UNIVERSIDAD DE EL SALVADOR
FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA
ESCUELA DE MATEMÁTICA

"TEORÍA DE MODELOS MULTINIVEL Y SUS APLICACIONES".

PRESENTADO POR:

WELMAN DEL CARMEN ROSA ALVARADO

PARA OPTAR AL GRADO DE:

LICENCIADO EN ESTADÍSTICA

ASESOR:

DR. JOSÉ NERYS FUNES TORRES

ASESOR ADJUNTO:

LIC. RENÉ ARMANDO PEÑA

CIUDAD UNIVERSITARIA, SEPTIEMBRE DE 2005

UNIVERSIDAD DE EL SALVADOR

RECTORA : Dra. María Isabel Rodríguez

SECRETARIA GENERAL : Licda. Alicia Margarita Rivas de Recinos.

FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA

DECANO EN FUNCIONES : M.Sc. José Héctor Elías Díaz

SECRETARIO : Lic. Víctor Manuel Durán Belloso

ESCUELA DE MATEMÁTICA

DIRECTOR : Lic. Mauricio Hernán Lovo Córdova

TRABAJO DE GRADUACIÓN APROBADO POR:

COORDINADOR : Lic. Mauricio Hernán Lovo Córdova

ASESOR : Dr. José Nerys Funes Torres.

ASESOR ADJUNTO : Lic. René Armando Peña

DEDICATORIA

Le agradezco y dedico a Dios, a la Universidad de El Salvador, a mis padres Maria Adela Alvarado y Ángel Fabricio Rosa, por darme la oportunidad de finalizar este trabajo de graduación. De igual manera deseo expresar mi agradecimiento a mis asesores Dr. José Nerys Funes Torres y el Lic. René Armando Peña, quienes por sus enseñanzas contribuyeron en gran parte para el desarrollo del presente trabajo de graduación.

Welman del Carmen Rosa Alvarado

ÍNDICE.

Contenido **Pág.**

INTRODUCCIÓN

Actualmente es frecuente encontrar en diversas investigaciones de las áreas de la educación y de la salud el uso de la Teoría de Modelos Multinivel como alternativa metodológica de acercar el contexto del individuo a la explicación de la causalidad. Lo que se pretende con esta metodología es representar de una manera precisa aquellos fenómenos en los que la recogida de datos presenta una estructura anidada. En ese sentido el modelo multinivel tiene en cuenta el agrupamiento de los individuos en otras unidades, es decir, donde no solo se han seleccionado una serie de sujetos, sino también una serie de unidades contextuales a los que éstos pertenecen, tales como: hospitales, clases, escuelas, municipios, empresas u otras instituciones; situaciones en las que se tienen estructuras particulares de los datos, las cuales no pueden ser considerados en un análisis de regresión clásico; en caso que se utilice este tipo de regresión se llegaría a las siguientes consecuencias: la producción de sesgos en los errores típicos de los estimadores y el aumento de la probabilidad de rechazar la hipótesis nula de no asociación, cuando ésta es cierta.

Conceptualmente, los modelos multinivel son básicamente un modelo de regresión de efectos mixtos, en donde se estudia una relación lineal entre dos o más variables en estudios realizados mediante un muestreo por agrupamiento, es decir, una técnica correlacional adecuada para analizar variaciones en las características de los sujetos que son miembros de un grupo que a su vez, hace parte de otra agrupación, o sea, mediciones que forman una estructura agrupada y jerárquica. El modelo permite la descomposición de la variación de una variable criterio (como por ejemplo, rendimiento) en sus componentes “dentro del grupo” (dentro-escuela, dentro-departamentos) y “entre grupo” (entre-escuela, entre-departamento) y el análisis de la asociación entre variables en esos niveles de agregación.

Todos estos aspectos serán abordados en el presente trabajo de graduación que está estructurado en cuatro capítulos: introducción a la teoría multinivel, teoría de modelos multinivel, aplicación y conclusiones. Como énfasis de una comprensión conceptual y práctica de cómo utilizar esta técnica estadística cuando se tiene una situación anidada en los datos.

En el primer capítulo se presentan los antecedentes de los modelos multinivel, el modelo de ecuación estructural y se define el modelo multinivel como un modelo de regresión de efectos mixtos, en donde se estudia una relación lineal entre dos o más variables en estudios realizados mediante un muestreo por agrupamiento.

En el segundo capítulo se estudian los supuestos tanto de un modelo de regresión simple como un modelo multinivel de dos y de tres niveles, además se presenta el método de estimación de los parámetros para el modelo de componentes de varianza y la estimación de los parámetros fijos y aleatorios del modelo multinivel de dos y tres niveles a partir de los enfoques ANOVA (Análisis de varianza) y ANCOVA (Análisis de covarianzas).

En el tercer capítulo se desarrolla una aplicación de los modelos multinivel, disponiendo de la base de datos de un estudio realizado en diversos países de Latinoamérica, el cual consistió en aplicar un instrumento de recolección de información a: alumnos, profesores, padres de familia y escuela. El objetivo de esta aplicación es identificar las condiciones en las que los alumnos de tercero y cuarto grados de educación básica, alcanzan los aprendizajes en Lenguaje y Matemática, según el resultado registrado en pruebas estandarizadas. Esto, para buscar un modelo óptimo de Lenguaje y Matemática, de tal manera que se logre encontrar cuales son los factores asociados al rendimiento de los estudiantes en la prueba de Lenguaje y Matemática.

En el cuarto capítulo se concluye a partir de la aplicación realizada con la técnica multinivel tanto para Lenguaje y Matemática. Se presenta además, un ejemplo de predicción de la puntuación que obtendría un estudiante en Lenguaje y Matemática basado en el contexto en el que vive y se desenvuelve académicamente.

Finalmente dentro de este trabajo se presentan los anexos de algunas corridas realizadas en el software HLM (Modelos Jerárquicos Lineales) y las referencias bibliográficas que se utilizó para el desarrollo de este trabajo de graduación.

CAPÍTULO I

INTRODUCCIÓN A LA TEORÍA DE MODELOS MULTINIVEL.

Introducción.

A lo largo de la historia los problemas del mundo real han sido abordados desde diferentes perspectivas. Muchas disciplinas han contribuido a su conceptualización, de hecho, sus características diferenciadoras en los distintos momentos históricos se han ido perfilando desde la filosofía hasta la política, pasando por las aportaciones de la psicología, de la economía y de la teoría de la organización. Pero no ha sido hasta el siglo XX cuando se ha realizado un intento sistemático en construir un cuerpo de conocimiento sobre los problemas y sobre los fenómenos que en el se dan.

Este capítulo esta estructurado de la siguiente manera:

- Antecedentes de los modelos multinivel: aquí se presentan las metodologías que fueron fundamentales para el nacimiento de esta teoría, tales como: el modelo de Cronbach, causal y de regresión múltiple.

- El modelo de ecuación estructural: este es un modelo que permite separar las relaciones para cada conjunto de variables dependientes. Describiendo además los problemas que este modelo permite atacar, como: validez de las relaciones causales, interrelaciones y medida.

- La naturaleza de los modelos multinivel: se define el concepto de lo que es en si la teoría multinivel, el cual hace referencia a un modelo de efectos mixtos, donde se estudia una relación lineal entre dos o más variables en estudios realizados mediante un muestro por agrupamiento, es decir, una técnica correlacional adecuada para analizar variaciones en las características de los sujetos que son miembros de un grupo que a su vez, hace parte de otra agrupación, o sea, mediciones que forma de una estructura agrupada y jerárquica. Mencionando además a diversos autores que han aplicado el análisis multinivel en áreas de la educación, salud, epidemiología, etc. Todo con el fin de buscar la conexión entre las características del individuo y el contexto social en que se desenvuelve.

- Conjeturaciones previas al análisis multinivel: en este apartado se describe a groso modo que la teoría multinivel es una metodología fundamentalmente estadística para el análisis de datos que presentan una estructura jerárquica, y principalmente cuando los datos se han obtenido a través de la técnica de muestro por conglomerado.

- Enfoques multinivel basado por el análisis de regresión múltiple, ANOVA y ANCOVA: se estudian los conceptos sobre el análisis de varianza y de covarianza, que son fundamentales para estimar los componentes de varianza, para luego estimar los parámetros tanto fijo y aleatorios de un modelo multinivel de dos y tres niveles.

- El error de Medida: se menciona lo relacionado al error de medida, además de describir algunas metodologías claves para tratar el error de medida en diversos estudios de investigación.

1.1 ANTECEDENTES DE LOS MODELOS MULTINIVEL.

En la actualidad los investigadores se ven obligados a recoger un gran número de medidas para poder captar de forma adecuada la complejidad de los fenómenos del mundo real, e indagar o buscar un modelo estadístico. De tal manera, que nos permita tomar una connotación específica sobre su estructura, las relaciones entre sus componentes, su funcionamiento y los cambios que experimenta el sistema en su totalidad o en sus componentes. Los conocimientos generales por una investigación en particular, se unen a otros conocimientos ya existentes, acumulados durante mucho tiempo por otros investigadores, sea en la forma de aporte original o como confirmatorio de hallazgos ya existentes. Cualquiera que sea la situación que se enfrente, la investigación es siempre la búsqueda de la solución a algún problema de conocimiento.

La construcción de modelos estadísticos o modelos matemáticos aplicados se incluye dentro de lo que ha venido a denominarse modelación estadística ("statistical modelling"). Como señala Lindsey (1993), con el modelo estadístico se pretende descubrir la variabilidad en los datos observados mediante procedimientos matemáticos. Además, se formula en términos del o los constructos hipotéticos que no se pueden observar o medir directamente. Ejemplo de tales contructos son: la confianza, autoestima, discriminación, rendimiento, motivación, capacidad, etc. Sin embargo, para definir los constructos hipotéticos se debe primero incluir la clasificación de esos constructos como dependientes (causado, criterio, endógeno) o independiente (causal, explicativo, exógeno). En segundo lugar, para cada constructo dependiente, la teoría especifica cuales de los otros constructos se postula para ser dependiente. La construcción puede también incluir una declaración sobre la muestra y/o el tamaño relativo del efecto directo de un constructo en otra.

Las dos relaciones claves para la construcción del modelo estadístico son: la primera es que las relaciones entre construcciones hipotéticas constituyen la pieza estructural del modelo y la segunda es que las relaciones entre indicadores o variables observables y constructos teóricos constituyen la pieza fundamental de la medida del modelo.

En todo caso la construcción de un modelo estadístico consistirá en una serie de procesos encaminados a explicar el comportamiento de una variable respuesta, relacionado con indicadores exógenos. Sin embargo, debido a que la explicación de la variable criterio en función de las variables exógenas no es perfecta, se incluye un término residual, es decir, el término del error que puede estar formado por los efectos de otros indicadores sobre las variables endógenas, por errores de medición, etc. Además, la variable o nivel independiente del tratamiento se supone que es fijado por el experimentador (valores conocidos).

Desde el punto de vista cualitativo, al explicar el comportamiento de la variable respuesta, se pueden tener tres posibles modelos diferentes, a saber:

a) Que la variabilidad observada en los datos sea totalmente explicada por el modelo, lo cual daría lugar a considerarlo como un modelo absolutamente determinístico y formal.

b) Que la variabilidad observada en los datos sea totalmente explicada por factores aleatorios, lo cual daría lugar a un modelo aleatorio, la base del cual puede encontrarse en la ejecución de errores de muestreo, de medida o debido a las características inherentes del sistema observado.

c) Que la variabilidad observada en los datos necesite de la participación de factores sistemáticos y aleatorios, lo que da lugar a los modelos probabilísticos o estocásticos.

La modelación estadística consiste en buscar un modelo que genere los datos con un mínimo error posible. Según Arnau, 1981; las etapas de la modelación estadística en esencia vienen a englobar los mismos contenidos, pudiéndose resumir como las siguientes: identificación, estimación, validación y metadiagnóstico.

Una vez entendido el modelo estadístico tradicional, lo que se pretende ahora es poder representar de una manera precisa aquellos fenómenos en los que la recogida de datos presenta una estructura anidada. En ese sentido el modelo multinivel tiene en cuenta el agrupamiento de los individuos en otras unidades, es decir, donde no solo se han seleccionado una serie de sujetos, sino también una serie de unidades contextuales a los que éstos pertenecen, tales como: hospitales, clases, escuelas, municipios, empresas u otras instituciones; situaciones en las que se tienen estructuras particulares de los datos, las cuales no pueden ser considerados en un análisis de regresión clásico; en caso que se utilice este tipo de regresión se llegaría a las siguientes consecuencias: la producción de sesgos en los errores típicos de los estimadores y el aumento de la probabilidad de rechazar la hipótesis nula de no asociación, cuando ésta es cierta.

A continuación se presentan ciertas metodologías que fueron fundamentales para el inicio del estudio del análisis multinivel.

i) Modelo de Cronbach (Cronbach y Webb, 1975):

Este modelo propone obviar el problema de multicolinealidad en los datos, logrando dar una mejor interpretación en los parámetros de dicho modelo. Es un modelo de regresión donde las variables individuales están centradas en la media de su grupo y variables contextuales centradas en la media global.

En términos matemáticos este modelo se resume de la siguiente forma:

$$y_{ij} = \alpha + \gamma x_{.i} + \beta_1 (x_{ij} - x_{.j}) + \beta_2 (x_{.j} - x_{..}) + \varepsilon_{ij}$$

Donde x_{ij} es la variable de i-ésimo individuo en el j-ésimo grupo, $x_{.j}$ es la media del grupo j-ésimo y $x_{..}$ es la media global.

ii) Análisis Causal.

La filosofía define el término causalidad como una ley en virtud de la cual se producen efectos. El principio de causalidad establece la conexión entre causa y efecto, y es uno de los fundamentos del conocimiento humano. Por ello la causalidad es considerada como una ley natural de absoluta validez.

La esencia del análisis causal reside en la unificación de los comportamientos de los componentes de un sistema para deducir el comportamiento del sistema global. Estos comportamientos deben ser comprendidos en el contexto de sistema global, atendiendo a las conexiones entre los componentes que constituyen el sistema. Por otra parte el comportamiento de cada parte del sistema permite la realización de una función, y subfunciones. Tal identificación de los subsistemas con funciones conocidas ayuda al análisis causal.

La utilización del modelo causal tiene cierta tradición en el área de la prevención de riesgos laborales de trabajos originales de Meliá (1998). En la construcción de este modelo, se parte de la distinción de los modelos causales de los accidentes en dos grandes categorías: los modelos secuenciales (que tratan la cadena de eventos que conducen al accidente) y los modelos estructurales (que desarrollan la interacción persona-máquina). El modelo causal (ver figura 1.1) psicosocial responde a un planteamiento integrador de los factores organizacionales y de naturaleza psicosocial. Las variables que incluye son las siguientes:

- Clima de seguridad: variable que el ambiente de seguridad existente en la empresa e impulsado por la dirección de la organización.

- Riesgo basal: concepto asociado a la actividad de la empresa, reflejando los riesgos inherentes y específicos a una determinada industria o tipo de tarea.

- Respuesta de los superiores: variable que recoge la conducta y actitud de los supervisores en materia de seguridad, así como los mecanismos de comunicación y de contingencia, en el contexto de la prevención.

- Respuesta de los compañeros: factor psicosocial que trata de medir la respuesta frente a problemas de seguridad por parte de los compañeros, en base al tipo de respuesta y su frecuencia.

- Conducta hacia la seguridad del trabajador: descrita mediante comportamientos específicos tales como orden, cumplimiento de las normas, utilización de maquinaria defectuosa, etc.

- Riesgo basal: variable que refleja el riesgo de una determinada tarea en su contexto, con independencia de las acciones en materia de prevención del riesgo que adopte el trabajador que la realice.

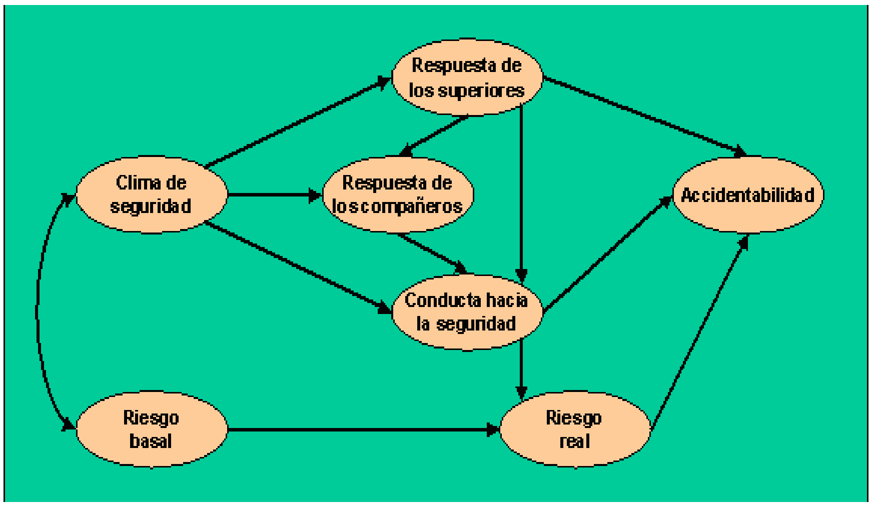
- Riesgo real: mide la probabilidad de ser víctima de un accidente de trabajo. Obsérvese que esta variable recoge los efectos tanto de las conductas como de las condiciones de trabajo (riesgo basal).

- Accidentalidad: variable criterio indicador de la existencia de un accidente. Medida mediante el número de accidentes laborales sufridos por el trabajador en los últimos cinco años.

Un análisis causal psicosocial de los accidentes laborales (Meliá, 1998) se aplica a una muestra de 316 trabajadores, con contrato mercantil o fijos, representando una variedad de sectores industriales, propiedad (públicas y privadas), características organizativas y ámbito de actuación (local, regional, nacional y multinacional). El modelo se ajusta correctamente, ofreciendo parámetros con el signo acorde con las hipótesis realizadas, una mayoría son estadísticamente significativos (sólo dos incumplen esta premisa).

Como conclusión, el autor indica que el modelo confirma que la variable clima de seguridad (impulsada por la gestión) influye sobre la conducta hacia la seguridad, directa e indirectamente a través de la respuesta de superiores y compañeros. De igual forma, el modelo muestra relaciones directas consistentes de la conducta y el riesgo basal sobre el riesgo real que, a su vez, influye sobre la accidentalidad.

Figura 1.1: Modelo causal psicosocial de los accidentes laborales.



Pascual (1993), dio otro gran aporte al análisis causal, al desarrollar en el área de la educación, el liderazgo transformacional de Bass (1985), sosteniendo así tres premisas claves de por qué el liderazgo es un indicador fundamental en la mejora de la eficacia de los centros educativos, distinguiendo teóricamente dos grandes estilos de liderazgo, lo que se denomina liderazgo transformacional y liderazgo transaccional, así mismo habla de un tercer comportamiento directivo al que da el nombre de no liderazgo. Así, define el liderazgo transformacional como aquel formado por carisma, consideración individual, estimulación intelectual, inspiración y en el caso de la organización escolar agrega un factor denominado tolerancia psicológica.

Según el modelo causal de liderazgo planteado por Pascual, a groso modo se clasifican tres niveles. Por un lado variables relativas al nivel de colaboradores (ejemplo: la satisfacción de los colaboradores con el líder, la disposición de realizar un mayor esfuerzo por incrementar la calidad de trabajo, etc.). En el segundo nivel, variables relativas a los efectos directos en la organización como puede ser la eficacia organizativa del centro, la capacidad de cambio en el centro, etc. Por último los efectos que puede tener el liderazgo en las actitudes de los alumnos. De modo que este modelo causal considera como unidad de análisis el centro educativo, pretendiendo así, saber cuales eran las variables relevantes a nivel de organización que aglutinan una serie de otras variables más específicas y que pudieran ser efectos directos relacionados con el ejercicio de un liderazgo transformacional.

Teniendo presente, por un lado el enfoque teórico de Bass, y por otro aquellas variables más relevantes que se asocian con el liderazgo en los centros educativos como variables mediadoras la participaron y la satisfacción con el trabajo docente en el centro, variables de control a nivel de centros (tamaño del centro, tipo del centro). Lo que se trata es establecer un modelo causal que sea una explicación plausible de los efectos del liderazgo sobre la eficacia escolar.

A continuación se presenta la especificación inicial del modelo relacional de liderazgo y eficacia escolar que pretendemos validar.

Tabla 1.1: Constructos y dimensiones del modelo de liderazgo transformacional de Bass.

Carisma personalizante	**Consideración Individual**	**Estimulación Intelectual**	**Inspiración**	**Tolerancia Psicológica**
Entusiasmo	Trato personal	Animación al cambio	Implicación	Humor
Credibilidad	Apoyo	Potenciación de un mayor esfuerzo	Identidad	

El principio de validez del modelo causal es a través del sistema de ecuaciones estructurales. Donde las variables que conforman un constructo complejo, como es el liderazgo, y que en nuestra propuesta se estructura como una variable independiente, se denomina variable latente exógena. Mientras que aquellas variables que aparezcan asociadas, tanto directa como indirectamente a los efectos de un liderazgo transformacional, y cuyas varianzas intentamos explicar, recibe el nombre de variables latentes endógenas.

Sin embargo, los estudios realizados mediante un análisis causal no son sencillos, el rigor científico requiere, entre otros, la realización de *"experimentos"*. Estos experimentos precisan la formación de dos o varios grupos (de tratamiento y de control) por asignación aleatoria.

Por tanto, para la realización de un modelo causal donde los diseños sean cuasiexperimentales y no experimentales [1], se necesita el uso de modelos causales no-experimentales (modelos de estructuras de covarianzas o de ecuaciones estructurales) ya que en su planteamiento tradicional, adolecen de ciertos problemas a la hora de asegurar la validez de la relación causal, interrelación integra efectos interrelacionados y tratar con los problemas de medida. Esto se tratará a profundidad en el siguiente apartado de este capítulo.

Murillo (1990) hace sus consideraciones sobre la utilización del analisis causal, destacando dos graves inconvenientes, uno de carácter práctico, que tiene gran dificultad para que se ajusten los datos a un modelo, y otro de carácter Técnico, que no tiene en cuenta la situación jerárquica de los datos. De hecho, es una alternativa muy poco utilizada en la investigación sobre organizaciones y más en el campo psicológico.

iii) Modelo de regresión múltiple.

La regresión múltiple es el método de análisis apropiado cuando el problema del investigador incluye una única variable métrica dependiente que se supone está relacionada con una o más variables métricas independientes. El objetivo del análisis de la regresión múltiple es predecir los cambios en la variable dependiente en respuesta a cambios en varias de las variables independientes. Este objetivo se consigue muy a menudo a través de la regla estadística de los mínimos cuadrados.

[1] Meliá JL. Un modelo causal psicosocial de los accidentes laborales. *Anuario de Psicología* 1998; 29(3): 25-43.

Aunque el uso de esta técnica ha sido común desde los inicios de la investigación sobre eficacia, con el tiempo ha experimentado algunos cambios en su aplicación. Como es, el utilizar estas técnicas a partir de distintos predictores, introduciendo la aplicación al cálculo de los residuales de regresión, tomando como base características individuales de los sujetos o medias ponderadas de las escuelas.

Por ejemplo, en estudios del área de la educación, el uso de las puntuaciones residuales se realiza como medida de la eficacia de la escuela a través de la diferencia entre las puntuaciones predicha a partir de las características individuales del sujeto (nivel socioeconómico, nivel cultural de los padres, rendimiento previo, etc.) y la puntuación obtenida realmente por el mismo. Tiene la ventaja de que evita los sesgos que se producen en la estimación de los efectos de los tratamientos cuado los grupos no son equivalentes, aunque es evidente que no se realiza un ajuste total para todas las diferencias entre las escuelas. Para Castejon (1994), la técnica de residuales más adecuado para la identificación de escuelas eficaces, estableciendo cuatro formas de llevar a cabo los análisis de datos y de regresión (modelo dentro de la escuela); por otra, el análisis puede ser "ponderado" o "no ponderado" según el número de sujetos pertenecientes a cada escuela. En el análisis empírico que realiza para comparar cada uno de los cuatro índices de eficacia, cada uno de ellos a partir de: a) puntuaciones residuales a partir de la ecuación de regresión múltiple con datos de los centros; b) puntuaciones residuales a partir de la ecuación de regresión múltiple con datos de los centros obtenidos a partir de la media ponderada de los sujetos; c) medias de las puntuaciones residuales para cada alumno a partir de la ecuación de regresión; d) igual que el anterior, con la media residual estandarizada para cada alumno a partir de la ecuación de regresión. Examinado además la consistencia entre dichos índices, concluye poniendo de manifiesto la existencia de centros cuyo rendimiento supera el predicho y, por tanto, muestran un efecto significativo, aunque moderado, de la escuela, coincidiendo con los resultados obtenidos por otros autores, y, por otra parte, el alto nivel de consistencia y concordancia entre los índices resultantes. Sin embargo, los cuatro índices comparados tienen serios problemas debido a la utilización del análisis de regresión, donde los datos presentan una estructura anidad. Llegando así, a poder concluir de forma errónea.

1.2 EL MODELO DE ECUACIÓN ESTRUTURAL.

El modelo de ecuaciones estructurales se podría definir como un modelo causal no-experimental que permite separar las relaciones para cada conjunto de variables dependientes. En su aceptación más simple, el modelo de ecuaciones estructurales proporciona la técnica de estimación más adecuada y eficiente para series de estimación de ecuaciones simultáneas mediante regresiones múltiples.

El uso de este tipo de modelos causales permite atacar los problemas de interrelación, de medida y validez de la relación causal.

- **Validez de las relaciones causales:** el interés y la peculiaridad de las relaciones causales radican en el elemento de "producción" o "fuerza" que la causa tiene sobre el efecto. En el estudio de este tipo de relaciones suelen considerarse como necesarias las tres condiciones de John Stuart Mill[2].

  - Covariación entre las presuntas causa y efecto.
  - Precedencia temporal de la causa.
  - Ausencia de explicaciones alternativas para la relación.

El cumplimiento de estas condiciones, y principalmente la tercera de las mismas, es el que lleva a la necesidad de realizar diseños experimentales, diseños en los que se aíslan las unidades a estudiar (ausencia de alternativas) y se realiza una prueba controlada (precedencia) para, posteriormente, examinar el efecto (covariación).

Sin embargo existe un enfoque alternativo al análisis experimental, que consiste en formular la teoría incluyendo todas las variables que son importantes a juicio del investigador. La formulación así realizada determina la estructura de la matriz de covarianzas de todas las variables, estructura que es contrastable.

---

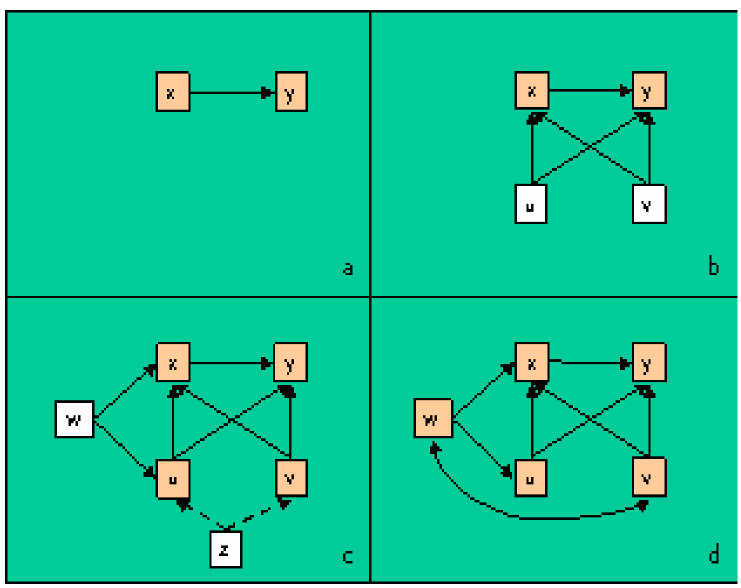
[2] Cook TD, Campbell DT. *Quasi-experimentation. Design and Analysis Issues for Field Settings*. Rand McNally College Publishing: Chicago, 1979.

De esta forma, se dispone de un instrumento que permite rechazar la teoría si los datos no soportan las conclusiones deducidas. Dicho de otra manera, se tiene un método de refutación en investigación no experimental.

- **Interrelaciones:** El problema de las interrelaciones resulta de menor importancia que los otros dos. Una vez que se ha especificado cuidadosamente la teoría, la técnica del análisis de la covarianza permite resolver el problema.

La especificación de la teoría requiere incluir todas las variables importantes, el problema es ¿dónde parar?. Un posible procedimiento consiste en ir incluyendo las variables que son causa común de cada par de variables causa-efecto, siempre que al menos una de las dos esté incluida en la hipótesis original.

Figura 1.2: Derivación de las teorías causales.



La figura 1.2 ilustra este procedimiento. La hipótesis inicial se muestra en la fig. 2.a. A partir de dicha hipótesis, se buscan otras variables que sean causa común de cada par de variables causa-efecto. En la ilustración aparecen dos nuevas variables (u, v), que aparecen como causa común de las variables originales (x,y), fig. 2.b. Como ambas variables x e y están representadas en la hipótesis original, las nuevas variables deben explicitarse en el modelo. A continuación volvemos a examinar cada par de variables causa-efecto; en este examen aparecen dos nuevas variables (w,z), fig. 2.c. La variable w es causa común de las variables x,u; como x está incluida en la hipótesis original, explicitamos w en el modelo. Sin embargo, la variable z es causa común de las variables u,v; como ninguna de estas variables aparece en la hipótesis inicial, el procedimiento indica que esta variable no es necesario explicitarla. Por último, fig. 2.d, todas las variables que no son efecto de ninguna otra (w,v), variables predeterminadas o exógenas, se unen con un arco con flechas en los dos extremos, indicando que pueden existir otros efectos, más lejanos, no explicitados en el modelo. Evidentemente, la explicación de la variable u será incompleta (no hemos incluido a z en el modelo), algo que resulta secundario para validar la hipótesis inicial (Fig. 2.a).

- **Problemas de medición:** según el modelo causal sobre la prevención de riesgos laborales[3]. Los problemas de medida en este tipo de modelos tienen dos orígenes: por un lado, la información sobre prevención de riesgos laborales puede ser interpretada de distinta manera por diferentes sujetos y, por lo tanto, estar sujeta a error; por otro lado, los conceptos organizacionales (tales como clima de seguridad, motivación, comportamiento, etc.) son difíciles de evaluar con una única variable. Una de las características de la metodología que se propone es que permite tratar fácilmente ambos problemas.

Para terminar de entender mejor la metodología, los modelos de ecuaciones estructurales utilizan una representación simbólica: los diagramas de trayectoria (path diagrams), gráficos en los que causa y efecto se unen mediante flechas dirigidas, a continuación presentamos como ejemplo un digrama path extraido de un estudio de factores asociados al rendimiento de los estudiantes que se sometieron a la PAES 2000[4].

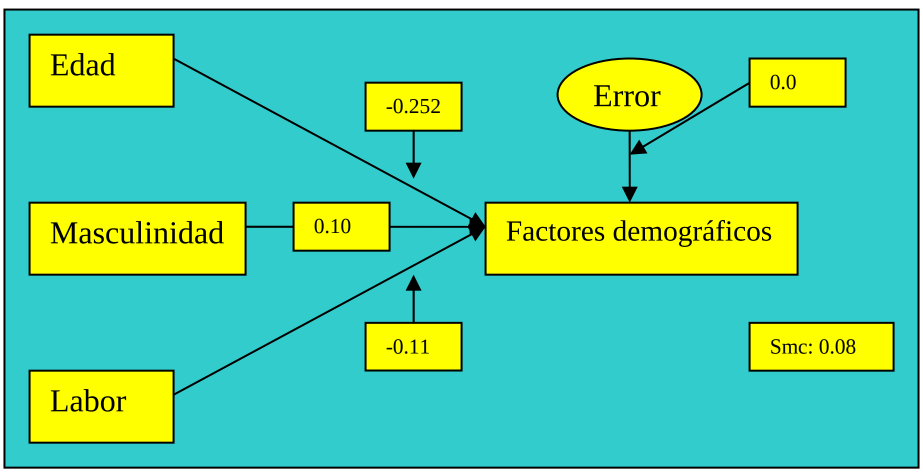
[3] Dalrymple H, Redinger C y otros. *Occupational Health and Safety Management Systems: Review and Analysis of International, National, and Regional Systems and Proposals for a New International Document*. Informe preparado por The International Occupational Hygiene Association para la OIT: Ginebra, 1998.

[4] Carlos Roberto Briones et al., Factores asociados al rendimiento de los estudiantes que se sometieron a la PAES 2000. Pág. 84-85. Ministerio de Educación, Dirección Nacional de Monitoreo y Evaluación. El Salvador.

En el diagrama de la figura 1.3, se consideran tres aspectos importantes que tienen relación con los factores demográficos: “edad”, en términos de si se está o no en el tramo “normal” para el momento de someterse a la prueba, o bien si hay atraso en el bachillerato o adelanto; “relación de masculinidad”, que da la proporción hombre/mujer en el conjunto de alumnos”, qué alumnos trabajan y estudian o solamente estudian. En dicho diagrama se observa que la edad aparece como el factor demográfico más importante (beta=-0.252): a mayor edad, menor resultado en la prueba, e indica la importancia de que no se produzcan retrasos en cursar el bachillerato y que se logre retener al alumnado.

En su conjunto, los factores demográficos considerados permiten explicar un 8% (smc=0.08) de las variaciones de los resultados globales de la PAES.

Figura 1.3: Diagrama Path de factores demográficos.



El modelo de ecuación estructural se caracteriza por dos componentes básicos: el modelo estructural y el modelo de medida. El modelo de medida permite al investigador usar varias variables (indicadores), para una única variable dependiente o independiente. Por ejemplo, la variable dependiente puede ser un concepto representado por una escala aditiva, tal como el amor propio.

En el modelo de medida el investigador puede evaluar la contribución de cada ítem de la escala así como incorporar como la escala mide el concepto (fiabilidad) en la estimación de las variables dependientes e independientes. Este procedimiento es similar al desarrollo del análisis factorial de los ítems de la escala y utiliza las cargas factoriales en la regresión.

En tal caso, el modelo estructural es el modelo “guía”, que relaciona variables independientes y dependientes. En tales situaciones, la teoría, antes que la experiencia u otras directrices, permitiera al investigador distinguir qué variables independientes predicen cada variable dependiente. Los modelos previamente discutidos que incluyen múltiples variables dependientes (análisis multivariante de la varianza y correlación canónica) no son apropiados en esta situación. Dado que permiten sólo una *única* relación entre variables dependientes e independientes.

1.3 LA NATURALEZA DE LOS MODELOS MULTINIVEL.

Actualmente es frecuente encontrar en los estudios más importantes del mundo científico y, en general, en los del área de la salud y educación. El uso de los modelos multinivel como alternativa metodológica de acercar el contexto del individuo a la explicación de la causalidad. Las causas que se asocian a la ocurrencia de un evento en salud, frecuentemente se estudian de manera estática en el sentido de que no se tiene en cuenta el análisis prospectivo o retrospectivo del ambiente social o cultural que rodea al individuo, debido a definiciones importantes, tales como ausencia de variables, imposibilidad de relacionar el contexto o el ambiente o dificultades en la aplicación metodológica, entre otras.

Esto ha implicado que el investigador llegue a sentirse potencialmente incapaz de explicar metodológicamente las causas de uno o varios eventos, puesto que frecuentemente se omiten características importantes que pertenecen a la dinámica ecológica donde interactúa el individuo, como el ambiente del trabajo, la comunidad, la familia y los aspectos culturales y sociales, entre otros.

Recientemente, las ciencias sociales y la epidemiología han intentado suplir la necesidad de una explicación más global de la morbilidad o la mortalidad, buscando la conexión entre las características del individuo y el contexto social en el que este se desenvuelve.

