



INFORME FINAL

RAZIEL ACEVEDO ÁLVAREZ

2015



CONTENIDO

INFORME FINAL 2015.....	6
1. INFORMACIÓN GENERAL.....	6
1.1. No del proyecto: Código.....	6
1.2. Nombre del proyecto.....	6
1.3. Unidad base del investigador.....	6
1.4. Unidad de adscripción.....	6
1.5. Programa al que pertenece.....	6
1.6. Nombre de investigadores y carga académica asignada y quien asigna la misma.....	6
1.7. Vigencia del proyecto.....	6
1.8. Resumen.....	6
1.9. Descriptores.....	7
2. ANTECEDENTES.....	7
2.1. Surgimiento y antecedentes del proyecto.....	7
2.2. Planteamiento del problema.....	8
2.3. Objetivo General.....	9
2.4. Objetivos Específicos, Metas.....	9
3. REFERENTE TEÓRICO.....	10
4. METODOLOGÍA.....	10
4.1. TIPO DE INVESTIGACIÓN.....	10
4.2. MÉTODO.....	10
4.3. TÉCNICA UTILIZADAS.....	10
4.4. POBLACIÓN A LA QUE SE DIRIGE LA INVESTIGACIÓN.....	16
4.5. SELECCIÓN DE LA MUESTRA O PARTICIPANTES.....	17
4.6. PROCEDIMIENTO DE RECOLECCIÓN DE DATOS.....	17
4.7. ANÁLISIS DE LA INFORMACIÓN.....	17
5. ANÁLISIS Y DISCUSIÓN DE LOS RESULTADOS.....	17
6. CONCLUSIONES Y RECOMENDACIONES.....	17



6.1. PRODUCCIÓN ACADÉMICA.....	19
6.2. ACCIÓN SOCIAL.....	19
6.3. VINCULACIÓN CON POSTGRADO.....	20
6.4. VINCULACIÓN CON OTRAS INSTANCIAS.....	20
7. BIBLIOGRAFÍA.....	20
8. ANEXOS.....	21
ANEXO 8.1. LIBRO DE METODOLOGÍA DE LA INVESTIGACIÓN.....	21
CAPITULO 1: LOS MODELOS DE ECUACIONES ESTRUCTURALES.....	21
1.1. DE LOS MODELOS DE ECUACIONES ESTRUCTURALES.....	21
1.2. ESPECIFICACIÓN DEL MODELO.....	22
1.2.1. Del modelo de medida y el modelo estructural.....	22
1.2.2. Tipos de variables.....	23
1.2.3. Variables no observables y observables.....	23
1.2.4. Representación de las ecuaciones.....	25
1.2.5. Parámetros estructurales.....	27
1.3. IDENTIFICACIÓN DEL MODELO.....	29
1.3.1. Condicionantes para la identificación.....	31
1.4. ESTIMACIÓN DE PARÁMETROS.....	31
1.4.1. Tamaño de la muestra.....	33
1.4.2. Procesos de estimación.....	33
1.5. EVALUACIÓN Y REESPECIFICACIÓN.....	36
1.5.1. Observación del gráfico estimado y sus valores.....	36
1.5.2. Índices de ajuste del modelo.....	37
1.5.3. Síntesis para evaluar modelos.....	38
1.5.4. Otros índices de ajustes.....	42
1.5.5. Medidas de parsimonia.....	43
1.5.6. La muestra mínima en la función de discrepancia.....	44
1.5.7. Medidas basadas en la población de discrepancia.....	46
1.5.8. Medidas de información teórica.....	49
1.5.9. Comparación con un modelo de fondo.....	50



CAPITULO 2: EJEMPLO DE MODELOS DE ECUACIONES ESTRUCTURALES.....	56
2.6. ESPECIFICACIÓN del EJEMPLO.....	56
2.6.1. Ecuaciones del modelo de medida.....	57
2.6.2. Ecuación del modelo estructural.....	58
2.7. IDENTIFICACIÓN DEL EJEMPLO.....	59
2.8. ESTIMACIÓN DEL MODELO.....	60
2.9. EVALUACIÓN DEL EJEMPLO.....	60
2.9.1. Modelo N°1.....	60
CAPÍTULO 3: MODELOS JERÁRQUICOS LINEALES.....	66
3.1. METODOLOGÍA MULTINIVEL.....	69
3.1.1. El nacimiento.....	69
3.1.2. Posibles aplicaciones.....	70
3.1.3. Importancia de aplicación.....	71
3.1.4. Utilización de niveles.....	72
3.1.5. Sus nombres a través del tiempo.....	73
3.2. ECUACIÓN DEL MODELO JERÁRQUICO LINEAL.....	74
3.3. ESTIMACIÓN DEL MODELO BASE O NULO.....	77
3.4. LOS PARÁMETROS.....	78
3.5. TRES NIVELES.....	80
3.6. VARIANZA DEL MODELO.....	83
3.7. CORRELACIÓN INTRACLASE.....	84
CAPÍTULO 4: EJEMPLO DE MODELOS JERÁRQUICOS.....	84
4.1. RECOMENDACIONES INICIALES.....	84
4.2. MODELO NULO.....	88
4.3. VARIABLES PREDICTORAS DE PRIMER NIVEL.....	89
4.4. VARIABLES DE SEGUNDO NIVEL los docentes.....	92
4.5. RAZÓN DE MÁXIMA VEROSIMILITUD.....	94
5. CAPÍTULO 5: REFERENCIAS CITADAS Y CONSULTADAS.....	95
8.2. ANEXO 2.....	99
I. DESCRIPCIÓN DEL TALLER.....	99



II.	OBJETIVOS.....	100
III	TEMÁTICAS GENERALES.....	100
IV	METODOLOGÍA DE TRABAJO.....	101
V.	EVALUACIÓN.....	101
VI.	TEXTOS BÁSICOS.....	101
8.3.	ANEXO 3 Ponencia en Chile.....	106
TÍTULO: EN LA BÚSQUEDA DEL SESGO: EVALUACIÓN DOCENTE UNIVERSITARIA		
.....		106
Nombre completo: Razziel Acevedo Álvarez.....		106
Nacionalidad: Costarricense.....		106
Grado académico: Doctor en Educación.....		106
Institución de pertenencia: Universidad de Costa Rica, Sede de Guanacaste.....		106
Puesto de trabajo: profesor e investigador.....		106
Dirección electrónica de contacto: razziel.acevedo@ucr.ac.cr.....		106
I. EL punto de partida.....		106
Tabla 1 Características contextuales de cada nivel.....		111
Tabla 2: Resumen de las estimaciones del nivel Uno.....		114
Tabla 3: Estimaciones del nivel Uno.....		115
Tabla 4: Estimaciones del nivel Dos.....		116
Tabla 5: Estimaciones del nivel Dos.....		116
Tabla 6: Estimaciones del nivel Dos.....		116
Tabla 7: Modelo Final.....		117
V. BIBLIOGRAFÍA CITADA.....		121



INFORME FINAL 2015

1. INFORMACIÓN GENERAL

1.1. NO DEL PROYECTO: CÓDIGO

724-B1-147

1.2. NOMBRE DEL PROYECTO

Los métodos y técnicas cuantitativas para análisis de datos en la investigación educativa universitaria

1.3. UNIDAD BASE DEL INVESTIGADOR

Sede De Guanacaste

1.4. UNIDAD DE ADSCRIPCIÓN

I.N.I.E.

1.5. PROGRAMA AL QUE PERTENECE

Programa Cambio, Desarrollo y Gestión de la Educación Superior

1.6. NOMBRE DE INVESTIGADORES Y CARGA ACADÉMICA ASIGNADA Y QUIEN ASIGNA LA MISMA.

Raziel Acevedo Álvarez, 1/4

1.7. VIGENCIA DEL PROYECTO

01/1/2011 hasta 31/12/2013

Ampliación solicitada 01/12/2013 hasta 31/12/2014

Ampliación solicitada por enfermedad 30/06/2015

1.8. RESUMEN

El Proyecto trabajó con el objetivo de elaborar un texto sobre métodos y técnicas cuantitativas para el análisis de datos en la investigación universitaria y probarlo con estudiantes de grado, post grado e investigadores en general. Durante el inicio del proyecto se desarrollaron lecturas y la elaboración de una base de datos, para generar la teoría en relación al tema de los modelos jerárquicos lineales y los modelos de ecuaciones estructurales. Juntamente con este trabajo, se implementaron varios talleres de práctica, para poner a prueba los textos y prácticas elaborados, pues la idea general consistía en que los investigadores con formación y con poca formación en estadística pudieran comprenderlos. Además, se pudo realizar un taller internacional



sobre los Modelos Jerárquicos Lineales aplicados a la investigación educativa, con el Dr. José Luis Gaviria, de la Universidad Complutense de Madrid.

1.9. DESCRIPTORES

Estadística, investigación cuantitativa, modelos de ecuaciones estructurales, modelos jerárquicos lineales, MIWIN, SEM.

2. ANTECEDENTES

2.1. SURGIMIENTO Y ANTECEDENTES DEL PROYECTO

Desde el año 2006 hemos desarrollado en el I.N.I.E., dos proyectos de investigación cuantitativa alrededor de la competencia docente universitaria. El primero, Un modelo causal implicatorio de la competencia docente universitaria (Nº 720-A7-054) y el segundo, los factores de sesgo asociados a la evaluación docente universitaria (724-A9-118).

Dichos proyectos han fundamentado el análisis de sus datos, en metodologías estadísticas de avanzada: modelos de ecuaciones estructurales y modelos jerárquicos lineales, los cuales integran procesos de análisis muy complejos, con ecuaciones matemáticas que solamente pueden ser resueltas por con los programas informáticos actuales. Desde esa óptica, permiten al investigador una profundización mayor, que con las técnicas y métodos tradicionales, que a veces han aportado resultados errados o incrementado los efectos y el error, cuando en realidad era menos de lo esperado. Por esa razón, la necesidad imperiosa de estudiar los fenómenos de investigación cuantitativa, con metodologías emergentes que posibiliten mejores estimaciones y por ende, reduzcan el error.

Junto con esos proyectos investigación, se ha implementado año con año en el I.N.I.E. y en Liberia, el taller de Acción Social: Introducción al Manejo de Técnicas y Métodos Cuantitativos de Análisis en Investigación, Código: ED-2426, dirigido a investigadores, estudiantes, profesionales e instituciones que deseen conocer las metodologías emergentes de análisis de datos. En su desarrollo se han realizado más de seis talleres con la participación de profesionales del: MEP; CIP GALLUP, ITC, UCR, UNA, IIP, entre otras. Llegando a abarcar una setenta personas y la visita del Dr. José Luis Gaviria en el 2012, para impartir un taller internacional sobre el tema. El Dr. Gaviria es una reconocida autoridad en la metodología y es director del departamento de



Métodos de Investigación, Diagnóstico y Evaluación de la Universidad Complutense de Madrid.

A su vez, para incidir con mayor eficacia en el esfuerzo, el I.N.I.E. adquirió el paquete estadístico MIWin para mejorar el apoyo a la investigación cuantitativa, dotando de un paquete para futuras investigaciones.

Pero, a pesar de las investigaciones, de los talleres y del interés de los profesionales, por conocer y desarrollar las nuevas metodologías de análisis de datos cuantitativos, se hace palpable la carencia y necesidad de textos facilitadores, en lenguaje comprensible y práctico, que posibiliten el entendimiento de estos procesos, tanto a conocedores de la estadística, como aquellos que no tienen mucha relación con ella, pues para todos es conocida la importancia de los métodos de investigación y la importancia de los textos en el campo de la investigación.

En este sentido, el argumento anterior justifica la valía de este proyecto de investigación, el cual pretende inicialmente documentar sistemáticamente las metodologías emergentes de análisis cuantitativo de datos, para elaborar un libro didáctico, escrito en un lenguaje atractivo que permita llegar a comprender y aplicar los complejos procesos de estimación y análisis de datos en los modelos de ecuaciones estructurales y los modelos jerárquico lineales. De la misma forma, tratar de mostrar el camino para la toma de decisiones en cuanto a las actividades iniciales de los procesos de investigación, relacionados con: tipos variables, técnicas y metodologías a aplicar y otros relacionados con las labores de investigación.

Por otra parte, a pesar de que la investigación mantiene un ritmo de trabajo normal, se ha visto afectada por la cantidad de material bibliográfico seleccionado, la elaboración de bases y el análisis de las lecturas para la elaboración de los capítulos del texto: ello ha causado atraso en el cumplimiento de las metas y el cronograma propuesto. Además, la incapacidad física de seis meses en cama, incidió mucho en el atraso y presentación del informe final, pues después de la lesión, debido a lo lento de la recuperación, la investigación avanzó a un nivel muy lento y la prórroga solicitada desde el 2013, para todo el 2014 no fue aceptada y se redujo a la entrega en junio de este mismo año.

2.2. PLANTEAMIENTO DEL PROBLEMA

Durante varios años de desarrollo de la metodología en los diferentes talleres en el INIE y otros centros educativos, nos hemos dado cuenta que las personas temen a los procesos cuantitativos de análisis de datos, porque no pueden dominar el lenguaje



utilizado en los texto. Y en este sentido hemos tratado de buscar referencias bibliográficas que posibiliten a los investigadores un aprendizaje básico, que no requiera de una formación muy profunda, pues muchos de ellos laboran en instituciones y cuentan con poco tiempo para una capacitación profunda y en lenguaje complicado, por esa razón, les es difícil aplicar nuevas metodologías a sus problemas de investigación, de ahí la necesidad de resolver este problema con un texto en lenguaje básico, pero con la profundidad dedicada para entender estos procesos de análisis de datos.

2.3. OBJETIVO GENERAL

Indagar sobre la metodología de la investigación educativa, enfocada en los modelos de ecuaciones estructurales y los modelos jerárquicos lineales, a fin de profundizar sobre el tema y posibilitar su aplicación a docentes, investigadores y estudiantes.

2.4. OBJETIVOS ESPECÍFICOS, METAS

1. Identificar en las diversas fuentes de datos los diferentes procesos de estimación y análisis de las metodologías mencionadas, a fin de reunir información pertinente sobre el tema en estudio.

Metas: 1) Leer más de 40 artículos científicos sobre el tema; 2) Información sobre los cinco procesos de trabajo de la metodología modelos de ecuaciones estructurales; 3) Cuatro procesos de los modelos de jerárquicos lineales.

1. Elaborar un libro de texto sobre la metodología de investigación cuantitativa, a fin posibilitar su aplicación entre investigadores, estudiantes y profesionales.

Metas: a) Dos capítulos sobre los modelos de ecuaciones estructurales; b) Un capítulo de ejemplos sobre los modelos de ecuaciones estructurales; c) Dos capítulos sobre los modelos jerárquicos lineales; d) Un capítulo de ejemplos sobre los modelos jerárquicos lineales.

2. Elaborar un taller sobre la metodología de análisis cuantitativo, para utilizar y aplicar el libro de texto de apoyo.



Metas: a) Dos talleres sobre la metodología; b) Aplicación de un libro de texto; c) Aplicación en cuatro trabajos de investigación

3. REFERENTE TEÓRICO

De acuerdo con los evaluadores, este apartado no procede por qué el objetivo de esta investigación fue la elaboración de un libro. Sin embargo, vale la pena anotar que el mismo libro tiene todo un referente teórico que es el soporte del mismo libro. Por recomendación, este elemento tan importante se envió a los anexos, aunque desde nuestra perspectiva un libro no puede ser un anexo, dado que es el objetivo mismo de la investigación.

4. METODOLOGÍA

4.1. TIPO DE INVESTIGACIÓN

Este tipo de investigación es una herramienta que permite la tanto la investigación básica, como la aplicada, pues facilitan al investigador acercarse a procedimientos profundos para la estimación de los datos cuantitativos en la investigación.

4.2. MÉTODO

La investigación se fundamenta en el paradigma cuantitativo, para implementar dos tipos de metodología que permiten la estimación de datos, reduciendo el error de estimación, con muestras pequeñas y grandes. El primero, modelos de ecuaciones estructurales y el segundo, modelos jerárquicos lineales.

4.3. TÉCNICA UTILIZADAS

Los texto se desarrollaron utilizando los Modelos jerárquicos lineales y modelos de ecuaciones estructurales. La primera se fundamenta sobre la suposición que todas las relaciones de investigación o sus fenómenos, se encuentran anidados de un nivel micro a uno macro, por esa razón, los primeros caben dentro de los segundos y así sucesivamente. Esta propuesta permite, contextualizar los problemas de investigación y reducir el error de medida. La segunda, modelos de ecuaciones estructurales, busca



relaciones plausibles entre las variables, presupone que una variable causa genera un efecto sobre otra y que es efecto, tiene un peso determinado y un error de medida. De esa forma se puede considerar la existencia de un modelo específico. No obstante, ese modelo debe ser evaluado por una serie de índices de ajuste, para determinar la calidad del modelo metodológico utilizado.

Se ha realizado diferentes actividades para tratar de lograr el 100% de los objetivos y sus metas, para ello:

Respecto al primer objetivo específico:

1. Identificar en las diversas fuentes de datos los diferentes procesos de estimación y análisis de las metodologías mencionadas, a fin de reunir información pertinente sobre el tema en estudio.

Metas: 1) Leer más de 40 artículos científicos sobre el tema; 2) Información sobre los procesos de trabajo de la metodología modelos de ecuaciones estructurales; 3) Información sobre procesos de los modelos de jerárquicos lineales, 4) Elaboración de una base de datos en File Maker, para ordenar todo el proceso.

- a) Sobre la primera meta, las lecturas: se realizó una revisión profunda de la metodología, abarcando los principales autores que han desarrollado toda su investigación sobre estos temas, para contar con material preciso. Para su implementación, se trabajó en línea y en bibliotecas tradicionales, leyendo más de cuarenta artículos sobre el tema.
- b) Respecto de las metas dos y tres, que eran el objetivo fundamental del nuestro trabajo, reunir información sobre los modelos de ecuaciones estructurales y modelos jerárquicos lineales, ello contenido en la meta uno. Toda esta información era necesaria ordenarla, o sea la meta cuatro, por ello se construyó una base de datos: a fin de organizar digitalmente todos los materiales. Para ello fue necesario el paquete FILE MAKER. Con este se construyó una base que atiende específicamente: autor, bibliografía, año de publicación, indicadores del tema, aportación teórica. Una vez ordenados los materiales bibliográficos, en cuanto a temas, citas y referencias, se pudo iniciar el segundo. Una vez construida la búsqueda de elementos pertinentes para la redacción del texto se hizo más simple, pues con este paquete informático, se pueden hacer búsquedas muy específicas de acuerdo con criterios que el mismo investigador preestablece.



Sobre el segundo objetivo específico:

2. Elaborar un libro de texto sobre la metodología de investigación cuantitativa, a fin posibilitar su aplicación entre investigadores, estudiantes y profesionales.

Metas: a) Dos capítulos sobre los modelos de ecuaciones estructurales; b) Un capítulo de ejemplos sobre los modelos de ecuaciones estructurales; c) Dos capítulos sobre los modelos jerárquicos lineales; d) Un capítulo de ejemplos sobre los modelos jerárquicos lineales.

- a) Respecto de la meta número uno, se construyeron los capítulos de los modelos de ecuaciones estructurales, atendiendo a los cuatro procesos de trabajo sobre los que se implementan: especificación, identificación, estimación y evaluación. Este elemento solamente necesita el ejemplos gráficos que están en proceso de construcción. Nos habíamos dedicado específicamente a la teoría, a construir y ordenar todos los elementos teóricos que soportan ambos modelos los de ecuaciones y los jerárquicos.
- b) Sobre los capítulos de los modelos jerárquicos, se avanzaba a ritmo normal hasta que una lesión me incapacitó por cuatro meses. No obstante, el ritmo permitió abarcar el modelo nulo, parámetro fijo y aleatorio, los elementos principales sobre los que se erigen la jerarquía de los datos. Posterior a la lesión se pudo concluir el capítulo.
- c) Respecto de la meta dos y cuatro, la que era específicamente construir los ejemplos gráficos, nuevamente una lesión, pero de hombro con ruptura del manguito superior del brazo derecho hizo que otra vez la investigación se tuviera que parar, en primera instancia no supimos qué era, por ello decidimos parar de hacer los gráficos hasta que se redujera el dolor. No obstante, el dolor nunca se redujo, pero en esa espera se perdió mucho tiempo. Posterior de ir al médico y realizarme los análisis y radiografías correspondientes, nos enviaron terapia y con ella pudimos continuar la construcción de los dibujos, gráficos, ecuaciones matemáticas, y otros que dan sustento a este libro. La meta se alcanzó, pero con muchísimo esfuerzo y en un inicio, con mucho dolor.

Respecto del objetivo tres:

3. Elaborar un taller sobre la metodología de análisis cuantitativo, para utilizar y aplicar el libro de texto de apoyo.



Metas: a) Dos talleres sobre la metodología; b) Aplicación de un libro de texto; c) Aplicación en cuatro trabajos de investigación

Se han implementado diversos talleres, atendiendo las necesidades del investigador, la cual era que otros investigadores leyeran y encontraran errores en los textos, o elementos que no entendieran. También se atendió a las necesidades de los investigadores, estudiantes de grado, estudiantes doctorales, empresa privada, gobierno, etc, pues muchos de ellos tienen problemas de investigación sin alcanzar los procesos metodológicos para concluirla, y el taller les suministró los elementos metodológicos para profundizar en su investigación. Estos han utilizando los materiales trabajados producidos por esta investigación, como eje sobre el cual se desarrollaron todos los pasos del taller. Vale la pena mencionar, que siempre hubo retroalimentación respecto del texto con los participantes del taller, dado que la construcción textual fue simultánea al taller.

Un caso de ejemplo, fue de una estudiante de la Maestría en Psicología, Hellen Picado Barrantes, quién nos solicitó ser lector, de su trabajo llamado Bienestar psicológico, motivación intrínseca, conciencia plena y factores asociados en Docentes e Investigadores Universitarios: un análisis basado en el modelo jerárquico lineal. Para esta investigadora el taller y los materiales redactados en la elaboración del libro, fueron fundamentales.

Además, en el 2012, se desarrolló el Taller Internacional sobre los Modelos Jerárquicos Lineales, iniciando en el mes de mayo, hasta concluirlo en Julio con la visita del Dr. José Luis Gaviria, quién es el director del Departamentos de Métodos de Investigación de la Universidad Complutense de Madrid. Participaron investigadores, estudiantes de grado y post grado, así como profesionales del TEC, IIP, UNA, MEP, CIP GALLUP, entre otros. El detalle se encuentra en el anexo 2. Este investigador también participó con la lectura del texto que se estaba desarrollando, y sus observaciones sobre los modelos jerárquicos lineales fueron precisas y de gran ayuda.

A su vez, parte de este material sirvió de fundamento para la disertación en el II congreso de investigación educativa, 2011, con la ponencia: los modelos jerárquicos lineales en la investigación educativa Universitaria, 2011.

Ello se puede ver en:

Acevedo, Raziél (2011). Modelos jerárquicos lineales y su aplicación en la investigación educativa universitaria. II Congreso Internacional de Investigación Educativa. En línea, recuperado el 4 de febrero de 2012, de:



<http://congreso.inie.ucr.ac.cr/memoria/documentos/3/MODELOS%20JERARQUICOS%20LINEALES%20Y%20SU%20APLICACION%20EN%20LA%20INVESTIGACION%20EDUCATIVA%20UNIVERSITARIA.PDF>

Acevedo, Raziel (2011). Introducción al manejo de los modelos jerárquicos lineales. II Congreso Internacional de Investigación Educativa. En línea, recuperado el 4 de febrero de 2012, de: <http://congreso.inie.ucr.ac.cr/memoria/documentos/8/Introduccion%20al%20manejo%20de%20los%20modelos%20jerarquicos%20lineales-Raziel%20Acevedo.pdf>

También se participó en una ponencia en 2014, en Santiago de Chile, en el congreso de la Sociedad Española de Pedagogía, en la Línea temática: III Avances en metodologías, diseños, técnicas y análisis de datos para la investigación, diagnóstico y evaluación educativa, con el título En la búsqueda del sesgo: Evaluación Docente Universitaria. La ponencia puede verse en el anexo tres.



Sobre la metodología de análisis cuantitativo de datos y el objetivo tres del taller, se logró avanzar más allá de las expectativas, pues la misma necesidad de los investigadores (as), ha puesto en evidencia la importancia de este objetivo y hemos tenido que realizar más talleres de los planificados, pues siempre existen solicitudes de tesarios, doctorandos y otros investigadores, quienes necesitan apoyo sobre los temas de estudio de esta investigación.

Una síntesis de la actividad y logro de toda la actividad de investigación puede verse en el siguiente cuadro:



Tabla N°2 Actividades realizadas, logros y limitaciones

OBJETIVOS ESPECIFICOS Y METAS	% LOGRO	ACTIVIDADES DESARROLLADAS	AVANCE DE LAS ACTIVIDADES (R,A, SC)*	DIFICULTADES Y FORMAS DE RESOLVERLAS	ACTIVIDADES PENDIENTES
1.- Identificar en las diversas fuentes de datos	100 %	Lecturas, elaboración de base de datos, redacción	Concluido		
Meta 1: Leer más de 40 artículos	100 %	Búsqueda on line, bases de datos	Concluido	Ninguna	Ninguna s
Meta 2: Cinco procesos metodología SEM	100 %	Organización de la teoría sobre SEM, redacción de texto sobre la utilización de la metodología en investigación	Concluido	Ninguna	Ninguna
Meta 3: Cuatro procesos de los modelos jerárquicos lineales	100 %	Organización de los procesos relacionados con modelo nulo, parámetro fijo y aleatorio	Concluido	Incapacitado por lesión de tendón durante cuatro meses, los cuales pasé en cama e imposibilitado de moverme.	Ninguna
2.- Elaborar un libro de texto sobre la metodología de investigación cuantitativa	100%	Elaboración de capítulos y ejemplos sobre la metodología de la investigación. Construcción de índice	Concluido	Enfermedad no hubo forma de trabajar	Ninguna
Meta 1: Dos capítulos SEM	100 %	Redacción texto, ordenamiento de ideas	Concluido	Ninguna	Ninguna
Meta 2: un capítulo de ejemplos SEM	100%	Construcción de ejemplos con el taller impartido	Concluido	Lesión de hombro	Ninguna
Meta 3: Dos capítulos Modelos Jerárquicos lineales	100 %	Redacción de un capítulo	Concluido	Enfermedad	Ninguna
Meta 4: un capítulo de ejemplos MJL	100 %	Más ejemplos	Concluido	Lesión de hombro	Ninguna



OBJETIVOS ESPECIFICOS Y METAS	% LOGRO	ACTIVIDADES DESARROLLADAS	AVANCE DE LAS ACTIVIDADES (R,A, SC)*	DIFICULTADES Y FORMAS DE RESOLVERLAS	ACTIVIDADES PENDIENTES
3.- Taller sobre la metodología de análisis cuantitativo	100 %	Taller sobre SEM y Modelos Jerárquicos lineales. Los primeros se desarrollaron en el 2011 y los segundos en el 2012. Coordinación con OAI y Rectoría para trasladar a profesor al país. INIE realizó la divulgación del evento y la gestión de los mismos	Concluida	Ninguna	Ninguna
Meta 1: Dos talleres	100 %	Organización de los talleres, distribución de materiales, matrícula, divulgación, solución a problemas de los investigadores participantes	Concluida	Ninguna	Ninguna
Meta 2: Aplicación de un libro de texto	100 %	Se utilizaron como material de trabajo los textos de los modelos de ecuaciones estructurales y los modelos jerárquicos lineales.	Concluida	Enfermedad en el segundo semestre	Ninguna
Meta 3: Aplicación en cuatro trabajos	100%	Una tesis de grado los está empleando. Funcionarios de la UCR, CIP GALLUP, llegaron a construir un SEM. Investigadores del MEP, aplicaron un Modelo Jerárquico.	Concluida	Ninguna, el trabajo se desarrolló según lo planeado, más bien hubo muchas aplicaciones en las investigaciones de los participantes	Ninguna

4.4. POBLACIÓN A LA QUE SE DIRIGE LA INVESTIGACIÓN

La investigación está dirigida a los diferentes profesionales, investigadores, estudiantes de post grado, evaluadores y Centros de Investigación, Empresas Evaluadoras como: CIPGALLUP, INTEL, entre otros. Debido a que son pocos los materiales existentes con lenguaje comprensible, accesible al manejo de estos paquetes estadísticos relacionados con los modelos de ecuaciones estructurales y modelos jerárquicos lineales.



4.5. SELECCIÓN DE LA MUESTRA O PARTICIPANTES

No aplica la selección de muestra.

4.6. PROCEDIMIENTO DE RECOLECCIÓN DE DATOS

No hubo recolección de datos empíricos, por ello no se cuenta con una muestra significativa. Únicamente recopilación de materiales por medio de: Bases de datos, textos y artículos en línea, los cuales fueron necesarios para la organización de una base de datos sobre las temáticas estudiadas.

4.7. ANÁLISIS DE LA INFORMACIÓN

No aplica.

5. ANÁLISIS Y DISCUSIÓN DE LOS RESULTADOS

Según el evaluar no aplica

6. CONCLUSIONES Y RECOMENDACIONES

Fundamentados en el primer objetivo: Identificar en las diversas fuentes bibliográficas de datos, se ha podido obtener la identificación de fuentes de datos muy valiosas, con los principales autores sobre los temas seleccionados. Con estos escritos se han podido separar una gran cantidad de temas pertinentes, que giran alrededor de la metodología y se han extraído citas importantes, sobre temas neurálgicos. Este objetivo se cumplió totalmente.

Respecto del segundo objetivo, elaborar un libro de texto sobre la metodología de investigación cuantitativa, se ha de mencionar que los modelos de ecuaciones estructurales y los modelos jerárquicos lineales, fueron logrados totalmente en su componente teórico y con los ejemplos. Tuvimos atraso por mi lesión de tendón de



Aquiles que me mantuvo cuatro meses fuera y la recuperación ha sido difícil. Posterior a este problema en 2014, sufrimos la ruptura del tendón del brazo derecho, llamado manguito rotador. Por ello no se pudo lograr un ritmo de trabajo constante, pues la lesión fue dolorosa en todo momento, hasta que entendimos que no era un problema pequeño y decidimos acudir al médico. No obstante, el objetivo se concluyó totalmente.

Respecto del tercero, Taller sobre la metodología de análisis cuantitativo, este ha logrado avanzar más allá de las expectativas, pues la misma necesidad de los investigadores (as), ha puesto en evidencia la importancia de este objetivo y hemos tenido que realizar más talleres de los planificados, pues siempre existen solicitudes de tesarios, doctorandos y otros investigadores, quienes necesitan apoyo sobre los temas de estudio de esta investigación. Este objetivo permitió que tanto la empresa privada SIP-GALLUP como Hellen Picado, lograran finalizar sus problemas de investigación. La primera utilizando los modelos de ecuaciones estructurales, para un problema de estimación en la proyección de votantes por cantón. La Segunda, concluyó su tesis fundamentada en los modelos jerárquicos lineales, llegando a los talleres, utilizando nuestro texto y sobre todo, apoyándola como lector de su trabajo de grado, dado que ninguno de los otros lectores conocía la metodología de investigación.

