Artículo original

Un modelo de supervivencia bivariante para eventos dependientes bajo el enfoque de funciones cópulas.

A bivariate survival model for dependent events under the copula functions approach.

Peña G Jesús A[1*], Ramoni P Josefa[2], Giampaolo Orlandoni[3].

[1]Departamento de Bioanálisis Clínico, Cátedra de Bioestadística, Facultad de Farmacia y Bioanálisis, Universidad de Los Andes, Mérida, 5101, República Bolivariana de Venezuela. [2]Facultad de Ciencias Económicas, Administrativas y Contables. [3]Departamento de Matemática y Estadística, Facultad de Ciencias Exactas, Físicas y Naturales, [2,3]Universidad de Santander, Bucaramanga, Santander, 4-72, Colombia.

Recibido: julio de 2018 – Aceptado: septiembre de 2018

RESUMEN

Se propone un modelo de supervivencia bivariante para analizar los tiempos de ocurrencia de dos eventos de distinto tipo a lo largo del periodo de seguimiento de cada individuo. Este modelo está basado en el enfoque de las variables aleatorias no latentes, las funciones cópulas Arquimedianas y de la función media acumulada para eventos de distinto tipo no recurrentes. A partir de un conjunto de datos simulados referidos a la ocurrencia de dos eventos de distinto tipo para muestras de individuos de distinto tamaño y diferentes estructuras de dependencia no necesariamente lineal, se evalúa el rendimiento del modelo propuesto. Los resultados muestran que si ambos tiempos de fracaso, para cada tipo de evento, son independientes el modelo propuesto se simplifica al modelo no paramétrico de Kaplan y Meier. En caso contrario, se describe una estructura de dependencia entre ambos tiempos de fracaso cuyos pronósticos de supervivencia estimados, mediante el modelo de supervivencia bivariante con base en las funciones cópulas de la familia Arquimediana, son completamente distintos al caso anterior.

PALABRAS CLAVE

Función de supervivencia, función acumulada media, función cópula, función cópula Arquimediana.

ABSTRACT

A bivariate survival model is proposed to analyze the times of occurrence of two events of different types throughout the follow-up period of each individual. This model is based on the approach of non-latent random variables, the Archimedian copula functions and the cumulative average function for no recurrent events of different types. From a set of simulated data referred to the occurrence of two events of different types for samples of individuals of different sizes and different structures of dependence not necessarily linear, the performance of the proposed model is evaluated. The results show that if both failure time, for each type of event, are independent, the proposed model is simplified to the nonparametric model of Kaplan and Meier. Otherwise, a dependency structure between the two failure times is described whose estimated survival predictions, using the bivariate survival model based on the copula functions of the Archimediana, are completely different from the previous case.

KEYWORDS

Survival function, cumulative average function, copula function, Archimedian copula function.

[*]*Correspondencia al autor: japg@ula.ve*

INTRODUCCIÓN

En el área de la salud la variable tiempo de ocurrencia de algún evento de interés clínico ha sido analizada mediante un modelo de supervivencia univariante. El evento generalmente es de naturaleza terminal, como muerte o cualquier otra condición que reste la posibilidad de que el individuo sea observado nuevamente. El tiempo es el momento en que ocurre el evento de interés llamado tiempo de fracaso o de supervivencia. A cada individuo i con $i = 1,2, \ldots, N$ le corresponde un periodo de seguimiento $[0, \delta_i]$ donde δ_i es el último instante de tiempo en que fue observado cada individuo. Sea T_i el instante de tiempo en que ocurrió el evento de interés para el i-ésimo individuo. Si $T_i < \delta_i$ el tiempo registrado es el verdadero instante de tiempo en que ocurrió el evento, pero si $T_i > \delta_i$ el evento de interés aún no ha ocurrido y el tiempo registrado es censurado a la derecha, debido a que el evento aún puede ocurrir. Por consiguiente, el tiempo registrado es $T = \min\{T_i, \delta_i\}$.

Ahora bien, al considerar la ocurrencia de dos o más eventos de distinto tipo k con $k = 1,2, \ldots, K$ sobre un mismo individuo, el tiempo registrado para cada i es el tiempo mínimo de ocurrencia de cada evento, $T = \min\{T_1, T_2, \ldots, T_k, \delta_i\}$, esto es, una vez observado el primer evento este censura la ocurrencia de los demás. Por ejemplo, distintas causas de muerte de cada paciente. Este tópico en la literatura estadística se conoce como riesgos en competencia [1].

Sin embargo, en el área clínica también se estudian eventos de naturaleza no terminal que pueden ser experimentados por cada individuo junto con un evento fatal, como muerte. Por ejemplo, recaída del paciente posterior a una intervención quirúrgica y el evento muerte. En este caso se puede analizar el evento que es más propenso a generar el evento fatal. Este enfoque en la literatura estadística se conoce como riesgos en semi-competencia [2]. Si el evento es recaída a partir de alguna causa en particular, él puede ser censurado por otros eventos como muerte por otra causa, recaída por otra causa, perdido en el estudio o mantenerse sano hasta finalizado el estudio. Estos eventos compiten entre sí para evitar la ocurrencia del evento principal, dando lugar a un tipo de censura conocida como censura informativa, esto es, lo tiempos de fracaso y los censurados pueden ser dependientes [3].

En el área clínica pueden ser de interés eventos de naturaleza no terminal que no llevan a la ocurrencia de un evento fatal, sino que el individuo puede sobrevivir bajo la ocurrencia de algunos de estos eventos. Por ejemplo, el síndrome metabólico, el virus H1N1, lesiones cervico-vaginales, infertilidad, entre otras patologías, son eventos compuestos debido a que son producto de la combinación de otros eventos no terminales. Así, algunos eventos no terminales pueden tener la propiedad de generar otro evento. La ocurrencia de este nuevo evento puede suceder inmediata al evento anterior o por el contario, el evento anterior puede retardar en cierta forma la ocurrencia del nuevo evento. Por ejemplo, el inicio del tratamiento clínico puede retardar la ocurrencia de alguna patología o recaída del paciente. Por consiguiente, los eventos producto de la combinación de otros eventos no terminales los hemos denominado evento compuesto. Este evento se genera a partir de la acumulación de eventos de distinto tipo a lo largo del periodo de seguimiento de cada paciente, lo que implica que, el investigador debe afrontar el problema de la dependencia entre los tiempos de fracaso de los diferentes tipos de eventos, pero no desde el punto de vista de los riesgos en competencia.