Resultados obtenidos mediante la técnica del análisis multinivel han sido importantes dentro de muchos campos de la investigación, por ejemplo Bryk AS (1992) dando un aporte a la teoría multinivel, relacionando en un primer momento una variable respuesta cuantitativa con las características individuales del estudiante y variables asociadas al ambiente escolar mediante un conjunto de variables exploratorias de uno y otro nivel de jerarquía. Este modelo lo definió como modelo lineal de coeficientes aleatorios.

Sandoval JJ. (2001), a dado otro aporte a epidemiología, al desarrollar un modelo multinivel con el objeto de estudiar la sociedad y las diferentes formas de organización social que afectan la salud de los colectivos humanos y su bienestar. En particular, estudia la frecuencia, la distribución de los estados de salud en las poblaciones y su relación con los determinantes sociales, de tal manera que va más allá del análisis de los factores de riesgo individuales e incluye también estudios en el contexto social de los individuos en el cual se produce el fenómeno salud-enfermedad. Como lo señala Dodge KA (2002), algunos estudios epidemiológicos han mostrado principios metodológicos de cómo estudiar las relaciones que los grupos poblacionales, el ambiente, las comunidades y la sociedad en general tienen con la agresividad en niños. Es por ello que el estudio realizado por Sandoval (2001) sobre el comportamiento agresivo y pro-social en niños, en concordancia con las características individuales y los factores sociales relacionados con las escuelas. Se estimaron las relaciones de las variables ajustado por los síntomas de hiperactividad con déficit de atención y la edad del niño. La variable endógena está conformada por los síntomas de agresividad indirecta, variable cuantitativa con nivel de medición de intervalo, constituida mediante las puntuaciones factoriales de los resultados del dominio “síntomas de agresividad indirecta”. El modelo relacionó la variable respuesta cuantitativa con las características individuales del estudiante y las asociadas al ambiente escolar mediante un conjunto de variables explorativas de uno y otro nivel de jerarquía. El modelo describió las variables exploratorias en cada nivel, controlando por los posibles efectos de confusión.

Otros autores como Maas, Cora J.M., and Snijders, Tom A.B., (2003) han dado un mayor acercamiento al análisis multinievel, especialmente al caso de medidas repetidas para datos completos e incompletos, donde, comúnmente son analizados a menudo por el análisis de variación multivariante (MANOVA). Dando así una mayor estimación mediante el modelo lineal jerárquico (HLM) mediante el uso de los modelos de efectos aleatorios.

El análisis de datos mediante estos modelos es capaz de apresar dicha estructura, como menciona Goldstein (1995) cada uno de los niveles de la estructura jerárquica es representado formalmente con su propio submodelo. Cada submodelo expresa las relaciones entre variables dentro de un determinado nivel y el conjunto de submodelos especifica de qué modo variables de un nivel influyen en las relaciones que ocurren a otro nivel distinto, quedando de este modo formalizada esta estructura anidada de los datos.

Definición de los modelos multinivel.

Los modelos multinivel son básicamente un modelo de regresión de efectos mixtos, en donde se estudia una relación lineal entre dos o más variables en estudios realizados mediante un muestro por agrupamiento, es decir, una técnica correlacional adecuada para analizar variaciones en las características de los sujetos que son miembros de un grupo que a su vez, hace parte de otra agrupación, o sea, mediciones que forman una estructura agrupada y jerárquica. El modelo permite la descomposición de la variación de una variable criterio (como por ejemplo, rendimiento) en sus componentes “dentro del grupo” (dentro-escuela, dentro-departamentos) y “entre grupo” (entre-escuela, entre-departamento) y el análisis de la asociación entre variables en esos niveles de agregación.

Los modelos multinivel permiten además estimar un conjunto separado de establecimientos de efectos de coeficientes para cada unidad organizacional, y luego modelar las variaciones en el conjunto de coeficientes de las organizaciones como productos multivariados a ser explicados por factores organizacionales (recursos de los centros educativos). Por ejemplo, a partir del nivel de los estudiantes se estiman puntajes promedios de logro de los centros, los que en un segundo nivel se correlacionan con diferentes características de éstos, es decir, los modelos multinivel ofrecen distintas alternativas que se traducen en ventajas respecto de los modelos tradicionales en una población cuya naturaleza sea anidada dentro de contexto, dando una versión más realista ya que modelan cada nivel de jerarquía y no requieren de la independencia entre los grupos.

El análisis multinivel se ha extendido y generalizado el análisis de varianza introducido por Fisher (1880-1962), surgiendo en contextos de aplicaciones diversas, ejemplo en educación un investigador educativo puede, ir a un grupo de escuelas y recoger datos sobre el nivel socioeconómico de los alumnos y su rendimiento académico, a fin de contrastar si existe una relación funcional entre ambas variables. Tendríamos así dos tipos de unidades experimentales: el sujeto y la escuela, correspondiente a lo que la literatura se denomina también como primer y segundo nivel de análisis. En el caso del sujeto, la variable de interés es el rendimiento académico. En el caso de las escuelas, las variables de interés son sus respectivas medias y pendientes de asociación entre el nivel socioeconómico y rendimiento, expresados mediante los estimadores de sus rectas de regresión.

Básicamente, como señala Burnstein et, al. (1981), el principal interés consistirá en modelizar los resultados intra-organización como funciones sistemáticas tanto de características individuales, como de variaciones entre contexto o entre grupos. Es decir, constatará si existe variabilidad en las rectas de regresión entre las distintas escuelas. De esa manera, y mediante una consideración integrada de ambos niveles de análisis, se pretende cuantificar y contrastar hipótesis estadísticas acerca de la incidencia de los distintos contextos académicos en el rendimiento escolar y en su relación con el nivel socioeconómico del alumno.

En la investigación multinivel como ya hemos mencionado se ocupa del análisis de datos donde las observaciones se jerarquizan dentro de grupos. Entonces, las variables se pueden definir en cualquier unidad de análisis de la jerarquía. Algunas de estas variables se pueden medir directamente en su nivel natural; por ejemplo, en los resultados académicos de un estudiante se identifican minimamente dos niveles: nivel 1, características del estudiante; nivel 2, características de la institución educativa.

Hay dos acercamientos a menudo criticados para analizar las características de los diversos niveles en un solo nivel. En el primer acercamiento es que todas las características de un nivel superior se desagregan a un nivel más bajo. Es decir, asignando a todos los individuos una variable que refleje la denominación de su grupo a que ellos pertenecen. En este acercamiento, todas las variables desagregadas se asumen ser independientes. El segundo acercamiento, los datos en el nivel individual se agregan al nivel más alto. La idea implícita en casi todos los usos del término “datos agregados” es que la variable agregada es simplemente un índice agregado de propiedades de unidades de nivel inferior, y no una medida directa de una propiedad del nivel superior. Pero no siempre es cierto ya que son un tipo de variables grupales que se construye “agregando” matemáticamente las características de los individuos del grupo. Se han empleados los términos “variables analíticas” y “variables agregadas” como sinómino de “variables derivadas”. También se ha recurrido al término “variables contextuales” como sinómino de “variables derivadas”.

El análisis multinivel puede representarse metafóricamente por la “teoría de la rana en el charco”, que se refiere a la idea que una rana pequeña de individuos puede estar en un charco grande o una rana en un charco pequeño. Aplicado al ejemplo del estudiante indica que un alumno moderadamente inteligente en un contexto sumamente inteligente puede llegar a desmotivarse y así llegar a no tener éxito. Mientras el mismo alumno en un contexto considerado menos inteligente puede llegar a tener éxito. Así, el efecto de una inteligencia individual de alumnos depende por término medio de la inteligencia de otro alumno.

Ahora bien, antes de presentar las definiciones teóricas de los modelos multinivel, se comenta el tipo de variables que se consideran en los modelos multinivel.

- **Variables Globales y absolutas:** son las que se refieren únicamente al nivel en que están definidas, sin referencia a ninguna otra unidad o nivel de análisis. Por ejemplo, la inteligencia podría ser una variable global (se mide a nivel individual).

- **Variables Relacionales:** son las que conforman un nivel simple y describen las relaciones de una unidad a otra siempre dentro del mismo nivel. Por ejemplo, los índices de popularidad de un profesor se consideran variables relacionales en el nivel individual.

- **Variables Analíticas y estructurales:** estas se refieren a variables de niveles inferiores que se agregan en unidades mayores. Las variables analíticas se refieren a la agregación de una variable global en un nivel más bajo. Por ejemplo, las características sociales de los individuos son variables analíticas cuando se agregan para toda una comunidad.

- **Variables contextuales:** estas definen a las superunidades; todas las unidades en los niveles más bajos reciben el mismo valor en la variable que se mide en la superunidad a la cual pertenecen. Por ejemplo, el área geográfica, comunidad, etc.

De las variables descritas anteriormente, se presenta el siguiente esquema, adaptado de Swanborn (1981).

Tabla 2: Esquema adaptado a Swanborn (1981).

Nivel	**1**	**2**	**3**	**…**
Tipo de variable	Global ⇒	Analítica		
	Relacionales ⇒	Estructurales		
	Contextual ⇐	Global ⇒	Analítica	
		Relacionales ⇒	Estructurales	
		Contextual ⇐	Global ⇒	…
			Relacionales ⇒	…
			Contextual ⇐	…

Idealmente, una teoría de multinivel debe especificar cuales variables pertenecen a cada nivel, es decir, investigar de qué manera las variables grupales (macrovariables), las individuales (microvariables) y sus interacciones se relacionan con los resultados a nivel individual. Los efectos entre el individuo y el nivel del contexto requieren la especificación de algún proceso dentro de los individuos que causa que esos individuos sean influidos diferencialmente por ciertos aspectos de su contexto.

Tratando el ejemplo que venimos explorando, observamos que se han definido dos niveles de agregación, el estudiante y la escuela. Y según el esquema dado anteriormente, podríamos agregar variables derivadas de un nivel superior a un nivel más bajo, tal es el caso como las características socioeconómicas del estudiante y familia, capital cultural y social de las familias, áreas urbanas o rurales, el grado de involucramiento de las familias de los alumnos en actividades escolares, y el apoyo de niveles superiores del sistema educativo.

Con el fin de profundizar en el concepto de agregado y desagregado, tomaremos siempre el ejemplo del estudiante. Lo que se pretende es medir la variable endógena (rendimiento del estudiante), en primera instancia agregamos la variable contextual nivel socioeconómico del estudiante. Esto lo hacemos para que conozcamos cual será el impacto real de los diferentes factores sobre el rendimiento de los estudiantes, lo cual eliminaremos el efecto que el nivel socioeconómico del estudiante y su familia tienen sobre los resultados académicos. Ahora bien si desagregamos la variable nivel socioeconómico de la escuela, entonces este modelo busca estimar el rendimiento promedio de los estudiantes en las diferentes pruebas eliminado el efecto que sobre los puntajes tiene el nivel socioeconómico medio de los alumnos del centro. En otras palabras, trata de depurar la media general del rendimiento, aislando el efecto que sobre ella tiene el nivel socioeconómico de cada establecimiento, expresado en términos del promedio del nivel socioeconómico de sus estudiantes.

Tipos de modelos multinivel.

En los ejemplos citados anteriormente y mediante el esquema de variables adaptado a Swanborn (1981), un modelo multinivel de dos niveles de jerarquía, pueden ser expandido a más niveles de jerarquía. Algunos investigadores pueden estar interesados en describir los cambios de las respuestas de una enfermedad con respecto al tiempo y, desde esta perspectiva, el análisis multinivel puede ser utilizado para mostrar el efecto temporal que se tiene sobre los grupos de estudio. Por ejemplo, en un estudio longitudinal en el nivel 1 se consideraría los individuos. Así, el nivel 2 representaría las mediciones repetidas en diferentes lugares, y el nivel 3 podrían ser los lugares. Algunos ejemplos de estructura de encuestas transversales repetidas, se presentan en los estudios multicéntricos[5].

Otro tipo es el de estructura de mediciones repetidas se podría considerar en un nivel 1 la medición de la ocasión, el nivel 2, el individuo y nivel 3, los lugares del país. Así, el nivel 1 representa la medición repetida del evento en mismo individuo. Tales estructuras permiten explorar el cambio individual dentro de un grupo.

Los modelos multinivel pueden utilizarse también para representar diferentes variables respuestas relacionadas de un mismo individuo. En la necesidad de examinar diversas variables simultáneamente para evaluar, por ejemplo, la salud física de un individuo, las diferentes mediciones del cuerpo formarían un conjunto de respuestas en el nivel 1, las cuales estarían anidadas entre individuos (nivel 2), que a su vez estarían anidados en

[5] Benavides FG, Benach J, Diez-Roux AV, Roman C. How do types of employement relate to health indicators? Findings from the second European survey on working conditions. J Epidemiol Community Health 2000.

diferentes instituciones de salud. Este tipo se conoce como estructura de clasificación cruzada. Estudios con este enfoque pueden ser ecológicos o espaciales[6].

1.4 CONJETURACIONES PREVIAS AL ANÁLISIS MULTINIVEL.

En diversas áreas de investigación se han observado la existencia de estructuras jerárquicas en los datos, producto de la agrupación de unidades dentro de otras unidades en diferentes niveles que conforma una jerarquía. Por ejemplo, una muestra de una población se puede describir como una muestra llamado pasos múltiples: primero tomamos una muestra de unidades del nivel más alto (por ejemplo, las escuelas), y luego probamos las sub-unidades de las unidades disponibles (por ejemplo, buscar características de los estudiantes y de las escuelas). En tales muestras, las observaciones individuales generalmente no son completamente independientes. Esto es, porque los estudiantes en la misma escuela tienden a ser semejantes unos de otros, a causa de procesos de selección (por ejemplo, algunos profesores pueden darle mérito a algunos estudiantes considerados talentosos, mientras que otros seria lo contrario) y a causa del historial común que ellos han compartido en la misma escuela. Como resultado de eso, la correlación mediana (expresado en el llamado intra-correlación de clase) entre variables medidas en estudiantes de la misma escuela, será más alto que la correlación mediana entre variables medidos en estudiantes de escuelas diferentes. Las diversas pruebas estadísticas uniformes se inclinan pesadamente en la suposición de independencia de las observaciones. Pero si esta suposición se viola, las estimaciones de los errores uniformes de pruebas estadísticas convencionales serán demasiado pequeñas, y dará resultados falsamente significativos. A todo esto se le considera un efecto de diseño, y el procedimiento para la corrección será calcular los errores uniformes por métodos ordinarios de análisis, es decir, estimar la intra-correlación de clase entre los encuestados dentro de grupos. Algunos de estos procedimientos de corrección son bastantes poderosos (Skinner, Holt & Herrero, 1989). De hecho, estos procedimientos de corrección se podrían aplicar en el análisis multinivel. Sin embargo, en la mayoría de los problemas multinivel es que no solo tenemos anidados los individuos dentro de grupos, sino que

[6] Bramm AW, Van den EP, Prince MJ, Beekman AT, Kivela SL, Lawlor BA et, al. Religión as a crosscultural determinant of depresión in elderly Europeans: results from the EURODEP collaboration. Psychol Med 2001.

también tenemos variables medidas en todos los niveles disponibles. Las variables que combinan los niveles diferentes en un modelo estadístico son un problema diferente que estimar y corregir el diseño. Es por ello que los modelos multinivel se diseñan para analizar las variables de niveles diferentes simultáneamente, usando un modelo estadístico que incluya varias dependencias.

Supongamos por ejemplo, si se quiere estudiar a escala nacional la calidad de vida de los ancianos que viven en una residencia de forma permanente. En este caso se puede establecer una estructura jerárquica de 2 niveles, donde los ancianos se sitúan en el nivel 1 y las residencias en el nivel 2. Podemos suponer que dentro de cada residencia los ancianos tienen a ser más similares en sus comportamientos y características, o más aun, las residencias presentan características propias, lo que puede afectar a la calidad de vida de los ancianos. Si la calidad de vida se considera una variable de respuesta de tipo continua, a la que se le ajusta un modelo de regresión, considerando una característica (exógeno) de los ancianos, podríamos encontrar como resultado una de estas tres situaciones[7], tal como se observa en la figura 4:

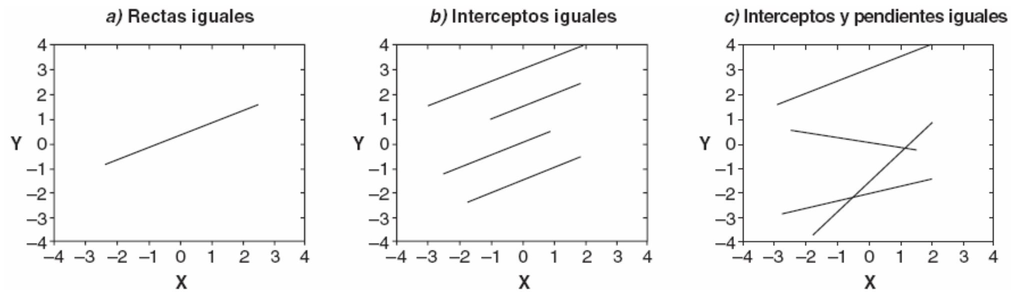
a) Las rectas de regresión para cada residencia son iguales, lo que llevaría a considerar un único nivel en el análisis de los datos.

b) Las rectas de regresión difieren solo en el intercepto.

c) difieren tanto en el intercepto como en la pendiente, lo que nos lleva a pensar en construir un modelo que considere los distintos niveles en la jerarquía de los datos.

[7] Catalán- Reyes MJ., et al. Utilización de los modelos multinivel en investigación sanitaria.

Figura 4: Ejemplo hipotético de un ajuste de regresión para la calidad de vida de los ancianos en cuatro residencias.



Para hacer frente a esta problemática, el análisis multinivel es una metodología fundamental para el análisis de datos que presentan una estructura jerárquica, y principalmente cuando los datos se han obtenido a través de la técnica de muestro por conglomerado, es decir, ocurre cuando en una investigación existen además de los sujetos, otras unidades de análisis tales como escuelas, municipios, empresas, hospitales, etc. La existencia de estas agrupaciones naturales queda reflejada en la estructura de los datos, hace que se incumpla el supuesto del muestreo aleatorio simple y que muchas técnicas estadísticas convencionales, tales como el análisis de varianza (ANOVA) efectos fijos o la regresión clásica sea inapropiada.

1.5 ENFOQUES MULTINIVEL BASADO POR EL ANÁLISIS DE REGRESIÓN MULTIPLE, ANOVA Y ANCOVA.

Históricamente, los problemas de multinivel han guiado a enfoques de análisis que basándose por un análisis de regresión múltiple, el análisis de variación, o de algún otro método uniforme de análisis. Una situación análoga en un contexto de análisis de regresión diferencia el análisis de covarianza clásico del modelo multinivel. Recordemos que el análisis de covarianza es una combinación de análisis de varianza y de regresión clásico. Se haría un análisis de regresión ordinario, por ejemplo, entre el número de horas de preparación para un examen y rendimiento académico a fin de estudiar la eficacia de los hábitos de estudio en una población de alumnos. En el análisis de covarianza clásico se incluiría el estudio de esta relación en un diseño de dos o más grupos, como por ejemplo, dos tipos de entrenamiento en hábitos de estudio con respecto a un grupo de control. Los sujetos en este caso podrían ser asignados aleatoriamente a uno de los grupos y se podrían contrastar hipótesis tanto sobre las diferencias entre los métodos de estudio, como sobre la eficacia del tiempo invertido.

Los efectos de interés en este caso son fijos, puesto que se han elegido premeditadamente dos programas de entrenamiento de estudios específicos frente a un grupo de control. Tendríamos análisis de covarianza de efectos mixtos o modelo multinivel si en lugar de estos programas concretos, el objetivo fuera, por ejemplo, el contrastar la eficacia de las horas de estudio en una selección de escuelas al azar. En este caso habría dos tipos de unidades experimentales, la escuela y el alumno, y las medidas de los alumnos dentro de cada escuela estarían correlacionadas, por lo que el modelo de efectos fijos no sería apropiado.

En el análisis multinivel, todo el conjunto de problemas es conceptual, es decir, si el analista no es muy cuidadoso en la interpretación de los resultados, puede cometer la falacia del nivel injustificadamente, que se compone de analizar los datos en un nivel, y las conclusiones que dibujan en otro nivel. Una de las falacias mejor conocidas es la falacia ecológica, que interpreta los datos agregados en el nivel individual, es decir, inferir conclusiones a escala individual a partir de datos grupales. La otra falacia que podría surgir es el hacer inferencias en un nivel más alto a partir de un análisis realizado en un nivel más bajo, se conoce también la falacia atomística, es decir, hacer inferencias sobre la variabilidad intergrupal (o relación entre variables grupales) a partir de datos individuales. A este concepto a veces se ha denominado falacia individualista.

Una manera más general de ver los datos multinivel es investigar las relaciones entre las variables que se miden en varios niveles jerárquicos diferentes. Por ejemplo, una pregunta común es cómo varias variables individuales y de grupo influyen una sola variable individual del resultado. Típicamente, parte de las variables individuales. En el pasado, tales datos se analizaban cuando generalmente un análisis de regresión múltiple convencional con una variable dependiente en el nivel más bajo (individuo) y una colección de forma explicativa de variables en todos los niveles disponibles. (Boyd y Iversen, 1979; Roberts y Burtein, 1980). Desde que este enfoque analiza todos los datos disponibles en un solo nivel, se sufre el problema conceptual y estadístico como el que producir sesgos en la estimación del error típico de medida y un aumento en la probabilidad de cometer el error de rechazar la hipótesis de asociación lineal cuando esta es cierta.

El análisis multinivel modela explícitamente estas relaciones jerárquicas, eliminando estos sesgos, proporcionando además estimaciones de interés psicológico sobre la variabilidad y replicabilidad de los coeficientes de regresión en los distintos contextos sociales (Robinson, 1950), y sobre la influencia de estos en el comportamiento del individuo.

En los últimos 50 años se han desarrollado diferentes métodos de estimación tal es el caso el de componente de varianza, aplicable a datos desequilibrados bajo modelos mixtos. Es decir, la existencia de correlación entre observaciones hace que el modelo de regresión por mínimos cuadrados ordinarios produzca sesgos en la estimación del error cuadrático medio, y por tanto producirá errores de inferencia. El grado de sesgo depende de la magnitud de la correlación entre los grupos tanto de la variable endógena como la exógena, así como del número de unidades experimentales.

En diversos estudios como en psicología, es frecuente que un contexto de regresión implicaría que no solamente el número de unidades de nivel inferior sea el mismo para cada unidad experimental del nivel superior, sino que también la distribución de los valores de las variables exógenas sea también la misma para todos los grupos del nivel superior. Tal como ocurre en el análisis de experimentos en estos casos, es aconsejable recurrir al procedimiento de Máxima Verosimilitud que es un método clásico de estimación de parámetros (no necesariamente varianza) asociada a funciones de densidad o probabilidades de variables aleatorias, con la restricción de que los parámetros han de estar en el espacio paramétrico. Produciendo estimadores con propiedades deseables con muestras grandes, tales como consistencia y eficacia, es decir, si se recoge gran cantidad de datos, el estimador será aproximadamente insesgado y con varianza mínima.

1.6 EL ERROR DE MEDIDA.

En muchas áreas de la investigación, los estudios pueden implicar variables que no se pueden observar directamente. Por ejemplo, la capacidad matemática de una persona no se puede medir directamente, solamente el funcionamiento en una serie de preguntas en una prueba de matemática. También, los datos recogidos de encuestadores contienen error de respuesta (es decir, hay variación de la respuesta en respuestas a la misma pregunta cuando está administrado en varias ocasiones a la misma persona). El error de medida puede ocurrir en variables explicativas, dependientes o independientes. La confiabilidad de variables explicativas es una pregunta metodológica importante. Cuando se sabe la confiabilidad, las correcciones se pueden hacer (Fuller, 1987), si las medidas repetidas están disponibles, la confiabilidad se puede incorporar en el modelo y estimar directamente. El uso de variables explicativas no fiables conduce a la valoración en polarización negativa de los coeficientes de la regresión y la inferencia estadística que resulta puede ser muy engañosa a menos que se hagan los ajustes cuidadosos.

Por otro lado, ya en el contexto de modelos de ecuaciones simultaneas o modelos de ecuaciones estructurales con variables observables, en ocasiones denominados path análisis, en Gillespie y Fox (1980) se señalan los posibles efectos en modelos recursivos, modelos en que la causalidad fluye en un único sentido, indicando que se produciría atenuación en los coeficientes gamma, aquellos que van de variables exógensas a endógenas, y sin embargo sesgo positivo en los coeficientes beta (aquellos que van de una variables endógenas a otras). Este último efecto podría llevar a una atenuación de las varianzas entre errores. Resultados para un modelo no recursivo (Fergusson y Horwood, 1986) mostraron que al aumentar el error de medida aleatorio, aumentaba el valor de los coeficientes gamma, mientras los coeficientes beta disminuían. Thompson y Getty (1994) ofrecen otro ejemplo de efectos de la fiabilidad de las medidas en un modelo no recursivo, modelo en que se plantea doble sentido en la causalidad donde una variable antecedente puede también ser consecuente. Es por ello que una forma alternativa de tratar el error de medida aleatoria es mediante modelos de ecuaciones estructurales con variables latentes. Estos permiten efectos, predicciones entre factores, en lugar de solo plantear relaciones entre variables observables. El concepto y la teoría estadística que permite las ecuaciones de regresión simultáneas entre factores fueron introducidos por Jöreskog (1970; 1973), y se basan en una idea simple, incorporar en un mismo modelo el análisis factorial confirmatorio y el modelo de ecuaciones estructurales con variables observables.

Aunque el asunto del modelo de error de medida ha recibido hoy en día la verdadera atención considerable en la literatura multinivel, ésta atención se ha centrado en los modelos lineales de error de medida, más específicamente, el modelo aditivo clásico del error de medida (Carroll et al., 1995; Fuller, 1987; Goldstein, 1987; Longford, 1993). Este modelo se basa en la asunción de la homocedasticidad, que exige la variación igual de los errores de medida condicionales en diversos niveles de la variable endógena. Es decir, en la terminología usado cuando se discuten las medidas específicas por estrato son “homogéneas” cuando son iguales y “heterogéneas” cuando son significativamente diferentes. Obviamente, una medida resumen es mejor en una situación en que la medida que esta siendo resumida es homogénea en los estratos. En caso habitual, para una medida de razón efecto, la homogeneidad entre los estratos es equivalente a tasas o razones que se adaptan a un modelo multiplicativo de efectos conjuntos (absolutos), la homogeneidad es equivalente a un modelo aditivo de efectos conjuntos. Modificación de la medida del efecto significa heterogeneidad para esa medida. Por ejemplo en la investigación educativa, el gasto del pre-test de los estudiantes, el estado socioeconómico o la inteligencia se utilizan a menudo como variables explicativas en los resultados del examen de los estudiantes. Además, los resultados del examen de los estudiantes o las capacidades se miden conforme a error o no se pueden observar directamente. Los errores de medida asociados a las variables explicativas o a las variables que no pueden ser observadas directamente no se toma en consideración.

La atención esta también en un modelo no lineal de error de medida y de un modelo estructural no lineal, esto es debido a que los modelos multinivel fueron desarrollados originalmente para variables con distribución normal y bajo los supuestos de una distribución normal de los errores en cada individuo, estos métodos han sido generalizados para situaciones en los que las variables de respuesta es binomial, nominal u ordinal y para procesos donde la probabilidad del evento es pequeña y se puede modelar con una distribución de Poisson. Se llama función vínculo de nivel 1 (link funtion) a la transformación de la variable dependiente de nivel 1 que se iguala a una combinación lineal de los coeficientes de las variables explicativas. Esta función puede ser una función logística binomial, ordinal, multinomial o una transformación de Poisson.

CAPÍTULO 2

TEORÍA DE LOS MODELOS MULTINIVEL

Introducción.

En este capítulo se presenta el fundamento teórico de modelos multinivel, el cual esta estructurado en los siguientes apartados:

- El modelo de dos niveles y su notación, en primer lugar se aborda este contenido a partir de un ejemplo desarrollado por Golstein en 1999, donde se hacen algunas conjeturaciones sobre el modelo de regresión clásica. Se define el modelo parcial con situación multigrupo, estableciendo así los supuestos para un modelo de regresión simple.

- El modelo de dos niveles, se define el modelo lineal mixto y se estudian los supuestos básicos.

- Estimación de los parámetros para el modelo de componentes de varianza, aquí se estiman los parámetros del modelo multinivel de dos niveles pero considerando el modelo donde no se incluye ninguna variable aleatoria al modelo (el enfoque ANOVA).

- El modelo de 2 niveles incluyendo coeficientes aleatorios, en este apartado se realiza la estimación de los parámetros del modelo multinivel de dos niveles, considerando una variable aleatoria al modelo (el enfoque ANCOVA).

- Estructura general y estimación para un modelo multinivel de dos niveles, aquí se estiman los parámetros del modelo multinivel general de dos niveles. Se parte con la incorporación de dos variables aleatorias al modelo lineal mixto, para poder así generalizarlo a tantas variables aleatorias que se deseen incorporar al modelo multinivel de dos niveles.

- Estimación de los parámetros del Modelo general de tres niveles. De igual manera como se desarrolló la estimación de los parámetros tanto fijo como aleatorios de un modelo multinivel general de dos niveles, se hace para un modelo multinivel general de tres niveles.

2.1 El modelo de 2 niveles y notación básica

Para entrar en materia, se presenta un ejemplo del estudio realizado por Goldstein (1999) a partir de los resultados obtenidos por alumnos en escuelas primarias (Junior School Proyect) en Londres, realizado por Mortimore et al (1988). Goldstein, utilizó una submuestra aleatoria de la data de Mortimore, considerando 728 alumnos en 50 escuelas y como unidad de medida a los alumnos que están en cuarto año de aprendizaje, en el cual los alumnos cumplen sus ocho años de vida. Por otra parte, dentro de este estudio se utilizaron las puntuaciones de la prueba de matemática administrada en dos momentos junto con la información recogida del contexto social de los alumnos y de su género.

Regresión lineal simple

Para introducirnos a la teoría multinivel se hacen algunas conjeturaciones sobre que tipo de relación seria de interés conocer a partir de la información de los gráficos.

En la figura 2.1 según el ejemplo tratado se muestra el diagrama de dispersión de las puntuaciones de la prueba de matemática en alumnos de 11 años de edad sobre las puntuaciones de la prueba de matemática en alumnos de 8 años de edad. En este diagrama no se hace ninguna distinción entre las escuelas a las cuales los alumnos pertenecen. Observamos que existe una dispersión estrecha de las puntuaciones de alumnos en edad de 11 años con el aumento de las puntuaciones de alumnos en edad de 8 años. Es importante recalcar que al no haber distinción entre escuelas, no podemos ver si la escuela influye sobre las puntuaciones de los alumnos.

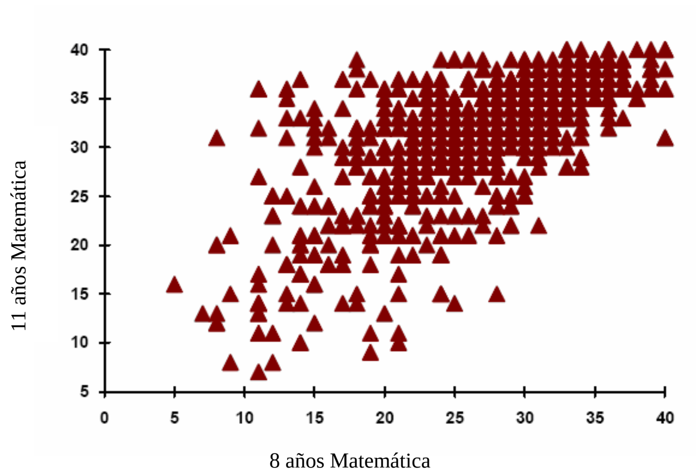


Figura 2.1: Diagrama de dispersión de las puntuaciones de la prueba de las matemáticas en alumnos de 11 años de edad sobre las puntuaciones de la prueba de matemática en alumnos de 8 años de edad.

En figura 2.2 se presenta para un caso particular de dos escuelas que han sido seleccionadas aleatoriamente, representadas por diversos símbolos. Observamos que conforme aumentan las puntuaciones de los alumnos de 8 años, las puntuaciones de alumnos de 11 años están entre 20 y 30 para la escuela 1 (símbolo círculo). Ahora bien, para esa misma escuela las puntuaciones de mayor edad se sobre pone a las puntuaciones de alumnos con menor edad. Sin embargo, si trazamos dos líneas de regresión para dichas escuelas, se tiene que las rectas no son paralelas, indicando que la escuela 2 (símbolo triángulo) tiene mejores puntuaciones en la prueba que la escuela 1. Además, hay un punto de intersección en las dos rectas o un balance de las puntuaciones obtenidas en alumnos de 8 y 11 años. Pero que el cambio surge después de ese punto de intersección, se observa que no solo la edad o variable explicativa del nivel alumno influye en su puntuación, sino que podemos pensar que existen otras características de la escuela, de tal modo que las características de la escuela estarían influyendo en las puntuaciones de los alumnos.

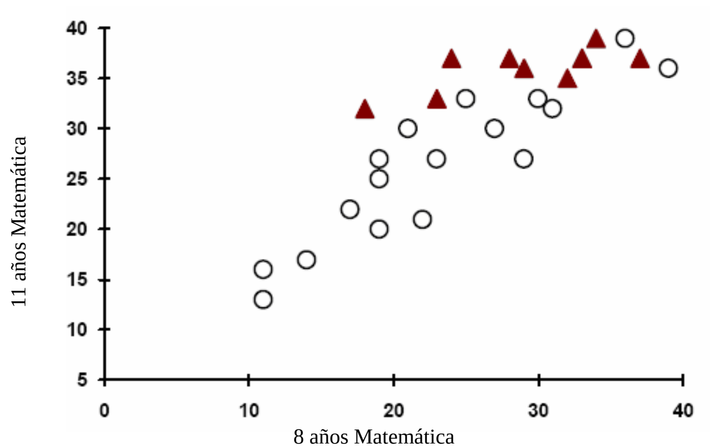


Figura 2.2: Diagrama de dispersión de las puntuaciones de la prueba de matemática para dos escuelas.

De las conjeturaciones hechas anteriormente, se escribe un modelo de regresión simple para una escuela, relacionando las puntuaciones de la prueba de matemática en alumnos de 11 años con las puntuaciones de 8 años, de la siguiente manera:

$$y_i = \alpha + \beta x_i + e_i \tag{2.1}$$

Donde y_i es la puntuación del i-ésimo alumno, x_i la edad del alumno, $\beta = (\beta_1, \beta_2, \ldots, \beta_n)'$ es el efecto que tiene la edad sobre la puntuación del i-ésimo alumno y $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)'$ es el promedio de las puntuaciones eliminando el efecto de la edad de los alumnos.

Pero en la ecuación 2.1 solamente tenemos una regresión que permite conocer el efecto de la edad sobre la puntuación del estudiante en una escuela en particular con una muestra de tamaño k_1 estudiantes. Ahora bien, si queremos conocer el efecto de la edad sobre la puntuación de los estudiantes de más de una escuela, con muestras de tamaño k_j en cada una, entonces tendríamos n modelos de regresión lineal, de tal modo que podamos conocer que tanto influye la edad sobre la puntuación del i-ésimo alumno en la j-ésima escuela. Esto implica el ajuste de n modelos de regresión, mediante una forma parcial para cada escuela con tamaño k_j, tenemos

Para la escuela 1: $y_{i1} = \alpha_1 + \beta_1 x_{i1} + e_{i1}$; $i = 1, 2, ..., k_1$

Para la escuela 2: $y_{i2} = \alpha_2 + \beta_2 x_{i2} + e_{i2}$; $i = 1, 2, ..., k_2$

$\vdots$

Para la escuela n: $y_{in} = \alpha_n + \beta_n x_{in} + e_{in}$; $i = 1, 2, ..., k_n$

El sistema de ecuaciones de los n modelos anteriores se puede simplificar con el modelo siguiente:

$$y_{ij} = \alpha_j + \beta_j x_{ij} + e_{ij} \qquad i = 1, 2, ..., k_j \ ; \qquad j = 1, 2, ..., n \tag{2.2}$$

Donde

y_{ij} es la puntuación del i-ésimo alumno en la j-ésima escuela.

x_{ij} es la edad sobre la puntuación del i-ésimo alumno en la j-ésima escuela.