Respecto de los talleres vale la pena mencionar dos elementos: en primer lugar el apoyo dado por el Dr. Gaviria, de la Universidad Complutense, quién participó en la investigación no solo impartiendo el taller, sino también, leyendo los materiales teóricos producidos. Su aporte fue significativo, dada su alta experiencia y capacitación sobre el tema, unido a su interés de construir un mejor texto para quienes enfrentan problemas de estimación. En segundo lugar, los talleres fueron de alimentación recíproca, pues cada vez que se llevó material escrito por este servidor, los y las participantes llevaron a sus casas para aportar sus observaciones sobre: redacción, ecuaciones, gráficos, estimaciones y formas de entender el texto, que en muchos casos hubo que cambiar totalmente el formato mostrado. Esta retroalimentación, fue muy lenta, porque muchos elementos se volvieron a construir más de cuatro o cinco veces.

Es importante anotar que fue un error tomar dos metodologías para hacer un libro, pues cada una de ellas es enorme y su contenido teórico es muy profundo, lo cual no es bueno mezclar en un texto. Máxime cuando se proponen dos años para concluirlo, teniendo tiempo completo de labores en la Universidad. Muchas veces el entusiasmo de este investigador, por aportar una síntesis tan necesaria sobre la metodología, se vio desbordado por la realidad laboral y por la profundidad de ambas metodologías, en una situación laboral en la que se debe de cumplir en tiempo completo en actividades docentes. Lo más razonable era dedicar dos años a una



metodología y dos años a la otra, para no tener la angustia de concluir en el límite señalado, aunado a las enfermedades que llegan cuando menos los investigadores lo imaginamos.

6.1. PRODUCCIÓN ACADÉMICA

Acevedo, Raziél (2014). *En la búsqueda del sesgo en la Evaluación Docente Universitaria*. En prensa: <http://congresodepedagogia.ucsh.cl/congreso.html>

Acevedo, Raziél (2011). Modelos jerárquicos lineales y su aplicación en la investigación educativa universitaria. II Congreso Internacional de Investigación Educativa. En línea, recuperado el 4 de febrero de 2012, de: <http://congreso.inie.ucr.ac.cr/memoria/documentos/3/MODELOS%20JERARQUICOS%20LINEALES%20Y%20SU%20APLICACION%20EN%20LA%20INVESTIGACION%20EDUCATIVA%20UNIVERSITARIA.PDF>

Acevedo, Raziél (2011). Introducción al manejo de los modelos jerárquicos lineales. II Congreso Internacional de Investigación Educativa. En línea, recuperado el 4 de febrero de 2012, de: <http://congreso.inie.ucr.ac.cr/memoria/documentos/8/Introduccion%20al%20manejo%20de%20los%20modelos%20jerarquicos%20lineales-Raziél%20Acevedo.pdf>

6.2. ACCIÓN SOCIAL

Esta investigación contiene un objetivo relacionado totalmente con Acción Social, dado que los componentes estudiados son muy necesarios para los investigadores, estudiantes y personas relacionadas con los métodos cuantitativos de análisis de datos. Por ello, la difusión por medio de talleres, les permite a ellos contar con material muy valioso y difícil de conseguir. Por esa razón, en el laboratorio del INIE, se han implementado gran cantidad de talleres, sobre los temas estudiados en esta investigación.

Esta actividad se encuentra inscrita con el nombre: introducción al manejo de técnicas y métodos cuantitativos de análisis de datos en la investigación Código: ED-2426. Ella ha permitido la difusión de la metodología, pero sobre todo los participantes son revisores y sujetos de práctica del texto a desarrollar.



6.3. VINCULACIÓN CON POSTGRADO

Por ser de componente metodológico, esta investigación está muy relacionada con el doctorado en Educación de nuestra universidad, en el cual hemos brindado diferentes talleres sobre esta metodología, a los estudiantes doctorales. Así mismo, en la Maestría en Psicología, también se han implementado los mismos.

Además, en este momento somos lectores de la Tesis de maestría de Hellen Picado Barrantes, lector, llamada Bienestar psicológico, motivación intrínseca, conciencia plena y factores asociados en Docentes e Investigadores Universitarios: un análisis basado en el modelo jerárquico lineal.

También Berman Siles al igual que Pietro Scaglioni, utilizan los modelos de ecuaciones estructurales, aprendidos en nuestro taller, para desarrollar sus tesis doctorales. El primero, trata sobre el Liderazgo Organizacional y el segundo sobre la Educación Física en la Tercera Edad.

6.4. VINCULACIÓN CON OTRAS INSTANCIAS

Esta investigación se encuentra muy vinculada con la Universidad Complutense de Madrid, Departamento de Métodos de Investigación, Evaluación y Diagnóstico, quienes por su componente metodológico, han invitado a este investigador para que desarrolle el proyecto en Europa, brindando todas las facilidades de estudio.

Por otra parte, diferentes instancias como. MEP, CIP GALLUP, TEC, UNA, entre otras organizaciones, han participado de los talleres y han mostrado su interés en que se les suministre, otros métodos de análisis a sus problemas cuantitativos, pues el grado de profundización y el lenguaje de los materiales, les ha facilitado la resolución de los problemas de estudio.

7. BIBLIOGRAFÍA

Por recomendación del evaluador se incluye en el anexo del libro, como producto de la investigación



8. ANEXOS

ANEXO 8.1. LIBRO DE METODOLOGÍA DE LA INVESTIGACIÓN

CAPITULO 1: LOS MODELOS DE ECUACIONES ESTRUCTURALES

Los modelos de ecuaciones estructurales son el resultado de la evolución y unión de la metodología desarrollada en el análisis de senderos y el análisis factorial, la cual comprende un amplio rango de métodos de análisis multivariante, incluyendo regresión, análisis factorial y análisis de varianza. Con ésta amalgama de técnicas de medida, se puede entre otras cosas elaborar estructuras parsimoniosas, aumentar la capacidad explicativa de un instrumento, confirmar la eficacia estadística y, obviamente, validar un constructo teórico tal y como el investigador lo plantea. Estructuralmente se erige sobre cuatro pasos: especificación, identificación, estimación de parámetros y evaluación del modelo. Este proceso permite pasar de una teoría expresada verbalmente a una formulada matemáticamente. Por esa razón la estructura teoría ha de ser sólida, para fundamentar las relaciones causales entre las variables, porque de no ser así, cualquier análisis realizado carece de validez.

Palabras claves:

Modelos causales, modelos de ecuaciones estructurales, análisis estadístico, path análisis, LISREL, AMOS, EQS.

1.1. DE LOS MODELOS DE ECUACIONES ESTRUCTURALES

Los modelos de ecuaciones estructurales son una metodología del paradigma cuantitativo, fundamentada en una teoría o profunda revisión teórica, para determinar las relaciones causales o el grado de asociación entre las variables de investigación. Su utilización permite elaborar modelos causales que sean explicaciones plausibles de los fenómenos sociales, con el objetivo de validarlos o rechazarlos empíricamente. No pretende bajo ninguna directriz, crear una doctrina o teoría, pues es un proceso secuencial que tiene como punto de partida la teoría misma sobre un tema de estudio. A partir de ella, se formulan las correspondencias causales, mediante un gráfico o matricialmente o escribiendo un sistema de ecuaciones simultáneas; por tanto, no es un método para descubrir causas, sino, para validar las relaciones hipotéticas entre las



variables basándonos en unas hipótesis previas. Esto supone esquematizar la realidad seleccionando las variables relevantes que mejor explican las relaciones entre éstas. Por tanto, si el modelo no está bien formulado o fundamentado sobre una teoría específica, todo el análisis realizado carece de validez.

Estos modelos están compuestos por un sistema de ecuaciones estructurales que no sólo representan una relación matemática, sino, como argumenta Page y otros (1990), también reflejan una relación teórica causa-efecto, puntualizados en función de la teoría subyacente. En este sentido, su objetivo consiste en describir de manera profunda, las posibles relaciones causales entre una serie de variables propuestas o hipotéticas, con el fin de explicar un problema o una realidad. Bizquerra (1989) afirma que el objetivo mismo es pasar de una teoría formulada verbalmente a un modelo gráfico expresado matemáticamente, sería entonces convertir las palabras de la teoría en procesos de estimación estadística.

Ahora bien, para implementar los modelos de ecuaciones estructurales es necesario llevar a cabo un proceso ordenado de cuatro, cinco u ocho pasos. Los autores no se ponen de acuerdo en su cantidad, no obstante, a partir de nuestra experiencia se pueden considerar básicamente cuatro pasos fundamentales: especificación, identificación, estimación de parámetros y evaluación. Estos involucran de alguna u otra forma todos los demás planteamientos propuestos por los teóricos.

1.2. ESPECIFICACIÓN DEL MODELO

La especificación del modelo se refiere al planteamiento y elaboración de una representación gráfica fundamentado en una teoría previa. En este gráfico o modelo se hipotetiza la realidad, estableciendo las relaciones causales entre las variables o factores que lo integran; con ello se esquematiza la propuesta teórica mediante un diagrama de paso, a partir del cual, se formulan las ecuaciones estructurales que van a describir las posibles relaciones entre las variables. Esta fase supone, sintetizar la realidad seleccionando cuidadosamente aquellas variables relevantes que mejor expliquen el fenómeno de estudio, formulando las hipótesis que van a servir de base al modelo estructural.

Esencialmente, este paso consiste en definir a través de qué vías, determinadas variables explican a otras, mediante la asociación entre ellas.

1.2.1. Del modelo de medida y el modelo estructural

Una propuesta integra dos tipos de modelos: el modelo de medida y el modelo estructural. El primero, define la relación entre la variable latente y sus variables observables. El segundo, mide las relaciones dadas entre todas las variables latentes



del modelo. Es decir, se especifica qué variables latentes influyen directa o indirectamente sobre otras.

Respecto del modelo general, la mayoría de los teóricos¹ lo consideran definido formalmente por tres ecuaciones, a saber:

Modelo de ecuación estructural:

$$\eta = \beta\eta + \Gamma\xi + \zeta$$

Modelo de medida para “X” (variable observable independiente o exógena)

$$X = \Lambda_x\xi + \delta$$

Modelo de medida para “Y” (variable observable dependiente o endógena):

$$Y = \Lambda_y\eta + \varepsilon$$

Con estas ecuaciones básicas se pueden estimar todos los modelos planteados por la teoría respectiva. De hecho, la primera de las ecuaciones corresponde a las relaciones entre las variables exógenas y endógenas. La segunda representa el modelo de medida de las variables exógenas y la tercera las endógenas.

1.2.2. Tipos de variables

Este análisis utiliza variables exógenas y endógenas. Las primeras son aquellas que no pretenden ser explicadas causalmente, pudiendo considerarse constructos hipotéticos, e influyen directa o indirectamente en las segundas. Las *endógenas* son variables dependientes que vienen explicadas por las otras variables del modelo. De hecho, su variabilidad es atribuida a causas externas a la misma variable.

1.2.3. Variables no observables y observables

En el modelo de estructuras de covarianza parte de las variables no observables², que son un constructo o concepto no medible por sí mismo, pues representa un pensamiento abstracto el cual se compone de diversos elementos. En otras palabras, las variables no observables son medibles solamente a través de otras.

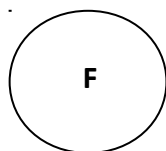
Para tal efecto, se hacen necesarias las variables observables que sí son cuantificables y se permite incluir tantas como sean necesarias para identificar un constructo, dado que no existe un límite en cuanto a la cantidad permitida.

¹ Ver por ejemplo: Darden, 1981; Dolan, 1996; Duncan, 1975; Hayduk, 1996; Jöreskog, 1969, 1970; Kenny, 1979; Werts y Linn, 1970

² Llámese también latentes o factoriales o constructos

La representación gráfica de las variables no observables es un círculo.

Figura 1 Variable no observable

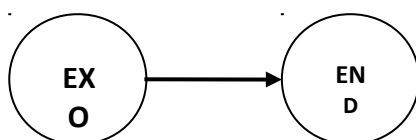


Por otra parte, las variables observables³ son materializadas por un rectángulo que la identifica⁴. Ésta es cuantificable en unidades y puede tratarse como una variable objetiva, por ejemplo la edad, el peso, la nota en matemática. Así mismo, puede verse como una variable subjetiva que mida las percepciones, las actitudes o los comportamientos de los estudiantes.

Figura 2: Variables observable \boxed{x} o \boxed{y}

La relación entre variables es simbolizada con flechas. Por esa razón, una flecha entre variable latente y otra similar, sea endógena o exógena, representa una relación causal directa. Pero, cuando en esa relación media otra variable interviniente, la asociación se considera indirecta, por no existir una relación gráfica directa. Como se ha señalado, el grado de asociación entre variables es definido por esta relación y por los residuos de la endógena. La suma de los efectos directos y de los indirectos se denomina efecto total o coeficiente de efecto, definiéndose como el efecto íntegro de una variable (causa) sobre otra (efecto).

Figura 3: Relación entre variables latentes

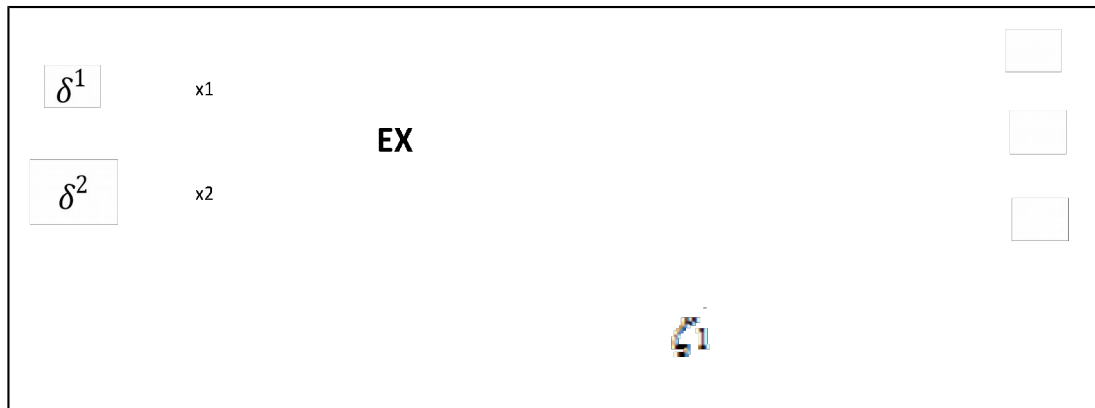


³ Existen variables observables independientes o sea exógenas y variables observables dependientes o endógenas. Las variables observables pueden también ser denominadas variables manifiestas por oposición a las variables latentes.

⁴ Por su relación con las ecuaciones, las ecuaciones definen la “x” para la variable observable de una variable exógena y la “y” para identificar las observables de la variable endógena. Los nuevos paquetes informáticos posibilitan ponerles un código o un nombre específico.

Una flecha direccionada partiendo de una variable no observable a la observable, representa una regresión y generalmente los paquetes estadísticos sobre el tema muestran ese valor muy cerca de la flecha.

Figura 4: Relación entre variables latentes y observables



Como se mencionó anteriormente, existirán cuantas variables observables se necesiten para identificar una no observable. Aún más, cada variable sea latente endógena u observable, tiene asignado un error de medida en el diagrama, el cual permite mejorar la medida de nuestras variables.

1.2.4. Representación de las ecuaciones

Las ecuaciones estructurales responden directamente a la organización del modelo y sus parámetros. Éste tendrá tantas, como parámetros contemple la propuesta. Si representamos la figura anterior, quedaría:

Modelo Estructural

$$\eta_1 = \gamma_{11}\xi_1 + \zeta_1$$

La ecuación expresa: La variable latente endógena (η), es igual al efecto de la exógena (ξ) sobre la endógena (η), simbolizado en gamma (γ). Por la variable exógena, más el error dseta (ζ) de la variable latente endógena (η). Este último es la parte de la variable latente endógena (η) no explicada por la variable exógena (ξ).

Modelo de medida:

$$x_1 = \lambda_{11}\xi_1 + \delta_1$$



$$x_2 = \lambda_{12}\xi_2 + \delta_2$$

El conjunto de ecuaciones se interpreta: La variable de medida (observada) X , es igual al peso de la variable latente exógena (ξ), sobre la observada X , representado por λ (λ). Por la variable latente exógena, más el error δ de X ; que representa lo específico de la variable de medida y los errores aleatorios.

$$y_1 = \lambda_{11}\eta_1 + \varepsilon_1$$

$$y_2 = \lambda_{12}\eta_1 + \varepsilon_2$$

$$y_3 = \lambda_{13}\eta_1 + \varepsilon_3$$

La interpretación correspondería: la variable de medida (observada) Y , es igual al peso de la variable latente exógena (η), sobre la observada Y , representado por λ (λ). Por la variable latente endógena, más el error ε (ε) de Y ; que representa lo específico de la variable de medida y los errores aleatorios.

Matricialmente este análisis factorial se podría presentar de la siguiente forma:

CARGAS FACTORIALES Y PARÁMETROS		
	ξ_1	η_1
x_1	X	
x_2	X	
y_1		X
y_2		X
y_3		X
η_1	X	

ERRORES DE MEDIDA Y DE REGRESIÓN						
	X1	X2	Y1	Y2	Y3	η_1
X1	δ_1					
X2		δ_2				
Y1			ε_1			
Y2				ε_2		
Y3					ε_3	
η_1						ζ_1

1.2.5. Parámetros estructurales

Los elementos que establecen la correspondencia entre la teoría, por un lado, y su expresión formal gráfica, por el otro, se denominan *parámetros estructurales*: a cada relación le corresponde un parámetro específico de identificación. En el paquete LISREL los parámetros tienen sus nombres en griego, aunque en los otros paquetes esta práctica ha caído en desuso. En la tabla N° 1, se observan estas relaciones entre variables y su referente en griego.

TABLA N° 1 PARÁMETROS Y MATRICES

DESDE	HASTA	PARÁMETRO	MATRIZ
ξ (xi)	ξ (xi)	ϕ (phi)	Φ (phi)
ξ (xi)	η (eta)	γ (gamma)	Γ (gamma)
η (eta)	η (eta)	β (beta)	β (beta)
ξ (xi)	χ	λ (lambda)	Λ (lambda)
η (eta)	Y	λ (lambda)	Λ (lambda)
<i>Error de medida en la variable endógenas</i>		ζ (dseta)	Ψ (psi)
<i>Error de medida en las variables observadas exógenas</i>		δ (delta)	θ (zeta-delta)
<i>Error de medida en las variables endógenas observables</i>		ε (épsilon)	θ (zeta-delta)
ζ (dseta)	ζ (dseta)	ψ (psi)	Ψ (psi)
X	χ	δ (delta)	θ (zeta-delta)
Y	Y	ε (épsilon)	θ (theta-épsilon)



Para Jöreskog (1991, 1993) y Page (1990) estos modelos cuentan con ocho matrices de parámetros para realizar las estimaciones. En algunos casos se señala de antemano cuáles son los parámetros libres y cuáles los fijos. Este último, tiene que ver con restricciones impuestas sobre el modelo y con el problema de la identificación.

Las ocho matrices de parámetros son:

1. Lambda “Y” (Λ_y)

Es una matriz de orden $p \times m$, donde p expresa el número de variables endógenas observables y m el número de variables endógenas latentes. Esta indica el peso de la variable endógena latente sobre la observable.

2. Lambda “X” (Λ_x)

Es una matriz de orden $q \times n$, donde q expresa el número de variables exógenas observables y n el número de variables exógenas. Este indica el peso directo de la variable exógena sobre la observable.

3. Beta (β)

Es una matriz cuadrada de orden $m \times m$, donde m es el número de variables endógenas latentes. Este indica el efecto directo de la variable endógena sobre otra de la misma categoría.

4. Gamma (Γ)

Es una matriz de orden $m \times n$, donde m es el número de variables latentes endógenas observables y n el número de variables latentes exógenas. Este indica el efecto directo de la variable exógena sobre la variable latente endógena.

5. Phi (Φ)

Es una matriz cuadrada de orden $n \times n$, donde n expresa el número de variables latentes exógenas. Indica la covarianza entre las variables latentes exógenas.

6. Psi (Ψ)

Es una matriz de orden $m \times m$, donde m expresa el número de residuales de las variables latentes endógenas. Este elemento (Ψ_{ij}) indica la covarianza entre los términos de perturbación o residuales (ζ_i) y (ζ_j), mientras el elemento (Ψ_{ii}) indica la varianza del i -ésimo término de perturbación de la variable endógena latente.

7. Theta - Epsilon ($\Theta\varepsilon$)



Es una matriz de orden $p \times p$, donde p expresa el número de los términos de error de las variables endógenas observables. Este elemento indica la covarianza de los errores (ε) de las variables endógenas observables “Y1” y “Y2”, y la varianza de error de la variable endógena observable.

8. Theta - Delta ($\Theta\delta$)

Es una matriz de orden $q \times q$, donde q expresa el número de términos de error de las variables exógenas observables. Este elemento indica la covarianza de los errores (δ) de las variables observables “X1” y “X2” y la varianza de error de la variable exógena observable.

Los ocho parámetros anteriores, especifican las relaciones existentes entre las distintas variables del modelo, posibilitando de esa forma, nominar cada interacción entre ellas.

1.3. IDENTIFICACIÓN DEL MODELO

La identificación es una de las operaciones más importantes y complejas del desarrollo de los modelos causales; sino queda claramente establecida, todo el proceso posterior puede carecer de sentido. Bentler (1980, 1986, 1988) considera es este contexto, cuando los modelos causales cobran vida y si el modelo no está identificado, no puede procederse al ajuste. Por tanto, habrá que modificarlo cuantas veces sea necesario hasta lograr la identificación del modelo; que se alcanza cuando se llega a obtener una solución única.

Un modelo estará identificado, si todos los parámetros lo están, es decir, si existe una solución algebraica única para cada uno de los parámetros estimados, expresándolo en función de las varianzas y covarianzas muestrales. A su vez, ciertas reglas generales pueden ser aplicables para identificar un modelo, una de ellas es la regla de los grados de libertad. Esta es una condición necesaria para la identificación del modelo, en la cual los grados de libertad, afirma Bentler (2001) deben ser iguales o superiores a cero ($gl \geq 0$)⁵.

⁵ Bizquerra (1989) explica, los grados de libertad (gl) vienen dados por la diferencia entre el número de ecuaciones (e) y el número de parámetros estructurales (π). Así la ecuación sería:

$$gl = e - \pi$$

El número de ecuaciones (e) viene dado por:
$$e = \frac{V(V+1)}{2}$$

Siendo (V) el número de variables observables, tanto endógenas como exógenas. La expresión anterior es, precisamente, el número de elementos distintos que tiene la matriz de varianzas-covarianzas entre las variables observables.

La ecuación de (π) sería la suma de todos los parámetros a estimar, que viene dado por:
$$\pi = \lambda + \delta + \phi + \dots$$



Los modelos pueden estar identificados, sobre-identificados o subidentificados. El modelo identificado, tiene equivalencia entre los momentos y los parámetros estructurales, donde el número de varianzas y covarianzas equivale al número de parámetros estimados, o sea, existe igual número de incógnitas que de ecuaciones. Ello se define con un valor cero o mayor de chi-cuadrado, esto ofrece un ajuste perfecto del modelo, pero la solución no tiene interés puesto que no se puede generalizar. Por otra parte, un modelo “sobre-identificado”, cuenta con un número de estimadores (parámetros) superior al número de momentos o estimaciones realizadas (correlaciones, covarianzas de las variables observables). Es el objetivo de todos los modelos de ecuaciones estructurales. O sea, tiene más información en la matriz de datos que el número de parámetros a estimar, obteniendo con ello un número positivo de grados de libertad ($gl > 0$).

Al igual que otras técnicas multivariantes, el investigador se esfuerza por conseguir un ajuste aceptable con el mayor grado de libertad posible. Esto asegura que el modelo sea tan generalizable como sea posible. El modelo “sub-identificado” es aquel donde el número de parámetros a estimar es superior al número de varianzas y covarianzas (momentos), por lo que, el modelo no posee suficiente información para poder alcanzar un solución aceptable⁶. Cuando el modelo de ecuaciones estructurales es recursivo, está también identificado.

Aunque esta condición necesaria se produzca, Hu (1999) advierte, ello no significa un modelo identificado pues pueden darse una serie de casos: a) el modelo de regresión de una sola ecuación está siempre identificado, b) los modelos recursivos no saturados tienen ecuaciones simultáneas sin efectos recíprocos, c) los saturados tienen un $V = 0$, están identificados por una sola solución.

1.3.1. Condicionantes para la identificación

Cabe añadir, un modelo identificado exige dos condiciones, por un lado el número de parámetros a estimar ha de ser igual al número de unidades de información de que se disponen. Por otro, cada ecuación del modelo es diferente a las demás o a cualquier combinación lineal entre ellas. Ahora bien, cuando cada parámetro del modelo está identificado, también lo estará el modelo completo; pero, Dolan (1996) y

⁶ De modo semejante se puede explicar por la regla del conteo, es decir, se denotará al número total de variables con $s = p + q$, siendo p las variables endógenas y q las exógenas. Luego, el número de elementos no redundantes en Σ es igual a $\frac{1}{2} s(s+1)$. Además, se toma el número total de parámetros a ser estimados en el modelo como t , entonces, para realizar la identificación del modelo se debe tener la siguiente condición necesaria $t \leq \frac{1}{2} s(s+1)$: a) si se tiene la igualdad, se dice que el modelo está *identificado*; b) si t es estrictamente menor que $\frac{1}{2} s(s+1)$, se dice que el modelo está *sobre identificado*; c) si t es mayor que $\frac{1}{2} s(s+1)$, entonces el modelo *no está identificado*.



Jöreskog (2001) afirman, basta uno solo de los parámetros que no esté identificado, para que el modelo total no lo esté.

La identificación del modelo también depende de las restricciones hechas sobre cada variable latente; es decir, por lo general se restringe a uno la primera variable observable de cada variable latente, o si se cuenta con valores de estimación de la saturación o correlación múltiple de las variables observables perteneciente a la variable latente, se debe seleccionar la de mayor valor, para restringirla a uno.

Insistiendo una vez más, la identificación puede plantearse de la siguiente forma: un modelo específico, con una estructura: $\Lambda_y, \Lambda_x, \beta, \Gamma, \Phi, \psi, \theta_\varepsilon, \theta_\delta$ entre otras cosas, genera una única Σ matriz de covarianza, obtenida a partir de los parámetros estimados, pero, podría haber varias estructuras que generen la misma matriz Σ . Evidentemente si dos o más estructuras generan la misma matriz Σ se dice que las estructuras son equivalentes. De esa forma, si en estructuras similares, un parámetro tiene el mismo valor, se dice que está identificado.

1.4. ESTIMACIÓN DE PARÁMETROS

Una vez identificados los parámetros del modelo estructural propiamente dicho, se procede con la estimación de éstos, por medio de diferentes técnicas implementadas a partir de la normalidad o no normalidad de los datos. Para ello, se puede usar como datos iniciales la matriz de varianzas-covarianzas⁷ o la matriz de correlaciones de las variables observadas, seleccionando el procedimiento de estimación que convenga según las necesidades de investigación: al respecto existen varias opciones, tanto para la estimación como para la elección del programa informático a utilizar⁸.

⁷ La utilización de matrices de correlación favorece la comparabilidad de los coeficientes estructurales y colabora en la interpretación de los resultados, en especial si la intención es comprender el modelo de relaciones entre los constructos. En cambio, si el objetivo de la investigación es comprar muestras o poblaciones diferentes en lugar de comparar distintas variables de una misma muestra, es más adecuado utilizar las matrices de varianza-covarianza.

⁸ Por lo amigable de su interfaz recomendamos AMOS, en el diseño de la estructura y en todas las estimaciones a realizar. Para Hox (1995), este programa ha ido ganando una popularidad creciente en los últimos años, debido a la sencillez del interfaz para el usuario, que permite acciones rápidas y simples. Además, por su capacidad de estimación se puede comparar con LISREL y EQS.



En las versiones de los programas informáticos LISREL⁹, EQS y AMOS, entre otros¹⁰, se pueden encontrar siguientes técnicas de estimación de parámetros: variables instrumentales (IV), mínimos cuadrados ordinarios (OLS), mínimos cuadrados en dos etapas (TSLS), escala libre de mínimos cuadrados (SFLS), asintóticamente libre de distribución (AGL), mínimos cuadrados generalizados (GLS), mínimos cuadrados ponderados (WLS), mínimos cuadrados no ponderados (ULS) y el método de máxima verosimilitud (ML).

Los primeros tres métodos no iterativos (IV, TSLS, SFLS), son más rápidos en alcanzar la estimación de los datos del modelo, logrando eficacia en la temporalidad del trabajo del investigador, pero, son menos precisos que los restantes o sea, los iterativos. Estos últimos (AGL, GLS, ULS, WLS y ML), parten de un conjunto de estimaciones iniciales, ejecutadas por repeticiones de aproximaciones sucesivas, hasta conseguir una estimación concreta, estable y precisa. El método de Máxima Verosimilitud (ML) es coherente, no sesgado, eficiente, invariante al tipo de escalas y normalmente distribuido si las variables observables responden a las condiciones de normalidad.

El método de Mínimos Cuadrados Generalizados (GLS) tiene las mismas propiedades que el método de Máxima Verosimilitud (ML), bajo consideraciones de normalidad multivariable menos rigurosas. Los estimadores obtenidos con el método de Mínimos Cuadrados Ponderados no responden a la hipótesis de normalidad de la distribución. Los estimadores varían con el tipo de escalas utilizada.

De todos, quizá el procedimiento que se ha convertido, en una técnica ampliamente utilizada por la mayoría de los investigadores, es la estimación de máxima verosimilitud (ML). En palabras de Hayduk (1996), es más eficiente y no sesgada que las otras, cuando se cumplen los supuestos de normalidad multivariante. De hecho, este procedimiento es lo suficiente robusto, como para no verse seriamente afectado por las ligeras oscilaciones de la distribución normal multivariada.

No obstante, debido a esa sensibilidad del ML respecto a la no normalidad de los datos, Wang y otros (1996) recomiendan utilizar alguna de las siguientes técnicas alternativas como: WLS, GLS y AGL. Indudablemente, la diferencia entre éstas y la ML,

⁹ El pionero de los programas informáticos ha sido LISREL (*Linear Structural RELations*) de Jöreskog y Sörbom (1988a, 1988b, 1993) y se utiliza para resolver problemas de investigación relacionados con: sección- cruzada, experimental, cuasi – experimental y estudios longitudinales. LISREL ha encontrado aplicaciones a lo largo de todos los campos de estudio y se ha convertido en sinónimo de la modelización de ecuaciones estructurales.

