UNIVERSIDAD DE LOS ANDES
FACULTAD DE CIENCIAS ECONÓMICAS Y SOCIALES
INSTITUTO DE ESTADÍSTICA APLICADA Y COMPUTACIÓN
MAESTRÍA EN ESTADÍSTICA

MODELACIÓN DEL RETRASO EN LOS ASCENSOS DE LOS PROFESORES DE LA UNIVERSIDAD DE LOS ANDES (2011)

AUTORA: Econ. Ana Carolina Alvarez Gubinelli
TUTOR: Douglas Rivas
CO-TUTOR: Rubén Delgadillo

TRABAJO DE GRADO

Presentado ante la Ilustre Universidad de Los Andes
como requisito final para optar al Grado Académico de
Magíster Scientiae en Estadística

Mérida, Junio de 2012

DEDICATORIA

A mi madre, Anna Gubinelli
A mi padre, Carlos E. Alvarez, y
A mi hermano, Carlos A. Alvarez

AGRADECIMIENTOS

A Dios, por concederme salud y guiar mis pasos.

A mis padres y hermano, por su apoyo incondicional.

A mis amigos, por estar presentes siempre que los necesito.

Al profesor Douglas Rivas, por su valiosa ayuda y orientación.

A los profesores, Elizabeth Torres, Prasad Sinha y Rubén Delgadillo, por dedicarme parte de su tiempo.

Al personal docente y administrativo del Instituto de Estadística Aplicada y Computación, por su continua asistencia.

ÍNDICE DE CONTENIDO

ÍNDICE DE TABLAS

ÍNDICE DE GRÁFICOS Y FIGURAS

UNIVERSIDAD DE LOS ANDES
FACULTAD DE CIENCIAS ECONÓMICAS Y SOCIALES
INSTITUTO DE ESTADÍSTICA APLICADA Y COMPUTACIÓN
MAESTRÍA EN ESTADÍSTICA

MODELACIÓN DEL RETRASO EN LOS ASCENSOS DE LOS PROFESORES DE LA UNIVERSIDAD DE LOS ANDES (2011)

Autora: Econ. Ana Carolina Alvarez Gubinelli
Tutor: Douglas Rivas
Co-Tutor: Rubén Delgadillo
Fecha: Mayo, 2012

RESUMEN

Para que un miembro ordinario del personal docente y de investigación de la Universidad de Los Andes ascienda a la categoría de profesor titular, además de cumplir con los requisitos restantes que establece el Estatuto del Personal Docente y de Investigación de la Universidad, deben transcurrir quince años. Por su parte, de quienes han alcanzado la categoría de profesores titulares, no todos lo han hecho en el tiempo reglamentario. Por lo tanto, resulta interesante conocer, si son significativas las diferencias existentes entre las escuelas que conforman al núcleo Mérida, respecto al tiempo promedio de retraso para alcanzar el máximo nivel del escalafón.

Los integrantes del personal docente y de investigación de la Universidad de Los Andes están agrupados por escuelas. Dicha estructura de agrupación podría influir sobre el comportamiento de los miembros de estos grupos, en lo relacionado al cumplimiento de los requisitos para ascender, pues las características de las escuelas son diferentes entre sí. Dada la estructura de agrupación que presenta la Universidad, se empleó la técnica conocida como modelos de regresión multinivel o modelos jerárquicos.

Los resultados indican que existen diferencias significativas entre las escuelas en cuanto al tiempo promedio de retraso, para alcanzar el máximo nivel del escalafón. Por otro lado, se comprobó que las variables del primer nivel que contribuyen en la explicación de la variable dependiente que es el retraso, son doctorado y dedicación. Cabe destacar, que la relación existente entre el retraso y la variable doctorado, es similar entre las diferentes escuelas, mientras que la relación entre la variable objeto de estudio y la dedicación no es la misma para todas las unidades del segundo nivel. En cuanto a las variables medidas sobre los grupos, resultó que ninguna contribuye en la explicación de las diferencias que existen entre dichos grupos respecto al retraso promedio.

Palabras claves: modelos jerárquicos, escuelas, profesores, ascensos, retraso.

CAPÍTULO I

EL PROBLEMA DE LA INVESTIGACIÓN

Planteamiento del Problema

Para que un miembro ordinario del personal docente y de investigación de la Universidad de Los Andes que ingresó como instructor, ascienda a la categoría de profesor titular, además de cumplir con los requisitos restantes que establece el Estatuto del Personal Docente y de Investigación de la Universidad de Los Andes, deben transcurrir quince años. Dos años como instructor, cuatro años como profesor asistente, cuatro años como profesor agregado y cinco años como profesor asociado.

Por su parte, de quienes han alcanzado la categoría de profesores titulares, así como los demás niveles del escalafón, no todos lo han hecho en el tiempo señalado en el reglamento de la Universidad. Según Torres y Torres (2001), un 40.7% de los miembros del personal docente y de investigación del núcleo Mérida, tuvieron entre 1 y 5 años de retraso, y un 21.5%, tardaron más de 5 años. Este estudio también reveló que no existen diferencias significativas en el tiempo que demoran para ascender los profesores de la Universidad de Los Andes, entre las distintas facultades del núcleo Mérida. Es decir, las facultades no son un elemento que influya en este aspecto. Los autores llegaron a ésta conclusión mediante un análisis de varianza.

Teniendo en cuenta que los resultados mencionados anteriormente corresponden a un estudio llevado a cabo hace más de diez años y se enfocó en las facultades que conforman el núcleo Mérida. Resulta interesante conocer, si son significativas las diferencias entre las escuelas que conforman al núcleo Mérida pero utilizando información un poco más reciente.

Los integrantes del personal docente y de investigación de la Universidad de Los Andes están agrupados por escuelas. Dicha estructura de agrupación podría influir sobre el comportamiento de los miembros de estos grupos, en lo relacionado al cumplimiento de los requisitos para los ascensos, pues las características de las escuelas son diferentes entre sí. Ahora bien, dada la estructura de agrupación que presenta la Universidad, se emplea la técnica conocida como modelos de regresión multinivel. La cual, además de evidenciar si existen diferencias significativas entre las escuelas en cuanto al promedio de años de retraso, permite también identificar posibles variables que influyan sobre el retraso que presenten los profesores para ascender a través de los diferentes niveles del escalafón.

Son objeto de estudio, aquellos profesores ordinarios pertenecientes al núcleo Mérida que para octubre del año 2011, tengan más de 15 años en la universidad y aún no hayan alcanzado la titularidad, y quienes alcanzaron la titularidad en un tiempo superior al que establece el reglamento de la universidad. Se excluyen de este estudio a los profesores contratados, jubilados y pensionados.

Formulación del problema

Las preguntas de investigación son las siguientes: ¿Existen diferencias significativas en el tiempo transcurrido por encima del reglamentario, para alcanzar los diferentes niveles del escalafón, entre los profesores de las escuelas que conforman al núcleo Mérida de la Universidad de Los Andes, para octubre del año 2011? ¿Qué variables del nivel de los profesores influyen sobre el retraso para alcanzar estos niveles de formación? ¿Qué variables del segundo nivel explican la variación entre grupos de la variable dependiente? ¿Existe interacción entre variables explicativas de los diferentes niveles?

Antecedentes

El tema de la demora en los ascensos por parte del personal docente y de investigación de la Universidad de Los Andes ha sido estudiado de diversas maneras. Algunas de ellas se presentan a continuación.

Torres y Torres (1996), partiendo de una muestra estratificada por facultades de los profesores ordinarios de la Universidad de Los Andes, determinaron la magnitud del retraso de cada profesor comparando el tiempo transcurrido desde su adscripción hasta el año 1995, con el tiempo de permanencia establecido para cada categoría del escalafón. Las facultades que presentaron cinco años o más de retraso fueron: Arquitectura (59.5%), Medicina (39.7%), Núcleo Universitario Rafael Rangel (38.8%), Ciencias Forestales (32.4%) y Ciencias (30.3%). En cuanto al total de profesores por categoría, las que reflejaron porcentajes mayores a los cinco años de retraso en la presentación de sus trabajos de ascenso fueron: Asistente (41,2%), Agregado (59,4%) y Asociado (36,9%). Los autores también señalan que el incumplimiento de los ascensos dentro de los lapsos establecidos, guarda relación con la sobrecarga de horas de docencia, el desempeño de funciones administrativas, el nivel académico y la experiencia en investigación.

Torres y Torres (2001), realizaron un análisis de varianza con el propósito de explicar la perturbación del funcionamiento del sistema de ascensos de la Universidad de Los Andes por la intervención de factores estructurales y de desempeño. Se tomó una muestra aleatoria estratificada por facultades de 172 profesores del nucleó Mérida. Los resultados reflejan que 37.8% de los profesores ascendieron a tiempo, 40.7% tuvieron entre 1 y 5 años de atraso, 21.5% presentaron más de 5 años de retraso. Lo anterior indica que un 62.2% de los profesores presentaron fuera del tiempo su último trabajo de ascenso. El análisis de correspondencia mostró interacción entre factores estructurales y de desempeño. El análisis de varianza reveló que no son significativas las diferencias en el promedio de años de retraso por facultad, pero sí, en cuanto al nivel académico y categoría.

Ramoni, Orlandoni, Sinha, y Rivas (2007), efectuaron una investigación sobre el sistema de ascensos y de fijación de sueldos a través de las normas de homologación, con el objetivo de evaluar el papel que juega la acumulación de capital humano sobre ellos. En dicho trabajo se estimó el peso que tienen la experiencia y la capacitación sobre los sueldos, en la Universidad de Los Andes, Núcleo Mérida. Se realizó un análisis de regresión relacionando el salario de los profesores ubicados en los diferentes escalafones, con variables como experiencia y nivel educativo. También se analizó la relación entre el salario y algunos índices de capacidad profesoral. Los resultados obtenidos indican que las universidades no brindan incentivos para la inversión en capital humano, más allá del mínimo requerido para pasar de un escalafón a otro. Entendiéndose por capital humano, las destrezas y conocimientos adquiridos por el trabajador a través del estudio, el entrenamiento y la experiencia.

Sinha, Ramoni, Orlandoni, Torres y Figueroa (2007), publicaron un artículo donde la idea principal fue proponer el concepto de Riesgo Académico Institucional (RAI), como el conjunto de factores que exponen a una institución al deterioro de su calidad. Se construyen índices como el de Capacitación Profesional (ICP), el cual incluye elementos tanto de formación académica como de antigüedad. De acuerdo con los resultados, el coeficiente de Gini refleja desigualdad por facultad en el nivel de riesgo, medido por ICP. Por su parte, teniendo en cuenta que los profesores que pertenecen a la categoría agregado y que tienen maestría, han alcanzado un nivel medio de formación, y que aquellos profesores que pertenezcan a la categoría titular y tengan un titulo de doctorado, han alcanzado un nivel alto de formación. El 84.90% de los profesores de la Universidad de Los Andes, para agosto del 2006, con 6 años de antigüedad, cumplió con el nivel medio de formación, y sólo un 54.05% alcanzó el nivel alto de formación en el tiempo reglamentario.

Alessio (2007), con el propósito de evaluar aquellos factores que influyen sobre el retraso que presenta el personal docente y de investigación de la Universidad de Los Andes para ascender en el escalafón, ajustó un modelo de regresión logística. Los resultados indican que las variables que resultaron ser significativas en la explicación del comportamiento del retraso en los ascensos de los profesores son las siguientes: falta de tiempo, de motivación y de información.

Justificación

Como se mencionó anteriormente, el estudio sobre la situación de retraso en los ascensos de los profesores de la Universidad de Los Andes llevado a cabo por Torres y Torres en el año 2001, se enfocó en los profesores agrupados en las diferentes facultades que conforman el núcleo Mérida, fue realizado hace más de diez años y se aplicó un análisis de varianza para estudiar el comportamiento del retraso entre las facultades.

El propósito ahora, es llevar a cabo un análisis por escuelas y utilizar otra técnica estadística para estudiar el problema del tiempo que demoran los profesores para ascender. Debido a la estructura de agrupación que presenta la Universidad de Los Andes se utiliza la herramienta estadística conocida como modelos de regresión multinivel o modelos jerárquicos, para estudiar el fenómeno del retraso en los ascensos. Estos modelos deben ser utilizados cuando los datos en una población presenten una jerarquía o estructura de agrupación. Una jerarquía es definida según Goldstein (1999), como el conjunto de unidades agrupadas en diferentes niveles. Las unidades del primer nivel serían los profesores, agrupados por escuelas, que serían las unidades del segundo nivel.

El uso de esta herramienta hace posible: ajustar un modelo de regresión que permita conocer si el tiempo promedio de retraso es diferente entre las escuelas; verificar la influencia de variables, medidas sobre los profesores y las escuelas, con respecto a la situación de retraso en los ascensos de los profesores; observar si las variables medidas sobre los profesores que resulten ser significativas, afectan en la misma medida a las diferentes escuelas; y si existe interacción entre variables explicativas de los diferentes niveles. Cabe destacar, que el conocer sí existen diferencias significativas entre los profesores de las diferentes escuelas en cuanto al retraso promedio, y sobre todo, saber cuáles son las posibles variables que explican el comportamiento del tal retraso. Facilita la búsqueda de soluciones en torno al problema del retraso que presentan los profesores de la Universidad de Los Andes para ascender en el escalafón.

Objetivos de la Investigación

Objetivo General

Estudiar el retraso que presentan los profesores de las diferentes escuelas que conforman el núcleo Mérida de la Universidad de Los Andes para ascender a través de los diferentes niveles del escalafón.

Objetivos Específicos

- Estimar la demora que presentan los profesores de las diferentes escuelas para ascender en el escalafón.

- Describir el comportamiento del retraso que presentan los profesores para ascender por escuela.

- Ajustar un modelo que permita observar si existen diferencias significativas entre las escuelas, en el tiempo que demoran los profesores para ascender a través de las distintas categorías, fuera del establecido en el reglamento de la Universidad de Los Andes.

- Identificar variables medidas sobre los profesores y escuelas que influyan en el comportamiento del retraso de los profesores para ascender.

- Determinar si existe interacción entre variables explicativas de los diferentes niveles.

CAPÍTULO II

MARCO TEÓRICO

Los Profesores Ordinarios y Su Ascenso en el Escalafón

Un profesor de la Universidad de Los Andes se sitúa en la condición ordinario una vez que ha ganado un concurso de oposición, mediante el traslado previsto en el Artículo 107 de la Ley de Universidades o por razón de reincorporación, así lo indica el Artículo 13 del Estatuto del Personal Docente y de Investigación de la Universidad de Los Andes. Por su parte, son miembros ordinarios: los instructores, los profesores asistentes, los profesores agregados, los profesores asociados, y los profesores titulares. Esto, según el Artículo 87 de la Ley de Universidades, y Artículo 9 del Estatuto del Personal Docente y de Investigación de la Universidad de Los Andes.

El Artículo 14, del Estatuto del Personal Docente y de Investigación de la Universidad de Los Andes establece que para ser miembro ordinario del personal docente y de investigación, es necesario poseer Título de Postgrado, a nivel de Maestría o su equivalente, en el área objeto de concurso, siempre que dicho título haya sido alcanzado en una universidad de reconocido prestigio académico, a juicio del Consejo Universitario (El Consejo Universitario en reunión ordinaria del 23.10.1991 y según Resolución No. 2469, conoció el siguiente planteamiento: “Cuando se exige como requisito específico un título de pregrado determinado para un Concurso, bien sea de Credenciales o de Oposición, y el aspirante no posee el título requerido, pero tiene una Maestría en la misma área objeto del Concurso. Preguntamos si el título de Postgrado puede suplir al de pregrado para obtener el derecho a concursar", y en consecuencia acordó “...que en efecto, una maestría o equivalente en el área objeto del concurso, cubre el requerimiento legal del título de pregrado exigido en el área”).

Los miembros Ordinarios del personal docente y de investigación se ubicarán y ascenderán en el escalafón de acuerdo con sus credenciales o méritos científicos y sus años de servicio. El escalafón de los miembros ordinarios del personal docente y de investigación comprende las siguientes categorías: Instructores, Profesores Asistentes, Profesores Agregados, Profesores Asociados, y Profesores Titulares. Esto, de acuerdo con los Artículos 161 y 162 del Estatuto del Personal Docente y de Investigación de la Universidad de Los Andes; 87 y 89 de la Ley de Universidades.

Los instructores

El Artículo 64 del Estatuto del Personal Docente y de Investigación de la Universidad de Los Andes establece, que de acuerdo a la Ley y a las características de las funciones desempeñadas, el instructor es un docente e investigador en proceso de formación y su preparación idónea es garantía de mejoramiento académico de la Institución. Según su Parágrafo Único. Se exceptúan de la condición de instructor en proceso de formación y por consiguiente de la reducción de la carga docente correspondiente:

a) A quienes habiendo cumplido dos años en tal condición, no se encuentren realizando cursos de nivel programados y aprobados por la unidad académica correspondiente.
b) Al instructor que ha cumplido dos años como profesor contratado y gane concurso de oposición en la misma área y no se encuentre en la condición antes señalada.
c) Al que ingrese con postgrado.

El instructor está obligado, antes de ascender a la categoría de asistente, a cumplir los programas de formación académica establecidos por la unidad docente y de investigación a la que esté adscrito. Esto, de acuerdo con el Artículo 65 del Estatuto del Personal Docente y de Investigación de la Universidad de Los Andes. Como se menciona en el Artículo 66, del mismo Estatuto. Las unidades académicas están obligadas a tomar las previsiones que posibiliten el cumplimiento de los planes de formación de los instructores.

Los profesores asistentes

Los requisitos que debe cumplir un instructor para ascender a profesor asistente, según el Artículo 163 del Estatuto del Personal Docente y de Investigación de la Universidad de Los Andes, se presentan a continuación:

1. Haber ejercido actividades docentes, de investigación o de extensión, con el carácter de dedicación exclusiva o tiempo completo, por lo menos durante dos años.
2. Haber cumplido los planes de formación aprobados por la unidad docente o de investigación a la cual esté adscrito, de acuerdo con la programación establecida por los Consejos de Departamento, de Escuela y de Facultad o de Núcleo.
3. Informe favorable del jefe de la unidad académica a que esté adscrito el aspirante sobre el cumplimiento de su plan de formación y su rendimiento académico.
4. Aprobar el trabajo al que se refiere el Artículo 89 de la Ley de Universidades.

Los profesores agregados

Como lo establece el Artículo 164. Para ascender a la categoría de agregado, el profesor asistente debe cumplir las condiciones siguientes:

1. Haber realizado actividades docentes, de investigación o de extensión como profesor asistente, durante un período no menor de cuatro años.
2. Aprobar un trabajo original como credencial de mérito.

Los profesores asociados

El profesor agregado que desee ascender a la categoría de profesor asociado requiere, conforme con el Artículo 165 del Estatuto del Personal Docente y de Investigación de la Universidad de Los Andes, lo sucesivo:

1. Haber ejercido actividades docentes, de investigación, o de extensión durante cuatro años como profesor agregado.
2. Poseer el título de doctor o en su defecto de maestría, o grados equivalentes, debidamente legalizados cuando hayan sido conferidos por una institución extranjera.
3. Aprobar un trabajo original como credencial de mérito. Este trabajo podrá consistir en un texto de comprobada utilidad docente en el campo en que profesa el aspirante.

De acuerdo con el Artículo 166. Aquellos profesores que no cumplan con el segundo requisito exigido en el artículo 165, para ascender a la categoría de profesor asociado, podrán sustituir el mismo por alguno de los méritos siguientes:

1. Haber recibido algún premio de una institución nacional o internacional, pública o privada, académica o de otra índole, de prestigio, en reconocimiento de méritos académicos sobresalientes, acumulados a lo largo de la trayectoria docente o de investigación y no de méritos inherentes a una actividad eventual o particular, a juicio del Consejo Universitario.
2. Haber publicado trabajos de indiscutible valor científico o humanístico en órganos de divulgación nacional o internacional, a juicio de una comisión ad hoc nombrada al efecto por el Consejo de Facultad o Núcleo respectivo. Calificarán sólo los trabajos producidos luego del último ascenso.
3. Elaborar un trabajo especial, además del trabajo de ascenso, que reúna todas las condiciones exigidas para este último y que haya sido aprobado por un jurado nombrado al efecto por el Consejo de Facultad o Núcleo,

cuyos integrantes deben ser diferentes de los del jurado que ha de conocer el trabajo de ascenso. Este trabajo no podrá ser realizado en equipo.

Los profesores titulares

Como lo establece el Artículo 166 del Estatuto del Personal Docente y de Investigación de la Universidad de Los Andes. Para alcanzar la categoría de titular es necesario:

1. Haber realizado actividades docentes, de investigación o de extensión como profesor asociado, por lo menos durante cinco años.

