

Instituto Nacional de Estadística y Geografía

Asociación Mexicana de Estadística

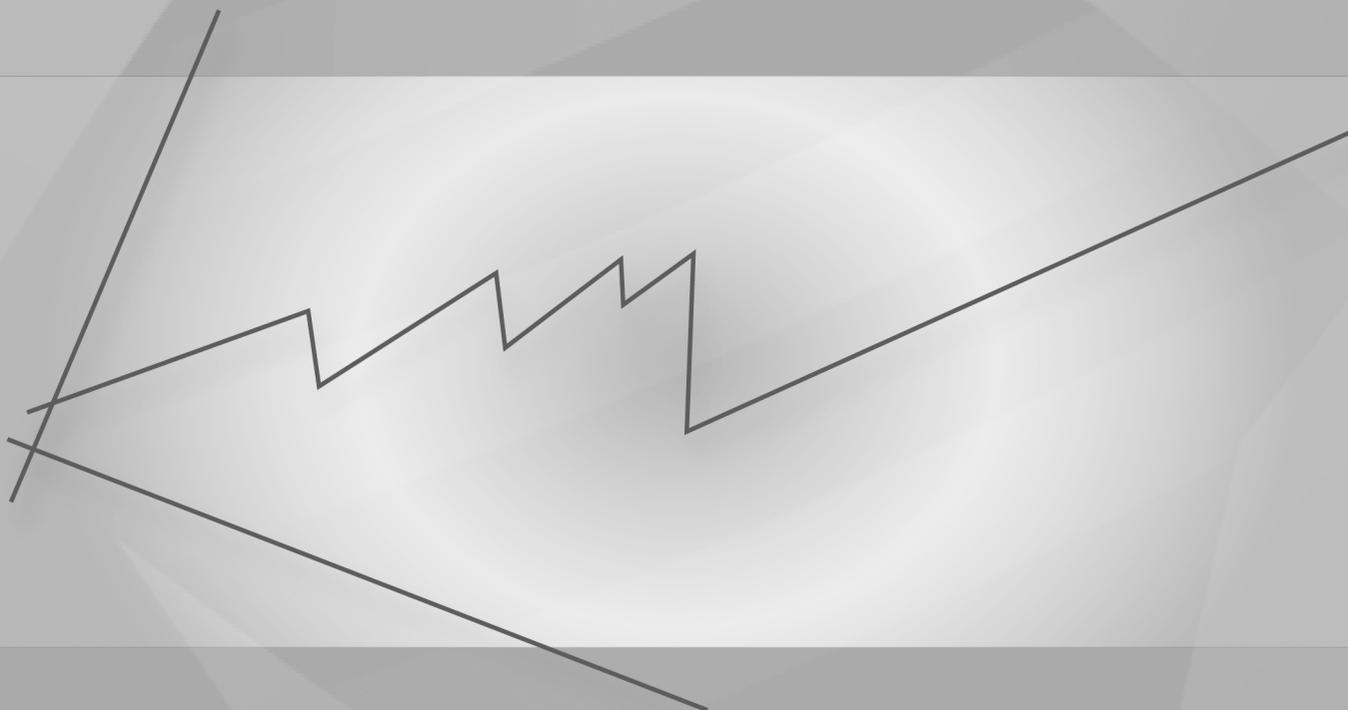
# Memoria del XXV Foro Nacional de Estadística



25  
años

Instituto Nacional de Estadística y Geografía

# Memoria del XXV Foro Nacional de Estadística



DR © 2011, **Instituto Nacional de Estadística y Geografía**  
Edificio Sede  
Av. Héroe de Nacozari Sur Núm. 2301  
Fracc. Jardines del Parque, CP 20276  
Aguascalientes, Ags.

[www.inegi.org.mx](http://www.inegi.org.mx)  
[atencion.usuarios@inegi.org.mx](mailto:atencion.usuarios@inegi.org.mx)

**Memoria del XXV Foro  
Nacional de Estadística**

Impreso en México

# Presentación

Este libro está constituido por 20 artículos que forman la Memoria del XXV Foro Nacional de Estadística, evento que se llevó a cabo en las instalaciones del Instituto Nacional de Salud Pública, en la ciudad de Cuernavaca, Morelos, del 22 al 24 de septiembre de 2010.

La Sección 1 contiene tres artículos y es acerca de temas generales. El primero, de Patricia Romero Mares y Raúl Rueda, invita a preguntarse si, en el caso del muestreo de poblaciones finitas para hacer inferencias, se debe escoger el enfoque basado en el modelo o en el diseño. El segundo, de Alexander von Eye y Patrick Mair, analiza la pérdida de información debido a la dicotomización de variables y presentan un ejemplo empírico donde este proceso no produce tal pérdida. El tercer trabajo, de María Guadalupe Russell Noriega y Enrique Villa Diharce, se refiere al uso de cartas de control en el área de salud.

La Sección 2 es de estadística bayesiana y contiene dos artículos. En el primero, Gabriel Núñez Antonio y Eduardo Gutiérrez Peña presentan “Un Modelo Bayesiano para Datos Longitudinales Circulares”. Por su parte, el trabajo de Olga Vladimirovna Panteleeva, Humberto Vaquera Huerta y Eduardo Gutiérrez González trata de los estimadores de máxima verosimilitud en “Modelos de Mezclas Finitas Univariadas”.

