

CELADE

DOCUMENTO  
MICROFILMADO

DOCPAL

CEPAL/EST/Borrador/151  
División de Estadística  
Carlos Cavallini  
Asesor Regional en Muestreo  
para Estadísticas Demográ-  
ficas adscrito a la CEPAL  
Diciembre de 1976

ESTIMADOR BASADO EN UN MODELO ADITIVO

76-12-2545 - 300

CELADE - SISTEMA DOCPAL  
DOCUMENTACION  
SOBRE POBLACION EN  
AMERICA LATINA

Handwritten text, possibly a list or notes, located in the upper left corner of the page. The text is faint and difficult to read.

Handwritten text, possibly a signature or a specific note, located in the middle of the page.

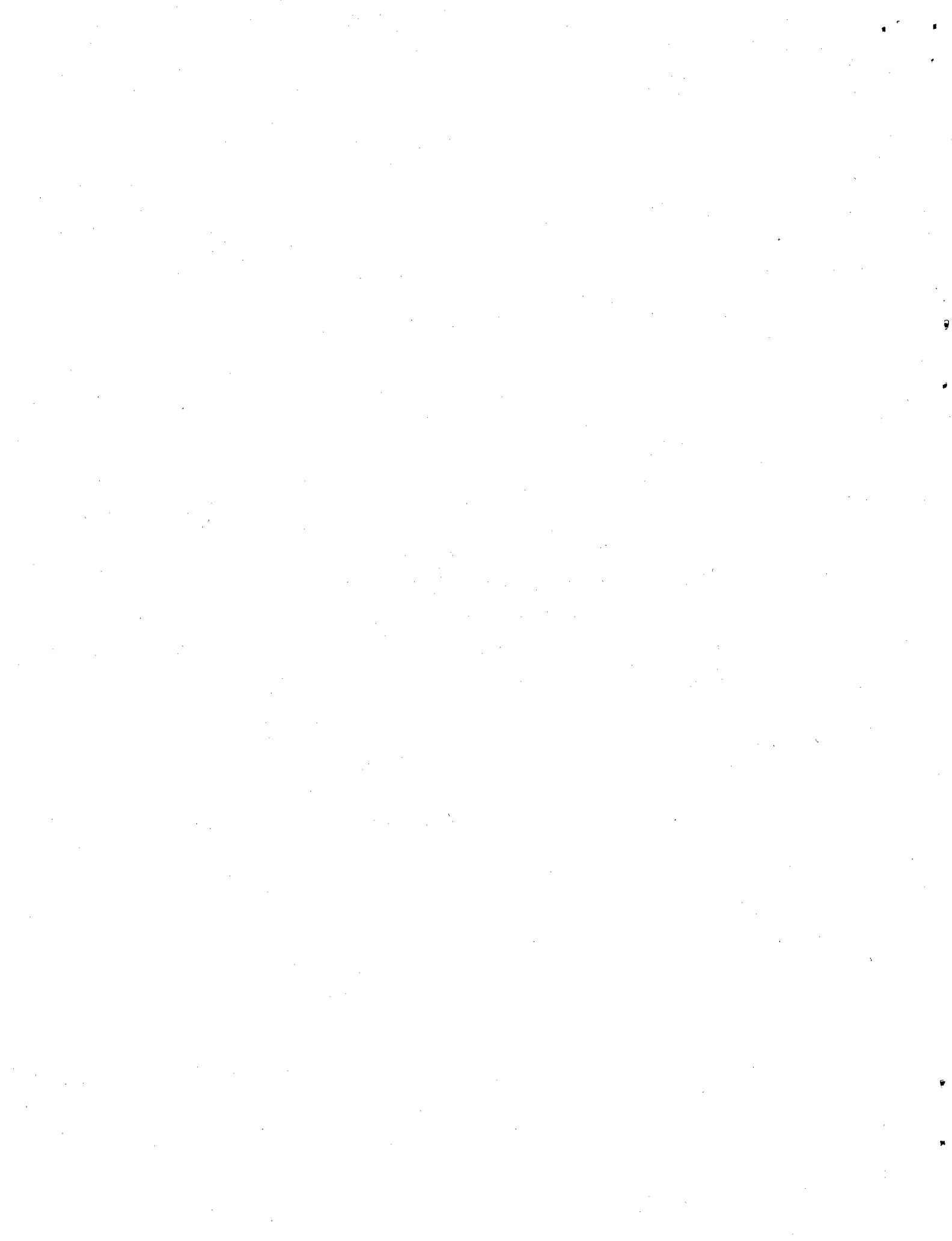
Handwritten text, possibly a date or a reference number, located in the lower right corner of the page.