¹⁰ También existen los programas: PROC CALIS de SAS (Hatcher, 1996); COSAN (Fraser, 1980) o LVPLS (Wold, 1981).



se fundamenta en que estas no asumen dicho supuesto, permitiendo una mayor aproximación a los datos sin distribución normal multivariada.

En los últimos años la técnica AGL, ha recibido una atención particular debido a su insensibilidad a la no normalidad de los datos. Sin embargo, su única objeción radica en la necesidad y exigencia de aumentar significativamente el tamaño muestral.

1.4.1. Tamaño de la muestra

Respecto al tamaño muestral en la estimación de los parámetros, en los restantes métodos de estimación Marsh y otros (1988), recomiendan que este deba estar comprendido entre 100 y 200 observaciones, sin importar cuál fuera el tamaño de la muestra original. En la misma línea se expresa Hoelter (1983) al señalar el tamaño, considera crítico rondar los 200, debido principalmente, si pasa de 400 a 500, el ML se hace más sensible y se reducen casi todas sus propiedades de análisis. Por su parte Tanaka (1987) advierte, en una muestra mayor de 500 observaciones, el método de máxima verosimilitud, casi no detecta ninguna diferencia, haciendo que todas las medidas de calidad de ajuste indiquen un ajuste muy pobre. Sin embargo, estas indicaciones en el tamaño de la muestra, difieren o no han de ser consideradas si se emplea el AGL, dado su primer requisito se fundamenta a una muestra grande, para compensar la dependencia sobre los supuestos de distribución mediante otros métodos.

1.4.2. Procesos de estimación

Ahora bien, además considerar la técnica para la estimación del modelo y el tamaño de la muestra, es imprescindible elegir alguno de los procesos de estimación. Estos van desde la estimación directa, la cual es similar a la utilizada en la mayoría de las técnicas multivariante, hasta los métodos que generan miles de estimaciones hasta alcanzar resultados finales. Se pueden nombrar entre otros: bootstrapping, la simulación y el análisis de jackknife.

La estimación directa ejecuta sus cálculos sobre el modelo propuesto, mediante uno de los procedimientos de estimación elegido sea IV, TSLS, SFLS, ADF, GLS, ULS o ML. Los últimos tres, no descansan en un único modelo de estimación, sino en estimaciones del parámetro calculado y sus intervalos de confianza basados en estimaciones múltiples.

Desde otra perspectiva, el objetivo de la estimación es obtener las estimaciones del vector de parámetros Ω , denotado como $\hat{\Omega}$, que minimice la función de discrepancia $F(S, \hat{\Sigma})$, donde $\hat{\Sigma} = \Sigma(\hat{\Omega})$ es la matriz de varianza-covarianza de las estimaciones del modelo llamada matriz de varianza-covarianza ajustada.



La función $F(S, \hat{\Sigma})$ es un escalar que mide la discrepancia (distancia) entre la matriz de varianza-covarianza muestral S y la matriz de varianza-covarianza ajustada $\hat{\Sigma}$.

Esta función de ajuste vendrá dada por la siguiente expresión:

$$F = [S - \Sigma(\Omega)]W^{-1}[S - \Sigma(\Omega)]$$

Donde:

S : es la matriz de varianzas-covarianzas de la muestra,

$\hat{\Sigma}(\Omega)$: es la matriz de varianzas-covarianzas predichas por el modelo,

W : es una matriz de ponderaciones que puede tomar diversas formas dependiendo de la distribución que tengan las variables observadas.

Si se asume que la distribución muestral de dichas variables es normal multivariante, entonces la función de ajuste tomará la siguiente forma:

$$F_{NORMAL} = 2^{-1} \text{Traza}[(S - \Sigma(\Omega))W_2]$$

Donde:

W_2 : es una matriz que puede tomar diversas formas en función del tipo de método de estimación que se escoja:

$$W_2 = \Sigma^{-1} \text{ Máxima verosimilitud,}$$

$$W_2 = S^{-1} \text{ Mínimos cuadrados generalizados,}$$

$$W_2 = I \text{ Mínimos cuadrados no ponderados.}$$

La función de discrepancia para los métodos más usuales son:



Máxima verosimilitud (ML):

$$F_{MV} = \log|\Sigma(\Omega)| + tr[S\Sigma^{-1}(\Omega)] - \log|S| - t$$

Mínimos cuadrados generalizados (GLS):

$$F_{MCG} = \frac{1}{2} tr(I - S^{-1} - \Sigma(\Omega))^2$$

Mínimos cuadrados no ponderados (ULS):

$$F_{MCNP} = \frac{1}{2} tr(S - \Sigma(\Omega))^2$$

En síntesis, la estimación de los parámetros representa esencialmente los coeficientes estructurales del sistema de ecuaciones y corresponden, aproximadamente al vector β en una ecuación de regresión, pudiendo ser interpretados para efectos de análisis, como pesos: entendiéndose como la magnitud de influencia que una variable “causa” provoca en la variable “efecto”. Por esa razón, una vez obtenidos los resultados, han de evaluarse para determinar si se modifica o no el modelo.

1.5. EVALUACIÓN Y REESPECIFICACIÓN

El cuarto paso de este proceso tiene relación con la evaluación y reespecificación del modelo, pues permite examinar con varios criterios e índices de ajuste, las relaciones causales destacadas por su peso y relevancia teórica; un hecho fundamental que permitirá al investigador tomar las decisiones correspondiente, para mantener o excluir la o las variables del modelo. Para tomar estas decisiones se distinguen dos actividades: la primera, analiza los valores de los parámetros visualizados en el modelo gráfico, los coeficientes de correlación y correlación múltiple al cuadrado, entre otros. La segunda, observa las medidas de bondad de ajuste ofrecidas por los programas informáticos¹¹.

¹¹ Los textos de Hair (1999), Bollen (1989), Jöreskog y Sörbom (1993), Arbuckle (1994, 1997) ofrecen una excelente revisión las medidas de calidad de ajuste, sus criterios de evaluación y su aplicación en múltiples situaciones.



1.5.1. Observación del gráfico estimado y sus valores

Una vez estimado el modelo, se recurre al gráfico ofrecido por el paquete informático para analizar valores: los pesos de la regresión de un factor sobre otro, pesos estandarizados o no, las correlaciones y las correlaciones múltiples al cuadrado para cada variable latente, entre otras. Este análisis busca identificar las estimaciones infractoras: o sea, valores no significativos, negativos o muy cercanos a uno. En este sentido Bentler (2001) y Hu (1999) afirman, son coeficientes estimados que exceden los límites aceptables. Los ejemplos más normales son: a) varianzas de error negativas o no significativas para cualquier constructo; b) coeficientes estandarizados que sobrepasan o están muy cerca de 1.0; c) errores de estimación elevados asociados con cualquier coeficiente estimado o sobrepasando el 0.05.

En caso de obtener estimaciones infractoras, en un ítem o una variable, ha de ser eliminada o borrada del modelo, para estimarlo nuevamente. Esta recomendación aplica en caso de ser una, pero, si son varias las variables infractoras y en diferentes factores, se recomienda revisar atentamente la teoría y el modelo, a fin de excluir las variables transgresoras de los ajuste, que por su condición teórica no afecten significativamente el modelo. Una vez borradas del modelo, se puede reestimar nuevamente.

1.5.2. Índices de ajuste del modelo

La evaluación de la calidad de ajuste del modelo¹², consiste en comprobar si el modelo se ajusta o no a los datos, si es parsimonioso o si reproduce la matriz de covarianza. Esta etapa de diagnóstico de la bondad del ajuste, se refiere a la exactitud de los supuestos del modelo especificado, para determinar si el modelo es correcto y si éste representa gráficamente una aproximación del fenómeno real de investigación, precisando así su poder de predicción. Por tanto, un ajuste perfecto se alcanza cuando en el ajuste global existe una correspondencia perfecta entre, la matriz reproducida por el modelo y la matriz de observaciones.

En ese mismo orden de ideas, a continuación se representan los tres tipos de índices que componen el ajuste global:

- a) Medidas absolutas de ajuste; evalúan el ajuste global del modelo, es decir, si los valores de los parámetros estimados reproducen o predicen la matriz observada de covarianza;

¹² La información es localizada en los resultados o “output” de los programas, referente al ajuste del modelo o “model fit”. La información no es ofrecida, en la primera imagen del modelo cuando es estimado.



-
- b) Medidas del ajuste incremental; comparan el modelo propuesto con un modelo base, llamado comúnmente: modelo nulo;
 - c) Medidas del ajuste de parsimonia: diagnostica la existencia o no de la sobre identificación de los datos o la presencia de demasiados¹³ parámetros¹⁴.

La idea de observar cada uno de los índices de ajuste, atendiendo directamente a los valores superiores y significativos, es simplemente considerar el buen o mal ajuste global de nuestro modelo. Pues, un modelo respetable ofrece valores por encima de lo aceptable en todos sus índices, anulando la posibilidad, de cumplir con unos y con otros no.

1.5.3. Síntesis para evaluar modelos

La literatura recomienda emplear múltiples indicadores para evaluar el ajuste del modelo (Hu & Bentler, 1995). Entre los más utilizados podemos destacar el estadístico chi-cuadrado, la razón de chi-cuadrado sobre los grados de libertad (CMIN/DF), el cambio en chi-cuadrado entre los modelos alternativos, el índice de ajuste comparativo (CFI), el índice de bondad de ajuste (GFI), y el error cuadrático medio de aproximación (RMSEA). Los valores de estos estadísticos de bondad del ajuste (CFI, GFI) varían por lo general entre 0 y 1, con 1 indicando un ajuste perfecto. Valores superiores a 0,9 sugieren un ajuste satisfactorio entre las estructuras teóricas y los datos empíricos, y valores de 0,95 o superiores, un ajuste óptimo. El chi cuadrado debe ser no significativo para indicar un buen ajuste de los datos. Esto es así porque un valor significativo de χ^2 implica que la estructura del modelo teórico propuesto es significativamente diferente de la indicada por la matriz de covarianza de los datos. No obstante, este último estadístico es sensible al tamaño muestral y debe interpretarse con precaución. Usualmente se interpreta también la razón de chi cuadrado sobre los grados de libertad, con valores inferiores a 2 indicando un buen ajuste. Cuando se comparan diferentes modelos teóricos, la reducción significativa en chi cuadrado de un modelo respecto a otro también sugiere un ajuste más adecuado a los datos (Tabachnick y

¹³ Sobre este tema Hair y otros (1999) advierten "...cuanto mejor ajuste pueda conseguirse con pocos coeficientes, mejor será el contraste del modelo y más confianza podemos tener en que los resultados no sean producto del sobreajuste de los datos" (pág. 637). Basados en ese argumento, se puede señalar la importancia de tratar de alcanzar un modelo robusto con pocos parámetros de estimación.

¹⁴ Estos índices desempeñan el mismo papel del R^2 en la regresión múltiple (Tanaka y Huba, 1989).



Fidell, 2001). El índice RMSEA es considerado óptimo cuando sus valores son inferiores a 0,06 (Hu y Bentler, 1995). Finalmente, además de considerar el ajuste del modelo, debe prestarse atención a la significación de los parámetros estimados que son análogos a los coeficientes de regresión. Al igual que en el análisis de regresión, un modelo que se ajusta bien a los datos, pero que posee pocos coeficientes significativos, no tendría mucho sentido.

AMOS ofrece en los resultados de sus estimados: los pesos de las regresiones, los pesos estandarizados de las regresiones, los interceptos, las varianzas, las correlaciones múltiples al cuadrado para cada variable latente y para las observables, las covarianzas, la correlación, los pesos totales de las variables latentes sobre las observables, los efectos totales, los efectos totales estandarizados, los efectos directos e indirectos, la matriz de varianza covarianza y la correlación de los estimados, y los ratios críticos entre las diferencias de los parámetros.

De todos estos estimados, la *correlación múltiple al cuadrado* para las variables latentes endógenas y las observables es muy empleada, dado que este valor indica la intensidad de relación especificada en el modelo para cada variable. En otras palabras, sería la cantidad de variabilidad de una variable determinada que queda explicada por el modelo.

Otro indicador es el *coeficiente de determinación total*, es la correlación múltiple al cuadrado para todas las variables conjuntas, e indica la intensidad de las relaciones especificadas en el modelo global. En los modelos con variables latentes, es una medida generalizada de cómo las variables “X” conjuntamente sirven para estimar la o las variables “xi” (ξ).

Todas estas medidas, mencionadas someramente, ofrecen una visión de su comportamiento y calidad en el modelo que se ha propuesto. Ahora bien, los coeficientes que determinan la calidad oscilan entre 0 y 1, en cuanto mayor es el valor, mejor es el modelo afirma Page (1990). No obstante, aunque no se ha fijado ningún límite que defina lo que se consideran correlaciones elevadas, los valores que exceden de 0.90, según Bentler y Chou (1987) siempre deberían ser examinados con cautela. Así mismo, son indicio claro de que el modelo propuesto no se ajusta a los datos, si los parámetros son superiores a uno o las varianzas son negativas.

Es *segundo* de los grupos corresponde a las medidas de bondad de ajuste, que están contenidas en los programas informáticos; al igual que en los párrafos anteriores, nos referiremos a ellos sucintamente.



La primera de estas medidas que se ofrecen es el Ji- cuadrado χ^2 . Esta es una medida del ajuste del modelo absoluto, que es sensible para comprobar el tamaño y la no normalidad en la distribución de las variables de entrada de los datos. No obstante, los investigadores no se basan solamente en este criterio, sino que tienen en consideración varias medidas ajuste para evaluar, el ajuste total de un modelo a los datos.

Los grados de libertad asociados al Ji- cuadrado son:

La utilización de esta medida se ofrece más como bondad de ajuste que como test, por eso los valores grandes de χ^2 corresponden por lo general a ajustes pobres y los pequeños a buenos ajustes. No hemos de olvidar, que una consideración importante al respecto en un texto clásico, Jöreskog (1967) se refiere al tema en el contexto del análisis factorial "...ha sido una visión ampliamente aceptada que un modelo es solamente la mejor aproximación, y que, afortunadamente, puede ser empleado sin que sea verdadero. Este punto de vista implica que los modelos nunca serán completamente correctos y por consiguiente pueden ser rechazados en contextos estadísticos si han sido probados con una muestra grande. Consecuentemente, el rechazo de un modelo es puramente un contexto estadístico particularmente cuando la muestra es grande, pero ello no significa necesariamente, que el modelo sea inservible" (p. 480).

Hu y Bentler (1999), al igual que otros investigadores, recomiendan que para el ajuste de los datos, se tomen en consideración otros índices de igual importancia como: el índice de Tucker y Lewis (*TLI*), índice ajuste comparado (*CFI*), el error de aproximación de raíz media cuadrada (*RMSEA*), entre otros. Para los dos primeros, los valores han de estar por encima de 0.95 para ser considerados como un ajuste excelente y para el último de los índices, los valores deben estar por debajo del 0.06.

Estos índices también comparan el ajuste absoluto del modelo especificado con el ajuste absoluto del modelo independiente. Más la discrepancia entre el ajuste total de los dos modelos.

De momento, las medidas de *ajuste general*, comúnmente empleados por LISREL, son: el índice de bondad de ajuste (*GFI*), el *GFI* ajustado a los grados de libertad (*AGFI*) y la raíz media cuadrática residual (*RMR*).

El *índice de bondad de ajuste (GFI)*, según Page (1990) es una medida de la variabilidad explicada por el modelo. Si se aplica el método de máxima verosimilitud, este vendrá expresado por:

$$GFI = 1 - \left[\frac{tr(\Sigma^{-1}S - I)^2}{tr(\Sigma^{-1}S)^2} \right]$$

Donde Σ^{-1} es la inversa de la matriz de varianzas reproducida a partir de los valores estimados en el modelo, S es la matriz de covarianza entre las variables observadas, tr es la traza que indica la suma de los elementos de la diagonal de la matriz que corresponda a la expresión que va entre paréntesis, e I la matriz de identidad.

Si se aplica el método de *mínimos cuadrados no ponderados (ULS)*, entonces la ecuación quedaría de la siguiente forma:

$$GFI = 1 - \frac{tr(S - \Sigma)^2}{trS^2}$$

El valor de GFI generalmente oscila entre 0 y 1, siendo la unidad o su cercanía el mejor ajuste. Este señala lo contrario si este coeficiente se acerca a cero. Una consideración importante que debemos tener en cuenta, nos la señala Hatcher (1996) quien afirma que este coeficiente es independiente del tamaño de la muestra y es menos sensible que el ji-cuadrado a las desviaciones de normalidad, por tanto se puede confiar más en su ajuste que en el de ji-cuadrado.

El *índice ajustado de bondad (AGFI)*, en función de los grados de libertad, se define como:

$$AGFI = 1 - \left[\frac{k(k+1)}{2d} \right] (1 - GFI)$$

Donde k es el número de variables observadas y d son los grados de libertad.

Al igual que el coeficiente anterior oscila entre 0 y 1, siendo los valores cercanos a la unidad los que nos indican un buen ajuste, y los bajos, lo contrario.

La *raíz de la media cuadrática residual (RMR)* viene definida por Rigdon (1997) como:

$$RMR = \sqrt{\frac{2 \sum_{i=1}^h \sum_{j=1}^i (s_{ij} - \sigma_{ij})^2}{k(k+1)}}$$

Donde k es el número de variables X o Y observables. Arbuckle (1997) define la expresión como la raíz cuadrada media de los cuadrados de las diferencias entre los valores de las varianzas y covarianzas de los parámetros estimados o, también, el promedio de la varianza-covarianzas residuales, en cuanto difieren de sus estimados. Es decir, una medida de discrepancia entre las covarianzas observables y las covarianzas de los parámetros estimados. Un RMR pequeño o cercano a cero es lo mejor y más recomendable; su cercanía a cero indica que el ajuste es perfecto.

El índice RMR nos sirve para comparar el ajuste de modelos alternativos para los mismos datos. De igual forma, puede ser empleado también el GFI , pero también lo utilizamos para comparar el ajuste de modelos para datos diferentes.

Los índices anteriores son medidas que se refieren al ajuste global del modelo, pero no indican nada sobre la adecuación o no de cada una de sus partes. Si queremos una valoración más detallada del ajuste, se han de inspeccionar los *residuales normalizados* a través del Q -plot y los *índices de modificación*. Estos ayudan a interpretar el ajuste del modelo y a detectar los desajustes. Un residual normalizado superior a 1.96 indica que el modelo no se ajusta en esta casilla, al nivel de significación del 0.05.

El Q -Plot es una gráfica que representa los residuales normalizados, donde una nube de puntos por encima de la pendiente de 45° representa un buen ajuste y los malos están por debajo de esa recta.

Los *índices de modificación* señalan la disminución mínima esperada en χ^2 si se liberara un parámetro determinado, manteniendo fijos todos los demás. De este modo, el modelo sobre el que se habían impuesto algunas restricciones queda algo más relajado.

Finalmente, se puede disponer de otras ayudas a la interpretación que vienen dadas por los errores típicos, la matriz de residuos y los valores t de cada parámetro. El valor t se obtiene dividiendo cada parámetro por su error típico; los valores obtenidos menores que 2 no son significativos.



1.5.4. Otros índices de ajustes

Este programa informático, además de los indicadores mencionados anteriormente, cuenta con otros índices para el ajuste del modelo. Cabe señalar que todas las medidas de ajuste en AMOS reporta para cada modelo específico, dos modelos adicionales llamados el *modelo saturado* y el *modelo independiente*. Arbuckle (1994, 1997) señala que el primero de ellos, es el más general de todos, es un modelo vacío en el sentido de que este garantiza el ajuste perfecto de cualquier base de datos con que se cuente. Cualquier modelo de AMOS es una versión con restricciones del modelo saturado. El segundo, *modelo independiente*, está en el extremo opuesto, pues se asume que las variables observadas están incorrelacionadas unas con otras, lo que significa que los estimados o las restricciones, tienen la media ajustada a cero. Estas severas e implausibles restricciones, son las que se podrían esperar que suministren un pobre ajuste del set de datos. Esto significa que, para cada uno de los modelos que se especifiquen, se pueden tener restricciones que son equivalentes con el modelo independiente. En este sentido, el modelo saturado y el modelo independiente han de ser vistos como los dos extremos entre los cuales nosotros proponemos el nuestro.

1.5.5. Medidas de parsimonia

Arbuckle (1994, 1997) expresa que los modelos con *pocos parámetros*, y relativamente con muchos grados de libertad, son algunas veces muy parsimoniosos o simples; y los modelos con *muchos parámetros* y pocos grados de libertad se dice que carecen de parsimonia. El uso de estos términos, *simple* y *complejo*, no siempre está entendido. Por ejemplo, el modelo saturado puede ser llamado complejo, mientras que un modelo con un elaborado patrón de dependencias lineales, pero con altas restricciones en los valores de los parámetros, puede ser llamado simple.

Al respecto, Mulaik (1989) destaca que un investigador puede preferir el modelo simple o el complejo, pero no parece existir ningún desacuerdo referente a que el modelo parsimonioso sea preferible al complejo. Lo importante es conseguir un buen ajuste de los datos. Para ello, existen muchas medidas que intentan balancear estos dos objetivos (simplicidad y buen ajuste de bondad). Steiger (1990) refiriéndose al tema, advierte:

“...en la conclusión de un análisis, podría ser, en un sentido, imposible definir en un solo índice numérico la mejor forma de combinar la complejidad y la medida de un mal ajuste, porque la naturaleza precisa del mejor equilibrio numérico entre la



complejidad y el ajuste es alguna prolongación de la experiencia personal en la materia de estudio. La selección del modelo es un problema clásico...” (pág. 179)

Por eso, como lo señalábamos anteriormente, es necesario contar con múltiples criterios de parsimonia. El primero que aparece en AMOS es el *NPAR*, que se refiere al *número de parámetros* distintos (q) que van a ser estimados. Aporta información sobre dos pesos de regresión, que se requiere que sean iguales y contabilizados como un solo parámetro y no como dos. En otras palabras, sería el χ^2 ji - cuadrado.

La *razón de parsimonia* (*PRATIO*), según Mulaik y otros (1989); Hu y Bentler (1999) expresa el número de restricciones en el modelo que va a ser evaluado como una fracción del número de restricciones en el modelo independiente. Su expresión sería la siguiente

$$PRATIO = \frac{d}{d_i}$$

Donde d son los grados de libertad del modelo que va a ser evaluado y d_i los grados de libertad de independencia del modelo. La *razón de parsimonia* (*PRATIO*), es empleado para calcular el *PNFI* y *PCFI*, pero estos índices los trataremos con más detalles en las páginas siguientes, en la sección de *medidas de ajuste de la parsimonia*.

1.5.6. La muestra mínima en la función de discrepancia

El *CMIN* es el *valor mínimo de la función de discrepancia*, C . Esta función es minimizada por AMOS en la ecuación:

$$C(\alpha, a) = [N - r] \left[\frac{\sum_{g=1}^G N^{(g)} \int \mu^g, \Sigma^{(g)}; \bar{x}^{(g)}, S^{(g)}}{N} \right] = [N - r] F(\alpha, a)$$



Se obtendría una función diferente de discrepancia, si cambiamos la definición de f . Si la media y el intercepto no tienen restricciones o no aparecen explícitos en los parámetros del modelo, $\bar{x}^{(g)}$ y $\mu^{(g)}$ estos pueden ser omitidos de f .

Para el método de máxima verosimilitud (ML), asintóticamente libre de distribución (ADF) y mínimos cuadrados generalizados (GLS), la estimación de la fórmula anterior tienen una distribución de ji-cuadrado χ^2 para los modelos correctamente especificados bajo una apropiada asunción distribucional.

Junto a este estimado, aparece una p que se refiere a la probabilidad de tener una discrepancia grande en el modelo. Este es un “ p value” se utiliza para probar que la hipótesis del modelo ajusta perfectamente en la población.

La prueba de la hipótesis ha sido un proceso muy bien aceptado y existe una amplia experiencia de su uso. Sin embargo, su inconveniencia como herramienta para la selección del modelo fue señalada en los primeros años del desarrollo de los modelos causales por Jöreskog (1969), como indicamos en los anteriores párrafos. Normalmente, se ha reconocido que la mayoría de los modelos que se utilizan no ajustan perfectamente a la población, lo que indica que la hipótesis nula del ajuste perfecto no es creíble. Arbuckle (1997) recomienda que puede aceptarse el modelo si la muestra reunida no es demasiado grande. Bentle y Bonett (1980) se expresan en la misma línea al afirmar que en muestras muy grandes, virtualmente todos los modelos que uno podría considerar tendrían que ser rechazados estadísticamente, por eso este procedimiento de la hipótesis nula no puede ser generalmente justificado.

Para concluir sobre este aspecto, tomamos las palabras de Browne y Mels (1992), “...en nuestra opinión la hipótesis nula, del ajuste perfecto, es implausible y no ayuda mucho a conocer si la prueba estadística hace algo para detectar que es nuestro modelo es falso...” (pág. 78)

Otro indicador que nos ofrece el programa informático AMOS 4.0 es el CMIN/DF, el cual proporciona un coeficiente comprendido por el mínimo de la discrepancia dividido por los grados de libertad:

$$CMIN / DF = \frac{\hat{C}}{d}$$

Diferentes autores sugieren que se puede emplear esta razón como una medida de ajuste del modelo. Un valor cercano a uno indica que el ajuste es correcto; no



obstante, no hay criterios muy definidos que nos señalen exactamente, cuál debe ser el valor para dar el modelo como inapropiado.

Algunas concepciones determinadas sobre este tema, se pueden observar en Carmines y McIver (1981), quienes, citando las palabras de Wheaton (1977) afirman que “el investigador también ha de calcular un ji-cuadrado *relativo*...este puede sugerir una razón de aproximadamente cinco o menos como un principio de los niveles aceptables. En nuestra experiencia, sin embargo, el χ^2 entre los grados de libertad debe estar en un rango de 2 o 1 o 3 los cuales indican un aceptable ajuste entre el modelo hipotético y la muestra de los datos” (pág. 80). Marsh y Hocervar (1985) ponen de ejemplo que los investigadores recomiendan que el uso de valores tan bajos como 2 o tan altos como 5 indican un razonable ajuste. Refiriéndose al mismo tema, Byrne (1989) reduce el rango al afirmar que valores mayores a 2 representan un ajuste inadecuado.

Realmente, es difícil encontrar un acuerdo completamente preciso, que ponga de manifiesto un valor máximo para que nos señale el desajuste del modelo, pues los investigadores recomiendan diverso rangos. No obstante, consideramos tomar en cuenta las recomendaciones de Marsh y Hocervar (1985) y de Wheaton (1977), citado por Carmines y McIver (1981), que los valores comprendidos entre 2 y 5 señalan un buen ajuste.

Un valor similar al anterior es el *FMIN*, que calcula el *valor mínimo de la función de discrepancia F*, con la fórmula:

$$F = \hat{C}/(N - G)$$

Cuando $G = 1$ y $r = 1$, equivale a *CMIN* y a la misma fórmula que no tiene restricciones, sobre la media y el intercepto, o no aparecen explícitos en los parámetros del modelo, $\bar{x}^{(g)}$ y $\mu^{(g)}$ y pueden ser omitidos de *f*. Fórmula que quedaría simplificada por:

$$C = (N^{(1)} - 1)F^{(1)} = (N - 1)F$$

1.5.7. Medidas basadas en la población de discrepancia

El término *población de discrepancia* fue introducido por Steiger y Lind (1980) como una medida de adecuación del modelo. El *estimado de la población de*



discrepancia F_0 es el valor de la función discrepancia obtenido por el ajuste del modelo a una población en lugar de la muestra.

Esta es

$$F_0 = \min_{\gamma} [F(\alpha(\gamma), \alpha_0)]$$

En contraste con

$$\hat{F} = \min_{\gamma} [F(\alpha(\gamma), \alpha_0)]$$

Steiger, Shapiro y Browne (1985) señalan que bajo ciertas condiciones $\hat{C} = n\hat{F}$ teniendo una distribución de ji-cuadrado no centralizada con d grados de libertad y parámetros no centralizados es $\delta = C = nF$. No obstante el enfoque de Steiger y Lin (1980) evalúa el modelo centrado alrededor de la estimación del F_0 y de sus cantidades relacionadas.

Uno de estos indicadores es un estimado de los *parámetros de no centralidad*, llamado el *NCP*, que es el resultado de la búsqueda de estadísticos para una medida alternativa, a la razón estadística de verosimilitud del ji-cuadrado, que se vea menos afectada por el tamaño muestral. Generalmente, se utiliza para comparar modelos alternativos. Este se refiere a:

$$NCP = \max(\hat{C} - d, 0)$$

Las columnas que aparecen al lado de este indicador señalan los límites bajo δ_1 y sobre δ_v , con un intervalo de un 90% de confianza en δ, δ_1 que se obtiene por :

$$\Phi(\hat{C}|\delta, d) = 0.95$$

El δ, δ_v se obtiene mediante la solución:

$$\Phi(\hat{C}|\delta, d) = 0.05$$

Para δ , donde $\Phi(x|\delta, d)$ es la distribución de no centralidad de la distribución del ji-cuadrado con parámetros de no centralidad δ , y d grados de libertad.



El $F_0 = \hat{F}_0 = \max \left(\frac{\hat{C} - d.0}{n} \right) = \frac{NCP}{n}$ es un estimado de $\frac{\delta}{n} = F_0$ la población

de discrepancia. Las columnas de la derecha, ofrecidas por los resultados muestran el límite superior e inferior con un 90% de confianza.

Otra medida que intenta corregir la tendencia estadística del ji-cuadrado para rechazar cualquier modelo especificado, con la muestra suficientemente grande, es el *error de aproximación de la raíz media cuadrática (RMSEA)*. Este índice divide la discrepancia entre los grados de libertad al cuadrado; por tanto, esta discrepancia se calcula, no solo en términos de la muestra utilizada para la estimación, sino en términos de población y se obtiene por:

$$poblaciónRMSEA = \sqrt{\frac{F_0}{d}}$$

$$estimadoRMSEA = \sqrt{\frac{\hat{F}_0}{d}}$$

Browne y Cudeck (1993) manifiestan que un valor cercano a 0.05 evidencia un buen ajuste. Aunque los autores advierten que este valor está fundamentado en un juicio subjetivo. No obstante, recomiendan basados en su experiencia, que estos exponentes suelen ser aceptables entre los rangos que van de 0.05 a 0.08. En última instancia, no se deben aceptar valores superiores a 0.1.

Junto a este resultado, aparecen dos valores más que indican el límite alto y bajo, con un 90% de confianza en el intervalo de la población del valor *RMSEA*. Con estos, también se localiza el *PCLOSE*, el cual es un “*p value*” que prueba la hipótesis nula de la cercanía del ajuste del *RMSEA*.