La Sección 3 está dedicada a problemas de inferencia estadística y lo forman cuatro artículos. En el primero, Félix Almendra Arao nos habla de “Una prueba de no inferioridad Basada en Estimadores de Proporción Contraídos” para la comparación de grupos vía la estimación de la desviación estándar de la diferencia de los estimadores de proporción. María D. Kantúm Chim y José A. Villaseñor Alva, se refieren a “La prueba de Bondad de Ajuste  $R$  para la Distribución Exponencial”, donde  $R$  es la razón de dos estimadores del parámetro de escala de la distribución exponencial. Por su parte, Agustín Jaime García Banda, Luis Cruz Kuri e Ismael Sosa Galindo, utilizan el programa Mathematica para mostrar con ejemplos sencillos “Una Aproximación Binomial con Tres parámetros”. El último trabajo de esta sección es el de Luis Cruz-Kuri, Agustín Jaime García Banda e Ismael Sosa Galindo, “Utilización de Procesos de Ramificación para el Estudio del Desarrollo de una Epidemia”.

La Sección 4 contiene cuatro artículos sobre el tema de muestreo. Martín H. Félix Medina discute la “Estimación de Totales y Medias en el Muestreo por Bola de Nieve en Presencia de Probabilidades de Nominación Heterogéneas”. A continuación, Alberto Manuel Padilla Terán nos habla de “Cotas para la varianza, efecto del diseño y coeficiente de variación de proporciones en el muestreo por conglomerados en dos etapas con tamaños iguales”. El tercer trabajo es de Javier Suárez Espinosa “Formulación natural del tamaño de muestra para el caso del Muestreo por Conglomerados en dos Etapas”. Esta sección termina con el artículo de Fernando Velasco Luna y Mario Miguel Ojeda Ramírez, “Caracterización del BLUP de la media poblacional finita  $Y_j$  en estimación en áreas pequeñas (Small Area Estimation)”.

La Sección 5 se dedica a las aplicaciones y está compuesta por siete artículos. El primero es de aplicaciones a las ciencias de la salud: Fidel Ulín Montejo, Jorge A. Pérez Chávez y Rosa Ma. Salinas Hernández discuten un “Análisis de Confiabilidad para Tiempos de Eficacia Analgésica en Pacientes con Cólico Renoureteral”. El siguiente trabajo es sobre cuestiones atmosféricas, en el que Nahun Israel Loya Monares, Hortensia J. Reyes Cervantes y Francisco J. Ariza Hernández hablan de la “Modelación de fenómenos atmosféricos usando Procesos de Poisson No Homogéneos”. El artículo de Soraida Nieto Murillo, Blanca Rosa Pérez Salvador y José Fernando Soriano Flores, “Credit Scoring: Una Aplicación de la Estadística”, presenta una aplicación en finanzas.

Posteriormente vienen dos trabajos de aplicaciones a cuestiones sociales: El primero, de Blanca Rosa Pérez Salvador, “Un modelo de series de tiempo para describir la demanda en grupos escolares con seriación estricta”. El segundo, de Alfredo Cuevas Sandoval, Flaviano Godínez Jaimes y Sulpicio Sánchez Tizapa, “Estudio de factores que influyen en la resistencia de los morteros formulados para reparación de vivienda de interés social en la zona costera de Guerrero”.

Finalmente, se discuten dos aplicaciones a la agronomía: Emilio Padrón Corral, Armando Muñoz Urbina, Haydée de la Garza Rodríguez e Ignacio Méndez Ramírez, en el trabajo “Relación entre Ecuaciones Estructurales y Correlación Canónica en un Experimento con Guayule”, y Lorena Alonso, Dante Covarrubias y Carlos N. Bouza, con “Muestreo por conjuntos ordenados (Ranked Set Sampling) y su aplicación en población de maguey silvestre”.

A nombre de la Asociación Mexicana de Estadística, agradecemos al Instituto Nacional de Salud Pública todo su entusiasmo y dedicación en la organización de este Foro. También estamos en deuda, una vez más, con el Instituto Nacional de Estadística y Geografía por su apoyo para la publicación del presente volumen. Finalmente, nuestro reconocimiento y gratitud a todos y cada uno de los colegas que amablemente accedieron a revisar los trabajos que aquí se presentan.