α representa el promedio de la puntuación muestral.

β_j representa los pesos asociados a la característica x_{ij} en la muestra completa.

e_{ij} es una variable aleatoria que representa el error de ajuste del modelo del i-ésimo alumno en la j-ésima escuela.

Los e_{ij} deben cumplir las siguientes hipótesis:

a) La perturbación tiene esperanza nula, es decir:

$$E\left(e_{ij}\right) = 0$$

b) La varianza de la perturbación es siempre constante, y no depende de x; conocido como homocedasticidad de la perturbación.

$$\operatorname{var}\left(e_{ij}\right) = \sigma_{eo}^{2}$$

c) La perturbación tiene una distribución normal. Esta hipótesis es consecuencia del teorema central del límite.

d) Las perturbaciones son independientes entre si.

Estas cuatro condiciones pueden expresarse igualmente respecto a la variable respuesta, o dependiente, como sigue:

a) La esperanza de la respuesta depende linealmente de x. Tomando esperanzas en (2.2), se tiene:

$$\begin{aligned} E\left(y_{ij}\right) &= E\left(\alpha_j + \beta_j x_{ij} + e_{ij}\right) \\ &= \alpha_j + \beta_j x_{ij} \end{aligned}$$

b) La varianza de la distribución de y_{ij} es constante.

$$\operatorname{var}\left(y_{ij}\right) = \sigma_{eo}^{2}$$

c) La distribución de y para cada x es normal.

d) Las observaciones y_{ij} son independientes entre sí.

Ahora, si utilizamos el modelo de regresión (2.2) se tendrían que estimar α, β y σ^2_{eo} que representa 2n+1 parámetros, suponiendo que α_j, β_j y σ^2_{eo} es fijo para cada escuela $j = 1, 2, ..., n$.

2.2 El modelo de dos niveles.

Con el fin de conocer otras variables aleatorias que no han sido medidas en el alumno se debe considerar los parámetros como variables aleatorias. Es por ello que para hacer la ecuación (2.2) más auténtica de dos niveles, dejamos α_j y β_j como variables aleatorias convertidas. Para la consistencia de la notación sustituiremos α_j por β_{0j} y β_j por β_{0j} y asumiremos que

$$\beta_{0j} = \beta_0 + u_{0j}, \qquad \beta_{1j} = \beta_1 + u_{1j}, \quad j = 1, 2, ..., n \tag{2.3}$$

Donde

β_{0j} es el logro promedio por escuela, está representado como una función de la gran media β_0 o media de todas las escuelas, más una variable aleatoria u_{0j} que captura la variación en la puntuación promedia a través de escuelas. Dicho de otra forma el u_{0j} contiene las variables no observables que hacen que la nota promedio de cada escuela no sean iguales. Por ejemplo, si todas las escuelas tuviesen media constante, entonces $u_{0j} = 0$.

β_{1j} es el efecto de la variable edad, está representado como una función de la estimación de la media de las pendientes relativas al efecto de la variable edad más una variable aleatoria u_{1j} que captura la variación de los pesos asociados a la característica edad a través de escuelas. La variable u_{1j} recoge todas las variables no observables en la escuela que influyen en la edad de los estudiantes y hace que el peso de la variable edad difiera entre escuelas. Si el peso de la variable edad de los estudiantes fuese el mismo en todas las escuelas, entonces $u_{1j} = 0$.

Sustituyendo (2.3) en (2.2) tenemos que,

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + \left(u_{0j} + u_{1j} x_{ij} + e_{0ij}\right) \tag{2.4}$$

Donde y_{ij} se ha expresado como la suma de una parte fija del modelo $\beta_0 + \beta_1 x_{ij}$ y una parte aleatoria $u_{0j} + u_{1j} x_{ij} + e_{0ij}$ dentro de las escuelas. Que al final la expresión (2.4) resulta ser un modelo lineal de efectos mixtos.

Para las perturbaciones se establecen los siguientes supuestos:

a) Las variables aleatorias u_{0j} y u_{1j} tiene esperanza nula, es decir:

$$E\left(u_{0j}\right) = E\left(u_{1j}\right) = 0 \tag{2.5}$$

b) La varianza de cada variable aleatoria u_{0j} y u_{1j} es siempre constante, y no depende de x.

$$\mathrm{var}\left(u_{0j}\right) = \sigma^2_{u0}, \quad \mathrm{var}\left(u_{1j}\right) = \sigma^2_{u1}, \quad \mathrm{cov}\left(u_{0j}, u_{1j}\right) = \sigma_{u01} \tag{2.6}$$

c) Las perturbaciones son independientes entre si.

Para la variable respuesta o dependiente, se tiene que:

a) La esperanza de la respuesta depende linealmente de x. Tomando esperanzas en (2.4), se tiene que:

$$E\left(y_{ij}\middle|\beta_0,\beta_1,x_{ij}\right)=E\left(\beta_0+\beta_1 x_{ij}+\left(u_{0j}+u_{1j}x_{ij}+e_{0ij}\right)\right)$$
$$=E\left(\beta_0\right)+E\left(\beta_1 x_{ij}\right)+E\left(u_{0j}\right)+E\left(u_{1j}x_{ij}\right)+E\left(e_{0ij}\right)$$

Y según el supuesto (2.5) la esperanza matemática de la variable respuesta resulta ser,

$$\hat{y}_j \cong E\left(y_{ij}\middle|\beta_0,\beta_1,x_{ij}\right)=\beta_0+\beta_1 x_{ij} \tag{2.7}$$

b) La distribución de y para cada x es normal.

c) La varianza de las perturbaciones del i-ésimo alumno en la j-ésima escuela e_{0ij} es constante.

$$\operatorname{var}\left(e_{0ij}\right)=\sigma_{eo}^2 \tag{2.8}$$

Hay que observar que la expresión (2.3) no existe una medición de una variable en la j-ésima escuela. Solamente hemos considerado el caso simple de variables del alumno (nivel 1).

2.3. Estimación de los parámetros para el modelo de componentes de varianza.

La ecuación (2.4) requiere la estimación de dos coeficientes fijos β_0, β_1, y de cuatro parámetros aleatorios como σ_{u0}^2, σ_{u1}^2, σ_{e0}^2 y σ_{e01}. Nos referimos a tales varianzas y covarianzas como parámetros aleatorios. Comenzaremos, sin embargo, considerando el modelo simple de componentes de varianza de 2 niveles porque incluye únicamente los parámetros aleatorios σ_{u0}^2, σ_{e0}^2, y que además no se han incluido los coeficientes aleatorios, así

$$y_{ij}=\beta_0+\beta_1 x_{ij}+\left(u_{0j}+e_{0ij}\right) \tag{2.9}$$

El modelo (2.9) puede ser estimado por el procedimiento de mínimos cuadrados. Para ver la diferencia más extrema de la ecuación (2.9) con (2.2) tenemos que obtener las varianzas para cada nivel. Para ello, encontremos la matriz de varianza-covarianza total (V) para $i = 1, 2, ..., k_j$ alumnos en $j = 1, 2, ..., n$ escuelas.

Bajo este esquema y con los supuestos antes mencionados para un modelo multinivel de dos niveles, la matriz de varianza-covarianza total (V) toma la siguiente forma:

$$V = E\left[\left(y_{ij} - E\left(y_{ij}\right)\right)\left(y_{ij} - E\left(y_{ij}\right)\right)'\right] \tag{2.10}$$

donde $\left(y_{ij} - E\left(y_{ij}\right)\right)$ es un vector de dimensión $\left(k_1 + k_2 + ... + k_n\right) \times 1$ que lo notaremos como $\left(y_{ij} - \hat{y}_j\right)$ para realizar los cálculos, así tenemos que:

$$V = E\left[\left(y_{ij} - \hat{y}_j\right)\left(y_{ij} - \hat{y}_j\right)'\right] \tag{2.11}$$

La matriz de varianza-covarianza total (V) resulta ser una matriz de orden $\left(k_1 + k_2 + ... + k_n\right) \times \left(k_1 + k_2 + ... + k_n\right)$, $j = 1, 2, ..., n$, así:

$$V = E\left[\begin{pmatrix} y_{11}-\hat{y}_1 \\ y_{21}-\hat{y}_1 \\ \vdots \\ y_{k_1 1}-\hat{y}_1 \\ \vdots \\ y_{1n}-\hat{y}_n \\ y_{2n}-\hat{y}_n \\ \vdots \\ y_{k_n n}-\hat{y}_n \end{pmatrix} \left(y_{11}-\hat{y}_1, y_{21}-\hat{y}_1, \ldots, y_{k_1 1}-\hat{y}_1, \ldots, y_{1n}-\hat{y}_n, y_{2n}-\hat{y}_n, \ldots, y_{k_n n}-\hat{y}_n\right)\right]$$

Sustituyendo (2.7) y (2.9) en (2.11), se tiene que

$$V = E\left[\begin{pmatrix} u_{01}+e_{011} \\ u_{01}+e_{021} \\ \vdots \\ u_{01}+e_{0k_1 1} \\ \vdots \\ u_{0n}+e_{01n} \\ u_{0n}+e_{02n} \\ \vdots \\ u_{0n}+e_{0k_n n} \end{pmatrix} \left(u_{01}+e_{011}, u_{01}+e_{021}, \ldots, u_{01}+e_{0k_1 1}; \ldots; u_{0n}+e_{01n}, \ldots, u_{0n}+e_{0k_n n}\right)\right]$$

Para ilustrar el procedimiento de cálculo, se presenta un caso particular, considerando dos escuelas y dos estudiantes. De tal modo que logremos encontrar la matriz de varianza-covarianza total (V) de dos alumnos, uno por escuela, es decir, sería realizar los cálculos solamente en un bloque de la matriz (V), así:

$$V_2 = E\left[\left(u_{0j}+e_{0ij}\right)\left(u_{0j}+e_{0ij}\right)'\right]; \quad j=1,2;\; k_1=1;\; k_2=1.$$

$$= E\left[\begin{pmatrix} u_{01}+e_{011} \\ u_{02}+e_{012} \end{pmatrix}\left(u_{01}+e_{011}\,;u_{02}+e_{012}\right)\right]$$

$$= E\begin{bmatrix} \left(u_{01}+e_{011}\right)^2 & \left(u_{01}+e_{011}\right)\left(u_{02}+e_{012}\right) \\ \left(u_{02}+e_{012}\right)\left(u_{01}+e_{011}\right) & \left(u_{02}+e_{012}\right)^2 \end{bmatrix}$$

Introduciendo el valor esperado para cada elemento de la matriz de V,

$$V_2 = \begin{bmatrix} E\left[\left(u_{01}+e_{011}\right)^2\right] & E\left[\left(u_{01}+e_{011}\right)\left(u_{02}+e_{012}\right)\right] \\ E\left[\left(u_{02}+e_{012}\right)\left(u_{01}+e_{011}\right)\right] & E\left[\left(u_{02}+e_{012}\right)^2\right] \end{bmatrix}$$

Ahora desarrollemos cada uno de los valores esperados de la matriz V_2 asumiendo los supuestos del modelo de dos niveles, de la manera siguiente:

$$E\left[\left(u_{01}+e_{011}\right)^2\right] = E\left[\left(u_{01}\right)^2\right] + E\left[\left(e_{011}\right)^2\right] + E\left[2u_{01}e_{011}\right]$$

$$= \sigma^2_{u01} + \sigma^2_{e011}$$

$$= \sigma^2_{u0} + \sigma^2_{e0}$$

$$E\left[\left(u_{02}+e_{012}\right)^2\right] = E\left[\left(u_{02}\right)^2\right] + E\left[\left(e_{012}\right)^2\right] + E\left[2u_{02}e_{012}\right]$$

$$= \sigma^2_{u02} + \sigma^2_{e012}$$

$$= \sigma^2_{u0} + \sigma^2_{e0}$$

$$E\left[\left(u_{01}+e_{011}\right)\left(u_{02}+e_{012}\right)\right] = E\left[u_{01}u_{02}+u_{02}e_{012}+e_{011}u_{02}+e_{011}e_{012}\right]$$

$$= E\left[u_{01}u_{02}\right] + E\left[u_{02}e_{012}\right] + E\left[e_{011}u_{02}\right] + E\left[e_{011}e_{012}\right]$$

$$= 0$$

Luego tenemos que,

$$V_2 = \begin{bmatrix} \sigma^2_{u0} + \sigma^2_{e0} & 0 \\ 0 & \sigma^2_{u0} + \sigma^2_{e0} \end{bmatrix} = \begin{bmatrix} \Omega_{k_1} & 0 \\ 0 & \Omega_{k_2} \end{bmatrix}$$

Como podemos observar de la matriz V_2 se desprende que la varianza de respuesta condicionada en los parámetros y sus predictores fijos es la suma de las varianzas del nivel 1 y del nivel 2, para n escuelas con muestras de k_j alumnos, así,

$$\operatorname{var}\left(y_{ij} \middle| \beta_0, \beta_1, x_{ij}\right) = \operatorname{var}\left(u_0 + e_{0ij}\right) = \sigma^2_{u0} + \sigma^2_{e0}$$

De este modo, se tiene que para $j = 1, 2, ..., n$, la matriz de varianza-covarianza total (V) toma la siguiente forma;

$$V_j = \begin{bmatrix} \Omega_{k_1} & 0 & \cdots 0 \\ 0 & \Omega_{k_2} & \cdots 0 \\ \vdots & \vdots & \ddots \vdots \\ 0 & 0 & \Omega_{k_n} \end{bmatrix}$$

Donde el bloque-diagonal Ω_{k_j} es una matriz de $k_j x k_j$ de la siguiente forma:

$$\Omega_{k_j} = \begin{bmatrix} \sigma^2_{u0} + \sigma^2_{e0} & \sigma^2_{u0} & \cdots & \sigma^2_{u0} \\ \sigma^2_{u0} & \sigma^2_{u0} + \sigma^2_{e0} & \cdots & \sigma^2_{u0} \\ \vdots & \vdots & & \vdots \\ \sigma^2_{u0} & \sigma^2_{u0} & \cdots & \sigma^2_{u0} + \sigma^2_{e0} \end{bmatrix}$$

Esta estructura 'bloque-diagonal' refleja el hecho de que la covariación entre los estudiantes en diversas escuelas es cero, y se extiende claramente a cualquier número de las unidades del nivel 2. Por otro parte, si en caso particular, se nos pidiera, cual es la distribución normal multivariante de las puntuaciones de 5 estudiantes en dos escuelas, una con tres estudiantes y otra con dos, se tiene lo siguiente:

$$Y \sim NM\left(X\beta,\ V\right)$$

$$\begin{pmatrix} y_{11} \\ y_{21} \\ y_{31} \\ y_{12} \\ y_{22} \end{pmatrix} \sim NM\left(\left(X\beta\right), \begin{pmatrix} \sigma_{u0}^2+\sigma_{e0}^2 & \sigma_{u0}^2 & \sigma_{u0}^2 & 0 & 0 \\ \sigma_{u0}^2 & \sigma_{u0}^2+\sigma_{e0}^2 & \sigma_{u0}^2 & 0 & 0 \\ \sigma_{u0}^2 & \sigma_{u0}^2 & \sigma_{u0}^2+\sigma_{e0}^2 & 0 & 0 \\ 0 & 0 & 0 & \sigma_{u0}^2+\sigma_{e0}^2 & \sigma_{u0}^2 \\ 0 & 0 & 0 & \sigma_{u0}^2 & \sigma_{u0}^2+\sigma_{e0}^2 \end{pmatrix} \right)$$

Una manera más compacta de presentar la matriz de varianza-covarianza total (V), se muestra a continuación:

$$V_2 = \begin{pmatrix} \sigma_{u0}^2 J_{(3)} + \sigma_{e0}^2 \mathrm{I}_{(3)} & 0 \\ 0 & \sigma_{u0}^2 J_{(2)} + \sigma_{e0}^2 \mathrm{I}_{(2)} \end{pmatrix} \tag{2.12}$$

Donde $\mathrm{I}_{(n)}$ es la matriz de identidad de $\left(n\,x\,n\right)$ y $J_{(n)}$ es una matriz de unos de $\left(n\,x\,n\right)$. Los 2 índices son indicados para un modelo de 2 niveles. Observamos que en esta matriz se aprecia mejor las muestras en cada escuela, es decir, el número de alumnos en cada escuela, de modo que las varianzas de los dos modelos de regresión, tanto el (2.2) como (2.5) son diferentes, reflejando así que este último explica mejor la variabilidad de la respuesta, separando dichas variaciones según los niveles que se considere (en este caso dos niveles, alumno y escuela) en el modelo.

Según el ejemplo que venimos tratando, el modelo de variación para cada estudiante es constante y la covariación entre dos estudiantes (denotados por i_1, i_2) en la misma escuela está dada por:

$$\operatorname{cov}\left(u_{0j}+e_{0i_1j},\ u_{0j}+e_{i_2j}\right)=\operatorname{cov}\left(u_{0j},u_{0j}\right)=\sigma_{u0}^{2}$$

Puesto que los residuales del nivel 1 se suponen ser independientes.

Por otra parte es de interés saber la proporción de varianza explicada por las escuelas. Está dado por:

$$\rho=\frac{\sigma_{u0}^{2}}{\sigma_{u0}^{2}+\sigma_{e0}^{2}}$$

Que esta referido como la correlación entre unidades del nivel 2; en este caso la correlación de intra-escuela (mide la proporción de la variación total que hay entre escuelas).

En un modelo con 3 niveles, la opinión con las escuelas, los salones de clase y los estudiantes, tendremos dos correlaciones; la correlación de la intra-escuela que mide la proporción de la variación total entre-escuelas y la correlación de la intra-salones de clase que mide eso entre los salones de clase.

2.4 El modelo de 2 niveles incluyendo coeficientes aleatorios.

Recordemos que el modelo de 2 niveles donde se incluye coeficientes aleatorios está dado por:

$$y_{ij}=\beta_0+\beta_1 x_{1ij}+\left(u_{0j}+u_{1j}x_{1ij}+e_{0ij}\right) \tag{2.13}$$

Las variables explicativas para la parte aleatoria del modelo son a menudo un subconjunto de las variables de la parte fija. Cualquiera de las variables explicativas se pueden medir en cualquiera de los niveles; por ejemplo podemos tener características del estudiante en el nivel 1 o características de la escuela en el nivel 2.

Ahora bien, así como estimamos los parámetros para el modelo de componentes de varianza de dos niveles (2.4). Necesitamos en este caso, estimar los parámetros ya incluyendo coeficientes aleatorios. Todo esto, para lograr encontrar la matriz de varianza-covarianza total (V) de $i=1,2,...,k_j$ alumnos en $j=1,2,...,n$ escuelas. Nos auxiliamos de la forma (2.10) para los cálculos, de la manera siguiente:

$$V = E\left[\left(y_{ij} - E\left(y_{ij}\right)\right)\left(y_{ij} - E\left(y_{ij}\right)\right)'\right] \tag{2.14}$$

La expresión anterior, es una matriz de orden kxk donde $k = k_1 + k_2 + ... + k_n$ para $j=1,2,...,n$. Consideremos además los mismos supuestos del modelo multinivel de dos niveles. Sustituyendo (2.4) y (2.7) en (2.13) obtenemos lo siguiente:

$$\begin{aligned} V &= E\left[\left(\beta_0+\beta_1 x_{ij}+\left(u_{0j}+u_{1j}x_{ij}+e_{0ij}\right)-\left(\beta_0+\beta_1 x_{ij}\right)\right)\left(\beta_0+\beta_1 x_{ij}+\left(u_{0j}+u_{1j}x_{ij}+e_{0ij}\right)-\left(\beta_0+\beta_1 x_{ij}\right)\right)'\right] \\ &= E\left[\left(u_{0j}+u_{1j}x_{ij}+e_{0ij}\right)\left(u_{0j}+u_{1j}x_{ij}+e_{0ij}\right)'\right] \end{aligned}$$

La matriz de varianza-covarianza total (V) toma la siguiente forma:

$$V = E\left[\begin{pmatrix} u_{01}+u_{11}x_{11}+e_{011} \\ u_{01}+u_{11}x_{21}+e_{021} \\ \vdots \\ u_{01}+u_{11}x_{k_1 1}+e_{0k_1 1} \\ \vdots \\ u_{0n}+u_{1n}x_{1n}+e_{01n} \\ u_{0n}+u_{1n}x_{2n}+e_{02n} \\ \vdots \\ u_{0n}+u_{1n}x_{k_n n}+e_{0k_n n} \end{pmatrix}\left(u_{01}+u_{11}x_{11}+e_{011},...,u_{01}+u_{11}x_{k_1 1}+e_{0k_1 1};...;u_{0n}+u_{1n}x_{1n}+e_{01n},...,u_{0n}+u_{1n}x_{k_n n}+e_{0k_n n}\right)\right]$$

Nuevamente, con fines ilustrativos, consideremos únicamente 2 estudiantes en $j=1,2,...,n$ escuelas, así:

$$V = E\left[\begin{pmatrix} u_{0j}+u_{1j}x_{1j}+e_{01j} \\ u_{0j}+u_{1j}x_{2j}+e_{02j} \end{pmatrix}\left(u_{0j}+u_{1j}x_{1j}+e_{01j};\ u_{0j}+u_{1j}x_{2j}+e_{02j}\right)\right]$$

$$V = E\begin{bmatrix} \left(u_{0j}+u_{1j}x_{1j}+e_{01j}\right)^2 & \left(u_{0j}+u_{1j}x_{1j}+e_{01j}\right)\left(u_{0j}+u_{1j}x_{2j}+e_{02j}\right) \\ \left(u_{0j}+u_{1j}x_{1j}+e_{01j}\right)\left(u_{0j}+u_{1j}x_{2j}+e_{02j}\right) & \left(u_{0j}+u_{1j}x_{2j}+e_{02j}\right)^2 \end{bmatrix}$$

$$V = \begin{bmatrix} E\left[\left(u_{0j}+u_{1j}x_{1j}+e_{01j}\right)^2\right] & E\left[\left(u_{0j}+u_{1j}x_{1j}+e_{01j}\right)\left(u_{0j}+u_{1j}x_{2j}+e_{02j}\right)\right] \\ E\left[\left(u_{0j}+u_{1j}x_{1j}+e_{01j}\right)\left(u_{0j}+u_{1j}x_{2j}+e_{02j}\right)\right] & E\left[\left(u_{0j}+u_{1j}x_{2j}+e_{02j}\right)^2\right] \end{bmatrix} = \begin{bmatrix} A & B \\ B & C \end{bmatrix}$$

Así tenemos que

$$\begin{aligned} A &= E\left[\left(u_{0j}+u_{1j}x_{1j}+e_{01j}\right)^2\right] \\ &= E\left[\left(u_{0j}\right)^2\right]+E\left[\left(u_{1j}x_{1j}\right)^2\right]+E\left[\left(e_{01j}\right)^2\right]+2x_{1j}\operatorname{cov}\left(u_{0j};u_{1j}\right) \\ &\quad +2\operatorname{cov}\left(u_{0j};e_{01j}\right)+2x_{1j}\operatorname{cov}\left(u_{1j};e_{01j}\right) \end{aligned}$$

$$A = \sigma^2_{u0}+2\sigma_{u01}x_{1j}+\sigma^2_{u1}x^2_{1j}+\sigma^2_{e0}$$

$$\begin{aligned} B &= E\left[\left(u_{0j}+u_{1j}x_{1j}+e_{01j}\right)\left(u_{0j}+u_{1j}x_{2j}+e_{02j}\right)\right] \\ &= E\left[\left(u_{0j}\right)^2+u_{0j}u_{1j}x_{2j}+u_{0j}e_{02j}+u_{1j}u_{0j}x_{1j}+\left(u_{1j}\right)^2x_{1j}x_{2j}+u_{1j}e_{02j}x_{1j}+\right. \\ &\quad \left.+e_{01j}u_{0j}+u_{1j}e_{01j}x_{2j}+e_{01j}e_{02j}\right] \\ &= E\left[\left(u_{0j}\right)^2\right]+E\left[u_{0j}u_{1j}x_{2j}\right]+E\left[u_{0j}e_{02j}\right]+E\left[u_{1j}u_{0j}x_{1j}\right]+E\left[\left(u_{1j}\right)^2x_{1j}x_{2j}\right]+ \\ &\quad +E\left[u_{1j}e_{02j}x_{1j}\right]+E\left[e_{01j}u_{0j}\right]+E\left[u_{1j}e_{01j}x_{2j}\right]+E\left[e_{01j}e_{02j}\right] \\ &= \sigma^2_{u0}+\sigma_{u01}\left(x_{1j}+x_{2j}\right)+\sigma^2_{u1}x_{1j}x_{2j} \end{aligned}$$

$$
\begin{aligned}
C &= E\left[\left(u_{0j}+u_{1j}x_{2j}+e_{02j}\right)^2\right] \\
&= E\left[\left(u_{0j}\right)^2\right]+E\left[\left(u_{1j}x_{2j}\right)^2\right]+E\left[\left(e_{02j}\right)^2\right]+2x_{2j}\operatorname{cov}\left(u_{0j};u_{1j}\right) \\
&\quad +2\operatorname{cov}\left(u_{0j};e_{02j}\right)+2x_{2j}\operatorname{cov}\left(u_{1j};e_{02j}\right) \\
C &= \sigma^2_{u0}+2\sigma_{u01}x_{2j}+\sigma^2_{u1}x^2_{2j}+\sigma^2_{e0}
\end{aligned}
$$

Ahora

$$
V=\begin{pmatrix} A & B \\ B & C \end{pmatrix}=\begin{pmatrix} \sigma^2_{u0}+2\sigma_{u01}x_{1j}+\sigma^2_{u1}x^2_{1j}+\sigma^2_{e0} & \sigma^2_{u0}+\sigma_{u01}\left(x_{1j}+x_{2j}\right)+\sigma^2_{u1}x_{1j}x_{2j} \\ \sigma^2_{u0}+\sigma_{u01}\left(x_{1j}+x_{2j}\right)+\sigma^2_{u1}x_{1j}x_{2j} & \sigma^2_{u0}+2\sigma_{u01}x_{2j}+\sigma^2_{u1}x^2_{2j}+\sigma^2_{e0} \end{pmatrix}
$$

Otra forma de escribir V es la siguiente:

$$
V=\begin{pmatrix} A & B \\ B & C \end{pmatrix}=X_j\Omega_2X_j^T+\begin{pmatrix} \Omega_1 & 0 \\ 0 & \Omega_1 \end{pmatrix}
$$

$$
X_j=\begin{pmatrix} 1 & x_{1j} \\ 1 & x_{2j} \end{pmatrix},\quad \Omega_2=\begin{pmatrix} \sigma^2_{u0} & \sigma_{u01} \\ \sigma_{u01} & \sigma^2_{u1} \end{pmatrix},\quad \Omega_1=\sigma^2_{e0} \tag{2.15}
$$

También vemos aquí el patrón general para construir la matriz de la covarianza de la respuesta que generaliza a los modelos de un orden más alto y, que veremos más adelante. En el siguiente apartado se precisa con más detalles los procedimientos para obtener estimaciones en los parámetros en un modelo multinivel general de dos niveles.

2.5 Estructura general y estimación para un modelo multinivel de dos niveles.

Ahora daremos una descripción de la iteración generalizada del procedimiento de estimación por mínimos cuadrados (IGLS). Para el proceso de estimación se utiliza la matriz de varianza-covarianza total (V_2) que se mostró en (2.15). Procedemos a encontrar V_2 para el caso de dos estudiantes en cada $j=1,2,\ldots,n$ escuelas. Considerando un nuevo coeficiente $\beta_{2j} = \beta_2 + u_{2j}$ al modelo (2.2), de tal modo que el nuevo modelo para este caso particular resulta ser:

$$y_{ij} = \beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij} + \left(u_{0j} + u_{1j} x_{1ij} + u_{2j} x_{2ij} + e_{0ij}\right) \tag{2.16}$$

Donde, para el nivel dos se han considerado tres coeficientes aleatorios y para el nivel 1 solamente una variable aleatoria.

Partiendo de la forma que toma la matriz V_2 para un modelo de 2 niveles donde incluye coeficientes aleatorios, y asumiendo los supuestos de dicho modelo, tenemos que el valor esperado de la expresión (2.16) es:

$$\hat{y}_j \cong E\left(y_{ij} \middle| \beta_0, \beta_1, \beta_2, x_{1ij}, x_{2ij}\right) = \beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij} \tag{2.17}$$

Sustituyendo (2.16) y (2.17) en (2.14) y simplificando, se tiene que:

$$V_2 = E\left[\left(u_{0j} + u_{1j} x_{1ij} + u_{2j} x_{2ij} + e_{0ij}\right)\left(u_{0j} + u_{1j} x_{1ij} + u_{2j} x_{2ij} + e_{0ij}\right)'\right] \tag{2.18}$$

Comenzaremos la prueba según el caso mencionado anteriormente, así:

$$V_2 = E\left[\begin{pmatrix} u_{0j} + u_{1j} x_{11j} + u_{2j} x_{21j} + e_{01j} \\ u_{0j} + u_{1j} x_{12j} + u_{2j} x_{22j} + e_{02j} \end{pmatrix}\left(u_{0j} + u_{1j} x_{11j} + u_{2j} x_{21j} + e_{01j},\ u_{0j} + u_{1j} x_{12j} + u_{2j} x_{22j} + e_{02j}\right)\right]$$

$$V_2 = E\begin{bmatrix} \left(u_{0j} + u_{1j} x_{11j} + u_{2j} x_{21j} + e_{01j}\right)^2 & \left(u_{0j} + u_{1j} x_{11j} + u_{2j} x_{21j} + e_{01j}\right)\left(u_{0j} + u_{1j} x_{12j} + u_{2j} x_{22j} + e_{02j}\right) \\ \left(u_{0j} + u_{1j} x_{12j} + u_{2j} x_{22j} + e_{02j}\right)\left(u_{0j} + u_{1j} x_{11j} + u_{2j} x_{21j} + e_{01j}\right) & \left(u_{0j} + u_{1j} x_{12j} + u_{2j} x_{22j} + e_{02j}\right)^2 \end{bmatrix}$$

Nuevamente, con fines ilustrativos, consideremos únicamente 2 estudiantes en $j = 1, 2, \ldots, n$ escuelas, así:

$$V_2=\begin{bmatrix} E\left(\left(u_{0j}+u_{1j}x_{11j}+u_{2j}x_{21j}+e_{01j}\right)^2\right) & E\left(\left(u_{0j}+u_{1j}x_{11j}+u_{2j}x_{21j}+e_{01j}\right)\left(u_{0j}+u_{1j}x_{12j}+u_{2j}x_{22j}+e_{02j}\right)\right) \\ E\left(\left(u_{0j}+u_{1j}x_{12j}+u_{2j}x_{22j}+e_{02j}\right)\left(u_{0j}+u_{1j}x_{11j}+u_{2j}x_{21j}+e_{01j}\right)\right) & E\left(\left(u_{0j}+u_{1j}x_{12j}+u_{2j}x_{22j}+e_{02j}\right)^2\right) \end{bmatrix}$$

Calculemos los valores esperados de cada elemento de la matriz, así:

$$\begin{aligned} E\left(\left(u_{0j}+u_{1j}x_{11j}+u_{2j}x_{21j}+e_{01j}\right)^2\right) &= E\left(\left(u_{0j}\right)^2\right)+E\left(\left(u_{1j}\right)^2\left(x_{11j}\right)^2\right)+E\left(\left(u_{2j}\right)^2\left(x_{21j}\right)^2\right)+E\left(\left(e_{01j}\right)^2\right)+ \\ &+2x_{11j}\operatorname{cov}\left(u_{0j},u_{1j}\right)+2x_{21j}\operatorname{cov}\left(u_{0j},u_{2j}\right)+2\operatorname{cov}\left(u_{0j},e_{01j}\right)+2x_{11j}x_{21j}\operatorname{cov}\left(u_{1j},u_{2j}\right)+ \\ &+2x_{11j}\operatorname{cov}\left(u_{1j},e_{01j}\right)+2x_{21j}\operatorname{cov}\left(u_{2j},e_{01j}\right) \end{aligned}$$

$$A=\sigma_{u0}^2+\sigma_{u1}^2x_{11j}^2+\sigma_{u2}^2x_{21j}^2+2\sigma_{u01}x_{11j}+2\sigma_{u02}x_{21j}+2\sigma_{u12}x_{11j}x_{21j}+\sigma_{e01}^2$$

$$\begin{aligned} E\left(\left(u_{0j}+u_{1j}x_{12j}+u_{2j}x_{22j}+e_{02j}\right)^2\right) &= E\left(\left(u_{0j}\right)^2\right)+E\left(\left(u_{1j}\right)^2\left(x_{12j}\right)^2\right)+E\left(\left(u_{2j}\right)^2\left(x_{22j}\right)^2\right)+E\left(\left(e_{02j}\right)^2\right)+ \\ &+2x_{12j}\operatorname{cov}\left(u_{0j},u_{1j}\right)+2x_{22j}\operatorname{cov}\left(u_{0j},u_{2j}\right)+2\operatorname{cov}\left(u_{0j},e_{02j}\right)+2x_{12j}x_{22j}\operatorname{cov}\left(u_{1j},u_{2j}\right)+ \\ &+2x_{12j}\operatorname{cov}\left(u_{1j},e_{02j}\right)+2x_{22j}\operatorname{cov}\left(u_{2j},e_{01j}\right) \end{aligned}$$

$$C=\sigma_{u0}^2+\sigma_{u1}^2x_{12j}^2+\sigma_{u2}^2x_{22j}^2+2\sigma_{u01}x_{12j}+2\sigma_{u02}x_{22j}+2\sigma_{u12}x_{12j}x_{22j}+\sigma_{e02}^2$$

$$\begin{aligned} E\left(\left(u_{0j}+u_{1j}x_{11j}+u_{2j}x_{21j}+e_{01j}\right)\left(u_{0j}+u_{1j}x_{12j}+u_{2j}x_{22j}+e_{02j}\right)\right) &= E\left(\left(u_{0j}\right)\left(u_{0j}+u_{1j}x_{12j}+u_{2j}x_{22j}+e_{02j}\right)\right)+ \\ &+E\left(\left(u_{1j}x_{11j}\right)\left(u_{0j}+u_{1j}x_{12j}+u_{2j}x_{22j}+e_{02j}\right)\right)+ \\ &+E\left(\left(u_{2j}x_{22j}\right)\left(u_{0j}+u_{1j}x_{12j}+u_{2j}x_{22j}+e_{02j}\right)\right)+ \\ &+E\left(\left(e_{02j}\right)\left(u_{0j}+u_{1j}x_{12j}+u_{2j}x_{22j}+e_{02j}\right)\right) \\ &=E\left(\left(u_{0j}\right)^2\right)+x_{12j}\operatorname{cov}\left(u_{0j},u_{1j}\right)+x_{22j}\operatorname{cov}\left(u_{0j},u_{2j}\right)+ \\ &+x_{11j}\operatorname{cov}\left(u_{0j},u_{1j}\right)+E\left(\left(u_{1j}\right)^2\right)x_{11j}x_{12j}+x_{11j}x_{22j}\operatorname{cov}\left(u_{1j},u_{2j}\right)+ \\ &+x_{21j}\operatorname{cov}\left(u_{0j},u_{2j}\right)+x_{21j}x_{12j}\operatorname{cov}\left(u_{1j},u_{2j}\right)+E\left(\left(u_{2j}\right)^2\right)x_{21j}x_{22j} \end{aligned}$$

$$B=\sigma_{u0}^2+\sigma_{u1}^2x_{11j}x_{12j}+\sigma_{u2}^2x_{21j}x_{22j}+\sigma_{u01}\left(x_{11j}+x_{12j}\right)+\sigma_{u02}\left(x_{21j}+x_{22j}\right)+\sigma_{u12}\left(x_{21j}x_{12j}+x_{11j}x_{22j}\right)$$

Ahora

$$V_2=\begin{pmatrix} A & B \\ B & C \end{pmatrix}$$

$$V_2 = \begin{pmatrix} \sigma_{u0}^2 + \sigma_{u1}^2 x_{11j}^2 + \sigma_{u2}^2 x_{21j}^2 + 2\sigma_{u01} x_{11j} + 2\sigma_{u02} x_{21j} + 2\sigma_{u12} x_{11j} x_{21j} + \sigma_{e01}^2 & \begin{array}{l} \sigma_{u0}^2 + \sigma_{u1}^2 x_{11j} x_{12j} + \sigma_{u2}^2 x_{21j} x_{22j} + \sigma_{u01}\left(x_{11j} + x_{12j}\right) + \sigma_{u02}\left(x_{21j} + x_{22j}\right) + \\ + \sigma_{u12}\left(x_{21j} x_{12j} + x_{11j} x_{22j}\right) \end{array} \\ \begin{array}{l} \sigma_{u0}^2 + \sigma_{u1}^2 x_{11j} x_{12j} + \sigma_{u2}^2 x_{21j} x_{22j} + \sigma_{u01}\left(x_{11j} + x_{12j}\right) + \sigma_{u02}\left(x_{21j} + x_{22j}\right) + \\ + \sigma_{u12}\left(x_{21j} x_{12j} + x_{11j} x_{22j}\right) \end{array} & \sigma_{u0}^2 + \sigma_{u1}^2 x_{12j}^2 + \sigma_{u2}^2 x_{22j}^2 + 2\sigma_{u01} x_{12j} + 2\sigma_{u02} x_{22j} + 2\sigma_{u12} x_{12j} x_{22j} + \sigma_{e02}^2 \end{pmatrix}$$

A continuación generalizamos V_2 para tantas variables explicativas tengamos en cada nivel. Para ello, considerando la siguiente estructura general de un modelo de dos niveles, así:

$$Y = X\beta + E \tag{2.19}$$

Donde

E es una matriz diseño que contiene la suma de los errores de cada nivel, así:

$$E = E_1 + E_2 = \{e_{ij}\}, \quad e_{ij} = e_{ij}^{(1)} + e_j^{(2)}, \quad e_{ij}^{(1)} = \sum_{h=0}^{q_1} z_{hij}^{(1)} e_{hij}^{(1)}, \quad e_j^{(2)} = \sum_{h=0}^{q_2} z_{hij}^{(2)} e_{hj}^{(2)}$$

Se escribe simplemente $e_{ij}^{(1)}$ como e_{ij} y $e_j^{(2)}$ como u_j, quedando la expresión (2.19) de la siguiente manera:

$$Y = X\beta + Z^{(2)}u + Z^{(1)}e \tag{2.20}$$

Donde

Y es el vector columna de las observaciones de la i-ésima unidad del nivel 1 en la j-ésima unidad del nivel dos.