Su expresión viene dada por

$$\text{El } PCLOSE = 1 - \Phi \left(\hat{C} | .05^2 nd, d \right) \text{ y prueba que } H_0 : RMSEA = \leq 0.05$$

Este ajuste según Arbuckle (1997), se diferencia del “*p value*” del ji-cuadrado, porque aquel prueba que la hipótesis en la población *RMSEA* es cero, por tanto:

$$H_0 : RMSEA = 0$$

En todo caso, un buen ajuste del *PCLOSE*, como indicador de la calidad del *RMSEA*, debe estar muy cercano a cero.

1.5.8. Medidas de información teórica

En el programa AMOS 4.0 se reportan varios estadísticos de la fórmula $\hat{C} + kp$ o $\hat{F} + kp$ donde k es una constante positiva. Cada uno de estos estadísticos crea una medida compuesta un mal ajuste $(\hat{F} \circ \hat{C})$ y la complejidad de q por la fórmula que suma el peso de las dos. Un modelo simple que ajusta recibe puntajes bajos y los complicados, con un pobre ajuste puntúan alto. Se entiende que estos estadísticos se utilizan para compara los modelos y no para evaluar un modelo aislado.

Uno de estos principios teóricos es el *criterio de información de Akaike* (Akaike, 1973; 1987). Al igual que el *PNFI*, el *AIC* es una medida comparativa entre modelos con diferente número de constructos. En otras palabras, es la bondad esperada en otra muestra del mismo tamaño. No hay rango establecido de valores aceptables y es utilizado en la comparación entre modelos

El mismo se expresa:

$$AIC : \hat{C} + 2q$$

Otro criterio, que se ha desarrollado específicamente para AMOS, es el *BCC* de Browne y Cudeck (1989). Este impone mayores penalizaciones para el modelo complejo que el de *AIC* y los autores sugieren que, el *BCC*, puede ser una medida superior a las medidas aplicadas normalmente. Viene dado por:

$$BCC : \hat{C} + 2q \frac{\sum_{g=1}^G b^{(g)} \frac{p^{(g)}(p^{(g)} + 3)}{N^{(g)} + P^{(g)} - 2}}{\sum_{g=1}^G p^{(g)}(p^{(g)} + 3)}$$

También, se halla el *criterio de información de Baysiano* o *BIC* (Schwarz, 1978; Raftery, 1993), dado por la fórmula:

$$BIC : \hat{C} + q \ln(N^{(1)} p^{(1)})$$



Comparado con los criterios *AIC*, *BCC* o *CAIC*, este asigna mayores penalizaciones al modelo complejo, y tiene una gran tendencia a seleccionar el modelo parsimonioso. El *BIC* se reporta en el programa, solamente cuando las medias y los interceptos no están explícitos en los parámetros del modelo.

El criterio de *consistencia de información de Akaike (CAIC)* de Bozdogan (1987) asigna también penalizaciones al modelo complejo, como *AIC* y *BCC*. Pero solamente es reportado por el programa para casos de un solo grupo, y al igual que el anterior, en aquellos parámetros del modelo donde no están explícitos las medias y los interceptos.

Su fórmula es:

$$CAIC : \hat{C} + q(\ln N^{(1)} + 1)$$

Un criterio que es similar al *AIC*, excepto por la constante, es el *índice esperado de validación cruzada o ECVI*.

$$ECVI : \frac{1}{n}(AIC) = \hat{F} + \frac{2q}{n}$$

Cuenta, además, con dos columnas que ofrecen información sobre el límite inferior y superior, con un 90% de confianza en la población del *ECVI*.

Finalmente, tenemos el *MECVI*, que es el *índice esperado de validación cruzada modificado*, el cual es idéntico al *BCC*, excepto por la magnitud del factor.

$$MECVI = \frac{1}{n}BCC = \hat{F} + 2q \frac{\sum_{g=1}^G a^{(g)} \frac{p^{(g)}(p^{(g)} + 3)}{N^{(g)} + P^{(g)} - 2}}{\sum_{g=1}^G p^{(g)}(p^{(g)} + 3)}$$

1.5.9. Comparación con un modelo de fondo

Existen varios ajustes de medida diferentes, que animan al investigador a reflexionar acerca del ajuste del modelo y demuestran que este no es tan malo, como se podría esperar, porque podría ser peor.

Bentle y Bonett (1980) y Tucker y Lewis (1973) refiriéndose a lo anterior, sugieren observar el ajuste del modelo independiente o algún otro mal ajuste con el modelo de



fondo, como un ejercicio para determinar el tamaño de la función de discrepancia. Si ninguno de los modelos propuestos ajusta excelentemente, podemos considerar una perspectiva deferente. Por ejemplo, ver tabla N° 26, tomando uno de los resultados de nuestros modelos de estudio, podemos ver que existe discrepancia ($\hat{C} : 1092,393$) en relación con los grados de libertad, pero no obstante, esa discrepancia no es tan mala si la comparamos con la discrepancia del modelo independiente ($\hat{C} : 40154,814$).

TABLA MODELO DE COMPARACIÓN

Model	NPAR	CMIN	DF	P	CMIN/D F
<i>Default model</i>	96	1092,393	401	0.000	2,724
<i>Saturated model</i>	495	0.000	0		
<i>Independence model</i>	30	40154,814	465	0.000	86354

En este caso, se ha escogido el modelo independiente como modelo de fondo, para compararlo con otro, pues es uno de los que frecuentemente se usan para ello. Sin embargo, Sobel y Bohrnstedt (1985) consideran que esta selección a menudo es inapropiada, y sugieren opciones mejores de comparación.

Uno de estos elementos alternativos que son sensibles para seleccionar el modelo de fondo, es el *índice de ajuste normado (NFI)*, de Bentler y Bonett (1980) que oscila entre los rangos de uno y cero; siendo los valores cercanos a la unidad los que muestran un buen ajuste. En la notación de Bollen (1989), se escribe:

$$NFI = \Delta_1 = 1 - \frac{\hat{C}}{\hat{C}_b} = 1 - \frac{\hat{F}}{\hat{F}_b}$$

Donde $\hat{C} = n\hat{F}$ es el mínimo de discrepancia el modelo que va a ser evaluado y

$\hat{C}_b = n\hat{F}_b$ es el mínimo de discrepancia del modelo base.



Ahora bien, si regresamos a observar la tabla N° 26, podemos señalar que nuestro modelo tiene una discrepancia de 97% entre el modelo independiente (el peor ajuste) y el modelo saturado (el mejor ajuste)

$$NFI = \frac{40154,814 - 1092,393}{40154,814} = 0.973$$

Bentler y Bonett (1980b) consideran que los índices de ajuste son difíciles de interpretar, y se requiere experiencia para establecer los valores de los índices que están asociados con varios grados de relevancia en los resultados. En la experiencia de los autores, los modelos de índices de ajuste totales con menos de 0.9 pueden ser mejorados substancialmente.

Bollen (1986) señala también el *índice de ajuste relativo (RFI)*, el cual se obtiene del *NFI* por la substitución de F/d por F . Pero, al igual que el anterior, un valor cercano a uno indica que existe un buen ajuste. La fórmula viene dada por.

$$RFI = p_1 = 1 - \frac{\hat{C}|d}{\hat{C}_b/d_b} = 1 - \frac{\hat{F}|d}{\hat{F}_b/d_b}$$

El mismo autor, señala también el *índice de ajuste incremental (IFI)*, dado por:

$$IFI = \Delta_2 = 1 - \frac{\hat{C}_b - \hat{C}}{\hat{C}_b - d} =$$

Donde \hat{C} y d son la discrepancia y los grados de libertad del modelo que va a ser evaluado y \hat{C}_b y d_b son la discrepancia y los grados de libertad para el modelo base. Unos valores cercanos a uno indican un buen ajuste.



En el contexto de los análisis causales, es muy empleado el *coeficiente de Tucker y Lewis o TLI*, como otro elemento más relacionado con el ajuste de los datos. Un valor muy cercano a uno indica un buen ajuste. No obstante, hay que recordar que Hu y Bentler (1999) recomiendan valores sobre el 0.95, para considerar un buen ajuste. La fórmula viene dada por:

$$TLI = p_2 = \frac{\frac{\hat{C}_h}{d_h} - \frac{\hat{C}}{d}}{\frac{\hat{C}_h}{d_h} - 1}$$

Bentler (1990) señala también el *índice de ajuste comparado o CFI*, dado por:

$$CFI = 1 - \frac{\max(\hat{C} - d, 0)}{\max(\hat{C}_h - d_h, 0)} = \frac{NCP}{NCP_h}$$

Donde \hat{C} , d y NCP son la discrepancia, los grados de libertad y la no centralidad de los parámetros estimados para el modelo que va a ser evaluado, y \hat{C}_b , d_b y NCP_b son la discrepancia, los grados de libertad y la no centralidad de los parámetros estimados para el modelo base. Según McDonald y Marsh (1990) este coeficiente es idéntico al *índice relativo de no centralidad o RNI*. Unos valores cercanos a uno indican un muy buen ajuste.

1.4.2.6. Medidas ajustadas de parsimoniosidad



Las medidas ajustadas de parsimoniosidad son el resultado de la multiplicación de los *NFI* y *CFI*, por un índice de parsimoniosidad, tomando en cuenta el número de grados de libertad para probar el modelo que va a ser evaluado y el modelo base. Estos fueron sugeridos por James (1982) y Mulaik (1989).

El índice de ajuste parsimonioso normado (PNFI) es:

$$PNFI = (NFI)(PRATIO) = NFI \frac{d}{d_h}$$

Donde d son los grados de libertad para el modelo que va a ser evaluado y d_h son los grados de libertad para el modelo base.

El índice de ajuste parsimonioso comparado (PCFI) es el resultado:

$$PCFI = (CFI)(PRATIO) = CFI \frac{d}{d_h}$$

En donde d son los grados de libertad para el modelo que va a ser evaluado y d_h son los grados de libertad para el modelo base.

Al igual que los índices anteriores, su rango es de cero a uno, siendo la cercanía a este último número, el mejor ajuste para el modelo.

1.4.2.7. Diversas medidas

AMOS ofrece el “N crítico” de Hoelter (1983) que se refiere al tamaño de la muestra para aceptar que la hipótesis del modelo es correcta. El autor no especifica un nivel de significancia para ser empleado en determinar el “N crítico”, no obstante, recomienda que valores superiores a 200 o más indican un ajuste satisfactorio. Arbuckle



(1997) por el contrario, no está muy convencido que el umbral determinado en 200 sea un buen argumento para elegir un modelo adecuado y recomienda para ello que empleemos más estándar.

En AMOS podemos ver dos rangos para el “N crítico” de Hoelter en los niveles de significancia de 0.05 y 0.01.

Un estándar que puede ser empleado para observar si el modelo es correcto o no, es la *raíz media cuadrática residual* o *RMR*. Este residual es la raíz cuadrada media por el cual la muestra difiere, en sus varianzas y covarianzas, de sus estimados obtenidos bajo la asunción que el modelo es correcto. Se puede expresar así:

$$RMR = \sqrt{\sum_{g=1}^G \left\{ \sum_{i=1}^{Pe} \sum_{j=1}^{j<1} \hat{S}_{ij}^{(g)} - \sigma_{ij}^{(g)} \right\} / \sum_{g=1}^G p^{(g)}}$$

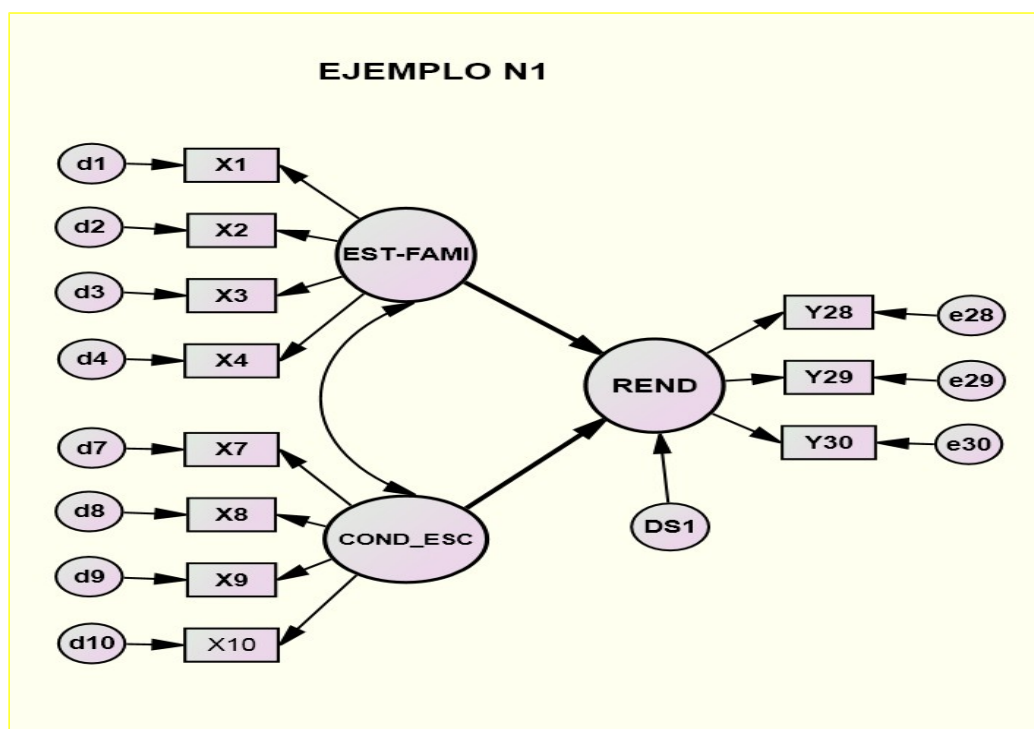
El *RMR* pequeño es lo más recomendado, sin embargo un *RMR* de cero indica que el ajuste es perfecto.

CAPITULO 2: EJEMPLO DE MODELOS DE ECUACIONES ESTRUCTURALES

En este apartado se presenta un ejemplo sobre el funcionamiento, de modelos de ecuaciones estructurales, con su correspondiente modelo de medida y los índices de ajuste de la estimación inicial.

2.6. ESPECIFICACIÓN DEL EJEMPLO

En la especificación del modelo, el investigador fundamenta su estructura en los conocimientos teóricos del fenómeno en estudio y lo traslada al planteamiento de las ecuaciones matemáticas relativas a los efectos causales entre las variables latentes y a las expresiones que las relacionan con los indicadores o variables observables. En este sentido, se plantea una estructura gráfica de una teoría y la relación entre sus variables.



Descripción del Ejemplo N°1:

1.- El ejemplo N°1 cuenta con tres variables: Dos exógenas, (Estatus Familiar-Contexto Escolar) y una endógena (Rendimiento Académico).



2.- Se hipotetiza, las variables exógenas Estatus Familiar y Contexto Escolar, afecta directamente a la endógena, Rendimiento Académico: este efecto directo se llama GAMMA.

3.- Las variables exógenas: a) Estatus Familiar conecta con los indicadores X1, X2, X3 y X4. Esa saturación de X en el Estatus Familiar se llama LAMBDA. Los residuales de estas variables son d1, d2, d3 y d4, o sea, DELTA. b) la variable Contexto Escolar conecta con los indicadores X7, X8, X9 y X10. Los residuales de esta son d7, d8, d9 y d10. El Delta representa lo específico de la variable observada X y los errores aleatorios de medida.

3.- La variable endógena Rendimiento, se conoce por Y28, Y29 y Y30. El residual de todos indicadores Y se llama EPSILON y el de la variable Rendimiento; Dzeta.

4.- Las variables exógenas Estatus Familiar y Contexto Escolar covarian juntas, eso se llama PHI y es un elemento no recursivo.

2.6.1. Ecuaciones del modelo de medida

Ejemplo N°1

Modelo de medida de las variables exógenas

$$\begin{aligned}X_1 &= \lambda_{11}\xi_1\delta_1 \\X_2 &= \lambda_{21}\xi_1\delta_2 \\X_3 &= \lambda_{31}\xi_1\delta_3 \\X_4 &= \lambda_{41}\xi_1\delta_4 \\X_7 &= \lambda_{72}\xi_2\delta_7 \\X_8 &= \lambda_{82}\xi_2\delta_8 \\X_9 &= \lambda_{92}\xi_2\delta_9 \\X_{10} &= \lambda_{102}\xi_2\delta_{10}\end{aligned}$$

El conjunto de ecuaciones se interpreta:

La variable medida (observada) X es igual; al peso de la variable latente (exógena), ξ , sobre la variable observada X , indicado por el λ . Por la variable latente



exógena, ξ ; más el Delta (δ) de X que representa lo específico de la variable medida (observada) y los errores aleatorios de medida.

Modelo de medida de la variable endógena

$$Y_{28} = \lambda_{281}\eta_1 + \varepsilon_{28}$$

$$Y_{29} = \lambda_{291}\eta_1 + \varepsilon_{29}$$

$$Y_{30} = \lambda_{301}\eta_1 + \varepsilon_{30}$$

Se interpreta: La variable medida (observada) Y es igual; al peso de la variable latente (endógena), η , sobre la variable observada Y , representado por λ . Más el Epsilon, (ε) de Y que define lo específico de la variable medida (observada) y los errores aleatorios.

2.6.2. Ecuación del modelo estructural

Modelo estructural ejemplo N°1

$$\eta_1 = \gamma_{11}\xi_1 + \gamma_{12}\xi_2 + \varsigma_1$$

La ecuación del modelo estructural explicita; la variable latente endógena η , es igual: al efecto de β_1 sobre η_1 , representado por γ_1 . Por el β_1 . Más el efecto de β_2 sobre η_1 , mediante el γ_2 . Por el β_2 . Más el ς_1 , considerado como la parte η no explicada por la variables exógenas.

2.7. IDENTIFICACIÓN DEL EJEMPLO

Esta parte representa el estudio de las condiciones para una solución única de los parámetros, cuando cumple ésta función el modelo estará identificado. La identificación propone que debe existir igual número de incógnitas que de ecuaciones, por tanto el sistema tiene una solución.

En este sentido si el modelo teórico es correcto, se procede a la identificación del modelo, en donde se asegura que se pueden estimar sus parámetros. El modelo está



identificado si todos los parámetros lo están, es decir, si existe una solución única para cada uno de los parámetros estimados.

Cabe añadir, que existen reglas básicas utilizadas para identificar un modelo: a) la regla de los grados de libertad, obtenidos como la diferencia entre el número de varianzas y covarianzas (ecuaciones) y el número de parámetros a estimar. Es una condición necesaria pero no suficiente. Cuando $g < 0$, serán modelos infraidentificados, ello indica que se intentan estimar más parámetros de los que permite la información disponible; cuando $g = 0$, los modelos están identificados logrando un ajuste perfecto, y cuando $g > 0$ el modelo está sobreidentificado, o sea, posee más información en la matriz de datos, que el número de parámetros a estimar, significando que tiene un número positivo de grados de libertad. Ha de mencionarse, al igual que en otras técnicas multivariantes, el investigador se esfuerza por conseguir un ajuste aceptable con el mayor grado de libertad posible; b) otra regla que es condición suficiente pero no necesaria, es que si el modelo es recursivo está identificado, siendo este tipo de modelos aquellos que no contienen efectos en sí misma entre sus variables.

Revisando la identificación del ejemplo N°1 tenemos:

The model is recursive.

Sample size = 374

Computation of degrees of freedom (Default model)

Number of distinct sample moments: 77

Number of distinct parameters to be
estimated: 36

Degrees of freedom (77 - 36): 41

Result (Default model)

Minimum was achieved

Chi-square = 110,577

Degrees of freedom = 41

Probability level = ,000



2.8. ESTIMACIÓN DEL MODELO

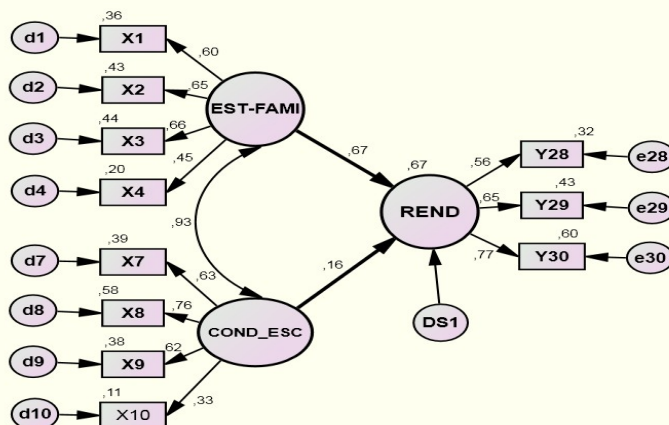
En este apartado se asume que el modelo propuesto está identificado y por lo tanto la función de la verosimilitud presenta un valor que no puede ser superado por ningún otro modelo con las mismas variables. El proceso de estimación consiste en la obtención de aquellos valores P de los parámetros π ajustados lo mejor posible a la matriz observada, por la que aquellos reproducen. La estimación de coeficientes se realiza mediante procedimientos iterativos de minimización de desviaciones, bajo la hipótesis de que el modelo es correcto

En la estimación, el criterio mayormente utilizados es el de máxima verosimilitud, bajo el supuesto de normalidad multivariante, en donde la matriz de ponderación es la inversa de la matriz implicada, $W = (S(p) \otimes \Sigma(p))^{-1}$. Por ello, EL modelo se estima con la razón de Máxima Verosimilitud, explicitado en las páginas anteriores, añadiendo las medias e interceptos, con la aproximación de emulación de LISREL.

2.9. EVALUACIÓN DEL EJEMPLO

2.9.1. MODELO N°1

EJEMPLO N1



Datos del gráfico

		Estimate
REDIMIENTO <---	ESTATUS FAMI	,667
REDIMIENTO <---	CONDICIÓN ESCOLAR	,157
RES28 <---	REDIMIENTO	,564
RES29 <---	REDIMIENTO	,654
RES30 <---	REDIMIENTO	,773
RES3 <---	ESTATUS FAMI	,662
RES2 <---	ESTATUS FAMI	,653
RES1 <---	ESTATUS FAMI	,601
RES9 <---	CONDICIÓN ESCOLAR	,617
RES8 <---	CONDICIÓN ESCOLAR	,759
RES7 <---	CONDICIÓN ESCOLAR	,627
RES4 <---	ESTATUS FAMI	,452
RES10 <---	CONDICIÓN ESCOLAR	,325

La variable Estatus de la Familia ejerce un peso de 0.67 sobre el Rendimiento Académico de su hijo, sin embargo la Condición de la Escuela no es un valor



significativo, con 0.16. No obstante la asociación entre las variables exógenas, Estatus de la Familia y la Condición de la Escuela, es altamente significativo con 0.93.

Ahora bien, cuando centramos la atención sobre las variables observables del modelo, todas parecen representar muy bien sus variables latentes, salvo la RES10 con 0.32. No obstante si tomamos en consideración la correlación múltiple al cuadrado, este mismo ítem solamente alcanza 0.10, un valor no significativo, al igual que el RES4 con 0.20. Ambos deben ser eliminados del modelo.

Correlación múltiple al cuadrado

	Estimate
REDIMIENTO	,665
RES10	,106
RES4	,204
RES7	,393
RES8	,576
RES9	,381
RES1	,362
RES2	,426
RES3	,439
RES30	,598
RES29	,428
RES28	,318

Una vez observados estos datos, ha de recurrirse a los índices de ajuste, que en este ejemplo son:

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	36	110,577	41	,000	2,697
Saturated model	77	,000	0		
Independence model	22	1182,291	55	,000	21,496

Cuando el Ji cuadrado no ajusta es mejor observar el CMIN/DF y el FMIN, tomando en cuenta la escala anotada anteriormente, o sea, entre 2 y 5 como ajuste aceptable.

Baseline Comparisons



Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	,906	,875	,939	,917	,938
Saturated model	1,000		1,000		1,000
Independence model	,000	,000	,000	,000	,000

Estos índices permiten comparar el modelo propuesto con el modelo nulo y sus valores deben superar el 0.90. Otros autores indican >0.95 . Pero son criterios meramente orientativos. Al respecto Bollen (1989) afirma, en área de estudio en el que los modelos anteriores generan: CFI= 0,70, NFI = 0,85 pueden ser considerados un avance y por lo tanto deberían ser aceptables.

Cabe anotar, tanto el índice de ajuste comparativo, como el IFI, NFI, TLI, y RFI, comparan el modelo propuesto con alguno alternativo, sea el modelo nulo o el independiente.

Así mismo, el CFI compara el ajuste de un modelo propuesto con el ajuste del modelo independiente, el cual supone que las variables no están correlacionadas. Por tanto, observa la relación de la discrepancia entre ambos modelos. Es decir CFI, representa el grado en que el modelo propuesto es mejor que el modelo independencia.

Parsimony-Adjusted Measures

Model	PRATIO	PNFI	PCFI
Default model	,745	,676	,699
Saturated model	,000	,000	,000
Independence model	1,000	,000	,000

Son los índices que buscan el ajuste de la parsimonia del modelo. Estos índices penalizan los modelos menos parsimonioso, para que los procesos teóricos más simples se vean favorecidos sobre los más complejos. Cuanto más complejo es el modelo, menor es el índice de ajuste. El índice de parsimonia Normed Fit Index (PNFI), es el producto de NFI y PRATIO. De igual forma, el PCFI es producto del CFI y PRATIO.

<NCP

Model	NCP	LO 90	HI 90
Default model	69,577	42,110	104,707



Model	NCP	LO 90	HI 90
Saturated model	,000	,000	,000
Independence model	1127,291	1019,160	1242,823

Medida de no centralidad.

FMIN

Model	FMIN	FO	LO 90	HI 90
Default model	,296	,187	,113	,281
Saturated model	,000	,000	,000	,000
Independence model	3,170	3,022	2,732	3,332

Cuando el ajuste del ji-cuadrado no ofrece un valor aceptable, puede utilizarse FMIN.

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	,067	,052	,083	,029
Independence model	,234	,223	,246	,000

El valor del RMSEA ha sido determinado como aceptable, entre 0.005 y 0.008

AIC

Model	AIC	BCC	BIC	CAIC
Default model	182,577	184,971		
Saturated model	154,000	159,119		
Independence model	1226,291	1227,753		

Schreiber, James y otros (2006) consideran que el AIC entre más pequeño el valor, es un mejor ajuste del modelo. Esta es una medida indicada para la comparación de modelos no anidados.

ECVI

Model	ECVI	LO 90	HI 90	MECVI
Default model	,489	,416	,584	,496



Model	ECVI	LO 90	HI 90	MECVI
Saturated model	,413	,413	,413	,427
Independence model	3,288	2,998	3,597	3,292

El ECVI es muy similar al AIC, excepto por la constante en el factor de la escala.

HOELTER

Model	HOELTER .05	HOELTER .01
Default model	193	220
Independence model	24	26

Execution time summary

Hoelter considera que un N crítico 200 o mayor, es señal de un ajuste satisfactorio. Este umbral puede ser utilizado en conjunción con un nivel de significación de 0,05.



CAPÍTULO 3: MODELOS JERÁRQUICOS LINEALES

La lógica de los modelos jerárquicos lineales, se fundamentan en la idea de que considerables grupos de datos, incluyendo los datos de observación recogidos en las ciencias humanas y biológicas, tienen una estructura jerárquica o anidada. Hox (1995c) afirma que, en términos generales, un individuo interactúa con su contexto social, lo que significa que las personas están influenciadas por los grupos que componen la sociedad y por las acciones que ese grupo realiza para llevar adelante a esa sociedad. En este sentido, podemos afirmar que los individuos y el grupo permiten ser conceptualizados como un sistema jerárquico, con características específicas que pueden ser definidas en niveles separados, lo que permite el estudio de poblaciones con estructuras complejas.

Para entender la lógica de estas propuestas, hemos de imaginarnos que toda sociedad se agrupa en niveles jerárquicos, de esa manera un individuo pertenece a un barrio, ese barrio a un pueblo, ese pueblo a una provincia, esa provincia a un comunidad autónoma y así sucesivamente. La sociedad se va construyendo y organizando por medio de diferentes niveles jerárquicos o anidados, donde cada unidad se agrupa en un nivel determinado, con características independientes que interactúan con uno u otro nivel.

Si tratamos de observar los datos de una comunidad estudiantil, podemos considerar que existen varios niveles jerárquicos en ella. Tomando a los estudiantes y la escuela, se podría explicar la presencia de dos niveles: el primer nivel correspondería al estudiante y la escuela al segundo. Evidentemente, las características de cada grupo pueden ser estudiadas en diferentes niveles jerárquicos y como resultado, se consigue analizar las características propias de cada nivel de análisis.

Ampliando esa primera aproximación a la lógica que consideran estos modelos, Plewis (1998) expone dos ejemplos más, uno en tres niveles y el otro en cuatro. En el primer ejemplo, se ha de observar que los estudiantes pertenecen a una familia y estas a su vez están insertas en la comunidad, ahí observamos los tres niveles: los estudiantes, sus familias y la comunidad. En el segundo ejemplo, los estudiantes asisten a una clase dentro de una escuela que depende de una autoridad de educación, aquí se determinan cuatro niveles: estudiantes, clases, escuelas y autoridades locales de educación.



El empleo de los niveles nos permiten analizar los resultados por estudiante, por la suma de sus efectos individuales, por clase, por escuela, por distrito o por estado. Cada uno de estos efectos se puede observar desde diferentes ángulos, como una colección cambiante de los efectos (p.e. todos los efectos del nivel de la escuela) determinados en una distribución de componentes de variación. De hecho puede haber también coeficientes de regresión en algunos o todos los niveles. Una vez que un modelo se especifique, las inferencias se pueden establecer con los datos disponibles son muchas y es posible estudiar los efectos en cualquier nivel y la interacción entre ellos. Gibbons y Hedeker (1997) destacan que esto es muy útil cuando hay que analizar los problemas en áreas muy pequeñas, para las cuales hay una cantidad de información muy limitada.

Si nos ponemos a reflexionar sobre la cantidad de niveles de agrupación que debemos considerar para construir un modelo, en principio y de acuerdo con la mayoría de los teóricos (ver: Leyland y Goldstein, 2001; Raudenbush y Bryk 2002; Raudenbush, 1999; Kasim y Raudenbush, 1998; Kreft y De Leeuw, 1998; Goldstein y Rasbash, 1996; Goldstein, 1995; Rodríguez y Goldman, 1995; Goldstein y otros 1993; Longford, 1993) no hay un límite establecido para ello que nos lo indique directamente, pero en la práctica raramente se utilizan más de cuatro o cinco niveles; debido principalmente a las complejidades técnicas que se deben de enfrentar y a la significación desmedida que implica una subdivisión mayor de niveles. Por ello es muy común encontrar estudios de dos o tres niveles, como los de Leyland y otros (2000); Rowan, Raudenbush, y Cheong (1993), Raudenbush y Bhumerat (1992), Bryk y otros (1998a), Selner-O'Hagan, y otros (1998), Sampson y otros (1997), Barnett y otros (1993), Fuller y otros (1993); Raudenbush (1997) y Kindlon y otros (1996).