2. Poseer credenciales científicas adecuadas a aquella categoría, la cual representa el más alto grado en el escalafón del personal docente y de investigación. Cabe destacar que, a los fines de la evaluación de estas credenciales se tomará en cuenta:

   a) Para los profesores dedicados fundamentalmente a la docencia: la responsabilidad, aplicación y eficiencia demostradas en el cumplimiento de su magisterio; la publicación de textos o monografías, la elaboración de otros materiales para la enseñanza, y su contribución en la formación del personal docente de jerarquía inferior y en los cursos de postgrado.
   b) Para los profesores dedicados fundamentalmente a la investigación: la realización efectiva y satisfactoria de los diversos estudios que le hubieren sido encomendados por el Centro o Instituto a que estén adscritos, las obras y trabajos publicados y su contribución a la docencia.

3. Aprobar el trabajo exigido en el artículo 89 de la Ley de Universidades. Este trabajo podrá consistir en un texto de comprobada calidad científica y utilidad docente en el campo en que profesa el aspirante.

Dedicación de los Profesores Ordinarios

El artículo 109 del Estatuto del Personal Docente y de Investigación de la Universidad de Los Andes establece que, según el tiempo que dediquen a las actividades correspondientes, los profesores ordinarios se clasificarán en:

1. Profesor a dedicación exclusiva.
2. Profesor a tiempo completo.
3. Profesor a medio tiempo.
4. Profesor a tiempo convencional.

Debe tenerse en cuenta, la participación en actividades docentes, administrativas, de investigación y de extensión, previstas en los planes y programas de la facultad o núcleo, y el tiempo dedicado al cumplimiento de las tareas asignadas, para la clasificación de un profesor en cualquiera de las categorías ya mencionadas. Ello, acorde con el artículo 110 del estatuto de la universidad.

Según los artículos 115, 116 y 118 del estatuto ya mencionado, un profesor a dedicación exclusiva deberá permanecer en ambientes universitarios durante cuarenta horas semanales cumpliendo las tareas que le hayan sido asignadas, no puede realizar actividades remuneradas fuera de la facultad sin permiso del Consejo Universitario, puede realizar labores de investigación, o cumplir labores de docencia, administración o extensión, o varias de estas actividades.

Todo aquel instructor en etapa de formación debe ser a dedicación exclusiva. La clasificación de un profesor bajo esta dedicación debe estar plenamente justificada en cuanto al cumplimiento de objetivos fundamentales para la buena marcha de la facultad o núcleo a que pertenezca. La solicitud de esta clasificación será hecha por la unidad académica a la que el profesor esté adscrito, como es el caso de departamentos, institutos o centros. Esto conforme con los artículos 114, 117 y 119 del estatuto de la universidad.

De acuerdo con los artículos 120 y 121 del mismo estatuto, las autoridades universitarias, decanos, directores de escuela, jefes de departamento, jefes y coordinadores de consejos y oficinas centrales, para desempeñar sus funciones, podrán pasar a dedicación exclusiva desde el momento en que asuman su cargo, haciendo una solicitud por escrito y sin sufrir desmejoras en las primas correspondientes a tales cargos, con excepción de los jefes de departamento. Los representantes de los profesores ante el Consejo Universitario y el representante del ministerio de educación pueden pasar también a dedicación exclusiva una vez que se incorporen al organismo, haciendo una solicitud por escrito.

Un profesor a tiempo completo deberá permanecer en ambientes universitarios, treinta y cinco horas semanales, dedicado al cumplimiento de las tareas que le hayan sido encomendadas. La clasificación bajo esta dedicación obedece a exigencias académicas, docentes, administrativo-docentes, de investigación o de extensión, que resultan de los programas y planes aprobados por los organismos competentes. Así lo establecen los artículos 123 y 124 del Estatuto del Personal Docente y de Investigación de la Universidad de Los Andes. Dicho estatuto también establece, en su artículo 125, que la condición de profesor a tiempo completo no es compatible con otra actividad remunerada dentro de la Universidad. Por el contrario, fuera de ésta, el profesor clasificado en tal condición podrá desempeñar otra actividad remunerada, únicamente si cumple con los siguientes requisitos:

1. Que las labores inherentes a la misma estén estrechamente vinculadas a las disciplinas a que se dedica el profesor, contribuyendo a ampliar y profundizar su experiencia, y, por tanto, al mejoramiento de la docencia y de la investigación.
2. Que el profesor realice esa otra actividad fuera del horario que debe cumplir en la Universidad y no dedique a la misma, en total, más de diez horas a la semana.
3. Que obtenga autorización del Consejo Universitario, a solicitud del Consejo de Facultad o Núcleo respectivo.

Según lo establecido en los artículos 126 y 127 del estatuto mencionado previamente, de la misma manera que la dedicación exclusiva, el tiempo completo puede otorgarse para realizar labores de investigación, cumplir principalmente labores docentes, administrativas o de extensión, o varias de esas actividades. Aquellos profesores a tiempo completo que desempeñen funciones académico-administrativas, tales como Jefaturas de Departamento y otras similares, se les reconocerá, a los efectos del cumplimiento de su dedicación, el tiempo que empleen en esas funciones, el cual será establecido por el Consejo de Facultad o Núcleo respectivo.

Los profesores a medio tiempo dedicarán a la Universidad dieciocho horas semanales, lapso durante el cual deben permanecer en ambientes universitarios, dedicados al cumplimiento de sus tareas especificas, y accesibles a los alumnos para laborares de consulta y orientación. Estos profesores se designarán, fundamentalmente, en aquellas disciplinas en que sea necesario o útil el aporte de la experiencia derivada del ejercicio profesional, para impartir una enseñanza adecuada.

La solicitud de clasificación de los profesores en la condición de medio tiempo la hará la unidad académica a que estén adscritos, Departamentos, Institutos o Centros, ante el Consejo de Facultad o Núcleo. Bajo esta dedicación, la remuneración es equivalente a la mitad de la establecida para los profesores a tiempo completo de igual categoría en el escalafón y no se podrá desempeñar otras actividades, ni dentro ni fuera de la Universidad, que exijan dedicación a tiempo completo. De esta manera lo establecen los artículos 130, 131 y 132 del estatuto universitario.

Los artículos 133, 134 y 135 disponen que un profesor a tiempo convencional sólo está obligado a permanecer en ambientes universitarios las horas dedicadas al cumplimiento de las labores docentes que se le haya asignado y de las obligaciones subsecuentes. Su carga horaria no podrá exceder de doce 12 horas semanales. Sera designado para aquellas cátedras donde se requiera personal docente calificado por su amplia experiencia profesional, vinculada con la materia que se enseña en las mismas y recibirá una remuneración acorde con su carga horaria y categoría en el escalafón, conforme con la escala establecida al efecto por el Consejo Universitario.

CAPÍTULO III

MARCO METODOLÓGICO

Nivel, Tipo y Diseño de la Investigación

Teniendo en cuenta la clasificación del paradigma holístico de Hurtado (2010), el planteamiento del problema, la formulación del problema, y los objetivos de la investigación. El nivel de la misma es perceptual, dado que el verbo que define al objetivo general es estudiar.

Debido a que las técnicas y métodos en una investigación dependen tanto del tipo de investigación como del objetivo de la misma. Antes de hablar de las técnicas debe explicarse el tipo de investigación (Hurtado, 2010). La presente investigación es del tipo confirmatoria de verificación empírica, ya que los objetivos implican formulación de hipótesis y relación entre variables.

Por su parte, el diseño de la investigación es de fuente mixta, transeccional retrospectivo y multivariable. De fuente mixta, porque las fuentes objeto de estudio son vivas y documentales. Transeccional retrospectivo, puesto a que se estudia el evento de interés en un único momento y en tiempo pasado, octubre del año 2011. Y Multivariable, debido a que la amplitud de foco comprende el estudio de diversas variables asociadas al evento de interés, que es el retraso que presentan los profesores de las diversas escuelas que conforman al núcleo Mérida de la Universidad de Los Andes para ascender a través de los diferentes niveles del escalafón.

Población e Instrumentos de Recolección

Las unidades objeto de estudio están conformadas por todos aquellos miembros ordinarios activos pertenecientes al núcleo Mérida de Universidad de Los Andes, que para octubre del año 2011 tienen más de 15 años siendo miembros ordinarios y aún no han alcanzado la titularidad, y por aquellos que para esa fecha ya alcanzaron el máximo nivel del escalafón pero lo hicieron en un tiempo superior al establecido en el reglamento de la universidad. Por lo tanto, se excluyen del análisis a los profesores contratados, pensionados y jubilados. Los datos provienen de una fuente secundaria, fueron proporcionados por la Dirección de Asuntos Profesorales (DAP), anteriormente Oficina de Asuntos Profesorales (OAP) de la Universidad de Los Andes.

Técnicas y Procedimiento para el Análisis de Datos

Técnicas a usar

Para el cumplimiento de los dos primeros objetivos se emplean técnicas descriptivas univariantes y bivariantes, mientras que para desarrollar los objetivos restantes se utiliza la técnica conocida como modelos de regresión multinivel, la cual de describe a continuación.

Modelos de Regresión Multinivel

Aspectos Generales

Cuando la estructura de los datos en una población presente una jerarquía debe ser usado un modelo multinivel. Goldstein (1999), define una jerarquía como el conjunto de unidades agrupadas en diferentes niveles, donde las unidades del primer nivel podrían ser estudiantes, agrupados por escuelas o unidades del segundo nivel. Salud, educación, encuestas de hogares, estudios de crecimiento de animales, son algunas de las áreas de aplicación de estos modelos.

Los modelos multinivel son conocidos como modelos jerárquicos. Esto se debe, por un lado a la estructura de los datos, por ejemplo, estudiantes agrupados en escuelas, y por otro, a que el modelo tiene su propia jerarquía, es decir, los parámetros de las regresiones en las escuelas están en el nivel 1, controlados por los hiperparámetros del modelo, ubicados en el nivel 2. También se refiere a ellos como modelos de efectos aleatorios, en el sentido de que los coeficientes de regresión son considerados resultados aleatorios, de un proceso identificado con el modelo que ellos predicen (Gelman y Hill, 2007). Por otra parte, son conocidos como modelos de coeficientes aleatorios, porque permiten estimar coeficientes de regresión diferentes para cada unidad del segundo nivel. Estos modelos pueden ser vistos como generalizaciones de los modelos lineales y ser extendidos a modelos no lineales.

El término efectos aleatorios está relacionado con el momento de analizar los resultados, es decir, si las unidades del segundo nivel fueron seleccionadas aleatoriamente de una población, se pueden generalizar los resultados a toda esa población, de lo contrario, los resultados se limitarían únicamente a las unidades que fueron objeto de estudio.

Es importante tener en cuenta que los individuos pertenecientes a un mismo grupo, llámese escuela, hospital, familia, etc., es probable que en la mayoría de los casos tiendan a ser más parecidos entre ellos que comparados con otros grupos. Posiblemente estas diferencias tengan que ver con la ubicación de estos grupos y con el hecho de que formen parte de diferentes culturas o políticas, por ejemplo, de fomento de la salud, educación, entre otras. Entonces, la pertenencia a un grupo puede tener influencia sobre los individuos de dichos grupos, y a su vez cada individuo podría ejercer influencia sobre el grupo al que pertenece. Esto trae como consecuencia que los integrantes de un grupo puede que no sean independientes entre sí, rompiendo con el supuesto de independencia en el que se basan los modelos de regresión tradicionales y ameritando procedimientos diferentes de estimación.

El hecho de que los individuos que forman parte de un mismo grupo no sean independientes entre ellos, indica que la correlación intraclase presente alrededor de las variables medidas sobre los integrantes de un mismo grupo, será superior que la de aquellas variables medidas entre los integrantes de los diferentes grupos.

En este tipo de modelos se pueden definir variables en cualquier nivel de la jerarquía. Dichas variables se pueden mover de un nivel a otro por agregación o desagregación. Cuando las variables del primer nivel se mueven al segundo nivel, este proceso se conoce como agregación. Entonces, si en el nivel 1 o nivel de los estudiantes estuviera definida la variable puntuación promedio de cada estudiante, al asignarle a las unidades del nivel 2, que son las escuelas, la variable puntuación promedio en cada escuela, se estaría moviendo una variable del primer nivel al segundo nivel. Por el contrario, cuando se mueven variables del nivel 2 al nivel 1, esto se define como desagregación. Por lo tanto, si se le asigna a cada estudiante de una escuela, una variable que indique una característica de la escuela a la que pertenece, se estaría moviendo una variable del segundo nivel al primer nivel. Tal característica podría ser, si la escuela es pública o privada.

Según Hox (1995), para cada nivel de la jerarquía se pueden construir diferentes tipos de variables:

- Variables globales o absolutas: son aquellas que hacen referencia sólo al nivel en el que están definidas. El género, es una variable global del nivel 1 o nivel donde en la mayoría de los casos se encuentran ubicados los individuos. El tamaño de los grupos o unidades del nivel 2, por ejemplo, el tamaño de las escuelas, es una variable absoluta de este nivel. Se denominan variables globales cuando hacen referencia a las unidades del primer nivel y variables absolutas cuando se refieren a unidades del segundo nivel.

- Variables relacionales: pertenecen únicamente a un nivel, pero describen las relaciones que existen entre una unidad y las otras unidades de un mismo grupo. Son variables relacionales algunos índices sociométricos, que son aquellos que estudian las relaciones que existen entre dos valores sociométricos o, entre un valor y los demás valores de un grupo, como es el caso de los índices de popularidad o reciprocidad (Fernández, 2000).

- Variables analíticas: son medidas teniendo como referencia a las variables del primer nivel, tienen que ver con la manera en que se distribuye una variable global. Por ejemplo, si se tiene a la puntuación promedio como una variable definida en el nivel 1 y se mueve mediante agregación al nivel 2, la variable resultante sería una variable analítica.

- Variables estructurales: también son establecidas partiendo de variables del primer nivel, hacen referencia a la distribución de variables relacionales en el nivel 1, como es el caso de algunos índices de red social. Entonces, al mover por agregación una variable relacional al nivel 2, se crea una variable estructural.

- Variables contextuales: hacen referencia a las unidades del segundo nivel. Resultan de mover una variable del nivel 2 al nivel 1. Por ejemplo, cuando a los estudiantes o unidades del primer nivel, se les asignan variables de las escuelas o del segundo nivel, como el tamaño de la escuela, tipo de escuela o la puntuación promedio de la escuela.

Un problema a ser abordado con un modelo multinivel es aquel donde el interés se centra en la relación que existe entre variables que son medidas en los diferentes niveles de la jerarquía. Por ejemplo, cómo un grupo de variables medidas en los diferentes niveles, influye sobre una variable respuesta. El principal propósito de un análisis de este tipo, es precisar el efecto directo tanto a nivel individual como grupal, de las variables explicativas, y determinar sí algunas de las variables explicativas en el nivel grupo o segundo nivel, sirven para moderar la relación que existe a un nivel individual o primer nivel. Entonces, si las variables del segundo nivel moderan la relación que existe a un nivel menor, esto indica una interacción entre variables explicativas de los diferentes niveles. Es importante destacar que los modelos multinivel se enfocan únicamente en aquellos casos donde la variable dependiente pertenece al nivel 1 (Hox, 1995).

En Goldstein (1999), se plantea que un estudio que se llevó a cabo a finales de los años 70 en el campo de la educación, sobre niños de primaria, en Bennett (1976), sostiene que aquellos niños expuestos a un estilo de enseñanza formal de lectura, manifestaron más progresos que los niños que no fueron expuestos. Estos datos se analizaron empleando técnicas de regresión múltiple tradicional, donde la unidad de análisis fue los niños de forma individual, no se tomó en cuenta la forma como ellos estaban agrupados por clases y maestros. En este estudio los resultados fueron significativos. El primer ejemplo relevante en el área de las ciencias sociales empleando un modelo de regresión multinivel, fue el llevado a cabo por Aitkin et al (1981), quienes realizaron el mismo estudio pero teniendo en cuenta cómo los niños estaban agrupados por clase y maestros. Se obtuvo como resultado que las diferencias entre los niños que fueron expuestos al estilo formal de enseñanza de lectura y los que no fueron expuestos, no eran significativas.

Modelos Multinivel Versus Modelos de Regresión Tradicionales

Para que los resultados de un análisis utilizando un modelo de regresión lineal tradicional sean confiables, se debe cumplir con los supuestos básicos del modelo. Por lo tanto, las observaciones o residuos deben ser homoscedásticos, independientes entre sí y deben distribuirse normalmente. Algunas consecuencias del incumplimiento de los supuestos del modelo de regresión lineal se presentan a continuación.

- Sí las varianza de las observaciones o términos de error no son iguales para todos los valores que toma la variable independiente, se incumple con el supuesto de homocedasticidad. Según Gujarati (2004), los resultados de la violación de este supuesto son los siguientes. Los coeficientes de regresión estimados empleando Mínimos Cuadrados Ordinarios son insesgados y consistentes pero no eficientes, cuando hay heteroscedasticidad las varianzas de los estimadores no deben obtenerse mediante las formulas usuales de Mínimos Cuadrados Ordinarios. Si esto ocurre, las pruebas t y F pueden conducir a resultados equivocados, debido a que produce valores de t más pequeños de lo apropiado.

Entonces en presencia de heterocedasticidad, los estimadores resultantes tienen varianzas amplias, lo que disminuye la confianza en los parámetros. Cabe destacar, que aunque las muestras sean grandes, las varianzas muestrales de los estimadores no serán las correctas, los contrastes de hipótesis para evaluar la significación de los parámetros no tendrán validez si se asume una distribución normal. Debido a que la heterocedasticidad ocasiona intervalos de confianza amplios y sesgados, produce valores de t pequeños y aumenta la posibilidad de aceptar una hipótesis que es falsa, es decir, de cometer el error tipo II.

- Cuando hay autocorrelación entre las observaciones se incumple con el supuesto de independencia, Gujarati (2004) establece que los estimadores resultantes presentan las siguientes características. Al igual que en presencia de heteroscedasticidad, los estimadores son insesgados, consistentes, lineales, se distribuyen de forma normal y asintótica, pero no son eficientes. Cuando se tiene en cuenta la presencia de la autocorrelación y se utiliza el método de Mínimos Cuadrados Ordinarios, aunque se incremente el tamaño de muestra indefinidamente, los intervalos de confianza son más amplios. Sí se ignora la autocorrelación, las pruebas usuales t, F y χ^2, no pueden ser usadas legítimamente. En cuanto a las pruebas de significación de los parámetros, aumenta la posibilidad de cometer el error tipo II, la verdadera varianza es subestimada y el coeficiente de determinación resultante puede estar sobre estimado.

- El modelo estadístico es robusto cuando las desviaciones del supuesto de normalidad son moderadas, y si la muestra es grande, esto no tiene mayor importancia (Sánchez, citado en Quintero, 2007). Ante la falta de normalidad, los estimadores mínimos cuadráticos no son eficientes, los intervalos de confianza y las pruebas de significación de los parámetros del modelo no son exactos.

Aunque existen diversas estrategias y métodos para suavizar las consecuencias de la violación de los supuestos de homoscedasticidad e independencia, la detección y corrección de la heteroscedasticidad es un proceso complicado. Adicional a esto, Sayrs, citado en Gujarati (2004), afirma que la autocorrelación se puede detectar únicamente luego de que se ha corregido la heteroscedasticidad.

Teniendo en cuenta que, ante la presencia de autocorrelación en los términos de error, el uso de los modelos tradicionales de regresión resulta complicado. A continuación se presentan algunas ventajas de utilizar los modelos de regresión multinivel en lugar de los enfoques tradicionales.

- Las estimaciones de los coeficientes de regresión son estadísticamente eficientes (Goldstein, 1999). Es decir, presentan una varianza menor.

- Corrige las estimaciones de los errores estándar, intervalos de confianza y pruebas de significación (Goldstein, 1999). Debido a que los modelos estadísticos estándar se fundamentan en el supuesto de que las observaciones son independientes entre sí, cuando esta suposición no se cumple, esto produce estimaciones de los errores estándar errados, es decir, mucho más pequeños, y estos conducen a resultados falsamente significativos o estimaciones incorrectas. Los modelos multinivel no requieren que este supuesto se cumpla, por lo tanto, estos problemas no se presentan.

- Al permitir que las variables independientes en la parte fija del modelo, puedan ser medidas en cualquier nivel de la jerarquía, el investigador puede identificar posibles variables que expliquen las diferencias que puedan existir entre unidades de los niveles superiores (Goldstein, 1999). Esto es, si se están comparando calificaciones promedio en determinada materia, entre escuelas, y las pendientes entre las escuelas son significativamente diferentes, un enfoque multinivel permite identificar las variables que influyen sobre las diferencias que existen en la pendiente de esa variable, entre escuelas.

Modelo Multinivel Lineal Básico de Dos Niveles

Antes de desarrollar el modelo multinivel lineal básico de dos niveles, resulta conveniente plantear el modelo clásico de regresión lineal simple, el cual describe una relación lineal entre dos variables, que puede ser visto como un modelo de un sólo nivel, y está definido de la siguiente manera

$$y_i = \beta_0 + \beta_1 x_i + e_i \qquad (1)$$

En este modelo, y, es la variable dependiente, la cual es aleatoria. El subíndice i denota el valor que toma ésta variable para el i-ésimo individuo dentro de un grupo, el cual puede ser una población o una muestra. x_i, es la variable independiente para el individuo i. β_0, es el intercepto o valor que toma la variable dependiente cuando la variable independiente es cero. β_1, es la pendiente o coeficiente de regresión que acompaña a la variable independiente, representa el cambio en la variable dependiente a medida que la variable independiente aumenta en una unidad. Ambos, β_0 y β_1, son llamados los parámetros del modelo y son desconocidos. e_i, es el término de error aleatorio asociado con el i-ésimo valor que toma la variable dependiente, tiene media cero y varianza σ^2 .

