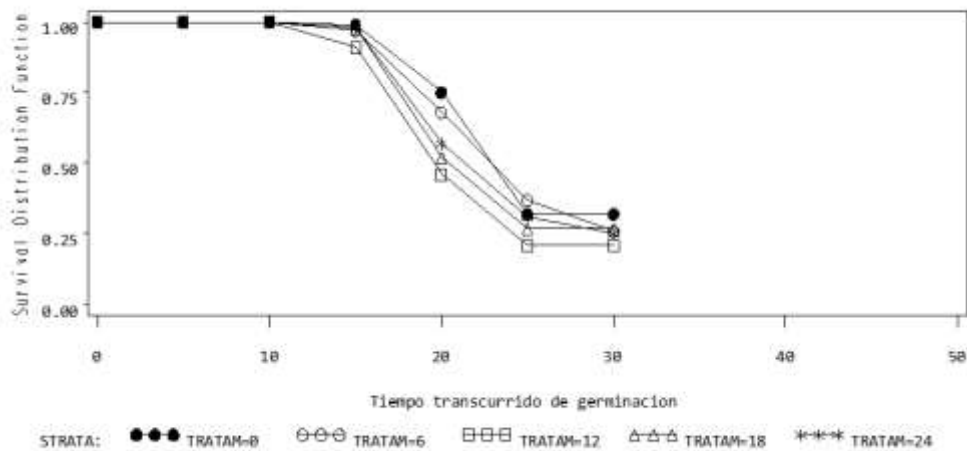
En las condiciones del ensayo, todos los tratamientos alcanzaron valores mayores al 50 % de germinación, donde el tratamiento de inmersión por 12 horas obtiene el menor tiempo medio de germinación (19 días). Los demás tratamientos presentaron tiempos medios de germinación muy cercanos, entre 20 – 22 días. Los valores de la media del tiempo de germinación resultaron variables, entre 18,7 – 21,6 días, donde inmersión durante 12 horas logró mayor velocidad de germinación, como expresión del número de días promedio empleados en la germinación (18,7 días) versus los demás tratamientos. Igualmente, presenta la mayor tasa de germinación (0,053 semillas germinadas/día), interpretada como valor recíproco del tiempo medio de germinación ( $1/T_{50}$ ). Según Ranal y García (2006), para los autores del índice, esta expresión debería interpretarse en el sentido de que altos valores deberían estar asociados con germinación concentrada en el tiempo. Además, para la primera fase de germinación ( $T_{25}$ ), el tratamiento de inmersión durante 12 horas, logró el menor tiempo para alcanzar el 25 % de germina-

ción, con un tiempo estimado de 16 días, seguido por las inmersiones durante 18 y 24 horas (17 días). Estos resultados sugieren que el tiempo de inmersión óptimo es 12 horas. Según Vale (2022), al evaluar la germinación en *H. crepitans* mediante índices de germinación (Germinación acumulada, VG, TMG e IVG) y por modelos de crecimiento no lineal, el tratamiento con más alta germinación resultó inmersión en agua durante 12 horas (79 %), los tratamientos con inmersión en agua durante 6 – 18 y 24 horas, también presentaron porcentajes de germinación arriba de 70 % (74,0 – 73,0 y 75,0 %, respectivamente), mientras que la germinación más baja la presentó el tratamiento sin inmersión (68,0 %).

### **a.1. METHOD = LT**

Las proporciones de supervivencia estiman la probabilidad,  $S(t)$  de que una semilla de la población estudiada permanezca sin germinar, desde el inicio de la siembra, durante un intervalo de tiempo no inferior a  $t$ . Las siguientes salidas de resultados de este método, describen las tasas de supervivencia a nivel de intervalos (0 – 5, 5 – 10, 10 – 15 días, etc). El primer intervalo comienza cuando la línea base del tiempo es ( $t = 0$ ) hasta el día 5, describiendo los primeros eventos de germinación. La tasa de supervivencia (semillas sin germinar) para este primer intervalo y los tres siguientes es 1, lo que significa que aún no ha germinado ninguna semilla en los tres primeros intervalos. Por ejemplo, para TRATAM=0, la probabilidad de no germinar después del intervalo entre 10 – 15 días cae de 1 a 0,99 en el intervalo 15 – 20 días, y a 0,75 en el intervalo 20 – 25 días y así sucesivamente. Indica que las semillas de este tratamiento tienen un 75% de posibilidades de no germinar durante el intervalo. Así se describe la germinación para cada tratamiento según los intervalos en la estratificación por tratamiento que se realizó con el programa. El programa también da origen a las curvas de supervivencia ( $S$ ) solicitadas con la opción plot (Figura 15). Además, también se origina un resumen del número de valores censurados y no censurados de cada tratamiento.





**Figura 15. Cambios en la supervivencia de la germinación de *H. crepitans* en los tratamientos de inmersión de semilla durante 30 días de evaluación.** Fuente. Elaboración propia.

Para cada intervalo, se calculan 14 parámetros estadísticos (Allison, 2010). Los primeros cuatro se requieren para calcular los demás estadísticos. La probabilidad condicional de falla o de germinación (Conditional Probability of Failure) es un estimado de la probabilidad de que una semilla germinará en el intervalo. Se calcula como  $\text{number failed} / \text{effective sample size}$ . Un valor estimado del error standard (Conditional Probability Standard Error) aparece en la siguiente columna. De esta manera, la probabilidad condicional de germinación para los tratamientos, en el intervalo 15 – 20 días, fue de  $0,2424 \pm 0,0431$  –  $0,2990 \pm 0,0465$  –  $0,4945 \pm 0,0524$  –  $0,4694 \pm 0,0504$  y  $0,4184 \pm 0,0498$ , para 0 – 6 – 12 – 18 y 24 horas de inmersión. Lo que sugiere que la mayor probabilidad condicional de germinación recayó en el tratamiento de 12 horas de inmersión. La columna Failure es 1 menos Survival. También aparecen los errores standard de las probabilidades de supervivencia (Survival standard errors). La columna Survival es el estimado de la tabla de vida de la función de sobreviviente, que significa la probabilidad de que ocurra el evento de germinación en un tiempo mayor o igual que el tiempo al inicio de cada intervalo. Por ejemplo, la probabilidad de que una semilla no germine entre 15 – 20 días, fue de  $0,99 \pm 0,00995$  –  $0,9700 \pm 0,0171$  –  $0,9100 \pm 0,0286$  –  $0,9800 \pm 0,0140$  y  $0,9800 \pm 0,0140$ , para 0 – 6 – 12 – 18 y 24 horas.

La columna del residual de la mediana (Median Residual Lifetime), es un estimado del tiempo restante hasta un evento para una semilla que sobrevivió al comienzo de cada intervalo. Para calcular éste estadístico para un intervalo dado, debe ser un intervalo posterior cuya probabilidad de sobrevivencia es menor que la mitad de la probabilidad de sobrevivencia asociada con el intervalo de interés (Allison, 2010). Esto significa que la mitad de las semillas por germinar en el tiempo  $t_i$  se espera que estén por germinar en el tiempo  $t_{mi}$ . Esto significa que para los tratamientos 0 – 6 – 12 – 18 y 24 horas, en el intervalo 15 – 20 días, con valores de  $7,9651 \pm 0,5785$  -  $8,1452 \pm 0,7943$  -  $5,1000 \pm 0,9539$  -  $5,6000 \pm 0,9899$  -  $6,5385 \pm 0,9519$  días, respectivamente, el tratamiento de 12 horas se espera que la mitad de las semillas por germinar en este intervalo lo hagan en 5,1 días en el siguiente. La columna PDF Standard Error indica los valores estimados de la función de densidad de probabilidad en el punto medio del intervalo. De gran interés, es la columna Hazard. El propósito general de la definición es cuantificar el riesgo instantáneo que un evento ocurra en un tiempo  $t$ . Para este caso, significa cuantificar los estimados de la función de riesgo en el punto medio de cada intervalo (número de germinaciones por unidad de tiempo). Para el intervalo 15 – 20 días fue de  $0,05517 \pm 0,01115$  -  $0,0703 \pm 0,01285$  -  $0,1314 \pm 0,0185$  -  $0,1227 \pm 0,01722$  y  $0,10581 \pm 0,01594$ , para 0 – 6 – 12 – 18 y 24 horas. Se observa que el mayor riesgo instantáneo en este intervalo es para 12 horas de inmersión. A continuación se describen los valores estimados para los 14 estadísticos, para los diferentes tratamientos.

### The LIFETEST Procedure

**Stratum 1: TRATAM = 0**

#### Life Table Survival Estimates

Interval Lower, Upper	Number Failed	Number Censored	Effective Sample Size	Conditional Probability of Failure	Conditional Probability Standard		Survival	Failure
					Error			
0	5	0	100.0	0	0	1.0000	0	
5	10	0	100.0	0	0	1.0000	0	
10	15	1	100.0	0.0100	0.00995	1.0000	0	
15	20	24	99.0	0.2424	0.0431	0.9900	0.0100	
20	25	43	75.0	0.5733	0.0571	0.7500	0.2500	
25	30	0	32.0	0	0	0.3200	0.6800	
30	35	0	32	16.0	0	0.3200	0.6800	

## Evaluated at the Midpoint of the Interval

Interval		Survival	Median	Median	PDF		Hazard	
Lower,	Upper	Standard	Residual	Standard	PDF	Standard	Hazard	Standard
		Error	Lifetime	Error		Error		Error
0	5	0	22.9070	0.5814	0	.	0	.
5	10	0	17.9070	0.5814	0	.	0	.
10	15	0	12.9070	0.5814	0.00200	0.00199	0.00201	0.00201
15	20	0.00995	7.9651	0.5785	0.0480	0.00854	0.05517	0.011154
20	25	0.0433	4.3605	0.5035	0.0860	0.00990	0.16074	0.022447
25	30	0.0466	.	.	0	.	0	.
30	35	0.0466	.	.	0	.	0	.

## Stratum 2: TRATAM = 6

## Life Table Survival Estimates

Interval		Number	Number	Effective	Conditional	Conditional			
Lower,	Upper					Failed	Censored	Sample	Probability
				Size	of Failure	Error			
0	5	0	0	100.0	0	0	1.0000	0	0
5	10	0	0	100.0	0	0	1.0000	0	0
10	15	3	0	100.0	0.0300	0.0171	1.0000	0	0
15	20	29	0	97.0	0.2990	0.0465	0.9700	0.0300	0.0300
20	25	31	0	68.0	0.4559	0.0604	0.6800	0.3200	0.3200
25	30	11	0	37.0	0.2973	0.0751	0.3700	0.6300	0.6300
30	35	0	26	13.0	0	0	0.2600	0.7400	0.7400

## Evaluated at the Midpoint of the Interval

Interval		Survival	Median	Median	PDF		Hazard	
Lower,	Upper	Standard	Residual	Standard	PDF	Standard	Hazard	Standard
		Error	Lifetime	Error		Error		Error
0	5	0	22.9032	0.8065	0	.	0	.
5	10	0	17.9032	0.8065	0	.	0	.
10	15	0	12.9032	0.8065	0.00600	0.00341	0.006091	0.003516
15	20	0.0171	8.1452	0.7943	0.0580	0.00908	0.070303	0.012852
20	25	0.0466	6.3636	1.8741	0.0620	0.00925	0.118095	0.020265
25	30	0.0483	.	.	0.0220	0.00626	0.069841	0.020734
30	35	0.0439	.	.	0	.	0	.

**Stratum 3: TRATAM = 12**

## Life Table Survival Estimates

Interval Lower, Upper	Number Failed	Number Censored	Effective Sample Size	Conditional Probability of Failure	Conditional Probability		Survival	Failure
					Standard Error			
0 5	0	0	100.0	0	0		1.0000	0
5 10	0	0	100.0	0	0		1.0000	0
10 15	9	0	100.0	0.0900	0.0286		1.0000	0
15 20	45	0	91.0	0.4945	0.0524		0.9100	0.0900
20 25	25	0	46.0	0.5435	0.0734		0.4600	0.5400
25 30	0	0	21.0	0	0		0.2100	0.7900
30 35	0	21	10.5	0	0		0.2100	0.7900

## Evaluated at the Midpoint of the Interval

Interval Lower, Upper	Survival Standard Error	Median Residual Lifetime	Median Standard Error	PDF		Hazard Standard Error	
				PDF	Standard Error	Hazard	Standard Error
0 5	0	19.5556	0.5556	0	.	0	.
5 10	0	14.5556	0.5556	0	.	0	.
10 15	0	9.5556	0.5556	0.0180	0.00572	0.018848	0.006276
15 20	0.0286	5.1000	0.9539	0.0900	0.00995	0.131387	0.018499
20 25	0.0498	4.6000	0.6782	0.0500	0.00866	0.149254	0.027695
25 30	0.0407	.	.	0	.	0	.
30 35	0.0407	.	.	0	.	0	.

**Stratum 4: TRATAM = 18**

## Life Table Survival Estimates

Interval Lower, Upper	Number Failed	Number Censored	Effective Sample Size	Conditional Probability of Failure	Conditional Probability		Survival	Failure
					Standard Error			
0 5	0	0	100.0	0	0		1.0000	0
5 10	0	0	100.0	0	0		1.0000	0
10 15	2	0	100.0	0.0200	0.0140		1.0000	0
15 20	46	0	98.0	0.4694	0.0504		0.9800	0.0200
20 25	25	0	52.0	0.4808	0.0693		0.5200	0.4800
25 30	0	0	27.0	0	0		0.2700	0.7300
30 35	0	27	13.5	0	0		0.2700	0.7300

## Evaluated at the Midpoint of the Interval

Interval		Survival	Median	Median	PDF		Hazard	
Lower, Upper		Standard	Residual	Standard	PDF	Standard	Hazard	Standard
		Error	Lifetime	Error		Error		Error
0	5	0	20.4000	1.0000	0	.	0	.
5	10	0	15.4000	1.0000	0	.	0	.
10	15	0	10.4000	1.0000	0.00400	0.00280	0.00404	0.002857
15	20	0.0140	5.6000	0.9899	0.0920	0.00997	0.12267	0.017215
20	25	0.0500	.	.	0.0500	0.00866	0.12658	0.024015
25	30	0.0444	.	.	0	.	0	.
30	35	0.0444	.	.	0	.	0	.

## Stratum 5: TRATAM = 24

## Life Table Survival Estimates

Interval		Number	Number	Effective	Conditional	Conditional			
Lower, Upper						Failed	Censored	Sample	Probability
				Size	of Failure	Error			
0	5	0	0	100.0	0	0	1.0000	0	
5	10	0	0	100.0	0	0	1.0000	0	
10	15	2	0	100.0	0.0200	0.0140	1.0000	0	
15	20	41	0	98.0	0.4184	0.0498	0.9800	0.0200	
20	25	26	0	57.0	0.4561	0.0660	0.5700	0.4300	
25	30	6	0	31.0	0.1935	0.0710	0.3100	0.6900	
30	35	0	25	12.5	0	0	0.2500	0.7500	

## Evaluated at the Midpoint of the Interval

Interval		Survival	Median	Median	PDF		Hazard	
Lower, Upper		Standard	Residual	Standard	PDF	Standard	Hazard	Standard
		Error	Lifetime	Error		Error		Error
0	5	0	21.3462	0.9615	0	.	0	.
5	10	0	16.3462	0.9615	0	.	0	.
10	15	0	11.3462	0.9615	0.00400	0.00280	0.00404	0.00286
15	20	0.0140	6.5385	0.9519	0.0820	0.00984	0.10581	0.01594
20	25	0.0495	7.0833	3.1458	0.0520	0.00877	0.11818	0.02214
25	30	0.0462	.	.	0.0120	0.00475	0.04286	0.01739
30	35	0.0433	.	.	0	.	0	.

En la salida de resultados siguiente (*Summary of the Number of Censored and Uncensored Values*) se observa que, en el caso de tratamientos (TRATAM), de las 500 evaluaciones realizadas en el tratamiento sin inmersión (0 horas), 32 de las semillas fueron censuradas (no germinaron), lo que representa un 32% de censuradas. Para los demás tratamientos las semillas censuradas fueron 26 – 21 – 27 y 25, para los tratamientos de 6 – 12 – 18 y 24 horas. De las 500 semillas bajo prueba de germinación germinaron 369 semillas, lo que representa un porcentaje de germinación de 73,8%, con el tratamiento de 12 horas con mayor porcentaje de germinación (79%).

Summary of the Number of Censored and Uncensored Values

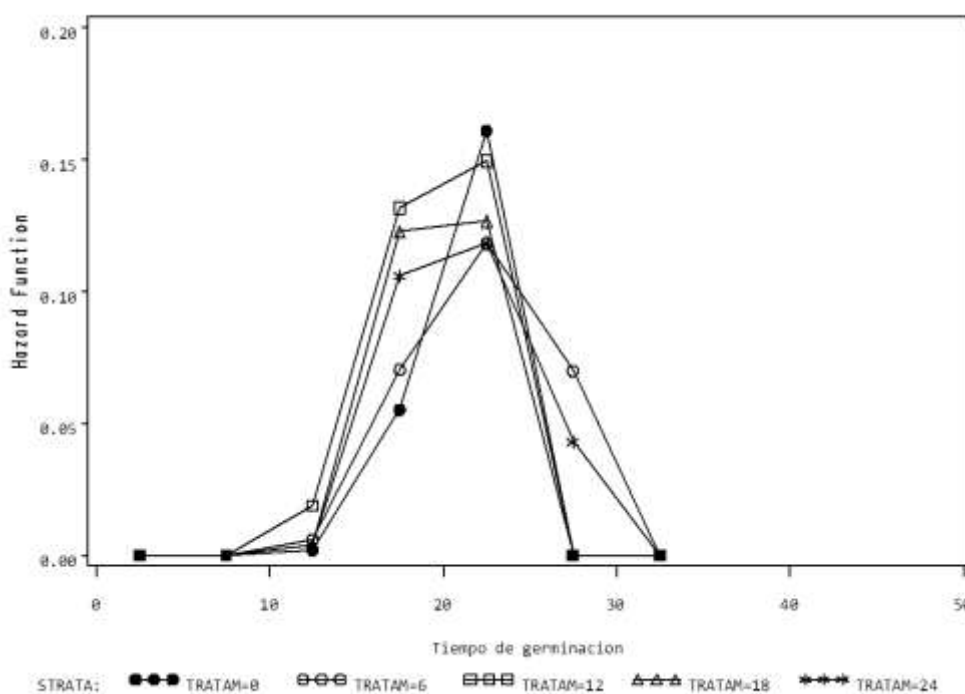
Stratum	TRATAM	Total	Failed	Censored	Percent Censored
1	0	100	68	32	32.00
2	6	100	74	26	26.00
3	12	100	79	21	21.00
4	18	100	73	27	27.00
5	24	100	75	25	25.00
-----					
Total		500	369	131	26.20

PROC LIFETEST también origina los estadísticos log-rank y Wilcoxon y un estimado de su matriz de covarianza para cada tratamiento, los cuales se utilizan para calcular el valor de Chi cuadrada. También se originan los valores de p con valores significativos ( $p = 0,0146$ ) en la prueba log-Rank y  $p = 0,0003$  en la prueba Wilcoxon. Entonces existe evidencia para rechazar la hipótesis nula que los cinco tratamientos tienen igual función de supervivencia, es decir, la misma distribución de probabilidad para los eventos de germinación en el tiempo. También se originan valores de p para la prueba razón de probabilidad ( $-2\log(LR)$ ), usualmente inferior a las anteriores ya que requiere la suposición que la función de riesgo es constante en cada tratamiento, implicando una distribución exponencial para los eventos en el tiempo.

The LIFETEST Procedure  
Test of Equality over Strata

Test	Chi-Square	Pr >	DF	Chi-Square
Log-Rank	12.3942		4	0.0146
Wilcoxon	20.8561		4	0.0003
-2Log (LR)	3.0095		4	0.5562

Estas pruebas generalizan para 3 o más tratamientos o grupos, con la hipótesis nula que todos los grupos tienen la misma función de supervivencia. Si la hipótesis nula es verdadera, todas las pruebas estadísticas tienen una distribución chi cuadrada con grados de libertad igual al número de grupos menos 1 y con valores de  $p < 0,05$ . El procedimiento STRATA provee gran flexibilidad en la definición de estos grupos. PROC LIFETEST construye un estrato correspondiente a cada valor de la variable definida en STRATA, la cual puede ser numérica o de carácter. Si STRATA contiene 2 o más variables, PROC LIFETEST construye un grupo para cada combinación de valores de las variables.



**Figura 16. Funciones de riesgo para datos de tratamientos de inmersión de semillas de jabillo (*Hura crepitans* L.) en agua.** Fuente. Elaboración propia.

Una tasa expresa el cociente entre la variación relativa de una magnitud  $Y$  respecto a la variación de otra magnitud  $X$ , que en este caso es el tiempo. Si en un instante  $t$  hay  $n$  semillas expuestas al acontecimiento terminal (germinación), la tasa de incidencia se define como la variación relativa de semillas (proporción de germinación) por unidad de tiempo en dicho instante; cuando se calcula en un cierto intervalo de tiempo se denomina tasa media de incidencia y se interpreta como una medida del riesgo de germinar en dicho intervalo. La función de riesgo  $h(t)$  (hazard function) representa la probabilidad condicionada por unidad de tiempo que tiene una semilla de germinar en un instante  $t$  si había sobrevivido hasta el instante anterior. Por este motivo se llama también tasa condicionada de germinación (conditional germination rate) o fuerza de germinación. Nótese que en la función de densidad  $f(t)$  la proporción de germinación es respecto al total de la muestra, mientras que en la función de riesgo es relativa



solo a las semillas que llegan sin germinar al instante  $t$ . Bajo las condiciones de los tratamientos, se tuvieron las más altas probabilidades de riesgo de germinación en el intervalo 20 – 25 días (Figura 16). Sin embargo, durante los intervalos comprendidos en los primeros 15 días y los últimos intervalos de seguimiento ( $> 25$  días), los riesgos de germinación fueron los más bajos. La Figura 16, indica que, para los diferentes tratamientos, la función de riesgo de germinar es ascendente, aproximadamente hasta 22 días de germinación, a partir del cual la razón de riesgo disminuye. Por ejemplo, el tratamiento de semillas sin inmersión, es el que tiene la razón de riesgo más alta (0,16074 germinaciones/día) en el intervalo 20 – 25 días, a partir del cual cae abruptamente hasta el final de la germinación. Por el contrario, en el mismo intervalo, los valores pico más bajos de razón de riesgo se presentan en los tratamientos de 6 y 24 horas de inmersión, con 0,118095 y 0,11818 germinaciones/día, respectivamente.

## a.2. METHOD=KM.

Por razones de espacio no se incluyen los resultados de los estimados de sobrevivencia Producto – límite (Product-Limit Survival Estimates) relacionados con los estratos de cada tratamiento (Stratum 1: TRATAM = 0; Stratum 2: TRATAM = 6; Stratum 3: TRATAM = 12; Stratum 4: TRATAM = 18; Stratum 5: TRATAM = 24). De esta información se utilizará solo los valores de los cuantiles, entre ellos la mediana, además de la media de cada estrato.

Uno de los principales propósitos de la germinación y emergencia de semillas para describir el proceso biológico, consiste en estimar algunos indicadores del lote de semillas, que permiten comparar lotes o cuantificar los efectos de factores experimentales sobre el proceso de germinación/emergencia. Entre ellos, la media y la mediana del tiempo de germinación (tiempo medio de germinación ( $T_{50}$ ): número de días transcurridos desde el inicio del experimento, para alcanzar el 50% del porcentaje de germinación final). En algunos casos también se requieren otros tiempos para caracterizar mejor la primera y última fase del proceso, los tiempos para el 25 % o 75 % de germinación ( $T_{25} - T_{75}$ ), (Onofri et al., 2010).

## The LIFETEST Procedure

**Stratum 1: TRATAM = 0**

## Product-Limit Survival Estimates

## Summary Statistics for Time Variable DÍAS

## Quartile Estimates

Percent	Point Estimate	95% Confidence Interval	
		[Lower	Upper)
75	.	24.0000	.
50	22.0000	21.0000	24.0000
25	19.5000	18.0000	20.0000

## Mean Standard Error

21.2900	0.3050
---------	--------

**Stratum 2: TRATAM = 6**

## Product-Limit Survival Estimates

## Summary Statistics for Time Variable DÍAS

## Quartile Estimates

Percent	Point Estimate	95% Confidence Interval	
		[Lower	Upper)
75	.	25.0000	.
50	21.0000	20.0000	23.0000
25	18.0000	17.0000	20.0000

## Mean Standard Error

21.5600	0.4003
---------	--------

**Stratum 3: TRATAM = 12**

## Product-Limit Survival Estimates

## Summary Statistics for Time Variable DÍAS

Quartile Estimates

Percent	Point Estimate	95% Confidence Interval [Lower Upper)	
75	22.0000	21.0000	.
50	19.0000	17.0000	20.0000
25	16.0000	15.0000	17.0000
	Mean	Standard Error	
	18.6500	0.3090	

**Stratum 4: TRATAM = 18**

Product-Limit Survival Estimates

Summary Statistics for Time Variable DÍAS  
Quartile Estimates

Percent	Point Estimate	95% Confidence Interval [Lower Upper)	
75	.	23.0000	.
50	20.0000	19.0000	21.0000
25	17.0000	16.0000	18.0000
	Mean	Standard Error	
	19.7300	0.2987	

**Stratum 5: TRATAM = 24**

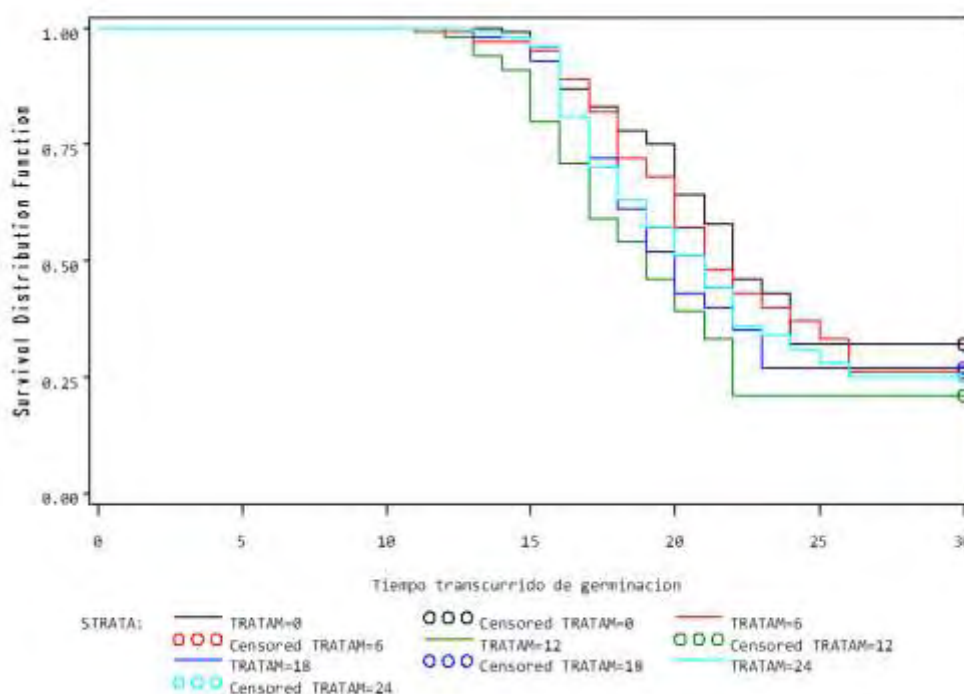
Product-Limit Survival Estimates

Summary Statistics for Time Variable DÍAS

Quartile Estimates			
Percent	Point Estimate	95% Confidence Interval [Lower Upper)	
75	.	23.0000	.
50	21.0000	19.0000	22.0000
25	17.0000	16.0000	18.0000
	Mean	Standard Error	
	20.8800	0.4050	

De los resultados anteriores, se observa que con el tratamiento de 12 horas de inmersión se obtiene el menor valor de la mediana o tiempo medio de germinación, correspondiente a 19 días y una media de días de germinación de 18,65 días. A este tratamiento le sigue el de 18 horas de inmersión, con mediana de 20 días y media de 19,73 días. El tratamiento con mayor valor del tiempo medio de germinación corresponde al testigo sin inmersión, con 22 días.

Las curvas escalonadas de Kaplan-Meier (KM) se muestran en la Figura 17. La naturaleza no continua de las curvas KM destaca el hecho de que no son funciones uniformes, sino estimaciones escalonadas. El valor de  $S(t)$  no es constante entre los tiempos de los eventos de germinación, y la probabilidad estimada cambia de valor solo en el momento de cada evento. Las distancias verticales entre horizontales representan el cambio en la probabilidad acumulada de no germinar a medida que avanza la curva. La pendiente de la curva está determinada por la duración de la supervivencia, representada por la longitud de las líneas horizontales.



**Figura 17. Estimados de Kaplan – Meier de las funciones de supervivencia para los diferentes tratamientos de inmersión de semillas de jabillo (*Hura crepitans* L.) en agua.**

*Fuente.* Elaboración propia.

Mientras que la probabilidad acumulada, vista en el eje Y de la curva, representa la probabilidad al principio y durante todo el intervalo, la tasa de supervivencia del intervalo representa la probabilidad de no germinar después del intervalo y al comienzo del siguiente (Romano y Stevanato, 2020).

Las funciones de sobrevivencia estimadas en la Figura 17, sugieren que el patrón temporal de germinación es diferente para la totalidad de tratamientos. Presentan variación en los eventos de germinación, con 12 eventos para los tratamientos de 0 y 12 horas, 15 eventos para los tratamientos 6 y 24 horas y 11 eventos para el tratamiento de 18 horas. Los cinco tratamientos no presentaron germinación durante los primeros intervalos (0 – 15 días). La tasa de supervivencia para este intervalo es 1, lo que significa que aún no ha germinado ninguna semilla. La primera semilla que germinó en el TRATAM= 0 horas fue el día 14 y la proporción de semillas sin germinar después del día 14 cae a 0,99 en el cuarto intervalo y a 0,75 en el quinto y así sucesivamente. A partir de estos primeros intervalos, los tratamientos presentaron probabilidades variables. Por ejemplo, para el tratamiento sin inmersión (TRATAM=0 horas), la incidencia de la probabilidad de germinación acumulada, expresada como  $1 - St$ , en el intervalo 15 – 20 días, es  $1 - 0,99$  (1%); por el contrario, para los tratamientos de 6-12-18 y 24 horas, la incidencia acumulada al mismo tiempo es de 3 – 9 – 2 – 2 %, respectivamente. Para el intervalo 20 – 25 días, considerando los mismos días, para todos los tratamientos, la probabilidad de germinación acumulada es  $1 - 0,75$  (25 %, TRATAM = 0 horas) y de 32 – 54 – 48 y 43 % para el resto de tratamientos (6 – 12 – 18 – 24 horas, respectivamente). Las probabilidades de germinación en los tratamientos, aumentaron en el intervalo final (30 – 35 días), a 0,68 - 0,74 - 0,79 – 0,73 y 0,75, respectivamente, con una razón de riesgo de 0,00 fallas/día, hasta el final del periodo de evaluación. Razonando en términos de la tasa de supervivencia del intervalo, el primer intervalo comienza cuando la línea base del tiempo es ( $t = 0$ ) y termina justo antes del primer evento de germinación. La tasa de supervivencia para este intervalo es 1, lo que significa que aún no ha germinado ninguna semilla. Por ejemplo, para los primeros 10 días del ensayo, ninguna semilla había germinado entre los tratamientos, lo que significa que la tasa de sobrevivencia de todos los tratamientos durante este intervalo es 1. La primera semilla germinó el día 11 en TRATAM= 6 y la probabilidad de no germinar después del día 11 cae a 0,99 en el segundo intervalo, a 0,97 en el tercer intervalo (día 13) y así sucesivamente.

Las tasas vienen dadas por la relación entre el número de semillas no germinadas (“en riesgo”) en un intervalo y el anterior. Las probabilidades acumuladas para un intervalo se calculan multiplicando las tasas de supervivencia hasta ese intervalo. Entonces, las probabilidades acumuladas antes mencionadas son el resultado de estas multiplicaciones. También es posible razonar en términos de tasas de riesgo no paramétricas y riesgo acumulativo. La primera representa la tasa de germinación instantánea en cualquier tiempo, dada por las relaciones entre semillas germinadas y semillas latentes, la función de riesgo acumulativo es la integral de las tasas de germinación desde el tiempo 0 y los siguientes 30 días de duración de la germinación, lo que representa la acumulación del riesgo a lo largo del tiempo. Es interesante observar que, para los datos de germinación, las diferencias en las curvas de supervivencia no parecen evidentes (Figura 17).

La intersección entre una línea horizontal asociada con la probabilidad de supervivencia igual a 0,5 y la curva de supervivencia en la Figura 17 nos permite encontrar la mediana del tiempo de germinación. Por ejemplo, fue de aproximadamente 19 días para TRATAM=12, 21 días para TRATAM=6 y de 22 días para TRATAM=0. Estos valores de mediana también se describen en los estimados de los cuartiles con intervalos de confianza (95 %) de las salidas de resultados. La mediana es una mejor medida de tendencia central que la media, debido al sesgo (skewness) de los tiempos de supervivencia. Además, no sabemos si las semillas que no germinaron, hubieran germinado en tiempo posterior al de evaluación (30 días). (Romano y Stevanato, 2020).

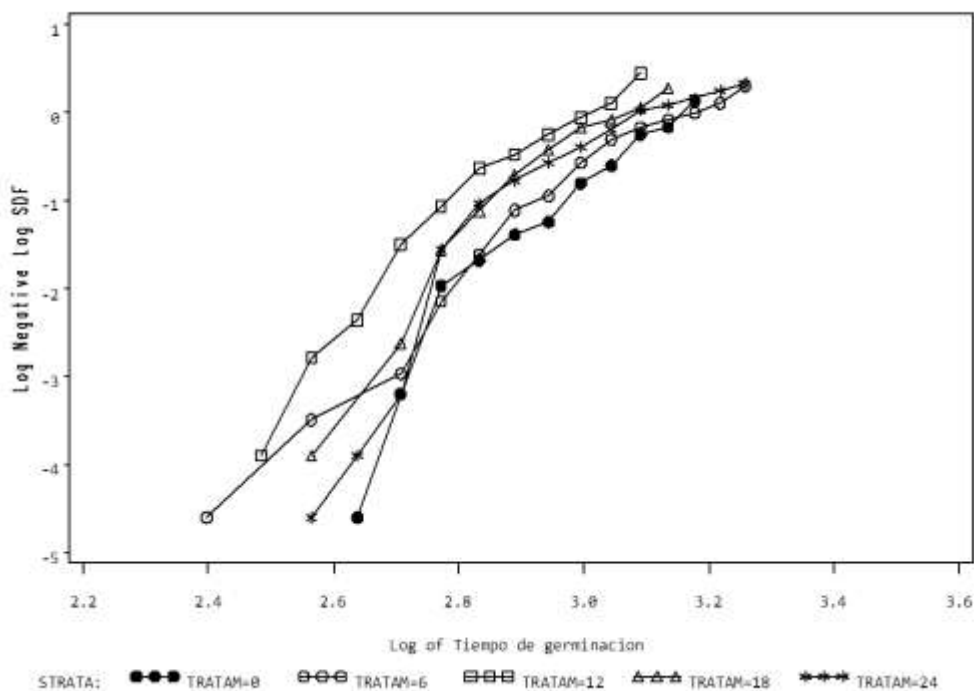
### **a.3. Comprobación de supuestos PH mediante la función de sobrevivencia Log – log, utilizando el procedimiento LIFETEST.**

La opción PLOTS= (S, LLS) produce curvas de sobrevivencia y log-log. Si la suposición PH se alcanza o logra, las curvas de los tratamientos o grupos formados con STRATA serán paralelas entre sí (sin cruzarse). Un método útil para evaluar gráficamente el supuesto de PH es trazar  $-\log(-\log(S(t)X))$  versus  $\log(t)$  para diferentes valores de las covariables X (restringido a valores de t tales que  $0 < S(t) < 1$ ), utilizando la tabla de vida o la estimación de Kaplan-Meier de la función de superviviente S (t). Las curvas aproximadamente paralelas para diferentes valores de las covariables indican que la suposi-

ción de PH es aceptable, mientras que las curvas que se cruzan de forma decisiva indican lo contrario. Dado que este método se basa en la comparación de curvas para grupos discretos, requiere categorizar las covariables cuantitativas.

Si las curvas de sobrevivencia log-log dan señales de líneas rectas con log (tiempo de sobrevivencia) sobre el eje horizontal, entonces existe evidencia que la “variable tiempo – evento” sigue una distribución Weibull. Por el contrario, si la pendiente de la línea es igual a 1, entonces hay evidencia que la variable sigue una distribución exponencial, un caso especial de distribución Weibull. Para estas situaciones se puede utilizar un modelo de sobrevivencia paramétrico (Kleinbaum y Klein, 2012).

Las variables TRATAM (tratamientos) y DÍAS (tiempo para germinar), son variables cuantitativas, cada una con cinco y 30 valores, respectivamente. Las gráficas de  $-\log(-\log(S(t)/X))$  versus  $\log(t)$  se muestran en la Figura 18 del texto. La figura muestra las curvas de los cinco tratamientos. El espacio entre las partes de nivel de las funciones de sobrevivencia no permanece similar a lo largo del tiempo, existiendo evidencia de un cruce decisivo que invierte claramente su orden, sobre todo entre los tratamientos de 6 y 24 horas, que cruzan las curvas de los tratamientos de 18 horas y el testigo. Esto indica que existe evidencia que no difieren entre ellos, caso contrario del tratamiento de imbibición de 12 horas, claramente diferente de ellos. Se concluye que la suposición de PH no es aceptable para todas las covariables.



**Figura 18.** Gráficas de diagnóstico para comprobar los supuestos del modelo de riesgo proporcional de Cox mediante la función de supervivencia –  $\text{Log}(-\log(S(t)))$  versus  $\text{log}(t)$  para tratamientos de inmersión de semillas de jabillo (*Hura crepitans* L.) en agua. Fuente. Elaboración propia.

Cuando se producen claras violaciones de la suposición de PH, están disponibles algunas soluciones en software estadístico que puede resolver el problema y permiten el uso del modelo Cox. Éstos incluyen convertir covariables con efectos no proporcionales en factores de estratificación o covariables dependientes del tiempo, y dividiendo el eje del tiempo en segmentos discretos y se analizan los datos para uno o más de los segmentos por separado (Therneau y Grambsch, 2000).

Según Allison (2010), la opción de graficación PLOTS= LLS, produce una gráfica de  $\text{log}(-\log S(t))$  vs  $\text{log } t$ , originando una línea recta si la verdadera distribución es Weibull. Por el contrario, PLOTS= LS produce una gráfica de  $-\log S(t)$  versus  $t$ . Si la verdadera distribución es exponencial, esta gráfica debe originar una línea con origen en cero (Figura 19), y puede ayudar en la selección de la distribución para analizar los datos con PROC LIFEREG.



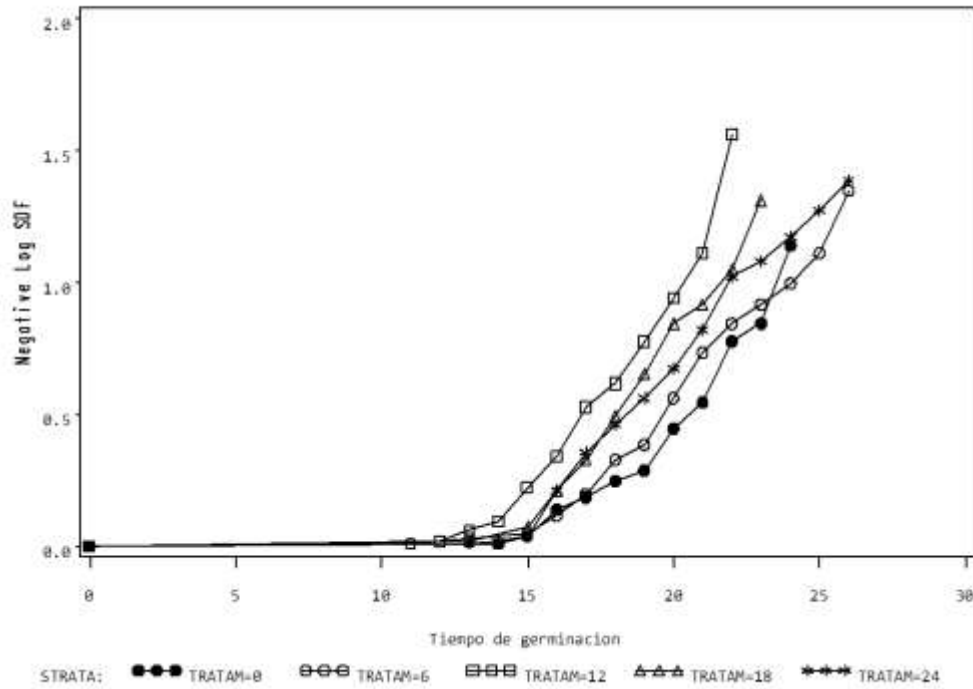


Figura 19. Funciones de sobrevivencia  $-\log S(t)$  versus  $t$  para tratamientos de inmersión de semillas de jabillo (*Hura crepitans* L.) en agua. Fuente. Elaboración propia.

**b. Métodos semiparametricos**

**b.1. Prueba del supuesto de proporcionalidad con la declaración ASSESS.**

The PHREG Procedure  
 Model Information

Data Set	WORK.JABILLO
Dependent Variable	DÍAS Tiempo de germinación
Censoring Variable	STATUS Status de germinación
Censoring Value(s)	0
Ties Handling	EFRON

Number of Observations Read	500
Number of Observations Used	500

## Summary of the Number of Event and Censored Values

Total	Event	Censored	Percent Censored
500	369	131	26.20

## Model Fit Statistics

Criterion	Without Covariates	With Covariates
-2 LOG L	4200.645	4197.239
AIC	4200.645	4199.239
SBC	4200.645	4203.149

## Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	3.4063	1	0.0649
Score	3.4154	1	0.0646
Wald	3.4081	1	0.0649

## The PHREG Procedure

## Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
TRATAM	1	0.01096	0.00594	3.4081	0.0649	1.011

El resultado anterior, muestra el parámetro estimado para el coeficiente de regresión de la variable tratamientos (TRATAM), el error estándar, una prueba de Chi cuadrado de Wald para la variable y su correspondiente valor de p. La columna etiquetada como Hazard Ratio indica la razón de riesgo estimada por una unidad de cambio en la variable predictora, al exponenciar el coeficiente de regresión.

## Supremum Test for Proportionals Hazards Assumption

Variable	Maximum Absolute Value	Replications	Seed	Pr > MaxAbsVal
TRATAM	1.2237	1000	214453001	0.9780

Con la declaración ASSESS, los valores de  $p$  se producen con la opción RESAMPLE, la cual genera una simulación de 1000 senderos. También produce una tabla resumen con los resultados para cada covariable. Al observar los resultados para las covariables, de la prueba para suposiciones de riesgo proporcional (*Supremum Test for Proportionals Hazards Assumption*), la covariable (TRATAM) no presenta evidencia de violación de supuestos de riesgo proporcional ( $p = 0,9780$ ,  $pr > 0,05$ ). Sin embargo, esta prueba no indica la naturaleza de la violación.

## b.2. Prueba se residuales de Schoenfeld

El uso principal de estos residuos es detectar posibles desviaciones del supuesto de riesgos proporcionales. Si se satisface la suposición, los residuos de Schoenfeld deben ser independientes del tiempo. La idea es probar si los residuos de Schoenfeld están correlacionados con el tiempo o con alguna función del tiempo (Allison, 2010).

## The PHREG Procedure

Data Set		Model Information	
	WORK.JABILLO		
Dependent Variable	DÍAS	Timeo de germinación	
Censoring Variable	STATUS	Status de germinación	
Censoring Value(s)	0		
Ties Handling	EFRON		
	Number of Observations Read		500
	Number of Observations Used		500

## Summary of the Number of Event and Censored Values

Total	Event	Censored	Percent Censored
500	369	131	26.20

## Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

## Model Fit Statistics

Criterion	Without Covariates	With Covariates
-2 LOG L	4200.645	4197.239
AIC	4200.645	4199.239
SBC	4200.645	4203.149

## Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	3.4063	1	0.0649
Score	3.4154	1	0.0646
Wald	3.4081	1	0.0649

## The PHREG Procedure

## Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
TRATAM	1	0.01096	0.00594	3.4081	0.0649	1.011

A continuación, el resultado de las correlaciones entre la covariable TRATAM, tiempo y dos transformaciones del tiempo para evaluar su efecto sobre los residuales de Schoenfeld, utilizando PROC RANK para la jerarquización de la variable de tiempo (DIASRANK):

**b.2.1. Resultado de las correlaciones entre los residuales de Schoenfeld para Tratamientos (Schtratam) y variable jerarquizada de tiempo de germinación (Diasrank) por PROC RANK:**

The CORR Procedure							
4 variables: DIASRANK DÍAS1 DÍAS2 SCHTRATAM							
Variable	N	Mean	Simple Statistics			Minimum	Maximum
			Std Dev	Sum			
DÍASRANK	500	250.50	142.88	125250	1.00	435.00	
DÍAS1	500	3.05	0.25057	1528	2.39	3.40	
DÍAS2	500	511.20	254.33634	255600	121.00	900.00	
SCHTRATAM	369	-0.03515	8.33159	-12.97	-12.79	12.23	

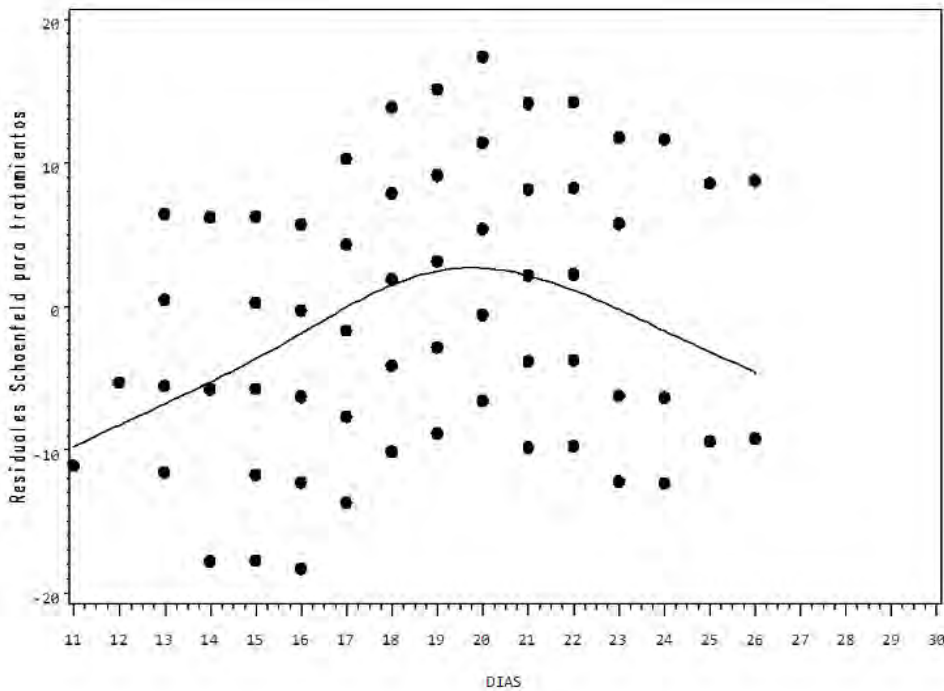
Variable	Label	Simple Statistics	
	DÍAS		Tiempo transcurrido de germinación
	DÍAS1		Transformación logarítmica
	DÍAS2		Transformación al cuadrado
	schTRATAM		Residuales de Schoenfeld para TRATAM

En los resultados mostrados a continuación, como producto de los residuales de Schoenfeld, se computan las correlaciones de los residuales con el tiempo (Diasrank) y las dos funciones transformadas (Días1 y Días2), coinciden con los resultados originados por ASSESS, en el sentido que la covariable (TRATAM) no presenta evidencia de violación de supuestos de riesgo proporcional solo con la variable DIAS1 ( $p = 0,069$ ), pero si existe violación de supuestos con DIASRANK y DIAS2 ( $pr < 0,05$ ). También se observa que el tamaño de muestra de estas correlaciones es de 369, ya que los residuales de Schoenfeld no están definidos para los casos censurados (Allison, 2010).

Pearson Correlation Coefficients				
Prob >  r  under H0: Rho=0				
Number of Observations				
	DIASRANK	DÍAS1	DÍAS2	SCHTRATAM
DIASRANK	1.00000	0.99129	0.95876	-0.10586
Rank for Variable DÍAS		<.0001	<.0001	0.0421
	500	500	500	369
DÍAS1	0.99129	1.00000	0.97846	-0.09463
	<.0001		<.0001	0.0694
	500	500	500	369
DÍAS2	0.95876	0.97846	1.00000	-0.10699
	<.0001	<.0001		0.0400
	500	500	500	369
SCHTRATAM	-0.10586	-0.09463	-0.10699	1.00000
Schoenfeld Residual TRATAM	0.0421	0.0694	0.0400	
	369	369	369	369

Las correlaciones mediante el procedimiento RANK y sus correspondientes valores de  $p$  se indican en el cuadro anterior. Los valores de  $p$  para la variable de tiempo jerarquizada (Diasrank) y DÍAS2 es  $p=0,04$  ( $p<0,05$ ), lo que puede llevar a la conclusión que la suposición de riesgo proporcional sea probablemente violada por ambas covariables. El uso principal de estos residuales es detectar posibles desviaciones del supuesto de riesgos proporcionales. Si se satisface la suposición, los residuales de Schoenfeld deben ser independientes del tiempo. Probando la suposición de riesgos proporcionales es equivalente, por lo tanto, a probar si los residuos de Schoenfeld están correlacionados con el tiempo o con alguna función del tiempo.

En la Figura 20, los residuales de Schoenfeld muestran una tendencia convexa, lo que indica que las suposiciones de riesgo proporcional no se satisfacen y sugieren violación de las suposiciones para tratamientos (TRATAM). Los residuales Schoenfeld están basados en el principio que, para una covariable regresora, la suposición restrictiva  $\beta(t_k) = \beta$  para todo valor de tiempo  $t_k$ . Esto implica que una gráfica de  $\beta(t_k)$  versus tiempo tendrá inclinación cero. La hipótesis nula es teniendo inclinación cero, la cual es equivalente a probar que el log de la función de razón de riesgo (Hr) es constante en el tiempo.



**Figura 20.** Gráfica de residuos de Schoenfeld para la interacción entre el tiempo (DÍAS) y la covariable tratamientos de inmersión de semillas de jabillo (*Hura crepitans* L.) en agua. Fuente. Elaboración propia.

El siguiente método para probar el cumplimiento de supuestos de riesgo proporcional, es mediante interacciones con covariables dependientes del tiempo. En este caso se ilustrará como estimar un modelo Cox que permita que el efecto de los tratamientos y las semillas cambie con el tiempo, originando una prueba definitiva de la suposición de proporcionalidad para esta variable.

### b.3. Análisis de multicolinealidad

La prueba de colinealidad diagnosticó que la covariable tratamientos, alcanzó valores menores de 1,00 ( $VIF < 5$ ), lo que indica que no existe evidencia de colinealidad en la variable.

#### **b.4. Agregación de la covariable Tratamientos individualmente y en interacción al modelo Cox con el Procedimiento PHREG**

Como en el caso del análisis de regresión múltiple, hay varios procedimientos disponibles para construir un modelo de Cox “mejor”, y la base para elegir entre ellos es algo subjetiva. Los programas estadísticos SAS permiten emplear, por ejemplo, el criterio de información de Akaike (AIC), valores p, pruebas de razón de verosimilitud o combinaciones de estos métodos. En este caso, se ilustra el proceso de construcción de modelos con el método del valor p, utilizando un procedimiento de selección hacia adelante. Se inserta individualmente la variable TRATAM en el modelo de Cox (usando PHREG) y se evalúa la significancia estadística. Insertando TRATAM (tratamientos) como una covariable y probando la hipótesis de que  $\beta_1 = 0$ , se origina que  $p < 0,0001$ . Por lo tanto, esta variable se consideran candidatas para el modelo de Cox multivariado.

Sin embargo, al graficar preventivamente la función de supervivencia versus el tiempo de supervivencia y el log (-log (sobrevivencia)) versus el logaritmo del tiempo de supervivencia, no se cumplieron los supuestos del modelo de riesgo proporcional de Cox, obteniendo curvas no paralelas para cada combinación de tratamiento, indicando que los supuestos de PH no se cumplen. Se continúa con los procedimientos como una forma de explicar los procedimientos. A continuación, la salida de resultados del procedimiento PHREG con la agregación individual de las covariables Tratamientos al modelo Cox:

##### **a. /\*Primer Modelo Cox de estimación de Interacciones con Variables de Tiempo Primeros 15 Días del proceso de germinación\*/**

The PHREG Procedure		
	Model Information	
Data Set	WORK.JABILLO	
Dependent Variable	DÍAS	Tiempo de germinacion
Censoring Variable	STATUS	Status de germinacion
Censoring Value(s)	0	
Ties Handling	BRESLOW	



Number of Observations Read	500
Number of Observations Used	500

## Class Level Information

Class	Value	Design Variables			
TRATAM	0	1	0	0	0
	6	0	1	0	0
	12	0	0	1	0
	18	0	0	0	1
	24	0	0	0	0

## Summary of the Number of Event and Censored Values

Total	Event	Censored	Percent Censored
500	369	131	26.20

## Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

## Model Fit Statistics

Criterion	Without Covariates	With Covariates
-2 LOG L	4200.645	4188.783
AIC	4200.645	4198.783
SBC	4200.645	4218.337

## The PHREG Procedure

## Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	11.8614	5	0.0367
Score	12.2599	5	0.0314
Wald	12.1147	5	0.0333

Effect	Type 3 Tests		
	DF	Chi-Square	Pr > ChiSq
TRATAM	4	12.0891	0.0167
TRATDIAS	1	0.0408	0.8398

La toma de agua por una semilla madura es trifásica: toma rápida inicial, fase de meseta ( $\psi$  entre 1 a 1,5 MPa) y nuevo incremento en la absorción de agua, que se corresponde con el período de elongación del embrión o de la radícula. La duración de cada fase dependerá de las características de la semilla (tamaño, contenido de sustratos hidratables, permeabilidad de la cubierta seminal, toma de  $O_2$ , etc.) y de las condiciones externas en las que se produce la imbibición (temperatura, composición del sustrato del suelo, contenido de humedad). Esto significa que cuatro de los tratamientos de tiempos de imbibición en agua, con excepción del tratamiento sin imbibición (0 horas), cumplen con las condiciones para presentar fases idénticas en la toma de agua por las semillas, imbibirse rápidamente y luego, germinar. Los tiempos de imbibición aportarían las diferencias en los tiempos de germinación (Bewley y Black, 1994; Azcón y Talón, 2013).

Según estos autores, una fase rápida de absorción de agua provocaría alteraciones temporales en la permeabilidad diferencial de las membranas de la semilla y, por consiguiente, una pérdida al medio circundante de los inhibidores de la germinación, acelerando o retrasando los tiempos de germinación.

Los coeficientes exponenciados,  $\exp(\beta_i)$  representan las razones de riesgo indicando el tamaño del efecto de cada covariable. Consideradas como el cambio predictivo en el riesgo para un aumento de una unidad en el predictor. Un signo positivo del coeficiente de regresión indica que para una covariable las probabilidades de germinación aumentan durante los primeros 15 días. También significa que la covariable aumenta el riesgo, lo que reduce el tiempo de falla o de germinación. El efecto contrario se produce con los coeficientes negativos. En este caso, para los primeros 15 días del proceso germinativo, los tratamientos de 0 y 6 horas disminuyen las probabilidades de germinación y, en consecuencia, aumentan el tiempo de falla o de germinación, aunque el análisis mostró un efecto no significativo en ellos ( $p= 0,1114$  y  $p= 0,4398$ ), indicando para ambos tratamientos que no existe relación entre ellos y el cambio decreciente de germinación, mientras que para los tratamientos

de 12 y 18 horas, por tener valores positivos sus parámetros aumentan las probabilidades de germinación y, en consecuencia, disminuyen el tiempo de falla, aunque el análisis también mostró un efecto no significativo en ellos ( $p= 0,0878$  y  $p= 0,8728$ ), indicando para ambos tratamientos que no existe relación entre ellos y el cambio creciente de germinación (Cuadro Analysis of Maximum Likelihood Estimates). En esta primera etapa del proceso de germinación, también se observa que la interacción de los tratamientos con el tiempo también tiene un impacto negativo en la germinación. Es decir, la interacción disminuye las probabilidades de germinación en el tiempo ( $\beta_i= - 0,00621$ ), y de manera no significativa ( $p=0,8398$ ), lo cual es un indicio de no violación de supuestos ( $p>0,05$ ) durante este primer período de germinación (0 -15 días).