Hox (1995c) insiste que en realidad el límite es algunas veces borroso y arbitrario, la asignación de las variables no es siempre tan obvio y simple como se piensa. Blalock, (1990) por su parte, destaca que en los problemas multinivel, las decisiones acerca de los miembros del grupo y su operacionalización involucra un amplio rango de asunciones teóricas y de problemas específicos. Cuando el número de variables tiene tamaños diferentes, existe un enorme número de interacciones cruzadas entre niveles.

Idealmente, la teoría multinivel puede especificar qué variables pertenecen a cada nivel y qué efectos son directos del individuo o del contexto y, por lo tanto, requieren la especificación de algún proceso dentro de los individuos que causan esas diferencias individuales, que influyen en ciertos aspectos del contexto. Ahora bien, una vez que se establezcan las agrupaciones, incluso si su establecimiento es al azar, se ha de observar que cada nivel se distingue plenamente. Esta diferenciación implica que el grupo y sus miembros se ven influenciados unos de otros. Si no se hace caso de esta



relación, se podría pasar por alto la importancia de los efectos del grupo y también invalidar muchas de las conclusiones y hallazgos encontrados debido especialmente, a que este método permite, no sólo el análisis de cada nivel, sino también el análisis de la interacción entre los niveles y sus características. Bryk y Raudenbush (1992) hablando sobre este tema advierten que la falta de profundidad en el análisis de los niveles y sus interacciones ha llevado a la acumulación de sesgo, a reducir la precisión del cálculo y a problemas con la unidad de análisis, lo que ha “fomentado el empobrecimiento de la conceptualización” (pág. 3), y a observar algunos problemas como irresolubles. Además los autores plantean que con el empleo de los modelos jerárquicos lineales, se ha logrado enfrentar los problemas con la “unidad de análisis” y con las “medidas de cambio” y ahora se puede plantear la hipótesis sobre las relaciones existentes en cada nivel y entre los niveles, a su vez también evaluar la cantidad de varianza de cada nivel.

En un planteamiento similar, Plewis (1998) concluye que los modelos multinivel tienen como objetivo balancear y entender tanto a los individuos como a su contexto social por ello, muchos estudios e investigaciones presentan problemas de jerarquía en los datos. Por ende, un investigador puede considerar o analizar las diferencias en la implementación de un tratamiento, estudiando las características de los sujetos y relacionándoles con algún tratamiento específico. Un ejemplo muy destacado por los autores lo constituye un estudio realizado en escuelas elementales en los años 70. Bennett (1976) encontró que los niños que de métodos formales de enseñanza en lectura exhibían un mejor progreso, que aquellos que no estaban sujetos a estas prácticas. Para poder llegar a tales resultados, los investigadores analizaron los datos utilizando las técnicas tradicionales de regresión múltiple, las cuales consideraron como unidades de análisis, exclusivamente a los niños y no observaron su contexto en agrupaciones, por ejemplo dentro de profesores y éstos en clases. De todas formas, los resultados obtenidos por la investigación eran estadísticamente significativos, lo que demostraba diferencia en la aplicación del método tradicional de lectura. Posteriormente, en un estudio con los mismos datos, Aitkin (1981) demostró que, cuando en el análisis se considera el agrupamiento de los niños en clases y se delimita la unidad de análisis, las diferencias entre ellos desaparecen. Este re-análisis es el primer ejemplo trascendental sobre la aplicación de los modelos multinivel a las ciencias sociales y sobre la importancia que tiene analizar los datos en estructuras jerárquicas.



3.1. METODOLOGÍA MULTINIVEL

La metodología multinivel, pertenece al paradigma cuantitativo y se fundamenta en que los datos de observación de las ciencias humanas y biológicas, poseen estructura jerárquica. Es decir se agrupa en niveles de jerarquía unos con otros. Por ejemplo hablando de votos electorales podríamos decir: un individuo pertenece a un barrio, ese barrio a un pueblo, ese pueblo a una provincia, esa provincia a una comunidad autónoma. Si pretendemos plantear una estructura anidada o jerárquica, el sujeto sería primer nivel, el barrio segundo, pueblo tercero y así sucesivamente.

La lógica de la sociedad está jerarquizada, sea en donde sea, la sociedad agrupa en niveles de agregación, de menor a mayor, y la teoría subyacente propone, cada grupo pertenece o se encuentra agrupado en otro superior. Pensando en organizaciones jerárquicas, podemos distinguir innumerables estructuras en los datos utilizados por las ramas del saber. Por ejemplo, en medicina se pueden organizar los datos por pacientes en clínicas y éstos en hospitales. En economía encontramos agrupaciones de ingreso por empresas, provincias o ingreso de las agrupaciones gremiales por lugares u otros.

3.1.1. El nacimiento

Aparecen a inicios de los años 70, como una nueva clase de estudios denominada *investigación multinivel*, supone una estructura jerárquica en la agrupación de los datos. Durante esa década, se sientan las bases de la metodología con el surgimiento de diversos textos y aproximaciones teóricas referentes a las estructuras jerárquicas de los datos en las ciencias sociales. No obstante, el punto culminante de todo este desarrollo, según Bryk y Raudenbush (1992) llega a ser esquematizado de forma muy significativa en los trabajos de Lindley y Smith (1972) y Smith (1973), con la formulación de un modelo general para tratar los datos con estructura jerárquica lineal. El trabajo de estos investigadores desarrolla un marco general para datos anidados con una estructura compleja de error, lo cual es un gran avance para el fortalecimiento de la metodología.

Es de reconocer, aunque la aportación de Lindley y Smith (1972) fue un hecho relevante, el estudio y el ajuste de estos modelos no fue operable durante esos años, debido principalmente a las herramientas básicas de cálculo utilizadas en las computadoras de la época, que no podían ejecutar las complicadas estimaciones requeridas por la metodología. Por esa razón, hubo que esperar unos años para que se desarrollen mejores computadoras para que la propuesta fuera operable totalmente.



3.1.2. Posibles aplicaciones

Actualmente, diversas ramas del saber utilizan las estructuras anidadas o multinivel, pues ellas se adaptan a muchos campos, incluidas las ciencias sociales. Este amplio se observa en las áreas más variadas, por ejemplo: educación³, psicología⁴, economía⁵, sociología⁶ y medicina⁷, entre otros. La metodología se difunde con mayor fuerza y en estos momentos existen una gran cantidad de campos del saber que la utilizan para sus investigaciones y estudios.

De los diferentes campos de aplicación, el sistema escolar es el más claro ejemplo de estructuras jerárquicas, con estudiantes agrupados en escuelas y estos en autoridades de educación. Dentro de estos sistemas educativos, Gaviria y otros (1997) destaca diferentes estudios:

- a- Aquellos que se analizan bajo la etiqueta genérica de calidad de los centros educativos, los datos obtenidos de las muestras de estudiantes, profesores, etc. y en la mayor parte de los casos se obtienen mediante muestreo por conglomerados, como unidades de muestreo, que son agrupaciones naturales de individuos;
- b- los estudios sobre la eficacia de la educación en determinadas áreas geográficas (p.e. en los estudios transculturales) y del sistema educativo en general se basan en datos obtenidos sobre variables obtenidas en diversos niveles de agregación: estudiantes, clases, centros escolares, distritos, regiones, países, etc., que deben integrarse y combinarse en un modelo único;
- c- en las evaluaciones de muchas intervenciones educativas y sociosanitarias, los datos suelen tratarse de forma agregada, tanto en los diseños clásicos pre/pos test, como en los de series temporales interrumpidas utilizando promedios de los grupos, a la hora de obtener estos promedios, el evaluador encuentra numerosos problemas derivados de las peculiaridades de los datos: líneas de base inestables, enorme variabilidad intragrupo, necesidad de obtener el mismo número de medidas para todos los sujetos, etc; las sucesivas medidas con las que se va registrando el cambio de los sujetos, están anidadas dentro de éstos y constituirían un primer nivel de análisis;
- d- situaciones genéricamente encuadradas bajo la etiqueta de *medida del cambio*, que dan lugar a modelos de *curvas de desarrollo*, donde uno de los mayores problemas es el de las diferencias individuales, este ha sido uno de



los temas más debatidos en la metodología de las ciencias del comportamiento.

Dentro de este entorno la metodología se puede utilizar en la asignación de pacientes a diversas clínicas, a la organización de un padrón electoral, cuya distribución se ordena por barrios, pueblos y ciudades, o en estudios económicos, fundamentados en el ingreso per cápita por ocupación profesional o actividad económica de la persona, su barrio y provincias. Otra posibilidad de aplicación puede ser en experimentos o ensayos clínicos realizados en varios centros con grupos de individuos elegidos aleatoriamente, ordenados por hospitales. También se puede incluir en estudios políticos relacionados con la influencia que ejercen los medios de comunicación masiva en el electorado.

3.1.3. Importancia de aplicación

La metodología multinivel ofrece a los investigadores cuantitativos la oportunidad de conocer la realidad de estudio por nivel y entre niveles, permitiendo con ello, una mayor profundidad en el análisis del objeto de estudio. Dicho de otra forma, se observa y analiza el comportamiento de las variables que integran el fenómeno de estudio, en forma individual, colectiva y cruzada; y se hace uso de las características de todas las variables en sus diferentes niveles (micro y macro); lo cual distingue plenamente las particularidades de los sujetos y los elementos contextuales donde se desarrollan.

Utilizar estructuras jerárquicas en las comunidades educativas y en otros sistemas, posibilita el análisis de los datos desde diferentes perspectivas; por estudiante, por la suma de sus efectos individuales, por clase, por escuela, por distrito o por estado. Cada uno de estos efectos es observable desde diferentes ángulos y son una colección de los múltiples efectos que produce el contexto y los individuos.

Además, los modelos multinivel cuentan con herramientas muy precisas que permiten el análisis cruzado, entre variables de nivel micro (individual) y macro (contextual), una característica única y fundamental de la metodología e imposible para las técnicas estadísticas tradicionales. La diferencia es notable y es su mayor logro. En este sentido, ninguna de las técnicas estadísticas anteriores cuenta con estimaciones tan exactas para todos los efectos (fijos y aleatorios). De ahí la importancia de su utilización y aplicación en las diferentes ramas del saber.



3.1.4. Utilización de niveles

Refiriéndose sobre la cantidad de niveles de agrupación considerados para elaborar un modelo determinado, en principio y de acuerdo con la mayoría de los teóricos¹⁵, no existe un límite establecido. Es más, en la literatura analizada por este autor no se encontró ninguna referencia en cuanto a la cantidad de niveles que deben ser incluidos en una investigación. Las propuestas estudiadas únicamente hacen referencia al análisis de los coeficientes, sus efectos y al modelado de las estructuras. Ninguna delimita o recomienda una determinada cantidad de niveles.

A pesar de ello, en la práctica raramente se encuentran modelos con más de cuatro o cinco niveles. Son desconocidos para este autor los ejemplos que superan estas cifras. Quizá esto se deba principalmente a las complejidades técnicas enfrentadas cuando un modelo tiene un mayor tamaño y a la significación desmedida que implica delimitar una gran subdivisión niveles. Así mismo, cuando el número de variables ofrece tamaños diferentes, existen un enorme número de interacciones cruzadas entre niveles, lo cual agudiza el problema de elaboración e interpretación el modelo. Por esa razón no imperan patrones con muchos niveles en la literatura especializada.

Los modelos más comunes en la literatura especializada, cuentan con dos o tres niveles para su desarrollo jerárquico. Esta cifra es observada en diversos trabajos realizados en el plano internacional¹⁶ por los más variados autores. Dicha cantidad de niveles domina el panorama de aplicación de esta metodología y se ha convertido en la guía base para los demás investigadores, quienes generalmente no sobrepasan la estructura de tres niveles.

Por ejemplo, Ting (2001) en un estudio sobre la evaluación del docente universitario en China, construye un modelo en dos niveles, tomando en el primer nivel las características del estudiante, como: interés, expectativas, edad, años en la universidad. En el segundo incluye los determinantes del curso, como: tipo (ciencias duras o blandas), dificultad del curso, nivel, tamaño de la clase, profesor (experiencia,

¹⁵ Ver: Leyland y Goldstein (2001); Raudenbush y Bryk (2002); Raudenbush (1999); Kasim y Raudenbush (1998); Kreft y De Leeuw (1998); Goldstein y Rasbash (1996); Goldstein (1995); Goldstein y Healy (1993); y Longford (1993).

¹⁶ Ver: Goldstein (2004); Leyland y otros (2000); Raudenbush y Bhumirat (1992); Bryk, Raudenbush, Seltzer y Congton (1998a); Ting (2001); Selner-O'Hagan, Kindlon, Buka, Raudenbush y Earls (1998); Sampson y otros (1997); Raudenbush (1997); y Kindlon, Wrigth, Raudenbush y Earls (1996).



años de servicio). También Gaviria, Martínez y Castro (2004) con su estudio en Brasil emplea tres niveles. El primero lo componen las características de los estudiantes: raza, edad y horas de televisión. El segundo, las escuelas caracterizadas por: desarrollo de contenidos del profesor, utilización de libro de texto, recursos didácticos, tiempo de servicio, centro privado o no, estudios del padre. En el tercero se incluyen los estados al que pertenece la escuela.

En los problemas multinivel, las decisiones acerca de los miembros del grupo y su operacionalización involucra un amplio rango de asunciones teóricas y de problemas específicos que muchas veces se complican cuando no se cuentan con propuestas teóricas muy sólidas. Por ello, es importante contar con buenos recursos teóricos, los cuales permiten una estructuración sólida del agrupamiento de las variables por niveles, de esa forma se considerarán todos elementos posibles de influencia e interacción.

Ahora bien, una vez establecidas las agrupaciones y sus niveles, incluso si su establecimiento es al azar, ha de observarse la distinción plena de cada nivel, es decir lo exclusivo. Ésta diferenciación implica que el grupo y sus miembros están completamente identificados en su nivel de pertenencia y sus componentes no pueden ser encontrados en otra parte del modelo. Hacer caso omiso de esta relación es arriesgarse a pasar por alto, la importancia de los efectos del grupo o del individuo y puede, también invalidar muchas de las conclusiones y hallazgos encontrados durante el proceso de análisis de datos; debido especialmente a que este método permite, no solo el análisis de cada nivel, sino también, el análisis de la interacción entre los niveles y sus características (Goldstein, 2003). Bryk y Raudenbush (1992) hablando sobre este tema advierten que la falta de profundidad en la organización, análisis de los niveles y sus interacciones ha llevado a la acumulación de sesgo, a reducir la precisión del cálculo y a problemas con la unidad de análisis, lo que “fomenta el empobrecimiento de la conceptualización” (p. 3); lo cual lleva al investigador a observar algunos problemas como insolubles, mientras en realidad lo que sucede tiene mayor relación con la falta de atención en la organización de las variables y sus niveles, que con la solución del problema en sí.

3.1.5. Sus nombres a través del tiempo

En la literatura sobre la metodología se utilizan una variedad de términos tomados por los autores a través de los años y referidos a sus especialidades y su preocupación por el modelo de investigación. Por esa razón, hay varios como: modelos



lineales multinivel¹⁷; modelos de efectos mixtos y aleatorios¹⁸; modelos de coeficientes aleatorios o al azar¹⁹; modelos aleatorios de coeficientes de regresión²⁰. Desde nuestra óptica, cada calificativo hace referencia a uno u otro componente de la metodología, pero ninguno de estos logra abarcarla en su totalidad como para decantarnos a utilizar un término específico.

El nombre *modelos jerárquicos lineales* fue introducido Lindley y Smith (1972), como una contribución a los métodos de estimación *bayesiana* de los modelos lineales. El término ha sido muy empleado por los más diversos investigadores, quienes reconocen la amplitud del concepto para abarcar todos los componentes incluidos en la metodología. Al respecto, Bryck y Raudenbush (1992) consideran que el término propuesto por Lindley “expresa un rasgo importante en la estructura de los datos que es normal en una amplia variedad de aplicaciones” (p. 3), por esa razón, recomiendan utilizar la expresión cuando se trabaje con los modelos multinivel.

Por nuestra parte, consideramos conveniente el empleo del nombre modelos jerárquicos lineales para referirnos a la metodología, debido en primer lugar, a la recomendación anterior. En segundo, al amplio rango del concepto que acoge la mayoría de los términos empleados por los investigadores. Y finalmente, por su uso generalizado en la investigación educativa internacional.

3.2. ECUACIÓN DEL MODELO JERÁRQUICO LINEAL

Los modelos jerárquicos lineales tienen diferentes formas de representación matemática, obviamente, a mayor profundidad teórica mayor complejidad de representación matemática y de estimación; por tanto, inicialmente es conveniente conocer el modelo más simple o modelo nulo para ir desarrollando otros más complejos.

El modelo simple o nulo no tiene predictores o variables independientes, únicamente cuenta con la variable dependiente y una constante para la estimación de los parámetros. En él se encuentran dos ecuaciones, atendiendo directamente a los niveles de trabajo, o sea dos niveles. De hecho no existen modelos con menos de dos niveles, pues la teoría se estructura a partir de dos.

¹⁷ En inglés “multilevel linear models” por Bock (1989) y Goldstein (1987a,1995)

¹⁸ En inglés “mixed- effects models y random models” por en Elston y Grizzle (1962) y Laird y Ware (1982).

¹⁹ En inglés “random coefficient models” por Longford (1993)

²⁰ En inglés “random coeffecient regresión models” por Rosenberg (1973)



Ahora se va a proceder a modelar estructuras jerárquicas, fundamentados en un ejemplo sobre el rendimiento académico de los estudiantes. El ejemplo cuenta con una estructura en dos niveles, el primero los estudiantes y en el segundo la escuela de pertenencia. El mismo se puede representar por medio de dos ecuaciones matemáticas.

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

$$\beta_{0j} = \beta_{00} + u_{0j}$$

La ecuación de primer nivel sería:

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

Y simboliza:

y_{ij} Es la variable dependiente observada en el nivel inferior o micro nivel, o sea, la unidad mínima definida por el investigador, en nuestro caso hipotético, el rendimiento del estudiante i (en la clase de matemática, lengua) dentro de la escuela j . Notemos que desde el inicio la representación de los dos niveles; i para el estudiante y j la escuela de pertenencia.

β_{0j} define al intercepto, el punto de partida de la recta en el plano. Para nuestro ejemplo, representa la media general de rendimiento de la escuela j . El valor del rendimiento medio de todas las escuelas de una muestra determinada puede ser un número negativo o positivo.

e_{ij} Es el residuo o varianza residual, cuya media es cero y tiene una varianza σ_e^2 . Representa las diferencias en el rendimiento de los estudiantes, entre el valor estimado por la regresión y el valor real. En su modelo más simple asume la varianza de error al azar, es la misma para todos los estudiantes.

Como se puede inferir, la ecuación hace referencia directa al nivel uno de la ecuación de regresión, o sea, los estudiantes.

Analizada la ecuación de nivel uno, proseguimos con la de nivel dos:

$$\beta_{0j} = \beta_{00} + u_{0j}$$

Es la ecuación de nivel macro, referente a la información o características de las escuelas.



β_{00} es el intercepto del nivel dos, la gran media general de todas las escuelas j incluidas en el estudio.

u_{0j} es el residuo o varianza residual del nivel dos, la desviación del valor estimado para la escuela de su valor real. Tiene media cero y varianza $\sigma_{\mu_0}^2$. Cuando el puntaje de la desviación es alto indica que existe una marcada diferencia entre las escuelas.

Como observamos, el segundo nivel hace referencia directa a las escuelas y ofrece estimaciones de su media, con el respectivo residuo.

La riqueza de los modelos jerárquicos lineales, es que integra las dos ecuaciones en una sola, de esa forma se obtiene, una representación. Esta se representa de la siguiente forma:

$$Y_{ij} = \beta_{00} + u_{0j} + e_{ij}$$

Donde:

Y_{ij} corresponde a la variable dependiente. El rendimiento del sujeto i en la escuela j .

β_{00} identifica la gran media general del rendimiento entre las escuelas.

u_{0j} es el residuo o varianza residual de segundo nivel (la escuela) y representa la desviación entre la media entre las escuelas; las diferencias de estimación entre la predicción y el valor real.

e_{ij} hace referencia a la varianza residual del primer nivel: los estudiantes.

El primero de los coeficientes el β_{00} es llamado la parte fija del modelo. Los otros dos coeficientes, u_{0j} y e_{ij} representan las varianzas residuales denominada la parte aleatoria del modelo. Como se deduce, la ecuación de regresión del modelo nulo ofrece información sobre dos elementos, la parte fija y la aleatoria en la estructura jerárquica de los datos.

En la parte aleatoria del modelo se encuentran las varianzas de los residuales. Los residuales no son parámetros del modelo, pero sus varianzas sí. Los residuales son las diferencias entre el valor que toma la variable dependiente para una unidad y lo que el modelo predice para esa misma unidad (sujeto o escuela). Hay tantos residuales de nivel uno como sujetos y tantos residuales de nivel dos como escuelas, pero esos no son parámetros, los parámetros son sus varianzas. Por tanto, cuando MLwiN ofrece los



resultados de la parte aleatoria se refiere directamente a la varianza residual, no al residual.

3.3. ESTIMACIÓN DEL MODELO BASE O NULO

Una vez conocido e identificado el modelo nulo y sus parámetros (fijos y aleatorios) se procede a estimarlo, a fin de observar si la proporción de varianza, sin explicar que queda entre los niveles, es estadísticamente significativa; lo cual permite continuar con la elaboración de modelos con mayor complejidad, en donde intervienen todas las variables independientes consideradas en nuestra propuesta de estudio. Cuando no se evidencien diferencias en el modelo nulo, estadísticamente significativas, éste carece de valor y de significado, como para profundizar en el análisis los mismos.

Para construir un modelo nulo y estimarlo, el primer paso es la elaboración de una constante, o sea, una nueva variable, cuyos valores son ajustados a uno (1). Redefinidos todos los valores en una nueva variable, se le ubica en cada uno de los niveles jerárquicos propuestos, en nuestro caso dos niveles. De esa forma ya contamos con una constante introducida en el modelo y se procede con su estimación.

Estimado el modelo nulo, contamos con tres productos de los parámetros de la ecuación:

β_{00} La media general de rendimiento

u_{0j} Varianza residual de nivel dos

ε_{ij} Varianza residual de nivel uno

Cuando se estima el modelo con el programa MLwiN, cada uno de estos símbolos matemáticos es acompañado de un valor y de su respectivo error estándar.

Para conocer la capacidad explicativa del modelo nulo y analizar la proporción de varianza que queda sin explicar, ha de dividirse el producto de cada uno de los estimadores anteriores (β_{00} , u_{0j} , ε_{ij}) entre su error estándar. El resultado de esta división ha de ser un valor mayor a ± 1.96 . En caso afirmativo, nuestro modelo de estudio cuenta con suficiente varianza sin explicar y ahora se puede introducir las variables predictoras y profundizar en el fenómeno de interés.

En caso contrario, cuando los resultados de la división propuesta son menores a ± 1.96 , es innecesario plantear el modelo. Ante esta situación, se debe cambiar o reconsiderar el modelo nulo, pues la inexistencia de variabilidad en los datos no permite seguir trabajando con este modelo o con esos datos.



Regresando a nuestro ejemplo, supongamos que todos los resultados encontrados en la división de los parámetros del modelo nulo son superiores a ± 1.96 . En este sentido, el modelo explicaría gran parte de la varianza en las medias generales de rendimiento entre las escuelas y entre los alumnos; expresando: a) las escuelas difieren entre sí en cuanto a su rendimiento general y b) los alumnos también difieren entre sí en cuanto a su rendimiento en la clase.

Ante este argumento, podríamos señalar entonces que los resultados demuestran la existencia significativa de varianza en el nivel uno y dos. A partir de este momento estamos sabiendo que el modelo deja mucha varianza sin explicar, justificando con ello el desarrollo de modelos más complejos, introduciendo las variables explicativas o independientes, de modo tal que expliquen las particularidades de cada uno de los niveles propuestos.

3.4. LOS PARÁMETROS

Los parámetros para hacer las estimaciones del objeto de estudio, son dos: a) los efectos fijos (intercepto, pendiente) y b) los efectos aleatorios y los componentes de varianza/covarianza. Ambos parámetros permiten obtener estimaciones del efecto observado y muestran información específica sobre cada una de las variables de estudio, el nivel de pertenencia y su grado de asociación con la variable dependiente.

El valor denominado *efectos fijos* representa en su forma inicial al intercepto, o sea, plantea el punto donde la recta de regresión corta al eje de ordenadas. En el caso que no haya variables independientes, es decir, en el modelo nulo, o cuando las variables independientes están centradas respecto a la media general, entonces el intercepto representa la media de la variable dependiente.

Ahora bien, cuando se introducen variables explicativas en el modelo aparecen otros efectos fijos denominados pendientes, los cuales representan, el incremento de la variable dependiente por cada unidad que aumenta la correspondiente variable independiente, *ceteris paribus*. Por ejemplo, podríamos decir que el rendimiento medio en la escuela j se encuentra asociado al nivel socioeconómico. En este caso podríamos estudiar el grado de asociación entre el rendimiento y el estatus socioeconómico, para ver si el mayor o menor rendimiento se debe a este factor.

Los efectos fijos se representan en la ecuación matemática con las letras β_{00} , referida la constante y $\beta_1 X_{ij}$, $\beta_2 X_{ij}$, $\beta_3 X_{ij}$ etc. para las pendientes del modelo, las cuales simbolizan las variables independientes determinadas en el estudio, por ejemplo: estatus socioeconómico, estudios del padre, edad del hijo, aspiraciones, entre



otros. Estos efectos cuentan con un componente que varía aleatoriamente en su respectivo nivel, llamado coeficientes aleatorios.

Los *efectos aleatorios o al azar* son estimadores de la parte no explicada por el modelo y reúnen información de: los residuales, término de varianza residual y las estructuras de covarianza.

Los residuales de estimación se refieren a las diferencias existentes entre lo explicado por el modelo y el valor real obtenido durante la aplicación medida. Por ejemplo, podemos suponer que con nuestro modelo predecimos una calificación para un estudiante i en una escuela j de 418.34 puntos. No obstante, al consultar los datos reales sobre su rendimiento, el sujeto se encuentra por encima de la predicción, alcanzando una nota de 676.79 puntos. ¿Qué sucede? Encontramos una diferencia de 100.76 entre el pronóstico hecho por el modelo y el valor real obtenido por el estudiante. ¿Qué pasó? A estas diferencias se les llama residuales y son efectos aleatorios empleados en la estructura jerárquica para observar y analizar las divergencias entre lo estimado por el modelo y el valor real de la medida.

El componente de *varianza-covarianza* es un estimador que puede introducirse o no en el modelo. Los programas estadísticos de la metodología brindan esta posibilidad para integrarlos o no en la especificación del modelo. De hecho, muchas veces los investigadores introducen el parámetro en el modelo, pero cuando se realiza la estimación su valor no es significativo y debe ser excluido del modelo.

Cuando el valor de la estimación de una covarianza es negativo, se interpreta que estas diferencias se dan en forma más acusada en aquellos sujetos con un rendimiento más bajo. Por ejemplo, una covarianza negativa y significativa entre el estatus socioeconómico y el rendimiento académico, significa que la importancia del estatus socioeconómico es más grande en las escuelas que tienen rendimiento medio más bajo y menos importante en aquellas escuelas que tienen rendimientos medios más altos. En los estudiantes de las escuelas con alto rendimiento medio, el estatus socioeconómico estaría menos asociado a su rendimiento que en los demás sujetos.

Para realizar los cálculos de los *efectos aleatorios* los investigadores han propuesto diferentes estimadores, como el de máxima verosimilitud²¹, el método de máxima verosimilitud restringida²² y el método de estimación bayesiana²³, pero no van a ser considerados en estas líneas.

²¹ Longford (1987) y Goldstein (1986)

²² Mason (1983) y Raudenbush y Bryk (1986)

²³ Dempster (1981)



Un hecho relevante a tomar en cuenta es que a cada variable del modelo le corresponde un residuo en la ecuación. Con ello se logran intervalos de confianza y pruebas significación, más precisas que las obtenidas con los métodos tradicionales.

3.5. TRES NIVELES

El modelo jerárquico lineal en tres niveles, es una extensión del modelo básico explicado en los párrafos anteriores, al cual se añade un nivel más. Esta inclusión permite aumentar la capacidad explicativa del modelo y a su vez, incrementa la complejidad matemática de las ecuaciones, las cuales involucran una serie de elementos de cálculo que solo es posible con las computadoras y los programas específicos de la metodología multinivel.

Por ejemplo, un modelo en tres niveles puede estar constituido por estudiantes (primer nivel) anidados en clases (segundo) y éstas en escuelas (tercero). El nivel uno del modelo representa las relaciones entre las variables de nivel del estudiante, el nivel dos captura de influencia de los factores determinantes de la clase y el tercero, incorpora los efectos del nivel de las escuelas. Hipotéticamente, podemos destacar que la existencia de varianza entre las escuelas indica que se diferencian en su rendimiento. La varianza en el segundo nivel establece que las clases también difieren en el rendimiento medio de cada una de ellas dentro de las escuelas y la varianza en el nivel uno (los estudiantes) indica que los estudiantes muestran diferente rendimiento en clase. Claro está, el modelo asume una distribución jerárquica de los datos y la existencia de una variable dependiente: el rendimiento del estudiante.

Matemáticamente se podría representar el modelo nulo de la siguiente forma:

$$\begin{aligned}y_{ijk} &= \beta_{0jk} + e_{ijk} \\ \beta_{0jk} &= \beta_{0k} + u_{0jk} \\ \beta_{0k} &= \beta_{00} + v_{0k}\end{aligned}$$

Observe, que los dos niveles iniciales son similares a los estudiados anteriormente, con la diferencia de la inclusión de una tercera letra, la cual indica un nivel más, o el tercero.

Las letras representativas de las características de nivel se pueden definir en nuestro ejemplo hipotético de la siguiente forma:

$i = 1, 2, 3, n_{jk}$ son las unidades de primer nivel, los estudiantes.



$j = 1, 2, 3, j_k$ indican las unidades de nivel dos, las clases.

$k = 1, 2, 3, k$ se refieren a las de nivel tres, las escuelas.

Conocidos los elementos de cada nivel, podemos detallar sus elementos específicos:

Y_{ijk} es la variable dependiente, el rendimiento medio del estudiante i de la clase j en la escuela k . La escritura de la variable pone de manifiesto la cantidad de niveles del modelo jerárquico, por ello recomendamos observar cuidadosamente sus componentes.

β_{0jk} representa la media general de rendimiento del estudiante i en la clase j de la escuela k . Este parámetro identifica el rendimiento medio de los estudiantes de una determinada clase dentro de una escuela.

El parámetro siguiente, β_{0k} incorpora la media de rendimiento de las clases j en la escuela k ; este elemento identifica el rendimiento general de las clases dentro de la escuela. Claramente se puede observar clases con un rendimiento general alto y otras, por el contrario, bajo. Nótese que aquí estamos hablando solamente del segundo nivel, las clases.

