UNIVERSIDAD VERACRUZANA

Doctorado en Matemáticas

FACULTAD DE MATEMÁTICAS

Caracterización del mejor predictor lineal insesgado de la media poblacional finita por medio de proyectores

Presenta

M. en C. Fernando Velasco Luna

Director

Dr. Mario Miguel Ojeda Ramírez

XALAPA, VER., SEPTIEMBRE 2011

CONTENIDO

Pag.

Pag.

Introducción

Considérese una población finita de la cual se obtienen una muestra, es decir un subconjunto de la población. La teoría del muestreo probabilístico para poblaciones finitas, proporciona los principios y procedimientos para la selección de muestras y a partir de ellas realizar inferencias acerca de ciertas características de la población bajo estudio. Sea N el número de unidades en la población finita $\{U_1, U_2, ..., U_N\}$. Asociado con cada unidad de la población finita sea Y_i, $i = 1, 2, ..., N$, la medición de U_i; entonces el vector de valores de las $Y's$ será denotado por $\mathbf{Y}$=$(Y_1, Y_2, ..., Y_N)$. Con el fin de obtener información acerca de alguna función de $\mathbf{Y}$, se selecciona una muestra $\mathbf{s}$ de tamaño n de $\{U_1, U_2, ..., U_N\}$. Ejemplos de funciones $\theta(\mathbf{Y})$ son: $\boldsymbol{\gamma}^t\mathbf{Y}$, donde $\boldsymbol{\gamma}$= $(\gamma_1, \gamma_2, ..., \gamma_N)$; y $\mathbf{Y}^t\mathbf{A}\mathbf{Y}$, donde $\mathbf{A}$ es una matriz conocida de orden N x N. Algunos casos especiales de funciones lineales son el total poblacional, $T = \sum_{i=1}^{N} Y_i = \boldsymbol{\gamma}^t\mathbf{Y}$ donde $\boldsymbol{\gamma} = \mathbf{1}_N$ y la media poblacional $\overline{Y} = T/N = \boldsymbol{\gamma}^t\mathbf{Y}$ donde $\boldsymbol{\gamma} = \mathbf{1}_N/N$; y un caso conocido de función cuadrática es la varianza poblacional S_y^2=$\sum_{i=1}^{N}\left(Y_i - \overline{Y}\right)^2/N = \mathbf{Y}^t\mathbf{A}\mathbf{Y}$ donde $\mathbf{A} = (\mathbf{I}_N - \mathbf{J}/N)/N$. Una vez que la muestra $\mathbf{s}$ de tamaño n ha sido obtenida y las mediciones realizadas se tiene la parte observada $\mathbf{Y_s}$ y la no observada $\mathbf{Y_r}$ del vector $\mathbf{Y}$, donde $\mathbf{Y_s} \in \mathbb{R}^n$ y donde $\mathbf{Y_r} \in \mathbb{R}^r$, con $r = N - n$. Así la cantidad $\theta(\mathbf{Y})$ se puede expresar como $\theta(\mathbf{Y})$=$\theta(\mathbf{Y_s})$+$\theta(\mathbf{Y_s}, \mathbf{Y_r})$, donde $\theta(\mathbf{Y_s})$ es una función de la parte observada $\mathbf{Y_s}$ y $\theta(\mathbf{Y_s}, \mathbf{Y_r})$ es una función de la parte observada $\mathbf{Y_s}$ y de la parte no observada $\mathbf{Y_r}$. Entonces el objetivo de inferencia se puede caracterizar como la obtención de una predicción de $\theta(\mathbf{Y_s}, \mathbf{Y_r})$ a partir de $\mathbf{Y_s}$; es decir, predecir los valores no observados a partir de los datos en la muestra $\mathbf{s}$.

El enfoque clásico del muestreo probabilístico considera que los valores de la característica de interés en cada unidad de una población finita son fijos (no aleatorios), aunque desconocidos, salvo para los elementos de la muestra una vez que ésta ha sido seleccionada y los valores medidos. En este enfoque la aleatoriedad se debe únicamente a la selección de la muestra y es definida por el diseño muestral. Esta concepción del muestreo se denomina enfoque de población fija y la teoría que se ha desarrollado para conducir las inferencias en este marco se conoce como la teoría clásica del muestreo.

Otro enfoque es el basado en modelos estadísticos, llamado el enfoque superpoblacional. En éste un dato es un número real que resulta de hacer una medición; así si un dato es desconocido se considera una variable aleatoria y entonces, cuando se conoce, será una realización de ella. Los modelos estadísticos se utilizan para caracterizar las variables y los procesos de inferencia en el muestreo de poblaciones finitas; en general, se asume que el valor de la variable de interés Y_i se puede expresar como la suma de un elemento determinista η_i y un elemento aleatorio e_i, por medio del modelo estadístico, $Y_i = \eta_i + e_i$, donde se asume que el vector $\mathbf{e} = (e_1, e_2, ..., e_N)$ tiene media cero y una matriz de varianzas y covarianzas $\mathbf{V}$. El uso del modelo superpoblacional para describir una población finita fue utilizado por Royall y Herson [41] quienes consideraron un enfoque de predicción para la estimación del total de una población finita. En este enfoque se supone que asociado con cada unidad en la población finita $\{U_1, U_2, ..., U_N\}$ existe un valor aleatorio Y_i, $i = 1, 2, ..., N$ y M cantidades conocidas, denotadas por $X_{i1}, X_{i2}, ..., X_{iM}$. Los vectores $\mathbf{X}_i = (X_{i1}, X_{i2}, ..., X_{iM})$ forman la matriz $\mathbf{X}$ de orden N x M. La relación entre la variable Y_k y las variables $X_{k1}, X_{k2}, ..., X_{kM}$

se expresa por un modelo lineal

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \quad E\left(\mathbf{e}\right) = 0, \quad Cov\left(\mathbf{e}\right) = \mathbf{V}.$$

donde $\mathbf{V}$ denota a la matriz de varianzas y covarianzas. Una vez que se tiene la muestra $\mathbf{s}$ el vector $\mathbf{Y}$, y las matrices $\mathbf{X}$ y $\mathbf{V}$ se pueden descomponer como

$$\begin{pmatrix} \mathbf{Y_s} \\ \mathbf{Y_r} \end{pmatrix}, \begin{pmatrix} \mathbf{X_s} \\ \mathbf{X_r} \end{pmatrix} \text{y} \begin{pmatrix} \mathbf{V_{ss}} & \mathbf{V_{sr}} \\ \mathbf{V_{rs}} & \mathbf{V_{rr}} \end{pmatrix},$$

donde las matrices tienen los ordenes: $\mathbf{Y_s}$ $n \times 1$; $\mathbf{Y_r}$ $r \times 1$; $\mathbf{X_s}$ $n \times p$; $\mathbf{X_r}$ $r \times p$; $\mathbf{V_{ss}}$ $n \times n$; $\mathbf{V_{sr}}$; $n \times r$; $\mathbf{V_{rs}}$ $r \times n$; $\mathbf{V_{rr}}$ $r \times r$. La estimación de la función $\theta\left(\mathbf{Y}\right)$ de interés se obtiene haciendo uso de la teoría de la predicción de Royall [39], quien propone un predictor de la cantidad de interés para poblaciones finitas, y demuestra que el mejor predictor lineal insesgado (*BLUP*, por sus siglas en Inglés Best Linear Unbiased Predictor, donde "mejor" significa que es el predictor lineal insesgado que tiene el mínimo Error Cuadrático Medio *MSE*, "lineal" significa que el predictor es una combinación lineal de los valores de la variable respuesta e "insesgado" significa que el valor esperado del error de predicción es cero), puede ser expresado como una combinación lineal de los valores observados de la muestra $\mathbf{Y_s}$ y un predictor de los valores no observados $\mathbf{Y_r}$. En el caso de funciones lineales $\theta\left(\mathbf{Y}\right)$ el vector $\boldsymbol{\gamma}$ involucrado en $\boldsymbol{\gamma}^t\mathbf{Y}$ se descompone en las partes $\boldsymbol{\gamma}_s$ y $\boldsymbol{\gamma}_\mathbf{r}$. El teorema general de predicción proporciona el *BLUP* para la cantidad poblacional de interés bajo un modelo superpoblacional a partir de la muestra. Tal predictor está dado por $\boldsymbol{\gamma}_s^t\mathbf{Y_s}+\boldsymbol{\gamma}_\mathbf{r}^t\left[\mathbf{X_r}\,\hat{\boldsymbol{\beta}}_\mathbf{s}+\mathbf{V_{rs}}\mathbf{V_{ss}}^{-1}\left(\mathbf{Y_s}-\mathbf{X_s}\,\hat{\boldsymbol{\beta}}_\mathbf{s}\right)\right]$, donde $\hat{\boldsymbol{\beta}}_\mathbf{s}$ es el estimador de mínimos cuadrados generalizados obtenido a partir de la muestra $\mathbf{s}$. En este predictor $\boldsymbol{\gamma}_{\mathbf{s}}^{t}\mathbf{Y}_{\mathbf{s}}$ es la parte calculada a partir de la muestra $\mathbf{s}$ y

$$\boldsymbol{\gamma}_{\mathbf{r}}^{t}\left[\mathbf{X}_{\mathbf{r}}\,\hat{\boldsymbol{\beta}}_{\mathbf{s}}+\mathbf{V}_{\mathbf{rs}}\mathbf{V}_{\mathbf{ss}}^{-1}\left(\mathbf{Y}_{\mathbf{s}}-\mathbf{X}_{\mathbf{s}}\,\hat{\boldsymbol{\beta}}_{\mathbf{s}}\right)\right],$$

es la parte que se debe predecir. Para una revisión de la teoría de la predicción para poblaciones finitas ver [39], [40], [41], [50], [51], entre otros. Un tratamiento monográfico de este tema se encuentra en [3], [4] y [46].

Los modelos lineales jerárquicos forman 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 (efectividad de escuela, logro escolar), biología (curvas de crecimiento, estudios genéticos), investigación social (análisis de encuestas, estudios de mercado), psicología (análisis de conducta), medicina (ajuste de datos de medidas repetidas), entre otras. Los modelos lineales jerárquicos tienen una larga historia, pero han recibido especial atención en los últimos años ([5], [7], [8], [10], [14], [15], [16], [27], [28], [36], [48], [52]). Los modelos lineales jerárquicos son también conocidos como modelos multinivel [10], [16], modelos de componentes de la varianza y covarianza [42], modelos de coeficientes aleatorios [27], o como modelos de efectos mixtos [24], [25], [48], [52]. Un tratamiento y abundantes referencias de estos modelos se pueden encontrar en [15], [16], [23], [24], [27], [29], [36], [47], y [48]. En la actualidad existe software para analizar datos con estructura jerárquica: MLwiN, [37], S-PLUS [32], SAS [26] y Stata [34].

Los datos con estructura jerárquica surgen en varias situaciones. Por ejemplo: como en las investigaciones educativas que se relacionan con problemas con la interacción entre los alumnos o los integrantes de un grupo. Generalmente los alumnos y el grupo de clase se conceptualizan como un sistema con estructura jerárquica, donde los alumnos y los grupos de clase son definidos en niveles separados de esta estructura. En general, supóngase que se tienen datos con 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; se tienen J unidades de nivel 2, y a cada unidad de las n_j unidades en cada grupo se le 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. Así también en investigaciones sociales se tratan problemas relacionados con la interacción entre los individuos en su contexto social, significando que las personas son influidas por los grupos sociales a los cuales pertenecen. Generalmente, los individuos y los grupos se conceptualizan como un sistema con estructura jerárquica, donde los individuos son las unidades de nivel 1 y los grupos sociales las unidades de nivel 2. En estudios sociales los miembros de una familia son las unidades de nivel 1 y las familias las unidades de nivel 2.

La teoría de espacios vectoriales de dimensión finita proporciona un marco de referencia para trabajar conceptos de inferencia en el modelo lineal general (*GLM*, por sus siglas en Inglés General Linear Model) que está dado por

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e},\ E(\mathbf{e}) = 0,\ Cov(\mathbf{e}) = \sigma^2\mathbf{V},$$

donde $\mathbf{V}$ es una matriz definida positiva. Conceptos como subespacio columna $\mathrm{S}(\mathbf{X})$; es decir, el espacio generado por los vectores columna de la matriz $\mathbf{X}$, operador proyector ortogonal u oblicuo, matriz inversa generalizada, entre otros, juegan un papel de suma importancia en la estadística teórica. En el criterio de mínimos cuadrados para definir el mejor ajuste, el estimador de mínimos cuadrados ordinarios del vector de coeficientes en el *GLM*, en el caso de que $\mathbf{V} = \mathbf{I}$, está relacionado con el operador proyector ortogonal $\mathbf{P_X}$ sobre el espacio $\mathrm{S}(\mathbf{X})$ o el estimador de mínimos cuadrados generalizados en términos del operador proyector oblicuo $\mathbf{P_{XV}}$. Para un estudio del operador oblicuo y de sus propiedades ver [43]. Así mismo las matrices inversas generalizadas juegan un papel importante en la representación de los estimadores de los parámetros de interés en el *GLM*. Otro parámetro de interés en el *GLM* es la varianza σ^2; así entonces las estimaciones insesgadas de la varianza del error se relacionan con el operador proyector ortogonal $\mathbf{P_X}$, o con el operador proyector oblicuo $\mathbf{P_{XV}}$ (ver [6]).

Las investigaciones en estimación en áreas pequeñas han recibido más atención en años recientes, debido a la creciente demanda de estadísticas confiables para este tipo de dominios de estudio, que son cualquier subdivisión de la población para la variable de interés, por ejemplo; regiones geográficas (estados, municipios, localidades), un grupo obtenido a partir del cruce de factores demográficos edad, sexo, grupo racial, etc. Los campos de aplicación de la teoría de estimación en áreas pequeñas son diversos. Ver, por ejemplo, [1], [11], [12], [49], entre otros. En la estimación en áreas pequeñas el interés es, por lo general, la estimación de alguna característica poblacional para el área pequeña, tal como la media, un total o una proporción. Las encuestas en general proporcionan poca información sobre áreas pequeñas, ya que están diseñadas para producir estadísticas para grandes poblaciones. Así entonces, los estimadores directos basados en el diseño no son confiables debido al tamaño de la muestra de la área pequeña. Más aún en ocasiones no se tiene información sobre el área pequeña de interés, y por lo tanto los estimadores no pueden ser calculados. La teoría de estimación en áreas pequeñas está relacionada con la solución de tales problemas. La idea principal para mejorar un estimador basado en el diseño es usar información suplementaria, ya sea de valores de la variable de interés en otra área similar, valores pasados de la misma área o valores de otras variables que están relacionados a la variable de interés. Se han sugerido muchos métodos estadísticos durante los últimos veinte años para abordar el problema de áreas pequeñas. Ghosh y Rao [17] presentan una revisión de métodos para estimación en áreas pequeñas, cubriendo la literatura al año 1993 y Rao [35] cubre de 1994 al 2003. Lo principal de tales métodos es la especificación de modelos usados para describir ligas entre áreas pequeñas, y el uso de información suplementaria.

El modelo lineal general mixto (*GLMM*, por sus siglas en Inglés General Linear Mixed Model) ha sido usado para tratar el problema de estimación en áreas pequeñas. Algunos investigadores han considerado tal modelo para problemas de predicción en áreas pequeñas [1], [9], [11], [12], [13], [22], [30], [31], [33], [45], [49], entre otros.

Para la predicción de la media poblacional finita $\overline{Y}_j$ en la j-ésima área pequeña, Battese *et al.* [1] propusieron un modelo de regresión jerárquico, el modelo intercepto aleatorio para la estimación (o predicción) de la media del número de hectáreas plantadas con maíz y soya para 12 ciudades en el norte de Iowa. Estos autores hacen uso de datos de satélite para obtener información del número de píxeles plantados con maíz y soya; su modelo está dado por $Y_{ij} = X_{ij}\boldsymbol{\beta} + u_j + e_{ij}$,

$j = 1, ..., t$; $i = 1, ..., n_j$. Los errores aleatorios u_j se suponen $N(0, \sigma_u^2)$, independientes entre si y de los e_{ij}, los cuales son asimismo asumidos independientes $N(0, \sigma_e^2)$.

Dempster *et al.* [11] consideran datos de estudiantes de leyes, donde se usan dos variables explicatorias para predecir el promedio en la calificación del primer año escolar. Estos autores consideran un modelo lineal jerárquico con pendientes aleatorias. Con el fin de estimar el ingreso per capita de pequeñas áreas con población menor de 1000 habitantes, Fay y Herriot [13] hacen uso de variables explicatorias a nivel 2 con el fin de modelar el promedio de la variable respuesta; ellos no hacen uso de variables explicatorias a nivel 1.

Prasad y Rao [33] derivan el *BLUP* para la media poblacional finita $\overline{Y}_j$ en la j-ésima área pequeña, para cada uno de los modelos de los trabajos de Fay y Herriot [13], Dempster *et al.* [11] y Battese *et al.* [1], usando la teoría de Henderson [18] para el *GLMM*. Estos autores obtienen una estimación del *BLUP* para la media poblacional μ_j de la j- ésima área pequeña y posteriormente el *BLUP* de la media poblacional finita $\overline{Y}_j$ de la j-ésima área pequeña haciendo uso de la teoría de la predicción de Royall [39].

Holt and Moura [21], [22] hicieron uso del modelo de regresión lineal jerárquico de dos niveles con una variable explicatoria a nivel 1, para obtener el *BLUP* de la media poblacional finita $\overline{Y}_j$ de la j-ésima área pequeña. Paddison [30] hace uso de un modelo de regresión intercepto aleatorio y llevan a cabo un estudio para simular los datos usados por Battese *et al.* [1], con el fin de obtener una estimación de la media para la j-ésima área pequeña. Petrucci *et al.* [31] hacen uso del modelo de Fay y Herriot [13] para estimar la producción promedio de aceitunas por granja. Fabrizi *et al.* [12] consideran extensiones del modelo usado por Battese *et al.* [1]; es decir, el modelo intercepto aleatorio. El objetivo del trabajo es la estimación del promedio de ingreso domestico en regiones dentro de países. Torabi y Rao [45] proponen un nuevo estimador de la media de un área pequeña bajo un modelo de dos niveles.

En el enfoque de inferencia basado en el *GLMM* se considera el siguiente modelo para la j-ésima área pequeña

$$\mathbf{Y}_j = \mathbf{X}_j\boldsymbol{\beta} + \mathbf{Z}_j\mathbf{u}_j + \mathbf{e}_j,$$

la cual cuenta con N_j unidades. Sea $\mathbf{s}_j$ la muestra de n_j unidades en el área pequeña j, y r_j el número de las correspondientes $N_j - n_j$ unidades no muestreadas. Una vez que la muestra $\mathbf{s}_j$ ha sido obtenida se tiene la descomposición de la información en la j-ésima área pequeña, de la siguiente manera:

$$\mathbf{Y}_j = \begin{pmatrix} \mathbf{Y}_{j\mathbf{s}} \\ \mathbf{Y}_{j\mathbf{r}} \end{pmatrix}, \mathbf{X}_j = \begin{pmatrix} \mathbf{X}_{j\mathbf{s}} \\ \mathbf{X}_{j\mathbf{r}} \end{pmatrix}, \mathbf{Z}_j = \begin{pmatrix} \mathbf{Z}_{j\mathbf{s}} \\ \mathbf{Z}_{j\mathbf{r}} \end{pmatrix} \text{ y } \mathbf{V}_j = \begin{pmatrix} \mathbf{V}_{j\mathbf{ss}} & \mathbf{V}_{j\mathbf{sr}} \\ \mathbf{V}_{j\mathbf{rs}} & \mathbf{V}_{j\mathbf{rr}} \end{pmatrix},$$

donde las matrices tienen los ordenes: $\mathbf{Y}_{j\mathbf{s}}$ $n_j \times 1$; $\mathbf{Y}_{j\mathbf{r}}$ $r_j \times 1$; $\mathbf{X}_{j\mathbf{s}}$ $n_j \times p$; $\mathbf{X}_{j\mathbf{r}}$ $r_j \times p$; $\mathbf{Z}_{j\mathbf{s}}$ $n_j \times q$; $\mathbf{Z}_{j\mathbf{r}}$ $r_j \times q$; $\mathbf{V}_{j\mathbf{ss}}$ $n_j \times n_j$; $\mathbf{V}_{j\mathbf{sr}}$ $n_j \times r_j$; $\mathbf{V}_{j\mathbf{rs}}$ $r_j \times n_j$; y $\mathbf{V}_{j\mathbf{rr}}$ $r_j \times r_j$. La media de la población finita en la j-ésima área $\overline{Y}_j = N_j^{-1} \sum_{i=1}^{N_j} Y_{ij}$ se puede descomponer en la media obtenida de la muestra $\overline{Y}_{j\mathbf{s}}$ más la media de las unidades no muestreadas $\overline{Y}_{j\mathbf{r}}$. Por la teoría de Royall [39], se obtiene el *BLUP* de la media poblacional finita $\overline{Y}_j = N_j^{-1} \sum_{i=1}^{N_j} Y_{ij}$ de la j-ésima área pequeña, que está dado por

$$BLUP\left(\overline{Y}_j\right) = f_j\overline{Y}_{j\mathbf{s}} + (1-f_j)\left[\overline{\mathbf{X}}_{j\mathbf{r}}\ \hat{\boldsymbol{\beta}} + \overline{\mathbf{Z}}_{j\mathbf{r}}\ \hat{\mathbf{u}}_j\right],$$

donde $\overline{\mathbf{X}}_{j\mathbf{r}}$ y $\overline{\mathbf{Z}}_{j\mathbf{r}}$ son los vectores de medias para las $N_j - n_j$ unidades no muestreadas en la j-ésima área pequeña y $f_j = n_j/N_j$, es la fracción de muestreo de la j-ésima área.

El modelo intercepto ha sido usado como el modelo de punto de partida para llevar a cabo las estimaciones de la media poblacional finita $\overline{Y}_j$ en el contexto de los modelos lineales jerárquicos por varios autores. Bajo el modelo intercepto aleatorio el $BLUP$ de la media poblacional finita $\overline{Y}_j$ de la j-ésima área pequeña en el contexto de la teoría de predicción de Royall está dado por

$$f_j\overline{Y}_{j\mathbf{s}} + (1-f_j)\left[\overline{\mathbf{X}}_{j\mathbf{r}}\ \hat{\boldsymbol{\beta}}_{\mathbf{s}} + \hat{u}_{j\mathbf{s}}\right].$$

Existe una gran cantidad de literatura respecto a la temática de la teoría de estimación en el modelo lineal general. En particular del uso del álgebra lineal y de la relación existente entre el operador proyector $\mathbf{P}_{\mathbf{XV}}$ y el estimador del vector de los coeficientes de regresión $\boldsymbol{\beta}$ o de la varianza σ^2. Sin embargo, respecto al uso del operador proyector en la teoría de la predicción en poblaciones finitas, en particular en la teoría de estimación en áreas pequeñas haciendo uso del modelo lineal jerárquico, no existen trabajos. En este trabajo el objetivo principal es la caracterización del $BLUP$ de la media poblacional finita $\overline{Y}_j$ correspondiente a la j-ésima unidad de nivel 2, en el contexto del modelo intercepto aleatorio, ya que como se ha mecionado anteriormente este modelo es ampliamente utilizado en la teoría de estimación en áreas pequeñas. Esta caracterización está dada en términos de los operadores proyector oblicuo $\mathbf{P}_{\mathbf{XV}}$ y ortogonal $\mathbf{P}_{\mathbf{Z}}$, y de una transformación lineal $\mathbf{T}_{j\mathbf{s}}$ definidos sobre los subespacios $S(\mathbf{X})$, $S(\mathbf{Z})$ y $S(\mathbf{X}_j)$, generados por las matrices de diseño $\mathbf{X}$, $\mathbf{Z}$ y $\mathbf{X}_j$, respectivamente. Se trabaja principalmente con el modelo de dos niveles con dos componentes de la varianza y en particular con el modelo intercepto aleatorio. Para el modelo intercepto aleatorio el *BLUP* de la media poblacional finita $\overline{Y}_j$ de la j-ésima unidad de nivel 2, está dado por $f_j\overline{Y}_{j\mathbf{s}} + (1 - f_j)\left[\overline{\mathbf{X}}_{j\mathbf{r}}\ \overset{\wedge}{\boldsymbol{\beta}}_{\mathbf{s}} + \hat{u}_{j\mathbf{s}}\right]$. En esta expresión interviene el término $\overline{\mathbf{X}}_{j\mathbf{r}}\ \overset{\wedge}{\boldsymbol{\beta}}_{\mathbf{s}} + \hat{u}_{j\mathbf{s}}$. Así para poder obtener la caracterización del *BLUP* de la media poblacional finita para el modelo intercepto aleatorio, es necesario obtener la caracterización del término $\overline{\mathbf{X}}_{j\mathbf{r}}\ \overset{\wedge}{\boldsymbol{\beta}}_{\mathbf{s}} + \hat{u}_{j\mathbf{s}}$, lo cual es posible al obtenerse la caracterización del *BLUP* del efecto mixto $\mathbf{X}\boldsymbol{\beta}+\mathbf{Z}\mathbf{u}$. En consecuencia, en primer lugar se obtendrá la caracterización del *BLUP* del efecto mixto $\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u}$, en segundo lugar la caracterización del término $\overline{\mathbf{X}}_{j\mathbf{r}}\ \overset{\wedge}{\boldsymbol{\beta}}_{\mathbf{s}} + \hat{u}_{j\mathbf{s}}$ correspondiente a la j-ésima unidad de nivel 2 y por último la caracterización del *BLUP* de la media poblacional finita $\overline{Y}_j$ correspondiente a la j-ésima unidad de nivel 2. En todas las situaciones se considera el caso balanceado y el desbalanceado.

Se espera que está caracterización de los predictores en términos de operadores proyector permita estudiar las propiedades del predictor de un efecto mixto, así como las propiedades del predictor de la media poblacional finita, tal como sucede con los estimadores de los coeficientes de regresión en el modelo lineal general.

Contenido de la tesis

La tesis consta de 4 capítulos además de la introducción, organizada como sigue:

El capítulo 1 presenta el modelo lineal general mixto. En este capítulo se presentan, brevemente, el modelo lineal jerárquico, casos particulares de éste, así como el modelo lineal general mixto. En segundo lugar se presenta la teoría de estimación y predicción en el modelo lineal general mixto, y por último se presenta la teoría de predicción de la media poblacional finita $\overline{Y}_j$ correspondiente a la j-ésima unidad de nivel 2.

En el capítulo 2 se presenta la teoría relacionada con el operador proyector, y sus principales aplicaciones en la estimación de parámetros en el modelo lineal general. Cabe mencionar que los resultados presentados en este capítulo son resultados ya establecidos, es decir, ningún resultado de este capítulo es contribución de la tesis.

En el capítulo 3 se presenta la contribución de este trabajo. Se brinda la caracterización del *BLUP* de la media poblacional finita correspondiente a la j-ésima unidad de nivel 2 en términos de operadores proyector y de una transformación lineal. Este capítulo está dividido en seis partes. En las primeras tres se presenta la caracterización del *BLUP* de un efecto mixto de la forma $\mathbf{X}\boldsymbol{\beta} + \mathbf{Zu}$, la caracterización del término $\overline{\mathbf{X}}_{j\mathbf{r}} \hat{\boldsymbol{\beta}}_{\mathbf{s}} + \hat{u}_{j\mathbf{s}}$, y la caracterización del *BLUP* de la media poblacional finita $\overline{Y}_j$. En las últimas tres partes se presentan casos particulares del modelo lineal mixto, los cuales cumplen la condición presentada en las secciones anteriores de este capítulo, se aplican los resultados relacionados con la caracterización del $BLUP$ desarrollados en este capítulo al modelo intercepto aleatorio y se presentan la factibilidad del cumplimiento de los resultados obtenidos en situasiones reales, así como las limitaciones de la teoría desarrollada.

En el capítulo 4 se presentan las conclusiones del trabajo desarrollado en la tesis, así también se presentan recomendaciones para trabajos a futuro en esta misma temática.