En consecuencia, el objetivo general de este artículo consiste en proponer un modelo de supervivencia para el análisis de la ocurrencia de eventos de distinto tipo experimentados durante distintos instantes de tiempo por el mismo individuo, cuyos tiempos de fracaso son dependientes. Este modelo está fundamentado en una función similar a la función acumulada media (MCF) [4], que en este caso es tratada para eventos de distinto tipo no recurrentes (FAME$\overline{R}$) y de las funciones cópulas de la familia Arquimediana. A partir de estas funciones se trata el problema tanto de dependencia como de independencia entre los tiempos de fracaso de distinto tipo experimentados por un mismo individuo a lo largo del tiempo. Por lo tanto, la combinación de ambas funciones genera un modelo de supervivencia muy particular cuyas probabilidades de supervivencia estimadas cambian considerablemente como función de las características de los tiempos de fracaso y en relación a que estos se consideren independientes.

La estructura de este artículo es como sigue, en la siguiente sección se muestra la descripción del modelo de supervivencia propuesto, luego se explica el modelo propuesto con base en las funciones cópulas y finalmente la evaluación del modelo propuesto a partir de un conjunto de datos simulados obtenidos mediante la elaboración de un algoritmo a través de las librerías de R para funciones cópulas.

Modelo de supervivencia propuesto

Se propone un modelo de supervivencia para analizar tiempos de fracaso de eventos de distinto tipo no recurrentes y de naturaleza no terminal experimentados por un mismo individuo i a lo largo de su propio periodo de seguimiento $[0, \delta_i]$. Sea T_{ik} los distintos instantes de tiempo en que fue observado un evento de tipo k para el i-ésimo individuo. Estos eventos o subeventos, con base en la definición operacional de cada uno, pueden tener la propiedad de constituir el evento compuesto. La composición del evento compuesto puede consistir de al menos un evento de tipo k con $k = 1,2, \dots, K$ y su significado dependerá del objetivo, aplicación e interés por parte del investigador, así como de la definición operacional del evento compuesto.

Supongamos que cada paciente está expuesto a la ocurrencia de K eventos de distinto tipo y no existe un orden particular de su ocurrencia, a menos que la naturaleza de cada evento o del evento compuesto así lo amerite. Además, una vez que ocurre el primer evento este no se repite (no es recurrente en el tiempo) hasta que se observa la ocurrencia del segundo evento y por consiguiente el evento compuesto. De acuerdo a la forma y secuencia acumulada en que van ocurriendo los eventos, para cada individuo se va generando un “proceso histórico acumulado” (PHA) durante $[0, \delta_i]$, siendo δ_i el tiempo máximo de seguimiento censurado a la derecha, en ese instante de tiempo no se observa ninguno de los eventos de interés, es simplemente el “último” registro de cada paciente. Cada PHA genera una “función de un proceso histórico acumulado” (FPHA) denotada como $Y_{ik}^*(t)$ con rango igual a 1, debido a que cada evento de distinto tipo k será observado sólo una vez en cada individuo i en el instante de tiempo t. En cambio, si cada evento tipo k es recurrente entonces $Y_{ik}^*(t) > 1$.

Sea $Y_i^*(t)$ la cantidad total de eventos de distinto tipo experimentados por el i-ésimo individuo en el instante de tiempo t, entonces

$$\sum_{k=1}^{K} Y_{ik}^*(t) = Y_i^*(t) \qquad (1)$$

Esto implica que el número total medio de eventos de distinto tipo k experimentados por los N individuos del estudio, denotado como $M_k^*(t)$, es

$$M_k^*(t) = \frac{1}{N}\left(Y_1^*(t) + Y_2^*(t) + \cdots + Y_N^*(t)\right) \qquad (2)$$

que llamaremos “función acumulada media para eventos de distinto tipo no recurrentes” (FAME$\overline{R}$).

Sea G un subconjunto de eventos de distinto tipo k observados en cada individuo i, que conforman el evento compuesto, $G = \{1,2, \dots, g\}$ con $g \leq K$.

Si $G = 1$, significa que de los K posibles eventos que podría experimentar el i-ésimo individuo, sólo uno constituye el evento compuesto. Esto es, basta que ocurra sólo un evento de tipo k y el evento compuesto está presente, así, el i-ésimo individuo sobrevive (no experimenta el evento compuesto) si supera cada uno de los K posibles eventos a los que está expuesto. Por ejemplo, en el área clínica existen patologías que están presentes si el paciente padece sólo uno de sus síntomas, como es el caso de las bacterias patógenas, basta que el paciente experimente sólo una de ellas y el síndrome diarreico infeccioso está presente [5].

Si $G = 2$, significa que de los K posibles eventos que podría experimentar el i-ésimo individuo, sólo dos constituyen el evento compuesto, así, el i-ésimo individuo sobrevive si supera al menos $K - 1$ posibles eventos. Por ejemplo, en el área clínica existen patologías compuestas por un subconjunto de eventos, como es el caso del síndrome metabólico, el dengue, el virus de H1N1, entre otras patologías. Cabe mencionar, que los eventos considerados son de naturaleza no terminal, así el i-ésimo individuo puede experimentar al menos uno de estos eventos y aun mantenerse con vida o no censurar la ocurrencia de los demás eventos, incluso estando presente el evento compuesto.

Por lo tanto, el evento compuesto puede ocurrir de $\binom{K}{g}$ diferentes maneras posibles, cualquiera de estas combinaciones constituye el evento compuesto. Cada subconjunto g de eventos genera una función FAME$\overline{R}$, denotada como $M_g^*(t)$. Además, en cada instante de tiempo t para el i-ésimo individuo se tiene una cantidad media de eventos de distinto tipo para todos los N valores poblacionales de la forma $Y_{ig}^*(t)$ con $Y_{ig}^*(t) = \sum_{k\epsilon G} Y_{ik}^*(t)$. Se puede probar a partir de las ecuaciones (1) y (2) que

$$M_g^*(t) = E_i\{Y_{ig}^*(t)\} = \sum_{k\epsilon G} M_k^*(t) \qquad (3)$$

siendo un estimador de $M_g^*(t)$ la media aritmética clásica, con $Y_{ig}^*(t)$ la cantidad de eventos de distinto tipo experimentados por el i-ésimo individuo en el instante de tiempo t en función del evento compuesto.

Para calcular los valores de la expresión (3), se ordenan los instantes de tiempo en que ocurrió cada evento. Supongamos tenemos L de esos instantes de tiempos de fracaso, $t_{(1)} < t_{(2)} < \cdots < t_{(L)}$, donde cada $t_{(l)}$ con $l = 1,2,\dots,L$ es el instante de tiempo ordenado en que ocurrió un evento de distinto tipo k.