**El Comité Editorial**

Juan González Hernández

Juan Morales Velasco

Elida Estrada Barragán



# Índice general

## Sección I. Temas Generales

<b>Inferencias basadas, ¿en modelo o en diseño? . . . . .</b>	<b>3</b>
<i>Patricia Romero Mares, Raúl Rueda Díaz del Campo</i>	
<b>On the effects of dichotomizing information . . . . .</b>	<b>11</b>
<i>Alexander von Eye, Patrick Mair</i>	
<b>Uso de cartas de control en el área de la salud . . . . .</b>	<b>21</b>
<i>María Guadalupe Russell Noriega, Enrique Villa Diharce</i>	

## Sección II. Estadística Bayesiana

<b>Un modelo bayesiano para datos longitudinales circulares . . . . .</b>	<b>31</b>
<i>Gabriel Núñez Antonio, Eduardo Gutiérrez Peña</i>	
<b>EMV en modelos de mezclas finitas univariadas . . . . .</b>	<b>37</b>
<i>Olga Vladimirovna Panteleeva, Humberto Vaquera Huerta, Eduardo Gutiérrez González</i>	

## Sección III. Inferencia Estadística

<b>Una prueba de no inferioridad basada en estimadores de proporción contraídos . . . . .</b>	<b>51</b>
<i>Félix Almendra Arao</i>	
<b>La prueba de bondad de ajuste <math>R</math> para la distribución exponencial . . . . .</b>	<b>59</b>
<i>María D. Kantún Chim, José A. Villaseñor Alva</i>	

**Una aproximación binomial con tres parámetros. . . . . 67**

*Agustín Jaime García Banda, Luis Cruz-Kuri, Ismael Sosa Galindo*

**Utilización de procesos de ramificación para el estudio del desarrollo de una epidemia . . . . . 75**

*Luis Cruz-Kuri, Agustín Jaime García Banda, Ismael Sosa Galindo*

## Sección IV. Muestreo

**Estimación de totales y medias en el muestreo por bola de nieve en presencia de probabilidades de nominación heterogéneas . . . . . 85**

*Martín H. Félix Medina*

**Cotas para la varianza, efecto del diseño y coeficiente de variación de proporciones en el muestreo por conglomerados en dos etapas con tamaños iguales . . . . . 91**

*Alberto Manuel Padilla Terán*

**Formulación natural del tamaño de muestra para el caso del muestreo por conglomerados en dos etapas . . . . . 99**

*Javier Suárez Espinosa*

**Caracterización del BLUP de la media poblacional finita  $\bar{Y}_j$  en predicción en áreas pequeñas (Small Area Estimation) . . . . . 105**

*Fernando Velasco Luna, Mario Miguel Ojeda Ramírez*

## Sección V. Aplicaciones

**Análisis de confiabilidad para tiempos de eficacia analgésica en pacientes con cólico renoureteral. . . . . 115**

*Fidel Ulín-Montejo, Jorge A. Pérez Chávez, Rosa Ma. Salinas-Hernández*

---

<b>Modelación de fenómenos atmosféricos usando procesos de Poisson no homogéneos . . . . .</b>	<b>123</b>
<i>Nahun Israel Loya Monares, Hortensia J. Reyes Cervantes, Francisco J. Ariza Hernández</i>	
<b>Credit Scoring: una aplicación de la estadística . . . . .</b>	<b>129</b>
<i>Soraida Nieto Murillo, Blanca Rosa Pérez Salvador, José Fernando Soriano Flores</i>	
<b>Un modelo de series de tiempo para describir la demanda en grupos escolares con seriación estricta. . . . .</b>	<b>137</b>
<i>Blanca Rosa Pérez Salvador</i>	
<b>Estudio de factores que influyen en la resistencia de los morteros formulados para reparación de vivienda de interés social en la zona costera de Guerrero</b>	<b>143</b>
<i>Alfredo Cuevas Sandoval, Flaviano Godínez Jaimes, Sulpicio Sánchez Tizapa</i>	
<b>Relación entre ecuaciones estructurales y correlación canónica en un experimento con guayule. . . . .</b>	<b>149</b>
<i>Emilio Padrón Corral, Haydée de la Garza Rodríguez, Armando Muñoz Urbina, Ignacio Méndez Ramírez</i>	
<b>Muestreo por conjuntos ordenados (Ranked Set Sampling) y su aplicación en población de maguey silvestre. . . . .</b>	<b>157</b>
<i>Lorena Alonso, Dante Covarrubias, Carlos N. Bouza</i>	

## Sección IV

### Muestreo



# Caracterización del BLUP de la media poblacional finita $\bar{Y}_j$ en predicción en áreas pequeñas (Small Area Estimation)

Fernando Velasco Luna<sup>a</sup>, Mario Miguel Ojeda Ramírez<sup>b</sup>  
*Facultad de Estadística e Informática. Universidad Veracruzana*