1

Modelo lineal general mixto en la teoría de predicción

En este capítulo se presentan casos particulares del modelo lineal jerárquico, así como el modelo lineal general mixto. En segundo lugar se presenta la teoría de estimación y predicción en el modelo lineal general mixto, y por último la teoría de predicción de la media poblacional finita.

1.1 Modelos lineales jerárquicos

Para analizar datos que presentan una estructura jerárquica se tienen que emplear técnicas estadísticas que tomen en cuenta dicha estructura. En esta situación es razonable postular un modelo de regresión que considere una posible diferencia entre las unidades de nivel 2, es decir, plantear un modelo de regresión tal que, para cada unidad de nivel 2, se tengan diferentes coeficientes de regresión. Bajo esta situación el modelo lineal jerárquico de dos niveles permite simultáneamente hacer un estudio de unidades de nivel 1 y un estudio de unidades de nivel 2, tomando en cuenta variables explicatorias para las unidades de nivel 1 y variables explicatorias para las unidades de nivel 2. En los modelos lineales jerárquicos cada uno de los niveles de la estructura jerárquica es representado formalmente con su propio submodelo. En un sistema con estructura jerárquica puede ser de interés estudiar la relación existente entre una variable respuesta, la cual se mide en las unidades de nivel 1, y variables explicatorias las cuales se pueden medir en cada uno de los niveles de la estructura jerárquica. Se tiene que para cada una de las n_j unidades de nivel 1 en la j-ésima unidad de nivel 2 se registraron mediciones sobre una variable respuesta Y_{ij}, y sobre m variables explicatorias $X_1, X_2, ..., X_m$; éstas se denominan variables explicatorias a nivel 1. Además, se puede medir otro conjunto de variables explicatorias $W_{1j}, W_{2j}, ..., W_{qj}$ en cada una de las unidades de nivel 2, las que se denominan variables explicatorias a nivel 2.

A continuación se presentan algunos modelos lineales jeráquicos los cuales se usarán posteriormente.

1.1.1 Modelo intercepto aleatorio

El caso más simple de un modelo lineal jerárquico es el denominado modelo intercepto aleatorio. El modelo para la i-ésima unidad de nivel 1, la cual se encuentra en la j-ésima unidad de nivel 2, está dado por:

$$Y_{ij} = \beta_{00} + u_{0j} + e_{ij}; \quad i = 1, \ldots, n_j; \quad j = 1, \ldots, J,$$

$$E(e_{ij}) = 0; \quad Var(e_{ij}) = \sigma_e^2, \quad e_{ij} \sim N\left(0, \sigma_e^2\right), \tag{1.1.1}$$

$$E(u_{0j}) = 0 \quad \text{y} \quad Var(u_{0j}) = \sigma_{u_0}^2.$$

Los parámetros en el modelo (1.1.1) son tres: El coeficiente β_{00} y los componentes de la varianza σ_e^2 y $\sigma_{u_0}^2$. En este modelo la varianza de la variable respuesta es descompuesta como la suma de las varianzas a nivel 1, σ_e^2 y a nivel 2, $\sigma_{u_0}^2$,

$$Var\left(Y_{ij}\right) = \sigma_e^2 + \sigma_{u_0}^2. \tag{1.1.2}$$

El modelo para el nivel 1 está dado por:

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

y el modelo para el nivel 2 está dado por:

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

1.1.2 Modelo intercepto aleatorio con una variable explicatoria

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 a nivel 1. Con una variable explicatoria a nivel 1 el modelo intercepto aleatorio está dado por:

$$Y_{ij} = \beta_{00} + \beta_1 X_{ij} + u_{0j} + e_{ij}; \quad i = 1, \ldots, n_j; \quad j = 1, \ldots, J,$$

$$E\left(e_{ij}\right) = 0; \quad Var\left(e_{ij}\right) = \sigma_e^2, \quad e_{ij} \sim N\left(0, \sigma_e^2\right), \tag{1.1.3}$$

$$E\left(u_{0j}\right) = 0 \quad \text{y} \quad Var\left(u_{0j}\right) = \sigma_{u_0}^2.$$

El modelo (1.1.3) se denomina modelo intercepto aleatorio con una variable explicatoria. Los parámetros en el modelo (1.1.3) son cuatro: Los coeficientes de regresión β_{00} y β_1, y los componentes de la varianza σ_e^2 y $\sigma_{u_0}^2$. En este modelo la varianza de la variable respuesta es descompuesta como la suma de las varianzas a nivel 1, σ_e^2 y a nivel 2, $\sigma_{u_0}^2$,

$$Var\left(Y_{ij}\right) = \sigma_e^2 + \sigma_{u_0}^2. \tag{1.1.4}$$

El modelo para el nivel 1 está dado por:

$$Y_{ij} = \beta_{0j} + \beta_1 x_{ij} + e_{ij},$$

y el modelo para el nivel 2 está dado por:

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

1.1.3 Modelo de pendientes aleatorias

En el modelo lineal jerárquico intercepto aleatorio con 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 cada una de 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 varibles explicatorias y la variable respuesta no es la misma en todas las unidades de nivel 2. Como los coeficientes varían entre las unidades de nivel 2 se les denomina coeficientes aleatorios. Para el caso de una variable explicatoria a nivel 1 lo anterior se expresa en el siguiente modelo:

$$Y_{ij} = \beta_{00} + \beta_{10} X_{ij} + u_{0j} + u_{1j} X_{ij} + e_{ij},$$

$$i = 1, \ldots, n_j; \quad j = 1, \ldots, J,$$

$$E\left(e_{ij}\right) = 0, \quad Var\left(e_{ij}\right) = \sigma_e^2, \quad e_{ij} \sim N\left(0, \sigma_e^2\right), \tag{1.1.5}$$

$$E\left(u_{0j}\right) = 0, \quad Var\left(u_{0j}\right) = \sigma_{u_0}^2, \quad E\left(u_{1j}\right) = 0, \quad Var\left(u_{1j}\right) = \sigma_{u_1}^2$$

$$\text{y Cov}\left(u_{0j}, u_{1j}\right) = \sigma_{u_{01}},$$

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

Los parámetros en el modelo (1.1.5) son seis: Los coeficientes de regresión β_{00} y β_{10}, y los componentes de la varianza σ_e^2, $\sigma_{u_0}^2$, $\sigma_{u_1}^2$ y $\sigma_{u_{01}}$. En este modelo la varianza de la variable respuesta es descompuesta como:

$$Var\left(Y_{ij}\right) = \sigma_{u_0}^2 + \sigma_{u_1}^2 X_{ij}^2 + 2\sigma_{u_{01}} X_{ij} + \sigma_e^2. \tag{1.1.6}$$

De la ecuación (1.1.6) se tiene que en el modelo de pendientes aleatorias con una variable explicatoria a nivel 1 la varianza de la variable respuesta depende de la variable explicatoria a nivel 1, X_{ij}.

El modelo para el nivel 1 está dado por:

$$Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + e_{ij},$$

y el modelo para el nivel 2 está dado por:

$$\beta_{0j} = \beta_{00} + u_{0j} \quad \text{y} \quad \beta_{1j} = \beta_{10} + u_{1j}.$$

1.1.4 Modelo lineal general mixto

El modelo lineal general mixto está dado por:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e},$$

$$E(\mathbf{e}) = \mathbf{0}, \quad Var(\mathbf{e}) = \mathbf{R}, \tag{1.1.7}$$

$$E(\mathbf{u}) = \mathbf{0}, \quad Var(\mathbf{u}) = \mathbf{G} \quad \text{y} \quad Cov\left(\mathbf{e}, \mathbf{u^t}\right) = \mathbf{0},$$

donde $\mathbf{Y}$ es un vector perteneciente a $\mathbb{R}^n$, $\mathbf{X}$ es una matriz conocida de orden $n \times p$, $\boldsymbol{\beta}$ es un vector perteneciente a $\mathbb{R}^p$, $\mathbf{Z}$ es una matriz conocida de orden $n \times q$, y $\mathbf{e}$ y $\mathbf{u}$ están distribuidos independientemente con media cero y matriz de varianza y covarianza $\mathbf{R}$ y $\mathbf{G}$, respectivamente, tales matrices dependen de parámetros desconocidos llamados los componentes de la varianza, los cuales serán denotados por σ. Los modelos intercepto aleatorio y el modelo de pendientes aleatorias, presentados en las secciones anteriores, son casos particulares del *GLMM* como se muestra a continuación.

Para el modelo intercepto aleatorio sin variables explicatoria (1.1.1), el modelo para la j-ésima unidad de nivel 2 está dado por:

$$\mathbf{Y}_j = \mathbf{1}_{n_j}\beta_{00} + \mathbf{1}_{n_j}u_{0j} + \mathbf{e}_j, \quad j = 1, ..., J, \tag{1.1.8}$$

tomando $\boldsymbol{\beta} = \beta_{00}$, $\mathbf{X} = \mathbf{1}_J \otimes \mathbf{1}_{n_j}$, $\mathbf{Z} = \mathbf{I}_J \otimes \mathbf{1}_{n_j}$ y $\mathbf{u} = (u_{01}, ..., u_{0J})^t$, el modelo

(1.1.8) es de la forma $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$.

Considerando el modelo intercepto aleatorio con una variable explicatoria (1.1.3), definiendo a $\mathbf{X}_j^*$ como la matriz de variables explicatorias para la j-ésima unidad de nivel 2, el modelo para la j-ésima unidad de nivel 2 está dado por:

$$\mathbf{Y}_j = \mathbf{1}_{n_j}\beta_{00} + \mathbf{X}_j^*\beta_1 + \mathbf{1}_{n_j}u_{0j} + \mathbf{e}_j, \quad j = 1, ..., J, \tag{1.1.9}$$

tomando $\boldsymbol{\beta} = (\beta_0, \beta_1)^t$, $\mathbf{X}_j = \left(\mathbf{1}_{n_j} : \mathbf{X}_j^*\right)$, $\mathbf{Z}_j = \mathbf{1}_{n_j}$ el modelo (1.1.9) está dado por

$$\mathbf{Y}_j = \mathbf{X}_j\boldsymbol{\beta} + \mathbf{Z}_j u_{0j} + \mathbf{e}_j. \tag{1.1.10}$$

Además, definiendo a $\mathbf{X}^*$ como la matriz de las variables explicatorias de nivel 1, $\mathbf{X} = \left(\mathbf{1}_J \otimes \mathbf{1}_{n_j} : \mathbf{X}^*\right)$, $\mathbf{Z} = \mathbf{I}_J \otimes \mathbf{1}_{n_j}$ y $\mathbf{u} = (u_{01}, \ldots, u_{0J})^t$, el modelo (1.1.10) es de la forma $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$.

Para el modelo de pendientes aleatorias dado por (1.1.5), definiendo a $\mathbf{X}_j^*$ como la matriz de variables explicatorias para la j-ésima unidad de nivel 2, el modelo para la j-ésima unidad de nivel 2 está dado por:

$$\mathbf{Y}_j = \mathbf{1}_{n_j}\beta_{00} + \mathbf{X}_j^*\beta_{10} + \mathbf{1}_{n_j}u_{0j} + \mathbf{X}_j^*u_{1j} + \mathbf{e}_j, \quad j = 1, ..., J. \tag{1.1.11}$$

tomando $\boldsymbol{\beta} = (\beta_{00}, \beta_{10})^t$, $\mathbf{u}_j = (u_{0j}, u_{1j})^t$, $\mathbf{X}_j = \left(\mathbf{1}_{n_j} : \mathbf{X}_j^*\right)$, y $\mathbf{Z}_j = \left(\mathbf{1}_{n_j} : \mathbf{X}_j^*\right)$, el modelo (1.1.11) está dado por:

$$\mathbf{Y}_j = \mathbf{X}_j\boldsymbol{\beta} + \mathbf{Z}_j\mathbf{u}_j + \mathbf{e}_j. \quad j = 1, ..., J. \tag{1.1.12}$$

Además, definiendo a $\mathbf{X}$ como matriz de las variables explicatorias, $\mathbf{u} = (\mathbf{u}_1, \ldots, \mathbf{u}_J)^t$, y $\mathbf{Z} = \oplus\mathbf{Z}_j = \oplus\mathbf{X}_j$, el modelo (1.1.12) es de la forma $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$.

1.2 Estimación y predicción en el modelo lineal general mixto

En está sección se presenta la teoría relacionada con la estimación y predicción en el *GLMM*.

El *GLMM* se divide en dos partes; la parte fija y la parte aleatoria. La parte fija está compuesta por los coeficientes de regresión, los cuales forman el parámetro $\boldsymbol{\beta}$, mientras que la parte aleatoria está compuesta por los efectos aleatorios $\mathbf{u}$.

De los supuestos del *GLMM* dado por la ecuación (1.1.7) se tiene

$$E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta} \quad \text{y} \quad \mathbf{V} = Var(\mathbf{Y}) = \mathbf{Z}\mathbf{G}\mathbf{Z}^t + \mathbf{R}.$$

Uno de los objetivos en el análisis de datos con estructura jerárquica es la estimación de los coeficientes de la parte fija y la estimación (predicción) de los efectos aleatorios. Los estimadores para efectos aletorios son conocidos como predictores. Predictor es un término usado para distinguirlo de estimador, ya que este último se usa para los coeficientes de la parte fija, mientras que predictor es para efectos aleatorios. La predicción de efectos aleatorios tiene una larga historia la cual data desde los primeros trabajos de Henderson sobre genética animal, [19].

En una serie de trabajos Henderson desarrolla el método del mejor predictor lineal insesgado para modelos mixtos. Ver Robinson [38] para un análisis del

BLUP con ejemplos y aplicaciones. Para una revisión de la teoría de la predicción ver [3], [6], [42] y [46]. Henderson *et al.* [20] desarrollaron un conjunto de ecuaciones que proporcionan simultáneamente el estimador de mínimos cuadrados generalizados de $\boldsymbol{\beta}$ y el *BLUP* de $\mathbf{u}$. Éstas son conocidas como las ecuaciones del modelo mixto. Estas ecuaciones son derivadas por la maximización de la densidad conjunta de $\mathbf{Y}$ y $\mathbf{u}$, la cual está dada para, $Var\left(\mathbf{e}\right) = \mathbf{R}$ y $Var\left(\mathbf{u}\right) = \mathbf{G}$, por:

$$f\left(\mathbf{Y}, \mathbf{u}\right) = \frac{\exp\left\{-\frac{1}{2}\left[\left(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{u}\right)^{t}\mathbf{R}^{-1}\left(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{u}\right) + \mathbf{u}^{t}\mathbf{G}^{-1}\mathbf{u}\right]\right\}}{\left(2\pi\right)^{\frac{n+q}{2}}\left|\mathbf{R}\right|^{\frac{1}{2}}\left|\mathbf{G}\right|^{\frac{1}{2}}}.$$

Igualando a cero las derivadas parciales de $log\left(f\left(\mathbf{Y}, \mathbf{u}\right)\right)$ con respecto a $\boldsymbol{\beta}$ y a $\mathbf{u}$, se obtienen las ecuaciones

$$\mathbf{X}^{t}\mathbf{R}^{-1}\mathbf{X}\boldsymbol{\beta} + \mathbf{X}^{t}\mathbf{R}^{-1}\mathbf{Z}\mathbf{u} = \mathbf{X}^{t}\mathbf{R}^{-1}\mathbf{Y},$$

$$\mathbf{Z}^{t}\mathbf{R}^{-1}\mathbf{X}\boldsymbol{\beta} + \left(\mathbf{Z}^{t}\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1}\right)\mathbf{u} = \mathbf{Z}^{t}\mathbf{R}^{-1}\mathbf{Y}.$$

Las cuales se pueden expresar de la siguiente manera:

$$\left[\begin{array}{cc} \mathbf{X}^{t}\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}^{t}\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}^{t}\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}^{t}\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{array}\right]\left[\begin{array}{c} \boldsymbol{\beta} \\ \mathbf{u} \end{array}\right] = \left[\begin{array}{c} \mathbf{X}^{t}\mathbf{R}^{-1}\mathbf{Y} \\ \mathbf{Z}^{t}\mathbf{R}^{-1}\mathbf{Y} \end{array}\right].$$

Éstas se denominan Ecuaciones del Modelo Mixto.

Henderson explica que el desarrollo se lleva a cabo bajo el supuesto de normalidad de $\mathbf{e}$ y $\mathbf{u}$. Henderson sugiere que las estimaciones no deben ser llamadas de Máxima Verosimilitud ya que la función a ser maximizada no es una función de verosimilitud.

Para obtener estimaciones de $\boldsymbol{\beta}$ y $\mathbf{u}$, el método estándar es resolver las ecuaciones del modelo mixto (Henderson [19]). Las estimaciones están dadas por:

$$\stackrel{\wedge}{\boldsymbol{\beta}} = \left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^t\mathbf{V}^{-1}\mathbf{Y},$$

$$\stackrel{\wedge}{\mathbf{u}} = \mathbf{G}\mathbf{Z}^t\mathbf{V}^{-1}\left(\mathbf{Y} - \mathbf{X}\stackrel{\wedge}{\boldsymbol{\beta}}\right).$$

Además, de la estimación del parámetro $\boldsymbol{\beta}$ y de la prediccón de $\mathbf{u}$, es necesaria la estimación de combinaciones lineales de éstos, es decir, funciones de la forma $\mathbf{k}^t\boldsymbol{\beta} + \mathbf{m}^t\mathbf{u}$, para vectores de constantes $\mathbf{k}$ y $\mathbf{m}$, estas funciones se denominan efectos mixtos ya que son combinaciones de efectos fijos y efectos aleatorios. Henderson [18] obtiene el *BLUP* del efecto mixto $\mathbf{k}^t\boldsymbol{\beta} + \mathbf{m}^t\mathbf{u}$ bajo el *GLMM*, el *BLUP* de este efecto mixto está dado por:

$$\mathbf{k}^t\stackrel{\wedge}{\boldsymbol{\beta}} + \mathbf{m}^t\stackrel{\wedge}{\mathbf{u}}. \qquad (1.2.1)$$

1.3 Predicción de la media poblacional finita $\overline{Y}_j$

En está sección se presenta el teorema que proporciona el *BLUP* de la media poblacional finita $\overline{Y}_j$ de la j-ésima unidad de nivel 2, este teorema se deduce del teorema general de predicción debido a Royall [39].

Como se mencionó al principio, sea N el número de unidades en la población finita $\{U_1, U_2, ..., U_N\}$. Asociado con cada unidad de la población finita sea Y_i, $i = 1, 2, ..., N$, la medición de U_i y denótese por $\mathbf{Y}=(Y_1, Y_2, ..., Y_N)$ el vector de valores Y_i. Con el fin de obtener información acerca de alguna función de

$\mathbf{Y}$, θ=$\theta\left(\mathbf{Y}\right)$, se selecciona una muestra $\mathbf{s}$ de tamaño n de $\{U_1, U_2, ..., U_N\}$. Una vez que la muestra $\mathbf{s}$ ha sido obtenida y las mediciones realizadas se tiene la parte observada $\mathbf{Y_s}$ y la no observada $\mathbf{Y_r}$ del vector $\mathbf{Y}$, donde $\mathbf{Y_s} \in \mathbb{R}^n$ y donde $\mathbf{Y_r} \in \mathbb{R}^r$, con $r = N - n$. Así la cantidad $\theta\left(\mathbf{Y}\right)$ se puede expresar como $\theta\left(\mathbf{Y}\right)$=$\theta\left(\mathbf{Y_s}\right)$+$\theta\left(\mathbf{Y_s}, \mathbf{Y_r}\right)$, donde $\theta\left(\mathbf{Y_s}\right)$ es una función de la parte observada $\mathbf{Y_s}$ y $\theta\left(\mathbf{Y_s}, \mathbf{Y_r}\right)$ es una función de la parte observada $\mathbf{Y_s}$ y de la parte no observada $\mathbf{Y_r}$. Entonces el objetivo de inferencia se puede caracterizar como la obtención de una predicción de $\theta\left(\mathbf{Y_s}, \mathbf{Y_r}\right)$ a partir de $\mathbf{Y_s}$; es decir, predecir los valores no observados a partir de los datos en la muestra $\mathbf{s}$.

Se supone que asociado con cada unidad en la población finita $\{U_1, U_2, ..., U_N\}$ existe un valor aleatorio Y_i, $i = 1, 2, ..., N$ y M cantidades conocidas, denotadas por $X_{i1}, X_{i2}, ..., X_{iM}$. Los vectores $\mathbf{X}_i = (X_{i1}, X_{i2}, ..., X_{iM})$ forman la matriz $\mathbf{X}$ de orden N x M. La relación entre la variable Y_k y las variables $X_{k1}, X_{k2}, ..., X_{kM}$ se expresa por un modelo lineal

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \quad E\left(\mathbf{e}\right) = 0, \quad Cov\left(\mathbf{e}\right) = \mathbf{V}. \tag{1.3.1}$$

donde $\mathbf{V}$ denota a la matriz de varianzas y covarianzas. De (1.3.1) se tiene que $E\left(\mathbf{Y}\right) = \mathbf{X}\boldsymbol{\beta}$ y $Cov\left(\mathbf{Y}\right) = \mathbf{V}$ considerada definidad positiva. En este proceso se supone que los valores auxiliares X_i para cada unidad en la población son conocidos.

Una vez que se tiene la muestra $\mathbf{s}$ el vector $\mathbf{Y}$, y las matrices $\mathbf{X}$ y $\mathbf{V}$ se pueden descomponer por

$$\begin{pmatrix} \mathbf{Y_s} \\ \mathbf{Y_r} \end{pmatrix}, \begin{pmatrix} \mathbf{X_s} \\ \mathbf{X_r} \end{pmatrix} \text{ y } \begin{pmatrix} \mathbf{V_{ss}} & \mathbf{V_{sr}} \\ \mathbf{V_{rs}} & \mathbf{V_{rr}} \end{pmatrix},$$

donde las matrices tienen los ordenes: $\mathbf{Y_s}$ $n \times 1$; $\mathbf{Y_r}$ $r \times 1$; $\mathbf{X_s}$ $n \times p$; $\mathbf{X_r}$ $r \times p$; $\mathbf{V_{ss}}$ $n \times n$; $\mathbf{V_{sr}}$ $n \times r$; $\mathbf{V_{rs}}$ $r \times n$; $\mathbf{V_{rr}}$ $r \times r$. En el caso de funciones lineales $\theta(\mathbf{Y})$ el vector $\boldsymbol{\gamma}^t$ involucrado en $\boldsymbol{\gamma}^t\mathbf{Y}$ se descompone en las partes $\boldsymbol{\gamma}_s$ y $\boldsymbol{\gamma}_{\mathbf{r}}$. La estimación de la función $\theta(\mathbf{Y})$ de interés se obtiene haciendo uso del teorema de la predicción de Royall [39], quien propone un predictor de cantidades de interés para poblaciones finitas. Este teorema se enuncia a continuación.

Teorema 1.3.1. ***Teorema general de predicción***. *Entre todos los predictores de $\theta(\mathbf{Y})$, el BLUP de $\theta(\mathbf{Y})$ bajo el modelo (1.3.1) está dado por*

$$\boldsymbol{\gamma}_s^t\mathbf{Y_s} + \boldsymbol{\gamma}_{\mathbf{r}}^t\left[\mathbf{X_r}\,\overset{\wedge}{\boldsymbol{\beta}}_{\mathbf{s}} + \mathbf{V_{rs}}\mathbf{V_{ss}}^{-1}\left(\mathbf{Y_s} - \mathbf{X_s}\,\overset{\wedge}{\boldsymbol{\beta}}_{\mathbf{s}}\right)\right] \tag{1.3.2}$$

donde $\overset{\wedge}{\boldsymbol{\beta}}_{\mathbf{s}} = \left(\mathbf{X_s^t V_s^{-1} X_s}\right)^{-1}\mathbf{X_s^t V_s^{-1} Y_s}$.

Para una demostración ver [3], [39], [46].

Así el teorema general de predicción proporciona el *BLUP* para la cantidad poblacional de interés $\theta(\mathbf{Y})$ a partir de la muestra. Tal predictor está dado por $\boldsymbol{\gamma}_s^t\mathbf{Y_s} + \boldsymbol{\gamma}_{\mathbf{r}}^t\left[\mathbf{X_r}\,\overset{\wedge}{\boldsymbol{\beta}}_{\mathbf{s}} + \mathbf{V_{rs}}\mathbf{V_{ss}}^{-1}\left(\mathbf{Y_s} - \mathbf{X_s}\,\overset{\wedge}{\boldsymbol{\beta}}_{\mathbf{s}}\right)\right]$, donde $\overset{\wedge}{\boldsymbol{\beta}}_{\mathbf{s}}$ es el estimador de mínimos cuadrados generalizados obtenido a partir de la muestra $\mathbf{s}$. En este predictor, $\boldsymbol{\gamma}_{\mathbf{s}}^t\mathbf{Y_s}$ es la parte calculada a partir de la muestra $\mathbf{s}$ y

$$\boldsymbol{\gamma}_{\mathbf{r}}^t\left[\mathbf{X_r}\,\overset{\wedge}{\boldsymbol{\beta}}_{\mathbf{s}} + \mathbf{V_{rs}}\mathbf{V_{ss}}^{-1}\left(\mathbf{Y_s} - \mathbf{X_s}\,\overset{\wedge}{\boldsymbol{\beta}}_{\mathbf{s}}\right)\right],$$

es la parte que se debe predecir.

Como se ha visto uno de los enfoques para obtener la predicción de la media poblacional finita $\overline{Y}_j$ en la j-ésima área pequeña es el basado en el modelo lineal jerárquico. De la sección anterior el modelo lineal general mixto está dado por

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e} \tag{1.3.3}$$

que es un caso particular del *MLG* dado por (1.3.1).