Los coeficientes de regresión se representan con letras griegas y cuando son estimaciones muestrales tienen un acento circunflejo sobre ellos. En Gujarati (2004), se define al término de error, como una variable aleatoria o estocástica que puede representar todos aquellos factores que posiblemente influyen sobre la variable dependiente y que por diversas razones no son incluidos en el modelo de manera explícita.

El modelo multinivel lineal básico de dos niveles, también conocido como modelo de pendiente aleatoria, en el nivel 1, es el siguiente

$$y_{ij} = \beta_{0j} + \beta_{1j} x_{ij} + e_{ij} \tag{2}$$

donde,

$$\begin{aligned} &\beta_{0j} = \beta_0 + u_{0j} \quad \text{y} \quad \beta_{1j} = \beta_1 + u_{1j} \\ &e_{ij} \sim N\left(0, \sigma^2\right) \\ &\begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix} \;,\; \begin{pmatrix} \sigma_{u0}^2 & \sigma_{u01} \\ \sigma_{u01} & \sigma_{u1}^2 \end{pmatrix}\right] \end{aligned} \tag{3}$$

En las ecuaciones anteriores, $i \in \{1,\ldots n_j\}$ hace referencia a las unidades del primer nivel, y $j \in \{1,\ldots m\}$ hace referencia a las unidades del segundo nivel. y_{ij}, es el valor que toma la variable dependiente para *i*-ésimo individuo o unidad del nivel 1, agrupada dentro de la *j*-ésima unidad del nivel 2. β_{0j}, es el intercepto, o valor promedio de la variable dependiente para cada grupo o unidad del nivel 2. Como lo indica el subíndice *j*, el intercepto es aleatorio, es decir, va a ser diferente para cada una de las unidades del segundo nivel. Por lo tanto, está conformado por β_0, que es la media total de la variable dependiente para todos los grupos, y por u_{0j}, que es el efecto de los grupos sobre el intercepto. β_{1j}, es la pendiente o coeficiente de regresión que acompaña a la variable independiente *x*, para cada una de las unidades del nivel 2, indica el cambio promedio en *y*, a través de los grupos, a medida que la variable independiente aumenta en una unidad. Al igual que el intercepto, la pendiente también va a ser diferente para cada una de las unidades del segundo nivel. Razón por la cual está conformada por β_1, que es la pendiente promedio para todos los grupos, y por u_{1j}, que es el efecto de los grupos sobre la pendiente.

Por su parte, e_{ij}, es el término de error aleatorio asociado con el i-ésimo valor que toma la variable dependiente para el i-ésimo individuo o unidad del nivel 1, agrupada dentro de la j-ésima unidad del nivel 2. Hox (1995), señala que al igual que en regresión lineal múltiple, los errores aleatorios en cada grupo o escuela tiene media cero y varianza σ^2_j; también expone que en la mayoría de los modelos multinivel se asume que la varianza del error aleatorio es la misma para todos los grupos o unidades del segundo nivel, la cual se denota como σ^2 ; y por último, que los términos de error del segundo nivel que son u_{0j} y u_{1j}, se asumen a ser independientes de los e_{ij} al nivel individual.

Según Goldstein (1999), para que un modelo como el presentado en la ecuación (1) se convierta en un modelo multinivel, β_0 y β_1, deben convertirse en variables aleatorias (ver ecuación 3). En consecuencia, los términos de error del segundo nivel, u_{0j} y u_{1j}, son variables aleatorias con parámetros

$$\begin{aligned} &E(u_{0j}) = E(u_{1j}) = 0 \\ &\mathrm{var}(u_{0j}) = \sigma^2_{u0}, \quad \mathrm{var}(u_{1j}) = \sigma^2_{u1}, \quad \mathrm{cov}(u_{0j}, u_{1j}) = \sigma_{u01} \end{aligned} \tag{4}$$

Como puede observarse, la varianza del intercepto y la de la pendiente que acompaña a la variable independiente son respectivamente σ^2_{u0} y σ^2_{u1}. Cuando tales varianzas son significativamente grandes, esto indica que los interceptos y las pendientes son diferentes para cada una de las unidades del segundo nivel. Por su parte, σ_{u01} es la covarianza entre u_{0j} y u_{1j}, señala si la relación existente entre dos unidades de un mismo grupo es positiva o negativa.

Ahora bien, sustituyendo las ecuaciones del intercepto y de la pendiente que acompaña a la variable independiente en la ecuación (2), en el segundo nivel, el modelo multinivel lineal básico de 2 niveles toma la forma

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + u_{1j} x_{ij} + u_{0j} + e_{ij} \tag{5}$$

La parte fija del modelo está compuesta por los coeficientes $\beta_0 + \beta_1 x_{ij}$, y la parte aleatoria por $u_{1j}x_{ij} + u_{0j} + e_{ij}$. Como puede observarse, el término de error aleatorio u_{1j}, está conectado a x_{ij}, debido a que u_{1j} se multiplica por la variable independiente o explicativa x_{ij}. Por lo tanto, el error total resultante será diferente para cada uno de los valores que tome tal variable. La parte fija del modelo presentado en la ecuación (5), en forma matricial, se presenta a continuación

$$
\begin{aligned}
&E(Y) = X\beta \\
&con \;\; Y = \{y_{ij}\} \\
&E(y_{ij}) = X_{ij}\beta = (X\beta)_{ij}, \quad X = \{x_{ij}\}
\end{aligned}
$$

El símbolo $\{\ \}$ representa una matriz. X, es la matriz de diseño para la variable independiente. X_{ij}, es la *ij*-ésima fila de X. Por último, β, es el vector de parámetros (Goldstein, 1999).

El modelo presentado en la ecuación (2) tiene una sola variable independiente en el primer nivel, es fácil extender este modelo a más de una variable independiente. Cuando se introducen variables independientes en el nivel 1, el efecto sobre las varianzas de un modelo lineal básico de dos niveles es el siguiente: la varianza del segundo nivel puede, aumentar, disminuir o permanecer igual; mientras que la varianza del primer nivel y la variación residual total, pueden permanecer igual o disminuir.

El intercepto y la pendiente que acompaña a la variable independiente son aleatorios, es decir, diferentes para cada unidad del nivel 2, como lo indica el subíndice *j*. Entonces, cuando el intercepto y la pendiente de una variable independiente son significativamente diferentes para cada una de las unidades del segundo nivel, esto se conoce como variación del segundo nivel, múltiple o compleja, y se introducen variables independientes en este nivel tanto en la ecuación del intercepto como en la de la pendiente, con la intención de explicar o conocer a qué se deben éstas diferencias, de la siguiente manera

$$\beta_{0j} = \gamma_{00} + \gamma_{01} z_j + u_{0j} \quad \text{y} \quad \beta_{1j} = \gamma_{10} + \gamma_{11} z_j + u_{1j} \qquad (6)$$

La variable independiente *z*, que se introdujo en la ecuación del intercepto, podría explicar ahora el comportamiento de la variable dependiente *y*. Mientras que, la variable independiente *z*, que se encuentra en la ecuación de la pendiente, expresa que la relación que existe entre la variable dependiente y la variable independiente del nivel 1 depende de la variable independiente del nivel 2. γ_{00}, es el intercepto o media total de la variable dependiente. γ_{01}, es la pendiente que acompaña a la variable independiente del segundo nivel, que intenta explicar las diferencias entre los interceptos. u_{0j}, es el efecto de los grupos sobre el intercepto, condicionado a la variable independiente *z*. γ_{10}, es la pendiente promedio de todos los grupos. γ_{11}, es la pendiente que acompaña a la variable independiente del segundo nivel que se introdujo con la intención de explicar las diferencias entre las pendientes del primer nivel. Por último, u_{1j}, es el efecto de los grupos sobre la pendiente, condicionado a *z*.

Sustituyendo las ecuaciones del intercepto y de la pendiente en la ecuación (2), el modelo es el siguiente

$$y_{ij} = \gamma_{00} + \gamma_{10} x_{ij} + \gamma_{01} z_j + \gamma_{11} z_j x_{ij} + u_{1j} x_{ij} + u_{0j} + e_{ij} \qquad (7)$$

La expresión $z_j x_{ij}$ que aparece en la ecuación (7), puede ser interpretada como un término de interacción que aparece en el modelo como consecuencia de modelar la pendiente de la regresión, β_{1j}, que acompaña a la variable independiente del nivel de los individuos, con una variable independiente del nivel de los grupos que es z_j. Entonces, la variable independiente *z*, actúa como moderadora de la relación entre la variable dependiente *y*, y la variable independiente x_{ij}. Es decir, hay una interacción entre variables explicativas de ambos niveles (Hox, 1995).

Asumiendo que existen $P \in \{1,\dots P\}$ variables independientes o explicativas x en el nivel 1, y $Q \in \{1,\dots Q\}$ variables explicativas z en el segundo nivel. El modelo general toma la forma

$$y_{ij} = \gamma_{00} + \sum_{p=1}^{P} \gamma_{p0} x_{pij} + \sum_{q=1}^{Q} \gamma_{0q} z_{qj} + \sum_{p=1}^{P} \sum_{q=1}^{Q} \gamma_{pq} z_{qj} x_{pij} + \sum_{p=1}^{P} u_{pj} x_{pij} + u_{0j} + e_{ij} \quad (8)$$

Los e_{ij} son los errores del nivel 1, se asume que tienen distribución normal con media cero y varianza común. Los términos u_{0j} y u_{Pj} representan los términos de error correspondientes al nivel 2, los cuales se asumen como independientes de los e_{ij} y tienen distribución normal multivariante con vector de medias cero. Como se mencionó anteriormente, la varianza de u_{0j} viene a ser la varianza del intercepto entre los grupos, que se conoce como σ^2_{u0}. Las varianzas de los errores residuales u_{Pj} son las varianzas de las pendientes entre los grupos σ^2_{uP}. Las covarianzas entre los términos residuales son generalmente diferentes de cero y son colectadas en la matriz de varianza y covarianza del nivel 2 (Σ).

Los coeficientes de regresión γ no varían para cada una de las unidades del nivel 2, por lo tanto, es posible referirse a ellos como coeficientes fijos. Toda la variación entre las unidades del segundo nivel que quede en los coeficientes β después de la predicción de estos con la variable independiente de este nivel, que es (z_j), se asume que es la variación del error residual, y es capturada por el termino de error u_j, donde el subíndice j indica la unidad del nivel 2 a la que pertenece.

Cuando el intercepto y la pendiente son diferentes para cada unidad del segundo nivel, como se señaló previamente, se espera poder explicar parte de esta variación introduciendo variables explicativas en este nivel, o niveles superiores.

Sin embargo, existen casos en los que estos coeficientes no explican tal variación, y resulta que luego de introducir variables explicativas en el nivel 2, quedaría alguna variación aleatoria por explicar. Por otro lado, el nombre de modelo de coeficientes aleatorios para este tipo de modelos, se debe a que los coeficientes de regresión, es decir, el intercepto y la pendiente, se supone que tienen cierta cantidad de variación aleatoria entre las unidades del segundo nivel; y el nombre de modelo de componentes aleatorios, se refiere al problema estadístico de estimar el monto de dicha variación aleatoria (Hox, 1995).

Según Goldstein (1999), lo que hace diferente a la ecuación (5) del modelo de regresión lineal estándar y del análisis de varianza, es que tiene más de un término residual, lo que implica procedimientos especiales para obtener estimaciones satisfactorias de los parámetros del modelo. Entonces, es la parte aleatoria del modelo el factor clave.

A continuación se presenta el modelo de intercepto aleatorio, el cual es conocido también como modelo de componentes de varianza. Se utiliza únicamente cuando los grupos son significativamente diferentes en cuanto a su intercepto.

Modelo de Intercepto Aleatorio

Como su nombre lo indica, el intercepto va a ser diferente para cada una de las unidades del segundo nivel y la pendiente sería un componente fijo, es decir, la misma pendiente para cada una de las unidades del segundo nivel. En el nivel 1, el modelo de intercepto aleatorio o modelo de componentes de varianza, es el siguiente

$$y_{ij} = \beta_{0j} + \beta_1 x_{ij} + e_{ij} \tag{9}$$

En el nivel 2, el modelo de componentes de varianza se presenta a continuación

$$y_{ij} = \gamma_{00} + \gamma_{01}x_{ij} + u_{0j} + e_{ij} \tag{10}$$

La parte fija del modelo está conformada por $\gamma_{00} + \gamma_{01}x_{ij}$, que son, el intercepto o media total de la variable dependiente, y la pendiente o coeficiente de regresión que acompaña a la variable independiente, respectivamente. La parte aleatoria está dada por dos variables aleatorias que son los términos de error del segundo y primer nivel, respectivamente, $u_{0j} + e_{ij}$.

Con el propósito de explicar a qué se debe la diferencia entre los interceptos de las unidades del nivel 2, se introducen variables explicativas en este nivel en la ecuación del intercepto. Entonces, un modelo de intercepto aleatorio con una sola variable independiente en el segundo nivel, tiene la siguiente forma

$$y_{ij} = \gamma_{00} + \gamma_{01}z_j + \gamma_{10}x_{ij} + u_{0j} + e_{ij} \tag{11}$$

Como se mencionó previamente, la variable independiente del segundo nivel que es z, viene a explicar el comportamiento de la variable dependiente y. La parte fija del modelo está dada por los términos $\gamma_{00} + \gamma_{01}z_j + \gamma_{10}x_{ij}$, que son, la media total de la variable dependiente, la pendiente que acompaña a la variable independiente del segundo nivel y la que acompaña a la variable independiente del primer nivel, respectivamente. La parte aleatoria, como se mencionó anteriormente, está dada por las variables aleatorias $u_{0j} + e_{ij}$. Asumiendo que existen $Q \in \{1,\ldots Q\}$ variables explicativas z en el nivel 2 y $P \in \{1,\ldots P\}$ variables independientes o explicativas x, fijas en el nivel 1. El modelo general se presenta a continuación

$$y_{ij} = \gamma_{00} + \sum_{q=1}^{Q} \gamma_{0q} z_j + \sum_{p=1}^{P} \gamma_{p0} x_{pij} + u_{0j} + e_{ij} \tag{12}$$

Modelo Nulo

El modelo nulo, también conocido como modelo incondicional de medias, permite observar la variación de la variable respuesta alrededor de las unidades del nivel 2 y calcular la correlación intraclase. En el nivel 1, el modelo nulo, se plantea de la siguiente manera

$$y_{ij} = \beta_{0j} + e_{ij} \tag{13}$$

Como se mencionó previamente, *ij* denota el *i*-ésimo individuo o unidad del primer nivel agrupado en la *j*-ésima unidad del segundo nivel, y_{ij} es la variable dependiente, e_{ij} es el término de error aleatorio del nivel 1, y β_{0j} es el intercepto o valor promedio de la variable dependiente para cada una de la unidades del nivel 2. Ahora, en el segundo nivel, teniendo en cuenta que el intercepto es aleatorio ($\beta_{0j} = \gamma_{00} + u_{0j}$) y sustituyendo dicho intercepto en la ecuación (13), el modelo sería

$$y_{ij} = \gamma_{00} + u_{0j} + e_{ij}, \qquad u_{0j} \sim N\left(0, \sigma_{u0}^2\right) \quad e_{ij} \sim N\left(0, \sigma^2\right) \tag{14}$$

En la ecuación anterior, γ_{00}, es el intercepto o media total de la variable dependiente, u_{0j}, es el término de error aleatorio asociado con las unidades del nivel 2, el cual es independiente de e_{ij}. El modelo nulo dado en (14) puede obtenerse también eliminando de las ecuaciones (7), (8) o (11) todos los términos que contengan *x* o *z*.

El modelo incondicional de medias puede ser visto como un modelo de análisis de varianza de una sola vía con efectos aleatorios (Singer, 1998). Esto, en cuanto a la estructura del modelo, es decir, dicho modelo está formado, como se muestra en la siguiente ecuación, a la izquierda, por la variable dependiente, a la derecha, por la media general de la variable respuesta, el efecto del tratamiento o factor de interés para el *i*-ésimo nivel, aplicado sobre n unidades experimentales, cuyas cantidades son variables aleatorias independientes e idénticamente distribuidas con esperanza cero y varianza σ_a^2, y por el término de error aleatorio.

$$y_{ij} = \mu + a_i + e_{ij} \tag{15}$$

donde $a_i = \mu_i - \mu$, i=1,2…,k y j=1,2,…,n

Entonces, y_{ij}, es la respuesta de la *j*-ésima observación de la unidad experimental n, sobre la cual se le aplicó el tratamiento o nivel del factor *i*. a_i y e_{ij}, son variables aleatorias que se distribuyen independientemente. Las observaciones se distribuyen de forma normal con media μ y var($y_{ij} = \mu + a_i + e_{ij}$)=$\sigma_a^2 + \sigma^2$. La varianza de la variable respuesta es el resultado de la suma de dos varianzas, que son, la varianza del factor, para todo *i*, y la varianza del término de error aleatorio.

Puede notarse que el modelo de análisis de varianza de una sola vía de efectos aleatorios, difiere del modelo incondicional de medias, en cuanto a que, el segundo elemento a la derecha de la igualdad del modelo nulo presentado en la ecuación (14), es el término de error correspondiente a las unidades del segundo nivel, mientras que en el modelo de análisis de varianza, mostrado en la ecuación (15), representa el término de error asociado con los niveles del factor tratamiento. El termino efectos aleatorios en el modelo nulo, tiene que ver con la forma como fueron seleccionadas las unidades del segundo nivel, mientras que en el modelo de análisis de varianza tiene que ver con la forma en que el investigador seleccionó los niveles del tratamiento, es decir, si la selección de los mismos fue de forma aleatoria o arbitraria.

El modelo incondicional de medias permite calcular la correlación intraclase o correlación intra unidades del segundo nivel, es decir, entre dos unidades del primer nivel, agrupadas dentro de la misma unidad del segundo nivel, la cual puede interpretarse como la proporción de la varianza total proveniente de las unidades del nivel 2. Como se mencionó anteriormente, los modelos multinivel pueden ser utilizados cuando las observaciones no son independientes, siendo la correlación intraclase la manera de medir la magnitud de esa dependencia. Dicha correlación es igual a la proporción estimada de la varianza entre grupos comparada con la varianza total estimada, como se muestra a continuación

$$\rho = \frac{\sigma_{u0}^2}{\left(\sigma_{u0}^2 + \sigma^2\right)} \tag{16}$$

Según Goldstein (1999), una correlación intra unidad, es decir, dentro una unidad del segundo nivel o grupo diferente de cero, se debe a que en el modelo existe más de un término residual. Esto ocasiona que los procedimientos tradicionales de estimación, como es el caso del método de Mínimos Cuadrados Ordinarios (Ordinary Least Squares), utilizados en modelos de regresión múltiple tradicionales, no deban utilizarse. Con la intención de explicar a qué se debe la diferencia entre los interceptos se introducen variables explicativas en el segundo nivel. En la ecuación (17) se presenta el modelo nulo con una sola variable independiente del nivel 2.

$$y_{ij} = \gamma_{00} + \gamma_{01} z_j + u_{0j} + e_{ij} \tag{17}$$

Ahora bien, si se introduce más de una variable independiente, el modelo tomaría la forma

$$y_{ij} = \gamma_{00} + \sum_{q=1}^{Q} \gamma_{0q} z_j + u_{0j} + e_{ij} \tag{18}$$

El modelo incondicional de medias se puede considerar como el primer paso o procedimiento exploratorio, para llevar a cabo un análisis multinivel, con él se puede determinar si los datos requieren un análisis de este tipo o si pueden ser analizados mediante modelos de regresión tradicionales. Este modelo proporciona una base con la que se pueden comparar modelos más complejos. Entonces, si resulta que la variación entre las unidades del nivel 2 es significativa, es decir, que el porcentaje de varianza debido a las unidades del segundo nivel es significativo, se puede proceder a ajustar un modelo por niveles.

Estimación de Parámetros

Hox (1995), señala que en el análisis de regresión multinivel, los estimadores frecuentemente usados son los conocidos como estimadores máximo verosímiles, los cuales se obtienen mediante el método de Máxima verosimilitud[1]. Estos estimadores calculan los parámetros de un modelo, proporcionando estimaciones para los valores poblacionales que maximicen la función de verosimilitud, la cual proporciona la probabilidad de observar los datos muestrales, es decir, los valores de *y*, dadas las estimaciones actuales de los parámetros.

Según Goldstein (1999), el procedimiento de máxima verosimilitud, debido a que no toma en cuenta la variación muestral de los parámetros fijos, produce estimaciones sesgadas de los parámetros aleatorios. Lo que puede ser importante cuando el tamaño de la muestra es pequeño. La solución es utilizar una modificación conocida como Máxima Verosimilitud Restringida (Maximum

[1] El Método de Máxima Verosimilitud (Maximum Likelihood) consiste en estimar los parámetros del modelo de manera tal, que la probabilidad de observar los valores dados de la variable dependiente (*y*), sea lo más grande posible.