## 1. Introducción

La teoría de muestreo para poblaciones finitas se encarga de la selección de muestras, de las que se observan y miden características de cada una de las unidades muestreadas; usando estas observaciones la teoría estadística, en este contexto, desarrolla mecanismos para conducir inferencias acerca de ciertas características de la población, como por ejemplo la media poblacional  $\bar{Y} = T/N$  Valliant et al. (2000), donde  $T$  denota el total poblacional y  $N$  denota el número de unidades en la población. Uno de los enfoques de inferencia en la teoría de muestreo de poblaciones finitas para estudiar los procesos de inferencia en el muestreo bietápico es el basado en el Modelo Lineal Mixto. En este enfoque se considera el modelo  $\mathbf{Y}_j = \mathbf{X}_j\boldsymbol{\beta} + \mathbf{Z}_j\mathbf{u}_j + \mathbf{e}_j$ , donde  $\mathbf{Y}_j$ ,  $\mathbf{X}_j$  y  $\mathbf{Z}_j$  denotan el vector respuesta, y las matrices de diseño (variables explicatorias a nivel 1 y nivel 2), respectivamente, en la  $j$ -ésima unidad de nivel 2 (área pequeña), la cual cuenta con  $N_j$  unidades,  $\boldsymbol{\beta}$  denota el vector de parámetros fijos, y  $\mathbf{u}_j$  y  $\mathbf{e}_j$  los efectos aleatorios de nivel 1 y nivel 2, respectivamente; sea  $\mathbf{s}_j$  la muestra de  $n_j$  unidades en la  $j$ -ésima área pequeña, la cual cuenta con  $N_j$  unidades en la población,  $\mathbf{r}_j$  denotando las unidades en la  $j$ -ésima área que no están en  $\mathbf{s}_j$  y  $r_j = N_j - n_j$  el número de unidades no muestreadas. Una vez que la muestra  $\mathbf{s}_j$  ha sido obtenida se tiene la descomposición del modelo para la parte observada, que está dado por:

$$\mathbf{Y}_{j\mathbf{s}} = \mathbf{X}_{j\mathbf{s}}\boldsymbol{\beta} + \mathbf{Z}_{j\mathbf{s}}\mathbf{u}_j + \mathbf{e}_{j\mathbf{s}} \quad (1)$$

---

<sup>a</sup>fvelasco@uv.mx

<sup>b</sup>mojeda@uv.mx

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

$$\mathbf{Y}_{jr} = \mathbf{X}_{jr}\boldsymbol{\beta} + \mathbf{Z}_{jr}u_j + \mathbf{e}_{jr}. \quad (2)$$

Entre todos los predictores, el mejor predictor lineal insesgado (*BLUP*) de la media poblacional finita  $\bar{Y}_j = N_j^{-1} \sum_{i=1}^{N_j} Y_{ij}$  en la  $j$ -ésima área pequeña está dado por  $f_j \bar{Y}_{js} + (1 - f_j) \left[ \bar{\mathbf{X}}_{jr} \hat{\boldsymbol{\beta}}_s + \bar{\mathbf{Z}}_{jr} \mathbf{G} \mathbf{Z}_{js}^t \mathbf{V}_{jss}^{-1} \left( \mathbf{Y}_{js} - \mathbf{X}_{js} \hat{\boldsymbol{\beta}}_s \right) \right]$ , donde  $f_j = n_j/N_j$ , y  $\bar{\mathbf{X}}_{jr}$  y  $\bar{\mathbf{Z}}_{jr}$  son los vectores de medias para las  $r_j$  unidades no muestreadas en la  $j$ -ésima unidad de nivel 2. La media de la población finita  $\bar{Y}_j$  se puede descomponer en la media obtenida de la muestra  $\bar{Y}_{js}$  más la media de las unidades no muestreadas  $\bar{Y}_{jr}$ . Para la parte no muestreada se debe de tener una estimación de la media poblacional  $\mu_j$  de la  $j$ -ésima área pequeña, la cual es un efecto mixto. Velasco y Ojeda (2010a) desarrollan la caracterización del *BLUP* de la media poblacional  $\mu_j$  en términos de los operadores proyector, ortogonal  $\mathbf{P}_Z = \mathbf{Z}(\mathbf{Z}^t\mathbf{Z})^{-1}\mathbf{Z}^t$  y oblicuo  $\mathbf{P}_{\mathbf{XV}} = \mathbf{X}(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{V}^{-1}$ , definidos sobre los subespacios generados por las matrices de diseño. Aunque en la literatura se conocen suficientes resultados acerca de la teoría del álgebra lineal relacionada con la teoría de estimación y prueba de hipótesis en el modelo lineal general (MLG), no existen resultados que caracterizen al *BLUP* de la media poblacional finita  $\bar{Y}_j$  de la  $j$ -ésima unidad de nivel 2 en términos de las matrices de proyección. En este trabajo se presenta la caracterización del *BLUP* de la media poblacional finita  $\bar{Y}_j$  de la  $j$ -ésima área pequeña en términos de los operadores proyector  $\mathbf{P}_Z$  y  $\mathbf{P}_{\mathbf{XV}}$ , aplicandose la caraterización obtenida al modelo intercepto aleatorio.

## 2. Efecto mixto

Se considera el modelo dado por:

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

$$\mathbf{u} \sim N(\mathbf{0}, \sigma_u^2 \mathbf{I}_q), \quad \mathbf{e} \sim N(\mathbf{0}, \sigma_e^2 \mathbf{I}_n), \quad \text{Cov}(\mathbf{e}, \mathbf{u}^t) = \mathbf{0}, \quad (3)$$

donde  $\mathbf{Y} \in \mathbb{R}^n$ ,  $\mathbf{X}$  y  $\mathbf{Z}$  son matrices de orden  $n \times p$  y  $n \times q$ , respectivamente y  $\boldsymbol{\beta} \in \mathbb{R}^p$ . En este caso la matriz de varianzas y covarianzas de  $\mathbf{Y}$  está dada por  $\mathbf{V} = \sigma_u^2 \mathbf{Z}\mathbf{Z}^t + \sigma_e^2 \mathbf{I}_n$ .

