Modelo lineal general jerárquico

Fernando Velasco Luna *

Laboratorio de Investigación y Asesoría Estadística, LINAE
Facultad de Estadística e Informática
Universidad Veracruzana
Av. Xalapa Esq. Manuel Avila Camacho s/n.
C.P. 91020. Xalapa Veracruz, México.

La Modelación Lineal Jerárquica es una técnica de regresión la cual toma en cuenta la estructura jerárquica de los datos. Existe gran diversidad de situaciones en las cuales es necesario hacer uso de los modelos jerárquicos dependiendo del objetivo del estudio y de la naturaleza de los datos. En este trabajo se presentan los modelos lineales jerarquicos de mayor aplicación así como el modelo lineal general jerárquico.

Hierarchical modeling is a regression technique that takes into account the hierarchical structure of the data. There exist a diversity of situations where there is the need of the use of hierarchical models, depending on the objectives of the study and the nature of the data. In this work some the most used hierarchical linear models are presented as well as the general hierarchical linear model.

Palabras clave: Datos con Estructura Jerárquica, Modelos Lineales Jerárquicos, Componentes de la Varianza.
Keywords: Structure Hierarchical, Hierarchical Linear Models, Variance Components.

1. Introducción

Los modelos lineales jerárquicos son una clase general de modelos que permiten la modelación en una gran variedad de situaciones en las cuales se tienen datos que presentan una estructura jerárquica. Estos modelos tienen una gran variedad de aplicaciones en diversas áreas, tales como: investigación educativa, biología, investigación social, psicología, medicina, entre otras. Los modelos lineales jerárquicos tienen una gran historia, pero han recibido especial atención en los últimos años y sus áreas de aplicación se han multiplicado considerablemente [13, 14, 5, 6, 17]. Recientes desarrollos en computo han hecho que se incremente la atención en el uso de modelos lineales jerárquicos en el análisis de datos con estructura jerárquica. En la actualidad existe software estadístico el cual permite analizar datos con estructura jerárquica de acuerdo al modelo apropiado, MLwiN, [16], S-PLUS [15], SAS [12, 20]. Una revisión exhaustiva sobre el tema puede encontrarse en [10]. Los modelos lineales jerárquicos son también conocidos en la literatura bajo una gran variedad de nombres, tales como modelos multinivel [4, 6], modelos de coeficientes aleatorios [14], modelos de componentes de la varianza y covarianza [18], o como modelos de efectos mixtos [11].

*fvelasco@uv.mx

2. Datos con estructura jerárquica

Los datos con estructura jerárquica surgen en varias situaciones. Por ejemplo: investigaciones educativas frecuentemente están relacionadas con problemas de investigación de relaciones existentes entre alumnos y el grupo de clase en el que éstos se desenvuelven. El concepto general es que el alumno interactúa con el grupo al cual éste pertenece, generalmente los alumnos y los grupos de clase son definidos en niveles separados de esta estructura jerárquica; es decir, se tienen J grupos con n_j unidades en el j-ésimo grupo, $j = 1, ..., J$. A cada grupo se le denomina unidad de nivel 2; así se tienen J unidades de nivel 2, y a cada unidad de las n_j unidades en cada grupo se les denomina unidad de nivel 1; con lo que se tienen n_j unidades de nivel 1 en la j-ésima unidad de nivel 2. El número n_j de unidades de nivel 1 no tiene que ser necesariamente igual en cada unidad de nivel 2. En general en un sistema con estructura jerárquica se pueden presentar varios niveles. Los datos con estructura jeráquica ocurren de manera natural en medicina [3, 19], educación [7, 8], estudios de datos longitudinales [22, 1], en psicología [2, 9], entre otros campos.