$X\beta$ es la parte fija del modelo multinivel general de dos niveles.

$Z^{(2)} = \left(z_{0ij} \ z_{1ij} \cdots z_{q_2 ij}\right)$ es la matriz diseño de las $h = 0,1,...,q_2$ variables aleatorias para el nivel dos.

$Z^{(1)} = \left(z_{0ij} \quad z_{1ij} \cdots z_{q_1 ij} \right)$ es la matriz diseño de las $h = 0,1,\ldots,q_1$ variables aleatorias para el nivel uno.

$u = \left(u_{0j}, u_{1j}, \ldots, u_{q_2 j} \right)'$ es el vector columna con $h = 0,1,\ldots,q_2$ parámetros aleatorios para el nivel dos.

$e = \left(e_{0ij}, e_{1ij}, \ldots, e_{q_1 ij} \right)'$ es el vector columna con $h = 0,1,\ldots,q_1$ parámetros aleatorios para el nivel 1.

Normalmente $z_{0ij} = x_{0ij} = 1$

Así, tenemos que la expresión (2.20) la parte aleatoria del modelo $Z^{(2)}u + Z^{(1)}e$ se supone que las variables aleatorias tanto para el nivel 1 como el nivel 2 se distribuyen de una forma normal con media cero y varianza constante. El valor esperado de la estructura general de un modelo multinivel general de 2 niveles es:

$$E\left(y_{ij} \middle| \beta_0, \beta_1, \beta_2, x_{1ij}, x_{2ij}, \ldots, x_{pij} \right) = \beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij} + \ldots + \beta_p x_{pij} = X\beta \qquad (2.21)$$

En consecuencia, la matriz de varianza-covarianza total V para un modelo multinivel general de 2 niveles, resulta ser:

$$V_2 = E\left[\left(Y - E(Y) \right) \left(Y - E(Y) \right)' \right] \qquad (2.22)$$

Sustituyendo (2.20) y (2.21) en (2.22), se tiene que:

$$
\begin{aligned}
V_2 &= E\left[\left(Z^{(2)}u + Z^{(1)}e\right)\left(Z^{(2)}u + Z^{(1)}e\right)'\right] \\
&= E\left[(E_1 + E_2)(E_1 + E_2)'\right] \\
&= E\left[(E_1 + E_2)\left(E_1' + E_2'\right)\right] \\
&= E\left[\left(E_1E_1'\right) + \left(E_1E_2'\right) + \left(E_2E_1'\right) + \left(E_2E_2'\right)\right] \\
&= E\left[\left(E_1E_1'\right)\right] + E\left[\left(\cancel{E_1E_2'}\right)\right] + E\left[\left(\cancel{E_2E_1'}\right)\right] + E\left[\left(E_2E_2'\right)\right] \\
&= E\left[\left(E_1E_1'\right)\right] + E\left[\left(E_2E_2'\right)\right]
\end{aligned}
$$

Donde $E\left[\left(E_1E_1'\right)\right]$ es la varianza para las variables aleatorias del nivel 1 y $E\left[\left(E_2E_2'\right)\right]$ es la varianza para las variables aleatorias del nivel 2.

Encontremos estas varianzas, calculando sus valores esperados, así:

$$E\left[\left(E_1E_1'\right)\right] = E\left[\left(Z^{(1)}e\right)\left(Z^{(1)}e\right)'\right] = E\left[\left(Z^{(1)}e\right)\left(e'\,Z^{(1)'}\right)\right] = E\left[Z^{(1)}e\,e'\,Z^{(1)'}\right]$$

$$E\left[Z^{(1)}e\,e'\,Z^{(1)'}\right] = Z^{(1)}E\left[e\,e'\right]Z^{(1)'} = V_{2(1)j}$$

Donde

$$E\left[e\,e'\right] = E\left[\begin{pmatrix} e_{0ij} \\ e_{1ij} \\ \vdots \\ e_{q_1ij} \end{pmatrix}\left(e_{0ij}, e_{1ij}, \ldots, e_{q_1ij}\right)\right]_{q_1 x q_1}$$

Desarrollemos introduciendo el valor esperado y considerando los supuestos del modelo multinivel general de dos niveles, así:

$$E\left[e\,e'\right]=E\left[\begin{pmatrix}\sigma^2_{e011} & 0 & \dots & 0 & \dots & 0 & 0 & \dots & 0\\ 0 & \sigma^2_{e021} & \dots & 0 & \dots & 0 & 0 & \dots & 0\\ \vdots & \vdots & \ddots & \vdots & \dots & \vdots & \vdots & & \vdots\\ 0 & 0 & \dots & \sigma^2_{e0k_11} & \dots & 0 & 0 & \dots & 0\\ \vdots & \vdots & & \vdots & \dots & \vdots & \vdots & & \vdots\\ 0 & 0 & \dots & 0 & \dots & \sigma^2_{eq_11n} & 0 & \dots & 0\\ 0 & 0 & \dots & 0 & \dots & 0 & \sigma^2_{eq_12n} & \dots & 0\\ \vdots & \vdots & \dots & \vdots & \dots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots. & 0 & \dots & 0 & 0 & \dots & \sigma^2_{eq_1k_nn}\end{pmatrix}\right]_{(k_1+\dots+k_n)x(k_1+\dots+k_n)}$$

$$E\left[e\,e'\right]=\Omega_e=\operatorname{cov}\left(e_h^{(1)}\right)$$

La matriz Ω_e es una matriz bloque-diagonal de orden q_1xq_1 bloques, donde, cada bloque representa la varianza del bloque h en el i-ésimo alumno de la j-ésima escuela.

$$Z^{(1)}=\begin{pmatrix}1 & x_{111} & x_{211} & \cdots & x_{q_111}\\ 1 & x_{121} & x_{221} & \cdots & x_{q_121}\\ \vdots & \vdots & \vdots & & \vdots\\ 1 & x_{1k_11} & x_{2k_11} & \cdots & x_{q_1k_11}\\ \vdots & \vdots & \vdots & & \vdots\\ 1 & x_{11n} & x_{21n} & \cdots & x_{q_11n}\\ 1 & x_{12n} & x_{22n} & \cdots & x_{q_12n}\\ \vdots & \vdots & \vdots & & \vdots\\ 1 & x_{1k_nn} & x_{2k_nn} & \cdots & x_{q_1k_nn}\end{pmatrix}_{(k_1+k_2+\dots+k_n)\times(q_1)}$$

$$Z^{(1)'} = \begin{pmatrix} 1 & 1 & \cdots & 1 & \cdots & 1 & 1 & \cdots & 1 \\ x_{111} & x_{121} & \cdots & x_{1k_11} & \cdots & x_{11n} & x_{12n} & \cdots & x_{1k_nn} \\ x_{211} & x_{221} & \cdots & x_{2k_11} & \cdots & x_{21n} & x_{22n} & \cdots & x_{2k_nn} \\ \vdots & \vdots & & \vdots & & \vdots & \vdots & & \vdots \\ x_{q_111} & x_{q_121} & \cdots & x_{q_1k_11} & \cdots & x_{q_11n} & x_{q_12n} & \cdots & x_{q_1k_nn} \end{pmatrix}_{(q_1)\times(k_1+k_2+\dots+k_n)}$$

Al realizar el producto de las matrices $Z^{(1)'}$, Ω_e y $Z^{(1)}$ nos da como resultado una matriz de orden $q_1 \times q_1$.

Ahora, de igual manera,

$$E\left[\left(E_2 E_2'\right)\right] = E\left[\left(Z^{(2)}u\right)\left(Z^{(2)}u\right)'\right] = E\left[\left(Z^{(2)}u\right)\left(u' Z^{(2)'}\right)\right] = E\left[Z^{(2)}u\, u' Z^{(2)'}\right]$$

$$E\left[Z^{(2)}u\, u' Z^{(2)'}\right] = Z^{(2)}E\left[u\, u'\right]Z^{(2)'} = V_{2(2)j}$$

Donde

$$E\left(u\, u'\right) = E\left[\begin{pmatrix} u_{0j} \\ u_{1j} \\ \vdots \\ u_{q_1 j} \end{pmatrix}\left(u_{0j}, u_{1j}, \dots, u_{q_1 j}\right)\right]$$

Así, tenemos que:

$$E\left(u\, u'\right) = E\left[\begin{pmatrix} u_{0j}^2 & u_{0j}u_{1j} & \cdots & u_{0j}u_{q_2 j} \\ u_{0j}u_{1j} & u_{1j}^2 & \cdots & u_{1j}u_{q_2 j} \\ \vdots & \vdots & \ddots & \vdots \\ u_{0j}u_{q_2 j} & u_{1j}u_{q_2 j} & \cdots & u_{q_2 j}^2 \end{pmatrix}\right]$$

Luego se tiene

$$E\left(u\,u'\right)=\begin{pmatrix}\sigma_{u0}^2 & \sigma_{u01} & \cdots & \sigma_{u0q_2}\\ \sigma_{u01} & \sigma_{u1}^2 & \cdots & \sigma_{u1q_2}\\ \vdots & & & \\ \sigma_{u0q_2} & \sigma_{u1q_2} & \cdots & \sigma_{uq_2}^2\end{pmatrix}_{q_2\times q_2}=\Omega_u=\operatorname{cov}\left(e_h^{(2)}\right)$$

Además,

$$Z^{(2)}=\begin{pmatrix}1 & x_{111} & x_{211} & \cdots & x_{q_211}\\ 1 & x_{121} & x_{221} & \cdots & x_{q_221}\\ \vdots & \vdots & \vdots & & \vdots\\ 1 & x_{1k_11} & x_{2k_11} & \cdots & x_{q_2k_11}\\ \vdots & \vdots & \vdots & & \vdots\\ 1 & x_{11n} & x_{21n} & \cdots & x_{q_21n}\\ 1 & x_{12n} & x_{22n} & \cdots & x_{q_22n}\\ \vdots & \vdots & \vdots & & \vdots\\ 1 & x_{1k_nn} & x_{2k_nn} & \cdots & x_{q_2k_nn}\end{pmatrix}_{(k_1+k_2+\ldots+k_n)\times(q_2)}$$

y

$$Z^{(2)'}=\begin{pmatrix}1 & 1 & \cdots & 1 & \cdots & 1 & 1 & \cdots & 1\\ x_{111} & x_{121} & \cdots & x_{1k_11} & \cdots & x_{11n} & x_{12n} & \cdots & x_{1k_nn}\\ x_{211} & x_{221} & \cdots & x_{2k_11} & \cdots & x_{21n} & x_{22n} & \cdots & x_{2k_nn}\\ \vdots & \vdots & & \vdots & & \vdots & \vdots & & \vdots\\ x_{q_211} & x_{q_221} & \cdots & x_{q_2k_11} & \cdots & x_{q_21n} & x_{q_22n} & \cdots & x_{q_2k_nn}\end{pmatrix}_{(q_2)\times(k_1+k_2+\ldots+k_n)}$$

Al realizar el producto de las matrices $Z^{(2)}$, Ω_u y $Z^{(2)'}$ nos da como resultado $V_{2(2)}$ una matriz bloque-diagonal con j-ésimos bloques.

El j-ésimo bloque de V_2, está dado por:

$$V_{2j} = V_{2(1)j} + V_{2(2)j} \tag{2.23}$$

Con la expresión (2.23) representemos el caso particular que desarrollamos anteriormente de dos estudiantes en cada $j = 1, 2, \ldots, n$ escuelas, de la siguiente manera:

$$\begin{aligned} V_{2j} &= E\left[\left(Z^{(2)}u + Z^{(1)}e\right)\left(Z^{(2)}u + Z^{(1)}e\right)'\right] \\ &= V_{2(1)j} + V_{2(2)j} \end{aligned}$$

Donde

$$V_{2(2)j} = Z^{(2)}E\left[u\,u'\right]Z^{(2)'}; \qquad V_{2(1)j} = Z^{(1)}E\left[u\,u'\right]Z^{(1)'} \tag{2.24}$$

Recordemos además que en este caso particular, se consideran dos variables aleatorias para el nivel dos y una única variable aleatoria para el nivel 1, así:

$$V_{2(2)j} = \left(\underset{(j)\times(2=q_2)}{\begin{pmatrix} 1 & x_{111} & x_{21j} \\ 1 & x_{121} & x_{22j} \end{pmatrix}} \underset{q_2\times q_2}{\begin{pmatrix} \sigma_{u0}^2 & \sigma_{u01} & \sigma_{u02} \\ \sigma_{u01} & \sigma_{u1}^2 & \sigma_{u12} \\ \sigma_{u02} & \sigma_{u12} & \sigma_{u2}^2 \end{pmatrix}} \underset{q_2\times j}{\begin{pmatrix} 1 & 1 \\ x_{11j} & x_{12j} \\ x_{21j} & x_{22j} \end{pmatrix}} \right)_{jxj} \tag{2.25}$$

Ahora, encontremos $V_{2(1)j}$. Según la expresión (2.24) se tendrán j-bloques, no considerando además variables aleatorias (h=0) en el modelo, así:

$$V_{2(1)j} = \begin{pmatrix} \sigma^2_{e01j} & 0 \\ 0 & \sigma^2_{e02j} \end{pmatrix}_{jxj} \quad (2.26)$$

Sumando las matrices (2.24) y (2.25) se tiene que,

$$V_{2j} = V_{2(1)j} + V_{2(2)j}$$

$$V_{2j} = \left(\underset{(j)\times(2=q_2)}{\begin{pmatrix} 1 & x_{111} & x_{21j} \\ 1 & x_{121} & x_{22j} \end{pmatrix}} \underset{q_2\times q_2}{\begin{pmatrix} \sigma^2_{u0} & \sigma_{u01} & \sigma_{u02} \\ \sigma_{u01} & \sigma^2_{u1} & \sigma_{u12} \\ \sigma_{u02} & \sigma_{u12} & \sigma^2_{u2} \end{pmatrix}} \underset{q_2\times j}{\begin{pmatrix} 1 & 1 \\ x_{11j} & x_{12j} \\ x_{21j} & x_{22j} \end{pmatrix}} \right)_{jxj} + \begin{pmatrix} \sigma^2_{e01j} & 0 \\ 0 & \sigma^2_{e02j} \end{pmatrix}_{jxj}$$

Por tanto, la matriz de varianza-covarianza total (V) para dos estudiantes y $j = 1, 2, \ldots, n$ escuelas es similar a la expresión obtenida a partir de (2.18).

Otra forma de escribir $V_{2(1)j}$ es expresándolo como la suma directa de esas matrices, dado que es una matriz diagonal con ij-ésimo elementos. Lo cual esta en términos de submatrices o bloques, así:

$$V_{2j} = \otimes_j \sigma^2_{eij} + V_{2(2)j} \quad (2.27)$$

donde $\otimes$ es el operador de Kronecker, el cual se aplica de la siguiente forma:

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

Ya que conocemos V_2 por conveniencia omitiremos el subíndice, podemos estimar los coeficientes fijos del modelo multinivel general de dos niveles, mediante el procedimiento generalizado de mínimos cuadrados, de la manera siguiente:

$$\hat{\beta} = \left(X' V^{-1} X \right)^{-1} X' V^{-1} Y \tag{2.28}$$

con matriz de la covariación $\left(X' V^{-1} X \right)^{-1}$

En este caso, consideremos:

$$X = \begin{pmatrix} 1 & x_{11} \\ 1 & x_{21} \\ \vdots & \vdots \\ 1 & x_{n_m m} \end{pmatrix} \qquad Y = \begin{pmatrix} y_{11} \\ y_{21} \\ \vdots \\ y_{n_m m} \end{pmatrix} \tag{2.29}$$

con m unidades en el nivel 2 y n_j unidades en el nivel 1 en la j-ésima unidad del nivel 2. Cuando los residuales tienen una distribución normal (2.28) también se estima mediante máxima verosimilitud.

Para conocer β formamos primero

$$Y^{*} = \tilde{Y} \tilde{Y}' \tag{2.30}$$

Donde

$$\tilde{Y} = Y - X \beta = E_1 + E_2$$

y tenemos además que,

$$E\left(Y^{*}\right)=V \tag{2.31}$$

Si formamos la matriz de producto-cruzado $\tilde{Y}\,\tilde{Y}'$ vemos que el valor esperado de este es simplemente *V*. Podemos volver a arreglar esta matriz de producto cruzado como un vector, reuniendo las columnas uno encima del otro, que se escribe de la siguiente manera:

$$Y^{**}=vec\left(\tilde{Y}\,\tilde{Y}'\right)=vec\left(\left(Y-X\beta\right)\left(Y-X\beta\right)'\right) \tag{2.32}$$

La notación vec() se aplica de la siguiente forma:

Por ejemplo:

Si M es una matriz simétrica de orden 2x2, entonces:

$\mathrm{M}=\begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}$ y el vector operador de M es,

$vec\left(\mathrm{M}\right)=\begin{pmatrix} a_{11} \\ a_{21} \\ a_{12} \\ a_{22} \end{pmatrix}$ de orden $4x1$.

Construiremos además el vector

$$vec\left(V\right)=vec\left(E\left(Y^{*}\right)\right)$$

Para comprender mejor las expresiones anteriores, lo haremos para la estructura dada en (2.12). Como vemos, esta expresión tiene $3^2+2^2=13$ elementos. Obtengamos dicha relación para este caso, así:

$$X=\begin{pmatrix} 1 & x_{11} \\ 1 & x_{21} \\ 1 & x_{31} \\ 1 & x_{12} \\ 1 & x_{22} \end{pmatrix} \qquad Y=\begin{pmatrix} y_{11} \\ y_{21} \\ y_{31} \\ y_{12} \\ y_{22} \end{pmatrix}$$

El vector Y^* resulta ser,

$$Y^*=\tilde{Y}\,\tilde{Y}'$$

$$Y^*=\begin{pmatrix} y_{11}-\hat{y}_1 \\ y_{21}-\hat{y}_1 \\ y_{31}-\hat{y}_1 \\ y_{12}-\hat{y}_2 \\ y_{22}-\hat{y}_2 \end{pmatrix}\left(y_{11}-\hat{y}_1; y_{21}-\hat{y}_1; y_{31}-\hat{y}_1; y_{12}-\hat{y}_2; y_{22}-\hat{y}_2\right)$$

o

$$Y^*=\begin{pmatrix} \tilde{y}_{11} \\ \tilde{y}_{21} \\ \tilde{y}_{31} \\ \tilde{y}_{12} \\ \tilde{y}_{22} \end{pmatrix}\left(\tilde{y}_{11}; \tilde{y}_{21}; \tilde{y}_{31}; \tilde{y}_{12}; \tilde{y}_{22}\right)=\begin{pmatrix} \tilde{y}_{11}^2 & \tilde{y}_{11}\tilde{y}_{21} & \cdots & \tilde{y}_{11}\tilde{y}_{22} \\ \tilde{y}_{21}\tilde{y}_{11} & \tilde{y}_{21}^2 & \cdots & \tilde{y}_{21}\tilde{y}_{22} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{y}_{22}\tilde{y}_{11} & \tilde{y}_{22}\tilde{y}_{21} & \cdots & \tilde{y}_{22}^2 \end{pmatrix}$$

Luego tenemos que el valor esperado de Y^{*} es la matriz de varianza-covarianza. Ahora encontremos el vector operador $\tilde{Y}\,\tilde{Y}'$ como el vector que reúne todas las columnas uno sobre otra, así:

$$Y^{**} = vec\left(\begin{pmatrix}\tilde{y}_{11}\\ \tilde{y}_{21}\\ \tilde{y}_{31}\\ \tilde{y}_{12}\\ \tilde{y}_{22}\end{pmatrix}\left(\tilde{y}_{11};\tilde{y}_{21};\tilde{y}_{31};\tilde{y}_{12};\tilde{y}_{22}\right)\right) = vec\begin{pmatrix}\tilde{y}_{11}^{2} & \tilde{y}_{11}\tilde{y}_{21} & \cdots & \tilde{y}_{11}\tilde{y}_{22}\\ \tilde{y}_{21}\tilde{y}_{11} & \tilde{y}_{21}^{2} & \cdots & \tilde{y}_{21}\tilde{y}_{22}\\ \vdots & \vdots & \ddots & \vdots\\ \tilde{y}_{22}\tilde{y}_{11} & \tilde{y}_{22}\tilde{y}_{21} & \cdots & \tilde{y}_{22}^{2}\end{pmatrix}$$

Que también (2.32) se puede expresar como el producto Kronecker, así:

$$\tilde{Y} \otimes \tilde{Y} = \begin{pmatrix}\tilde{y}_{11}\tilde{Y}\\ \tilde{y}_{21}\tilde{Y}\\ \vdots\\ \tilde{y}_{22}\tilde{Y}\end{pmatrix} = \left(\begin{matrix}\begin{pmatrix}\tilde{y}_{11}^{2}\\ \tilde{y}_{21}\tilde{y}_{11}\\ \tilde{y}_{31}\tilde{y}_{11}\\ \tilde{y}_{21}\tilde{y}_{11}\\ \tilde{y}_{22}\tilde{y}_{11}\end{pmatrix}\\ \vdots\\ \begin{pmatrix}\tilde{y}_{11}\tilde{y}_{22}\\ \tilde{y}_{21}\tilde{y}_{22}\\ \tilde{y}_{31}\tilde{y}_{22}\\ \tilde{y}_{12}\tilde{y}_{22}\\ \tilde{y}_{22}^{2}\end{pmatrix}\end{matrix}\right) = Y^{**} = vec\left(\tilde{Y}\,\tilde{Y}'\right)$$

Así, tenemos que,

$$E\left(Y^{**}\right) = E\left(vec\left(\begin{pmatrix}\tilde{y}_{11}\\ \tilde{y}_{21}\\ \tilde{y}_{31}\\ \tilde{y}_{12}\\ \tilde{y}_{22}\end{pmatrix}\left(\tilde{y}_{11};\tilde{y}_{21};\tilde{y}_{31};\tilde{y}_{12};\tilde{y}_{22}\right)\right)\right) = E\begin{pmatrix}\tilde{y}_{11}^{2}\\ \tilde{y}_{21}\tilde{y}_{11}\\ \vdots\\ \tilde{y}_{22}^{2}\end{pmatrix}$$

Desarrollando lo anterior tenemos que,

$$E\left(Y^{**}\right)=\begin{pmatrix}\sigma_{u0}^{2}+\sigma_{e0}^{2}\\ \sigma_{u0}^{2}\\ \vdots\\ \sigma_{u0}^{2}\\ \sigma_{u0}^{2}+\sigma_{e0}^{2}\end{pmatrix} \tag{2.33}$$

Así tenemos que la relación (2.33) se puede expresar como el modelo lineal siguiente:

$$Y^{**}=E\left(Y^{**}\right)+R$$

En consecuencia,

$$\begin{pmatrix}\tilde{y}_{11}^{2}\\ \tilde{y}_{21}\tilde{y}_{11}\\ \vdots\\ \tilde{y}_{22}^{2}\end{pmatrix}=\begin{pmatrix}\sigma_{u0}^{2}+\sigma_{e0}^{2}\\ \sigma_{u0}^{2}\\ \vdots\\ \sigma_{u0}^{2}+\sigma_{e0}^{2}\end{pmatrix}+R=\sigma_{u0}^{2}\begin{pmatrix}1\\ 1\\ \vdots\\ 1\end{pmatrix}+\sigma_{e0}^{2}\begin{pmatrix}1\\ 0\\ \vdots\\ 1\end{pmatrix}+R \tag{2.34}$$

donde R es un vector residual. El lado de la mano izquierda de (2.34) es el vector de la respuesta en el modelo lineal y el lado derecho contiene dos variables explicativas, con los coeficientes σ_{u0}^{2} y σ_{e0}^{2} que deben ser estimados.

Podemos ahora escribir un modelo lineal que implica los parámetros aleatorios de los elementos de Ω_u, Ω_e, como sigue

$$E\left(Y^{**}\right)=Z^{*}\theta \tag{2.35}$$

Donde Z^* es la matriz de diseño para los parámetros aleatorios del modelo. Un ejemplo de tal matriz de diseño para un modelo simple de los componentes de variación se da en el apartado 2.3 de este capítulo. Ahora procederemos a estimar θ por medio del procedimiento generalizado de estimación de mínimos cuadrados, a saber:

$$\hat{\theta} = \left(Z^{*\prime} V^{*-1} Z^*\right)^{-1} Z^{*\prime} V^{*-1} Y^{**}, \qquad V^* = V \otimes V \tag{2.36}$$

Donde $\otimes$ es el operador del producto de Kronecker.

La matriz de la covariación del estimador $\hat{\theta}$ resulta ser:

$$\operatorname{cov}\left(\hat{\theta}\right) = \left(Z^{*\prime} V^{*-1} Z^*\right)^{-1} Z^{*\prime} V^{*-1} \operatorname{cov}\left(Y^{**}\right) V^{*-1} Z^* \left(Z^{*\prime} V^{*-1} Z^*\right)^{-1} \tag{2.37}$$

Como

$$Y^{**} = vec\left(\tilde{Y}\,\tilde{Y}'\right) = \tilde{Y} \otimes \tilde{Y}$$

Y usando el resultado estándar de Searle (1992) tenemos que,

$$\operatorname{cov}\left(\tilde{Y} \otimes \tilde{Y}\right) = \left(V \otimes V\right)\left(I + S_N\right) \tag{2.38}$$

donde $V \otimes V = V^*$ y S_N es la matriz de permutación del vector vec.

En notas de Goldstein y de Rasbash (1992) se tiene que,

$$V^{*-1} = \left(V \otimes V\right)^{-1} = V^{-1} \otimes V^{-1}$$

En consecuencia

$$V^{*-1}Z^{*}=\left(V^{-1}\otimes V^{-1}\right)vec\left(A\right)$$

Donde $vec\left(A\right)$ es igual al vector operador de la matriz de diseño para los parámetros aleatorios del modelo. Según propiedades de los productos de Kronecker se tiene que,

$$V^{*-1}Z^{*}=\left(V^{-1}\otimes V^{-1}\right)vec\left(A\right)=vec\left(V^{-1}AV^{-1}\right) \tag{2.39}$$

Y por ser $V^{-1}AV^{-1}$ simétrico, llegamos a obtener que,

$$S_{N}V^{*^{-1}}Z^{*}=V^{*^{-1}}Z^{*} \tag{2.40}$$

Sustituyendo (2.38), (2.39) y (2.40) en (2.37) tenemos que la covarianza de $\hat{\theta}$ es:

$$\begin{aligned}
\operatorname{cov}\left(\hat{\theta}\right)&=\left(Z^{*\prime}V^{*-1}Z^{*}\right)^{-1}Z^{*\prime}V^{*^{-1}}\left(\left(V\otimes V\right)\left(I+S_{N}\right)\right)V^{*^{-1}}Z^{*}\left(Z^{*\prime}V^{*-1}Z^{*}\right)^{-1}\\
&=\left(Z^{*\prime}V^{*-1}Z^{*}\right)^{-1}Z^{*\prime}V^{*^{-1}}\left(V^{*}I+S_{N}V^{*}\right)V^{*^{-1}}Z^{*}\left(Z^{*\prime}V^{*-1}Z^{*}\right)^{-1}\\
&=\left(Z^{*\prime}V^{*-1}Z^{*}\right)^{-1}\left(Z^{*\prime}+S_{N}Z^{*\prime}\right)V^{*^{-1}}Z^{*}\left(Z^{*\prime}V^{*-1}Z^{*}\right)^{-1}\\
&=\left(Z^{*\prime}V^{*-1}Z^{*}\right)^{-1}\left(Z^{*\prime}V^{*^{-1}}Z^{*}+S_{N}Z^{*\prime}V^{*^{-1}}Z^{*}\right)\left(Z^{*\prime}V^{*-1}Z^{*}\right)^{-1}\\
&=(I+I)\left(Z^{*\prime}V^{*-1}Z^{*}\right)^{-1}\\
&=2\left(Z^{*\prime}V^{*-1}Z^{*}\right)^{-1} \qquad (2.41)
\end{aligned}$$

Mediante la iteración generalizada de estimación de mínimos cuadrados (IGLS) se producen estimaciones polarizadas y en general esto puede ser importante en muestras pequeñas. Goldstein (1989) demuestra una modificación simple sobre la iteración general restringida de mínimos cuadrados o por máxima verosimilitud restringida (REML) de estimaciones que son insengadas. Si reescribimos $Y^{*}=\tilde{Y}\,\tilde{Y}'$ y usando las estimaciones de los parámetros fijos $\hat{\beta}$ obtenemos que,

$$E\left(Y^{*}\right)=V_{2}-X\ \mathrm{cov}\left(\hat{\beta}\right)X'=V_{2}-X\left(X'\,V_{2}^{-1}X\right)^{-1}X' \tag{2.42}$$

Tomando en cuenta la variación del estimador $\hat{\beta}$ podemos obtener una estimación insesgada de V_{2} agregando el segundo término dentro de (2.42), la matriz sombrero, de Y^{*} en cada iteración hasta poder converger. En el caso donde estamos estimando una varianza de una muestra aleatoria simple este se convierte en el procedimiento estándar para usar el divisor $n-1$ más bien que n producirá una estimación sesgada.

2.6 Estimación de los parámetros del Modelo general de tres niveles.

En el apartado anterior hemos explorado la estimación de los parámetros del modelo general de dos niveles de componentes de varianza, además, del modelo incluyendo coeficientes aleatorios. En este apartado, se explora el modelo general de tres niveles. Recordemos que hemos partido del modelo multinivel simple definido como (2.4), pero que ahora lo expresamos de la siguiente forma,

$$y_{ij}=\mathrm{X}_{(f)}\beta+\mathrm{X}_{(2)}u+e_{0ij};\quad j=1,2,\ldots,n \tag{2.43}$$

Donde

$X_{(f)}=\left(1\ \ x_{1ij}\ x_{1ij}\cdots x_{pij}\right)$ es la matriz de diseño para la parte fija del modelo.

$\beta=\begin{pmatrix}\beta_{0}\\ \beta_{1}\\ \vdots\\ \beta_{p}\end{pmatrix}$ el vector de coeficientes para la parte fija del modelo.

$X_{(2)}$ es la matriz de diseño de la parte aleatoria del modelo de dos niveles.

$u = \begin{pmatrix} u_{oj} \\ u_{1j} \\ \vdots \\ u_{q_2 j} \end{pmatrix}$ la matriz de diseño para las variables aleatorias del modelo de dos niveles.

Permitiremos así, considerar todas las suposiciones que utilizamos en el apartado 2.2. Consideramos además la situación donde una variable de respuesta "y" puede depender de un sistema p variables explicativas $x_1, \ldots, x_p$. Definiéndose así el modelo general de tres niveles como:

$$y_{ijk} = \mathrm{X}'_{(f)ijk}\beta + \mathrm{X}'_{(3)ijk}v_i + \mathrm{X}'_{(2)ijk}u_i + \mathrm{X}'_{(1)ijk}e_{ijk}, \qquad (2.44)$$

Donde

$i = 1,2,....,N$: denota las unidades del nivel 3 (ejemplo, escuelas en departamentos),
$j = 1,2,....,n_i$: denota las unidades nivel 2 (ejemplo, escuelas), y
$k = 1,2,....,n_{ij}$: denota las unidades del nivel 1 (ejemplo, alumnos).
$\mathrm{X}'_{(f)ijk}:1xs$, es una fila típica de la matriz del diseño de la parte fija del modelo, los elementos que son un subconjunto de p variables predictoras.
$\mathrm{X}'_{(3)ijk}:1xq$, es una fila típica de la matriz del diseño para la parte aleatoria en el nivel 3, los elementos que son un subconjunto de p variables predictoras.
$\mathrm{X}'_{(2)ijk}:1xm$, es una fila típica de la matriz del diseño para la parte aleatoria en el nivel 2, los elementos que son un subconjunto de p variables predictoras.
$\mathrm{X}'_{(1)ijk}:1xr$, es una fila típica de la matriz del diseño para la parte al azar en el nivel 1, los elementos que son un subconjunto de p variables predictoras.
$\beta : s x1$, es un vector de los parámetros fijos, pero desconocidos que se estimarán.

Se supone además que $v_1,\ldots,v_N$ están distribuidos independientemente e idénticamente con media 0 y matriz de varianza-covarianza constante $\Phi_{(3)}$. Se supone más a fondo que $u_{i1},\ldots,u_{in_i}$ son $i.i.d.$ con media 0 y matriz de varianza-covarianza $\Phi_{(2)}$, mientras que $e_{ij1},\ldots,e_{ijn_{ij}}$ son $i.i.d.$ con media 0 y matriz de varianza-covarianza $\Phi_{(1)}$. Finalmente, se supone que v_i, u_{ij}, y e_{ijk} son independientes.