β_{00} define el rendimiento general de las escuelas e indican la media general de rendimiento entre las escuelas k . Este elemento informa sobre el nivel macro de la ecuación.

Los parámetros identificados hasta el momento hacen referencia a la parte fija del modelo, a la medida sin error.

La parte aleatoria de la ecuación, o sea, los elementos de error en la medida son representados por:

v_{0k} es un efecto al azar de tercer nivel, el residuo o varianza entre el valor estimado para la escuela y su valor real, asumiendo que el efecto tiene una distribución normal cuya media es 0 y varianza $\sigma_{\mu_{0k}}^2 [v_{0k} \approx N(0, \tau_\beta)]$.

u_{0jk} es el efecto al azar de la clase, o sea segundo nivel, identifica la desviación de la clase j de escuela a escuela k , con media cero y $\sigma_{\mu_{0jk}}^2$. Al igual que el parámetro anterior se asume que el efecto $u_{j0k} \approx N(0, \tau_\pi)$

Finalmente, el residual de nivel uno: los estudiantes. Es un efecto aleatorio que determina la desviación del rendimiento del estudiante i en la clase j , respecto a su



valor real. El efecto asume que tiene una distribución normal cuya media de 0 y varianza σ_e^2 .

Identificados los componentes anteriores, la ecuación de tercer nivel puede ser resumida en una sola definida por:

$$y_{ijk} = \beta_{0jk} + v_{0k} + u_{0jk} + e_{ijk}$$

La ecuación se describe como:

Y_{ijk} es la variable dependiente, el rendimiento medio del estudiante i de la clase j en la escuela k . La escritura de la variable pone de manifiesto la cantidad de niveles del modelo jerárquico.

La parte fija del modelo β_{0jk} representa la media general de rendimiento del estudiante i en la clase j de la escuela k . Este parámetro determina el rendimiento medio de los estudiantes de una clase dentro de una escuela.

La parte aleatoria de la ecuación, o sea, los elementos de error en la medida son representados por:

v_{0k} es un efecto al azar de tercer nivel, el residuo o varianza entre el valor estimado para la escuela y su valor real, asumiendo que el efecto tiene una distribución normal cuya media es 0 y varianza $\sigma_{\mu_{0k}}^2$ [$v_{0k} \approx N(0, \tau_\beta)$].

u_{0jk} es el efecto al azar de la clase, o sea segundo nivel, identifica la desviación de la clase j de escuela a escuela k , con media cero y $\sigma_{\mu_{0jk}}^2$. Al igual que el parámetro anterior se asume que el efecto $u_{j0k} \approx N(0, \tau_\pi)$

Finalmente, e_{ijk} el residual de nivel uno: los estudiantes. Es un efecto aleatorio que determina la desviación del rendimiento del estudiante i en la clase j , respecto a su valor real. El efecto asume que tiene una distribución normal cuya media de 0 y varianza σ_e^2 .

Cada vez que se incluye un nivel más, aumenta la cantidad de parámetros a estimar, volviendo más compleja su explicación matemática.

3.6. VARIANZA DEL MODELO



Para conocer la proporción de varianza explicada por el modelo nulo o de otros más complejos con variables explicativas, se pueden realizar unos cálculos básicos con los resultados obtenidos en los efectos aleatorios. Simplemente, para dicho propósito ha de guiarse por el conjunto de formulas recomendadas en Bryk y Raudenbush (1992).

En la partición de la varianza, los autores consideran importante iniciar los cálculos con los índices del modelo nulo. Uno con tres niveles tendría los siguientes: $\nu_{0k}, \mu_{0jk}, e_{ijk}$. Sustituyendo los símbolos por los valores relativos, se puede proceder con la estimación de la varianza en cada nivel del modelo empleado; de esa forma obtendremos diferentes coeficientes por nivel jerárquico del modelo de estudio.

Para analizar la proporción de varianza explicada por nivel:

$\sigma^2 / (\sigma^2 \tau_\pi \tau_\beta)$ es la varianza explicada entre los estudiantes, o sea, primer nivel.

$\tau_\pi / (\sigma^2 \tau_\pi \tau_\beta)$ corresponde a la proporción de varianza de las clases entre las escuelas, segundo nivel.

$\tau_\beta / (\sigma^2 \tau_\pi \tau_\beta)$ es la referida entre las escuelas, o sea, el tercer nivel.

El símbolo σ^2 corresponde al valor obtenido en la estimación de e_{ijk} ; τ_π a μ_{0jk} ; y finalmente τ_β a μ_{0k} .

El producto de cada una de las ecuaciones precisa el porcentaje de varianza explicada en el nivel correspondiente, permitiendo con ello determinar la cantidad de varianza aportada por cada nivel al modelo de propuesto.

3.7. CORRELACIÓN INTRACLASE

La correlación intraclase denominada con “ ρ ” puede ser conocida por medio de la ecuación:

$$\rho = \sigma_{00} / (\sigma_{00} + \sigma^2)$$

Leyland y Goldstein (2001) indican que la correlación ρ intraclase es un estimado de población de la varianza explicada por la estructura del grupo, ello es igual a la proporción estimada del nivel de varianza del grupo comparada con el estimado total de varianza.

CAPÍTULO 4: EJEMPLO DE MODELOS JERÁRQUICOS

4.1. RECOMENDACIONES INICIALES

Antes de ingresar a en la construcción de un modelo en dos o tres niveles, es importante considerar la base de datos y su ordenamiento jerárquico, debido a la estructura del software, el cual requiere un ordenamiento sistemático y riguroso agrupado en niveles jerárquicos, es decir, por ejemplo: estudiantes agrupados por profesor; ese sería un modelo en dos niveles.

Antes que nada, veremos el comportamiento de la base de datos. Eso se puede observar en “data manipulation” el segundo apartado “names”, al pinchar se obtiene:

Figura N°1. Base de datos

MLwiN - C:\Users\Dircción\Documents\Base tesis.WS

FileEditOptionsModelEstimationData ManipulationBasic StatisticsGraphsWindowHelp

StartMoreStopIGLS

Estimation control..

Names

Column

NameDescriptionToggle Categorical

Data

ViewCopyPasteDelete

Categories

ViewCopyPasteRegenerate

Window

Used columns

Help

Name	Cn	n	missing	min	max	categorical	description
ESTUDIAN	1	10562	0	1	374	False	
PREGUN	2	10562	0	1	30	False	
CARRERA	3	10562	0	1	22	False	
PROFESO	4	10562	0	1	27	False	
RESPUES	5	10562	0	0	4	False	
ANOUNI	6	10562	0	0	4	False	
ESTSEX	7	10562	0	0	1	False	
OBLIGA	8	10562	0	0	1	False	
CANTEST	9	10562	0	0	3	False	
DIFICULT	10	10562	0	0	4	False	
CANTRAB	11	10562	0	0	4	False	
ESPNOTA	12	10562	0	1	4	False	
VALORH	13	10562	0	0	4	False	
EDAEST	14	10562	0	17	41	False	
ASITE	15	10562	0	0	4	False	
DPTOCARR	16	10562	0	1	3	False	
DPTOPROF	17	10562	0	1	3	False	
PROEDAD	18	10562	0	0	3	False	
NOMBRA	19	10562	0	0	4	False	
INVESTG	20	10562	0	0	4	False	
SALRIO	21	10562	0	0	3	False	
---	---	---	---	---	---	---	---

randomfixediteration 7NamesHierarchy viewerEquationsCalculate

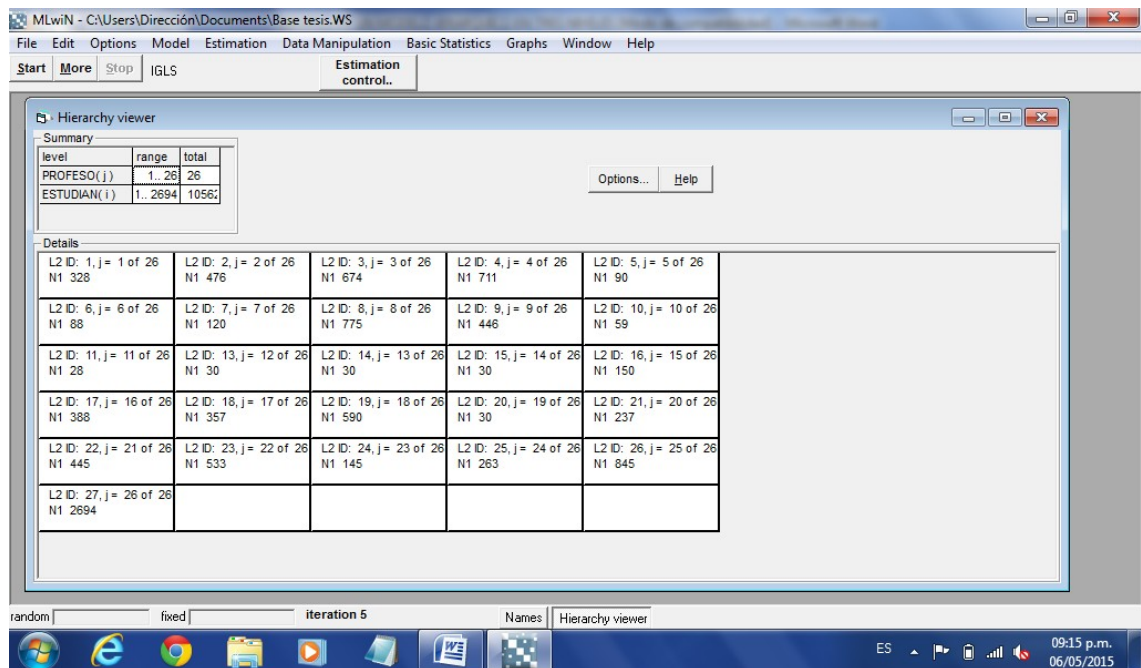
Entre otras cosas puede verse que son 10562 estudiantes, agrupados de 1 hasta 374, en profesores de 1 hasta 27. Como elemento de contexto, estos también están agrupados en las 22 carreras. También puede verse que la gran mayoría de variables, tanto del estudiante como del profesor, se agrupan en categorías de 1 a 4.

Ahora se debe pinchar “Model”, en el apartado de “Hierarchy Viewer”, el cual brinda una panorámica del ordenamiento jerárquico que tiene la base de datos. En nuestro ejemplo aparece en la primera caja de diálogo: profesor (j), de 1 a 26; estudiante (i) de 1 a 2694. para un total de 10562. Luego ofrece un detalle de



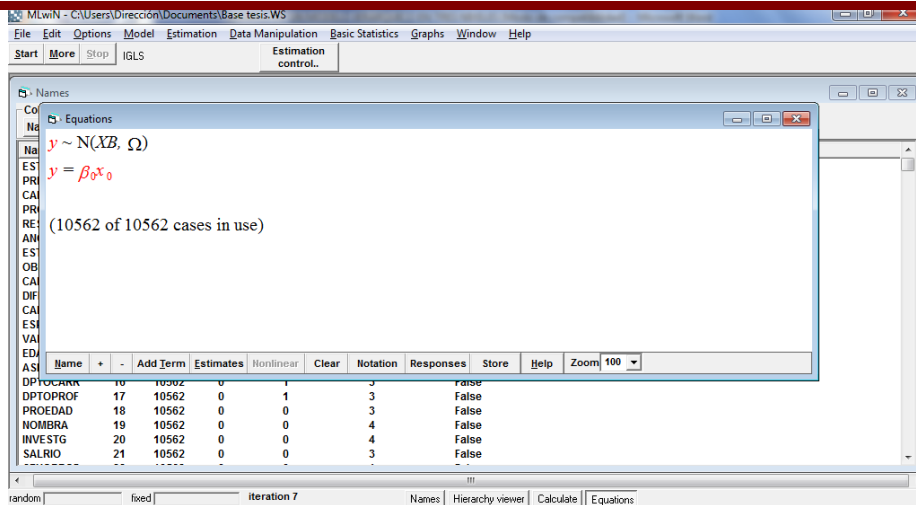
agrupación de los estudiantes en profesores. Al profesor uno tiene 328 estudiantes, el dos 476, y así sucesivamente hasta llegar a los 26 docentes. Todo ello se puede ver en el siguiente figura N°2.

Figura N°2 Visor Jerárquico



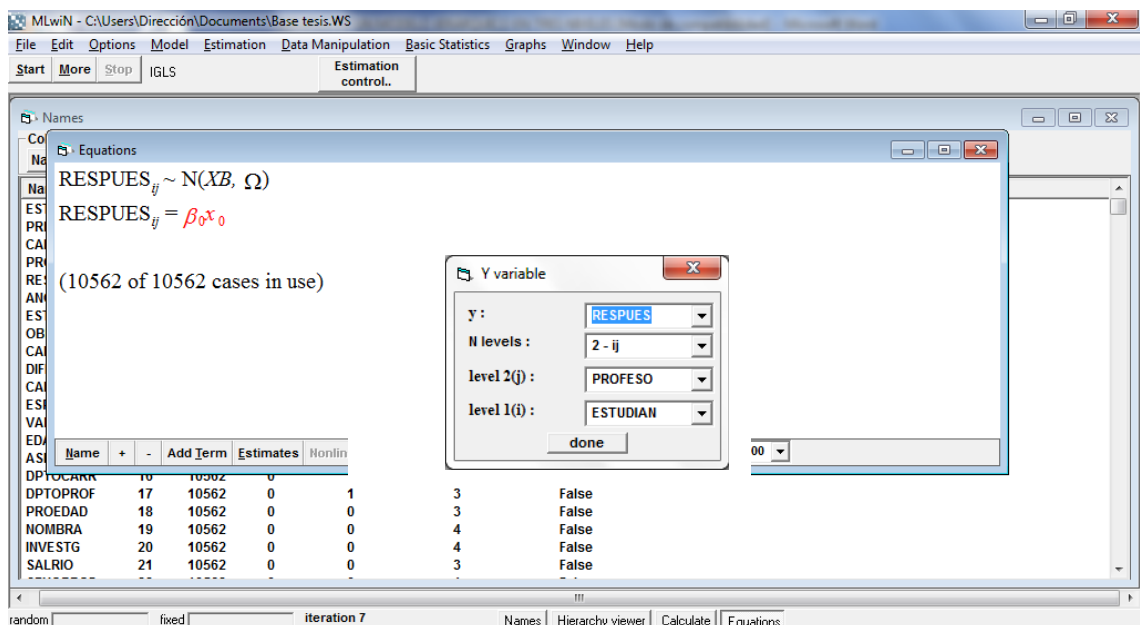
Una vez observada la base de datos y su ordenamiento, se pincha en “Model”, en su primera opción “Equations”. Aparece un cuadro con la ecuación por desarrollar, pero en color rojo; indicando con ello que no se han introducido las variables y los niveles para estimarlo.

Figura N°3 Equations



Lo primero será definir la variable dependiente “Y”, se pincha sobre ella. Aparece un cuadro de diálogo para identificar la variable dependiente y cuantos niveles tiene el modelo. En este ejemplo la variable dependiente será RESPUES. Luego se definen los niveles, o sea dos, ordenados en segundo nivel por el PROFESO y el primero ESTUDIAN.

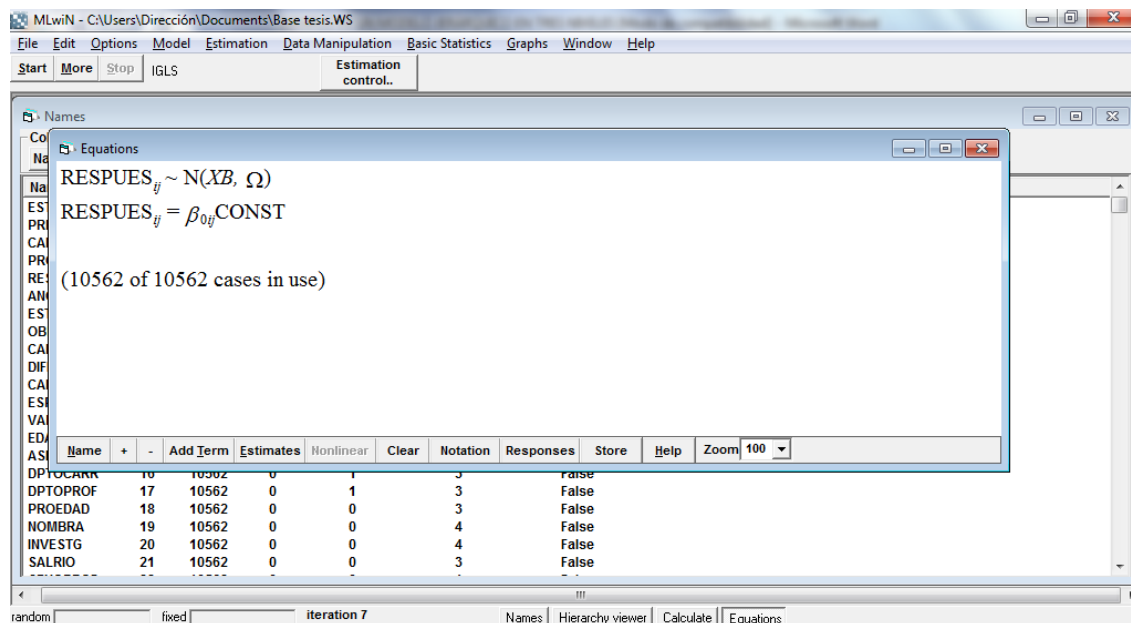
Figura N°4 Variable dependiente y niveles del modelo



Como lo notan, todavía se tiene en la ecuación un elemento con el color rojo, eso señala la ausencia de definición del parámetro. Este beta, corresponde a la CONSTAN o sea, la constante del modelo. Incluyendo este, también se debe definir el parámetro

fijo, la varianza a nivel de PROFESO (j), también el nivel de ESTUDIAN (i). Incluidos estos últimos elementos, ya no se cuenta con el color rojo en la ecuación.

Figura N°4 Constante, varianza fija y aleatoria

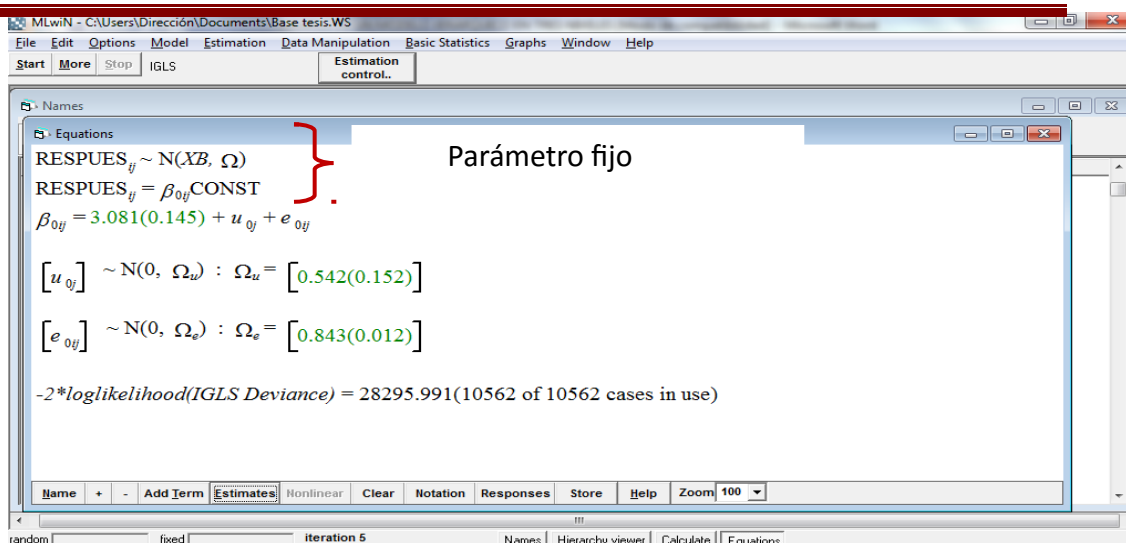


4.2. MODELO NULO

Realizado el proceso, se ha concluido con la construcción del modelo nulo, el cual permite conocer si el modelo ajusta a los datos o no. Para estimar el modelo, solamente debemos de hacer click en START, y con ello se realiza la estimación de los parámetros.

Estimados los parámetros, es necesario observarlos y analizarlos para ello es necesario hacer click en ESTIMATES, en la parte inferior del software. Cuando se ha alcanzado un nivel de estimación aceptable, los valores se ponen en color verde, como lo muestra la figura siguiente.

Figura N°6 Modelo nulo estimado



Estimado el modelo nulo es importante determinar la significatividad de sus parámetros, pues si no logran alcanzar el umbral, el modelo no ajusta a los datos, lo cual puede afectar la investigación debido a que podría ser necesario aplicar otro instrumento.

Para analizar la capacidad explicativa del modelo nulo y la proporción de varianza que queda sin explicar, se divide el producto de la estimación entre el parámetro fijo o aleatorio (β_{00} , u_{0j} , ε_{ij}). Cuando los resultados de la división propuesta, en cada uno de sus parámetros, son menores a ± 1.96 , es innecesario replantear el modelo o cambiar de datos. Veamos un ejemplo con sus valores del de la figura anterior: El parámetro fijo es $3.081/0.145 = 21.24$, el producto es superior a ± 1.96 , por lo tanto, es significativo. Los parámetros aleatorios de segundo y primer nivel son $0.542/0.152 = 3.56$ y $0.843/0.012 = 70.25$, estos también son superiores a ± 1.96 . Al ser todos los productos superiores al valor de aceptación, se puede inferir que los datos ajustan al modelo en dos niveles. Ahora es posible proceder a desarrollar modelos más complejos.

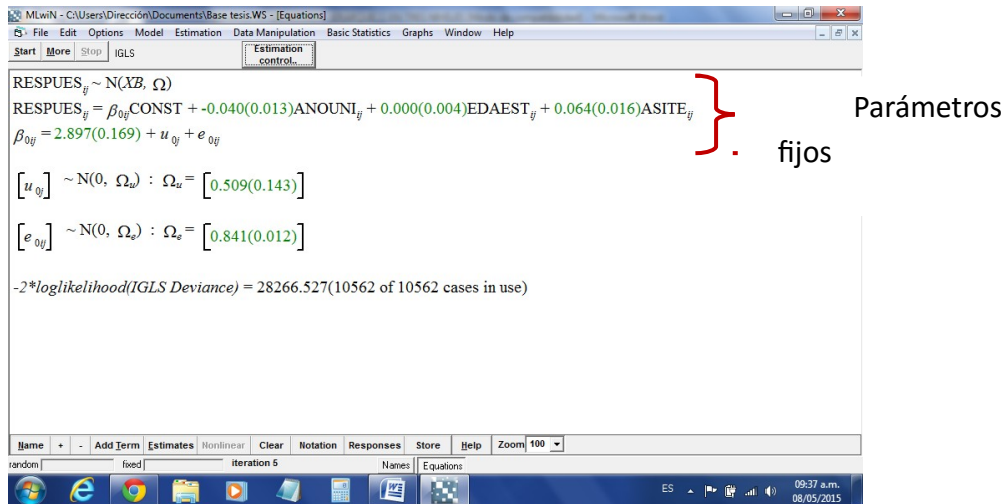
4.3. VARIABLES PREDICTORAS DE PRIMER NIVEL

Contando con un modelo nulo significativo, se introducirán las variables predictoras, según las variables con que cuenta la base de datos para este ejemplo. De primer nivel, pueden considerar: años de universidad, sexo del estudiante, asistencia al curso.

Ahora se pincha en “ADD TERM” (se encuentra abajo del escritorio). Aparece un cuadro de diálogo, e introducimos la variable de elección. Cada vez que se introduce

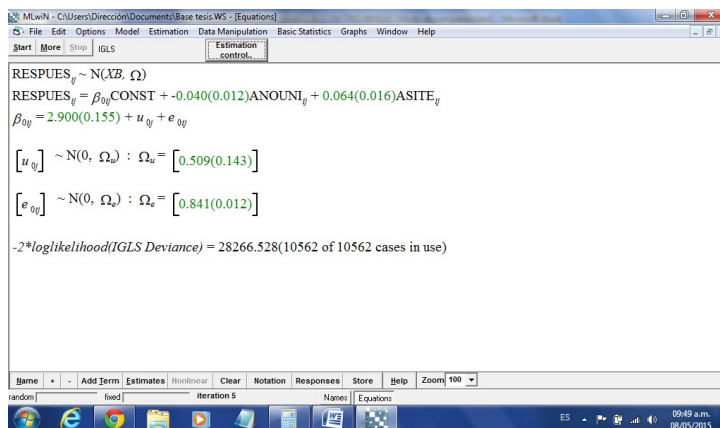
una variable, se debe pinchar en “ADD TERM”. Añadidas las variables, procedemos a estimar el modelo, en “START”. El resultado se puede observar en la siguiente figura:

Figura N°7 Modelo con variables de primer nivel



Notarán la reducción del valor de la media, debido a las otras variables, pasó de 3.081 a 2.897, de un puntaje máximo es cinco. Cada una de las variables debe ser analizada su significatividad: AÑOUNI -0.040/0.013= -3.07; ASITE 0.064/0.016=4. Ambos valores son superiores a ± 1.96, son significativas y deben continuar en el modelo. EDAEST 0.000/0.004= 0 no es superior a ± 1.96 y debe eliminarse del modelo. Se realiza la estimación nuevamente.

Figura N°8 Modelo con una variable eliminada

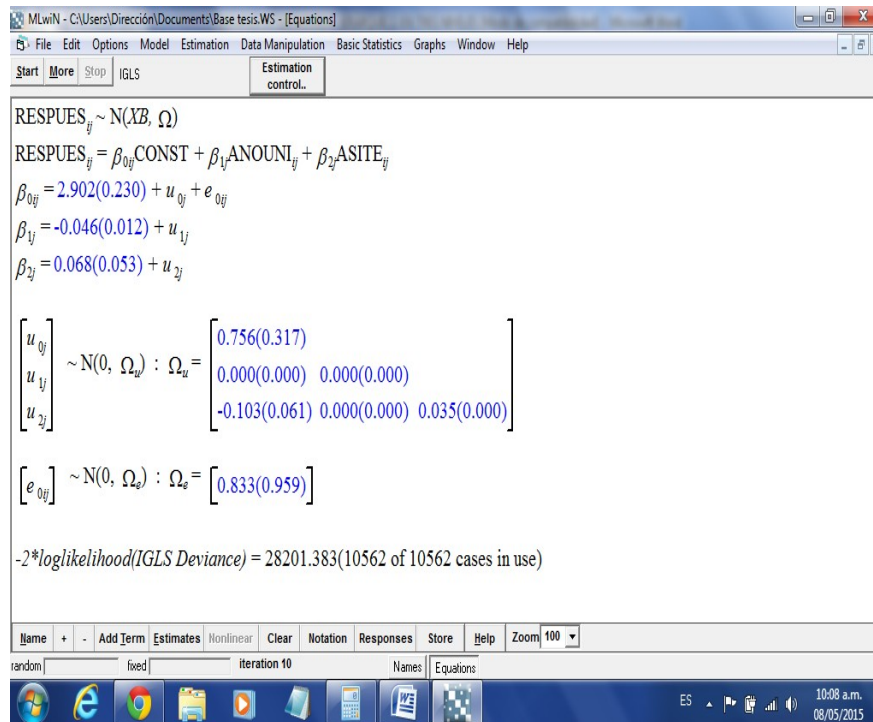




El modelo expresa, en esta evaluación de la docencia universitaria por parte de los estudiantes, que la media general de todos los profesores de la muestra es 2,90, de cinco puntos máximo. Señala además, que por cada año en la universidad el estudiante reduce su opinión del profesor en un 0.040, significa que conforme pasan los años, se siente menos satisfecho con su profesor (a) o la madurez de los años le permiten una evaluación más rigurosa de sus docentes. Caso contrario, los estudiantes que asisten poco a clases valoran positivamente al docente en un 0.064, pero quienes asisten en un 90% o sea en el valor 4 de la categoría, su evaluación del docente se incrementa en 0.256 ($0.064 \times 4 = 0.256$).

Ahora se introducirán los parámetros aleatorios de primer nivel, se pincha directamente en la variable y un cuadro de diálogo muestra los niveles donde varía. En este caso, los estudiantes varían a nivel del profesor, por ello han de incluirse en el segundo nivel.

Figura N°9 Parámetros aleatorios del modelo

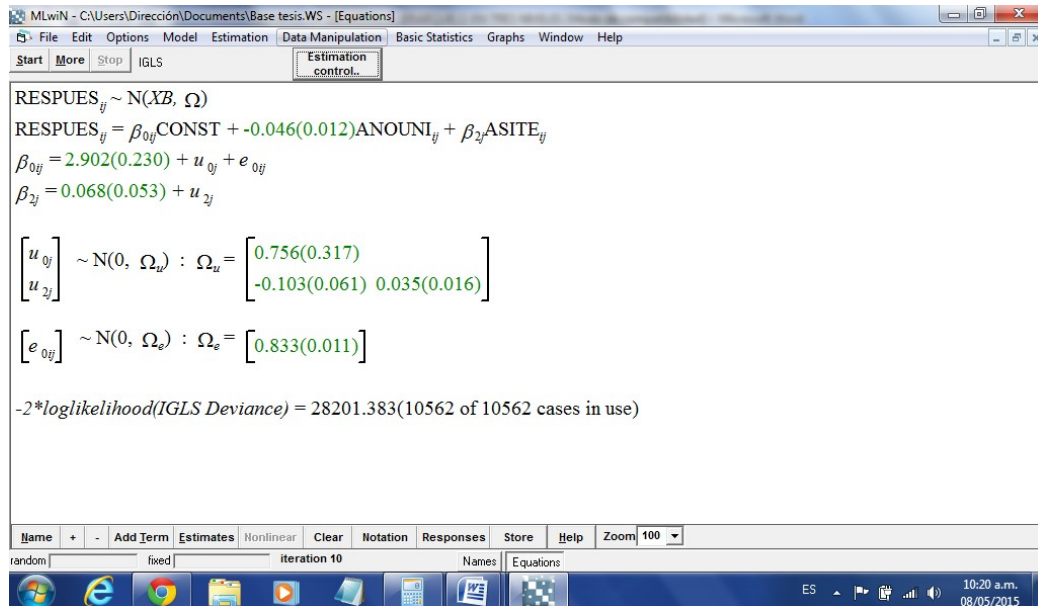


Parámetros aleatorios

Observé el parámetro aleatorio u_{1j} de AÑOUNI su valor es cero; por tanto se elimina del modelo. Probaremos si el otro se comporta igual o diferente. Se realiza la estimación otra vez.