3. Algunos modelos lineales jerárquicos

Para analizar datos con estructura jerárquica se tiene que emplear técnicas estadísticas que tomen en cuenta dicha estructura. En esta situación, es razonable postular un modelo que considere una posible diferencia entre las unidades de nivel 2, es decir, plantear un modelo tal que, para cada unidad en el nivel 2, se tengan diferentes coeficientes. Bajo esta situación el modelo lineal jerárquico permite simultáneamente hacer un estudio en los niveles de la estructura jerárquica, tomando en cuenta variables para las unidades en cada uno de los niveles. En los modelos lineales jerárquicos cada uno de los niveles de la estructura jerárquica es representado formalmente con su propio submodelo. Un tratamiento y abundantes referencias acerca de estos modelos se puede encontrar en [5, 6, 13, 14, 10, 21].

A continuación se describen algunos modelos lineales jerárquicos

3.1 Modelo con intercepto aleatorio

El caso más simple de un modelo lineal jerárquico es el denominado intercepto aleatorio, el cual no contiene ni variables explicatorias en el nivel 1, ni variables explicatorias en el nivel 2. En este modelo solamente se tiene variabilidad entre las unidades de nivel 2 y dentro de las unidades de nivel 2. Este modelo puede ser expresado como un modelo donde la variable respuesta, y_{ij}, es la suma de una media general dada por β_{00}, un efecto aleatorio a nivel 2 dado por u_{0j}, y un efecto aleatorio a nivel 1 dado por e_{ij}. El modelo para la i-ésima unidad de nivel 1, la cual se encuentra en la j-ésima unidad de nivel 2, tiene la forma

$$\begin{aligned} y_{ij} &= \beta_{00} + u_{0j} + e_{ij}, \\ E(e_{ij}) &= 0,\ \operatorname{Var}(e_{ij}) = \sigma_e^2, \\ E(u_{0j}) &= 0,\ \operatorname{Var}(u_{0j}) = \sigma_{u0}^2. \end{aligned} \tag{1}$$

Los parámetros en el modelo (1) son tres: El coeficiente β_{00} y las varianzas σ_e^2 y

σ_{u0}^2, las cuales se denominan componentes de la varianza. En el modelo intercepto aleatorio la varianza de la variable respuesta es descompuesta como la suma de las varianzas a nivel 1, σ_e^2, y a nivel 2, σ_{u0}^2,

$$\mathrm{Var}\,(y_{ij}) = \sigma_e^2 + \sigma_{u0}^2. \tag{2}$$

El modelo para el nivel 1 tienen la forma:

$$y_{ij} = \beta_{0j} + e_{ij}, \tag{3}$$

y el modelo para el nivel 2 tiene la forma:

$$\beta_{0j} = \beta_{00} + u_{0j}. \tag{4}$$

3.2 Modelo con intercepto aleatorio con varias explicatorias a nivel 1

En el modelo intercepto aleatorio el valor esperado de la variable respuesta puede ser explicado en términos de variables explicatorias a nivel 1. Así la siguiente etapa es la inclusión de variables explicatorias en el nivel 1, esto con el objetivo de tratar de explicar el comportamiento de la variable respuesta. Con una variable explicatoria en el nivel 1 el modelo intercepto aleatorio tiene la forma:

$$\begin{aligned} y_{ij} &= \beta_{00} + \beta_1 x_{ij} + u_{0j} + e_{ij}, \\ E\,(e_{ij}) &= 0,\ \mathrm{Var}\,(e_{ij}) = \sigma_e^2, \\ E\,(u_{0j}) &= 0,\ \mathrm{Var}\,(u_{0j}) = \sigma_{u0}^2. \end{aligned} \tag{5}$$

El modelo (5) se denomina modelo intercepto aleatorio con una variable explicatoria a nivel 1. Los parámetros en el modelo (5) son cuatro: Los coeficientes de regresión β_{00} y β_1, y los componentes de la varianza σ_e^2 y σ_{u0}^2.