## Contenido

	<u>Página</u>	<u>Párrafo</u>
Objetivo .....	1	1
Definiciones .....	1	2
Estimador Aditivo (1) .....	2	3
Modelo lineal (5) .....	2	4
Muestras independientes de periodo a periodo .....	3	6.1
Varianzas (20) .....	3	6.1
Varianzas comunes para las $i$ y para las $j$ (21) .....	6	6.2.1
Varianzas comunes para las $i$ pero no para las $j$ (24) ..	6	6.2.2
Sesgo del estimador (27) .....	7	6.3
Estimación del desvío medio cuadrático del estimador (38)	8	6.4
Muestras correlacionadas de periodo a periodo .....	10	7.1
Varianzas comunes para las $i$ pero no para las $j$ (47) ..	10	7.1
Varianzas comunes para las $i$ y para las $j$ (50) .....	12	7.2
Estimación del desvío medio cuadrático del estimador (56)	13	7.3
Estimación de la eficiencia .....	14	7.7.1
Ejemplo práctico .....	15	8.1
Resumen .....	18	9.1
Bibliografía .....	19	

$i$ : periodos

$j$ : estratos



## Estimador Basado en un Modelo Aditivo

1. El objeto de este trabajo está relacionado con las estadísticas continuas y con los procedimientos de combinar los datos.

En las estadísticas sobre población, empleo, producción, precios, y todas aquellas que se elaboran en forma periódica, puede resultar eficiente utilizar estimadores compuestos. El estimador compuesto que aquí se propone hace uso de la combinación de otros estimadores referidos a la misma variable y utiliza la correlación que puede existir entre estos estimadores a través de los distintos períodos.

La utilización de este estimador compuesto no incrementa el costo operativo de la investigación. El trabajo es de laboratorio y se puede llegar a aumentar la eficiencia relativa (eficiencia por unidad de costo) de las estimaciones.

2. Consideremos el caso de una población distribuida en  $k$  estratos donde periódicamente se dan estimaciones por estrato. Designando con  $t$  el número de períodos definimos,

$y_{ijh}$

$h$  - observación en el  $j$  - estrato  
del  $i$  - período, donde

$i = 1, 2, \dots, t$  períodos

$j = 1, 2, \dots, k$  estratos

$h = 1, 2, \dots, n_{ij}$  observaciones en la  
( $ij$ ) - muestra

$N_{ij}$

número de unidades muestrales en la  
( $ij$ ) - calda

$n_{ij}$

número de unidades muestrales en la  
( $ij$ ) - muestra

$$\hat{Y}_{ij} = \frac{N_{ij}}{n_{ij}} \sum_h^{n_{ij}} y_{ijh}$$

estimador regular insesgado del total de la población en la (ij)-celda

3. Establezcamos el siguiente estimador

$$y_{ij}^o = \bar{Y}_{i.} + \bar{Y}_{.j} - \bar{Y}_{..} \quad (1)$$

que tiene la propiedad de ajustar las estimaciones muestrales de las celdas de tal manera que las nuevas estimaciones no modifiquen los totales marginales estimados originalmente.