Una vez que la muestra $\mathbf{s}$ ha sido obtenida se tiene la descomposición

$$\mathbf{Y} = \begin{pmatrix} \mathbf{Y_s} \\ \mathbf{Y_r} \end{pmatrix}, \mathbf{X} = \begin{pmatrix} \mathbf{X_s} \\ \mathbf{X_r} \end{pmatrix}, \mathbf{Z} = \begin{pmatrix} \mathbf{Z_s} \\ \mathbf{Z_r} \end{pmatrix} \text{ y } \mathbf{V} = \begin{pmatrix} \mathbf{V_{ss}} & \mathbf{V_{sr}} \\ \mathbf{V_{rs}} & \mathbf{V_{rr}} \end{pmatrix},$$

donde las matrices tienen los ordenes: $\mathbf{Y_s}$ $n \times 1$; $\mathbf{Y_r}$ $r \times 1$; $\mathbf{X_s}$ $n \times p$; $\mathbf{X_r}$ $r \times p$; $\mathbf{Z_s}$ $n \times q$; $\mathbf{Z_r}$ $r \times q$; $\mathbf{V_{ss}}$ $n \times n$; $\mathbf{V_{sr}}$ $n \times r$; $\mathbf{V_{rs}}$ $r \times n$; y $\mathbf{V_{rr}}$ $r \times r$. Así el modelo (1.3.3) se puede descomponer en dos partes: el modelo para la parte observada, dado por:

$$\mathbf{Y_s} = \mathbf{X_s}\boldsymbol{\beta} + \mathbf{Z_s}\mathbf{u} + \mathbf{e_s} \tag{1.3.4}$$

y el modelo para la parte no observada, dado por:

$$\mathbf{Y_r} = \mathbf{X_r}\boldsymbol{\beta} + \mathbf{Z_r}\mathbf{u} + \mathbf{e_r}. \tag{1.3.5}$$

Así también el modelo correspondiente a la j-ésima unidad de nivel 2 está dado por

$$\mathbf{Y}_j = \mathbf{X}_j\boldsymbol{\beta} + \mathbf{Z}_j\mathbf{u}_j + \mathbf{e}_j \tag{1.3.6}$$

la cual cuenta con N_j unidades en la población, sea $\mathbf{s}_j$ la muestra de n_j unidades y r_j el número de las correspondientes $N_j - n_j$ unidades no muestreadas. Una vez que la muestra $\mathbf{s}_j$ ha sido obtenida se tiene la descomposición en la j-ésima unidad de nivel 2

$$\mathbf{Y}_j = \begin{pmatrix} \mathbf{Y}_{j\mathbf{s}} \\ \mathbf{Y}_{j\mathbf{r}} \end{pmatrix}, \mathbf{X}_j = \begin{pmatrix} \mathbf{X}_{j\mathbf{s}} \\ \mathbf{X}_{j\mathbf{r}} \end{pmatrix}, \mathbf{Z}_j = \begin{pmatrix} \mathbf{Z}_{j\mathbf{s}} \\ \mathbf{Z}_{j\mathbf{r}} \end{pmatrix} \text{ y } \mathbf{V}_j = \begin{pmatrix} \mathbf{V}_{j\mathbf{ss}} & \mathbf{V}_{j\mathbf{sr}} \\ \mathbf{V}_{j\mathbf{rs}} & \mathbf{V}_{j\mathbf{rr}} \end{pmatrix},$$

donde las matrices tienen los ordenes: $\mathbf{Y}_{j\mathbf{s}}$ $n_j \times 1$; $\mathbf{Y}_{j\mathbf{r}}$ $r_j \times 1$; $\mathbf{X}_{j\mathbf{s}}$ $n_j \times p$; $\mathbf{X}_{j\mathbf{r}}$ $r_j \times p$; $\mathbf{Z}_{j\mathbf{s}}$ $n_j \times q$; $\mathbf{Z}_{j\mathbf{r}}$ $r_j \times q$; $\mathbf{V}_{j\mathbf{ss}}$ $n_j \times n_j$; $\mathbf{V}_{j\mathbf{sr}}$ $n_j \times r_j$; $\mathbf{V}_{j\mathbf{rs}}$ $r_j \times n_j$; y $\mathbf{V}_{j\mathbf{rr}}$ $r_j \times r_j$. Así el modelo (1.3.6) se descompone en dos partes: el modelo para la parte observada, dado por:

$$\mathbf{Y}_{j\mathbf{s}} = \mathbf{X}_{j\mathbf{s}}\boldsymbol{\beta} + \mathbf{Z}_{j\mathbf{s}}\mathbf{u}_j + \mathbf{e}_{j\mathbf{s}} \tag{1.3.7}$$

y el modelo para la parte no observada, dado por:

$$\mathbf{Y}_{j\mathbf{r}} = \mathbf{X}_{j\mathbf{r}}\boldsymbol{\beta} + \mathbf{Z}_{j\mathbf{r}}\mathbf{u}_j + \mathbf{e}_{j\mathbf{r}}. \tag{1.3.8}$$

En este trabajo la cantidad de interés $\boldsymbol{\theta}(\mathbf{Y})$ es la media poblacional finita $\overline{Y}_j$ correspondiente a la j-ésima unidad de nivel 2, el siguiente resultado proporciona el mejor predictor de la media poblacional finita $\overline{Y}_j$.

Teorema 1.3.2. *Entre todos los predictores de $\overline{Y}_j$, bajo el modelo (1.3.3), el BLUP de $\overline{Y}_j$ está dado por:*

$$f_j\overline{Y}_{j\mathbf{s}} + (1 - f_j)\left[\overline{\mathbf{X}}_{j\mathbf{r}}\ \hat{\boldsymbol{\beta}}_{\mathbf{s}} + \overline{\mathbf{Z}}_{j\mathbf{r}}\ \hat{\mathbf{u}}_{j\mathbf{s}}\right], \tag{1.3.9}$$

donde $\overline{\mathbf{X}}_{j\mathbf{r}}$ y $\overline{\mathbf{Z}}_{j\mathbf{r}}$ son los vectores de medias para las $N_j - n_j$ unidades no muestreadas en la j-ésima unidad de nivel 2 y $f_j = n_j/N_j$.

Demostración. La media poblacional finita de la j-ésima unidad de nivel 2, está dada por $\overline{Y}_j = \sum_{i=1}^{N_j} Y_{ij}$, ésta se puede descomponer como

$$\overline{Y}_j = \overline{Y}_{j\mathbf{s}} + \overline{Y}_{j\mathbf{r}} \tag{1.3.10}$$

donde $\overline{Y}_{j\mathbf{s}}$ y $\overline{Y}_{j\mathbf{r}}$ denotan la media muestral y la media correspondiente a las unidades no muestreadas, respectivamente.

En este caso $\boldsymbol{\gamma} = \frac{\mathbf{1}_N^{*j}}{N_j}$, donde $\mathbf{1}_N^{*j}$ denota un vector de 0's de orden $N \times 1$ con un 1 en las posiciones de las unidades de nivel 1 correspondientes a la j-ésima unidad de nivel 2. Definiendo

$$\boldsymbol{\gamma}_s = \frac{\mathbf{1}_n^{*j\mathbf{s}}}{N_j} = \frac{n_j}{N_j}\frac{\mathbf{1}_n^{*j\mathbf{s}}}{n_j} \text{ y } \boldsymbol{\gamma}_{\mathbf{r}} = \frac{\mathbf{1}_{N-n}^{*j\mathbf{r}}}{N_j} = \frac{r_j}{N_j}\frac{\mathbf{1}_{N-n}^{*j\mathbf{r}}}{r_j}, \tag{1.3.11}$$

donde $\mathbf{1}_n^{*j\mathbf{s}}$ denota un vector de 0's de orden $n \times 1$ con un 1 en las posiciones de las unidades de nivel 1 correspondientes a la j-ésima unidad de nivel 2 y que se encuentran en la muestra, mientras que $\mathbf{1}_{N-n}^{*j\mathbf{r}}$ denota un vector de 0's de orden $(N-n) \times 1$ con un 1 en las posiciones de las unidades de nivel 1 correspondientes a la j-ésima unidad de nivel 2 que no están en la muestra.

Además de los modelos dados por (1.3.4) y (1.3.5)

$$\mathbf{V}_{\mathbf{rs}} = \mathbf{Z}_{\mathbf{r}}\mathbf{G}\mathbf{Z}_{\mathbf{s}}^t. \tag{1.3.12}$$

Así sustituyendo (1.3.11) y (1.3.12) en (1.3.2), que está dada por

$$\boldsymbol{\gamma}_s^t \mathbf{Y_s} + \boldsymbol{\gamma}_\mathbf{r}^t \left[\mathbf{X_r} \stackrel{\wedge}{\boldsymbol{\beta}}_\mathbf{s} + \mathbf{V_{rs}} \mathbf{V_{ss}}^{-1} \left(\mathbf{Y_s} - \mathbf{X_s} \stackrel{\wedge}{\boldsymbol{\beta}}_\mathbf{s} \right) \right], \tag{1.3.13}$$

se tiene

$$\frac{n_j}{N_j} \frac{\mathbf{1}_n^{*j\mathbf{s}t}}{n_j} \mathbf{Y_s} + \frac{r_j}{N_j} \frac{\mathbf{1}_{N-n}^{*j\mathbf{r}t}}{r_j} \left[\mathbf{X_r} \stackrel{\wedge}{\boldsymbol{\beta}}_\mathbf{s} + \mathbf{Z_r} \mathbf{G} \mathbf{Z_s}^t \mathbf{V_{ss}}^{-1} \left(\mathbf{Y_s} - \mathbf{X_s} \stackrel{\wedge}{\boldsymbol{\beta}}_\mathbf{s} \right) \right], \tag{1.3.14}$$

donde

$$\mathbf{G} \mathbf{Z_s}^t \mathbf{V_{ss}}^{-1} \left(\mathbf{Y_s} - \mathbf{X_s} \stackrel{\wedge}{\boldsymbol{\beta}}_\mathbf{s} \right) = \stackrel{\wedge}{\mathbf{u}}_\mathbf{s} . \tag{1.3.15}$$

Se cumple

$$\frac{\mathbf{1}_n^{*j\mathbf{s}t}}{n_j} \mathbf{Y_s} = \overline{Y}_{j\mathbf{s}}, \quad \frac{\mathbf{1}_{N-n}^{*j\mathbf{r}t}}{r_j} \mathbf{X_r} = \overline{\mathbf{X}}_{j\mathbf{r}} \ \text{ y } \ \frac{\mathbf{1}_{N-n}^{*j\mathbf{r}t}}{r_j} \mathbf{Z_r} = \overline{\mathbf{Z}}_{j\mathbf{r}}^{\,*}, \tag{1.3.16}$$

donde $\overline{\mathbf{Z}}_{j\mathbf{r}}^{\,*}$ denota un vector de 0's de orden $1 \times q$ que tiene al vector $\overline{\mathbf{Z}}_{j\mathbf{r}}$ en las entradas correspondientes a la j-ésima unidad de nivel 2. Sustituyendo los términos en (1.3.16) en (1.3.14) se obtiene

$$\frac{n_j}{N_j} \overline{Y}_{j\mathbf{s}} + \frac{r_j}{N_j} \left[\overline{\mathbf{X}}_{j\mathbf{r}} \stackrel{\wedge}{\boldsymbol{\beta}}_\mathbf{s} + \overline{\mathbf{Z}}_{j\mathbf{r}}^{\,*} \stackrel{\wedge}{\mathbf{u}}_\mathbf{s} \right]. \tag{1.3.17}$$

Además se cumple

$$\overline{\mathbf{Z}}_{j\mathbf{r}}^{\,*} \stackrel{\wedge}{\mathbf{u}}_\mathbf{s} = \overline{\mathbf{Z}}_{j\mathbf{r}} \stackrel{\wedge}{\mathbf{u}}_{j\mathbf{s}} . \tag{1.3.18}$$

Definiendo $f_j = n_j / N_j$ en (1.3.17) se obtiene

$$f_j \overline{Y}_{j\mathbf{s}} + (1 - f_j) \left[\overline{\mathbf{X}}_{j\mathbf{r}} \stackrel{\wedge}{\boldsymbol{\beta}} + \overline{\mathbf{Z}}_{j\mathbf{r}} \stackrel{\wedge}{\mathbf{u}}_{j\mathbf{s}} \right]. \tag{1.3.19}$$

■

Como se ha mencionado el interés principal es la caracterización de la media poblacional finita $\overline{Y}_j$ correspondiente a la j-ésima unidad de nivel 2 para el modelo intercepto aleatorio, en este modelo la matriz $\mathbf{Z}_j$ es de la forma $\mathbf{1}_{N_j}$.

Teorema 1.3.3. *Bajo el modelo intercepto aleatorio, el BLUP de la media poblacional finita $\overline{Y}_j$ está dado por:*

$$f_j \overline{Y}_{j\mathbf{s}} + (1 - f_j) \left[\overline{\mathbf{X}}_{j\mathbf{r}} \overset{\wedge}{\boldsymbol{\beta}}_{\mathbf{s}} + \hat{u}_{j\mathbf{s}} \right]. \tag{1.3.20}$$

Demostración. Por el teorema 1.3.2, el *BLUP* de la media poblacional finita $\overline{Y}_j$ está dado por:

$$f_j \overline{Y}_{j\mathbf{s}} + (1 - f_j) \left[\overline{\mathbf{X}}_{j\mathbf{r}} \overset{\wedge}{\boldsymbol{\beta}}_{\mathbf{s}} + \overline{\mathbf{Z}}_{j\mathbf{r}} \overset{\wedge}{\mathbf{u}}_{j\mathbf{s}} \right], \tag{1.3.21}$$

donde $\overline{\mathbf{X}}_{j\mathbf{r}}$ y $\overline{\mathbf{Z}}_{j\mathbf{r}}$ son los vectores de medias para las $N_j - n_j$ unidades no muestreadas en la j-ésima unidad de nivel 2 y $f_j = n_j/N_j$.

Bajo el modelo intercepto aleatorio la matriz $\mathbf{Z}_j$ es de la forma $\mathbf{1}_{N_j}$, por lo tanto $\overline{\mathbf{Z}}_{j\mathbf{r}} = 1$, así (1.3.21) toma la forma:

$$f_j \overline{Y}_{j\mathbf{s}} + (1 - f_j) \left[\overline{\mathbf{X}}_{j\mathbf{r}} \overset{\wedge}{\boldsymbol{\beta}}_{\mathbf{s}} + \hat{u}_{j\mathbf{s}} \right]. \tag{1.3.22}$$

■

2

Teoría de proyectores

En este capítulo se presenta la teoría relacionada con el operador proyector. Se presenta como los estimadores de los parámetros que intervienen en el modelo lineal general se relacionan con el operador proyector. Cabe mencionar que la teoría presentada en este capítulo en ningún momento es contribución de este trabajo de tesis.

En la sección 2.1 se presenta la teoría relacionada con el operador proyector, así como sus propiedades, además se presenta la matriz inversa generalizada. En la sección 2.2 se presentan las principales aplicaciones del operador proyector, ortogonal u oblicuo, en la caracterización de los estimadores de los parámetros $\boldsymbol{\beta}$ y σ^2 que intervienen en el modelo lineal general.

2.1 Operador proyector y sus propiedades

Sean V_1 y W_1 subespacios de $\mathbb{R}^n$, tales que su suma directa es $\mathbb{R}^n$. Cualquier elemento $\alpha \in \mathbb{R}^n$, tiene descomposición única $\boldsymbol{\alpha} = \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2$, donde $\boldsymbol{\alpha}_1 \in V_1$ y $\alpha_2 \in W_1$. Denotemos por $\mathbf{P}$ el mapeo de $\mathbb{R}^n$ a V_1 dado por $\mathbf{P}(\alpha) = \alpha_1$. $\mathbf{P}$ es un operador lineal sobre $\mathbb{R}^n$.

Definición 2.1.1. *Al operador* $\mathbf{P}$ *se le denomina el* ***operador proyector*** *sobre* V_1 *a lo largo de* W_1.

A continuación presentamos algunas propiedades del operador proyector.

Teorema 2.1.1. *Un operador* $\mathbf{P}$ *sobre* $\mathbb{R}^n$ *es un operador proyector sobre algún subespacio* V_1 *si, y solo si, es idempotente, es decir, si* $\mathbf{PP} = \mathbf{P}$.

Demostración. Sea $\mathbf{P}$ el operador proyector sobre V_1 a lo largo de W_1. Si $\boldsymbol{\alpha}_1 \in V_1$, la proyección de $\boldsymbol{\alpha}_1$ sobre V_1 a lo largo de W_1 es $\boldsymbol{\alpha}_1$. Para cualquier $\boldsymbol{\alpha} \in \mathbb{R}^n$, se tiene $\boldsymbol{\alpha} = \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2$, donde $\boldsymbol{\alpha}_1 \in V_1$ y $\boldsymbol{\alpha}_2 \in W_1$. Tenemos

$$\mathbf{PP}(\boldsymbol{\alpha}) = \mathbf{P}(\boldsymbol{\alpha}_1) = \boldsymbol{\alpha}_1 = \mathbf{P}(\boldsymbol{\alpha}) \tag{2.1.1}$$

por lo que $\mathbf{P}$ es un operador idempotente. Inversamente, sea $\mathbf{P}$ un operador idempotente. Definamos V_1 como el conjunto de todos los vectores $\boldsymbol{\alpha} \in \mathbb{R}^n$, tales que $\mathbf{P}(\boldsymbol{\alpha}) = \boldsymbol{\alpha}$ y sea W_1 el conjunto de todos los vectores $\boldsymbol{\alpha} \in \mathbb{R}^n$ tales que $\mathbf{P}(\boldsymbol{\alpha}) = \mathbf{0}$; es decir,

$$V_1 = \{\boldsymbol{\alpha} \in \mathbb{R}^n \,|\mathbf{P}(\boldsymbol{\alpha}) = \boldsymbol{\alpha}\} \tag{2.1.2}$$

y

$$W_1 = \{\boldsymbol{\alpha} \in \mathbb{R}^n \,|\mathbf{P}(\boldsymbol{\alpha}) = \mathbf{0}\}. \tag{2.1.3}$$

Ahora supóngase $\boldsymbol{\alpha} \in V_1 \cap W_1$, entonces $\boldsymbol{\alpha} = \mathbf{0}$, así $V_1 \cap W_1 = \{\mathbf{0}\}$, además para cualquier $\boldsymbol{\alpha} \in \mathbb{R}^n$ se tiene que $\boldsymbol{\alpha} = \mathbf{P}(\boldsymbol{\alpha}) + (\mathbf{I} - \mathbf{P})\boldsymbol{\alpha}$. Definamos $\boldsymbol{\alpha}_1 = \mathbf{P}(\boldsymbol{\alpha})$ y $\boldsymbol{\alpha}_2 = (\mathbf{I} - \mathbf{P})\boldsymbol{\alpha}$, entonces $\boldsymbol{\alpha} = \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2$, y se tiene

$$\mathbf{P}\boldsymbol{\alpha}_1 = \mathbf{PP}(\boldsymbol{\alpha}) = \mathbf{P}(\boldsymbol{\alpha}) = \boldsymbol{\alpha}_1$$

por lo que $\boldsymbol{\alpha}_1 \in V_1$, y

$$\mathbf{P}(\boldsymbol{\alpha}_2) = \mathbf{0},$$

lo cual implica que $\boldsymbol{\alpha}_2 \in W_1$. Así $V_1 + W_1 = \mathbb{R}^n$. Así se tiene $V_1 \bigoplus W_1 = \mathbb{R}^n$. Además, como $\boldsymbol{\alpha} = \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2$, ésto implica que

$$\mathbf{P}(\boldsymbol{\alpha}) = \mathbf{P}(\boldsymbol{\alpha}_1) + \mathbf{P}(\boldsymbol{\alpha}_2) = \boldsymbol{\alpha}_1 + \mathbf{0} = \boldsymbol{\alpha}_1. \tag{2.1.4}$$

Entonces por la definición 2.1.1 de operador proyector se tiene de (2.1.4) que $\mathbf{P}$ es el operador proyector sobre V_1 a lo largo de W_1 ■

Teorema 2.1.2. *Sean V_1 y W_1 dos subespacios de $\mathbb{R}^n$, tales que $V_1 \bigoplus W_1 = \mathbb{R}^n$, y sea $\mathbf{P}$ un operador lineal sobre $\mathbb{R}^n$. Tenemos que $\mathbf{P}$ es el operador proyector sobre V_1 a lo largo de W_1, si, y solo si, $\mathbf{P}$ satisface*

$$\mathbf{P}(\boldsymbol{\alpha}) = \begin{cases} \boldsymbol{\alpha} \ si \ \boldsymbol{\alpha} \in V_1 \\ \mathbf{0} \ si \ \boldsymbol{\alpha} \in W_1. \end{cases} \tag{2.1.5}$$

Demostración. Supóngase se cumple (2.1.5) y denotemos con $\mathbf{P}_{V.W}$ al operador proyector sobre V_1 a lo largo de W_1 y con $\mathbf{P}_{W.V}$ al operador proyector sobre W_1 a lo largo de V_1. Para cualquier $\boldsymbol{\alpha} \in \mathbb{R}^n$, así $\boldsymbol{\alpha} = \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2$, donde $\boldsymbol{\alpha}_1 \in V_1$ y $\boldsymbol{\alpha} \in W_1$. Ahora $\boldsymbol{\alpha} = \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2 = \mathbf{P}_{V.W}\boldsymbol{\alpha}_1 + \mathbf{P}_{W.V}\boldsymbol{\alpha}_2$, de lo cual obtenemos

$$\mathbf{P}(\boldsymbol{\alpha}) = \mathbf{P}\mathbf{P}_{V.W}(\boldsymbol{\alpha}_1) + \mathbf{P}\mathbf{P}_{W.V}(\boldsymbol{\alpha}_2).$$

Además, tenemos que $\mathbf{P}_{V.W}(\boldsymbol{\alpha}_1) \in V_1$ y $\mathbf{P}_{W.V}(\boldsymbol{\alpha}_2) \in W_1$, así por la condición (2.1.5) tenemos

$$\mathbf{P}\boldsymbol{\alpha} = \mathbf{P}_{V.W}(\boldsymbol{\alpha}_1) = \boldsymbol{\alpha}_1,$$

así $\mathbf{P}(\boldsymbol{\alpha}) = \boldsymbol{\alpha}_1$, con lo que se tiene por definición de operador proyector, que $\mathbf{P}$ es el operador proyector sobre V_1 a lo largo de W_1. Inversamente, sea $\mathbf{P}$ el operador proyector sobre V_1 a lo largo de W_1, y sean los vectores $\boldsymbol{\alpha}_1 \in V_1$ y $\boldsymbol{\alpha}_2 \in W_1$, así se tiene que

$$\mathbf{P}(\boldsymbol{\alpha}_1) = \mathbf{P}(\boldsymbol{\alpha}_1 + \mathbf{0}) = \boldsymbol{\alpha}_1 \quad \text{y} \quad \mathbf{P}(\boldsymbol{\alpha}_2) = \mathbf{P}(\mathbf{0} + \boldsymbol{\alpha}_2) = \mathbf{0}.$$

Entonces $\mathbf{P}$ satisface

$$\mathbf{P}(\boldsymbol{\alpha}) = \begin{cases} \boldsymbol{\alpha} \text{ si } \boldsymbol{\alpha} \in V_1 \\ \mathbf{0} \text{ si } \boldsymbol{\alpha} \in W_1 \end{cases}$$

■

Se denotará por $\mathbf{I}$ el operador identidad sobre el espacio vectorial $\mathbb{R}^n$.

Nota 2.1.1. *Si* $\mathbf{P}$ *es el operador proyector sobre* V_1 *a lo largo de* W_1*, entonces por el teorema 2.1.2 tenemos que* $(\mathbf{I} - \mathbf{P})$ *es el operador proyector sobre* W_1 *a lo largo de* V_1*. Además,* $\mathbf{P}(\mathbf{I} - \mathbf{P}) = (\mathbf{I} - \mathbf{P})\mathbf{P} = \mathbf{0}$.

Definición 2.1.2. *Sea* V *un subespacio del espacio* $\mathbb{R}^n$*, y sea* $V^\perp$ *su complemento ortogonal. Entonces el operador proyector sobre* V *a lo largo de* $V^\perp$*, se denomina* ***proyector ortogonal*** *sobre* V*.*

Definición 2.1.3. *Sea* $\mathbf{X}$ *una matriz de orden* $n \times m$*, se denotará por* $S(\mathbf{X})$ *al espacio generado por los vectores columna de la matriz* $\mathbf{X}$*.*

Teorema 2.1.3. $\mathbf{P}$ *es el operador proyector ortogonal sobre* $S(\mathbf{X})$ *sí y sólo sí*

a) $\mathbf{P}(\boldsymbol{\alpha})=\boldsymbol{\alpha}$ *para cualquier* $\boldsymbol{\alpha} \in S(\mathbf{X})$ *y* *b)* $\mathbf{P}\left(\boldsymbol{\beta}\right)=\mathbf{0}$ *para cualquier* $\boldsymbol{\beta}\in S\left(\mathbf{X}\right)^{\perp}$

Demostración. Se tiene que $S\left(\mathbf{X}\right)\oplus\ S\left(\mathbf{X}\right)^{\perp}\ =\mathbb{R}^{n}$, así por la definición de proyector ortogonal y por el teorema 2.1.2 se cumple el resultado $\blacksquare$

Matriz inversa generalizada y sus propiedades

Definición 2.1.4. *Dada una matriz* $\mathbf{A}$ *una matriz inversa generalizada de la matriz* $\mathbf{A}$ *es cualquier matriz* $\mathbf{G}$ *la cual satisface* $\mathbf{AGA}=\mathbf{A}$.

La notación usada para denotar una inversa generalizada de la matriz $\mathbf{A}$ es $\mathbf{A}^{-}$. Para un estudio de la teoría y aplicaciones de la inversa generalizada de una matriz ver [2].

Teorema 2.1.4. *Si* $\mathbf{G}$ *es una matriz inversa generalizada de* $\mathbf{X}^{t}\mathbf{X}$*, entonces* $\mathbf{GX}^{t}$ *es una matriz inversa generalizada de* $\mathbf{X}$*, es decir, se cumple* $\mathbf{XGX}^{t}\mathbf{X}=\mathbf{X}$.

Demostración. Sea $\boldsymbol{\alpha}\in\mathbb{R}^{n}$, éste tiene descomposición única $\boldsymbol{\alpha}=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2$, donde $\boldsymbol{\alpha}_1\in S\left(\mathbf{X}\right)$ y $\boldsymbol{\alpha}_2\in S\left(\mathbf{X}\right)^{\perp}$. Así $\boldsymbol{\alpha}_1=\mathbf{X}\boldsymbol{\gamma}_1$ para algún $\boldsymbol{\gamma}_1$. Entonces

$$\begin{aligned}\boldsymbol{\alpha}^{t}\mathbf{XGX}^{t}\mathbf{X} &= \boldsymbol{\alpha}_1^{t}\mathbf{XGX}^{t}\mathbf{X}\\ &= \boldsymbol{\gamma}_1^{t}\left(\mathbf{X}^{t}\mathbf{X}\right)\mathbf{G}\left(\mathbf{X}^{t}\mathbf{X}\right)\\ &= \boldsymbol{\gamma}_1^{t}\left(\mathbf{X}^{t}\mathbf{X}\right)\\ &= \boldsymbol{\alpha}^{t}\mathbf{X}.\end{aligned}$$

Así se cumple $\mathbf{XGX}^{t}\mathbf{X}=\mathbf{X}$ $\blacksquare$

Teorema 2.1.5. *Sea* $(\mathbf{X}^t\mathbf{X})^-$ *una inversa generalizada de la matriz* $(\mathbf{X}^t\mathbf{X})$*, entonces* $\mathbf{X}(\mathbf{X}^t\mathbf{X})^-\mathbf{X}^t$ *es el operador proyector ortogonal sobre* $S(\mathbf{X})$*.*

Demostración. Si $\boldsymbol{\alpha} \in S(\mathbf{X})$ entonces $\boldsymbol{\alpha} = \mathbf{X}\boldsymbol{\gamma}$ y por el teorema 2.1.4 se tiene

$$
\begin{aligned}
\mathbf{X}\left(\mathbf{X}^t\mathbf{X}\right)^-\mathbf{X}^t(\boldsymbol{\alpha}) &= \mathbf{X}\left(\mathbf{X}^t\mathbf{X}\right)^-\mathbf{X}^t\mathbf{X}\boldsymbol{\gamma} \\
&= \mathbf{X}\boldsymbol{\gamma} = \boldsymbol{\alpha}.
\end{aligned}
$$

Ahora si $\boldsymbol{\beta} \in S(\mathbf{X})^\perp$ se tiene que $\mathbf{X}(\mathbf{X}^t\mathbf{X})^-\mathbf{X}^t(\boldsymbol{\beta})=\mathbf{0}$. Así se cumple que

a) $\mathbf{X}(\mathbf{X}^t\mathbf{X})^-\mathbf{X}^t(\boldsymbol{\alpha})=\boldsymbol{\alpha}$ para cualquier $\boldsymbol{\alpha} \in S(\mathbf{X})$ y

b) $\mathbf{X}(\mathbf{X}^t\mathbf{X})^-\mathbf{X}^t(\boldsymbol{\beta})=\mathbf{0}$ para cualquier $\boldsymbol{\beta} \in S(\mathbf{X})^\perp$

y por el teorema 2.1.3 se tiene el resultado ■

Sea $\mathbf{V}$ una matriz definida positiva. Así existe una matriz no singular $\mathbf{R}$ para la cual se cumple $\mathbf{V} = \mathbf{R}\mathbf{R}^t$. Es de interés el operador proyector ortogonal $\mathbf{P}_\mathbf{R}$ sobre $S(\mathbf{R}^{-1}\mathbf{X})$, el cual, por el teorema 2.1.5 está dado por

$$
\left(\mathbf{R}^{-1}\mathbf{X}\right)\left[(\mathbf{R}^{-1}\mathbf{X})^t(\mathbf{R}^{-1}\mathbf{X})\right]^-\left(\mathbf{R}^{-1}\mathbf{X}\right)^t.
$$

Teorema 2.1.6. *Sea* $\mathbf{V}$ *una matriz definida positiva, entonces*

$$
\mathbf{X}\left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)^-\mathbf{X}^t\mathbf{V}^{-1}
$$

es un operador proyector sobre $S(\mathbf{X})$*.*

Demostración. Denotaremos por $\mathbf{P_{XV}}$ al operador $\mathbf{X}\left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}^t\mathbf{V}^{-1}$. En primer lugar veamos que efectivamente es un operador proyector. Se tiene

$$\begin{aligned}
\mathbf{P_{XV}}^2 &= \left[\mathbf{X}\left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}^t\mathbf{V}^{-1}\right]\left[\mathbf{X}\left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}^t\mathbf{V}^{-1}\right] \\
&= \left[\mathbf{X}\left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)^{-}\left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)\left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}^t\mathbf{V}^{-1}\right] \\
&= \mathbf{X}\left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}^t\mathbf{V}^{-1} = \mathbf{P_{XV}},
\end{aligned}$$

se tiene que se cumple la idempotencia, así por el teorema 2.1.1 $\mathbf{P_{XV}}$ es un operador proyector. Ahora veamos que es un proyector sobre $S\left(\mathbf{X}\right)$. Se tiene que el operador proyector ortogonal $\mathbf{P_R}$ sobre $S\left(\mathbf{R}^{-1}\mathbf{X}\right)$, cumple la relación

$$\left(\mathbf{R}^{-1}\mathbf{X}\right)\left[(\mathbf{R}^{-1}\mathbf{X})^t(\mathbf{R}^{-1}\mathbf{X})\right]^{-}\left(\mathbf{R}^{-1}\mathbf{X}\right)^t\left(\mathbf{R}^{-1}\mathbf{X}\right) = \left(\mathbf{R}^{-1}\mathbf{X}\right),$$

la cual puede ser escrita como:

$$\mathbf{R}^{-1}\mathbf{X}\left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X} = \mathbf{R}^{-1}\mathbf{X},$$

de lo que se obtiene

$$\mathbf{X}\left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X} = \mathbf{X}. \tag{2.1.6}$$

Si $\boldsymbol{\alpha} \in S\left(\mathbf{X}\right)$ entonces $\boldsymbol{\alpha} = \mathbf{X}\boldsymbol{\gamma}$. Entonces

$$\begin{aligned}
\mathbf{P_{XV}}\left(\boldsymbol{\alpha}\right) &= \mathbf{X}\left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}^t\mathbf{V}^{-1}\boldsymbol{\alpha} \\
&= \mathbf{X}\left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\boldsymbol{\gamma} \\
&= \mathbf{X}\boldsymbol{\gamma} = \boldsymbol{\alpha}.
\end{aligned}$$

Así se cumple $\mathbf{P_{XV}}\left(\boldsymbol{\alpha}\right) = \boldsymbol{\alpha}$ para $\boldsymbol{\alpha} \in S\left(\mathbf{X}\right)$ ■

2.2 Aplicaciones del operador proyector en la estadística

En está sección se presentan las principales aplicaciones del operador proyector en el *GLM*. Tanto en la estimación del parámetro $\boldsymbol{\beta}$, como en la estimación de la varianza σ^2.