En la siguiente figura observaremos el parámetro aleatorio de ASISTE, u_{2j} . Este muestra dos valores, la covarianza y la varianza. El primero, $-0.103/0.61=1.68$, un valor inferior a ± 1.96 , no es significativo y debe eliminarse del modelo. El segundo, $0.035/0.016= 2,18$, es superior a ± 1.96 , e indica que existen diferencias opinión entre los estudiantes, respecto del profesor, dentro de cada uno de los criterios de la variable asistencia a clases. Las opiniones de los que asisten un 90% a clases, no es igual. Lo mismo sucede con las demás.

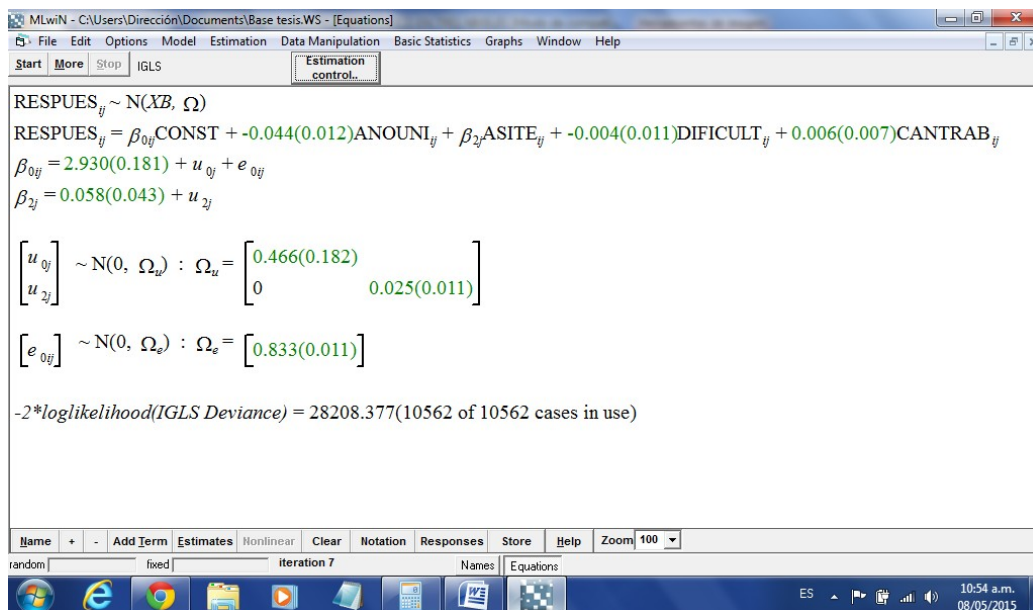
Figura N°10 Parámetro aleatorio del modelo



4.4. VARIABLES DE SEGUNDO NIVEL LOS DOCENTES

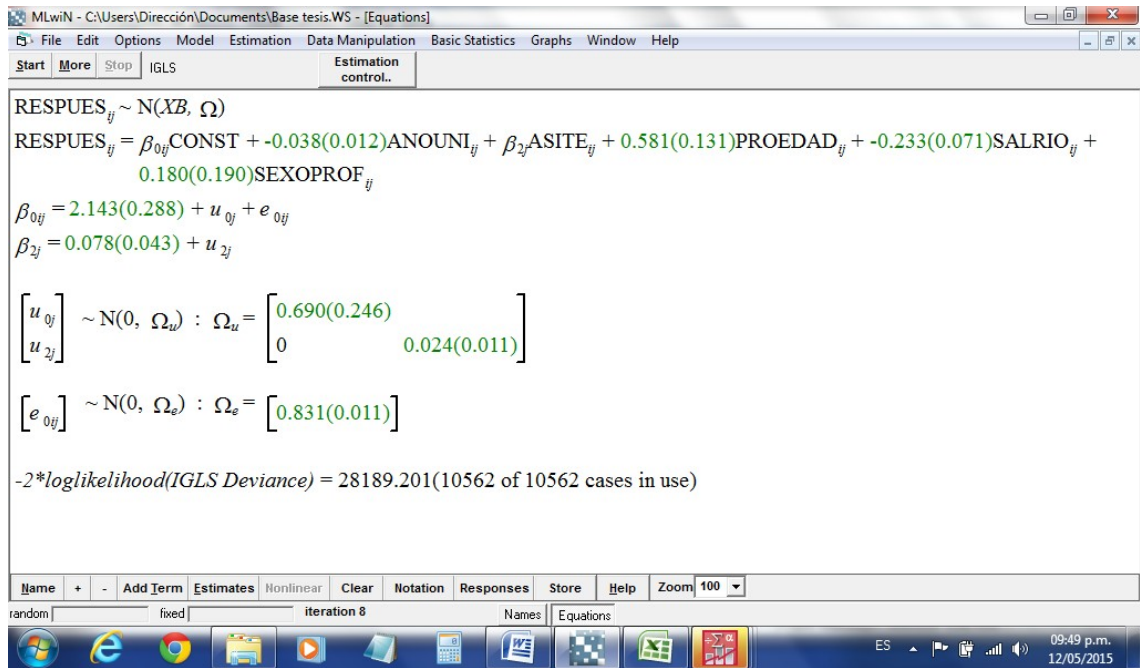
Ahora se verá el comportamiento de las variables de segundo nivel DIFICULT (dificultad del curso) y CANTRAB (cantidad de trabajo en el curso). La primera -0.004/0.011=0.363 y la segunda 0.006/0.007=0.857, como notarán ambas no son significativas.

Figura N°11 Variables de segundo nivel



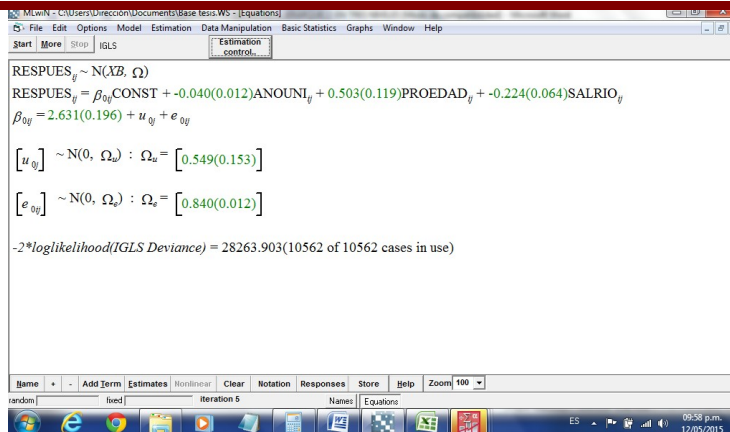
Añadiremos tres nuevas variables de segundo nivel: edad del profesor, salario y su sexo. Notarán que controlando otras variables en el modelo, la Asistencia al curso, por parte de los estudiantes ya no es significativa. Ha de eliminarse. Como también el sexo del profesor.

Figura N°12 Variables de segundo nivel



Eliminadas las variables no significativas del modelo, se procede con la estimación de: años en la universidad, edad del profesor y salario. Los datos mantienen valoración que los estudiantes con mayor cantidad de años están menos satisfecho con sus docentes, quizá han madurado y sus criterios para evaluar la docencia han cambiado. Además, tienen una mejor evaluación aquellos profesores de más edad, pues por cada punto en la escala, la evaluación del profesor sube un 0.503. Caso contrario, los docentes con mejores salarios, puntúan -0.224, por cada valor de la escala. Es decir aquellos docentes cuyos salarios se ubican al inicio de la escala, la valoración de sus estudiantes es 2.63, pero si se ubican en el número tres de la escala salarios, estos tienen bajan en un -0.672

Figura N°13 Variables significativas de primer y segundo nivel

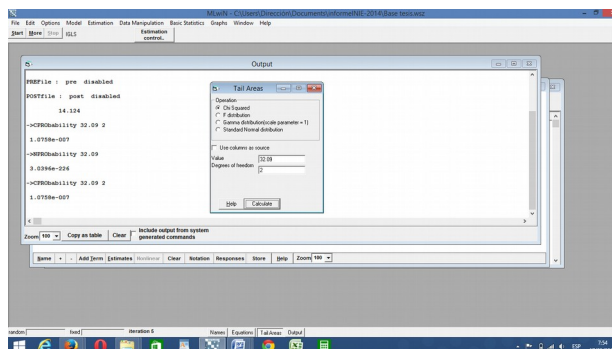


Finalizada la introducción de variables de primero y segundo nivel, se observa la razón de máxima verosimilitud, para determinar el modelo que mejor ajusta a los datos.

4.5. RAZÓN DE MÁXIMA VEROSIMILITUD

La razón de máxima verosimilitud es utilizado para la comparación de los modelos y es necesario para determinar el que mejor ajusta. Por ejemplo: el modelo nulo de los ejemplos obtuvo un valor de 28295.99 (ver la imagen del modelo nulo) y el último de la figura N°13 28263.90. Sería entonces $28295.99 - 28263.90 = 32.09$. El resultado muestra que ha decrecido el valor estadístico con le nuevo modelo. Este nuevo modelo tiene tres grados de libertad, bajo la hipótesis nula de que los parámetros extras tienen un valor de población de cero. Se puede manifestar que $\rho(\chi^2 \geq 32.09) \leq 0.001$. Estimado el valor 32.09 establece un producto de $1.0758e-007$, un número inferior a 0.001, por esa razón el nuevo modelo es superior al modelo nulo

Figura 14 Razón de máxima verosimilitud





Del ejemplo anterior se puede señalar la existencia de un modelo ajustado a los datos y que el último modelo con las variables significativas de nivel y dos, es mejor que el modelo nulo.

5. CAPÍTULO 5: REFERENCIAS CITADAS Y CONSULTADAS

Bentler, Peter; Satorra, Albert & Yuan, Ke-Hai. (2009) "Smoking and Cancers: Case-Robust Analysis of a Classic Data Set". *Structural Equation Modeling*, 16, p382–390

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.

Bollen, K.A. & Long, J.S. [Eds.] (1993). *Testing structural equation models*. Newbury Park, CA: Sage.

Cañas Madueño, Juan Antonio y Fernández Sánchez Javier (1994). *Metodología de las Ciencias Sociales*. España: Universidad de Córdoba.

Cea D'Ancona, María de los Ángeles (2001). *Metodología cuantitativa, estrategias y técnicas de investigación social*. Madrid: Editorial Síntesis, S.A.

Cronbach, Lee. (1990). *Essentials of psychological testing (5th ed.)* New York: Harper Collins

DeLuca, Christopher. Interpretive validity theory: mapping a methodology of validating educational assessments. *Educational Research*. 2011, 53 (3): 303-320.

Fan, Xitao; Thompson, Bruce & Wang, Lin. Effects of Sample Size, Estimation Methods, and Model Specification on Structural Equation Modeling Fit Indexes. *Structural Equation Modeling*, 1999, 6 (1): 56-83.



Gliner, Jeffrey; Morgan, George & Leech, Nancy (2009). *Research Methods in Applied Settings*. New York: Taylor & Francis Group.

Hu, Li-tze & Bentler, Peter. "Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives". *Structural Equation Modeling* 1999, 6, 1-55.

Jacob, H. (1994): "Using published data: errors and remedies". En Lewis-Beck, M.S. (comp): *Research Practice*. California, Sage, pp. 339-390

López-Barajas, David y Ruiz Carrascosa, Juan. La evaluación de la docencia Universitaria. Dimensiones y variables mas relevantes. *Revista de Investigación Educativa RIE*, 2005, 23 (1): 57-84

Losada, José Luis y López-Feal, Rafael (2003). *Métodos de Investigación en Ciencias Humanas y Sociales*. Madrid: Thomson Editores.

Marsh, H. W., Balla, J. R., & Hau, K. T. (1996). An evaluation of incremental fit indices: A clarification of mathematical and empirical properties. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques* (pp. 315-353). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Marsh, Herbert; Arief D., Gregory; Liem, Martin. Andrew; Morin, Alexandre J. S. & Nagengast, Benjamin. Methodological Measurement Fruitfulness of Exploratory Structural Equation Modeling (ESEM): New Approaches to Key Substantive Issues in Motivation and Engagement. *Journal of Psychoeducational Assessment*, 2011, 29, 322-346.



Marsh, Herberth., Hau, K.-T., & Wen, Z. L. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and danger in overgeneralizing Hu and Bentler's 1999 findings. *Structural Equation Modeling*, 11, 320-341.

Messick, Samuel. Test validity: A matter of consequence. *Social Indicators Research*, 1998, 45: 35-44.

Schmitt, Thomas. "Current Methodological Considerations in Exploratory and Confirmatory Factor Analysis" *Journal of Psychoeducational Assessment*, 2011, 29(4), 304-321.

Schreiber, James y otros (2006). Reporting structural equation modeling confirmatory factor analysis results: a review. *The Journal of Educational Research*, 99, 6, 323-337.

Sivo, Stephen; Fan, XITAO; Lea, Witta & Willse, John. "The Search for "Optimal" Cutoff Properties: Fit Index Criteria in Structural Equation Modeling". *The Journal of Experimental Education*, 2006, 74(3), 267-288.

Sun, Jun. "Assessing Goodness of Fit in Confirmatory Factor Analysis". *Measurement on Evaluation in Counseling and Development*, 2005 (37), 240-256.

Thompson, B. (Ed) (2003). Score reliability: Contemporary thinking about reliability issues. Thousand Oaks, CA: Sage.

Vanhoff, Jan; Verhaeghe, Goedeke; Verhaeghe, Jean Pierre, Valcke, Martin & Van Petergem, Peter. "The influence of competences and support on scholl performance feedback use". *Educational Studies*, 2011, 37(2): 141-154.



Vanhoof, Jan; De Maeyer, Sven & Van Petergem, Peter (2011): "Variation in the conduct and the quality of self-evaluations: a multi-level path analysis". *Educational Studies*, 2011, 37 (3). 277-289.

Yuan, Ke-Hai Yuan & Bentler, Peter. "On Chi-Square Difference and z Tests in Mean and Covariance Structure Analysis when the Base Model is Misspecified". *Educational and Psychological Measurement*, 2004, 64, 737-757.



8.2. ANEXO 2

3 de agosto del 2012

Señora
Dra. Alicia Vargas
Directora
I.N.I.E.

Estimada señora:

El presente es un informe de las actividades realizadas en el Taller de Modelos Jerárquicos Lineales, módulo 2, impartido por el Dr. José Luis Gaviria Soto, del 28 de julio al 3 de agosto del 2012.

El taller comprendió tres horas diarias de trabajo directo con los estudiantes. Del 30, 31 de julio y 1 de agosto. De 9 a 12md se atendieron consultas individuales de los participantes y se resolvieron problemáticas de los campos de investigación.

La temática teórica desarrollada en el laboratorio de informática del INIE, fue la siguiente:

I. DESCRIPCIÓN DEL TALLER

En la actualidad, la metodología de la investigación exige diferentes herramientas para el desarrollo y análisis de una investigación cuantitativa. apoyados en la informática, para el análisis de los fenómenos educativos, psicológicos, económicos, etc. Por ejemplo: los modelos de ecuaciones estructurales (AMOS, EQS, LISREL) se utilizan para estudiar las relaciones causales entre las variables y los modelos jerárquicos lineales (MlWin, HML), para observar los complejos fenómenos de investigación a nivel contextual y a nivel particular, una actividad posible únicamente con esta metodología, la cual ha demostrado los fuertes errores cometidos por el empleo de los métodos tradicionales análisis.

Ante la existencia de estas propuestas metodológicas informáticas, surge la idea de implementa en el I.N.I.E. un taller sobre la utilización de técnicas y métodos cuantitativos de análisis en la investigación, para fomentar el uso de de Mlwin, con el fin de desarrollar nuevas destrezas en el complejo mundo del análisis de datos.



El taller ofrece a los (as) estudiantes e investigadores (as) un panorama básico acerca de las técnicas utilizadas para analizar los diferentes fenómenos de estudio, las propiedades psicométricas en los instrumentos de medida; llámese cuestionarios, test de aptitudes u otros. Durante el período de trabajo se estudiarán, analizarán y observarán los diversos elementos teórico-prácticos, utilizados por los paquetes informáticos para realizar los diversos tipos de análisis.

Para participar en el taller se requiere conocer algunas herramientas de informática y la posibilidad de asistir a todas las sesiones, dado que los elementos básicos de estadística, como los paquetes, serán tratados durante la implementación del taller, por ello no es necesario poseer conocimientos profundos en estadística para entender la globalidad de los procesos de trabajo.

Finalmente, se espera que los (las) participantes conozcan y desarrollen los principios básicos de los modelos jerárquicos lineales, ello le brindará nuevas armas para enfrentar el complejo mundo de la investigación.

II. OBJETIVOS

Brindar herramientas cuantitativas de investigación y análisis, fundamentadas en Mlwin, con el fin de desarrollar nuevas destrezas en los (las) participantes, que le faciliten la comprensión de este paradigma.

Generar actitudes positivas hacia las técnicas y metodologías cuantitativas, para su implementación en las áreas de trabajo de los (las) participantes.

III TEMÁTICAS GENERALES

3.2. MODELOS JERÁRQUICOS LINEALES (MIWin)

3.2.1.- Elaboración de la base de datos

3.2.2.- Construcción del modelo nulo

3.2.3.- Creación de constante

3.2.4.- Elaboración de los niveles a estimar

3.2.5.- Parámetros a estimar

3.2.6.- Estimación y reelaboración del modelo

3.2.7.- Análisis del modelo jerárquico lineal



IV. METODOLOGÍA DE TRABAJO

La metodología del taller considera diferentes técnicas de trabajo como la disertación sobre temas o áreas problemáticas, el empleo de las computadoras y los paquetes informáticos. Así como otros elementos de apoyo: lecturas, trabajo en casa, trabajos extra clase.

Para lograr un nivel de profundización importante, es necesaria la participación activa del estudiante en los componentes relacionados con: lecturas, discusiones, elaboración de bases de datos y la construcción de un instrumento de medición, fundamentado en una temática de interés.

V. EVALUACIÓN

Asistencia	55%
Envíos de trabajos extra clase	35%
Progreso en el manejo de los tutoriales informáticos	10%
TOTAL	100%

VI. TEXTOS BÁSICOS

Acevedo Alvarez, R. & Rodríguez, N. M. (2006). Factores de sesgo asociados a la validez de la evaluación docente universitaria: un modelo jerárquico lineal. *Archivos Analíticos de Políticas Educativas*, 14 (34). Recuperado [20 julio del 2007] de <http://epaa.asu.edu/epaa/>

Acevedo, R. (2009): *Los Modelos Jerárquicos Lineales: Fundamentos Básicos para su Uso y Aplicación*. Instituto de Investigaciones Psicológicas. U.C.R. San José: Costa Rica.

Acevedo, R. y Fernández, M. J. (2004): "La percepción de los estudiantes universitarios en la medida de la competencia docente: validación de una escala". En *Revista de Educación*. Nº 28 (2): 145-166. San José: Editorial de la Universidad de Costa Rica.

Acevedo, R.(2003): *La evaluación de la competencia docente universitaria: un modelo jerárquico lineal*. Madrid: Editorial de la Universidad Complutense de Madrid. Recuperado [13 diciembre del 2007] de <http://www.ucm.es/BUCM/tesis/edu/ucm-t26870.pdf>

Hox, J.J. (1994). Hierarchical regression models for interviewer and respondent effects. *Sociological Methods and Research*, 22(3), 300-318.

Gaviria, José Luis (2008) *Los Modelos Jerárquicos Lineales*. Madrid: Editorial La Muralla.



Gaviria, J. Martínez-Arias, R. & Castro, M. (2004, Mayo 5). Un Estudio Multinivel Sobre los Factores de Eficacia Escolar en Países en Desarrollo: El Caso de los Recursos en Brasil. *Education Policy Analysis Archives*, 12 (20). Recuperado [11-03-2008] de <http://epaa.asu.edu/epaa/v12n20/>.

Martínez, Rosario y Gaviria, José Luis (2008). *Concepto y evolución de los modelos de valor añadido en Educación*. Revista de Educación, 348. Enero-abril 2009, pp. 15-45

Ting, K. F. (2001). A Multilevel Perspective On Student Ratings of Instruction. Lessons From the Chinese Experience. *Research in Higher Education*. Vol. 41, 5, pp. 637-653.

Aparte del taller impartido por las tardes, se realizaron las siguientes actividades:

Lunes 30 de Julio

Revisión en línea y atención de problemas de investigación de los (as) participantes del taller. El horario de trabajo fue de 9 a 12:00 md.

Martes 31 de Julio

10:00 Reunión con la directora del CEA, para tratar temas referentes a la medición.

13:00 Reunión con la MSc. Lillian Aguilar, subdirectora del MEP, en el departamento de dirección de Gestión y Evaluación de la Calidad. Ahí se trataron temas relacionados con bases de datos, evaluación y medición. Además, sobre los problemas que tiene el país en relación al análisis de datos cuantitativos.

Miércoles 1 de agosto

10:00 a.m. Conversatorio, realizado en el Mini Auditorio de la Facultad de Ciencias Sociales, en este nos acompañaron en laa mesa principal la Dra. Alicia Vargas, Directora del INIE, Dr. Raziel Acevedo, investigador. Al evento asistieron 23 personas, de diferentes instituciones: CONARE, INIE, CEA y otros, quienes dedicaron a plantear y desarrollar diferentes pensamientos acerca de la problemática de la evaluación, la educación y la gestión.

14:00 Reunión Dra. Eilyana Montero, para atender conceptos y pensamientos sobre acciones propias, destinadas a establecer y buscar puntos de encuentro, los cuales generen alternativas para futuras actividades, donde se involucre la gestión, evaluación y análisis de datos.

Sin otro particular, se despide

Raziel Acevedo Álvarez. Gestor



ANEXOS DE NOTICIAS DIGITALES DEL EVENTO



Universidad de Costa Rica
Facultad de Educación
Instituto de Investigación en Educación



Invitan al:

Conversatorio: Propuestas medulares sobre medición en la educación superior

Invitado especial:

Dr. José Luis Gaviria Soto

Catedrático, Universidad Complutense de Madrid

Miércoles 1 de agosto de 2012 a las 10:00 a.m.

Miniauditorio Facultad de Ciencias Sociales



INIE INFORMATIVO

Boletín N° 018

NUEVA DIRECTORA DEL INIE

La **Dra. Alicia E. Vargas Porras** asumió la Dirección del Instituto de Investigación en Educación para el período comprendido entre el 1 de julio de 2012 y el 30 de junio de 2016. Nuestra nueva Directora es Doctora en Educación, graduada del Doctorado Latinoamericano en Educación de la Universidad de Costa Rica. Ha laborado por 22 años en la Universidad de Costa Rica, primero en el Centro de Evaluación Académica y posteriormente en el Sistema de Estudios de Posgrado. Es docente de la Maestría de Evaluación Educativa y de la Maestría en Evaluación de Programas y Proyectos de Desarrollo y ha sido investigadora de este Instituto por varios años, siendo coordinadora del Programa Cambio, Desarrollo y Gestión de la Educación Superior durante los últimos tres años. Auguramos los mayores éxitos para doña Alicia.

CONVERSATORIO

PROPUESTAS MEDULARES SOBRE MEDICIÓN EN LA EDUCACIÓN SUPERIOR

Invitado especial: **Dr. José Luis Gaviria Soto**, Catedrático, Universidad Complutense de Madrid

- Fecha: miércoles 1 de agosto de 2012
- Hora: 10:00 a.m.
- Lugar: Miniauditorio Facultad de Ciencias Sociales

CLAUSURA DE CURSOS DE FECOOPSE

El pasado viernes 13 de julio de 2012 se realizó la clausura y entrega de certificados de las capacitaciones "Microcrédito y el Programa de Capacitación Dirigencial en cumplimiento con el Reglamento del Gobierno Corporativo", a las 6 p.m. en el Auditorio de la Escuela de Arquitectura. En esta oportunidad se contó con la asistencia del Lic. Ramsés Ramírez Flores en Representación de la



UNIVERSIDAD DE
COSTA RICA



Instituto de Investigación
en Educación

Tiene el gusto de invitar al taller:

“Taller Internacional: Los modelos jerárquicos lineales aplicados a la investigación”



Módulo 1:

Impartido por el profesor:
Dr. Raziel Acevedo,
Universidad de Costa Rica
18 y 25 de Junio / 16 y 23 de Julio de 2012
De 4:00 p.m. a 6:00 p.m.
Lab. de cómputo del INIE

Módulo 2:

Impartido por el profesor:
Dr. José Luis Gaviria Soto,
Universidad Complutense de Madrid
30 y 31 de Julio y 1 de agosto de 2012
De 3:00 p.m. a 6:00 p.m.
Lab. de cómputo de la Facultad de Educación

Información:

2511-1412
inie@ucr.ac.cr
facebook.com/inie.ucr



8.3. ANEXO 3 PONENCIA EN CHILE

Modalidad comunicación: ponencia

Línea temática: III Avances en metodologías, diseños, técnicas y análisis de datos para la investigación, diagnóstico y evaluación educativa

Palabras claves: Evaluación docente universitaria, estadística, modelos jerárquicos lineales, metodología de investigación.

TÍTULO: **EN LA BÚSQUEDA DEL SESGO: EVALUACIÓN DOCENTE UNIVERSITARIA**

NOMBRE COMPLETO: RAZZIEL ACEVEDO ÁLVAREZ

NACIONALIDAD: COSTARRICENSE

GRADO ACADÉMICO: DOCTOR EN EDUCACIÓN

INSTITUCIÓN DE PERTENENCIA: UNIVERSIDAD DE COSTA RICA, SEDE DE GUANACASTE

PUESTO DE TRABAJO: PROFESOR E INVESTIGADOR

DIRECCIÓN ELECTRÓNICA DE CONTACTO: **RAZZIEL.ACEVEDO@UCR.AC.CR**

I. EL punto de partida

Conocer la actividad docente ha sido y es, una preocupación constante en la gestión de las instituciones de educación superior, debido a la idea de que una parte de la calidad de éstas, descansa en las manos de su equipo docente. En efecto, la preocupación de la influencia en el aprendizaje de los estudiantes, parece haber dirigido y motivado la propuesta acerca de la necesidad de valorar su actividad, para determinar los altos o bajos logros institucionales con su acción.

En este sentido Greenwald, Hedges y Laine (1996) afirma “los profesores, después de los estudiantes, constituyen el elemento más importante en el sistema educativo” (p. 366), por consiguiente alrededor de estas dos figuras, gira todo lo demás: planes de carrera, objetivos, contenidos, normas, actividades, matrícula, infraestructura, entre otras. De ahí la necesidad e importancia por reunir información sistemática para la toma de decisiones y ello no es otra cosa que el proceso de



evaluación docente, iniciado según Good y Mulryan (1990)²⁴, hace más de cien años por Kratz en 1896.

Dichosamente, a pesar de la tendencia que responsabilizó a la docencia de los éxitos o fracasos institucionales, con la llegada de la evaluación institucional se supera esta idea, al integrar la evaluación docente como un elemento más del complejo engranaje valorativo universitario, compartiendo así, los aciertos o desaciertos institucionales.

Dadas las consideraciones anteriores, no es posible dotar de significación la evaluación del profesorado, si no es insertada en el proceso de evaluación institucional y en las gestiones de mejora de la calidad de la docencia en el contexto de la Universidad. Donde estos conozcan de primera línea los procedimientos, compartan las acciones y no sean un ente aislado de la actividad. Al respecto Lapotin y Haessig (1998) destacan que la participación de los profesores, es esencial en el proceso de evaluación institucional. Es necesario demostrarles la importancia de la evaluación y como ellos pueden ser parte de la instrucción y de la significación de su rol, como gestores de las acciones educativas.

Con este propósito, han surgido diferentes técnicas para su evaluación: cuestionarios de opinión de estudiantes, entrevistas abiertas, grupos focales, evaluación por pares, la autoevaluación, el portafolio, la recolección electrónica, entre otras. Sin embargo el tema que nos preocupa es los cuestionarios de opinión de estudiantes para evaluar la actividad de la docencia universitaria.

1.1. Los cuestionarios de opinión de estudiantes

Son un recurso recogida de información para evaluar la docencia universitaria a partir de la opinión de los estudiantes. Cohen (1981), Fedman (1988) y Marsh (1984) concuerdan al señalarla como la técnica más utilizada por las Universidades, la más profundamente estudiada, para conseguir datos sobre el nivel de calidad de la enseñanza universitaria, desde la visión estudiantil.

²⁴ Luego se desarrollaron otras evaluaciones docentes a saber Elliot (1915), Boyce (1915), Ruggs (1920) y Barr (1931). McKeachie (1990) cita otros estudios de principios del siglo XX como Seashore y otros (1910) y Klaper (1920), pero afirma, la investigación sobre la docencia universitaria se desarrolla hasta los años veinte. Marsh (1987) destaca también que es en esos años cuando los programas de evaluación docente por parte de los estudiantes se incluyen formalmente en las universidades de Harvard, Washington, Purdue, Texas y otras instituciones de Educación Superior.



Al mismo tiempo del progreso y aplicación de la evaluación docente, surgieron dos corrientes de investigación: la primera intenta relacionar los resultados de estas evaluaciones con los factores de competencia o eficacia docente y la segunda, observa los factores potenciales de sesgo²⁵.

En referencia a la primera línea, desde el inicio a tratado de identificar las características o competencias de un “buen profesor”, en la transformación e incidencia sobre el pensamiento de sus estudiantes. Dichas características, en los últimos años han sido influenciadas por el proyecto Tuning Europa y Tuning América Latina (Beneitone y otros, 2007), OCDE (2005) y la UNESCO (2005).

Sobre la base de las consideraciones anteriores Zabalza (2005) propone las siguientes competencias: planificación del proceso enseñanza aprendizaje, seleccionar y presentar los contenidos, ofrecer informaciones y explicaciones comprensibles, alfabetización tecnológica y dominar el manejo didáctico de las TIC’S, gestionar las metodologías de trabajo didáctico, relacionarse constructivamente con los alumnos, ofrecer tutorías y acompañamiento a los estudiantes, reflexionar e investigar sobre la enseñanza e involucrarse institucionalmente.

Por otra parte Marsh (1987) en uno de los instrumentos con mayor aplicación alrededor de mundo plantea: aprendizaje y valor, interacción con el grupo, rapport individual, exámenes calificaciones, carga de trabajo dificultad, organización y claridad, entusiasmo, amplitud de enfoque, trabajo extra clase- lecturas. En la misma línea, García Ramos (1998), con el CEDA, enumera seis dimensiones implicadas en la evaluación de la competencia del docente universitario, que se describen como: Programa – organización de la enseñanza, dominio de contenidos – claridad expositiva, motivación de aprendizaje – incremento del interés del alumno, interacción con el grupo de clase, atención individual al alumno, evaluación – exámenes.

Inmerso en este cuerpo teórico, la segunda línea de investigación analiza la incidencia o no, en la evaluación docente de las variables contextuales del estudiante, docente, e institución como: el sexo (de los profesores y estudiantes), los años de estudio, el interés previo, tamaño del grupo, entre otros. Estas otras más, causan gran preocupación en los cuerpos docentes, porque según ellos, podrían distorsionar la valoración de los estudiantes respecto a su actividad profesional y son características sobre las cuales no tienen ningún dominio.