Henderson (1975) obtiene el mejor estimador lineal insesgado (*BLUE*) de  $\boldsymbol{\beta}$  y el *BLUP* de  $\mathbf{u}$ , que están dados por  $\hat{\boldsymbol{\beta}} = (\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{V}^{-1}\mathbf{Y}$  y  $\hat{\mathbf{u}} = \mathbf{G}\mathbf{Z}^t\mathbf{V}^{-1} \left( \mathbf{Y} - \mathbf{X} \hat{\boldsymbol{\beta}} \right)$ , respecti-

vamente. Además obtiene el *BLUP* del efecto mixto  $\mathbf{k}^t\boldsymbol{\beta} + \mathbf{m}^t\mathbf{u}$  que está dado por:

$$\mathbf{k}^t \hat{\boldsymbol{\beta}} + \mathbf{m}^t \hat{\mathbf{u}}. \quad (4)$$

Velasco y Ojeda (2010b) desarrollan la caracterización del *BLUP* del efecto mixto  $\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u}$  en términos de los operadores  $\mathbf{P}_{\mathbf{XV}}$  y  $\mathbf{P}_{\mathbf{Z}}$ .

**Teorema 2.1.** *Bajo el modelo (3), si se cumple la condición  $n_j\mathbf{P}_{\mathbf{Z}_j} = \mathbf{Z}_j\mathbf{Z}_j^t$ , entonces el *BLUP* del efecto mixto  $\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u}$  se expresa en términos de los operadores, proyector oblicuo  $\mathbf{P}_{\mathbf{XV}}$  sobre  $S(\mathbf{X})$  y proyector ortogonal  $\mathbf{P}_{\mathbf{Z}}$  sobre  $S(\mathbf{Z})$  por:*

$$[\mathbf{P}_{\mathbf{XV}} + \mathbf{P}_{\mathbf{Z}}\mathbf{B}\mathbf{Q}_{\mathbf{XV}}] \mathbf{Y}, \quad (5)$$

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

*Demostración.* Para una demostración ver Velasco y Ojeda (2010b).

**Observación.** Dada una matriz  $\mathbf{Z}_j$  de orden  $n_j \times q$ , la condición  $n_j\mathbf{P}_{\mathbf{Z}_j} = \mathbf{Z}_j\mathbf{Z}_j^t$  se cumple si  $\mathbf{Z}_j^t\mathbf{Z}_j = n_j\mathbf{I}_q$ , lo cual ocurre si las columnas de la matriz  $\mathbf{Z}_j$  son ortogonales y  $\sum_{i=1}^{n_j} z_{ij}^2 = n_j$ . En términos prácticos, la condición se cumple cuando todos los niveles  $j$  de los efectos aleatorios asociados a  $\mathbf{Z}$ , tienen el mismo número de observaciones (repeticiones).

### 3. Media poblacional $\mu_j$

En esta sección se presenta la caracterización del *BLUP* de la media poblacional  $\mu_j$ , de la  $j$ -ésima área pequeña, en términos de  $\mathbf{P}_{\mathbf{XV}}$ ,  $\mathbf{P}_{\mathbf{Z}}$  y  $\mathbf{T}_{j\mathbf{s}}$ .

Una vez que la muestra  $\mathbf{s}$ , de tamaño  $n$ , ha sido obtenida el vector  $\mathbf{Y}$ , las matrices  $\mathbf{X}$  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 medio de  $\mathbf{Y}_{\mathbf{s}}$ ,  $\mathbf{X}_{\mathbf{s}}$ ,  $\mathbf{V}_{\mathbf{s}}$ ,  $\mathbf{Q}_{\mathbf{X}_{\mathbf{s}}\mathbf{V}_{\mathbf{s}}}$ ,  $\mathbf{P}_{\mathbf{Z}_{\mathbf{s}}}$ ,  $\hat{\boldsymbol{\beta}}_{\mathbf{s}}$  y  $\hat{u}_{j\mathbf{s}}$ , respectivamente.

**Teorema 3.1.**  $\mathbf{T}_{j\mathbf{s}}$  dada por  $\mathbf{T}_{j\mathbf{s}} = \mathbf{X}_j(\mathbf{X}_{\mathbf{s}}^t\mathbf{V}_{\mathbf{s}}^{-1}\mathbf{X}_{\mathbf{s}})^{-1}\mathbf{X}_{\mathbf{s}}^t\mathbf{V}_{\mathbf{s}}^{-1}$  define una transformación lineal de  $\mathbb{R}^n$  a  $\mathbb{R}^{N_j}$ .

La media poblacional  $\mu_j$  se define como  $E(\bar{Y}_j | u_j)$ , que bajo el modelo (2) con  $\mathbf{Z}_j = \mathbf{1}_{n_j}$  está dada por  $\bar{X}_{j\mathbf{r}}^t\boldsymbol{\beta} + u_j$ , que es un efecto mixto.