Likelihood Restricted), mediante la cual se obtienen estimaciones insesgadas de los parámetros aleatorios.

Otros procedimientos de estimación para modelos de regresión multinivel se señalan en Goldstein (1999), y son los siguientes: el método de Mínimos Cuadrados Generalizados Esperados (Expected Generalized Least Squares), el Algoritmo de Puntajes de Fisher (Fisher Scoring), el método de Estimación completamente Bayesiano (The Full Bayes Estimation), el método Empírico de Bayes (Empirical Bayes Estimation) y el método de Cadenas de Markov Monte Carlo (Markov Chain Monte Carlo).

El método de Mínimos Cuadrados Esperados, el Algoritmo de Puntajes de Fisher, y el método de Máxima Verosimilitud Restringida, están basados en verisimilitud. Mientras que el método de Estimación Completamente Bayesiano, el método Empírico de Bayes, y el método de Cadenas de Markov Monte Carlo, se basan en estadística Bayesiana.

Se requiere un procedimiento iterativo para obtener los valores de los parámetros mediante el método de máxima verosimilitud. El programa genera estimaciones de partida para los diferentes parámetros, en el análisis multinivel estos se basan generalmente en la regresión tradicional. Entonces, mediante un procedimiento de computación ingenioso se mejoran los valores de partida, esto se realiza una serie de veces. Luego de cada iteración, el programa reconoce cuantas estimaciones cambiaron comparadas con las anteriores, si el cambio es muy pequeño, el programa determina que el procedimiento ha convergido y que ha terminado. Cabe destacar, que este procedimiento no garantiza el detenerse, hay modelos y grupos de datos para los que el programa produce un sinfín de iteraciones. Esto se debe a que muchos programas no tienen un límite máximo. La interpretación usual para un modelo que no converge, es que no es un buen modelo, aunque el problema también puede deberse a los datos (Hox, 1995).

Teniendo en cuenta la preferencia de Goldstein (1999) por el método de Máxima Verosimilitud Restringida, sobre el método de Máxima Verosimilitud, cuando la muestra es pequeña; las evidencias obtenidas en los resultados de los estudios llevados a cabo por Swallow y Monahan (1984), mencionados en la Guía de Usuario del SAS en su Versión 9.1, quienes basándose en simulaciones, favorecen al método de Máxima Verosimilitud y el Método de Máxima Verosimilitud Restringida; y teniendo especialmente en cuenta que el método de Máxima Verosimilitud no considera la variación muestral de los parámetros fijos, es decir, la estructura jerárquica, lo que es importante cuando la muestra es pequeña, se favorece el uso del método de Máxima Verosimilitud Restringida, que es, el que por defecto emplean en la estimación de los parámetros, software como SAS y R.

Ahora bien, para cualquier modelo multinivel es necesario conocer los parámetros, que son las varianzas y covarianzas, lo coeficientes fijos y los residuos. Estos se presentan en las secciones siguientes.

Parámetros aleatorios

El modelo de intercepto aleatorio, también conocido como modelo de componentes de varianza, presentado en las ecuaciones (9) y (10), contiene dos parámetros aleatorios a ser estimados, que son σ_{u0}^2 y σ^2, los cuales conforman la varianza de la variable respuesta, como se muestra a continuación

$$\operatorname{var}\left(y_{ij} / \beta_0, \beta_1, x_{ij}\right) = \operatorname{var}\left(u_{0j} + e_{ij}\right) = \sigma_{u0}^2 + \sigma^2 \tag{19}$$

Como puede observarse en la ecuación anterior, la varianza de la variable dependiente para el modelo de componentes de varianza, está compuesta por la suma de las varianzas del nivel 2 y del nivel 1. La varianza total para cada unidad del primer nivel es constante, y la covarianza entre dos unidades de dicho nivel pertenecientes al mismo grupo, denotada por i_1, i_2, se representa de la siguiente manera

$$\operatorname{cov}\left(u_{0j} + e_{i_1 j}; u_{0j} + e_{i_2 j}\right) = \operatorname{cov}\left(u_{0j}, u_{0j}\right) = \sigma_{u0}^2 \tag{20}$$

Una matriz de covarianza de un modelo de intercepto aleatorio, para tres unidades del nivel 1, que podrían ser tres estudiantes de una escuela o tres pacientes de un hospital, agrupados en la misma unidad del nivel 2, se deriva de la expresión anterior, y se muestra a continuación

$$\begin{pmatrix} \sigma_{u0}^2+\sigma^2 & \sigma_{u0}^2 & \sigma_{u0}^2 \\ \sigma_{u0}^2 & \sigma_{u0}^2+\sigma^2 & \sigma_{u0}^2 \\ \sigma_{u0}^2 & \sigma_{u0}^2 & \sigma_{u0}^2+\sigma^2 \end{pmatrix}$$

Figura 1.
Matriz de covarianza para tres unidades del nivel 1 agrupadas dentro de la misma unidad del nivel 2, para un modelo de componentes de varianza. Fuente: Goldstein, 1999.

Ahora, una matriz de covarianza total, para dos unidades del segundo nivel, digamos *A* y *B*, donde *A* tiene tres unidades del nivel 1 y *B* sólo dos unidades, puede observarse en la figura (2). Su estructura de bloque diagonal refleja, que la diagonal entre las covarianzas de las unidades del segundo nivel, es cero. Esto puede extenderse a cualquier número de unidades del nivel 2.

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

Donde

$$A=\begin{pmatrix} \sigma_{u0}^2+\sigma^2 & \sigma_{u0}^2 & \sigma_{u0}^2 \\ \sigma_{u0}^2 & \sigma_{u0}^2+\sigma^2 & \sigma_{u0}^2 \\ \sigma_{u0}^2 & \sigma_{u0}^2 & \sigma_{u0}^2+\sigma^2 \end{pmatrix}, \qquad B=\begin{pmatrix} \sigma_{u0}^2+\sigma^2 & \sigma_{u0}^2 \\ \sigma_{u0}^2 & \sigma_{u0}^2+\sigma^2 \end{pmatrix}$$

Figura 2.

Matriz de covarianza bloque diagonal del vector respuesta (*Y*), para un modelo de componentes de varianza de dos niveles con dos unidades del nivel 2. Fuente: Goldstein, 1999.

En la siguiente figura, se presenta la matriz que se encuentra en la figura (2), de forma más compacta

$$V_2 = \begin{bmatrix} \sigma_{u0}^2 J_{(3)} + \sigma^2 I_{(3)} & 0 \\ 0 & \sigma_{u0}^2 J_{(2)} + \sigma^2 I_{(2)} \end{bmatrix}$$

Figura 3.
Notación general de la matriz de covarianza bloque diagonal. Fuente: Goldstein, 1999.

En la figura anterior. $I_{(n)}$ es la matriz identidad nxn, formada por unos en su diagonal y tiene la propiedad de ser el elemento neutro, $J_{(n)}$ es la matriz de unos nxn, el subíndice 2, en $V_{(2)}$ indica que es un modelo de dos niveles. Para una estructura de datos de un sólo nivel, σ_{u0}^2 de los modelos (MCO) sería cero y la matriz de covarianza volvería a ser de la forma estándar $\sigma^2 I$, donde la única varianza residual estaría dada por σ^2.

Para el modelo básico de dos niveles, donde la pendiente o coeficiente de la variable explicativa *X* es aleatorio en el nivel 2, corresponde la siguiente estructura típica de bloques, para un bloque de dos niveles con dos unidades del nivel 1. Entonces, en la figura (4) se puede observar la matriz de covarianza de la variable respuesta para un modelo de dos niveles donde el intercepto y el coeficiente de regresión son aleatorios.

La matriz de covarianza para los coeficientes aleatorios del primer nivel es Ω_1, está formada por un sólo término; la matriz de covarianza del intercepto y de la pendiente, ambos aleatorios en el segundo nivel, es Ω_2. Para el grupo de esas matrices de covarianza también puede escribirse $\Omega = \{\Omega_i\}$

$$\begin{pmatrix} A & B \\ B & C \end{pmatrix}$$

$$A = \left(\sigma_{u0}^2 + 2\sigma_{u01}x_{1j} + \sigma_{u1}^2 x_{1j}^2 + \sigma^2\right)$$
$$B = \left(\sigma_{u0}^2 + \sigma_{u01}\left(x_{1j} + x_{2j}\right) + \sigma_{u1}^2 x_{1j}^2 x_{2j}\right)$$
$$C = \left(\sigma_{u0}^2 + 2\sigma_{u01}x_{2j} + \sigma_{u1}^2 x_{2j}^2 + \sigma^2\right)$$

Donde,

$$\begin{pmatrix} A & B \\ B & C \end{pmatrix} = X_j \Omega_2 X_j^T + \begin{pmatrix} \Omega_1 & \\ & \Omega_1 \end{pmatrix}$$

$$X_j = \begin{pmatrix} 1 & x_{1j} \\ 1 & x_{2j} \end{pmatrix}, \qquad \Omega_2 = \begin{pmatrix} \sigma_{u0}^2 & \sigma_{u01} \\ \sigma_{u01} & \sigma_{u1}^2 \end{pmatrix}, \quad \Omega_1 = \sigma^2$$

Figura 4.
Matriz de covarianza de la respuesta para una unidad del nivel 2 con dos unidades del nivel 1, para un modelo de dos niveles con un intercepto aleatorio y un coeficiente de regresión aleatorio en el nivel 2. Fuente: Goldstein, 1999.

Coeficientes fijos

Teniendo en cuenta el modelo de componentes de varianza presentado en la ecuación (10) y suponiendo que los valores de las varianzas son conocidos, se aplica el procedimiento de Mínimos Cuadrados Generalizados, mostrado en (21), para obtener el estimador de los coeficientes fijos.

$$\hat{\beta} = \left(X^T V^{-1} X\right)^{-1} X^T V^{-1} Y \tag{21}$$

con matriz de covarianza $\left(X^T V^{-1} X\right)^{-1}$, y donde

$$X = \begin{pmatrix} 1 & x_{11} \\ 1 & x_{21} \\ & \vdots \\ 1 & x_{n_m m} \end{pmatrix} \qquad Y = \begin{pmatrix} y_{11} \\ y_{21} \\ \vdots \\ y_{n_m m} \end{pmatrix} \tag{22}$$

Residuos

La estimación del término residual e_i de un modelo de un sólo nivel, tal como el que se muestra en la ecuación (1), es $\tilde{y}_i$, que se denomina residuo bruto. En un modelo multinivel lineal de dos niveles existen varios residuos que son, el término residual del nivel 1 (e_{ij}) y los términos residuales del nivel 2 (u_{0j} y u_{pj}), asociados con el intercepto y con las pendientes o coeficientes de regresión, respectivamente.

Cuando los residuos se distribuyen de forma normal, la ecuación (21) produce estimaciones de Máxima Verosimilitud (Goldstein, 1999). Del procedimiento de estimación iterativo, donde las estimaciones razonables de los parámetros fijos, normalmente son ajustadas por Mínimos Cuadrados Ordinarios asumiendo que la varianza del intercepto es cero, se forman los residuos brutos, como se observa a continuación

$$\tilde{y}_{ij} = y_{ij} - \hat{\beta}_0 - \hat{\beta}_1 x_{ij} \tag{23}$$

donde, el vector de residuos brutos es

$$\tilde{Y} = \{\tilde{y}_{ij}\}$$

Ahora, una vez conocidos los parámetros, se puede predecir un residuo específico, esto se debe a que para cada vector residual, los residuos en cualquier nivel son independientes de los residuos en cualquier otro nivel. Por ejemplo u_{0j}

para un modelo de componentes de varianza, donde se requiere por cada nivel dos unidades, sería

$$\hat{u}_{0j} = E\left(u_{0j} / Y, \hat{\beta}, \hat{\Omega}\right) \tag{24}$$

Según Goldstein (1999), la expresión (24) viene a ser el cálculo o predicción de los residuos. Es como una regresión lineal de u_{0j} sobre el grupo de $\{\tilde{y}_{ij}\}$ para la j-ésima unidad del segundo nivel. En terminología bayesiana esta expresión se conoce como estimación residual posterior. Ahora, cuando no se tiene en cuenta la variación muestral asociada a las estimaciones de los parámetros en la expresión (24), se plantea lo siguiente

$$\begin{aligned} &\operatorname{cov}\left(\tilde{y}_{ij}, u_{0j}\right) = \operatorname{var}\left(u_{0j}\right) = \sigma_{u0}^2 \\ &\operatorname{cov}\left(\tilde{y}_{ij}, e_{ij}\right) = \sigma^2 \\ &\operatorname{var}\left(\tilde{y}_{ij}\right) = \sigma_{u0}^2 + \sigma^2 \end{aligned} \tag{25}$$

La cantidad requerida para estimar los coeficientes de regresión y por tanto $\hat{u}_{0j}$, está dada en (25). Detalles sobre la estimación de los residuos del segundo nivel asociados con las pendientes, u_{pj}, pueden encontrarse en el apéndice 2.2 de (Goldstein, 1999). Para el modelo de componentes de varianza

$$\hat{u}_{0j} = \frac{n_j \sigma_{u0}^2}{\left(n_j \sigma_{u0}^2 + \sigma_{e0}^2\right)} \tilde{y}_j$$

$$\tilde{e}_{ij} = \tilde{y}_{ij} - \hat{u}_{0j} \qquad (26)$$

$$\tilde{y}_j = \frac{\sum_i \tilde{y}_{ij}}{n_j}$$

donde n_j es el número de unidades del primer nivel que hay en la j-ésima unidad del segundo nivel. Los residuos estimados son consistentes pero no son incondicionalmente insesgados. El factor multiplicador de la media $\tilde{y}_j$ de los residuos brutos para la unidad j, también es conocido como factor de reducción debido a que toma valores menores o iguales a 1. Entonces, cuando n_j aumenta, dicho factor de reducción tiende a uno y cuando n_j disminuye, tiende a cero.

Estos residuos se pueden interpretar de dos maneras. Por un lado su interpretación básica como variables aleatorias, con su propia distribución, donde los valores de los parámetros nos dan información sobre la variación existente entre unidades del nivel 2, también proporcionan estimadores eficientes para los coeficientes fijos de la regresión. Y por otro lado, es una estimación individual para cada una de las unidades del segundo nivel, partiendo del supuesto de que pertenecen a una población de unidades para predecir sus valores. Para las unidades del nivel 2 que tienen pocas unidades del nivel 1, se obtienen estimaciones más precisas, partiendo de que vienen de una población, que cuando se ignora la suposición de que son una muestra de una población, y se utilizara la información sólo de esas unidades (Goldstein, 1999).

Bondad de Ajuste

Cuando se ajustan diferentes modelos sobre un conjunto de datos, surge la necesidad de comparar los ajustes y así poder seleccionar un modelo. Entonces, a la hora de comparar la bondad de ajuste de varios modelos con los mismos efectos fijos pero diferentes efectos aleatorios los criterios más utilizados son el criterio de información de Akaike (*AIC*) y el criterio de Schwarz (*BIC*), (Singer, 1998). Los cuales están definidos en la Guía de Usuario del SAS 9.1 de la siguiente manera

$$AIC=-2l+2d$$

$$BIC=-2l+d\ ln(n)$$

donde, l denota el máximo valor del logaritmo de verosimilitud posiblemente restringido, d es la dimensión del modelo, q es el número de parámetros de covarianza estimados, y n es el número de observaciones. Entonces, se considera un mejor aquel modelo arroje el menor valor de *AIC* y *BIC*.

Balzarini, Macchiavelli y Casanoves (2005) plantean, que la dimensión del modelo, d, se compone de la suma del número de parámetros de covarianza estimados y el rango de la matriz de diseño, y que para comparar dos modelos que están anidados, por ejemplo, con igual estructura de medial pero diferente estructura de covarianza, o con igual covarianza pero diferente estructura de medias, puede utilizarse la prueba del cociente de verisimilitud, la cual está basada en la relación

$$-2\ln(\lambda) = -2\ln\left(\frac{l\left(\hat{\theta}, reducido\right)}{l\left(\hat{\theta}, completo\right)}\right)$$

La cantidad resultante, teniendo en cuenta que el logaritmo de un cociente, es la diferencia entre el logaritmo del numerador y el logaritmo del denominador, se contrasta con una chi-cuadrado. Los grados de libertad están dados por la diferencia entre el número de parámetros que tiene cada modelo. Esta prueba del cociente de verosimilitud es análoga a la que se conoce como prueba deviance y se presenta a continuación. Los procedimientos de máxima verosimilitud también producen el estadístico llamado deviance, el cual indica qué tan bien el modelo se ajusta a los datos. El modelo que arroje un valor de la deviance menor, ajusta mejor a los datos.

En esta prueba se contrasta una hipótesis nula que establece que el modelo M_0 (con m_0 parámetros) ajusta mejor a los datos, contra una hipótesis alternativa que establece que M_1 (con m_1 parámetros) ajusta mejor a los datos. Como el método de máxima verisimilitud (ML) provee la verosimilitud de los parámetros, es posible observar la deviance para cada modelo, que son

$$D_0=-2log(verosimilitud\ de\ M_0\ )$$

$$D_1=-2log(verosimilitud\ de\ M_1\ )$$

Entonces, si el modelo M_0 está anidado en el modelo M_1, es decir, M_1 incluye a M_0 . Se puede demostrar que si D_0-D_1> $\chi^2_{m_1-m_0;\alpha}$ el modelo M_1 se ajusta mejor a los datos.

Esta prueba puede realizarse cuando las estimaciones de los parámetros se hacen mediante el uso de estimadores de máxima verosimilitud y estimadores de máxima verosimilitud restringida. Sólo es recomendable el uso de estimadores de máxima verosimilitud restringida cuando los modelos a comparar tienen igual media y diferente estructura de covarianza (Balzarini, Macchiavelli y Casanoves, 2006).

La diferencia de las desviaciones para estos dos modelos tiene una distribución chi-cuadrado, los grados de libertad se obtienen de la diferencia entre el número de parámetros estimados en los dos modelos. Esto puede usarse para probar, mediante una prueba chi-cuadrado, si el modelo general ajusta mejor que el modelo simple. La prueba chi-cuadrado de las desviaciones puede utilizarse también para explorar la importancia de los efectos aleatorios, comparando un modelo con efectos aleatorios y un modelo sin efectos aleatorios. Cuando los modelos a comparar no están anidados, el principio de que los modelos deben ser tan simples como sea posible, indica que debe escogerse el modelo más sencillo.

Las salidas o resultados al ajustar un modelo variarán dependiendo Software que se emplee. Por ejemplo, cuando se usa R, en la salida se encontrará el estadístico conocido como deviance o desviación. Ahora, si se utiliza alguna versión del SAS, en los resultados se observará el estadístico llamado logaritmo de la verosimilitud o Log Likelihood.

Pruebas de Significación

Para la mayoría de las estimaciones, los procedimientos de máxima verosimilitud, producen los errores estándar respectivos. Estos pueden utilizarse en pruebas de significación, para contrastar hipótesis sobre los parámetros del modelo. La prueba de Wald, donde el estadístico Z asociado con la distribución normal estándar, se obtiene dividiendo el valor del parámetro entre su error estándar, establece en la hipótesis nula que determinado parámetro es igual a cero. Los errores estándar son asintóticos, es decir, validos para muestras grandes[2]. En el análisis de regresión tradicional se requieren diez observaciones por cada coeficiente de regresión. En regresión multinivel, los coeficientes del nivel superior y los componentes de varianza, son estimados sobre una muestra de los grupos, que por lo general no es muy grande (Hox, 1995).

Bryk and Raudenbush, citado en Goldstein (1999), argumentan que para los efectos fijos es mejor utilizar una prueba basada en la distribución *t*, que la prueba de Wald no es apropiada para las varianzas porque la distribución muestral de las varianzas es sesgada, y proponen utilizar una prueba basada en la distribución chi-cuadrado para los residuos.

[2] Cuando la muestra es pequeña la prueba de Wald es poco confiable.

Para los componentes de varianza, que serían la varianza del primer nivel, y las varianzas del segundo nivel. En cuanto a la significación, se utiliza la prueba *Z* de Wald, donde la hipótesis nula establece que las varianzas no son significativas y la alternativa que las varianzas del modelo son significativas. Como prueba adicional para los componentes de varianza se puede utilizar la prueba de la razón de la verosimilitud.

En cuanto a la independencia de la varianza del nivel 2, se contrasta con una chi-cuadrado, los grados de libertad resultan de restarle uno a la cantidad de unidades que hay en el segundo nivel, la hipótesis nula establece la independencia de la varianza del nivel 2, y la alternativa, la no independencia. Para los parámetros fijos se utiliza una prueba *t*, la hipótesis nula establece que los parámetros son significativos versus que los parámetros no son significativos, los grados de libertad están dados por las unidades del segundo nivel.