En el modelo intercepto aleatorio con una variable explicatoria a nivel 1 la varianza de la variable respuesta es descompuesta como la suma de las varianzas a nivel 1 σ_e^2, y a nivel 2, σ_{u0}^2, de la siguiente manera:

$$\mathrm{Var}\,(y_{ij}) = \sigma_e^2 + \sigma_{u0}^2. \tag{6}$$

El modelo para el nivel 1 tienen la forma:

$$y_{ij} = \beta_{0j} + \beta_1 x_{ij} + e_{ij}, \tag{7}$$

y el modelo para el nivel 2 tiene la forma:

$$\beta_{0j} = \beta_{00} + u_{0j}. \tag{8}$$

Al igual que en el modelo de regresión múltiple, más de una variable explicatoria a nivel 1 puede ser usada en el modelo intercepto aleatorio. La generalización del modelo (5) para incluir más variables explicatorias a nivel 1; es decir, el modelo intercepto aleatorio con m variables explicatorias a nivel 1 tiene la forma:

$$\begin{aligned} y_{ij} &= \beta_{00} + \beta_1 x_{1ij} + \beta_2 x_{1ij} + \cdots + \beta_m x_{mij} + u_{0j} + e_{ij}, \\ E\,(e_{ij}) &= 0,\ \mathrm{Var}\,(e_{ij}) = \sigma_e^2, \\ E\,(u_{0j}) &= 0,\ \mathrm{Var}\,(u_{0j}) = \sigma_{u0}^2. \end{aligned} \tag{9}$$

El modelo (9) se denomina modelo intercepto aleatorio con m variables explicatorias a nivel 1. Los parámetros de este modelo son $m+3$; los $m+1$ coeficientes de regresión β_{00},β_1,β_2,...,β_m y los componentes de la varianza σ_e^2 y σ_{u0}^2.

El modelo para el nivel 1 tiene la forma:

$$y_{ij} = \beta_{0j} + \beta_1 x_{1ij} + \beta_2 x_{2ij} + \cdots + \beta_m x_{mij} + e_{ij}, \tag{10}$$

y el modelo para el nivel 2 tiene la forma:

$$\beta_{0j} = \beta_{00} + u_{0j}. \tag{11}$$

3.3 Modelo con pendientes aleatorias

En el modelo lineal jerárquico intercepto aleatorio con una o más variables explicatorias a nivel 1, sólo el intercepto se supone aleatorio, mientras que los demás coeficientes de regresión se suponen fijos para todas las unidades de nivel 2. En ocasiones la relación entre las variables explicatorias y la variable respuesta puede ser diferente en las unidades de nivel 2. Lo anterior da surgimiento al modelo de pendientes aleatorias. En este modelo los coeficientes de algunas o de todas las variables explicatorias están variando entre las unidades de nivel 2, es decir, la relación existente entre cada una de las variables explicatorias y la variable respuesta no es la misma en todas las unidades de nivel 2. Para el caso de una variable explicatoria a nivel 1 lo anterior se expresa en el siguiente modelo:

$$\begin{aligned} y_{ij} &= \beta_{00} + \beta_{10} x_{1ij} + u_{0j} + u_{1j} x_{1ij} + e_{ij}, \\ E\left(e_{ij}\right) &= 0, \quad \mathrm{Var}\left(e_{ij}\right) = \sigma_e^2, \\ E\left(u_{0j}\right) &= 0, \quad \mathrm{Var}\left(u_{0j}\right) = \sigma_{u0}^2, \\ E\left(u_{1j}\right) &= 0, \;\; \mathrm{Var}\left(u_{1j}\right) = \sigma_{u1}^2, \\ Cov\left(u_{0j}, u_{1j}\right) &= 0, \quad Cov\left(u_{kj}, e_{ij}\right) = 0, k = 0, 1. \end{aligned} \tag{12}$$

el cual se denomina modelo de pendientes aleatorias con una variable explicatoria a nivel 1.