**Teorema 3.2.** *Bajo el modelo (3) si  $n_j \mathbf{P}_{\mathbf{Z}_{js}} = \mathbf{Z}_{js} \mathbf{Z}_{js}^t$ , entonces el BLUP de  $\mu_j$  se expresa en términos de los operadores proyector  $\mathbf{Q}_{\mathbf{X}_s \mathbf{V}_s}$  y  $\mathbf{P}_{\mathbf{Z}_s}$ , y de la transformación lineal  $\mathbf{T}_{js}$  por:*

$$\left[ \frac{\mathbf{1}_{N_j}^{*jr^t}}{r_j} \mathbf{T}_{js} + \frac{\mathbf{1}_n^{*jst}}{n_j} (\mathbf{P}_{\mathbf{Z}_s} \mathbf{B}_s \mathbf{Q}_{\mathbf{X}_s \mathbf{V}_s}) \right] \mathbf{Y}_s,$$

donde  $\mathbf{1}_n^{*js}$  es un vector de 0's en  $\mathbb{R}^n$  con un 1 en las posiciones correspondientes a las unidades que pertenecen a la  $j$ -ésima área pequeña,  $\mathbf{B}_s = \bigoplus_{j=1}^J (b_j \mathbf{I}_{n_j})$  y  $b_j = n_j \sigma_{u_0}^2 / (n_j \sigma_{u_0}^2 + \sigma_e^2)$ .

*Demostración.* Para una demostración ver Velasco y Ojeda (2010a).

## 4. Caracterización del *BLUP* de la media poblacional finita $\bar{Y}_j$

En esta sección se presenta la caracterización del *BLUP* de  $\bar{Y}_j$  en términos de  $\mathbf{P}_{\mathbf{XV}}$ ,  $\mathbf{P}_{\mathbf{Z}}$  y  $\mathbf{T}_{js}$ .

$\mathbf{1}_{N_j}^{*jr}$  es un vector de 0's en  $\mathbb{R}^{N_j}$  con un 1 en las posiciones correspondientes a las unidades de nivel 1, que pertenecen a la  $j$ -ésima área pequeña, que no están en la muestra.

**Teorema 4.1.** *Bajo el modelo (3), con  $\mathbf{Z}_j = \mathbf{1}_{n_j}$ , el BLUP de la media poblacional finita  $\bar{Y}_j$  está dado por:*

$$f_j \bar{Y}_{js} + (1 - f_j) \left[ \bar{\mathbf{X}}_{jr} \hat{\boldsymbol{\beta}}_s + \hat{u}_{js} \right]. \quad (6)$$

*Demostración.* El *BLUP* de la media poblacional finita  $\bar{Y}_j$  está dado por:

$$f_j \bar{Y}_{js} + (1 - f_j) \left[ \bar{\mathbf{X}}_{jr} \hat{\boldsymbol{\beta}}_s + \bar{\mathbf{Z}}_{jr} \mathbf{G} \mathbf{Z}_{js}^t \mathbf{V}_{jss}^{-1} \left( \mathbf{Y}_{js} + \mathbf{X}_{js} \hat{\boldsymbol{\beta}}_s \right) \right] \quad (7)$$

en (7) interviene el término

$$\bar{\mathbf{Z}}_{jr} \mathbf{G} \mathbf{Z}_{js}^t \mathbf{V}_{jss}^{-1} \left( \mathbf{Y}_{js} + \mathbf{X}_{js} \hat{\boldsymbol{\beta}}_s \right) \quad (8)$$

Bajo el modelo (3), si  $\mathbf{Z}_j = \mathbf{1}_{n_j}$ , entonces  $\bar{\mathbf{Z}}_{jr} = 1$ , por lo que (8) toma la forma:

$$\mathbf{G} \mathbf{Z}_{js}^t \mathbf{V}_{jss}^{-1} \left( \mathbf{Y}_{js} + \mathbf{X}_{js} \hat{\boldsymbol{\beta}}_s \right) \quad (9)$$

que corresponde a  $\hat{u}_{js}$ . De lo cual se sigue (6).

**Teorema 4.2.** *Bajo el modelo (3), con  $\mathbf{Z}_j = \mathbf{1}_{n_j}$ , si  $n_j \mathbf{P}_{\mathbf{Z}_{j_s}} = \mathbf{Z}_{j_s} \mathbf{Z}_{j_s}^t$ ,  $n_j \in \mathbb{R}$ , entonces el BLUP de la media poblacional finita  $\bar{Y}_j$ , se expresa en términos de los operadores proyector  $\mathbf{Q}_{\mathbf{X}_s \mathbf{V}_s}$  y  $\mathbf{P}_{\mathbf{Z}_s}$ , y de la transformación lineal  $\mathbf{T}_{j_s}$  por medio de:*

$$\left( \frac{\mathbf{1}_n^{*j_s t}}{N_j} + \frac{r_j}{N_j} \left[ \frac{\mathbf{1}_{N_j}^{*j_r t}}{r_j} \mathbf{T}_{j_s} + \frac{\mathbf{1}_n^{*j_s t}}{n_j} (\mathbf{P}_{\mathbf{Z}_s} \mathbf{B}_s \mathbf{Q}_{\mathbf{X}_s \mathbf{V}_s}) \right] \right) \mathbf{Y}_s,$$

donde  $\mathbf{B}_s = \bigoplus_{j=1}^J (b_j \mathbf{I}_{n_j})$  y  $b_j = n_j \sigma_u^2 / (n_j \sigma_u^2 + \sigma_e^2)$ .