2.2.1 Estimación por mínimos cuadrados

Considérese el modelo

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \qquad E(\mathbf{e}) = \mathbf{0}, \quad Cov(\mathbf{e}) = \sigma^2\mathbf{I}, \tag{2.2.1}$$

donde $\mathbf{Y} \in \mathbb{R}^n$, $\mathbf{X}$ es una matriz de constantes de orden $n \times p$, $\boldsymbol{\beta} \in \mathbb{R}^p$ es un vector de parámetros desconocidos, y $\mathbf{e} \in \mathbb{R}^n$ es un vector de errores aleatorios no observables.

El interés es la estimación del valor esperado de la variable respuesta, $E(\mathbf{Y})$. Se tiene que $E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}$, pero $\boldsymbol{\beta}$ es desconocido, lo que se sabe es que $E(\mathbf{Y}) \in S(\mathbf{X})$. Para la estimación de $E(\mathbf{Y})$ se toma el vector perteneciente a $S(\mathbf{X})$ que este más cercano a $\mathbf{Y}$. La distancia está dada en términos de $(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^t(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$. Así se tiene la siguiente definición.

Definición 2.2.1. *Una estimación $\hat{\boldsymbol{\beta}}$ se denomina una estimación por mínimos cuadrados de $\boldsymbol{\beta}$, si $\mathbf{X}\hat{\boldsymbol{\beta}}$ es el vector perteneciente a $S(\mathbf{X})$ que esta más cercano a $\mathbf{Y}$. En otras palabras, $\hat{\boldsymbol{\beta}}$ es una estimación mínimos cuadrados si*

$$\left(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}\right)^t \left(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}\right) = \min_{\boldsymbol{\beta}} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^t(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}). \tag{2.2.2}$$

Teorema 2.2.1. *Sea* $\mathbf{P_X}$ *el operador proyector ortogonal sobre* $S(\mathbf{X})$*, entonces*

$$(\mathbf{Y}-\mathbf{X}\boldsymbol{\beta})^t(\mathbf{Y}-\mathbf{X}\boldsymbol{\beta}) = (\mathbf{Y}-\mathbf{P_X}\mathbf{Y})^t(\mathbf{Y}-\mathbf{P_X}\mathbf{Y})+(\mathbf{P_X}\mathbf{Y}-\mathbf{X}\boldsymbol{\beta})^t(\mathbf{P_X}\mathbf{Y}-\mathbf{X}\boldsymbol{\beta}).$$

Demostración. Lo anterior se demuestra en base a las propiedades del operador proyector ortogonal, $(\mathbf{I}-\mathbf{P_X})\mathbf{P_X}=\mathbf{0}$ y $(\mathbf{I}-\mathbf{P_X})\mathbf{X}=\mathbf{0}$ ■

Teorema 2.2.2. $\overset{\wedge}{\boldsymbol{\beta}}$ *es un estimador por mínimos cuadrados de* $\boldsymbol{\beta}$ *sí y sólo sí* $\mathbf{X}\overset{\wedge}{\boldsymbol{\beta}}=\mathbf{P_X}\mathbf{Y}$*, donde* $\mathbf{P_X}$ *es el operador proyector ortogonal sobre* $S(\mathbf{X})$*, es decir si y sólo si* $\mathbf{X}\overset{\wedge}{\boldsymbol{\beta}}=\mathbf{X}(\mathbf{X}^t\mathbf{X})^-\mathbf{X}^t\mathbf{Y}$*.*

Demostración. Por el teorema 2.2.1 se cumple la relación

$$\begin{aligned}(\mathbf{Y}-\mathbf{X}\boldsymbol{\beta})^t(\mathbf{Y}-\mathbf{X}\boldsymbol{\beta}) &= (\mathbf{Y}-\mathbf{P_X}\mathbf{Y})^t(\mathbf{Y}-\mathbf{P_X}\mathbf{Y})\\ &+ (\mathbf{P_X}\mathbf{Y}-\mathbf{X}\boldsymbol{\beta})^t(\mathbf{P_X}\mathbf{Y}-\mathbf{X}\boldsymbol{\beta}).\end{aligned}$$

Ambos términos de la derecha son no negativos, y el primer término no depende de $\boldsymbol{\beta}$. $(\mathbf{Y}-\mathbf{X}\boldsymbol{\beta})^t(\mathbf{Y}-\mathbf{X}\boldsymbol{\beta})$ es minimizado cuando se minimiza

$$(\mathbf{P_X}\mathbf{Y}-\mathbf{X}\boldsymbol{\beta})^t(\mathbf{P_X}\mathbf{Y}-\mathbf{X}\boldsymbol{\beta}).$$

la cual es mínima sí y sólo sí $\mathbf{P_X}\mathbf{Y}=\mathbf{X}\boldsymbol{\beta}$, lo cual prueba el teorema ■

Teorema 2.2.3. $\boldsymbol{\beta}^*$ *dado por* $\boldsymbol{\beta}^*=(\mathbf{X}^t\mathbf{X})^-\mathbf{X}^t\mathbf{Y}$ *es un estimador mínimos cuadrados de* $\boldsymbol{\beta}$*.*

Demostración. Por el teorema 2.2.2 se tiene que $\boldsymbol{\beta}^*$ es un estimador por mínimos cuadrados de $\boldsymbol{\beta}$ sí y sólo sí $\mathbf{X}\boldsymbol{\beta}^* = \mathbf{P_X}\mathbf{Y}$, donde $\mathbf{P_X}$ es el operador proyector ortogonal sobre $S(\mathbf{X})$. En este caso se tiene

$$\mathbf{X}\boldsymbol{\beta}^* = \mathbf{X}\left[\left(\mathbf{X}^t\mathbf{X}\right)^{-}\mathbf{X}^t\mathbf{Y}\right] = \left[\mathbf{X}\left(\mathbf{X}^t\mathbf{X}\right)^{-}\mathbf{X}^t\right]\mathbf{Y} = \mathbf{P_X}\mathbf{Y}.$$

Por lo que $\boldsymbol{\beta}^* = \left(\mathbf{X}^t\mathbf{X}\right)^{-}\mathbf{X}^t\mathbf{Y}$ es un estimador mínimos cuadrados de $\boldsymbol{\beta}$ ■

Otro parámetro de interés en el modelo lineal general es la varianza σ^2. Ahora se verá como el operador proyector ortogonal interviene en la caracterización de la estimación del parámetro σ^2.

Teorema 2.2.4. *Sea el modelo*

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \qquad E(\mathbf{e}) = \mathbf{0}, \quad Cov(\mathbf{e}) = \sigma^2\mathbf{I}. \tag{2.2.3}$$

Denótese por $r(\mathbf{X})$ el rango de la matriz $\mathbf{X}$. Entonces

$$\frac{\mathbf{Y}^t(\mathbf{I} - \mathbf{P_X})\mathbf{Y}}{n - r(\mathbf{X})} \tag{2.2.4}$$

es un estimador insesgado de σ^2, respecto al modelo (2.2.1).

Demostración. Para la función cuadrática $\mathbf{Y}^t\mathbf{B}\mathbf{Y}$ se cumple la siguiente relación

$$E[\mathbf{Y}^t\mathbf{B}\mathbf{Y}] = tr(\mathbf{B}[Cov(\mathbf{Y})]) + [E(\mathbf{Y})]^t\mathbf{B}[E(\mathbf{Y})].$$

En el caso de interés se tiene $\mathbf{B} = \mathbf{I} - \mathbf{P_X}$, $E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}$ y $Cov(\mathbf{Y}) = \sigma^2$, así se tiene

$$E\left[\mathbf{Y}^t(\mathbf{I}-\mathbf{P_X})\mathbf{Y}\right] = tr(\sigma^2(\mathbf{I}-\mathbf{P_X})) + \boldsymbol{\beta}^t\mathbf{X}^t(\mathbf{I}-\mathbf{P_X})\mathbf{X}\boldsymbol{\beta}.$$

Además se cumple

$$tr(\sigma^2(\mathbf{I}-\mathbf{P_X})) = \sigma^2[tr(\mathbf{I}-\mathbf{P_X})] = \sigma^2(n - r\,(\mathbf{X}))$$

y

$$(\mathbf{I}-\mathbf{P_X})\mathbf{X} = \mathbf{0},$$

así $\boldsymbol{\beta}^t\mathbf{X}^t(\mathbf{I}-\mathbf{P_X})\mathbf{X}\boldsymbol{\beta} = \mathbf{0}$. De lo anterior

$$E[\mathbf{Y}^t(\mathbf{I}-\mathbf{P_X})\mathbf{Y}] = \sigma^2(n - r\,(\mathbf{X})).$$

Así

$$E\left[\frac{\mathbf{Y}^t(\mathbf{I}-\mathbf{P_X})\mathbf{Y}}{n - r\,(\mathbf{X})}\right] = \sigma^2$$

■

$\mathbf{Y}^t(\mathbf{I}-\mathbf{P_X})\mathbf{Y}$ se denomina la suma de cuadrados para el error (SCE). Al rango de $(\mathbf{I}-\mathbf{P_X})$ se le denomina los grados de libertad para el error (gl).

De lo anterior se tiene que una estimación insesgada de σ^2 está dada por la razón de la suma de cuadrados del error dividida entre sus grados de libertad.

2.2.2 Mínimos cuadrados generalizados

Ahora considérese el modelo

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \qquad E(\mathbf{e}) = \mathbf{0}, \quad Cov(\mathbf{e}) = \sigma^2\mathbf{V}, \tag{2.2.5}$$

donde $\mathbf{V}$ es una matriz definida positiva. Así existe una matriz no singular $\mathbf{R}$ para la cual se cumple $\mathbf{V} = \mathbf{R}\mathbf{R}^t$. Del modelo (2.2.5), se obtiene el modelo

$$\mathbf{R}^{-1}\mathbf{Y} = \mathbf{R}^{-1}\mathbf{X}\boldsymbol{\beta} + \mathbf{R}^{-1}\mathbf{e}, \quad E\left(\mathbf{R}^{-1}\mathbf{e}\right) = \mathbf{0}, \quad Cov\left(\mathbf{R}^{-1}\mathbf{e}\right) = \sigma^2\mathbf{I}. \tag{2.2.6}$$

Para el modelo (2.2.6) las estimaciones de mínimos cuadrados minimizan

$$\left(\mathbf{R}^{-1}\mathbf{Y} - \mathbf{R}^{-1}\mathbf{X}\boldsymbol{\beta}\right)^t \left(\mathbf{R}^{-1}\mathbf{Y} - \mathbf{R}^{-1}\mathbf{X}\boldsymbol{\beta}\right) = (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^t\, \mathbf{V}^{-1}\,(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})\,. \tag{2.2.7}$$

Está distancia es una generalización de la distancia $(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^t\,(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$, por lo que los estimadores del parámetro β que minimizan tal distancia se conocen como estimadores por **mínimos cuadrados generalizados**.

Teorema 2.2.5. *Bajo el modelo* (2.2.6), $\overset{\wedge}{\boldsymbol{\beta}}$ *es un estimador de mínimos cuadrados de* $\boldsymbol{\beta}$ *si y sólo si*

$$\mathbf{X}\overset{\wedge}{\boldsymbol{\beta}} = \mathbf{X}\left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}^t\mathbf{V}^{-1}\mathbf{Y}. \tag{2.2.8}$$

En términos del operador proyector $\mathbf{P_{XV}}$ del teorema 2.2.2, el teorema 2.2.5 se expresa como: Bajo el modelo (2.2.6), $\overset{\wedge}{\boldsymbol{\beta}}$ es un estimador por mínimos cuadrados de $\boldsymbol{\beta}$ sí y sólo sí $\mathbf{X}\overset{\wedge}{\boldsymbol{\beta}} = \mathbf{P_{XV}}\mathbf{Y}$ donde

$$\mathbf{P_{XV}} = \mathbf{X}\left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}^t\mathbf{V}^{-1} \tag{2.2.9}$$

es un proyector sobre $S\,(\mathbf{X})$.

En la sección correspondiente a mínimos cuadrados y estimación de σ^2 se vió que una estimación insesgada de σ^2 está dada por la razón de SCE dividida entre los grados de libertad. Lo anterior bajo el modelo propuesto.

Bajo el modelo (2.2.1) la SCE estaba dada en términos del operador proyector ortogonal $\mathbf{P_X}$ sobre $S\,(\mathbf{X})$ por $\mathbf{Y}^t(\mathbf{I}-\mathbf{P_X})\mathbf{Y}$.

En este caso el modelo de interés es el modelo (2.2.6) por lo cual se tiene que la SCE está dada en términos del operador proyector ortogonal $\mathbf{P_R}$ sobre $S\,(\mathbf{R}^{-1}\mathbf{X})$. El operador proyector ortogonal $\mathbf{P_R}$ sobre $S\,(\mathbf{R}^{-1}\mathbf{X})$ está dado por medio de

$$\left(\mathbf{R}^{-1}\mathbf{X}\right)\left[(\mathbf{R}^{-1}\mathbf{X})^t(\mathbf{R}^{-1}\mathbf{X})\right]^{-}\left(\mathbf{R}^{-1}\mathbf{X}\right)^t.$$

Así la SCE está dada por medio de $\left(\mathbf{R}^{-1}\mathbf{Y}\right)^t\left(\mathbf{I}-\mathbf{P_R}\right)\left(\mathbf{R}^{-1}\mathbf{Y}\right)$.

$$SCE = \left(\mathbf{R}^{-1}\mathbf{Y}\right)^t\left[\mathbf{I}-\mathbf{R}^{-1}\mathbf{X}\left[(\mathbf{R}^{-1}\mathbf{X})^t(\mathbf{R}^{-1}\mathbf{X})\right]^{-}\left(\mathbf{R}^{-1}\mathbf{X}\right)^t\right]\left(\mathbf{R}^{-1}\mathbf{Y}\right).$$

Teorema 2.2.6. *Bajo el modelo* (2.2.6)*, la SCE se puede expresar como*

$$\mathbf{Y}^t\left(\mathbf{I}-\mathbf{P_{XV}}\right)^t\mathbf{V}^{-1}\left(\mathbf{I}-\mathbf{P_{XV}}\right)\mathbf{Y}.$$

Demostración.

$$
\begin{aligned}
SCE &= \left(\mathbf{R}^{-1}\mathbf{Y}\right)^{t}\left[\mathbf{I}-\mathbf{R}^{-1}\mathbf{X}\left[(\mathbf{R}^{-1}\mathbf{X})^{t}(\mathbf{R}^{-1}\mathbf{X})\right]^{-}\left(\mathbf{R}^{-1}\mathbf{X}\right)^{t}\right]\left(\mathbf{R}^{-1}\mathbf{Y}\right)\\
&= \mathbf{Y}^{t}\mathbf{V}^{-1}\mathbf{Y}-\mathbf{Y}^{t}\mathbf{V}^{-1}\mathbf{X}\left(\mathbf{X}^{t}\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}^{t}\mathbf{V}^{-1}\mathbf{Y}\\
&= \mathbf{Y}^{t}\left(\mathbf{I}-\mathbf{P}_{\mathbf{XV}}\right)^{t}\mathbf{V}^{-1}\left(\mathbf{I}-\mathbf{P}_{\mathbf{XV}}\right)\mathbf{Y}
\end{aligned}
$$

■

Se tiene que $\mathbf{R}$ es una matriz no singular, así se cumple $r\left(\mathbf{R}^{-1}\mathbf{X}\right)=r\left(\mathbf{X}\right)$ y se tiene que los gl para el error bajo el modelo (2.2.6) están dados por medio de $n-r\left(\mathbf{X}\right)$.

De lo anterior se tiene que bajo el modelo (2.2.6) una estimación insesgada para σ^2 está dada por medio de

$$
\frac{\mathbf{Y}^{t}\left(\mathbf{I}-\mathbf{P}_{\mathbf{XV}}\right)^{t}\mathbf{V}^{-1}\left(\mathbf{I}-\mathbf{P}_{\mathbf{XV}}\right)\mathbf{Y}}{n-r\left(\mathbf{X}\right)}. \tag{2.2.10}
$$

3

Caracterización del mejor predictor lineal insesgado de la media poblacional finita $\overline{Y}_j$

En este capítulo se presenta la caracterización del *BLUP* de la media poblacional finita $\overline{Y}_j$, de la j-ésima unidad de nivel 2. Tal predictor está basado en la teoría de la predicción de Royall y en el *GLMM*, la caracterización se da en términos de operadores proyector y de una transformación lineal, definidas sobre los subespacios $S(\mathbf{X_s})$, $S(\mathbf{Z_s})$ y $S(\mathbf{X}_j)$ generados por las matrices de diseño $\mathbf{X_s}$, $\mathbf{Z_s}$ y $\mathbf{X}_j$, respectivamente.

Como se mencionó, en este trabajo el objetivo principal es la caracterización del *BLUP* de la media poblacional finita $\overline{Y}_j$ asociada a la j-ésima unidad de nivel 2, en el contexto del modelo intercepto aleatorio, ya que este modelo es utilizado en la teoría de estimación en áreas pequeñas. Esta caracterización está dada en términos del operador oblicuo $\mathbf{P_{XV}}$, del operador proyector ortogonal $\mathbf{P_Z}$, y de una transformación lineal $\mathbf{T}_{j\mathbf{s}}$ definidos sobre los subespacios $S(\mathbf{X})$, $S(\mathbf{Z})$ y $S(\mathbf{X}_j)$, generados por las matrices de diseño $\mathbf{X}$, $\mathbf{Z}$ y $\mathbf{X}_j$, respectivamente. Para este modelo el *BLUP* de la media poblacional finita $\overline{Y}_j$ de la j-ésima unidad de nivel 2, está dado por $f_j\overline{Y}_{j\mathbf{s}} + (1-f_j)\left[\overline{\mathbf{X}}_{j\mathbf{r}}\,\hat{\boldsymbol{\beta}}_{\mathbf{s}} + \hat{u}_{j\mathbf{s}}\right]$. En esta expresión interviene el término $\overline{\mathbf{X}}_{j\mathbf{r}}\,\hat{\boldsymbol{\beta}}_{\mathbf{s}} + \hat{u}_{j\mathbf{s}}$, así para poder obtener la caracterización del *BLUP* de la media poblacional finita, es necesario obtener la caracterización del término $\overline{\mathbf{X}}_{j\mathbf{r}}\,\hat{\boldsymbol{\beta}}_{\mathbf{s}} + \hat{u}_{j\mathbf{s}}$, lo cual es posible al obtenerse la caracterización del *BLUP* del efecto mixto $\mathbf{X}\boldsymbol{\beta}+\mathbf{Zu}$. En consecuencia, en primer lugar se obtendrá la caracterización del *BLUP* del efecto mixto $\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u}$, en segundo lugar la caracterización del término $\overline{\mathbf{X}}_{j\mathbf{r}}\, \overset{\wedge}{\boldsymbol{\beta}}_{\mathbf{s}} + \hat{u}_{j\mathbf{s}}$ correspondiente a la j-ésima unidad de nivel 2 y por último la caracterización del *BLUP* de la media poblacional finita $\overline{Y}_j$ de la j-ésima unidad de nivel 2.

En la sección 3.1 se presenta la caracterización del *BLUP* del efecto mixto $\mathbf{X}\boldsymbol{\beta}+\mathbf{Z}\mathbf{u}$ en términos de los operadores proyector definidos sobre los subespacios generados por las matrices de diseño $\mathbf{X}$ y $\mathbf{Z}$; en primer lugar se considera el caso balanceado y en segundo lugar se presenta el caso desbalanceado. En la sección 3.2 se presenta la caracterización del *BLUP* del efecto mixto $\overline{\mathbf{X}}_{j\mathbf{r}}\boldsymbol{\beta} + u_j$ de la j-ésima unidad de nivel 2, igualmente se considera el caso balanceado y el caso desbalanceado. En la sección 3.3 se presenta la caracterización del *BLUP* de la media poblacional finita $\overline{Y}_j$, de la j-ésima unidad de nivel 2, considerando el caso balanceado y el caso desbalanceado. En la sección 3.4 se presentan casos particulares del *GLMM* que satisfacen la condición $\mathbf{Z}\mathbf{Z}^t = d\mathbf{P}_{\mathbf{Z}}$ o la condición $n_j\mathbf{P}_{\mathbf{Z}_j} = \mathbf{Z}_j\mathbf{Z}_j^t$, por lo que es posible la aplicación a estos modelos de los resultados obtenidos en la secciones anteriores de este capítulo. En la sección 3.5 se presentan la caracterización del *BLUP* del efecto mixto $\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u}$, del *BLUP* del efecto mixto $\overline{\mathbf{X}}_{j\mathbf{r}}\boldsymbol{\beta} + u_j$ y de la *BLUP* de la media poblacional finita $\overline{Y}_j$, para el modelo intercepto aleatorio. Finalmente en la sección 3.6 se presentan la factibilidad del cumplimiento de los resultados obtenidos en este capítulo en situasiones reales, así como las limitaciones de la teoría desarrollada.

3.1 Caracterización del *BLUP* del efecto mixto $\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u}$

En está sección se presenta la condición que debe cumplir la matriz de diseño $\mathbf{Z}$ bajo la cual el *BLUP* del efecto mixto $\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u}$ se expresa en términos de los operadores proyector $\mathbf{P}_{\mathbf{XV}}$ y $\mathbf{P}_{\mathbf{Z}}$. Además, se presenta la caracterización del *BLUP* del efecto mixto en términos de los operadores $\mathbf{P}_{\mathbf{XV}}$ y $\mathbf{P}_{\mathbf{Z}}$.

En primer lugar, en la sección 3.1.1 se presenta la caracterización del *BLUP* del efecto mixto $\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u}$ considerando el caso balanceado y en la sección 3.1.2 considerando el caso desbalanceado.

En este capítulo se considera el modelo dado por:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e},$$

$$\mathbf{u} \sim N\left(\mathbf{0}, \sigma_{u_0}^2 \mathbf{I}_q\right), \qquad \mathbf{e} \sim N\left(\mathbf{0}, \sigma_e^2 \mathbf{I}_n\right), \tag{3.1.1}$$

$$Cov\left(\mathbf{e}, \mathbf{u^t}\right) = \mathbf{0},$$

donde $\mathbf{Y} \in \mathbb{R}^n$, $\mathbf{X}$ es una matriz conocida de orden $n \times p$, $\boldsymbol{\beta} \in \mathbb{R}^p$ es un vector de efectos fijos, $\mathbf{Z}$ es una matriz conocida de orden $n \times q$, y $\mathbf{u}$ es un vector de efectos aleatorios. En este caso la matriz de varianzas y covarianzas $\mathbf{V}$ de $\mathbf{Y}$ está dada por:

$$\sigma_{u_0}^2 \mathbf{Z}\mathbf{Z}^t + \sigma_e^2 \mathbf{I}_n.$$

Del epígrafe 1.2, el *BLUP* del efecto mixto $\mathbf{k}^t\boldsymbol{\beta} + \mathbf{m}^t\mathbf{u}$ bajo el *GLMM* está dado por:

$$\mathbf{k}^t \stackrel{\wedge}{\boldsymbol{\beta}} + \mathbf{m}^t \stackrel{\wedge}{\mathbf{u}} . \tag{3.1.2}$$

3.1.1 Caso balanceado

El siguiente resultado presenta la condición que debe de cumplir la matriz de diseño $\mathbf{Z}$ para que la matriz de varianzas y covarianzas $\mathbf{V}$ se exprese en términos del operador proyector ortogonal $\mathbf{P_Z}$ y de su complemento ortogonal $\mathbf{Q_Z}$, y se da la caracterización de la matriz $\mathbf{V}$ en términos de los proyectores mencionados.

Teorema 3.1.1. *Bajo el modelo (3.1.1), si* $\mathbf{ZZ}^t = d\mathbf{P_Z}$*,* $d \in \mathbb{R}$*, entonces la matriz de varianzas y covarianzas de* $\mathbf{Y}$ *se expresa en términos del operador proyector ortogonal* $\mathbf{P_Z}$ *y de su complemento ortogonal* $\mathbf{Q_Z}$ *por:*

$$\mathbf{V} = \left(d\sigma_{u_0}^2 + \sigma_e^2\right)\mathbf{P_Z} + \sigma_e^2\mathbf{Q_Z}.$$

Demostración. Bajo el modelo (3.1.1), si $\mathbf{ZZ}^t = d\mathbf{P_Z}$, entonces

$$\begin{aligned}
\mathbf{V} &= \sigma_{u_0}^2\mathbf{ZZ}^t + \sigma_e^2\mathbf{I}_n \\
&= d\sigma_{u_0}^2\mathbf{P_Z} + \sigma_e^2\left(\mathbf{P_Z} + \mathbf{Q_Z}\right) \\
&= d\sigma_{u_0}^2\mathbf{P_Z} + \sigma_e^2\mathbf{P_Z} + \sigma_e^2\mathbf{Q_Z} \\
&= \left(d\sigma_{u_0}^2 + \sigma_e^2\right)\mathbf{P_Z} + \sigma_e^2\mathbf{Q_Z}
\end{aligned}$$

■

El siguiente resultado presenta la caracterización de la matriz $\mathbf{V}^{-1}$, inversa de la matriz de varianzas y covarianzas, en términos del operador proyector ortogonal $\mathbf{P_Z}$ y de su complemento ortogonal $\mathbf{Q_Z}$.