Los parámetros del modelo de pendientes aleatorias con una variable explicatoria a nivel 1 son seis: los coeficientes de regresión β_{00} y β_{10}, y los componentes de la varianza σ_e^2, σ_{u0}^2, σ_{u1}^2 y σ_{u01}.

En el modelo de pendientes aleatorias con una variable explicatoria a nivel 1 la varianza de la variable respuesta se descompone de la siguiente forma:

$$\mathrm{Var}\left(y_{ij}\right) = \sigma_{u0}^2 + \sigma_{u1}^2 x_{ij}^2 + 2\sigma_{u01} x_{ij} + \sigma_e^2. \tag{13}$$

De la ecuación (13) se tiene que en el modelo de pendientes aleatorias con una variable explicatoria nivel 1 la varianza de la variable respuesta depende de la variable explicatoria a nivel 1, x_{ij}.

El modelo para el nivel 1 tiene la forma:

$$y_{ij} = \beta_{0j} + \beta_{1j} x_{ij} + e_{ij}, \tag{14}$$

y el modelo para el nivel 2 tiene la forma:

$$\beta_{0j} = \beta_{00} + u_{0j}, \quad \beta_{1j} = \beta_{10} + u_{1j}, \tag{15}$$

Aquí se observa que los coeficientes β_{0j} y β_{1j} son aleatorios, es decir cambian de unidad de nivel 2 a unidad de nivel 2.

El modelo con pendientes aleatorias con m varias variables explicatorias a nivel 1 tiene la forma:

$$\begin{aligned} y_{ij} &= \beta_{00} + \sum_{k=1}^{m} \beta_{k0} x_{kij} + u_{0j} + \sum_{k=1}^{m} u_{kj} x_{kij} + e_{ij}, \\ E(e_{ij}) &= 0,\ \mathrm{Var}(e_{ij}) = \sigma_e^2,\ \ E(u_{kj}) = 0,\ \mathrm{Var}(u_{kj}) = \sigma_{uk}^2, \\ Cov(u_{kj,} u_{lj}) &= \sigma_{ukl},\ \ Cov(u_{kj}, e_{ij}) = 0,\ \ k, l = 0, ..., m. \end{aligned} \tag{16}$$

el cual se denomina modelo de pendientes aleatorias con m variables explicatorias a nivel 1.

Los parámetros del modelo de pendientes aleatorias con m variables explicatorias a nivel 1 son $\frac{(m+2)(m+3)}{2}$; los $m+1$ coeficientes de regresión β_{00},β_{10},β_{20},...,β_{m0} y los $\frac{(m+2)(m+1)}{2}$ +1 componentes de la varianza σ_e^2, σ_{uk}^2, y σ_{ukl}, $k, l = 0, ..., m$.

En el modelo de pendientes aleatorias con m variables explicatorias a nivel 1 la varianza de la variable respuesta se descompone de la siguiente forma:

$$\mathrm{Var}(y_{ij}) = \sum_{k=0}^{m} \sigma_{uk}^2 x_{kij}^2 + 2 \sum_{k,l=0}^{m} \sigma_{ukl} x_{kij} x_{lij} + \sigma_e^2. \tag{17}$$

donde $x_{0ij} = 0$.

De la ecuación (17) se tiene que en el modelo de pendientes aleatorias con m variables explicatorias a nivel 1 la varianza de la variable respuesta depende de las m variables explicatorias a nivel 1, $x_{1ij}, x_{2ij}, ..., x_{mij}$.

3.4 Modelo jerárquico con variables explicatorias a nivel 1 y a nivel 2

Se han tratado hasta el momento modelos jerárquicos en los cuales se registraron mediciones sobre una variable respuesta y y sobre una o más variables explicatorias a nivel 1, $x_1, x_2, ..., x_m$. Además se puede medir otro conjunto de variables, w_1,w_2,...,w_q para cada una de las unidades de nivel 2, las cuales se denominan variables explicatorias a nivel 2. Lo anterior da surgimiento al modelo jerárquico con variables explicatorias a nivel 1 y a nivel 2.