*Demostración.* Por el teorema 4.1,  $BLUP(\bar{Y}_j) = f_j \bar{Y}_{j_s} + (1 - f_j) \left[ \bar{\mathbf{X}}_{j_r} \hat{\boldsymbol{\beta}}_s + \hat{u}_{j_s} \right]$ . Además, por el teorema 3.2,  $BLUP(\mu_j) = \left[ \frac{\mathbf{1}_{N_j}^{*j_r t}}{r_j} \mathbf{T}_{j_s} + \frac{\mathbf{1}_n^{*j_s t}}{n_j} (\mathbf{P}_{\mathbf{Z}_s} \mathbf{B}_s \mathbf{Q}_{\mathbf{X}_s \mathbf{V}_s}) \right] \mathbf{Y}_s$  y recordando que  $\mu_j$  está dado por  $\bar{\mathbf{X}}_{j_r}^t \boldsymbol{\beta} + u_j$ , se tiene

$$\begin{aligned} BLUP(\bar{Y}_j) &= \left( \frac{n_j}{N_j} \right) \bar{Y}_{j_s} + \left( \frac{r_j}{N_j} \right) BLUP(\mu_j) \\ &= \left( \frac{n_j}{N_j} \right) \left( \frac{\mathbf{1}_n^{*j_s t}}{n_j} \right) \mathbf{Y}_s + \left( \frac{r_j}{N_j} \right) BLUP(\mu_j) \\ &= \frac{\mathbf{1}_n^{*j_s t}}{N_j} \mathbf{Y}_s + \frac{r_j}{N_j} \left[ \frac{\mathbf{1}_{N_j}^{*j_r t}}{r_j} \mathbf{T}_{j_s} + \frac{\mathbf{1}_n^{*j_s t}}{n_j} (\mathbf{P}_{\mathbf{Z}_s} \mathbf{B}_s \mathbf{Q}_{\mathbf{X}_s \mathbf{V}_s}) \right] \mathbf{Y}_s \\ &= \left( \frac{\mathbf{1}_n^{*j_s t}}{N_j} + \frac{r_j}{N_j} \left[ \frac{\mathbf{1}_{N_j}^{*j_r t}}{r_j} \mathbf{T}_{j_s} + \frac{\mathbf{1}_n^{*j_s t}}{n_j} (\mathbf{P}_{\mathbf{Z}_s} \mathbf{B}_s \mathbf{Q}_{\mathbf{X}_s \mathbf{V}_s}) \right] \right) \mathbf{Y}_s. \end{aligned}$$

## 5. Caracterización en el modelo intercepto aleatorio

Se presenta la caracterización del BLUP de la media poblacional finita  $\bar{Y}_j$ , considerando el caso balanceado, es decir  $n_j = d \forall j = 1, \dots, k$ , bajo el modelo intercepto aleatorio, ya que este modelo es ampliamente usado en la teoría de estimación en áreas pequeñas. Este modelo está dado para el nivel 1 por:

$$Y_{ij} = \mu_j + e_{ij}, \quad i = 1, \dots, d, \quad j = 1, \dots, k.$$

y para el nivel 2 por:

$$\mu_j = \mu + u_j, \quad j = 1, \dots, k.$$

lo cual lleva al modelo combinado:

$$Y_{ij} = \mu + u_j + e_{ij}, \quad i = 1, \dots, d, \quad j = 1, \dots, k. \quad (10)$$

donde  $\mu$  es un parámetro fijo;  $u_j$  es el efecto aleatorio;  $u_j$  y  $e_{ij}$  son independientes, con  $u_j \sim N(0, \sigma_{u0}^2)$  y  $e_{ij} \sim N(0, \sigma_e^2)$ . El modelo para la  $j$ -ésima unidad de nivel 2 tiene la forma  $\mathbf{Y}_j = \mathbf{1}_d \mu + \mathbf{1}_d u_j + \mathbf{e}_j, j = 1, \dots, k$ . En este caso  $\mathbf{X} = \mathbf{1}_k \otimes \mathbf{1}_d$  y  $\mathbf{Z} = \mathbf{I}_k \otimes \mathbf{1}_d$ , por lo que

$$\mathbf{P}_{\mathbf{X}_s \mathbf{V}_s} = \frac{\mathbf{1}_{kd} \mathbf{1}_{kd}^t}{kd} \text{ y } \mathbf{P}_{\mathbf{Z}_s} = \frac{(\mathbf{I}_k \otimes \mathbf{1}_d \mathbf{1}_d^t)}{d}. \quad (11)$$

**Teorema 5.1.** *Bajo el modelo intercepto aleatorio (10). El BLUP de la media poblacional finita  $\bar{Y}_j$  está dado por:*

$$\frac{n_j}{N_j} \bar{Y}_{js} + \frac{r_j}{N_j} \left[ \bar{Y}_s + \frac{c(k-1)}{k} [\bar{Y}_{js} - \bar{Y}_{(-j)s}] \right] \quad (12)$$

donde  $\bar{Y}_s, \bar{Y}_{js}$  y  $\bar{Y}_{(-j)s}$  denotan la media muestral, la media muestral de la  $j$ -ésima unidad de nivel 2, y la media muestral de las unidades de nivel 2 restantes, respectivamente.