Teorema 3.1.2. *Bajo el modelo (3.1.1), si* $\mathbf{ZZ}^t = d\mathbf{P_Z}$*,* $d \in \mathbb{R}$*, entonces la inversa* $\mathbf{V}^{-1}$ *de la matriz de varianzas y covarianzas de* $\mathbf{Y}$ *se expresa en términos del operador proyector ortogonal* $\mathbf{P_Z}$ *y de su complemento ortogonal* $\mathbf{Q_Z}$ *por:*

$$\mathbf{V}^{-1} = \frac{\mathbf{P_Z}}{\left(d\sigma_{u_0}^2 + \sigma_e^2\right)} + \frac{\mathbf{Q_Z}}{\sigma_e^2}.$$

Demostración. Por el teorema 3.1.1, si $\mathbf{ZZ}^t = d\mathbf{P_Z}$, entonces

$$\mathbf{V} = \left(d\sigma_{u_0}^2 + \sigma_e^2\right)\mathbf{P_Z} + \sigma_e^2\mathbf{Q_Z}.$$

Así la inversa $\mathbf{V}^{-1}$ se expresa por:

$$\mathbf{V}^{-1} = \frac{\mathbf{P_Z}}{\left(d\sigma_{u_0}^2 + \sigma_e^2\right)} + \frac{\mathbf{Q_Z}}{\sigma_e^2}$$

■

En el siguiente resultado se presenta la caracterización del *BLUP* del efecto mixto $\mathbf{X}\boldsymbol{\beta} + \mathbf{Zu}$ bajo el modelo (3.1.1), considerando el caso balanceado, en términos de operadores proyector definidos sobre los subespacios $S(\mathbf{X})$ y $S(\mathbf{Z})$.

Teorema 3.1.3. *Bajo el modelo (3.1.1), si* $\mathbf{ZZ}^t = d\mathbf{P_Z}$*,* $d \in \mathbb{R}$*, entonces el BLUP del efecto mixto* $\mathbf{X}\boldsymbol{\beta} + \mathbf{Zu}$ *se expresa en términos de los operadores* $\mathbf{P_{XV}}$ *y* $\mathbf{P_Z}$ *por:*

$$\mathbf{P_{XV}Y} \;+\; c\mathbf{P_Z}\left(\mathbf{I} - \mathbf{P_{XV}}\right)\mathbf{Y},$$

donde $c = d\sigma_{u_0}^2 / \left(d\sigma_{u_0}^2 + \sigma_e^2\right)$*.*

Demostración. Por (3.1.2), el *BLUP* del efecto mixto $\mathbf{X}\boldsymbol{\beta}+\mathbf{Zu}$ está dado por $\mathbf{X}\overset{\wedge}{\boldsymbol{\beta}}+\mathbf{Z}\overset{\wedge}{\mathbf{u}}$. Por el teorema 2.2.5, se tiene $\mathbf{X}\overset{\wedge}{\boldsymbol{\beta}}=\mathbf{P_{XV}Y}$. Además, del epígrafe 1.2, $\overset{\wedge}{\mathbf{u}}=\mathbf{GZ}^t\mathbf{V}^{-1}\mathbf{Q_{XV}Y}$. Por el teorema 3.1.2, bajo el modelo (3.1.1), si $\mathbf{ZZ}^t=d\mathbf{P_Z}$, entonces

$$\mathbf{V}^{-1}=\frac{\mathbf{P_Z}}{\left(d\sigma_{u_0}^2+\sigma_e^2\right)}+\frac{\mathbf{Q_Z}}{\sigma_e^2}.$$

Así

$$\begin{aligned}
BLUP\left(\mathbf{X}\boldsymbol{\beta}+\mathbf{Zu}\right) &= \mathbf{X}\overset{\wedge}{\boldsymbol{\beta}}+\mathbf{Z}\overset{\wedge}{\mathbf{u}}\\
&= \mathbf{P_{XV}Y}\ +\ \mathbf{ZGZ}^t\mathbf{V}^{-1}\mathbf{Q_{XV}Y}\\
&= \mathbf{P_{XV}Y}\ +\ \sigma_{u_0}^2\mathbf{ZZ}^t\mathbf{V}^{-1}\mathbf{Q_{XV}Y}\\
&= \mathbf{P_{XV}Y}\ +\ \sigma_{u_0}^2\mathbf{ZZ}^t\left[\frac{\mathbf{P_Z}}{\left(d\sigma_{u_0}^2+\sigma_e^2\right)}+\frac{\mathbf{Q_Z}}{\sigma_e^2}\right]\mathbf{Q_{XV}Y}\\
&= \mathbf{P_{XV}Y}\ +\ d\sigma_{u_0}^2\mathbf{P_Z}\left[\frac{\mathbf{P_Z}}{\left(d\sigma_{u_0}^2+\sigma_e^2\right)}+\frac{\mathbf{Q_Z}}{\sigma_e^2}\right]\mathbf{Q_{XV}Y}\\
&= \mathbf{P_{XV}Y}\ +\ \frac{d\sigma_{u_0}^2}{\left(d\sigma_{u_0}^2+\sigma_e^2\right)}\mathbf{P_ZQ_{XV}Y}\\
&= \mathbf{P_{XV}Y}\ +\ c\mathbf{P_Z}\left(\mathbf{I}-\mathbf{P_{XV}}\right)\mathbf{Y},
\end{aligned}$$

donde $c=d\sigma_{u_0}^2/\left(d\sigma_{u_0}^2+\sigma_e^2\right)$ ■

Corolario 3.1.1. *Bajo el modelo (3.1.1), si* $\mathbf{ZZ}^t=d\mathbf{P_Z}$, $d\in\mathbb{R}$, *entonces el BLUP del efecto aleatorio* $\mathbf{Zu}$ *se expresa en términos de los operadores* $\mathbf{P_{XV}}$ *y* $\mathbf{P_Z}$ *por:*

$$c\mathbf{P_Z}\left(\mathbf{I}-\mathbf{P_{XV}}\right)\mathbf{Y},$$

donde $c=d\sigma_{u_0}^2/\left(d\sigma_{u_0}^2+\sigma_e^2\right)$.

3.1.2 Caso desbalanceado

En está sección se presenta la caracterización del *BLUP* del efecto mixto $\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u}$ considerando el caso desbalanceado.

En el siguiente resultado se presenta la condición que debe de cumplir la matriz de diseño $\mathbf{Z}_j$ para que la matriz de varianzas y covarianzas $\mathbf{V}$ se exprese en términos de los operadores proyector $\mathbf{P}_{\mathbf{Z}_j}$ y $\mathbf{Q}_{\mathbf{Z}_j}$, así mismo se da la caracterización de la matriz $\mathbf{V}$ en términos de los operadores proyector $\mathbf{P}_{\mathbf{Z}_j}$ y $\mathbf{Q}_{\mathbf{Z}_j}$.

Teorema 3.1.4. *Bajo el modelo (3.1.1), si* $n_j\mathbf{P}_{\mathbf{Z}_j} = \mathbf{Z}_j\mathbf{Z}_j^t$, $n_j \in \mathbb{R}$, *entonces la matriz* $\mathbf{V}$ *se expresa en términos de los operadores proyector* $\mathbf{P}_{\mathbf{Z}_j}$ *y* $\mathbf{Q}_{\mathbf{Z}_j}$ *por* $\mathbf{V} = \oplus_{j=1}^{J}\left[\left(n_j\sigma_{u_0}^2 + \sigma_e^2\right)\mathbf{P}_{\mathbf{Z}_j} + \sigma_e^2\mathbf{Q}_{\mathbf{Z}_j}\right]$. *Donde* $\oplus$ *denota la suma directa.*

Demostración. Bajo el modelo (3.1.1), si $n_j\mathbf{P}_{\mathbf{Z}_j} = \mathbf{Z}_j\mathbf{Z}_j^t$, entonces

$$\begin{aligned}
\mathbf{V} &= \sigma_{u_0}^2\left(\oplus_{j=1}^{J}\mathbf{Z}_j\right)\left(\oplus_{j=1}^{J}\mathbf{Z}_j\right)^t + \sigma_e^2\mathbf{I}_n \\
&= \sigma_{u_0}^2\left(\oplus_{j=1}^{J}\mathbf{Z}_j\right)\left(\oplus_{j=1}^{J}\mathbf{Z}_j^t\right) + \sigma_e^2\mathbf{I}_n \\
&= \sigma_{u_0}^2\left(\oplus_{j=1}^{J}\mathbf{Z}_j\mathbf{Z}_j^t\right) + \sigma_e^2\mathbf{I}_n \\
&= \sigma_{u_0}^2\left(\oplus_{j=1}^{J}n_j\mathbf{P}_{\mathbf{Z}_j}\right) + \sigma_e^2\mathbf{I}_n \\
&= \oplus_{j=1}^{J}\left[n_j\sigma_{u_0}^2\mathbf{P}_{\mathbf{Z}_j} + \sigma_e^2\mathbf{I}_{n_j} + \sigma_e^2\mathbf{P}_{\mathbf{Z}_j} - \sigma_e^2\mathbf{P}_{\mathbf{Z}_j}\right] \\
&= \oplus_{j=1}^{J}\left[\left(n_j\sigma_{u_0}^2 + \sigma_e^2\right)\mathbf{P}_{\mathbf{Z}_j} + \sigma_e^2\mathbf{I}_{n_j} - \sigma_e^2\mathbf{P}_{\mathbf{Z}_j}\right] \\
&= \oplus_{j=1}^{J}\left[\left(n_j\sigma_{u_0}^2 + \sigma_e^2\right)\mathbf{P}_{\mathbf{Z}_j} + \sigma_e^2\mathbf{Q}_{\mathbf{Z}_j}\right]
\end{aligned}$$

■

Teorema 3.1.5. *Bajo el modelo (3.1.1), si* $n_j\mathbf{P}_{\mathbf{Z}_j} = \mathbf{Z}_j\mathbf{Z}_j^t$, $n_j \in \mathbb{R}$, *entonces* $\mathbf{V}^{-1}$ *se expresa en términos de los operadores proyector* $\mathbf{P}_{\mathbf{Z}_j}$ *y* $\mathbf{Q}_{\mathbf{Z}_j}$ *por:*

$$\mathbf{V}^{-1} = \oplus_{j=1}^{J}\left[\frac{\mathbf{P}_{\mathbf{Z}_j}}{\left(n_j\sigma_{u_0}^2+\sigma_e^2\right)} + \frac{\mathbf{Q}_{\mathbf{Z}_j}}{\sigma_e^2}\right].$$

Demostración. Del teorema 3.1.4, $\mathbf{V} = \oplus_{j=1}^{J}\left[\left(n_j\sigma_{u_0}^2+\sigma_e^2\right)\mathbf{P}_{\mathbf{Z}_j} + \sigma_e^2\mathbf{Q}_{\mathbf{Z}_j}\right]$. Así

$$\begin{aligned}
\mathbf{V}^{-1} &= \left\{\oplus_{j=1}^{J}\left[\left(n_j\sigma_{u_0}^2+\sigma_e^2\right)\mathbf{P}_{\mathbf{Z}_j} + \sigma_e^2\mathbf{Q}_{\mathbf{Z}_j}\right]\right\}^{-1} \\
&= \left\{\oplus_{j=1}^{J}\left[\left(n_j\sigma_{u_0}^2+\sigma_e^2\right)\mathbf{P}_{\mathbf{Z}_j} + \sigma_e^2\mathbf{Q}_{\mathbf{Z}_j}\right]^{-1}\right\} \\
&= \oplus_{j=1}^{J}\left[\frac{\mathbf{P}_{\mathbf{Z}_j}}{\left(n_j\sigma_{u_0}^2+\sigma_e^2\right)} + \frac{\mathbf{Q}_{\mathbf{Z}_j}}{\sigma_e^2}\right]
\end{aligned}$$

■

Lema 3.1.1. *Bajo el modelo (3.1.1), si* $n_j\mathbf{P}_{\mathbf{Z}_j} = \mathbf{Z}_j\mathbf{Z}_j^t$, $n_j \in \mathbb{R}$, *entonces*

$$\sigma_{u_0}^2\mathbf{Z}\mathbf{Z}^t\mathbf{V}^{-1} = \oplus_{j=1}^{J}\left[\frac{n_j\sigma_{u_0}^2\mathbf{P}_{\mathbf{Z}_j}}{\left(n_j\sigma_{u_0}^2+\sigma_e^2\right)}\right].$$

Demostración. Por el teorema 3.1.5, $\mathbf{V}^{-1} = \oplus_{j=1}^{J}\left[\frac{\mathbf{P}_{\mathbf{Z}_j}}{\left(n_j\sigma_{u_0}^2+\sigma_e^2\right)} + \frac{\mathbf{Q}_{\mathbf{Z}_j}}{\sigma_e^2}\right]$. Así

$$\begin{aligned}
\sigma_{u_0}^2\mathbf{Z}\mathbf{Z}^t\mathbf{V}^{-1} &= \sigma_{u_0}^2\mathbf{Z}\mathbf{Z}^t\left\{\oplus_{j=1}^{J}\left[\frac{\mathbf{P}_{\mathbf{Z}_j}}{\left(n_j\sigma_{u_0}^2+\sigma_e^2\right)} + \frac{\mathbf{Q}_{\mathbf{Z}_j}}{\sigma_e^2}\right]\right\} \\
&= \sigma_{u_0}^2\left\{\oplus_{j=1}^{J}n_j\mathbf{P}_{\mathbf{Z}_j}\right\}\left\{\oplus_{j=1}^{J}\left[\frac{\mathbf{P}_{\mathbf{Z}_j}}{\left(n_j\sigma_{u_0}^2+\sigma_e^2\right)} + \frac{\mathbf{Q}_{\mathbf{Z}_j}}{\sigma_e^2}\right]\right\} \\
&= \oplus_{j=1}^{J}\left[\frac{n_j\sigma_{u_0}^2\mathbf{P}_{\mathbf{Z}_j}}{\left(n_j\sigma_{u_0}^2+\sigma_e^2\right)}\right]
\end{aligned}$$

■

Teorema 3.1.6. *Bajo el modelo (3.1.1), si* $n_j\mathbf{P}_{\mathbf{Z}_j} = \mathbf{Z}_j\mathbf{Z}_j^t$*,* $n_j \in \mathbb{R}$*, entonces el BLUP del efecto mixto* $\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u}$ *se expresa en términos de los operadores proyector* $\mathbf{P}_{\mathbf{XV}}$ *y* $\mathbf{P}_{\mathbf{Z}}$*, por:*

$$[\mathbf{P}_{\mathbf{XV}} + \mathbf{P}_{\mathbf{Z}}\mathbf{B}\mathbf{Q}_{\mathbf{XV}}]\,\mathbf{Y},$$

donde $\mathbf{B} = \oplus_{j=1}^{J}\left(b_j\mathbf{I}_{n_j}\right)$ *y* $b_j = n_j\sigma_{u_0}^2/\left(n_j\sigma_{u_0}^2 + \sigma_e^2\right)$*.*

Demostración. Por el lema 3.1.1, si $n_j\mathbf{P}_{\mathbf{Z}_j} = \mathbf{Z}_j\mathbf{Z}_j^t$, entonces

$$\sigma_{u_0}^2\mathbf{Z}\mathbf{Z}^t\mathbf{V}^{-1} = \oplus_{j=1}^{J}\left[\frac{n_j\sigma_{u_0}^2\mathbf{P}_{\mathbf{Z}_j}}{\left(n_j\sigma_{u_0}^2 + \sigma_e^2\right)}\right].$$

De lo cual se tiene

$$\begin{aligned}
BLUP\left(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u}\right) &= \mathbf{X}\,\overset{\wedge}{\boldsymbol{\beta}} + \mathbf{Z}\,\overset{\wedge}{\mathbf{u}} \\
&= \mathbf{P}_{\mathbf{XV}}\mathbf{Y} + \mathbf{Z}\mathbf{G}\mathbf{Z}^t\mathbf{V}^{-1}\left(\mathbf{Y} - \mathbf{X}\,\overset{\wedge}{\beta}\right) \\
&= \mathbf{P}_{\mathbf{XV}}\mathbf{Y} + \sigma_{u_0}^2\mathbf{Z}\mathbf{Z}^t\mathbf{V}^{-1}\left(\mathbf{Y} - \mathbf{X}\,\overset{\wedge}{\beta}\right) \\
&= \mathbf{P}_{\mathbf{XV}}\mathbf{Y} + \left(\oplus_{j=1}^{J}\left[\frac{n_j\sigma_{u_0}^2\mathbf{P}_{\mathbf{Z}_j}}{\left(n_j\sigma_{u_0}^2 + \sigma_e^2\right)}\right]\right)\left(\mathbf{Y} - \mathbf{X}\,\overset{\wedge}{\beta}\right) \\
&= \mathbf{P}_{\mathbf{XV}}\mathbf{Y} + \mathbf{P}_{\mathbf{Z}}\left(\oplus_{j=1}^{J}\left[\frac{n_j\sigma_{u_0}^2}{\left(n_j\sigma_{u_0}^2 + \sigma_e^2\right)}\right]\right)\left(\mathbf{Y} - \mathbf{X}\,\overset{\wedge}{\beta}\right) \\
&= [\mathbf{P}_{\mathbf{XV}} + \mathbf{P}_{\mathbf{Z}}\mathbf{B}\mathbf{Q}_{\mathbf{XV}}]\,\mathbf{Y},
\end{aligned}$$

donde $\mathbf{B} = \oplus_{j=1}^{J}\left(b_j\mathbf{I}_{n_j}\right)$ y $b_j = n_j\sigma_{u_0}^2/\left(n_j\sigma_{u_0}^2 + \sigma_e^2\right)$ ■

Corolario 3.1.2. *Bajo el modelo (3.1.1), si* $n_j\mathbf{P}_{\mathbf{Z}_j} = \mathbf{Z}_j\mathbf{Z}_j^t$*,* $n_j \in \mathbb{R}$*, entonces el BLUP del efecto aleatorio* $\mathbf{Z}\mathbf{u}$ *se expresa en términos de los operadores* $\mathbf{P}_{\mathbf{XV}}$ *y* $\mathbf{P}_{\mathbf{Z}}$*, por* $\mathbf{P}_{\mathbf{Z}}\mathbf{B}\mathbf{Q}_{\mathbf{XV}}\mathbf{Y}$*, donde* $\mathbf{B} = \oplus_{j=1}^{J}\left(b_j\mathbf{I}_{n_j}\right)$ *y* $b_j = n_j\sigma_{u_0}^2/\left(n_j\sigma_{u_0}^2 + \sigma_e^2\right)$*.*

3.2 Caracterización del *BLUP* del efecto mixto $\overline{\mathbf{X}}_{j\mathbf{r}}\boldsymbol{\beta} + u_j$

Se presenta la caracterización del $BLUP$ del efecto mixto $\overline{\mathbf{X}}_{j\mathbf{r}}\boldsymbol{\beta} + u_j$, correspondiente a la j-ésima unidad de nivel 2, en términos de los operadores $\mathbf{P}_{\mathbf{XV}}$ y $\mathbf{P}_{\mathbf{Z}}$, y de una transformación lineal definida sobre el subespacio $S(\mathbf{X}_j)$ generado por la matriz de diseño $\mathbf{X}_j$. En la sección 3.2.1 se presenta la caracterización considerando el caso balanceado y en la sección 3.2.2 el caso desbalanceado.

Una vez que la muestra $\mathbf{s}$ ha sido obtenida el vector $\mathbf{Y}$, las matrices $\mathbf{X}$, $\mathbf{Z}$ y $\mathbf{V}$, los operadores $\mathbf{Q}_{\mathbf{XV}}$ y $\mathbf{P}_{\mathbf{Z}}$, la estimación del parámetro $\boldsymbol{\beta}$ y la predicción del efecto aleatorio u_j, se denotarán por $\mathbf{Y}_\mathbf{s}$, $\mathbf{X}_\mathbf{s}, \mathbf{Z}_\mathbf{s}, \mathbf{V}_\mathbf{s}$, $\mathbf{Q}_{\mathbf{X}_\mathbf{s}\mathbf{V}_\mathbf{s}}$, $\mathbf{P}_{\mathbf{Z}_\mathbf{s}}$, $\hat{\boldsymbol{\beta}}_\mathbf{s}$ y $\hat{u}_{j\mathbf{s}}$, respectivamente. A partir de está sección se considera que $\mathbf{Z}_j = \mathbf{1}_{N_j}$.

Asociada a la j-ésima unidad de nivel 2 está su matriz de diseño $\mathbf{X}_j$. Defínase $\mathbf{T}_{j\mathbf{s}}$ por

$$\mathbf{T}_{j\mathbf{s}} = \mathbf{X}_j \left(\mathbf{X}_\mathbf{s}^t \mathbf{V}_\mathbf{s}^{-1} \mathbf{X}_\mathbf{s}\right)^{-1} \mathbf{X}_\mathbf{s}^t \mathbf{V}_\mathbf{s}^{-1}. \tag{3.2.1}$$

Teorema 3.2.1. $\mathbf{T}_{j\mathbf{s}}$ *dada por (3.2.1), define una transformación lineal de* $\mathbb{R}^n$ *a* $\mathbb{R}^{N_j}$.

Demostración. Sean $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2 \in \mathbb{R}^n$ y $c \in \mathbb{R}$,

$$\begin{aligned}
\mathbf{T}_{j\mathbf{s}}(c\boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2) &= \mathbf{X}_j \left(\mathbf{X}_\mathbf{s}^t \mathbf{V}_\mathbf{s}^{-1} \mathbf{X}_\mathbf{s}\right)^{-1} \mathbf{X}_\mathbf{s}^t \mathbf{V}_\mathbf{s}^{-1} (c\boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2) \\
&= \mathbf{X}_j \left(\mathbf{X}_\mathbf{s}^t \mathbf{V}_\mathbf{s}^{-1} \mathbf{X}_\mathbf{s}\right)^{-1} \mathbf{X}_\mathbf{s}^t \mathbf{V}_\mathbf{s}^{-1} (c\boldsymbol{\alpha}_1) \\
&\quad + \mathbf{X}_j \left(\mathbf{X}_\mathbf{s}^t \mathbf{V}_\mathbf{s}^{-1} \mathbf{X}_\mathbf{s}\right)^{-1} \mathbf{X}_\mathbf{s}^t \mathbf{V}_\mathbf{s}^{-1} (\boldsymbol{\alpha}_2) \\
&= c\mathbf{T}_{j\mathbf{s}}(\boldsymbol{\alpha}_1) + \mathbf{T}_{j\mathbf{s}}(\boldsymbol{\alpha}_2)
\end{aligned}$$

■

$\mathbf{1}_n^{*j\mathbf{s}}$ denota un vector en $\mathbb{R}^n$ de 0's con un 1 en las posiciones correspondientes a las unidades de nivel 1 pertenecientes a la j-ésima unidad de nivel 2 en la muestra $\mathbf{s}$, mientras que $\mathbf{1}_{N_j}^{*j\mathbf{r}}$ denota un vector en $\mathbb{R}^{N_j}$ de 0's con un 1 en las posiciones correspondientes a las unidades de nivel 1 pertenecientes a la j-ésima unidad de nivel 2 que no están en la muestra.

3.2.1 Caso balanceado

En está sección se presenta la caracterización del *BLUP* del efecto mixto $\overline{\mathbf{X}}_{j\mathbf{r}}\boldsymbol{\beta} + u_j$, considerando el caso balanceado.

Teorema 3.2.2. *Bajo el modelo (3.1.1), con* $\mathbf{Z}_j = \mathbf{1}_{N_j}$*, si* $d\mathbf{P}_{\mathbf{Z}_\mathbf{s}} = \mathbf{Z}_\mathbf{s}\mathbf{Z}_\mathbf{s}^t$*,* $d \in \mathbb{R}$*, entonces el BLUP del efecto aleatorio* u_j *se expresa en términos de* $\mathbf{Q}_{\mathbf{X}_\mathbf{s}\mathbf{V}_\mathbf{s}}$ *y* $\mathbf{P}_{\mathbf{Z}_\mathbf{s}}$ *por:*

$$\frac{\mathbf{1}_n^{*j\mathbf{s}\,t}}{d}\left(c\mathbf{P}_{\mathbf{Z}_\mathbf{s}}\mathbf{Q}_{\mathbf{X}_\mathbf{s}\mathbf{V}_\mathbf{s}}\right)\mathbf{Y}_\mathbf{s},$$

donde $c = d\sigma_{u_0}^2 / \left(d\sigma_{u_0}^2 + \sigma_e^2\right)$*.*

Demostración. Por el corolario 3.1.1, si $d\mathbf{P}_{\mathbf{Z}_\mathbf{s}} = \mathbf{Z}_\mathbf{s}\mathbf{Z}_\mathbf{s}^t$, entonces el *BLUP* de $\mathbf{Z}_\mathbf{s}\mathbf{u}$ se expresa en términos de los operadores proyector $\mathbf{P}_{\mathbf{X}_\mathbf{s}\mathbf{V}_\mathbf{s}}$ y $\mathbf{P}_{\mathbf{Z}_\mathbf{s}}$ por:

$$c\mathbf{P}_{\mathbf{Z}_\mathbf{s}}\left(\mathbf{I} - \mathbf{P}_{\mathbf{X}_\mathbf{s}\mathbf{V}_\mathbf{s}}\right)\mathbf{Y}_\mathbf{s}, \tag{3.2.2}$$

donde $c = d\sigma_{u_0}^2 / \left(d\sigma_{u_0}^2 + \sigma_e^2\right)$. Además, la relación entre el *BLUP* de $\mathbf{Z}_\mathbf{s}\mathbf{u}$ y el *BLUP* del efecto aleatorio u_j está dada por:

$$\hat{u}_{j\mathbf{s}} = \ BLUP\left(u_j\right) = \frac{\mathbf{1}_n^{*j\mathbf{s}\,t}}{d} BLUP\left(\mathbf{Z}_\mathbf{s}\mathbf{u}\right). \tag{3.2.3}$$

Así de (3.2.2) y (3.2.3), el *BLUP* del efecto aleatorio u_j se expresa en términos de los operadores proyector $\mathbf{Q}_{\mathbf{X_s V_s}}$ y $\mathbf{P}_{\mathbf{Z_s}}$ por:

$$\frac{\mathbf{1}_n^{*j\mathbf{s}^t}}{d}\left(c\mathbf{P}_{\mathbf{Z_s}}\mathbf{Q}_{\mathbf{X_s V_s}}\right)\mathbf{Y_s}$$

■

Teorema 3.2.3. *Bajo el modelo (3.1.1), con* $\mathbf{Z}_j = \mathbf{1}_{N_j}$*, si* $d\mathbf{P}_{\mathbf{Z_s}} = \mathbf{Z_s}\mathbf{Z_s}^t$*,* $d \in \mathbb{R}$*, entonces el BLUP del efecto mixto* $\overline{\mathbf{X}}_{j\mathbf{r}}\boldsymbol{\beta} + u_j$*, de la* j*-ésima unidad de nivel 2, se expresa en términos de los operadores proyector* $\mathbf{Q}_{\mathbf{X_s V_s}}$ *y* $\mathbf{P}_{\mathbf{Z_s}}$*, y de la transformación lineal* $\mathbf{T}_{j\mathbf{s}}$ *por:*

$$\left[\frac{\mathbf{1}_{N_j}^{*j\mathbf{r}^t}}{r_j}\mathbf{T}_{j\mathbf{s}} \;+\; \frac{\mathbf{1}_n^{*j\mathbf{s}^t}}{d}\left(c\mathbf{P}_{\mathbf{Z_s}}\mathbf{Q}_{\mathbf{X_s V_s}}\right)\right]\mathbf{Y_s},$$

donde $c = d\sigma_{u_0}^2 / \left(d\sigma_{u_0}^2 + \sigma_e^2\right)$*.*

Demostración. El *BLUP* de $\overline{\mathbf{X}}_{j\mathbf{r}}\boldsymbol{\beta} + u_j$ está dado por $\overline{\mathbf{X}}_{j\mathbf{r}}\,\hat{\boldsymbol{\beta}}_{\mathbf{s}} + \hat{u}_{j\mathbf{s}}$, y por el teorema 3.2.2, se tiene

$$\begin{aligned}
\overline{\mathbf{X}}_{j\mathbf{r}}\,\hat{\boldsymbol{\beta}}_{\mathbf{s}} \;+\; \hat{u}_{j\mathbf{s}} &= \frac{\mathbf{1}_{N_j}^{*j\mathbf{r}^t}\mathbf{X}_j}{r_j}\,\hat{\boldsymbol{\beta}}_{\mathbf{s}} \;+\; \hat{u}_{j\mathbf{s}} \\
&= \frac{\mathbf{1}_{N_j}^{*j\mathbf{r}^t}}{r_j}\mathbf{X}_j\left(\mathbf{X_s}^t\mathbf{V_s}^{-1}\mathbf{X_s}\right)^{-1}\mathbf{X_s}^t\mathbf{V_s}^{-1}\mathbf{Y_s} + \hat{u}_{j\mathbf{s}} \\
&= \frac{\mathbf{1}_{N_j}^{*j\mathbf{r}^t}}{r_j}\mathbf{T}_{j\mathbf{s}}\mathbf{Y_s} \;+\; \frac{\mathbf{1}_n^{*j\mathbf{s}^t}}{d}\left(c\mathbf{P}_{\mathbf{Z_s}}\mathbf{Q}_{\mathbf{X_s V_s}}\right)\mathbf{Y_s} \\
&= \left[\frac{\mathbf{1}_{N_j}^{*j\mathbf{r}^t}}{r_j}\mathbf{T}_{j\mathbf{s}} \;+\; \frac{\mathbf{1}_n^{*j\mathbf{s}^t}}{d}\left(c\mathbf{P}_{\mathbf{Z_s}}\mathbf{Q}_{\mathbf{X_s V_s}}\right)\right]\mathbf{Y_s}
\end{aligned}$$

■

3.2.2 Caso desbalanceado

En está sección se presenta la caracterización del $BLUP$ del efecto mixto $\overline{\mathbf{X}}_{j\mathbf{r}}\boldsymbol{\beta}+u_j$ considerando el caso desbalanceado.