Es decir

$$\sum_i^t y_{ij}^o = \sum_i^t \hat{Y}_{ij} \quad (2)$$

$$\sum_j^k y_{ij}^o = \sum_j^k \hat{Y}_{ij} \quad (3)$$

$$\sum_i^t \sum_j^k y_{ij}^o = \sum_i^t \sum_j^k \hat{Y}_{ij} \quad (4)$$

4. El estimador establecido se obtiene ajustando a los datos el siguiente modelo lineal

$$Y_{ij} = u + a_i + b_j + e_{ij} \quad (5)$$

con

$$\hat{u} = \bar{Y}_{..} \quad (6)$$

$$\hat{a}_i = \bar{Y}_{i.} - \bar{Y}_{..} \quad (7)$$

$$\hat{b}_j = \bar{Y}_{.j} - \bar{Y}_{..} \quad (8)$$

siendo

$$\bar{Y}_{..} = \frac{\sum_{i=1}^k \sum_{j=1}^k \hat{Y}_{ij}}{t \cdot k} \quad (9)$$

$$\bar{Y}_{i.} = \frac{\sum_{j=1}^k \hat{Y}_{ij}}{k} \quad (10)$$

$$\bar{Y}_{.j} = \frac{\sum_{i=1}^t \hat{Y}_{ij}}{t} \quad (11)$$

5. Los dos esquemas muestrales que a continuación se tratan están basados en el supuesto que la selección de las muestras en cada período es independiente de estrato a estrato. Es decir la probabilidad de seleccionar una unidad muestral en el estrato  $j$  no está condicionada a la probabilidad de seleccionar una unidad muestral en el estrato  $j'$ .

6.1 Muestras independientes de período a período.

Bajo el supuesto que las muestras son independientes de período a período, la varianza del estimador  $\hat{Y}_{ij}$  viene dada por

$$\begin{aligned} \text{Var} (\hat{Y}_{ij}) &= \text{Var} (\bar{Y}_{i.} + \bar{Y}_{.j} - \bar{Y}_{..}) \\ &= \text{Var} (\bar{Y}_{i.}) + \text{Var} (\bar{Y}_{.j}) + \text{Var} (\bar{Y}_{..}) \\ &\quad + 2 \text{Cov} (\bar{Y}_{i.}; \bar{Y}_{.j}) - 2 \text{Cov} (\bar{Y}_{i.}; \bar{Y}_{..}) \\ &\quad - 2 \text{Cov} (\bar{Y}_{.j}; \bar{Y}_{..}) \end{aligned} \tag{12}$$

donde

$$\begin{aligned} \text{Var} (\bar{Y}_{i.}) &= \frac{1}{k^2} \text{Var} \left( \sum_j^k \hat{Y}_{ij} \right) \\ &= \frac{1}{k^2} \left[ \sum_j^k \text{Var} (\hat{Y}_{ij}) + 2 \sum_{j < j'}^k \text{Cov} (\hat{Y}_{ij}; \hat{Y}_{ij'}) \right] \end{aligned} \tag{13}$$

Pero de acuerdo con el supuesto establecido en el esquema muestral, todas las covarianzas valen cero

Por tanto

$$\text{Var} (\bar{Y}_{i.}) = \frac{1}{k^2} \sum_j^k \text{Var} (\hat{Y}_{ij}) \tag{14}$$



Asimismo

$$\text{Var} (\bar{Y}_{.j}) = \frac{1}{t^2} \sum_i^t \text{Var} (\hat{Y}_{ij}) \quad (15)$$

$$\text{Var} (\bar{Y}_{..}) = \frac{1}{t^2 k^2} \sum_i^t \sum_j^k \text{Var} (\hat{Y}_{ij}) \quad (16)$$

y

$$\begin{aligned} \text{Cov} (\bar{Y}_{i.}; \bar{Y}_{.j}) &= \frac{1}{t k} \text{Cov} \left( \sum_j^k \hat{Y}_{ij}; \sum_i^t \hat{Y}_{ij} \right) \\ &= \frac{1}{t k} \text{Var} (\hat{Y}_{ij}) \end{aligned} \quad (17)$$

$$\text{Cov} (\bar{Y}_{i.}; \bar{Y}_{..}) = \frac{1}{t k^2} \sum_j^k \text{Var} (\hat{Y}_{ij}) \quad (18)$$

$$\text{Cov} (\bar{Y}_{.j}; \bar{Y}_{..}) = \frac{1}{t^2 k} \sum_i^t \text{Var} (\hat{Y}_{ij}) \quad (19)$$