Sea $Y^*(t)$ la cantidad de individuos en riesgo de experimentar un nuevo evento en cada $t_{(l)}$. Si $t_{(l)}$ es censurado, $Y^*(t) = N - 1$; en caso contrario, $Y^*(t) = N$. En cada $t_{(l)}$ se tiene la cantidad de veces que ocurrió el evento tipo k, que es equivalente a la cantidad de individuos que experimentaron el evento tipo k en $t_{(l)}$, denotado como $Y_k^*(t)$ o $Y_g^*(t)$ dependiendo del caso. Para cada $t_{(l)}$ se tiene el incremento medio de nuevos eventos de distinto tipo por individuo, así

$$m_k^*(t) = \frac{Y_k^*(t)}{Y^*(t)} \qquad (4)$$

que con base en la ecuación (3) se tiene que $M_k^*(t)$ se obtiene sumando los incrementos precedentes, $m_k^*(t)$, esto es

$$M_k^*(t) = \sum_{\{t_{(l)} < t\}} m_k^*\left(t_{(l)}\right) \qquad (5)$$

lo que implica que la población de valores tiene una curva media K-variante, dada por

$$[M_1^*(t), M_2^*(t), \dots, M_K^*(t)] \qquad (6)$$

En particular, las expresiones (4), (5) y (6) también se cumplen con base en el evento compuesto g, lo que implica que a partir de la expresión (3), que

$$M_g^*(t) = \sum_{k\epsilon G} M_k^*(t) = \sum_{k\epsilon G} \sum_{\{t_{(l)} < t\}} m_k^*\left(t_{(l)}\right) \qquad (7)$$

La expresión (7), también se cumplen cuando el evento compuesto g consta de todos los K eventos que puede experimentar cada individuo, similar a la construcción de la función MCF indicada en [4].

Ahora bien, en el análisis de supervivencia clásico univariante, como se indica en [6], una forma de obtener la función de supervivencia es mediante

$$S_T(t) = P[T > t] = \exp(-\Lambda(t)) \qquad (8)$$

donde $\Lambda(t)$ es la función de riesgo acumulada, cuyo evento de interés es de naturaleza terminal. Entre las funciones $M_k^*(t)$ y $\Lambda(t)$ existe cierta similitud, aunque conceptualmente son diferentes. El dominio de ambas funciones es $[0, \infty)$, $\Lambda(t)$ mide el riesgo acumulado de experimentar el evento terminal como máximo en el instante de tiempo t, y $M_k^*(t)$ es la cantidad media acumulada de eventos de distinto tipo experimentados en el instante de tiempo t. Por consiguiente, se propone que la función de supervivencia también puede ser determinada mediante

$$S_T(t) = \exp(-M_k^*(t)\,) \qquad (9)$$

La expresión (9) coincide con el estimador no paramétrico de [7] cuando los tiempos de fracaso son independientes. Además, es un estimador no paramétrico debido a que no involucra supuestos acerca de la población de curvas acumuladas ni parámetros a ser estimados. En consecuencia, se sugieren las siguientes proposiciones:

Proposición 1: Sea g el conjunto de eventos de tipo k que conforman el evento compuesto, g puede contener como máximo K eventos de distinto tipo. Si $M_g^*(t)$ es la cantidad media acumulada de eventos de distinto tipo k experimentados por el mismo individuo durante el instante de tiempo t, entonces $M_g^*(t) \equiv \Lambda(t)$ donde $\Lambda(t)$ es el riesgo acumulado obtenido a partir del riesgo simple $\lambda(t)$ para el análisis de la ocurrencia de sólo un evento de naturaleza terminal.

Proposición 2: El evento compuesto g es producto de la combinación de otros eventos. Si tenemos una muestra aleatoria de N individuos y en cada uno se observó el evento compuesto, entonces las funciones $m_g^*(t)$ y $\lambda(t)$ son equivalentes, esto es, $m_g^*(t) \equiv \lambda(t)$.

Proposición 3: Sea una muestra aleatoria de N individuos y se considera que cada uno experimentó K eventos de distinto tipo durante el periodo de seguimiento $[0, \delta_i]$ con $i = 1,2, \dots, N$, entonces la función de supervivencia conjunta $S_{12\dots K}(t_1, t_2, \dots, t_K)$ está dada por

$$S_{12\dots K}(t_1, t_2, \dots, t_K) = \exp\left(-\sum_{k=1}^{K} M_k^*(t_k)\right) \qquad (10)$$

de donde se obtiene que

$$S_{12\dots K}(t_1, t_2, \dots, t_K) = \prod_{k=1}^{K} S_k(t_k) \qquad (11)$$

En relación con la expresión (9) y la Proposición 3, como se indica en [8], la función de supervivencia global es igual al producto de las funciones de supervivencia marginales, pero bajo el esquema de riesgos en competencia. Sin embargo, en este caso también se cumple dicha igualdad. La demostración de estas proposiciones se muestran en el Apéndice de este artículo.

Si los tiempos de supervivencia, T_{ik}, de los diferentes tipos de eventos son dependientes el modelo de supervivencia propuesto está basado en las funciones cópulas. Estas funciones permiten obtener un modelo de supervivencia multivariante a partir de las funciones de supervivencia marginales para cada tipo de evento.

Modelo propuesto a partir de las funciones cópulas

Sean $T_1, T_2, \dots, T_K$ variables aleatorias continuas definidas como $T_k: \Omega \longrightarrow \mathbb{R}$, $k = 1,2, \dots, K$ sobre el espacio de probabilidad $(\Omega, \mathcal{A}, \mathcal{P})$. Las funciones $F_1, F_2, \dots, F_K$ son funciones de distribución inducidas por las variables aleatorias sobre el espacio de probabilidad $(\mathbb{R}, \mathcal{B}, \mathcal{P}_T)$, con $F_{T_k}: \mathbb{R} \longrightarrow [0,1]$ tal que $F_{T_k}(t_k) = P[T_k \leq t_k]$, $k = 1,2, \dots, K$ donde $\mathcal{B}$ son los conjuntos de Borel [9].

Supongamos que $K = 2$, entonces a cada par de números reales $(t_1, t_2) \in \mathbb{R}^2$ le corresponde la terna $(F_{T_1}(t_1), F_{T_2}(t_2), F_{12}(t_1, t_2)) \in \mathbb{R}^3$ con cada función definida en el intervalo $[0,1]$, esto es, a cada par de números reales (t_1, t_2) le corresponde un punto de la forma $(F_{T_1}(t_1), F_{T_2}(t_2))$ en el espacio producto $[0,1]^2$ y a partir de este par ordenado se obtiene otro número de la forma $F_{12}(t_1, t_2) \in [0,1]$. Así, podemos definir la función C, dada por

$$\begin{aligned} C: [0,1]\times[0,1] &\longrightarrow [0,1] \\ (F_{T_1}(t_1), F_{T_2}(t_2)) &\longrightarrow F_{12}(t_1, t_2) \end{aligned} \qquad (11)$$

lo que implica que $C\left(F_{T_1}(t_1), F_{T_2}(t_2)\right) = F_{12}(t_1, t_2)$ que es equivalente a escribir

$$C\left(F_{T_1}, F_{T_2}\right) = F(t_1, t_2) \qquad (12)$$

Por ser $F(t_1, t_2)$ función de distribución conjunta implica que la función $C\left(F_{T_1}, F_{T_2}\right)$ también lo sea, con la diferencia de que ahora sus componentes son funciones de distribución marginales univariantes. Por consiguiente, la función C como función de distribución bivariante recibe el nombre de función cópula [10].