²⁵ Ver por ejemplo: Watchel, 1999; Brodile, 1999; Centra, 1998; Kolitch, 1999; Cranton y Smith, 1986; March, 1987; Aleamori, 1999; Marsh, 2000; Marsh y Roche, 2000; Abrami, 1997; Ting, 2001.



1.2. Las consideradas fuentes de sesgo

Existen características externas a la docencia universitarias influyentes en las evaluaciones? Como lo mencionamos, es una interrogante de intranquilidad para el claustro académico, por ser un elemento ajeno al control de este y que podría incidir en su evaluación.

Cabe agregar, Abrami y otros (1990, p. 225-226) agrupan estas características explicativas de la evaluación docente, en cuatro: del instructor (rango, sexo, tipo de curso, experiencia, horas para la enseñanza), del estudiantado (sexo, años en la Universidad, tipo de carrera, personalidad, razones de matrícula, nota esperada), del cursos (tipo de curso, duración, electiva o no, cantidad de trabajo, práctica profesional o no) y del ambiente (estación del año, horario de clase, tamaño de la clase, ambiente psíquico).

Una de las variables con cierto peso ha sido la expectativa de nota del estudiante, pero eso no lleva a que el profesor obtenga buenas calificaciones simplemente por dar altas notas a los estudiantes. Watchel (1998) brinda explicaciones alternativas, incluyendo la posibilidad de que la acción docente inspire a los estudiantes a trabajar más fuerte y a obtener mejores resultados. Otros han señalado la cantidad de estudiantes en clase y la interacción del profesor, pero es obvio, clases con muchos alumnos hace no posibilita tanto la interacción, como cuando hay pocos.

Contrariamente a lo esperado, quizá la influencia encontrada se deba menos a las características o variables mencionadas, y más a una metodología no adecuada para su estudio, lo cual conduce a interpretaciones erróneas de los datos. Al respecto Villa y Morales (1993, p. 67) destacan algunas de las irregularidades metodológicas o interpretativas observadas en este tipo de investigaciones: 1) se utiliza el dato del alumno de forma individual y no la media de clase como unidad de análisis; 2) en la interpretación de los datos se pueden confundir “relación entre” con “causalidad”; también suele confundirse la “significación estadística” con “relevancia”; 3) La relación de alguna de las características de contexto de evaluación docente, no es un dato en contra de la validez. Los autores recomiendan examinar si la relación esperada es plausible, si es lo que lógicamente se podría esperar; 4) generalización de los resultados, aún cuando el estudio se ha hecho en unas pocas clases o en una sola institución. El dato descriptor de una situación limitada, no se puede extrapolar indebidamente a todas las instituciones.



Comprensiblemente, se requieren investigaciones más cuidadosas para determinar la naturaleza de la relación entre las variables de contexto y las valoraciones de estudiantes de la docencia universitaria.

1.3. Acercamiento a la Fiabilidad y Validez

A pesar de las inconsistencias y los temas sin resolver en la voluminosa literatura existente sobre los cuestionarios de opinión de los estudiantes en la evaluación de la docencia universitaria, la mayoría de los teóricos²⁶ se inclinan en diverso grado a considerar algunas conclusiones, más o menos aceptadas, resumidas por Marsh:

"...las encuestas a los estudiantes son: (1) multidimensionales; (2) fiables y estables; (3) su función primaria es mostrarle al instructor como está enseñando su curso; (4) son relativamente válidas sobre una variedad de indicadores de eficacia docente, (5) se ven parcialmente poco afectados por la cantidad de variables hipotetizadas como sesgo potencial, tal como expectativas de la nota del curso, tamaño de la clase, cantidad de trabajo e interés previo del estudiante y (6) se ha demostrado la utilidad en la mejora de la eficacia docente, cuando existe un aumento concreto de estrategias institucionales en áreas específicas para que los profesores se dirijan hacia la mejora..." (2001, pág. 184).

Abrami, d'Apollonia y Cohen (1990) apuntan, las evaluaciones de los alumnos son válidas en cuanto reflejan su satisfacción con la enseñanza que reciben y, en este sentido no puede criticarse su validez. Lo que se trata de demostrar es que no sólo reflejan la satisfacción de los alumnos, sino están relacionadas con la eficacia de la instrucción. El problema de la validez de las opiniones de los estudiantes en la evaluación de la docencia podría estar relacionado con la mejora del aprendizaje. La clave en este caso, es poder determinar en qué medida los estudiantes son capaces de percibir mejoras en su aprendizaje, como resultado de una práctica docente adecuada.

II. Método.

El estudio de la literatura sobre el tema, conduce a varias interrogantes, pero principalmente a una: ¿la evaluación de la docencia, por parte de la Universidad de Costa Rica, está asociada a las variables de contexto del estudiante, el profesor y la unidad académica de adscripción? En respuesta, se propone como objetivo determinar el grado de asociación entre el instrumento del CEA²⁷ y las variables contextuales del estudiante, profesor y unidad académica, mediante un modelo jerárquico lineal en tres

²⁶ Abrami, d'Apollonia y Cohen, 1990; Cashin, 1988; Cohen, 1987; Feldman, 1989, 1997, 1998; Marsh, 1984, 1987, 2001; y McKeachie 1979.

²⁷ CEA, siglas del Centro de Evaluación Académica de la Universidad de Costa Rica, ente encargado de evaluar la docencia en la institución.



niveles: en el nivel uno, las características del estudiante, en el dos las del profesor y en el nivel tres la unidad académica de adscripción, sea Escuela o Facultad.

El uso de los modelos jerárquicos se propone, en primer lugar, porque la estructura de los datos está anidada en un nivel de jerarquía en tres niveles: estudiantes, profesores y Escuela o Facultad. En segundo lugar, la naturaleza de la muestra y la variedad de las relaciones entre cada nivel, solamente pueden estudiarse con este tipo de metodología, la estadística tradicional es de poca ayuda en este caso.

2.1. Muestra de estudio

El instrumento del CEA fue aplicado a 12319 estudiantes, de los cuales el 57% eran mujeres y el 43% hombres, con una media de edad de 22 años, matriculados el 29% en Ciencias Sociales, 20% en Ciencias de la Salud y 14% en Ingeniería. Así mismo, 260 docentes fueron evaluados, con una edad media de 44 años, 47 % hombres y el 53% mujeres, perteneciente el 39% a la primera categoría salarial, de cuatro posibles en la Universidad de Costa Rica. Estos se encuentran anidados en 18 unidades académicas o escuelas, siendo el 25% de la escuela de Estudios Generales, 24% de Ingenierías y 14% de Sedes Regionales.

2.2. Características de Nivel

Las variables de modelo, han sido ordenadas de acuerdo con cada uno de los niveles de correspondencia: el primero las características del estudiante, el segundo, las del profesor y tercero la Escuela o Facultad. La siguiente tabla agrupa las variables según el nivel de correspondencia:

Tabla 1 Características contextuales de cada nivel

NIVEL 1	NIVEL 2	NIVEL 3
Estudiante	Profesor	Escuelas o Facultades
Edad Sexo Carrera de pertenencia Área del conocimiento Interés previo Repetiría curso Condición laboral Años de estudio Conocimientos previos Grado de dificultad evaluación Dificultad del curso Interés de salida	Edad Sexo Área Académica Horas de nombramiento Proyectos de investigación Proyectos de Acción social Cargos administrativos Otros cargos Grado obtenido Categoría salarial Puntaje Carga laboral	Ciclo Área Académica Horario Sede Regional



2.3. Estimación y análisis

2.3.1. Descripción del modelo

Conocidas las características de cada uno de los niveles, se procede con la descripción del modelo y su correspondiente ecuación de regresión:

$i = 1, 2, 3, n_{jk}$ son las unidades de primer nivel, los estudiantes.

$j = 1, 2, 3, j_k$ son las unidades de segundo nivel, los docentes.

$k = 1, 2, 3, k$ son las de tercer nivel, las escuelas o facultades.

Se propone y asume la distribución jerárquica de los datos, procediendo con la inclusión de una variable dependiente, en este caso el puntaje dado por los estudiantes a la actividad docente. Ésta se escribe de la siguiente forma: y_{ijk}

Ahora bien la ecuación básica del modelo sería:

$$y_{ijk} = \beta_{0jk} + e_{ijk}$$

$$\beta_{0jk} = \beta_{00k} + u_{0jk}$$

$$\beta_{00k} = \beta_{000} + v_{0k}$$

Donde:

y_{ijk} Es la variable dependiente, simboliza la nota dada por el estudiante (i) respecto del profesor (j) en la Escuela o Facultad (k).

β_{0jk} Incorpora la media de evaluación del profesor (j) respecto de la Escuela o Facultad (k).

β_{00k} la gran media general entre las unidades académicas (k).

$e_{ijk}, u_{0jk}, v_{00k}$ Constituyen individualmente, la varianza residual de su correspondiente nivel. Donde $\sigma_e^2, \sigma_{\mu_{0jk}}^2, \sigma_{\mu_{0k}}^2$ son los efectos al azar que determina la desviación del puntaje observado del estimado, asumiendo una distribución normal cuya media es de 0 y su varianza σ^2 .

La ecuación de tercer nivel se integra en una sola de la siguiente forma:

$$y_{ijk} = \beta_{0ijk} + v_{0k} + u_{0jk} + e_{ijk}$$

2.3.2. Modelo nulo



La estimación del modelo nulo o no condicional, es un paso preparatorio muy conveniente en el análisis de los modelos jerárquicos lineales, porque permite obtener una estimación de la gran media general β_{00k} , ofreciendo información sobre la variabilidad del resultado en cada uno de los tres niveles.

Al estimar el modelo nulo, la parte fija, o sea la gran media de competencia docente corresponde un valor β_{00k} 8,722 (de diez puntos posibles) y 0,069 de error estándar, con un 95% de confianza en el intervalo. El puntaje obtenido, indica la buena opinión de los estudiantes, con respecto a la labor de los docentes en la Universidad de Costa Rica. La parte aleatoria del modelo, llamado varianza residual obtiene: en primer nivel $e_{ijk} = 1,507$ (0,022); en segundo $u_{0jk} = 0,639$ (0,031) y en tercero $v_{0k} = 0,062$ (0,027).

Interesa ahora conocer la significatividad de los valores obtenidos, tanto de la parte fija, como de la parte aleatoria. Para ello se divide el valor obtenido entre su error estándar y el producto, ha de ser mayor o igual a ± 1.96 . Continuando con el procedimiento, β_{00k} 8,722/0,069=126,40. El valor es significativo porque es mayor a ± 1.96 .

Una vez conocida la parte fija de modelo, se indaga la parte aleatoria o varianza residual. En el primer nivel e_{ijk} 1,507/0,022=68,5: es un valor significativo; en el segundo u_{0jk} 0,639/0,031=20,61, es significativo y v_{0k} 0,062 /0,027=2,29, también es significativo. Estos resultados manifiestan la existencia de diferencias de opinión, entre los estudiantes cuando evalúan a un profesor (primer nivel). La varianza residual de segundo nivel, expresa que profesores difieren entre sí y que las Escuelas o Facultades, poseen evaluaciones diferentes, en las valoraciones hechas por los estudiantes de sus profesores. Estos valores indican el ajuste de los datos al modelo y por tanto, se puede proceder a desarrollar modelos más complejos, los cuales incluyan las variables contextuales de la docencia universitaria.

Por otra parte, la varianza explicada de primer nivel es de un 68,25%; la de segundo 28,94% y tercero 2,8%. La mayor proporción de varianza, es explicada por las diferencias entre los estudiantes al evaluar la acción de un docente universitario en el aula de una facultad o escuela específica.

2.3.3. Variables de primer nivel, estudiantes:



Siendo significativos los valores obtenidos por el modelo nulo, se procede a construir un modelo con las variables explicativas de primer nivel, las cuales se observan en los siguientes resultados:

Tabla 2: Resumen de las estimaciones del nivel Uno

NIVEL 1 estudiante	β_{00k}	sd	Significativa>1.96
Edad	0.017	0.003	Si
Sexo	0.001	0.025	No
Dum-Chicos_1	0.000	0.000	No
Carrera de pertenencia	0.002	0.003	No
Área de conocimiento	-0.025	0.064	No
Interés previo	0.893	0.030	Si
Repetiría curso	2.554	0.065	Si
Condición laboral	0.011	0.031	No
Años de estudio	-0.048	0.043	No
Conocimientos previos	0.718	0.032	Si
Grado de dificultad evaluación	0.335	0.041	Si
Dificultad del curso	0.594	0.022	Si
Interés de salida	0.170	0.011	Si

Conocidas las variables significativas, se procede nuevamente con la estimación para observar el comportamiento conjunto de las mismas. Las estimaciones resultantes se observan a continuación:

Tabla 3: Estimaciones del nivel Uno

NIVEL 1 estudiante	β_{00k}	sd	Significativa>1.96
Edad	0.001	0.003	No
Interés previo	0.298	0.031	Si
Repetiría curso	2.365	0.077	Si
Conocimientos previos	0.198	0.028	Si
Grado de dificultad evaluación	0.190	0.032	Si
Dificultad del curso	0.292	0.020	Si
Interés de salida	0.065	0.009	Si



Cuando incluimos en el modelo, las variables significativas vemos que la media general se reduce 5.401 (0.116) y la varianza de nivel tres se reduce 0.006 (0.003). Respecto a la media, se puede considerar que al existir otras variables explicativas, estas comparten valores con la media general. Además, la reducción de varianza en nivel tres, significa que las medias entre profesores se diferencian más en sí y que las medias entre facultades o escuelas.

En este sentido, considerando la codificación realizada, la constante o intercepto indica la media general de evaluación docente, cuyos estudiantes no tienen interés previo por la clase, nunca repetiría otra clase con ese docente, asisten a la clase sin conocimientos previos, considera nula la dificultad de la evaluación y de la clase, y al finalizar la misma no le generó ningún interés.

Ahora bien, observando la tabla 3 podemos señalar que las variables del estudiantes relacionadas con: el interés previo por la clase, los conocimientos anteriores sobre la materia de clase, el grado de dificultad de la evaluación y la dificultad del curso, son elemento de valoración positiva del profesor, por parte de los estudiantes. En este contexto, hemos de resaltar la variable relacionada con la habilidad del docente por mantener o incrementar el interés del estudiante por la clase, la cual se asocia positivamente en 2,36 puntos de diez posibles; siendo la variable con un puntaje mayor, de todas las mencionadas. Sin embargo, el valor de la covarianza es de -2.55 (0.19) señalando que esas diferencias son mayormente acentuadas, en los estudiantes con bajas calificaciones.

2.3.4. Variables de nivel dos, los docentes

Interesa conocer ahora las variables de la docencia, cuyas estimaciones se encuentran en la siguiente tabla

Tabla 4: Estimaciones del nivel Dos

NIVEL 2 docente	β_{00k}	<i>sd</i>	Significativa>1.96
Edad	0.02	0.02	No
Sexo	0.169	0.042	Si
Área Académica	0.025	0.061	No
Horas de nombramiento	0.012	0.05	Si
Proyectos de investigación	-0.007	0.003	Si
Proyectos de Acción social	0.001	0.003	No
Cargos administrativos	0.003	0.006	No
Otros cargos	0.007	0.005	No
Grado obtenido	0.005	0.036	No
Categoría salarial	-0.008	0.016	No



Puntaje	0.001	0.001	No
---------	-------	-------	----

Una vez estimadas las variables docentes, se procede a trabajar solamente con aquellas que fueron significativas.

Tabla 5: Estimaciones del nivel Dos

NIVEL 2 docente	β_{00k}	sd	Significativa>1.96
Sexo	0.27	0.169	No
Horas de nombramiento	0.013	0.05	Si
Proyectos de investigación	-0.007	0.003	Si

Considerando las estimaciones de la tabla cinco, el valor de sexo del profesor, no está asociado a la evaluación docente cuando se controlan dos variables: tiempo dedicado a la investigación y cantidad de horas nombrado en la Universidad. Los datos expresan, por cada hora dedicada a la investigación, los estudiantes valoran negativamente en un 0.007, o sea, un investigador dedicado de 40 horas, vería reducida su evaluación en 0.28, de diez posibles. Por otra parte, un valor positivo se obtiene cuando la Universidad contrata más tiempo al docente, pues se incrementa en 0.013 por hora laborada, es decir un docente con mayor cantidad de tiempo contratado será mejor valorado.

2.3.5. Variables de nivel tres, las Facultades o Escuelas

El último paso será observar las variables de las unidades académicas contenidas en la tabla 6:

Tabla 6: Estimaciones del nivel Dos

NIVEL 3 Facultades o Escuelas	β_{00k}	sd	Significativa>1.96
Ciclo	-0.120	0.042	Si
Área Académica	-0.211	0.154	No
Horario	0.008	0.007	No

Respecto al nivel tres de Facultad o Escuela, podemos señalar los valores no significativos el horario de trabajo y el área académica. Sin embargo, del ciclo o semestre de pertenencia es significativo, señalando una reducción en la valoración de los estudiantes en 0.120, a los docentes que laboran en el segundo semestre o ciclo.

2.3.6. Modelo final



Conociendo los valores significativos de cada uno de los niveles se procede con una última estimación, integrando un modelo con todas las variables. La estimación se observa en la tabla 7.

Tabla 7: Modelo Final

Modelo Final	β_{00k}	sd	Significativa>1.96
Ciclo	-0.54	0.32	No
Horas de nombramiento	0.004	0.004	No
Proyectos de investigación	-0.003	0.002	No
Interés previo	0.320	0.035	Si
Repetiría curso	2.020	0.046	Si
Conocimientos previos	0.189	0.032	Si
Grado de dificultad evaluación	0.221	0.034	Si
Dificultad del curso	0.258	0.023	Si
Interés de salida	0.734	0.039	Si

Observando los valores obtenidos en la estimación conjunta, de las variables significativas de cada nivel, continúan manteniendo su valor las de nivel uno o sea el estudiante, dado que el ciclo, las horas de nombramiento y los proyectos de investigación, no son significativos.

III. Resultados Y DISCUSIÓN

Evaluar la Universidad y sus Facultades o Escuelas, no es una señal de cambio o mejora. La evaluación no transforma por sí misma los avances cualitativos internos, referentes a la ética, la moral y el trabajo, estos deben ir acompañados con las modificaciones normativas y legales necesarias.

En este sentido, la evaluación docente ha de estar integrada en un proceso de evaluación institucional, donde se mida científicamente la estructura universitaria en



su complejidad y donde la docencia será una parte de la compleja estructura; nunca un ente aislado y único responsable de las debilidades o desaciertos del sistema.

Por otra parte, en la evaluación de la docencia se ha realizado una búsqueda profunda por definir exactamente las competencias que ha de tener un “buen profesor”, pero, pese a la existencia de una voluminosa literatura, los esfuerzos de Tuning Europa y Tuning América Latina, OCDE (2005) y la UNESCO (2005), es notorio todavía la falta de acuerdo y estas varían de un país a otro, de una Universidad a otra, aún dentro de un mismo país. Todavía no se ha logrado definir cuantas y cuales se pueden utilizar, pues las condiciones contextuales profundizan las diferencias entre unas y otras.

Ahora bien, refiriéndonos a las encuestas de estudiantes podemos señalar que no son el mejor procedimiento para el control del cumplimiento del profesorado, ni tampoco lo son para la toma de decisiones de tipo fiscalizador. Sin duda, pueden aportar información en tal sentido, pero no podemos olvidar la perspectiva del alumno es fundamentalmente la visión del cliente, que reclama calidad del servicio señalando lo que le satisface o no y en ninguna circunstancia es la perspectiva del administrador.

Este tipo de evaluación tiene muchos adversarios, contradicciones y no pocas objeciones en cuanto a su validez y fiabilidad. Pero las evaluaciones de los alumnos son válidas en cuanto reflejan su satisfacción con la enseñanza que reciben y, en este sentido no puede criticarse su validez. Lo que se trata de demostrar es que no sólo reflejan la satisfacción de los alumnos, sino que están relacionadas con la eficacia de la instrucción. Además, guste o no guste, las universidades tendrán que atender la voz de sus estudiantes, por lo que resulta crítico analizarlas con atención y profundidad, evitando fenómenos de evaluación espontánea descontrolada, que son bastante inútiles y, además, irritantes para sus afectados.

Aún más, pese a la importancia de los estudiante como evaluadores de la competencia docente universitaria, las investigación recomienda tener en cuenta la existencia de otros elementos intervinientes en el aprendizaje del alumno y que no dependen solamente del profesor, como el interés en la materia, la motivación personal, el tiempo dedicado al estudio, la inteligencia general, los recursos y otras tantas variables que influyen en el aprendizaje.

Mediante esta investigación se pudo comprobar en la Universidad de Costa Rica, mediante un modelo jerárquico lineal que, pese a las características de las Facultades o de los docentes, son significativas aquellas relacionadas con los estudiantes como: el interés y los conocimientos previos por la clase, los cuales valoran positivamente al docente.



Este importante proceso integrada como un elemento de lo que denominaríamos evaluación institucional, y no aislarlo como si se tratara de algo al margen, aunque paralelo, del resto de acciones evaluadoras. Aún más, los datos deben ser analizados profundamente por medio de los procesos metodológicos emergentes, para obtener resultados de análisis con el mínimo de error y para ello los modelos jerárquicos lineales son una potente herramienta de trabajo, que permite observar la actividad docente en diferentes niveles de agregación, atendiendo a las características específicas de cada uno de los involucrados. Este componente únicamente se puede atender desde esta metodología, de ahí la importancia por trabajar con ella.

Mediante los modelos jerárquicos se pudo observar que algunas variables contextuales, pueden ejercer influencia en la opinión de los estudiantes respecto de su profesor.

Hoy día nadie puede negar la importancia que tiene todo este proceso de recogida de la opinión valorativa de los alumnos, aunque pensamos que debe ser

En consecuencia entendemos, que la evaluación de los estudiantes tendrá sentido, en la medida en que la propia institución universitaria en su conjunto, o los departamentos en su especificidad, sean capaces de utilizar provechosamente la información que se desprende de estas evaluaciones.

Consecuentemente, la evaluación hecha por los alumnos es importante porque:

- a) Él es el único que conoce su ambiente de aprendizaje, incluyendo los aspectos como, la habilidad del docente para motivar hacia el continuo aprendizaje y la calidad de compenetración entre el estudiante y su profesor.
- b) El estudiante es el más apto evaluador de la calidad, la eficacia y la satisfacción en relación con los contenidos del curso, métodos de instrucción, libros de textos, tareas y su propio interés.
- c) Esta evaluación puede ser usada por otros estudiantes para seleccionar cursos y profesores. Ello puede, según Aleamori (1981), incrementar la excelencia del profesor, la cual podría ser reconocida o premiada.
- d) Aleamoni (1981, 1998) los estudiantes pueden



diferenciar claramente entre la enseñanza eficaz y otras dimensiones afectivas como aptitud, interés y amistad del profesor.

Nosotros consideramos, a pesar de las muchas críticas, la investigación sobre estas evaluaciones indica que son razonablemente válidas y no se ven contaminadas en un grado apreciable por factores distintos a la misma calidad docente. Muchas veces, el peso de algunas de las variables potenciales de sesgo es mínimo, a pesar de su significación estadística. Además el estudiante es una parte de la evaluación, sus juicios no deben ser los únicos para valorar al profesor, debido a la importancia de diversos instrumentos y evaluadores para acercarnos de manera adecuada a la estimación de un proceso tan complicado. La opinión de ellos, no debe ser la única vía de información para tomar decisiones laborales o que puedan afectar a la estabilidad de este.

El problema, como sugiere De la Orden (1987) es que no existen procedimientos totalmente satisfactorios para evaluar a los profesores. En este sentido una amplia mayoría de expertos en evaluación institucional defienden la diversificación y complementación de procedimientos y fuentes en la evaluación, algo que está incidiendo de manera decisiva en los modelos de evaluación que emplean las universidades. Sin embargo y hasta el momento, las encuestas de valoración de los estudiantes son el procedimiento más empleado y en demasiadas ocasiones, como afirma Romainville (1999), es el único que se emplea de manera sistemática.

De Miguel (1991) y Casanova (1992, 1997) sostienen que el gran reto de la evaluación no está tanto en seguir profundizando en el debate sobre bondades y maldades del procedimiento, sino en centrar adecuadamente su utilización y de manera especial en los temas relacionados con el estatus sociolaboral de los profesores.

Las encuestas de estudiantes no son el mejor procedimiento de control del cumplimiento del profesorado, ni tampoco lo son para la toma de decisiones de tipo fiscalizador. Sin duda, pueden aportar información en tal sentido, pero no podemos olvidar como bien afirman Biggs (2001) y Kwan (1999), la perspectiva del alumno es



fundamentalmente la visión del cliente, que reclama calidad del servicio señalando lo que le satisface o no y en ninguna circunstancia es la perspectiva del administrador.

V. BIBLIOGRAFÍA CITADA

- BARR, A.S. (1931). *An Introduction to the Specific Study of Classroom Supervision*. New York: D. Appleton.
- BIGGS, J. (2001): "The reflective institution: Assuring and enhancing the quality of teaching and learning". En *Higher Education*, 41 (3), pp. 221-238.
- BOYCE, A.C. (1915). *Method for Measuring Teacher's Efficiency*. Bloomington, ILL.: Public School Publishing Co.
- GREENWALD, R., HEDGES, L.V., AND LAINE, R.D. (1996). "The effect of school resources on student achievement". En *Review of Educational Research*, 66, 361-396.
- KRATZ, H.F. (1896). "Characteristic of the Best Teachers as Recognized by Children". En *Pedagogical Seminary*, 3, pp. 413-418.
- KWAN, K.P. (1999): "How Fair Are Rating in Assessings The Teaching Performance Of University Teacher?". En *Assessment & Evaluation In Higher Education*, 24 (June), 181-195.
- LAPOTIN, A.S. & HAESSIG, C.J. (1998-99): "Fostering Leadership in The Institutional Assessment Process" . En *Journal Staff, Program & Organization Development*, 16 (4), PP. 213-220
- McKEACHIE, W.J. (1990). "Research on College Teaching: The Historical Background". En *Journal of Educational Psychology*, 82 (2), pp. 189- 200.
- RUGGS, H.O. (1920). "Self – improvement of teachers through self – rating: A new scale for rating teacher's efficiency". En *Elementary School Journal*, 20, pp. 670-684.
- ZABALZA, M. A. (2005, febrero). *Competencias docentes*. Documento presentado en la Pontificia Universidad Javeriana de Cali, Colombia. Consultado el 12 de julio de 2014, en: <http://portales.puj.edu.co/didactica/Archivos/Competencias%20docentes.pdf>
- ORGANIZACIÓN PARA LA COOPERACIÓN Y EL DESARROLLO ECONÓMICO (2005). *La definición y selección de competencias clave: Resumen ejecutivo*. Consultado el 01 de agosto de 2012, en: <http://www.deseco.admin.ch/bfs/deseco/en/index/03/02.parsys.78532.downloadList.94248.DownloadFile.tmp/2005.dscexecutivesummary.sp.pdf>
- BENEITONE P., ESQUETINI C., GONZÁLEZ J., MALETÁ, M. M., SIUFI, G., WAGENAAR, R. (Eds.). (2007). *Tuning América Latina. Reflexiones y perspectivas de la educación superior en América Latina*. Informe Final. Deusto, España: Universidad de Deusto Publicaciones.
- UNESCO, (2005). *Informe de seguimiento de la educación para todos, 2006. La alfabetización, un sector vital*. París: UNESCO. Consultado el 01 de junio de 2014, en: <http://unesdoc.unesco.org/images/0014/001470/147000s.pdf>



MLwiN - C:\Users\Dirección\Documents\PONENCIA CHILE\Modelo NULO-sabado.wsz - [Equations]

File Edit Options Model Estimation Data Manipulation Basic Statistics Graphs Window Help

Start More Stop IGLS Estimation control..

Criterio_{ijk} ~ N(XB, Ω)

Criterio_{ijk} = β_{0ijk}constante + β_{1j}repet-curso_{ijk} + 0.198(0.028)cono-previo_{ijk} + 0.190(0.032)Dif-evalua_{ijk} + 0.292(0.020)dif-curso_{ijk} + 0.065(0.009)interes-salida_{ijk} + -0.000(0.003)in_edad_{ijk} + 0.298(0.031)interes-ini_{ijk}

β_{0ijk} = 5.401(0.116) + v_{0ik} + u_{0jk} + e_{0ijk}

β_{1j} = 2.365(0.077) + u_{1jk}

$\begin{bmatrix} v_{0ik} \end{bmatrix} \sim N(0, \Omega_v) : \Omega_v = \begin{bmatrix} 0.006(0.003) \end{bmatrix}$

$\begin{bmatrix} u_{0jk} \\ u_{1jk} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 2.685(0.197) \\ -2.522(0.192) \ 2.366(0.193) \end{bmatrix}$

$\begin{bmatrix} e_{0ijk} \end{bmatrix} \sim N(0, \Omega_e) : \Omega_e = \begin{bmatrix} 0.732(0.015) \end{bmatrix}$

-2*loglikelihood(IGLS Deviance) = 16107.855(5968 of 12319 cases in use)

Name + - Add Term Estimates Nonlinear Clear Notation Responses Store Help Zoom 100

random fixed iteration 81 Names Equations

ES 09:15 a.m. 24/08/2014

MLwiN - C:\Users\Dirección\Documents\PONENCIA CHILE\BASE-CEA-2012.sav.wsz - [Equations]

File Edit Options Model Estimation Data Manipulation Basic Statistics Graphs Window Help

Start More Stop IGLS Estimation control..

puntaje_{ijk} = β_{0ik}constante + β_{1i}in_sexo_{ijk} + β_{3k}Catedra-prof_{ijk} + 9.756(0.335)grado_{ijk} + e_{2ijk}chicos_1_{ijk}

β_{0ik} = 36.498(2.142) + v_{0k} + u_{0ik}

β_{1i} = -0.059(0.027) + e_{1iik}

β_{3k} = 25.026(5.759) + v_{3k}

$\begin{bmatrix} v_{0k} \\ v_{3k} \end{bmatrix} \sim N(0, \Omega_v) : \Omega_v = \begin{bmatrix} 60.277(22.580) \\ -154.384(61.245) \ 457.548(188.923) \end{bmatrix}$

$\begin{bmatrix} u_{0ik} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 307.955(7.938) \end{bmatrix}$

$\begin{bmatrix} e_{1iik} \\ e_{2iik} \end{bmatrix} \sim N(0, \Omega_e) : \Omega_e = \begin{bmatrix} 0.914(0.020) \\ 0 \ 1.771(0.043) \end{bmatrix}$

-2*loglikelihood(IGLS Deviance) = 53452.683(11217 of 12319 cases in use)

Name + - Add Term Estimates Nonlinear Clear Notation Responses Store Help Zoom 75

random fixed iteration 10 Names Hierarchy viewer Equations

ES 05:54 p.m. 09/08/2014