Dejando que

$$y_i = \begin{bmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{in_i} \end{bmatrix} \tag{2.45}$$

Donde $n_{ij}\ x\ 1$ denota el vector de las respuestas para la unidad de la i-ésima unidad del nivel 3 en la j-ésima unidad del nivel 2. Observemos que y_{ij} se puede expresar como

$$y_{ij} = X_{(f)ij}\beta + X_{(3)ij}v_i + X_{(2)ij}u_{ij} + \begin{bmatrix} X'_{(1)ij1}e_{ij1} \\ X'_{(1)ij2}e_{ij2} \\ \vdots \\ X'_{(1)ijn_{ij}}e_{ijn_{ij}} \end{bmatrix}, \tag{2.46}$$

Donde

$$X_{(f)ij} = \begin{bmatrix} X'_{(f)ij1} \\ X'_{(f)ij2} \\ \vdots \\ X'_{(f)ijn_{ij}} \end{bmatrix}, \quad X_{(3)ij} = \begin{bmatrix} X'_{(3)ij1} \\ X'_{(3)ij2} \\ \vdots \\ X'_{(3)ijn_{ij}} \end{bmatrix}, \quad y \;\; X_{(2)ij} = \begin{bmatrix} X'_{(2)ij1} \\ X'_{(2)ij2} \\ \vdots \\ X'_{(2)ijn_{ij}} \end{bmatrix}.$$

Bajo las suposiciones distribucionales dadas arriba, se sigue que,

$$E(y_i) = \mathrm{X}_{(f)i}\,\beta, \qquad (2.47)$$

donde

$$\mathrm{X}_{(f)i} = \begin{bmatrix} \mathrm{X}'_{(f)i1} \\ \mathrm{X}'_{(f)i2} \\ \vdots \\ \mathrm{X}'_{(f)in_i} \end{bmatrix}$$

También

$$Cov\left(y_i, y_i'\right) = \mathrm{X}_{(3)i}\Phi_{(3)}\mathrm{X}'_{(3)i} + \Lambda_i\,, \qquad (2.48)$$

Donde

$$\Lambda_i = \begin{bmatrix} \Lambda_{i1} + D_{i1} & 0 & \cdots & 0 \\ 0 & \Lambda_{i2} + D_{i2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Lambda_{in_i} + D_{in_i} \end{bmatrix} \qquad (2.49)$$

Con

$$\Lambda_{ij} = \mathrm{X}_{(2)ij}\Phi_{(3)}\mathrm{X}'_{(2)ij}$$

y

$$D_{ij} = \begin{bmatrix} \lambda_{ij1} & 0 & \cdots & 0 \\ 0 & \lambda_{ij1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_{ijn_{ij}} \end{bmatrix} \qquad (2.50)$$

con

$$\lambda_{ijk} = X'_{(1)ijk}\Phi_{(1)}X_{(1)ijk}.$$

El modelo dado en (2.46) se introduce el modelo general de 3 niveles en el apartado 2.5 y se puede escribir como

$$y_i = X_{(f)i}\beta + X_{(3)i}v_i + \sum_{j=1}^{n_i} Z_{(2)ij}u_{ij} + \sum_{j=1}^{n_i}\sum_{k=1}^{n_{ij}} U_{(1)ijk}e_{ijk} \tag{2.51}$$

donde

$$X_{(3)i} = \begin{bmatrix} X_{(3)i1} \\ X_{(3)i2} \\ \vdots \\ X_{(3)in_i} \end{bmatrix}, \tag{2.52}$$

$$Z_{(2)ij} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ X_{(2)ij} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \tag{2.53}$$

$$U_{(1)ijk} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ X_{(1)ijk} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \tag{2.54}$$

y donde v_1, u_{ij} , y e_{ijk} se denotan como vectores de los parámetros aleatorios en el nivel 3, 2, y nivel 1 del modelo. Será conveniente sustituir el subíndice doble jk solamente con el subíndice l donde l = 1.2…, n_i^*

$$n_i^* = \sum_{j=1}^{n_i} n_{ij}$$

Así, (2.51) puede ser reescrito como

$$y_i = \mathrm{X}_{(3)i} v_i + \sum_{j=1}^{n_i} Z_{(2)ij} u_{ij} + \sum_{l=1}^{n_i^*} U_{(1)il} e_{il}$$

Bajo las suposiciones distribucionales dadas en el modelo general de 2 niveles del apartado 2.5, se sigue que

$$E(y_i) = \mathrm{X}_{(f)i}\,\beta,$$

donde

$$\mathrm{X}_{(f)i} = \begin{bmatrix} \mathrm{X}_{(f)i1} \\ \mathrm{X}_{(f)i2} \\ \vdots \\ \mathrm{X}_{(f)in_i} \end{bmatrix} \qquad (2.55)$$

También se sigue que

$$Cov\left(y_i, y_i'\right) = \sum i$$

$$= \mathrm{X}_{(3)i} \Phi_{(3)} \mathrm{X}'_{(3)i} + \sum_{j=1}^{n_i} Z_{(2)ij} \Phi_{(2)} Z'_{(2)ij} + \sum_{j=1}^{n_i^*} U_{(1)il} \Phi_{(1)} U'_{(1)il}.$$

Suponiendo que $\hat{\Phi}_{(3)}, \hat{\Phi}_{(2)}$, y $\hat{\Phi}_{(1)}$ son estimadores consistentes de $\Phi_{(3)}, \Phi_{(2)}$, y $\Phi_{(1)}$, respectivamente, de modo que

$$V_i = \mathrm{X}_{(3)i}\Phi_{(3)}\mathrm{X}'_{(3)i} + \sum_{j=1}^{n_i} Z_{(2)ij}\Phi_{(2)}Z'_{(2)ij} + \sum_{j=1}^{n_i^*} U_{(1)il}\Phi_{(1)}U'_{(1)il}$$

es un estimador consistente de $\sum i$

El procedimiento de estimación generalizado por mínimos cuadrados $\hat{\beta}$ de β se obtiene mediante la mínima función cuadrática

$$Q_f = \sum_{i=1}^{N}\left[y_i - X_{(f)i}\beta\right]' V_i^{-1}\left[y_i - X_{(f)i}\beta\right]$$

Con la solución

$$\hat{\beta} = \left[\sum_{i=1}^{N} \mathrm{X}'_{(f)i} V_i^{-1} \mathrm{X}'_{(f)i}\right]^{-1}\left[\sum_{i=1}^{N} \mathrm{X}'_{(f)i} V_i^{-1} y_i\right] \tag{2.56}$$

Para poder estimar $\Phi_{(3)}, \Phi_{(2)}$, y $\Phi_{(1)}$ dejamos que

$$y_i^* = vecs\left(y_i - \mathrm{X}_{(f)i}\beta\right)\left(y_i - \mathrm{X}_{(f)i}\beta\right)' = vecs(\mathrm{M}), \tag{2.57}$$

Donde $vecs(\mathrm{M})$ es el vector operador de M (vector mitad de M), esto porque M es una matriz simétrica. El vector operador se obtiene al reunir todas las columnas en elementos únicos y no duplicados en la columna de M hasta obtener un único vector.

Por ejemplo:

Si M es una matriz simétrica de orden 2x2, entonces:

$\mathrm{M} = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}$ y el vector operador medio de M es,

$$vecs(\mathrm{M})=\begin{pmatrix} a_{11} \\ a_{12} \\ a_{22} \end{pmatrix} \text{ de orden } \frac{2(2+1)}{2}x1=3x1.$$

En consecuencia,

$$\mathrm{E}\left(y_i^*\right)=\mathrm{E}\left(vecs(\mathrm{M})\right)=vecs\sum i$$

Donde $\sum i$ es la covarianza de la respuesta y_i del modelo multinivel general de tres niveles.

Usando el resultado (Browne, 1974), en operaciones de vectores operadores vecs,

$$vec\left(C\mathrm{A}C'\right)=\left(C\otimes C\right)vec\mathrm{A},$$

Donde se sigue que

$$vecV_i=vec\left(\mathrm{X}_{(3)i}\Phi_{(3)}\mathrm{X}'_{(3)i}+\sum_{j=1}^{n_i}Z_{(2)ij}\Phi_{(2)}Z'_{(2)ij}+\sum_{j=1}^{n_i^*}U_{(1)il}\Phi_{(1)}U'_{(1)il}\right)$$

$$vecV_i=\mathrm{X}_{(3)i}\otimes\mathrm{X}_{(3)i}vec\Phi_{(3)}+\sum_{i=1}^{n_i}\left(Z_{(2)ij}\otimes Z_{(2)ij}\right)vec\Phi_{(2)}+\sum_{i=1}^{n_i^*}\left(U_{(1)ij}\otimes U_{(1)ij}\right)vec\Phi_{(1)}.$$

De la expresión anterior tenemos que V_i es una matriz simétrica, de tal modo que existe una matriz única que depende de V_i pero solo a través de las dimensiones.

Por ejemplo:

Sea S_2 una matriz única que duplica los elementos fuera de la diagonal en el vector operador mitad de $\Phi_{(3)}$ ($vecs\Phi_{(3)}$). De modo que podamos así crear el vector operador $vec\Phi_{(3)}$.

Para conocer el vector operador $vec\Phi_{(3)}$ consideremos un caso particular, dos únicas variables aleatorias para el nivel 3, así:

$$\Phi_{(3)} = \begin{pmatrix} \sigma_{v0}^{2} & \sigma_{v01} \\ \sigma_{v01} & \sigma_{v1}^{2} \end{pmatrix}$$

Además, consideramos la matriz duplicada como

$$D_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}_{4x3}$$

Ahora tenemos que,

$$vec\Phi_{(3)} = D_2 vecs\Phi_{(3)}$$

$$vec\Phi_{(3)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}_{4x3} \begin{pmatrix} \sigma_{v0}^{2} \\ \sigma_{v01} \\ \sigma_{v1}^{2} \end{pmatrix}_{3x1} = \begin{pmatrix} \sigma_{v0}^{2} \\ \sigma_{v01} \\ \sigma_{v01} \\ \sigma_{v1}^{2} \end{pmatrix}_{4x1}$$

En general esa matriz única (ver Robert J. Boik, Lecture notes, sección 3.1.11.2) la notaremos por G_p de orden $p^2 \times \frac{1}{2} p(p+1)$ donde p es el número de variables en ese nivel depende de A.

Así, tenemos que:

$$vec\mathrm{A} = G_p vecs\mathrm{A}$$

con A una matriz simétrica de orden $p \times p$.

Hay también una matriz no-única H_p de orden $\frac{1}{2}p(p+1) \times p^2$, así:

$$vecs\mathrm{A} = \mathrm{H}_p vec\mathrm{A}$$

Donde H_p depende de A solo a través de las dimensiones. Una forma de expresar H_p es $\left(G_p{}'G_p\right)^{-1}G'_p$.

Por ejemplo:

Según el ejemplo que venimos tratando, encontremos el vector operador media $vecs\Phi_{(3)}$, de la siguiente manera:

$$vecs\Phi_{(3)} = \left(\left(G_2{}'G_2\right)^{-1}G'_2\right)vec\Phi_{(3)}$$

Donde

$$D'_2 = G'_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}_{3x4} \;;\; D'_2 D_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}_{3x3} \;;\; \left(D'_2 D_2\right)^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}_{3x3}$$

Así, tenemos que,

$$vecs\Phi_{(3)} = \left(\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}_{3x3} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}_{3x4}\right)_{3x4} \begin{pmatrix} \sigma^2_{v0} \\ \sigma_{v01} \\ \sigma_{v01} \\ \sigma^2_{v1} \end{pmatrix}_{4x1} = \begin{pmatrix} \sigma^2_{v0} \\ \sigma_{v01} \\ \sigma^2_{v1} \end{pmatrix}_{3x1}$$

De lo anterior, se tiene que el vector $vecs\sum i$ esta conteniendo los elementos no-duplicados de $\sum i$, se puede entonces escribir como

$$vecs\sum i = vecs\left(\mathrm{X}_{(3)i} \otimes \mathrm{X}_{(3)i} vec\Phi_{(3)} + \sum_{i=1}^{n_i}\left(Z_{(2)ij} \otimes Z_{(2)ij}\right) vec\Phi_{(2)} + \sum_{i=1}^{n_i^*}\left(U_{(1)ij} \otimes U_{(1)ij}\right) vec\Phi_{(1)} \right)$$

Otra forma de escribirlo es,

$$vecs\sum i = \mathrm{X}_i^* \tau \tag{2.58}$$

donde

$$\mathrm{X}_i^* = H_{n_i^*}\begin{bmatrix} \left(\mathrm{X}_{(3)i} \otimes \mathrm{X}_{(3)i}\right) G_q \\ \left(\sum_{j=1}^{n_i} Z_{(2)ij} \otimes Z_{(2)ij}\right) G_m \\ \left(\sum_{j=1}^{n_i^*} U_{(1)il} \otimes U_{(1)il}\right) G_r \end{bmatrix}$$

y

$$\tau = \begin{bmatrix} vecs\Phi_{(3)} \\ vecs\Phi_{(2)} \\ vecs\Phi_{(1)} \end{bmatrix}.$$

La expresión (2.58) es el vector operador medio de la matriz de varianza-covarianza del modelo general de tres niveles. Consideremos un ejemplo sencillo, tratado en el apartado 2.4. Utilizaremos la expresión (2.58) para llegar a la forma (2.15).

Encontremos el vector operador medio de la matriz de varianza-covarianza V, así:

$$vecs\left(V_2\right) = vecs\left(\mathrm{X}_j \Phi_{(2)} \mathrm{X}'_j + \left(\Phi_1\right)\right) = vecs\left(\mathrm{X}_j \Phi_{(2)} \mathrm{X}'_j\right) + vecs\left(\Phi_1\right)$$

$$vecs\left(V_2\right) = \mathrm{X}_j \otimes \mathrm{X}_j vecs\left(\Phi_{(2)}\right) + vecs\left(\Phi_{(1)}\right) \tag{2.59}$$

Para el término derecho de la expresión (2.59) existe una matriz no única, de tal manera que podemos escribirlo de la forma siguiente:

$$vecs\left(V_2\right) = \mathrm{H}_P vec\left(\Phi_{(2)}\right) + vecs\left(\Phi_{(1)}\right) \tag{2.60}$$

Para encontrar el $vec\left(\Phi_{(2)}\right)$ de (2.60), existe en este una matriz única, quedando de este modo como:

$$vecs\left(V_2\right) = \mathrm{H}_2\left(\mathrm{X}_j \otimes \mathrm{X}_j G_2\right) vecs\left(\Phi_{(2)}\right) + vecs\left(\Phi_{(1)}\right)$$

Desarrollemos la expresión anterior, encontrando H_2 y $\mathrm{X}_j \otimes \mathrm{X}'_j G_2$, de la siguiente manera:

$$\mathrm{H}_2 = \left(G_2' G_2\right)^{-1} G'_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}_{3x3} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}_{3x4} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}_{3x4}$$

$$\mathrm{X}_j = \begin{pmatrix} 1 & x_{1j} \\ 1 & x_{2j} \end{pmatrix}$$

Luego

$$\mathrm{X}_j \otimes \mathrm{X}_j = \begin{pmatrix} 1 & x_{1j} \\ 1 & x_{2j} \end{pmatrix}_{2xj} \otimes \begin{pmatrix} 1 & x_{1j} \\ 1 & x_{2j} \end{pmatrix}_{2xj} = \begin{pmatrix} 1 & x_{1j} & x_{1j} & x_{1j}^2 \\ 1 & x_{2j} & x_{1j} & x_{1j}x_{2j} \\ 1 & x_{1j} & x_{2j} & x_{1j}x_{2j} \\ 1 & x_{2j} & x_{2j} & x_{2j}^2 \end{pmatrix}$$

Así,

$$\left(\mathrm{X}_j\otimes\mathrm{X}_j\right)G_2=\begin{pmatrix}1 & x_{1j} & x_{1j} & x_{1j}^2\\ 1 & x_{2j} & x_{1j} & x_{1j}x_{2j}\\ 1 & x_{1j} & x_{2j} & x_{1j}x_{2j}\\ 1 & x_{2j} & x_{2j} & x_{2j}^2\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}=\begin{pmatrix}1 & 2x_{1j} & x_{1j}^2\\ 1 & x_{1j}+x_{2j} & x_{1j}x_{2j}\\ 1 & x_{1j}+x_{2j} & x_{1j}x_{2j}\\ 1 & 2x_{2j} & x_{2j}^2\end{pmatrix} \tag{2.61}$$

Ahora multipliquemos (2.61) con el vector operador medio, así,

$$\left(\left(\mathrm{X}_j\otimes\mathrm{X}_j\right)G_2\right)vecs\left(\Phi_{(2)}\right)=\begin{pmatrix}1 & 2x_{1j} & x_{1j}^2\\ 1 & x_{1j}+x_{2j} & x_{1j}x_{2j}\\ 1 & x_{1j}+x_{2j} & x_{1j}x_{2j}\\ 1 & 2x_{2j} & x_{2j}^2\end{pmatrix}\begin{pmatrix}\sigma_{u0}^2\\ \sigma_{u01}\\ \sigma_{u1}^2\end{pmatrix}$$

$$\left(\left(\mathrm{X}_j\otimes\mathrm{X}_j\right)G_2\right)vecs\left(\Phi_{(2)}\right)=\begin{pmatrix}\sigma_{u0}^2+2\sigma_{u01}x_{1j}+\sigma_{u1}^2x_{1j}^2\\ \sigma_{u0}^2+\sigma_{u01}\left(x_{1j}+x_{2j}\right)+\sigma_{u1}^2x_{1j}x_{2j}\\ \sigma_{u0}^2+\sigma_{u01}\left(x_{1j}+x_{2j}\right)+\sigma_{u1}^2x_{1j}x_{2j}\\ \sigma_{u0}^2+2\sigma_{u01}x_{2j}+\sigma_{u1}^2x_{2j}^2\end{pmatrix}$$

En consecuencia, tenemos que,

$$vecs\left(V_2\right)=\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1/2 & 1/2 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix}\sigma_{u0}^2+2\sigma_{u01}x_{1j}+\sigma_{u1}^2x_{1j}^2\\ \sigma_{u0}^2+\sigma_{u01}\left(x_{1j}+x_{2j}\right)+\sigma_{u1}^2x_{1j}x_{2j}\\ \sigma_{u0}^2+\sigma_{u01}\left(x_{1j}+x_{2j}\right)+\sigma_{u1}^2x_{1j}x_{2j}\\ \sigma_{u0}^2+2\sigma_{u01}x_{2j}+\sigma_{u1}^2x_{2j}^2\end{pmatrix}+\begin{pmatrix}\sigma_{e0}^2\\ 0\\ \sigma_{e0}^2\end{pmatrix}$$

$$vecs\left(V_2\right)=\begin{pmatrix}\sigma_{u0}^2+2\sigma_{u01}x_{1j}+\sigma_{u1}^2x_{1j}^2+\sigma_{e0}^2\\ \sigma_{u0}^2+\sigma_{u01}\left(x_{1j}+x_{2j}\right)+\sigma_{u1}^2x_{1j}x_{2j}\\ \sigma_{u0}^2+2\sigma_{u01}x_{2j}+\sigma_{u1}^2x_{2j}^2+\sigma_{e0}^2\end{pmatrix}$$

Como podemos observar, hemos llegado al mismo resultado de (2.15).

Ahora consideramos la forma cuadrática

$$Q_{\mathrm{T}} = \left\{ \sum_{i=1}^{N} \left[y_i^* - \mathrm{X}_i^* \tau \right]' W_i^{-1} \left[y_i^* - \mathrm{X}_i^* \tau \right] \right\}$$

de donde W_i es un estimador consistente de la covariación de y_i^*.

Por ejemplo, ha sido demostrado eso por Browne (1974) y Goldstein (1989), si

$$W_i^{-1} = \frac{1}{2} G'_{n_i} \left(V_i^{-1} \otimes V_i^{-1} \right) G_{n_i} \text{ ,} \tag{2.62}$$

entonces W_i es un estimador consistente de la covariación de y_i^*.

Minimización Q_{T} con respecto a τ obtendremos lo siguiente

$$\hat{\tau} = \left[\sum_{i=1}^{N} \mathrm{X}_i^{*\prime} W_i^{-1} \mathrm{X}_i^* \right]^{-1} \left[\sum_{i=1}^{N} \mathrm{X}_i^{*\prime} W_i^{-1} y_i^* \right] \tag{2.63}$$

Para asegurar eficacia de cómputo, los componentes de $\hat{\tau}$ se simplifican más a fondo según lo discutido en du Toit (1995).

2.7 Residuales.

Una vez estimado el modelo es posible calcular los residuos brutos (*raw residuals*). Consideraremos el modelo multinivel general de dos niveles, así:

$$\tilde{y}_{ij} = y_{ij} - \hat{y}_{ij} \tag{2.64}$$

donde en general al considerar el modelo (2.13),

$$\hat{y}_{ij} = \hat{\beta}_0 + \hat{\beta}_1 x_{ij} \tag{2.65}$$

De esta forma, para estimar los residuos de las unidades en el nivel 2, dadas las estimaciones de los parámetros $\hat{\beta}$, se calcula

$$\hat{u}_{oj} = E\left(u_{oj} \middle| Y , \hat{\beta} , \hat{\Omega} \right) \tag{2.66}$$

Considerando la regresión del conjunto de residuos $\hat{u}_{oj}$ sobre $\tilde{Y} = Y - X \hat{\beta}$ se obtiene el siguiente estimador consistente,

$$\hat{u} = \frac{n_j \sigma_{u0}^2}{n_j \sigma_{u0}^2 + \sigma_{e0}^2} \tilde{y}_j \tag{2.67}$$

Cuando se desea hacer inferencia acerca del verdadero $\hat{u}_{oj}$ despreciando la variabilidad muestral de los parámetros aleatorios, se tiene (Goldstein, 1995):

$$\hat{u}_{oj} = \frac{n_j \sigma_u^2}{\left(n_j \sigma_u^2 + \sigma_{e0}^2\right)} y_j$$

$$\tilde{e}_{0ij} = \tilde{y}_{ij} - \hat{u}_{oj} \tag{2.68}$$

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

donde n_j es el número de las unidades del nivel 1 en la j-ésima unidad del nivel 2. Las estimaciones residuales no son, incondicionale ni insesgadas sino que son constantes. El factor que multiplica el promedio $\left(\tilde{y}_j\right)$ de los residuales β parámetros de la j-ésima unidad con frecuencia se refiere a menudo como 'factor de contracción' puesto que es siempre menor o igual a uno. Cuando n_j aumenta este factor tiende a ser uno, y cuando el número de las unidades del nivel 1 y en el nivel 2 disminuye el 'el estimador de contracción' de u_{oj} se convierte más cercano a cero. En muchos aplicaciones el nivel más alto de los residuales son de interés en la su propio derecho y la contracción creciente para una unidad pequeña del nivel 2 se puede ver como expresar la carencia relativa de la información en la unidad, de modo que la mejor estimación ponga la prediccion residual cercano al valor total de la población según lo dado por la parte fija.

Estos residuales por lo tanto pueden tener dos papeles. Su interpretación básica está como variables aleatorias con una distribución que valores de los parametros nos digan sobre la variación entre las unidades del nivel 2, y cuáles proporcionan las estimaciones eficientes para los coeficientes fijos. Una segunda interpretación es como las estimaciones individuales para cada unidad del nivel 2 donde utilizamos la suposicion que pertenecen a una población de unidades para predecir sus valores. En detalle, para las unidades que tienen solamente algunas unidades en el nivel 1, podemos obtener estimaciones más exactas que si no hiciéramos caso de la suposicion de la calidad de miembro de la población y utilizáramos solamente la información de esas unidades. Esto llega a ser especialmente importante para las estimaciones de las residuales y para los coeficientes aleatorios, donde en el caso extremo solamente una unidad del nivel 1 en una unidad del nivel 2 carecemos de la información para formar una estimación independiente.

Como en modelos de un solo nivel podemos utilizar los residuales estimados para ayudar a comprobar en las suposiciones del modelo. Las dos suposiciones particulares que pueden ser estudiadas fácilmente son la suposicion de normalidad y que son constantes las variaciones en el modelo. Porque las variaciones de las estimaciones residuales dependen en general de los valores de los coeficientes fijos es común estandardizar los residuales dividiéndose por los errores estandares apropiados.

2.8 Estadística inferencial.

En el análisis del apartado 2.6 presentamos la estimación de los parámetros de la parte fija junto con sus errores estándar. Estos son adecuados para probar hipótesis o construir intervalos de confianza separadamente para cada parámetro. En muchos casos, sin embargo, estamos interesados en combinaciones de parámetros. Para probar las hipótesis, muy a menudo surge para variables explicativas agrupadas o clasificadas donde los efectos de n grupos se definen en términos de n-1 contraste de variables falsas y desear porbar simultáneamente si estos contrastes son ceros.

Después, en contrastes, discutimos las hipótesis de la forma $c_1\beta_1 + c_2\beta_2 + \ldots + c_q\beta_q = k$ sobre los elementos del vector fijo del parámetro β.

2.8.1 Error estándar

De la expresión (2.56) se sigue que la matriz de covarianza de $\hat{\beta}$ en un modelo multinivel de tres niveles está dada por

$$\operatorname{cov}\left(\hat{\beta}, \hat{\beta}'\right) = \left[\mathrm{X}_{(f)}^{*\prime} \Sigma^{-1} \mathrm{X}_{(f)}^{*} \right]^{-1} \tag{2.69}$$

En la práctica, $\sum_i$ es desconocido y es substituido por el estimador de máxima verosimilitud $\sum_i(\hat{\tau}) = V_i$. Por lo tanto, un estimador consistente de la matriz de covarianza de $\hat{\beta}$ se da por

$$\operatorname{cov}\left(\hat{\beta}, \hat{\beta}'\right) = \left[\mathrm{X}_{(f)}^{*\prime} \Sigma^{-1} \mathrm{X}_{(f)}^{*} \right]^{-1} \tag{2.70}$$

Semejantemente, puede ser demostrado que un estimador consistente de la matriz de covarianza de $\hat{\tau}$ está dado por

$$\operatorname{cov}\left(\hat{\tau}, \hat{\tau}'\right) = \left[\sum_{i=1}^{N} \mathrm{X}_i^{*\prime} W_i^{-1} \mathrm{X}_i^{*} \right]^{-1} \tag{2.71}$$

Los elementos diagonales de la matriz de la covarianza (2.70) se pueden utilizar para obtener las estimaciones con muestras grandes de los errores estándar para las estimaciones fijas del parámetro. Para las muestras grandes, $\hat{\beta}$ y $\hat{\tau}$ tinen distribuciones aproximadamente normales multivariante.

2.8.2 Contrastes

La construccion de hipótesis o de funciones lineales de los parámetros es un análisis estadístico útil también que permite al investigador realizar las pruebas de hipótesis referente a la igualdad de subconjuntos de parámetros. En este apartado un resumen de los resultados requeridos para la prueba del contraste se da a conituación.

Sea A una matriz de contraste de orden $p \times q$, donde p denota el número de contrastes, se puede utilizar para formular hipótesis complejas sobre varios elementos de β.

La hipótesis se escribe en la forma $C\beta = k$, donde está un vector k conocido de orden $p \times 1$.

Consideremos como ejemplo el caso donde $q = 3$ y ésta puede ser probada de la siguiente forma:

$$\beta_1 - \beta_2 = 0$$
$$\beta_3 - \beta_2 = 0.$$

La hipótesis nula se puede formular como

$$\begin{bmatrix} 1 & -1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Para las muestras grandes, puede ser demostrado que la varianza aleatoria del vector $C\hat{\beta}$ será distribuida aproximadamente como $N\left(C\beta, C\left(X'_{(f)}\hat{V}^{-1}X_{(f)}\right)^{-1}C'\right)$.

Por lo tanto, si H_0 es verdadera,

$$M = \left(C\hat{\beta} - k\right)' \left\{ C\left(X'_{(f)} \hat{V}^{-1} X_{(f)} \right)^{-1} \right\} \left(C\hat{\beta} - k\right) \tag{2.70}$$

Que tiene una distribución χ^2 aproximada con p grados de libertad.

Denotando c' como la i-ésima fila de C y $\chi^2_{q,\alpha}$ el valor crítico de la distribución χ^2 con q grados de libertad. Un sistema de $100(1-\alpha)\%$ intervalos de confianza simultáneos para los elementos de p intervalos de confianza de $C\beta$ que es dado por los p intervalos

$$c'_i \hat{\beta} \pm \left\{ c'_i \left(X'_{(j)} \hat{V}^{-1} X_{(j)} \right)^{-1} c'_i \chi^2_{q,\alpha} \right\}^{0.5}, \quad p<\text{q} \tag{2.71}$$

La hipótesis nula $H_0 : \hat{\beta}_j = 0, \quad j = 1, 2, ..., q$ es probada usando la estadística de prueba

$$z = \frac{\hat{\beta}_k}{S.E\left(\hat{\beta}_k\right)}$$

La cual, para muestras grandes, tiene una distribución aproximadamente $N(0,1)$ si H_0 es verdadera.

CAPÍTULO III

APLICACIÓN DE MODELOS MULTINIVEL

Introducción.

Para la aplicación de los modelos multinivel se dispone de una base de datos de un estudio que ha sido realizado en diversos países de Latinoamérica el cual consistió en aplicar un instrumento de recolección de información a: alumnos, profesores, padres de familia y escuela.

Con el fin de aplicar la teoría desarrollada en los capítulos anteriores se va a trabajar en este documento con los datos que se obtuvieron para un país centroamericano; en adelante se le llamará "base de datos".

Este capítulo está estructurado de la siguiente manera:

- En la primera parte se presentan los objetivos que se pretende alcanzar a partir de la aplicación de la teoría multinivel. Además, descripción de variables, técnica de análisis, test estadístico y la estrategia de análisis.

- En la segunda parte, se analiza primero la descomposición de la varianza tanto del nivel individual como grupal a partir del modelo nulo o no condicional. Seguidamente se identifican los efectos significativos del modelo geográfico (tomando como ubicación aquellas escuelas que pertenecen al sector urbano público y privado) a partir de la aplicación de los modelos jerárquicos lineales (HLM).

- En la tercera parte, se pretende responder a la pregunta ¿Cuál es el efecto que tienen las características demográficas, socioeconómicas y culturales referidas a: alumno individual y a su familia (factores extra-escolares individuales); y al contexto de la escuela (factores extra-escolares contextuales). Con el objeto de saber que factores son estadísticamente significativos para los resultados obtenidos por los alumnos en Lenguaje y Matemática, de tal manera que estas variables puedan ser agregadas y desagregadas para definir el modelo multinivel, (A: efectos que tienen las características del alumno, familia y contexto de la escuela sobre el rendimiento de los estudiantes en Lenguaje y Matemática) y poder así, determinar que indicadores inciden significativamente sobre el resultado de los estudiantes.

- La cuarta parte, tiene como objetivo identificar cuales de las variables pertenecientes a los niveles de agregación demuestran asociación importante con el logro de los estudiantes sobre el resultado en Lenguaje y Matemática. Primero se busca un modelo multinivel, (B: efecto de las características cultura institucional y práctica en el aula) para determinar que variables están influyendo estadísticamente significativa sobre el resultado de los estudiantes. Seguidamente clasificar estas variables para agregarlas y desagregarlas ya cuando definamos el modelo óptimo para Lenguaje y Matemática.

- La quinta parte, tiene como objetivo determinar el modelo óptimo en Lenguaje y Matemática a partir de los indicadores que en la tercera y cuarta parte de este capítulo incidieron de una manera significativa al resultado de Lenguaje y Matemática para contabilizar la variabilidad entre escuelas (ejemplo comprender porque algunas tienes más altos promedios que otras) y evaluar el efecto que tienen dichos indicadores como: ubicación sociodemográfica, el estatus socioeconómico y el clima en el aula, sobre el resultado de los estudiantes.

- Finalmente se hace un diagnóstico del modelo óptimo para los residuos del nivel uno y dos en Lenguaje y Matemática. Con el propósito de validar dicho modelo, a partir las hipótesis importantes como la homocedasticidad y normalidad. El primero, para saber si se cumple el supuesto de que las perturbaciones del nivel uno y dos tienen varianza constante. Y, el segundo, para saber si los residuos presentan una distribución normal. De tal manera, que podemos utilizar el modelo multinivel para generar predicciones.

3.1 Objetivos de la aplicación.

El objetivo general es aplicar la teoría de los modelos multinivel a los datos obtenidos en una muestra de centros educativos y alumnos que se tienen en la base de datos para identificar las condiciones en las que los alumnos de tercero y cuarto grados de educación básica, alcanzan los aprendizajes en Lenguaje y Matemática, según el resultado registrado en pruebas estandarizadas.

Este objetivo general se desagrega en los siguientes objetivos específicos para Lenguaje y Matemática:

i) Determinar la importancia relativa de cada uno de los dos componentes de la variación total del rendimiento escolar en el conjunto de datos: la desigualdad de los rendimientos promedios de las escuelas (inter-escuela) en la base de datos, y la desigualdad de rendimiento entre los alumnos dentro de la escuela (intra-escuela).

ii) Identificar los factores demográficos, socioeconómicos y culturales propios del alumno individual y su familia que afectan el nivel y la distribución del logro escolar.

iii) Determinar el efecto de la composición demográfica, socioeconómica y cultural de la escuela sobre el logro de aprendizaje.

iv) Determinar el efecto de las características de la escuela y del aula sobre el nivel y la distribución de los rendimientos promedios institucionales.

v) Determinar la variación de los rendimientos promedios de las escuelas y de la variación de los rendimientos de los alumnos dentro de la escuela.

vi) Identificar factores extraescolares y escolares que expliquen las posibles diferencias en el nivel de rendimiento de los alumnos del sector urbano público y privado.

3.2 Descripción de los datos.

El presente estudio es de tipo “correlacional” porque la determinación del efecto factor (Ej. Nivel socioeconómico familiar) sobre el logro escolar del alumno se basa en el grado de asociación entre ambas variables. Expresiones como efecto, explicación, determinación, etc., están referidas y restringidas a la asociación (correlación) estadística entre variables o a la capacidad predictiva de algunas variables respecto de los resultados escolares.

La información de la base de datos se encuentra clasificada por escuelas de la zona urbana y escuelas de la zona rural. Sin embargo, para realizar la aplicación de los modelos multinivel únicamente se trabajará con la información de las escuelas de la zona urbana[8] para el sector público y privado.

A continuación se presenta el número de alumnos en la muestra final de pruebas de Lenguaje y Matemática, por estrato:

Lenguaje		
Público	Privado	Total
900	166	1066
Matemática		
Público	Privado	Total
887	164	1051

[8] Los modelos que determinaremos sobre el rendimiento de los estudiantes será solamente para el sector urbano (la técnica también se puede aplicar para determinar los modelos óptimos sobre el resultado de Lenguaje y Matemática considerando el sector rural).

3.2.1 Descripción de Variables.

a) Variables-criterio.

Se comprenden dos variables criterio: Puntuaciones en la prueba de Lenguaje y Matemática; indicadoras de la calidad de la educación. Con la primera el alumno construye y desarrolla conocimientos y aprendizaje, le da significado a sus experiencias y sentido a los conocimientos de los otros, constituye la base del desarrollo de la capacidad de pensar y la puerta de acceso al conocimiento y la capacidad crítica, sin él todos los demás aprendizajes están comprometidos. La segunda, se desarrolla la capacidad para resolver situaciones problemáticas que requieren la puesta en práctica de habilidades matemáticas de complejidad superior.

b) Variables individuales del alumno y familia.