Sustituyendo en (12) obtenemos

$$\begin{aligned} \text{Var} (Y_{ij}^o) &= \frac{1}{k^2} \sum_j^k \text{Var} (\hat{Y}_{ij}) + \frac{1}{t^2} \sum_i^t \text{Var} (\hat{Y}_{ij}) \\ &+ \frac{1}{t^2 k^2} \sum_i^t \sum_j^k \text{Var} (\hat{Y}_{ij}) + \frac{2}{t k} \text{Var} (\hat{Y}_{ij}) \\ &- \frac{2}{t k^2} \sum_j^k \text{Var} (\hat{Y}_{ij}) - \frac{2}{t^2 k} \sum_i^t \text{Var} (\hat{Y}_{ij}) \quad (20) \end{aligned}$$

2.2.1 Para comparar la  $\text{Var} (Y_{ij}^o)$  con la  $\text{Var} (\hat{Y}_{ij})$  consideremos primero el caso especial que las  $\text{Var} (\hat{Y}_{ij})$  son comunes para todas las  $i$  y para todas las  $j$ . En este caso

$$\text{Var} (Y_{ij}^o) = \frac{t+k-1}{t k} \text{Var} (\hat{Y}_{ij}) \quad (21)$$

$$\text{Var} (Y_{ij}^o) < \text{Var} (\hat{Y}_{ij}) \quad (22)$$

para  $t ; k > 1$ .

2.2.2 Consideremos ahora el caso más real, que las  $\text{Var} (\hat{Y}_{ij})$  son comunes para un estrecho  $i$  a través de los periodos. Es decir  $\text{Var} (\hat{Y}_{ij})$  es común para las  $i$  pero no para las  $j$ .

Entonces

$$\sum_i^t \text{Var} (Y_{ij}^o) = t \text{Var} (\hat{Y}_{ij}) \quad (23)$$

obteniéndose

$$\text{Var} (Y_{ij}^{\circ}) = \frac{(t-1)}{tk^2} \sum_j^k \text{Var} (Y_{ij}^{\wedge}) + \frac{1}{t} \text{Var} (Y_{ij}^{\wedge}) \quad (24)$$

De aquí se obtiene que

$$\text{Var} (Y_{ij}^{\circ}) < \text{Var} (Y_{ij}^{\wedge}) \quad (25)$$

si

$$\text{Var} (Y_{ij}^{\wedge}) > \frac{1}{k^2} \sum_j^k \text{Var} (Y_{ij}^{\wedge}) \quad (26)$$

Esto significa que el estimador aditivo será más eficiente que el estimador regular cuando la varianza de la celda que se esté considerando sea mayor que la media de las varianzas de todos los estratos para ese periodo, dividido por el número de estratos.

6.3 El sesgo del estimador aditivo es

$$\begin{aligned} \text{Sesgo} (Y_{ij}^{\circ}) &= E (Y_{ij}^{\circ}) - Y_{ij} \\ &= \frac{\sum_j^k Y_{ij}}{k} + \frac{\sum_i^t Y_{ij}}{t} - \frac{\sum_i^t \sum_j^k Y_{ij}}{tk} - Y_{ij} \quad (27) \end{aligned}$$

Este sesgo puede atribuirse a la falla del modelo lineal, o "interacción". Sin embargo la suma de los sesgos sobre las j y la suma de los sesgos sobre las i suman cero. Es decir

$$\sum_j^k \text{Sesgo} (Y_{ij}^{\circ}) = \sum_i^t \text{Sesgo} (Y_{ij}^{\circ}) = 0 \quad (28)$$