Teorema 3.2.4. *Bajo el modelo (3.1.1), con* $\mathbf{Z}_j = \mathbf{1}_{N_j}$*, si* $n_j\mathbf{P}_{\mathbf{Z}_{j\mathbf{s}}} = \mathbf{Z}_{j\mathbf{s}}\mathbf{Z}_{j\mathbf{s}}^t$*,* $n_j \in \mathbb{R}$*, entonces el BLUP del efecto aleatorio* u_j*, de la* j*-ésima unidad de nivel 2, se expresa en términos de los operadores* $\mathbf{Q}_{\mathbf{X_s V_s}}$ *y* $\mathbf{P}_{\mathbf{Z_s}}$ *por:*

$$\frac{\mathbf{1}_n^{*j\mathbf{s}t}}{n_j}\left(\mathbf{P}_{\mathbf{Z_s}}\mathbf{B_s}\mathbf{Q}_{\mathbf{X_s V_s}}\right)\mathbf{Y_s}, \tag{3.2.4}$$

donde $\mathbf{B_s} = \oplus_{j=1}^{J}\left(b_j\mathbf{I}_{n_j}\right)$ *y* $b_j = n_j\sigma_{u0}^2/\left(n_j\sigma_{u_0}^2+\sigma_e^2\right)$.

Demostración. Por el corolario 3.1.2, si $n_j\mathbf{P}_{\mathbf{Z}_{j\mathbf{s}}} = \mathbf{Z}_{j\mathbf{s}}\mathbf{Z}_{j\mathbf{s}}^t$ entonces el $BLUP$ de $\mathbf{Z_s u}$ se expresa en términos de $\mathbf{P}_{\mathbf{X_s V_s}}$ y $\mathbf{P}_{\mathbf{Z_s}}$ por

$$BLUP\left(\mathbf{Zu}\right) = \left(\mathbf{P}_{\mathbf{Z_s}}\mathbf{B_s}\left(\mathbf{I}-\mathbf{P}_{\mathbf{X_s V_s}}\right)\right)\mathbf{Y_s}, \tag{3.2.5}$$

donde $\mathbf{B_s} = \oplus_{j=1}^{J}\left(b_j\mathbf{I}_{n_j}\right)$ y $b_j = n_j\sigma_{u0}^2/\left(n_j\sigma_{u_0}^2+\sigma_e^2\right)$. Además, la relación entre el $BLUP$ de $\mathbf{Z_s u}$ y el $BLUP$ del efecto aleatorio u_j está dada por

$$\hat{u}_{j\mathbf{s}} = BLUP\left(u_j\right) = \frac{\mathbf{1}_n^{*j\mathbf{s}t}}{n_j}BLUP\left(\mathbf{Z_s u}\right). \tag{3.2.6}$$

Así de (3.2.5) y (3.2.6), el $BLUP$ del efecto aleatorio u_j se expresa en términos de los operadores proyector $\mathbf{Q}_{\mathbf{X_s V_s}}$ y $\mathbf{P}_{\mathbf{Z_s}}$ por

$$\frac{\mathbf{1}_n^{*jst}}{n_j}\left(\mathbf{P}_{\mathbf{Z_s}}\mathbf{B_s}\mathbf{Q}_{\mathbf{X_s V_s}}\right)\mathbf{Y_s}$$

■

Teorema 3.2.5. *Bajo el modelo (3.1.1), con* $\mathbf{Z}_j = \mathbf{1}_{N_j}$*, si* $n_j \mathbf{P}_{\mathbf{Z}_{j\mathbf{s}}} = \mathbf{Z}_{j\mathbf{s}} \mathbf{Z}_{j\mathbf{s}}^t$*,* $n_j \in \mathbb{R}$*, entonces el BLUP del efecto mixto* $\overline{\mathbf{X}}_{j\mathbf{r}} \boldsymbol{\beta} + u_j$*, de la* j*-ésima unidad de nivel 2, se expresa en términos de los operadores proyector* $\mathbf{Q}_{\mathbf{X_s V_s}}$ *y* $\mathbf{P}_{\mathbf{Z_s}}$*, y de la transformación lineal* $\mathbf{T}_{j\mathbf{s}}$ *por:*

$$\left[\frac{\mathbf{1}_{N_j}^{*j\mathbf{r}^t}}{r_j} \mathbf{T}_{j\mathbf{s}} \; + \; \frac{\mathbf{1}_{n}^{*j\mathbf{s}^t}}{n_j} \left(\mathbf{P}_{\mathbf{Z_s}} \mathbf{B_s} \mathbf{Q}_{\mathbf{X_s V_s}} \right) \right] \mathbf{Y_s},$$

donde $\mathbf{B_s} = \oplus_{j=1}^{J} \left(b_j \mathbf{I}_{n_j} \right)$ *y* $b_j = n_j \sigma_{u_0}^2 / \left(n_j \sigma_{u_0}^2 + \sigma_e^2 \right)$*.*

Demostración. Del hecho de que $\overline{\mathbf{X}}_{j\mathbf{r}} \boldsymbol{\beta} + u_j$ sea un efecto mixto se tiene que su *BLUP* está dado por $\overline{\mathbf{X}}_{j\mathbf{r}} \hat{\boldsymbol{\beta}}_{\mathbf{s}} + \hat{u}_{j\mathbf{s}}$, y por el teorema 3.2.4, se tiene

$$\begin{aligned}
BLUP \left(\overline{\mathbf{X}}_{j\mathbf{r}} \boldsymbol{\beta} + u_j \right) &= \overline{\mathbf{X}}_{j\mathbf{r}} \hat{\boldsymbol{\beta}}_{\mathbf{s}} \; + \; \hat{u}_{j\mathbf{s}} \\
&= \frac{\mathbf{1}_{N_j}^{*j\mathbf{r}^t} \mathbf{X}_j}{r_j} \hat{\boldsymbol{\beta}}_{\mathbf{s}} \; + \; \hat{u}_{j\mathbf{s}} \\
&= \frac{\mathbf{1}_{N_j}^{*j\mathbf{r}^t}}{r_j} \mathbf{X}_j \left(\mathbf{X}_{\mathbf{s}}^t \mathbf{V}_{\mathbf{s}}^{-1} \mathbf{X}_{\mathbf{s}} \right)^{-1} \mathbf{X}_{\mathbf{s}}^t \mathbf{V}_{\mathbf{s}}^{-1} \mathbf{Y_s} \; + \; \hat{u}_{j\mathbf{s}} \\
&= \frac{\mathbf{1}_{N_j}^{*j\mathbf{r}^t}}{r_j} \mathbf{T}_{j\mathbf{s}} \mathbf{Y_s} \; + \; \frac{\mathbf{1}_{n}^{*j\mathbf{s}^t}}{n_j} \left(\mathbf{P}_{\mathbf{Z_s}} \mathbf{B_s} \mathbf{Q}_{\mathbf{X_s V_s}} \right) \mathbf{Y_s} \\
&= \left[\frac{\mathbf{1}_{N_j}^{*j\mathbf{r}^t}}{r_j} \mathbf{T}_{j\mathbf{s}} \; + \; \frac{\mathbf{1}_{n}^{*j\mathbf{s}^t}}{n_j} \left(\mathbf{P}_{\mathbf{Z_s}} \mathbf{B_s} \mathbf{Q}_{\mathbf{X_s V_s}} \right) \right] \mathbf{Y_s}
\end{aligned}$$

■

3.3 Caracterización del *BLUP* de la media poblacional finita $\overline{Y}_j$

En está sección se presenta la caracterización del $BLUP$ de la media poblacional finita $\overline{Y}_j$, correspondiente a la j-ésima unidad de nivel 2, en términos de los operadores $\mathbf{P}_{\mathbf{X_s V_s}}$ y $\mathbf{P}_{\mathbf{Z_s}}$, y de la transformación lineal $\mathbf{T}_{j\mathbf{s}}$.

En la sección 3.3.1 se presenta la caracterización de la media poblacional finita $\overline{Y}_j$ considerando el caso balanceado y en la sección 3.3.2 el caso desbalanceado.

La relación entre el vector $\mathbf{Y_s}$ y la media poblacional finita $\overline{Y}_{j\mathbf{s}}$, de las unidades de la j-ésima unidad de nivel 2 en la muestra, está dada por:

$$\overline{Y}_{j\mathbf{s}} = \frac{\mathbf{1}_n^{*j\mathbf{s}^t}}{n_j}\mathbf{Y_s}.$$

3.3.1 Caso balanceado

En está sección se presenta la caracterización de la media poblacional finita $\overline{Y}_j$ para el caso balanceado.

Teorema 3.3.1. *Bajo el modelo (3.1.1), con $\mathbf{Z}_j = \mathbf{1}_{N_j}$, si $d\mathbf{P}_{\mathbf{Z_s}} = \mathbf{Z_s}\mathbf{Z_s}^t$, $d \in \mathbb{R}$, entonces el BLUP de la media poblacional finita $\overline{Y}_j$, de la j-ésima unidad de nivel 2, se expresa en términos de los operadores proyector $\mathbf{Q}_{\mathbf{X_s V_s}}$ y $\mathbf{P}_{\mathbf{Z_s}}$, y de la transformación lineal $\mathbf{T}_{j\mathbf{s}}$ por:*

$$\left(\frac{\mathbf{1}_n^{*j\mathbf{s}^t}}{N_j} + \frac{r_j}{N_j}\left[\frac{\mathbf{1}_{N_j}^{*j\mathbf{r}^t}}{r_j}\mathbf{T}_{j\mathbf{s}} + \frac{\mathbf{1}_n^{*j\mathbf{s}^t}}{d}\left(c\mathbf{P}_{\mathbf{Z_s}}\mathbf{Q}_{\mathbf{X_s V_s}}\right)\right]\right)\mathbf{Y_s},$$

donde $c = d\sigma_{u_0}^2 / \left(d\sigma_{u_0}^2 + \sigma_e^2\right)$.

Demostración. Por el teorema 1.3.3, el *BLUP* de $\overline{Y}_j$, bajo el modelo (3.1.1) con $\mathbf{Z}_j = \mathbf{1}_{N_j}$, está dado por

$$f_j \overline{Y}_{j\mathbf{s}} + (1 - f_j) \left[\overline{\mathbf{X}}_{j\mathbf{r}} \overset{\wedge}{\boldsymbol{\beta}}_{\mathbf{s}} + \hat{u}_{j\mathbf{s}} \right]. \tag{3.3.1}$$

Además, por el teorema 3.2.3,

$$\overline{\mathbf{X}}_{j\mathbf{r}} \overset{\wedge}{\boldsymbol{\beta}}_{\mathbf{s}} + \hat{u}_{j\mathbf{s}} = \left[\frac{\mathbf{1}_{N_j}^{*j\mathbf{r}t}}{r_j} \mathbf{T}_{j\mathbf{s}} + \frac{\mathbf{1}_{n}^{*j\mathbf{s}t}}{d} \left(c \mathbf{P}_{\mathbf{Z}_{\mathbf{s}}} \mathbf{Q}_{\mathbf{X}_{\mathbf{s}} \mathbf{V}_{\mathbf{s}}} \right) \right] \mathbf{Y}_{\mathbf{s}},$$

donde $c = d\sigma_{u_0}^2 / \left(d\sigma_{u_0}^2 + \sigma_e^2 \right)$. Así

$$\begin{aligned}
BLUP\left(\overline{Y}_j\right) &= \frac{n_j}{N_j} \overline{Y}_{j\mathbf{s}} + \frac{r_j}{N_j} \left[\overline{\mathbf{X}}_{j\mathbf{r}} \overset{\wedge}{\boldsymbol{\beta}}_{\mathbf{s}} + \hat{u}_{j\mathbf{s}} \right] \\
&= \frac{n_j}{N_j} \frac{\mathbf{1}_{n}^{*j\mathbf{s}t}}{n_j} \mathbf{Y}_{\mathbf{s}} + \frac{r_j}{N_j} \left[\overline{\mathbf{X}}_{j\mathbf{r}} \overset{\wedge}{\boldsymbol{\beta}}_{\mathbf{s}} + \hat{u}_{j\mathbf{s}} \right] \\
&= \frac{\mathbf{1}_{n}^{*j\mathbf{s}t}}{N_j} \mathbf{Y}_{\mathbf{s}} + \frac{r_j}{N_j} \left[\frac{\mathbf{1}_{N_j}^{*j\mathbf{r}t}}{r_j} \mathbf{T}_{j\mathbf{s}} + \frac{\mathbf{1}_{n}^{*j\mathbf{s}t}}{d} \left(c \mathbf{P}_{\mathbf{Z}_{\mathbf{s}}} \mathbf{Q}_{\mathbf{X}_{\mathbf{s}} \mathbf{V}_{\mathbf{s}}} \right) \right] \mathbf{Y}_{\mathbf{s}} \\
&= \left(\frac{\mathbf{1}_{n}^{*j\mathbf{s}t}}{N_j} + \frac{r_j}{N_j} \left[\frac{\mathbf{1}_{N_j}^{*j\mathbf{r}t}}{r_j} \mathbf{T}_{j\mathbf{s}} + \frac{\mathbf{1}_{n}^{*j\mathbf{s}t}}{d} \left(c \mathbf{P}_{\mathbf{Z}_{\mathbf{s}}} \mathbf{Q}_{\mathbf{X}_{\mathbf{s}} \mathbf{V}_{\mathbf{s}}} \right) \right] \right) \mathbf{Y}_{\mathbf{s}},
\end{aligned}$$

donde $c = d\sigma_{u_0}^2 / \left(d\sigma_{u_0}^2 + \sigma_e^2 \right)$ ■

3.3.2 Caso desbalanceado

En está sección se presenta la caracterización de la media poblacional finita $\overline{Y}_j$ para el caso desbalanceado.

Teorema 3.3.2. *Bajo el modelo (3.1.1), con* $\mathbf{Z}_j = \mathbf{1}_{N_j}$, *si* $n_j \mathbf{P}_{\mathbf{Z}_{j\mathbf{s}}} = \mathbf{Z}_{j\mathbf{s}} \mathbf{Z}_{j\mathbf{s}}^t$, $n_j \in \mathbb{R}$, *entonces el BLUP de la media poblacional finita* $\overline{Y}_j$, *se expresa en* *términos de los operadores proyector* $\mathbf{Q}_{\mathbf{X_s V_s}}$ *y* $\mathbf{P}_{\mathbf{Z_s}}$, *y de la transformación lineal* $\mathbf{T}_{j\mathbf{s}}$ *por:*

$$\left(\frac{\mathbf{1}_n^{*j\mathbf{s}t}}{N_j} + \frac{r_j}{N_j} \left[\frac{\mathbf{1}_{N_j}^{*j\mathbf{r}t}}{r_j} \mathbf{T}_{j\mathbf{s}} + \frac{\mathbf{1}_n^{*j\mathbf{s}t}}{n_j} \left(\mathbf{P}_{\mathbf{Z_s}} \mathbf{B_s} \mathbf{Q}_{\mathbf{X_s V_s}} \right) \right] \right) \mathbf{Y_s},$$

donde $\mathbf{B_s} = \oplus_{j=1}^{J} \left(b_j \mathbf{I}_{n_j} \right)$ *y* $b_j = n_j \sigma_{u0}^2 / \left(n_j \sigma_{u0}^2 + \sigma_e^2 \right)$.

Demostración. Por el teorema 1.3.3, el *BLUP* de $\overline{Y}_j$ está dado por

$$f_j \overline{Y}_{j\mathbf{s}} + (1 - f_j) \left[\overline{\mathbf{X}}_{j\mathbf{r}} \stackrel{\wedge}{\boldsymbol{\beta}}_{\mathbf{s}} + \stackrel{\wedge}{u}_{j\mathbf{s}} \right].$$

Además, por el teorema 3.2.5,

$$\overline{\mathbf{X}}_{j\mathbf{r}} \stackrel{\wedge}{\boldsymbol{\beta}}_{\mathbf{s}} + \stackrel{\wedge}{u}_{j\mathbf{s}} = \left[\frac{\mathbf{1}_{N_j}^{*j\mathbf{r}t}}{r_j} \mathbf{T}_{j\mathbf{s}} + \frac{\mathbf{1}_n^{*j\mathbf{s}t}}{n_j} \left(\mathbf{P}_{\mathbf{Z_s}} \mathbf{B_s} \mathbf{Q}_{\mathbf{X_s V_s}} \right) \right] \mathbf{Y_s},$$

donde $\mathbf{B_s} = \oplus_{j=1}^{J} \left(b_j \mathbf{I}_{n_j} \right)$ y $b_j = n_j \sigma_{u0}^2 / \left(n_j \sigma_{u0}^2 + \sigma_e^2 \right)$. Así

$$\begin{aligned}
BLUP\left(\overline{Y}_j\right) &= \frac{n_j}{N_j} \overline{Y}_{j\mathbf{s}} + \frac{r_j}{N_j} \left[\overline{\mathbf{X}}_{j\mathbf{r}} \stackrel{\wedge}{\boldsymbol{\beta}}_{\mathbf{s}} + \stackrel{\wedge}{u}_{j\mathbf{s}} \right] \\
&= \frac{n_j}{N_j} \frac{\mathbf{1}_n^{*j\mathbf{s}t}}{n_j} \mathbf{Y_s} + \frac{r_j}{N_j} \left[\overline{\mathbf{X}}_{j\mathbf{r}} \stackrel{\wedge}{\boldsymbol{\beta}}_{\mathbf{s}} + \stackrel{\wedge}{u}_{j\mathbf{s}} \right] \\
&= \frac{\mathbf{1}_n^{*j\mathbf{s}t}}{N_j} \mathbf{Y_s} + \frac{r_j}{N_j} \left[\frac{\mathbf{1}_{N_j}^{*j\mathbf{r}t}}{r_j} \mathbf{T}_{j\mathbf{s}} + \frac{\mathbf{1}_n^{*j\mathbf{s}t}}{n_j} \left(\mathbf{P}_{\mathbf{Z_s}} \mathbf{B_s} \mathbf{Q}_{\mathbf{X_s V_s}} \right) \right] \mathbf{Y_s} \\
&= \left(\frac{\mathbf{1}_n^{*j\mathbf{s}t}}{N_j} + \frac{r_j}{N_j} \left[\frac{\mathbf{1}_{N_j}^{*j\mathbf{r}t}}{r_j} \mathbf{T}_{j\mathbf{s}} + \frac{\mathbf{1}_n^{*j\mathbf{s}t}}{n_j} \left(\mathbf{P}_{\mathbf{Z_s}} \mathbf{B_s} \mathbf{Q}_{\mathbf{X_s V_s}} \right) \right] \right) \mathbf{Y_s},
\end{aligned}$$

donde $\mathbf{B_s} = \oplus_{j=1}^{J} \left(b_j \mathbf{I}_{n_j} \right)$ y $b_j = n_j \sigma_{u0}^2 / \left(n_j \sigma_{u0}^2 + \sigma_e^2 \right)$ ■

3.4 Modelos

El punto de partida para aplicar la teoría desarrollada en las secciones anteriores de este capítulo es que la matriz $\mathbf{Z}$ del modelo bajo estudio cumpla la condición $\mathbf{Z}\mathbf{Z}^t = d\mathbf{P}_{\mathbf{Z}}$ en el caso balanceado o la condición $n_j\mathbf{P}_{\mathbf{Z}_j} = \mathbf{Z}_j\mathbf{Z}_j^t$ en el caso desbalanceado. En esta sección se presentan tres modelos que cumplen la condición mencionada.

3.4.1 Caso balanceado

Casos particulares del modelo (3.1.1), que cumplen la condición $\mathbf{Z}\mathbf{Z}^t = d\mathbf{P}_{\mathbf{Z}}$ son:

(i) Modelo intercepto aleatorio. De la sección 1.1.1, el modelo para la i-ésima unidad de nivel 1 en la j-ésima unidad de nivel 2 está dado por:

$$Y_{ij} = \mu + u_j + e_{ij}, \quad i = 1, ..., d, \quad j = 1, ..., k, \tag{3.4.1}$$

donde μ es un parámetro fijo; u_j es el efecto aleatorio; u_j y e_{ij} son independientes, con $u_j \sim N\left(0, \sigma_{u0}^2\right)$ y $e_{ij} \sim N\left(0, \sigma_e^2\right)$. El modelo para la j-ésima unidad de nivel 2 está dado por:

$$\mathbf{Y}_j = \mathbf{1}_d\mu + \mathbf{1}_d u_j + \mathbf{e}_j, \quad j = 1, ..., k, \tag{3.4.2}$$

tomando $\boldsymbol{\beta} = \mu$, $\mathbf{X} = \mathbf{1}_k \otimes \mathbf{1}_d$, $\mathbf{Z} = \mathbf{I}_k \otimes \mathbf{1}_d$ y $\mathbf{u} = (u_1, ..., , u_k)^t$, el modelo (3.4.2) es de la forma $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$. En este caso se tiene:

$$\begin{aligned}
\mathbf{Z}\mathbf{Z}^t &= (\mathbf{I}_k \otimes \mathbf{1}_d)(\mathbf{I}_k \otimes \mathbf{1}_d)^t = (\mathbf{I}_k \otimes \mathbf{1}_d)\left(\mathbf{I}_k^t \otimes \mathbf{1}_d^t\right) \\
&= \left(\mathbf{I}_k\mathbf{I}_k^t \otimes \mathbf{1}_d\mathbf{1}_d^t\right) = \left(\mathbf{I}_k \otimes \mathbf{1}_d\mathbf{1}_d^t\right).
\end{aligned}$$

Así

$$
\begin{aligned}
\mathbf{P}_{\mathbf{Z}} &= \mathbf{Z}\left(\mathbf{Z}^t\mathbf{Z}\right)^{-1}\mathbf{Z}^t \\
&= \left(\mathbf{I}_k \otimes \mathbf{1}_d\right)\left[\left(\mathbf{I}_k \otimes \mathbf{1}_d\right)^t\left(\mathbf{I}_k \otimes \mathbf{1}_d\right)\right]^-\left(\mathbf{I}_k \otimes \mathbf{1}_d\right)^t \\
&= \left(\mathbf{I}_k \otimes \mathbf{1}_d\right)\left[\left(\mathbf{I}_k^t \otimes \mathbf{1}_d^t\right)\left(\mathbf{I}_k \otimes \mathbf{1}_d\right)\right]^-\left(\mathbf{I}_k \otimes \mathbf{1}_d\right)^t \\
&= \left(\mathbf{I}_k \otimes \mathbf{1}_d\right)\left[\left(\mathbf{I}_k^t\mathbf{I}_k \otimes \mathbf{1}_d^t\mathbf{1}_d\right)\right]^-\left(\mathbf{I}_k \otimes \mathbf{1}_d\right)^t \\
&= \left(\mathbf{I}_k \otimes \mathbf{1}_d\right)\left[\left(\mathbf{I}_k \otimes 1/d\right)\right]\left(\mathbf{I}_k \otimes \mathbf{1}_d\right)^t = \left(\mathbf{I}_k \otimes \mathbf{1}_d/d\right)\left(\mathbf{I}_k \otimes \mathbf{1}_d\right)^t \\
&= \mathbf{I}_k \otimes \frac{\mathbf{1}_d\mathbf{1}_d^t}{d} = \frac{1}{d}\left(\mathbf{I}_k \otimes \mathbf{1}_d\mathbf{1}_d^t\right) \\
&= \frac{1}{d}\mathbf{Z}\mathbf{Z}^t.
\end{aligned}
$$

(ii) Modelo intercepto aleatorio con una variable explicatoria. De la sección 1.1.2, el modelo para la i-ésima unidad de nivel 1 en la j-ésima unidad de nivel 2 está dado por:

$$
Y_{ij} = \beta_0 + \beta_1 X_{ij} + u_j + e_{ij}, \quad i = 1, ..., d, \quad j = 1, ..., k, \tag{3.4.3}
$$

donde β_0 y β_1 son parámetros fijos; u_j es el efecto aleatorio; u_j y e_{ij} son independientes, con $u_j \sim N\left(0, \sigma_{u_0}^2\right)$ y $e_{ij} \sim N\left(0, \sigma_e^2\right)$. Definiendo a $\mathbf{X}_j^*$ como la matriz de variables explicatorias para la j-ésima unidad de nivel 2, el modelo para la j-ésima unidad de nivel 2 está dado por:

$$
\mathbf{Y}_j = \mathbf{1}_d\beta_0 + \mathbf{X}_j^*\beta_1 + \mathbf{1}_d u_j + \mathbf{e}_j, \quad j = 1, ..., k, \tag{3.4.4}
$$

tomando $\boldsymbol{\beta} = \left(\beta_0, \beta_1\right)^t$, $\mathbf{X}_j = \left(\mathbf{1}_d : \mathbf{X}_j^*\right)$, y $\mathbf{Z}_j = \mathbf{1}_d$, el modelo (3.4.4) está dado por:

$$
\mathbf{Y}_j = \mathbf{X}_j\boldsymbol{\beta} + \mathbf{Z}_j u_j + \mathbf{e}_j. \tag{3.4.5}
$$

Además, definiendo a $\mathbf{X}^*$ como la matriz de las variables explicatorias de nivel 1, $\mathbf{X} = (\mathbf{1}_k \otimes \mathbf{1}_d : \mathbf{X}^*)$, $\mathbf{Z} = \mathbf{I}_k \otimes \mathbf{1}_d$ y $\mathbf{u} = (u_1, \ldots, u_k)^t$, el modelo (3.4.5) es de la forma $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$. En este caso también se cumple la condición $d\mathbf{P}_{\mathbf{Z}} = \mathbf{Z}\mathbf{Z}^t$.