Características del estudiante	**Notación para software HLM**	**Notación para presentación de Modelos.**
Grado (cuarto grado=1, tercer grado=0)	MUDA	X_{ij_grado}
Sexo (niñas=1, niños=0)	FEMALEC	X_{ij_female}
Antecedentes académicos del alumno.		
El padre le lee a menudo al niño	READLOTS	$X_{ij\_radlots}$
El padre le lee a veces al niño	READSOME	$X_{ij\_readsome}$
El Padre le ayuda al hijo en la preparación académica	HELP_HW	$X_{ij\_help\_hw}$
Características familia	**Notación para software HLM**	**Notación para presentación de Modelos.**
El padre está de 3 a 4 horas al día	HOME34	$X_{ij\_\mathrm{hom}e34}$
Material didáctico en el hogar	TENBOOKS	$X_{ij\_tenbooks}$
Estatus sociocultural	SES_1	$X_{ij\_ses\_1}$
Educación de los padres	PARED_1	$X_{ij\_pared\_1}$
Práctica en el aula		
Implicación del padre (estudiante-nivel)	PAR_INV	$X_{ij\_par\_inv}$
Implicación del padre (escuela-nivel)	MNPARINV	$W_{ij\_mnparinv}$
Clima en el aula	MNDISCIP	$W_{ij\_mndiscip}$

c) **Variables contextuales institucionales.**

Recursos de la Escuela	Notación para software HLM	Notación para presentación de Modelos.
Relación alumno-profesor	PTR_ALL	$X_{ij\_ptr\_all}$
Material didáctico	MATERIAL	$W_{0k\_material}$
Tamaño de biblioteca	BOOK1000	$W_{0k\_book1000}$
Profesor en servicio	INSERV	X_{ij_inserv}
La escuela ofrece el desayuno libre	BREKFAST	$X_{ij\_brekfast}$
La escuela ofrece la comida libre	FREEMEAL	$X_{ij\_freemeal}$
Tamaño de la clase	CLSIZE	$W_{0k\_book1000}$
Efecto contexto		
Nivel Socioeconómico de la Escuela	MEANSES	$W_{0k\_meanses}$
Ubicación geográfica		
Escuela urbano pública	URB_PUB	$W_{0k\_urb\_pub}$
Escuela urbano privada	URB_PR	$W_{0k\_urb\_pr}$

Nota: k=1,2,...,p (p: número de variables en el nivel 2)

3.3 Técnica de Análisis.

Se utiliza la técnica de análisis estadístico por niveles múltiples. La cual es una técnica correlacional adecuada para analizar variaciones en las características de los individuos (Ej. Rendimiento en Lenguaje) que son miembros de un grupo (Ej. Escuela), mediciones que forman parte de una estructura agrupada y anidada. La técnica permite la descomposición de la variación de una variable (Ej. Rendimiento) en sus componentes dentro de grupo y entre grupo, y el análisis de asociación entre variables en esos diferentes niveles de agregación. El modelo se compone de dos partes una aleatoria y otra fija, la primera muestra en cada nivel de agregación la estimación de la varianza de los parámetros determinados en la parte fija; la segunda con los parámetros que definen una línea promedio para todos los alumnos de todas las escuelas. Es importante resaltar que la inclusión de más variables en un modelo, permite estimar con mayor precisión el efecto de cada una de las variables, controlando por el efecto de las otras.

3.3.1 Test Estadístico.

Generalmente la técnica estadística que se utiliza para determinar si una variable es un factor asociado al resultado escolar de los estudiantes, es analizar el grado de asociación lineal, y para decidir si la asociación es estadísticamente significativa se usa el test de la razón de máxima verosimilitud. Trabajando con la hipótesis de nulidad de diferencia igual a cero, la diferencia entre los valores de máxima verosimilitud de dos modelos sigue la distribución de Chi-cuadrado, con grados de libertad igual al número de nuevos parámetros. Habitualmente para indicar el nivel de significación de cada estimación se usan como referencia el límite de probabilidad propuestos o utilizados por Fisher (0.01, 0.05, 0.0025). En este trabajo se utilizará el límite de 0.05.

3.3.2 Estrategia de Análisis: niveles de agregación.

La base de datos que disponemos permite trabajar con dos niveles: alumno y escuela (es importante destacar que a pesar que el estudio posee variables a nivel del maestro y del aula, el diseño muestral no permite incluir el nivel aula como tal. Por lo tanto, las variables referidas al aula y el maestro deben agregarse a nivel escuela).

3.4 Análisis multinivel para Lenguaje y Matemática.

En este apartado, se analizan los resultados de la base de datos en la prueba de Lenguaje y Matemática. Primero se descompone la varianza de resultados de los estudiantes de tal manera que se pueda examinar si las diferencias en el rendimiento se dan entre escuelas o al interior de éstas. Finalmente, se aplican modelos lineales jerárquicos (HLM)) para identificar los factores que explican los resultados. En estos modelos se incorporan una serie de variables de los alumnos (y sus familias) y de los centros escolares, para calcular el efecto de cada una sobre los resultados en la base de datos.

3.4.1 Las diferencias geográficas y del rendimiento escolar para Lenguaje y Matemática.

En este apartado se abordan algunos de los objetivos específicos propuestos, los de interés son:

i) Las variaciones del rendimiento escolar entre escuelas y alumnos.
ii) El efecto del sector geográfico (urbano público y urbano privado).

Para darle respuesta a estos dos literales, se consideran dos bases de datos, una del alumno y otra de la escuela, esto, debido al nivel de agregación que intervienen en el análisis.

La técnica que se utilizará para este análisis es el "análisis estadístico multinivel", la cual se desarrolla a través de los siguientes pasos:

i) Modelo "One-Way ó Modelo Nulo": significa descomponer la varianza, media y error estándar de los niveles (alumno y escuela).

ii) Modelo geográfico: determinar las diferencias urbano público y urbano privado del los resultados en Lenguaje y Matemática.

3.4.2 Modelo Nulo para Lenguaje.

La estimación de los modelos nulos representa el punto de partida de todo análisis multinivel y presenta las siguientes formas funcionales y supuestos:

Tabla 3.2: Características de los modelos nulos.

Definición	Supuesto	Varianza
$y_{ij} = \beta_{0j} + e_{0ij}$ $\beta_{0j} = \beta_0 + u_{0j}$ $y_{ij} = \beta_0 + \left(u_{0j} + e_{0ij}\right)$	$u_{0j} \sim N\left(0, \sigma_{u0}^2\right)$ $e_{0ij} \sim N\left(0, \sigma_{e0}^2\right)$	$\operatorname{var}\left(y_{ij} \middle\| \beta_0, \beta_1, x_{ij}\right) = \operatorname{var}\left(u_0 + e_{0ij}\right) = \sigma_{u0}^2 + \sigma_{e0}^2$ $\rho = \dfrac{\sigma_{u0}^2}{\sigma_{u0}^2 + \sigma_{e0}^2}$

En la tabla 3.2, se observa lo siguiente:

- Los modelos nulos no contienen ningún tipo de predictor (variable exógena), bien individual o grupal, excluyendo incluso aquellas variables que pertenecen al nivel inferior o superior.

- El coeficiente β_0 equivale al intercepto o media global conformada por la parte fija del modelo.

- Los términos de error del grupo u_{0j} (escuela) y del individuo e_{0ij} (alumnos) siguen una distribución normal con media igual a cero y varianzas iguales σ_{e0}^2 y σ_{u0}^2, respectivamente.

- La varianza total es igual a la suma de las varianzas de u_{0j} y e_{0ij} y la importancia del grupo, la cual es la proporción de la varianza total atribuible a ese nivel (ρ), aunque sin ningún control respecto del efecto de las variables de contexto.

El primer parámetro que estimáremos es la variación alrededor de la media global de rendimiento, involucrando simultáneamente los dos niveles con que estamos trabajando: alumno y escuelas.

Variación entre-escuelas.

Aquí se estima una media global del resultado en Lenguaje y se conoce como varían los promedios de las escuelas y los puntajes de los alumnos en torno a ella.

El modelo global nulo para lenguaje.

Modelo nivel 1

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

Modelo nivel 2

$$\beta_{0j} = \beta_0 + u_{0j}$$

Modelo con una sola ecuación.

$$y_{ij\_Lenguaje} = \beta_0 + \left(u_{0j} + e_{0ij}\right)$$

Donde $y_{ij\_Lenguaje}$ es el rendimiento escolar en Lenguaje del i-ésimo alumno en la j-ésima escuela y está representada como una función de la gran media β_0 o media de todos los centros más dos errores aleatorios del nivel escuela u_{0j} y del nivel estudiante e_{0ij}.

Tabla 3.3: Estimación final de componentes de varianza para Lenguaje.

Efectos aleatorios	**Desviación estándar**	**Componentes de varianza**	**Df**	χ^2 **(chi-cuadrada)**	**p-valor**
u_{0j}	24.39242	594.99027	54	932.45386	0.000
e_{0ij}	39.70943	1576.83865			

$$\mathrm{var}\left(e_{0ij}\right) = \hat{\sigma}^2_{eo} = 1576.84$$

$$\mathrm{var}\left(u_{0j}\right) = \hat{\sigma}^2_{u0} = 594.99$$

$$\mathrm{var}\left(y_{ij} \middle| \beta_0, \beta_1, x_{ij}\right) = \mathrm{var}\left(u_0 + e_{0ij}\right) = \hat{\sigma}^2_{u0} + \hat{\sigma}^2_{e0} = 594.99027+1576.83865=2171.83$$

El coeficiente de correlación intra-clase, representa en este modelo la proporción de varianza de "y" entre escuelas, el cual es estimado de la manera siguiente:

$$\hat{\rho} = \frac{\hat{\sigma}^2_{u0}}{\hat{\sigma}^2_{u0} + \hat{\sigma}^2_{e0}} = \frac{594.99027}{594.99027+1576.83865} = 0.27$$

De la variación total del resultado escolar en lenguaje el 27.4% se debe a las diferencias entre las escuelas, o correlación intra-escuelas.

Representa el peso que tienen las características grupales o internas al sistema escolar en la explicación de las variaciones totales de los resultados. Por otro lado, la estimación del nivel 1, alumno (72.6%) representa el peso de los factores externos al sistema escolar.

El estimador de la fiabilidad de la media muestral en cualquier escuela j-ésima es el siguiente:

$$\hat{\lambda}_j = \frac{\hat{\sigma}^2_{u0}}{\left(\hat{\sigma}^2_{u0} + \left(\frac{\hat{\sigma}^2_{e0}}{n_j}\right)\right)}$$

y un indicador global de la fiabilidad es el promedio de la fiabilidad de las escuelas, el cual es:

$$\hat{\lambda}_j = \sum \frac{\hat{\lambda}_j}{j} = 0.94$$

Lo cual nos indica que la media muestral[9] tiende a ser un buen estimador de la medida verdadera de la escuela, es decir, con un nivel de confianza del 94% se concluye que la escuela hace un aporte del 27% sobre el resultado del alumno en Lenguaje.

[9] Anthony S. Bryk, Stephen W Raundenbush. "Hierarchical Linear Models: Applications and Data Analysis Methods". Advanced Quantitative Techniques in the Sciences Series, 1992, pags. 61-62

3.4.3 Modelo Multinivel geográfico para Lenguaje.

i) Análisis para el sector público en Lenguaje:

Modelo nivel 1

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

Modelo nivel 2

$$\beta_{0j} = \beta_0 + \beta_1 W_{01\_Urb\_pub} + u_{0j}$$

Modelo con una sola ecuación.

$$y_{ij\_Lenguaje} = \beta_0 + \beta_1 W_{01\_Urb\_pub} + \left(u_{0j} + e_{0ij}\right)$$

Donde

$y_{ij\_Lenguaje}$ es el rendimiento escolar en Lenguaje del i-ésimo alumno en la j-ésima escuela.

β_0 es la media global de las escuelas.

β_1 es el peso esperado entre el rendimiento de los alumnos del sector urbano público.

u_{0j} es el error aleatorio para el nivel dos (escuelas).

e_{0ij} es el error aleatorio para el nivel uno (alumnos).

Tabla 3.4: Estimación final de los efectos fijos.

Efectos fijos	Coeficiente	Error estándar	T-ratio	d.f. aprox	p-valor
Intercepto 1, β_{0j}					
Intercepto 2, β_0	230.291678	3.229264	71.314	53	0.000
Pendiente, β_1	22.717739	8.688224	2.615	53	0.012

Tabla 3.5: Estimación final de componentes de varianza.

Efectos aleatorios	**Desviación estándar**	**Componentes de varianza**	**Df**	χ^2 **(chi-cuadrada)**	**p-valor**
u_{0j}	23.09823	533.52843	53	842.89410	0.000
e_{0ij}	39.70901	1576.80545			

$$\operatorname{var}\left(e_{0ij}\right)=\hat{\sigma}_{eo}^{2}=1576.80545$$

$$\operatorname{var}\left(u_{0j}\right)=\hat{\sigma}_{u0}^{2}=533.52843$$

$$\operatorname{var}\left(y_{ij}\middle|\beta_0,\beta_1,x_{ij}\right)=\operatorname{var}\left(u_0+e_{0ij}\right)=\hat{\sigma}_{u0}^{2}+\hat{\sigma}_{e0}^{2}=533.52843+1576.80545=2110.33$$

El coeficiente de correlación intra-clase, representa en este modelo la proporción de varianza de “y” entre escuelas, el cual es estimado de la manera siguiente:

$$\hat{\rho}=\frac{\hat{\sigma}_{u0}^{2}}{\hat{\sigma}_{u0}^{2}+\hat{\sigma}_{e0}^{2}}=\frac{533.52843}{533.52843+1576.80545}=0.25$$

De la variación total del resultado escolar en lenguaje es 25.3% y se debe a las diferencias entre las escuelas, o correlación intra-escuelas.

Representa el peso que tienen las características grupales o internas al sistema escolar en la explicación de las variaciones totales de los resultados. Por otro lado, la estimación del nivel 1, alumno (74.7%) representa el peso de los factores externos al sistema escolar.
El estimador de la fiabilidad de la media muestral en cualquier escuela j-ésima es el siguiente:

$$\hat{\lambda}_j=\frac{\hat{\sigma}_{u0}^{2}}{\left(\hat{\sigma}_{u0}^{2}+\left(\frac{\hat{\sigma}_{e0}^{2}}{n_j}\right)\right)}$$

y un indicador global de la fiabilidad es el promedio de la fiabilidad de las escuelas el cual está dado por:

$$\hat{\lambda}_j = \sum \frac{\hat{\lambda}_j}{j} = 0.93$$

Lo cual nos indica que la media muestral tiende a ser un buen estimador de la medida verdadera de la escuela, es decir, que con un nivel de confianza del 93% se concluye que la escuela hace un aporte sobre el resultado del alumno en Lenguaje.

ii) Análisis para el sector privado en Lenguaje:

Modelo nivel 1

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

Modelo nivel 2

$$\beta_{0j} = \beta_0 + \beta_1 W_{01\_Urb\_priv} + u_{0j}$$

Modelo con una sola ecuación.

$$y_{ij\_Lenguaje} = \beta_0 + \beta_1 W_{01\_Urb\_priv} + \left(u_{0j} + e_{0ij}\right)$$

Donde

$y_{ij\_Lenguaje}$ es el rendimiento escolar en Lenguaje del i-ésimo alumno en la j-ésima escuela.

β_0 es la media global de las escuelas.

β_1 es el peso esperado entre el rendimiento de los alumnos del sector urbano privado.

u_{0j} es el error aleatorio para el nivel dos (escuelas).

e_{0ij} es el error aleatorio para el nivel uno (alumnos).

Tabla 3.6: Estimación final de los efectos fijos. (Con errores estándares robustos)

Efectos fijos	**Coeficiente**	**Error estándar**	**T-ratio**	**d.f. aprox**	**p-valor**
Intercepto 1, β_{0j}					
Intercepto 2, β_0	230.343986	3.319902	69.383	53	0.000
Pendiente, β_1	16.043198	5.825011	2.754	53	0.008

d.f. aprox. (grados de libertad)

Tabla 3.7: Estimación final de componentes de varianza.

Efectos aleatorios	**Desviación estándar**	**Componentes de varianza**	**Df**	χ^2 **(chi-cuadrada)**	**p-valor**
u_{0j}	24.27904	589.47161	53	911.19716	0.000
e_{0ij}	39.70914	1576.81562			

$$\mathrm{var}\left(e_{0ij}\right) = \hat{\sigma}^2_{eo} = 1576.82$$

$$\mathrm{var}\left(u_{0j}\right) = \hat{\sigma}^2_{u0} = 589.47$$

$$\mathrm{var}\left(y_{ij} \middle| \beta_0, \beta_1, x_{ij}\right) = \mathrm{var}\left(u_0 + e_{0ij}\right) = \hat{\sigma}^2_{u0} + \hat{\sigma}^2_{e0} = 589.47161+1576.81562=2166.29$$

El coeficiente de correlación intra-clase, representa en este modelo la proporción de varianza de "y" entre escuelas, el cual es estimado de la manera siguiente:

$$\hat{\rho} = \frac{\hat{\sigma}^2_{u0}}{\hat{\sigma}^2_{u0} + \hat{\sigma}^2_{e0}} = \frac{589.47161}{589.47161+1576.81562} = 0.2721$$

De la variación total del resultado escolar en Lenguaje es 27.21% y se debe a las diferencias entre las escuelas, o correlación intra-escuelas.

Representa el peso que tienen las características grupales o internas al sistema escolar en la explicación de las variaciones totales de los resultados. Por otro lado, la estimación del nivel 1, alumno (72.79%) representa el peso de los factores externos al sistema escolar.

Y un indicador global de la fiabilidad es el promedio de la fiabilidad de las escuelas el cual está dado por:

$$\hat{\lambda}_j = \sum \frac{\hat{\lambda}_j}{j} = 0.94$$

Lo cual nos indica que la media muestral tiende a ser un buen estimador de la medida verdadera de la escuela.

Tabla 3.8: Resumen para lenguaje.

Nivel	Global	Privado	Público
Escuela	27.4%	27.21%	25.3%
Alumnos	72.6%	72.79%	74.7%
Indicador global de fiabilidad	0.94	0.94	0.93

En la tabla 3.8, se muestra el resumen de los modelos multinivel tanto nulo como para el área geográfica (urbano público y privado). Los resultados reflejan que de la variación global del resultado en Lenguaje el 27.4% se debe a las diferencias entre escuelas. Este porcentaje se interpreta como el peso que tienen las características que son parte de las escuelas o internas al sistema escolar en la explicación de las variaciones totales de los resultados. Por otro lado, la estimación del nivel 1, alumno (72.6%) representa el peso de los factores externos al sistema escolar. Se observa además que las diferencias entre centros privados y públicos son muy significativas. Esto, debido a que los centros urbanos privados logran un nivel de explicación de las variaciones en el resultado en Lenguaje a partir de las características entre el sistema, de 27.21%, mientras que los centros públicos llegan solo a 25.3%. Además, la media muestral tiende a ser un buen estimador de la medida verdadera de la escuela, es decir, con un nivel de confianza del 94% se concluye que la escuela de sector urbano privado hace un aporte sobre el resultado del alumno en Lenguaje. Mientras que para las escuelas del sector urbano público es del 93%.

3.4.4 Modelo Nulo Matemática.

Modelo nivel 1

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

Modelo nivel 2

$$\beta_{0j} = \beta_0 + u_{0j}$$

Modelo con una sola ecuación.

$$y_{ij\_Matemática} = \beta_0 + \left(u_{0j} + e_{0ij}\right)$$

Donde $y_{ij\_Matemática}$ es el rendimiento escolar en matemática del i-ésimo alumno en j-ésima escuela está representada como una función de la gran media β_0 o media de todos los centros y dos errores aleatorios del nivel escuela u_{0j} y nivel estudiante e_{0ij}.

3.4.5 Modelo Multinivel geográfico para Matemática.

i) Análisis para el sector público en Matemática:

Modelo nivel 1

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

Modelo nivel 2

$$\beta_{0j} = \beta_0 + \beta_1 W_{01\_Urb\_pub} + u_{0j}$$

Modelo con una sola ecuación.

$$y_{ij\_Matematica} = \beta_0 + \beta_1 W_{01\_Urb\_pub} + \left(u_{0j} + e_{0ij}\right)$$

Donde

$y_{ij\_Matemática}$ es el rendimiento escolar en Matemática del i-ésimo alumno en la j-ésima escuela.

β_0 es la media global de las escuelas.

β_1 es el peso esperado entre el rendimiento de los alumnos del sector urbano público.

u_{0j} es el error aleatorio para el nivel dos (escuelas).

e_{0ij} es el error aleatorio para el nivel uno (alumnos).

ii) Análisis para el sector privado en Matemática:

Modelo nivel 1

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

Modelo nivel 2

$$\beta_{0j} = \beta_0 + \beta_1 W_{01\_Urb\_priv} + u_{0j}$$

Modelo con una sola ecuación.

$$y_{ij\_Matematica} = \beta_0 + \beta_1 W_{01\_Urb\_priv} + (u_{0j} + e_{0ij})$$

Donde

$y_{ij\_Matemática}$ es el rendimiento escolar en Matemática del i-ésimo alumno en la j-ésima escuela.

β_0 es la media global de las escuelas.

β_1 es el peso esperado entre el rendimiento de los alumnos del sector urbano privado.

u_{0j} es el error aleatorio para el nivel dos (escuelas).

e_{0ij} es el error aleatorio para el nivel uno (alumnos).

En la tabla 3.9, se muestra el resumen de los modelos nulos tanto global como para el área geográfica como urbano privado y público. Los resultados reflejan que de la variación total del resultado en matemática el 31.25% se debe a las diferencias entre escuelas, o dicho de otra manera correlación intra-escuelas. Este porcentaje se interpreta como el peso que tienen las características que son parte de las escuelas o internas al sistema escolar en la explicación de las variaciones totales de los resultados.

Por otra parte, la estimación referida al nivel individual (alumno), el 67.75% representa el peso de los factores externos al sistema escolar, es decir, indicadores que son parte del estudiante o familia.

Tabla 3.9: Resumen para Matemática.

Nivel	Global	Privado	Público
Escuela	31.25%	30.86%	29.06%
Alumnos	67.75%	69.14%	70.94%
Indicador global de fiabilidad	0.95	0.95	0.94

Se observa además que las diferencias entre centros privados y públicos son muy significativas. Esto, debido a que los centros urbanos privados logran un nivel de explicación de las variaciones en el resultado en matemática a partir de las características entre el sistema, de 30.86%, mientras que lo centros públicos llegan sólo a 29.06%.

Por tanto, la variación total tanto para privado y público se centra más en el aporte del alumno y aquellos indicadores de la familia.

3.5 La familia y el contexto de la escuela.

En esta sección corresponde evaluar ¿Cuál es el efecto que tienen las características demográficas, socioeconómicas y culturales? referidas a:

i) Alumno individual y a su familia (factores extra-escolares individuales).
ii) Al contexto de la escuela (factores extra-escolares contextuales).

Además de las variables que analizamos en las secciones anteriores (sector geográfico urbano público y privado), analizaremos otro conjunto de variables individuales o familiares del estudiante, a saber:

i) Características del alumno y de su familia.
ii) Involucramiento educativo familiar.
iii) Características contextuales de la escuela (como: indicadores de tamaño y disponibilidad de recursos del centro escolar).

De acuerdo a los objetivos propuestos en esta sección, la secuencia de análisis es la siguiente:

i) Efecto de cada variable individual: aquí se evaluará el efecto que tiene cada una de las variables individuales y familiares del alumno sobre su resultado académico.

Nota técnica: determinación del efecto de las variables individuales. Para efecto del análisis, el objetivo consiste en determinar que variables del nivel individual y grupal resultan ser estadísticamente significativas. Para fin de ejemplificación, se toman el resultado en Lenguaje (variable criterio) y la educación de los padre (factor), la operación para lograr el objetivo consiste en controlar este factor ($X_{ij\_pared\_1}$) y analizar los efectos que tiene sobre el resultado del alumno en lenguaje. Para ello, se introduce la variable en la parte fija del modelo "One-Way" (ver sección 3.4.1) y se estiman nuevamente los parámetros. La operación se expresa así:

Modelo nivel 1

$$y_{ij\_Lenguaje} = \beta_{0j} + \beta_{1j} X_{ij\_pared\_1} + e_{0ij}$$

Modelo nivel 2

$$\beta_{0j} = \beta_{00} + u_{0j}$$
$$\beta_{1j} = \beta_{10}$$

Donde $y_{ij\_Lenguaje}$ es el rendimiento del alumno i en la escuela j; β_{0j} es el rendimiento promedio estimado (parte fija); u_{0j} y e_{0ij} son los residuos del nivel 1 y 2, respectivamente, y cantidades aleatorias (parte aleatoria), cuyas varianzas σ^2_{e0} y σ^2_{u0}, serán estimadas. El indicador agregado a la parte fija, $X_{ij\_pared\_1}$ indica la educación de los padres del alumno i en la escuela j y β_{1j} es el efecto de la línea promedio que representa la relación entre Lenguaje y educación de los padres, parámetro que también será estimado.

Se supone que la intensidad de la asociación entre educación de los padres y el resultado de Lenguaje es constante para todas las escuelas del sistema. Esto es así, porque en la parte fija del modelo se estiman:

i) El valor promedio de Lenguaje, ajustado por la educación de los padres, éste valor está indicado por el punto donde la línea que representa la relación entre ambas variables corta a la ordenada Lenguaje.

ii) El valor promedio de la fuerza de la relación entre lenguaje y educación de los padres, es decir, la pendiente de la línea que representa la relación entre ambas variables.

Tabla 3.10: Efecto de las características del alumno, familia y contexto de la escuela para Matemática y Lenguaje

Indicador	Lenguaje	Matemática
Educación de los padres, $X_{ij\_pared\_1}$	0.419203 (prob $\leq$ 0.294) (ns)	0.521692 (ns)
Atención del padre hacia el hijo, X_{ij_home34}	-5.390128 (*)	-1.563281 (ns)
Material didáctico en la casa, $X_{ij\_tenbooks}$	-4.713633 (*)	-0.629336 (ns)
El padre le lee al niño a menudo, $X_{ij\_radlots}$	6.562683 (ns)	1.833544 (ns)
El padre le lee al niño a veces, $X_{ij\_readsome}$	6.813717 (*)	6.502449 (*)
Padre ayuda al hijo en la preparación académica, $X_{ij\_help\_hw}$	-2.380996 (ns)	-1.819584 (ns)
Relación alumno-profesor, $X_{ij\_ptr\_all}$	2.274139 (ns)	1.506581 (ns)
Material didáctico, $W_{0k\_material}$	3.010718 (*)	1.539578 (ns)
Tamaño de biblioteca, $W_{0k\_book1000}$	4.029063 (ns)	-5.899542 (ns)
Profesor en servicio, X_{ij_inserv}	-1.778224 (*)	-0.838229 (ns)
La escuela da desayuno libre, $X_{ij\_brekfast}$	3.345463 (ns)	4.739660 (ns)
La escuela da comida libre, $X_{ij\_freemeal}$	3.345463 (ns)	4.739660 (ns)

(*) prob $\leq$ 0.05; ns: no significativo

En la tabla 3.10, se muestra el efecto individual de las características del alumno, familia y contexto de la escuela para Lenguaje y Matemática. Se observa que la mayoría de los efectos que son parte de la escuela resultaron ser no significativos tales como: la escuela da comida libre, el tamaño de la biblioteca; además, características propias de la familia como: educación de los padres, el padre le lee al niño a veces y cuando el padre ayuda al hijo en la preparación académica. Indicando que el mayor aporte al resultado reside en aquellas características que son propias de la familia y del alumno.

Por otra parte, la mayor parte de los indicadores que son parte de la escuela resultaron ser no significativos sobre el resultado de Matemática, aunque los valores de los coeficientes tanto para Lenguaje como Matemática coinciden en la variable relación alumno y profesor. Concluyendo que gran parte de estos indicadores referidos a la escuela se van a considerar importantes al momento de buscar el modelo multinivel óptimo, donde incluya características de la cultura institucional y práctica en el aula. [10]

ii) Modelo multinivel: efecto de las características del alumno, familia y contexto de la escuela para Lenguaje y Matemática.

Esta secuencia permitirá conocer el peso relativo de los subconjuntos de variables que resultaron ser significativos, a saber: características académicas y socioeconómicas de los estudiantes, las percepciones y actitudes de la familia y la composición institucional (contexto de la escuela), además se agregaran variables demográficas y el grado del alumno.

A continuación se muestra el modelo A para lenguaje, agregando y desagregando en los diferentes niveles, aquellos indicadores que resultaron estadísticamente significativas.

Modelo Nivel 1:

$$y_{ij\_lenguaje} = \beta_{0j} + \beta_{1j} X_{ij\_female} + \beta_{2j} X_{ij\_tenbooks} + \beta_{3j} X_{ij\_pared\_1} + \beta_{4j} X_{ij\_grado} + e_{0ij}$$

Modelo Nivel 2:

$$\beta_{0j} = \beta_{00} + \beta_{01} W_{01\_urb\_pub} + \beta_{02} W_{02\_urb\_pr} + \beta_{03} W_{03\_meanses} + \beta_{04} W_{04\_ptr\_all} + u_{0j}$$

---

10 Se contrastó la hipótesis siguiente:

$$H_0 : \beta_{1j} = 0$$

$$H_1 : \beta_{1j} \neq 0$$

Se acepta la hipótesis nula (no estadísticamente significativo) cuando el p-valor sea mayor que 0.05 caso contrario se rechazará la hipotes nula.

$\beta_{1j} = \beta_{10}$
$\beta_{2j} = \beta_{20}$
$\beta_{3j} = \beta_{30}$
$\beta_{4j} = \beta_{40} + u_{4j}$

donde

$y_{ij\_lenguaje}$: es el resultado de Lenguaje del alumno i-ésimo en la escuela j-ésima.

β_{00}: es el promedio del rendimiento en Lenguaje de las escuelas a través de la población de escuelas, pero controlado por el centro.

β_{01}: es el efecto de la variable $W_{01\_urb\_pub}$ (urbano público) más una variable aleatoria u_{1j} que captura la variación de los pesos asociados a la característica urbano público a través de escuelas.

β_{02}: es el efecto de la variable $W_{02\_urb\_pr}$ (urbano privado) más una variable aleatoria u_{1j} que captura la variación de los pesos asociados a la característica urbano privado a través de escuelas.

β_{03}: es el efecto de la variable $W_{03\_meanses}$ (nivel socioeconómico) más una variable aleatoria u_{1j} que captura la variación de los pesos asociados al nivel socioeconómico a través de escuelas.

β_{04}: es el efecto de la variable $W_{04\_ptr\_all}$ (relación alumno-profesor) más una variable aleatoria u_{1j} que captura la variación de los pesos asociados a la relación alumno-profesor a través de escuelas.

β_{10}: es el efecto de la variable X_{ij_female} (sexo del alumno) que se mantiene fijo.

β_{20}: es el efecto de la variable $X_{ij\_tenbooks}$ (material didáctico en casa) que se mantiene fijo.

β_{30}: es el efecto de la variable $X_{ij\_pared\_1}$ (educación de los padres) que se mantiene fijo.

β_{40}: es el efecto de la variable X_{ij_grado} (grado del alumno) sobre el resultado en lenguaje.

u_{4j}: es el efecto aleatorio sobre el resultado en lenguaje de la j-ésima escuela una vez controlado por las variables correspondientes.

Tabla 3.11: Modelo multinivel A- Efecto de las características de alumno, familia y contexto de la escuela. Estimación final de los efectos fijos para Lenguaje.

Efectos fijos	Coeficiente	Error estándar	T-ratio	d.f. aprox	p-valor
Para el intercepto1 β_{0j} Intercepto 2, β_{00}	220.088050	4.389207	50.143	50	0.000 (*)
$W_{01\_urb\_pub}$, β_{01}	18.125244	8.160932	2.221	50	0.031 (*)
$W_{02\_urb\_pr}$, β_{02}	8.596398	12.016116	0.715	50	0.478 (ns)
$W_{03\_meanses}$, β_{03}	20.219968	6.477875	3.121	50	0.003 (*)
$W_{04\_ptr\_all}$, β_{04}	0.288914	0.249357	1.159	50	0.253 (ns)
Para el efecto X_{ij_female}, β_{1j} Intercepto β_{10}	4.726493	1.725074	2.740	2342	0.007 (*)
Para el efecto $X_{ij\_tenbooks}$, β_{2j} Intercepto β_{20}	-3.742453	1.906409	-1.963	2342	0.049 (*)
Para el efecto $X_{ij\_pared\_1}$, β_{3j} Intercepto β_{30}	0.555684	0.406254	1.368	2342	0.172 (ns)
Para el efecto X_{ij_grado} β_{4j} Intercepto β_{40}	15.794619	2.571715	6.142	54	0.000 (*)

(*) prob $\leq$ 0.05; ns: no significativo

En la tabla 3.11, se muestra el resultado final del análisis multinivel. Se observa que para las características relativas al alumno y familia se mantuvieron significativos, a excepción de la educación de los padres ($X_{ij\_pared\_1}$), y relación entre alumno-profesor.

Por otro lado, las características relativas al contexto escolar tal como: de carácter sociodemográfico ($W_{01\_urb\_pub}$ y $W_{02\_urb\_pr}$) solo $W_{01\_urb\_pub}$ (urbano público) tiene promedio significativo que es superior al $W_{02\_urb\_pr}$ (urbano privado), aunque por ser variables indicadoras no se retiraran del modelo; en cuanto al nivel socioeconómico ($W_{03\_meanses}$) se tiene que resultó ser estadísticamente significativo.

Es de resaltar que el efecto de la educación de los padres ($X_{ij\_pared\_1}$) experimenta sólo un breve descenso (ver tabla 3.10) y por tanto, continúa siendo también significativo. Para las características como recursos de casa (material didáctico) resultó ser estadísticamente significativo, es decir, que influye sobre el resultado del estudiante en Lenguaje.

A continuación se muestra el modelo A para Matemática, agregando y desagregando en los diferentes niveles, aquellos indicadores que resultaron estadísticamente significativas.

Modelo Nivel 1:

$$y_{ij\_matematica} = \beta_{0j} + \beta_{1j} X_{ij\_female} + \beta_{2j} X_{ij\_pared\_1} + \beta_{3j} X_{ij\_readsome} + \beta_{4j} X_{ij\_grado} + e_{0ij}$$

Modelo Nivel 2:

$$\beta_{0j} = \beta_{00} + \beta_{01} W_{01\_urb\_pub} + \beta_{02} W_{02\_urb\_pr} + \beta_{03} W_{03\_meanses} + \beta_{04} W_{04\_material} + u_{0j}$$

$$\beta_{1j} = \beta_{10}$$

$$\beta_{2j} = \beta_{20}$$

$$\beta_{3j} = \beta_{30}$$

$$\beta_{4j} = \beta_{40} + u_{4j}$$

donde

$y_{ij\_matemática}$: es el resultado en matemática del alumno i-ésimo en la escuela j-ésima.

β_{00} : es el promedio del rendimiento en lenguaje de los establecimientos a través de la población de establecimientos, pero controlado por el centro.

β_{01} : es el efecto de la variable $W_{01\_urb\_pub}$ (urbano público) más una variable aleatoria u_{1j} que captura la variación de los pesos asociados a la característica urbano público a través de escuelas.

β_{02} : es el efecto de la variable $W_{02\_urb\_pr}$ (urbano privado) más una variable aleatoria u_{1j} que captura la variación de los pesos asociados a la característica urbano privado a través de escuelas.

β_{03} : es el efecto de la variable $W_{03\_meanses}$ (nivel socioeconómico) más una variable aleatoria u_{1j} que captura la variación de los pesos asociados al nivel socioeconómico a través de escuelas.

β_{04} : es el efecto de la variable $W_{04\_material}$ (material didáctico en el establecimiento) más una variable aleatoria u_{1j} que captura la variación de los pesos asociados al material didáctico a través de escuelas.

β_{10} : es el efecto de la variable X_{ij_female} (sexo del alumno) que se mantiene fijo.

β_{20} : es el efecto de la variable $X_{ij\_pared\_1}$ (educación de los padres) que se mantiene fijo.

β_{30} : es el efecto de la variable $X_{ij\_readsome}$ (el padre le lee al niño a veces) que se mantiene fijo.

β_{40} : es el efecto de la variable X_{ij_grado} (grado del alumno) sobre el resultado en lenguaje.

u_{4j} : es el efecto aleatorio sobre el resultado en lenguaje de la j-ésima escuela una vez controlado por las variables correspondientes.