#### Analysis of Maximum Likelihood Estimates

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limit
TRATAM 0	1	-0.27129	0.17042	2.5339	0.1114	0.762	0.546 1.065
TRATAM 6	1	-0.12791	0.16557	0.5969	0.4398	0.880	0.636 1.217
TRATAM 12	1	0.27675	0.16212	2.9141	0.0878	1.319	0.960 1.812
TRATAM 18	1	0.02636	0.16461	0.0256	0.8728	1.027	0.744 1.418
TRATDIAS	1	-0.00621	0.03072	0.0408	0.8398	0.994	0.936 1.055

Además de la comparación de los tratamientos con el tratamiento de referencia de 24 horas, se realizaron pruebas de hipótesis para comparar entre los demás tratamientos- Para la primera etapa, con un valor de chi cuadrado de 12,09 y 4 grados de libertad, la prueba resulto estadísticamente significativa ( $p=0,0167$ ). Entre los tratamientos comparados no se presentaron diferencias significativas entre 0 y 6 horas ( $p=0,394$ ), entre 6 y 18 horas ( $p=0,352$ ) y entre 12 y 18 horas ( $p=0,124$ ). Para la ocurrencia de diferencias entre tratamientos (posterior a inmersión entre 0 – 6 horas), debe existir una diferencia de tiempo de inmersión entre 12 y más horas. La absorción de agua por las semillas es esencial para el comienzo de la germinación y define cada una de las etapas del patrón trifásico de absorción de agua, que caracteriza el proceso germinativo (Bewley, 1997).

**Analysis of Maximum Likelihood Estimates**

Parameter		Variable Label
TRATAM	0	Tiempos de inmersión, horas 0
TRATAM	6	Tiempos de inmersión, horas 6
TRATAM	12	Tiempos de inmersión, horas 12
TRATAM	18	Tiempos de inmersión, horas 18
TRATDIAS		

**Linear Hypotheses Testing Results**

Label	Wald		Pr > ChiSq
	Chi-Square	DF	
Test 1	0.7265	1	0.3940
Test 2	10.8464	1	0.0010
Test 3	3.0510	1	0.0807
Test 4	6.2279	1	0.0126
Test 5	0.8652	1	0.3523
Test 6	2.3701	1	0.1237

**b. /\*Segundo Modelo Cox de estimación de Interacciones con Variables de Tiempo en el intervalo 16 – 30 días del proceso de germinación\*/**

En esta segunda etapa del proceso de germinación, la interacción de los tratamientos con el tiempo tiene un impacto positivo en la germinación. Es decir, la interacción de los tratamientos con el tiempo, aumentaron las probabilidades de germinación ( $\beta_i = 0,00743$ ) en el tiempo. De manera individual, el signo negativo de los coeficientes de regresión para los tratamientos de inmersión de 0 y 6 horas, indica que las probabilidades de germinación también disminuyeron durante este período, con más baja probabilidad de germinación que el tratamiento de referencia (inmersión en agua durante 24 horas), a diferencia de los tratamientos de 12 y 18 horas de inmersión, donde el signo positivo de los coeficientes indica que las probabilidades de germinación también aumentaron durante este período y, en consecuencia, disminuyen el tiempo de germinación, aunque el análisis también mostró un efecto no significativo en ellos ( $p = 0,189$  y  $p = 0,732$ ), indicando para ambos tratamientos que no existe relación entre ellos y el cambio creciente de germinación (Tabla de Análisis de máxima probabilidad).

	Model Information	Data Set
WORK.JABILLO		
Dependent Variable	DÍAS	Tiempo de germinación
Censoring Variable	STATUS	Status de germinación
Censoring Value(s)	0	
Ties Handling	EFRON	

Number of Observations Read	500
Number of Observations Used	500

## Class Level Information

Class	Value	Design Variables			
TRATAM	0	1	0	0	0
	6	0	1	0	0
	12	0	0	1	0
	18	0	0	0	1
	24	0	0	0	0

## Summary of the Number of Event and Censored Values

Total	Event	Censored	Percent Censored
500	369	131	26.20

## Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

## Model Fit Statistics

Criterion	Without Covariates	With Covariates
-2 LOG L	4200.645	4188.694
AIC	4200.645	4198.694
SBC	4200.645	4218.248

## The PHREG Procedure

## Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	11.9513	5	0.0355
Score	12.3457	5	0.0303
Wald	12.1980	5	0.0322

## Type 3 Tests

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRATAM	4	9.0145	0.0607
TRATDIAS	1	0.1308	0.7176

## Analysis of Maximum Likelihood Estimates

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limit
TRATAM	0	1	-0.10470	0.47359	0.0489	0.8250	0.901 0.356 2.279
TRATAM	6	1	-0.00314	0.37046	0.0001	0.9932	0.997 0.482 2.060
TRATAM	12	1	0.35978	0.27393	1.7249	0.1891	1.433 0.838 2.451
TRATAM	18	1	0.06808	0.19887	0.1172	0.7321	1.070 0.725 1.581
TRATDIAS	1	0.00743	0.02054	0.1308	0.7176	1.007	0.968 1.049

El valor de la razón de riesgo para los tratamientos de 0 y 6 horas, resultó mayor que durante el primer período, con valores de 0,901 y 0,997 ( $HR < 1$ ), revelando que la probabilidad de experimentar germinación se incrementa por estos factores con el tratamiento inmersión en agua durante 24 horas. Es decir, el riesgo de germinar de las semillas es alrededor de 90 % y 99 %, de la razón de riesgo de las semillas del tratamiento de referencia. Al comparar con el primer período, el tratamiento de referencia, comparado con los tratamientos sin inmersión en agua y 6 horas de inmersión, presentó un aumento de las probabilidades de germinación por el orden de 14 % y 11 %. Este resultado también deja ver que las semillas de estos tratamientos tuvieron tendencia a germinar posteriormente que las semillas con inmersión en agua durante 24 horas, durante el segundo período (Cuadro anterior).

En el caso de los tratamientos (12 – 18 horas), que presentaron signos positivos en sus coeficientes, significa que, en este período, también presentan más alta probabilidad de germinación que el tratamiento de referencia (inmersión de semillas en agua durante 24 horas). Sus valores de la razón de riesgo ( $HR > 1$ ), de 1,43 y 1,07, respectivamente, revelan que experimentan mayor riesgo de germinación que el tratamiento de referencia, es decir, la probabilidad de experimentar germinación disminuye por estos factores con el tratamiento inmersión de semillas en agua durante 24 horas, comparado con ambos tratamientos. Por lo tanto, aumentaran la función de riesgo a través del periodo de germinación por 43 % y 7 %, en comparación con las semillas del tratamiento de referencia. Este resultado revela que las semillas de estos tratamientos continúan con tendencia a germinar primero que las semillas con inmersión en agua durante 24 horas.

Igual que en el primer periodo, todos los tratamientos incluyen en el intervalo de confianza el valor nulo (1), indicando que el riesgo de germinación fue similar en los dos grupos (0=24; 6=24; 12=24 y 18=24 horas). Por lo tanto, tampoco existen diferencias significativas del riesgo temporal de germinación entre las semillas de los tratamientos bajo comparación. La interpretación correcta sería que las semillas de los tratamientos 0 – 6 – 12 y 18 horas que no hubiesen germinado durante este periodo, tenían igual probabilidad de presentarla en el siguiente intervalo de tiempo que las semillas del grupo control o tratamiento de referencia (con inmersión en agua durante 24 horas).

En esta segunda etapa del proceso de germinación, también se observa que la interacción de los tratamientos con el tiempo tiene un impacto positivo en la germinación. Es decir, la interacción aumenta las probabilidades de germinación en el tiempo ( $\beta_i = 0,00743$ ), aunque no significativamente ( $p = 0,7176$ ), lo cual es un indicio de no violación de supuestos ( $p > 0,05$ ) durante este segundo período de germinación (16 -30 días). Para la segunda etapa, la comparación entre tratamientos, sin incluir sus comparaciones con el tratamiento de referencia (con inmersión en agua durante 24 horas), con un valor de chi cuadrado de 9,0145 y 4 grados de libertad, la prueba no resulto estadísticamente significativa ( $p = 0,061$ ). Entre los tratamientos comparados no se presentaron diferencias significativas, a diferencia de la primera etapa del proceso de germinación (0 – 15 días) (Analysis of Maximum Likelihood Estimates, Linear Hypotheses Testing Results).

**Analysis of Maximum Likelihood Estimates**

Parameter	Variable	Label
TRATAM	0	Tiempos de inmersión, horas 0
TRATAM	6	Tiempos de inmersión, horas 6
TRATAM	12	Tiempos de inmersión, horas 12
TRATAM	18	Tiempos de inmersión, horas 18
TRATDIAS		

**Linear Hypotheses Testing Results**

Label	Wald Chi-Square	DF	Pr > ChiSq
Test 1	0.2547	1	0.6138
Test 2	2.8181	1	0.0932
Test 3	0.2161	1	0.6420
Test 4	3.4189	1	0.0645
Test 5	0.0669	1	0.7959
Test 6	2.2170	1	0.1365

**c. /\*Tercer Modelo Cox de estimación de Interacciones con Variables de Tiempo en el intervalo total de días del proceso de germinación\*/**

The PHREG Procedure

Model Information

Data Set	WORK.JABILLO		
Dependent Variable	DÍAS	Tiempo de germinación	
Censoring Variable	STATUS	Status de germinación	
Censoring Value(s)	0		
Ties Handling	EFRON		
Number of Observations Read			500
Number of Observations Used			500

Class Level Information

Class	Value	Design Variables			
TRATAM	0	1	0	0	0
	6	0	1	0	0
	12	0	0	1	0
	18	0	0	0	1
	24	0	0	0	0



## Summary of the Number of Event and Censored Values

Total	Event	Censored	Percent Censored
500	369	131	26.20

## Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

## Model Fit Statistics

Criterion	Without Covariates	With Covariates
-2 LOG L	4200.645	4184.294
AIC	4200.645	4194.294
SBC	4200.645	4213.848

## The PHREG Procedure

## Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	16.3509	5	0.0059
Score	16.2393	5	0.0062
Wald	15.9204	5	0.0071

## Type 3 Tests

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRATAM	4	13.9326	0.0075
TRATDIAS	1	4.4391	0.0351

La Tabla de Análisis de máxima probabilidad (Analysis of Maximum Likelihood Estimates) muestra los parámetros calculados por la aplicación del modelo de Cox, con las interacciones de los tratamientos de inmersión y el tiempo, considerando los días totales del proceso de germinación. El signo negativo de los coeficientes de regresión para los diferentes tratamientos ( $\beta_1 = -2,24755$ ;  $\beta_2 = -$

1,5962;  $\beta_3 = -0,69297$  y  $\beta_4 = -0,45836$ ) y su interacción con el tiempo ( $\beta_i = -0,0043$ ), indica bajas probabilidades de germinación y, en consecuencia, aumentan el tiempo de falla. Al compararlos con el tratamiento de referencia (*inmersión de semillas en agua durante 24 horas*).

Analysis of Maximum Likelihood Estimates								
Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limit	
TRATAM 0	1	-2.24755	0.95647	5.5217	0.0188	0.106	0.016	0.689
TRATAM 6	1	-1.59620	0.71431	4.9935	0.0254	0.203	0.050	0.822
TRATAM 12	1	-0.69297	0.48368	2.0527	0.1519	0.500	0.194	1.290
TRATAM 18	1	-0.45836	0.27985	2.6827	0.1014	0.632	0.365	1.094
TRATDIAS	1	-0.00430	0.00204	4.4391	0.0351	0.996	0.992	1.000

El análisis mostró un efecto significativo en los tratamientos de inmersión para los tratamientos sin inmersión (0 horas) y para 6 horas ( $p = 0,0188$  y  $p = 0,0254$ ) indicando para ambos tratamientos que existe relación entre ellos y el cambio decreciente de germinación, mientras que, para los tratamientos de 12 y 18 horas, con valores de  $p > 0,05$  (0,1519 y 0,1014), no existe relación entre ellos y el cambio decreciente de germinación. Para todo el proceso de germinación, también se observa que la interacción de los tratamientos con el tiempo tiene un impacto negativo en la germinación. Es decir, la interacción disminuye las probabilidades de germinación en el tiempo ( $\beta_i = -0,0043$ ), y de manera significativa ( $p = 0,0351$ ), lo cual es un indicio de violación de supuestos ( $p < 0,05$ ) cuando se incluye la duración de todo el proceso de germinación (30 días).

El valor de la razón de riesgo para los diferentes tratamientos, resultó de 0,106; 0,203; 0,500 y 0,632, revelando que la probabilidad de experimentar germinación se incrementa por estos factores con el tratamiento *inmersión de semillas en agua durante 24 horas*, comparado con los demás tratamientos. Es decir, el riesgo de germinar de las semillas es alrededor de 10,6 %, 20,3 %, 50,0 % y 63,2 % de la razón de riesgo de las semillas del tratamiento de referencia. Este resultado revela que las semillas de estos tratamientos tendrán tendencia a germinar posteriormente que las semillas con *inmersión en agua durante 24 horas* (Analysis of Maximum Likelihood Estimates). El intervalo de confianza de la razón de riesgo (HR) de los tratamientos de 12 y 18 horas de inmersión, incluye dentro del intervalo el valor nulo (1) que indica la igualdad entre ellos y el tratamiento de referencia. Determina que el riesgo de germinación fue similar en los dos grupos bajo comparación (12=24 horas;

18=24 horas). Por lo tanto, no existen diferencias significativas del riesgo temporal de germinación entre las semillas de los tratamientos bajo comparación. La interpretación correcta sería que las semillas de los tratamientos 12 y 18 horas que no hubiesen germinado tenían igual probabilidad de presentarla que las semillas del grupo control o tratamiento de referencia (24 horas). Por el contrario, en los tratamientos sin inmersión (0 horas) e inmersión durante 6 horas, al no incluir el intervalo de confianza el valor nulo (1), se determina que el riesgo de germinación fue diferente en los dos grupos bajo comparación (0 ≠ 24 horas; 6 ≠ 24 horas). Por lo tanto, existen diferencias significativas del riesgo temporal de germinación entre las semillas de los tratamientos bajo comparación. Significa que las semillas de los tratamientos 0 y 6 horas que no hubiesen germinado no tenían igual probabilidad que las semillas del grupo control o tratamiento de referencia (24 horas).

Las pruebas de hipótesis para comparar entre los demás tratamientos, originando una prueba, donde todos los coeficientes asociados con esta covariable tienen valor de 0. Para la totalidad de datos, con un valor de chi cuadrado de 13,933 y 4 grados de libertad, la prueba resultó estadísticamente significativa ( $p=0,0075$ ). Entre los tratamientos comparados, la única diferencia no significativa (Test 6,  $p=0,4041$ ) se presentó entre los tratamientos de inmersión de 12 y 18 horas. En las demás comparaciones se presentaron diferencias significativas a altamente significativas (Analysis of Maximum Likelihood Estimates, Linear Hypotheses Testing Results).

#### Analysis of Maximum Likelihood Estimates

Parameter		Variable Label
TRATAM	0	Tiempos de inmersion, horas 0
TRATAM	6	Tiempos de inmersion, horas 6
TRATAM	12	Tiempos de inmersion, horas 12
TRATAM	18	Tiempos de inmersion, horas 18
TRATDIAS		

#### Linear Hypotheses Testing Results

Label	Wald		
	Chi-Square	DF	Pr > ChiSq
Test 1	4.7691	1	0.0290
Test 2	9.1767	1	0.0025
Test 3	5.9277	1	0.0149
Test 4	9.7716	1	0.0018
Test 5	5.2408	1	0.0221
Test 6	0.6961	1	0.4041

### c. Resultado de métodos paramétricos con el procedimiento LIFEREG

En el análisis de residuos de Schoenfeld de los modelos para verificar los supuestos de proporcionalidad, las covariables y sus interacciones, satisfacen los supuestos de riesgo proporcional. Por lo tanto, el siguiente análisis paramétrico se realiza para complementar la información de la germinación y para mejorar la interpretación de los resultados generales.

La aplicación del Criterio de Información Akaike (Akaike, 1974), permitió seleccionar la distribución Lognormal por mejor ajuste. La distribución log-normal para la función de riesgo no siempre se considera apropiada en otros campos, ya que su patrón considera un valor inicial igual a 0, aumenta hasta un máximo con el aumento del tiempo y eventualmente disminuye a cero a medida que  $t$  aumenta, y en muchos análisis la función de riesgo generalmente no se acercará a cero en un tiempo prolongado, ya que todos los objetos eventualmente fallarán. Sin embargo, este patrón parece ajustarse a la tendencia ordinaria del proceso de germinación de la presente investigación, lo que hace que este modelo sea adecuado para ensayos de germinación (Lee y Wenyu, 2003; Lawless, 2003)

La salida de resultados del Análisis de Parámetros, muestra las estimaciones de máxima verosimilitud (MLE) de los parámetros calculados mediante la aplicación del modelo Lognormal, expresado como regresión de tiempo de falla acelerada (AFT). La primera observación, es que el valor estimado de  $\sigma$  (Escala) es  $< 1$ , lo que implica que la función de riesgo estimada sigue la forma de una “U” invertida. Para el modelo Lognormal, cuando el parámetro de escala,  $\sigma < 1$ , el riesgo comienza en 0, asciende a un pico y luego decrece a 0. El parámetro de escala para la distribución Lognormal resultó  $0,1566 < 1$  (Allison, 2010).

El coeficiente de regresión para el tratamiento *inmersión en agua durante 24 horas*, es 0. Su razón es que se designó como nivel de referencia para comparar con los demás tratamientos. El análisis de la Tabla de Análisis de Parámetros (Analysis of Parameter Estimates), indica que cada coeficiente, e manera independiente, difiere significativamente del tratamiento *inmersión en agua durante 24 horas* ( $p < 0,05$ ).

### The LIFEREG Procedure

#### Model Information

Data Set	WORK.JABILLO	
Dependent Variable	Log(DÍAS)	Tiempo de germinación
Censoring Variable	STATUS	Status de germinación
Censoring Value(s)	0	
Number of Observations	500	
Noncensored Values	369	
Right Censored Values	131	
Left Censored Values	0	
Interval Censored Values	0	
Name of Distribution	Lognormal	
Log Likelihood	79.547555253	

Number of Observations Read	500
Number of Observations Used	500

#### Class Level Information

Name	Levels	Values
TRATAM	5	0 6 12 18 24

Algorithm converged.

#### Type III Analysis of Effects

Effect	DF	Chi-Square	Pr > ChiSq
TRATAM	4	788.8803	<.0001
TRATDIAS	1	666.6866	<.0001

#### Analysis of Parameter Estimates

Parameter	DF	Estimate	Standard Error	95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	1.1548	0.0715	1.0146	1.2950	260.78	<.0001
TRATAM	0	1.9844	0.0735	1.8404	2.1285	729.12	<.0001
TRATAM	6	1.4323	0.0551	1.3243	1.5402	676.14	<.0001
TRATAM	12	0.9061	0.0423	0.8232	0.9890	458.61	<.0001
TRATAM	18	0.4640	0.0312	0.4029	0.5250	221.77	<.0001
TRATAM	24	0.0000	.	.	.	.	.
TRATDIAS	1	0.0039	0.0002	0.0036	0.0042	666.69	<.0001
Scale	1	0.1566	0.0060	0.1452	0.1688		

En la estimación de los parámetros para el modelo Lognormal, la intersección está representado por  $\beta_0 = 1,1548$ , tratamiento 0 horas ( $\beta_2 = 1,9844$ ), tratamiento 6 horas ( $\beta_3 = 1,4323$ ), tratamiento 12 horas ( $\beta_4 = 0,9061$ ) y tratamiento 18 horas ( $\beta_5 = 0,4640$ ). Según Romano y Stevanato (2020), estos valores se utilizan para calcular el tiempo medio de germinación ( $T_{50}$ ). El  $T_{50}$  (mediana o tiempo medio de germinación) para referencia de cada parámetro, se calcula como  $\exp(\beta)$ . Para todo el lote (intercepto =  $\beta_0 = 1,1548$ ), fue de  $\exp(1,1548) = 3,2$  días, mientras que para el efecto de tratamiento 0 horas es,  $\exp(\beta_0 + \beta_1) = \exp(1,1548 + 1,9844) = 23,1$  días; para tratamiento 6 horas,  $\exp(\beta_0 + \beta_2) = 13,3$  días; para tratamiento 12 horas,  $\exp(\beta_0 + \beta_3) = 7,9$  días; para tratamiento 18 horas,  $\exp(\beta_0 + \beta_4) = 5,0$  días. De esta manera, se calcula el tiempo medio de germinación para cada tratamiento, según su parámetro. Cuanto más negativo sea el coeficiente, más temprano será el tiempo medio de germinación (Romano y Stevanato, 2020). Los resultados de la prueba chi cuadrada indican diferencias altamente significativas entre estas comparaciones, con el tratamiento de referencia (Tabla de análisis de parámetros). De manera similar, las comparaciones del análisis de las tablas de vida (LIFETES) también sugieren que existen diferencias entre los tratamientos ( $p=0,0003$ ). Sin embargo, para Fox (2001), el modelo de tiempo de falla acelerada (LIFEREG) tiene mayor poder estadístico y resulta razonable tener mayor confianza en sus resultados.

En términos de la razón de tiempo ( $\gamma$ ), para tratamiento 0 horas  $\gamma = 23,1/3,2 = 7,2$ ; de manera similar, los demás tratamientos tienen razones de tiempo de 4,2; 2,5; 1,6 y un valor de 1 para el tratamiento inmersión en agua durante 24 horas. Estos valores, interpretados como razones estimadas de los tiempos de sobrevivencia esperados para dos grupos, según la fórmula  $(100\% (1 - \gamma))$ , se puede observar que los tratamientos 0 – 6 – 12 y 18 horas, retardan o prolongan la germinación en 620,0 %, 320,0 %, 150,0 % y en 60,0 %, en comparación con el tratamiento inmersión en agua durante 24 horas, considerado el tratamiento de referencia, respectivamente. Esto significa que el tratamiento que prolonga u origina una germinación más lenta, es el tratamiento 0 horas (sin inmersión en agua). El tratamiento 24 horas (inmersión en agua durante 24 horas) se considera de referencia, por lo que aparece con un valor neutro (0,00). En general, en los modelos AFT, una razón de tiempo  $> 1$  para una covariable, implica que prolonga el tiempo de germinación, mientras que una razón de tiempo  $< 1$  implica que es más probable un evento de germinación más temprano (Romano y Stevanato 2020). Según esta referencia, el

tratamiento 0 horas (sin inmersión en agua) es el que más prolonga el tiempo de germinación. En los tratamientos de inmersión en agua (6 – 12 – 18 y 24 horas) es notable que, a mayor tiempo de inmersión, la germinación tiene menor prolongación de la germinación.

El hecho que cada coeficiente de regresión de manera independiente es significativo, indica que cada uno difiere del tratamiento de referencia. Estas comparaciones múltiples, son análogas a las del análisis LIFETEST, excepto que en este caso se utilizaron las pruebas de Wilcoxon y en el ATF, los coeficientes de regresión. El análisis LIFETEST también mostró heterogeneidad entre tratamientos, con diferencias altamente significativas entre ellos. Los coeficientes de regresión llegan a una conclusión semejante: la comparación de los demás tratamientos con el tratamiento de referencia, es significativamente diferente del tratamiento de inmersión en agua durante 24 horas, y las comparaciones entre los tratamientos restantes permanecen variables, es decir, dependiendo del periodo o etapa de germinación, resultan significativamente diferentes o no.

Según Romano y Stevanato (2020), mientras que el modelo de Cox expresa el efecto multiplicativo de las variables explicativas sobre el riesgo (escala de riesgo), el modelo AFT expresa el mismo efecto sobre el tiempo de supervivencia (escala de tiempo). Esta característica permite una interpretación más fácil de los resultados porque los parámetros miden la fuerza y el efecto de la covariable correspondiente en la mediana del tiempo de supervivencia. Además, por la aplicación de tiempo de falla acelerado, el efecto de las covariables sobre la supervivencia se describe en términos absolutos (por. ejemplo, número de días) en lugar de términos relativos como el índice de riesgo. Además, las razones de tiempo representan índices útiles que describen los efectos de las variables explicativas sobre la germinación y más interpretable que una relación de dos riesgos.

Para Onofri *et al.* (2010), los modelos AFT asumen que cada individuo de la cohorte bajo investigación, experimentará el evento tarde o temprano. Significa que la curva acumulada de proporción de semillas germinadas debe acercarse a una asíntota, con el tiempo tendiendo al infinito, lo que resulta poco realista para las semillas, considerando que muchas de ellas podrían haber perdido la capacidad de germinar debido a condiciones ambientales desfavorables.

### 3.3.1.2. Análisis de supervivencia en germinación de *Cassia moschata* Kunth, con aplicación de hormonas vegetales (Vale, 2023)

#### Materiales y métodos

##### Material vegetal

Se utilizaron semillas extraídas de frutos colectados de árboles plantados en los jardines de la Villa Universitaria del Núcleo Rafael Rangel, Universidad de Los Andes, Trujillo, Venezuela, que se encuentra a 392,49 msnm. Para lo cual fue necesario utilizar un descopador. Las vainas o legumbres se colocaron al aire libre para culminar su secado y liberación de semillas, las cuales se almacenaron en envase de plástico a temperatura ambiente y protegidas de la luz hasta el establecimiento del ensayo (2 – 3 meses).

##### Hormonas sintéticas utilizadas

Giber Grop®.

Es un concentrado soluble en agua como regulador fisiológico de las plantas, para floración y enguese de flores y frutos. Es un producto que contiene auxinas y giberelinas (Ácido alfa-naftalenacético 17,2% y Giberelina (AG3) 10%).

##### Diseño experimental y tratamientos pregerminativos

El experimento fue realizado mediante un completamente aleatorizado con cuatro repeticiones, cada una contentiva de veinte semillas (20), para ochenta (80) semillas por tratamiento y un total de cuatrocientas (400) semillas, sembradas a una profundidad de 0,5 cm. Antes de la siembra, las semillas se sumergieron en diferentes concentraciones del producto Giber Grop a temperatura ambiente, durante 12 y 24 horas. Los tratamientos pregerminativos a evaluar fueron:

T1: Inmersión de semillas en solo agua (testigo, 0 mg/L) durante 24 horas;

T2: Inmersión de semillas en Giber Grop® a 200 mg/L durante 24 horas;

T3: Inmersión de semillas en Giber Grop® a 400 mg/L durante 24 horas;

T4: Inmersión de semillas en Giber Grop® a 600 mg/L durante 24 horas;



T5: Inmersión de semillas en Giber Grop® a 600 mg/L durante 12 horas;

El procedimiento a seguir fue el siguiente:

En un envase plástico limpio, se mezcló un litro de agua, con un surfactante, en dosis de 4 cc por litro de agua y la concentración del tratamiento correspondiente y en cada uno de ellos se colocaron las semillas, cubiertas completamente por la solución. Luego de la inmersión en el tratamiento correspondiente, se extendieron las semillas en una superficie limpia y se dejaron secar durante 30 minutos. Este procedimiento se realizó en un lugar sombreado, a primeras horas de la mañana.

### **Siembra de semillas y evaluación de la germinación**

La siembra de semillas se hizo por siembra directa en gaveras para germinación, colocando una (01) semillas por abertura, sembradas a una profundidad de 1,0 cm. Las gaveras para germinación se mantuvieron bajo condiciones a la sombra (sin luz solar directa) durante la evaluación del experimento. Mientras las semillas estuvieron en germinación, los riegos con agua se continuaron 2 veces por día. Para efectos prácticos, se consideró que una semilla había germinado al emerger el talluelo de la plántula, llevando un registro diario de semillas germinadas por tratamiento durante 30 días después de la siembra (DDS). Cada semilla individual seleccionada para la germinación por tratamiento se consideró una unidad experimental. Con esta información, se realizaron las determinaciones del número de semillas germinadas diariamente y acumuladas por tratamiento.

### **Hipótesis y nivel de significación**

Afortunadamente, los métodos utilizados para analizar los datos de tiempo de falla o germinación pueden manejar correctamente los datos censurados y no requieren datos distribuidos normalmente (Fox, 2001; Allison, 2010).

El procedimiento LIFETEST permite comparar pares de curvas de supervivencia y detectar si existen diferencias estadísticamente significativas entre ellas y probar la hipótesis nula de que las funciones de supervivencia son iguales entre curvas. El nivel de significancia del análisis de la función de supervivencia se ajustó al nivel  $p < 0,05$ . Para comparar las curvas de supervivencia, la prueba t de Student y la prueba de rangos de Wilcoxon en algunos casos no resultan apropiados. Para la prueba t de Stu-

dent, la razón es que los datos de germinación no siguen una distribución normal, lo cual es un requisito necesario para realizar esta prueba. Por el contrario, la prueba de Wilcoxon, se podría utilizar en los casos en que los tratamientos tengan mayor efecto en las fases iniciales de la germinación donde la prueba de Wilcoxon dará una aproximación válida (McNair *et al.*, 2012; Scherm y Ojiambo, 2004; Fox, 2001).

### **Estructura y análisis de datos**

Para solucionar las preguntas planteadas, se utilizaron tres métodos estadísticos diferentes: análisis de las tablas de vida (LIFETEST o métodos de Kaplan-Meier), modelo de riesgo proporcional de Cox y por el modelo de tiempo de falla acelerada (ATF).

Se adoptó un esquema de “observación continua” durante la duración del ensayo, asumiendo que los tiempos de germinación tuvieran un “dato exacto” para cada semilla (McNair *et al.*, 2012). El tiempo de germinación para la semilla no germinada se consideró como “censurada a la derecha”.

Las variables tratamientos de inmersión en hormona sintética, se consideraron como covariables categóricas. Para el análisis, en cada medición se asignaron valores de 1 o 0 para las semillas germinadas y no germinadas, respectivamente. Luego de cada conteo diario, cada semilla germinada fue removida del experimento. Se utilizaron códigos para discriminar entre los tratamientos (T1=1, T2=2, T3=3, T4=4 y T5=5).

Los datos de germinación se analizaron por análisis de supervivencia con el programa SAS (SAS, 2020; Fox, 2001), de la manera siguiente:

Data Canafis;

Input Días Tratam Status;

Label Tratam='Tiempos De Inmersion'

Días='Tiempo De Germinacion'

Status='Status De Germinacion'

Tratam 1='Inmersión en agua durante 24 horas'

Tratam 2='Inmersión en Giber Grop a 200 ppm 24 horas'

Tratam 3='Inmersión en Giber Grop a 400 ppm 24 horas'

Tratam 4='Inmersión en Giber Grop a 600 ppm 24 horas'

Tratam 5='Inmersión en Giber Grop a 600 ppm 12 horas';

Tratdias=Tratam\*Días;

Datalines;

5 2 1

5 4 1

5 5 1

6 3 1

6 4 1

.....

.....

30 5 0

30 5 0

30 5 0

Proc Print Data=Canafis;

Run;

Proc Sort Data=Canafis;

By Descending Tratam;

Run;

## Análisis no paramétrico

Los datos de germinación fueron descritos y modelados bajo dos probabilidades relacionadas: supervivencia y función de riesgo, ambas dependientes del tiempo. Este análisis se hizo con el procedimiento PROC LIFETEST, que imprime por defecto, los estadísticos Log-rank y Wilcoxon, así como valores de chi-cuadrado para las comparaciones entre tratamientos. Las diferencias de supervivencia entre covariables, se analizaron a partir de curvas de supervivencia construidas por el método Kaplan – Meier (Kaplan y Meier, 1958; Klein *et al.*, 2001; Gómez *et al.*, 2009); para ello, se tomó en cuenta el estatus de cada semilla (status 1 = germinación, status 0 = censura debido a pérdida durante el experimento o semilla no germinada al final del experimento), así como el tiempo de germinación de la misma en días. Los valores de media del tiempo, tiempo medio de germinación y otros índices del tiempo de germinación, se obtienen con este procedimiento. Según Onofri *et al.* (2010), los tiempos de germinación se usan para calcular la probabilidad de “germinación/emergencia”, es decir, la probabilidad que una semilla individual germine después de un tiempo específico  $t$ , considerando el evento inicial (imbibición) en tiempo  $t = 0$ .

Este análisis compara la respuesta de las curvas de germinación por el método de máxima verosimilitud o probabilidad, utilizando la prueba de Wilcoxon en caso de presentarse cruzamientos entre curvas de germinación y, en caso contrario, la prueba de Log-Rank (Bárcenas *et al.*, 2013; SAS, 2020; Fox, 2001). La prueba Log-Rank otorga igual peso al principio y final de la germinación y asume que las funciones de riesgo para los tratamientos son paralelas, sin cruzamiento entre curvas, indicando que las razones de riesgo son constantes o proporcionales en todos los períodos de tiempo. Solo en esta situación es una prueba eficaz (Fleming *et al.*, 1987; Lee, 1996; Buyske *et al.*, 2000; Lee y Wenyu, 2003). Cuando la razón de riesgo no es constante, las pruebas de Gehan – Wilcoxon y la prueba Tarone – Ware pueden ser más eficaces (Tarone y Ware, 1977; Pepe y Fleming, 1989). La prueba Peto-Peto también es eficiente cuando se viola la suposición de riesgo proporcional (Kleinbaum y Klein, 2005).

El procedimiento para el análisis no paramétrico fue el siguiente:

```
Legend1 Frame Cframe=Ligr Cborder=Black
```

```
Position=Center Value=(Justify=Center);
```

```
Axis1 Label=(Angle=90 Rotate=0 'Estimated Cdf') Minor=None;
```

```
Axis2 Minor=None;
```

```
Symbol1 Value=Dot C=Black H=0.15in I=None R=1;
```

```
Symbol2 Value=Circle C=Black H=0.15in I=None R=1;
```

```
Symbol3 Value=Square C=Black H=0.15in I=None R=1;
```

```
Symbol4 Value=Triangle C=Black H=0.15in I=None R=1;
```

```
Symbol5 Value=Star C=Black H=0.15in I=None R=1;
```

```
Proc Lifetest Data=Canafis Method=Lt Plots=(S,H);
```

```
Time Días*Status (0);
```

```
Strata Tratam;
```

```
Run;
```

```
Proc Lifetest Data=Canafis Method=Km Plots=(Ls, Lls);
```

```
Time Días*Status (0);
```

```
Strata Tratam;
```

```
Run;
```

### **Análisis semiparamétrico.**

En la práctica, es relativamente sencillo probar si se cumple el supuesto de riesgo proporcional desarrollando un modelo de riesgo proporcional que incluye un término de interacción entre el tiempo y la covariable de interés, además de la propia covariable; si el término de interacción es estadísticamente significativo, el modelo AFT es más apropiado (Lee y Wang, 2003). Se utilizaron dos métodos adicionales para verificar y avalar las suposiciones de riesgo proporcional. El primero, un análisis gráfico de Kaplan – Meier, utilizando las curvas de  $\log(-\log(S(t)))$  versus  $\log(t)$ . El segundo, mediante el análisis de residuales de Schoenfeld. También se corrió una prueba para analizar la multicolinealidad entre las variables, basado en el diagnóstico VIF. Un valor de  $VIF > 5$  se considera evidencia de multicolineali-

dad y un valor VIF  $> 10$  es indicio de fuerte multicolinealidad. Para determinar el efecto de los factores de riesgo de las covariables tratamientos y semillas sobre la función de germinación, se abordó el análisis de sobrevivencia de forma multivariada, utilizando la función de regresión de riesgos proporcionales de Cox, que relaciona varios factores de riesgo de forma simultánea al tiempo de sobrevivencia (Cox, 1972).

Para ajustar un modelo de PH (riesgo proporcional) a los datos, se utilizó el procedimiento PHREG en SAS, con los términos de interacción entre ambas covariables y el tiempo. El valor nulo para la HR es el uno, que indica igualdad de probabilidad de producción del suceso en los dos grupos en el siguiente intervalo de tiempo. Una  $HR > 1$  indica más riesgo de producción en el tratamiento y una  $HR < 1$  menor riesgo en el tratamiento de intervención que en el control (tratamiento de referencia). En este sentido, una  $HR = 2$  no significa que las semillas bajo un tratamiento desarrollan el suceso dos veces más rápido, sino que, si una semilla no ha presentado todavía el suceso (germinación), tiene el doble de probabilidad de presentarlo que un control (tratamiento de referencia) durante el siguiente periodo de tiempo.

```
/*Prueba De Multicolinealidad*/
```

```
Proc Reg Data = Canafis;
```

```
Model Días = Tratam / Vif;
```

```
Run;
```

```
/*Proporcionabilidad Con Assess*/
```

```
Proc Phreg Data=Canafis;
```

```
Model Días*Status (0) =Tratam / Ties=Efron;
```

```
Assess Ph / Resample;
```

```
Run;
```

```
/*Residuales Schoenfeld*/
```

```
Proc Phreg Data=Canafis;
```

```
Model Días*Status (0) =Tratam / Ties=Efron;
```

```
Output Out=Schoen Rensch=Schtratam;
```

```
Run;
```

```
/*Generación De Gráfica De Residuales Schoenfeld Para Tratamientos Como Función Del Tiempo*/
```

```
Proc Gplot Data =Schoen;
```

```
Symbol1 V=Dot C=Black Width=1 I=Sm80s;
```

```
Plot Schtratam*Días/Haxis=Axis1 Vaxis=Axis2;
```

```
Axis1 Label=('Días');
```

```
Axis2 Label = (A=90 'Residuales Schoenfeld Para Tratamientos');
```

```
Run;
```

El procedimiento para la jerarquización de la variable de tiempo y su correlacion es el siguiente:

```
Proc Print Data=Schoen;
```

```
Run;
```

```
Data Event;
```

```
Set Schoen;
```

```
Run;
```

```
Proc Rank Data=Event Out=Resrank Ties=Mean;
```

```
Var Días;
```

```
Ranks Diasrank;
```

```
Run;
```

```
Proc Print Data=Resrank;
```

```
Run;
```

Data C;

Set Resrank;

Días1=Log (Días);

Días2=Días\*\*2;

Proc Corr;

Var Diasrank Días1 Días2 Schtratam;

Run;

**/\*Primer Modelo Cox De Estimación De Interacciones Con Variables De Tiempo Primeros 15 Días del proceso de germinacion\*/**

Proc Phreg Data=Canafis;

Class Tratam;

Model Días\*Status (0) = Tratam Tratdías / Ties=Efron Rl;

Test Tratam1=Tratam2; Test Tratam1=Tratam3; Test Tratam1=Tratam4;

Test Tratam2=Tratam3; Test Tratam2=Tratam4; Test Tratam3=Tratam4;

Tratdías = Tratam\*(Días< 15);

Run;

**/\*Segundo Modelo Cox De Estimación De Interacciones Con Variables De Tiempo para el periodo 15-30 Días del proceso de germinación\*/**

Proc Phreg Data=Canafis;

Class Tratam;

Model Días\*Status (0) = Tratam Tratdías / Ties=Efron Rl;

Test Tratam1=Tratam2; Test Tratam1=Tratam3; Test Tratam1=Tratam4;

Test Tratam2=Tratam3; Test Tratam2=Tratam4; Test Tratam3=Tratam4;



```
Tratdias = Tratam*(Días> 15);
```

```
Run;
```

### **Análisis paramétrico**

En este caso, para la distribución de supervivencia se asumió el modelo de tiempo de falla acelerado (AFT, por sus siglas en inglés), que postula una relación directa entre predictores y tiempo de supervivencia (George et al., 2014). Para determinar la distribución más apropiada y mejorar el potencial del modelo AFT, se ajustaron y compararon con PROC LIFEREG, cuatro distribuciones paramétricas (Weibull, Lognormal, Log-Logística y Exponencial). Para establecer el mejor ajuste entre las distribuciones, se utilizó el Criterio de Máxima Verosimilitud (ML).

```
Proc Sort Data= Canafis;
```

```
By Descending Tratam;
```

```
Run;
```

```
Proc Lifereg Data= Canafis;
```

```
Class Tratam;
```

```
Model Días* Status (0)= Tratam Tratdias /Dist= Weibull;
```

```
Probplot;
```

```
Run;
```

### **Resultados**

#### **a. Análisis no paramétrico**

Los datos de germinación presentan un especial problema estadístico: no se conoce si las semillas remanentes (no germinadas) son capaces de germinar, están latentes, germinaron y luego murieron sin emerger el talluelo de la plántula o siempre fueron inviables. En el presente caso, las muestras presentaron baja a mediana germinación, sin saber en cual condición de las anteriores se encontraban. La solución alternativa es examinar las semillas no germinadas mediante una prueba de viabilidad con tetrazo-

lio, y luego clasificar correctamente cada semilla. En el presente caso, al proceso de germinación se le hizo seguimiento durante 30 días continuos y las semillas que no germinaron al final de este período se trataron como no censuradas (censura a la derecha).

### **a.1. Comparación entre tratamientos de inmersión**

Uno de los principales propósitos de la germinación para describir el proceso biológico, consiste en estimar algunos indicadores del lote de semillas, que permiten comparar tratamientos o cuantificar los efectos de factores experimentales sobre el proceso de germinación/emergencia. Entre ellos, la media y la mediana del tiempo de germinación (tiempo medio de germinación ( $T_{50}$ ): número de días transcurridos después de la incubación, para alcanzar el 50% del porcentaje de germinación final). En algunos casos también se requieren otros tiempos para caracterizar mejor la primera y última fase del proceso, los tiempos para el 25 % o 75 % de germinación ( $T_{25} - T_{75}$ ), (Onofri *et al.*, 2010).

Son indicadores utilizados para comparar la energía de germinación al registrar la “tasa de germinación”, es decir, el número de días que se necesitan para conseguir el 50 por ciento de la capacidad de germinación. Cuanto más breve sea ese período, tanto mayor será la energía de germinación.

A continuación, se presentan los cuartiles (indicadores) para caracterizar los tratamientos de inmersión. Es un lote de semillas con germinación mediana a baja, dependiendo del tratamiento. En publicación anterior, en la que se analizó la información mediante índices de germinación y análisis de varianza, los tratamientos con Giber Grop® a 400 mg/L y el tratamiento de inmersión solo en agua (testigo, 0 mg/L), lograron los mayores porcentajes de germinación promedio (45,0 – 43,8 %). Al aumentar la concentración por arriba de 400 ppm y con menor tiempo de inmersión, disminuye la germinación (Vale-Montilla, 2022).

### **a.2. Cuartiles estimados para cada tratamiento de inmersión de semillas de *Cassia moschata* Kunth., en días de germinación**

En las condiciones del ensayo, los únicos tratamientos que alcanzaron el 50 % de germinación, fueron los tratamientos de inmersión en agua y Giber Grop a 400 ppm, con tiempo medio de germinación de 22 y 24 días, respectivamente. Por este motivo, los valores estimados para los tiempos  $T_{50}$  y  $T_{75}$ , no se reflejan para los demás tratamientos. Los valores de la media del tiempo de germinación son muy

similares, con inmersión en agua con la menor media (21,1 días), lo cual indica mayor velocidad de germinación, como expresión del número de días promedio empleados en la germinación (21,1 días) versus los demás tratamientos, con excepción de Giber Grop 400 y 600 ppm con un valor de media muy similar (21,8 días). El tratamiento Giber Grop 600 ppm (12 horas), con mayor media del tiempo de germinación ( $\geq 22,9$  días). También se aprecia que para la primera fase de germinación ( $T_{25}$ ), el tratamiento de inmersión en agua, estableció el menor tiempo para alcanzar el 25 % de germinación, con un tiempo estimado de 14 días. Los valores de la media del tiempo sugieren una concentración entre 400 - 600 ppm, pero los valores estimados para  $T_{25}$  son menores para 600 ppm que para 400 ppm (16,5 < 18,0), solo que al disminuir el tiempo de inmersión en Giber Grop 600 ppm, aumenta la media del tiempo de germinación de 21,8 días a 22,9 días. Este resultado genera una hipótesis: parece indicar que se requiere mayor tiempo de inmersión, para disminuir el tiempo de germinación. Sería recomendable una futura evaluación manteniendo los tratamientos de inmersión en agua, inmersión en Giber Grop 600 ppm por 24 horas y comparar con tiempos de inmersión por 48 horas.

**The LIFETEST Procedure**  
Summary Statistics for Time Variable DÍAS  
Quartile Estimates

**Tratamiento 1:**

Percent	Point Estimate	95% Confidence Interval	
		[Lower	Upper)
75	.	.	.
50	.	22.0000	.
25	14.0000	12.0000	21.0000
	Mean	Standard Error	
	21.1266	0.	

**Tratamiento 2:**

Percent	Point Estimate	95% Confidence Interval	
		[Lower	Upper)
75	.	.	.
50	.	.	.
25	24.0000	17.0000	.
	Mean	Standard Error	
	22.1235	0.6322	

**Tratamiento 3:**

Percent	Point Estimate	95% Confidence Interval [Lower Upper)	
75	.	.	.
50	.	24.0000	.
25	18.0000	14.0000	23.0000
	Mean	Standard Error	
	21.8000	0.7043	

**Tratamiento 4:**

Percent	Point Estimate	95% Confidence Interval [Lower Upper)	
75	.	.	.
50	.	.	.
25	16.5000	11.0000	.
	Mean	Standard Error	
	21.7625	0.7974	

**Tratamiento 5:**

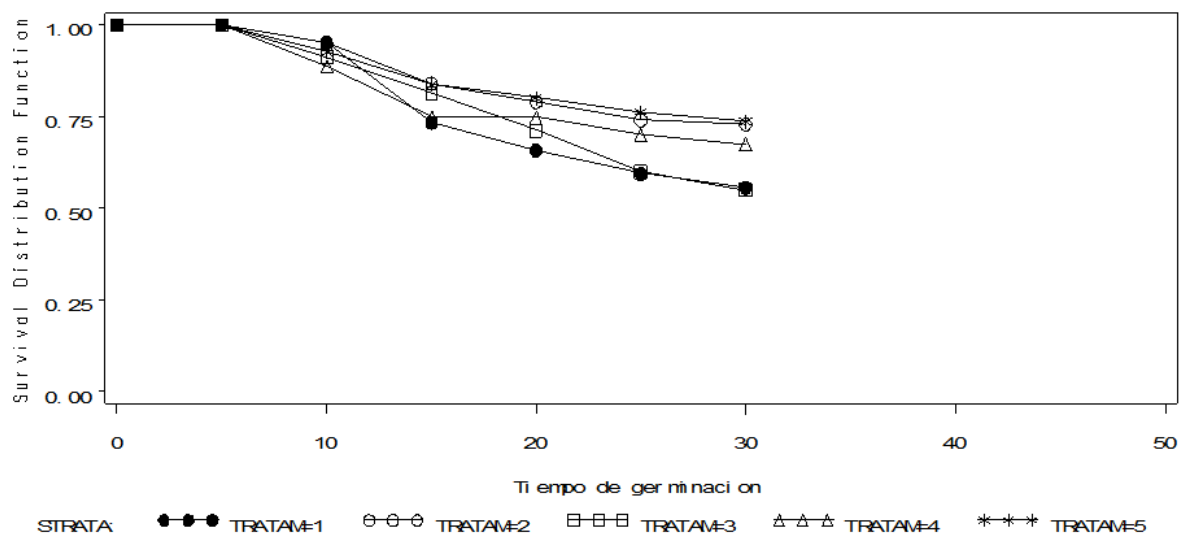
Percent	Point Estimate	95% Confidence Interval [Lower Upper)	
75	.	.	.
50	.	.	.
25	26.0000	15.0000	.
	Mean	Standard Error	
	22.9000	0.6984	

En la tabla Summary of the Number of Censored and Uncensored Values, que sigue a continuación, se observa que 65,00% de la cohorte de semillas fue censurada. Esto es equivalente a  $100\% - 65,00\% = 35,00\%$  de las semillas presentaron el evento de germinación (Failed). La germinación más alta corresponde al tratamiento 3 (inmersión en Giber Grop a 400 ppm 24 horas) con 36% de germinación y a continuación el tratamiento 1 (inmersión en agua durante 24 horas) con 35%.

Summary of the Number of Censored and Uncensored Values

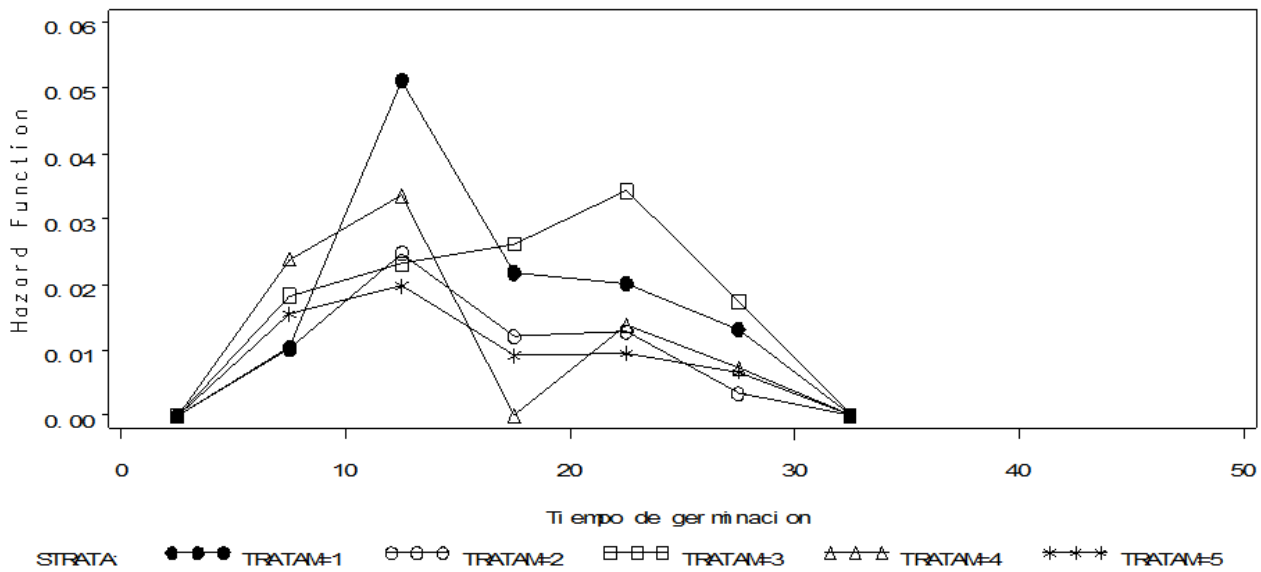
Stratum	TRATAM	Total	Failed	Censored	Percent Censored
1	1	79	35	44	55.70
2	2	81	22	59	72.84
3	3	80	36	44	55.00
4	4	80	26	54	67.50
5	5	80	21	59	73.75
-----					
Total		400	140	260	65.00

Las curvas de supervivencia de la germinación o proporción de semillas sin germinar, bajo los cinco tratamientos de inmersión se muestran en la Figura 21. Al final del período de evaluación de la germinación (30 días), la proporción de las semillas sin germinar para los tratamientos fue de 0,56 – 0,73 – 0,55 – 0,68 y 0,74 para los tratamientos de inmersión en agua durante 24 horas, en Giber Grop a 200 ppm 24 horas, en Giber Grop a 400 ppm 24 horas, en Giber Grop a 600 ppm 24 horas y Giber Grop a 600 ppm 12 horas, respectivamente. Las curvas de supervivencia bajo los cinco tratamientos no fueron diferentes (Wilcoxon:  $\chi^2 = 8,33$ ; gl = 4, p = 0,08).



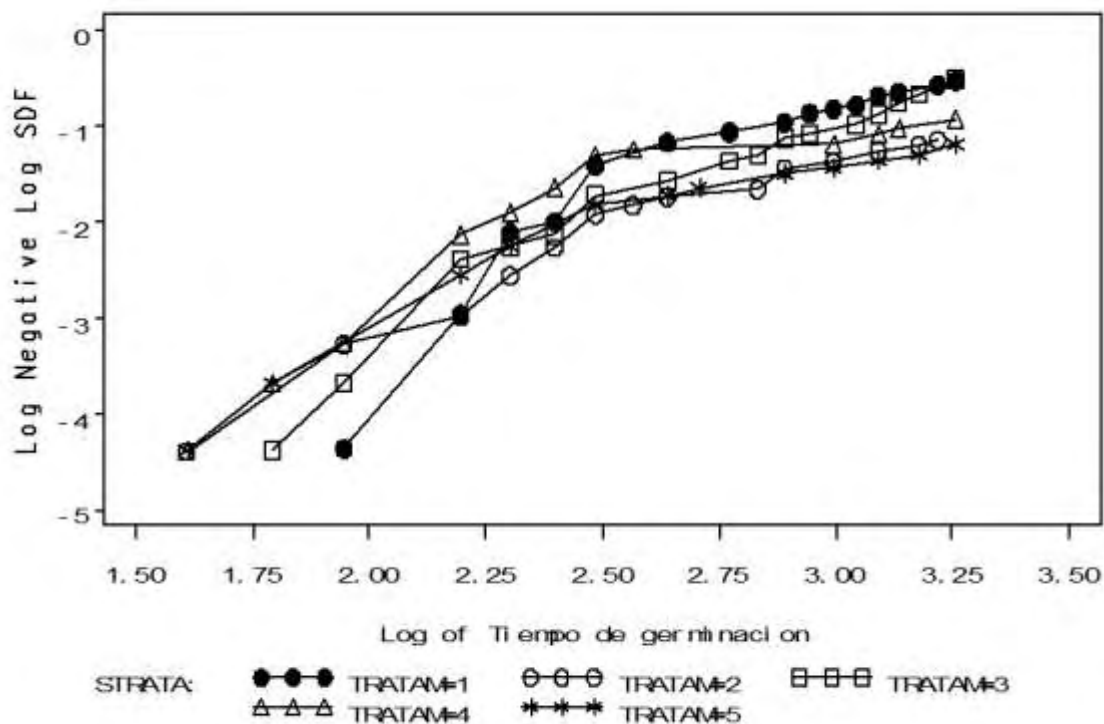
**Figura 21. Cambios en la supervivencia de la germinación de *C. moschata* en los tratamientos de inmersión de semilla durante 30 días de evaluación.** Fuente. Elaboración propia.

En la Figura 22 se observa que los tratamientos inmersión en agua durante 24 horas, Giber Grop a 200 ppm 24 horas, Giber Grop a 600 ppm 24 horas y Giber Grop a 600 ppm 12 horas, presentaron las más altas razones de riesgo en el intervalo 10 – 15 días (0,051 – 0,0248 – 0,0335 y 0,0198 germinaciones/día, respectivamente). En el caso del tratamiento Giber Grop a 400 ppm 24 horas, se observa un riesgo de germinación sostenido en el tiempo, con pico de riesgo en el intervalo 20 – 25 días (0,0343 germinaciones/día).



**Figura 22. Valores estimados de la función de riesgo para germinación de *C moschata* Kunth., en cinco tratamientos de inmersión.** Fuente. Elaboración propia.

Al comparar las respuestas de las curvas de germinación individual, por efecto de los tratamientos, se observan cruzamientos de las curvas de germinación durante el proceso de germinación entre todos los tratamientos. Por esta razón, se utiliza la prueba de Wilcoxon para la comparación de tratamientos, lo que indica que no existen diferencias entre tratamientos (Wilcoxon  $\chi^2 = 8,3349$ ;  $gl=4$ ;  $p=0,0801$ ; Figura 23).



**Figura 23.** Función de Log (-Log (S (t))) versus log de tiempo con suposición PH violada de las funciones de germinación de *C. moschata* Kunth., en cinco tratamientos de inmersión. Fuente. Elaboración propia.

### a.3. Resultado de las pruebas de comparación para las funciones de sobrevivencia.

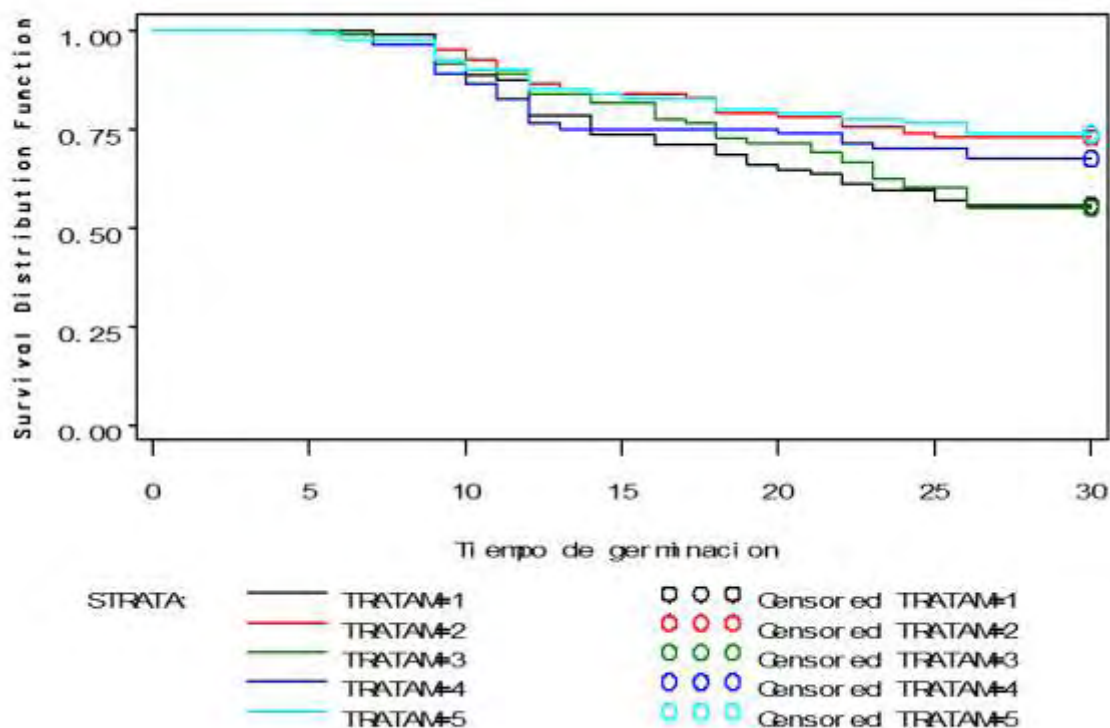
A continuación, se indican los resultados de las pruebas que comparan las respuestas de las curvas de germinación. Las pruebas de Wilcoxon indicaron que no existen diferencias significativas entre tratamientos ( $p = 0,0801$ ). Las pruebas indican que los cinco tratamientos son homogéneos, pero no indican cuales difieren entre ellos.