(iii) Modelo de pendientes aleatorias. De la sección 1.1.3, el modelo para la i-ésima unidad de nivel 1 en la j-ésima unidad de nivel 2 está dado por:

$$Y_{ij} = \beta_{00} + \beta_{10}X_{ij} + u_{0j} + u_{1j}X_{ij} + e_{ij}, \quad i = 1, ..., d, \quad j = 1, ..., k, \tag{3.4.6}$$

donde β_{00} y β_{10} son parámetros fijos; u_{0j} y u_{1j} denotan a los efectos aleatorios; u_{0j}, u_{1j} y e_{ij} son independientes, con $u_{lj} \sim N\left(0, \sigma^2_{u_l}\right)$, $l = 0, 1$, y $e_{ij} \sim N\left(0, \sigma^2_e\right)$. Definiendo a $\mathbf{X}^*_j$ como la matriz de variables explicatorias para la j-ésima unidad de nivel 2, el modelo para la j-ésima unidad de nivel 2 tiene la forma:

$$\mathbf{Y}_j = \mathbf{1}_d\beta_{00} + \mathbf{X}^*_j\beta_{10} + \mathbf{1}_d u_{0j} + \mathbf{X}^*_j u_{1j} + \mathbf{e}_j, \quad j = 1, ..., k. \tag{3.4.7}$$

tomando $\boldsymbol{\beta} = (\beta_{00}, \beta_{10})^t$, $\mathbf{u}_j = (u_{0j}, u_{1j})^t$, $\mathbf{X}_j = \left(\mathbf{1}_d : \mathbf{X}^*_j\right)$, y $\mathbf{Z}_j = \left(\mathbf{1}_d : \mathbf{X}^*_j\right)$, el modelo (3.4.7) está dado por:

$$\mathbf{Y}_j = \mathbf{X}_j\boldsymbol{\beta} + \mathbf{Z}_j\mathbf{u}_j + \mathbf{e}_j. \quad j = 1, ..., k. \tag{3.4.8}$$

Además, definiendo a $\mathbf{X}$ como matriz de las variables explicatorias, $\mathbf{u} = (\mathbf{u}_1, \ldots, \mathbf{u}_k)^t$, y $\mathbf{Z} = \oplus\mathbf{Z}_j = \oplus\mathbf{X}_j$, el modelo (3.4.8) es de la forma $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$. Así:

$$\begin{aligned} \mathbf{Z}\mathbf{Z}^t &= (\oplus\mathbf{X}_j)(\oplus\mathbf{X}_j)^t \\ &= (\oplus\mathbf{X}_j)\left(\oplus\mathbf{X}^t_j\right) = \oplus\left(\mathbf{X}_j\mathbf{X}^t_j\right). \end{aligned}$$

Por lo que:

$$\begin{aligned}
\mathbf{P}_{\mathbf{Z}} &= \mathbf{Z}\left(\mathbf{Z}^t\mathbf{Z}\right)^{-1}\mathbf{Z}^t \\
&= \left(\oplus\mathbf{X}_j\right)\left[\left(\oplus\mathbf{X}_j\right)^t\left(\oplus\mathbf{X}_j\right)\right]^{-}\left(\oplus\mathbf{X}_j\right)^t \\
&= \left(\oplus\mathbf{X}_j\right)\left[\left(\oplus\mathbf{X}_j^t\right)\left(\oplus\mathbf{X}_j\right)\right]^{-}\left(\oplus\mathbf{X}_j^t\right) \\
&= \left(\oplus\mathbf{X}_j\right)\left(\oplus\mathbf{X}_j^t\mathbf{X}_j\right)^{-}\left(\oplus\mathbf{X}_j^t\right) \\
&= \oplus\left[\mathbf{X}_j\left(\mathbf{X}_j^t\mathbf{X}_j\right)^{-}\mathbf{X}_j^t\right].
\end{aligned}$$

Así, si $\mathbf{X}_j^t\mathbf{X}_j = d\mathbf{I}$, entonces $d\mathbf{P}_{\mathbf{Z}} = \mathbf{Z}\mathbf{Z}^t$.

3.4.2 Caso desbalanceado

Casos particulares del modelo lineal mixto, para los cuales se cumple la condición $n_j\mathbf{P}_{\mathbf{Z}_j} = \mathbf{Z}_j\mathbf{Z}_j^t$, son:

(iv) Modelo intercepto aleatorio. De la sección 1.1.1, el modelo para la i-ésima unidad de nivel 1 en la j-ésima unidad de nivel 2 está dado por:

$$Y_{ij} = \mu + u_j + e_{ij}, \quad i = 1, ..., n_j, \quad j = 1, ..., k, \tag{3.4.9}$$

donde μ es un parámetro fijo; u_j es el efecto aleatorio; u_j y e_{ij} son independientes, con $u_j \sim N\left(0, \sigma_{u0}^2\right)$ y $e_{ij} \sim N\left(0, \sigma_e^2\right)$. El modelo para la j-ésima unidad de nivel 2 está dado por:

$$\mathbf{Y}_j = \mathbf{1}_{n_j}\mu + \mathbf{1}_{n_j}u_j + \mathbf{e}_j, \quad j = 1, ..., k, \tag{3.4.10}$$

tomando $\boldsymbol{\beta} = \mu$, $\mathbf{X} = \mathbf{1}_k \otimes \mathbf{1}_{n_j}$, $\mathbf{Z} = \mathbf{I}_k \otimes \mathbf{1}_{n_j}$ y $\mathbf{u} = (u_1, ..., , u_k)^t$, el modelo

(3.4.10) es de la forma $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$. En este caso se tiene:

$$\begin{aligned}
\mathbf{P}_{\mathbf{Z}_j} &= \mathbf{Z}_j\left(\mathbf{Z}_j^t\mathbf{Z}_j\right)^{-1}\mathbf{Z}_j^t \\
&= \mathbf{1}_{n_j}\left(\mathbf{1}_{n_j}^t\mathbf{1}_{n_j}\right)^{-1}\mathbf{1}_{n_j}^t \\
&= \mathbf{1}_{n_j}\left(n_j\right)^{-1}\mathbf{1}_{n_j}^t \\
&= \frac{1}{n_j}\left(\mathbf{1}_{n_j}\mathbf{1}_{n_j}^t\right) \\
&= \frac{1}{n_j}\mathbf{Z}_j\mathbf{Z}_j^t,
\end{aligned}$$

de lo cual se cumple $n_j\mathbf{P}_{\mathbf{Z}_j} = \mathbf{Z}_j\mathbf{Z}_j^t$.

(v) Modelo intercepto aleatorio con una variable explicatoria. De la sección 1.1.2, el modelo para la i-ésima unidad de nivel 1 en la j-ésima unidad de nivel 2 está dado por:

$$Y_{ij} = \beta_0 + \beta_1 X_{ij} + u_j + e_{ij}, \quad i = 1, ..., n_j, \quad j = 1, ..., k, \tag{3.4.11}$$

donde β_0 y β_1; u_j es el efecto aleatorio; u_j y e_{ij} son independientes, con $u_j \sim N\left(0, \sigma_{u_0}^2\right)$ y $e_{ij} \sim N\left(0, \sigma_e^2\right)$. Definiendo a $\mathbf{X}_j^*$ como la matriz de variables explicatorias para la j-ésima unidad de nivel 2, el modelo para la j-ésima unidad de nivel 2 está dado por:

$$\mathbf{Y}_j = \mathbf{1}_{n_j}\beta_0 + \mathbf{X}_j^*\beta_1 + \mathbf{1}_{n_j}u_j + \mathbf{e}_j, \quad j = 1, ..., k, \tag{3.4.12}$$

tomando $\boldsymbol{\beta} = (\beta_0, \beta_1)^t$, $\mathbf{X}_j = \left(\mathbf{1}_{n_j} : \mathbf{X}_j^*\right)$, $\mathbf{Z}_j = \mathbf{1}_{n_j}$ el modelo (3.4.12) está dado por

$$\mathbf{Y}_j = \mathbf{X}_j\boldsymbol{\beta} + \mathbf{Z}_j u_j + \mathbf{e}_j. \tag{3.4.13}$$

Además, definiendo a $\mathbf{X}^*$ como la matriz de las variables explicatorias de nivel 1, $\mathbf{X} = \left(\mathbf{1}_k \otimes \mathbf{1}_{n_j} : \mathbf{X}^*\right)$, $\mathbf{Z} = \mathbf{I}_k \otimes \mathbf{1}_{n_j}$ y $\mathbf{u} = (u_1, \ldots, u_k)^t$, el modelo (3.4.13) es de la forma $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$. En este caso también $\mathbf{Z} = \mathbf{I}_k \otimes \mathbf{1}_{n_j}$, por lo que se cumple la condición $n_j \mathbf{P}_{\mathbf{Z}_j} = \mathbf{Z}_j \mathbf{Z}_j^t$.

(vi) Modelo de pendientes aleatorias. De la sección 1.1.3, el modelo para la i-ésima unidad de nivel 1 en la j-ésima unidad de nivel 2 está dado por:

$$Y_{ij} = \beta_{00} + \beta_{10} X_{ij} + u_{0j} + u_{1j} X_{ij} + e_{ij}, \quad i = 1, ..., n_j, \quad j = 1, ..., k, \quad (3.4.14)$$

donde β_{00} y β_{10} son parámetros fijos; u_{0j} y u_{1j} denotan a los efectos aleatorios; u_{0j}, u_{1j} y e_{ij} son independientes, con $u_{lj} \sim N\left(0, \sigma^2_{u_l}\right)$, $l = 0, 1$, y $e_{ij} \sim N\left(0, \sigma^2_e\right)$. Definiendo a $\mathbf{X}^*_j$ como la matriz de variables explicatorias para la j-ésima unidad de nivel 2, el modelo para la j-ésima unidad de nivel 2 está dado por:

$$\mathbf{Y}_j = \mathbf{1}_{n_j} \beta_{00} + \mathbf{X}^*_j \beta_{10} + \mathbf{1}_{n_j} u_{0j} + \mathbf{X}^*_j u_{1j} + \mathbf{e}_j, \quad j = 1, ..., k. \quad (3.4.15)$$

tomando $\boldsymbol{\beta} = (\beta_{00}, \beta_{10})^t$, $\mathbf{u}_j = (u_{0j}, u_{1j})^t$, $\mathbf{X}_j = \left(\mathbf{1}_{n_j} : \mathbf{X}^*_j\right)$, y $\mathbf{Z}_j = \left(\mathbf{1}_{n_j} : \mathbf{X}^*_j\right)$, el modelo (3.4.15) está dado por:

$$\mathbf{Y}_j = \mathbf{X}_j \boldsymbol{\beta} + \mathbf{Z}_j \mathbf{u}_j + \mathbf{e}_j. \quad j = 1, ..., k. \quad (3.4.16)$$

Además, definiendo a $\mathbf{X}$ como matriz de las variables explicatorias, $\mathbf{u} = (\mathbf{u}_1, \ldots, \mathbf{u}_k)^t$, y $\mathbf{Z} = \oplus \mathbf{Z}_j = \oplus \mathbf{X}_j$, el modelo (3.4.16) es de la forma $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$. Así:

$$\begin{aligned} \mathbf{Z}\mathbf{Z}^t &= (\oplus \mathbf{X}_j)(\oplus \mathbf{X}_j)^t \\ &= (\oplus \mathbf{X}_j)\left(\oplus \mathbf{X}^t_j\right) = \oplus \left(\mathbf{X}_j \mathbf{X}^t_j\right). \end{aligned}$$

Por lo que:

$$\begin{aligned}\mathbf{P}_{\mathbf{Z}_j} &= \mathbf{Z}_j\left(\mathbf{Z}_j^t\mathbf{Z}_j\right)^{-1}\mathbf{Z}_j^t\\ &= \mathbf{X}_j\left(\mathbf{X}_j^t\mathbf{X}_j\right)^{-1}\mathbf{X}_j^t.\end{aligned}$$

Así, si $\mathbf{X}_j^t\mathbf{X}_j = n_j\mathbf{I}$, entonces $n_j\mathbf{P}_{\mathbf{Z}_j} = \mathbf{Z}_j\mathbf{Z}_j^t$.

3.5 Modelo intercepto aleatorio

Como se ha mencionado a lo largo de este trabajo el modelo punto de partida de las investigaciones sobre la predicción de la media poblacional finita $\overline{Y}_j$, bajo el enfoque de modelos, es el modelo intercepto aleatorio. Por lo anterior en está sección se aplican los resultados obtenidos a este modelo. En la sección 3.5.1 se presenta la caracterización para el efecto mixto $\mathbf{X}\boldsymbol{\beta}+\mathbf{Z}\mathbf{u}$, en la sección 3.5.2 para el efecto mixto $\overline{\mathbf{X}}_{j\mathbf{r}}\boldsymbol{\beta}+u_j$, y en la sección 3.5.3 para la media poblacional finita $\overline{Y}_j$. Lo anterior se lleva a cabo considerando el caso balanceado.

3.5.1 *BLUP* del efecto mixto $\mathbf{X}\boldsymbol{\beta}+\mathbf{Z}\mathbf{u}$

En esta sección se presenta el desarrollo de la caracterización del *BLUP* del efecto mixto $\mathbf{X}\boldsymbol{\beta}+\mathbf{Z}\mathbf{u}$ bajo el modelo intercepto aleatorio (3.4.1).

Para obtener la caracterización del *BLUP* del efecto mixto $\mathbf{X}\boldsymbol{\beta}+\mathbf{Z}\mathbf{u}$ son necesarios los siguientes resultados:

Lema 3.5.1. *Para el modelo intercepto aleatorio (3.4.1), el operador proyector oblicuo está dado por:*

$$\mathbf{P_{XV}} = \frac{\mathbf{1}_{kd}\mathbf{1}_{kd}^t}{kd}. \tag{3.5.1}$$

Demostración. Bajo el modelo

$$Y_{ij} = \mu + u_j + e_{ij}, \quad i = 1, ..., d, \quad j = 1, ..., k, \tag{3.5.2}$$

con $u_j \sim N\left(0, \sigma_{u_0}^2\right)$ y $e_{ij} \sim N\left(0, \sigma_e^2\right)$, se tiene $\mathbf{Z} = \mathbf{I}_k \otimes \mathbf{1}_d$, de lo cual se cumple

$$\mathbf{P_Z} = \frac{\mathbf{Z}\mathbf{Z}^t}{d} = \frac{\left(\mathbf{I}_k \otimes \mathbf{1}_d\mathbf{1}_d^t\right)}{d}. \tag{3.5.3}$$

Del teorema 3.1.2, la matriz $\mathbf{V}^{-1}$ está dada por:

$$\mathbf{V}^{-1} = \frac{\left(\mathbf{I}_k \otimes \mathbf{1}_d\mathbf{1}_d^t\right)}{d\left(d\sigma_{u_0}^2 + \sigma_e^2\right)} + \frac{\left(d\mathbf{I}_{kd} - \left(\mathbf{I}_k \otimes \mathbf{1}_d\mathbf{1}_d^t\right)\right)}{d\sigma_e^2}. \tag{3.5.4}$$

En este caso $\mathbf{X} = \mathbf{1}_k \otimes \mathbf{1}_d = \mathbf{1}_{kd}$, así de (3.5.4)

$$\begin{aligned}
\mathbf{X}^t\mathbf{V}^{-1} &= \mathbf{1}_{kd}^t\left[\frac{\left(\mathbf{I}_k \otimes \mathbf{1}_d\mathbf{1}_d^t\right)}{d\left(d\sigma_{u_0}^2 + \sigma_e^2\right)} + \frac{\left(d\mathbf{I}_{kd} - \left(\mathbf{I}_k \otimes \mathbf{1}_d\mathbf{1}_d^t\right)\right)}{d\sigma_e^2}\right] \\
&= \mathbf{1}_{kd}^t\frac{\left(\mathbf{I}_k \otimes \mathbf{1}_d\mathbf{1}_d^t\right)}{d\left(d\sigma_{u_0}^2 + \sigma_e^2\right)} + \mathbf{1}_{kd}^t\frac{\left(d\mathbf{I}_{kd} - \left(\mathbf{I}_k \otimes \mathbf{1}_d\mathbf{1}_d^t\right)\right)}{d\sigma_e^2},
\end{aligned}$$

desarrollando cada uno de los términos involucrados en $\mathbf{X}^t\mathbf{V}^{-1}$

$$\mathbf{1}_{kd}^t\frac{\left(\mathbf{I}_k \otimes \mathbf{1}_d\mathbf{1}_d^t\right)}{d\left(d\sigma_{u_0}^2 + \sigma_e^2\right)} = \frac{d\mathbf{1}_{kd}^t}{d\left(d\sigma_{u_0}^2 + \sigma_e^2\right)} = \frac{\mathbf{1}_{kd}^t}{\left(d\sigma_{u_0}^2 + \sigma_e^2\right)} \tag{3.5.5}$$

y

$$\mathbf{1}_{kd}^t\frac{\left(d\mathbf{I}_{kd} - \left(\mathbf{I}_k \otimes \mathbf{1}_d\mathbf{1}_d^t\right)\right)}{d\sigma_e^2} = \mathbf{1}_{kd}^t\frac{\left(d\mathbf{I}_{kd}\right)}{d\sigma_e^2} - \mathbf{1}_{kd}^t\frac{\left(\left(\mathbf{I}_k \otimes \mathbf{1}_d\mathbf{1}_d^t\right)\right)}{d\sigma_e^2} = \mathbf{0}. \tag{3.5.6}$$

Así de (3.5.5) y (3.5.6), se tiene

$$\mathbf{X}^t\mathbf{V}^{-1} = \frac{\mathbf{1}_{kd}^t}{\left(d\sigma_{u_0}^2 + \sigma_e^2\right)} \tag{3.5.7}$$

Considerando (3.5.7), el producto $\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}$ está dado por:

$$\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X} = \frac{\mathbf{1}_{kd}^t\mathbf{1}_{kd}}{\left(d\sigma_{u_0}^2 + \sigma_e^2\right)} = \frac{kd}{\left(d\sigma_{u_0}^2 + \sigma_e^2\right)}$$

y su inversa por:

$$\left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)^{-1} = \frac{d\sigma_{u_0}^2 + \sigma_e^2}{kd}. \tag{3.5.8}$$

De (3.5.7) y (3.5.8), se tiene

$$\left(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^t\mathbf{V}^{-1} = \frac{d\sigma_{u_0}^2 + \sigma_e^2}{kd}\frac{\mathbf{1}_{kd}^t}{\left(d\sigma_{u_0}^2 + \sigma_e^2\right)} = \frac{\mathbf{1}_{kd}^t}{kd}. \tag{3.5.9}$$

Así el operador proyector oblicuo $\mathbf{P_{XV}}$, bajo el modelo intercepto aleatorio (3.4.1), está dado por:

$$\mathbf{P_{XV}} = \frac{\mathbf{1}_{kd}\mathbf{1}_{kd}^t}{kd}$$

■

Nota 3.5.1. *Bajo el modelo intercepto aleatorio (3.4.1),* $\mathbf{P_Z P_{XV}} = \mathbf{P_{XV}}$.

Demostración. De (3.5.1) y (3.5.3), se tiene

$$\mathbf{P_Z P_{XV}} = \frac{\left(\mathbf{I}_k \otimes \mathbf{1}_d\mathbf{1}_d^t\right)}{d}\frac{\mathbf{1}_{kd}\mathbf{1}_{kd}^t}{kd} = \frac{d\mathbf{1}_{kd}\mathbf{1}_{kd}^t}{dkd} = \mathbf{P_{XV}}$$

■

Teorema 3.5.1. *Bajo el modelo intercepto aleatorio (3.4.1), el BLUP del efecto mixto* $\mathbf{X}\boldsymbol{\beta}+\mathbf{Zu}$ *está dado por*

$$(1-c)\,\frac{\mathbf{1}_{kd}\mathbf{1}_{kd}^{t}}{kd}\mathbf{Y}+c\frac{(\mathbf{I}_k\otimes\mathbf{1}_d\mathbf{1}_d^{t})}{d}\mathbf{Y},$$

donde $c=d\sigma_{u_0}^2/\left(d\sigma_{u_0}^2+\sigma_e^2\right).$

Demostración. Por el teorema 3.1.3, la caracterización del *BLUP* del efecto mixto $\mathbf{X}\boldsymbol{\beta}+\mathbf{Zu}$ está dada por

$$\mathbf{P_{XV}Y}\ +\ c\mathbf{P_Z}\left(\mathbf{I}-\mathbf{P_{XV}}\right)\mathbf{Y},$$

donde $c=d\sigma_{u_0}^2/\left(d\sigma_{u_0}^2+\sigma_e^2\right)$. Así por el lema 3.5.1 y la nota 3.5.1, se tiene

$$\begin{aligned}BLUP\left(\mathbf{X}\boldsymbol{\beta}+\mathbf{Zu}\right) &= \mathbf{P_{XV}Y}+c\mathbf{P_ZY}-c\mathbf{P_ZP_{XV}Y}\\ &= \mathbf{P_{XV}Y}+c\mathbf{P_ZY}-c\mathbf{P_{XV}Y}\\ &= (1-c)\,\mathbf{P_{XV}Y}+c\mathbf{P_ZY}\\ &= (1-c)\,\frac{\mathbf{1}_{kd}\mathbf{1}_{kd}^{t}}{kd}\mathbf{Y}+c\frac{(\mathbf{I}_k\otimes\mathbf{1}_d\mathbf{1}_d^{t})}{d}\mathbf{Y}\end{aligned}$$

donde $c=d\sigma_{u_0}^2/\left(d\sigma_{u_0}^2+\sigma_e^2\right)$ ■

3.5.2 *BLUP* del efecto mixto $\overline{\mathbf{X}}_{j\mathbf{r}}\boldsymbol{\beta}+u_j$

En esta sección se presenta el desarrollo de la caracterización del *BLUP* del efecto mixto $\overline{\mathbf{X}}_{j\mathbf{r}}\boldsymbol{\beta}+u_j$ bajo el modelo intercepto aleatorio (3.4.1).

Teorema 3.5.2. *Bajo el modelo intercepto aleatorio (3.4.1)*

$$\mathbf{T}_{j\mathbf{s}} = \frac{\mathbf{1}_{N_j}\mathbf{1}_{kd}^t}{kd}. \tag{3.5.10}$$

Demostración. Por definición $\mathbf{T}_{j\mathbf{s}} = \mathbf{X}_j\left(\mathbf{X}_\mathbf{s}^t\mathbf{V}_\mathbf{s}^{-1}\mathbf{X}_\mathbf{s}\right)^{-}\mathbf{X}_\mathbf{s}^t\mathbf{V}_\mathbf{s}^{-1}$. De (3.5.7),

$$\mathbf{X}_\mathbf{s}^t\mathbf{V}_\mathbf{s}^{-1} = \frac{\mathbf{1}_{kd}^t}{\left(d\sigma_{u_0}^2 + \sigma_e^2\right)}$$

y por (3.5.8)

$$\left(\mathbf{X}_\mathbf{s}^t\mathbf{V}_\mathbf{s}^{-1}\mathbf{X}_\mathbf{s}\right)^{-1} = \frac{d\sigma_{u_0}^2 + \sigma_e^2}{kd}.$$

Así

$$\mathbf{T}_{j\mathbf{s}} = \mathbf{X}_j\left(\mathbf{X}_\mathbf{s}^t\mathbf{V}_\mathbf{s}^{-1}\mathbf{X}_\mathbf{s}\right)^{-}\mathbf{X}_\mathbf{s}^t\mathbf{V}_\mathbf{s}^{-1} \; = \mathbf{1}_{N_j}\frac{d\sigma_{u_0}^2 + \sigma_e^2}{kd}\frac{\mathbf{1}_{kd}^t}{\left(d\sigma_{u_0}^2 + \sigma_e^2\right)} \; = \frac{\mathbf{1}_{N_j}\mathbf{1}_{kd}^t}{kd}$$

■

Teorema 3.5.3. *Bajo el modelo intercepto aleatorio (3.4.1),*

$$c\mathbf{P}_{\mathbf{Z}_s}\mathbf{Q}_{\mathbf{X}_s\mathbf{V}_s} = \frac{c}{kd}\left(k\left(\oplus\mathbf{1}_d\mathbf{1}_d^t\right) - \mathbf{1}_{kd}\mathbf{1}_{kd}^t\right). \tag{3.5.11}$$

Demostración. De (3.5.1) y (3.5.3), se tiene

$$\begin{aligned}
c\mathbf{P}_{\mathbf{Z}_s}\mathbf{Q}_{\mathbf{X}_s\mathbf{V}_s} &= c\,\mathbf{P}_{\mathbf{Z}_s}(\mathbf{I} - \mathbf{P}_{\mathbf{X}_s\mathbf{V}_s}) \; = c\left(\mathbf{P}_{\mathbf{Z}_s} - \mathbf{P}_{\mathbf{X}_s\mathbf{V}_s}\right) \\
&= c\left(\frac{(\mathbf{I}_k \otimes \mathbf{1}_d\mathbf{1}_d^t)}{d} - \frac{\mathbf{1}_{kd}\mathbf{1}_{kd}^t}{kd}\right) \\
&= \frac{c}{kd}\left(k\left(\oplus\mathbf{1}_d\mathbf{1}_d^t\right) - \mathbf{1}_{kd}\mathbf{1}_{kd}^t\right)
\end{aligned}$$

■

Teorema 3.5.4. *Bajo el modelo intercepto aleatorio (3.4.1), el BLUP del efecto mixto* $\overline{\mathbf{X}}_{j\mathbf{r}}\boldsymbol{\beta} + u_j$ *está dado por:*

$$\overline{Y}_{\mathbf{s}} \ + \ \frac{c\,(k-1)}{k}\left[\overline{Y}_{j\mathbf{s}} - \overline{Y}_{(-j)\mathbf{s}}\right] \tag{3.5.12}$$

donde $\overline{Y}_{\mathbf{s}}$, $\overline{Y}_{j\mathbf{s}}$ *y* $\overline{Y}_{(-j)\mathbf{s}}$ *denotan la media muestral, la media muestral correspondiente a la* j*-ésima unidad de nivel 2, y la media muestral de las unidades de nivel 2 restantes, respectivamente.*

Demostración. Por el teorema 3.2.3, el *BLUP* de $\overline{\mathbf{X}}_{j\mathbf{r}}\boldsymbol{\beta} + u_j$ está dado por

$$\left[\frac{\mathbf{1}_{N_j}^{*j\mathbf{r}^t}}{r_j}\mathbf{T}_{j\mathbf{s}} + \frac{\mathbf{1}_n^{*j\mathbf{s}^t}}{d}\left(c\mathbf{P}_{\mathbf{Z_s}}\mathbf{Q}_{\mathbf{X_sV_s}}\right)\right]\mathbf{Y}_{\mathbf{s}},$$

donde $c = d\sigma_{u_0}^2/\left(d\sigma_{u_0}^2 + \sigma_e^2\right)$. Así por (3.5.10) y (3.5.11)

$$\begin{aligned}
BLUP\left(\overline{\mathbf{X}}_{j\mathbf{r}}\boldsymbol{\beta} + u_j\right) &= \left[\frac{\mathbf{1}_{N_j}^{*j\mathbf{r}^t}}{r_j}\mathbf{T}_{j\mathbf{s}} + \frac{\mathbf{1}_n^{*j\mathbf{s}^t}}{d}\left(c\mathbf{P}_{\mathbf{Z_s}}\mathbf{Q}_{\mathbf{X_sV_s}}\right)\right]\mathbf{Y}_s \\
&= \left[\frac{\mathbf{1}_{N_j}^{*j\mathbf{r}^t}}{r_j}\frac{\mathbf{1}_{N_j}\mathbf{1}_{kd}^t}{kd} \ + \ \frac{\mathbf{1}_n^{*j\mathbf{s}^t}}{d}\left[\frac{c}{kd}\left(k\left(\oplus\mathbf{1}_d\mathbf{1}_d^t\right) - \mathbf{1}_{kd}\mathbf{1}_{kd}^t\right)\right]\right]\mathbf{Y}_s \\
&= \left[\frac{r_j}{r_j}\frac{\mathbf{1}_{kd}^t}{kd} \ + \ \frac{c}{kdd}\left[\mathbf{1}_n^{*j\mathbf{s}^t}\left(k\left(\oplus\mathbf{1}_d\mathbf{1}_d^t\right) - \mathbf{1}_{kd}\mathbf{1}_{kd}^t\right)\right]\right]\mathbf{Y}_s \\
&= \frac{\mathbf{1}_{kd}^t}{kd}\mathbf{Y}_{\mathbf{s}} \ + \ \frac{c}{kdd}\left[d\left(k\mathbf{1}_n^{*j\mathbf{s}^t} - \mathbf{1}_{kd}^t\right)\right]\mathbf{Y}_s \\
&= \overline{Y}_{\mathbf{s}} \ + \ \frac{c}{kd}\left[(k-1)\,d\overline{Y}_{j\mathbf{s}} - (kd-d)\,\overline{Y}_{(-j)\mathbf{s}}\right] \\
&= \overline{Y}_{\mathbf{s}} \ + \ \frac{c\,(k-1)}{k}\left[\overline{Y}_{j\mathbf{s}} - \overline{Y}_{(-j)\mathbf{s}}\right]
\end{aligned}$$

■

3.5.3 *BLUP* de la media poblacional finita $\overline{Y}_j$

En esta sección se presenta la caracterización del *BLUP* de la media poblacional finita $\overline{Y}_j$ bajo el modelo intercepto aleatorio (3.4.1).