6.4 Para estimar el desvío medio cuadrático de  $\overset{\circ}{Y}_{ij}$  partimos de la expresión

$$\begin{aligned} E \left( \overset{\circ}{Y}_{ij} - \hat{Y}_{ij} \right)^2 &= E \left[ \overset{\circ}{Y}_{ij} - E \left( \overset{\circ}{Y}_{ij} \right) + E \left( \overset{\circ}{Y}_{ij} \right) - \hat{Y}_{ij} \right]^2 \\ &= \text{Var} \left( \overset{\circ}{Y}_{ij} \right) - 2 \text{Cov} \left( \overset{\circ}{Y}_{ij}; \hat{Y}_{ij} \right) + E \left[ E \left( \overset{\circ}{Y}_{ij} \right) - \hat{Y}_{ij} \right]^2 \end{aligned} \quad (29)$$

Pero

$$\begin{aligned} E \left[ E \left( \overset{\circ}{Y}_{ij} \right) - \hat{Y}_{ij} \right]^2 &= E \left[ E \left( \overset{\circ}{Y}_{ij} \right) - Y_{ij} + Y_{ij} - \hat{Y}_{ij} \right]^2 \\ &= E \left[ E \left( \overset{\circ}{Y}_{ij} \right) - Y_{ij} \right]^2 + E \left( Y_{ij} - \hat{Y}_{ij} \right)^2 \\ &\quad + 2 E \left\{ \left[ E \left( \overset{\circ}{Y}_{ij} \right) - Y_{ij} \right] \left[ Y_{ij} - \hat{Y}_{ij} \right] \right\} \\ &= \left[ \text{Sesgo} \left( \overset{\circ}{Y}_{ij} \right) \right]^2 + \text{Var} \left( \hat{Y}_{ij} \right) \end{aligned} \quad (30)$$

dado que  $E \left( \hat{Y}_{ij} \right) = Y_{ij}$

Por tanto sustituyendo en (29) queda

$$\begin{aligned} E \left( \overset{\circ}{Y}_{ij} - \hat{Y}_{ij} \right)^2 &= \text{Var} \left( \overset{\circ}{Y}_{ij} \right) - 2 \text{Cov} \left( \overset{\circ}{Y}_{ij}; \hat{Y}_{ij} \right) \\ &\quad + \left[ \text{Sesgo} \left( \overset{\circ}{Y}_{ij} \right) \right]^2 + \text{Var} \left( \hat{Y}_{ij} \right) \end{aligned} \quad (31)$$

Fero por definición, el desvío medio cuadrático de  $\overset{\circ}{Y}_{ij}$  es igual a la varianza de  $\overset{\circ}{Y}_{ij}$  más el sesgo al cuadrado de  $\overset{\circ}{Y}_{ij}$ . Luego de la expresión (31) podemos escribir

$$DMC(\overset{\circ}{Y}_{ij}) = E(\overset{\circ}{Y}_{ij} - \hat{Y}_{ij})^2 = \text{Var}(\hat{Y}_{ij}) + 2 \text{Cov}(\overset{\circ}{Y}_{ij}; \hat{Y}_{ij}) \quad (32)$$

6.5 Para hallar la covarianza de (32) sustituimos en ella el estimador aditivo por los sumandos del modelo. Obtenemos así

$$\begin{aligned} \text{Cov}(\overset{\circ}{Y}_{ij}; \hat{Y}_{ij}) &= \text{Cov}(\bar{Y}_{i.}; \hat{Y}_{ij}) + \text{Cov}(\bar{Y}_{.j}; \hat{Y}_{ij}) \\ &\quad - \text{Cov}(\bar{Y}_{..}; \hat{Y}_{ij}) \end{aligned} \quad (33)$$