Procedimiento general

Análisis Descriptivo

Luego de identificar aquellos profesores de la Universidad de Los Andes que se encuentren en situación de retraso para octubre del año 2011, se estima la demora que presenta cada uno de ellos. Esta demora en avanzar a través los diferentes niveles del escalafón es la variable objeto de estudio o variable dependiente.

Para describir el comportamiento por escuela del retraso que presentan los profesores para ascender en el escalafón, haciendo uso de la estadística descriptiva, se procede a estimar y graficar por un lado el tiempo promedio del retraso por escuelas, y por otro la proporción de profesores que están en situación de retraso, con respecto al total de profesores pertenecientes a cada escuela. Luego se estudia la correlación entre la variable objeto de estudio y las variables de los diferentes niveles de la jerarquía.

Modelos Multinivel

Cuando no se cuenta con argumentos teóricos, o procedentes de otras investigaciones se puede utilizar un procedimiento exploratorio para definir un modelo. Hox (1995), plantea como alternativa para seleccionar el modelo adecuado para un análisis, utilizar un procedimiento exploratorio. Comenzando con un modelo que contenga sólo el intercepto e ir agregando varios tipos de parámetros paso a paso. En cada paso debe examinarse el resultado para ver qué parámetros son significativos y ver cuánto error corresponde a cada nivel. Los pasos a seguir para llevar a cabo tal procedimiento exploratorio son los siguientes.

Paso 1. Analizar el modelo sin variables explicativas o modelo que contiene sólo el intercepto, este modelo además de ser usado para estimar la correlación intraclase, da un valor de la desviación (deviance) o de la verosimilitud, que es una medida del grado de ajuste del modelo.

Paso 2. Analizar un modelo con todas las variables explicativas fijas en el nivel 1. Lo que significa que los componentes de varianza correspondientes a las pendientes son cero. El modelo es el siguiente

$$y_{ij} = \gamma_{00} + \sum_{p=1}^{P} \gamma_{p0} x_{pij} + u_{0j} + e_{ij} \tag{1}$$

Aquí se evalúa la contribución de cada una de las variables independientes, se puede comprobar si el modelo mejoró con la inclusión de las variables, mediante la diferencia de la verosimilitud de ambos modelos, esta diferencia se aproxima a una chi-cuadrado, donde los grados de libertad están dados por la diferencia entre el número de parámetros de ambos modelos, que viene a ser el número de variables explicativas que se definieron en el paso anterior.

Paso 3. Evaluar si la pendiente de cualquiera de las variables independientes tiene un componente de varianza significativo entre los grupos. El modelo se describe como

$$y_{ij} = \gamma_{00} + \sum_{p=1}^{P} \gamma_{p0} x_{pij} + \sum_{p=1}^{P} u_{pj} x_{pij} + u_{0j} + e_{ij} \qquad (2)$$

La prueba de variación de las pendientes aleatorias debe hacerse una por una. Variables que se omitieron en el paso dos pueden analizarse de nuevo en este paso, puesto que la media del coeficiente de regresión para una variable explicativa puede que no sea significativo, pero tenga un componente de varianza que sí lo es. Luego de decidir cuales coeficientes de regresión tienen componentes de varianza significativos entre grupos se puede formar un modelo con todos esos componentes y usar la prueba chi-cuadrado basada en las desviaciones para probar si el modelo final del paso tres es mejor que el modelo propuesto en el paso dos.

Paso 4. Agregar variables explicativas en el segundo nivel como se muestra en la siguiente ecuación

$$y_{ij} = \gamma_{00} + \sum_{p=1}^{P} \gamma_{p0} x_{pij} + \sum_{q=1}^{Q} \gamma_{0q} z_{qj} + \sum_{p=1}^{P} u_{pj} x_{pij} + u_{0j} + e_{ij} \qquad (3)$$

Este modelo permite observar si esas variables explican la variación entre grupos presente en la variable dependiente.

Paso 5. Agregar interacciones entre niveles, es decir, interacción entre variables explicativas del nivel grupo y variables explicativas del nivel individual, que tuvieron significante variación en la pendiente en el paso tres. La ecuación es la siguiente

$$y_{ij} = \gamma_{00} + \sum_{p=1}^{P} \gamma_{p0} x_{pij} + \sum_{q=1}^{Q} \gamma_{0q} z_{qj} + \sum_{p=1}^{P} \sum_{q=1}^{Q} \gamma_{pq} z_{qj} x_{pij} + \sum_{p=1}^{P} u_{pj} x_{pij} + u_{0j} + e_{ij} \qquad (4)$$

En cada paso se debe decidir qué coeficientes de regresión, varianza o covarianza mantener, basándose en pruebas de significación, en el cambio en la verosimilitud y en el cambio que presentan los componentes de varianza. Cuando se introducen variables explicativas en el primer nivel se espera que disminuya la varianza de ese nivel. Si la composición de los grupos con respecto a las variables explicativas es diferente para todos los grupos, esperamos que disminuya la varianza del segundo nivel. En consecuencia, las variables independientes en el primer nivel, explican parte de la varianza tanto en el nivel 1 como en el nivel 2; y las variables explicativas en el nivel 2, sólo explican la varianza en el segundo nivel o nivel de los grupos.

Variables

Las variables fueron creadas partiendo de la información recolectada y proporcionada por la Dirección de Asuntos Profesorales de la Universidad de Los Andes. El tiempo transcurrido para que un profesor ascienda a través de los diferentes niveles del escalafón, por encima del establecido en el Estatuto de la Universidad de Los Andes, es la variable objeto de estudio.

Variables del Primer Nivel

- Retraso: es el tiempo en años por encima del reglamentario, que un profesor tarda en ascender a través de los diferentes niveles del escalafón.

- Edad: es la edad del profesor para octubre del año 2011.

- Dedicación: hace referencia a la dedicación del profesor. Toma el valor 0 si es un profesor a dedicación exclusiva y 1 si es de tiempo completo, medio tiempo o tiempo convencional.

- Categoría: tiene que ver con la categoría a la que pertenece el profesor. Toma el valor 0 si el profesor pertenece a la categoría asistente, agregado o asociado y 1 si es titular.

- Contratado: tiempo en años, que el profesor estuvo como contratado antes de pasar a ser miembro ordinario.

- Graduado: es el tiempo en años, desde que el profesor obtuvo su título universitario hasta ser miembro de la universidad.

- Maestría: si el profesor tiene o no Maestría. Toma el valor uno si tiene, y cero en el caso contrario.

- Doctorado: toma el valor uno si el profesor tiene doctorado, y cero en el caso contrario.

Variables del Segundo Nivel

- PRetrasados: es la proporción de profesores en situación de retraso por escuela.

- PTitulares: indica la proporción de profesores que alcanzaron el máximo nivel del escalafón para cada escuela.

- PContratados: hace referencia a la proporción de profesores contratados por escuela.

- PMaestría: es la proporción de profesores que tienen maestría por escuela.

- PDoctorado: tiene que ver con la proporción de profesores que tienen doctorado por escuela.

CAPÍTULO IV

RESULTADOS

Análisis Descriptivo

Las medidas descriptivas del retraso que presentan los profesores de la Universidad de Los Andes para ascender a través de los diferentes niveles del escalafón por escuela, se muestran en la tabla 1. Puede apreciarse que Medicina, Odontología, Ingeniería Forestal, Historia, Departamento de Física y Arquitectura, son las escuelas que tienen mayor número de profesores en condición de retraso, es decir, ya tienen más de 15 años como profesores ordinarios y aún no son titulares, o alcanzaron esta categoría en un tiempo superior al reglamentario. Por el contrario, Ciencias Políticas, Enfermería, Escuela Técnica Superior Forestal, Estadística, Ingeniería Geológica y Medios Audiovisuales son las escuelas que revelan menor cantidad de profesores en tal situación.

Ahora, las escuelas que exhiben un retraso promedio mayor son Arquitectura, Ciencias Políticas, Enfermería, Geografía e Ingeniería Mecánica. En cambio, la Escuela Técnica Superior Forestal, Ingeniería Química y Medios Audiovisuales son las que presentan menor retraso promedio. Al mismo tiempo, Ciencias Políticas y Enfermería son las escuelas que manifiestan una desviación estándar mayor. De otro modo Ingeniería Química, Nutrición y Enseñanza Básica reflejan los menores valores para este coeficiente. Lo cual implica que los datos están más concentrados alrededor del valor promedio en estas tres últimas escuelas.

Tabla 1
Medidas descriptivas del retraso para ascender en el escalafón

ESCUELAS	NÚMERO DE OBSERVACIONES	MEDIA	DESVIACIÓN ESTÁNDAR	MÍN.	MÁX.
1. ADMINISTRACIÓN	7	3.2857143	3.9880775	1	12
2. ARQUITECTURA	15	10.2666667	6.9123768	1	21
3. BIOANÁLISIS	14	2.2857143	2.3673605	1	10
4. CIENCIAS POLÍTICAS	2	14.0000000	18.3847763	1	27
5. DEPARTAMENTO DE BIOLOGÍA	10	4.4000000	4.5018515	1	14
6. DEPARTAMENTO DE FÍSICA	15	6.8666667	5.4098675	1	14
7. DEPARTAMENTO DE MATEMÁTICAS	10	6.2000000	4.1846280	1	15
8. DEPARTAMENTO DE QUÍMICA	10	5.2000000	3.7947332	1	11
9. DERECHO	13	2.8461538	2.1926450	1	8
10. ECONOMÍA	10	3.1000000	1.9692074	1	8
11. EDUCACIÓN	12	6.2500000	5.9256760	1	19
12. ENFERMERÍA	2	9.0000000	11.3137085	1	17
13. ENSEÑANZA BÁSICA	9	2.8888889	1.8333333	1	7
14. ESCUELA TÉCNICA SUPERIOR FORESTAL	1	2.0000000	.	2	2
15. ESTADÍSTICA	2	2.5000000	2.1213203	1	4
16. FARMACIA	12	4.3333333	3.9619401	1	15
17. GEOGRAFÍA	3	8.0000000	6.5574385	1	14
18. HISTORIA	14	5.7142857	4.3928013	1	14
19. IDIOMAS MODERNOS	3	6.0000000	7.0000000	1	14
20. INGENIERÍA CIVIL	8	5.6250000	8.0345237	1	25
21. INGENIERÍA DE SISTEMAS	9	3.2222222	2.0480343	1	7
22. INGENIERÍA ELÉCTRICA	6	4.8333333	4.7504386	1	13
23. INGENIERÍA FORESTAL	14	6.2142857	5.1317799	1	15
24. INGENIERÍA GEOLÓGICA	1	4.0000000	.	4	4
25. INGENIERÍA MECÁNICA	6	7.3333333	4.0331956	2	13
26. INGENIERÍA QUÍMICA	2	2.0000000	1.4142136	1	3
27. LETRAS	5	2.8000000	2.4899799	1	7
28. MEDICINA	44	6.2500000	4.5958778	1	20
29. MEDIOS AUDIOVISUALES	1	2.0000000	.	2	2
30. NUTRICIÓN	5	2.4000000	1.6733201	1	5
31. ODONTOLOGÍA	17	3.5294118	3.5727729	1	16

Nota. Cálculos propios con base en los datos proporcionados por la DAP para octubre del 2011. Cifras en años.

En el gráfico 1 se presenta un diagrama de cajas por escuelas, para el tiempo que demoran los profesores en ascender a través del escalafón. Como puede observarse, los conjuntos de datos correspondientes a Ciencias Políticas, Educación, Estadística, Ingeniería Química, Letras, Nutrición y Odontología, son simétricos. Esto quiere decir, que hay aproximadamente la misma cantidad de valores tanto a la derecha como a la izquierda, del valor promedio.

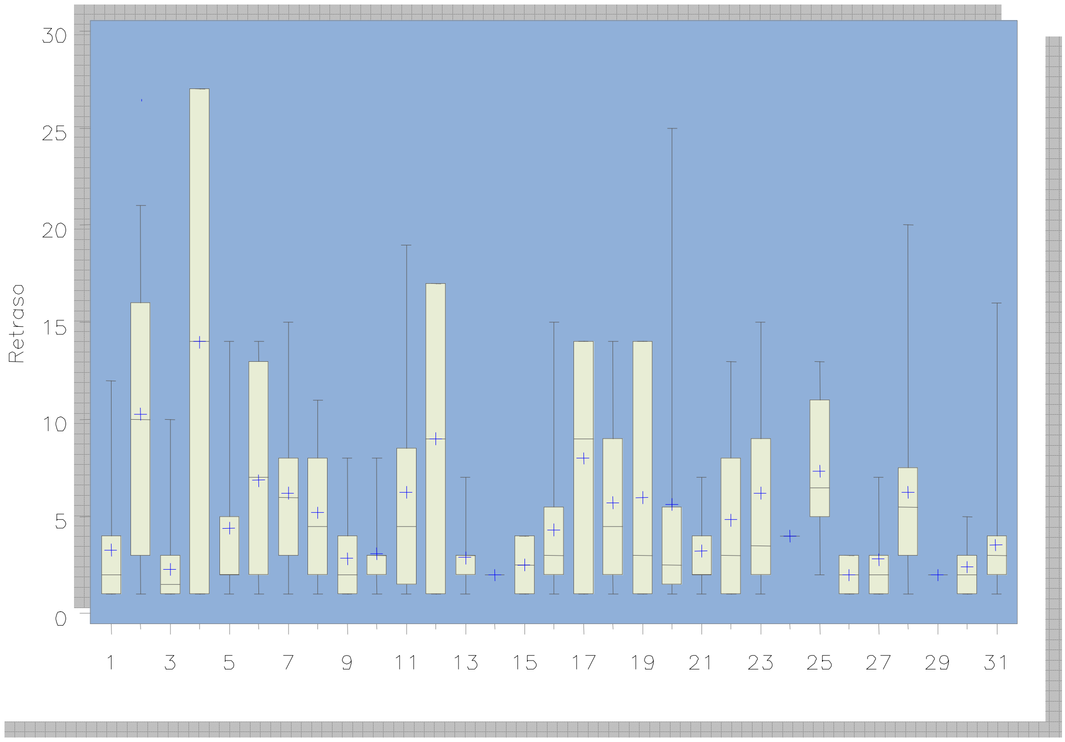

Gráfico 1
Diagrama de cajas por escuelas para el retraso que presentan los profesores de la Universidad de Los Andes al ascender en el escalafón. Cálculos propios.

Al mismo tiempo, las observaciones pertenecientes a las escuelas Administración, Bioanálisis, Departamento de Física, Departamento de Química, Derecho, Educación, Farmacia, Historia, Idiomas Modernos, Ingeniería Civil, Ingeniería Eléctrica, Ingeniería Forestal e Ingeniería Mecánica, están sesgadas a la derecha. Esto indica que los datos se agrupan en el extremo bajo de la escala. Por lo tanto, el 75% de ellos se encuentra ubicado entre el brazo izquierdo y el cuartil 3, y el 25% restante se localiza en el brazo derecho.

Ahora bien, los datos correspondientes a las escuelas Arquitectura, Departamento de Matemáticas, Geografía y Medicina, están sesgados a la izquierda. Lo cual implica que existe un denso agrupamiento de los datos en el extremo alto de la escala. Es decir, el 75% de los valores de los datos se encuentra ubicado entre el cuartil 1 y el extremo del brazo derecho, mientras que el 25% restante, se sitúa en el brazo izquierdo que es el de mayor longitud.

En la tabla 2 se presenta el porcentaje de profesores en condición de retraso, por escuela. En la primera columna es posible apreciar la cantidad de profesores pertenecientes a cada escuela para octubre del año 2011, en la columna siguiente se puede observar la cantidad de profesores en situación de retraso, y en la última columna se señala el porcentaje correspondiente. Cabe destacar que Historia, Economía, Medicina, Nutrición e Ingeniería Civil, son las escuelas que reflejan mayor proporción de profesores que superan el tiempo establecido en el reglamento para alcanzar el máximo nivel del escalafón. Por el contrario, las que revelan menor proporción son Ingeniería Geológica e Ingeniería Química. Dicho comportamiento se ilustra fácilmente en el gráfico 2.

Tabla 2
Porcentaje de retrasados por escuela

ESCUELA	NÚMERO DE OBSERVACIONES	RETRASADOS	PORCENTAJE
1. ADMINISTRACIÓN	70	7	10.0%
2. ARQUITECTURA	75	15	20.0%
3. BIOANÁLISIS	53	14	26.4%
4. CIENCIAS POLÍTICAS	20	2	10.0%
5. DEPARTAMENTO DE BIOLOGÍA	39	10	25.6%
6. DEPARTAMENTO DE FÍSICA	56	15	26.7%
7. DEPARTAMENTO DE MATEMÁTICAS	42	10	23.8%
8. DEPARTAMENTO DE QUÍMICA	52	10	19.2%
9. DERECHO	75	13	17.3%
10. ECONOMÍA	30	10	33.3%
11. EDUCACIÓN	92	12	13.0%
12. ENFERMERÍA	19	2	10.5%
13. ENSEÑANZA BÁSICA	37	9	24.3%
14. ESCUELA TÉCNICA SUPERIOR FORESTAL	13	1	7.6%
15. ESTADÍSTICA	26	2	7.6%
16. FARMACIA	54	12	22.2%
17. GEOGRAFÍA	18	3	16.6%
18. HISTORIA	38	14	36.8%
19. IDIOMAS MODERNOS	31	3	9.6%
20. INGENIERÍA CIVIL	29	8	27.5%
21. INGENIERÍA DE SISTEMAS	37	9	24.3%
22. INGENIERÍA ELÉCTRICA	39	6	15.3%
23. INGENIERÍA FORESTAL	52	14	26.9%
24. INGENIERÍA GEOLÓGICA	20	1	5.0%
25. INGENIERÍA MECÁNICA	29	6	20.6%
26. INGENIERÍA QUÍMICA	29	2	6.8%
27. LETRAS	39	5	12.8%
28. MEDICINA	154	44	28.5%
29. MEDIOS AUDIOVISUALES	12	1	8.3%
30. NUTRICIÓN	18	5	27.7%
31. ODONTOLOGÍA	108	17	15.7%

Nota. Cálculos propios con base en los datos proporcionados por la DAP para octubre del 2011.

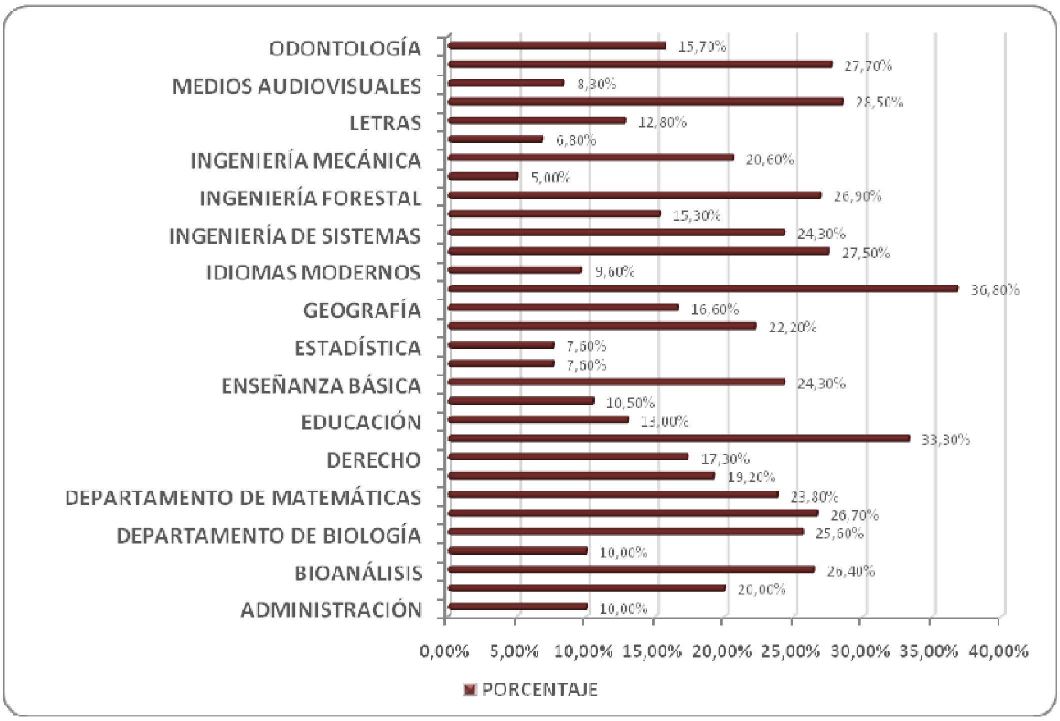

Gráfico 2
Porcentaje de profesores por escuela que están en situación de retraso para octubre del año 2011. Cálculos propios.

En la tabla 3 es posible apreciar la correlación o grado de asociación lineal existente entre la variable objeto de estudio, que es el retraso, y el resto de las variables del primer nivel definidas en el capítulo III, que son: edad, categoría, dedicación, contratado, graduado, maestría y doctorado. Según el coeficiente de correlación de Pearson, con un nivel de significación del 1%, existe correlación positiva y significativa entre la variable objeto de estudio y la edad del profesor para octubre del año 2011. Lo quiere decir que a medida que aumenta la edad del profesor aumenta el tiempo de retraso y viceversa. También existe correlación negativa y significativa entre las variables, retraso y doctorado. Esto significa que los profesores con menor retraso son los que tienen doctorado y que aquellos que tienen mayor retraso no poseen doctorado.