Por facilidad de comprensión se comenzara con los modelos para cada uno de los niveles de la jerárquia.

Modelo en el nivel 1: El modelo para el nivel 1, con una variable explicatoria a nivel 1, x_{1ij}, tiene la forma:

$$\begin{aligned} y_{ij} &= \beta_{0j} + \beta_{1j} x_{1ij} + e_{ij}, \\ i &= 1, 2, ..., n_j;\ \ j = 1, 2, ..., J, \\ E(e_{ij}) &= 0,\ \mathrm{Var}(e_{ij}) = \sigma_e^2, \end{aligned} \tag{18}$$

Modelo en el nivel 2: El modelo para el nivel 2, con una variable explicatoria a nivel 2, w_{1j}, tiene la forma:

$$\begin{aligned}\beta_{0j} &= \beta_{00} + \beta_{01} w_{1j} + u_{0j},\\ \beta_{1j} &= \beta_{10} + \beta_{11} w_{1j} + u_{1j},\\ E(u_{0j}) &= 0, \quad \mathrm{Var}(u_{0j}) = \sigma^2_{u0},\\ E(u_{1j}) &= 0, \ \mathrm{Var}(u_{1j}) = \sigma^2_{u1},\\ Cov(u_{0j}, u_{1j}) &= 0. \end{aligned} \tag{19}$$

El modelo combinado para la j-ésima unidad de nivel 2, del modelo nivel 1 con una variable explicatoria a nivel 1; y del modelo nivel 2 con una variable explicatoria a nivel 2, tiene la forma:

$$y_{ij} = (\beta_{00} + \beta_{01} w_{1j} + u_{0j}) + (\beta_{10} x_{1ij} + \beta_{11} w_{1j} x_{1ij} + u_{1j} x_{1ij}) + e_{ij}, \tag{20}$$

reordenando términos se obtiene el modelo

$$y_{ij} = \beta_{00} + \beta_{10} x_{1ij} + \beta_{01} w_{1ij} + \beta_{11} w_{1j} x_{1ij} + u_{0j} + u_{1j} x_{1ij} + e_{ij}. \tag{21}$$

La generalización del modelo (21) para incluir más variables explicatorias a nivel 1 y a nivel 2 se presenta a continuación:

Modelo en el nivel 1: El modelo para el nivel 1, con m variables explicatorias a nivel 1 x_{1j}, x_{2j},...,x_{mj} tiene la forma

$$\begin{aligned} y_{ij} &= \beta_{0j} + \beta_{1j} x_{1ij} + \beta_{2j} x_{2j} + \cdots + \beta_{mj} x_{mij} + e_{ij},\\ & i = 1, 2, ..., n_j; \ \ j = 1, 2, ..., J,\\ E(e_{ij}) &= 0, \ \mathrm{Var}(e_{ij}) = \sigma^2_e. \end{aligned} \tag{22}$$

Modelo en el nivel 2: El modelo para el nivel 2, con q variables explicatorias a nivel 2 $w_{1j}, w_{2j}, ..., w_{qj}$, tiene la forma:

$$\begin{aligned} \beta_{0j} &= \beta_{00} + \beta_{01} w_{1j} + \beta_{02} w_{2j} + \cdots + \beta_{0q} w_{qj} + u_{0j}\\ \beta_{1j} &= \beta_{10} + \beta_{11} w_{1j} + \beta_{12} w_{2j} + \cdots + \beta_{1q} w_{qj} + u_{1j}\\ &\vdots\\ \beta_{mj} &= \beta_{m0} + \beta_{m1} w_{1j} + \beta_{m2} w_{2j} + \cdots + \beta_{mq} w_{qj} + u_{mj}. \end{aligned} \tag{23}$$