Definición 1: Una función cópula es una función C, $C: [0,1]^2 \longrightarrow [0,1]$ tal que

(a) Para cada u, v en $[0,1]$, $C(u, 0) = 0 = C(0, v)$ y $C(u, 1) = u; C(1, v) = v$.
(b) Para cada u_1, u_2, v_1, v_2 en $[0,1]$ con $u_1 \leq u_2$ y $v_1 \leq v_2$, se tiene que

$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0$$

En relación con la propiedad (a) indica que la función C posee una cota inferior y superior, y posee marginales distribuidas uniformemente en el intervalo $[0,1]$. Con respecto a la propiedad (b) dice que la función C es 2-creciente. Por lo tanto, para cada $(u, v) \in DomC$ se tiene que $0 \leq C(u, v) \leq 1$, de ahí su interés en el estudio de las funciones de probabilidad para el caso de las variables aleatorias dependientes [10].

La importancia de las funciones cópulas se describe en el Teorema de Sklar (1959) [11]. Las funciones cópulas tienen la propiedad de conectar las funciones de distribución conjuntas a sus marginales. Además, representan una forma de estudiar medidas de dependencia de escala libre y permiten construir familias de distribuciones bivariantes a partir de marginales dadas. [12].

A través de la forma funcional de la función cópula, estas se clasifican en cópulas fundamentales, Elípticas, Arquimedianas, de valor extremo y Arquimax. La familia de cópulas fundamentales permite estudiar estructuras de dependencia casi perfecta positiva (comonotonicidad), casi perfecta negativa (contramonotonicidad) e independencia (cópula del producto), por lo cual sirven de referencia para las demás estructuras. En este artículo, el modelo de supervivencia bivariante propuesto está basado en la familia de cópulas Arquimedianas, debido a que estas cópulas comparten algunas de sus propiedades con la función de supervivencia y también resuelve el problema de identificabilidad del modelo tratado en [3], a partir del modelo de supervivencia no paramétrico (9) propuesto en la sección anterior.

Además, el modelo propuesto tiene la particularidad de que la ocurrencia de un evento no censura la ocurrencia de otro u otros eventos, debido a la composición del evento compuesto, esto es, todos los instantes de tiempo son observados. Sin embargo, aun estando presente el evento compuesto, este no censura la ocurrencia del resto de los eventos del conjunto g, debido a que son eventos de naturaleza no terminal.

Cópulas Arquimedianas

Las cópulas Arquimedianas son todas las funciones $C(u,v)$ expresadas como

$$C(u,v) = \phi^{-1}(\phi(u) + \phi(v)) \qquad (13)$$

La función ϕ es la función generadora de la cópula Arquimediana, definida como $\phi\colon[0,1] \longrightarrow [0,\infty]$ con $\phi(1) = 0$, es continua, decreciente $(\phi'(t) < 0)$ y convexa $(\phi''(t) > 0)$ para todo $0 < t < 1$, con $\phi(0) = \infty$. Estas condiciones aseguran la existencia de la inversa ϕ^{-1}. Cualquier función ϕ con estas características es capaz de generar una función de distribución bivariante [13]. Las funciones cópulas con estas propiedades reciben el nombre de cópulas Arquimedianas estrictas y son las que tienen mejor ajuste con las funciones de supervivencia. Dentro de esta familia se encuentran las cópulas de Clayton, Frank, Gumbel y Joe, así como otras cópulas que se generan a partir de la combinación de estas como BB1, BB2, BB7 y BB8 [10].

Función cópula de supervivencia bivariante propuesta

Sean T_1 y T_2 los tiempos de fracaso de dos eventos de distinto tipo que experimenta un mismo individuo. A su vez, suponemos que la ocurrencia de al menos uno de los eventos constituye el evento compuesto, lo que implica que la ocurrencia de uno de los eventos no censura la ocurrencia del otro evento. Por consiguiente el i-ésimo individuo sobrevive si no experimenta ambos eventos. De acuerdo a [14], la función de supervivencia bivariante está dada por

$$S_{12}(t_1,t_2) = P[T_1 > t_1, T_2 > t_2] \qquad (14)$$

Sea C la función cópula definida entre las variables aleatorias T_1 y T_2, entonces con base en (14) y [10] se tiene

$$\begin{aligned} S_{12}(t_1,t_2) = S_1(t_1) + S_2(t_2) - 1 \\ + C(1 - S_1(t_1), 1 - S_2(t_2)) \end{aligned} \qquad (15)$$

Así, con base en las funciones cópulas se define la función C_θ^S como $C_\theta^S\colon[0,1]^2 \longrightarrow [0,1]$ tal que,

$$C_\theta^S(u,v) = u + v - 1 + C(1-u, 1-v) \qquad (16)$$

lo que implica que

$$S_{12}(t_1,t_2) = C_\theta^S\big(S_1(t_1), S_2(t_2)\big) \qquad (17)$$

La función C_θ^S es la cópula de supervivencia de T_1 y T_2, $C(\cdot)$ en la expresión (16) es cualquier función cópula de la familia Arquimediana, $S_k(t_k)$ con $k = 1{,}2$ son las funciones de supervivencia marginales dadas por (9) para cada T_k con $k = 1{,}2$ y θ es el parámetro desconocido de la cópula, que indica el grado de dependencia entre ambas marginales, respectivamente.

Con base en el Teorema de Sklar [11], la función C_θ^S une a través de la función de supervivencia conjunta sus funciones de supervivencia marginales. Además, por el Teorema de la Transformación Integral [9] se tiene que $S_{T_k}(T_k) \sim U(0{,}1)$, $k = 1{,}2$; así, cada componente de (17) converge a la distribución uniforme $U(0{,}1)$.

Al considerar la ocurrencia simultánea de dos eventos de distinto tipo sobre el mismo individuo, existe la posibilidad de que los instantes de tiempo en que ocurrieron estos eventos sean dependientes o independientes.