Test of Equality over Strata			
Test	Chi-Square	DF	Pr > Chi-Square
Log-Rank	10.0398	4	0.0398
Wilcoxon	8.3349	4	0.0801
-2Log (LR)	10.1403	4	0.0381

Las funciones de supervivencia estimada en la Figura 24, sugieren que el patrón temporal de germinación es diferente para la totalidad de tratamientos. Presentan variación en los eventos de germinación, con 15 – 14 – 16 eventos para T1, T2 y T3 y 12 eventos para los tratamientos T4 y T5. Los cinco tratamientos no presentaron eventos de germinación durante los primeros intervalos (0 – 10 días). A partir de estos primeros intervalos, los tratamientos presentaron probabilidades semejantes. Por ejemplo, para el intervalo 15 – 20 días, el tratamiento de inmersión en agua (T<sub>1</sub>) y el tratamiento T<sub>4</sub> (inmersión en Giber Grop a 600 ppm, durante 24 horas), la proporción estimada de semillas que permanecieron sin germinar fue de 0,73 – 0,75, respectivamente (73 - 75 %). Para los tratamientos T<sub>2</sub>, T<sub>3</sub> y T<sub>5</sub> la proporción durante el mismo período fue de 0,84 – 0,81 – 0,84 (84 – 81 – 84 %), respectivamente. Significa que, en todos los tratamientos las proporciones estimadas de semillas que permanecieron sin germinar, fueron altas. Representa que una estimación de la función de distribución de probabilidad acumulada de germinación, expresada como 1 – St para el intervalo 15 – 20 días fue 1 – 0,73 (27%) para inmersión en agua y de 16 – 19 – 25 y 16% para T<sub>2</sub> – T<sub>3</sub> – T<sub>4</sub> y T<sub>5</sub>. Lo que sugiere que, en este intervalo, inmersión en agua e inmersión en Giber Grop a 600 ppm, durante 24 horas, mantuvieron mayores probabilidades de germinación que los demás tratamientos, aunque bajas.

Para el intervalo final (30 – 35 días), los valores resultantes fueron 0,56 (0,44) – 0,73 (0,27) – 0,55 (0,45) – 0,68 (0,32) y de 0,74 (0,26), respectivamente, con las más altas probabilidades de presentar el evento de germinación en los tratamientos inmersión en Giber Grop® a 400 ppm durante 24 horas e inmersión en agua (0,45 vs 0,44).



**Figura 24.** Estimados de Kaplan–Meyer de las funciones de germinación de *C. moschata* Kunth., en cinco tratamientos de inmersión. *Fuente.* Elaboración propia.

## b. Análisis semiparamétrico.

### b.1. Análisis de multicolinealidad.

La prueba de colinealidad diagnosticó que la covariable tratamientos, alcanzó valores menores de 1,00 ( $VIF < 5$ ), lo que indica que no existe evidencia de colinealidad en la variable.

### b.2. Análisis de supuestos del modelo.

Las curvas de  $\log(-\log(S(t)))$  versus  $\log(t)$  para tratamientos, presentaron líneas razonablemente paralelas para algunos de ellos (Figura 23). Además, los resultados de los residuos de Schoenfeld, para satisfacer la suposición, deben ser independientes del tiempo, es decir, no estar correlacionados con el

tiempo o con alguna función del tiempo (Allison, 2010). Los resultados mostrados a continuación, indican que no existen problemas de violación de supuestos de riesgo proporcional con la covariable tratamientos para las funciones de tiempo (Diasrank, Días1 y Días2), con valores de  $p > 0,05$ , tanto por el procedimiento ASSESS ( $p=0,9460$ ), como con los residuales de Schoenfeld para las tres variables de tiempo ( $p=0,1530$ ,  $p=0,140$  y  $p=0,4315$ ). Por lo tanto, se ajustan al modelo, sin problemas de violación de supuestos. El número de observaciones es de 140, ya que los residuales de Schoenfeld no están definidos para casos censurados (260 valores censurados, de un total de 400 valores), donde el 65,0 % resultó censurado (semilla sin germinar).

**The PHREG Procedure**  
Model Information

```

Data Set                WORK.CANAFIS
Dependent Variable      DIAS          Tiempo de germinacion
Censoring Variable      STATUS      Status de germinacion
TRATAM 1=Inmersión en agua durante 24 horas
24 horas TRATAM 2=Inmersión en Giber Grop 200 ppm
24 horas TRATAM 3=Inmersión en Giber Grop 400 ppm
24 horas TRATAM 4=Inmersión en Giber Grop 600 ppm
24 horas TRATAM 5=Inmersión en Giber Grop a
Censoring Value(s)     0
Ties Handling           EFRON

```

```

Number of Observations Read      400
Number of Observations Used      400

```

Summary of the Number of Event and Censored Values

Total	Event	Censored	Percent Censored
400	140	260	65.00

Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Without Covariates	With Covariates
-2 LOG L	1622.048	1619.254
AIC	1622.048	1621.254
SBC	1622.048	1624.196

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	2.7934	1	0.0947
Score	2.7876	1	0.0950
Wald	2.7755	1	0.0957

The PHREG Procedure

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
TRATAM	1	-0.10041	0.06027	2.7755	0.0957	0.904

Supremum Test for Proportionals Hazards Assumption

Variable	Maximum Absolute Value	Replications	Seed	Pr > MaxAbsVal
TRATAM	0.9097	1000	175444001	0.9460

**b.2.1. Correlaciones entre las variables de tiempo y los residuales de Schoenfeld.**

The CORR Procedure

4 Variables: DIASRANK DÍAS1 DÍAS2 schTRATAM

Variable	N	Mean	Std Dev	Simple Statistics		
				Sum	Minimum	Maximum
DIASRANK	400	200.500	98.44157	80200	2.00000	270.50000
DÍAS1	400	3.125	0.45064	1250	1.60944	3.40120
DÍAS2	400	674.562	328.74505	269825	25.00000	900.00000
schTRATAM	140	0.00189	1.38890	0.26517	-1.90775	2.22214

Simple Statistics

Variable	Label
DIASRANK	Rank for Variable DÍAS
DÍAS1	
DÍAS2	

		schTRATAM Schoenfeld Residual for TRATAM		
		Pearson Correlation Coefficients		
		Prob >  r  under H0: Rho=0		
		Number of Observations		
	DIASRANK	DÍAS1	DÍAS2	sch TRATAM
DIASRANK	1.00000	0.94208	0.98855	-0.12140
Rank for Variable DIAS	<.0001	<.0001	<.0001	0.1530
	400	400	400	140
DÍAS1	0.94208	1.00000	0.96313	-0.12535
	<.0001	<.0001	<.0001	0.1400
	400	400	400	140
DÍAS2	0.98855	0.96313	1.00000	-0.06701
	<.0001	<.0001	<.0001	0.4315
	400	400	400	140
schTRATAM	-0.12140	-0.12535	-0.06701	1.00000
Schoenfeld TRATAM	0.1530	0.1400	0.4315	
	140	140	140	140

### b.3. Análisis del modelo con variables tiempo dependientes

En este modelo, la variable de respuesta es el "riesgo". El modelo asume que la tasa de riesgo subyacente, en lugar del tiempo de supervivencia, es una función de las variables independientes o covariables. Significa que un cambio de una unidad en una variable explicativa, causa cambios proporcionales en el riesgo. Para un análisis detallado, la germinación se analizó por etapas. En una primera etapa, se analizaron los primeros 15 días del proceso germinativo (entre 0 y 15 días), luego el período entre 16 y 30 días. Para cada una de estas evaluaciones, se consideró al tratamiento T<sub>5</sub> como grupo de referencia (Giber Grop® a 600 ppm, 12 horas de inmersión) y todos los tratamientos se compararon con él (Romano y Stevanato, 2020; McNair et al., 2012).

#### b.3.1. Primera etapa del proceso de germinación (0 – 15 días)

Los coeficientes exponenciados,  $\exp(\beta_i)$  representan las razones de riesgo y dan el tamaño del efecto de cada covariable. Pueden considerarse como el cambio predictivo en el riesgo para un aumento de una unidad en el predictor. El signo positivo del coeficiente de regresión indica que para una

covariable las probabilidades de germinación aumentan durante los primeros 15 días. También significa que la covariable aumenta el riesgo, lo que reduce el tiempo de falla o germinación. El efecto contrario se produce con los coeficientes negativos. En este caso, los tratamientos disminuyen las probabilidades de germinación y, en consecuencia, aumentan el tiempo de germinación. En este caso, para los primeros 15 días del proceso germinativo, todos los tratamientos ( $T_1 - T_2 - T_3 - y T_4$ ), en comparación con el tratamiento de referencia ( $T_5 =$  Inmersión en Giber Grop® a 600 ppm durante 12 horas) aumentaron las probabilidades de germinación y, en consecuencia, disminuyeron el tiempo de falla o de germinación, de manera significativa para los tratamientos  $T_1$  y  $T_3$ , y no significativamente para  $T_2$  y  $T_4$ . Para evaluar los efectos de los tratamientos y sus cambios en el tiempo (interacciones), al definir una variable tiempo dependiente (TRAT\*DÍA), la interacción también tuvo un impacto positivo en la germinación, aumentando las probabilidades de germinación en el tiempo ( $\beta_i = 0,11946$ ), aunque no significativamente ( $p = 0,3371$ ). Ya que la interacción no resultó significativa, se concluye que la suposición de riesgo proporcional no se viola en este período (Resultado siguiente).

**Análisis de covariables con interacciones con el tiempo por el método de máxima verosimilitud (ML) para la primera etapa del proceso de germinación (0 – 15 días).**

```

The PHREG Procedure
Model Information

Data Set          WORK.CANAFIS
Dependent Variable  DÍAS           Tiempo de germinacion
Censoring Variable STATUS       Status de germinacion TRATAM 1=Inmersión en agua duran-
te 24 horas
                24 horas TRATAM 2=Inmersión en Giber Grop a 200 ppm
                24 horas TRATAM 3=Inmersión en Giber Grop a 400 ppm
                24 horas TRATAM 4=Inmersión en Giber Grop a 600 ppm
                12 horas TRATAM 5=Inmersión en Giber Grop a 600 ppm
Censoring Value(s)  0
Ties Handling       EFRON

                Number of Observations Read      400
                Number of Observations Used      400

```

## Class Level Information

Class	Value	Design Variables			
TRATAM	1	1	0	0	0
	2	0	1	0	0
	3	0	0	1	0
	4	0	0	0	1
	5	0	0	0	0

## Summary of the Number of Event and Censored Values

Total	Event	Censored	Percent Censored
400	140	260	65.00

## Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

## Model Fit Statistics

Criterion	Without Covariates	With Covariates
-2 LOG L	1622.048	1611.171
AIC	1622.048	1621.171
SBC	1622.048	1635.879

## The TPHREG Procedure

## Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	10.8768	5	0.0539
Score	10.9338	5	0.0527
Wald	10.6491	5	0.0588

## Type 3 Tests

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRATAM	4	10.1093	0.0386
TRATDIAS	1	0.9214	0.3371

## Analysis of Maximum Likelihood Estimates

Parameter	Parameter	Standard				Hazard	95% H Ratio	
	DF	Estimate	Error	Chi-Square	Pr >	ChiSq	Ratio	ConfidLimits
TRATAM 1	1	0.90522	0.40505	4.9945	0.0254	2.472	1.118	5.469
TRATAM 2	1	0.24452	0.37976	0.4146	0.5196	1.277	0.607	2.688
TRATAM 3	1	0.74737	0.31609	5.5905	0.0181	2.111	1.136	3.923
TRATAM 4	1	0.35521	0.30449	1.3609	0.2434	1.426	0.785	2.591
TRATDIAS	1	0.11946	0.12446	0.9214	0.3371	1.127	0.883	1.438

El valor de la razón de riesgo para los diferentes tratamientos, resultó de 2,476; 1,277; 2,111 y 1,426 y representan valores de  $HR > 1$ , que además presentaron signos positivos en sus coeficientes, significa que, en este período, tienen más alta probabilidad de germinación que el tratamiento de referencia. Estos valores significan que los tratamientos  $T_1$ ,  $T_2$ ,  $T_3$  y  $T_4$  aumentarían la función de riesgo a través del periodo de germinación por 147,6 %, 27,7 %, 111,1 % y 42,6 % con relación al tratamiento de referencia (Analysis of Maximum Likelihood Estimates). También puede significar para  $T_1$  (inmersión en agua durante 24 horas) con una HR de 2,476, con un intervalo de confianza del 95% comprendido entre 1,12 a 5,47 que al no incluir el intervalo de confianza el valor nulo (1), se concluye que el riesgo de germinación no fue similar en los dos grupos ( $T_1 \neq T_5$ ). Igual situación se presenta con  $T_3$  (inmersión en Giber Grop a 400 ppm durante 24 horas), con una HR de 2,11 y un intervalo de confianza del 95% de 1,14 a 3,92, de manera que el riesgo de germinación tampoco es similar entre los tratamientos a comparar ( $T_3 \neq T_5$ ), por lo que existen diferencias significativas del riesgo temporal de germinación entre las semillas de los tratamientos bajo comparación. La interpretación correcta sería que las semillas de los tratamientos  $T_1$  y  $T_3$  que no hubiesen germinado tenían casi el doble de probabilidad de presentarla en el siguiente intervalo de tiempo que las semillas del grupo control o tratamiento de referencia. Un escenario diferente se presenta con  $T_2$  (inmersión en Giber Grop a 200 ppm durante 24 horas) y  $T_4$  (inmersión en Giber Grop a 600 ppm durante 24 horas), tratamientos que incluyen el valor nulo (1) entre sus intervalos de confianza en las comparaciones con el tratamiento de referencia ( $T_5$ ). De esta manera, el riesgo de germinación resultó similar en los dos grupos ( $T_2 = T_5$ ;  $T_4 = T_5$ ), por lo tanto, no existen diferencias significativas del riesgo temporal de germinación entre las semillas de los tratamientos bajo comparación. La interpretación correcta sería que las semillas de los tratamientos  $T_2$  y  $T_4$  que no hubie-



sen germinado tenían igual probabilidad de presentarla en el siguiente intervalo de tiempo que las semillas del grupo control o tratamiento de referencia (T<sub>5</sub>).

El valor inverso  $1/\exp(\beta_i)$  en los tratamientos a comparar con el tratamiento de referencia, iguales a 0,404 – 0,783 – 0,474 y 0,70, indica que las probabilidades de experimentar germinación se incrementan por estos factores para T<sub>1</sub>, T<sub>2</sub>, T<sub>3</sub> y T<sub>4</sub>, comparados con el tratamiento de referencia.

Además de la comparación de los demás tratamientos con Giber Grop® a 600 ppm (12 horas de inmersión), se realizaron otras pruebas de hipótesis para comparar entre tratamientos, originando una prueba muy útil, donde todos los coeficientes asociados con esta covariable tienen valor de 0 (Analysis of Maximum Likelihood Estimates). Para la primera etapa, con un valor de chi cuadrado de 10,6491 y 5 grados de libertad, la prueba resultó estadísticamente significativa ( $p=0,05$ ). Entre los tratamientos comparados solo se presentaron diferencias significativas en la comparación entre inmersión en agua e inmersión en Giber Grop 200 ppm. Entre los tratamientos Giber Grop 200 ppm y 400 ppm se presentaron diferencias significativas al 0,1 % ( $p < 0,10$ ).

### **Resultados de las pruebas de hipótesis entre tratamientos, para la primera etapa del proceso de germinación (0 – 15 días).**

#### **Analysis of Maximum Likelihood Estimates**

##### **Linear Hypotheses Testing Results**

Label	Wald		Pr > ChiSq
	Chi-Square	DF	
Test 1	5.5252	1	0.0187
Test 2	0.3281	1	0.5668
Test 3	2.6706	1	0.1022
Test 4	3.2362	1	0.0720
Test 5	0.1168	1	0.7325
Test 6	2.1386	1	0.1436

### b.3.2. Segunda etapa del proceso de germinación (16 – 30 días)

En esta segunda etapa del proceso de germinación, la interacción de los tratamientos con el tiempo, a diferencia del período anterior, tiene un impacto negativo en la germinación. Es decir, la interacción de los tratamientos con el tiempo, disminuyeron las probabilidades de germinación ( $\beta_i = -0,15390$ ) en el tiempo, aunque de manera no significativa ( $p=0,2189$ ), lo cual también sugiere que no existió violación de supuestos de riesgo proporcional. De manera individual, el signo negativo del coeficiente de regresión para T<sub>2</sub> (inmersión en Giber Grop a 200 ppm durante 24 horas) y positivo para los demás tratamientos, indica que las probabilidades de germinación para T<sub>2</sub> disminuyeron durante este período y tienen más baja probabilidad de germinación que el tratamiento de referencia (T<sub>5</sub>, Giber Grop® a 600 ppm, 12 horas) a diferencia de los tratamientos T<sub>1</sub>, T<sub>3</sub> y T<sub>4</sub>. El valor de razón de riesgo para T<sub>2</sub> de HR=0,86 (HR<1) revela que experimentó menor riesgo de germinación que el tratamiento de referencia, es decir, la probabilidad de experimentar germinación se incrementa por este factor para el tratamiento de referencia. Por lo tanto, el riesgo de germinar de las semillas es alrededor de 14 % de la razón de riesgo de las semillas del tratamiento de referencia. Este resultado revela que las semillas del tratamiento T<sub>2</sub> tendrán tendencia a germinar posteriormente que las semillas del tratamiento de referencia.

El valor de la razón de riesgo para los demás tratamientos (T<sub>1</sub> – T<sub>3</sub> y T<sub>4</sub>, resultó de 1,457; 1,625 y 1,253 (HR>1). Al interpretar estos valores, significan que los riesgos de germinar de las semillas aumentarán la función de riesgo a través del período de germinación por 45,7 %, 62,5 % y 25,3 % con relación al tratamiento de referencia y revela que experimentaron mayor riesgo de germinación que el tratamiento de referencia, es decir, la probabilidad de experimentar germinación disminuye por estos factores con el tratamiento de referencia. Otra interpretación para el valor de HR de T<sub>1</sub> (inmersión en agua durante 24 horas), con una HR de 1,457, en un intervalo de confianza del 95%, comprendido dentro de 0,75 a 2,84, tiene por resultado que al incluir el intervalo de confianza el valor nulo (1) en las comparaciones con el tratamiento de referencia (T<sub>5</sub>), se concluye que el riesgo de germinación fue similar en los dos grupos (T<sub>1</sub>=T<sub>5</sub>). Igual situación se presenta con T<sub>2</sub> (inmersión en Giber Grop a 200 ppm durante 24 horas), con una HR de 0,86, comprendido entre 0,44 a 1,66, por tanto, T<sub>2</sub>=T<sub>5</sub>; para T<sub>3</sub> (inmersión en Giber Grop a 400 ppm durante 24 horas), con HR de 1,625 (0,92 a 2,86), indica que

$T_3=T_5$  y para  $T_4$  (inmersión en Giber Grop a 600 ppm durante 24 horas) con HR de 1,253 (0,70 – 2,24), también sugiere que  $T_4=T_5$ . Al incluir dentro del intervalo de confianza el valor nulo (1), se concluye que el riesgo de germinación para este periodo fue similar en todos los tratamientos ( $T_1=T_2=T_3=T_4=T_5$ ). Por lo tanto, no existen diferencias significativas del riesgo temporal de germinación entre las semillas de los tratamientos bajo comparación. La interpretación correcta sería que las semillas de los tratamientos  $T_1$ ,  $T_2$ ,  $T_3$  y  $T_4$  que no hubiesen germinado tenían igual probabilidad de presentarla en el siguiente intervalo de tiempo que las semillas del grupo control o tratamiento de referencia ( $T_5$ ).

El valor inverso  $1/\exp(\beta_i)$  en los tratamientos a comparar con el tratamiento de referencia, iguales a 0,686 – 0,615 y 0,798, indica que las probabilidades de experimentar germinación se incrementan por estos factores para  $T_1$ ,  $T_3$  y  $T_4$ , comparados con el tratamiento de referencia. En el caso del tratamiento  $T_2$ , sucede lo contrario, el valor inverso  $1/\exp(\beta_i)$  de 1,16 indica que las probabilidades de experimentar germinación se incrementan por un factor de 1,16 con el tratamiento de referencia, comparado con  $T_2$ .

### **Análisis de covariables con interacciones con el tiempo por el método de máxima verosimilitud (ML), para la segunda etapa del proceso de germinación (16 – 30 días)**

The PHREG Procedure

Model Information

Data Set	WORK.CANAFIS	
Dependent Variable	DÍAS	Tiempo de germinacion
Censoring Variable	STATUS	Status de germinacion TRATAM 1=Inmersión en agua du-
rante 24 horas		
	24 horas TRATAM 2=Inmersión en Giber Grop a 200 ppm	
	24 horas TRATAM 3=Inmersión en Giber Grop a 400 ppm	
	24 horas TRATAM 4=Inmersión en Giber Grop a 600 ppm	
	12 horas TRATAM 5=Inmersión en Giber Grop a 600 ppm	
	Censoring Value(s)	0
	Ties Handling	EFRON
	Number of Observations Read	400
	Number of Observations Used	400

## Class Level Information

Class	Value	Design Variables			
TRATAM	1	1	0	0	0
	2	0	1	0	0
	3	0	0	1	0
	4	0	0	0	1
	5	0	0	0	0

## Summary of the Number of Event and Censored Values

Total	Event	Censored	Percent Censored
400	140	260	65.00

## Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

## Model Fit Statistics

Criterion	Without Covariates	With Covariates
-2 LOG L	1622.048	1610.573
AIC	1622.048	1620.573
SBC	1622.048	1635.281

## The PHREG Procedure

## Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	11.4748	5	0.0427
Score	11.5163	5	0.0421
Wald	11.2049	5	0.0475

## Type 3 Tests

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRATAM	4	7.2355	0.1240
TRATDIAS	1	1.5119	0.2189

### Analysis of Maximum Likelihood Estimates

Parameter	Parameter DF	Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% H Ratio ConfidLimits
TRATAM 1	1	0.37611	0.34003	1.2235	0.2687	1.457	0.748 2.836
TRATAM 2	1	-0.15087	0.33681	0.2007	0.6542	0.860	0.444 1.664
TRATAM 3	1	0.48566	0.28797	2.8442	0.0917	1.625	0.924 2.858
TRATAM 4	1	0.22538	0.29606	0.5796	0.4465	1.253	0.701 2.238
TRATDIAS	1	-0.15390	0.12517	1.5119	0.2189	0.857	0.671 1.096

Para la segunda etapa, la comparación entre tratamientos, sin incluir sus comparaciones con el tratamiento de referencia (Giber Grop 600 ppm), con un valor de chi cuadrado de 11,2049 y 5 grados de libertad, la prueba resulto estadísticamente significativa ( $p=0,0475$ ). Entre los tratamientos comparados se presentaron diferencias significativas al 0,1 % ( $p = 0,0579$ ) entre inmersión en agua e inmersión en Giber Grop 200 ppm; diferencias significativas ( $p = 0,0213$ ) entre Giber Grop 200 ppm y 400 ppm.

### Resultados de las pruebas de hipótesis entre tratamientos, para la segunda etapa del proceso de germinación (16 – 30 días)

#### Linear Hypotheses Testing Results

Label	Wald Chi-Square	DF	Pr > ChiSq
Test 1	3.5976	1	0.0579
Test 2	0.1743	1	0.6763
Test 3	0.2460	1	0.6199
Test 4	5.3027	1	0.0213
Test 5	1.4968	1	0.2212
Test 6	0.9890	1	0.3200

El análisis de residuos del modelo definitivo para verificar los supuestos de proporcionalidad del modelo, las covariables y sus interacciones, satisfacen los supuestos de riesgo proporcional. Por lo tanto, el siguiente análisis paramétrico se realiza para complementar la información de la germinación y para mejorar la interpretación de los resultados generales.

### c. Análisis paramétrico

La aplicación del Criterio de Información Akaike (Akaike, 1974), permitió seleccionar la distribución Weibull por mejor ajuste. En la presente investigación, este patrón parece ajustarse a la tendencia ordinaria del proceso de germinación, lo que hace que este modelo sea adecuado a los resultados. En una distribución Weibull, cuando el parámetro de escala  $\sigma > 1$ , el riesgo disminuye con el tiempo. Cuando  $0,5 < \sigma < 1$ , el riesgo es creciente en una tasa decreciente; cuando  $0 < \sigma < 0,5$ , el riesgo es creciente a una tasa creciente y cuando  $\sigma = 0,5$ , la función de riesgo es una línea recta creciente con origen en 0 (Allison 2010).

La salida de resultados, muestra las estimaciones de máxima verosimilitud (MLE) de los parámetros calculados mediante la aplicación del modelo Weibull, expresado como regresión de tiempo de falla acelerada (AFT). La primera observación es que el valor estimado de  $\sigma$  (Escala) se encuentra entre  $0 < \sigma < 0,5$  ( $\sigma = 0,1588$ ), lo que implica que la función de riesgo es creciente a una tasa creciente (Allison, 2010).

El coeficiente de regresión para el tratamiento T5 (Giber Grop 600 ppm), es 0. Su razón es que se designó como nivel de referencia para comparar con los demás tratamientos. El análisis indica que cada coeficiente, de manera independiente, difiere significativamente del tratamiento Giber Grop 600 ppm, 12 horas de inmersión ( $p < 0,05$ ).

### Parámetros estimados de las distribuciones por el método de máxima verosimilitud (ML) para la distribución Weibull.

#### The LIFEREG Procedure Model Information

Data Set	WORK.CANAFIS
Dependent Variable	Log(DÍAS) Tiempo de germinación
Censoring Variable	STATUS Status de germinación
	TRATAM 1=Inmersión en agua durante 24 horas
	TRATAM 2=Inmersión en Giber Grop a 200 ppm 24 horas
	TRATAM 3=Inmersión en Giber Grop a 400 ppm 24 horas
	TRATAM 4=Inmersión en Giber Grop a 600 ppm 24 horas
	TRATAM 5=Inmersión en Giber Grop a 600 ppm 12 horas

```

Censoring Value(s)          0
Number of Observations      400
Noncensored Values          140
Right Censored Values       260
Left Censored Values        0
Interval Censored Values    0
Name of Distribution         Weibull
Log Likelihood              -56.5873031

```

```

Number of Observations Read  400
Number of Observations Used  400

```

#### Class Level Information

```

Name      Levels  Values
TRATAM    5      1 2 3 4 5

```

Algorithm converged.

#### Type III Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRATAM	4	1620.2210	<.0001
TRATDIAS	1	579.8638	<.0001

#### The LIFEREG Procedure

#### Analysis of Parameter Estimates

Parameter	DF	Estimate	Standard Error	95% Confidence Limits		Chi- Square	Pr > ChiSq
Intercept	1	0.8592	0.0651	0.7316	0.9867	174.21	<.0001
TRATAM	1	1.7935	0.0517	1.6922	1.8949	1203.52	<.0001
TRATAM	2	1.1471	0.0523	1.0447	1.2495	481.79	<.0001
TRATAM	3	0.5719	0.0443	0.4851	0.6587	166.85	<.0001
TRATAM	4	0.3006	0.0474	0.2077	0.3935	40.21	<.0001
TRATAM	5	0.0000	.	.	.	.	.
TRATDIAS	1	0.0270	0.0011	0.0248	0.0292	579.86	<.0001
Scale	1	0.1588	0.0120	0.1369	0.1843		
Weibull Shape	1	6.2964	0.4777	5.4264	7.3058		

En la estimación de los parámetros ATF para el modelo Weibull, la intersección está representado por  $\beta_0 = 0,8592$ , tratamiento 1 ( $\beta_1 = 1,7935$ ), tratamiento 2 ( $\beta_2 = 1,1471$ ), tratamiento 3 ( $\beta_3 = 0,5719$ ) y tratamiento 4 ( $\beta_4 = 0,3006$ ). Según Romano & Stevanato (2020), estos valores se utilizan para calcular el tiempo medio de germinación ( $T_{50}$ ). El  $T_{50}$  (mediana o tiempo medio de germinación) para referencia de cada parámetro, se calcula como  $\exp(\beta)$ . Para todo el lote (intercepto =  $\beta_0 = 0,8592$ ), fue de  $\exp^{(0,8592)} = 2,36$  días, mientras que para el efecto de tratamiento 1 es,  $\exp(\beta_0 + \beta_1) = \exp(0,8592 + 1,7935) = 14,2$ ; para tratamiento 2,  $\exp(\beta_0 + \beta_2) = 7,4$  días; para tratamiento 3,  $\exp(\beta_0 + \beta_3) = 4,2$ ; para tratamiento 4,  $\exp(\beta_0 + \beta_4) = 3,2$ . De esta manera, se calcula el tiempo medio de germinación para cada tratamiento, según su parámetro. Cuanto más negativo sea el coeficiente, más temprano será el tiempo medio de germinación (Romano y Stevanato 2020). Los resultados de la prueba chi cuadrada indican diferencias altamente significativas entre estas comparaciones, con el tratamiento de referencia. Por contraste, las comparaciones del análisis de las tablas de vida (LIFETES) sugieren que no existen diferencias entre los tratamientos. ¿Cuál análisis es el correcto? Para Fox (2001), el modelo de tiempo de falla acelerada (LIFEREG) tiene mayor poder estadístico y resulta razonable tener mayor confianza en sus resultados.

En términos de la razón de tiempo ( $\gamma$ ), para tratamiento 1  $\gamma = 14,2/2,36 = 6,02$ ; de manera similar, los demás tratamientos tienen razones de tiempo de 3,14; 1,78; 1,36 y un valor de 1 para el tratamiento Giber Grop 600 ppm. Estos valores, interpretados como razones estimadas de los tiempos de sobrevivencia esperados para dos grupos, según la fórmula ( $100\% (1 - \gamma)$ ), se puede observar que los tratamientos  $T_1$ ,  $T_2$ ,  $T_3$  y  $T_4$ , retardan o prolongan la germinación en 502,0 %, 214,0 %, 78,0 % y en 36,0 %, en comparación con inmersión en Giber Grop® a 600 ppm, 12 horas (tratamiento de referencia), respectivamente. Esto significa que el tratamiento que prolonga u origina una germinación más lenta, es el tratamiento  $T_1$  (inmersión en agua). El tratamiento 5 (inmersión en Giber Grop® a 600 ppm, 12 horas) se considera de referencia, por lo que aparece con un valor neutro (0,00). En general, en los modelos AFT, una razón de tiempo  $> 1$  para una covariable, implica que prolonga el tiempo de germinación, mientras que una razón de tiempo  $< 1$  implica que es más probable un evento de germinación más temprano (Romano y Stevanato, 2020). Según esta referencia, el tratamiento de inmersión en agua es el que más prolonga el tiempo de germinación y el tratamiento que menos prolonga la germinación es el de inmersión en Giber Grop a 600 ppm, durante 24 horas.



El hecho que cada coeficiente de regresión de manera independiente es significativo, indica que cada uno difiere del tratamiento de referencia. Estas comparaciones múltiples, son análogas a las del análisis LIFETEST, excepto que en este caso se utilizaron las pruebas de Wilcoxon y en el ATF, los coeficientes de regresión. El análisis LIFETEST mostró homogeneidad entre tratamientos, sin diferencias entre ellos. Los coeficientes de regresión llegan a una conclusión diferente: la comparación de los demás tratamientos con el tratamiento de referencia, es significativamente diferente del tratamiento Giber Grop 600 ppm y las comparaciones entre los tratamientos restantes permanecen invariables y no significativamente diferente.

Según Romano y Stevanato (2020), mientras que el modelo de Cox expresa el efecto multiplicativo de las variables explicativas sobre el riesgo (escala de riesgo), el modelo AFT expresa el mismo efecto sobre el tiempo de supervivencia (escala de tiempo). Esta característica permite una interpretación más fácil de los resultados porque los parámetros miden la fuerza y el efecto de la covariable correspondiente en la mediana del tiempo de supervivencia. Además, por la aplicación de tiempo de falla acelerado, el efecto de las covariables sobre la supervivencia se describe en términos absolutos (por ejemplo, número de días) en lugar de términos relativos como el índice de riesgo. Además, las razones de tiempo representan índices útiles que describen los efectos de las variables explicativas sobre la germinación y más interpretable que una relación de dos riesgos.

### **3.3.1.3. Análisis de supervivencia en germinación de *Swietenia macrophylla* King con aplicación de hormonas vegetales (Vale-Montilla, 2022)**

#### **Materiales y métodos**

El ensayo se realizó bajo condiciones a cielo abierto, con semillas colectadas de árboles ubicados en la Plaza Sucre (Trujillo), ubicada a 800 msnm. La siembra de semillas se hizo en gaveras para germinación, con tierra de textura franco arenosa.

#### **Hormonas sintéticas utilizadas.**

Giber Grop®.

Es un concentrado soluble en agua como regulador fisiológico de las plantas, para floración y engrosamiento de flores y frutos. Es un producto que contiene auxinas y giberelinas (Ácido alfa-naftalenacético 17,2% y Giberelina (AG3) 10%).

ProGibb®.

Es un regulador de crecimiento de las plantas, cuyo principio activo es el ácido giberélico 10 % p.p e ingredientes inertes 90 % p.p. Esta hormona produce incremento en el crecimiento de la planta, la interrupción de la latencia de la semilla, tubérculo y/o bulbos, la inducción de la floración, el retraso en la maduración de algunos frutos, mejora la calidad, etc.

### **Diseño experimental y tratamientos pregerminativos.**

El experimento fue realizado mediante un diseño de bloques al azar. Antes de la siembra, las semillas se sumergieron en diferentes concentraciones y tipos de hormonas vegetales y a temperatura ambiente, durante 24 horas. Se remojaron 50 semillas en cada tratamiento. Los tratamientos pregerminativos a evaluar fueron:

T<sub>1</sub>: inmersión de semillas en agua durante 24 horas;

T<sub>2</sub>: inmersión de semillas en ácido giberélico a 200 ppm (ProGibb®) durante 24 horas;

T<sub>3</sub>: inmersión de semillas en ácido giberélico a 400 ppm (ProGibb®) durante 24 horas;

T<sub>4</sub>: inmersión de semillas en ácido giberélico a 800 ppm (ProGibb®) durante 24 horas;

T<sub>5</sub>: inmersión de semillas en Giber Grop® a 600 ppm durante 24 horas;

El procedimiento a seguir fue el siguiente:

En un envase plástico limpio, se mezcló un litro de agua, con un surfactante, en dosis de 4 cc por litro de agua y la hormona del tratamiento correspondiente y en cada uno de ellos se colocaron las semillas, cubiertas completamente por la solución. Antes de sumergir las semillas en los tratamientos, se eliminó el “ala” de cada semilla. Luego de la inmersión en el tratamiento correspondiente, se extendieron las semillas en una superficie limpia y se dejaron secar durante 30 minutos. Este procedimiento se realizó en un lugar sombreado, a primeras horas de la mañana.

## Siembra de semillas

La siembra de semillas se hizo por siembra directa en gaveras para germinación, colocando una (01) semillas por abertura, sembradas a una profundidad de 1,0 cm. Las gaveras para germinación se mantuvieron bajo iluminación permanente a campo abierto (intemperie) durante la evaluación del experimento. Mientras las semillas estuvieron en germinación, los riegos con agua se continuaron 2 veces por día.

## Hipótesis y nivel de significación

Se discuten tres preguntas muy relacionadas de esta investigación: 1) ¿Cuáles son las distribuciones de los tiempos de germinación entre tratamientos de imbibición en diferentes concentraciones y tipo de hormona sintética; 2) ¿estas difieren entre sí?, y 3) relación entre covariables.

Los procedimientos de análisis de supervivencia permiten comparar pares de curvas de supervivencia y detectar si existen diferencias estadísticamente significativas entre ellas y probar la hipótesis nula ( $H_0$ ) de que las funciones de supervivencia son iguales entre curvas mientras que la hipótesis alternativa ( $H_1$ ) indica que las funciones de supervivencia no son iguales. El nivel de significancia para las diferentes pruebas se consideró como  $p < 0,05\%$ .

## Estructura y análisis de datos

Para efectos prácticos, se consideró que una semilla había germinado al emerger el talluelo de la plántula, llevando un registro diario de semillas germinadas por tratamiento durante 30 días después de la siembra (DDS). Con esta información, se realizaron las determinaciones del número de semillas germinadas diariamente y acumuladas por tratamiento. Los valores de media del tiempo, tiempo medio de germinación y otros índices del tiempo de germinación, se obtienen con la salida de resultados del procedimiento PROC LIFETEST.

Para la germinación diaria (número de semillas que germinan o no diariamente), se utilizó un análisis de supervivencia. Cada semilla individual seleccionada para la germinación por tratamiento se consideró una unidad experimental. Se adoptó un esquema de “observación continua” durante la duración del ensayo, asumiendo que los tiempos de germinación tuvieran un “dato exacto” para cada semilla

(McNair *et al.*, 2012). Además, no existió pérdida de semilla, pero el tiempo de germinación para la semilla no germinada se consideró como “censurada a la derecha”.

Las variables tratamientos de imbibición, tipo de hormona sintética y semilla, se consideraron como covariables categóricas. Para el análisis, en cada medición se asignaron valores de 1 o 0 para las semillas germinadas y no germinadas, respectivamente. Luego de cada conteo diario, cada semilla germinada fue removida del experimento. Se utilizaron códigos para discriminar entre los tratamientos ( $T_1=1$ ,  $T_2=2$ ,  $T_3=3$ ,  $T_4=4$  y  $T_5=5$ ). También se utilizaron códigos para discriminar entre el testigo y tipo de hormona (testigo = 0; ProGibb = 1; Giber Grop = 2).

Para el análisis semiparamétrico y paramétrico se incorporó la covariable semillas (SEM), por considerarla importante para el análisis. No fue incluida en el análisis LIFETEST, ya que es un análisis a base de curvas de germinación y ameritaba hacer demasiadas curvas de sobrevivencia.

Los datos de germinación se analizaron por análisis de supervivencia con el programa SAS (SAS, 2020; Fox, 1993), de la manera siguiente:

### **Análisis no paramétrico**

Los datos de germinación fueron descritos y modelados bajo dos probabilidades relacionadas: supervivencia y función de riesgo, ambas dependientes del tiempo. El procedimiento PROC LIFETEST, imprime por defecto, los estadísticos Log-rank y Wilcoxon, así como valores de chi-cuadrado para las comparaciones entre tratamientos. Se utiliza la prueba de Wilcoxon en caso de presentarse cruzamientos entre curvas de germinación y, en caso contrario, la prueba de Log-Rank (Bárceñas *et al.*, 2013; SAS, 2020; Fox, 2001; Fleming *et al.*, 1987; Lee, 1996; Buyske *et al.*, 2000; Lee y Wang, 2003).

Las diferencias de supervivencia entre covariables, se analizaron a partir de curvas de supervivencia construidas por el método Kaplan – Meier (Kaplan y Meier, 1958; Klein *et al.*, 2001; Gómez *et al.*, 2009); para ello, se tomó en cuenta el estatus de cada semilla (status 1 = germinación, status 0 = censura debido a semilla no germinada al final del experimento), así como el tiempo de germinación de la misma en días. Según Onofri *et al.* (2010), los tiempos de germinación se usan para calcular la probabilidad de germinación.

```
Proc Print Data=Caoba;
```

```
Run;
```

```
Proc Sort Data=CAOBA;
```

```
By Descending Tratam Descending Tipo Descending Sem;
```

```
Run;
```

```
Legend1 frame cframe=ligr cborder=black position=center value=(justify=center);
```

```
axis1 label=(angle=90 rotate=0 'Estimated CDF') minor=none;
```

```
axis2 minor=none;
```

```
symbol1 value=dot c=black h=0.15in i=none r=1;
```

```
symbol2 value=circle c=black h=0.15in i=none r=1;
```

```
symbol3 value=square c=black h=0.15in i=none r=1;
```

```
symbol4 value=triangle c=black h=0.15in i=none r=1;
```

```
symbol5 value=star c=black h=0.15in i=none r=1;
```

```
Proc Lifetest Data=Caoba Method=Lt Plots= (S, H);
```

```
Time Días*Status (0);
```

```
Strata Tratam;
```

```
Run;
```

```
Proc Lifetest Data=Caoba Method=Km Plots= (Ls, Lls);
```

```
Time Días*Status (0);
```

```
Strata Tratam;
```

```
Run;
```

### **Análisis semiparamétrico**

En la práctica, es relativamente sencillo probar si se cumple el supuesto de riesgo proporcional desarrollando un modelo de riesgo proporcional que incluye un término de interacción entre el tiempo y la covariable de interés, además de la propia covariable; si el término de interacción es estadísticamente significativo, el modelo AFT es más apropiado (Lee y Wang, 2003). Se utilizaron dos métodos para verificar y avalar las suposiciones de riesgo proporcional. El primero, un análisis gráfico de Kaplan – Meier, utilizando las curvas de  $\log(-\log(S(t)))$  versus  $\log(t)$ . El segundo, mediante el análisis de residuales de Schoenfeld. También se corrió una prueba para analizar la multicolinealidad entre las variables, basado en el diagnóstico VIF. Un valor de  $VIF > 5$  se considera evidencia de multicolinealidad y un valor  $VIF > 10$  es indicio de fuerte multicolinealidad. Para determinar el efecto de los factores de riesgo de la covariable tratamientos sobre la función de germinación, se abordó el análisis de sobrevivencia utilizando la función de regresión de riesgos proporcionales de Cox, que relaciona varios factores de riesgo de forma simultánea al tiempo de sobrevivencia (Cox, 1972).

*/\*Prueba De Multicolinealidad\*/*

Proc Reg Data = Caoba;

Model Días = Tratam Sem/ Vif;

Run;

Para ajustar un modelo de PH (riesgo proporcional) a los datos, se utilizó el procedimiento PHREG en SAS, con los términos de interacción entre ambas covariables y el tiempo. El valor nulo para la HR es el uno, que indica igualdad de probabilidad de producción del suceso en los dos grupos en el siguiente intervalo de tiempo. Una  $HR > 1$  indica más riesgo de producción en el tratamiento y una  $HR < 1$  menor riesgo en el tratamiento de intervención que en el control (tratamiento de referencia). En este sentido, una  $HR = 2$  no significa que las semillas bajo un tratamiento desarrollan el suceso dos veces más rápido, sino que, si una semilla no ha presentado todavía el suceso (germinación), tiene el doble de probabilidad de presentarlo que un control (tratamiento de referencia) durante el siguiente periodo de tiempo.

```
/*Proporcionabilidad con ASSESS*/  
Proc Phreg Data=Caoba;  
Model Días*Status (0) =Tratam Sem / Ties=Efron;  
Assess Ph / Resample;  
Run;  
/*Residuales Schoenfeld*/  
Proc Phreg Data=Caoba;  
Model Días*Status (0) =Tratam Sem / Ties=Efron;  
Output Out=Schoen Rensch=Schtratam Schsem;  
Run;  
Proc Print Data=Schoen;  
Run;  
Data Event;  
Set Schoen;  
Run;  
Proc Rank Data=Event Out=Resrank Ties=Mean;  
Var Días;  
Ranks Diasrank;  
Run;  
Proc Print Data=Resrank;  
Run;  
Data C;
```

```
Set Resrank;
```

```
Días1=Log (Días);
```

```
Días2=Días**2;
```

```
Proc Corr;
```

```
Var Diasrank Días1 Días2 Schtratam Schsem;
```

```
Run;
```

```
/*Generación de gráfica de residuales Schoenfeld para tratamientos como función del tiempo*/
```

```
Proc Gplot Data =Schoen;
```

```
Symbol1 V=Dot C=Black Width=1 I=Sm80s;
```

```
Plot Schtratam*Días/Haxis=Axis1 Vaxis=Axis2;
```

```
Axis1 Label=('Días');
```

```
Axis2 Label = (A=90 'Residuales Schoenfeld Para Tratamientos');
```

```
Run;
```

```
/*Generación De Gráfica De Residuales Schoenfeld Para Sem Como Función Del Tiempo*/
```

```
Proc Gplot Data =Schoen;
```

```
Symbol1 V=Dot C=Black Width=1 I=Sm80s;
```

```
Plot Schsem*Días/Haxis=Axis1 Vaxis=Axis2;
```

```
Axis1 Label=('Días');
```

```
Axis2 Label = (A=90 'Residuales Schoenfeld Para Sem');
```

```
Run;
```

```
/*Primer Modelo Cox de estimación con interacciones con variables de tiempo*/
```

```
Proc Phreg Data=Caoba;
```



```

Class Tratam;

Model Días*Status (0) = Tratam Sem Tratdias Semdias / Ties=Efron R1;

Test Tratam1=Tratam2; Test Tratam1=Tratam3; Test Tratam1=Tratam4;

Test Tratam2=Tratam3; Test Tratam2=Tratam4; Test Tratam3=Tratam4;

Tratdias=Tratam*(Días<15);

Semdias=Sem*(Días<15);

Run;

/*Segundo Modelo Cox de estimación con interacciones con variables de tiempo*/

Proc Phreg Data=Caoba;

Class Tratam;

Model Días*Status (0) = Tratam Sem Tratdias Semdias / Ties=Efron R1;

Test Tratam1=Tratam2; Test Tratam1=Tratam3; Test Tratam1=Tratam4;

Test Tratam2=Tratam3; Test Tratam2=Tratam4; Test Tratam3=Tratam4;

Tratdias=Tratam*(Días>15);

Semdias=Sem*(Días>15);

Run;

```

### **Análisis paramétrico**

En este caso, para la distribución de sobrevivencia se asumió el modelo de tiempo de falla acelerado (AFT, por sus siglas en inglés). Este modelo postula una relación directa entre predictores y tiempo de sobrevivencia (George, Seals y Aban, 2014).

Con el fin de establecer la distribución más apropiada y mejorar el potencial del modelo AFT, se ajustaron y compararon cuatro distribuciones paramétricas (Weibull, Log-Normal, Log-Logística y Exponencial), cuyos modelos se implementaron con PROC LIFEREG en SAS para describir la distribución del tiempo de supervivencia subyacente. Para establecer el mejor ajuste entre las distribuciones, se utilizó el Criterio de Máxima Verosimilitud (ML).

```
Proc Sort Data=Caoba;
```

```
By Tratam;
```

```
Run;
```

```
Proc Sort Data=Caoba;
```

```
By Sem;
```

```
Run;
```

```
Proc Lifereg Data=Caoba;
```

```
Class Tratam;
```

```
Model Días* Status (0) = Tratam Sem Tratdias Semdias/Dist= Weibull;
```

```
Probplot;
```

```
Run;
```

## Resultados.

### a. Análisis no paramétrico.

Uno de los principales propósitos de la germinación y emergencia de semillas para describir el proceso biológico, consiste en estimar algunos indicadores del lote de semillas, que permiten comparar lotes o cuantificar los efectos de factores experimentales sobre el proceso de germinación/emergencia. Entre ellos, la media y la mediana del tiempo de germinación (tiempo medio de germinación ( $T_{50}$ ): número de días transcurridos después de la incubación, para alcanzar el 50% del porcentaje de germi-

nación final). En algunos casos también se requieren otros tiempos para caracterizar mejor la primera y última fase del proceso, los tiempos para el 25 % o 75 % de germinación ( $T_{25} - T_{75}$ ), (Onofri *et al.*, 2010).

### **a.1. Comparación entre tratamientos de inmersión**

En las condiciones del ensayo, no existen grandes diferencias en el tiempo  $T_{50}$ , entre el tratamiento testigo y las hormonas sintéticas, con excepción de la inmersión en ácido giberélico a 200 ppm (ProGibb®), que obtuvo un tiempo medio de germinación de 18 días y los demás tratamientos 14,0 – 14,5 días. Otra observación importante es que, al aumentar la concentración de ProGibb®, disminuye el  $T_{50}$  o tiempo medio de germinación (18,0 – 14,5 – 14,0 días, respectivamente).

A continuación, se presentan estos indicadores para caracterizar los tratamientos de inmersión. Es un lote de semillas con germinación mediana a alta, dependiendo del tratamiento. Los tratamientos con más alta germinación fueron inmersión en Giber Grop® a 600 ppm e inmersión solo en agua (95,8 % vs 75,0 %, respectivamente; los tratamientos con ProGibb® la germinación más baja (62,5 – 66,7 %). Los valores de la media del tiempo de germinación resultaron variables, entre 14 – 17 días, en los que tratamientos de inmersión en ácido giberélico a 800 ppm (ProGibb®) durante 24 horas e inmersión en GiberGrop® a 600 ppm durante 24 horas, lograron mayor velocidad de germinación, como expresión del número de días promedio empleados en la germinación (13,9 y 14,3 días) versus los demás tratamientos. Igualmente, presentan la mayor tasa de germinación (0,071 semillas germinadas/día), interpretada como valor recíproco del tiempo medio de germinación ( $1/T_{50}$ ). Según Ranal y García, (2006), para los autores del índice, esta expresión debería interpretarse en el sentido de que altos valores deberían estar asociados con germinación concentrada en el tiempo. También se observa un detalle importante en el producto ProGibb®: al aumentar la concentración de 200 ppm a 800 ppm, disminuye la media del tiempo de germinación, de 17,3 a 14,3 días. También se aprecia que para la primera fase de germinación ( $T_{25}$ ), todos los tratamientos alcanzaron el 25 % de germinación en 12 días.

### a.1.1. Cuartiles estimados para cada tratamiento de inmersión de semillas de *Swietenia macrophylla* King, en días de germinación

#### Tratamiento 1:

Summary Statistics for Time Variable DÍAS  
Quartile Estimates

Percent	Point Estimate	95% Confidence Interval [Lower Upper)	
75	.	15.0000	.
50	14.5000	12.0000	19.0000
25	12.0000	12.0000	14.0000
	Mean	Standard Error	
	15.5833	0.6964	

#### Tratamiento 2:

Summary Statistics for Time Variable DÍAS  
Quartile Estimates

Percent	Point Estimate	95% Confidence Interval [Lower Upper)	
75	.	21.0000	.
50	18.0000	15.0000	.
25	13.0000	12.0000	16.0000
	Mean	Standard Error	
	17.2917	0.8103	

#### Tratamiento 3:

Summary Statistics for Time Variable DÍAS  
Quartile Estimates

Percent	Point Estimate	95% Confidence Interval [Lower Upper)	
75	.	16.0000	.
50	14.5000	12.0000	.
25	12.0000	12.0000	14.0000
	Mean	Standard Error	
	15.6667	0.7593	

**Tratamiento 4:**

## Summary Statistics for Time Variable DÍAS

Quartile Estimates			
Percent	Point Estimate	95% Confidence Interval [Lower Upper)	
75	.	16.0000	.
50	14.0000	14.0000	.
25	12.0000	12.0000	14.0000
Mean		Standard Error	
14.3333		0.3583	

**Tratamiento 5:**

## Summary Statistics for Time Variable DÍAS

Quartile Estimates			
Percent	Point Estimate	95% Confidence Interval [Lower Upper)	
75	15.0000	14.0000	17.0000
50	14.0000	12.0000	15.0000
25	12.0000	.	.
Mean		Standard Error	
13.8750		0.4011	

En la tabla Summary of the Number of Censored and Uncensored Values, que sigue a continuación, se observa que 26,67% de la cohorte de semillas fue censurada. Esto es equivalente a  $100\% - 26,67\% = 73,33\%$  de las semillas presentaron el evento de germinación (Failed). La germinación más alta corresponde al tratamiento 5 (inmersión de semillas en Giber Grop® a 600 ppm durante 24 horas;) con 95,83% de germinación y a continuación el tratamiento 1 (inmersión en agua durante 24 horas) con 75%.

Summary of the Number of Censored and Uncensored Values

Stratum	TRATAM	Total	Failed	Censored	Percent Censored
1	1	24	18	6	25.00
2	2	24	15	9	37.50
3	3	24	16	8	33.33
4	4	24	16	8	33.33
5	5	24	23	1	4.17
-----					
Total		120	88	32	26.67

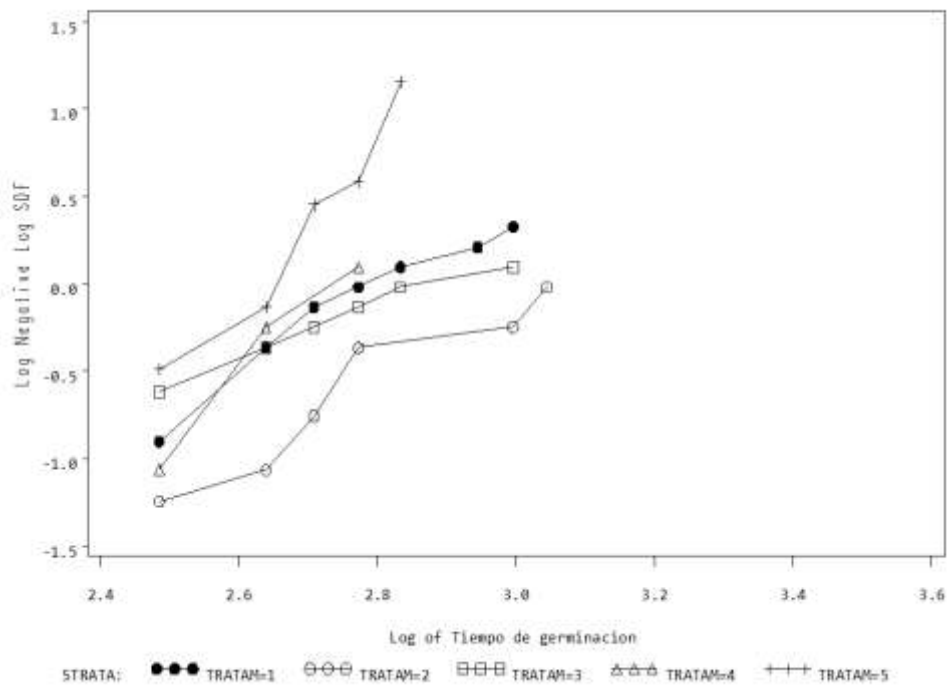
Uno de los principales propósitos de la germinación y emergencia de semillas para describir el proceso biológico, consiste en estimar algunos indicadores del lote de semillas, que permiten comparar tratamientos o cuantificar los efectos de factores experimentales sobre el proceso de germinación/emergencia. Entre ellos, la media y la mediana del tiempo de germinación (tiempo medio de germinación ( $T_{50}$ ): número de días transcurridos después de la incubación, para alcanzar el 50% del porcentaje de germinación final). En algunos casos también se requieren otros tiempos para caracterizar mejor la primera y última fase del proceso, los tiempos para el 25 % o 75 % de germinación ( $T_{25} - T_{75}$ ) (Onofri, Gresta y Tei, 2010).

Son indicadores utilizados para comparar la energía de germinación al registrar la “tasa de germinación”, es decir, el número de días que se necesitan para conseguir el 50 por ciento de la capacidad de germinación. Cuanto más breve sea ese período, tanto mayor será la energía de germinación. También se observa en los resultados, que la interacción hormonal entre giberelinas y auxinas favorece la capacidad de germinación de caoba, indicado por el tratamiento  $T_5$  (Giber Grop® a 600 ppm), producto que contiene ácido giberélico al 10% y auxina al 17,2%.

El ácido giberélico ( $AG_3$ ) se utiliza en semillas con bajo porcentaje de germinación, siendo útil en la determinación de los tiempos de germinación (25 – 50 – 75 %). Por ejemplo, López *et al.* (2017) en semillas de *Asclepias Subulata* Decne., una especie originaria de zona desértica, probaron seis concentraciones de  $AG_3$  (0,125, 250, 500, 1000 y 2000 ppm) a dos tiempos de remojo (24 y 48 horas) y un testigo sin ningún tratamiento. Encuentran que la escarificación con  $AG_3$  ejerce influencia positiva

en la germinación. El mayor porcentaje de germinación lo obtuvieron en concentración de 250 a 2000 ppm, a 24 y 48 horas de remojo. El tiempo de germinación se vio afectado por la concentración de AG<sub>3</sub>, siendo favorecido a una concentración de 1000 y 2000 ppm a 24 y 48 horas de remojo; en tanto que en los demás tratamientos no se alcanzó el 75% de germinación.

Al comparar las respuestas de las curvas de germinación individual, por efecto de los tratamientos, se observan cruzamientos de las curvas de germinación entre los tratamientos T<sub>1</sub>, T<sub>3</sub> y T<sub>4</sub>, alrededor de los 13 – 15 días. Los únicos tratamientos sin cruces son T<sub>2</sub> y T<sub>5</sub>. Por esta razón, se utiliza la prueba de Wilcoxon para la comparación de tratamientos, lo que indica que no existen diferencias entre tratamientos (Wilcoxon  $\chi^2 = 7,0502$ ; gl=4; p=0,1333; Figura 25).



**Figura 25. Función de Ln (-Ln(S(t))) versus log de tiempo con suposición PH violada de las funciones de germinación de *S. macrophylla* King, en cinco tratamientos de inmersión en hormonas. Fuente. Elaboración propia.**

A continuación, se indican los resultados de las pruebas que comparan las respuestas de las curvas de germinación. Las pruebas de Wilcoxon indicaron que no existen diferencias significativas entre tratamientos ( $p = 0,1333$ ).