Tabla 3.12: Modelo multinivel A- Efecto de las características del alumno, familia y contexto de la escuela. Estimación final de los efectos fijos para Matemática.

Efectos fijos	Coeficiente	Error estándar	T-ratio	d.f. aprox	p-valor
Para el intercepto1 β_{0j} Intercepto 2, β_{00}	219.512386	3.548862	61.854	50	0.000 (*)
$W_{01\_urb\_pub}$, β_{01}	19.338964	8.246946	2.345	50	0.023 (*)
$W_{02\_urb\_pr}$, β_{02}	12.571496	11.565768	1.087	50	0.283 (ns)
$W_{03\_meanses}$, β_{03}	7.974525	7.403129	1.077	50	0.287 (ns)
$W_{04\_ptr\_all}$, β_{04}	-0.102047	1.867935	-0.055	50	0.957 (ns)
Para el efecto X_{ij_female} β_{1j} Intercepto β_{10}	1.505206	1.368292	1.100	2342	0.272 (ns)
Para el efecto $X_{ij\_pared\_1}$ β_{2j} Intercepto β_{20}	0.531445	0.322295	1.649	2342	0.099 (*)
Para el efecto $X_{ij\_readsome}$ β_{3j} Intercepto β_{30}	6.009107	2.419395	2.484	2342	0.013 (*)
Para el efecto X_{ij_grado} β_{4j} Intercepto β_{40}	10.990327	2.462676	4.463	54	0.000 (*)

(*) prob ≤ 0.05; ns: no significativo

En la tabla 3.12, se observa que de las características pertenecientes al alumno, las que más están influyendo significativamente sobre el resultado de matemática son: grado escolar del alumno (X_{ij_grado}), el padre le lee al niño a veces ($X_{ij\_readsome}$) y educación de los padres ($X_{ij\_pared\_1}$). Parece ser que la educación de los padres tiene mayor incidencia sobre el logro en Matemática que sobre en Lenguaje (aunque p-valor sea mayor que el nivel de significancia se considera importante). El único indicador que tiene menor incidencia sobre el resultado matemática es el sexo del alumno que sobre el resultado en Lenguaje. Por otro lado, en este modelo A, el efecto de las características referidas al contexto de la escuela no tienen incidencia significativa sobre el resultado, pero si, para el sector urbano público. Se observa además que el nivel socioeconómico de la escuela no influye mucho sobre el resultado en Matemática para este modelo, pero dada la importancia en investigaciones la consideramos para encontrar el modelo siguiente.

3.6 Cultura institucional y práctica en el aula.

El objetivo de la presente sección es identificar cuales de las variables pertenecientes a los niveles de agregación demuestran asociación importante con el logro de los estudiantes sobre el resultado en Lenguaje y Matemática, y por tanto, poder inferir consecuencias para las políticas del sistema educativo.

Para determinar este modelo, se introducen aquellas variables pertenecientes a la cultura institucional e indicadores del alumno que fueron agregados a nivel de alumnos que resultaron significativas (ver sección 3.5) como:

i) Indicadores sociodemográficos y clima en el aula.
ii) Efecto contexto de la escuela

A continuación se muestra el modelo B para lenguaje, agregando y desagregando en los diferentes niveles, aquellos indicadores que resultaron estadísticamente significativas.

Modelo Nivel 1:

$$y_{ij\_lenguaje} = \beta_{0j} + \beta_{1j} X_{ij\_female} + \beta_{2j} X_{ij\_tenbooks} + \beta_{3j} X_{ij\_grado} + e_{0ij}$$

Modelo Nivel 2:

$$\beta_{0j} = \beta_{00} + \beta_{01} W_{01\_urb\_pub} + \beta_{02} W_{02\_urb\_pr} + \beta_{03} W_{03\_meanses} + \beta_{04} W_{04\_mndiscip} + u_{0j}$$

$$\beta_{1j} = \beta_{10}$$

$$\beta_{2j} = \beta_{20}$$

$$\beta_{3j} = \beta_{30} + u_{3j}$$

donde

$y_{ij\_lenguaje}$: es el resultado de lenguaje del alumno i-ésimo en la escuela j-ésima.

β_{00}: es el promedio del rendimiento en lenguaje de los establecimientos a través de la población de establecimientos, pero controlado por el centro.

β_{01}: es el efecto de la variable $W_{01\_urb\_pub}$ (urbano público) más una variable aleatoria u_{1j} que captura la variación de los pesos asociados a la característica urbano público a través de escuelas.

β_{02}: es el efecto de la variable $W_{02\_urb\_pr}$ (urbano privado) más una variable aleatoria u_{1j} que captura la variación de los pesos asociados a la característica urbano privado a través de escuelas.

β_{03}: es el efecto de la variable $W_{03\_meanses}$ (nivel socioeconómico de la escuela) más una variable aleatoria u_{1j} que captura la variación de los pesos asociados al nivel socioeconómico a través de escuelas.

β_{04}: es el efecto de la variable $W_{04\_mndiscip}$ (clima en el aula) más una variable aleatoria u_{1j} que captura la variación de los pesos asociados a la relación alumno-profesor a través de escuelas.

β_{10}: es el efecto de la variable X_{ij_female} (sexo del alumno) que se mantiene fijo.

β_{20}: es el efecto de la variable $X_{ij\_tenbooks}$ (material didáctico en casa) que se mantiene fijo.

β_{30}: es el efecto de la variable X_{ij_grado} (grado del alumno) sobre el resultado en lenguaje.

u_{3j}: es el efecto aleatorio sobre el resultado en lenguaje de la j-ésima escuela una vez controlado por las variables correspondientes.

Tabla 3.13: Modelo multinivel B- Efecto de las características cultura institucional y práctica en el aula. Estimación final de los efectos fijos para Lenguaje.

Efectos fijos	Coeficiente	Error estándar	T-ratio	d.f. aprox	p-valor
Para el intercepto1 β_{0j} Intercepto 2, β_{00}	223.435224	3.648291	61.244	50	0.000 (*)
$W_{01\_urb\_pub}$, β_{01}	17.813535	7.926552	2.247	50	0.029 (*)
$W_{02\_urb\_pr}$, β_{02}	7.050427	11.514963	0.612	50	0.543 (ns)
$W_{03\_meanses}$, β_{03}	15.860972	6.062378	2.616	50	0.012 (*)
$W_{04\_mndiscip}$, β_{04}	41.225457	21.278536	1.937	50	0.058 (ns)
Para el efecto X_{ij_female} β_{1j} Intercepto β_{10}	4.764018	1.724695	2.762	2343	0.006 (*)
Para el efecto $X_{ij\_tenbooks}$ β_{2j} Intercepto β_{20}	-3.456277	1.887391	-1.831	2343	0.067 (ns)
Para el efecto X_{ij_grado} β_{3j} Intercepto β_{30}	15.696049	2.577830	6.089	54	0.000 (*)

(*) prob $\leq$ 0.05; ns: no significativo

En la tabla 3.13, se muestra que la característica perteneciente al alumno y familia tal como la educación de los padres ($X_{ij\_pared\_1}$), no se consideró en esta estimación debido a que en el modelo A resultó ser no significativo. La variable que si consideramos fue material didáctico en la casa del alumno ($X_{ij\_tenbooks}$), esto, para ver su incidencia sobre el resultado. Como observamos, dentro del conjunto de variables que son parte del nivel de agregación escuela, las que no inciden estadísticamente significativa sobre el resultado en Lenguaje son: la ubicación geográfica urbano privado ($W_{02\_urb\_pr}$) y clima en el aula ($W_{04\_mndiscip}$). Parece ser que el clima en el aula tiende a influir más sobre el resultado al igual que el nivel socioeconómico del estudiante.

Por otra parte, para los indicadores que se agregaron al nivel estudiante, la que tuvo menor incidencia a la prueba fue el material didáctica en casa del alumno ($X_{ij\_tenbooks}$), esto, se debe porque la mayoría de características se están extrayendo de la cultura institucional y práctica en el aula.

A continuación se muestra el modelo B para Matemática, agregando y desagregando en los diferentes niveles, aquellos indicadores que resultaron estadísticamente significativas.

Modelo Nivel 1:

$$y_{ij\_matematica} = \beta_{0j} + \beta_{1j} X_{ij\_female} + \beta_{2j} X_{ij\_readsome} + \beta_{3j} X_{ij\_grado} + e_{0ij}$$

Modelo Nivel 2:

$$\beta_{0j} = \beta_{00} + \beta_{01} W_{01\_urb\_pub} + \beta_{02} W_{02\_urb\_pr} + \beta_{03} W_{03\_meanses} + \beta_{04} W_{04\_mndiscip} + u_{0j}$$

$$\beta_{1j} = \beta_{10}$$

$$\beta_{2j} = \beta_{20}$$

$$\beta_{3j} = \beta_{30}$$

$$\beta_{4j} = \beta_{40} + u_{4j}$$

donde

$y_{ij\_matemática}$: es el resultado en Matemática del alumno i-ésimo en la escuela j-ésima.

β_{00} : es el promedio del rendimiento en lenguaje de los establecimientos a través de la población de establecimientos, pero controlado por el centro educativo.

β_{01} : es el efecto de la variable $W_{01\_urb\_pub}$ (urbano público) más una variable aleatoria u_{1j} que captura la variación de los pesos asociados a la característica urbano público a través de escuelas.

β_{02} : es el efecto de la variable $W_{02\_urb\_pr}$ (urbano privado) más una variable aleatoria u_{1j} que captura la variación de los pesos asociados a la característica urbano privado a través de escuelas.

β_{03}: es el efecto de la variable $W_{03\_meanses}$ (nivel socioeconómico de la escuela) más una variable aleatoria u_{1j} que captura la variación de los pesos asociados al nivel socioeconómico a través de escuelas.

β_{04}: es el efecto de la variable $W_{04\_mndiscip}$ (clima en el aula) más una variable aleatoria u_{1j} que captura la variación de los pesos asociados a la relación alumno-profesor a través de escuelas.

β_{10}: es el efecto de la variable X_{ij_female} (sexo del alumno) que se mantiene fijo.

β_{20}: es el efecto de la variable $X_{ij\_readsome}$ (el padre le lee al niño a veces) que se mantiene fijo.

β_{30}: es el efecto de la variable X_{ij_grado} (grado del alumno) sobre el resultado en lenguaje.

u_{3j}: es el efecto aleatorio sobre el resultado en lenguaje de la j-ésima escuela una vez controlado por las variables correspondientes.

En la tabla 3.14, observamos que las variables desagregadas al nivel escuela, las que no son estadísticamente significativas sobre el resultado en Matemática son: la ubicación geográfica urbano privado y el nivel socioeconómico del alumno. Es curioso observar que el nivel socioeconómico no influye mucho sobre el resultado, pero, la consideraremos significativa para analizarla cuando determinemos el modelo óptimo para Matemática.

En la parte que corresponde al alumno y familia, se generó un cambio drástico para la característica (el padre le lee al niño a veces); esto, porque en el modelo A resultó ser significativa con p-valor igual a 0.013, y en esta estimación a disminuido ese valor, reflejando que en este modelo hay mayor influencia cuando el padre le lee al niño a veces sobre el resultado de Matemática.

Tabla 3.14: Modelo multinivel B- Efecto de las características cultura institucional y práctica en el aula. Estimación final de los efectos fijos para Matemática.

Efectos fijos	**Coeficiente**	**Error estándar**	**T-ratio**	**d.f. aprox**	**p-valor**
Para el intercepto1 β_{0j} Intercepto 2, β_{00}	222.960023	2.623312	84.992	50	0.000 (*)
$W_{01\_urb\_pub}$, β_{01}	19.469253	8.975149	2.169	50	0.035 (*)
$W_{02\_urb\_pr}$, β_{02}	13.304215	7.079590	1.879	50	0.066 (ns)
$W_{03\_meanses}$, β_{03}	5.853759	5.452635	1.074	50	0.289 (ns)
$W_{04\_mndiscip}$, β_{04}	47.130015	18.568581	2.538	50	0.015 (*)
Para el efecto X_{ij_female} β_{1j} Intercepto β_{10}	1.463841	1.638495	0.893	2343	0.372 (ns)
Para el efecto $X_{ij\_readsome}$ β_{2j} Intercepto β_{20}	6.570093	1.784336	3.682	2343	0.000 (*)
Para el efecto X_{ij_grado} β_{3j} Intercepto β_{40}	10.851335	2.432481	4.461	54	0.000 (*)

(*) prob $\leq$ 0.05; ns: no significativo

Finalmente, la mayoría de los indicadores de este modelo B, demográfico y socioeconómico, ha tenido mayor incidencia sobre el resultado de Lenguaje que Matemática.

3.7 Modelo optimal de factores asociados al rendimiento en Lenguaje y Matemática.

El objetivo de esta sección es determinar el modelo óptimo en Lenguaje y Matemática a partir de los indicadores que en los modelos A y B incidieron de una manera significativa al resultado de Lenguaje y Matemática para contabilizar la variabilidad entre escuelas (ejemplo comprender porque algunas tienes más altos promedios que otras) y evaluar el efecto que tienen dichos indicadores como: ubicación sociodemográfica, el estatus socioeconómico, y el clima en el aula sobre el resultado.

3.7.1 Las variables.

En las secciones anteriores investigamos variables que algunas de ellas no resultaron ser estadísticamente significativas, pero dada la importancia que tienen en estudios de este tipo se consideran en el análisis. Dichas variables las agregamos y desagregamos en los niveles (alumno y escuela), es decir, por subconjunto:

ii) Características sociodemográficas, y clima en el aula (nivel escuela).
iii) Estatus sociocultural (nivel alumno).
iv) Características pertenecientes a la familia (se agrega a nivel alumno).

Parece oportuno mencionar algunos comentarios:

- La inclusión del sexo nos permitirá distinguir que tanto influye el sexo del alumno sobre el resultado de la prueba de Lenguaje y Matemática.

- El indicador de apoyo educativo familiar (el padre le lee al niño a veces) nos servirá para determinar la incidencia que tiene sobre todo la eficacia positiva cuando el alumno no siente que sus problemas de aprendizaje en la escuela se deben a que el padre de familia no ayuda al hijo a realizar tareas escolares o no le toma interés.

3.7.2 Modelo óptimo para Lenguaje.

Hemos hecho el análisis de agregar todas las variables que se demostraron significativos: sexo del alumno, estatus sociocultural, clima en el aula, grado del estudiante y ubicación geográfica.

Modelo Nivel 1:

$$y_{ij\_lenguaje} = \beta_{0j} + \beta_{1j} X_{ij\_female} + \beta_{2j} X_{ij\_ses\_1} + \beta_{3j} X_{ij\_discip} + \beta_{4j} X_{ij\_grado} + e_{0ij}$$

Modelo Nivel 2:

$$\beta_{0j} = \beta_{00} + \beta_{01} W_{01\_urb\_pub} + \beta_{02} W_{02\_urb\_pr} + \beta_{03} W_{03\_meanses} + \beta_{04} W_{04\_mndiscip} + u_{0j}$$

$$\beta_{1j} = \beta_{10}$$

$$\beta_{2j} = \beta_{20}$$

$$\beta_{3j} = \beta_{30}$$

$$\beta_{4j} = \beta_{40} + u_{4j}$$

donde

$y_{ij\_lenguaje}$: es el resultado de lenguaje del alumno i-ésimo en la escuela j-ésima.

β_{00}: es el promedio del rendimiento en Lenguaje de las escuelas a través de la población de escuelas, pero controlado por el centro educativo.

β_{01}: es el efecto de la variable $W_{01\_urb\_pub}$ (urbano público) más una variable aleatoria u_{1j} que captura la variación de los pesos asociados a la característica urbano público a través de escuelas.

β_{02}: es el efecto de la variable $W_{02\_urb\_pr}$ (urbano privado) más una variable aleatoria u_{1j} que captura la variación de los pesos asociados a la característica urbano privado a través de escuelas.

β_{03}: es el efecto de la variable $W_{03\_meanses}$ (nivel socioeconómico de la escuela) más una variable aleatoria u_{1j} que captura la variación de los pesos asociados al nivel socioeconómico a través de escuelas.

β_{04}: es el efecto de la variable $W_{04\_mndiscip}$ (clima en el aula) más una variable aleatoria u_{1j} que captura la variación de los pesos asociados a la relación alumno-profesor a través de escuelas.

β_{10}: es el efecto de la variable X_{ij_female} (sexo del alumno) que se mantiene fijo.

β_{20}: es el efecto de la variable $X_{ij\_ses\_1}$ (estatus sociocultural) que se mantiene fijo.

β_{30}: es el efecto de la variable X_{ij_discip} (clima en el aula) que se mantiene fijo.

β_{40}: es el efecto de la variable X_{ij_grado} (grado del alumno) sobre el resultado en Lenguaje.

u_{4j}: es el efecto aleatorio sobre el resultado en lenguaje de la j-ésima escuela una vez controlado por las variables correspondientes.

Al incluir estas variables[11] la variación entre-escuelas (ver tabla 3.16) ha caído para el caso Lenguaje en 61.6769 puntos, comparando con 594.99027 en el modelo nulo, es decir en 10%. Mientras que para la variación entre-alumnos ha caído en 122.819 puntos.

En la tabla 3.15, observamos que ninguna de las variables agregadas tanto para el nivel alumno como el nivel escuela existe un promedio de la frecuencia negativamente relacionado con la media del rendimiento en Lenguaje de las escuelas, es decir, es

[11] La manera tradicional de estimar un modelo lineal jerárquico es a través de un análisis sistemática, en que se parte del modelo simple (que sólo incluye las constantes y los términos aleatorios) para luego ir agregando variables explicativas en ambos niveles (alumno y escuela). En general, el criterio para incorporar una o varias variables en el modelo es que su inclusión reduzca la varianza total σ^2_{e0}.

razonable que el involucramiento de factores en el contexto escolar y del alumno no empeora el rendimiento de los estudiantes.

Tabla 3.15: Modelo multinivel óptimo - Estimación final de los efectos fijos para Lenguaje.

Efectos fijos	**Coeficiente**	**Error estándar**	**T-ratio**	**d.f. aprox**	**p-valor**
Para el intercepto1 β_{0j} Intercepto 2, β_{00}	220.841140	3.220584	68.572	50	0.000 (*)
$W_{01\_urb\_pub}$, β_{01}	17.615386	5.595715	3.148	50	0.003 (*)
$W_{02\_urb\_pr}$, β_{02}	7.403279	4.771457	1.552	50	0.127 (ns)
$W_{03\_meanses}$, β_{03}	15.553226	6.023593	2.582	50	0.013 (*)
$W_{04\_mndiscip}$, β_{04}	40.789613	18.606450	2.192	50	0.033 (*)
Para el efecto X_{ij_female} β_{1j} Intercepto β_{10}	5.040774	2.287249	2.204	2342	0.028 (*)
Para el efecto $X_{ij\_ses\_1}$ β_{2j} Intercepto β_{20}	2.192339	2.019762	1.085	2342	0.278 (ns)
Para el efecto X_{ij_discip} β_{3j} Intercepto β_{30}	2.805326	3.507539	0.800	2342	0.424 (ns)
Para el efecto X_{ij_grado} β_{4j} Intercepto β_{40}	15.916197	2.545566	6.253	54	0.000 (*)

(*) prob $\leq$ 0.05; ns: no significativo

Por otro lado, se observa que el involucramiento de los padres no fue incluido en este modelo debido a que en el modelo B tampoco fue considerado, a consecuencia de no ser estadísticamente significativo, cuando fue agregada en conjunto con las variables familia y contexto de la escuela (ver tabla 3.11, Modelo A).

Al analizar la conducta del clima en el aula, se observa que está positivamente relacionada con el promedio del rendimiento en Lenguaje de las escuelas, con un p-valor de 0.33 (lo cual es menor que el nivel de significancia dado en el test estadístico) resulta ser significativo.

Se muestra además que el nivel socioeconómico, tuvo un aumento insignificante del valor de probabilidad comparado con el modelo B. Parece ser que no fue afectado al momento de asociarla con el resto de indicadores.

Con respecto a las variables que son parte del alumno solamente el sexo y grado del alumno tienen incidencia sobre el rendimiento en Lenguaje.

Por tanto, el rendimiento de los estudiantes en Lenguaje se expresa de la siguiente manera:

$$\begin{aligned}\hat{y}_{ij\_lenguaje} = &\left(220.84 + (17.62)W_{01\_urb\_pub} + (15.55)W_{03\_meanses} + (40.79)W_{04\_mndiscip}\right)\\ &+ (5.04)\,X_{ij\_female} + (39.01)\,X_{ij\_grado}\end{aligned}$$

Significa que un cambio de la ubicación urbano público a la urbano privado, hará que aumente en 17.62 sobre el resultado de Lenguaje del estudiante. Parece ser, además que el nivel socioeconómico de la escuela esta influyendo, esto, porque un estudiante al pasar de un bajo nivel socioeconómico a un alto, el resultado del logro del estudiante en Lenguaje aumentará en 15.55 puntos en promedio.

Con base al modelo ajustado, un alumno/a que:

- Estudia en una escuela urbano público.
- Pertenece a una clase socioeconómica alta.
- El clima en el aula sea armonioso.
- El sexo sea femenino.
- El grado sea 4º.

Obtendrá una nota promedio de

$$\hat{y}_{ij\_lenguaje} = 220.84 + (17.62)(1) + (15.55)(0.55) + (40.79)(0.27) + (5.04)(1) + (39.01)(1)$$
$$= 302.08$$

Ahora bien, un estudiante que cumpla las características anteriores excepto: ahora que pertenezca a una clase de extrema pobreza, obtendrá una nota promedio de:

$$\hat{y}_{ij\_lenguaje} = 220.84 + (17.62)(1) + (15.55)(-0.55) + (40.79)(0.27) + (5.04)(1) + (39.01)(1)$$
$$= 284.97$$

En consecuencia hay una disminución de 17.11 puntos en promedio sobre el resultado en Lenguaje.

Tabla 3.16: Estimación final de componentes de varianza para Lenguaje.

Efectos aleatorios	**Desviación estándar**	**Componentes de varianza**	**Df**	χ^2 **(chi-cuadrada)**	**p-valor**
u_{0j}	23.09358	533.31340	50	447.71797	0.000
Pendiente X_{ij_grado}, u_{4j}	14.75238	217.63279	54	139.62557	0.000
e_{0ij}	38.13160	1454.01926			

En la tabla 3.16, se muestra la estimación final de componentes de varianza para Lenguaje. Estas estimaciones indican que la mayor parte de la variación de los resultados reside en los alumnos (1454.02), para las escuelas una variación de 533.31. Por otro parte, la perturbación que recoge las variables no observables en la escuela que inciden en la característica (grado del alumno) hace que el peso de esa variable difiera entre escuelas con una variación de 217.63, es decir, si el peso de la variable grado de los estudiantes fuese el mismo en todas las escuelas, entonces la varianza fuera igual a 0.

Variación explicada en el nivel 1[12]

La varianza explicada en este nivel se calculará comparándola con la del modelo nulo así:

$$\text{var\_exp\_}nivel\text{\_1} = \frac{\left(\hat{\sigma}^2_{e0}(Modelo\,nulo) - \hat{\sigma}^2_{e0}(Modelo\,ajustado)\right)}{\hat{\sigma}^2_{e0}(Modelo\,nulo)}$$

$$\text{var\_exp\_}nivel\text{\_1} = \frac{(1576.83865 - 1454.01926)}{1576.83865} = 0.078$$

Por tanto la proporción de varianza explicada por el modelo del nivel 1 es igual al 7.8%, y significa que un 7.8% de la variabilidad debida al estudiante se explica por el sexo del alumno (X_{ij_female}), el estatus socioeconómico ($X_{ij\_ses\_1}$ **)**, el clima en el aula **(** X_{ij_discip} **)** y el grado al que pertenece el estudiante **(** X_{ij_grado} **)** lo cual es superior al aporte en Lenguaje**.**

Proporción de varianza explicada en β_{0j} [13]

$$propr\text{\_var\_exp\_}\beta_{0j} = \frac{\left(\hat{\sigma}^2_{u0}(Modelo\,coeficiente-aleatorio) - \hat{\sigma}^2_{u0}(Modelo\,ajustado)\right)}{\hat{\sigma}^2_{u0}(Modelo\,coeficiente-aleatorio)}$$

$$propr\text{\_var\_exp\_}\beta_{0j} = \frac{(727.33183 - 533.31340)}{727.33183} = 0.27$$

Significa que un 27% de la variabilidad entre escuelas es explicada por las variables sociodemográficas, el estatus socioeconómico y el clima en el aula.

[12] Para mayor información sobre este concepto ver Anthony S. Bryk, Stephen W Raundenbush. "Hierarchical Linear Models: Applications and Data Analysis Methods". Advanced Quantitative Techniques in the Sciences Series, 1992, pág. 70

[13] Ver corrida completa en Anexo 4.

3.7.3 Modelo óptimo para Matemática.

Hemos hecho el análisis de agregar todas las variables que se demostraron significativos: estatus sociocultural, atención al alumno (el padre le lee a veces al niño), grado del estudiante y ubicación geográfica.

Modelo Nivel 1:

$$y_{ij\_matematica} = \beta_{0j} + \beta_{1j} X_{ij\_ses\_1} + \beta_{2j} X_{ij\_readsome} + \beta_{3j} X_{ij\_discip} + \beta_{4j} X_{ij\_grado} + e_{0ij}$$

Modelo Nivel 2:

$$\beta_{0j} = \beta_{00} + \beta_{01} W_{01\_urb\_pub} + \beta_{02} W_{02\_urb\_pr} + \beta_{03} W_{03\_meanses} + \beta_{04} W_{04\_mndiscip} + u_{0j}$$

$$\beta_{1j} = \beta_{10} + u_{1j}$$

$$\beta_{2j} = \beta_{20} + u_{2j}$$

$$\beta_{3j} = \beta_{30} + u_{3j}$$

$$\beta_{4j} = \beta_{40} + u_{4j}$$

donde

$y_{ij\_matemática}$: es el resultado de lenguaje del alumno i en la escuela j.

β_{00}: es el promedio del rendimiento en lenguaje de los establecimientos a través de la población de establecimientos, pero controlado por el centro.

β_{01}: es el efecto de la variable $W_{01\_urb\_pub}$ (urbano público) más una variable aleatoria u_{1j} que captura la variación de los pesos asociados a la característica urbano público a través de escuelas.

β_{02}: es el efecto de la variable $W_{02\_urb\_pr}$ (urbano privado) más una variable aleatoria u_{1j} que captura la variación de los pesos asociados a la característica urbano privado a través de escuelas.

β_{03}: es el efecto de la variable $W_{03\_meanses}$ (nivel socioeconómico de la escuela) más una variable aleatoria u_{1j} que captura la variación de los pesos asociados al nivel socioeconómico a través de escuelas.

β_{04}: es el efecto de la variable $W_{04\_mndiscip}$ (clima en el aula) más una variable aleatoria u_{1j} que captura la variación de los pesos asociados a la relación alumno-profesor a través de escuelas.

β_{10}: es el efecto de la variable $X_{ij\_ses\_1}$ (estatus sociocultural) que se mantiene fijo.

β_{20}: es el efecto de la variable $X_{ij\_readsome}$ (el padre le lee al niño a veces) que se mantiene fijo.

β_{30}: es el efecto de la variable X_{ij_discip} (clima en aula) que se mantiene fijo.

β_{40}: es el efecto de la variable X_{ij_grado} (grado del alumno) sobre el resultado en lenguaje.

u_{4j}: es el efecto aleatorio sobre el resultado en lenguaje de la j-ésima escuela una vez controlado por las variables correspondientes.

Al incluir estas variables la variación entre-escuelas (ver tabla 3.16) ha caído para el caso Lenguaje en 89.08 puntos, comparando con 452.55 en el modelo nulo, es decir en 20%.

En la tabla 3.17, se muestra los resultados finales de los análisis de las variables demográficas, nivel socioeconómico y clima del aula del rendimiento en Matemática. Observamos que el nivel socioeconómico y clima del aula están positivamente relacionados con el promedio del rendimiento en Matemática de las escuelas, aunque solo el clima en el aula resultó ser estadísticamente significativos.

Por otro lado, donde no hay mucha incidencia sobre el rendimiento, tal es el caso del estatus sociocultural del alumno. Aunque si se puede observar que el grado en el que está el alumno influye estadísticamente sobre el rendimiento en Matemática.

Tabla 3.17: Modelo multinivel óptimo - Estimación final de los efectos fijos para Matemática.

Efectos fijos	Coeficiente	Error estándar	T-ratio	d.f. aprox	p-valor
Para el intercepto1 β_{0j} Intercepto 2, β_{00}	222.854175	2.653179	83.995	50	0.000 (*)
$W_{01\_urb\_pub}$, β_{01}	21.184078	9.231150	2.295	50	0.026 (*)
$W_{02\_urb\_pr}$, β_{02}	14.375516	6.125741	2.347	50	0.023 (*)
$W_{03\_meanses}$, β_{03}	5.284402	5.668667	0.932	50	0.356 (ns)
$W_{04\_mndiscip}$, β_{04}	48.986292	18.008678	2.720	50	0.009 (*)
Para el efecto $X_{ij\_ses\_1}$ β_{1j} Intercepto β_{10}	1.477074	1.673280	0.883	54	0.382 (ns)
Para el efecto $X_{ij\_readsome}$ β_{2j} Intercepto β_{20}	3.247181	1.234127	2.631	54	0.011 (*)
Para el efecto X_{ij_discip} β_{3j} Intercepto β_{30}	2.537821	2.424762	1.047	54	0.300 (ns)
Para el efecto X_{ij_grado} β_{4j} Intercepto β_{40}	10.975756	2.439764	4.499	54	0.000 (*)

(*) prob ≤ 0.05; ns: no significativo

Por tanto, el rendimiento de los estudiantes en Matemática se expresa de la siguiente manera:

$$\hat{y}_{ij\_matemática} = \left(222.85 + (21.18)W_{01\_urb\_pub} + (14.37)W_{02\_urb\_pr} + (48.99)W_{04\_mndiscip}\right) + (10.98)X_{ij\_grado}$$

Significa que los padres de familia al atender al alumno en su casa por medio de la lectura de libros ayuda a que su resultado aumente 14.37 puntos promedio en Matemática. Se observa además que el clima en el aula es verdaderamente importante sobre el rendimiento del alumno, ya que, en un 48.99 puntos promedio aumenta sobre el resultado de Matemática. Por ejemplo:

Con base al modelo ajustado, un alumno/a que:

- Estudia en una escuela urbano público.
- El clima en el aula sea armonioso.
- El grado sea 4º.

Obtendrá una nota promedio de:

$$\begin{aligned}\hat{y}_{ij\_matemática} &= 222.85+(21.18)(1)+(14.37)(0)+(48.99)(0.25)+(10.98)(1)\\ &= 268.24\end{aligned}$$

Ahora bien, un estudiante que cumpla las características anteriores excepto que el clima del aula sea menos armonioso, obtendrá una nota promedio de:

$$\begin{aligned}\hat{y}_{ij\_matemática} &= 222.85+(21.18)(1)+(14.37)(0)+(48.99)(-0.25)+(10.98)(1)\\ &= 241.78\end{aligned}$$

En consecuencia hay una disminución de 26.46 puntos en promedio sobre el resultado de Matemática.

En la tabla 3.18, se muestra la estimación final de componentes de varianza para Matemática. Estas estimaciones indican que la mayor parte de la variación de los resultados reside en los estudiantes (894.87), para las escuelas una variación de 363.47. Por otro parte, la perturbación que recoge las variables no observables en la escuela que inciden en la característica (sexo del alumno) hace que el peso de esa variable difiera entre escuelas con una variación de 53.74, es decir, si el peso de la variable sexo de los estudiantes fuese el mismo en todas las escuelas, entonces la varianza fuera igual a 0. De igual forma, la perturbación que recoge las variables no observables en la escuela determinada por la característica (estatus sociocultural) hace que el peso de esa variable difiera entre escuelas con una variación de 5.55, es decir, si el peso de la variable estatus sociocultural de los estudiantes fuese el mismo en todas las escuelas, entonces la varianza fuera igual a 0, aquí se observa que esta varianza es menor comparada con las demás varianza.

Tabla 3.18: Estimación final de componentes de varianza para Matemática.

Efectos aleatorios	Desviación estándar	Componentes de varianza	Df	χ^2 (chi-cuadrada)	p-valor
u_{0j}	19.06488	363.46950	49	319.00167	0.000
X_{ij_female}, u_{1j}	7.33092	53.74238	53	78.47514	0.013
$X_{ij\_ses\_1}$, u_{2j}	2.35560	5.54887	53	41.88302	>.500
$X_{ij\_discip\_1}$, u_{3j}	2.80811	7.88549	53	48.60653	>.500
X_{ij_grado}, u_{4j}	15.59441	243.18568	53	177.68969	0.000
e_{0ij}	29.91444	894.87386			

Variación explicada en el nivel 1

La varianza explicada en este nivel se calculará comparándola con la del modelo nulo así:

$$\text{var\_exp\_}nivel\_1 = \frac{\left(\hat{\sigma}^2_{e0}(Modelo\,nulo) - \hat{\sigma}^2_{e0}(Modelo\,ajustado)\right)}{\hat{\sigma}^2_{e0}(Modelo\,nulo)}$$

$$\text{var\_exp\_}nivel\_1 = \frac{(991.46760 - 894.87386)}{991.46760} = 0.097$$

Por tanto la proporción de varianza explicada es igual al 9.7%, y significa que un 9.7% de la variabilidad debida al estudiante se explica por el sexo del alumno (X_{ij_female}), el estatus socioeconómico ($X_{ij\_ses\_1}$**)**, el clima en el aula **(**X_{ij_discip}**)** y el grado al que pertenece el estudiante **(**X_{ij_grado}**)** lo cual es superior al aporte en Lenguaje**.**

Proporción de varianza explicada en β_{0j}

$$propr\_\text{var}\_\exp\_\beta_{0j} = \frac{\left(\hat{\sigma}_{u0}^{2}(Modelo\,coeficiente - aleatorio) - \hat{\sigma}_{u0}^{2}(Modelo\,ajustado)\right)}{\hat{\sigma}_{u0}^{2}(Modelo\,coeficiente - aleatorio)}$$

$$propr\_\text{var}\_\exp\_\beta_{0j} = \frac{(475.19861 - 363.46950)}{475.19861} = 0.24$$

Significa que un 24% de la variabilidad entre escuelas es explicada por las variables sociodemográficas, el estatus socioeconómico y el clima en el aula.

3.7.4 La diagnosis del modelo mediante los residuos para el nivel 1.

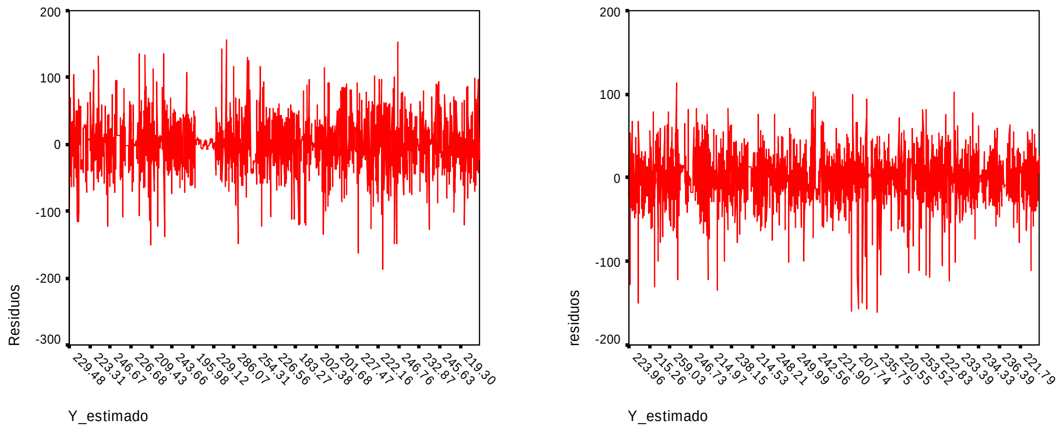
El siguiente proceso tiene como propósito comprobar si las hipótesis que se han utilizado para construir los modelos óptimos para Lenguaje y Matemática no están en contradicción con nuestros datos. Si las hipótesis son adecuadas, podemos utilizar el modelo multinivel para generar predicciones. Si no lo son, habrá que modificar el modelo para adaptarlo a los datos observados. A continuación comentamos las dos hipótesis básicas del modelo:

Homocedasticidad.