Si los tiempos de fracaso son dependientes, entonces la función de supervivencia bivariante está dada por

$$S_{12}(t_1,t_2) = C_\theta^S\left(e^{-M_1^*(t_1)}, e^{-M_2^*(t_2)}\right) \qquad (18)$$

Si los tiempos de fracaso son independientes, la función de supervivencia bivariante de (18) se simplifica al producto de las funciones de supervivencia marginales, que es equivalente a la función cópula del producto y coincide con el modelo no paramétrico de Kaplan y Meier [7], esto es

$$S_{12}(t_1,t_2) = C_\theta^S\left(e^{-M_1^*(t_1)}, e^{-M_2^*(t_2)}\right) = \prod_{k=1}^{2} e^{-M_k^*(t_k)} \qquad (19)$$

Otra situación que podemos considerar es que el evento compuesto consta de todos los $K = 2$ posibles eventos que experimenta cada individuo. Esto es, durante el periodo de seguimiento, $[0, \delta_i]$, cada individuo experimenta K fracasos. Por consiguiente, el i-ésimo individuo sobrevive si no experimenta al menos una de las componentes de K. En este caso, la función de supervivencia bivariante está dada por

$$S_{12}(t_1,t_2) = P[T_1 > t_1 \text{ ó } T_2 > t_2] \qquad (20)$$

que en términos de funciones cópulas está dada por

$$S_{12}(t_1,t_2) = 1 - C_\theta(1 - S_1(t_1), 1 - S_2(t_2)) \qquad (21)$$

La expresión (21) en [10] recibe el nombre de co-copula, y para tiempos de fracaso independientes la expresión (21) se simplifica a

$$S_{12}(t_1,t_2) = 1 - \prod_{K=1}^{2}(1 - S_k(t_k)) \qquad (22)$$

La ecuación (22) no coincide con el estimador no paramétrico de Kaplan y Meier.

Por lo tanto, a partir de las expresiones (17) y (21) se estima la probabilidad de supervivencia como función de la estructura del evento compuesto, lo que hace que la probabilidad de supervivencia estimada sea diferente, con base en cada caso a ser analizado.

Considerar la ocurrencia de tres o más eventos donde al menos dos de los tiempos de fracaso de cada tipo de evento son dependientes, no se trata de una simple generalización, se deben tomar en consideración algunas restricciones a partir de las funciones cópulas Arquimedianas multivariantes, que en este momento no es tema de interés.

MATERIALES Y MÉTODOS

Con el propósito de probar el modelo de supervivencia propuesto, se presenta un esquema de simulación para comparar las curvas de supervivencia cuando los tiempos de fracaso son dependientes versus tiempos de fracaso independientes.

Supongamos que cada individuo está expuesto a la ocurrencia de dos eventos, denotados como A y B, respectivamente. Primero vamos a considerar que el evento compuesto consta de al menos un evento y luego suponemos que consta de dos eventos. Así, para el primer caso la probabilidad de supervivencia está dada por la expresión (14) y para el segundo caso está dada por la expresión (20). En relación con la ocurrencia de los eventos, no existe un orden particular de ocurrencia entre ellos (a menos que la variable bajo estudio, tiempo de ocurrencia o permanencia, lo amerite, como en los casos donde se analiza el instante de tiempo en que el paciente padece alguna recaída en relación con el tiempo de intervención quirúrgica o muerte). Además, se tiene que $Y_{iA}^* = Y_{iB}^* = 1, i = 1,2,\dots,N$.

Sean T_{iA} y T_{iB} los tiempos de fracaso de cada tipo de evento experimentados por el i-ésimo individuo. Las distribuciones bivariantes se generan a partir de marginales Weibull con parámetros de forma y escala $\alpha_1 = 2, \beta_1 = 1$ y $\alpha_2 = 2.1, \beta_2 = 1.1$, para los tiempos de fracaso T_{iA} y T_{iB}, haciendo uso de la cópula de Frank de la familia Arquimediana.

Se simularon muestras de diferentes tamaños, $N = \{600,300,100,50,20,10\}$, observando que la forma funcional de la cópula no varía en función del tamaño de la muestra. Por esta razón se muestra un ejemplo para una muestra de $N = 600$ individuos. El periodo de seguimiento de cada individuo es aleatorio, $[0, \delta_i]$, con δ_i como el último instante de tiempo en que fue observado cada individuo, pero en ese momento no se observa ninguno de los eventos A o B, esto es, δ_i es un tiempo de fracaso censurado a la derecha, lo que implica que $T_{ik} \leq \delta_i$. Para este caso particular, se consideró que la distribución de probabilidad de δ_i es exponencial de parámetro 1 ($\theta = 1$), que es independiente del resto de los instantes de tiempo registrados para cada individuo.

En cada instante de tiempo $t_{(l)}$ donde se observa la ocurrencia de al menos uno de los eventos, se puede tener una cierta cantidad de individuos cuyos tiempos de fracaso son censurados. La suma de esta cantidad de individuos es igual a N, tal como se indica en [4].

En cuanto al grado de asociación entre los tiempos de fracaso, se consideraron desde independencia hasta un alto grado de dependencia para cada tamaño de N, cuyos valores son 0.00, 0.20, 0.60 y 0.85, obtenidos mediante el coeficiente de correlación tau de Kendall para funciones cópulas [10].

Cuando las variables aleatorias T_{iA} y T_{iB} son independientes, la función de supervivencia bivariante para estimar la probabilidad de superar ambos eventos, está dada por la ecuación (19) y la probabilidad de superar al menos uno de los eventos está dada por la ecuación (22). De lo contrario, usamos las ecuaciones (18) y (21), dependiendo del caso. En la sección siguiente se presentan los resultados, mediante la aplicación de estos modelos para estimar la probabilidad de supervivencia bivariante. Se elaboró para ello un algoritmo desarrollado con el software libre R [15] para funciones cópulas, mediante las librerías Copula y VineCopula [16], que puede ser solicitado vía e-mail.

RESULTADOS Y DISCUSIÓN

La siguiente matriz de gráficos muestra las diferentes estructuras de dependencia que se pueden describir entre los tiempos de fracaso de los eventos A y B.

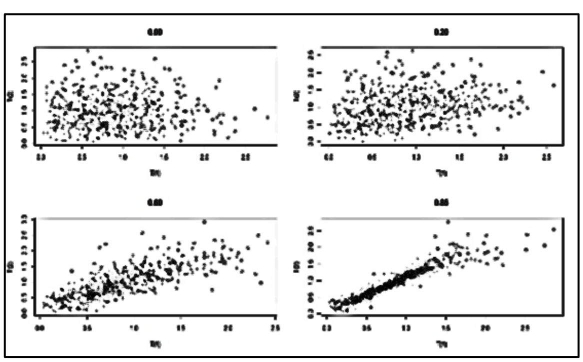

Fig. 1. Estructuras de Dependencia con tendencia positiva

El diagrama de dispersión ubicado en la parte superior izquierda, es una estructura con tiempos de ocurrencia de los eventos A y B independientes. Los demás casos son estructuras de dependencia con tendencia positiva que oscilan entre 0.20, 0.60 y 0.85. Esto significa que a medida que ocurre un evento, la ocurrencia del siguiente evento es casi inmediata. Sin embargo, la estructura de dependencia que se describa entre los tiempos de fracaso de ambos eventos también puede mostrar una tendencia negativa (0.00, -0.20, -0.60, -0.85), como se muestra en la figura 2.