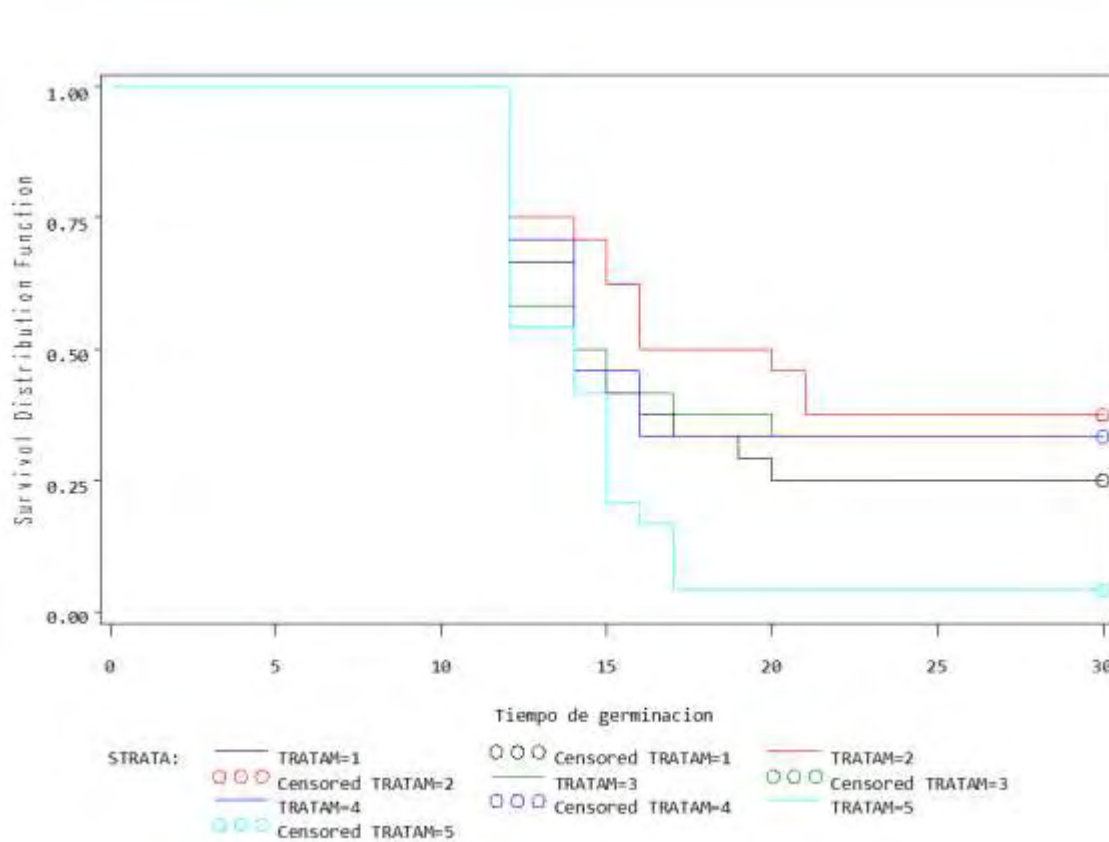
#### a.1.2. Resultado de las pruebas de comparación para las funciones de sobrevivencia entre tratamientos de inmersión de semillas de *S. macrophylla* King

##### Test of Equality over Strata

Test	Chi-Square	DF	Pr >
			Chi-Square
Log-Rank	10.3376	4	0.0351
Wilcoxon	7.0502	4	0.1333
-2Log(LR)	6.9302	4	0.1396

Los tratamientos presentaron distribuciones de probabilidad de germinación/emergencia semejante en algunos de ellos, con variación en los eventos de germinación, con 8 eventos para el tratamiento T<sub>1</sub>, 7 eventos para los tratamientos T<sub>2</sub> y T<sub>3</sub>, 4 eventos para el tratamiento T<sub>4</sub> y 6 eventos para el tratamiento T<sub>5</sub>. Los cinco tratamientos no presentaron eventos de germinación durante los primeros intervalos (0 – 15 días). A partir de estos primeros intervalos, los tratamientos presentaron probabilidades variables (Figura 26).

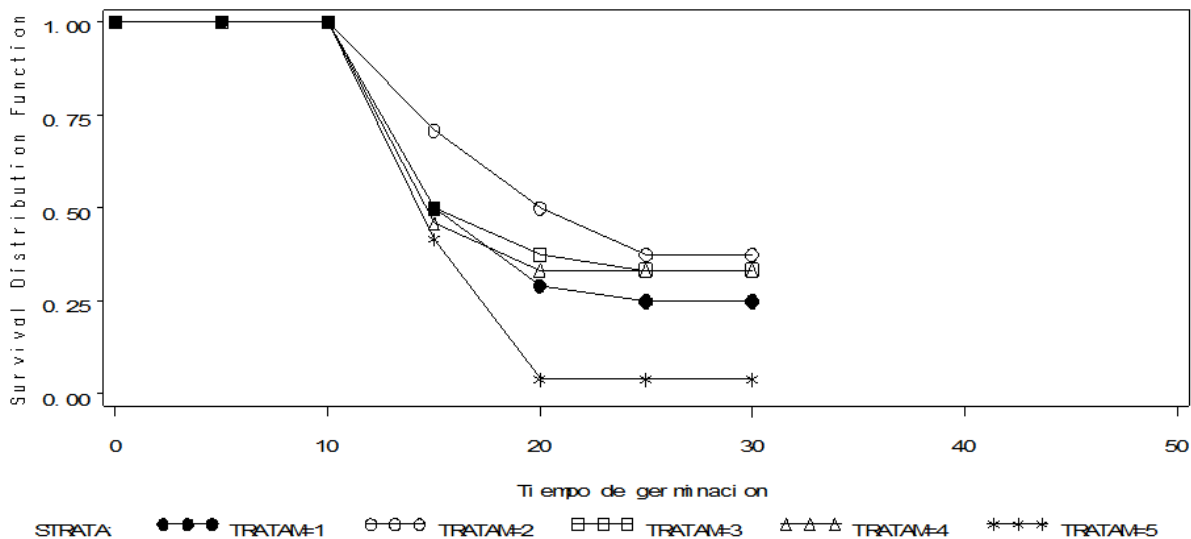




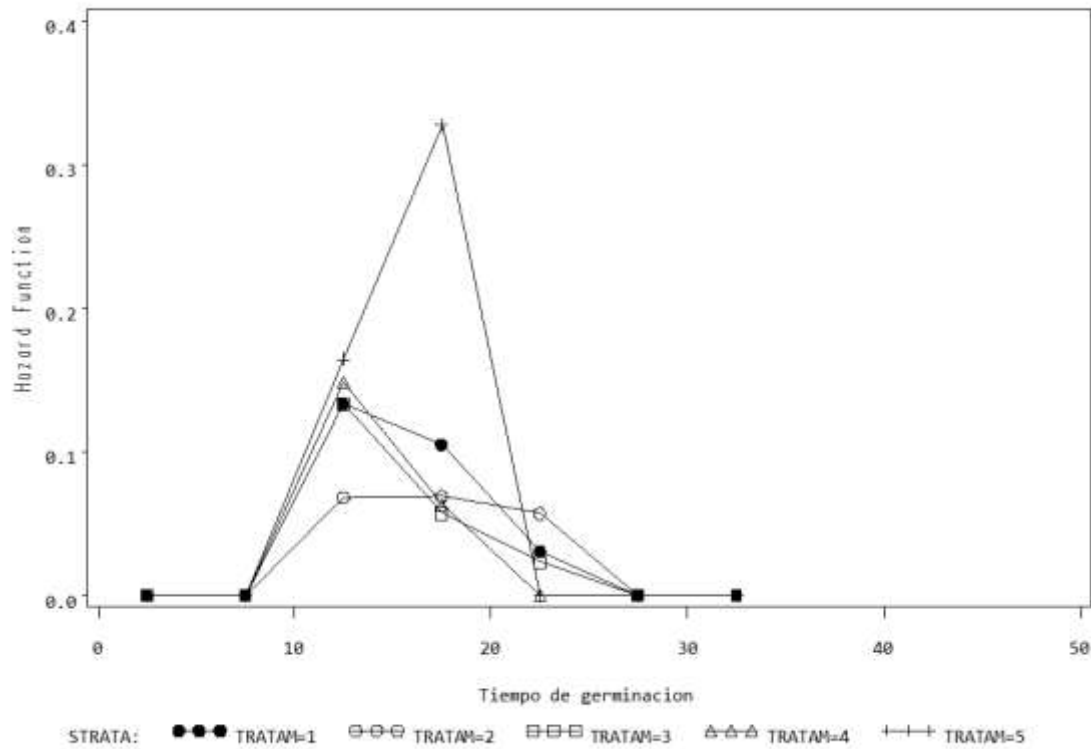
**Figura 26. Estimados de Kaplan–Meyer de las funciones de germinación de *S. macrophylla* King, en cinco tratamientos de inmersión en hormonas.** Fuente. Elaboración propia.

Las curvas de supervivencia de las semillas bajo las condiciones de los tratamientos de imbibición se muestran en la Figura 27. La proporción estimada de semillas que permanecieron sin germinar en el siguiente intervalo (15 – 20 días), para los tratamientos de inmersión en agua e inmersión en ácido giberélico a 400 ppm, fue de 0,50 (50 %) aumentando sus probabilidades de germinación, con una razón de riesgo de 0,10526 germinaciones/día y 0,05714 germinaciones/día, respectivamente (Figura 28). Para el intervalo final (30 – 35 días), las proporciones estimadas de semillas que permanecieron sin germinar en ambos tratamientos, además del tratamiento inmersión en ácido giberélico a 800 ppm, disminuyeron a 0,25, 0,33 y 0,33 (25 % y 33 %, respectivamente), con una razón de riesgo de 0,00 germinaciones/día. Significa que para el intervalo final sus probabilidades de germinación, aumentaron

a 0,75% y 0,67%. Por otra parte, el tratamiento inmersión en Giber Grop® a 600 ppm, luego de los primeros intervalos, disminuyó las probabilidades estimadas de que una semilla permaneciera sin germinar en los intervalos intermedios y finales (20 – 35 días) a 0,042 (4 %) de probabilidad de no germinar, aumentando la probabilidad de germinación (0,96%), presentando una razón de riesgo de 0,000 germinaciones/día. El tratamiento inmersión en ácido giberélico a 200 ppm, luego de los primeros intervalos (0 – 15 días), mantuvo una proporción de semillas sin germinar entre 0,71 y 0,50, en el intervalo 15 – 25 días, con una probabilidad de germinación entre 0,29 y 0,50% y una razón de riesgo entre 0,0571 germinaciones/día y 0,000 germinaciones/día. Finalmente, en el intervalo 25 – 35 días, mantuvo una proporción de semillas sin germinar de 0,38 (38 %), lo que significa una probabilidad de germinar de 0,62%, con 0,000 germinaciones/día, hasta el final del periodo de evaluación. En resumen, para el intervalo final (30 – 35 días), la proporción de semillas germinadas fue de 0,75 – 0,63 – 0,67 – 0,67 y 0,96%, para T<sub>1</sub> – T<sub>2</sub> – T<sub>3</sub> – T<sub>4</sub> y T<sub>5</sub>, respectivamente. Las curvas de supervivencia bajo las condiciones de los tratamientos no fueron diferentes (Wilcoxon:  $\chi^2 = 7,05$ , gl = 4, p = 0,1333).



**Figura 27.** Cambios en la supervivencia de la germinación de *S. macrophylla* King, en cinco tratamientos de inmersión en hormonas durante 30 días de evaluación. Fuente. Elaboración propia.



**Figura 28. Valores estimados de la función de riesgo para germinación de *S. macrophylla* King, en cinco tratamientos de inmersión en hormonas.** Fuente. Elaboración propia.

En la Figura 28 se observa que los tratamientos presentaron las más altas razones de riesgo en el intervalo 10 – 20 días, con el tratamiento T<sub>5</sub> con mayor razón de riesgo (0,3273 fallas/día, en el intervalo 15 – 20 días); T<sub>4</sub> (0,1487 germinación/día, en el intervalo 10 -15 días); T<sub>1</sub> y T<sub>3</sub>, con 0,1333 germinaciones/día, en el intervalo 10 – 15 días. El tratamiento T<sub>2</sub> logró la más baja razón de riesgo (0,068 – 0,0571 germinaciones/día, en el intervalo 10 – 25 días).

## b. Análisis semiparamétrico.

### b.1. Análisis de multicolinealidad.

La prueba de colinealidad diagnosticó que las covariables semillas y tratamientos, lograron valores de 1,00 ( $VIF < 5$ ), lo que indica que no existe evidencia de colinealidad en las variables.

The REG Procedure  
Model: MODEL1  
Dependent Variable: DÍAS Tiempo de germinación

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	232.98104	116.49052	2.16	0.1195
Error	117	6298.48562	53.83321		
Corrected Total	119	6531.46667			

Root MSE	7.33711	R-Square	0.0357
Dependent Mean	18.23333	Adj R-Sq	0.0192
Coeff Var	40.24010		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr >  t	Variance Inflation
Intercept	Intercept	1	19.90942	1.98248	10.04	<.0001	0
TRATAM		1	-0.90000	0.47361	-1.90	0.0599	1.00000
SEM		1	0.08191	0.09676	0.85	0.3990	1.00000

### b.2. Análisis de supuestos del modelo.

Las curvas de  $\log(-\log(S(t)))$  versus  $\log(t)$  para tratamientos e interacciones, presentaron líneas razonablemente paralelas para algunas de las covariables y sus interacciones. Los resultados del procedimiento ASSESS y los residuos de Schoenfeld, indican que, para satisfacer la suposición, ambos resultados deben ser independientes del tiempo, es decir, no estar correlacionados con el tiempo o con alguna función del tiempo (Allison, 2010).

**b.2.1. Prueba de suposiciones de riesgo proporcional con el procedimiento ASSESS.**

## The PHREG Procedure

## Model Information

Data Set	WORK.CAOBA	
Dependent Variable	DÍAS	Tiempo de germinación
Censoring Variable	STATUS	Status de germinación
Censoring Value(s)	0	
Ties Handling	EFRON	

Number of Observations Read	120
Number of Observations Used	120

## Summary of the Number of Event and Censored Values

Total	Event	Censored	Percent Censored
120	88	32	26.67

## Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

## Model Fit Statistics

Criterion	Without Covariates	With Covariates
-2 LOG L	752.509	747.383
AIC	752.509	751.383
SBC	752.509	756.337

## Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	5.1262	2	0.0771
Score	5.0982	2	0.0782
Wald	5.0725	2	0.0792

## Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
TRATAM	1	0.15596	0.07982	3.8177	0.0507	1.169
SEM	1	-0.02001	0.01565	1.6348	0.2010	0.980

## Supremum Test for Proportionals Hazards Assumption

Variable	Maximum Absolute Value	Replications	Seed	Pr > MaxAbsVal
TRATAM	0.4761	1000	339803000	0.9990
SEM	0.7181	1000	339803000	0.9930

Los resultados del procedimiento ASSESS, para la prueba de supuestos (Supremum Test for Proportionals Hazards Assumption) indica que no existen problemas de violación de supuestos, ya que  $p = 0,99$  ( $p > 0,05$ ), tanto en tratamientos como semillas.

### b.2.2. Prueba de suposiciones de riesgo proporcional con residuales de Schoenfeld, indicando coeficientes de correlación de Pearson, $\text{Prob} > |r|$ .

The CORR Procedure  
5 Variables: DIASRANK DÍAS1 DÍAS2 schTRATAM schSEM

## Simple Statistics

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
DIASRANK	120	60.50000	33.63128	7260	21.50000	104.50000
DÍAS1	120	2.83042	0.37114	339.64986	2.48491	3.40120
DÍAS2	120	386.88333	316.98533	46426	144.00000	900.00000
schTRATAM	88	-0.04272	1.48213	-3.75951	-2.30868	2.10833
schSEM	88	0.23652	6.93723	20.81337	-10.90430	11.33349

Pearson Correlation Coefficients  
 Prob > |r| under H0: Rho=0  
 Number of Observations

	DIASRANK	DÍAS1	DÍAS2	schTRATAM	schSEM
DIASRANK	1.00000	0.95372	0.88682	0.04391	0.20831
Rank Variable DÍAS		<.0001	<.0001	0.6846	0.0515
	120	120	120	88	88
DÍAS1	0.95372	1.00000	0.98377	-0.00013	0.18834
	<.0001		<.0001	0.9991	0.0789
	120	120	120	88	88
DÍAS2	0.88682	0.98377	1.00000	-0.03050	0.16677
	<.0001	<.0001		0.7778	0.1204
	120	120	120	88	88

Los resultados anteriores indican que no existen problemas de violación de supuestos de riesgo proporcional para los residuales de Schoenfeld, con las covariables semillas y tratamientos para las funciones de tiempo (Diasrank, Días1 y Días2) con valores de  $p > 0,05$ , coincidiendo con el resultado del procedimiento ASSESS. Por lo tanto, se ajustan al modelo, sin problemas de violación de supuestos.

### b.3. Análisis del modelo con variables tiempo dependientes

En este modelo, la variable de respuesta es el "riesgo". El modelo asume que la tasa de riesgo subyacente, en lugar del tiempo de supervivencia, es una función de las variables independientes o covariables. Significa que un cambio de una unidad en una variable explicativa causa cambios proporcionales en el riesgo. Para un análisis pormenorizado, la germinación se analizó por períodos. Es decir, se analizaron en un primer período para los primeros 15 días del proceso germinativo (entre 0 y 15 días) y luego en un segundo período para los siguientes 15 días (entre 16 y 30 días).

#### b.3.1. Para el primer período del proceso de germinación (0 – 15 días)

El signo positivo del coeficiente de regresión indica que para una covariable las probabilidades de germinación aumentan durante el proceso. También significa que la covariable aumenta el riesgo, lo que reduce el tiempo de falla o germinación. El efecto contrario se produce con los coeficientes negativos. Es decir, los tratamientos disminuyen las probabilidades de germinación y, en consecuencia, aumentan el tiempo de germinación.

**Análisis de covariables con interacciones con el tiempo por el método de máxima verosimilitud (ML) para la primera etapa del proceso de germinación (0 – 15 días)**

The PHREG Procedure

Model Information

Data Set	WORK.CAOBA	
Dependent Variable	DÍAS	Tiempo de germinación
Censoring Variable	STATUS	Status de germinación
Censoring Value(s)	0	
Ties Handling	EFRON	

Number of Observations Read	120
Number of Observations Used	120

Class Level Information

Class	Value	Design Variables			
TRATAM	1	1	0	0	0
	2	0	1	0	0
	3	0	0	1	0
	4	0	0	0	1
	5	0	0	0	0

Summary of the Number of Event and Censored Values			
Total	Event	Censored	Percent Censored
120	88	32	26.67

Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Without Covariates	With Covariates
-2 LOG L	752.509	736.989
AIC	752.509	750.989
SBC	752.509	768.330

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
------	------------	----	------------



Likelihood Ratio	15.5203	7	0.0299
Score	16.7383	7	0.0192
Wald	16.1789	7	0.0235

## Type 3 Tests

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRATAM	4	8.2037	0.0844
SEM	1	0.7347	0.3914
TRATDIAS	1	0.3481	0.5552
SEMDIAS	1	4.1595	0.0414

## Analysis of Maximum Likelihood Estimates

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% HR Confi.Limits
TRATAM 1	1	-0.85333	0.55332	2.3783	0.1230	0.426	0.144 1.260
TRATAM 2	1	-1.17408	0.47634	6.0752	0.0137	0.309	0.122 0.786
TRATAM 3	1	-0.85735	0.40579	4.4640	0.0346	0.424	0.192 0.940
TRATAM 4	1	-0.79591	0.35821	4.9367	0.0263	0.451	0.224 0.910
SEM	1	0.02289	0.02671	0.7347	0.3914	1.023	0.971 1.078
TRATDIAS	1	-0.09365	0.15873	0.3481	0.5552	0.911	0.667 1.243
SEMDIAS	1	-0.06728	0.03299	4.1595	0.0414	0.935	0.876 0.997

En este caso, para los primeros 15 días del proceso germinativo, todos los tratamientos ( $T_1 - T_2 - T_3 - T_4$ , en comparación con el tratamiento de referencia ( $T_5$ , inmersión en Giber Grop® a 600 ppm) disminuyeron las probabilidades de germinación y, en consecuencia, aumentaron el tiempo de falla o de germinación, de manera no significativa para  $T_1$  y significativamente para  $T_2$ ,  $T_3$  y  $T_4$ . En esta primera etapa del proceso de germinación, se observa que las interacciones de semillas y de los tratamientos con el tiempo también tuvieron un impacto negativo en la germinación. Es decir, ambas interacciones disminuyeron las probabilidades de germinación en el tiempo ( $\beta_i = -0,09365$  y  $\beta_i = -0,06728$ ), significativamente para la interacción entre semillas y tiempo ( $p = 0,041$ ), lo cual sugiere una violación de supuestos y no significativa para tratamientos y tiempo ( $p = 0,555$ ).

El valor de la razón de riesgo para los diferentes tratamientos, resultó de 0,426 – 0,309 – 0,424 y 0,451 y representan valores de  $HR < 1$ , que además presentaron signos negativos en sus coeficientes, significa que, en este período, tienen más baja probabilidad de germinación que el tratamiento de

referencia. Estos valores significan que los tratamientos T<sub>1</sub>, T<sub>2</sub>, T<sub>3</sub> y T<sub>4</sub> disminuyeron la función de riesgo a través del periodo de germinación por 57,4 %, 69,1 %, 57,6 % y 54,9 % con relación al tratamiento de referencia. Al observar los valores de HR, T<sub>1</sub> incluye el valor nulo (HR=1) entre su intervalo de confianza (0,14 – 1,26), indicando que el riesgo de germinación resultó similar en los dos grupos (T<sub>1</sub>=T<sub>5</sub>), por lo tanto, no existen diferencias significativas del riesgo temporal de germinación entre las semillas de ambos tratamientos. Para T<sub>2</sub>, T<sub>3</sub> y T<sub>4</sub> se presenta un escenario diferente, con valores HR que no incluyen en sus intervalos de confianza del 95% el valor nulo (HR=1), se concluye que el riesgo de germinación no fue similar en los tres grupos (T<sub>2</sub>≠T<sub>5</sub>; T<sub>3</sub>≠T<sub>5</sub>; T<sub>4</sub>≠T<sub>5</sub>), con diferencias significativas del riesgo temporal de germinación entre las semillas de los tratamientos bajo comparación.

Por otra parte, el valor inverso  $1/\exp(\beta_i)$  en los tratamientos a comparar con el tratamiento de referencia, iguales a 2,34 – 3,23 – 2,36 y 2,22, indica que las probabilidades de experimentar germinación disminuyen por estos factores para T<sub>1</sub>, T<sub>2</sub>, T<sub>3</sub> y T<sub>4</sub>, comparados con el tratamiento de referencia.

Además de la comparación de los demás tratamientos con Giber Grop® a 600 ppm, se realizaron otras pruebas de hipótesis para comparar entre tratamientos, originando una prueba muy útil. Para la primera etapa, con un valor de chi cuadrado de 8,2037 y 4 grados de libertad, la prueba resultó no significativa ( $p=0,08$ ). Entre los tratamientos comparados no se presentaron diferencias significativas:

#### Analysis of Maximum Likelihood Estimates

##### Linear Hypotheses Testing Results

Label	Wald Chi-Square	DF	Pr > ChiSq
Test 1	0.7463	1	0.3877
Test 2	0.0001	1	0.9921
Test 3	0.0153	1	0.9017
Test 4	0.7144	1	0.3980
Test 5	0.8413	1	0.3590
Test 6	0.0281	1	0.8670

### **b.3.2. Para el segundo período del proceso de germinación (16 - 30 días).**

#### **Análisis de covariables con interacciones con el tiempo por el método de máxima verosimilitud (ML) para la segunda etapa del proceso de germinación (16 – 30 días)**

En este caso, para el segundo período del proceso de germinación, todos los tratamientos ( $T_1 - T_2 - T_3 -$  y  $T_4$ , en comparación con el tratamiento de referencia ( $T_5$ , Inmersión en Giber Grop® a 600 ppm) disminuyeron las probabilidades de germinación y, en consecuencia, aumentaron el tiempo de falla o de germinación, de manera no significativa para  $T_1$ , alta significancia con  $T_2$  y significativamente para  $T_3$  y  $T_4$ .

En esta segunda etapa del proceso de germinación, a diferencia del período anterior, las interacciones de semillas y de los tratamientos con el tiempo tuvo un impacto positivo en la germinación. Es decir, la interacción aumentó las probabilidades de germinación en el tiempo ( $\beta_i=0,0758$  y  $\beta_i=0,0158$ ), significativamente para la interacción entre semillas y tiempo ( $p=0,042$ ) y no significativa para tratamientos y tiempo ( $p=0,933$ ).

Los valores de la razón de riesgo para los diferentes tratamientos, resultaron mayores que en el período anterior (0,563 – 0,382 – 0,491 y 0,495), continúan con valores de  $HR < 1$  y signos negativos en sus coeficientes, significa que, también en este período, tienen más baja probabilidad de germinación que el tratamiento de referencia. Estos valores significan que los tratamientos  $T_1$ ,  $T_2$ ,  $T_3$  y  $T_4$  disminuyeron la función de riesgo a través del periodo de germinación por 43,7%, 61,8%, 50,9% y 50,5 %, con relación al tratamiento de referencia. Representa una disminución en relación al período anterior de 13,7%, 7,3%, 6,7% y 4,4% para cada tratamiento. Al observar los valores de HR,  $T_1$  incluye el valor nulo ( $HR=1$ ) entre su intervalo de confianza (0,29 – 1,05), indicando que el riesgo de germinación resultó similar en los dos grupos ( $T_1=T_5$ ), aunque sin diferencias significativas del riesgo temporal de germinación entre las semillas de ambos tratamientos. Para  $T_2$ ,  $T_3$  y  $T_4$  se presenta un escenario diferente, con valores HR que no incluyen en sus intervalos de confianza del 95% el valor nulo ( $HR=1$ ), se concluye que el riesgo de germinación no fue similar en los tres grupos ( $T_2 \neq T_5$ ;  $T_3 \neq T_5$ ;  $T_4 \neq T_5$ ), con diferencias significativas del riesgo temporal de germinación entre las semillas de los tratamientos bajo comparación.

Por otra parte, el valor inverso  $1/\exp(\beta_i)$  en los tratamientos a comparar con el tratamiento de referencia, iguales a 1,77 – 2,62 – 2,03 y 2,02, indica que las probabilidades de experimentar germinación disminuyen por estos factores para T<sub>1</sub>, T<sub>2</sub>, T<sub>3</sub> y T<sub>4</sub>, comparados con el tratamiento de referencia.

## The PHREG Procedure

## Testing Global Null Hypothesis: BETA=0

Likelihood Ratio	15.0271	7	0.0357
Score	16.1486	7	0.0238
Wald	15.6290	7	0.0287

## Type 3 Tests

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRATAM	4	9.5175	0.0494
SEM	1	4.6012	0.0319
TRATDIAS	1	0.0071	0.9328
SEMDIAS	1	4.1216	0.0423

## Analysis of Maximum Likelihood Estimates

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% HR Confi.Limits
TRATAM 1	1	-0.57534	0.34461	2.7874	0.0950	0.563	0.286 1.105
TRATAM 2	1	-0.96201	0.34930	7.5852	0.0059	0.382	0.193 0.758
TRATAM 3	1	-0.71058	0.33198	4.5815	0.0323	0.491	0.256 0.942
TRATAM 4	1	-0.70268	0.32912	4.5584	0.0328	0.495	0.260 0.944
SEM	1	-0.03837	0.01789	4.6012	0.0319	0.962	0.929 0.997
TRATDIAS	1	0.01580	0.18729	0.0071	0.9328	1.016	0.704 1.467
SEMDIAS	1	0.07587	0.03737	4.1216	0.0423	1.079	1.003 1.161

Además de la comparación de los demás tratamientos con Giber Grop® a 600 ppm, las pruebas de hipótesis para comparar entre tratamientos, con un valor de chi cuadrado de 9,5175 y 4 grados de libertad, la prueba resultó significativa ( $p=0,04$ ), aunque sin diferencias significativas entre los tratamientos comparados.

## The PHREG Procedure

## Linear Hypotheses Testing Results

Label	Wald		Pr > ChiSq
	Chi-Square	DF	
Test 1	1.1808	1	0.2772
Test 2	0.1439	1	0.7044
Test 3	0.1163	1	0.7331
Test 4	0.4744	1	0.4910
Test 5	0.4734	1	0.4914
Test 6	0.0005	1	0.9824

**c. Análisis paramétrico**

La aplicación del Criterio de Información Akaike (Akaike, 1974), permitió seleccionar la distribución Weibull por mejor ajuste. En una distribución Weibull, cuando el parámetro de escala  $\sigma > 1$ , el riesgo disminuye con el tiempo. Cuando  $0,5 < \sigma < 1$ , el riesgo es creciente en una tasa decreciente; cuando  $0 < \sigma < 0,5$ , el riesgo es creciente a una tasa creciente y cuando  $\sigma = 0,5$ , la función de riesgo es una línea recta creciente con origen en 0 (Allison, 2010).

La salida de resultados, muestra las estimaciones de máxima verosimilitud (MLE) de los parámetros calculados mediante la aplicación del modelo Weibull, expresado como regresión de tiempo de falla acelerada (AFT). La primera observación es que el valor estimado de  $\sigma$  (Escala) se encuentra entre  $0 < \sigma < 0,5$  ( $\sigma = 0,0771$ ), lo que implica que la función de riesgo es creciente a una tasa creciente (Allison 2010).

El coeficiente de regresión para el tratamiento T<sub>5</sub> (Giber Grop 600 ppm), es 0. Su razón es que se designó como nivel de referencia para comparar con los demás tratamientos. El análisis de resultados, indica que cada coeficiente, de manera independiente, difiere significativamente del tratamiento Giber Grop 600 ppm ( $p < 0,0001$ ).

**Parámetros estimados de las distribuciones por el método de máxima verosimilitud (ML) para la distribución Weibull.**

The LIFEREG Procedure							
Model Information							
Data Set	WORK.CAوبا						
Dependent Variable	Log (DÍAS)		Tiempo de germinación				
Censoring Variable	STATUS		Status de germinación				
Censoring Value(s)	0						
Number of Observations	120						
Noncensored Values	88						
Right Censored Values	32						
Left Censored Values	0						
Interval Censored Values	0						
Name of Distribution	Weibull						
Log Likelihood	83.37549349						
Number of Observations Read							120
Number of Observations Used							120
Class Level Information							
Name	Levels	Values					
TRATAM	5	1	2	3	4	5	
Algorithm converged.							
Type III Analysis of Effects							
Effect	DF	Wald		Pr > ChiSq			
TRATAM	4	100.0935		<.0001			
SEM	1	34.5827		<.0001			
TRATDIAS	1	79.7502		<.0001			
SEMDIAS	1	28.7385		<.0001			
Analysis of Parameter Estimates							
Parameter	DF	Estimate	Standard Error	95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	1.3755	0.1439	1.0935	1.6575	91.40	<.0001
TRATAM	1	1.1157	0.1161	0.8882	1.3433	92.36	<.0001
TRATAM	2	0.7696	0.0787	0.6153	0.9239	95.55	<.0001
TRATAM	3	0.5053	0.0574	0.3928	0.6178	77.54	<.0001
TRATAM	4	0.2339	0.0361	0.1632	0.3046	42.02	<.0001
TRATAM	5	0.0000	.	.	.	.	.
SEM	1	-0.0371	0.0063	-0.0494	-0.0247	34.58	<.0001
TRATDIAS	1	0.0193	0.0022	0.0150	0.0235	79.75	<.0001
SEMDIAS	1	0.0024	0.0005	0.0015	0.0033	28.74	<.0001
Scale	1	0.0771	0.0062	0.0659	0.0901		
Weibull Shape	1	12.9764	1.0377	11.0939	15.1784		

En la estimación de los parámetros ATF para el modelo Weibull, la intersección está representado por  $\beta_0 = 1,3755$ , semillas ( $\beta_1 = -0,037$ ), tratamiento 1 ( $\beta_2 = 1,1157$ ), tratamiento 2 ( $\beta_3 = 0,7696$ ), tratamiento 3 ( $\beta_4 = 0,505$ ) y tratamiento 4 ( $\beta_5 = 0,2339$ ), etc. Según Romano y Stevanato (2020), estos valores se utilizan para calcular el tiempo medio de germinación ( $T_{50}$ ,  $\exp(\beta_i)$ ). El  $T_{50}$  (mediana o tiempo medio de germinación) para referencia de cada parámetro, se calcula como  $\exp(\beta)$ . Para todo el lote (intercepto) fue de  $\exp(1,3755) = 4,0$  días, mientras que para el efecto de semillas  $\exp(\beta_0 + \beta_1) = 3,8$ ; para tratamiento 1,  $\exp(\beta_0 + \beta_2) = 12,1$  días; para tratamiento 2,  $\exp(\beta_0 + \beta_3) = 8,5$ ; tratamiento 3,  $\exp(\beta_0 + \beta_4) = 6,6$ ; tratamiento 4,  $\exp(\beta_0 + \beta_5) = 5,0$  y tratamiento 5,  $\exp(\beta_0 + \beta_6) = 4,0$ . El tratamiento 5 (inmersión en Giber Grop® a 600 ppm) se considera de referencia, por lo que aparece con un valor neutro (0,00). De esta manera, se calcula el tiempo medio de germinación para cada variable, según su parámetro (Romano y Stevanato, 2020). Según estos valores el tratamiento con sólo inmersión en agua tuvo el mayor tiempo medio de germinación en comparación con el tratamiento inmersión en Giber Grop® a 600 ppm. En general, todos los tratamientos que se compararon con inmersión en Giber Grop® a 600 ppm, lograron mayor tiempo medio de germinación ( $T_{50}$ ).

En términos de la razón de tiempo ( $\gamma$ ), para sem  $\gamma = 3,8/4,0 = 0,95$ ; para tratamiento 1,  $\gamma = 12,1/4,0 = 3,03$ ; tratamiento 2,  $\gamma = 8,5/4,0 = 2,13$ ; tratamiento 3,  $\gamma = 6,6/4,0 = 1,65$ ; tratamiento 4,  $\gamma = 5,0/4,0 = 1,25$ . Interpretados como razones estimadas de los tiempos de sobrevivencia esperados para dos grupos ( $100\%(1 - \gamma)$ ), se puede observar que las semillas prolongan la germinación por 5 %. Por el contrario, los tratamientos  $T_1$ ,  $T_2$ ,  $T_3$  y  $T_4$ , prolongan la germinación en 203,0 %, 113 %, 65% y en 25 %, respectivamente, en comparación con el tratamiento de referencia. Esto significa que el tratamiento que prolonga u origina una germinación más lenta, es el tratamiento  $T_1$  (inmersión en agua). En general, en los modelos AFT, una razón de tiempo  $> 1$  para una covariable, implica que prolonga el tiempo de germinación, mientras que una razón de tiempo  $< 1$  implica que es más probable un evento de germinación más temprano (Romano y Stevanato, 2020).

Según Romano y Stevanato (2020), mientras que el modelo de Cox expresa el efecto multiplicativo de las variables explicativas sobre el riesgo (escala de riesgo), el modelo AFT expresa el mismo efecto sobre el tiempo de supervivencia (escala de tiempo). Según ello, esta característica permite una interpretación más fácil de los resultados porque los parámetros miden la fuerza y el efecto de la covariable

correspondiente en la mediana del tiempo de supervivencia. Además, por la aplicación de tiempo de falla acelerado, el efecto de las covariables sobre la supervivencia se describe en términos absolutos (por ejemplo, número de días) en lugar de términos relativos como el índice de riesgo. Además, las razones de tiempo representan índices útiles que describen los efectos de las variables explicativas sobre la germinación y más interpretable que una relación de dos riesgos.

### **3.3.2. Entomología y otros experimentos de laboratorio**

#### **3.3.2.1. Efecto de dos dietas alternantes vegetales en el desarrollo larval de dos especies de mariposas diurnas; *Morpho menelaus* y *Morpho helenor* bajo condiciones de cautiverio (Peña, 2017).**

##### **Metodología**

Este estudio, tuvo como objetivo principal evaluar el desarrollo larval en cautiverio de dos especies de mariposas diurnas mediante el suministro de dietas vegetales. Para evaluar el ciclo biológico de las dos especies, se emplearon 20 larvas recién emergidas, mientras que, para determinar el efecto en el desarrollo larval, se empleó un diseño completamente al azar, que consistió en la aplicación de dos tratamientos con diferentes repeticiones ( $n_1=31$  y  $n_2=14$ ). Entre los resultados obtenidos, se reporta que la dieta con *Arachis pintoii* fue óptima para las especies *Morpho menelaus occidentalis* y *Morpho helenor theodorus*. *M. menelaus occidentalis* completó su ciclo biológico en  $97,51 \pm 7,98$  días ( $n=15$ ), teniendo una supervivencia del 75 %. En cuanto a *M. helenor theodorus*, el ciclo de vida lo completó en  $73,22 \pm 7,26$  días ( $n=14$ ), con una sobrevivencia del 70 %; mientras que, con la dieta de *Platymiscium stipulare* completó su ciclo biológico en mayor tiempo,  $75,84 \pm 10,19$  días ( $n=10$ ), en contraste presentó una baja tasa de sobrevivencia (50 %). Respecto al uso de las plantas hospederas *Vigna candida* af. y *Desmodium adscendens*, no resultaron las especies preferidas por las dos *Morpho* spp., ya que las larvas murieron durante el primer estadio de desarrollo.

##### **Objetivo general.**

Determinar el ciclo biológico y el efecto de dos dietas alternantes vegetales en el desarrollo larval de dos especies de mariposas diurnas; *Morpho menelaus occidentalis* y *Morpho helenor theodorus*, en condiciones de cautiverio.



### Objetivos específicos

Evaluar la duración del ciclo biológico de *M. menelaus occidentalis* y *M. helenor theodorus* bajo condiciones de cautiverio de acuerdo a los dos tipos de dietas vegetales administradas.

Conocer el comportamiento de la dieta alternante vegetal, en la tasa de desarrollo larvario y tasa de supervivencia en los diferentes estadios de desarrollo del ciclo biológico y tamaño del adulto de ambas especies de mariposas.

### Estructura y análisis de datos

Para el análisis de datos se utilizará el ciclo biológico con *M. helenor theodorus* bajo el efecto de dos dietas: *Arachis pintoi* (Dieta 1) y *Platymiscium stipulare* (Dieta 2). Se hace uso de los datos del autor por considerarlos ejemplo típico de tablas de vida en insectos, lo cual puede contribuir al uso principal del Manual, como libro de enseñanza utilizando las aplicaciones en SAS, es decir, analizar los ciclos biológicos mediante análisis no paramétrico, semiparamétrico y paramétrico con PROC LIFETEST, PHREG y LIFEREG para describir las diferencias entre las dietas en los diferentes estadios de crecimiento.

#### a. Análisis no paramétrico

Los datos de ciclo biológico fueron descritos y modelados bajo dos probabilidades relacionadas: sobrevivencia y función de riesgo, ambas dependientes del tiempo. Las diferencias de supervivencia entre dietas, fases y estadios, se analizaron mediante las pruebas de Log-Rank, Wilcoxon y razón de probabilidad ( $-2\log(LR)$ ), como pruebas predeterminadas del análisis a partir de curvas de supervivencia construidas por los métodos Kaplan-Meier y método tabla de vida o actuarial (LT). Se tomó en cuenta el estatus de cada fase y estadio (completada o no) al final del periodo de evaluación, así como el tiempo de duración de la misma en días.

Ya que se ha aplicado un tratamiento de dieta a cada grupo, la pregunta obvia es ¿cual tratamiento marcó una diferencia en la supervivencia de los dos grupos? Un enfoque natural para responder a esta pregunta es probar la hipótesis nula de que las funciones de superviviente son las mismas en los dos grupos (es decir,  $S_1(t) = S_2(t)$  para todo  $t$ ), ya que PROC LIFETEST puede calcular varias estadísticas diferentes para probar esta hipótesis nula. PROC LIFETEST por defecto, imprime los estadísticos log-

rank y Wilcoxon para cada grupo o tratamiento, seguido por un estimado de la matriz de covarianza. Estos estadísticos se utilizan para calcular el valor de Chi cuadrada. La opción METHOD= especifica los métodos para computar las funciones de sobrevivencia. En este ejemplo, se utilizarán las opciones KM (Kaplan – Meier), y la opción LT, para especificar la tabla de vida (o actuarial). Se utilizarán para las dietas, fases y estadios de las especies.

Para los datos de dietas de alimentación, fases y estadios, es deseable probar si individualmente, las variables de alimentación (DIETA), las fases y los estadios tienen algún efecto sobre la experiencia de supervivencia de los dos grupos. Para hacer eso con PROC LIFETEST, simplemente se agrega una declaración STRATA después de la declaración TIME y con la opción PLOT se originan las diferentes gráficas:

```
Proc Lifetest Data=Morpho Method=Km Plots= (S, Ls, Lls);
```

```
Time Duración*Status (0);
```

```
Strata Dieta;
```

```
Run;
```

```
Proc Lifetest Data=Morpho Method=Lt Plots= (S, H);
```

```
Time Duración*Status (0);
```

```
Strata Dieta;
```

```
Run;
```

```
Proc Lifetest Data=Morpho Method=Km Plots= (S, Ls, Lls);
```

```
Time Duración*Status (0);
```

```
By Dieta;
```

```
Strata Fase;
```

```
Run;
```

```
Proc Lifetest Data=Morpho Method=Lt Plots= (S, H);
```

```

Time Duración*Status (0);

By Dieta;

Strata Fase;

Run;

Proc Lifetest Data=Morpho Method=Km Plots= (S, Ls, Lls);

Time Duración*Status (0);

By Dieta;

Strata Estadio;

Run;

Proc Lifetest Data=Morpho Method=Lt Plots= (S, H);

Time Duración*Status (0);

By Dieta;

Strata Estadio;

Run;

```

En estos programas, la primera variable (DURACIÓN), correspondiente al tiempo de duración de los individuos en cada fase o estadio, también conocida como tiempo del evento o censura, la segunda variable (Status) contiene la información sobre las observaciones que fueron censuradas o no (si completo la muda o no) y el número entre paréntesis (0) son valores de observaciones censuradas.

La declaración STRATA tiene tres consecuencias:

1. Primero, en lugar de una sola tabla con estimaciones de LT, se producen tablas separadas para cada uno de los dos grupos de tratamientos a base de dietas.
2. En segundo lugar, correspondientes a las tablas para cada estrato (Dieta, Fase y Estadio) hay tres gráficos de la función de superviviente, superpuesta en los mismos ejes para facilitar la comparación.

3. En tercer lugar, PROC LIFETEST informa varias estadísticas relacionadas con la prueba de diferencias entre los tres grupos.

Para las gráficas de sobrevivencia se utiliza la opción PLOTS= para definir o crear gráficas de los estimados de sobrevivencia u observaciones censuradas para dieta, fase y estadio. En estos casos, se utiliza Plots= S, para crear una gráfica de la función de sobrevivencia versus tiempo. Además, de la creación de gráficas con Plots=H. para especificar una gráfica de la función de riesgo versus tiempo, gráfica solo para métodos de tablas de vida (LT). Las opciones Plots=LS para una gráfica de  $-\log$  (estimated SDF) versus tiempo. y Plots=LLS para especificar una gráfica de la función  $\log(-\log$  (estimated SDF)) versus  $\log(\text{tiempo})$ , ambas gráficas solo para métodos de KM.

#### **b. Análisis semiparamétrico.**

Para determinar el efecto de los factores de riesgo (dietas de alimentación) sobre la función de fases y estadios al mismo tiempo, se abordó el análisis de sobrevivencia de forma multivariada, utilizando la función de regresión de riesgos proporcionales de Cox, que relaciona varios factores de riesgo de forma simultánea al tiempo de sobrevivencia (Cox, 1972). El modelo definido estima un coeficiente  $\beta$  para cada factor o covariable y prueba la hipótesis nula que  $\beta = 0$  con el estadístico  $\text{Chi}^2$ . Dicho coeficiente explica el efecto de un factor o una covariable en la función de riesgo; es decir, si el coeficiente  $\beta$  es negativo, significa que el riesgo de no completar la fase y estadio se reduce con el incremento de la covariable, mientras que un coeficiente  $\beta$  positivo indica lo contrario (Williams, 2008). El análisis se efectuó mediante el procedimiento PHREG (SAS, 2020).

Uno de los objetivos del análisis de sobrevivencia consiste en evaluar si el tiempo de sobrevivencia está relacionado a caracteres distintivos. Los modelos de regresión se pueden utilizar para evaluar los efectos de covariables sobre el evento, de manera similar a modelos de regresión lineal o logística. Dos tipos de modelos de regresión aplicados a datos de sobrevivencia, con los que se asume que las covariables afectan el tiempo de sobrevivencia son, el *modelo de riesgos proporcionales de Cox* y el *modelo de tiempo de falla acelerado* (Kartsonak, 2016).

La interpretación del parámetro  $\beta$  es que por cada unidad que aumenta la covariable X, el riesgo se multiplica por  $e^\beta$ . En el caso especial de X tomando el valor 0 o 1 para representar a grupos, tal como A y B, el grupo B tiene  $e^\beta$  veces el riesgo del grupo A. Como en todo modelo de regresión, más de una variable se puede incluir en el modelo de Cox. Para el cálculo se utilizará el siguiente programa, con la Dieta como covariable:

```
Proc Phreg Data=Morpho;
```

```
Model Duración*Status (0) =Dieta/Ties=Efron RL;
```

```
Run;
```

### **Análisis semiparamétrico y pruebas de supuestos de riesgo no proporcional (PH)**

El principal beneficio de los métodos semiparamétricos es que permiten evaluar efectos potenciales de múltiples covariables. Una importante ventaja del modelo de Cox sobre métodos no paramétricos es que permite covariables cuantitativas. También se permiten covariables categóricas, y estas son codificadas usando variables indicadoras (o "ficticias"), como en regresión ordinaria de mínimos cuadrados. Las covariables dependientes del tiempo son aquellas que pueden cambiar de valor durante el curso de la observación.

Como ejemplo, se pueden crear dos variables "ficticias", con el nombre de X e Y. Es decir, en este ejemplo, no se miden originalmente otras variables relacionadas con los organismos experimentales, sino como "creación de nuevas variables de pronóstico", para ser utilizadas en PROC PHREG, por la siguiente explicación: la desventaja de la regresión de Cox es la poca o ninguna habilidad para ajustar otras variables e incluirlas en el modelo. Por ejemplo, en caso de que se requiera expandir el modelo incluyendo una nueva variable que contiene el peso inicial de las unidades experimentales (variable X). Cuando se introducen covariables dependientes del tiempo en un modelo de regresión de Cox, no es preciso llamarlo un modelo de riesgos proporcionales (PH). ¿Por qué? Porque las covariables dependientes del tiempo cambiarán a diferentes velocidades para diferentes individuos, por lo que las proporciones de sus riesgos no pueden permanecer constantes. Pero suponga que no tiene ninguna covariable dependiente del tiempo. ¿Cómo saber si sus datos satisfacen la suposición de PH y qué sucede si se viola el supuesto? Cada modelo incorpora muchas suposiciones, algunas más cuestionables o conse-

cuentas que otras. Al mismo tiempo, a menudo ignoran cuestiones críticas como: ¿Están incluidas todas las covariables relevantes? ¿El error de medición en las covariables es aceptablemente bajo? (Allison, 2010).

Para poner este problema en perspectiva, se debe comprender que violaciones de la suposición de PH son equivalentes a interacciones entre una o más covariables y tiempo. Es decir, el modelo PH asume que el efecto de cada covariable es el mismo en todos los momentos. Si el efecto de una variable varía con el tiempo, la suposición de PH se viola para esa variable. Es poco probable que la suposición de PH se cumpla exactamente alguna vez, pero eso es cierto para casi todos los supuestos estadísticos. Si estimamos un modelo de PH cuando se viola la suposición para alguna variable (suprimiendo la interacción), entonces el coeficiente que estimamos para esa variable es una especie del efecto promedio en el rango de tiempos observados en los datos. ¿Es esto así de complicado? De hecho, los investigadores suprimen las interacciones todo el tiempo cuando estiman modelos de regresión. En modelos con un número de variables moderadamente grande, nadie prueba todas las posibles interacciones bidireccionales: hay demasiados de ellos (Allison, 2010).

En algunos casos, sin embargo, existen motivos para creer que las interacciones con el tiempo son tan fuertes que resultaría inconveniente reprimirlos. En otras situaciones, podría existir un fuerte interés en esas interacciones, incluso cuando no son muy fuertes. Para esos casos, necesitamos examinar métodos para probar y modelar riesgos no proporcionales. De esta manera, se comienza a probar los riesgos proporcionales usando residuales. Luego se pueden implementar dos métodos adicionales. Uno de ellos, incorpora explícitamente las interacciones en el modelo. El otro, llamado estratificación, incorpora las interacciones en la función arbitraria del tiempo.

### **Prueba del supuesto de proporcionalidad con los residuales de Schoenfeld**

También puede realizar pruebas de riesgos no proporcionales con los residuales de Schoenfeld. A diferencia de la sentencia ASSESS, este método funciona incluso si tienen covariables cambiantes en el tiempo definidas por el método del proceso de conteo o variables evaluadas en la investigación, por ejemplo, temperatura (pero no con el método de declaraciones de programación). Los residuos de Schoenfeld no se definen para observaciones censuradas (faltan en el conjunto de datos de salida). El

uso principal de estos residuos es detectar posibles desviaciones del supuesto de riesgos proporcionales. Si se satisface la suposición, los residuos de Schoenfeld deben ser independientes del tiempo. Probando la suposición de riesgos proporcionales es equivalente, por lo tanto, a probar si los residuos de Schoenfeld están correlacionados con el tiempo o con alguna función del tiempo (Allison, 2010). La opción RL se utiliza para la definición de los intervalos de confianza al 95% de las razones de riesgo. Para los datos del ejemplo, se produce un conjunto de datos que contiene los residuos de Schoenfeld utilizando estas declaraciones:

```
/*Residuales Schoenfeld*/
```

```
Proc Phreg Data= Morpho;
```

```
Model Duración*Status (0) =Dieta / Ties=Efron RL;
```

```
Output Out= Schoen Resch=Schdieta;
```

```
Run;
```

El siguiente paso es modificar el conjunto de datos para incorporar las funciones del tiempo, el LOG y el cuadrado de la variable relacionada con el tiempo (Duración), además de una función de tiempo jerarquizada (Diasrank), seguido de PROC CORR para calcular las correlaciones de los residuos con el tiempo y sus dos funciones.

```
Proc Print Data=Schoen;
```

```
Run;
```

```
Data Event;
```

```
Set Schoen;
```

```
Run;
```

```
Proc Rank Data=Event Out=Resrank Ties=Mean;
```

```
Var Duración;
```

```
Ranks Duracionrank;
```

```
Run;
Proc Print Data=Resrank;
Run;
Data C;
Set Resrank;
Duración1=Log (Duración);
Duración2=Duración*2;
Proc Corr;
Var Duracionrank Duración1 Duración2 Schdieta;
Run;
```

### **Generacion de gráficas de residuales Schoenfeld para Grupos y covariable X como función del tiempo.**

*/\*Generación de gráfica de residuales Schoenfeld para DIETA como función del tiempo\*/*

```
Proc Gplot Data =Schoen;
Symbol1 V=Dot C=Black Width=1 I=Sm80s;
Plot Schdieta*Duración/Haxis=Axis1 Vaxis=Axis2;
Axis1 Label= ('Dias');
Axis2 Label= (A=90 'Residuales Schoenfeld Para Dieta');
```

```
Run;
```

*/\*Modelo Cox de estimación inicial simple, solo con la covariable DIETA (sin interacciones) \*/.*

```
Proc Tphreg Data=Morpho;
Class Dieta;
```



```
Model Duración*Status (0) =Dieta/Ties=Efron RI;
```

```
Run;
```

```
/*Modelo Cox con interacción de la covariable DIETA y el tiempo (DURACIÓN) */.
```

```
Proc Tphreg Data=Morpho;
```

```
Class Dieta;
```

```
Model Duración*Status (0) =Dieta Dietdur/ Ties=Efron RL;
```

```
Dietdur=Dieta*Duración;
```

```
Run;
```

### c. Análisis paramétrico

En este caso, para la distribución de sobrevivencia se asumió una forma específica, llamada modelo de tiempo de falla acelerado (AFT, por sus siglas en inglés). Este modelo postula una relación directa entre predictores y tiempo de sobrevivencia (George, Seals y Aban, 2014).

Con el fin de establecer la distribución más apropiada y mejorar el potencial del modelo AFT, se ajustaron y compararon cuatro distribuciones paramétricas (Weibull, Log-normal, Log-logistic y Exponencial). Para establecer el mejor ajuste, los modelos en PROC LIFEREG son estimados por el método de máxima verosimilitud (ML), la magnitud más baja corresponde al mejor ajuste (Allison 2010). Se corrieron las distribuciones Llogistic – Exponencial – Weibull – Lognormal – Gamma, resultando la distribución Llogistic con la mejor convergencia de los datos:

```
Proc Sort Data=Morpho;
```

```
By Dieta;
```

```
Run;
```

```
Proc Lifereg Data=Morpho;
```

```
Class Dieta;
```

```
Model Duración*Status (0) = Dieta Dietdur/Dist=Llogistic;
```

Probplot;

By Dieta;

Run;

### **Generación de gráficas de la distribución Llogística para DIETAS como función del tiempo (DURACIÓN)**

```
Symbol1 Value=STAR C=Black H=0.15in I=None R=1;
```

```
Proc Lifereg Data=Morpho;
```

```
Class Dieta;
```

```
Model Duración*Status (0) = Dieta Dietdur/Dist=Llogistic;
```

```
Probplot Cencolor = Red Cframe = Ligr Cfit = Blue Ppout Npintervals=Simul;
```

```
Inset / Cfill = White Ctext = Blue;
```

```
BY DIETA;
```

```
Run;
```

### **Resultados.**

#### **Análisis no paramétrico**

A continuación, los resultados de tablas de sobrevivencia por el método Kaplan – Meier para dietas, donde cada línea corresponde a uno de los 140 casos relacionados con la metamorfosis del insecto, partiendo de la muestra inicial de 20 larvas, arreglados en orden ascendente, con excepción de la primera línea que es referida al tiempo 0. Las observaciones censuradas aparecen marcadas con asterisco. La segunda columna, llamada *Survival* es la principal, la cual indica el estimado valor de KM. Por ejemplo, el día 12 el valor de KM para la Dieta 1 (*Arachis pintoi*) es de 0,1981 y para la Dieta 2 (*Platymiscium stipulare*) de 0,1477. Esto significa que la probabilidad estimada de sobrevivencia para que un individuo sobreviva hasta el día 12 o más, es mayor para la Dieta 1. Después de 21 días, correspondiente al mayor tiempo de sobrevivencia en la Dieta 1, los valores de KM se hacen indefinidos, mientras que

para la Dieta 2 estos valores indefinidos de KM son a partir del día 27. La quinta columna, etiquetada como *Number Failed*, se refiere al número acumulado de casos que experimentaron eventos. La columna *Number Left* es el número de casos que no experimentaron eventos o haber sido censurados en cada punto en el tiempo. Significa el tamaño del riesgo para cada punto en el tiempo.

### a. Resultados de dietas alimenticias

#### a.1. Análisis por método Kaplan – Meier para dietas:

Por razones de espacio, se obviarán algunas salidas de resultado por estrato (Dieta 1 y Dieta 2) correspondiente a los estimados de sobrevivencia producto – limite (Product – Limit Survival Estimates), se incluyen percentiles (25 – 50 – 75 %) y valores de la media para cada dieta (Strato).