El modelo combinado para la j-ésima unidad de nivel 2, del modelo nivel 1 con m variables explicatorias a nivel 1; y del modelo nivel 2 con q variables explicatorias a nivel 2, tiene la forma:

$$\begin{aligned} y_{ij} = \beta_{00} &+ \beta_{01} w_{1j} + \beta_{02} w_{2j} + \cdots + \beta_{0q} w_{qj} + u_{0j} +\\ &+ \beta_{10} x_{1ij} + \beta_{11} w_{1j} x_{1ij} + \cdots + \beta_{1q} w_{qj} x_{1ij} + u_{1j} x_{1ij}\\ &+ \cdots + \beta_{m0} x_{mij} + \cdots + \beta_{mq} w_{qj} x_{mij} + u_{mj} x_{mij} + e_{ij}, \end{aligned} \tag{24}$$

reordenando términos se obtiene el modelo

$$\begin{aligned} y_{ij} = \beta_{00} &+ \beta_{10} x_{1j} + \cdots + \beta_{m0} x_{mij} +\\ &+ \beta_{01} w_{1j} + \beta_{11} w_{1j} x_{1ij} + \cdots + \beta_{m1} w_{1j} x_{mij} +\\ &+ \cdots + \beta_{0q} w_{qj} + \beta_{1q} w_{qj} x_{1ij} \cdots + \beta_{mq} w_{qj} x_{mij} +\\ &+ u_{0j} + u_{1j} x_{1ij} + \cdots + u_{mj} x_{mij} + e_{ij}. \end{aligned} \tag{25}$$

4. Modelo lineal general jerárquico

Existen muchas variantes de los modelos anteriores, ya sea teniendo algunas variables explicatorias a nivel 1, con coeficientes fijos y otras con coeficientes aleatorios, o más aún añadiendo más niveles a la estructura jerárquica. Todos estos modelos son casos especiales del modelo lineal general jerárquico el cual se presenta a continuación. Definiendo

$$\mathbf{Y}_j = \begin{bmatrix} y_{1j} \\ y_{2j} \\ \vdots \\ y_{n_j j} \end{bmatrix} ; \quad \mathbf{X}_j = \begin{bmatrix} 1 & x_{11j} & x_{21j} & \cdots & x_{m1j} \\ 1 & x_{12j} & x_{22j} & \cdots & x_{m2j} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n_j j} & x_{2n_j j} & \cdots & x_{mn_j j} \end{bmatrix} ;$$

$$\mathbf{e}_j = \begin{bmatrix} e_{1j} \\ e_{2j} \\ \vdots \\ e_{n_j j} \end{bmatrix} ; \quad \beta_j = \begin{bmatrix} \beta_{0j} \\ \beta_{1j} \\ \vdots \\ \beta_{mj} \end{bmatrix}$$

En forma matricial el modelo nivel 1 (22) toma la forma

$$\mathbf{Y}_j = \mathbf{X}_j \beta_j + \mathbf{e}_j; \quad j = 1, ..., J, \tag{26}$$

donde $\mathbf{Y}_j$ es el vector respuesta n_jx1, $\mathbf{X}_j$ es la matriz de variables explicatorias a nivel 1 de orden n_jx$(m+1)$, β_j es el vector de parámetros de orden $(m+1)$x1 y $\mathbf{e}_j$ es un vector de errores aleatorios n_jx1. Se supone $E(\mathbf{e}_j) = \mathbf{0}$, $\operatorname{Var}(\mathbf{e}_j) = \sigma_e^2 \mathbf{I}_{n_j}$. Definiendo

$$\mathbf{W}_j = \begin{bmatrix} 1 & w_{1j} & w_{2j} & \cdots & w_{qj} & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & \cdots & 1 & w_{1j} & w_{2j} & \cdots & w_{qj} \end{bmatrix}$$

y

$$\begin{aligned} \beta &= [\beta_{00}, \beta_{01}, ..., \beta_{0q}, \beta_{10}, \beta_{11}, ..., \beta_{1q}, ..., \beta_{m0}, \beta_{m1}, ..., \beta_{mq}]^T ; \\ \mathbf{u}_j &= [u_{0j}, u_{1j}, ..., u_{mj}]^T \end{aligned}$$