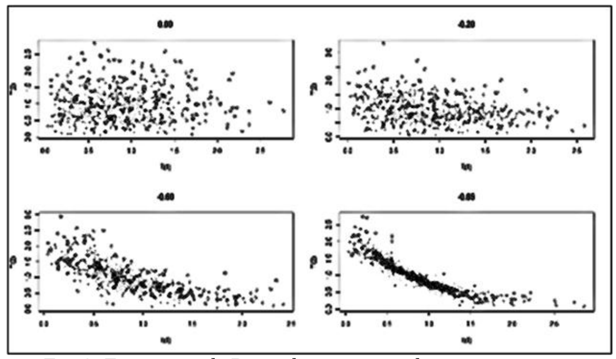

Fig. 2. Estructuras de Dependencia con tendencia negativa.

En este caso, la ocurrencia de un evento tiene la propiedad de retardar en cierta forma la ocurrencia del siguiente evento, dado que el grado de asociación entre los tiempos de fracaso sea menor a -0.60, aproximadamente. Por ejemplo, desde el punto de vista clínico, el inicio de un tratamiento médico en un instante de tiempo "temprano" puede retardar la ocurrencia de otro evento diferente como la reaparición de alguna patología, pero el inicio "tardío" del tratamiento puede tener como consecuencia la reaparición "temprana" del siguiente evento.

En vista de los diferentes grados de dependencia entre los tiempos de fracaso de los diferentes tipos de eventos, se tienen las siguientes curvas de supervivencia generadas a partir de los modelos de supervivencia propuestos (Figura 3), con base en los grados y estructuras de dependencia mostradas en la Figura 1. En particular se comparan dos estimadores, para dos diferentes composiciones del evento compuesto. Cuando los tiempos de fracaso son independientes, el modelo propuesto 1 (curva azúl) se simplifica al producto de las funciones de supervivencia marginales (cópula del producto) que coincide con el modelo de Kaplan y Meir (curva gris). En caso contrario, a medida que el grado de dependencia entre los tiempos de fracaso de ambos eventos, se incrementa, la separación entre las curvas (azul y gris) es proporcional.

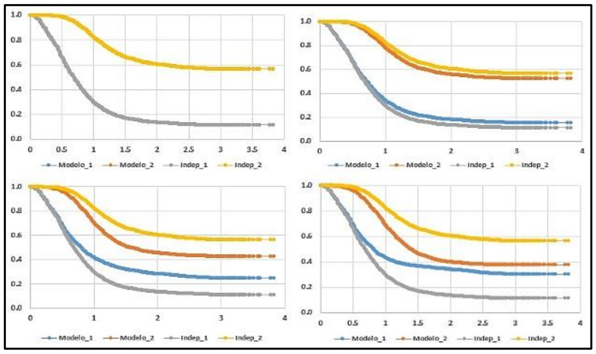

Fig. 3. Funciones de supervivencia estimadas mediante los modelos propuestos para tiempos de fracaso dependientes con tendencia positiva versus tiempos de fracaso independientes.

A partir de la Figura 3, las curvas de supervivencia identificadas de color amarillo (tiempos de fracaso independientes) y color ladrillo (tiempos de fracaso dependientes), estiman la probabilidad de supervivencia bajo la ocurrencia de al menos uno de los eventos A o B (ecuaciones 20, 21 y 22), respectivamente. La curva de color amarillo, en todos los casos se muestra por encima de resto de los modelos de supervivencia, indicando un mejor pronóstico de supervivencia a partir de ella. Sin embargo, esta probabilidad de supervivencia puede ser sobreestimada al no tomar en consideración el verdadero grado de dependencia que exista entre los tiempos de fracaso de cada tipo de evento. Si las estructuras de dependencia entre los tiempos de fracaso de los diferentes tipos de eventos tienen tendencia negativa (Figura 2) entonces las curvas de supervivencia estimadas a partir de los modelos bivariantes propuestos, están dados por (Figura 4):

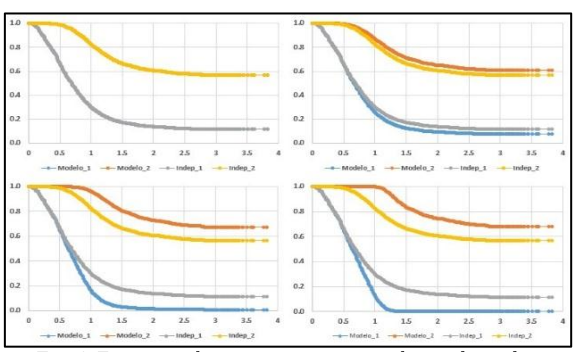

Fig. 4. Funciones de supervivencia estimadas mediante los modelos propuestos para tiempos de fracaso dependientes con tendencia negativa versus tiempos de fracaso independientes.

Si la relación de dependencia entre los tiempos de fracaso es negativa, la forma de las curvas de supervivencia estimada cambia completamente, como se muestra en la Figura 4.

A menor grado de dependencia la separación entre las curvas de supervivencia es proporcional, pero en sentido contrario a lo mostrado en la Figura 3. La probabilidad de sobrevivir a la ocurrencia de al menos un evento (curva color azul) es menor que si los tiempos de fracaso fueran independientes (curva color gris). A pesar de que la curva de color gris muestra un mejor pronóstico de supervivencia esta es una probabilidad de supervivencia sobrestimada.

En relación con la ocurrencia de ambos eventos, la probabilidad de supervivencia es mayor (curva color ladrillo) que la probabilidad de supervivencia en relación con la ocurrencia de al menos uno de los eventos dado que los tiempos de fracaso son independientes (curva color amarillo). En este caso, la curva color amarillo es una subestimación de la verdadera probabilidad de supervivencia.

En cuanto a la estimación de los parámetros de la cópula Frank para cada uno de los casos considerados, se usó el método de máxima verosimilitud, obteniendo los siguientes pares de estimaciones: $(\tau,\theta)=\{(0.0;0.0),(0.2;1.7),(0.6;7.7),(0.85;24.2)\}$ y $(-\tau,\theta)=\left\{\begin{matrix}(0.0;0.0),(-0.2;-1.7),(-0.6;-7.7),\\(-0.85;-24.2)\end{matrix}\right\}$.