##### a.1.1. Análisis por método Kaplan – Meier para Dieta 1:

#### The LIFETEST Procedure

#### Stratum 1: DIETA = 1

#### Product-Limit Survival Estimates

DURACIÓN	Survival	Failure	Survival Standard Error	Number Failed	Number Left
0.0000	1.0000	0	0	0	140
0.0000*	.	.	.	0	139
0.0000*	.	.	.	0	138
0.0000*	.	.	.	0	137
-----					
12.0000	0.1981	0.8019	0.0387	85	21
13.0000	.	.	.	86	20
-----					
18.0000	0.0377	0.9623	0.0185	102	4
19.0000	.	.	.	103	3
19.0000	0.0189	0.9811	0.0132	104	2
20.0000	0.00943	0.9906	0.00939	105	1
21.0000	0	1.0000	0	106	0

**The LIFETEST Procedure**

**Stratum 1: DIETA = 1**

Summary Statistics for Time Variable DURACIÓN

Quartile Estimates

Percent	Point Estimate	95% Confidence Interval	
		[Lower	Upper)
75	12.0000	11.0000	13.0000
50	8.0000	7.0000	10.0000
25	7.0000	6.0000	7.0000
Mean		Standard Error	
9.3113		0.4143	

**a.1.2. Análisis por método Kaplan – Meier para Dieta 2:**

**Stratum 2: DIETA = 2**

Product-Limit Survival Estimates

DURACIÓN	Survival	Failure	Survival Standard Error	Number Failed	Number Left
0.0000	1.0000	0	0	0	140
0.0000*	.	.	.	0	139
0.0000*	.	.	.	0	138
0.0000*	.	.	.	0	137
-----					
12.0000	0.1477	0.8523	0.0378	75	13
13.0000	.	.	.	76	12
13.0000	0.1250	0.8750	0.0353	77	11
-----					
18.0000	0.0795	0.9205	0.0288	81	7
19.0000	0.0682	0.9318	0.0269	82	6
20.0000	0.0568	0.9432	0.0247	83	5
22.0000	.	.	.	84	4
22.0000	0.0341	0.9659	0.0193	85	3
24.0000	.	.	.	86	2
24.0000	0.0114	0.9886	0.0113	87	1
27.0000	0	1.0000	0	88	0

## The LIFETEST Procedure

## Stratum 2: DIETA = 2

## Summary Statistics for Time Variable DURACIÓN

## Quartile Estimates

Percent	Point Estimate	95% Confidence Interval	
		[Lower	Upper)
75	11.0000	10.0000	12.0000
50	8.0000	7.0000	9.0000
25	7.0000	6.0000	7.0000
	Mean	Standard Error	
	9.3636	0.5405	

**a.1.3. Análisis conjunto de resultados del método Kaplan - Meier (KM) para ambas Dietas.**

Los percentiles 25<sup>th</sup> y 50<sup>th</sup> (7 y 8 para ambas dietas) son los tiempos menores de los eventos que la probabilidad de sobrevivencia antes del evento es mayor que 0,25 y 0,50. De mayor interés es el percentil 50<sup>th</sup>, referido como la mediana del tiempo de duración durante el crecimiento, indica el momento que en cada dieta sobrevive el 50 % de los individuos. En ambas dietas es de 8 días, con intervalos de confianza del 95% de 7 – 10 días para la Dieta 1 y de 7 – 9 días para la Dieta 2.

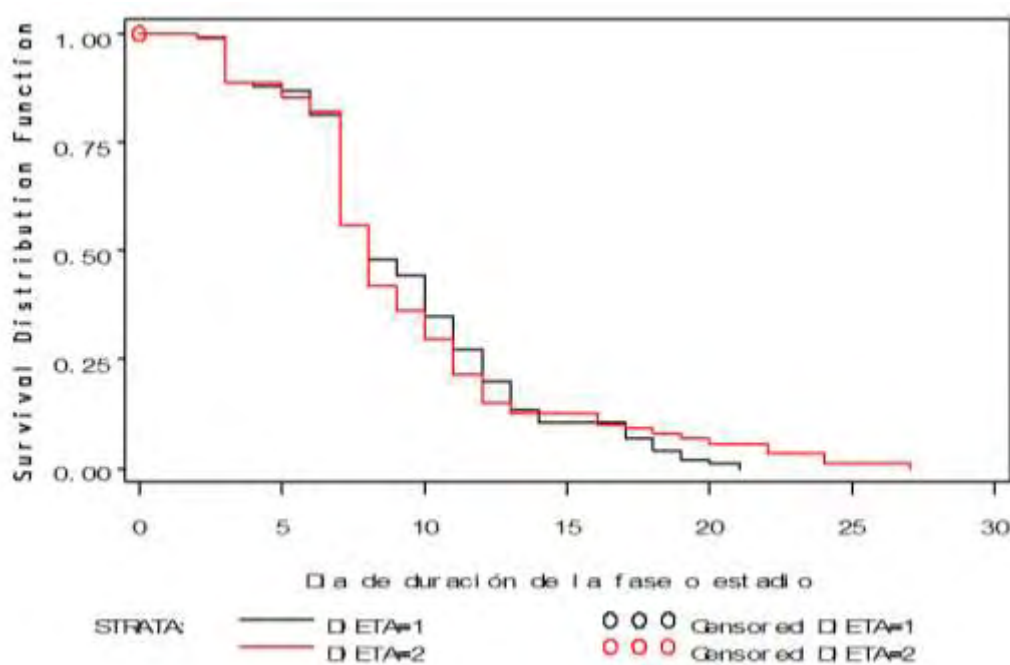
El percentil 75<sup>th</sup> es mayor para la Dieta 1 que para la Dieta 2 (12 versus 11), son los tiempos mayores de los eventos que la probabilidad de sobrevivir antes del evento es mayor de 0,75. También se registra la media estimada del tiempo de duración, que para la Dieta 1 es de 9,3113 días y para la Dieta 2 de 9,3636 días, lo cual indica que los insectos alimentados con la Dieta 1 tienen menor promedio de sobrevivencia (completan el ciclo en menor tiempo). La media tiene predisposición a bajar cuando existen tiempos de censura mayores que el tiempo mayor del evento. Consecuentemente, la mediana es una mejor medida de tendencia central para datos de sobrevivencia censurados (Allison, 2010).

En la salida de resultados siguiente (*Summary of the Number of Censored and Uncensored Values*) se observa que, en el caso de las dietas, de las 140 evaluaciones realizadas en la DIETA 1 a los diferentes estadios, 34 de ellas fueron censuradas (no completaron el ciclo siguiente), lo que representa un 24,3% de censuradas. Para la DIETA 2, 52 de ellas no completaron el estadio, lo que representa el 37,1%.

Summary of the Number of Censored and Uncensored Values

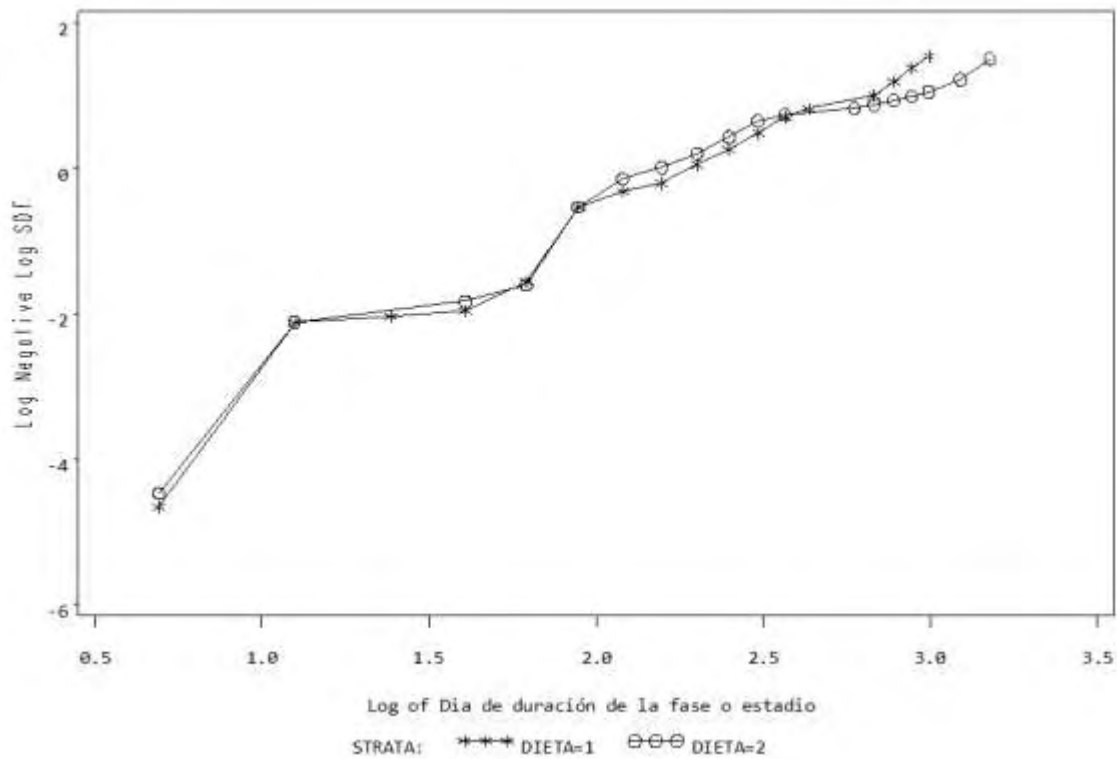
Stratum	DIETA	Total	Failed	Censored	Percent Censored
1	1	140	106	34	24.29
2	2	140	88	52	37.14
-----					
Total		280	194	86	30.71

A continuación, se muestran las gráficas con los valores estimados de la función de sobrevivencia, producto de los métodos KM y LT. Antes de los 8 días, las dos curvas de sobrevivencia de ambas dietas, son indistinguibles, muy parecidas y con poca evidencia de un efecto de las dietas sobre la sobrevivencia. La diferencia, favorable para la Dieta 1, se observa a partir de este tiempo hasta los 12 días, aproximadamente. A los 14 días, los efectos son más favorables para la Dieta 2, hasta el final del ensayo. Se observa que el ciclo de vida de las especies alimentadas con Dieta 1 culmina a los 21 días, mientras que bajo Dieta 2 el ciclo se prolonga hasta 27 días (Figura 29).



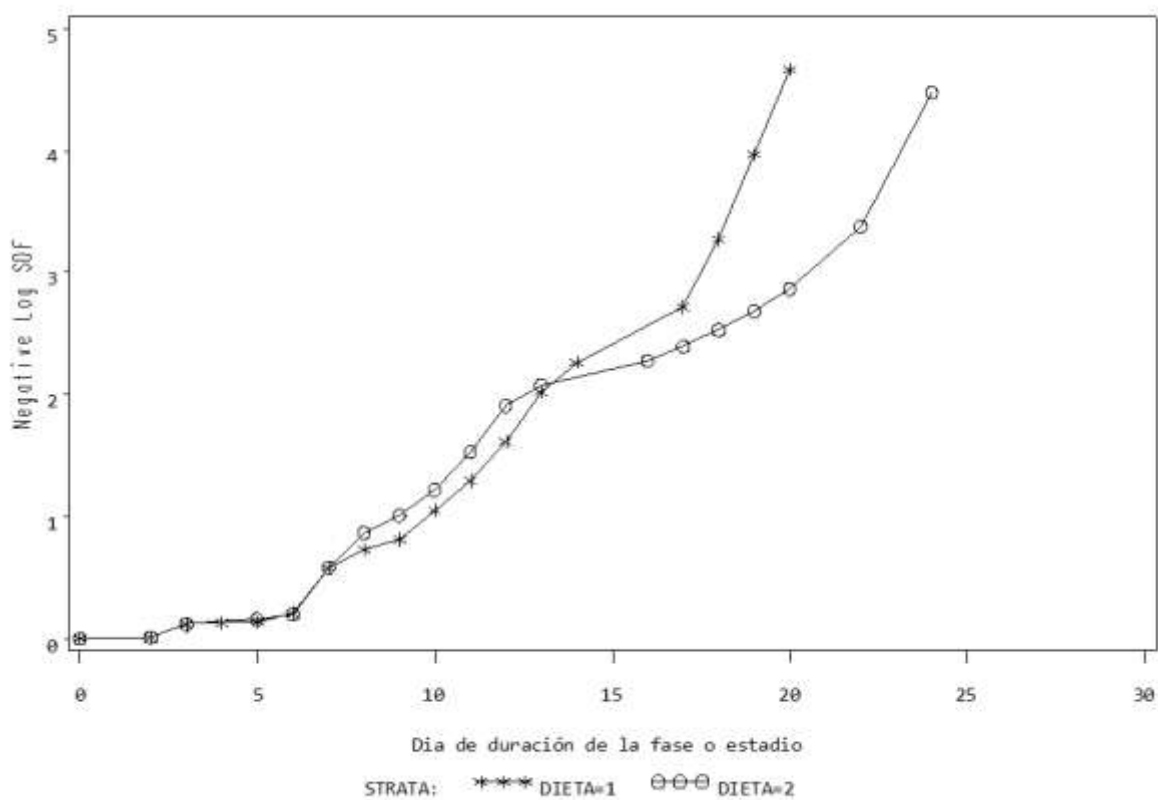
**Figura 29.** Estimados de la función de sobrevivencia por método Kaplan-Meier para dietas alimenticias con el conjunto de fases de crecimiento y total de individuos de *M. menelaus occidentalis* y *M. helenor theodorus* bajo condiciones de cautiverio. Fuente. Elaboración propia.

La opción de PLOT= (LLS) origina la Figura 30 para tratar de establecer un método gráfico común para probar el supuesto PH. Consiste en trazar  $-\log(-\log(S(t/x)))$  versus  $\log(t)$  para diferentes valores de la covariable DIETAS (restringido a valores de  $t$  tal que  $0 < S(t) < 1$ ). Bajo suposición PH, diferentes valores de covariables producirán funciones con la misma forma, pero diferentes elevaciones. El modelo de Cox es robusto, por lo que solo es necesario preocuparse sobre claras desviaciones de la proporcionalidad, indicado por el cruce de las funciones para dos o más covariables en el gráfico de diagnóstico.



**Figura 30.** Funciones de supervivencia  $\log(-\log S(t))$  vs  $\log t$  para dietas alimenticias de *M. menelaus occidentalis* y *M. helenor theodorus* bajo condiciones de cautiverio, estimadas por estratos (STRATA). Fuente. Elaboración propia.

En la Figura 30, se muestran varios cruces decisivos en las curvas de sobrevivencia entre DIETA 1 y DIETA 2, lo que sugiere que la suposición de PH es probablemente violada para ambos grupos. La opción de graficación PLOTS= (LLS), originaría una línea recta si la verdadera distribución es Weibull. Por el contrario, PLOTS= (LS), produce una gráfica de  $-\log S(t)$  versus  $t$ . Si la verdadera distribución es exponencial, esta gráfica debe originar una línea con origen en cero (Figura 31), y puede ayudar en la selección de la distribución para analizar los datos con PROC LIFEREG (Allison, 2010).



**Figura 31. Funciones de sobrevivencia  $-\log S(t)$  versus  $t$  para dietas alimenticias estimadas por estratos (STRATA) de *M. menelaus occidentalis* y *M. helenor theodorus*, bajo condiciones de cautiverio. Fuente. Elaboración propia.**



## a.2. Análisis por método tabla de vida o actuarial (LT) para dietas

Una de las formas de presentar los datos de sobrevivencia es utilizar el método de tablas de vida (LT), en el cual los eventos del tiempo son agrupados en intervalos. Además, este método puede producir estimados y gráficas para la función de riesgo. PROC LIFETEST construye cinco intervalos, comenzando en 0 y se incrementa por periodos de 5 días. Esto se puede obviar especificando WIDTH=W, y luego los intervalos comenzaran con (0, w) incrementándose por el valor de w. También puede elegir INTERVALS=a b c ... donde a, b, y c son puntos de corte. Por ejemplo, INTERVALS= 15 20 30 50 produce los intervalos [0, 15), [15, 20), [20, 30), [30, 50), [50, ∞). A continuación, los resultados para ambas dietas.

Para cada intervalo, se calculan 14 parámetros estadísticos (Allison, 2010). Los primeros cuatro se requieren para calcular los demás estadísticos. El tamaño efectivo de muestra (*Effective Sample Size*) es directo para los últimos cuatro o cinco intervalos, ya que ellos contienen casos no censurados en ambas dietas. El tamaño efectivo de muestra para el primer intervalo se refiere al número de individuos (insectos) no censurados o que completaron el ciclo (106 vs 88), lo que significa que los tamaños de muestra son de 123 y 114 durante los primeros 5 días. La probabilidad condicional de falla (*Conditional Probability of Failure*) es un estimado de la probabilidad de que un insecto mude o cambie de estadio en el intervalo. Se calcula como *number failed / effective sample size*. Un valor estimado del error standard (*Conditional Probability Standard Error*) aparece en la siguiente columna. De esta manera, la probabilidad condicional de muda para ambas Dietas, durante el intervalo 10 – 15 días, fue de  $0,766 \pm 0,0618$  para Dieta 1 y de  $0,6563 \pm 0,084$  para Dieta 2.

La columna *Survival* es el estimado de la tabla de vida de la función de sobreviviente, que significa la probabilidad de que ocurra el evento de muda en un tiempo mayor o igual que el tiempo al inicio de cada intervalo. Por ejemplo, para el intervalo 10 – 15 días, esta probabilidad fue de  $0,4520 \pm 0,0484$  y de  $0,3743 \pm 0,0520$  para Dieta 1 y Dieta 2, respectivamente.

Continuando con los resultados de los 14 estadísticos, la columna *Failure* es 1 menos *Survival*. La columna del residual de la mediana (*Median Residual Lifetime*), es un estimado del tiempo restante hasta un evento para un individuo que sobrevivió al comienzo de cada intervalo. Para calcular éste es-

tadístico para un intervalo dado, debe ser un intervalo posterior cuya probabilidad de sobrevivencia es menor que la mitad de la probabilidad de sobrevivencia asociada con el intervalo de interes. Esto significa que para el intervalo 0 – 5 días para Dieta 1 este valor fue de  $9,457 \pm 0,5096$  y para Dieta 2 de  $8,8315 \pm 0,4352$  días. A partir de este intervalo comenzó a disminuir, hasta el intervalo 15 – 20 días. Para el último intervalo, en ambas dietas, no se pudo calcular. La columna PDF Standard Error indica los valores estimados de la funcion de densidad de probabilidad en el punto medio del intervalo. De gran interes, es la columna Hazard, la cual da los estimados de la funcion de riesgo evaluados en el punto medio de cada intervalo.

A continuacion se describen los valores estimados para los 14 estadísticos, tanto para la Dieta 1 como para la Dieta 2.

### a.2.1. Análisis por método tabla de vida o actuarial (LT) para Dieta 1:

Funcion de sobrevivencia versus tiempo  
The LIFETEST Procedure  
Stratum 1: DIETA = 1  
Life Table Survival Estimates

Interval [Lower, Upper)	Number Failed	Number Censored	Effective Sample Size	Conditional Probability of Failure	Conditional Probability Standard Error	Survival	Failure
0 5	13	34	123.0	0.1057	0.0277	1.0000	0
5 10	46	0	93.0	0.4946	0.0518	0.8943	0.1057
10 15	36	0	47.0	0.7660	0.0618	0.4520	0.5480
15 20	9	0	11.0	0.8182	0.1163	0.1058	0.8942
20 25	2	0	2.0	1.0000	0	0.0192	0.9808

Evaluated at the Midpoint of the Interval

Interval [Lower, Upper)	Survival Standard Error	Median Residual Lifetime	Median Standard Error	PDF PDF	PDF Standard Error	Hazard Hazard	Hazard Standard Error
0 5	0	9.4570	0.5096	0.0211	0.00554	0.022318	0.00618
5 10	0.0277	5.0694	0.6697	0.0885	0.00967	0.131429	0.01830
10 15	0.0484	3.2639	0.4761	0.0692	0.00929	0.248276	0.03244
15 20	0.0301	3.0556	0.9213	0.0173	0.00551	0.276923	0.06661
20 25	0.0135	.	.	0.00385	0.00269	0.4	0

### a.2.2. Análisis por método tabla de vida o actuarial (LT) para Dieta 2:

Funcion de sobrevivencia versus tiempo  
The LIFETEST Procedure  
Stratum 2: DIETA = 2  
Life Table Survival Estimates

Interval [Lower, Upper)		Number Failed	Number Censored	Effective Sample Size	Conditional Probability of Failure	Conditional Probability Standard Error	Survival	Failure
0	5	10	52	114.0	0.0877	0.0265	1.0000	0
5	10	46	0	78.0	0.5897	0.0557	0.9123	0.0877
10	15	21	0	32.0	0.6563	0.0840	0.3743	0.6257
15	20	5	0	11.0	0.4545	0.1501	0.1287	0.8713
20	25	5	0	6.0	0.8333	0.1521	0.0702	0.9298
25	30	1	0	1.0	1.0000	0	0.0117	0.9883

Evaluated at the Midpoint of the Interval

Interval [Lower,Upper)		Survival Standard Error	Median Residual Lifetime	Median Standard Error	PDF PDF	PDF Standard Error	Hazard Hazard	Hazard Standard Error
0	5	0	8.8315	0.4352	0.0175	0.00530	0.018349	0.005796
5	10	0.0265	4.2391	0.4800	0.1076	0.0106	0.167273	0.022403
10	15	0.0520	3.8095	0.6734	0.0491	0.00927	0.195349	0.037199
15	20	0.0361	5.5000	1.6583	0.0117	0.00507	0.117647	0.050286
20	25	0.0276	3.0000	1.2247	0.0117	0.00507	0.285714	0.089424
25	30	0.0116	.	.	0.00234	0.00233	0.4	0

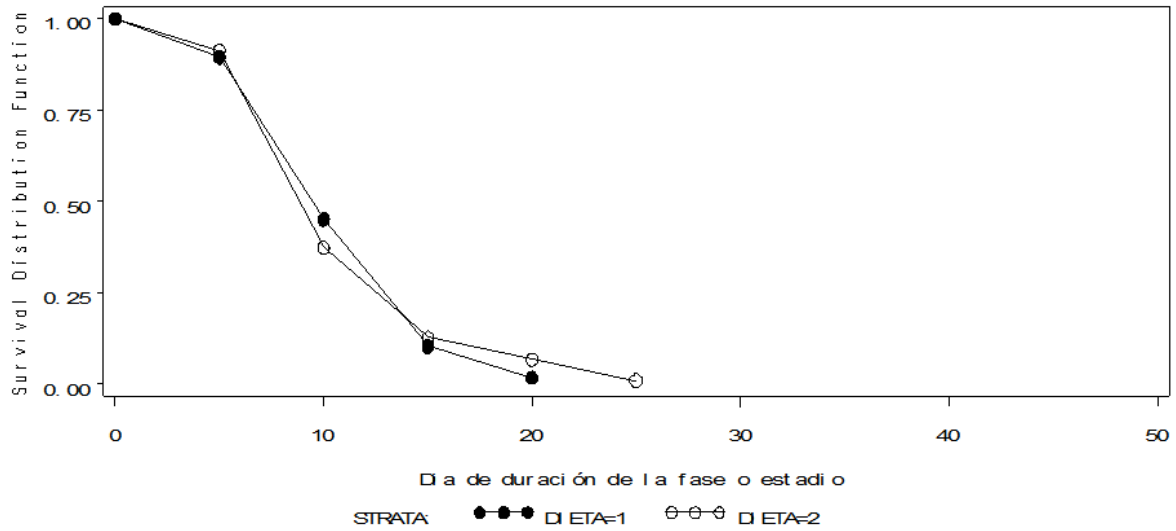
### a.2.3. Análisis conjunto de resultados del método tabla de vida o actuarial (LT) para ambas Dietas.

La probabilidad condicional de falla para la Dieta 1 en el intervalo 10 – 15 días fue de  $0,766 \pm 0,0618$ , que representa la probabilidad estimada de que las especies muden o cambien de estadio en este intervalo. Para la Dieta 2, en el mismo intervalo, fue de  $0,6563 \pm 0,0840$ , lo que significa que en la Dieta 2 las especies presentaron menor probabilidad de mudar en el intervalo. En relación a la probabilidad estimada que un insecto cambie de estadio o muda en un tiempo mayor o igual que el tiempo al inicio del intervalo entre 10 – 15 días (Survival), es de  $0,4520 \pm 0,0484$  para los insectos sometidos a la Dieta 1 y, de  $0,3743 \pm 0,0520$  para la Dieta 2 (Figura 32). La columna del residual de la mediana (Me-

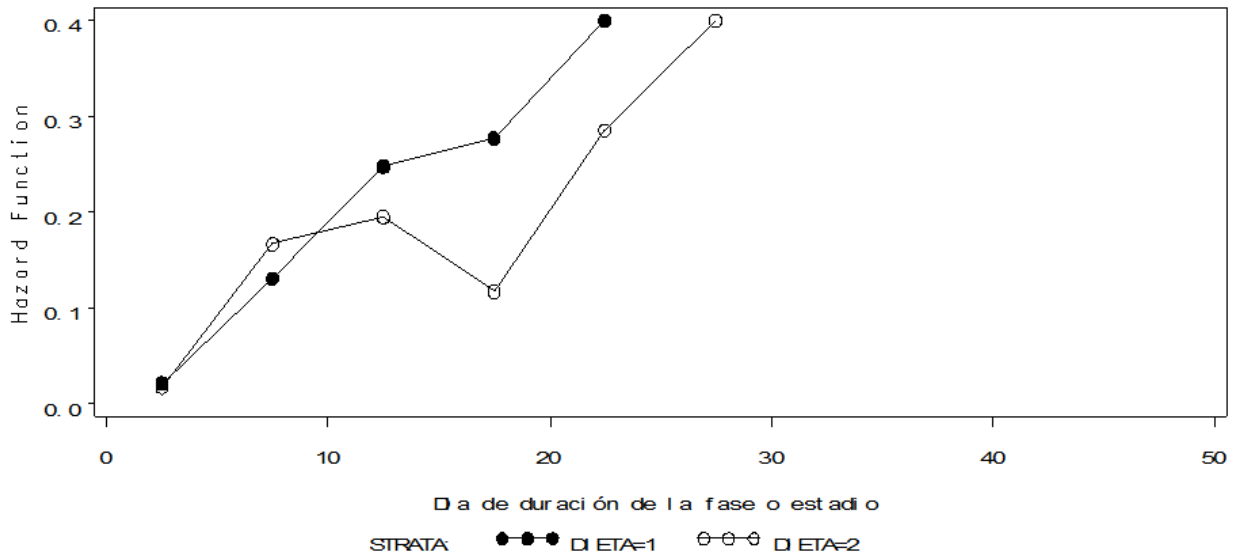
dian Residual Lifetime), como un estimado del tiempo restante hasta un evento para un individuo que sobrevivió al comienzo del intervalo 10 – 15 días, resultó de  $3,2639 \pm 0,4761$  para Dieta 1 comparado con  $3,8095 \pm 0,6734$  para Dieta 2. Este valor significa que para la Dieta 1, las especies tienen un desarrollo en menor tiempo. La columna Hazard, la cual da los estimados de la función de riesgo evaluados en el punto medio de cada intervalo, para el intervalo 10 – 15 días, bajo Dieta 1 el riesgo de mudar representó  $0,248276 \pm 0,032444$  mudas/día y para Dieta 2,  $0,195349 \pm 0,037199$  mudas/día.

En la Dieta 1, se experimentó un aumento de la función de riesgo desde el inicio del experimento. En ella se observa que se incrementa de forma lineal en el intervalo 0 – 10 días, pasando de 0,0223 mudas/día a 0,1314 mudas/día, para continuar con un aumento creciente en los intervalos de tiempo posteriores, duplicando la falla de riesgo en el intervalo 15 – 20 días (0,2769 mudas/día). En la Dieta 2, también ocurre un aumento de la función de riesgo desde el inicio del experimento. En ella se observa que se incrementa de forma lineal en el intervalo 0 – 15 días, pasando de 0,0183 mudas/día en el primer intervalo a 0,1953 mudas/día en el intervalo 10 – 15 días. Luego se presenta un decrecimiento de 0,117 mudas/día en el intervalo 15 – 20 días, pasando a casi triplicar la función de riesgo a 0,2857 mudas/día en el intervalo 20 – 25 días.

Ambas especies criadas bajo la Dieta 2, se mantuvieron en desarrollo en el intervalo 25 – 30 días, lo que significa un tiempo de desarrollo más lento, experimentando una probabilidad de sobrevivencia de  $0,0117 \pm 0,0$ . Las curvas de supervivencia bajo las dos condiciones de alimentación no fueron diferentes (Wilcoxon:  $\chi^2 = 0,1791$ ;  $gl = 1$ ,  $p = 0,8005$ ). Bajo condiciones de alimentación con Dieta 1 se tuvieron las más altas probabilidades de riesgo de sobrevivencia, experimentando un riesgo lineal y progresivo en el tiempo. No se presentó esta característica para las especies alimentadas con Dieta 2, con disminución de la probabilidad de sobrevivencia en el intervalo 10 – 20 días, con aumento en los intervalos siguientes (Figura 33).



**Figura 32. Cambios en la supervivencia de ambas especies durante todas las etapas de desarrollo, criadas y mantenidas bajo cada dieta alimenticia.** Fuente. Elaboración propia.



**Figura 33. Fluctuación en los riesgos de mortalidad de ambas especies criadas con Dieta 1 (círculo negro) y Dieta 2 (círculo blanco) durante todas las etapas de desarrollo.** Fuente. Elaboración propia.

### a.3. Pruebas estadísticas para la comparación entre dietas alimenticias

Luego del resultado de los 14 estadísticos anteriores, PROC LIFETEST imprime los 6 estadísticos solicitados con la opción TEST=ALL (log-rank, Wilcoxon, Tarone-Ware, Peto-Peto, Peto-Peto modificado y Fleming-Harrington) para cada tratamiento o Dieta, seguido de los estimados de la matriz de covarianza. Estos son utilizados para calcular el estadístico Chi cuadrada.

Funcion de sobrevivencia versus tiempo  
The LIFETEST Procedure  
Testing Homogeneity of Survival Curves for DURACIÓN over Strata  
Rank Statistics

DIETA	Log-Rank	Wilcoxon	Tarone	Peto	Modified Peto	Fleming
1	1.5327	-326.00	-19.47	-1.3838	-1.3888	-1.6804
2	-1.5327	326.00	19.47	1.3838	1.3888	1.6804

Por ejemplo, el estadístico *Long – Rank* para la Dieta 1 es 1,5327. Para calcular el valor de Chi cuadrada, este valor se eleva al cuadrado y se divide por la varianza estimada ( $1,5327^2 = 2,3492/36,79 = 0,06385$ ). Así se obtiene que el valor de Chi cuadrada para este estadístico es de, 0,06385. Igual procedimiento se cumple con los demás estadísticos. De esta manera, se prueban las diferencias entre dos grupos (Dieta 1 y Dieta 2). En caso que la hipótesis nula sea verdadera, todas las pruebas estadísticas tienen una distribución de Chi cuadrada con grados de libertad igual al número de grupos menos 1.

#### Covariance Matrix for the Log-Rank Statistics

DIETA	1	2
1	36.7876	-36.7876
2	-36.7876	36.7876

#### Covariance Matrix for the Wilcoxon Statistics

DIETA	1	2
1	593339	-593339
2	-593339	593339

#### Covariance Matrix for the Tarone Statistics

DIETA	1	2
1		
2		

1	4141.87	-4141.87
2	-4141.87	4141.87

## Covariance Matrix for the Peto Statistics

DIETA	1	2
1	10.8174	-10.8174
2	-10.8174	10.8174

## Covariance Matrix for the Modified Peto Statistics

DIETA	1	2
1	10.6713	-10.6713
2	-10.6713	10.6713

## Covariance Matrix for the Fleming Statistics

DIETA	1	2
1	15.7652	-15.7652
2	-15.7652	15.7652

## Test of Equality over Strata

Test	Chi-Square	DF	Pr > Chi-Square
Log-Rank	0.0639	1	0.8005
Wilcoxon	0.1791	1	0.6721
Tarone	0.0915	1	0.7623
Peto	0.1770	1	0.6739
Modified Peto	0.1808	1	0.6707
Fleming(1)	0.1791	1	0.6721

Finalmente, se observa que los seis estadísticos y sus valores de Chi cuadrada están lejos de ser significativos (valores de p no significativos), lo que sugiere que existe evidencia para aceptar la hipótesis nula que las dos DIETAS tienen la misma función de sobrevivencia, o sea, exactamente la misma distribución de probabilidad para el evento en el tiempo. En esta última tabla, al observar el valor de la prueba *Log Rank* con valor de  $p = 0,8005$ , mayor que el valor predefinido de  $0,05$ , sugiere que no existe evidencia de diferencias significativas entre las dos curvas de sobrevivencia.

Para tiempos discretos (días, meses, años), la tasa de riesgo es la probabilidad que un individuo experimente un evento en el tiempo  $t$ , momento de riesgo para tener el evento, en el caso de los insectos, etapas de la metamorfosis. Si la tasa de riesgo es constante en el tiempo, por ejemplo, igual a 1,5 indica

que se debe esperar 1,5 eventos para que ocurra en un intervalo de tiempo igual. Además, si un organismo tuvo una tasa de riesgo de 1,2 en tiempo  $t$  y un segundo organismo tuvo una tasa de 2,4 en tiempo  $t$ , la tasa de riesgo de este organismo es dos veces mayor de experimentar un evento. La tasa de riesgo es una variable no observada que controla la ocurrencia y tiempo del evento, por lo tanto, es la variable fundamental del análisis de sobrevivencia.

## b. Resultados para la fase de larva

### b.1. Análisis por método Kaplan – Meier para larva:

#### b.1.1. Análisis para Dieta 1:

**Stratum 1: FASE = LARVA**

Summary Statistics for Time Variable DURACIÓN

Quartile Estimates

Percent	Point Estimate	95% Confidence Interval	
		[Lower	Upper)
75	11.0000	10.0000	14.0000
50	8.0000	7.0000	10.0000
25	7.0000	.	.
	Mean	Standard Error	
	9.9125	0.4473	

#### b.1.2. Análisis para Dieta 2:

**Stratum 1: FASE = LARVA**

Summary Statistics for Time Variable DURACIÓN

Quartile Estimates

Percent	Point Estimate	95% Confidence Interval	
		[Lower	Upper)
75	10.0000	9.0000	13.0000
50	8.0000	7.0000	9.0000
25	7.0000	.	.
	Mean	Standard Error	
	9.9853	0.6175	



## b.2. Análisis por método tabla de vida o actuarial (LT) para larva:

### b.2.1. Análisis para Dieta 1:

Funcion de sobrevivencia versus tiempo

The LIFETEST Procedure

Stratum 1: FASE = LARVA

Life Table Survival Estimates

Interval [Lower, Upper)	Number Failed	Number Censored	Effective Sample Size	Conditional Probability of Failure	Conditional Probability Standard Error	Survival	Failure
0 5	0	20	90.0	0	0	1.0000	0
5 10	46	0	80.0	0.5750	0.0553	1.0000	0
10 15	23	0	34.0	0.6765	0.0802	0.4250	0.5750
15 20	9	0	11.0	0.8182	0.1163	0.1375	0.8625
20 25	2	0	2.0	1.0000	0	0.0250	0.9750

Evaluated at the Midpoint of the Interval

Interval [Lower,Upper)	Survival Standard Error	Median Residual Lifetime	Median Standard Error	PDF PDF	PDF Standard Error	Hazard Hazard	Hazard Standard Error
0 5	0	9.3478	0.4583	0	.	0	.
5 10	0	4.3478	0.4861	0.1150	0.0111	0.16140	0.02177
10 15	0.0553	3.6957	0.6338	0.0575	0.0101	0.20444	0.03664
15 20	0.0385	3.0556	0.9213	0.0225	0.0071	0.27692	0.06661
20 25	0.0175	.	.	0.0050	0.0035	0.4	0

**b.2.2. Análisis para Dieta 2:**

Funcion de sobrevivencia versus tiempo  
The LIFETEST Procedure  
**Stratum 1: FASE = LARVA**  
Life Table Survival Estimates

Interval [Lower, Upper)	Number Failed	Number Censored	Effective Sample Size	Conditional Probability of Failure	Conditional Probability Standard Error	Survival	Failure	
0	5	0	32	84.0	0	1.0000	0	
5	10	46	0	68.0	0.6765	0.0567	1.0000	0
10	15	11	0	22.0	0.5000	0.1066	0.3235	0.6765
15	20	5	0	11.0	0.4545	0.1501	0.1618	0.8382
20	25	5	0	6.0	0.8333	0.1521	0.0882	0.9118
25	30	1	0	1.0	1.0000	0	0.0147	0.9853

Evaluated at the Midpoint of the Interval

Interval [Lower,Upper)	Survival Standard Error	Median Residual Lifetime	Median Standard Error	PDF	PDF Standard Error	Hazard	Hazard Standard Error	
0	5	0	8.6957	0.4032	0	.	0	
5	10	0	3.6957	0.4482	0.1353	0.0113	0.2044	0.025909
10	15	0.0567	5.0000	1.0660	0.0324	0.0089	0.1333	0.037902
15	20	0.0447	5.5000	1.6583	0.0147	0.0063	0.1176	0.050286
20	25	0.0344	3.0000	1.2247	0.0147	0.0063	0.2857	0.089424
25	30	0.0146	.	.	0.0029	0.0029	0.4	0

**b.2.3. Análisis conjunto de resultados de los métodos Kaplan-Meier (KM) y tabla de vida o actuarial (LT) para la fase larva.**

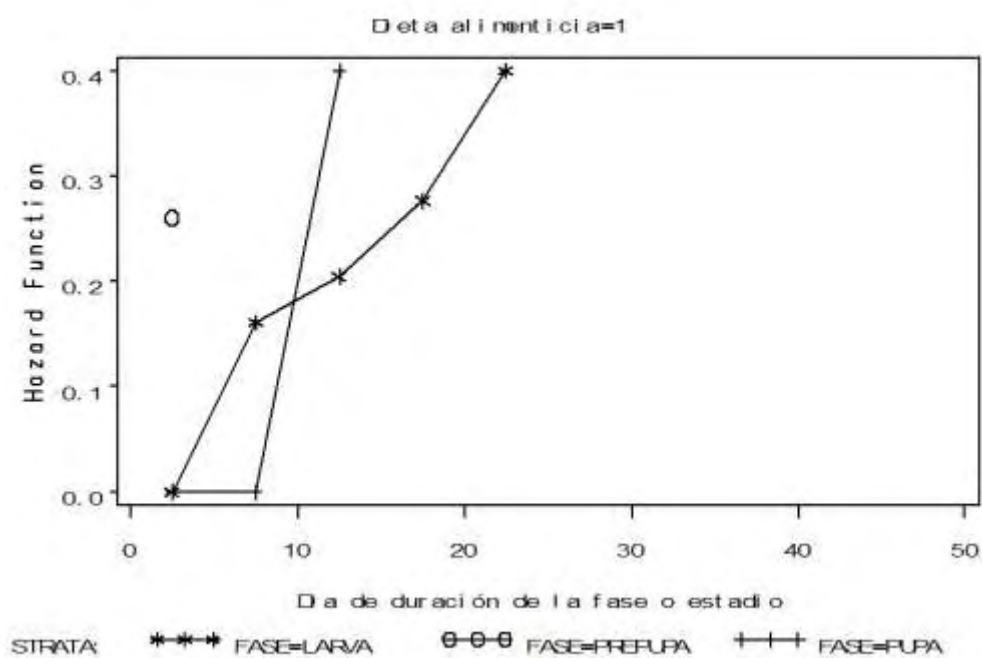
En el caso de las larvas, alimentadas con la DIETA =1, el percentil 50<sup>th</sup> (mediana del tiempo de sobrevivencia), el valor es de 8 días, con intervalos de confianza del 95% de 7 – 10 días. El percentil 75<sup>th</sup>, el cual representa que la probabilidad de sobrevivir al evento antes, es mayor de 0,75, tiene un valor promedio de 11 días, con un intervalo de confianza entre 10 – 14 días. La media estimada del tiempo de duración es de 9,9125 días.

Las larvas, alimentadas con la DIETA=2, el percentil 50<sup>th</sup> (mediana del tiempo de sobrevivencia), el valor es de 8 días, con intervalos de confianza del 95% de 7 – 9 días. El percentil 75<sup>th</sup> tiene un valor promedio de 10 días, con un intervalo de confianza entre 9 – 13 días. La media estimada del tiempo de duración en la fase larval es de 9,9853 días. En consecuencia, al comparar ambas dietas en el desarrollo larval, su tiempo promedio de duración es menor en las larvas alimentadas con la DIETA=1. Sin embargo, las diferentes pruebas de chi cuadrado indican que no existen diferencias significativas entre ambas dietas:

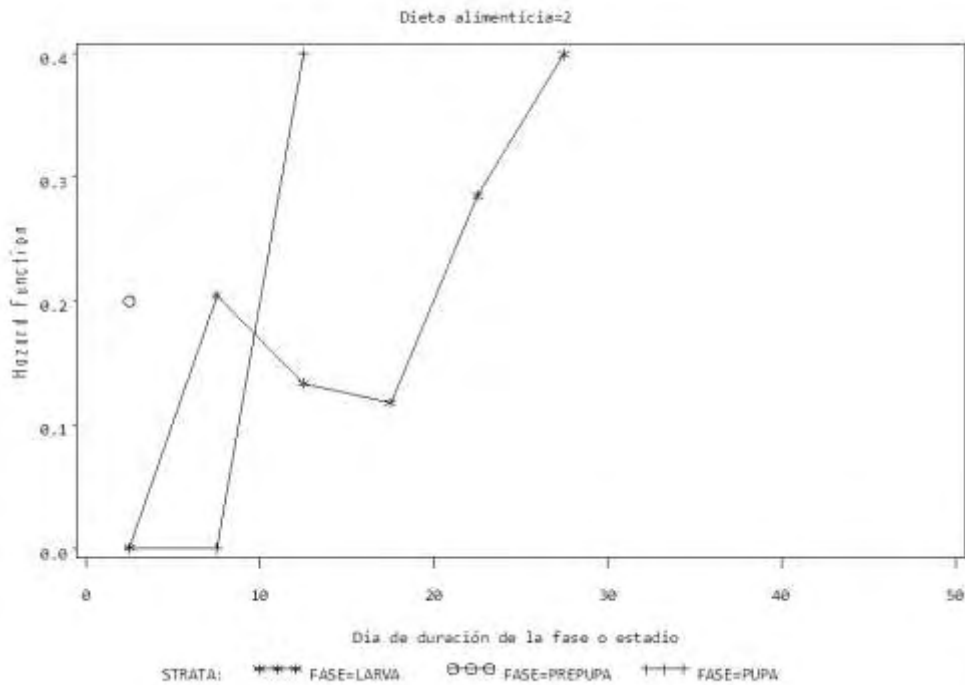
**Test of Equality over Strata**

Test	Chi-Square	DF	Pr > Chi-Square
Log-Rank	0.0639	1	0.8005
Wilcoxon	0.1791	1	0.6721
-2Log(LR)	0.0015	1	0.9690

Las Figuras 34 y 35 corresponden a las gráficas de la función de riesgo de experimentar los eventos de la metamorfosis para las tres fases (larva, prepupa y pupa), para DIETA 1 y DIETA 2, respectivamente. En la Dieta 1, se experimentó un aumento de la función de riesgo desde el inicio del experimento. En ella se observa que se incrementa de forma lineal en el intervalo 0 – 10 días, pasando de 0,1614 mudas/día a 0,2044 mudas/día en el intervalo 10-15 días, para continuar con un aumento creciente en los intervalos de tiempo posteriores, duplicando la falla de riesgo en el intervalo 20 – 25 días (0,400 mudas/día). En la Dieta 2, ocurre una disminución de la función de riesgo desde el inicio del experimento. En ella se observa que en el intervalo 0 – 15 días, pasa de 0,2044 mudas/día a 0,1333 mudas/día en el intervalo 10 – 15 días y a 0,1176 mudas/día entre 15 – 20 días. Luego presenta un crecimiento durante los últimos intervalos del desarrollo, de 0,2857 mudas/día en el intervalo 20 – 25 días, a 0,400 mudas/día en el intervalo 25 – 30 días. En general, las larvas bajo DIETA 1 experimentaron un riesgo ascendente durante el experimento, mientras que bajo DIETA 2, experimentaron decrecimiento en el periodo o intervalo entre 5 – 20 días, a partir del cual asciende el riesgo hasta el final.



**Figura 34.** Funcion de riesgo para la Dieta 1 para las fases de crecimiento de *M. menelaus occidentalis* y *M. helenor theodorus* bajo condiciones de cautiverio. Fuente. Elaboración propia.



**Figura 35. Funcion de riesgo para la Dieta 2 para las fases de crecimiento de *M. menelaus occidentalis* y *M. helenor theodorus* bajo condiciones de cautiverio. Fuente. Elaboración propia.**

La probabilidad condicional de falla para larva bajo la Dieta 1 en el intervalo 10 – 15 días fue de  $0,6765 \pm 0,0802$ , que representa la probabilidad estimada de que las especies muden o cambien de estadio en este intervalo. Para las larvas en Dieta 2, en el mismo intervalo, fue de  $0,5000 \pm 0,1066$ , lo que significa que en la Dieta 2 las larvas de ambas especies presentaron menor probabilidad de mudar en el intervalo. En relación a la probabilidad estimada que una larva cambie de estadio o muda en un tiempo mayor o igual que el tiempo al inicio del intervalo entre 10 – 15 días (Survival), es de  $0,4250 \pm 0,0553$  para las larvas sometidas a la Dieta 1 y, de  $0,3235 \pm 0,0567$  para la Dieta 2 (Figuras 34 y 35). La columna del residual de la mediana (Median Residual Lifetime), como un estimado del tiempo restante hasta un evento para una larva que sobrevivió al comienzo del intervalo 10 – 15 días, resulto de  $3,6957 \pm 0,6358$  días para Dieta 1, comparado con  $5,000 \pm 1,0660$  días para Dieta 2. Este valor significa que para la Dieta 1, las especies tienen un desarrollo larval en menor tiempo.

### c. Resultados para la fase de Pupa

#### c.1. Análisis por método Kaplan – Meier para Pupa:

##### c.1.1. Análisis para Dieta 1:

The LIFETEST Procedure  
**Stratum 3: FASE = PUPA**  
 The LIFETEST Procedure  
 Summary Statistics for Time Variable DURACIÓN  
 Quartile Estimates

Percent	Point Estimate	95% Confidence Interval [Lower Upper)	
75	12.0000	12.0000	13.0000
50	12.0000	.	.
25	12.0000	11.0000	12.0000

Mean	Standard Error
11.9231	0.2392

##### c.1.2. Análisis para Dieta 2:

**Stratum 3: FASE = PUPA**  
 The LIFETEST Procedure  
 Summary Statistics for Time Variable DURACIÓN  
 Quartile Estimates

Percent	Point Estimate	95% Confidence Interval [Lower Upper)	
75	12.0000	11.0000	13.0000
50	11.0000	11.0000	12.0000
25	11.0000	11.0000	12.0000

Mean	Standard Error
11.6000	0.2211

En el caso de las pupas, el percentil 50<sup>th</sup> (mediana del tiempo de sobrevivencia), el valor promedio para la DIETA 1 es de 12 días y de 11 días para la DIETA 2, con intervalo de confianza del 95% de 11 – 12 días. El percentil 75<sup>th</sup>, tiene un valor promedio para ambas dietas, de 12 días, con intervalos de confianza del 95% entre 12 – 13 días para la DIETA 1 y entre 11 – 13 días, para la DIETA 2. La media estimada del tiempo de sobrevivencia es de 11, 9231 días para la DIETA 1 y 11,6000 días para la DIETA 2.

## c.2. Análisis por método tabla de vida o actuarial (LT) para Pupa:

### c.2.1. Análisis para Dieta 1:

Funcion de sobrevivencia versus tiempo  
The LIFETEST Procedure  
Stratum 3: FASE = PUPA  
Life Table Survival Estimates

Interval [Lower, Upper)	Number Failed	Number Censored	Effective Sample Size	Conditional Probability of Failure	Conditional Probability Standard Error	Survival	Failure
0 5	0	7	16.5	0	0	1.0000	0
5 10	0	0	13.0	0	0	1.0000	0
10 15	13	0	13.0	1.0000	0	1.0000	0

Evaluated at the Midpoint of the Interval

Interval [Lower, Upper)	Survival Standard Error	Median Residual Lifetime	Median Standard Error	PDF PDF	PDF Standard Error	Hazard Standard Error
0 5	0	.	.	0	.	0
5 10	0	.	.	0	.	0
10 15	0	.	.	0.2000	0	0.4

### c.2.2. Análisis para Dieta 2:

Funcion de sobrevivencia versus tiempo  
The LIFETEST Procedure  
Stratum 3: FASE = PUPA  
Life Table Survival Estimates

Interval [Lower, Upper)	Number Failed	Number Censored	Effective Sample Size	Conditional Probability of Failure	Conditional Probability Standard Error	Survival	Failure
0 5	0	10	15.0	0	0	1.0000	0
5 10	0	0	10.0	0	0	1.0000	0
10 15	10	0	10.0	1.0000	0	1.0000	0

Evaluated at the Midpoint of the Interval

Interval [Lower, Upper)	Survival Standard Error	Median Residual Lifetime	Median Standard Error	PDF PDF	PDF Standard Error	Hazard Standard Error	Hazard Standard Error
0 5	0	.	.	0	.	0	.
5 10	0	.	.	0	.	0	.
10 15	0	.	.	0.2000	0	0.4	0

Las pupas en ambas dietas, experimentan bajo riesgo de mortalidad durante los primeros días de la metamorfosis, incrementándose en etapas posteriores, con un valor de 0,4 mudas/día (Figuras 34 y 35).

En la salida de resultados siguiente (*Summary of the Number of Censored and Uncensored Values*) se observa que, en el caso de las larvas, durante los 5 instares larvales de las 40 larvas iniciales (origina 200 evaluaciones), 52 de ellas fueron censuradas (no completaron el instar siguiente), lo que representa un 26% de censuradas. Para las prepupas, 17 de ellas no completaron el estadio, lo que representa el 42,5%. Igual situación para las pupas.



## Summary of the Number of Censored and Uncensored Values

Stratum	FASE	Total	Failed	Censored	Percent Censored
1	LARVA	200	148	52	26.00
2	PREPUPA	40	23	17	42.50
3	PUPA	40	23	17	42.50
-----					
	Total	280	194	86	30.71

## Funcion de sobrevivencia versus tiempo

## The LIFETEST Procedure

## Testing Homogeneity of Survival Curves for DURACIÓN over Strata

## Rank Statistics

FASE	Log-Rank	Wilcoxon	Tarone	Peto	Modified Peto	Fleming
LARVA	-9.805	-1528.0	-109.8	-8.318	-8.271	-7.876
PREPUPA	20.570	3933.0	284.4	18.425	18.329	20.273
PUPA	-10.764	-2405.0	-174.6	-10.107	-10.058	-12.397

## Covariance Matrix for the Log-Rank Statistics

FASE	LARVA	PREPUPA	PUPA
LARVA	21.9387	-1.7034	-20.2353
PREPUPA	-1.7034	1.9681	-0.2647
PUPA	-20.2353	-0.2647	20.5000

## Covariance Matrix for the Wilcoxon Statistics

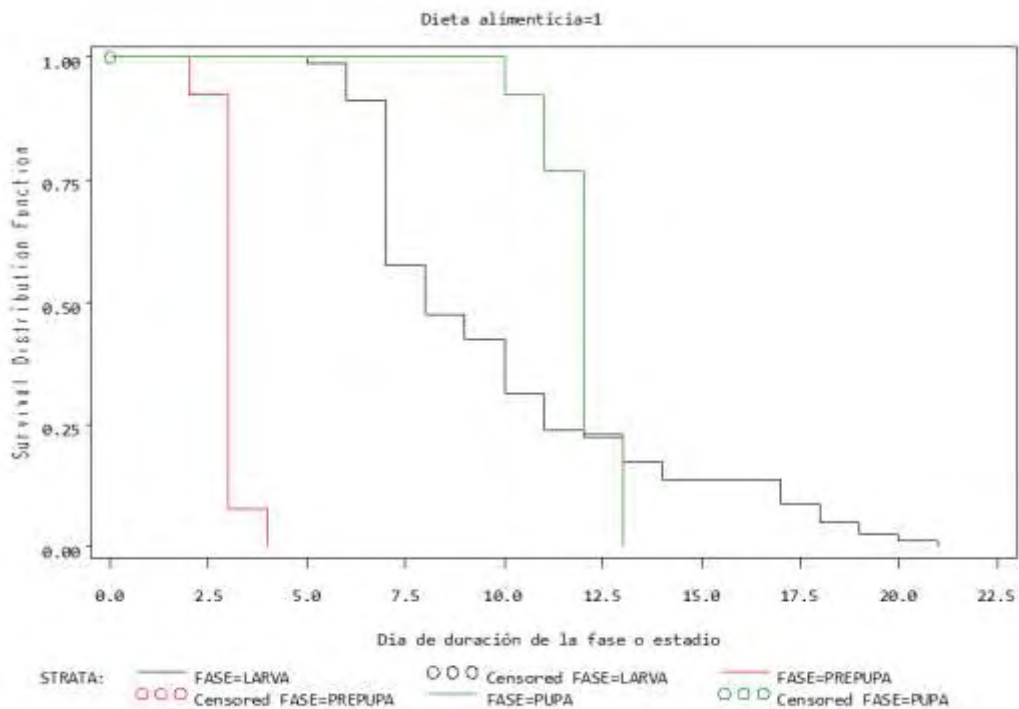
FASE	LARVA	PREPUPA	PUPA
LARVA	360251	-62897	-297354
PREPUPA	-62897	72672	-9775
PUPA	-297354	-9775	307128

Finalmente, en la salida de resultados “*Test of Equality over Strata*” se observa que los seis estadísticos y sus valores de Chi cuadrada son altamente significativos (valores de  $p < 0,0001$ ), lo que sugiere que existe evidencia para soportar el rechazamiento de la hipótesis nula que para las tres fases en con-

junto (larva – prepupa – pupa) no tienen la misma función de sobrevivencia, o sea, no tienen exactamente la misma distribución de probabilidad para el evento en el tiempo en ambas dietas. Por ejemplo, en esta última tabla, al observar el valor de la prueba *Log Rank* que el valor de  $p < 0,0001$ , el cual es menor que el valor predefinido de 0,05, sugiere que existe evidencia de diferencias significativas entre las curvas de sobrevivencia de las tres fases en ambas dietas. Las Figuras 36 y 37 se presentan para separar los efectos de cada dieta alimenticia en los estimados de la función de sobrevivencia de cada fase de crecimiento.

Test of Equality over Strata

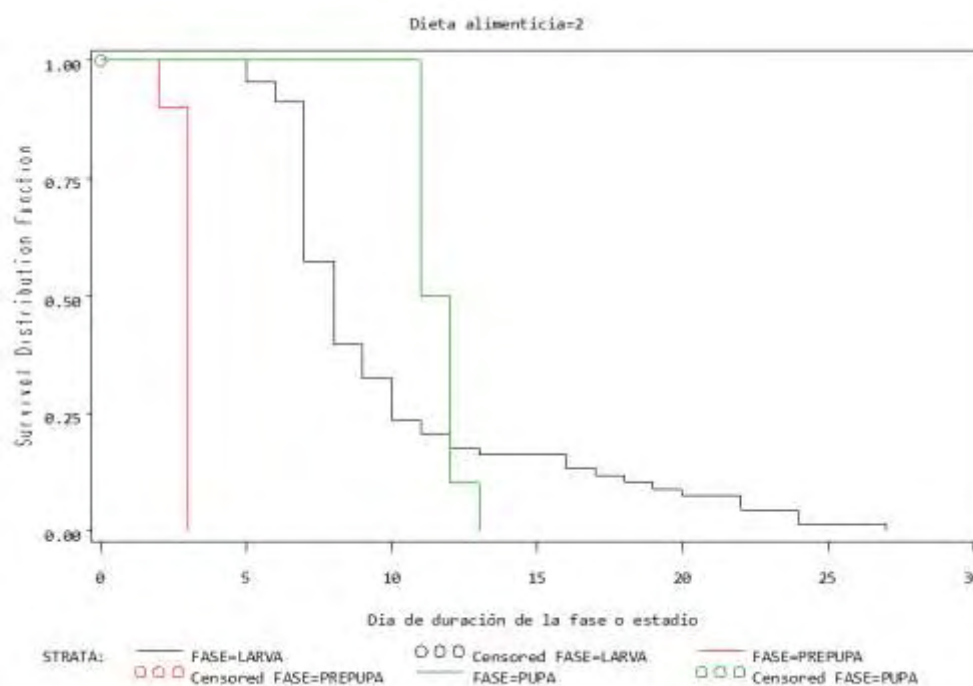
Test	Chi-Square	DF	Pr > Chi-Square
Log-Rank	218.1041	2	<.0001
Wilcoxon	224.3622	2	<.0001
Tarone	222.0287	2	<.0001
Peto	224.2279	2	<.0001
Modified Peto	224.2708	2	<.0001
Fleming(1)	224.3622	2	<.0001



**Figura 36.** Estimados de la función de sobrevivencia por método Kaplan-Meier para individuos alimentados con Dieta 1 y su efecto en las fases de crecimiento de *M. menelaus occidentalis* y *M. helenor theodorus* bajo condiciones de cautiverio. Fuente. Elaboración propia.

Al comparar las Figuras 36 y 37, correspondientes a las funciones de crecimiento, de ambas dietas, se observa que el crecimiento de las larvas tiene mayor duración en la Dieta 2 (el periodo larvario dura más de 25 días) que en la Dieta 1 (el periodo larvario dura menos de 25 días). Por otra parte, la duración de los intervalos de ambas dietas son similares durante los primeros siete días y a partir de aquí, la altura de los escalones o líneas verticales entre intervalos es mayor en la DIETA 1. En relación a las prepupas, en la Dieta 2 se observan dos intervalos de crecimiento, mientras que en la Dieta 1, tres intervalos. Para las pupas se presenta un fenómeno parecido, ya que la Dieta 1 presenta más intervalos que la Dieta 2 (4 vs. 3).

Las distancias verticales entre horizontales representan el cambio en la probabilidad acumulada de no mudar o pasar al intervalo siguiente a medida que avanza la curva. La pendiente de la curva está determinada por la duración de la supervivencia, representada por la longitud de las líneas horizontales.



**Figura 37.** Estimados de la función de supervivencia por método Kaplan-Meier para individuos alimentados con Dieta 2 y su efecto en las fases de crecimiento de *M. menelaus occidentalis* y *M. helenor theodorus* bajo condiciones de cautiverio. Fuente. Elaboración propia.