Así mismo, al 5% hay correlación positiva entre el retraso y la dedicación. Lo cual indica, que los profesores con mayor retraso son aquellos cuya dedicación es medio tiempo o tiempo convencional y que los profesores con menor retraso son aquellos que están a dedicación exclusiva. Para ilustrar esto, en el grafico 3 se presenta un diagrama de dispersión que exhibe la correlación existente entre la variable objeto de estudio y las variables edad, doctorado y dedicación.

Tabla 3
Correlación entre el retraso y las variables del primer nivel

	Retraso
Edad	0.51935 <.0001 282
Categoría	-0.03704 0.5356 282
Dedicación	0.12634 0.0339 282
Contratado	0.06585 0.2774 274
Graduado	-0.09306 0.1399 253
Maestría	-0.10990 0.0653 282
Doctorado	-0.21068 0.0004 282

Nota. Cálculos propios con base en los datos proporcionados por la DAP para octubre del 2011.

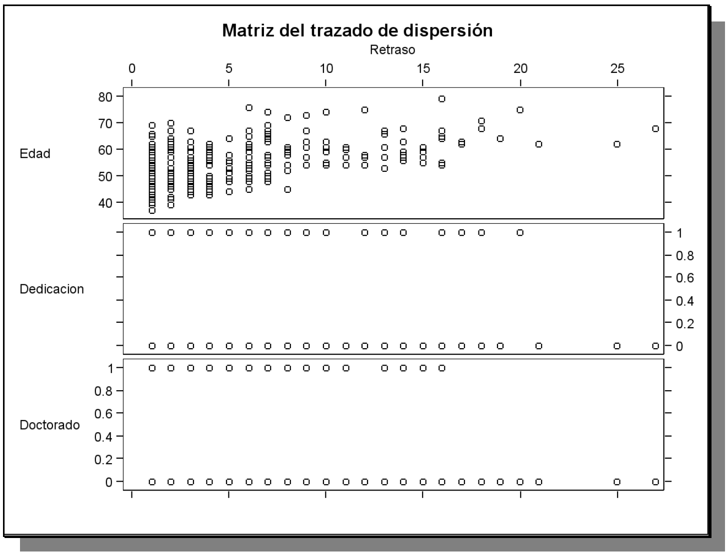

Gráfico 3
Diagrama de dispersión de la correlación entre el retraso y las variables del primer nivel que resultaron ser significativas. Cálculos propios.

Como se puede observar en la tabla 4, de acuerdo con el coeficiente de correlación de Pearson, no existe correlación estadísticamente significativa entre el retraso y las variables medidas sobre las escuelas o unidades del segundo nivel. El diagrama de dispersión que ilustra estos resultados se presenta en el gráfico 4.

Tabla 4
Correlación entre el retraso y las variables del segundo nivel

	Retraso
PRetrasados	0.02546
	0.6703
PTitulates	-0.05225
	0.3820
PContratados	0.05798
	0.3320
PMaestría	-0.07956
	0.1828
PDoctorado	-0.07275
	0.2233

Nota. Cálculos propios con base en los datos proporcionados por la DAP para octubre del 2011.

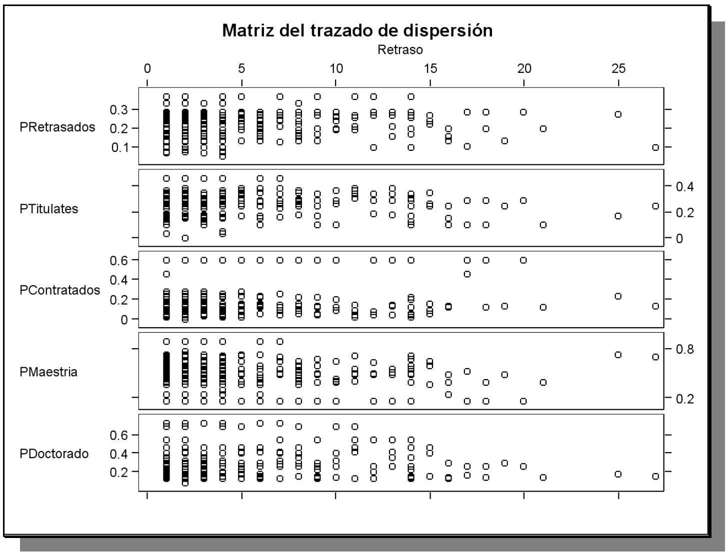

Gráfico 4
Diagrama de dispersión de la correlación entre el retraso y las variables del segundo nivel. Cálculos propios.

El Modelo

Paso 1. Modelo Nulo

En el nivel 1, el modelo nulo, también conocido como modelo incondicional de medias toma la forma

$$y_{ij} = \beta_{0j} + e_{ij}$$

Ahora bien, en el nivel 2, teniendo en cuenta que el intercepto es aleatorio ($\beta_{0j} = \gamma_{00} + u_{0j}$) y sustituyéndolo en la ecuación anterior, el modelo es el siguiente

$$y_{ij} = \gamma_{00} + u_{0j} + e_{ij}, \qquad u_{0j} \sim N\left(0, \sigma_{u0}^2\right) \quad e_{ij} \sim N\left(0, \sigma^2\right)$$

El modelo incondicional de medias está formado por dos componentes aleatorios, uno asociado con la varianza en el primer nivel o nivel de los profesores, y el otro con el intercepto aleatorio o varianza del segundo nivel, que es el nivel de las escuelas, cuyos valores son respectivamente $\hat{\sigma}^2 = 21.7649$ y $\hat{\sigma}_{u0}^2 = 2.3839$ (ver tabla 5). Estos valores indican que hay mayor variabilidad dentro de las escuelas que entre ellas**.**

Por otro lado, contrastando el valor de z asociado a la varianza del nivel 2, con una χ^2 con 30 grados de libertad y un nivel de significación del 5 %, a la que corresponde un valor crítico de 43.77 (ver apéndice B.1), puede comprobarse que se cumple con el supuesto de independencia de la varianza del segundo nivel. La hipótesis a contrastar es la independencia como hipótesis nula versus la no independencia.

Según la prueba de Wald, en la cual se establece como hipótesis nula que el valor de un determinado parámetro es igual a cero y empleando un nivel de significación del 5 por ciento, se rechaza la hipótesis nula que establece que las varianzas no son significativas para los componentes de variación de los diferentes niveles. Como prueba adicional a la de Wald para verificar si en realidad el componente de varianza del segundo nivel es significativo se procede a utilizar la prueba de la razón de la verosimilitud, en la cual se compara la verosimilitud del modelo incondicional de medias con la de un modelo de regresión tradicional sin variables independientes.

Tabla 5
Componentes de varianza del modelo nulo

Parm Cov	Asunto	Estimador	Error estándar	Valor Z	Pr Z
Intercepto	ID	2.3839	1.2995	1.83	0.0333
Residual		21.7649	1.9188	11.34	<.0001

Nota. Cálculos propios con base en los datos proporcionados por la DAP para octubre del 2011.

En la tabla 6 se ilustra el valor de la verosimilitud para el modelo nulo, para el modelo de regresión lineal simple y la diferencia entre ambas. Dado que la diferencia entre la verosimilitud de la regresión simple y la del modelo nulo es mayor a 3.841, que es el valor critico para una χ^2con 1 grado de libertad y un nivel de significación del 5%, el modelo nulo se ajusta mejor a los datos. Es decir, existen diferencias significativas entre las escuelas en cuanto al retraso promedio que presentan los profesores de la Universidad de Los Andes al ascender a través de los diferentes niveles del escalafón.

Tabla 6
Verosimilitud para el modelo nulo y el de regresión tradicional

Modelo	Reg. Tradicional (l_0)	Nulo (l_1)	(l_0)- (l_1)
Verosimilitud -2 Res Log	1696.8	1687.0	9.8

Nota. Cálculos propios con base en los datos proporcionados por la DAP para octubre del 2011.

Como se mencionó en el capítulo II, la correlación intraclase refleja la proporción de la varianza total que se debe a las unidades del segundo nivel.

Entonces, para estudiar la variación de la variable respuesta alrededor de las unidades del nivel 2, se procedió a estimar la correlación intraclase, como se muestra a continuación

$$\rho = \frac{\sigma_{u0}^2}{\left(\sigma_{u0}^2 + \sigma^2\right)} = \frac{2.3839}{\left(2.3839 + 21.7649\right)} = 0.0987$$

El valor que arroja la correlación intraclase indica que el 9,87% de la variación total se debe a las escuelas o unidades del segundo nivel. Ahora, teniendo en cuenta que el modelo incondicional de medias, es el primer paso o procedimiento exploratorio para llevar a cabo un análisis multinivel y sabiendo que dicho modelo proporciona una base con la que se pueden comparar modelos más complejos. Luego de comprobar que la variación entre las unidades del segundo nivel es significativa, se procede a ajustar un modelo por niveles.

El modelo nulo contiene sólo un componente fijo, $\hat{\gamma}_{00}$ = 5.0718, que es el intercepto o media total del retraso que presentan los profesores del núcleo Mérida para ascender, el cual es significativo debido a que el valor de t asociado es superior a 2 (ver tabla 7). La ecuación correspondiente al modelo nulo ajustado en este paso toma la siguiente forma

$$\text{Retraso}_{ij} = 5.0718 + u_{0j} + e_{ij}$$

donde $\hat{\sigma}^2 = 21.7649$ y $\hat{\sigma}_{u0}^2 = 2.3839$

Tabla 7
Efectos fijos del modelo nulo

Efecto	Estimador	Error estándar	DF	Valor t	Pr > |t|
Intercepto	5.0718	0.4241	30	11.96	<.0001

Nota. Cálculos propios con base en los datos proporcionados por la DAP para octubre del 2011.

En la tabla 8 se muestra información que es de interés a la hora de comparar la bondad de ajuste de varios modelos con los mismos efectos fijos pero diferentes efectos aleatorios. Como se mencionó en el capítulo II, Singer 1998, plantea que a la hora de comparar la bondad de ajuste de varios modelos con los mismos efectos fijos pero diferentes efectos aleatorios, los criterios más utilizados son el criterio de información de Akaike (*AIC*) y el criterio de Schwarz (*BIC*). Entonces, ajusta mejor a los datos aquel modelo que presente menores valores de estos criterios. Otro método de comparación es la verosimilitud que presenta el modelo, como se mencionó en el capítulo II.

Tabla 8
Estadísticos de ajuste para el modelo nulo

Verosimilitud -2 Res Log	1687.0
AIC (mejor más pequeño)	1691.0
AICC (mejor más pequeño)	1691.0
BIC (mejor más pequeño)	1693.9

Nota. Cálculos propios con base en los datos proporcionados por la DAP para octubre del 2011.

Paso 2. Incluyendo Variables Explicativas Fijas en el Primer Nivel

En el segundo nivel, el modelo incluyendo variables explicativas fijas se representa de la siguiente manera

$$y_{ij} = \gamma_{00} + \sum_{p=1}^{P} \gamma_{p0} x_{pij} + u_{0j} + e_{ij}, \qquad u_{0j} \sim N\left(0, \sigma_{u0}^2\right) \quad e_{ij} \sim N\left(0, \sigma^2\right)$$

En la ecuación anterior, γ_{00}, es la media total del retraso o variable dependiente, el término $\sum_{p=1}^{P} \gamma_{p0} x_{pij}$, hace referencia a las pendientes de las variables independientes fijas en el nivel 1, por último, u_{0j} y e_{ij} hacen referencia a los términos aleatorios del segundo y primer nivel, respectivamente.

Luego de incluir a las variables, edad, dedicación y doctorado, que fueron las que reflejaron correlación estadísticamente significativa con la variable objeto de estudio, la única variable que resultó ser significativa en la predicción del retraso fue la edad, pero el signo que acompaña al intercepto para este modelo resultó ser negativo, lo que no tiene sentido en el contexto del presente análisis (ver tabla 9). Por tal razón, se elimina del modelo a la variable edad.

Tabla 9

Efectos fijos incluyendo a las variables edad, dedicación y doctorado

Efecto	Estimador	Error estándar	DF	Valor t	Pr > |t|
Intercepto	-11.1495	1.9437	30	-5.74	<.0001
Edad	0.3054	0.03454	248	8.84	<.0001
Dedicación	-0.1553	0.9010	248	-0.17	0.8633
Doctorado	-0.3607	0.5573	248	-0.65	0.5181

Nota. Cálculos propios con base en los datos proporcionados por la DAP para octubre del 2011.

Ahora, el modelo resultante revela que la variable doctorado es estadísticamente significativa y que la dedicación no lo es (ver tabla 10). Por lo tanto se elimina del modelo a la variable dedicación.

Tabla 10
Efectos fijos incluyendo a las variables dedicación y doctorado

Efecto	Estimador	Error estándar	DF	Valor t	Pr > |t|
Intercepto	5.6406	0.4821	30	11.70	<.0001
Dedicación	1.6155	1.0367	249	1.56	0.1204
Doctorado	-1.7421	0.6037	249	-2.89	0.0042

Nota. Cálculos propios con base en los datos proporcionados por la DAP para octubre del 2011.

Posteriormente, se ajustó un modelo incluyendo únicamente a la variable doctorado, la cual resultó explicar mejor el comportamiento del retraso[3]. Los resultados se exponen a continuación.

Debido a que la variable explicativa es fija, el modelo está conformado por los mismos componentes aleatorios que el modelo nulo, los valores para estos coeficientes son $\hat{\sigma}^2 = 21.2562$ y $\hat{\sigma}^2_{u0} = 1.9210$ (ver tabla 11). Tales coeficientes tienen en este momento diferente significado, ya no son incondicionales, ahora están condicionados por la variable incluida en el modelo. Empleando la prueba de Wald, para un nivel de significación del 5 por ciento, se rechaza la hipótesis nula que establece que las varianzas no son significativas para los componentes de varianza de ambos niveles.

Así mismo, contrastando el valor de z asociado a la varianza del segundo nivel, con una χ^2 con 30 grados de libertad y un nivel de significación del 5%, a la

[3] Además de evaluar la contribución de las variables que presentaron correlación estadísticamente significativa con el retraso, también se estudiaron el resto de las variables del nivel 1 y como era de esperarse, no aportaron ninguna información en la explicación de la variable de interés.

que corresponde un valor crítico de 43.77 (ver apéndice B.1), puede comprobarse que en este paso también se cumple con el supuesto de independencia de las varianzas del segundo nivel.

Tabla 11
Componentes de varianza del modelo del paso 2

Parm Cov	Asunto	Estimador	Error estándar	Valor Z	Pr Z
Intercepto	ID	1.9210	1.1467	1.68	0.0469
Residual		21.2562	1.8753	11.33	<.0001

Nota. Cálculos propios con base en los datos proporcionados por la DAP para octubre del 2011.

El valor de la verosimilitud para el modelo propuesto en este paso, el valor correspondiente al modelo de regresión lineal simple tradicional y la diferencia entre ambos se presenta en la tabla 12. Dado que la diferencia entre la verosimilitud de la regresión tradicional y la del modelo del paso 2 es mayor a 3.841, que es el valor critico para una χ^2 con 1 grado de libertad y un nivel de significación del 5%, se comprueba que existen diferencias significativas entre escuelas en cuanto al retraso promedio que presentan los profesores para ascender en el escalafón luego de la inclusión de la variable doctorado.

Tabla 12
Verosimilitud para el modelo del paso 2 y el de regresión tradicional

Modelo	Reg. Tradicional (l_0)	Paso 2 (l_1)	(l_0)- (l_1)
Verosimilitud -2 Res Log	1683.2	1676.2	7

Nota. Cálculos propios con base en los datos proporcionados por la DAP para octubre del 2011.

Para el modelo incondicional de medias, las estimaciones de la varianza dentro de las escuelas y entre ellas, son respectivamente 21.7649 y 2.3839 (ver tabla 5). Ahora, para el modelo condicionado por una variable explicativa fija en el primer nivel, los valores estimados de las varianzas de los niveles 1 y 2 son

21.2562 y 1.9210 (ver tabla 11). La inclusión de tal variable contribuye a explicar (21.7649 – 21.2562) / 21.7649 = 0.023 o el 2.3% de la variación existente dentro de las escuelas o variación del nivel 1, y (2.3839 - 1.9210) / 2.3839 = 0.19 o 19% de la variación que hay entre ellas o variación del nivel 2.

El modelo contiene dos coeficientes fijos: en primer lugar $\hat{\gamma}_{00}$ = 5.8247, que es el intercepto o media total del retraso que presentan los profesores de la Universidad de Los Andes para ascender; y en segundo lugar $\hat{\gamma}_{10}$ = -1.9082, que es la pendiente que acompaña a la variable doctorado, refleja el cambio en el retraso cuando la variable toma el valor 1, esto es, cuando el profesor tiene doctorado. Entonces, para aquellos profesores que tienen doctorado el retraso disminuye en 1.9082. Cabe destacar que dichos coeficientes son significativos al 5 por ciento (ver tabla 13). La ecuación correspondiente al modelo ajustado en este paso se presenta a continuación

$$\text{Retraso}_{ij} = 5.8247 - 1.9082 * Doctorado_{ij} + u_{0j} + e_{ij}$$

donde $\hat{\sigma}^2 = 21.2562$ y $\hat{\sigma}^2_{u0} = 1.9210$

Tabla 13
Efectos fijos del modelo del paso 2

Efecto	Estimador	Error estándar	DF	Valor t	Pr > |t|
Intercepto	5.8247	0.4628	30	12.59	<.0001
Doctorado	-1.9082	0.5963	250	-3.20	0.0016

Nota. Cálculos propios con base en los datos proporcionados por la DAP para octubre del 2011.

Los estadísticos de ajuste para comparar el modelo de este paso con el modelo incondicional de medias o modelo nulo, se ilustran en la tabla 14. Ya que la diferencia entre la verosimilitud del modelo del paso 1 y la del modelo del paso 2 es mayor a 3.841, que es el valor crítico para una χ^2 con 1 grado de libertad y un nivel de significación del 5%, el modelo del paso 2 se ajusta mejor a los datos (ver tabla 15).

Tabla 14
Estadísticos de ajuste para el modelo del paso 2

Verosimilitud -2 Res Log	1676.2
AIC (mejor más pequeño)	1680.2
AICC (mejor más pequeño)	1680.3
BIC (mejor más pequeño)	1683.1

Nota. Cálculos propios con base en los datos proporcionados por la DAP para octubre del 2011.

Tabla 15
Verosimilitud para el modelo nulo y el modelo propuesto en el paso 2

Modelo	Nulo (l_0)	Paso 2 (l_1)	(l_0)- (l_1)
Verosimilitud -2 Res Log	1687.0	1676.2	10.8

Nota. Cálculos propios con base en los datos proporcionados por la DAP para octubre del 2011.

Paso 3. Evaluación de los Componentes de Varianza de las Pendientes de las Variables del Primer Nivel

En el segundo nivel, el modelo con intercepto aleatorio y variables explicativas aleatorias en el primer nivel se muestra a continuación

$$y_{ij} = \gamma_{00} + \sum_{p=1}^{P} \gamma_{p0} x_{pij} + \sum_{p=1}^{P} u_{pj} x_{pij} + u_{0j} + e_{ij}, \qquad u_{Pj} \sim N\left(0, \sigma_{uP}^{2}\right)$$

Los resultados de las pruebas de variación de las pendientes correspondientes a las variables del nivel 1 se muestran en la tabla 16. Es posible observar que sólo el componente de variación de la pendiente de la variable dedicación resultó ser significativa, puesto que la diferencia entre la verosimilitud del modelo del paso 2 y el modelo que incluye a la variable dedicación aleatoria en el primer nivel, es superior a 5.99 que es el valor del estadístico de prueba para una χ^2 con 2 grados de libertad. La variable graduado a pesar de que presenta una verosimilitud significativamente pequeña no tiene un componente de varianza estadísticamente significativo, ya que la estimación del componente de varianza asociado a ella es cero, razón por la cual no presenta resultado para el p-valor y no se incluye en el modelo.

Tabla 16
Evaluación de la varianza de las pendientes de las variables del primer nivel

Variable	Modelo paso 2 (l_0)		Modelo paso 3 (l_1)		(l_0)- (l_1)	gl	p-valor
	Verosimilitud	Nº de parámetros estimados	Verosimilitud	Nº de parámetros estimados			
Edad	El modelo no converge						
Categoría	1676.2	4	1673.8	6	2.4	2	.
Dedicación	1676.2	4	1669.6	6	6.6	2	0.4796
Contratado	El modelo no converge						
Graduado	1676.2	4	1491.2	6	184.4	2	.
Maestría	1676.2	4	1671.8	6	4.4	2	.
Doctorado	1676.2	4	1673.9	5	2.3	1	0.4283

Nota. Cálculos propios con base en los datos proporcionados por la DAP para octubre del 2011.