Teorema 3.5.5. *Bajo el modelo intercepto aleatorio (3.4.1), el BLUP de la media poblacional finita $\overline{Y}_j$ está dado por:*

$$\frac{n_j}{N_j}\overline{Y}_{j\mathbf{s}} + \frac{r_j}{N_j}\left[\overline{Y}_{\mathbf{s}} \ + \ \frac{c\,(k-1)}{k}\left[\overline{Y}_{j\mathbf{s}} - \overline{Y}_{(-j)\mathbf{s}}\right]\right] \tag{3.5.13}$$

donde $\overline{Y}_{\mathbf{s}}$, $\overline{Y}_{j\mathbf{s}}$ y $\overline{Y}_{(-j)\mathbf{s}}$ denotan la media muestral, la media muestral correspondiente a la j-ésima unidad de nivel 2, y la media muestral de las unidades de nivel 2 restantes.

Demostración. Por el teorema 3.3.1,

$$BLUP\left(\overline{Y}_j\right) = \left(\frac{\mathbf{1}_n^{*j\mathbf{s}^t}}{N_j} + \frac{r_j}{N_j}\left[\frac{\mathbf{1}_{N_j}^{*j\mathbf{r}^t}}{r_j}\mathbf{T}_{j\mathbf{s}} + \frac{\mathbf{1}_n^{*j\mathbf{s}^t}}{d}\left(c\mathbf{P}_{\mathbf{Z_s}}\mathbf{Q}_{\mathbf{X_s V_s}}\right)\right]\right)\mathbf{Y_s},$$

donde $c = d\sigma_{u_0}^2 / \left(d\sigma_{u_0}^2 + \sigma_e^2\right)$. Por la demostración del teorema 3.5.4,

$$\begin{aligned}
BLUP\left(\overline{Y}_j\right) &= \left(\frac{\mathbf{1}_n^{*j\mathbf{s}^t}}{N_j} + \frac{r_j}{N_j}\left[\frac{\mathbf{1}_{N_j}^{*j\mathbf{r}^t}}{r_j}\mathbf{T}_{j\mathbf{s}} + \frac{\mathbf{1}_n^{*j\mathbf{s}^t}}{d}\left(c\mathbf{P}_{\mathbf{Z_s}}\mathbf{Q}_{\mathbf{X_s V_s}}\right)\right]\right)\mathbf{Y_s} \\
&= \frac{\mathbf{1}_n^{*j\mathbf{s}^t}}{N_j}\mathbf{Y_s} + \frac{r_j}{N_j}\left[\frac{\mathbf{1}_{N_j}^{*j\mathbf{r}^t}}{r_j}\mathbf{T}_{j\mathbf{s}} + \frac{\mathbf{1}_n^{*j\mathbf{s}^t}}{d}\left(c\mathbf{P}_{\mathbf{Z_s}}\mathbf{Q}_{\mathbf{X_s V_s}}\right)\right]\mathbf{Y_s} \\
&= \frac{\mathbf{1}_n^{*j\mathbf{s}^t}}{N_j}\mathbf{Y_s} + \frac{r_j}{N_j}\left[\overline{Y}_{\mathbf{s}} \ + \ \frac{c\,(k-1)}{k}\left[\overline{Y}_{j\mathbf{s}} - \overline{Y}_{(-j)\mathbf{s}}\right]\right] \\
&= \frac{n_j}{N_j}\overline{Y}_{j\mathbf{s}} + \frac{r_j}{N_j}\left[\overline{Y}_{\mathbf{s}} \ + \ \frac{c\,(k-1)}{k}\left[\overline{Y}_{j\mathbf{s}} - \overline{Y}_{(-j)\mathbf{s}}\right]\right]
\end{aligned}$$

■

3.6 Factibilidad y limitaciones

Como se mencionó en la introducción, los modelos lineales jerárquicos han sido usados para tratar el problema de estimación en áreas pequeñas. Algunos investigadores hacen uso del modelo intercepto aleatorio, otros hacen uso del modelo de pendientes aleatorias.

Como se vio en las secciones anteriores de este capítulo, para llevar a cabo la caracterización del $BLUP$ de la media poblacional finita $\overline{Y}_j$, en términos del operador proyector, se parte del hecho de que el modelo bajo estudio cumple tres condiciones:

1. Se tienen dos componentes de la varianza; el primero correspondiente al nivel 1, σ_e^2 y el segundo correspondiente al nivel 2, $\sigma_{u_0}^2$,

2. La matriz $\mathbf{Z}_j$ cumple la condición $n_j\mathbf{P}_{\mathbf{Z}_j} = \mathbf{Z}_j\mathbf{Z}_j^t$,

3. La matriz $\mathbf{Z}_j$ está dada por $\mathbf{1}_{N_j}$.

Si las primeras dos condiciones se cumplen entonces la matriz de varianzas y covarianzas se puede expresar por $\mathbf{V} = \oplus_{j=1}^{J}\left[\left(n_j\sigma_{u_0}^2 + \sigma_e^2\right)\mathbf{P}_{\mathbf{Z}_j} + \sigma_e^2\mathbf{Q}_{\mathbf{Z}_j}\right]$. Mientras que la tercera condición permite la caracterización del término $\overline{\mathbf{X}}_{j\mathbf{r}}\boldsymbol{\beta} + u_j$.

Factibilidad. Modelo intercepto aleatorio

La factibilidad de que se cumplan las tres condiciones es alta, ya que éstas se dan siempre que el único parámetro aleatorio en el modelo bajo estudio sea β_0, es decir, de que no exista variabilidad entre los coeficientes correspondientes a cada una de las variables explicatorias, por lo cual sólamente hay dos componentes de la varianza. Cabe destacar que en muchos casos de estudio el único parámetro aleatorio es el intercepto. Al ser β_0 el único parámetro aleatorio la matriz $\mathbf{Z}_j$ está dada por $\mathbf{1}_{N_j}$ y por ende se cumple la condición $n_j\mathbf{P}_{\mathbf{Z}_j} = \mathbf{Z}_j\mathbf{Z}_j^t$.

En el desarrollo del trabajo se supone la existencia de únicamente variables explicatorias a nivel 1, por lo que en el modelo intercepto aleatorio considerado no intervienen variables explicatorias a nivel 2, así β_{0j} está dado por $\beta_{0j} = \beta_{00} + u_{0j}$. En el caso de existencia de variables explicatorias a nivel 2 y suponiendo un modelo intercepto aleatorio, β_{0j} estaría dado por $\beta_{0j} = \beta_{00} + \beta_{01}w_j + u_{0j}$, pero también en este modelo sólo hay dos componentes de la varianza y la matriz $\mathbf{Z}_j$ está dada por $\mathbf{1}_{N_j}$.

Así en el caso del modelo intercepto aleatorio, ya sea sin o con variables explicatorias y no dependiendo de que se tengan variables explicatorias a nivel 1 y/o a nivel 2, se cumplen las tres condiciones. Afortunadamente el modelo intercepto aleatorio es el modelo usado o el modelo de punto de partida de varios investigadores como se puede ver en las investigaciones realizadas por los autores mencionados [1], [9], [12], [13], [22], [30], [33], [49], entre otros.

Limitantes. Modelo pendientes aleatorias y derivados

En el caso del modelo de pendientes aleatorias, si el parámetro β_0 es aleatorio, el modelo está dado por $Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + e_{ij}$, en este modelo intervienen cuatro componentes de la varianza σ_e^2, σ_{u0}^2, $\sigma_{u_1}^2$ y $\sigma_{u_{01}}$, por lo que la teoría desarrollada en este trabajo no se puede aplicar.

Supongamos que el parámetro β_0 es fijo y el parámetro β_1 es aleatorio, el modelo para la variable respuesta Y_{ij} está dado por

$$Y_{ij} = \beta_0 + \beta_{1j} x_{ij} + e_{ij},$$

suponiendo la existencia únicamente de variables explicatorias a nivel 1, β_{1j} estaría dado por $\beta_{1j} = \beta_{10} + u_{1j}$, lo cual deviene en el modelo

$$Y_{ij} = \beta_0 + \beta_{10} x_{ij} + u_{1j} X_{ij} + e_{ij}$$

y la varianza de Y_{ij} está dada por $Var\left(Y_{ij}\right) = \sigma_{u_1}^2 X_{ij}^2 + \sigma_e^2$. En este caso el modelo para la j-ésima unidad de nivel 2 está dado por $\mathbf{Y}_j = \mathbf{X}_j \boldsymbol{\beta} + \mathbf{Z}_j \mathbf{u}_j + \mathbf{e}_j$, donde

$$\mathbf{X}_j = \begin{pmatrix} 1 & x_{1j} \\ \vdots & \vdots \\ 1 & x_{n_j j} \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_{10} \end{pmatrix} \text{ y } \mathbf{Z}_j = \begin{pmatrix} x_{1j} \\ \vdots \\ x_{n_j j} \end{pmatrix}$$

y aunque hay sólo dos componentes de la varianza $\sigma_{u_1}^2$ y σ_e^2, la matriz $\mathbf{Z}_j$ no está dada por $\mathbf{1}$.

Algunos trabajos, donde el objetivo es la predicción de la media poblacional finita $\overline{Y}_j$ y se hace uso del modelo de pendientes aleatorias son los realizados por [11], [13], [33], [45].

4

Conclusiones y recomendaciones

4.1 Conclusiones

Como se ha venido mencionando, el objetivo principal fue la caracterización del $BLUP$ de la media poblacional finita $\overline{Y}_j$ correspondiente a la j-ésima unidad de nivel 2, en el contexto del modelo intercepto aleatorio, ya que este modelo es ampliamente utilizado en la teoría de estimación en áreas pequeñas. Esta caracterización está dada en términos de los operadores proyector oblicuo $\mathbf{P_{XV}}$ y ortogonal $\mathbf{P_Z}$, y de una transformación lineal $\mathbf{T}_{j\mathbf{s}}$ definidos sobre los subespacios $S(\mathbf{X})$, $S(\mathbf{Z})$ y $S(\mathbf{X}_j)$, generados por las matrices de diseño $\mathbf{X}$, $\mathbf{Z}$ y $\mathbf{X}_j$, respectivamente. Se trabajó principalmente con el modelo de dos niveles con dos componentes de la varianza y en particular con el modelo intercepto aleatorio. Para el modelo intercepto aleatorio el $BLUP$ de la media poblacional finita $\overline{Y}_j$ de la j-ésima unidad de nivel 2, está dado por $f_j\overline{Y}_{j\mathbf{s}} + (1-f_j)\left[\overline{\mathbf{X}}_{j\mathbf{r}}\,\hat{\boldsymbol{\beta}}_{\mathbf{s}} + \hat{u}_{j\mathbf{s}}\right]$. En esta expresión interviene el término $\overline{\mathbf{X}}_{j\mathbf{r}}\,\hat{\boldsymbol{\beta}}_{\mathbf{s}} + \hat{u}_{j\mathbf{s}}$. Así para poder obtener la caracterización del mejor predictor lineal insesgado de la media poblacional finita para el modelo intercepto aletorio, fue necesario obtener la caracterización del término $\overline{\mathbf{X}}_{j\mathbf{r}}\,\hat{\boldsymbol{\beta}}_{\mathbf{s}} + \hat{u}_{j\mathbf{s}}$, lo cual fue posible al obtenerse la caracterización del $BLUP$ del efecto mixto $\mathbf{X}\boldsymbol{\beta} + \mathbf{Zu}$.

En primer lugar, para poder dar la caracterización del $BLUP$ del efecto mixto $\mathbf{X}\boldsymbol{\beta} + \mathbf{Zu}$ la matriz de diseño $\mathbf{Z}$ del modelo considerado debe de cumplir la condición $d\mathbf{P_Z} = \mathbf{Z}\mathbf{Z}^t$.

Respecto a la caracterización del *BLUP* del efecto mixto $\overline{\mathbf{X}}_{j\mathbf{r}}\boldsymbol{\beta} + u_j$ por medio de proyectores esta caracterización además de estar en función de los proyectores $\mathbf{P_{XV}}$ y $\mathbf{P_Z}$, depende de una transformación lineal definida sobre el espacio $S\left(\mathbf{X}_j\right)$.

Respecto a la caracterización del *BLUP* de la media poblacional finita $\overline{Y}_j$, al igual que para el efecto mixto $\overline{\mathbf{X}}_{j\mathbf{r}}\boldsymbol{\beta} + u_j$, esta caracterización además de estar en función de los proyectores $\mathbf{P_{XV}}$ y $\mathbf{P_Z}$, depende de una transformación lineal definida sobre el espacio $S\left(\mathbf{X}_j\right)$.

Al aplicarse las caraterizaciones obtenidas al modelo intercepto aleatorio, considerando el caso balanceado, se expresa el *BLUP* de la media poblacional finita $\overline{Y}_j$ como la suma de la media muestral, y un múltiplo de la diferencia entre las medias muestrales $\overline{Y}_{j\mathbf{s}}$ y $\overline{Y}_{(-j)\mathbf{s}}$.

Así las conclusiones de este trabajo son:

1. Sí es posible llevar a cabo la caracterización del *BLUP* de la media poblacional finita $\overline{Y}_j$, de la j-ésima unidad de nivel 2, en términos de los proyectores definidos sobre los subespacios generados por las matrices de diseño que intervienen en el modelo bajo estudio.

2. Además, de los operadores proyector interviene una transformacion lineal definida sobre el espacio $S(\mathbf{X}_j)$.

3. Para poder tener esta caracterización, la matriz de diseño $\mathbf{Z}_j$ asociada a los efectos aleatorios debe cumplir la condición $n_j\mathbf{P}_{\mathbf{Z}_j} = \mathbf{Z}_j\mathbf{Z}_j^t$. Cabe destacar

que la caracterización obtenida está restringida a que la matriz de diseño $\mathbf{Z}_j$ sea de la forma $\mathbf{1}_{N_j}$.

4. Al aplicarse la caracterización en el modelo intercepto aleatorio, se tiene que el *BLUP* de la media poblacional finita es la suma de múltiplos de la media muestral, de la media muestral de la j-ésima unidad de nivel 2, y de la media muestral de las unidades que no se encuentran en la j-ésima unidad de nivel 2.

4.2 Recomendaciones

Durante el desarrollo de este trabajo fueron surgiendo preguntas de investigación en varios aspectos. En esta sección se exponen algunas de éstas, relacionadas a situaciones de modelación.

4.2.1 Estimación en áreas pequeñas

En este trabajo el objetivo principal fue la caracterización del *BLUP* de la media poblacional finita $\overline{Y}_j$ en términos de los operadores proyector definidos sobre los subespacios de las matrices de diseño. Se hizo uso del modelo intercepto aleatorio, que es el modelo base en la teoría de estimación en áreas pequeñas, para llevar a cabo la predicción de la media poblacional finita. Tal modelo cumple la condición $d\mathbf{P}_{\mathbf{Z}} = \mathbf{Z}\mathbf{Z}^t$, que permite la caracterización deseada. Mientras que el modelo de pendientes aleatorias

$$Y_{ij} = \beta_{00} + \beta_{10}X_{ij} + u_{0j} + u_{1j}X_{ij} + e_{ij}, \quad i = 1, ..., d, \quad j = 1, ..., k, \qquad (4.2.1)$$

no necesariamente cumple la condición. Así queda abierta la carateriazación del $BLUP$ en términos de los operadores bajo este modelo o modelos en los que la matriz de diseño $\mathbf{Z}$ no cumpla la condición $d\mathbf{P_Z} = \mathbf{ZZ}^t$.

4.2.2 Componentes de la varianza

En el desarrollo de la tesis se trabajó con dos componentes de la varianza, uno en cada nivel. Cuando en el modelo bajo estudio, se tienen más componentes de la varianza no se puede llevar a cabo la caracterización del $BLUP$ de la media poblacional finita $\overline{Y}_j$ en términos de proyectores, por lo menos basados en lo presentado en este trabajo. Por ejemplo en el modelo de pendientes aleatorias intervienen además de los componentes de la varianza σ_e^2 y $\sigma_{u_0}^2$ otros componentes, dependiendo del número de pendientes aleatorias. Por ejemplo en el caso de una pendiente aleatoria, digamos β_1 intervienen los parámetros $\sigma_{u_1}^2$ y $\sigma_{u_{01}}$. Lo que implica que la matriz de varianzas y covarianzas $\mathbf{V}$ no se puede expresar como $\left(d\sigma_{u_0}^2 + \sigma_e^2\right)\mathbf{P_Z} + \sigma_e^2\mathbf{Q_Z}$, lo cual se debe de cumplir. Así queda por desarrollar la caracterización cuando en el modelo bajo estudio intervienen más de dos componentes de la varianza.

4.2.3 Matriz de efectos aleatorios

En este trabajo el objetivo principal fue el de obtener la caracterización de la media poblacional finita bajo el modelo intercepto aleatorio, ya que este modelo es ampliamente usado en estimación en áreas pequeñas. Debido a lo anterior el trabajo desarrollado se restringió a que la matriz $\mathbf{Z}_j$ asociada a los efectos aleatorios fuera de la forma $\mathbf{1}_{N_j}$, queda por desarrollar la caracterización del $BLUP$ en el caso de que la matriz $\mathbf{Z}_j$ no necesariamente sea el vector $\mathbf{1}_{N_j}$.

4.2.4 Modelo ANOVA de dos criterios de clasificación

En este trabajo se presentó la caracterización del *BLUP* de efectos mixtos bajo el contexto de los modelos lineales jerárquicos haciendo uso de la teoría de predicción bajo el *GLMM.* Un modelo ampliamente usado en investigación y en aplicaciones es el modelo ANOVA de dos criterios de clasificación:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ijk}, \quad i = 1, ..., a; \; j = 1, ..., b, \; k = 1, ..., n.$$

Queda abierta la caracterización de los predictores y estimadores de los parámetros que intervienen en este modelo en términos de los operadores proyector o de transformaciones lineales, enfocando los esfuerzos al modelo de tipo III, es decir, al modelo donde intervienen tanto efectos fijos como efectos aleatorios.

Referencias

[1] Battese, G.E., Harter, R.M. and Fuller, W.A. (1988). An error-component model for prediction of county crop areas using survey and satellite data, **Journal of the American Statistical Association**, 83, 28-36.

[2] Ben-Israel, A. and Greville, T.N.E. (2003). **Generalized Inverses. Theory and Applications**. Second Edition. Springer-Verlag. New York.

[3] Bolfarine, H. and Zacks, S. (1992). **Prediction Theory for Finite Populations**, Springer-Verlag, New York.

[4] Bolfarine, H., Zacks, S., Elian, S.N. and Rodrigues, J. (1994). Optimal prediction of the finite population regression coefficient. **Sankhya**, 56, 1-10.

[5] Cervantes, V.H., Santana, A.C., Guilera, G. and Gómez-Benito, J. (2009). Hierarchical linear models in psychiatric: A bibliometric study. **Scientometrics**, 80, 797-808.

[6] Christensen, R. (2001). **Plane Answers to Complex Questions. The Theory of Linear Models**. 3rd Ed Springer, New York, USA.

[7] Coutiño-Estrada, B. y Vidal-Martnez, V.A. (2003). Grain yield statibility of corn hybrids using linear unbiased predictors. **Agrociencia**, 37. 605-617.

[8] Coutiño-Estrada, B. y Vidal-Martnez, V.A. (2006). Componentes de varianza de híbridos de maíz evaluados en la faja maicera de los Estados Unidos. **Agrociencia**, 40. 89-98.

[9] Datta, G.S. and Lahiri, P. (2000). A Unified measure of uncertainty of estimated best linear unbiased predictor in small-area estimation problems. **Statistica Sinica**, 10, 613-627.

[10] De Leeuw, J. and Meijer, E. (2008). **Handbook of Multilevel Analysis**. Springer. New York, USA.

[11] Dempster, A.P., Rubin, D.B., and Tsutakawa, R.K. (1981). Estimation in covariance components models. **Journal of the American Statistical Association**, 76, 341-353.

[12] Fabrizi, E., Ferrante, M.R. and Pacei, S. (2007). Small area astimation of average household income based on unit level models for panel data. **Survey Methodology**, 33, 187-198.

[13] Fay, R.E. and Herriot, R.A. (1979). Estimation of income from small places: An application of James-Stein procedures to census data. **Journal of the American Statistical Association**, 74, 269-277.

[14] Fujimoto, T. and Koga, S. (2009). An application of mixed-effects model to evaluate the effects of initial spancing on radial variation in wood density in Japanese larch (Larix Kaempferi). **Journal of Wood Science**, 0, 1-8.

[15] Gelman, A. and Hill, J. (2007). **Data Analysis Using Regression and Multilevel/Hierarchical Models**. Cambridge University Press.

[16] Goldstein, H. (1995) **Multilevel Statistical Models**, Third Edition, Halsted Press, New York, USA.

[17] Ghosh, M. and Rao, J.N.K. (1994). Small area estimation: an appraisal. **Statistical Science**, 9, 55-93.

[18] Henderson, C.R. (1975). Best linear unbiased estimation and prediction under a selection model, **Biometrics**, 31, 423-447.

[19] Henderson, C.R. (1984). **Applications of Linear Models in Animal Breeding**, University of Guelph, Guelph, Ontario.

[20] Henderson, C.R., Kempthorne, O., Searle, S.R. and von Krosigk, C.N. (1959). Estimation of environmental and genetic trends from records subject to culling, **Biometrics**, 15, 192-218.

[21] Holt, D. and Moura, F.A.S. (1993). Mixed models for making small area estimates. In Kalton, G, Kordos, J and Platek, R, eds: **Small Area Statistics and Survey Designs**. 1, 221-231. Central Statistical Office, Warsaw.

[22] Holt, D. and Moura, F.A.S. (1999). Small area estimation using multi-level models. **Survey Methodology**, 25, 73-80.

[23] Hox, J. (2002). **Multilevel Analysis; Techniques and Applications**. Lawrence Erlbaum Associates, Publishers, London.

[24] Jiang, J. (2007). **Linear and Generalized Linear Mixed Models and Their Applications**. Springer, New York.

[25] Laird, N. and Ware, J. (1982). Random effects models for longitudinal data. **Biometrics**, 38, 963-974.

[26] Littell, R.C., Milliken, G.A., Stroup, W.W., Wolfinger, R.D. and Schabenberger, O. (2006). **SAS for Mixed Models**, Cary, NC: SAS Institute, Inc.

[27] Longford, N.T. (1995). (eds. Arminger, G., Clogg, C.C. and Sobel, M.E.), Random coefficient models. In: **Handbook of Statistical Models for the**

Social and Behavioral Sciences, pp. 519-577, Plenum Press, New York, USA.

[28] McArdle, J.J. and Prescott, C.A. (2005). Mixed-effects variance components models for biometric family analyses. **Behavior Genetics**, 35, 631-652.

[29] Molengberghs, G. and Verbeke, G. (2005). **Models for Discrete Longitudinal Data**. Springer. New York. USA.

[30] Paddison, C. F. (2001). Small area estimation based on a nested-error regression model. **Proceedings of the Annual Meeting of the American Statistical Association**, August 5-9.

[31] Petrucci, A., Pratesi, M., and Salvati, N. (2005). Geografic information in small area estimation: small area models and spatially correlated random area effects. **Statistics in Transition**, 7, 609-623.

[32] Pinheiro J.C. and Bates, D.M. (2009). **Mixed Effects Models in S and S-PLUS**. Springer, New York, USA.

[33] Prasad, N.G.N. and Rao, J.N.K. (1990). The estimation of the mean squared error of small-area estimators. **Journal of the American Statistical Association**, 85, 163-171.

[34] Rabe-Hesketh, S. and Srondal, A. (2008). **Multilevel and Longitudinal Modeling Using Stata**. Second Edition, Stata Press.

[35] Rao, J.N.K. (2003). Some news development in small area estimation. **Proceedings of the Survey Methods Section**, SSC Annual Meerting, June 2003.

[36] Raudenbush, S.W. and Bryk, A.S (2002). **Hierarchical Linear Models Applications and Data Analysis Methods**. 2nd. Ed. Sage Publications, Thousand Oaks, USA.

[37] Rasbash, J., Steele, F., Browne, W.J. and Goldstein, H. (2009). **A user's guide to MLwiN version 2.10**. Center for multilevel modeling. University of Bristol. London. England.

[38] Robinson, G.K. (1991). That BLUP is a good thing: The estimation on random effects. **Statistical Science**, 6, 15-51.

[39] Royall, R.M. (1976) The linear least squares prediction. Approach to two-stage sampling. **Journal of the American Statistical Association**, 71, 657-664.

[40] Royall, R.M. and Cumberland, W.G. (1978). An Empirical Study of Prediction Theory in Finite Population Sampling: Simple Random Sampling and the Ratio Estimator. Ch. 18 in **Survey Sampling and Measurement**, N.K. Namboodiri, ed. New York: Academic Press.

[41] Royall, R.M. and Herson, J. (1973). Robust estimation in finite population I. **Journal of the American Statistical Association**, 68, 880-889.

[42] Searle, S.R., Casella, G. and McCulloch, C.E. (2006). **Variance Component**. 2nd Ed., John Wiley, New York, USA.

[43] Takane, Y. and Yanai, H. (1999). On oblique projectors, **Linear Algebra and its Applications**, 289, 297-310, 1999.

[44] Takane, Y., Tian, Y. and Yanai, H. (2007). On constrained generalized inverses of matrices and their properties. **Annals of the Institute of Statistical Mathematics**, 59, 807-820.

[45] Torabi, M. and Rao, J.N.K. (2008). Small area estimation under a two-level model. **Survey Methodology**, 34, 11-17.

[46] Valliant, R., Dorfman, A.H. and Royall, R.M. (2000). **Finite Population Sampling and Inference: A Prediction Approach**. John Wiley, New York.

[47] Velasco-Luna, F., López S.L. (2003). Modelo de regresión lineal en dos niveles. **Revista de Ciencias Básicas**, Vol. 2, Num. 1, 3-8.

[48] West, B.T., Welch, K.B., and Galecki, A.T. (2007). **Linear Mixed Models. A Practical Guide Using Statistical Software**, Chapman Hall/CRC. Boca Raton, Florida, USA.

[49] You, Y. and Rao, J.N.K. (2000). Hierarchical Bayes estimation of small area means using multilevel models. **Survey Methodology**, 26, 173-181.

[50] Zadlo, T. (2006a) On accuracy of EBLUP under random regression coefficient model. **Statistics in Transition**, 7, 1299-1245.

[51] Zadlo, T. (2006b) On prediction of total value in incompletely specified domains. **Australian & New Zealand Journal of Statistics**, 48, 269-283.

[52] Zuur, A.F., Ieno, A.N., Walker, N.J., Saveliev, A.A. and Smith, G.M., (2009). **Mixed Effects Models and Extensions in Ecology with R**. Springer, New York.