#### d. Resultados para los estadios larvales.

En razón de que el procedimiento para el análisis de los estadios larvales es idéntico, se procederá obviando los resultados del primer y segundo estadio larval, haciendo el análisis para tercer, cuarto y quinto estadio larval.

## d.1. Análisis por método Kaplan – Meier para tercer estadio larval:

### d.1.1. Análisis para Dieta 1:

Stratum 3: ESTADIO = L3  
The LIFETEST Procedure  
Summary Statistics for Time Variable DURACIÓN  
Quartile Estimates

Percent	Point Estimate	95% Confidence Interval [Lower Upper)	
75	10.0000	9.0000	10.0000
50	8.0000	8.0000	10.0000
25	8.0000	.	.
	Mean	Standard Error	
	8.7037	0.2605	

### d.1.2. Análisis para Dieta 2:

Stratum 3: ESTADIO = L3  
The LIFETEST Procedure  
Summary Statistics for Time Variable DURACIÓN  
Quartile Estimates

Percent	Point Estimate	95% Confidence Interval [Lower Upper)	
75	10.0000	9.0000	10.0000
50	9.0000	8.0000	10.0000
25	8.0000	8.0000	9.0000
	Mean	Standard Error	
	9.0000	0.2928	

En el caso del tercer instar larval, el percentil 50<sup>th</sup>, el valor promedio para la Dieta 1 es de 8 días y de 9 días para la Dieta 2, con un intervalo de confianza de 95% entre 8 y 10 días, para ambas dietas. El percentil 75<sup>th</sup>, tiene un valor promedio para ambas dietas, de 10 días, con un intervalo entre 9 – 10 días. La media estimada del tiempo de sobrevivencia es de 8,7 días para la Dieta 1 y de 9 para la Dieta 2.

**d.2. Análisis por método de tabla de vida o actuarial (LT) para tercer estadio larval**

**d.2.1. Análisis para Dieta 1:**

Funcion de sobrevivencia versus tiempo

The LIFETEST Procedure

Stratum 3: ESTADIO = L3

Life Table Survival Estimates

Interval [Lower, Upper)	Number Failed	Number Censored	Effective Sample Size	Conditional Probability of Failure	Conditional Probability Standard Error	Survival	Failure
0 5	0	5	17.5	0	0	1.0000	0
5 10	9	0	15.0	0.6000	0.1265	1.0000	0
10 15	6	0	6.0	1.0000	0	0.4000	0.6000

Evaluated at the Midpoint of the Interval

Interval [Lower, Upper)	Survival Standard Error	Median Residual Lifetime	Median Standard Error	PDF	PDF Standard Error	Hazard Standard Error
0 5	0	9.1667	0.9960	0	.	0
5 10	0	4.1667	1.0758	0.1200	0.0253	0.17143
10 15	0.1265	.	.	0.0800	0.0253	0.4

**d.2.2. Análisis para Dieta 2:**

Stratum 3: ESTADIO = L3

Life Table Survival Estimates

Interval [Lower, Upper)	Number Failed	Number Censored	Sample Size	Probability of Failure	Standard Error	Survival	Failure
0 5	0	8	16.0	0	0	1.0000	0
5 10	9	0	12.0	0.7500	0.1250	1.0000	0
10 15	3	0	3.0	1.0000	0	0.2500	0.7500

Evaluated at the Midpoint of the Interval								
Interval [Lower, Upper)		Survival Standard Error	Median Residual Lifetime	Median Standard Error	PDF PDF	PDF Standard Error	Hazard Standard Error	Hazard Standard Error
0	5	0	8.3333	0.8333	0	.	0	.
5	10	0	3.3333	0.9623	0.1500	0.0250	0.24	0.064
10	15	0.1250	.	.	0.0500	0.0250	0.4	0

Las larvas del tercer instar también presentan tres intervalos de vida en ambas dietas. La función de riesgo de mudar es menor para la Dieta 1 que para la Dieta 2 en los que experimentan aumento de la función de riesgo después del primer intervalo (0 – 5 días), pasando de 0,0 mudas/día en el primer intervalo, a 0,1714 mudas/día en el segundo intervalo (5 – 10 días), para llegar a 0,4 mudas/día en el tercer intervalo, comprendido entre 10 y 15 días. La Dieta 2 presenta una falla de riesgo de 0,24 mudas/día en el segundo intervalo.

### d.3. Análisis por método de Kaplan - Meier para cuarto estadio larval:

#### d.3.1. Análisis para Dieta 1:

The LIFETEST Procedure				
Stratum 4: ESTADIO = L4				
Summary Statistics for Time Variable DURACIÓN				
Quartile Estimates				
Percent	Point Estimate	95% Confidence Interval [Lower Upper)		
75	13.0000	11.0000	13.0000	
50	11.0000	10.0000	12.0000	
25	10.0000	10.0000	11.0000	
	Mean	Standard Error		
	11.2667	0.3712		

### d.3.2. Análisis para Dieta 2

Stratum 4: ESTADIO = L4 Summary Statistics  
for Time Variable DURACIÓN  
Quartile Estimates

Percent	Point Estimate	95% Confidence Interval [Lower Upper)	
75	12.0000	10.0000	16.0000
50	10.0000	9.0000	12.0000
25	9.0000	8.0000	10.0000
	Mean	Standard Error	
	10.9091	0.6803	

En el cuarto instar larval, el percentil 50<sup>th</sup> (mediana del tiempo de sobrevivencia), el valor promedio para la Dieta 1 es de 11 días, con un intervalo de confianza de 95% entre 10 y 12 días. A diferencia de la Dieta 2 de 10 días (9 – 12) El percentil 75<sup>th</sup>, tiene un valor promedio de 13 días (11 – 13 días) y de 12 días (10 – 16 días), para las Dieta 1 y Dieta 2, respectivamente. La media estimada del tiempo de sobrevivencia es mayor para la Dieta 1 que para la Dieta 2 (11,27 días versus 10,91 días, respectivamente).

### d.4. Análisis por método de tabla de vida o actuarial (LT) para cuarto estadio larval:

#### d.4.1. Análisis para Dieta 1:

Funcion de sobrevivencia versus tiempo  
The LIFETEST Procedure  
Stratum 4: ESTADIO = L4  
Life Table Survival Estimates

Interval [Lower, Upper)	Number Failed	Number Censored	Effective Sample Size	Conditional Probability of Failure	Conditional Probability Standard Error	Survival	Failure
0 5	0	5	17.5	0	0	1.0000	0
5 10	1	0	15.0	0.0667	0.0644	1.0000	0
10 15	14	0	14.0	1.0000	0	0.9333	0.0667



Evaluated at the Midpoint of the Interval

Interval [Lower, Upper)		Survival Standard Error	Median Residual Lifetime	Median Standard Error	PDF PDF	PDF Standard Error	Hazard Standard Error
0	5	0	.	.	0	.	0
5	10	0	.	.	0.0133	0.0129	0.0138 0.01379
10	15	0.0644	.	.	0.1867	0.0129	0.4 0

#### d.4.2. Análisis para Dieta 2:

Stratum 4: ESTADIO = L4  
Life Table Survival Estimates

Interval [Lower, Upper)		Number Failed	Number Censored	Effective Sample Size	Conditional Probability of Failure	Conditional Probability Standard Error	Survival	Failure
0	5	0	9	15.5	0	0	1.0000	0
5	10	3	0	11.0	0.2727	0.1343	1.0000	0
10	15	7	0	8.0	0.8750	0.1169	0.7273	0.2727
15	20	1	0	1.0	1.0000	0	0.0909	0.9091

Evaluated at the Midpoint of the Interval

Interval [Lower, Upper)		Survival Standard Error	Median Residual Lifetime	Median Standard Error	PDF PDF	PDF Standard Error	Hazard Standard Error
0	5	0	11.7857	0.9979	0	.	0
5	10	0	6.7857	1.1845	0.0545	0.0269	0.06316 0.036007
10	15	0.1343	2.8571	1.0102	0.1273	0.0290	0.31111 0.073909
15	20	0.0867	.	.	0.0182	0.0173	0.4 0

Las larvas del cuarto instar presentan tres y cuatro intervalos de vida, para Dieta 1 y Dieta 2, en los que experimentan aumento de la función de riesgo de mudar después del primer intervalo (0 – 5 días), pasando de 0,0 mudas/día en el primer intervalo, a 0,0138 mudas/día en el segundo intervalo (5 – 10 días) en la Dieta 1, a 0,06316 mudas/día en Dieta 2 y de 0,311 mudas/día en el tercer intervalo (10 – 15 días), para la Dieta 2, para llegar a 0,4 mudas/día en el tercer y cuarto intervalo, para Dieta 1 y Dieta 2.

**d.5. Análisis por método de Kaplan - Meier para quinto estadio larval:****d.5.1. Análisis para Dieta 1:**

The LIFETEST Procedure  
Stratum 5: ESTADIO = L5  
Summary Statistics for Time Variable DURACIÓN  
Quartile Estimates

Percent	Point Estimate	95% Confidence Interval	
		[Lower	Upper)
75	19.0000	17.0000	20.0000
50	17.5000	17.0000	19.0000
25	17.0000	14.0000	18.0000
	Mean	Standard Error	
	17.2857	0.6151	

**d.5.2. Análisis para Dieta 2:**

Summary Statistics for Time Variable DURACIÓN  
Quartile Estimates

Percent	Point Estimate	95% Confidence Interval	
		[Lower	Upper)
75	24.0000	20.0000	27.0000
50	21.0000	18.0000	24.0000
25	18.0000	16.0000	22.0000
	Mean	Standard Error	
	20.9000	1.1101	

En el quinto instar larval, el percentil 50<sup>th</sup>, el valor promedio para la Dieta 1 es de 17,5 días (17 – 19 días) y el percentil 75<sup>th</sup> con un valor promedio de 19 días (17 – 20 días), con una media estimada del tiempo de sobrevivencia es de 17,29 días. La Dieta 2 continúa con mayores valores de crecimiento, alcanzando para el percentil 50<sup>th</sup> un valor promedio de 21 días (18 – 24 días) y para el percentil 75<sup>th</sup>, un valor promedio de 24 días (20 -27 días), además de una media estimada de sobrevivencia de 20,9 días.

## d.6. Análisis por método de tabla de vida o actuarial (LT) para quinto estadio larval:

### d.6.1. Análisis para Dieta 1:

Funcion de sobrevivencia versus tiempo

The LIFETEST Procedure

Stratum 5: ESTADIO = L5

Life Table Survival Estimates

Interval [Lower, Upper)	Number Failed	Number Censored	Effective Sample Size	Conditional Probability of Failure	Conditional Probability Standard Error	Survival	Failure
0 5	0	6	17.0	0	0	1.0000	0
5 10	0	0	14.0	0	0	1.0000	0
10 15	3	0	14.0	0.2143	0.1097	1.0000	0
15 20	9	0	11.0	0.8182	0.1163	0.7857	0.2143
20 25	2	0	2.0	1.0000	0	0.1429	0.8571

Evaluated at the Midpoint of the Interval

Interval [Lower, Upper)	Survival Standard Error	Median Residual Lifetime	Median Standard Error	PDF	PDF Standard Error	Hazard	Hazard Standard Error
0 5	0	17.2222	0.9432	0	.	0	.
5 10	0	12.2222	1.0393	0	.	0	.
10 15	0	7.2222	1.0393	0.0429	0.0219	0.048	0.027513
15 20	0.1097	3.0556	0.9213	0.1286	0.0256	0.277	0.066609
20 25	0.0935	.	.	0.0286	0.0187	0.4	0

### d.6.2. Análisis para Dieta 2:

Stratum 5: ESTADIO = L5

Life Table Survival Estimates

Interval [Lower, Upper)	Number Failed	Number Censored	Effective Sample Size	Conditional Probability of Failure	Conditional Probability Standard Error	Survival	Failure
0 5	0	10	15.0	0	0	1.0000	0
5 10	0	0	10.0	0	0	1.0000	0
10 15	0	0	10.0	0	0	1.0000	0
15 20	4	0	10.0	0.4000	0.1549	1.0000	0
20 25	5	0	6.0	0.8333	0.1521	0.6000	0.4000
25 30	1	0	1.0	1.0000	0	0.1000	0.9000

Interval		Evaluated at the Midpoint of the Interval						
[Lower,	Upper)	Survival Standard Error	Median Residual Lifetime	Median Standard Error	PDF PDF	PDF Standard Error	Hazard Standard Hazard	Hazard Standard Error
0	5	0	21.0000	1.2910	0	.	0	.
5	10	0	16.0000	1.5811	0	.	0	.
10	15	0	11.0000	1.5811	0	.	0	.
15	20	0	6.0000	1.5811	0.0800	0.0310	0.1	0.048412
20	25	0.1549	3.0000	1.2247	0.1000	0.0316	0.2857	0.089424
25	30	0.0949	.	.	0.0200	0.0190	0.4	0

La probabilidad condicional de falla para el quinto instar larval bajo la Dieta 1 en el intervalo 10 – 15 días fue de  $0,2143 \pm 0,1097$ , que representa la probabilidad estimada de que las larvas muden o cambien de estadio en este intervalo. Para las larvas en Dieta 2, en el mismo intervalo, fue de  $0,000 \pm 0,000$ , lo que significa que en la Dieta 2 las larvas no presentaron probabilidad de mudar en el intervalo. En relación a la probabilidad estimada que una larva cambie de estadio o muda en un tiempo mayor o igual que el tiempo al inicio del intervalo entre 10 – 15 días (Survival), es de  $1,000 \pm 0,000$  para las larvas sometidas a ambas Dietas. La columna del residual de la mediana (Median Residual Lifetime), como un estimado del tiempo restante hasta un evento para una larva que sobrevivió al comienzo del intervalo 10 – 15 días, resulto de  $7,2222 \pm 1,0393$  días para Dieta 1, comparado con  $11,0000 \pm 1,5811$  días para Dieta 2. Este valor significa que para la Dieta 1, las especies tienen un desarrollo larval en menor tiempo.

Las larvas del quinto instar presentan cinco y seis intervalos de vida, para la Dieta 1 y Dieta 2. La Dieta 1 en los dos primeros intervalos (0 – 10 días) no presenta riesgos de muda. Experimentan aumento de la función de riesgo de mudar a partir del tercer intervalo (10 – 15 días), pasando de 0,0 mudas/día, a 0,048 mudas/día, aumentando a 0,277 mudas/día en el cuarto intervalo (15 – 20 días), y a 0,4 mudas/día en el quinto intervalo (20 – 25 días). Por su parte, la Dieta 2 no presenta riesgo de muda hasta el cuarto intervalo (15 – 20 días), pasando de 0,0 mudas/día, a 0,1 mudas/día. Aumenta a 0,2857 mudas/día en el quinto intervalo (20 – 25 días), y a 0,4 mudas/día en el sexto.

## Análisis semiparamétrico y pruebas de supuestos de riesgo no proporcional (PH)

### a. Prueba del supuesto de proporcionalidad con los residuales de Schoenfeld.

El uso principal de estos residuos es detectar posibles desviaciones del supuesto de riesgo proporcional. Si se satisface la suposición, los residuos de Schoenfeld deben ser independientes del tiempo. La idea es probar si los residuos de Schoenfeld están correlacionados con el tiempo o con alguna función del tiempo (Allison, 2010).

```

The PHREG Procedure
Model Information
Data Set          WORK.MORPHO
Dependent Variable  DURACIÓN      Duración de la fase o estadio
Censoring Variable  Status
Censoring Value(s)  0
Ties Handling       EFRON

Number of Observations Read  280
Number of Observations Used  280

```

#### Summary of the Number of Event and Censored Values

Total	Event	Censored	Percent Censored
280	194	86	30.71

#### Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

#### Model Fit Statistics

Criterion	Without Covariates	With Covariates
-2 LOG L	1663.036	1662.974
AIC	1663.036	1664.974
SBC	1663.036	1668.242

## Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	0.0613	1	0.8045
Score	0.0612	1	0.8046
Wald	0.0609	1	0.8050

## Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
DIETA	1	-0.03646	0.14771	0.0609	0.8050	0.964

## The CORR Procedure

4 Variables: DURACIÓN DURACION1 DURACION2 schDIETA

## Simple Statistics

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
DURACIONRANK	280	140.50	79.50	39340	43.50	280.00
DURACION1	194	2.11	0.514	409.54	0.693	3.296
DURACION2	280	12.94	11.572	3622	0	54.000
schDIETA	194	0.00074	0.49456	0.14511	-0.82821	0.62622

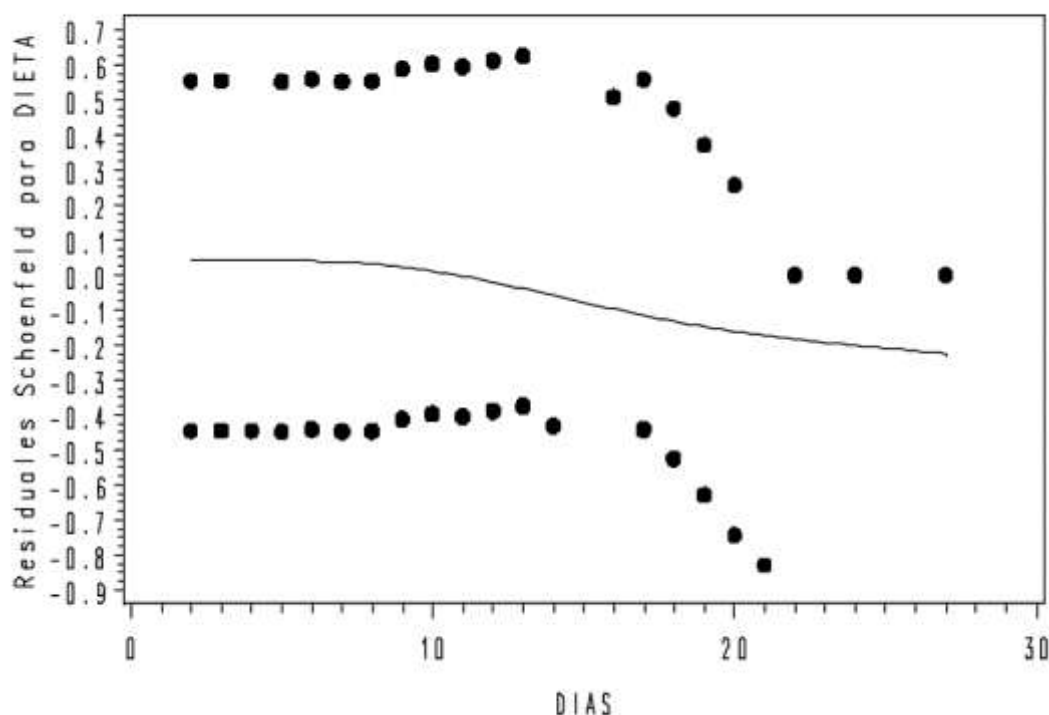
## Simple Statistics

Variable	Label
DURACIONRANK	Día JERARQUIZADO
DURACION1	Logaritmo de DURACIÓN
DURACION2	DURACION al cuadrado
schDIETA	Schoenfeld Residual for DIETA

Los resultados mostrados a continuación, como producto de los residuales de Schoenfeld, indican que no se presentan problemas de riesgo proporcional con la covariable DIETAS, con valores de  $p$  comprendidos entre 0,1105 y 0,2324 en las tres funciones relacionadas con el tiempo ( $p > 0,05$ ), lo que indica un resultado no significativo. También se observa que el tamaño de muestra de estas correlaciones es de 194, ya que los residuales de Schoenfeld no están definidos para los casos censurados (Allison, 2010).

Pearson Correlation Coefficients				
Prob >  r  under H0: Rho=0				
Number of Observations				
	DURACIONRANK	DURACION1	DURACION2	schDIETA
DURACIONRANK	1.00000	0.95219	0.96504	-0.08548
		<.0001	<.0001	0.2360
	280	194	280	194
DURACION1	0.95219	1.00000	0.93803	-0.08614
	<.0001		<.0001	0.2324
	194	194	194	194
DURACION2	0.96504	0.93803	1.00000	-0.11495
	<.0001	<.0001		0.1105
	280	194	280	194

Sin embargo, en la Figura 38, los residuales de Schoenfeld muestran una tendencia a la inclinación, con una pendiente de  $15 - 20^\circ$ , lo que indica que las suposiciones de riesgo proporcional no se satisfacen y puede sugerir que existe violación de las suposiciones para DIETAS. Los residuales Schoenfeld están basados en el principio que, para una covariable regresora, la suposición restrictiva  $\beta(t_k) = \beta$  para todo valor de tiempo  $t_k$ . Esto implica que una gráfica de  $\beta(t_k)$  versus tiempo tendrá cero inclinaciones. La hipótesis nula es teniendo inclinación cero, la cual es equivalente a probar que el log de la función de razón de riesgo ( $H_r$ ) es constante en el tiempo.



**Figura 38.** Gráfica de residuos de Schoenfeld para la interacción entre el tiempo (DURACIÓN) y la covariable DIETAS de *M. menelaus occidentalis* y *M. helenor theodorus* bajo condiciones de cautiverio. Fuentes. Elaboración propia.

**b. Modelo Cox de estimación inicial simple, solo con la covariable DIETA (Sin interacciones).**

El uso de la función loglogística asegura que el riesgo es positivo, es un valor relativo y no absoluto. Se asume que las covariables tienen un efecto aditivo sobre el logaritmo natural de razón de riesgo. La interpretación del parámetro  $\beta$  es que por cada unidad que aumenta la covariable DIETAS, el riesgo se multiplica por  $e^\beta$ . En el caso especial de DIETAS tomando el valor 0 o 1 para representar a grupos, tal como DIETA 1 y DIETA 2, el grupo DIETA 2 tiene  $e^\beta$  veces el riesgo del grupo DIETA 1.



## The PHREG Procedure

## Model Information

Data Set	WORK.MORPHO
Dependent Variable	DURACIÓN Duración de la fase oestadio
Censoring Variable	Status
Censoring Value(s)	0
Ties Handling	EXACT

Number of Observations Read	280
Number of Observations Used	280

## Class Level Information

Class	Value	Design Variables
DIETA	1	1
	2	0

## Summary of the Number of Event and Censored Values

Total	Event	Censored	Percent Censored
280	194	86	30.71

## Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

## Model Fit Statistics

Criterion	Without Covariates	With Covariates
-2 LOG L	1663.036	1662.974
AIC	1663.036	1664.974
SBC	1663.036	1668.242

En los resultados del análisis PROC PHREG, ya que la variable DIETA toma dos valores, la hipótesis nula para ninguna diferencia entre las dos dietas es idéntica a la hipótesis nula que el coeficiente de regresión para DIETAS es cero. Las tres pruebas que aparecen en la tabla indicada como "*Testing Global Null Hypothesis: BETA=0*" sugieren que las curvas de sobrevivencia para las dos dietas alimenticias pueden resultar iguales ( $p > 0,05$ ). En este modelo, el signo negativo del coeficiente de regression  $\beta_i = -0,03646$  para DIETA, indica que la Dieta 1 tuvo menor tiempo de desarrollo que la Dieta 2. El coe-

ficiente exponenciado,  $\exp(\beta_i)$  representa la razón de riesgo y da el tamaño del efecto de la covariable. Puede considerarse como el cambio pronosticado en el riesgo para una unidad de incremento en el predictor. El valor estimado de la razón de riesgo de 0,964 ( $<1$ ), con signo negativo en su coeficiente indica que el riesgo de sobrevivencia para los insectos que recibieron la DIETA 1 es alrededor del 96% del riesgo de los insectos que recibieron la DIETA 2. Por ello, bajo la DIETA 1, tiene mas baja probabilidad de sobrevivencia. El inverso  $1/\exp(\beta_i)$  igual a 1,04, indica que la probabilidad de experimentar la sobrevivencia o desarrollo de los estadios, disminuirá por un factor de 1,04 cuando es la DIETA 1 (*Platymiscium stipulare*) comparada con la DIETA 2 (*Arachis pintoii*). En otras palabras, los insectos en DIETA=2 tienen un ciclo de vida máas prolongado que los de DIETA=1. El valor de  $p = 0,8050$ , indica que no existe relacion entre la covariables DIETA y el cambio decreciente de la sobrevivencia. Además, al incluir el valor nulo (1) en HR, en el intervalo de confianza al 95% se concluye que el riesgo de sobrevivencia fue similar en los dos grupos (DIETA 1 = DIETA 2).

## Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	0.0613	1	0.8045
Score	0.0612	1	0.8046
Wald	0.0609	1	0.8050

## Type 3 Tests

Effect	DF	Wald	
		Chi-Square	Pr > ChiSq
DIETA	1	0.0609	0.8050

## Analysis of Maximum Likelihood Estimates

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits	
DIETA 1	1	-0,03646	0,14771	0,0609	0,8050	0,964	0,722	1,288

### c. Modelo Cox con interacción de la covariable DIETA y el tiempo (DURACIÓN).

En este modelo, la variable de respuesta es el "riesgo". El modelo asume que la tasa de riesgo subyacente, en lugar del tiempo de supervivencia, es una función de las variables independientes o covariables. Significa que un cambio de una unidad en una variable explicativa, causa cambios proporcionales en el riesgo. Para el análisis, se consideró la DIETA 2 como grupo de referencia para comparar con DIETA 1 (Romano y Stevanato, 2020; McNair et al., 2012).

```

The PHREG Procedure
Model Information
Data Set          WORK.MORPHO
Dependent Variable  DURACIÓN          Duración de la fase o estadio
Censoring Variable  Status
Censoring Value(s)  0
Ties Handling       EFRON

Number of Observations Read  280
Number of Observations Used  280

Class Level Information

Class      Value      Design
          Variables
DIETA      1          1
           2          0

Summary of the Number of Event and Censored Values

Total      Event      Censored      Percent
          Censored
280        194        86          30.71

Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

```

Criterion	Without Covariates	With Covariates
-2 LOG L	1663.036	1658.736
AIC	1663.036	1662.736
SBC	1663.036	1669.271

## Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	4.2998	2	0.1165
Score	4.3771	2	0.1121
Wald	4.2716	2	0.1182

## Type 3 Tests

Effect	DF	Wald Chi-Square	Pr > ChiSq
DIETA	1	3.1453	0.0761
DIETDUR	1	4.2164	0.0400

## Analysis of Maximum Likelihood Estimates

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits
DIETA 1	1	0.63778	0.35962	3.1453	0.0761	1.892	0.935 3.829
DIETDUR	1	-0.07619	0.03711	4.2164	0.0400	0.927	0.862 0.997

Para esta segunda evaluación, la interacción de las DIETAS con el tiempo, a diferencia de la evaluación anterior (sin interacción), tiene un impacto negativo en el desarrollo. Es decir, la interacción disminuye las probabilidades de desarrollo ( $\beta_i = -0,07619$ ) en el tiempo. De manera similar, para evaluar los efectos de los tratamientos y sus cambios en el tiempo (interacciones), al definir la variable tiempo dependiente (DIET\*DUR), para el período de evaluación, la interacción produjo valores significativos ( $p = 0,04$ ), indicando efectos en el período de evaluación. Ya que la interacción fue significativa se concluye que la suposición de riesgo proporcional se viola en el período. Por este motivo se rechaza el modelo de Cox y se acepta el modelo paramétrico. De manera individual, el signo

positivo del coeficiente de regresión para DIETA 1, indica que sus probabilidades de desarrollo aumentaron durante el período y tienen más alta probabilidad de desarrollo que la DIETA 2 (dieta de referencia).

El valor de la razón de riesgo para la DIETA 1, resultó de 1,892. Al interpretar este valor, significa que el riesgo de desarrollo de los insectos es alrededor de 89,2% de la razón de riesgo de los insectos del tratamiento de referencia (DIETA 2). Otra interpretación para el valor de HR de DIETA 1, con una HR de 1,892, con un intervalo de confianza del 95% de 0,935 a 3,829, tiene por resultado que al incluir el intervalo de confianza el valor nulo (1) en las comparaciones con la DIETA 2 (dieta de referencia), se concluye que el riesgo de desarrollo fue similar en ambas dietas (DIETA 1 = DIETA 2). Por lo tanto, no existen diferencias significativas del riesgo temporal de desarrollo entre los insectos de los tratamientos bajo comparación. La interpretación correcta sería que los insectos alimentados con la DIETA 1 que no hubiesen completado el desarrollo tenían igual probabilidad de presentarlo en el siguiente intervalo de tiempo que los insectos del grupo control o tratamiento de referencia (DIETA 2).

De manera equivalente, los valores inversos de  $1 / \exp(\beta_i)$  igual a 1,89, revela que la probabilidad de experimentar desarrollo disminuyó por este factor para el tratamiento con DIETA 1 (*Arachis pintoii*), al compararse con DIETA 2 (*Platymiscium stipulare*), aunque no significativamente ( $p = 0,0761$ ).

#### **g. Resultados del Análisis Paramétrico**

Se corrieron las distribuciones Llogistic – Exponential – Weibull – Lognormal, resultando la distribución Lognormal con la mejor convergencia de los datos, siguiendo la LLogística. Para establecer el mejor ajuste los modelos en PROC LIFEREG son estimados por el método de máxima verosimilitud (ML), la magnitud más baja corresponde al mejor ajuste (Allison, 2010).

A continuación, los resultados para el modelo Lognormal. La salida de resultados primero provee información preliminar: nombres de las variables de tiempo y censura, valores correspondientes a censurados y números de casos por tipo de censura. Luego la salida muestra que el log de probabilidad para el modelo es  $-7,9259$ , valor estadístico que se utilizará posteriormente para probar varias hipótesis.

La salida “Type III Analysis of Effects” del modelo Lognormal, tendrá importancia al considerar variables de clasificación (CLASS). También se reportan los coeficientes estimados, los errores estándar, chi cuadrado para la hipótesis nula que cada coeficiente es 0 y valores de p asociados con este estadístico.

```

The LIFEREG Procedure
Model Information
Data Set                WORK.MORPHO
Dependent Variable      Log(DURACIÓN)        Días de duración de
                                                                la fase o estadio
Censoring Variable      Status
Censoring Value(s)      0
Number of Observations  194
Noncensored Values      194
Right Censored Values   0
Left Censored Values    0
Interval Censored Values 0
Zero or Negative Response 86
Name of Distribution     Lognormal
Log Likelihood          -7.925991367

Number of Observations Read  280
Number of Observations Used  194

Class Level Information
Name      Levels  Values
DIETA     2      1 2

Algorithm converged.

Type III Analysis of Effects
Effect    DF      Wald
          Chi-Square  Pr > ChiSq
DIETA     1      179.0844   <.0001
DIETDUR   1      607.3855   <.0001

```

Analysis of Parameter Estimates							
Parameter	DF	Estimate	Standard Error	95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	0.9876	0.0527	0.8843	1.0908	351.53	<.0001
DIETA	1	0.5741	0.0429	0.4900	0.6582	179.08	<.0001
DIETA	2	0.0000	.	.	.	.	.
DIETDUR	1	0.0596	0.0024	0.0549	0.0644	607.39	<.0001
Scale	1	0.2521	0.0128	0.2282	0.2784		

El coeficiente de regresión para el tratamiento bajo DIETA 2, es 0. Su razón es que se designó como nivel de referencia para comparar con el tratamiento DIETA 1. El análisis de los parámetros estimados, indica que el coeficiente de DIETA 1, de manera independiente, difiere significativamente del tratamiento bajo DIETA 2 ( $p < 0,0001$ ). Es decir, en el modelo Lognormal, el efecto es significativo y satisface el criterio de  $p=0,05$  de significancia estadística, además de presentar un coeficiente positivo. Los signos de los coeficientes para DIETAS indican la dirección de la relación. El signo positivo logrado por la variable DIETA 1 indica que a medida que transcurre el tiempo (DURACIÓN) aumenta el periodo de crecimiento o ciclo de vida de los insectos (con el tiempo, el periodo de crecimiento o ciclo de vida es más rápido). Significa que los insectos alimentados con esta dieta tienen un desarrollo más rápido. Específicamente, el coeficiente de 0,5741 para la DIETA 1, significa que, por cada unidad de incremento en tiempo, el crecimiento aumenta, siendo multiplicado por  $\exp(0,5741) = 1,78$ . Es decir, una unidad de duración (tiempo) produce un aumento del crecimiento en 1,78 unidades de tiempo (días) que representa un crecimiento de 78,0 % de la Dieta 1 en relación a la Dieta 2. Para la DIETA 2, el tiempo de duración es definido por  $\exp(0,000) = 1,00$ , es decir una unidad de tiempo por unidad de duración.

En la estimación de los parámetros ATF para el modelo Lognormal, la intersección está representado por  $\beta_0 = 0,9876$  y DIETA 1, por  $\beta_1 = 0,5741$ . Estos valores se utilizan para calcular el tiempo medio de desarrollo ( $T_{50}$ ). El  $T_{50}$  (mediana o tiempo medio de sobrevivencia) para referencia de cada parámetro, se calcula como  $\exp(\beta)$ . Para todo el lote (intercepto =  $\beta_0 = 0,9876$ ), fue de  $\exp^{(0,9876)} = 2,7$  días, mientras que para el efecto de DIETA 1 es,  $\exp^{(0,9876 + 0,5741)} = 4,8$  días. Para DIETA 2 es,  $\exp^{(0,9876 + 0,000)} = 2,7$  días. Cuanto más positivo sea el coeficiente, más corto será el tiempo medio de desarrollo. En términos de la razón de tiempo ( $\gamma$ ), para DIETA 1  $\gamma = 4,8/2,7 = 1,8$  y un valor de 1 para DIETA 2 ( $\gamma$

=  $2,7/2,7 = 1,000$ . Estos valores, interpretados como razones estimadas de los tiempos de desarrollo esperados para dos grupos, según la fórmula  $(100 \% (1 - \gamma))$ , significa que DIETA 1 ( $100 \% (1 - 1,8) = 80,0 \%$ ), disminuye el tiempo de desarrollo en 80,0 %, en comparación con la DIETA 2. En general, en los modelos AFT, una razón de tiempo  $> 1$  para una covariable, implica que prolonga el tiempo de sobrevivencia, mientras que una razón de tiempo  $< 1$  implica que es más probable un evento de sobrevivencia más temprano (Romano y Stevanato, 2020).

Entre los resultados obtenidos por Peña (2017), la dieta con el hospedero *Arachis pintoi* (DIETA 1) fue óptima para la especie *Morpho helenor theodorus*. Su ciclo de vida lo completó en  $73,22 \pm 7,26$  días ( $n= 14$ ), con una sobrevivencia del 70%; mientras que, con el hospedero *Platymiscium stipulare* (DIETA 2) completó su ciclo biológico en un tiempo más prolongado de  $75,84 \pm 10,19$  días ( $n= 10$ ), en contraste presentó una baja tasa de sobrevivencia (50%).

### 3.3.2.2. Experimentos de laboratorio

El objetivo de este tipo de datos es determinar la duración de la remisión y sobrevivencia, además de comparar su distribución en cada grupo o tratamiento aplicado. Después de comparar la remisión y distribuciones de sobrevivencia se determina si los tratamientos aplicados son comparables con respecto a factores o variables de pronóstico. Se siguen los siguientes pasos:

1. Estimar y graficar las funciones de sobrevivencia de los tratamientos. Las curvas resultantes se llaman curvas de sobrevivencia. Los puntos sobre la curva estiman la proporción de individuos que sobrevivirán al menos un periodo dado de tiempo. Para muestras pequeñas, con observaciones censuradas, es apropiado el método Producto - Limite (PL) de Kaplan – Meier (Kaplan – Meier, 1958) para la estimación de las funciones de sobrevivencia, la cual no requiere de suposiciones acerca de la forma de las funciones que están siendo estimadas (Lee y Wenyu, 2003). Además de PL se puede especificar la opción LT para estimar resultados de la tabla de vida (o actuarial). Para las gráficas se especifican las gráficas SURVIVAL (S) y LOGLOGS (LLS) de la función  $\log(-\log(\text{estimada SDF}))$  versus  $\log(\text{tiempo})$ .



2. Examinar los pronósticos de homogeneidad de los grupos. ¿La siguiente interrogante es conocer si las diferencias en supervivencia entre los grupos de tratamientos son estadísticamente significativas o son producto de variaciones al azar? Para ello es necesaria una prueba de significancia. Sin embargo, una prueba estadística sin considerar características de los individuos tiene sentido sólo si son homogéneos con respecto a factores de pronóstico. Se asume entonces, que los individuos de los grupos son comparables y que la única diferencia entre ellos son los tratamientos. Entonces, luego de realizar una prueba estadística es necesario examinar la homogeneidad entre los grupos para las diferentes variables de pronóstico (edad, sexo sanidad, etc).

3. Comparar las distribuciones de supervivencia. Existen varias pruebas paramétricas y no paramétricas para comparar las distribuciones de supervivencia. Al no disponer de información de las distribuciones que siguen los datos, se utilizan los métodos no paramétricos para la comparación de las distribuciones. En adición a las pruebas log-rank y Wilcoxon, producidas por defecto, la opción STRATA también origina otras cuatro pruebas no paramétricas: Tarone-Ware, Peto-Peto, modificada Peto-Peto y Fleming-Harrington.

Al probar la significancia de la diferencia entre dos distribuciones, la hipótesis es que la distribución de supervivencia de ambos grupos es igual. La hipótesis nula es:  $H_0: S_1(t) = S_2(t)$ . Por el contrario, la hipótesis alternativa es:  $H_1: S_1(t) \neq S_2(t)$ , ya que no se dispone de información anterior sobre la superioridad de uno u otro tratamiento. La diferencia entre las curvas estimadas de supervivencia debe ser por una variación al azar. La hipótesis alternativa  $H_1: S_1(t) < S_2(t)$  debe considerarse inapropiada. Por ejemplo, si en la prueba Wilcoxon, la diferencia en las distribuciones de supervivencia de los tratamientos se establece como no significativa ( $p=0,33$ ). Además, no se rechaza la hipótesis nula que las distribuciones son iguales. Aunque la conclusión es que los datos no proveen una evidencia fuerte para rechazar la hipótesis, “no rechazar la hipótesis nula” de manera automática no indica “aceptar la hipótesis nula”. La diferencia entre las dos afirmaciones es que la probabilidad de error de la última afirmación es mucho mayor que la primera.

### **3.3.2.2.1. Relación entre dieta alimenticia y desarrollo de tumores en ratas de laboratorio (King *et al.*, 1979)**

#### **Metodología**

Un laboratorio de investigación interesado en obtener información sobre la relación entre dietas alimenticias y desarrollo de tumores, dividió 90 ratas en tres grupos, uno de ellos con una dieta baja en grasa, otro con grasas saturadas y un tercer grupo con grasas insaturadas (King *et al.*, 1979). Las ratas eran de la misma especie y edad e igual condición física. Una cantidad igual de células tumorales se inyectaron dentro de cada rata y fueron observadas durante 200 días. La rata 16 del grupo bajo en grasa y la rata 24 del grupo grasas saturadas, murieron accidentalmente luego de 140 y 170 días, respectivamente, sin evidencia de tumores. El Cuadro 14, describe el tiempo libre de tumores, tiempo desde la inyección al tiempo que se desarrolla el tumor o la finalización del estudio. Quince (15) de las 30 ratas alimentadas con dieta baja de grasa, desarrollaron tumores antes de terminar la investigación. Las ratas que murieron presentaron un tiempo libre de tumores de al menos 140 días. Las otras 14 ratas no desarrollaron algún tumor hasta el final del experimento; su tiempo libre de tumores fue de al menos, 200 días. Entre las ratas del grupo dietas saturadas, 23 de ellas desarrollaron tumor, una murió libre de tumores después de 170 días y 6 estuvieron libre de tumores hasta el final del experimento. En el grupo de dietas con grasas insaturadas, las 30 ratas desarrollaron tumores dentro de los 200 días.

**Cuadro 14. Tiempo libre de tumores (días) de 90 ratas en tres dietas diferentes.**

Rata N°	Baja grasa	Rata N°	Grasa satura- da	Rata N°	Grasa insatu- rada
1	140	1	124	1	112
2	177	2	58	2	68
3	50	3	56	3	84
4	65	4	68	4	109
5	86	5	79	5	153
6	153	6	89	6	143
7	181	7	107	7	60
8	191	8	86	8	70
9	77	9	142	9	98
10	84	10	110	10	164
11	87	11	96	11	63
12	56	12	142	12	63
13	66	13	86	13	77
14	73	14	75	14	91
15	119	15	117	15	91
16	140 <sup>+</sup>	16	98	16	66
17	200 <sup>+</sup>	17	105	17	70
18	200 <sup>+</sup>	18	126	18	77
19	200 <sup>+</sup>	19	43	19	63
20	200 <sup>+</sup>	20	46	20	66
21	200 <sup>+</sup>	21	81	21	66
22	200 <sup>+</sup>	22	133	22	94
23	200 <sup>+</sup>	23	165	23	101
24	200 <sup>+</sup>	24	170 <sup>+</sup>	24	105
25	200 <sup>+</sup>	25	200 <sup>+</sup>	25	108
26	200 <sup>+</sup>	26	200 <sup>+</sup>	26	112
27	200 <sup>+</sup>	27	200 <sup>+</sup>	27	115
28	200 <sup>+</sup>	28	200 <sup>+</sup>	28	126
29	200 <sup>+</sup>	29	200 <sup>+</sup>	29	161
30	200 <sup>+</sup>	30	200 <sup>+</sup>	30	178

Fuente: King et al. (1979).

### Manejo de datos

A continuación, se describe el programa utilizado para la disposición y arreglo de los datos especificados en el Cuadro 14.

### Comparación de K (K>2) muestras.

En este caso, situaciones en las cuales se presenta una muestra de cada uno de los tratamientos asignados (K>2). El problema es decidir si las K muestras independientes pueden ser consideradas como provenientes de la misma población, o en la práctica, si los datos de sobrevivencia de individuos que reciben los K tratamientos proporcionan suficientes evidencias para concluir que los K tratamientos no son igualmente efectivos. Este problema se puede resolver utilizando las pruebas log-rank, Wilcoxon y la prueba razón de probabilidad (-2log (LR)), que origina por defecto PROC LIFETEST para K > 2 muestras.

### El procedimiento STRATA con PROC LIFETEST para K > 2 muestras y las salidas que genera.

PROC LIFETEST puede ejecutar las tres pruebas citadas anteriormente para comparar 2 o más grupos con respecto al tiempo de sobrevivencia. Simplemente se agrega a PROC LIFETEST la declaración STRATA var(s):

```
Data Ratas2;
```

```
Input Días Status Grupo @@;
```

```
Label Días='Día De exposicion a la mortalidad' Status='Variable indicadora de Censura'
```

```
Grupo='Pretratamientos colocados a las ratas' Grupo 1='Baja grasa, Bg' Grupo 2='Grasa saturada, Gs'
```

```
Grupo 3='Grasa insaturada, Gi';
```

```
Grupdias = Grupo * (Días < 100);
```

```
Grupdias = Grupo * (Días > 100);
```

```
Grupdias=Grupo*Días;
```

```
Datalines;
```

140	1	1	124	1	2	112	1	3
177	1	1	58	1	2	68	1	3
50	1	1	56	1	2	84	1	3
65	1	1	68	1	2	109	1	3
86	1	1	79	1	2	153	1	3
153	1	1	89	1	2	143	1	3
181	1	1	107	1	2	60	1	3
191	1	1	86	1	2	70	1	3
77	1	1	142	1	2	98	1	3
84	1	1	110	1	2	164	1	3
87	1	1	96	1	2	63	1	3
56	1	1	142	1	2	63	1	3
66	1	1	86	1	2	77	1	3
73	1	1	75	1	2	91	1	3
119	1	1	117	1	2	91	1	3
140	0	1	98	1	2	66	1	3
200	0	1	105	1	2	70	1	3
200	0	1	126	1	2	77	1	3
200	0	1	43	1	2	63	1	3
200	0	1	46	1	2	66	1	3
200	0	1	81	1	2	66	1	3
200	0	1	133	1	2	94	1	3

200	0	1	165	1	2	101	1	3
200	0	1	170	0	2	105	1	3
200	0	1	200	0	2	108	1	3
200	0	1	200	0	2	112	1	3
200	0	1	200	0	2	115	1	3
200	0	1	200	0	2	126	1	3
200	0	1	200	0	2	161	1	3
200	0	1	200	0	2	178	1	3

;

Proc Print Data=Ratas2;

Run;

### 1. Estimación y graficación de las funciones de sobrevivencia para los 3 grupos de tratamientos

Para muestras pequeñas con observaciones censuradas, el método producto – limite (PL) de Kaplan – Meier es apropiado para estimar las funciones de sobrevivencia. El método no requiere de suposiciones acerca de la forma de la función que se estima.

En los resultados aparecen los estimados PL de las funciones de sobrevivencia  $S(t)$  para los grupos de tratamientos. En los resultados también se indican los errores estándar de  $S(t)$  y los percentiles 75,50 y 25%. Para la obtención de estos resultados se utilizan las opciones METHOD=KM y METHOD=LT, con la opción de graficación PLOTS=(H), especifica una gráfica de estimados de la función de riesgo versus el tiempo (sólo para el método LT), con las que se producen las funciones de riesgo para cada grupo. Con la opción de PLOT= (LLS) se trata de establecer un método gráfico común para probar el supuesto PH. Consiste en trazar  $-\log(-\log(S(t/x)))$  versus  $\log(t)$ .

## 2. Estimación de la significancia estadística de la diferencia entre tratamientos

¿La siguiente interrogante es conocer si las diferencias en sobrevivencia entre los grupos de tratamientos son estadísticamente significativas o son producto de variaciones al azar? Para ello es necesaria una prueba de significancia. Sin embargo, una prueba estadística sin considerar características de los individuos tiene sentido sólo si son homogéneos entre ellos. Se asume entonces, que los individuos de los grupos son comparables y que la única diferencia entre ellos son los tratamientos. Estos métodos se describen a continuación:

```
Proc Lifetest Data=Ratas2 Method=Km Plots= (S, LS, Lls);
```

```
Time Días*Status (0);
```

```
Strata Grupo;
```

```
Run;
```

```
Proc Lifetest Data=Ratas2 Method=Lt Plots= (S, H);
```

```
Time Días*Status (0);
```

```
Strata Grupo;
```

```
Run;
```

```
Proc sort data=RATAS2;
```

```
by Descending GRUPO;
```

```
Run;
```

```
Legend1 frame cframe=ligr cborder=black Position=center value= (justify=center);
```

```
Axis1 label= (angle=90 rotate=0 'Estimated CDF') minor=none;
```

```
Axis2 minor=none;
```

```
Symbol1 value=dot c=black h=0.15in i=none r=1;
```

```
Symbol2 value=circle c=black h=0.15in i=none r=1;
```

```
Symbol3 value=square c=black h=0.15in i=none r=1;
```

### **Análisis semiparamétrico y pruebas de supuestos de riesgo no proporcional (PH).**

Para el ejemplo se crearon dos variables “ficticias”, correspondientes a las interacciones con el tiempo. Es decir, en este ejemplo, no se midieron originariamente otras variables relacionadas con las ratas experimentales, sino que fueron “creación de nuevas variables de pronóstico”, para ser utilizadas en PROC PHREG e incluirlas en el modelo. Por ejemplo, dos variables relacionadas con el tiempo y el grupo, durante los primeros 100 días del experimento,  $GRUPDIAS=Grupo*(Días<100)$  y el tiempo para los últimos 100 días del experimento,  $GRUPDIAS= Grupo*(Días>100)$ . Además de otra variable de tiempo para relacionarla con la totalidad de días del experimento ( $GRUPDIAS=Grupo*días$ ). Con ellas se evalúa el contagio de las ratas en estos períodos. El modelo PH asume que el efecto de cada covariable es el mismo en todos los momentos. Si el efecto de una variable varía con el tiempo, la suposición de PH se viola para esa variable.

Comenzamos con métodos para probar riesgos no proporcionales usando residuales. Luego examinamos otros dos métodos para evaluar los riesgos proporcionales. Un método incorpora explícitamente las interacciones en el modelo. El otro, llamado estratificación, incorpora las interacciones en la función arbitraria del tiempo.

### **Prueba del supuesto de proporcionalidad con la declaración ASSESS**

Los residuos de martingala, se pueden utilizar para probar la proporcionalidad. Este método está incorporado en la declaración ASSESS en PROC PHREG. Desafortunadamente, esta declaración no funcionará si el modelo contiene covariables dependientes del tiempo. A continuación, se explica cómo utilizar la declaración ASSESS:

```
Proc Phreg Data=Ratas2;
```

```
Model Días*Status (0) =Grupo / Ties=Efron RL;
```

```
Assess Ph / Resample;
```

```
Run;
```



En la declaración ASSESS, la opción PH solicita una evaluación del supuesto de riesgos proporcionales. Para cada covariable, esta declaración produce una representación gráfica del proceso de puntuación empírica, basado en los residuos de martingala (para versiones superiores a SAS 9.2). La opción RL es para definir los límites de confianza al 95% de las razones de riesgo.

### **Prueba del supuesto de proporcionalidad con los residuales de Schoenfeld**

A diferencia de la sentencia ASSESS, los residuales de Schoenfeld funcionan incluso si tienen covariables cambiantes en el tiempo, definidas por el método del proceso de conteo (pero no con el método de declaraciones de programación). Los residuos de Schoenfeld tienen la propiedad de incluir un residual para cada individuo y para cada covariable. Si se satisface la suposición, los residuos de Schoenfeld deben ser independientes del tiempo. La suposición de riesgos proporcionales es equivalente a probar si los residuos de Schoenfeld están correlacionados con el tiempo o con alguna función del tiempo (Allison, 2010). Para los datos del ejemplo, se produce un conjunto de datos que contiene los residuos de Schoenfeld utilizando las siguientes declaraciones:

```
Proc Phreg Data=Ratas2;
Model Días*Status (0) =Grupo / Ties=Efron RL;
Output Out=Schoen Ressch=SchGrupo;
Run;
Proc Print Data=Schoen;
Run;
Data Event;
Set Schoen;
Run;
Proc Rank Data=Event Out=Resrank Ties=Mean;
Var Días;
```

Ranks Diasrank;

Run;

Proc Print Data=Resrank;

Run;

El siguiente paso es modificar el conjunto de datos para incorporar dos funciones del tiempo, el LOG y el cuadrado de la variable relacionada con el tiempo (DÍAS), seguido de PROC CORR para calcular las correlaciones de los residuos con el tiempo y sus dos funciones.

Data C;

Set Resrank;

Días1=Log (Días);

Días2=Días\*\*2;

Proc Corr;

Var Diasrank Días1 Días2 Schtratam Schsem;

Run;

### **Generacion de gráficas de residuales Schoenfeld para Grupos como función del**

**tiempo.** /\*Generación de gráfica de residuales Schoenfeld para Grupos como función del tiempo\*/ Proc Gplot Data =Schoen;

Symbol1 V=Dot C=Black Width=1 I=Sm80s;

Plot Schgrupo\*Días/Haxis=Axis1 Vaxis=Axis2;

Axis1 Label=('Días');

Axis2 Label = (A=90 'Residuales Schoenfeld Para Grupos');

Run;

### Generacion de Razon de Riesgo de estimación inicial (sin interacciones).

Para pruebas simples que no requieran la inclusion de nuevas variables, la validez del supuesto del modelo de riesgo proporcional, consiste en comparar los tres grupos de tratamientos, de la siguiente manera:

```
Proc Phreg Data=Ratas2;
```

```
Class GRUPOS;
```

```
Model Días*Status (0) =Grupo /Ties= EFRON RL;
```

```
Run;
```

### Interacciones con covariables dependientes del tiempo.

Una forma común de representar la interacción entre dos variables en un modelo de regresión lineal es incluir una nueva variable que es el producto de las dos variables en cuestión. Para representar la interacción entre una covariable  $x$  y tiempo en un modelo de Cox, podemos escribir

$$\text{Log } h(t) = \alpha(t) + \beta_1 x + \beta_2 x t.$$

Factorizando la  $x$ , podemos reescribir esto como

$$\text{Log } h(t) = \alpha(t) + (\beta_1 + \beta_2 t) x$$

En esta ecuación, el efecto de  $x$  es  $\beta_1 + \beta_2 t$ . Si  $\beta_2$  es positivo, entonces el efecto de  $x$  aumenta linealmente con el tiempo; si es negativo, el efecto disminuye linealmente con el tiempo.  $\beta_1$  se puede interpretar como el efecto de  $x$  en el tiempo 0, el origen del proceso (Allison, 2010).

```
/*Prueba de Multicolinealidad*/
```

```
Proc Reg data = RATAS2;
```

```
Model Días = GRUPO / vif;
```

```
Run;
```

**a. Primer Modelo Cox de estimación de interacciones con variables de tiempo primeros 100 días del experimento.**

Puede estimarse este modelo definiendo una covariable tiempo dependiente  $z = xt$ . Por ejemplo, en caso de que tanto la declaración ASSESS como los residuales de Schoenfeld, aporten evidencia de que la variable GRUPOS viola la suposición de riesgos proporcionales, lo que implica que el efecto de GRUPOS varía con el tiempo desde el inicio de la prueba. Para incorporar esa interacción usando el método de declaraciones de programación, se puede ejecutar el siguiente código SAS:

```
Proc TPhreg Data=Ratas2;
Class GRUPOS;
Model Días*Status (0) =Grupo GrupDias/ Ties=Efron RL;
TEST GRUPO1=GRUPO2;
Grupdias = Grupo*(Días< 100);
Run;
```

**b. Segundo Modelo Cox de estimación de interacciones con variables de tiempo más de 100 días del experimento.**

```
Proc Tphreg Data=Ratas2;
Class GRUPO;
Model Días*Status (0) = Grupo Grupdias / Ties=Efron RL;
Test Grupo1=Grupo2;
Grupdias = Grupo*(Días> 100);
Run;
```

**c. Tercer Modelo Cox de estimación de interacciones con variables de tiempo total de días del experimento.**

```
Proc Tphreg Data=Ratas2;
```

Class GRUPO;

Model Días\*Status (0) = Grupo Grupdias / Ties=Efron RL;

Test Grupo1=Grupo2;

Grupdias=Grupo\*Días;

Run;

La idea es la siguiente:

- Agregar la interacción de una covariable para la cual la suposición de riesgo proporcional es violada con la variable tiempo (o alguna función del tiempo)
- Si la interacción es estadísticamente significativa, el efecto de esta covariable no es constante en el tiempo;
- Incluyendo la interacción en el modelo permite interpretar los parámetros adaptando este hecho dentro de la estimación.

Se puede probar si el efecto sobre los GRUPOS es significativamente diferente de 0 durante cualquier día, utilizando un procedimiento TEST. La alternativa anterior consistió en tratar el tiempo como continuo en la interacción y se definió como dicotómico con la siguiente declaración de programación:

GrupDias = Grupo \* (Día < 100);

GrupDias = Grupo \* (Día > 100);

Otro procedimiento para probar el efecto durante los días 50 – 100 y 200, es el siguiente:

Proc Tphreg Data=Ratas2;

Class GRUPO;

Model Días\*Status (0) = Grupo Grupdias / Ties=Efron RL;

Test Grupo + 50\*Grupox=0

Test Grupo + 100\*Grupox=0;

Test Grupo + 200\*Grupox=0;

```
Grupdias=Grupo*Días;
```

```
Run;
```

### **Análisis paramétrico**

Se corrieron las distribuciones Llogistic – Exponential – Weibull – Lognormal –, resultando la distribución Lognormal con la mejor convergencia de los datos:

```
Proc Sort Data=Ratas2;
```

```
By Grupo;
```

```
Run;
```

```
Proc Lifereg Data=Ratas2;
```

```
Class Grupo;
```

```
Model Días*Status (0) = Grupo GrupDias /Dist=Lognormal;
```

```
Run;
```

El programa para graficar las curvas de cada grupo es el siguiente:

```
Symbol1 Value=STAR C=Black H=0.15in I=None R=1;
```

```
Proc Lifereg data=Ratas2;
```

```
Class Grupo;
```

```
Model DÍAS*Status (0) = Grupo Grupdias /Dist=lnormal;
```

```
Probplot cencolor = red cframe = ligr cfit = blue ppout npintervals = simul;
```

```
inset / cfill = white ctext = blue;
```

```
By Grupo;
```

```
Run;
```

## Resultados

### a. Resultados por método Kaplan – Meier (METHOD=KM)

Las dos primeras muertes pueden considerarse como pérdidas por el esfuerzo inicial. Los datos son censurados individualmente si son excluidas las dos primeras muertes. El objetivo del investigador consistió en comparar la habilidad de las tres dietas para mantener las ratas libres de tumores. Para obtener información acerca de la distribución del tiempo libre de tumores, primero se estima la función de sobrevivencia (libre de tumores) de los tres grupos, utilizando el método de Kaplan – Meier PL, obteniendo la gráfica de distribución de sobrevivencia de los tres grupos, la media y los percentiles (25 – 50 – 75 %). Si los tres grupos son homogéneos, se puede saltar el paso que chequea la homogeneidad y comparar las tres distribuciones del tiempo libre de tumores, utilizando la prueba de K muestras para probar la significancia de la diferencia entre los tres grupos.