*Demostración.* Por el teorema 4.2, considerando el caso balanceado

$$BLUP(\bar{Y}_j) = \left( \frac{\mathbf{1}_n^{*jst}}{N_j} + \frac{r_j}{N_j} \left[ \frac{\mathbf{1}_{N_j}^{*jr^t}}{r_j} \mathbf{T}_{js} + \frac{\mathbf{1}_n^{*jst}}{d} (c \mathbf{P}_{\mathbf{Z}_s} \mathbf{Q}_{\mathbf{X}_s \mathbf{V}_s}) \right] \right) \mathbf{Y}_s.$$

Además,  $\mathbf{T}_{js} = \frac{\mathbf{1}_{N_j} \mathbf{1}_{kd}^t}{kd}$  y de (11)  $c \mathbf{P}_{\mathbf{Z}_s} \mathbf{Q}_{\mathbf{X}_s \mathbf{V}_s} = \frac{c}{kd} (k(\oplus \mathbf{1}_d \mathbf{1}_d^t) - \mathbf{1}_{kd} \mathbf{1}_{kd}^t)$ , por lo que

$$\begin{aligned} BLUP(\bar{Y}_j) &= \left( \frac{\mathbf{1}_n^{*jst}}{N_j} + \frac{r_j}{N_j} \left[ \frac{\mathbf{1}_{N_j}^{*jr^t}}{r_j} \mathbf{T}_{js} + \frac{\mathbf{1}_n^{*jst}}{d} (c \mathbf{P}_{\mathbf{Z}_s} \mathbf{Q}_{\mathbf{X}_s \mathbf{V}_s}) \right] \right) \mathbf{Y}_s \\ &= \frac{\mathbf{1}_n^{*jst}}{N_j} \mathbf{Y}_s + \frac{r_j}{N_j} \left[ \frac{\mathbf{1}_{N_j}^{*jr^t}}{r_j} \mathbf{T}_{js} + \frac{\mathbf{1}_n^{*jst}}{d} (c \mathbf{P}_{\mathbf{Z}_s} \mathbf{Q}_{\mathbf{X}_s \mathbf{V}_s}) \right] \mathbf{Y}_s \\ &= \frac{\mathbf{1}_n^{*jst}}{N_j} \mathbf{Y}_s + \frac{r_j}{N_j} \left[ \bar{Y}_s + \frac{c(k-1)}{k} [\bar{Y}_{js} - \bar{Y}_{(-j)s}] \right] \\ &= \frac{n_j}{N_j} \bar{Y}_{js} + \frac{r_j}{N_j} \left[ \bar{Y}_s + \frac{c(k-1)}{k} [\bar{Y}_{js} - \bar{Y}_{(-j)s}] \right]. \end{aligned}$$

## 6. Conclusiones

En este trabajo se expresó el *BLUP* de la media poblacional finita  $\bar{Y}_j$  de la  $j$ -ésima área pequeña como la suma ponderada de un elemento en  $S(\mathbf{X}_j)$  y un elemento en el espacio  $S(\mathbf{Z}_s)$ . Lo anterior al aplicarlo al modelo intercepto aleatorio sin variables explicatorias, considerando el caso balanceado, permitió expresar el *BLUP* de la media poblacional finita  $\bar{Y}_j$  como la suma de múltiplos de  $\bar{Y}_s$ ,  $\bar{Y}_{js}$  y  $\bar{Y}_{(-j)s}$ , que denotan la media muestral, la media muestral de las unidades en la  $j$ -ésima área pequeña y la media muestral de las unidades que no pertenecen a la  $j$ -ésima área pequeña, respectivamente. Se espera que esta caracterización en términos de los proyectores permita una mejor comprensión de las propiedades del *BLUP* de  $\bar{Y}_j$  tal como sucede en la caracterización del estimador de parámetros  $\beta$  en el MLG.

## Bibliografía

- Henderson, C. (1975), “Best linear unbiased estimation and prediction under a selection model”, *Biometrics* **31**, 423–447.
- Valliant, R., Dorfman, A. y Royall, R. (2000), *Finite Population Sampling and Inference: A Prediction Approach*, New York: John Wiley.
- Velasco, L. y Ojeda, R. (2010a), “Caracterización del blup de la media poblacional en el modelo lineal general mixto”, *Memorias del XXIV Foro nacional de Estadística, INEGI: México* pp. 81–87.
- Velasco, L. y Ojeda, R. (2010b), “Caracterización del blup del efecto mixto  $x\beta+zu$ ”, *Editores: Taponar S.F.S. and Cruz S.H.A. and Reyes C.H. and Zacarías F.J.D. Aportaciones y Aplicaciones de la Probabilidad y la Estadística, Puebla, México: Benemérita Universidad Autónoma de Puebla* .