En forma matricial el modelo nivel 2 (23) tiene la forma

$$\beta_j = \mathbf{W}_j \beta + \mathbf{u}_j; \quad j = 1, ..., J, \tag{27}$$

donde $\mathbf{W}_j$ es la matriz de variables explicatorias a nivel 2, de orden $(m+1)$x$(q+1)(m+1)$, β es el vector $(q+1)(m+1)$x1 de coeficientes fijos, y $\mathbf{u}_j$ es el vector de errores aleatorios nivel 2 de orden $(m+1)$x1. Supongase $E(\mathbf{u}_j) = \mathbf{0}$ y

$$\operatorname{Var}(\mathbf{u}_j) = \mathbf{\Omega} = \begin{bmatrix} \sigma_{uo}^2 & \sigma_{u01} & \cdots & \sigma_{u0m} \\ \sigma_{u01} & \sigma_{u1}^2 & \cdots & \sigma_{u1m} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{u0m} & \sigma_{u1m} & \cdots & \sigma_{um}^2 \end{bmatrix},$$

En forma matricial el modelo combinado para la j-ésima unidad de nivel 2 tiene la forma

$$\begin{aligned} \mathbf{Y}_j &= \mathbf{X}_j \mathbf{W}_j \beta + \mathbf{X}_j \mathbf{u}_j + \mathbf{e}_j; \quad j = 1, ..., J, \\ E(\mathbf{Y}_j) &= \mathbf{X}_j \mathbf{W}_j \beta, \quad \mathbf{V}_j = \mathrm{Var}(\mathbf{Y}_j) = \mathbf{X}_j \boldsymbol{\Omega} \mathbf{X}_j^T + \sigma_e^2 \mathbf{I}_{n_j}. \end{aligned} \tag{28}$$

Definiendo

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \\ \vdots \\ \mathbf{Y}_J \end{bmatrix}; \ \mathbf{u} = \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \vdots \\ \mathbf{u}_J \end{bmatrix}; \quad \mathbf{e} = \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \vdots \\ \mathbf{e}_J \end{bmatrix}; \quad \mathbf{X} = \mathrm{diag}(\mathbf{X}_j); \ \mathbf{W} = \begin{bmatrix} \mathbf{W}_1 \\ \mathbf{W}_2 \\ \vdots \\ \mathbf{W}_J \end{bmatrix}$$

donde $\mathrm{diag}(\mathbf{A}_j)$ representa los términos diagonales por matriz bloque, con $\mathbf{A}_j$ en el j-ésimo bloque de la diagonal. Del modelo (28) se tiene

$$\mathbf{Y} = \mathbf{X}\mathbf{W}\beta + \mathbf{X}\mathbf{u} + \mathbf{e}, \tag{29}$$

el cual se denomina modelo lineal general jerárquico. La matriz de varianzas y covarianzas tiene la forma

$$\mathbf{V} = \mathrm{Var}(\mathbf{Y}) = \mathbf{X}\mathrm{diag}(\boldsymbol{\Omega})\mathbf{X}^T + \mathrm{diag}(\sigma_e^2 \mathbf{I}_{n_j}) \tag{30}$$

Definiendo

$$\mathrm{Var}(\mathbf{e}) = \mathbf{R} = \mathrm{diag}(\sigma_e^2 \mathbf{I}_{n_j}) \quad \text{y} \quad \mathrm{Var}(\mathbf{u}) = \mathbf{G} = \mathrm{diag}(\Omega) \tag{31}$$

la matriz de varianzas y covarianzas para el modelo lineal general jerárquico (29) tiene la forma

$$\mathbf{V} = \mathrm{Var}(\mathbf{Y}) = \mathbf{X}\mathbf{G}\mathbf{X}^T + \mathbf{R}. \tag{32}$$

Referencias

[1] Arnau, J. y Balluerka, N. "Análisis de Datos Longitudinales y de Curvas de Crecimiento. Enfoque clásico y propuestas actuales". *Psicuthema*, **16(1),** 2004, 156-162.