A continuación, cada línea corresponde a uno de los 30 casos, dispuestos en orden ascendente, con excepción de la primera línea, la cual es para reflejar tiempo 0. Las observaciones censuradas se indican por un asterisco. La columna principal es etiquetada como *Survival*, que indica los valores estimados de KM (Kaplan – Meier). Por ejemplo, a 84 días, el valor de  $KM=0,7667$ , lo que indica que la probabilidad estimada que una rata del GRUPO=1 sobreviviera por 84 días o más es 0,7667. En la quinta columna, etiquetada como *Number Failed*, representa el número de casos acumulados que experimentaron eventos previos, incluyendo cada punto en el tiempo. El GRUPO=1, incluye 15, el GRUPO=2, incluye a 23 y el GRUPO=3, a 30. Luego, la columna *Number Left* refleja el número de casos sin eventos experimentados o haber sido censurados antes de cada punto en el tiempo. Representa el tamaño del riesgo en cada punto, para cada uno de los grupos se incluye a 30 ratas.

Seguidamente, para el GRUPO=1, se detallan los percentiles 75th, 50th y 25th (etiquetados *Quartile Estimates*). De esta manera, el percentil 25th (86 en este caso) es el menor evento en el tiempo, donde la probabilidad de morir antes es mayor que 0,25. Por otra parte, el percentil 50th, representa la mediana del tiempo de morir. En este caso es de 191 días, con intervalo de confianza (95%) con el valor bajo de 119 días. El valor estimado de la media para morir es de 148,9 días.

## The LIFETEST Procedure

Stratum 1: GRUPO = 1

## Product-Limit Survival Estimates

DÍAS	Survival	Failure	Survival Standard Error	Number Failed	Number Left
0.000	1.0000	0	0	0	30
50.000	0.9667	0.0333	0.0328	1	29
56.000	0.9333	0.0667	0.0455	2	28
65.000	0.9000	0.1000	0.0548	3	27
66.000	0.8667	0.1333	0.0621	4	26
73.000	0.8333	0.1667	0.0680	5	25
77.000	0.8000	0.2000	0.0730	6	24
84.000	0.7667	0.2333	0.0772	7	23
86.000	0.7333	0.2667	0.0807	8	22
87.000	0.7000	0.3000	0.0837	9	21
119.000	0.6667	0.3333	0.0861	10	20
140.000	0.6333	0.3667	0.0880	11	19
140.000*	.	.	.	11	18
153.000	0.5981	0.4019	0.0899	12	17
177.000	0.5630	0.4370	0.0912	13	16
181.000	0.5278	0.4722	0.0920	14	15
191.000	0.4926	0.5074	0.0924	15	14
200.000*	.	.	.	15	13
200.000*	.	.	.	15	12
200.000*	.	.	.	15	11
200.000*	.	.	.	15	10
200.000*	.	.	.	15	9
200.000*	.	.	.	15	8
200.000*	.	.	.	15	7
200.000*	.	.	.	15	6
200.000*	.	.	.	15	5
200.000*	.	.	.	15	4
200.000*	.	.	.	15	3
200.000*	.	.	.	15	2
200.000*	.	.	.	15	1
200.000*	.	.	.	15	0

NOTE: The marked survival times are censored observations.



The LIFETEST Procedure  
Summary Statistics for Time Variable DÍAS

Quartile Estimates

Percent	Point Estimate	95% Confidence Interval [Lower Upper)	
75	.	.	.
50	191.000	119.000	.
25	86.000	66.000	177.000
	Mean	Standard Error	
	148.885	10.137	

NOTE: The mean survival time and its standard error were underestimated because the largest observation was censored and the estimation was restricted to the largest event time.

The LIFETEST Procedure

Stratum 2: GRUPO = 2

Product-Limit Survival Estimates

DÍAS	Survival	Failure	Survival Standard Error	Number Failed	Number Left
0.000	1.0000	0	0	0	30
43.000	0.9667	0.0333	0.0328	1	29
46.000	0.9333	0.0667	0.0455	2	28
56.000	0.9000	0.1000	0.0548	3	27
58.000	0.8667	0.1333	0.0621	4	26
68.000	0.8333	0.1667	0.0680	5	25
75.000	0.8000	0.2000	0.0730	6	24
79.000	0.7667	0.2333	0.0772	7	23
81.000	0.7333	0.2667	0.0807	8	22
86.000	.	.	.	9	21
86.000	0.6667	0.3333	0.0861	10	20
89.000	0.6333	0.3667	0.0880	11	19
96.000	0.6000	0.4000	0.0894	12	18
98.000	0.5667	0.4333	0.0905	13	17
105.000	0.5333	0.4667	0.0911	14	16
107.000	0.5000	0.5000	0.0913	15	15
110.000	0.4667	0.5333	0.0911	16	14
117.000	0.4333	0.5667	0.0905	17	13
124.000	0.4000	0.6000	0.0894	18	12

126.000	0.3667	0.6333	0.0880	19	11
133.000	0.3333	0.6667	0.0861	20	10
142.000	.	.	.	21	9
142.000	0.2667	0.7333	0.0807	22	8
165.000	0.2333	0.7667	0.0772	23	7
170.000*	.	.	.	23	6
200.000*	.	.	.	23	5
200.000*	.	.	.	23	4
200.000*	.	.	.	23	3
200.000*	.	.	.	23	2
200.000*	.	.	.	23	1
200.000*	.	.	.	23	0

NOTE: The marked survival times are censored observations.

The LIFETEST Procedure

Summary Statistics for Time Variable DÍAS

Quartile Estimates

Percent	Point Estimate	95% Confidence Interval	
		[Lower	Upper)
75	165.000	124.000	.
50	108.500	86.000	142.000
25	81.000	58.000	98.000
Mean		Standard Error	
112.900		7.466	

NOTE: The mean survival time and its standard error were underestimated because the largest observation was censored and the estimation was restricted to the largest event time.

The LIFETEST Procedure

Stratum 3: GRUPO = 3

Product-Limit Survival Estimates

DÍAS	Survival	Failure	Survival Standard Error	Number Failed	Number Left
0.000	1.0000	0	0	0	30
60.000	0.9667	0.0333	0.0328	1	29
63.000	.	.	.	2	28
63.000	.	.	.	3	27
63.000	0.8667	0.1333	0.0621	4	26
66.000	.	.	.	5	25

66.000	.	.	.	6	24
66.000	0.7667	0.2333	0.0772	7	23
68.000	0.7333	0.2667	0.0807	8	22
70.000	.	.	.	9	21
70.000	0.6667	0.3333	0.0861	10	20
77.000	.	.	.	11	19
77.000	0.6000	0.4000	0.0894	12	18
84.000	0.5667	0.4333	0.0905	13	17
91.000	.	.	.	14	16
91.000	0.5000	0.5000	0.0913	15	15
94.000	0.4667	0.5333	0.0911	16	14
98.000	0.4333	0.5667	0.0905	17	13
101.000	0.4000	0.6000	0.0894	18	12
105.000	0.3667	0.6333	0.0880	19	11
108.000	0.3333	0.6667	0.0861	20	10
109.000	0.3000	0.7000	0.0837	21	9
112.000	.	.	.	22	8
112.000	0.2333	0.7667	0.0772	23	7
115.000	0.2000	0.8000	0.0730	24	6
126.000	0.1667	0.8333	0.0680	25	5
143.000	0.1333	0.8667	0.0621	26	4
153.000	0.1000	0.9000	0.0548	27	3
161.000	0.0667	0.9333	0.0455	28	2
164.000	0.0333	0.9667	0.0328	29	1
178.000	0	1.0000	0	30	0

## Summary Statistics for Time Variable DÍAS

## Quartile Estimates

Percent	Point Estimate	95% Confidence Interval [Lower Upper)	
75	112.000	101.000	153.000
50	92.500	70.000	109.000
25	68.000	63.000	84.000

## The LIFETEST Procedure

Mean	Standard Error
98.467	6.193

## b. Resultados de la media, mediana y percentiles de cada grupo

En la salida de resultados reflejados anteriormente, se resalta que los valores de la media del tiempo libre de tumores de las dietas baja en grasas, grasas saturadas e insaturadas, fueron 148,9 – 112,9 y 98,5, respectivamente. De igual manera, los valores de la mediana de los tres grupos alcanzaron valores de 191 – 108,5 y 92,5, respectivamente. Esto significa que las ratas sometidas a una dieta baja en grasas presentaron el mayor tiempo libre de tumores de las tres dietas.

## c. Resultados por método de tabla de vida (METHOD=LIFE)

Una solución alternativa al método Kaplan – Meier es utilizar el método de tabla de vida, también conocido como método actuarial, en el que los tiempos de los eventos se agrupan en intervalos. Este método puede producir estimados y gráficas de la función de riesgo (Allison, 2010).

La siguiente salida muestra los resultados donde ROC LIFETEST construye cinco intervalos, comenzando en 0 e incrementando en periodos de 5 días. Estos intervalos se originaron por defecto, pero se pueden elegir con la opción INTERVALS=. Por ejemplo, INTERVALS= 15 20 30 50 produce los intervalos [0, 15), [15, 20), [20, 30), [30, 50), [50, ∞).

Para cada intervalo, se calculan 14 parámetros estadísticos (Allison, 2010). Los primeros cuatro se requieren para calcular los demás estadísticos. El tamaño efectivo de muestra (*Effective Sample Size*) es directo para los primeros cuatro o cinco intervalos, ya que ellos contienen casos no censurados en los tres grupos. El tamaño efectivo de muestra para el primer intervalo se refiere al número de individuos (ratas) sin ser censurados en cada grupo (15 – 23 – 30), lo que significa que los tamaños de muestra son de 30 ratas durante los primeros 5 días. La probabilidad condicional de falla (*Conditional Probability of Failure*) es un estimado de la probabilidad de que una rata se contagie en el intervalo. Se calcula como  $number\ failed / effective\ sample\ size$ . Un valor estimado del error standard (*Conditional Probability Standard Error*) aparece en la siguiente columna. De esta manera, la probabilidad condicional de contagio para los tres Grupos, durante los primeros 100 días fue de  $0,30 \pm 0,084$  –  $0,393 \pm 0,0923$  –  $0,567 \pm 0,0905$ , respectivamente. La columna Failure es 1 menos Survival. También aparecen los errores standard de las probabilidades de sobrevivencia (*Survival standard errors*). La columna *Survival* es el estimado de la tabla de vida de la función

de sobreviviente, que significa la probabilidad de que ocurra el evento de contagio en un tiempo mayor o igual que el tiempo al inicio de cada intervalo. Por ejemplo, entre 100 – 150 días, esta probabilidad fue de  $0,70 \pm 0,084$  –  $0,567 \pm 0,0905$  –  $0,433 \pm 0,0905$  para los tres Grupos, respectivamente.

La columna del residual de la mediana (*Median Residual Lifetime*), es un estimado del tiempo restante hasta un evento para un individuo que sobrevivió al comienzo de cada intervalo. Para calcular este estadístico para un intervalo dado, debe ser un intervalo posterior cuya probabilidad de sobrevivencia es menor que la mitad de la probabilidad de sobrevivencia asociada con el intervalo de interes. Esto significa que para el Grupo 1 el ultimo intervalo para esta evaluación fue 50 – 100 días, con valor de  $146,9 \pm 32,5$ ; para el Grupo 2, el intervalo 100 – 150 días, con valor de  $47,2 \pm 11,5$  días y para el Grupo 3, el intervalo 100 – 150 días, con valor de  $36,1 \pm 10$  días. A partir de estos intervalos, en los tres grupos, no se pudo calcular. La columna PDF Standard Error indica los valores estimados de la funcion de densidad de probabilidad en el punto medio del intervalo. De gran interes, es la columna Hazard, la cual da los estimados de la funcion de riesgo evaluados en el punto medio de cada intervalo.

A continuacion se describen los valores estimados para los 14 estadisticos, para los tres grupos.

The LIFETEST Procedure

Stratum 1: GRUPO = 1

Life Table Survival Estimates

Interval [Lower,Upper)	Number Failed	Number Censored	Effective Sample Size	Conditional Probability of Failure	Conditional Probability		Survival	Failure
					Standard Error			
0	50	0	0	30.0	0	0	1.0000	0
50	100	9	0	30.0	0.3000	0.0837	1.0000	0
100	150	2	1	20.5	0.0976	0.0655	0.7000	0.3000
150	200	4	0	18.0	0.2222	0.0980	0.6317	0.3683
200	250	0	14	7.0	0	0	0.4913	0.5087

Evaluated at the Midpoint of the Interval

Interval [Lower,Upper)		Survival Standard Error	Median Residual Lifetime	Median Standard Error	PDF PDF	PDF Standard Error	Hazard Hazard	Hazard Standard Error
0	50	0	196.9	32.5144	0	.	0	.
50	100	0	146.9	32.5144	0.00600	0.00167	0.007059	0.002316
100	150	0.0837	.	.	0.00137	0.00093	0.002051	0.001449
150	200	0.0883	.	.	0.00281	0.00130	0.005	0.00248
200	250	0.0925	.	.	0	.	0	.

**Stratum 2: GRUPO = 2**

Life Table Survival Estimates

Interval [Lower,Upper)		Number Failed	Number Censored	Effective Sample Size	Conditional Probability of Failure	Conditional Probability Standard Error	Survival	Failure
0	50	2	0	30.0	0.0667	0.0455	1.0000	0
50	100	11	0	28.0	0.3929	0.0923	0.9333	0.0667
100	150	9	0	17.0	0.5294	0.1211	0.5667	0.4333
150	200	1	1	7.5	0.1333	0.1241	0.2667	0.7333
200	250	0	6	3.0	0	0	0.2311	0.7689

Evaluated at the Midpoint of the Interval

Interval [Lower,Upper)		Survival Standard Error	Median Residual Lifetime	Median Standard Error	PDF PDF	PDF Standard Error	Hazard Hazard	Hazard Standard Error
0	50	0	111.1	15.2145	0.00133	0.00091	0.001379	0.000975
50	100	0.0455	66.6667	14.6986	0.00733	0.00176	0.009778	0.002859
100	150	0.0905	47.2222	11.4531	0.00600	0.00167	0.0144	0.004478
150	200	0.0807	.	.	0.00071	0.00069	0.002857	0.00285
200	250	0.0774	.	.	0	.	0	.

**Stratum 3: GRUPO = 3**

Life Table Survival Estimates

Interval [Lower,Upper)		Number Failed	Number Censored	Effective Sample Size	Conditional Probability of Failure	Conditional Probability Standard Error	Survival	Failure
0	50	0	0	30.0	0	0	1.0000	0
50	100	17	0	30.0	0.5667	0.0905	1.0000	0
100	150	9	0	13.0	0.6923	0.1280	0.4333	0.5667
150	200	4	0	4.0	1.0000	0	0.1333	0.8667

		Evaluated at the Midpoint of the Interval						
Interval [Lower,Upper)	Survival Standard Error	Median	Median	PDF		Hazard		
		Residual Lifetime	Standard Error	PDF	Standard Error	Hazard	Standard Error	
0	50	0	94.1176	8.0547	0	.	0	.
50	100	0	44.1176	8.0547	0.0113	0.00181	0.01581	0.003523
100	150	0.0905	36.1111	10.0154	0.00600	0.00167	0.02117	0.005988
150	200	0.0621	.	.	0.00267	0.00124	0.04	0

#### d. Resultados por método Kaplan – Meier.

Ambos métodos (KM o LIFE) originan las mismas pruebas. PROC LIFETEST imprime los estadísticos log-rank y Wilcoxon para cada grupo o tratamiento, seguido por un estimado de su matriz de covarianza, los cuales se utilizan para calcular el valor de chi cuadrada y los valores de p. El resultado indica alta significancia, lo cual sugiere que existe evidencia para rechazar la hipótesis nula que los tres grupos tienen la misma función de sobrevivencia (distribuciones de probabilidad iguales para los eventos en el tiempo). Entre los resultados también se reportan valores p para la prueba razón de probabilidad (-2log (LR)). Esta prueba se considera inferior a las otras dos (log-rank y Wilcoxon), ya que requiere la innecesaria suposición que la función de riesgo es constante para cada grupo, implicando una distribución exponencial para los eventos en el tiempo (Allison, 2010).

En la salida de resultados siguiente (*Summary of the Number of Censored and Uncensored Values*) se observa que, en el caso de GRUPOS, de las 30 evaluaciones realizadas en el Grupo 1, 15 de ellas fueron censuradas (no completaron la investigación o murieron), lo que representa un 50% de censuradas. Para Grupo 2, 7 de ellas no completaron la investigación, lo que representa el 23,33%, mientras que en el Grupo 3, la totalidad de ratas sobrevivió.

Summary of the Number of Censored and Uncensored Values						
Stratum	GRUPO	Total	Failed	Censored	Percent Censored	
1	1	30	15	15	50.00	
2	2	30	23	7	23.33	
3	3	30	30	0	0.00	
-----						
Total		90	68	22	24.44	

The LIFETEST Procedure  
Testing Homogeneity of Survival Curves for DÍAS over Strata

## Rank Statistics

GRUPO	Log-Rank	Wilcoxon
1	-14.768	-714.00
2	0.969	37.00
3	13.799	677.00

## Covariance Matrix for the Log-Rank Statistics

GRUPO	1	2	3
1	15.6801	-9.3772	-6.3029
2	-9.3772	14.6563	-5.2791
3	-6.3029	-5.2791	11.5821

## Covariance Matrix for the Wilcoxon Statistics

GRUPO	1	2	3
1	55839.7	-30196.9	-25642.8
2	-30196.9	53607.4	-23410.5
3	-25642.8	-23410.5	49053.3

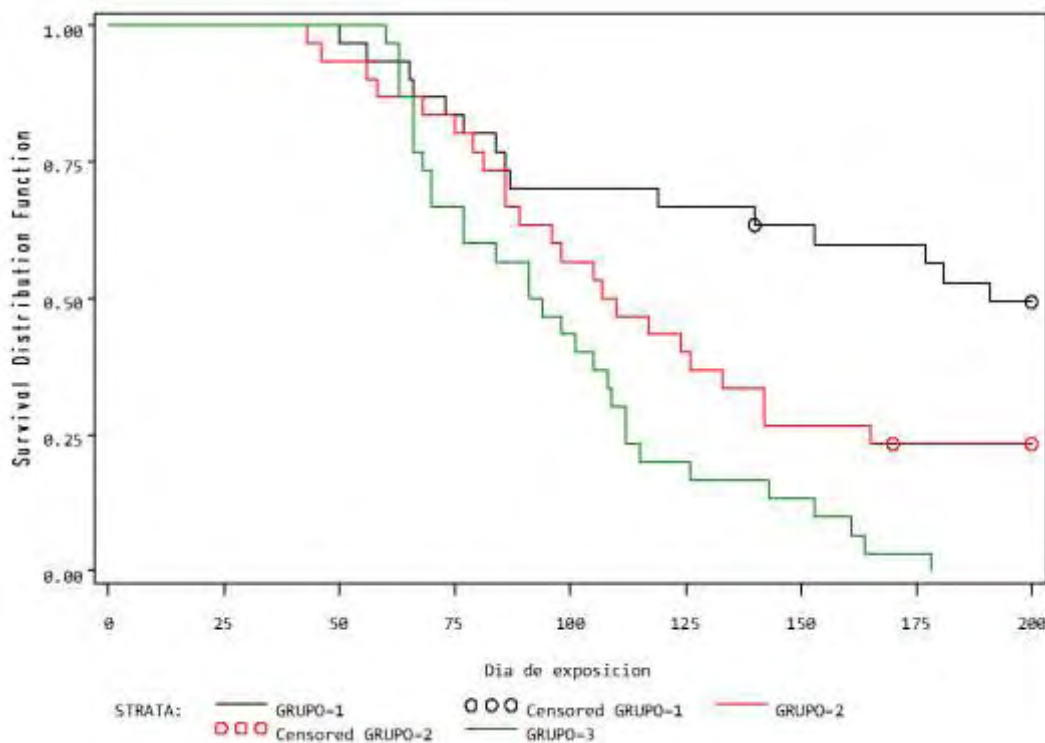
## Test of Equality over Strata

Test	Chi-Square	DF	Pr > Chi-Square
Log-Rank	20.7404	2	<.0001
Wilcoxon	12.3992	2	0.0020
-2Log(LR)	13.6909	2	0.0011

Esta prueba de K muestras, sólo le indica al investigador la significancia entre los grupos, pero no le indica cuales de los grupos contribuyen con la mayoría de la diferencia, es decir, si la dieta baja en grasa produce una diferencia significativa del tiempo libre de tumores comparado con las grasas saturadas o si las grasas saturadas son significativamente diferentes de la dieta grasas insaturadas. Todo lo que se puede concluir es que los datos por la prueba de Wilcoxon, muestran una diferencia altamente significativa ( $p = 0,002$ ) entre los tiempos libres de tumor de las tres dietas.



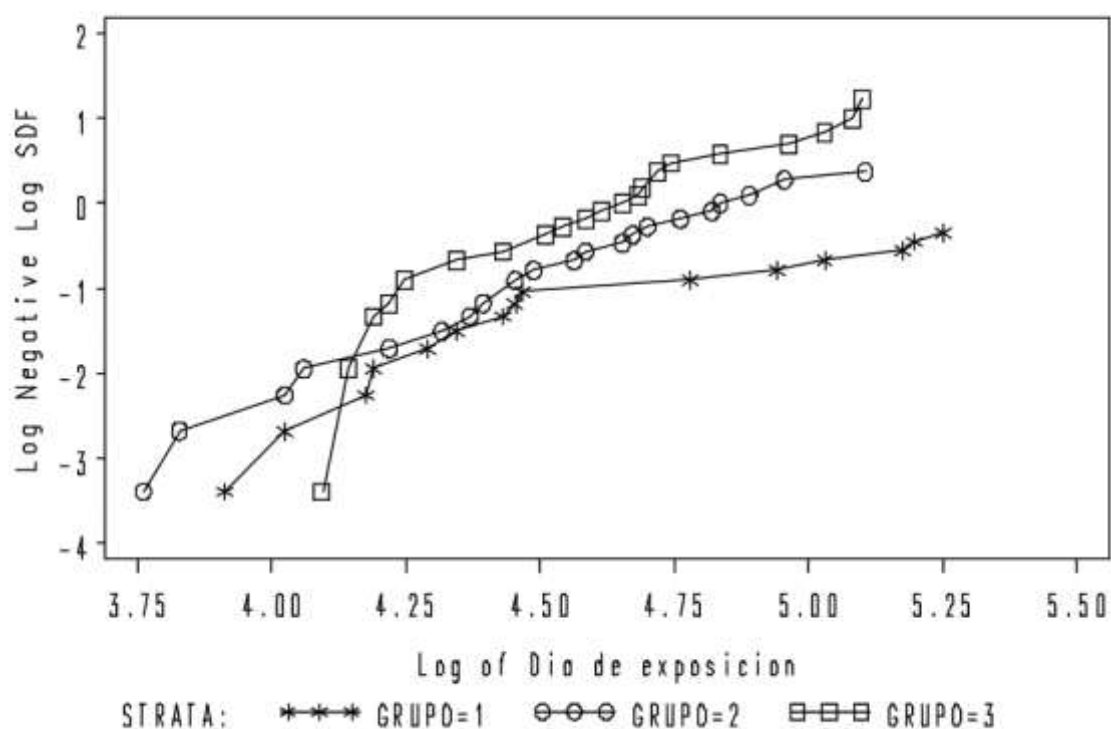
La opción de PLOT= (S) origina la Figura 39, en ella aparecen los estimados de la función de sobrevivencia  $S(t)$  para los individuos que recibieron los tres tratamientos. En esta figura, aparecen los estimados de la función de sobrevivencia  $S(t)$  para los individuos que recibieron los tratamientos. El valor de la mediana del tiempo de sobrevivencia (percentil 50 %) para los grupos también puede determinarse gráficamente. Por ejemplo, las curvas de sobrevivencia muestran que los individuos tratados con el tratamiento Grupo=1 tuvieron mejor experiencia de sobrevivencia (191 días) que los individuos de los tratamientos Grupo=2 (108,5 días) y Grupo=3 (92,5 días).



**Figura 39. Curvas de sobrevivencia para datos de ratas que recibieron dietas bajas en grasas (Grupo=1), grasas saturadas (Grupo=2) y grasas insaturadas (Grupo=3).** Fuente. Elaboración propia.

La opción de PLOT= (LLS) origina la Figura 40 para tratar de establecer un método gráfico común para probar el supuesto PH. Consiste en trazar  $-\log(-\log(S(t/x)))$  versus  $t$  o  $\log(t)$  para diferentes valores del vector covariable  $x$  (restringido a valores de  $t$  tal que  $0 < S(t) < 1$ ), donde  $S(t)$  es una tabla de vida o Función de superviviente de Kaplan-Meier. Bajo suposición PH, diferentes valores de covaria-

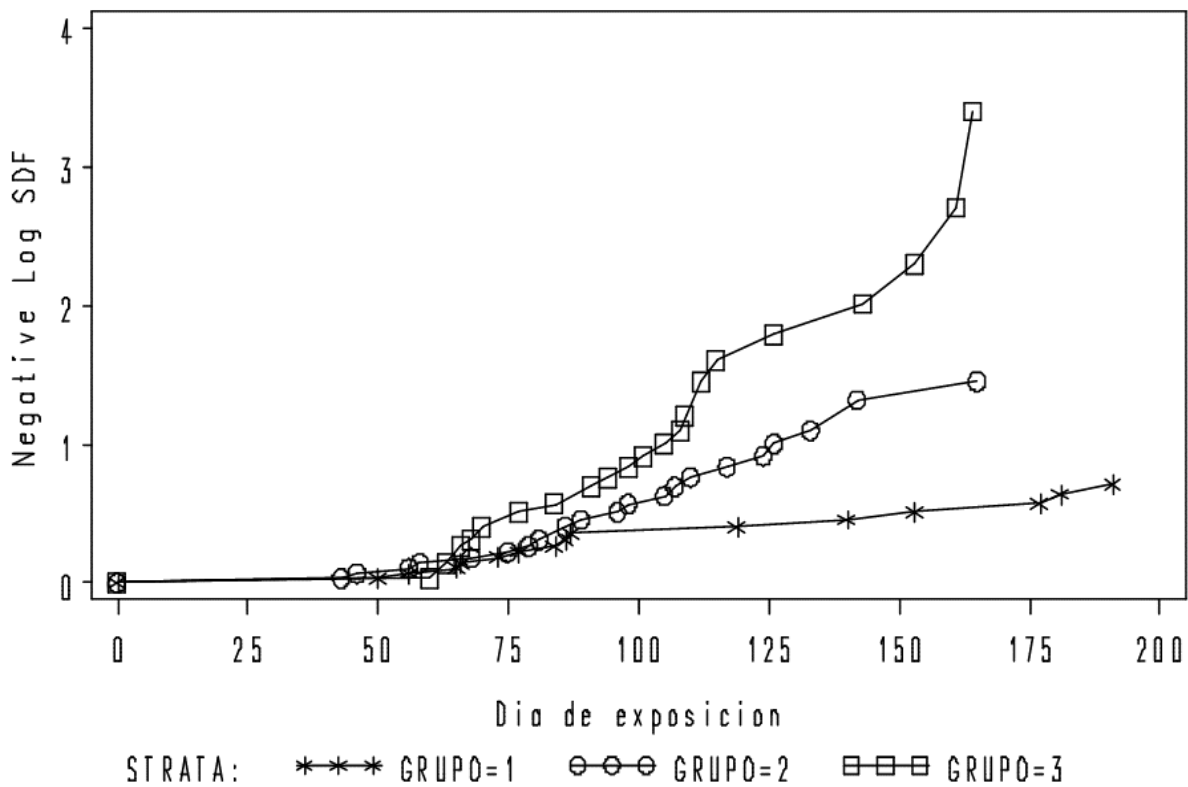
bles producirán funciones con la misma forma, pero diferentes elevaciones. El modelo de Cox es razonablemente robusto, por lo que solo es necesario preocuparse sobre claras desviaciones de la proporcionalidad, indicado por el cruce de las funciones para dos o más covariables en el gráfico de diagnóstico.



**Figura 40.** Curvas de supervivencia  $\log(-\log S(t))$  vs  $\log t$  para datos de ratas que recibieron dietas bajas en grasas (Grupo=1), grasas saturadas (Grupo=2) y grasas insaturadas (Grupo=3). Fuente. Elaboración propia.

La Figura 40 no muestran evidencia de cruce decisivo de las funciones de superviviente transformadas entre el Grupo=1 y Grupo=2, por lo que la suposición de PH es plausible para ambas covariables. Se observa un cruce del Grupo=3 sobre los Grupos 1 y 2, ocurrido antes de 70 días de exposición a la mortalidad.

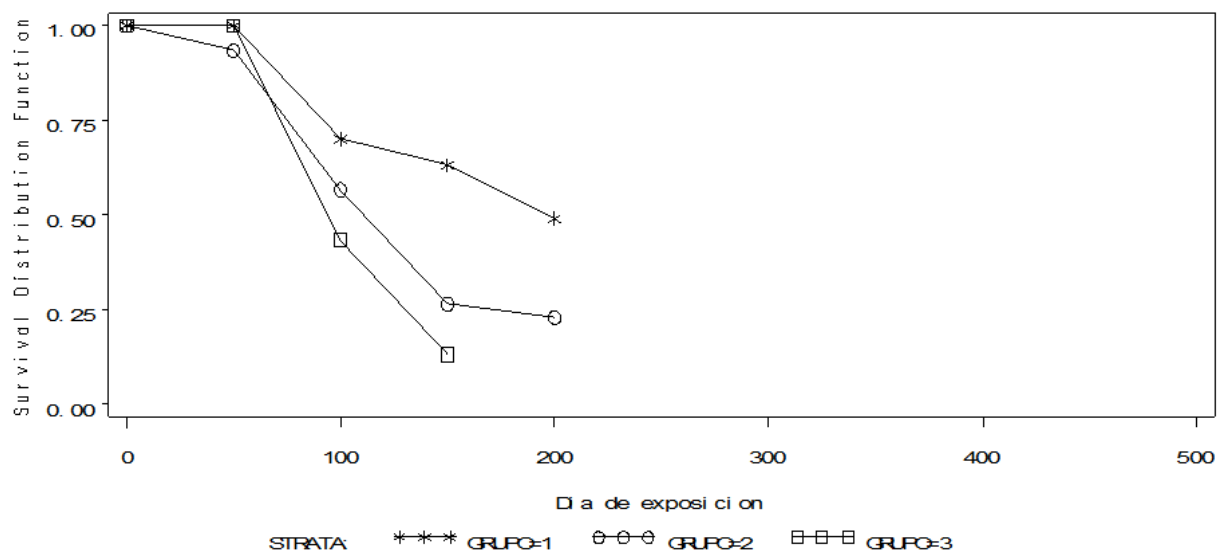
Según Allison (2010), la opción de graficación PLOTS= LLS, produce una gráfica de  $\log(-\log S(t))$  vs  $\log t$ , originando una línea recta si la verdadera distribución es Weibull. Por el contrario, PLOTS= LS produce una gráfica de  $-\log S(t)$  versus  $t$ . Si la verdadera distribución es exponencial, esta gráfica debe originar una línea con origen en cero (Figura 41), y puede ayudar en la selección de la distribución para analizar los datos con PROC LIFEREG.



**Figura 41. Funciones de supervivencia  $-\log S(t)$  versus  $t$  para datos de ratas que recibieron dietas bajas en grasas (Grupo=1), grasas saturadas (Grupo=2) y grasas insaturadas (Grupo=3).** Fuente. Elaboración propia.

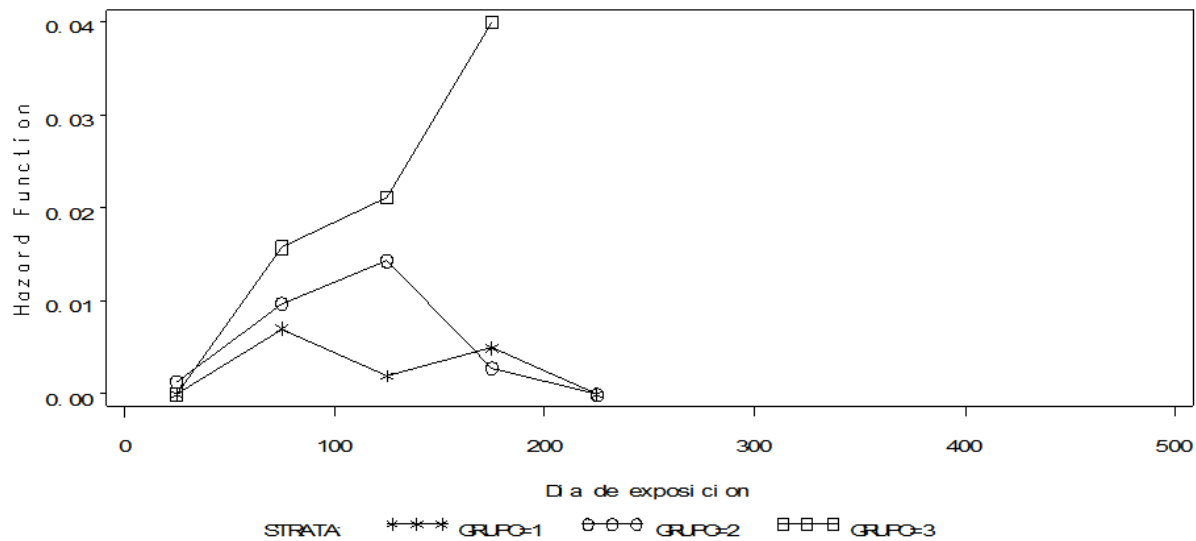
Las curvas de probabilidad de supervivencia de las ratas que recibieron dietas bajas en grasas (Grupo=1), grasas saturadas (Grupo=2) y grasas insaturadas (Grupo=3) se muestran en la Figura 42. Las ratas del Grupo 1, presentaron mayor probabilidad de supervivencia a lo largo del experimento, con pro-

babilidades de 0,70 y 0,63, en los intervalos 100 – 150 y 150 – 200 días, Al final del período de evaluación, en el intervalo 200 – 250 días, la probabilidad disminuyó a 0,49, pero superior a Grupos 2 y Grupo 3, con valores de probabilidad de 0,23 y 0,13, respectivamente. Las curvas de supervivencia bajo los cinco tratamientos fueron distintas (Wilcoxon:  $\chi^2 = 12,69$ ; gl = 2, p = 0,002).



**Figura 42. Tasas de supervivencia de ratas ( $\pm$  error estándar) que recibieron dietas bajas en grasas (Grupo=1), grasas saturadas (Grupo=2) y grasas insaturadas (Grupo=3). Fuente. Elaboración propia.**

La opción de graficación PLOTS= (H), produce la Figura 43, que indica que para GRUPO=3 la tasa de contagios es Exponencial desde el inicio del tratamiento y todas las ratas desarrollaron tumores dentro de los 200 días (el día 178 ocurrió el contagio final). Para GRUPO=2, durante los primeros 150 días la tasa de contagios fue más alta (Exponencial) y a partir de este tiempo disminuyó hasta el fin del experimento. Para GRUPO=1 la tasa de contagios fue menor durante el tiempo del experimento. Su tasa más alta se presentó antes de los 100 días permaneciendo relativamente constante hasta el final del experimento.



**Figura 43. Funciones de riesgo para datos de ratas que recibieron dietas bajas en grasas (Grupo=1), grasas saturadas (Grupo=2) y grasas insaturadas (Grupo=3).** Fuente. Elaboración propia.

#### e. Resultado de la prueba del supuesto de proporcionalidad con la declaración ASSESS.

Los valores de  $p$  se producen con la opción RESAMPLE, la cual genera una simulación de 1000 senderos. También produce una tabla resumen con los resultados para cada covariable. Al observar los resultados de la prueba para suposiciones de riesgo proporcional (*Supremum Test for Proportional Hazards Assumption*), de las dos covariables, solo la covariable GRUPO no presenta clara evidencia de violación de supuestos de riesgo proporcional ( $Pr = 0,1970$ ). Sin embargo, esta prueba no indica la naturaleza de la violación.

## The PHREG Procedure

## Model Information

Data Set	WORK.RATAS2	
Dependent Variable	DÍAS	Día de exposición
Censoring Variable	Status	Variable de censura
Censoring Value(s)	0	
Ties Handling	EFRON	

Number of Observations Read	90
Number of Observations Used	90

## Summary of the Number of Event and Censored Values

Total	Event	Censored	Percent Censored
90	68	22	24.44

## Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

## Model Fit Statistics

Criterion	Without Covariates	With Covariates
-2 LOG L	538.144	517.966
AIC	538.144	519.966
SBC	538.144	522.185

## Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	20.1777	1	<.0001
Score	20.5696	1	<.0001
Wald	19.2050	1	<.0001

Analysis of Maximum Likelihood Estimates						
Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
GRUPO	1	0.70616	0.16114	19.2050	<.0001	2.026

Supremum Test for Proportionals Hazards Assumption				
Variable	Maximum Absolute Value	Replications	Seed	Pr > MaxAbsVal
GRUPO	1.4769	1000	705491000	0.1970

#### f. Resultado de la prueba del supuesto de proporcionalidad con los residuales de Schoenfeld.

El uso principal de estos residuos es detectar posibles desviaciones del supuesto de riesgos proporcionales. Si se satisface la suposición, los residuos de Schoenfeld deben ser independientes del tiempo. La idea es probar si los residuos de Schoenfeld están correlacionados con el tiempo o con alguna función del tiempo (Allison, 2010).

The PHREG Procedure			
Model Information			
Data Set	WORK.RATAS2		
Dependent Variable	DÍAS	Día de exposición	
Censoring Variable	Status	Variable de censura	
Censoring Value(s)	0		
Ties Handling	EFRON		
Number of Observations Read			90
Number of Observations Used			90
Summary of the Number of Event and Censored Values			
Total	Event	Censored	Percent Censored
90	68	22	24.44

Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Without Covariates	With Covariates
-2 LOG L	538.144	517.966
AIC	538.144	519.966
SBC	538.144	522.185

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	20.1777	1	<.0001
Score	20.5696	1	<.0001
Wald	19.2050	1	<.0001

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
GRUPO	1	0.70616	0.16114	19.2050	<.0001	2.026

The CORR Procedure

4 Variables: DÍAS DÍAS1 DÍAS2 schGRUPO

Simple Statistics

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
DIASrank	90	45.50	25.98	4095	1.00000	80.50
DÍAS1	90	4.72	0.4495	424.68	3.76120	5.29832
DÍAS2	90	17975	14069	1617777	1849	40000
schGRUPO	68	-0.00367	0.78567	-0.24945	-1.4507	1.3686

Simple Statistics

Variable	Label
DIASrank	Día de exposición jerarquizado
DÍAS1	Logaritmo de días
DÍAS2	Días al cuadrado

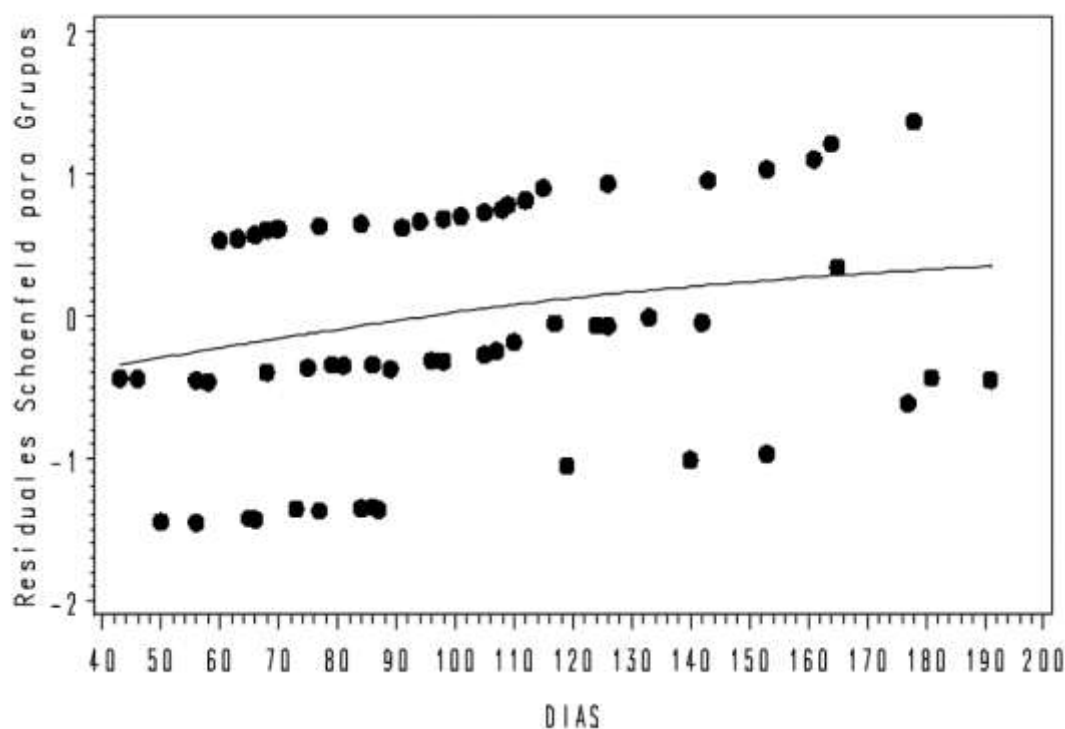


schGRUPO      Schoenfeld Residual para GRUPO  
SCHX            Schoenfeld Residual para X

Los resultados mostrados a continuacion, como producto de los residuales de Schoenfeld, coinciden con los resultados originados por ASSESS. Nuevamente, no se presentan problemas de riesgo no proporcional con la covariable GRUPO (con valores de  $p > 0,05$  en dos de las funciones relacionadas con el tiempo,  $p=0,0608$  y  $p=0,1033$ ). También se observa que el tamaño de muestra de estas correlaciones es de 68, ya que los residuales de Schoenfeld no están definidos para los casos censurados (Allison, 2010).

Pearson Correlation Coefficients				
Prob >  r  under H0: Rho=0				
Number of Observations				
	DIASRANK	DÍAS1	DÍAS2	schGRUPO
DIASRANK	1.00000	0.99579	0.95324	0.25282
		<.0001	<.0001	0.0608
	90	90	90	68
DÍAS1	0.99579	1.00000	0.94885	0.25236
	<.0001		<.0001	0.0379
	90	90	90	68
DÍAS2	0.95324	0.94885	1.00000	0.19925
	<.0001	<.0001		0.1033
	90	90	90	68

En la Figura 44, los residuales de Schoenfeld muestran una tendencia recta, línea recta que tiene una pendiente de  $15 - 20^\circ$ , lo que indica que las suposiciones de riesgo proporcional no se satisfacen completamente y son dependientes de la variable de tiempo. Por ejemplo, se satisfacen con las variables DIASRANK y DÍAS2 ( $p=0,06$  y  $p=0,10$ ), pero no se satisface para la variable DÍAS1 ( $p=0,037$ ) y puede sugerir poca violación de las suposiciones para GRUPOS.



**Figura 44.** Gráfica de residuos de Schoenfeld para la interacción entre el tiempo (DÍAS) y la covariable GRUPOS. Fuente. Elaboración propia.

El siguiente método para probar el cumplimiento de supuestos de riesgo proporcional es mediante interacciones con covariables dependientes del tiempo. En este caso se ilustrará como estimar un modelo Cox que permita que el efecto de los GRUPOS cambie con el tiempo, originando una prueba definitiva de la suposición de proporcionalidad para esta variable.

#### **g. Resultado inicial sin las Interacciones con el tiempo para los primeros 100 días del experimento**

El primer resultado analiza el modelo de estimación del tiempo de mortalidad, incluyendo el modelo simple que involucra solo los efectos simples sobre la covariables GRUPO. Las tres pruebas de la salida de resultados "Testing Global Null Hypothesis: BETA=0", indican que las curvas de contagios para los tres grupos de tratamientos pueden no resultar iguales entre los GRUPO 1 y GRUPO 2, comparados con el GRUPO 3 ( $p < 0,05$ ). En este modelo, la razón de riesgo para GRUPO, definido como el expo-

nencial de su coeficiente de regresión, es la razón de la función de riesgo entre los tres grupos. En el ejemplo, el GRUPO=1 tiene un valor de razón de riesgo (HR) de 0,240 y GRUPO=2 de 0,526. Cuando hay múltiples niveles de una variable de clase (en este caso, GRUPOS), SAS toma un nivel (el último) como nivel de referencia (GRUPO = 3), con un coeficiente con valor de cero (0) y todos los demás se comparan con este nivel. Un valor significativo de chi cuadrada para un nivel particular indica que solo este nivel es significativamente diferente del nivel de referencia y puede o no, ser significativamente diferente de otros niveles, requiriendo múltiples comparaciones (Allison, 2010).

The PHREG Procedure

Model Information

Data Set	WORK.RATAS2	
Dependent Variable	DÍAS	Día de exposición
Censoring Variable	Status	Variable de censura
Censoring Value(s)	0	
Ties Handling	EFRON	

Number of Observations Read	90
Number of Observations Used	90

Summary of the Number of Event and Censored Values

Total	Event	Censored	Percent Censored
90	68	22	24.44

Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Without Covariates	With Covariates
-2 LOG L	538.144	517.893
AIC	538.144	521.893
SBC	538.144	526.332

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	20.2508	2	<.0001
Score	20.8666	2	<.0001
Wald	18.8504	2	<.0001

Type 3 Tests									
		Effect	DF	Wald					
				Chi-Square	Pr >	ChiSq			
		GRUPO	2	18.8504	<.0001				
Analysis of Maximum Likelihood Estimates									
Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% HR Confid Limit		
GRUPO 1	1	-1.42865	0.33170	18.5505	<.0001	0.240	0.125	0.459	
GRUPO 2	1	-0.64326	0.28246	5.1862	0.0228	0.526	0.302	0.914	

En este modelo inicial, el signo negativo de los coeficientes de regresión para GRUPO =1 y GRUPO =2 ( $\beta_1 = -1,42865$  y  $\beta_2 = -0,64326$ ), indica que ambos grupos tuvieron menor probabilidad de contagios que el GRUPO=3. El coeficiente exponenciado,  $\exp(\beta_i)$  representa la razón de riesgo y da el tamaño del efecto de la covariable. Los valores de la razón de riesgo de ambos grupos (0,240 y 0,526) resultaron  $<1$  e indica que el riesgo de contagios de las ratas del GRUPO 1 es solo alrededor del 24 % del riesgo de las ratas del GRUPO 3, y para el GRUPO 2 es alrededor del 52,6 % del riesgo de las ratas del GRUPO 3.

El análisis de los parámetros estimados se muestran a continuación. El valor del estadístico chi cuadrado de Wald para probar la hipótesis nula que  $\beta_1$  y  $\beta_2 = 0$ , son 18,6 y 5,2, respectivamente. Ambos valores son estadísticamente significativos ( $p < 0,05$ ) al compararlo con la distribución chi cuadrada con un grado de libertad ( $p < 0,0001$  y  $p = 0,0228$ ). Por lo tanto, se puede concluir que existe evidencia de una tendencia a disminuir la razón de riesgo en el tiempo y en consecuencia no presentar una razón de riesgo constante.

#### **h. Resultado con las interacciones con el tiempo para los primeros 100 días del experimento (Días < 100).**

Las covariables dependientes del tiempo son aquellas que pueden cambiar sus valores en el curso de las observaciones. El segundo resultado analiza el modelo de estimación del tiempo de contagio, incluyendo el modelo que involucra los efectos de la interacción con el tiempo sobre la covariable GRUPO, durante los primeros 100 días del experimento ( $\text{GrupDias} = \text{Grupo} * (\text{Día} < 100)$ ). Las tres pruebas de

la salida de resultados "Testing Global Null Hypothesis: BETA=0", indican que las curvas de contagio para los tres grupos de tratamientos pueden resultar significativamente diferentes entre GRUPO 1 y GRUPO 2, comparados con GRUPO 3 ( $p < 0,0001$ ). En este modelo, la razón de riesgo para los grupos, el GRUPO=1 tiene un valor de razón de riesgo (HR) de 0,103, GRUPO=2 de 0,327 y la interacción GRUPDIAS de 0,479. Con este resultado se observa que al introducir la interacción la razón de riesgo disminuye un 50 % en relación al procedimiento anterior (sin interacción).

The TPHREG Procedure

Model Information

Data Set	WORK.RATAS2	
Dependent Variable	DÍAS	Día de exposición
Censoring Variable	Status	Variable de censura
Censoring Value(s)	0	
Ties Handling	EFRON	

Number of Observations Read	90
Number of Observations Used	90

Class Level Information

Class	Value	Design Variables	
		1	0
GRUPO	1	1	0
	2	0	1
	3	0	0

Summary of the Number of Event and Censored Values

Total	Event	Censored	Percent Censored
90	68	22	24.44

Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Without Covariates	With Covariates
-2 LOG L	538.144	512.762
AIC	538.144	518.762
SBC	538.144	525.420

Testing Global Null Hypothesis: BETA=0								
Test		Chi-Square	DF	Pr > ChiSq				
Likelihood Ratio		25.3817	3	<.0001				
Score		28.7146	3	<.0001				
Wald		23.6358	3	<.0001				
Type 3 Tests								
Effect		DF	Wald Chi-Square	Pr > ChiSq				
GRUPO		2	19.7618	<.0001				
GRUPDIAS		1	5.0962	0.0240				
Analysis of Maximum Likelihood Estimates								
Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% HR Confid Limit	
GRUPO 1	1	-2.27653	0.51508	19.5343	<.0001	0.103	0.037	0.282
GRUPO 2	1	-1.11840	0.34406	10.5663	0.0012	0.327	0.167	0.641
GRUPDIAS	1	-0.73670	0.32634	5.0962	0.0240	0.479	0.253	0.907

Durante los primeros 100 días del experimento, el signo negativo de los coeficientes de regresión para GRUPO =1 y GRUPO =2 ( $\beta_1 = -2,27653$  y  $\beta_2 = -1,1184$ ), indica que ambos grupos tuvieron menor probabilidad de contagio que el GRUPO=3. El coeficiente exponenciado,  $\exp(\beta_i)$  representa la razón de riesgo y da el tamaño del efecto de la covariable. Los valores de la razón de riesgo de ambos grupos (0,103 y 0,327) resultaron <1 e indica que el riesgo de contagios de las ratas del GRUPO 1 es de alrededor del 10 % del riesgo de las ratas del GRUPO 3, y para el GRUPO 2 es alrededor del 32,7 % del riesgo de las ratas del GRUPO 3.

En la comparación del GRUPO 1 y GRUPO 2 con el nivel de referencia (GRUPO 3), indica diferencias altamente significativas ( $p < 0,001$ ). Un valor significativo de chi cuadrada para un nivel particular indica que solo este nivel es significativamente diferente del nivel de referencia y puede o no, ser significativamente diferente de otros niveles, requiriendo múltiples comparaciones (Allison 2010). A continuación, el resultado de la prueba que compara el GRUPO 1 con el GRUPO 2, indicando que entre ambos grupos se presentan diferencias altamente significativas ( $p = 0,003$ ).

Linear Hypotheses Testing Results			
Label	Wald Chi-Square	DF	Pr > ChiSq
Test 1	8.8271	1	0.0030

Para averiguar el efecto de los tratamientos en el contagio de las ratas durante los primeros 100 días, se agregó la interacción con el tiempo al efecto GRUPOS. El término de interacción (GRUPDIAS) es significativamente diferente ( $p = 0,024$ ) durante los primeros 100 días, confirmando que durante este período existen problemas de violación de supuestos. Ahora se dispone de una vía para probar las violaciones de riesgo proporcional: simplemente se agregó al modelo una variable dependiente del tiempo que representa la interacción de la variable original con el tiempo. Si la interacción no tiene un coeficiente significativo, se concluye que la suposición de riesgo proporcional no es violada por esta variable. Por otro lado, si la interacción tiene un coeficiente significativo se tiene evidencia de no proporcionalidad, tal como se presenta durante los primeros 100 días.

**i. Resultado con las interacciones con el tiempo para los últimos 100 días del experimento (Días > 100).**

El siguiente resultado analiza el modelo de estimación del tiempo de contagio, incluyendo el modelo que involucra los efectos de la interacción con el tiempo sobre la covariable GRUPO, durante los últimos 100 días del experimento ( $\text{GrupDias} = \text{Grupo} * (\text{Día} > 100)$ ). Las tres pruebas de la salida de resultados "Testing Global Null Hypothesis: BETA=0", indican diferencias altamente significativas entre las curvas de los grupos, además que las curvas de contagio para el GRUPO 1 puede resultar diferentes a la curva del GRUPO 3 ( $p = 0,05$ ) y las del GRUPO 2, igual a la curva de contagio del GRUPO 3 ( $p = 0,2178$ ). En este modelo, la razón de riesgo aumentó para los grupos, con el GRUPO=1 con un valor de razón de riesgo (HR) de 0,448 y GRUPO=2 de 0,683.

## The TPHREG Procedure

## Model Information

Data Set	WORK.RATAS2	
Dependent Variable	DÍAS	Día de exposición
Censoring Variable	Status	Variable de censura
Censoring Value(s)	0	
Ties Handling	EFRON	

Number of Observations Read	90
Number of Observations Used	90

## Class Level Information

Class	Value	Design Variables	
GRUPO	1	1	0
	2	0	1
	3	0	0

## Summary of the Number of Event and Censored Values

Total	Event	Censored	Percent Censored
90	68	22	24.44

## Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

## Model Fit Statistics

Criterion	Without Covariates	With Covariates
-2 LOG L	538.144	512.762
AIC	538.144	518.762
SBC	538.144	525.420

## Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	25.3817	3	<.0001
Score	28.7146	3	<.0001
Wald	23.6358	3	<.0001



## Type 3 Tests

Effect	DF	Wald Chi-Square	Pr > ChiSq
GRUPO	2	3.8798	0.1437
GRUPDIAS	1	5.0962	0.0240

## Analysis of Maximum Likelihood Estimates

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% HR Confid Limit
GRUPO 1	1	-0.80313	0.40981	3.8406	0.0500	0.448	0.201 1.000
GRUPO 2	1	-0.38170	0.30970	1.5191	0.2178	0.683	0.372 1.253
GRUPDIAS	1	0.73670	0.32634	5.0962	0.0240	2.089	1.102 3.960

Durante los últimos 100 días del experimento, el signo negativo de los coeficientes de regresión para GRUPO =1 y GRUPO =2 ( $\beta_1 = -0,80313$  y  $\beta_2 = -0,3817$ ), indica que ambos grupos continuaron con menor probabilidad de contagio que el GRUPO=3. Los valores de la razón de riesgo de ambos grupos (0,448 y 0,683), indican que ambos grupos tienen  $HR < 1$ , que significa menor riesgo en el tratamiento de intervención que en el control (tratamiento de referencia) e indica que el riesgo de contagios de las ratas del GRUPO 1 es alrededor del 44,8 % del riesgo de las ratas del GRUPO 3, y para el GRUPO 2 es alrededor del 68,3 % del riesgo de las ratas del GRUPO 3.

En la comparación del GRUPO 1 con el GRUPO 3, indica diferencias significativas ( $p = 0,05$ ) y la comparación del GRUPO 2 con el GRUPO 3, indica que no existen diferencias significativas ( $p = 0,2178$ ). A continuación, el resultado de la prueba que compara el GRUPO 1 con el GRUPO 2, indicando que, en este período, entre ambos grupos no existen diferencias significativas ( $p = 0,2346$ ).

## Linear Hypotheses Testing Results

Label	Wald Chi-Square	DF	Pr > ChiSq
Test 1	1.4125	1	0.2346

**j. Resultado con las interacciones con el tiempo para el total de días del experimento (200 días)**

El siguiente resultado analiza el modelo de estimación del tiempo de contagio, incluyendo el modelo que involucra los efectos de la interacción con el tiempo sobre la covariable GRUPO, durante todo el período del experimento (200 días). Las tres pruebas de la salida de resultados "Testing Global Null Hypothesis: BETA=0", indican que las curvas de contagio para el GRUPO 1 y las del GRUPO 2, pueden resultar iguales a la curva del GRUPO 3 ( $p = 0,6733$  y  $p = 0,7135$ , respectivamente). En este modelo, la razón de riesgo aumentó para los grupos, con el GRUPO=1 con un valor de razón de riesgo (HR) de 1,485 contagios/día y GRUPO=2 de 1,205 contagios/día.

```

The TPHREG Procedure
Model Information
Data Set          WORK.RATAS2
Dependent Variable  DÍAS           Día de exposición
Censoring Variable  Status         Variable de censura
Censoring Value(s)  0
Ties Handling       EFRON

Number of Observations Read      90
Number of Observations Used     90

Class Level Information
Class      Value      Design
                Variables
GRUPO      1          1      0
            2          0      1
            3          0      0

Summary of the Number of Event and Censored Values
Total      Event      Censored      Percent
                Censored
90          68          22          24.44

```

## Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

## Model Fit Statistics

Criterion	Without Covariates	With Covariates
-2 LOG L	538.144	513.878
AIC	538.144	519.878
SBC	538.144	526.537

## Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	24.2655	3	<.0001
Score	28.3272	3	<.0001
Wald	23.7377	3	<.0001

## Type 3 Tests

Effect	DF	Wald Chi-Square	Pr > ChiSq
GRUPO	2	0.1779	0.9149
GRUPDIAS	1	4.1227	0.0423

## Analysis of Maximum Likelihood Estimates

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% HR Confid Limit
GRUPO 1	1	0.39537	0.93763	0.1778	0.6733	1.485	0.236 9.329
GRUPO 2	1	0.18662	0.50819	0.1349	0.7135	1.205	0.445 3.263
GRUPDIAS	1	0.00914	0.00450	4.1227	0.0423	1.009	1.000 1.018

Para todo el período del experimento, los coeficientes de regresión presentan signos positivos, siendo para GRUPO =1 ( $\beta_1 = 0,39537$ ) y para GRUPO =2 ( $\beta_2 = 0,18662$ ), indica que ambos grupos presentan igual probabilidad de contagio que el GRUPO=3. Para todo el periodo de observación (200 días), se observa que la interacción con el tiempo también tuvo un impacto positivo. Es decir, la interacción aumento significativamente las probabilidades de contagio en el tiempo ( $p=0,0423$ ). Los valores de la razón de riesgo de ambos grupos (1,485 y 1,205). Una HR > 1 indica más riesgo de contagio en el tratamiento que en el nivel de referencia e indica que el riesgo de contagios de las ratas del GRUPO 1 es solo alrededor del 48,5 % del riesgo de las ratas del GRUPO 3, y para el GRUPO 2 es alrededor del 20,5 % del riesgo de las ratas del GRUPO 3.