Tratar la ocurrencia simultánea de dos o más eventos sobre un mismo individuo, en diferentes instantes de tiempo no es algo novedoso. Sin embargo, se pueden considerar algunas aristas que marcan la diferencia en relación con la metodología clásica. El método estadístico para tratar el problema de la ocurrencia de dos eventos sobre un mismo individuo es a través del método de riesgos en competencia mediante $T=\min(T_1,T_2,\ldots,T_k)$ o aplicando el método de combinación de riesgos propuesto por [17], que consiste en una generalización del estimador cópula gráfico [18] para más de dos riesgos en competencia. El estimador cópula gráfico resuelve el problema de identificabilidad y estima las distribuciones marginales para cada riesgo asumiendo una estructura de dependencia conocida usando una cópula específica para la distribución conjunta de los modos de fracaso [19]. Sin embargo, considerar $T=\min(T_1,T_2,\ldots,T_k)$ no siempre es el método más eficiente. Existen situaciones como las mostradas en este artículo, vía simulación, donde es necesario conocer los diferentes instantes de tiempo donde ocurrió cada evento de distinto tipo debido a que ellos pueden ser dependientes y además, constituir lo que hemos denominado evento compuesto.

Si T_1 y T_2 son independientes y $T = \min(T_1, T_2)$ el método de riesgos en competencia proporciona el mejor pronóstico de supervivencia en relación con la probabilidad de supervivencia global. Sin embargo, algunos investigadores consideran que el supuesto de independencia entre los T_{ik} de distinto tipo de eventos debe ser cuidadosamente considerado debido a que podemos estimar probabilidades de supervivencia sesgadas si la dependencia es ignorada. Al respecto, en [8] establece que al considerar la ocurrencia de distintos tipos de eventos durante el periodo de seguimiento de cada paciente, se observa una historia parcial del evento, lo que implica que independencia entre tiempos de fracaso de diferentes tipos de eventos rara vez puede ser asumida. Ahora bien, si la dependencia es ignorada el método de riesgos en competencia podría mostrar mejor pronóstico de supervivencia pero puede ser sesgada y por consiguiente se hace necesario plantear un modelo de supervivencia bivariante para tratar el problema de la dependencia entre los tiempos de fracaso de eventos de distinto tipo.

La probabilidad de supervivencia estimada varía de acuerdo a las características propias de los tiempos de fracaso de los diferentes tipos de eventos. Si se estudian distintas causas de muerte por paciente, el método de riesgos en competencia es el más apropiado. En cambio, si no se trata de causas de muerte sino de diferentes tipos de eventos no terminales que pudieran por ejemplo causar alguna patología en particular, el método propuesto puede ser una buena opción, aunque en todos los casos no proporcione el mejor pronóstico de supervivencia, en relación a un método clásico, debido a que lo importante es que el modelo usado capte la verdadera relación de dependencia que existe entre los diferentes tiempos de fracaso.

Por lo tanto, si los tiempos de fracaso realmente son dependientes pero los tratamos como si no lo fueran, la probabilidad de supervivencia puede ser subestimada o sobrestimada, dependiendo del caso. De lo contrario, la probabilidad de superar la ocurrencia de ambos eventos sigue estando dada por el estimador no paramétrico de Kaplan y Meier. Este estimador es pionero en el análisis de supervivencia y actualmente aún continúa siendo uno de los de mayor uso en el campo de la bioestadística y otras áreas de la medicina donde se aplica la estadística como herramienta del análisis de datos clínicos.

Cabe señalar que a partir del modelo de supervivencia propuesto se pueden tratar problemas de interés práctico en relación a la supervivencia de los pacientes mediante el análisis de eventos de naturaleza no terminal. Puede ser de interés práctico para un Ginecólogo estudiar los instantes de tiempo en que ocurren las diferentes lesiones cervico-vaginales en una misma paciente, que pueden traer como consecuencia a lo largo del tiempo, el desarrollo del cáncer de cuello uterino (visto como evento compuesto). El virus de inmunodeficiencia humana (VIH) es una patología que también podemos considerar como un evento compuesto, debido a que se requiere hacer un seguimiento a cada paciente acerca de la evolución de los biomarcadores relacionados con esta patología, como son, los instantes de tiempo para cada paciente en que la carga viral plasmática sea mayor a 100000 copias o que la población linfocitaria sea menor de 200 células/mm^3 [20]. El modelo propuesto permite medir la probabilidad de supervivencia a partir de la ocurrencia de a lo sumo ambos eventos de naturaleza no terminal, lo que constituye la estructura y ocurrencia del evento compuesto. Así, estos y otros casos clínicos pueden ser tratados en un siguiente trabajo de investigación donde se muestre la aplicación del modelo de supervivencia propuesto.

CONCLUSIONES

En el artículo se mostró que la probabilidad de supervivencia puede ser determinada a partir del conteo de eventos sobre un mismo individuo, analizados como variables aleatorias no latentes. De acuerdo al objetivo, aplicación e interés por parte del investigador los eventos pueden ser agrupados de tal manera que se defina un nuevo evento que hemos denominado evento compuesto.

El evento compuesto puede ser de interés particular en el área de la salud, debido a que muchas de las patologías estudiadas hasta hoy día son producto de una combinación de otras patologías. Esta situación puede generar que los distintos tipos de eventos que conforman el evento compuesto, que por tener características similares entre sí, los instantes de tiempo en que ocurra cada evento sean dependientes, así como también, por el hecho de que comparten un elemento en común, el paciente o individuo bajo estudio.

Puede suceder que sobre un mismo individuo actúan varios eventos de distinto tipo y sin embargo, sus tiempos de fracaso sean independientes. Si ese es el caso, la probabilidad de supervivencia se determina mediante un modelo clásico del análisis de supervivencia, como el estimador no paramétrico de Kaplan y Meier, a partir del producto de las funciones de supervivencia marginales, que es equivalente a la función cópula del producto. Si se determina que los tiempos de fracaso son dependientes, la probabilidad de supervivencia se estima a través de una función cópula, cuya elección depende en parte de la forma funcional de la estructura de dependencia entre los tiempos de fracaso, y de la función acumulada media para eventos de distinto tipo no recurrentes (FAME$\overline{R}$). Por consiguiente, una ventaja del modelo propuesto es que puede ser usado para cualquier caso (T_{ik} independientes o dependientes) debido a que, si los T_{ik} realmente son independientes, el modelo propuesto se simplifica al estimador no paramétrico de Kaplan y Meier (1958).

Ahora bien, si los tiempos de fracaso de los distintos tipos de eventos son dependientes y estos son tratados como independientes, la probabilidad de supervivencia estimada es sesgada y las decisiones tomadas a partir de estos resultados serán erróneas e ineficientes. Por lo tanto, dependiendo del área de aplicación de esta metodología estadística propuesta, el hecho de tener un mejor pronóstico de supervivencia podrá resultar a favor o en contra de cada una de las unidades bajo observación.