El modelo final de este paso está conformado entonces por las variables doctorado y dedicación, la primera fija y la segunda aleatoria en el nivel 1 por presentar un componente de varianza significativo. El hecho de que la pendiente que acompaña a la variable dedicación, presente un componente de varianza significativo, refleja que la relación que existe entre la dedicación y la variable dependiente que es el retraso, no es igual en todas las escuelas. Es decir, la influencia de la variable dedicación sobre el retraso va a ser diferente para cada una de las escuelas o unidades del nivel 2. Los resultados de dicho modelo se exhiben a continuación.

Los componentes de varianza que se presentan en la tabla 17, corresponden a las estimaciones de la varianza dentro de las escuelas o varianza del primer nivel, que es $\hat{\sigma}^2 = 21.0337$; la varianza entre escuelas o varianza del segundo nivel asociada con el intercepto aleatorio, la cual es $\hat{\sigma}_{u0}^2 = 1.7567$; la varianza asociada con la pendiente de la variable dedicación, aleatoria en el nivel 1, que es $\hat{\sigma}_{u1}^2 = 0.2118$, y la covarianza entre los componentes de error del nivel 2, $\hat{\sigma}_{u01} = 3.7901$, la cual indica que la relación entre dos unidades del mismo grupo es positiva. Es posible observar que el componente de varianza del primer nivel es estadísticamente significativo al 1 %. Esto quiere decir que aún existe variación aleatoria por explicar en ese nivel. En cambio, los componentes de variación asociados con el intercepto y con la variable dedicación no son estadísticamente significativos según la prueba de Wald.

Al contrastar el valor de z asociado a la varianza del segundo nivel, con una χ^2 con 30 grados de libertad y un nivel de significación del 5%, a la que corresponde un valor crítico de 43.77 (ver apéndice B.1), puede comprobarse que en este paso también se cumple con el supuesto de independencia de los componentes varianza del nivel 2.

Tabla 17

Componentes de varianza del modelo del paso 3

Parm Cov	Asunto	Estimador	Error estándar	Valor Z	Pr Z
UN(1,1)	ID	1.7567	1.1515	1.53	0.0636
UN(2,1)	ID	3.7901	2.4704	1.53	0.1250
UN(2,2)	ID	0.2118	4.1410	0.05	0.4796
Residual		21.0337	1.8599	11.31	<.0001

Nota. Cálculos propios con base en los datos proporcionados por la DAP para octubre del 2011.

Para verificar si los componentes de variación asociados al intercepto y a la variable dedicación son significativos, debido a que la prueba de Wald no siempre es acertada, se procede a realizar la prueba de razón de verosimilitud. En la tabla 18 se aprecia la verosimilitud para un modelo tradicional que contiene a las variables, doctorado y dedicación, y la verosimilitud del modelo del paso 3 que contiene a las mismas variables pero en este caso la dedicación es aleatoria en el primer nivel. Al comparar la diferencia entre ambas con una χ^2 con 2 grados de libertad se comprueba que ambos componentes de varianza son significativos.

Tabla 18
Verosimilitud para el modelo tradicional y el modelo del paso 3

Modelo	Tradicional (l_0)	Paso 3 (l_1)	(l_0)- (l_1)
Verosimilitud -2 Res Log	1679.4	1669.6	9.8

Nota. Cálculos propios con base en los datos proporcionados por la DAP para octubre del 2011.

Para el modelo incondicional de medias, las estimaciones de las varianzas dentro y entre escuelas, son respectivamente 21.7649 y 2.3839 (ver tabla 5). Ahora, para el modelo propuesto en el paso 3, los valores estimados de las varianzas del primer y segundo nivel son 21.0337 y 1.7567 (ver tabla 17). Por lo tanto, el modelo hasta ahora contribuye a explicar (21.7649 – 21.0337) / 21.7649 = 0.033 o el 3.3% de la variación existente dentro de las escuelas o unidades del primer nivel, y (2.3839 - 1.7567) / 2.3839 = 0.26 o 26% de la variación que hay entre las escuelas o unidades del segundo nivel.

El modelo de este paso contiene tres coeficientes fijos: en primer lugar $\hat{\gamma}_{00}$ = 5.6032, que es la media total del retraso que presentan los profesores de la Universidad de Los Andes para ascender; en segundo lugar $\hat{\gamma}_{10}$ = 2.8790, que es la pendiente que acompaña a la variable dedicación, indica el cambio en el retraso cuando la variable toma el valor 1, esto es, cuando el profesor es de tiempo completo, medio tiempo o tiempo convencional. Esto quiere decir, que cuando el profesor es a tiempo completo, medio tiempo o a tiempo convencional, el retraso aumenta en 2.8790; y por último $\hat{\gamma}_{20}$ = -1.7314 que es la pendiente que acompaña a la variable doctorado, ilustra el cambio en el retraso cuando la variable toma el valor 1, es decir, cuando el profesor tiene doctorado. Entonces, para aquellos profesores que tienen doctorado, el retraso disminuye en 1.7314. Por otra parte, se puede observar que dichos coeficientes son significativos al 5 por ciento (ver tabla 19). La ecuación correspondiente al modelo ajustado en este paso tiene la siguiente forma

$$\text{Retraso}_{ij} = 5.6032 + (2.8790 * Dedicación_{ij}) - (1.7314 * Doctorado_{ij}) + u_{0j} + (u_{1j} * Dedicación_{j}) + e_{ij}$$

donde $\hat{\sigma}^2 = 21.2562$, $\hat{\sigma}^2_{u0} = 1.7567$, $\hat{\sigma}^2_{u1} = 0.2118$ y $\hat{\sigma}_{u01} = 3.7901$

Tal ecuación sugiere lo siguiente. En primer lugar, que el tiempo esperado de retraso para que un profesor a dedicación exclusiva que no tiene doctorado alcance el máximo nivel del escalafón es de 5.6032 años. En segundo lugar, que existe relación lineal entre el retraso y las variables dedicación y doctorado. La variable dedicación contribuye a aumentar el valor estimado del retraso en 2.879 años, para aquellos profesores a tiempo completo, medio tiempo o a tiempo convencional. Cabe resaltar, que la relación entre la dedicación y el retraso no es igual en todas las escuelas. Por su parte, el valor estimado para el retraso disminuye en 1.7314 para aquellos profesores que tienen doctorado. A diferencia de lo que ocurre con la variable dedicación, la relación que existe entre el retraso y el doctorado es similar en todas las escuelas.

Tabla 19
Efectos fijos del modelo del paso 3

Efecto	Estimador	Error estándar	DF	Valor t	Pr > |t|
Intercepto	5.6032	0.4614	30	12.14	<.0001
Dedicación	2.8790	0.4095	249	7.03	<.0001
Doctorado	-1.7314	0.5951	249	-2.91	0.0039

Nota. Cálculos propios con base en los datos proporcionados por la DAP para octubre del 2011.

Los estadísticos de ajuste para comparar el modelo de este paso con el modelo propuesto en el paso 2, se presentan en la tabla 20. Debido a que la diferencia entre la verosimilitud del modelo del paso 2 y la del modelo del paso 3 es mayor a 5.99, que es el valor crítico para una χ^2 con 2 grado de libertad y un nivel de significación del 5%, el modelo del paso 3 ajusta mejor a los datos (ver tabla 21).

Tabla 20
Estadísticos de ajuste para el modelo del paso 3

Verosimilitud -2 Res Log	1669.6
AIC (mejor más pequeño)	1677.6
AICC (mejor más pequeño)	1677.7
BIC (mejor más pequeño)	1683.3

Nota. Cálculos propios con base en los datos proporcionados por la DAP para octubre del 2011.

Tabla 21
Verosimilitud para el modelo del paso 2 y el modelo propuesto en el paso 3

Modelo	Paso 2 (l_0)	Paso 3 (l_1)	(l_0)- (l_1)
Verosimilitud -2 Res Log	1676.2	1669.6	6.6

Nota. Cálculos propios con base en los datos proporcionados por la DAP para octubre del 2011.

Paso 4. Agregando Variables en el segundo Nivel

A pesar de que ninguna de las variables del segundo nivel reflejó correlación estadísticamente significativa con el retraso, fueron introducidas en la ecuación del intercepto para probar su contribución al modelo. Como era de esperarse, ninguna resultó ser significativa explicando las diferencias con respecto al retraso que presentan los profesores en sus ascensos, entre las diferentes escuelas que conforman al núcleo Mérida de la Universidad de Los Andes.

Ahora, teniendo en cuenta que la única variable del primer nivel que presentó un componente de variación estadísticamente significativo entre los grupos fue la dedicación, la cual toma el valor cero si el profesor es, a dedicación exclusiva y uno en otro caso. Por la naturaleza de esta variable, no tiene sentido agregar covariables en la ecuación de esta pendiente para explicar tales diferencias, pues en esta institución la pertenencia de un profesor a determinada dedicación depende entre otras cosas, de la categoría a la que pertenezca y de las necesidades de cada escuela.

Por lo expuesto anteriormente, de los modelos propuestos en el intento de explicar el comportamiento del retraso que presentan los profesores para ascender a través de los diferentes niveles del escalafón, el modelo final es el que se presenta en al paso 3. A continuación se verifican los supuestos de dicho modelo.

Verificando Los Supuestos Del Modelo Final

Para que los resultados del modelo final sean confiables, además de que las varianzas del nivel 2 deben ser independientes, este debe cumplir con los supuestos de normalidad e igualdad de varianzas. En las figuras 5 y 6 se aprecian pruebas gráficas para verificar estos supuestos partiendo de residuos marginales y condicionales, respectivamente. Es posible observar que los valores de los residuos no se alinean completamente sobre la recta, lo que parece indicar que no están normalmente distribuidos. Por otro lado, la forma como se agrupan los valores predichos podría reflejar que las varianzas no son homogéneas.

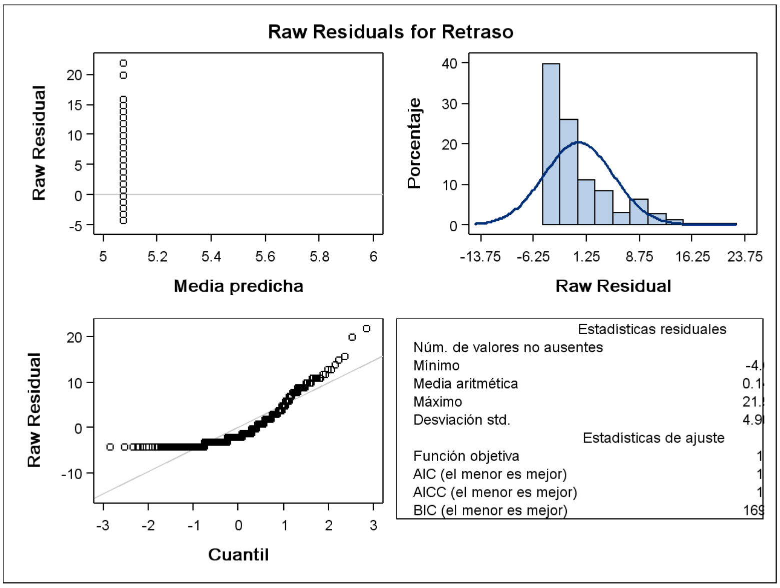

Figura 5.
Pruebas gráficas para evaluar los supuestos de normalidad y homogeneidad de varianzas para la variable Retraso utilizando los residuos marginales. Fuente: cálculos propios.

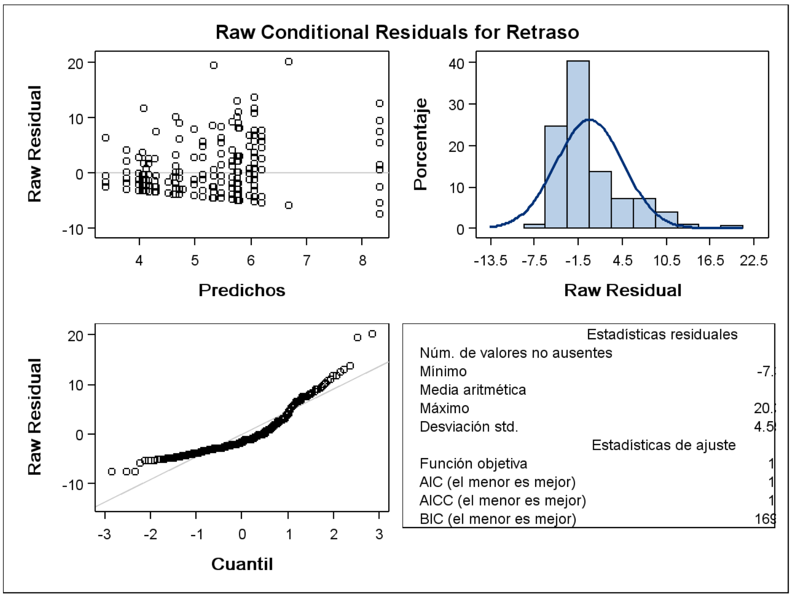

Figura 6.
Pruebas gráficas para evaluar los supuestos de normalidad y homogeneidad de varianzas para la variable Retraso utilizando los residuos condicionales. Fuente: cálculos propios.

Para verificar si tales supuestos se cumplen o no, se procede a realizar pruebas adicionales. Con un nivel de significación del 5 por ciento, luego de realizar diferentes pruebas de normalidad, se rechaza la hipótesis nula que establece que los errores se distribuyen normalmente. Por lo tanto, queda confirmado que los residuos para el modelo final ajustado para el retraso, no están normalmente distribuidos (ver tabla 23).

Tabla 22
Pruebas de normalidad para el modelo final

Test	Estadístico		P-valor	
Shapiro-Wilk	#11 X	0.80652	Pr < W	0.000
Kolmogorov-Smirnov	D	0.21140	Pr > D	<0.010
Cramer-von Mises	W-Sq	3.19154	Pr > W-Sq	<0.005
Anderson-Darling	A-Sq	18.04507	Pr > A-Sq	<0.005

Nota. Cálculos propios con base en los datos proporcionados por la DAP para octubre del 2011.

Ahora bien, como el valor de *p* (0,0698) es mayor que α (0,05) no se rechaza la hipótesis nula que establece que las varianzas son iguales. Por lo tanto, no se cumple el supuesto de normalidad pero sí se cumple el de igualdad de varianzas (ver tabla 23).

Tabla 23
Prueba Levene para homogeneidad de varianzas del modelo final

Fuente	DF	Suma de cuadrados	Cuadrado de la media	F-Valor	Pr > F
Escuela	23	42066.0	1829.0	1.50	0.0698
Error	247	300713	1217.5		

Nota. Cálculos propios con base en los datos proporcionados por la DAP para octubre del 2011.

CAPÍTULO V

CONCLUSIONES

Se consideran en situación de retraso a todos aquellos profesores ordinarios que tienen más de 15 años perteneciendo a la universidad y aún no han alcanzado la titularidad, y a quienes alcanzaron el máximo nivel del escalafón en un tiempo superior al que establece el Estatuto de la Universidad de Los Andes y la Ley de Universidades.

El análisis descriptivo sobre el retraso que presentan los profesores del núcleo Mérida para ascender a través de los diferentes niveles del escalafón revela que las escuelas que tienen mayor número de profesores en situación de retraso son Medicina, Odontología, Ingeniería Forestal, Historia, Departamento de Física y Arquitectura. Mientras que Ciencias Políticas, Enfermería, Escuela Técnica Superior Forestal, Estadística, Ingeniería Geológica y Medios Audiovisuales, son las que presentan menor cantidad de profesores en esta situación.

Asimismo, las escuelas que reflejan un retraso promedio superior son Arquitectura, Ciencias Políticas, Enfermería, Geografía e Ingeniería Mecánica. Por el contrario, la Escuela Técnica Superior Forestal, Ingeniería Química y Medios Audiovisuales son las que presentan menor retraso promedio. En cuanto a la proporción de profesores en condición de retraso por cada escuela, cabe destacar que Historia, Economía, Medicina, Nutrición e Ingeniería Civil, son las que muestran mayores proporciones. Ahora, las que presentan proporciones inferiores son Ingeniería Geológica e Ingeniería Química.

Con respecto a la correlación entre la variable retraso y las variables independientes, empleando el coeficiente de correlación de Pearson y partiendo de un nivel de significación del 1%, se observó que existe correlación positiva entre la variable objeto de estudio y la edad del profesor para octubre del año 2011, y correlación negativa entre las variables, retraso y doctorado. Por su parte, al 5 % hay correlación positiva entre el retraso y la dedicación. Al mismo tiempo, no hay evidencia de correlación entre el retraso y ninguna de las variables medidas sobre las unidades del segundo nivel.

De acuerdo con el diagrama de cajas por escuelas para el retraso, los conjuntos de datos correspondientes a Ciencias Políticas, Educación, Estadística, Ingeniería Química, Letras, Nutrición y Odontología, son simétricos. Por su parte, las observaciones pertenecientes a las escuelas Administración, Bioanálisis, Departamento de Física, Departamento de Química, Derecho, Educación, Farmacia, Historia, Idiomas Modernos, Ingeniería Civil, Ingeniería Eléctrica, Ingeniería Forestal e Ingeniería Mecánica, están sesgadas a la derecha. Mientras que, los valores correspondientes a las escuelas Arquitectura, Departamento de Matemáticas, Geografía y Medicina, están sesgados a la izquierda.

Los resultados del modelo nulo reflejan que el componente de variación del nivel 1 es estadísticamente significativo al 1%, que la varianza asociada con el intercepto es significativa tanto al 5% como empleando la prueba de razón de la verosimilitud, y que el coeficiente fijo que es el retraso promedio total, es significativo por presentar un valor de t superior a 2. Por su parte, se obtuvo un valor para la correlación intra clase de 0.0987, este valor indica que 9.87% de la variación presente en la variable respuesta se debe a las escuelas. El hecho de que el componente de varianza del segundo nivel sea significativo demuestra que existen diferencias significativas en cuanto al retraso promedio entre las escuelas, y justifica el haber ajustado un modelo de regresión por niveles para estudiar el comportamiento del retraso en los ascensos.

Al evaluar la contribución de cada una de las variables independientes fijas del primer nivel, la única que resultó ser significativa fue la covariable doctorado. Los resultados del modelo que incluye a esta variable fija en el nivel 1 reflejan que el componente de variación de dicho nivel es estadísticamente significativo al 1%, que el componente de variación del nivel 2 es significativo tanto al 5% como aplicando la prueba de razón de verosimilitud, así como también que los coeficientes fijos, que son el retraso promedio total y la pendiente que acompaña a la variable independiente doctorado son significativo por presentar un valor de t superior a 2.

Por otro lado, la inclusión de la variable doctorado contribuye a explicar un 2.3% de la variación existente dentro las escuelas, y un 19% de la variación presente entre ellas. Al comparar la verosimilitud de los modelos propuestos en los dos primeros pasos, con un nivel de significación del 5% se comprobó que el modelo del paso 2 se ajusta mejor a los datos.

De la evaluación de los componentes de variación de las pendientes de las variables del primer nivel resaltó que la única variable que mostró un componente de variación significativo entre las escuelas fue la dedicación. Los resultados del modelo que incluye a las variables doctorado y dedicación, la primera fija en el nivel 1. Reflejan que el componente de variación del primer nivel es estadísticamente significativo al 1%, que el componente de variación del segundo nivel es significativo según la prueba de razón de verosimilitud, y que los coeficientes fijos que son: el intercepto, la pendiente que acompaña a la variable dedicación y la pendiente que acompaña a la variable doctorado son significativos al 5%.

El modelo propuesto en el tercer paso explica un 3.3% de la variación que existe dentro de las escuelas y un 26% de la variación presente entre las mismas. Al contrastar la diferencia entre la verosimilitud del modelo del paso 2 y la del modelo del paso 3, con una χ^2 con 2 grado de libertad y un nivel de significación del 5%, se demuestra que el modelo del paso 3 se ajusta mejor a los datos.

A la hora de agregar variables en el nivel 2, ninguna resultó ser significativa explicando las diferencias que existen entre las escuelas en cuanto al retraso promedio. Por otra parte, no hay interacción entre variables explicativas de los diferentes niveles, porque dada la naturaleza de la variable dedicación no tiene sentido agregar covariables en la ecuación de ésta pendiente para explicar las diferencias entre los grupos.

Por lo expuesto anteriormente, el modelo que mejor explica el comportamiento del retraso que presentan los profesores pertenecientes a las diferentes escuelas que conforman al Núcleo Mérida de la Universidad de Los Andes es el ajustado en el tercer paso. Dicho modelo está formado en primer lugar, por el retraso promedio total para alcanzar el máximo nivel del escalafón, que toma el valor 5.6032; en segundo lugar, por la variable dedicación aleatoria en el primer nivel, la cual contribuye aumentando el valor del retraso en 2.8790 cuando el profesor es a medio tiempo o a tiempo convencional; en tercer lugar, por la variable doctorado fija, quién contribuye a disminuir en 1.7324 el retraso para aquellos profesores que tienen doctorado; luego, por el termino de error asociado al intercepto aleatorio; posteriormente, por el término de error asociado con la pendiente aleatoria de la variable dedicación multiplicado por el valor que toma la misma y por último, por el error aleatorio del nivel 1.