En la comparación del GRUPO 1 y GRUPO 2 con el GRUPO 3, indica que no existen diferencias significativas ( $p = 0,6733$  y  $p = 0,7135$ , respectivamente). El resultado de la prueba que compara el GRUPO 1 con el GRUPO 2, indica que entre ambos grupos no existen diferencias significativas ( $p = 0,7122$ ).

#### Linear Hypotheses Testing Results

Label	Wald Chi-Square	DF	Pr > ChiSq
Test 1	0.1361	1	0.7122

#### k. Resultado del análisis paramétrico.

Se corrieron las distribuciones Llogística – Exponential – Weibull – Lognormal y Gamma, resultando la distribución Lognormal con la mejor convergencia de los datos, seguida de la distribución Llogística. Con ello se confirma la curva definida con PLOTS= LS, la cual produce la gráfica de  $-\log S(t)$  versus  $t$ , indicando que, si la verdadera distribución es exponencial, esta gráfica debe originar una línea con origen en cero (Figura 40).

A continuación, los resultados para el modelo Lognormal ajustados para cada grupo. La salida de resultados primero provee información preliminar: nombres de las variables de tiempo y censura, valores correspondientes a censurados y números de casos por tipo de censura. Luego la salida muestra que el log de probabilidad para el modelo es 1,99, valor estadístico que se utilizará posteriormente para probar varias hipótesis.

La salida “Type III Analysis of Effects” del modelo Exponencial, tendrá importancia al considerar variables de clasificación (CLASS). También se reportan los coeficientes estimados, los errores estándar, chi cuadrado para la hipótesis nula que cada coeficiente es 0 y valores de  $p$  asociados con este estadístico.

## The LIFEREG Procedure

## Model Information

Data Set	WORK.RATAS2	
Dependent Variable	Log (DÍAS)	Día de exposición
Censoring Variable	Status	Variable de censura
Censoring Value(s)	0	
Number of Observations	90	
Noncensored Values	68	
Right Censored Values	22	
Left Censored Values	0	
Interval Censored Values	0	
Name of Distribution	<b>Lognormal</b>	
Log Likelihood	1.9927359721	
Number of Observations Read		90
Number of Observations Used		90

## Class Level Information

Name	Levels	Values
GRUPO	3	1 2 3

Algorithm converged.

## Type III Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
GRUPO	2	314.4798	<.0001
GRUPDIAS	1	358.4100	<.0001

## Analysis of Parameter Estimates

Parameter	DF	Estimate	Standard Error	95% Confidence Limits		Chi- Square	Pr > ChiSq
Intercept	1	3.0905	0.0839	2.9261	3.2550	1356.52	<.0001
GRUPO	1	1.1593	0.0657	1.0305	1.2880	311.35	<.0001
GRUPO	2	0.4897	0.0561	0.3799	0.5996	76.34	<.0001
GRUPO	3	0.0000	.	.	.	.	.
GRUPDIAS	1	0.0049	0.0003	0.0044	0.0054	358.41	<.0001
Scale	1	0.1903	0.0169	0.1600	0.2265		

Los signos de los coeficientes para GRUPOS indican la dirección de la relación. El signo positivo logrado por estas variables indica que el efecto del tiempo libre de tumores se alargó en el tiempo y, en consecuencia, con el desarrollo de tumores. El coeficiente de regresión para el tratamiento GRUPO 3, es 0. Su razón es que se designó como nivel de referencia para comparar con los demás grupos. El análisis del resultado de los parámetros estimados, indica que cada coeficiente de GRUPO 1 y GRUPO 2, de manera independiente, difiere significativamente del GRUPO 3 ( $p < 0,0001$ ).

En la estimación de los parámetros ATF para el modelo Lognormal, la intersección está representado por  $\beta_0 = 3,0905$ , GRUPO 1 ( $\beta_1 = 1,1593$ ) y GRUPO 2 ( $\beta_2 = 0,4897$ ). Según Romano & Stevanato (2020), estos valores se utilizan para calcular el tiempo medio de contagio ( $T_{50}$ ). El  $T_{50}$  (mediana o tiempo medio de contagio) para referencia de cada parámetro, se calcula como  $\exp(\beta)$ . Para todo el lote (intercepto =  $\beta_0 = 3,0905$ ), fue de  $\exp(3,0905) = 21,98$  días, mientras que para el efecto de GRUPO 1 es,  $\exp(\beta_0 + \beta_1) = \exp(3,0905 + 1,1593) = 70,1$  d; para GRUPO 2,  $\exp(\beta_0 + \beta_2) = \exp(3,0905 + 0,4897) = 35,88$  días y para GRUPO 3,  $\exp(\beta_0 + \beta_3) = \exp(3,0905 + 0,000) = 21,98$  días. De esta manera, se calcula el tiempo medio de contagios para cada grupo, según su parámetro. Cuanto más negativo sea el coeficiente, más temprano será el tiempo medio de contagio (Romano y Stevanato, 2020). Los resultados de la prueba chi cuadrada indican diferencias altamente significativas entre estas comparaciones, con el tratamiento de referencia. Por coincidencia, las comparaciones del análisis de las tablas de vida (LIFE-TES) sugieren que existen diferencias altamente significativas entre los grupos ( $p = 0,002$ ).

Para interpretar los parámetros de regresión, hay que recordar que corresponden a los efectos multiplicativos de las covariables sobre la probabilidad de contagios. El tiempo promedio de contagio de cada grupo se calcula por  $\exp(\beta_0 + \beta_i) * P_0(t)$ , donde  $\beta_0$  corresponde al intercepto,  $\beta_i$  al coeficiente de regresión y  $P_0(t)$  al parámetro de escala de la distribución (Fox 2001). Así, para GRUPO 1 es,  $\exp(3,0905 + 1,1593) = 70,1 * P_0(t) = 70,1 * 0,1903 = 13,3$  días, mientras que para GRUPO 2 es,  $\exp(3,0905 + 0,4897) = 35,88$  días \*  $P_0(t) = 35,9 * 0,1903 = 6,8$  días. Para el GRUPO 3, el tiempo promedio de contagio sería  $\exp(3,0905 + 0,000) * P_0(t) = 21,9 * 0,1903 = 4,2$  días. El análisis de esta información indica que, individualmente, las ratas del GRUPO 3 lograron un contagio más temprano.

En términos de la razón de tiempo ( $\gamma$ ), para GRUPO 1  $\gamma = 70,1/21,98 = 3,192$ ; para GRUPO 2,  $\gamma = 35,88/21,98 = 1,6323$  y para GRUPO 3,  $\gamma = 21,98/21,98 = 1,000$ . Estos valores, interpretados como razones estimadas de los tiempos de contagio esperados para dos grupos, según la fórmula  $(100 \% (1 - \gamma))$ , indica que los GRUPO 1 y GRUPO 2, retardan o prolongan el contagio en 219,2 % y en 63,23 %, en comparación con el GRUPO 3 (tratamiento de referencia), respectivamente. Esto significa que el tratamiento que prolonga u origina un contagio más lento, es el GRUPO 1. El GRUPO 3 se considera de referencia, por lo que aparece con un valor neutro (0,00). En general, en los modelos AFT, una razón de tiempo  $> 1$  para una covariable, implica que prolonga el tiempo de contagio, mientras que una razón de tiempo  $< 1$  implica que es más probable un evento de contagio más temprano (Romano & Stevanato 2020).

El hecho que cada coeficiente de regresión de manera independiente es significativo, indica que cada uno difiere del tratamiento de referencia. Estas comparaciones múltiples, son análogas a las del análisis LIFETEST, excepto que en este caso se utilizaron las pruebas de Wilcoxon y en el ATF, los coeficientes de regresión. El análisis LIFETEST mostró heterogeneidad entre grupos, con diferencias entre ellos. Los coeficientes de regresión llegan a una conclusión semejante: la comparación de los demás tratamientos con el tratamiento de referencia, es significativamente diferente.

# EPÍLOGO

En el manual se utilizan tres procedimientos para el análisis de crecimiento de organismos biológicos. Estos procedimientos son: análisis no paramétrico, semiparamétrico y paramétrico.

Su utilidad, en el caso de las plantas, parten desde el proceso de germinación hasta cualquier variable relacionada o que afecte su crecimiento. Por ejemplo, tratamientos aplicados para combatir una enfermedad. En el caso de insectos se puede evaluar aplicación de insecticidas y su efecto en la mortalidad o dietas alimenticias y su efecto en el crecimiento. Es decir, a cualquier organismo biológico se le puede evaluar el crecimiento relacionado con tratamientos o variables ambientales.

Para el análisis se hace uso del concepto de *seguimiento*, ya sea de pruebas de germinación en muestras de semilla con diferentes tratamientos pregerminativos o en otros estudios de investigación, donde lo que se valora es el tiempo transcurrido desde un momento inicial, como el inicio de los tratamientos hasta un tiempo final en el que finaliza la recolección de los datos, donde se puede demostrar o no la aparición de un suceso o evento de interés. A este tiempo habitualmente se lo conoce como “*tiempo de supervivencia*”: la ocurrencia o no de la germinación o de la muda o la mortalidad en el caso de insectos o la muerte o no, en el caso de otros organismos biológicos. El conjunto de técnicas estadísticas que se usan para analizar estos datos se conoce como “*análisis de supervivencia*” porque se diseñaron inicialmente para estudiar el tiempo transcurrido hasta el fallecimiento de un paciente, principalmente en el campo de la oncología, para proponer tratamientos terapéuticos y su seguimiento en un tiempo determinado o hasta la muerte del paciente.

La característica más importante de este tipo de datos (tiempo transcurrido hasta la aparición del suceso) es que, muy probablemente, al final del período de observación, no todos los individuos habrán presentado el suceso objeto de estudio y por lo tanto el tiempo hasta su ocurrencia es desconocido. A este fenómeno se lo denomina *censura* (casos censurados).



En el manual, se presentan algunos de los procedimientos utilizados para el cálculo y desarrollo de análisis de sobrevivencia en algunas especialidades biológicas. En vista de la utilidad de los análisis de sobrevivencia para casos particulares de la investigación biológica, el manual contribuye con su uso a través del paquete estadístico SAS, con el cual se analizan y agilizan estructuras de datos que involucran diferentes variables del crecimiento, tomando como ejemplo la germinación de semillas, utilizando diferentes tratamientos pregerminativos, evaluados durante el proceso de la germinación (tiempo). Se analizan otros casos de investigación biológica relacionados con el uso de dietas alimenticias y su efecto en el desarrollo de fases e instares larvales en insectos. Además, un estudio para investigar la relación entre dieta alimenticia y desarrollo de tumores en ratas de laboratorio. En el manual, como introducción al manejo de datos, se hizo una revisión bibliográfica de artículos científicos relacionados con el análisis de sobrevivencia en plantas y otros organismos biológicos. En todos los ejemplos señalados, se percibe la aplicación práctica del análisis de sobrevivencia como una herramienta ideal para la exploración y descripción de los datos en el ajuste de curvas de sobrevivencia. En el texto del manual, se incluye la aplicación práctica de las metodologías utilizadas para el desarrollo de tablas de vida, considerando los siguientes datos como producto del desarrollo metodológico: la edad o fase inicial de los miembros de la población o muestra, el número de individuos vivos o semilla germinada, el número de individuos muertos o semilla sin germinar, el número de individuos vivos en el intervalo, el número de individuos vivos el intervalo más los individuos vivos en los intervalos inferiores (conteo acumulado) y, por último, la *esperanza media de vida*, que se estima al dividir el número de individuos vivos en el intervalo más los individuos vivos en los intervalos inferiores entre el número de individuos en la muestra original. La esperanza media de vida informa de cuantas unidades de tiempo le quedan de vida a los individuos en cada intervalo por considerar. Con estos cálculos es posible predecir cuál segmento de la población, de acuerdo con su edad, tiene mayores expectativas de vida.

Se pudo calcular la *mediana de supervivencia*, que no requiere conocer el tiempo de supervivencia de todos los individuos y que se estima con el *percentil 50* de la distribución, que corresponde al primer tiempo con una proporción de supervivencia (probabilidad acumulada de supervivencia) igual o inferior a 0,5: sólo se trata de proyectar la probabilidad de supervivencia del 0,5 sobre la curva y comprobar a qué tiempo le corresponde.

Las curvas de supervivencia se originaron usando los dos métodos siguientes: el método del límite de producto de Kaplan-Meier o el análisis actuarial de Cutler-Ederer. 1) *Método de Kaplan-Meier*: si se asume que el evento terminal es independiente para cada individuo (semilla, larva, ratón, planta, etc.), las probabilidades de sobrevivir en un tiempo  $t$  determinado se calcularon por la ley multiplicativa de probabilidades. La característica distintiva es que la proporción acumulada que sobrevivió se calculó para el tiempo de supervivencia individual de cada individuo y no se agruparon los tiempos de supervivencia en intervalos. Debido a ello, es especialmente útil en estudios con número pequeño de individuos.

Se conoce que existen distintos tipos de curva de supervivencia, dependientes de la condición evaluada. Por ejemplo, mortalidad, germinación, presencia o no de una enfermedad, etc. La curva *Tipo I* describe la situación en que la mortalidad, supervivencia, presencia o no de algún evento de interés, según sea el caso, se halla concentrada al final del tiempo de análisis de supervivencia: representa una baja mortalidad o supervivencia y es la curva ideal de cualquier tratamiento. En la curva *Tipo II* hay un número constante de fallecidos o sobrevivientes desde el inicio hasta el final del seguimiento. En la curva *Tipo III* se indica una mortalidad o supervivencia temprana intensa, mientras los individuos que sobreviven tienen una elevada tasa de supervivencia. Estas curvas son modelos representativos de una curva de supervivencia, aunque en la práctica se presentan combinaciones de ellas.

2) *Método actuarial*: los tiempos de supervivencia se agrupan en intervalos. La longitud del intervalo depende de la frecuencia con que ocurre el evento de interés y puede no ser de la misma longitud. La desventaja principal de este método es que da estimaciones poco precisas cuando el número de individuos de la muestra es pequeño. En muestras grandes, tiene una importancia fundamental porque influye poco sobre las estimaciones y además permite obtener una función de riesgo. Se asume que las retiradas y pérdidas se distribuyen homogéneamente en el intervalo, pero

puede establecer un sesgo muy importante cuando los intervalos son muy grandes o hay muchos abandonos al estudio.

Es interesante conocer en qué momento del tiempo existe el máximo riesgo de presentar, por ejemplo, una germinación, una muda o una recidiva tumoral. En caso de utilizar la curva de Kaplan-Meier o la curva actuarial, se observarían escasas diferencias. Sin embargo, si se requiere la función de riesgo de defunción (el evento que nos interesa, se podría determinar cuál es el momento de mayor riesgo de que el individuo germine, mude o fallezca que sólo puede ser calculado con el método actuarial). Sólo observando la curva con el empleo de los datos de la curva de sobrevivencia, y sin necesidad de la búsqueda de “p” (probabilidad), se ve que momento es el de mayor riesgo de defunción.

Por lo tanto, la función de riesgo marca la dinámica del proceso estudiado, al dar sus valores una adecuada aproximación a la tasa de incidencia del evento de interés (germinación, muerte de plantas o aparición de enfermedades). Si se quieren comparar 2 curvas de sobrevivencia, la prueba no paramétrica de log-rank (o logaritmo-rango) es la de mayor uso. Las curvas pueden corresponder a dos grupos de individuos con distintos tratamientos o ser grupos con pronóstico diferente. Este método calcula para cada tiempo el número de eventos terminales que se esperarían y el realmente observado, asumiendo que no hay diferencias entre grupos. Se obtiene un valor final que puede compararse mediante una distribución de “ji” cuadrado y da un valor de “p” que demuestra si existen diferencias estadísticamente significativas o no.

En el caso de germinación de semillas toda la información a obtener se realizó mediante los procedimientos de Kaplan-Meier y por el método actuarial, adaptados para evaluar si existen diferencias entre las curvas de germinación de los tratamientos. Igual para los experimentos con dietas alimenticias y su efecto en el desarrollo de fases e instares larvales en insectos y para investigar la relación entre dieta alimenticia y desarrollo de tumores en ratas de laboratorio.

Luego de la determinación de las curvas de sobrevivencia utilizando LIFETEST, se pueden utilizar dos tipos de modelos de regresión: el modelo de Cox de riesgo proporcional, disponible en PHREG y el modelo de tiempo de falla acelerada (AFT), disponible en LIFEREG. El procedimiento

PHREG proporciona métodos de selección de modelos escalonados que incluyen modelos hacia atrás (backward), hacia adelante (forward) y escalonamiento (stepwise). PROC PHREG permite especificar niveles de significancia para la entrada y eliminación de efectos, agregar efectos en un orden secuencial, especificar el número de variables en el modelo para la selección hacia adelante o hacia atrás y seleccionar los mejores subconjuntos (Gordon, 2020).

El modelo de riesgos proporcionales (Modelo de Cox) permitió evaluar el efecto de covariables en el tiempo de vida hasta la ocurrencia del evento de interés. Esta prueba semiparamétrica mide el efecto de un conjunto de variables en el tiempo que ocurre un evento. En nuestro caso, este modelo fue adaptado para evaluar el efecto de los tratamientos pregerminativos en la germinación o las dietas alimenticias en la sobrevivencia de larvas o ratones. Cuando el valor del exponente del coeficiente (coef) es mayor a 1 se consideró que la covariable tiene un mayor riesgo asociado con la ocurrencia del evento. En nuestro caso la condición anterior significó un efecto positivo en la germinación o en la sobrevivencia de las larvas o las ratas. Su interpretación se basa en la razón de riesgo (HR), en donde un valor igual a 1 indica que la variable no ejerce ningún efecto y un valor mayor a 1 indica que la variable modifica la incidencia de la germinación o la mortalidad de las larvas o ratones.

El modelo AFT asume una forma paramétrica para los efectos de las variables explicatorias en la función de sobrevivencia. De igual manera, el modelo de Cox también asume una forma paramétrica para los efectos de las variables explicatorias, pero sigue una forma no especificada para la función de sobrevivencia.

El procedimiento LIFEREG ajusta los modelos de tiempo de falla acelerado a los datos de tiempo de falla o evento, incluidos los datos censurados por la izquierda, censurados por la derecha o censurados por intervalos. Los modelos para la variable de respuesta consisten en un efecto lineal que se compone de las covariables y un término de perturbación aleatorio. Las opciones de distribución para el término de perturbación aleatoria incluyen las distribuciones de valor extremo, normal, logística, exponencial, Weibull, Lognormal, log-logístico y gamma de tres parámetros (Gordon, 2020).

El efecto de las variables intrínsecas y ambientales sobre los tiempos de sobrevivencia se estimaron con el modelo de máxima verosimilitud (LIFEREG). Este procedimiento ajusta modelos paramétricos a datos de tiempos de germinación, muerte, etc., que se registran bajo el esquema de censos en intervalos y permite analizar variables categóricas con valores múltiples. Los coeficientes de regresión y sus errores estándar ( $\beta_i \pm ee$ ) se mejoraron con base en una prueba de bondad de ajuste, para establecer el tipo de distribución de los tiempos en los que ocurren germinaciones, muertes, etc. La distribución observada se compara con las distribuciones teóricas Weibull, Exponencial, Gamma, Log-logística y Log-normal. Los coeficientes de regresión de las variables ( $\beta_i$ ) se transforman con la función  $100(e^{\beta_i} - 1)$ , que estima el porcentaje esperado de aumento en los tiempos de supervivencia por cada unidad de incremento en la covariable si el resto de las variables permanecen sin cambios (Allison 2010).

De esta manera, en el desarrollo y análisis de los procedimientos estadísticos utilizados en el manual, se trata de dar a entender las relaciones funcionales entre los componentes ambientales y biológicos que permiten formalizar modelos e hipótesis en el crecimiento de los organismos. Con esta base, se pueden formular una serie de consideraciones tanto conceptuales como matemáticas para generar otras investigaciones relacionadas con el crecimiento de los organismos.

Hay que señalar la utilidad práctica del análisis de sobrevivencia para la germinación de semilla o para cualquier experimento, utilizando los métodos no paramétricos, semiparamétricos y paramétricos. Es una técnica adecuada para datos en los que el tiempo hasta la ocurrencia de un evento se considera la variable dependiente, conteniendo datos censurados. Las curvas de sobrevivencia de Kaplan-Meier (KM) proporcionan hipótesis exploratorias de la forma de la función de sobrevivencia para cada tratamiento pregerminativo. Una vez que la forma de la línea base del riesgo de germinación se conoce, existe una buena posibilidad de utilizar la forma de distribución adecuada. El análisis de sobrevivencia puede revelar un incremento significativo o no en el riesgo de la germinación por efecto de los tratamientos, los cuales pueden retardar o acelerar el proceso de germinación en el tiempo, con una germinación prolongada o no en el tiempo, variable evaluada en el análisis. El análisis de la germinación también puede evaluarse por etapas, permitido evidenciar la importancia de cada etapa del proceso. El modelo ATF se considera una herramienta útil para el potencial de germinación en

semillas con diferentes tratamientos, permitiendo estimar los tratamientos más importantes y determinar la distribución de la germinación con el tiempo. De igual manera, el modelo ATF permite interpretar la velocidad de germinación en respuesta a tratamientos evaluados, además de revelar las diferencias entre ellos y proporcionar mayor claridad en estudios sobre tiempos de falla, no identificadas en el análisis LIFETEST.

Como se ha visto en el manual, existen muchas formas de comunicar e interpretar la sobrevivencia de diferentes organismos: plantas, semillas o animales de experimentación. Es por ello que hay que definir muy bien qué se busca cuando se habla de sobrevivencia para evitar confusiones. Por lo tanto, se advierte la necesidad de un método que resuma resultados de datos, muchos de ellos censurados, en los cuales el tiempo de sobrevivencia de los que no tuvieron el evento terminal sea un tiempo desconocido. Recordar que el tiempo total define la presencia del evento terminal, no que el individuo haya sido dado de alta por no necesitar más seguimiento (es un sujeto retirado vivo, que cumplió un tiempo incompleto), concepto difícil de entender y que confunde a iniciados en este tema. Vale la pena comentar que los tiempos censurados donde hay retirados vivos (censura por derecha) subestiman la verdadera sobrevivencia del sujeto definido y único que nos interesa y sobreestimando el riesgo al cual fue expuesto y, por tanto, los estudios de sobrevivencia normalmente subestiman la sobrevivencia de la población. Por el contrario, si existen tiempos censurados cuando se incluye a un individuo que fue tratado en fecha muy reciente con muy poco tiempo de seguimiento (censura por izquierda), implicaría subestimar el riesgo de exposición.

Por último, y no menos importante, poner especial atención en la confección y manejo de datos para estos análisis debido a su complejidad, así como a la correcta codificación de las variables intervinientes. El análisis de la sobrevivencia, ya sea en la germinación o de investigación experimental, resultó una herramienta muy útil, con numerosas aplicaciones en el campo biológico porque permite evaluaciones efectivas con diferentes tratamientos de aplicación.

Para la aplicación del manual no se requiere de conocimientos avanzados en estadística o cálculo matemático. Aunque para las discusiones y ejemplos se utiliza el paquete estadístico SAS®, se requiere algún conocimiento de su programación, de interés general en el ajuste de las curvas.



# BIBLIOGRAFÍA

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans. Autom. Control*, 19: 716 – 723.

Alvarado, D. y Hernández, T. (2002). Decline of sacred fir in the Desierto de Los Leones National Park. In: Fenn M, de Bauer LI, Hernández-Tejeda T, eds. *Urban air pollution and forests: resources at risk in the Mexico City Air Basin*. New York: Springer-Verlag. pp. 243-260. DOI: [https://doi.org/10.1007/978-0-387-22520-3\\_10](https://doi.org/10.1007/978-0-387-22520-3_10)

Alves, C. (2017). Impacto de diferentes temperaturas nas características biológicas de *Helicoverpa armigera* (Hübner, 1808) (Lepidóptera: Noctuidae) durante três gerações. Faculdade de Ciências Agrônomicas da UNESP – Campus de Botucatu, para obtenção do título de Mestre em Agronomia (Proteção de Plantas). 51 p.

Allison, P. (1995). *Survival Analysis using SAS®: A practical Guide, Second Edition*. Carry, NC: SAS Institute Inc. 292 p.

Allison, P. (2010). *Survival Analysis using SAS®: A practical Guide, Second Edition*. Carry, NC: SAS Institute Inc. 324 p

Ángeles, E. y López, L. (2009). Supervivencia de una cohorte de plántulas de *Abies religiosa* bajo diferentes condiciones postincendio. *Boletín de la Sociedad Botánica de México*, 84: 25 – 33.

Arrieta, S. y Suárez, F. (2005). Spatial patterns of seedling emergence and survival as a critical phase in holly (*Ilex aquifolium* L.) woodland recruitment in Central Spain. *Forest Ecology and Management* 205: 267-282. DOI: <https://doi.org/10.1016/j.foreco.2004.10.009>

Ayala, G.; Terrazas, T.; López, L. y Trejo C. (2004). Variación en el tamaño y peso de la semilla y su relación con la germinación en una población de *Stenocereus beneckeii*. *Interciencia*, 29 (12): 692 – 697.

Azcón, J y Talón M. (2013). *Fundamentos de Fisiología Vegetal*. 2.<sup>a</sup> Edición. McGraw-Hill - Interamericana de España, S. L. Madrid, España. 651 p.



Bacci, L., Crespo, A., Galván, T., Pereira, E., Bacci, L., Pican, M. y Chediak, M. (2007). Toxicity of insecticides to the sweet potato whitefly (Hemiptera: Aleyrodidae) and its natural enemies. *Pest Management Science*, 706: 699 – 706.

Barak, R., Lichtenberger, T., Wellman, A., Kramer, A. y Larkin, D. (2018). Cracking the case: Seed traits and phylogeny predict time to germination in prairie restoration species. *Ecol. Evol.*, 8: 5551 – 5562.

Baraloto, C., Goldberg, D. y Bonal, D. (2005). Performance trade-offs among tropical tree seedlings in contrasting microhabitats. *Ecology* 86: 2461-2472. DOI: <https://doi.org/10.1890/04-1956>

Bárceñas, M., López, L., Terrazas, T. y García, E. (2013). Germinación de tres especies de *Cephalocereus* (Cactaceae) endémicas del istmo de Tehuantepec, México. *Polibotanica* 36: 105 – 116.

Barracough, H., Simms, L. y Govindan, R. (2011). Biostatistics primer: what a clinician ought to know: Hazard ratios. *J. Thorac. Oncol.*, 6, 978 – 982.

Baskin, C. y Baskin, J. (2001). *Seeds: Ecology, Biogeography, and Evolution of Dormancy and Germination*. Academic Press, San Diego.

Beckage, B., Lavine, M. y Clark, J. (2005). Survival of tree seedlings across space and time: estimates from long-term count data. *Journal of Ecology* 93: 1177-1184. DOI: <https://doi.org/10.1111/j.1365-2745.2005.01053.x>

Bewley J. (1997). Seed germination and dormancy. *Plant Cell* 9:1055.

Bewley, J. y Black, M. (1994). *Seeds. Physiology of development and germination*. Second Edition. Plenum Press, New York – London. 445 p.

Blennow, K. y Lindkvist, L. (2000). Models of low temperature and high irradiance and their application to explaining the risk of seedling mortality. *Forest Ecology and Management* 135: 289-301. DOI: [https://doi.org/10.1016/S0378-1127\(00\)00287-5](https://doi.org/10.1016/S0378-1127(00)00287-5)

Bolfarine, H., Rodrigues, J. y Achcar, J. (1991). *Análise de sobrevivência*. Rio de Janeiro: Instituto de Matemática da Universidade Federal do Rio de Janeiro / Associação brasileira de Estadística. 112p.

Bonfil, C. y Soberón, J. (1999). *Quercus rugosa* Seedling Dynamics in Relation to Its Re-Introduction in a Disturbed Mexican Landscape. *Applied Vegetation Science* 2 (2): 189 – 200.

Booth, J. y Hirschl, T. (2005). Life Table analysis using weighted survey data. Documento en línea, recuperado de: <http://bscb.cornell.edu/~booth/papers/lifetable.pdf>.

Brannen, P. M., Scherm, H., y Bruorton, M. D. (2003). Fungicidal control of *Septoria* leaf spot of blueberry. *Fung. Nemat. Tests* 58: SMF019.

- Buyske, S., Fagerstrom, R. y Ying, Z. (2000). A class of weighted log-rank tests for survival data when the event is rare. *Journal of the American Statistical Association*, 95 (449): 249 – 258.
- Canham, C. y Marks, P. (1985). The response of woody plants to disturbance patterns of establishment and growth. In: Pickett TA, White PS. eds. *The Ecology of Natural Disturbance and Patch Dynamics*. Londres: Academic Press, Inc. pp 197-216. DOI: <https://doi.org/10.1016/B978-0-12-554520-4.50016-2>
- Cantor, A. (2003). *SAS® Survival Analysis Techniques for Medical Research*, Second Edition. Cary, NC: SAS Institute Inc. 230 p.
- Caspersen, J. y Kobe, R. (2001). Interspecific variation in sapling mortality in relation to growth and soil moisture. *Oikos* 92: 160-168. DOI: <https://doi.org/10.1034/j.1600-0706.2001.920119.x>
- Castro, A., Corrêa, A., Legaspi, J., Guedes, R., Serrão, J. y Zanuncio, J. (2013). Survival and behavior of the insecticide-exposed predators *Podisus nigrispinus* and *Supputius cincticeps* (Heteroptera: Pentatomidae). *Chemosphere* 93: 1043 – 1050.
- CATIE. (2000). Técnicas para la germinación de semillas forestales. Serie Técnica. Manual Técnico No. 39. CATIE-PROSEFOR-DFSC. Turrialba, Costa Rica. 54 p.
- Challenger, A. (1998). Utilización y conservación de los ecosistemas terrestres de México. Pasado, presente y futuro. México, DF: Comisión Nacional para el Conocimiento y Uso de la Biodiversidad. ISBN 970-9000-02-0.
- Cline, W. O. (2002). Blueberry bud set and yield following the use of fungicides for leaf spot control in North Carolina. *Acta Hort.* 574: 71-74.
- Collett, D. (2004). *Modelling survival data in medical research*, Second Edition. Chapman & Hall. London
- Collett, D. (2003). *Modelling Survival Data in Medical Research*, 2nd ed. Chapman & Hall/CRC, Boca Raton, FL.
- Corpas, E. y Lara, A. (2009). Aplicación del Modelo de Riesgos Proporcionales de Cox a Pacientes con Sida En España. *Investigación Operacional* 30 (3): 214 – 22.
- Cox, D. y Oakes, D. (1984). *Analysis of survival data*. London, UK, Chapman and Hall.
- Cox, D. (1972). Regression models and life-tables. *J R Stat Soc B*; 34: 187 – 220.
- Cruz, J. y López, L. (2004). Demography of the seedling bank of *Manilkara zapota* (L.) Royen, in a subtropical rain forest of Mexico. *Plant Ecology* 172: 227-235. DOI: <https://doi.org/10.1023/B:VEGE.0000026341.46440.f5>

- Cruz, G., Guerra, E., Valderrábano, J. y Campo, J. (2020). Indicadores de calidad de suelos en bosques templados de la Reserva de la Biosfera los Volcanes, México. *Revista Terra Latinoamericana* 38: 781-793. DOI: <https://doi.org/10.28940/terra.v38i4.421>
- Curt, T., Coll, L., Prévosto, B., Balandier, P. y Kunstler, G. (2005). Plasticity in growth, biomass allocation and root morphology in beech seedlings as induced by irradiance and herbaceous competition. *Annals of Forest Science* 62: 51-60. DOI: <https://doi.org/10.1051/forest:2004092>
- Diggle, P., Heagerty, P., Liang, K. y Zeger, S. (2002). *Analysis of Longitudinal Data*. (Second edition). Oxford: Oxford University Press.
- Dixon, P. y Newman, M. (1991). Analyzing toxicity data using statistical models for time-to-death: An introduction. In: *Metal Ecotoxicology: Concepts and Application*. (M. C. Newman and A. W. McIntosh, eds.), pp. 207-242. Lewis Publishers, Chelsea, MI.
- Dubrovsky, J. (1998). Discontinuous hydration as a facultative requirement for seed germination in two cactus species of the Sonoran Desert. *Bull. Torrey Bot. Club* 125: 33-39.
- Dungan, R., Duncan, R., y Whitehead, D. (2003). Investigating leaf lifespans with interval-censored failure time analysis. *New Phytol.* 158: 593-600.
- Emmert, F. y Dehmer, M. (2019). Introduction to Survival Analysis in Practice. *Mach. Learn. Knowl. Extr.* 1, 1013–1038.
- Esker, P., Gibb, K., Padovan, A., Dixon, P. y Nutter, F. (2006). Use of survival analysis to determine the postincubation time-to-death of papaya due to yellow crinkle disease in Australia. *Plant Dis.*, 90: 102 – 107.
- Fenn, de Bauer M., Hernández, T. eds. (2002). *Urban air pollution and forests: resources at risk in the Mexico City Air Basin*. New York: Springer-Verlag. <https://doi.org/10.1007/978-0-387-22520-3>
- Fleming, T., Harrington, D. y O'sullivan, M. (1987). Superior versions of the log-rank and generalized Wilcoxon statistics. *Journal of the American Statistical Association*, 82 (397): 312 – 320.
- Fox, G. (2001) *Failure-time Analysis: Studying Times to Events and Rates at Which Events Occur* (pp. 235–266). En: Scheiner, S.M. y Gurevich, J. (Eds). *Design and analysis of ecological experiments* (2nd edition). New York, Oxford University Press.
- Fox, G. (1990a). Components of flowering time variation in a desert annual. *Evolution* 44: 1404 – 1423.
- Fox, G. (1990b). Perennation and the persistence of annual life histories. *American Naturalist* 135, 829 – 840.

- Garrett, K., Madden, L., Hughes, G. y Pfender, W. (2004). New applications of statistical tools in plant pathology. *Phytopathology*, 94: 999 – 1003.
- George, B., Seals, S. y Aban, I. (2014). Survival analysis and regression models. *J. Nucl. Cardiol.* 21, 686 – 694.
- Gómez, G., Calle, M., Oller, R. y Langohr, K. (2009). Tutorial on methods for interval-censored data and their implementation in R. *Stat. Model.*, 9, 259 –297.
- Gordon, G. (2020). Recent Developments in Survival Analysis with SAS® Software. In Proceedings of the SAS Global Forum 2020 Conference. Cary, NC: SAS Institute Inc. Paper SAS4286-2020. 1 – 9.
- Greene, G., Leppla, N. y Dickerson, W. (1976). Velvetbean caterpillar: a rearing procedure and artificial medium. *Journal of Economic Entomology*, 69: 487 – 488.
- Grubb, P. (1977). The maintenance of species richness in plant communities: the importance of the regeneration niche. *Biological Reviews* 52: 107-145. DOI: <https://doi.org/10.1111/j.1469-185X.1977.tb01347.x>
- Hallé, F., Oldeman, R. y Tomlinson P. (1978). *Tropical Trees and Forests. An architectural analysis.* Berlin-Heidelberg-New York: Springer-Verlag. pp. 441. ISBN: 978-3-642-81190-6.
- He, Z., Liu, J., Su, S., Zheng, S., Xu, D., Wu, Z., Hong, W. y Wang, J. (2015). Effects of Forest Gaps on Soil Properties in *Castanopsis kawakamii* Nature Forest. *Plos one* 10: e0141203. DOI: <https://doi.org/10.1371/journal.pone.0141203>
- Hernández Ramírez, V., López-Mata, L., Cruz-Rodríguez, J. y Luna Cavazos, M. (2022). Nicho de regeneración de *Abies religiosa* (Kunth) Schldl. & Cham. en el Monte Tláloc, Parque Nacional Iztacuíhuatl-Popocatepetl, México. *Botanical Sciences* 100 (2): 331 – 344.
- Hirsch, H., Wypior, C., vonWehrden, H., Wesche, K., Renison, D. y Hensen, I. (2012). Germination performance of native and non-native *Ulmus pumila* populations. *NeoBiota*, 15: 53 – 68.
- Hosmer, D. y Lemeshow, S. (2001). *Applied Logistic Regression.* 2nd ed. New York. Wiley Interscience, 375 p.
- Ives, M., Funk, R. y Dennis, M. (2007). *Survival Analysis/Life Tables.* Documento en línea, recuperado de: [http://www.chestnut.org/li/downloads/training\\_memos/survival\\_analysis.pdf](http://www.chestnut.org/li/downloads/training_memos/survival_analysis.pdf).
- Johnson, D. y Smith, W. (2005). Refugial forests of the southern Appalachians: photosynthesis and survival in current-year *Abies fraseri* seedlings. *Tree Physiology* 25: 1379-1387. DOI: <https://doi.org/10.1093/treephys/25.11.1379>

Kalbfleisch, J. y Prentice, R. (1980). *The Statistical Analysis of Failure Time Data*, New York: John Wiley & Sons, Inc.

Kaplan, E. y Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association* 53 (282): 457 – 481.

Kartsonak, C. (2016). Survival analysis. Mini-Symposium: Medical Statistics. *Diagnostic Histopathology* 22 (7): 263 – 270.

King, M., Bailey, D., Gibson, D., Pitha, J. y McCay, P. (1979). Incidence and growth of mammary tumors induced by 7,12-Dimethylbenz( $\alpha$ )anthracene as related to the dietary content of fat and antioxidant. *Journal of the National Cancer Institute*, 63, 656 – 664.

Kitajima, K. y Fenner M. (2000). Ecology of Seedling Regeneration. In: Fenner M, ed. *Seeds: The Ecology of Regeneration in Plant Communities*. London: CAB Publishing, pp 331-359. ISBN: 9780851994321. <http://dx.doi.org/10.1079/9780851994321.0331>

Kitajima, K. y Myers, J. (2008). Seedling ecophysiology: strategies toward achievement of positive net carbon balance. In: Leck MA, Parker VT, Simpson RL. eds. *Seedling Ecology and Evolution*. Cambridge: Cambridge University Press. Pp. 172-188. ISBN 978-0-521-87305-5

Klein, J., Rizzo, J., Zhang, M. y Keiding, N. (2001). Statistical methods for the analysis and presentation of the results of bone marrow transplants. Part I: Unadjusted analysis. *Bone Marrow Transpl.* 28, 909 – 915.

Kleinbaum, D. y Klein, M. (2012). *Survival Analysis: A Self-Learning Text, Third Edition, Statistics for Biology and Health*, DOI 10.1007/978-1-4419-6646-9, ©Springer Science+Business Media, LLC.

Kleinbaum, D. & Klein, M. (2005). *Survival Analysis a Self-Learning Text*. New York: Springer Science+Business Media, LLC.

Kleinbaum, D. (1996). *Survival Analysis, A Self-Learning Text*. New York. Springer-Verlag.

Kunstler, G., Curt, T., Bouchaud, M. y Lepart, J. (2005). Growth, mortality, and morphological response of European beech and downy oak along a light gradient in sub-Mediterranean forest. *Canadian Journal of Forest Research* 35: 1657-1668. DOI: <https://doi.org/10.1139/x05-097>

Lara, R., Sánchez, L. y Corral, J. (2009). Regeneración de *Abies religiosa* en claros del dosel versus sotobosque, Parque Nacional Cofre de Perote, México. *Agrociencia* 43: 739-747.

Lawless, J. (2003). *Statistical Models and Methods for Lifetime Data*. Second Edition. New York. Wiley Series In Probability And Statistics. Wiley Interscience. John Wiley & Sons, Inc., Publication, 630 p.

- Leck, M., Parker, V., Simpson, R. (2008). *Seedling Ecology and Evolution*. Cambridge: Cambridge University Press. ISBN: 978-0-521-87305-5
- Lee, E. y Wenyu, J. (2003). *Statistical methods for survival data analysis*. 3rd ed. New Jersey. Wiley series in probability and statistics. 513 p.
- Lee, J. (1996). Some versatile tests based on the simultaneous use of weighted log-rank statistics. *Biometrics*, 52 (2): 721 – 725.
- Lee, E. (1992). *Statistical Methods for Survival Data Analysis, Second Edition*, New York: John Wiley & Sons.
- Leishman, M. y Westoby, M. (1994). The role of large seed size in shaded conditions: experimental evidence. *Functional Ecology* 8: 205-214. DOI: <https://doi.org/10.2307/2389903>
- Lewicki, P. y Hill, T. (2006). *Statistics: Methods and Applications*, 1st ed. Tulsa, OK, USA. StatSoft, Inc.: 578 p. ISBN 1884233597.
- López, J., Jiménez, J., Huez, M., Dávila, J., Ávila, E. y Robles, R. (2017). Germinación de semillas de *Asclepias subulata* en condiciones de casa sombra utilizando ácido giberélico. *European Scientific Journal*.13 (15): 58 – 68.
- López, P., Demuner, A., Alvarenga, E., Almeida, L., Marques, A., Farias, E. y Coutinho, M. (2018). Bioactivity of essential oils from *Artemisia* against *Diaphania hyalinata* and its selectivity to beneficial insects. *Sci. Agric.*, 75 (6): 519 – 525.
- Madrigal, S. (1967). Contribución al conocimiento de la ecología de los bosques de oyamel *Abies religiosa* (HBK) Schl. et Cham) en el Valle de México. Boletín técnico 18, México, DF: Instituto Nacional de Investigaciones Forestales.
- Marañón, T., Camarero, J., Castro, J., Díaz, M., Espelta, J., Hampe, A., Jordano, P., Valladares, F., Verdú, M. y Zamora, R. (2004). Heterogeneidad ambiental y nicho de regeneración. In: Valladares F, ed. *Ecología del Bosque Mediterráneo en un Mundo Cambiante*. Madrid: Ministerio de Medio Ambiente, pp. 69-49. ISBN: 84-8014-552-8
- Marinho, J., Lourencao, A., Oliveira, J., Guedes, R. y Oliveira, M. (2012). Survival and feeding avoidance of the eucalyptus defoliator *Thyrinteina arnobia* exposed to the proteinase inhibitor berenil. *J. Appl. Entomol.* 135: 763 – 777.
- Matiz, J., Rodríguez, G. y Alfaro, C. (2019). Modelos de temperatura del suelo a partir de sondeos superficiales y sensores remotos para el área geotérmica de Paipa, Boyacá-Colombia. *Boletín de Geología* 41: 71-88. DOI: <https://doi.org/10.18273/revbol.v41n2-2019004>



- McNair, J., Sunkara, A. y Frobish, D. (2012). How to analyse seed germination data using statistical time to event analysis: Nonparametric and semiparametric methods. *Seed Sci. Res.*, 22, 77–95.
- Moles, A. y Leishman, M. (2008). The seedling as part of a plant's life history strategy. In: Leck MA, Parker VT, Simpson RL. eds. *Seedling Ecology and Evolution*. Cambridge: Cambridge University Press. pp. 189-1238. ISBN 978-0-521-87305-5
- Montoya, E., Guzmán, R. y López, L. (2020). Fragmentation dynamics in an *Abies religiosa* forest of central Mexico. *Canadian Journal of Forest Research* 50: 680-688. DOI: <https://doi.org/10.1139/cjfr-2019-0235>
- Muenchow, G. (1986). Ecological use of failure time analysis. *Ecology* 67:246-250.
- Nelder, J. y Wedderburn, R. (1972). Generalized linear models. *Journal of the Royal Statistical Society Series A*, 135 (3): 370 – 384.
- Ojiambo, P. S., Scherm, H., y Brannen, P. M. (2002). Septoria leaf spot intensity, defoliation, and yield loss relationships in southern blueberries. (Abstr.) *Phytopathology* 92 (suppl.): S1025.
- Onofri, A., Gresta, F. y Tei, F. (2010). A new method for the analysis of germination and emergence data of weed species. *Weed Research* 50, 187–198.
- Osunkjoya, O., Ash, J., Hopkins, M. y Graham, A. (1992). Factors affecting survival of tree seedlings in North Queensland rainforests. *Oecologia* 91: 569-578. DOI: <https://doi.org/10.1007/BF00650333>
- Packer, A. y Clay, K. (2003). Soil pathogens and *Prunus serotina* seedling and sapling growth near conspecific trees. *Ecology* 84: 108-119. DOI: [https://doi.org/10.1890/0012-9658\(2003\)084](https://doi.org/10.1890/0012-9658(2003)084) [0108: SPAPSS]2.0.CO;2
- Padovan, A. y Gibb, K. (2001). Epidemiology of phytoplasma diseases in papaya in Northern Australia. *J. Phytopathol.*, 149: 649- 658.
- Parra, J. (2001). Técnicas de criação de insetos para programas de controle biológico. *Esalq/Fealq*. 134.
- Pelz, C. y Klein, J. (1996). Analysis of survival data: A comparison of three major statistical packages (SAS, SPSS, BMDP). Tech. Rep. 17, Division of Biostatistics, Medical College of Wisconsin, Milwaukee.
- Pepe, M. y Fleming, T. (1989). Weighted Kaplan-Meier statistics: a class of distance tests for censored survival data. *Biometrics*, 45: 497 – 507.
- Peña, R. (2017). Efecto de dos dietas alternantes vegetales en el desarrollo larval de dos especies de mariposas diurnas; *Morpho menelaus* y *Morpho helenor* bajo condiciones de cautiverio. Universidad Científica del Perú, Facultad de Ciencias e Ingeniería. Iquitos. Loreto. 77 p.

- Peto, R., Pike, M., Armitage, P., Breslow, N., Cox, D., Howard, S., Mantel, N., McPherson, K., Peto, J. y Smith, P. (1977). Design and analysis of randomized clinical trials requiring prolonged observation of each patient. Part II: Analysis and Examples. *Brit. J. Cancer.*, 35: 1 – 39.
- Platt, W. y Streng, D. (1989). Gaps in forest ecology. Special Feature: Treefall Gaps and Forest Dynamics. *Ecology* 70: 535-576. DOI: <https://doi.org/10.2307/1940194>
- Portilla, M., Snodgrass, G. y Luttrell, R. (2017). Lethal and sub-lethal effects of *Beauveria bassiana* (Cordycipitaceae) strain NI8 on *Chrysoperla rufilabris* (Neuroptera: Chrysopidae). *Florida Entomologist*, 100 (3): 627 – 633.
- Portilla, M., Walker, J., Perera, O., Seiter, N. y Greene, J. (2016). Estimation of median lethal concentration of three isolates of *Beauveria bassiana* for control of *Megacopta cribaria* (Heteroptera: Plataspidae) bioassayed on solid *Lygus* spp. diet. *Insects* 7: 1–13.
- Portilla, M., Snodgrass, G. y Luttrell, R. (2014). A novel bioassay to evaluate the potential of *Beauveria bassiana* NI8 and the insect growth regulator novaluron against *Lygus lineolaris* on a nonautoclaved solid artificial diet. *Journal of Insect Science* 14: 1–13.
- Prentice, R., Williams, B. y Peterson, A. (1981). On the Regression Analysis of Multivariate Failure Time Data. *Biometrika*, 68: 373 – 379.
- Prinja, S., Gupta, N. y Varma, R. (2010). Censoring in clinical trials: Review of survival analysis techniques. *Indian J Community Med.*; 35:217–221
- Quero, A., Hernández, F., Pérez, P., Pool, D., Landa, P. y Nieto, R. (2017). Germinación y emergencia diaria de cariósides y diásporas de pastos nativos e introducidos. *Rev. Fitotec. Mex.* 40 (1): 35 – 44.
- Ranal, M. y Garcia, D. (2006). How and why to measure the germination process? *Revista Brasil. Bot.*, 29 (1): 1 – 11.
- Rey, P. y Alcántara, J. (2000). Recruitment dynamics of a fleshy- fruited plant (*Olea europaea*): connecting patterns of seed dispersal to seedling establishment. *Journal of Ecology* 88: 622-633. DOI: <https://doi.org/10.1046/j.1365-2745.2000.00472.x>
- Rich, P., Wood, J., Vieglais, D., Burek, K. y Webb, N. (1999). Hemiview Manual, version 2.1. Delta-T Devices, Ltd. [https://delta-t.co.uk/wp-content/uploads/2020/02/HemiView-User-Manual-\\_2.1.pdf](https://delta-t.co.uk/wp-content/uploads/2020/02/HemiView-User-Manual-_2.1.pdf) (revisado agosto 20, 2020)
- Rodríguez N., E. (2019). Factores que influyen en la viabilidad, germinación y establecimiento de *Carpinus caroliniana* Walt. Universidad Michoacana de San Nicolás de Hidalgo, Programa Institucional de Maestría en Ciencias Biológicas, Facultad de Biología, Morelia Michoacán, México. 66 p.



Roloff, I., Scherm, H., y van Iersel, M. W. (2004). Photosynthesis of blueberry leaves as affected by Septoria leaf spot and abiotic leaf damage. *Plant Dis.* 88:397-401.

Román R. (2002). Ecología de semillas y plántulas de *Abies religiosa* (HBK) Schl. et Cham. en el parque nacional “Cumbres del Ajusco”. BSc. Thesis. Universidad Nacional Autónoma de México.

Romano, A. y Stevanato, P. (2020). Germination Data Analysis by Time-to-Event Approaches. *Plants* 9, 617: 1 – 15 p.

Rzedowski, J. (2006). Bosque de Coníferas. In: Rzedowski J, ed. *Vegetación de México*. Ciudad de México, México: 1ra. Edición digital, Comisión Nacional para el Conocimiento y Uso de la Biodiversidad. Págs. 295-327. ISBN: 9681800028.

Sánchez, A. y López, L. (2003). Clasificación y ordenación de la vegetación del norte de la Sierra Nevada, a lo largo de un gradiente altitudinal. *Anales del Instituto de Biología, Universidad Nacional Autónoma de México, Serie Botánica* 74: 47-71.

SAS Institute Inc. 2020. SAS/STAT Guide for personal computers, version 9.4 edition. Ed. SAS Institute, Cary, NC. U S A.

SAS. 2012. Statistical Analysis System. V.9.0 ed. SAS. Institute. Inc. N.C. USA: Cary.

Scharenbroch, B. y Bockheim, J. (2007). Impacts of forest gaps on soil properties and processes in old growth northern hardwood-hemlock forests. *Plant Soil* 294: 219-233. DOI: <https://doi.org/10.1007/s11104-007-9248-y>

Scherm, H. y Ojiambo, P. (2004). Applications of survival analysis in botanical epidemiology. *Phytopathology*, 94: 1022 – 1026.

Scherm, H., Brannen, P. M., Ojiambo, P. S., Savelle, A. T., Krewer, G., y Bruorton, M. D. (2003). Blueberry leaf spots: Epidemiology, yield losses, and control. Pages 57-66 in: *Proc. Southeast. Blueberry Conf.*, Savannah, GA.

Schmidt L. (2000). Handling of tropical and subtropical forest tree seed. DFSC. Humleback, Denmark. 511 p.

Scott, S. y Jones, R. (1990). Generation means analysis of right censored response time traits: low temperature seed germination in tomato. *Euphytica* 48: 239 – 244.

Scott, S., Jones, R. y Williams, W. (1984). Review of data analysis for seed germination. *Crop Science* 4, 1192–1199.

Scott, S. y Jones, R. (1982). Low temperature seed germination of *Lycopersicon* species evaluated by survival analysis. *Euphytica*, 31, 869 – 883.

---

SEMARNAT. (2000). Norma Oficial Mexicana NOM-021-RECNAT-2000, que establece las especificaciones de fertilidad, salinidad y clasificación de suelos. Estudios, muestreo y análisis. Diario Oficial, 31 de diciembre de 2002.

Setti, B., Bencheikh, M., Henni, J. y Claire, N. (2010). Survival analysis to determine the length of latent period of *Mycosphaerella pinodes* on peas (*Pisum sativum* L.). African Journal of Microbiology Research. 4(18): 1897 – 1903.

Sheng, A. y Ghosh, S. (2020). Effects of Proportional Hazard Assumption on Variable Selection Methods for Censored Data. Stat Biopharm Res., 12 (2): 199–209.

Sigala, J., González, M. y Prieto, J. (2015). Supervivencia en plantaciones de *Pinus pseudostrobus* Lindl. en función del sistema de producción y preacondicionamiento en vivero. Revista Mexicana de Ciencias Forestales, 6 (30): 20 – 31.

Yamamoto, S. (2000). Forest gap dynamics and tree regeneration. Journal of Forest Research 5: 223-229. DOI: <https://doi.org/10.1007/BF02767114>



## EL AUTOR



**César C. Vale Montilla.** Ingeniero Forestal, egresado de la Universidad de Los Andes, Facultad de Ciencias Forestales, Mérida, Venezuela, 1979. Msc. en Entomología, egresado de la Universidad Central de Venezuela, Facultad de Agronomía, Caracas, Venezuela, 1991. Ha cursado numerosos y diversos Cursos de Postgrado no conducentes a grado académico: Gerencia de Investigación y Desarrollo, Instituto de Estudios Superiores de Administración – IESA 1991; Fisiología de Cultivos, Universidad de Los Andes, Facultad de Ciencias Forestales y Ambientales, Instituto de Investigaciones Agropecuarias, 1998; SAS para Investigadores, Universidad de Los Andes, Facultad de Ciencias Económicas y Sociales, Instituto de Estadística Aplicada y Computación, 1999; Diseño y Análisis de Experimentos usando SAS, Universidad de Los Andes, Facultad de Ciencias Económicas y Sociales, Instituto de Estadística Aplicada y Computación, 1999.

Profesor Titular e investigador del Centro de Investigaciones Agrícolas, Biológicas, Educativas y Sociales (CIABES), de la Universidad de Los Andes, Núcleo Universitario Rafael Rangel –NURR-, Trujillo, Venezuela. Asesor de 20 Tesis de Pregrado, 02 Tesis de Postgrado. Autor de 12 artículos de investigación en Revistas arbitradas e indexadas. Línea de investigación: Entomología y Control Biológico, Análisis de crecimiento de plantas.

Miembro Fundador del Grupo de Investigación en Fitopatología y Control Biológico “Dr. Carlos Díaz Polanco”, Universidad de Los Andes, Núcleo Universitario Rafael Rangel, Trujillo, Venezuela. 1996. Miembro del Centro de Investigaciones Agrícolas, Biológicas, Educativas y Sociales (CIABES), Universidad de Los Andes, Núcleo Universitario Rafael Rangel, Trujillo, Venezuela. Miembro del Consejo de Departamento de Biología y Química, Universidad de Los Andes, Núcleo Universitario Rafael Rangel, Trujillo, Venezuela. 1999 – 2016. Jefe de Departamento de Biología y Química, Universidad de Los Andes, Núcleo Universitario Rafael Rangel, Trujillo, Venezuela. 2016 – 2020. Secretario de Asuntos Gremiales de la Asociación de Profesores de la Universidad de Los Andes – APULA, Núcleo Universitario Rafael Rangel, Trujillo, Venezuela. 2000. Miembro del Consejo Superior del NURR, Universidad de Los Andes, Núcleo Universitario Rafael Rangel, Trujillo, Venezuela, desde 2010 al presente. Coordinador de la Unidad Delegada de Asuntos Profesorales – UDAP, Universidad de Los Andes, Núcleo Universitario Rafael Rangel, Trujillo, Venezuela. 2014 – 2016.





ISBN: 978-980-11-2190-9



Este manual, presenta algunos de los procedimientos utilizados para el cálculo y desarrollo de análisis de sobrevivencia en algunas especialidades biológicas. El propósito general es ofrecer a profesionales y estudiantes dedicados a la investigación biológica, una herramienta para la aplicación de análisis de sobrevivencia. Su objetivo consiste en manejar los detalles de exploración y descripción de datos, y servir de guía práctica para ilustrar las posibilidades de las técnicas de análisis de datos en diferentes actividades relacionadas con las plantas, insectos u otras investigaciones biológicas. Discute el uso de procedimientos con PROC LIFETEST, PROC PHREG y PROC LIFEREG, como herramientas para el ajuste de curvas de sobrevivencia.

Incluye la aplicación práctica de las metodologías utilizadas para el desarrollo de tablas de vida. Se detallan algunos procedimientos para el análisis de germinación de semillas de las especies *Hura crepitans* L, *Cassia moschata* Kunth y *Swietenia macrophylla* King bajo tratamientos pregerminativos a base de tiempos de imbibición en agua y en hormonas vegetales.

El texto del Manual se ha organizado de la siguiente manera:

**Capítulo 1:** Consideraciones generales sobre el Análisis de Sobrevivencia.

**Capítulo 2:** Características y procedimientos del análisis.

**Capítulo 3:** Metodologías y Procedimientos Aplicados al Crecimiento de Plantas, la Entomología y el Control Biológico