APÉNDICE

Demostración de Proposición 1

Sea $Y_g^*(t)$ la cantidad de eventos de distinto tipo que conforman el evento compuesto experimentados por el mismo individuo en el instante de tiempo t y $Y^*(t)$ el número total de individuos en el instante t en riesgo de experimentar un nuevo evento. La variable aleatoria $Y^*(t)$ está sujeta a que los instantes de tiempo t sean o no censurados, esto es, si t es censurado, $Y^*(t) = N - 1$, en caso contrario, $Y^*(t) = N$, así algunos individuos tienen chance de ser observados varias veces. Para que cada individuo $i, i = 1,2, \ldots, N$ experimente el evento compuesto, requiere ser observado en diferentes instantes de tiempo a lo largo de su periodo de seguimiento $[0, \delta_i]$, donde δ_i es el último instante de tiempo en que fue observado el i-ésimo individuo, pero en ese momento no ocurre ninguno de los eventos de interés. Así, la proporción de individuos que completan el evento compuesto en el instante de tiempo t está dada por, $m_k^*(t) = Y_k^*(t)/Y^*(t)$, lo que implica que la proporción de los que no experimentan este evento, esto es, los que sobreviven al instante de tiempo t, está dada por, $1 - m_k^*(t)$. Si los tiempos de fracaso en que ocurren los diferentes tipos de eventos son independientes, entonces la probabilidad de supervivencia está dada por el estimador no paramétrico de Kaplan y Meier (1958), esto es, $S_{KM}(t) = \prod_{t_{(l)}<t}[1 - m_k^*(t)]$. Por otro lado se tiene que, $M_k^*(t) = \sum_{t_{(l)}<t} m_k^*(t)$, es el número acumulado medio de eventos de tipo k en t, lo que implica que $-\ln[S_{KM}(t)] \cong M_k^*(t)$, que a su vez podemos escribir como $-\ln[S_T(t)] \cong M_k^*(t)$, de donde se tiene que $S_T(t) = \exp(-M_k^*(t))$. En relación con el evento compuesto, podemos concluir que $S_T(t) = \exp\left(-M_g^*(t)\right) = \exp(-\sum_{t_{(l)}<t} M_k^*(t))$. Por lo tanto, calcular $S_T(t) = \exp(-\Lambda(t))$ con base en la definición de función de supervivencia, es equivalente a calcular $S_T(t) = \exp\left(-M_g^*(t)\right)$, de donde se tiene que, $M_g^*(t) \approx \Lambda(t)$.

Demostración de Proposición 2:

A partir de la Proposición 1, $S_T(t) = e^{-\Lambda(t)} \equiv e^{-M_g^*(t)}$. Además, por definición de función de riesgo (con base en la ocurrencia de sólo un evento), se tiene que $\lambda(t) = \frac{f(t)}{S(t)} = -\frac{\partial}{\partial t} lnS(t) = -\frac{\partial}{\partial t}\left(-M_g^*(t)\right) = m_g^*(t)$.

Demostración de Proposición 3:

Con base en la Proposición 1, $S_{12\ldots K}(t_1, t_2, \ldots, t_K) = \exp(-\sum_{k=1}^{K} M_k(t)) = \exp\left(-M_1^*(t)\right) \exp\left(-M_2^*(t)\right) \ldots \exp\left(-M_K^*(t)\right) = S_1(t_1)S_2(t_2) \ldots . S_K(t_K) = \prod_{k=1}^{K} S_k(t_k)$

Esta igualdad siempre se cumple para la función de supervivencia global, que es igual al producto de las funciones de supervivencia marginales.

REFERENCIAS BIBLIOGRÁFICAS

[1] Crowder, M. Classical Competing Risks. New York: Chapman & Hall/CRC; 2001.

[2] Fine J, Jiang H, Chappell R. On semi-competing risk data. Biometrika 2001. 88(4): 907-919.

[3] Wang A, Chandra K, Xu R, Sun J. The identifiability of dependent competing risks models induced by bivariate frailty models. Scandinavian Journal of Statistics: Theory and Applications 2015. 42:427-437.

[4] Nelson W. Recurrent events data analysis for product repairs, disease recurrences, and other applications. New York: Asa Siam; 2002.

[5] Velazco J, González F, Diaz T, Peña J, Araque M. Profiles of enteropathogens in asymptomatic children from indegenous communities of Mérida, Venezuela. Journal Infection Dev Ctries 2011. 5(4):278-285.

[6] Lee E, Wang J. Statistical methods for survival data analysis. New Jersey: Wiley-Interscience; 2003.

[7] Kaplan E, Meier P. Nonparametric estimation from incomplete observations. Journal of the American Statistical Association 1958. 53:457-481.

[8] Boracchi P, Orenti A. Survival functions in the presence of several events and competing risk: Estimation and interpretation beyond Kaplan-Meier. International Journal of Statistics in Medical Research, 2015. 4:121-139.

[9] Resnick S. A Probability Path. New York: Springer-Verlag; 1999.

[10] Nelsen R. An Introduction to Copulas. New York: Springer; 2006.

[11] Schweizer B, Sklar D. Statistical metric spaces. Pacific J. Math 1960. 10:131-134.

[12] Rodríguez J, Úbeda M. A new class of bivariate copulas. Statistics & Probability 2004. 66: 315-325.

[13] Trivedi P, Zimmer D. Copula Modelling: An introduction for practitioners. Journal American Statistical Association 2005. 1:1-111.

[14] Eland R, Johnson N. Survival models and data analysis. New York: Wiley-Interscience; 1999.

[15] R Core Team. A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria, 2016. URL https://www.R-project.org/.

[16] Kojadinovic I, Yan J. Modeling Multivariate Distributions with Continuous Margins Using the copula R Package. Journal of Statistical Software, 2010. 34(9): 1-20.

[17] Lo S, Wilke R. A copula model for dependent competing risks. Journal of the Royal Statistical Society, 2010. 59(2), 359-376.

[18] Zheng M y Klein J. Estimates of marginal survival for dependent competing risk base on an assumed copula. Biometrika, 1995. 82: 12-38.

[19] Paz M, Yáñez S, Lopera C. Estudio comparativo del efecto de la dependencia en modelos de riesgos competitivos con tres modos de falla vía estimadores basados en cópulas. Ingeniería y Competitividad, 2014. 16(1), 169-183.

[20] Timaure R. Método unificado para el estudio de la sensibilidad a los datos faltantes y las desviaciones en los supuestos de la distribución en modelos de efectos mixtos. Tesis de Doctorado en Estadística, Mérida-Venezuela: Instituto de Estadística Aplicada y Computación. Universidad de Los Andes, 2017.