En cuanto a los supuestos de dicho modelo, se comprobó mediante una prueba χ^2 con 30 grados de libertad que se cumple con la suposición de independencia de los componentes de varianza del segundo nivel. También se evidenció que al 5% se cumple con el supuesto de igualdad de varianzas. Por el contrario, no se cumple con el de de normalidad de los residuos. Lo cual no es un problema serio teniendo en cuenta que el modelo estadístico es robusto ante desviaciones moderadas de supuesto de normalidad (Sánchez, citado en Quintero, 2007), y que el estudio se realizó sobre la población y no sobre una muestra. Según Murillo (2008), el error debe tener una distribución normal para que se puedan inferir los resultados de la muestra a la población.

En resumen, existen diferencias significativas entre las escuelas que conforman al núcleo Mérida, en cuanto al retraso promedio de los profesores al ascender en el escalafón para octubre del año 2011. Las variables del primer nivel que resultaron ser significativas son doctorado y dedicación. La relación entre el retraso y la variable doctorado, resultó ser similar entre las diferentes escuelas, mientras que la relación entre la variable objeto de estudio y la dedicación no es la misma para todas las unidades del segundo nivel. Por su parte, ninguna de las variables medidas sobre los grupos contribuye a explicar las diferencias que existen entre ellos respecto al retraso promedio y no hay interacción entre variables explicativas de los diferentes niveles de la jerarquía. En cuanto a los supuestos del modelo, se cumple con el de independencia de la varianza del nivel 2 y con el de igualdad de varianza.

Cabe destacar, que en la investigación llevada a cabo por Torres y Torres en el año 2001, el análisis de varianza demostró que no son significativas las diferencias entre las facultades que conforman al núcleo Mérida, en cuanto al retraso promedio para ascender en el escalafón. Mientras que al realizar un estudio por niveles, los resultados indican que las diferencias en cuanto al retraso promedio para ascender entre las diversas escuelas que conforman al núcleo Mérida, sí son significativas.

REFERENCIAS BIBLIOGRÁFICAS

Alessio, M. (2007). *Análisis Estadístico de los Tiempos en los Ascensos del Profesor Universitario: Caso ULA.* Trabajo de grado no publicado. Mérida: Universidad de Los Andes.

Balzarini, M., Macchiavelli, R. y Casanoves, F. (2005). *Aplicaciones de Modelos Mixtos en Agricultura y Forestaría.* Costa Rica: Centro Agronómico Tropical de Investigación y Enseñanza.

Fernández P. (2000). *Sociología de los Grupos Escolares: Sociometría y Dinámica de Grupos*. Universidad de Almería.

Gelman, A. y Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models.* New York, Cambridge University Press.

Goldstein, H. (1999). *Multilevel Statistical Models*. London: Institute of Education.

González, A. (2007). *Evaluación del Programa Fray Juan Ramos de Lora con la Aplicación de Métodos Multivariantes*. Trabajo de grado de maestría no publicado. Mérida: Universidad de Los Andes.

Gujarati, D. (2004). *Econometría*. México: McGraw-Hill.

Hox, J. (1995). *Applied Multilevel Analysis*. Amsterdam: TT-Publikaties.

Hurtado, J. (2010). *El Proyecto de Investigación. Comprensión Holística de la Metodología y la Investigación.* Caracas: Ediciones Quirón.

Murillo, F. (2008). *Los Modelos Multinivel Como Herramienta Para La Investigación Educativa*. Magis, Revista Internacional de Investigación en Educación, Vol. 1, Nº 1, pp. 45-62. Bogóta-Colombia.

Prasad, S., Ramoni, J., Orlandoni, G., Torres, E. y Figueroa, M. (2007). *Conceptuación y análisis descriptivo del riesgo académico institucional en las universidades nacionales venezolanas: El caso de la ULA*. Educere, Vol.11, Nº 39, pp.653-663. ISSN 1316-4910.

Quintero, H. (2007). *Modelo Lineal Multinivel: Aplicación a Dos Problemas en el Campo de las Ciencias Sociales*. Trabajo de grado de maestría no publicado. Universidad de Los Andes. Mérida.

Ramoni, J., Orlandoni, G., Prasad, S. y Rivas, D. (2007). *El factor Capital Humano en la determinación de los sueldos de los profesores universitarios en Venezuela*. Análisis de Coyuntura, Vol. 13, Nº 2, pp.165-180. ISSN 1315-3617.

República de Venezuela. *Ley de Universidades (1970).* Gaceta Oficial de la República de Venezuela, 1.429 (Extraordinario).

SAS Institute Inc. (2004). *SAS/STAT® 9.1 User's Guide*. Cary, NC: SAS Institute Inc.

Singer, J. (1998). *Using SAS PROC MIXED to Fit Multilevel Models, Hierarchical Models, and Individual Growth Models*. Journal of Educational and Behavioral Statistics, Vol. 23, Nº. 4. pp. 323-355.

Torres, E y Torres, J. (1996). *Factores Estructurales e Individuales del Retraso de los Ascensos. La Situación de la Universidad de Los Andes.* Instituto de Investigaciones Económicas. Universidad de Los Andes. Extraído el 01 de Julio de 2011 desde http://www.saber.ula.ve/handle/123456789/17147

Torres, E y Torres, J. (2001). *Ascensos Retrasados: Un Fenómeno Extendido entre Profesores de una Universidad Venezolana*. FERMENTUM, AÑO 11, Nº 31, pp. 279-294, Mérida – Venezuela.

Universidad de Los Andes (1990). *Estatuto del Personal Docente y de Investigación de la Universidad de Los Andes*. Gaceta Universitaria, 1 (Extraordinario). Mérida, Venezuela.

APÉNDICES

APÉNDICE A

Rutinas en SAS

[Apéndice A.1]

[Descriptivos]

```
ods rtf
file="D:\Tesis\Resultados\Escuelas\Descriptivos.rtf";
ods graphics on;
proc import datafile="D:\Tesis\datos\Data.xls"
out=Data REPLACE;
sheet='Escuelas';
getnames=yes;
run;
proc means data=Data mean std n min max;
   var Retraso;
   class ESCUELAS;
run;
*** Create graphics for analysis variables ***;
goptions ftext=SIMPLEX ctext=BLACK htext=1 cells;
symbol1  v=plus c=blue w=1;
pattern1 v=SOLID;
proc boxplot data=Data;
   plot (RETRASO)*ID
                                               / caxis=GRAY
                                                 cframe=vligb ctext=BLACK
cboxes=dagr
                                                 cboxfill=GRAY idcolor=WHITE
npanel=31
                                                 boxstyle=SKELETAL
name='BOX'
                                                 cboxfill = ywh
                                                 description='Boxplot of
RERTASO by ESCUELAS';
run; quit;
ods graphics off;
ods rtf close;
```

[Apéndice A.2]

[Correlación entre el retraso y las variables del primer nivel]

```
ods rtf file="D:\Tesis\Resultados\Escuelas\Correlacion.rtf";
ods graphics on;
proc import datafile="D:\Tesis\datos\Data.xls"
out=Data replace;
sheet='Variables$';
getnames=yes;
run;
TITLE 'Correlacion entre en retraso y las variables del
primer nive';
PROC CORR DATA=Data PEARSON plots;
VAR Retraso;
WITH Categoria Dedicacion Contratado Graduado Maestria
Doctorado;
TITLE ' Correlacion entre el retraso y las variables del
primer nivel';
PROC CORR DATA=Data spearman plots;
VAR Retraso;
WITH Categoria Dedicacion Contratado Graduado Maestria
Doctorado;
ods graphics off;
ods rtf close;
```

[Apéndice A.3]

[Correlación entre el retraso y las variables del segundo nivel]

```
ods rtf
file="D:\Tesis\Resultados\Escuelas\Correlacion2.rtf";
ods graphics on;
proc import datafile="D:\Tesis\datos\Data.xls"
out=Data replace;
sheet='Variables$';
getnames=yes;
run;
TITLE 'Correlacion entre el retraso y las variables del
segundo nivel ';
PROC CORR DATA=Data PEARSON plots;
VAR Retraso;
WITH PRetrasados PTitulates PContratados PMaestria
PDoctorado;
ods graphics off;
ods rtf close;
```

[Apéndice A.4]

[Modelo incondicional de medias]

```
ods rtf
file="D:\Tesis\Resultados\Escuelas\ModeloNuloEscuelas.rtf";
title'Modelo Incondicional de Medias';
proc import datafile="D:\Tesis\datos\Data.xls"
out=Data replace;
sheet='Escuelas$';
getnames=yes;
run;
proc mixed;
 class ID;
 model Retraso = ;
 random ID;
 run;
proc mixed noclprint covtest;
 class ID;
 model Retraso = /solution;
 random intercept/sub=ID;
 run;
 title 'Regresión Simple';
 proc mixed noclprint covtest;
 class ID;
 model Retraso = /solution;
 repeated;
 run;
 ods rtf close;
```

[Apéndice A.5]

[Modelo incluyendo variables explicativas fijas en el primer nivel]

```
ods rtf file="D:\Tesis\Resultados\Escuelas\ModeloPaso2.rtf";
title'Modelo Incluyendo Variables Explicativas Fijas en el
Nivel Inferior';
proc import datafile="D:\Tesis\datos\Data.xls"
out=Data replace;
sheet='Variables$';
getnames=yes;
run;
proc mixed noclprint covtest;
class ID;
model Retraso = Graduado /solution ddfm=bw notest;
random intercept/sub=ID;
run;
title 'Regresión Simple';
proc mixed noclprint covtest;
class ID;
model Retraso = Doctorado  /solution;
repeated;
run;
ods rtf close;
```

[Apéndice A.6]

[Modelo incluyendo variables explicativas aleatorias en el primer nivel]

```
ods rtf file="D:\Tesis\Resultados\Escuelas\ModeloPaso3.rtf";
ods graphics on;
title'Modelo Incluyendo Variables Explicativas Aleatorias en
el Nivel Inferior';
proc import datafile="D:\Tesis\datos\Data.xls"
out=Data replace;
sheet='Variables$';
getnames=yes;
run;
title'Modelo Incluyendo Edad Aleatoria en el Nivel
Inferior';
proc mixed noclprint covtest method=REML COVTEST;
 class ID;
 model Retraso = Doctorado Edad/solution ddfm=bw notest;
 random intercept Edad /sub=ID  type=un;
 run;
 title'Modelo Incluyendo Categoría Aleatoria en el Nivel
Inferior';
 proc mixed noclprint covtest noitprint;
 class ID;
 model Retraso = Doctorado Categoria/solution ddfm=bw
notest;
 random intercept Categoria /sub=ID  type=un;
 run;
  title'Modelo Incluyendo Dedicación Aleatoria en el Nivel
Inferior';
  proc mixed noclprint covtest method=REML COVTEST;
 class ID;
 model Retraso = Doctorado Dedicacion/solution ddfm=bw
notest;
 random intercept Dedicacion /sub=ID  type=un;
 run;
  title'Modelo Incluyendo Contratado Aleatoria en el Nivel
Inferior';
   proc mixed noclprint covtest method=REML COVTEST;
 class ID;
 model Retraso = Doctorado Contratado/solution ddfm=bw
notest;
 random intercept Contratado /sub=ID  type=un;
 run;
 title'Modelo Incluyendo Graduado Aleatoria en el Nivel
Inferior';
   proc mixed noclprint covtest method=REML COVTEST;
 class ID;
 model Retraso = Doctorado Graduado/solution ddfm=bw notest;
 random intercept Graduado /sub=ID  type=un;
 run;
```

```
 title'Modelo Incluyendo Maestria Aleatoria en el Nivel
Inferior';
   proc mixed noclprint covtest method=REML COVTEST;
 class ID;
 model Retraso = Doctorado Maestria/solution ddfm=bw notest;
 random intercept  Maestria /sub=ID type=un;
 run;
 title'Modelo Incluyendo Doctorado Aleatoria en el Nivel
Inferior';
    proc mixed noclprint covtest method=REML COVTEST;
 class ID;
 model Retraso = Doctorado/solution ddfm=bw notest;
 random intercept Doctorado /sub=ID  type=un;
 run;
  title'MODELO DEFINITIVO PASO 3';
    proc mixed noclprint covtest method=REML COVTEST;
 class ID;
 model Retraso = Doctorado Dedicacion/solution ddfm=bw
notest;
 random intercept Dedicacion /sub=ID  type=un;
 run;
  ***Regresión lineal Simple***;
 title 'Regresión Simple';
  proc mixed noclprint covtest;
 class ID;
 model Retraso = Doctorado Dedicacion  /solution;
 repeated;
 run;
 ods graphics off;
 ods rtf close;
```

[Apéndice A.7]

[Modelo incluyendo variables explicativas en el segundo nivel]

```
ods rtf file="D:\Tesis\Resultados\Escuelas\ModeloPaso4.rtf";
ods graphics on;
title'Modelo Incluyendo Variables Explicativas en el Nivel
Superior';
proc import datafile="D:\Tesis\datos\Data.xls"
out=Data replace;
sheet='Variables$';
getnames=yes;
run;
proc mixed noclprint covtest method=REML COVTEST;
 class ID;
 model Retraso = Doctorado Dedicacion PRetrasados/solution
ddfm=bw notest;
 random intercept Dedicacion /sub=ID  type=un;
 run;
proc mixed noclprint covtest method=REML COVTEST;
 class ID;
 model Retraso = Doctorado Dedicacion PTitulares/solution
ddfm=bw notest;
 random intercept Dedicacion /sub=ID  type=un;
 run;
proc mixed noclprint covtest method=REML COVTEST;
 class ID;
 model Retraso = Doctorado Dedicacion PContratados /solution
ddfm=bw notest;
 random intercept Dedicacion /sub=ID  type=un;
 run;
proc mixed noclprint covtest method=REML COVTEST;
 class ID;
 model Retraso = Doctorado Dedicacion PMaestria /solution
ddfm=bw notest;
 random intercept Dedicacion /sub=ID  type=un;
 run;
proc mixed noclprint covtest method=REML COVTEST;
 class ID;
 model Retraso = Doctorado Dedicacion PDoctorado /solution
ddfm=bw notest;
 random intercept Dedicacion /sub=ID  type=un;
 run;
proc mixed noclprint covtest method=REML COVTEST;
 class ID;
 model Retraso = Doctorado Dedicacion PTitulares PDoctorado
/solution ddfm=bw notest;
 random intercept Dedicacion /sub=ID  type=un;
 run;
 ods graphics off;
 ods rtf close;
```

[Apéndice A.8]

[Chequeando los supuestos de normalidad e igualdad de varianzas]

```
ods rtf file="D:\Tesis\Resultados\Escuelas\Supuestos.rtf";
title'verificando los supuestos';
ods graphics on;
proc import datafile="D:\Tesis\datos\Data.xls"
out=Data replace;
sheet='Variables$';
getnames=yes;
run;
proc mixed noclprint covtest method=REML COVTEST;
 class ID;
 model Retraso = Doctorado Dedicacion/solution ddfm=bw
notest;
 random intercept Dedicacion /sub=ID  type=un;
 run;
   proc mixed;
      class ID;
      model Retraso = / residual;
      random ID ;
   run;
*Pruebas de normalidad de los residuos mediante 2
procedimientos Univariate y Capability*;
PROC UNIVARIATE DATA=Data NORMAL PLOT;
  VAR Retraso;
  QQPLOT Retraso /NORMAL(MU=EST SIGMA=EST COLOR=RED L=1);
RUN;
PROC CAPABILITY DATA=Data NORMAL;
  VAR Retraso;
  QQPLOT Retraso /NORMAL(MU=EST SIGMA=EST COLOR=RED L=1);
  PPPLOT Retraso /NORMAL(MU=EST SIGMA=EST COLOR=RED L=1);
HISTOGRAM /NORMAL(COLOR=MAROON W=4) CFILL  = BLUE CFRAME =
LIGR;
  INSET MEAN STD /CFILL=BLANK FORMAT=5.2 ;
RUN;
*Prueba levene y welch de homogeneidad de varianzas*;
title'prueba levene';
   proc glm data=Data;
      class ID;
      model Retraso = ID;
      means ID / hovtest welch;
   RUN;
 ods graphics off;
 ods rtf close;
```

APÉNDICE B

Tablas

[Apéndice B.1]

[Tabla χ^2]

k \ P	0,01	0,05	0,10	0,20	0,25	0,30	0,40	0,50	0,60	0,70	0,75	0,80	0,90	0,95	0,99
1	0,000	0,004	0,016	0,064	0,102	0,148	0,275	0,455	0,708	1,074	1,323	1,642	2,706	3,841	6,635
2	0,020	0,103	0,211	0,446	0,575	0,713	1,022	1,386	1,833	2,408	2,773	3,219	4,605	5,991	9,210
3	0,115	0,352	0,584	1,005	1,213	1,424	1,869	2,366	2,946	3,665	4,108	4,642	6,251	7,815	11,34
4	0,297	0,711	1,064	1,649	1,923	2,195	2,753	3,357	4,045	4,878	5,385	5,989	7,779	9,488	13,28
5	0,554	1,145	1,610	2,343	2,675	3,000	3,656	4,351	5,132	6,064	6,626	7,289	9,236	11,07	15,09
6	0,872	1,635	2,204	3,070	3,455	3,828	4,570	5,348	6,211	7,231	7,841	8,558	10,64	12,59	16,81
7	1,239	2,167	2,833	3,822	4,255	4,671	5,493	6,346	7,283	8,383	9,037	9,803	12,02	14,07	18,48
8	1,647	2,733	3,490	4,594	5,071	5,527	6,423	7,344	8,351	9,524	10,22	11,03	13,36	15,51	20,09
9	2,088	3,325	4,168	5,380	5,899	6,393	7,357	8,343	9,414	10,66	11,39	12,24	14,68	16,92	21,67
10	2,558	3,940	4,865	6,179	6,737	7,267	8,295	9,342	10,47	11,78	12,55	13,44	15,99	18,31	23,21
11	3,053	4,575	5,578	6,989	7,584	8,148	9,237	10,34	11,53	12,90	13,70	14,63	17,28	19,68	24,73
12	3,571	5,226	6,304	7,807	8,438	9,034	10,18	11,34	12,58	14,01	14,85	15,81	18,55	21,03	26,22
13	4,107	5,892	7,041	8,634	9,299	9,926	11,13	12,34	13,64	15,12	15,98	16,98	19,81	22,36	27,69
14	4,660	6,571	7,790	9,467	10,17	10,82	12,08	13,34	14,69	16,22	17,12	18,15	21,06	23,68	29,14
15	5,229	7,261	8,547	10,31	11,04	11,72	13,03	14,34	15,73	17,32	18,25	19,31	22,31	25,00	30,58
16	5,812	7,962	9,312	11,15	11,91	12,62	13,98	15,34	16,78	18,42	19,37	20,47	23,54	26,30	32,00
17	6,408	8,672	10,09	12,00	12,79	13,53	14,94	16,34	17,82	19,51	20,49	21,61	24,77	27,59	33,41
18	7,015	9,390	10,86	12,86	13,68	14,44	15,89	17,34	18,87	20,60	21,60	22,76	25,99	28,87	34,81
19	7,633	10,12	11,65	13,72	14,56	15,35	16,85	18,34	19,91	21,69	22,72	23,90	27,20	30,14	36,19
20	8,260	10,85	12,44	14,58	15,45	16,27	17,81	19,34	20,95	22,77	23,83	25,04	28,41	31,41	37,57
21	8,897	11,59	13,24	15,44	16,34	17,18	18,77	20,34	21,99	23,86	24,93	26,17	29,62	32,67	38,93
22	9,542	12,34	14,04	16,31	17,24	18,10	19,73	21,34	23,03	24,94	26,04	27,30	30,81	33,92	40,29
23	10,20	13,09	14,85	17,19	18,14	19,02	20,69	22,34	24,07	26,02	27,14	28,43	32,01	35,17	41,64
24	10,86	13,85	15,66	18,06	19,04	19,94	21,65	23,34	25,11	27,10	28,24	29,55	33,20	36,42	42,98
25	11,52	14,61	16,47	18,94	19,94	20,87	22,62	24,34	26,14	28,17	29,34	30,68	34,38	37,65	44,31
26	12,20	15,38	17,29	19,82	20,84	21,79	23,58	25,34	27,18	29,25	30,43	31,79	35,56	38,89	45,64
27	12,88	16,15	18,11	20,70	21,75	22,72	24,54	26,34	28,21	30,32	31,53	32,91	36,74	40,11	46,96
28	13,56	16,93	18,94	21,59	22,66	23,65	25,51	27,34	29,25	31,39	32,62	34,03	37,92	41,34	48,28
29	14,26	17,71	19,77	22,48	23,57	24,58	26,48	28,34	30,28	32,46	33,71	35,14	39,09	42,56	49,59
30	14,95	18,49	20,60	23,36	24,48	25,51	27,44	29,34	31,32	33,53	34,80	36,25	40,26	43,77	50,89
31	15,66	19,28	21,43	24,26	25,39	26,44	28,41	30,34	32,35	34,60	35,89	37,36	41,42	44,99	52,19
32	16,36	20,07	22,27	25,15	26,30	27,37	29,38	31,34	33,38	35,66	36,97	38,47	42,58	46,19	53,49