Esta hipótesis es importante, esto, debido a que al graficar los residuos frente a valores estimados se puede ver la variabilidad cuando tiende a crecer o a decrecer con las variables predictoras del modelo.

En la figura 3.1, se presenta un gráfico de residuos que muestra un crecimiento lineal de la variabilidad con los valores de los predictores. Observamos que la varianza permanece constante para todos las puntuaciones de los estudiantes en Lenguaje y Matemática, con lo cual se cumple la homocedasticidad.

Figura 3.1: Representación gráfica de la serie de los residuos del nivel 1 para Lenguaje y Matemática.

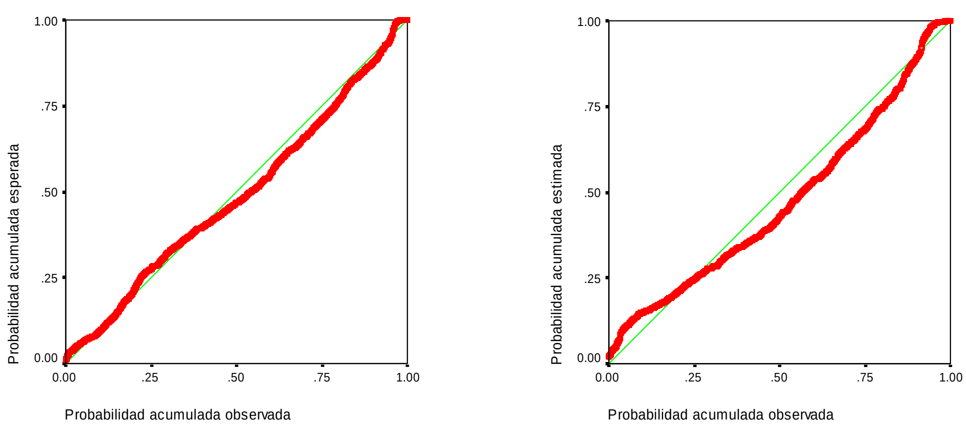


Normalidad.

Esta hipótesis de normalidad es importante para justificar el método de estimación y las distribuciones de los estimadores. La normalidad de los residuos puede contrastarse a partir del gráfico normal de lo residuos, y, si la distribución de los residuos es normal, el gráfico tiene que mostrar aproximadamente una línea recta. Por consiguiente también se cumple el supuesto de normalidad para el modelo óptimo tanto Lenguaje y Matemática

La figura 3.2, muestra que, los residuos están aproximadamente en una linea recta, por lo que las perturbaciones parecen normales tanto para Lenguaje como para Matemática.

Figura 3.2: Gráfico de probabilidad normal de los residuos del nivel 1 para Lenguaje y Matemática.



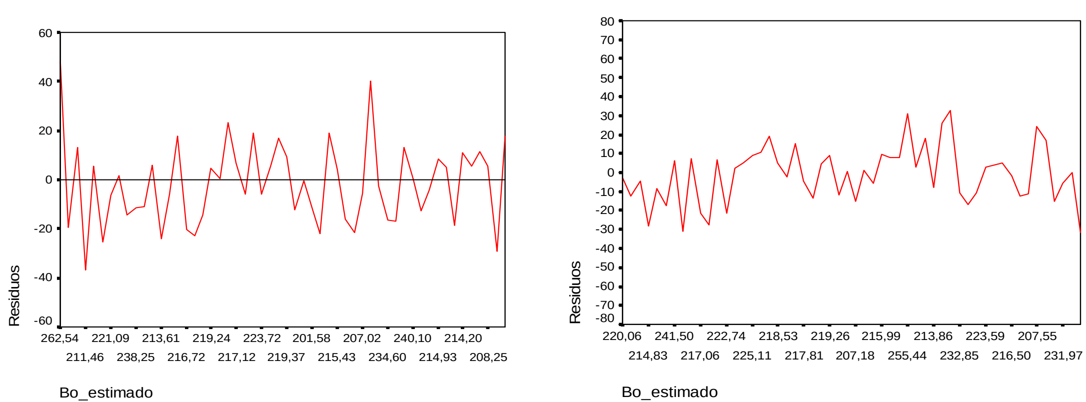
3.7.5 La diagnosis del modelo mediante los residuos nivel 2 para Lenguaje y Matemática.

De igual manera, así como se hizo el diagnóstico de los residuos en el nivel uno, lo haremos para los residuos del nivel 2 del modelo óptimo para Lenguaje y Matemática.

Homocedasticidad.

En la figura 3.3, se presenta el gráfico de los residuos del nivel 2 frente a las estimaciones del intercepto (promedio de los centros). El gráfico tanto para Lenguaje (izquierda) y Matemática (derecha) muestra que existe linealidad en los datos. Observando, ademas, que los valores de las estimaciones del intercepto están alrededor de la media de los datos y que la variabilidad permanece constante a través de todas las estimaciones. Por tanto, se cumple el supuesto de homocedasticidad para el modelo óptimo de Lenguaje y Matemática.

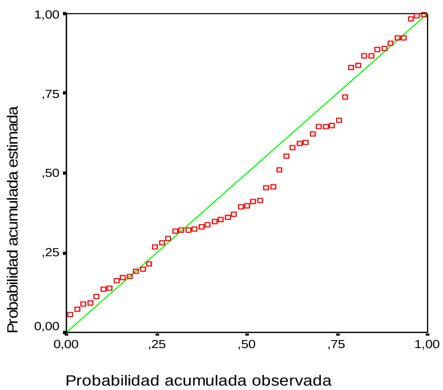
Figura 3.3: Representación gráfica de la serie de los residuos del nivel 2 para Lenguaje y Matemática.

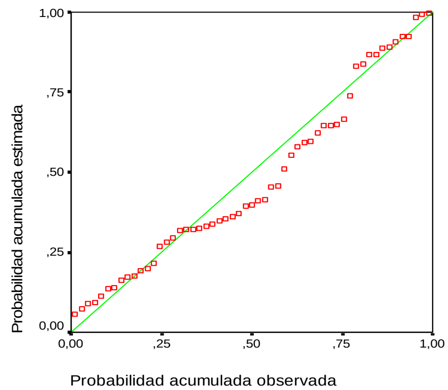


Normalidad

En la figura 3.4, se presenta una gráfica idealizada de probabilidad normal de los residuos del nivel dos para Lenguaje y Matemática. Se observa que los puntos caen aproximadamente sobre la línea recta, lo cual se podría suponer normalidad. Por consiguiente también se cumple el supuesto de normalidad para el modelo óptimo tanto Lenguaje y Matemática.

Figura 3.4: Gráfico de probabilidad normal de los residuos del nivel 2 para Lenguaje y Matemática.





Por tanto, una vez hecho la diagnosis del modelo, concluimos que el modelo óptimo para Lenguaje y Matemática es válido para hacer predicciones y/o tomar decisiones.

CAPÍTULO IV

CONCLUSIONES

4.1 Conclusiones para Lenguaje.

- Al aplicar el análisis multinivel jerárquico y examinar cómo se distribuyen las variaciones en el resultado escolar entre factores internos y externos al sistema escolar, se constató que el sistema escolar examinado responde por el 27.4% de la variación en el resultado en Lenguaje.

- La ubicación urbano publico de los centros educativos, el nivel socioeconómico y el clima en el aula predicen significativamente el resultado en Lenguaje entre escuelas.

- Al efectuar el análisis comparando el comportamiento de lo centros pertenecientes al urbano público y privado, en Lenguaje las diferencias son muy significativas, respondiendo en mayor proporción al resultado a partir de las características entre el sistema escolar, 27.21%, mientras que los centros urbanos públicos responden en forma baja por las variaciones de los resultados escolares.

- Al aplicar un modelo optimal de factores asociados ajustando según sexo de los alumnos, estatus sociocultural, clima en el aula, grado escolar, ubicación geográfica y nivel socioeconómico. Las variaciones del resultado entre escuelas se reduce en aproximadamente 8% (porcentaje de varianza explicada en el nivel 1) para Lenguaje, es decir, que un 8% de la variabilidad debida al estudiante se explica por el sexo del alumno, estatus socioeconómico, clima en el aula y grado escolar al que pertenece.

- La magnitud del aporte al rendimiento en Lenguaje y a los efectos de las variaciones del alumno y familia, usando la ubicación geográfica, el estatus sociocultural, y el clima en el aula como predictores es del 10% de la variabilidad entre centros. Parece ser, que estos aspectos como la ubicación urbano público, nivel socioeconómico, clima en el aula, sexo y grado del estudiante tienen un peso importante en el rendimiento de los estudiantes en Lenguaje.

- Por tanto, el rendimiento de los estudiantes en lenguaje se expresa de la siguiente manera:

$$\hat{y}_{ij\_lenguaje} = \left(220.84 + (17.62) W_{01\_urb\_pub} + (15.55) W_{03\_meanses} + (40.79) W_{04\_mndiscip}\right)$$
$$+ (5.04) X_{ij\_female} + (39.01) X_{ij\_grado}$$

Con base al modelo óptimo ajustado para Lenguaje, un alumno/a que:

- Estudia en una escuela urbano público ($W_{01\_urb\_pub}$).
- Pertenece a una clase socioeconómica baja ($W_{03\_meanses}$).
- El clima en el aula sea armonioso ($W_{04\_mndiscip}$).
- El sexo sea femenino (X_{ij_female}).
- El grado sea 4º (X_{ij_grado}).

Obtendrá una nota promedio de

$$\hat{y}_{ij\_lenguaje} = 220.84 + (17.62)(1) + (15.55)(-0.25) + (40.79)(0.27) + (5.04)(1) + (39.01)(1)$$
$$= 289.64$$

Ahora bien, un estudiante que cumpla las características anteriores excepto que su nivel socioeconómico sea alto y que el clima en el aula no sea armonioso, obtendrá una nota promedio de:

$$\hat{y}_{ij\_lenguaje} = 220.84 + (17.62)(1) + (15.55)(0.55) + (40.79)(-0.27) + (5.04)(1) + (39.01)(1)$$
$$= 280.04$$

En consecuencia hay una disminución de 9.6 puntos en promedio sobre el resultado en Lenguaje.

4.2 Conclusiones para Matemática.

- Las escuelas del país centroamericano considerado en este estudio varían significativamente en sus logros promedios de Matemática, en un 31.25% de la variación total y resultado superior al obtenido en Lenguaje. Este resultado refleja que los padres son menos competentes para apoyar a sus hijos en Matemática que en Lenguaje por esa razón la escuela se involucra mayormente en los resultados, que con Lenguaje.

- Al efectuar el análisis comparando el comportamiento de los centros urbanos públicos y privados, en Matemática las diferencias son muy significativas, respondiendo en mayor proporción al resultado a partir de las características entre el sistema escolar, 30.86%, mientras que los centros urbanos público responden en forma baja por las variaciones de los resultados escolares.

- Al aplicar un modelo optimal de factores asociados ajustado según estatus sociocultural, clima en el aula, grado escolar, ubicación geográfica, nivel socioeconómico y atención de los padres al hijo. Las variaciones del resultado entre escuelas se reduce aproximadamente en un 20% para Matemática.

- Al aplicar el análisis multinivel jerárquico y examinar como se distribuyen las variaciones en el resultado entre factores internos y externos al sistema escolar, se llegó a tener que el 24% de la variabilidad entre escuelas es explicada por las variables sociodemográficas, el estatus socioeconómico y el clima en el aula.

- Por tanto, el rendimiento de los estudiantes en Matemática se expresa de la siguiente manera:

$$\hat{y}_{ij\_matemática} = \left(222.85 + (21.18)W_{01\_urb\_pub} + (14.37)W_{02\_urb\_pr} + (48.99)W_{04\_mndiscip}\right) + (10.98)X_{ij\_grado}$$

Con base al modelo óptimo ajustado para Matemática, un alumno/a que:

- Estudia en una escuela urbano público ($W_{01\_urb\_pub}$).
- El clima en el aula no sea armonioso ($W_{04\_mndiscip}$).
- El grado sea 4º (X_{ij_grado}).

Obtendrá una nota promedio de:

$$\begin{aligned}\hat{y}_{ij\_matemática} &= 222.85+(21.18)(1)+(14.37)(0)+(48.99)(-0.29)+(10.98)(1)\\ &= 240.80\end{aligned}$$

Ahora bien, un estudiante que cumpla las características anteriores excepto que el clima del aula sea armonioso y que estudie en una escuela urbano privado, obtendrá una nota promedio de:

$$\begin{aligned}\hat{y}_{ij\_matemática} &= 222.85+(21.18)(0)+(14.37)(1)+(48.99)(0.29)+(10.98)(1)\\ &= 262.41\end{aligned}$$

En consecuencia hay un aumento de 21.61 puntos en promedio sobre el resultado de Matemática.

REFERENCIAS BIBLIOGRÁFICAS

- A. Naderiand J. Mace (2002). Education and Earnings; A Multilevel Analysis A Case Study of the Manufacturing Sector in Iran. Management and Planning Organisation, Tehran, Iran. University of London, London, U.K.

- Browne, W.J. And J. Rasbash, Multilevel Modelling (1999). Institute of Education, University of London.

- Cervini, Rubén (2002). Desigualdades Socioculturales en el Aprendizaje de Matemática y Lengua de la Educación Secundaria en Argentina.

- Ferrão. Leite. Beltrão (2001). Introdução à modelagem multinível em Avaliação Educacional. Ministério do Planejamento, Orçamento e Gestão Instituto Brasileiro de Geografia e Estatística – IBGE Escola Nacional de Ciências Estatísticas. Rio de Janeiro.

- Montgomery. Peck. Vining (2002). Introducción al Análisis de Regresión Lineal. Primera edición Mexico.

- Goldstein, Harvey. Rasbash. Yang. Geoffrey, Woodhouse (1993). Sally Thomas. A multilevel analysis of school examination results. Oxford review of education, vol 19, No. 4.

- Goldstein. Tutorial in biostatistics Multilevel modelling of medical data. Institute of Education; University of London; London; U.K.

- Goldstein, H. (1997). Methods in school effectiveness research. School effectiveness and school improvement.

- Gujarati, Damodar N. Econometría Básica. Tercera edición, United States Military Academy, West Point, 1997.

- Huáscar Taborga 1980. Como hacer una tesis. Buenos Aires.

- Jean-Paul Fox and Cees A.W. Glas. Modeling Measurement Error in Structural Multilevel Models. University of Twente, The Netherlands

- J.J. Hox. 1995. Applied Multilevel Análisis. TT-Publikaties, Amsterdam.

- Jöreskog. Sörbom. Toit. LISREL 8. New Statistical Features. 2000.

- Murillo, F.J. (1999). Los Modelos Jerárquicos Lineales aplicados a la Investigación sobre Eficacia Escolar. Revista de Investigación Educativa.

- Oliver Juan C, Jesús Rosel y Pilar Jara. Modelos de regresión multinivel: aplicación en psicología escolar. Universitat Jaume I.

- Peña, Daniel. Analisis de Datos Multivariantes.

- Peña, Daniel. Regresión y diseño de experimentos. Alianza editorial, Madrid. 2002.

- Carlos Roberto Briones et al., Factores asociados al rendimiento de los estudiantes que se sometieron a la PAES 2000. Pág. 84-85. Ministerio de Educación, Dirección Nacional de Monitoreo y Evaluación. El Salvador.

ANEXOS

ANEXO 1: Resultados del modelo multinivel familia y contexto de la escuela-Lenguaje.

```
Program:                HLM 6 Hierarchical Linear and Nonlinear Modeling
Authors:                Stephen Raudenbush, Tony Bryk, & Richard Congdon
Publisher:              Scientific Software International, Inc. (c) 2000
                                  techsupport@ssicentral.com
                                       www.ssicentral.com
-------------------------------------------------------------------------------
Module:     HLM2S.EXE (6.00.24282.2)
Date:       19 August 2005, Friday
Time:        11:22:32
-------------------------------------------------------------------------------


SPECIFICATIONS FOR THIS HLM2 RUN


 Problem Title: MODELO A FINAL FAMILIA Y CONTEXTO DE LA ESCUELA

 The data source for this run  = C:\Mis documentos\TESIS DE
WELMAN\MODELOS\FAMILIA Y EL CONTEXTO DE LA ESCUELA\logros.mdm
 The command file for this run = C:\ MODELOS\FAMILIA Y EL CONTEXTO DE LA
ESCUELA\MODEL FINAL FAMILIA Y CONTEXTO DE ESCUELA
LENGUAJE.hlm
 Output file name              = C:\\MODELOS\FAMILIA Y EL CONTEXTO DE LA
ESCUELA\MODELO FINAL FAMILIA Y CONTEXTO ESCUELA_LENGUAJE.out
 The maximum number of level-1 units = 2351
 The maximum number of level-2 units = 55
 The maximum number of iterations = 100
 Method of estimation: restricted maximum likelihood

Weighting Specification
-----------------------
                  Weight
                  Variable
          Weighting?   Name        Normalized?
Level 1        no
Level 2        no
Precision      no
```

The outcome variable is   LANGR

The model specified for the fixed effects was:

```
   Level-1                Level-2
   Coefficients           Predictors
 ----------------------   ---------------
       INTRCPT1, B0      INTRCPT2, G00
$                        URB_PUB, G01
$                         URB_PR, G02
$                        MEANSES, G03
$                        PTR_ALL, G04
#    FEMALE slope, B1      INTRCPT2, G10
#  TENBOOKS slope, B2      INTRCPT2, G20
#   PARED_1 slope, B3      INTRCPT2, G30
      MUDA slope, B4      INTRCPT2, G40
```

'#' - The residual parameter variance for this level-1 coefficient has been set
  to zero.
'$' - This level-2 predictor has been centered around its grand mean.

The model specified for the covariance components was:

Sigma squared (constant across level-2 units)

Tau dimensions
    INTRCPT1
        MUDA slope

Summary of the model specified (in equation format)

Level-1 Model

Y = B0 + B1*(FEMALE) + B2*(TENBOOKS) + B3*(PARED_1) + B4*(MUDA) + R

Level-2 Model
B0 = G00 + G01*(URB_PUB) + G02*(URB_PR) + G03*(MEANSES) + G04*(PTR_ALL) + U0
B1 = G10

B2 = G20
B3 = G30
B4 = G40 + U4

Iterations stopped due to small change in likelihood function
******* ITERATION 11 *******

Sigma_squared =   1452.21195

Tau
INTRCPT1,B0   538.38335   -177.06049
MUDA,B4   -177.06049   217.51319

Tau (as correlations)
INTRCPT1,B0 1.000 -0.517
MUDA,B4 -0.517 1.000

Random level-1 coefficient	Reliability estimate
INTRCPT1, B0	0.881
MUDA, B4	0.599

The value of the likelihood function at iteration 11 = -1.196958E+004

The outcome variable is   LANGR

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard Error	T-ratio	Approx. d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	220.088050	4.389207	50.143	50	0.000
URB_PUB, G01	18.125244	8.160932	2.221	50	0.031
URB_PR, G02	8.596398	12.016116	0.715	50	0.478
MEANSES, G03	20.219968	6.477875	3.121	50	0.003
PTR_ALL, G04	0.288914	0.249357	1.159	50	0.253
For FEMALE slope, B1					
INTRCPT2, G10	4.726493	1.725074	2.740	2342	0.007
For TENBOOKS slope, B2					
INTRCPT2, G20	-3.742453	1.906409	-1.963	2342	0.049
For PARED_1 slope, B3					
INTRCPT2, G30	0.555684	0.406254	1.368	2342	0.172

For MUDA slope, B4

INTRCPT2, G40	15.794619	2.571715	6.142	54	0.000

The outcome variable is LANGR

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Coefficient	Standard Error	T-ratio	Approx. d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	220.088050	4.076711	53.987	50	0.000
URB_PUB, G01	18.125244	6.444183	2.813	50	0.007
URB_PR, G02	8.596398	5.297601	1.623	50	0.111
MEANSES, G03	20.219968	6.018700	3.360	50	0.002
PTR_ALL, G04	0.288914	0.183253	1.577	50	0.121
For FEMALE slope, B1					
INTRCPT2, G10	4.726493	2.240159	2.110	2342	0.035
For TENBOOKS slope, B2					
INTRCPT2, G20	-3.742453	2.041979	-1.833	2342	0.067
For PARED_1 slope, B3					
INTRCPT2, G30	0.555684	0.373714	1.487	2342	0.137
For MUDA slope, B4					
INTRCPT2, G40	15.794619	2.562451	6.164	54	0.000

Final estimation of variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, U0	23.20309	538.38335	50	452.83022	0.000
MUDA slope, U4	14.74833	217.51319	54	139.56196	0.000
level-1, R	38.10790	1452.21195			

Statistics for current covariance components model

Deviance = 23939.167009
Number of estimated parameters = 4

ANEXO 2: Resultados del modelo multinivel cultura institucional y práctica en el aula-Lenguaje y Matemática.

```
Program:              HLM 6 Hierarchical Linear and Nonlinear Modeling
Authors:              Stephen Raudenbush, Tony Bryk, & Richard Congdon
Publisher:            Scientific Software International, Inc. (c) 2000
                                techsupport@ssicentral.com
                                     www.ssicentral.com
--------------------------------------------------------------------------------
Module:    HLM2S.EXE (6.00.24282.2)
Date:      19 August 2005, Friday
Time:       14:16:38
--------------------------------------------------------------------------------


SPECIFICATIONS FOR THIS HLM2 RUN


 Problem Title: MODELO FINAL B CULTURA INSTITUCIONAL Y CONTEXTO
DEL CENTRO ESCO

 The data source for this run  = logros.mdm
 The command file for this run = C:\ \MODELOS\CULTURA INSTITUCIONAL Y
PRACTICA EN EL AULA\Modelo final B para lenguaje.hlm
 Output file name             = C:\ MODELOS\CULTURA INSTITUCIONAL Y
PRACTICA EN EL AULA\Modelo final B para lenguaje.out
 The maximum number of level-1 units = 2351
 The maximum number of level-2 units = 55
 The maximum number of iterations = 100
 Method of estimation: restricted maximum likelihood

Weighting Specification
-----------------------
                  Weight
                  Variable
        Weighting?   Name       Normalized?
Level 1       no
Level 2       no
Precision      no

 The outcome variable is    LANGR

 The model specified for the fixed effects was:
----------------------------------------------------
```

```
  Level-1                Level-2
  Coefficients           Predictors
 ----------------------  ---------------
       INTRCPT1, B0      INTRCPT2, G00
$                        URB_PUB, G01
$                         URB_PR, G02
$                        MEANSES, G03
$                       MNDISCIP, G04
#    FEMALE slope, B1      INTRCPT2, G10
#  TENBOOKS slope, B2      INTRCPT2, G20
      MUDA slope, B3      INTRCPT2, G30

'#' - The residual parameter variance for this level-1 coefficient has been set
     to zero.
'$' - This level-2 predictor has been centered around its grand mean.

 The model specified for the covariance components was:
 ---------------------------------------------------------

        Sigma squared (constant across level-2 units)

        Tau dimensions
             INTRCPT1
                MUDA slope


 Summary of the model specified (in equation format)
 ---------------------------------------------------

Level-1 Model

          Y = B0 + B1*(FEMALE) + B2*(TENBOOKS) + B3*(MUDA) + R

Level-2 Model
          B0 = G00 + G01*(URB_PUB) + G02*(URB_PR) + G03*(MEANSES) +
G04*(MNDISCIP) + U0
          B1 = G10
          B2 = G20
          B3 = G30 + U3

Iterations stopped due to small change in likelihood function
```

******* ITERATION 8 *******

Sigma_squared = 1452.86148

Tau

INTRCPT1,B0	529.85542	-188.55399
MUDA,B3	-188.55399	219.59376

Tau (as correlations)

INTRCPT1,B0	1.000	-0.553
MUDA,B3	-0.553	1.000

Random level-1 coefficient	Reliability estimate
INTRCPT1, B0	0.879
MUDA, B3	0.602

The value of the likelihood function at iteration 8 = -1.196400E+004
The outcome variable is LANGR

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard Error	T-ratio	Approx. d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	223.435224	3.648291	61.244	50	0.000
URB_PUB, G01	17.813535	7.926552	2.247	50	0.029
URB_PR, G02	7.050427	11.514963	0.612	50	0.543
MEANSES, G03	15.860972	6.062378	2.616	50	0.012
MNDISCIP, G04	41.225457	21.278536	1.937	50	0.058
For FEMALE slope, B1					
INTRCPT2, G10	4.764018	1.724695	2.762	2343	0.006
For TENBOOKS slope, B2					
INTRCPT2, G20	-3.456277	1.887391	-1.831	2343	0.067
For MUDA slope, B3					
INTRCPT2, G30	15.696049	2.577830	6.089	54	0.000

The outcome variable is LANGR

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Coefficient	Standard Error	T-ratio	Approx. d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	223.435224	3.845826	58.098	50	0.000
URB_PUB, G01	17.813535	5.556967	3.206	50	0.003
URB_PR, G02	7.050427	4.742431	1.487	50	0.143
MEANSES, G03	15.860972	5.993452	2.646	50	0.011
MNDISCIP, G04	41.225457	18.738836	2.200	50	0.032
For FEMALE slope, B1					
INTRCPT2, G10	4.764018	2.238599	2.128	2343	0.033
For TENBOOKS slope, B2					
INTRCPT2, G20	-3.456277	1.966498	-1.758	2343	0.078
For MUDA slope, B3					
INTRCPT2, G30	15.696049	2.574998	6.096	54	0.000

Final estimation of variance components:

Random Effect		Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1,	U0	23.01859	529.85542	50	446.26152	0.000
MUDA slope,	U3	14.81870	219.59376	54	140.28136	0.000
level-1,	R	38.11642	1452.86148			

Statistics for current covariance components model

Deviance = 23927.997411

Number of estimated parameters = 4

ANEXO 3: Resultados del modelo optimo multinivel -Lenguaje.

Program: HLM 6 Hierarchical Linear and Nonlinear Modeling
Authors: Stephen Raudenbush, Tony Bryk, & Richard Congdon
Publisher: Scientific Software International, Inc. (c) 2000
techsupport@ssicentral.com
www.ssicentral.com

Module: HLM2S.EXE (6.00.24282.2)
Date: 22 August 2005, Monday
Time: 14:23:27

SPECIFICATIONS FOR THIS HLM2 RUN

Problem Title: MODELO OPTIMAL PARA LENGUAJE

The data source for this run = logros.mdm
The command file for this run = C:\ MODELO OPTIMO PARA LENGUAJE Y MATEMATICA\modelo optimal para lenguaje.hlm
Output file name = C:\\MODELOS\MODELO OPTIMO PARA LENGUAJE Y MATEMATICA\MODELO OPTIMAL PARA LENGUAJE.out
The maximum number of level-1 units = 2351
The maximum number of level-2 units = 55
The maximum number of iterations = 100
Method of estimation: restricted maximum likelihood

Weighting Specification

	Weighting?	Weight Variable Name	Normalized?
Level 1	no		
Level 2	no		
Precision	no		

The outcome variable is LANGR

The model specified for the fixed effects was:

Level-1 Coefficients	Level-2 Predictors

```
 ----------------------   ---------------
        INTRCPT1, B0      INTRCPT2, G00
$                          URB_PUB, G01
$                           URB_PR, G02
$                          MEANSES, G03
$                         MNDISCIP, G04
#    FEMALE slope, B1      INTRCPT2, G10
#*    SES_1 slope, B2      INTRCPT2, G20
#* DISCIP_1 slope, B3      INTRCPT2, G30
      MUDA slope, B4      INTRCPT2, G40
```

'#' - The residual parameter variance for this level-1 coefficient has been set
to zero.
'*' - This level-1 predictor has been centered around its group mean.
'$' - This level-2 predictor has been centered around its grand mean.

The model specified for the covariance components was:

Sigma squared (constant across level-2 units)

Tau dimensions
INTRCPT1
MUDA slope

Summary of the model specified (in equation format)

Level-1 Model

Y = B0 + B1*(FEMALE) + B2*(SES_1) + B3*(DISCIP_1) + B4*(MUDA) + R

Level-2 Model
B0 = G00 + G01*(URB_PUB) + G02*(URB_PR) + G03*(MEANSES) + G04*(MNDISCIP) + U0
B1 = G10
B2 = G20
B3 = G30
B4 = G40 + U4

Iterations stopped due to small change in likelihood function

******* ITERATION 8 *******

Sigma_squared = 1454.01926

Tau

INTRCPT1,B0	533.31340	-185.55464
MUDA,B4	-185.55464	217.63279

Tau (as correlations)

INTRCPT1,B0	1.000	-0.545
MUDA,B4	-0.545	1.000

Random level-1 coefficient	Reliability estimate
INTRCPT1, B0	0.880
MUDA, B4	0.599

The value of the likelihood function at iteration 8 = -1.196345E+004
The outcome variable is LANGR

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard Error	T-ratio	Approx. d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	220.841140	3.382570	65.288	50	0.000
URB_PUB, G01	17.615386	7.991056	2.204	50	0.032
URB_PR, G02	7.403279	11.606049	0.638	50	0.526
MEANSES, G03	15.553226	6.108256	2.546	50	0.014
MNDISCIP, G04	40.789613	21.445899	1.902	50	0.062
For FEMALE slope, B1					
INTRCPT2, G10	5.040774	1.728633	2.916	2342	0.004
For SES_1 slope, B2					
INTRCPT2, G20	2.192339	1.976350	1.109	2342	0.268
For DISCIP_1 slope, B3					
INTRCPT2, G30	2.805326	3.283337	0.854	2342	0.393
For MUDA slope, B4					
INTRCPT2, G40	15.916197	2.570142	6.193	54	0.000

The outcome variable is LANGR

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Coefficient	Standard Error	T-ratio	Approx. d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	220.841140	3.220584	68.572	50	0.000
URB_PUB, G01	17.615386	5.595715	3.148	50	0.003
URB_PR, G02	7.403279	4.771457	1.552	50	0.127
MEANSES, G03	15.553226	6.023593	2.582	50	0.013
MNDISCIP, G04	40.789613	18.606450	2.192	50	0.033
For FEMALE slope, B1					
INTRCPT2, G10	5.040774	2.287249	2.204	2342	0.028
For SES_1 slope, B2					
INTRCPT2, G20	2.192339	2.019762	1.085	2342	0.278
For DISCIP_1 slope, B3					
INTRCPT2, G30	2.805326	3.507539	0.800	2342	0.424
For MUDA slope, B4					
INTRCPT2, G40	15.916197	2.545566	6.253	54	0.000

Final estimation of variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, U0	23.09358	533.31340	50	447.71797	0.000
MUDA slope, U4	14.75238	217.63279	54	139.62557	0.000
level-1, R	38.13160	1454.01926			

Statistics for current covariance components model

Deviance = 23926.895571
Number of estimated parameters = 4

ANEXO 4: Resultados para la varianza explicada por la media global de las escuelas.

Program: HLM 6 Hierarchical Linear and Nonlinear Modeling
Authors: Stephen Raudenbush, Tony Bryk, & Richard Congdon
Publisher: Scientific Software International, Inc. (c) 2000
techsupport@ssicentral.com
www.ssicentral.com

Module: HLM2S.EXE (6.00.24282.2)
Date: 28 August 2005, Sunday
Time: 11: 6:16

SPECIFICATIONS FOR THIS HLM2 RUN

Problem Title: VARIANZA EXPLICATIVA DE Bo

The data source for this run = C:\DOCUMENTOS PARA TESIS ANALISIS DE REGRESION MULTINIVEL\CORRIDAS EN HLM\Modelos\MODELO OPTIMO PARA LENGUAJE Y MATEMATICA\logros.mdm
The command file for this run = C:\DOCUMENTOS PARA TESIS ANALISIS DE REGRESION MULTINIVEL\CORRIDAS EN HLM\Modelos\MODELO OPTIMO PARA LENGUAJE Y MATEMATICA\varianza explicativa lenguaje.hlm
Output file name = C:\DOCUMENTOS PARA TESIS ANALISIS DE REGRESION MULTINIVEL\CORRIDAS EN HLM\Modelos\MODELO OPTIMO PARA LENGUAJE Y MATEMATICA\VARIANZANZA EXPLICATIVA LENGUAJE.out
The maximum number of level-1 units = 2351
The maximum number of level-2 units = 55
The maximum number of iterations = 100
Method of estimation: restricted maximum likelihood

Weighting Specification

	Weighting?	Weight Variable Name	Normalized?
Level 1	no		
Level 2	no		
Precision	no		

The outcome variable is LANGR

The model specified for the fixed effects was:

Level-1 Coefficients	Level-2 Predictors
INTRCPT1, B0	INTRCPT2, G00
# FEMALE slope, B1	INTRCPT2, G10
#* SES_1 slope, B2	INTRCPT2, G20
#* DISCIP_1 slope, B3	INTRCPT2, G30
MUDA slope, B4	INTRCPT2, G40

'#' - The residual parameter variance for this level-1 coefficient has been set to zero.
'*' - This level-1 predictor has been centered around its group mean.

The model specified for the covariance components was:

Sigma squared (constant across level-2 units)

Tau dimensions
INTRCPT1
MUDA slope

Summary of the model specified (in equation format)

Level-1 Model

Y = B0 + B1*(FEMALE) + B2*(SES_1) + B3*(DISCIP_1) + B4*(MUDA) + R

Level-2 Model
B0 = G00 + U0
B1 = G10
B2 = G20
B3 = G30
B4 = G40 + U4

Iterations stopped due to small change in likelihood function

******* ITERATION 6 *******

Sigma_squared = 1454.07561

Tau

INTRCPT1,B0	727.33183	-198.10294
MUDA,B4	-198.10294	217.28680

Tau (as correlations)

INTRCPT1,B0	1.000	-0.498
MUDA,B4	-0.498	1.000

Random level-1 coefficient	Reliability estimate
INTRCPT1, B0	0.909
MUDA, B4	0.599

The value of the likelihood function at iteration 6 = -1.198774E+004
The outcome variable is LANGR

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard Error	T-ratio	Approx. d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	220.903120	3.869411	57.090	54	0.000
For FEMALE slope, B1					
INTRCPT2, G10	4.912514	1.730124	2.839	2346	0.005
For SES_1 slope, B2					
INTRCPT2, G20	2.193876	1.976323	1.110	2346	0.268
For DISCIP_1 slope, B3					
INTRCPT2, G30	2.733785	3.283302	0.833	2346	0.405
For MUDA slope, B4					
INTRCPT2, G40	15.965565	2.569328	6.214	54	0.000

The outcome variable is LANGR

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Coefficient	Standard Error	T-ratio	Approx. d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	220.903120	3.746048	58.970	54	0.000
For FEMALE slope, B1					
INTRCPT2, G10	4.912514	2.287743	2.147	2346	0.032
For SES_1 slope, B2					
INTRCPT2, G20	2.193876	2.022580	1.085	2346	0.279
For DISCIP_1 slope, B3					
INTRCPT2, G30	2.733785	3.490483	0.783	2346	0.434
For MUDA slope, B4					
INTRCPT2, G40	15.965565	2.546253	6.270	54	0.000

Final estimation of variance components:

Random Effect		Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1,	U0	26.96909	727.33183	54	654.80119	0.000
MUDA slope,	U4	14.74065	217.28680	54	139.66176	0.000
level-1,	R	38.13234	1454.07561			

Statistics for current covariance components model

Deviance = 23975.476608
Number of estimated parameters = 4