Desarrollando cada covarianza del segundo miembro y recordando el supuesto de muestras independientes de período a período y varianzas comunes para un estrato a través de los períodos, obtenemos

$$\text{Cov}(\bar{Y}_{i.}; \hat{Y}_{ij}) = \frac{1}{k} \quad \text{Var}(\hat{Y}_{ij}) \quad (34)$$

$$\text{Cov}(\bar{Y}_{.j}; \hat{Y}_{ij}) = \frac{1}{t} \quad \text{Var}(\hat{Y}_{ij}) \quad (35)$$

$$\text{Cov}(\bar{Y}_{..}; \hat{Y}_{ij}) = \frac{1}{t k} \quad \text{Var}(\hat{Y}_{ij}) \quad (36)$$

Reemplazando en (32) y operando el álgebra

$$\begin{aligned} \text{DMC} (\hat{Y}_{ij}) &= E (\hat{Y}_{ij} - \hat{Y}_{ij})^2 - \text{Var} (\hat{Y}_{ij}) \\ &+ 2 \text{Var} (\hat{Y}_{ij}) \left[ \frac{1}{k} + \frac{1}{t} - \frac{1}{tk} \right] \\ &= E (\hat{Y}_{ij} - \hat{Y}_{ij})^2 - \frac{t(k-2) - 2(k-1)}{tk} \text{Var} (\hat{Y}_{ij}) \end{aligned}$$

6.6 La estimación del desvío medio cuadrático es

$$\hat{\text{DMC}} (\hat{Y}_{ij}) = (\hat{Y}_{ij} - \hat{Y}_{ij})^2 - \frac{t(k-2) - 2(k-1)}{tk} \hat{\text{Var}} (\hat{Y}_{ij}) \quad (38)$$

Se observa que si hay un solo estrato, es decir  $k = 1$ , el  $\hat{\text{DMC}} (\hat{Y}_{ij})$  es igual a la  $\hat{\text{Var}} (\hat{Y}_{ij})$  por ser en este caso  $\hat{Y}_{ij} = \hat{Y}_{ij}$ .

7.1 Muestras correlacionadas de periodo a periodo.

Dejamos ahora de lado el supuesto que las muestras son independientes de periodo a periodo pero mantenemos el supuesto que son independientes de estrato a estrato. El supuesto de varianzas comunes para un estrato a través de los periodos se mantiene.

Luego

$$\text{Var} (\bar{Y}_{i.}) = \frac{1}{k^2} \sum_j^k \text{Var} (\hat{Y}_{ij}) \quad (39)$$

$$\begin{aligned} \text{Var} (\bar{Y}_{.j}) &= \frac{1}{t^2} \sum_i^t \sum_{i'}^t \text{Cov} (\hat{Y}_{ij}; \hat{Y}_{i'j}) \\ &= \frac{1}{t^2} \left[ t \text{Var} (\hat{Y}_{ij}) + \sum_{i \neq i'}^t \text{Cov} (\hat{Y}_{ij}; \hat{Y}_{i'j}) \right] \quad (40) \end{aligned}$$

Pero

$$\sum_{i \neq i'}^t \text{Cov} (\hat{Y}_{ij} ; \hat{Y}_{i'j}) = \text{Var} (\hat{Y}_{ij}) \sum_{i \neq i'}^t r (i i') \quad (41)$$

donde  $r (i i')$  es el coeficiente de correlación entre  $\hat{Y}_{ij}$  e  $\hat{Y}_{i'j}$ .