[2] Boyle, M. H. and Douglas, J. "Multilevel Modeling of Hierarchical Data in Developmental Studies". *The Journal of Child Psychology and Psychiatry*, **42**(1), 2001, 141-162.

[3] Brown, H. *Applied Mixed Models in Medicine.* Wiley, New York, 1999.

[4] Goldstein, H. "Efficient Statistical Modeling of Longitudinal Data". *Annals of Human Biology*, **13**(2), 1986, 129-141.

[5] Goldstein, H. *Multilevel Models in Educational and Social Research.* Griffin, London, 1987.

[6] Goldstein, H. *Multilevel Statistical Models.* Second Edition. Halsted Press, New York, 1995.

[7] Goldstein, H. "Methods in School Effectiveness Research". *School Effectiveness and school Improvement*, **8,** 1997, 369-39.

[8] Goldstein, H., Rasbash, J., Yang, M. y otros "A Multilevel Analysis of School Examination Results". *Oxford Review of Education*, **19**(4), 1993, 425-433.

[9] Hernández, L. M. V., Colmeras, F. y Martínez, A. R. "Modelos Jerárquicos por Piezas en el Análisis de la Relación entre Discontinuidad Conductual y Discontinuidad en Procesos Subyacentes". *Anales de Psicología*, **19**(1), 2003, 159-171.

[10] Kreft, I. and De Lleuw, J. *Introducing Multilevel Modelin*g. SAGE publications, 1998.

[11] Laird, N. M. and Ware, J. H. "Random Effects Models for Longitudinal Data". *Biometrics*, **38**, 1982 963-974.

[12] Little, R. C. Stroup, W. W. and Freund, R. J. *SAS for Linear Models*. Cary, NC: SAS Institute, Inc. 2002.

[13] Longford, N. T. *Random Coefficient Models*. Oxford: University Press, New York. 1993.

[14] Longford, N. T. *Random Coefficient Models. In Handbook of Statistical Models for the Social and Behavioral Sciences*, Plenum Press, New York, 1995.

[15] Pinheiro J. C. and Bates, D. M. *Mixed Effcets Models in S and S-PLUS*. Springer Verlag, New York, 2000.

[16] Rasbash, J., Steele, F., Browne, W., y otros. *A users Guide to MLwiN version 2.0*. Instirute of Education, University of London, London, 2004.

[17] Raudenbush, S. W. and Bryk, A. S. *Hierarchical Linear Models: Applications and Data Analysis Methods*. Sage Publications, NewburyPark, 2002.

[18] Searle, S. R., Casella, G., and Mcculloch, C. E. *Variance Components*. John Wiley, New York, 1992.

[19] Sullivan, L. M., Dukaes K. A. and Losina, E. "Tutorial in Biostatistics an Introduction to Hierarchical Linear modelling". *Statistics in Medicine* **18**, 1999, 855-888.

[20] Singer, J. D. "Using SAS PROC MIXED to fit Multilevel Models, Hierarchical Models, and Individual Growth Models". *Journal of Educational and Behavioral Statistics*, **24**(4), 1998, 323-355.

[21] Snijders, T. A. B. and Bosher, R. *Multilevel analysis: An Introduction to Basic and Advanced Multilevel Modeling*. Sage Publications, London, 1999.

[22] Verbeke, G. and Molenbergs, G. *Linear Mixed Models for Longitudinal Data*. Springer-Verlag, New York, 2000.