Por tanto

$$\text{Var} (\bar{Y}_{.j}) = \frac{\text{Var} (\hat{Y}_{ij})}{t^2} \left[ t + \sum_{i \neq i'}^t r (i i') \right] \quad (42)$$

y

$$\text{Var} (\bar{Y}_{..}) = \frac{1}{t^2 k^2} \sum_j^k \left\{ \text{Var} (\hat{Y}_{ij}) \left[ t + \sum_{i \neq i'}^t r_j (i i') \right] \right\} \quad (43)$$

$$\text{Cov} (\bar{Y}_{i.} ; \bar{Y}_{.j}) = \frac{1}{t k} \text{Var} (\hat{Y}_{ij}) \sum_{i'}^t r (i i') \quad (44)$$

$$\text{Cov} (\bar{Y}_{.j} ; \bar{Y}_{..}) = \frac{\text{Var} (\hat{Y}_{ij})}{t^2 k} \left[ t + \sum_{i \neq i'}^t r (i i') \right] \quad (45)$$

$$\text{Cov} (\bar{Y}_{i.} ; \bar{Y}_{..}) = \frac{1}{t k^2} \sum_j^k \left[ \text{Var} (\hat{Y}_{ij}) \sum_{i'}^t r_j (i i') \right] \quad (46)$$

Reemplazando en (12) se tiene

$$\begin{aligned}
 \text{Var} ( \hat{Y}_{ij} ) &= \frac{1}{k^2} \sum_j^k \text{Var} ( \hat{Y}_{ij} ) + \frac{(k-2)}{t^2 k} \left[ t + \sum_{i \neq i'}^t r (i i') \right] \text{Var} ( \hat{Y}_{ij} ) \\
 &+ \frac{1}{t^2 k^2} \sum_j^k \left\{ \text{Var} ( \hat{Y}_{ij} ) \left[ t + \sum_{i \neq i'}^t r_j (i i') \right] \right\} \\
 &+ \frac{2}{t k} \text{Var} ( \hat{Y}_{ij} ) \sum_{i'}^t r (i i') \\
 &- \frac{2}{t k^2} \sum_j^k \left\{ \text{Var} ( \hat{Y}_{ij} ) \sum_{i'}^t r_j (i i') \right\} \tag{47}
 \end{aligned}$$

7.2 Suponiendo varianzas comunes a través de los estratos e igual suma de los coeficientes de correlación a través de los estratos, es decir aceptando

$$\sum_j^k \text{Var} ( \hat{Y}_{ij} ) = k \text{Var} ( \hat{Y}_{ij} ) \tag{48}$$

y

$$\sum_j^k \sum_{i \neq i'}^t r_j (i i') = k \sum_{i \neq i'}^t r (i i') \tag{49}$$

se obtiene, luego de simplificar



$$\text{Var} ( \overset{\circ}{Y}_{ij} ) = \frac{t^2 + (k-1) \left[ t + \sum_{\substack{i=1 \\ i' \neq i}}^t r(i i') \right]}{t^2 k} \text{Var} ( \hat{Y}_{ij} ) \quad (50)$$

Haciendo  $r(i i') = 0$  queda por supuesto la (21)

7.3 Para hallar el desvío medio cuadrático de  $\overset{\circ}{Y}_{ij}$  consideremos el valor esperado de  $(\overset{\circ}{Y}_{ij} - \hat{Y}_{ij})^2$ . Haciendo un desarrollo similar al establecido en punto 6.4 obtenemos la expresión (32). La covarianza de esta expresión en función del modelo aditivo está escrita en (33). Bajo los supuestos de muestras correlacionadas de período a período y de varianzas comunes a través de los estratos y a través de los períodos, obtenemos

$$\text{Cov} ( \bar{Y}_{i.} ; \hat{Y}_{ij} ) = \frac{1}{k} \text{Var} ( \hat{Y}_{ij} ) \quad (51)$$

$$\text{Cov} ( \bar{Y}_{.j} ; \hat{Y}_{ij} ) = \frac{\text{Var} ( \hat{Y}_{ij} )}{t} \left[ 1 + \sum_{\substack{i=1 \\ i' \neq i}}^t r(i i') \right] \quad (52)$$

$$\text{Cov} ( \bar{Y}_{..} ; \hat{Y}_{ij} ) = \frac{\text{Var} ( \hat{Y}_{ij} )}{t k} \left[ 1 + \sum_{\substack{i=1 \\ i' \neq i}}^t r(i i') \right] \quad (53)$$

7.4 Reemplazando en (32) y operando el álgebra

$$\text{DMC} ( \overset{\circ}{Y}_{ij} ) = E ( \overset{\circ}{Y}_{ij} - \hat{Y}_{ij} )^2 = \frac{t(k-2) - 2(1+pt)(k-1)}{t k} \text{Var} ( \hat{Y}_{ij} )$$

donde

$$p = \frac{\sum_{\substack{i=1 \\ i' \neq i}}^t r(i i')}{t} \quad (55)$$

7.5 La estimación del desvío medio cuadrático es

$$\widehat{DMC}(\overset{\circ}{Y}_{ij}) = (\overset{\circ}{Y}_{ij} - \hat{Y}_{ij})^2 - \frac{t(k-2) - 2(1+pt)(k-1)}{tk} \widehat{Var}(\hat{Y}_{ij})$$

Para muestras independientes entre periodos o sea para  $p = 0$  se obtiene la expresión (38).

Para  $k = 1$  el desvío medio cuadrático estimado de  $\overset{\circ}{Y}_{ij}$  se hace igual a la varianza estimada de  $\hat{Y}_{ij}$ .

7.7.1 La eficiencia estimada del estimador regular  $\hat{Y}_{ij}$  con respecto al estimador editivo  $\overset{\circ}{Y}_{ij}$  para la celda  $(ij)$  se puede medir con

$$\frac{\widehat{DMC}(\overset{\circ}{Y}_{ij})}{\widehat{Var}(\hat{Y}_{ij})} \quad (57)$$

7.7.2 La eficiencia estimada obtenida en el  $j$ -estrato se puede conocer con

$$\frac{\sum_i^t \widehat{DMC}(\overset{\circ}{Y}_{ij})}{\sum_i^t \widehat{Var}(\hat{Y}_{ij})} \quad (58)$$

7.7.3 Para conocer la eficiencia estimada de todo el esquema se puede usar el criterio de la media de los cocientes

$$\begin{array}{c}
 k \\
 \text{---} \\
 | \\
 \text{---} \\
 | \\
 \text{---} \\
 J
 \end{array}
 \frac{\sum_1^t \widehat{DMC} (Y_{1j}^*)}{\sum_1^t \widehat{Var} (Y_{1j}^*)}$$

(59)

o el criterio del cociente de las medias

$$\frac{\sum_J^k \sum_1^t \widehat{DMC} (Y_{1j}^*)}{\sum_J^k \sum_1^t \widehat{Var} (Y_{1j}^*)}$$

(60)

B.1 El ejemplo que a continuación se presenta supone estimaciones de la población nacida en el extranjero en 4 regiones distintas de un país a través de 3 años consecutivos. Los datos fueron obtenidos a través de investigaciones muestrales. La media de los 3 coeficientes de correlación obtenido a través de los 3 períodos se supuso aproximadamente igual a .20 y este número se usó para los cálculos en todos los estratos.

Cuadro 1. Estimación de la población nacida en el extranjero usando el estimador regular  $\hat{Y}_{1j}$  y el estimador aditivo  $\overset{\circ}{Y}_{1j}$ .

Región	AÑOS			Total
	1	2	3	
1 $\hat{Y}_{11}$	26792	28571	31326	$\bar{Y}_{.1} = 28896$
$\overset{\circ}{Y}_{11}$	27277	28580	30832	
2 $\hat{Y}_{12}$	19206	20012	20897	$\bar{Y}_{.2} = 20038$
$\overset{\circ}{Y}_{12}$	18419	19722	21974	
3 $\hat{Y}_{13}$	9586	12361	17057	$\bar{Y}_{.3} = 13009$
$\overset{\circ}{Y}_{13}$	11390	12693	14945	
4 $\hat{Y}_{14}$	2298	2131	2822	$\bar{Y}_{.4} = 2417$
$\overset{\circ}{Y}_{14}$	798	2101	4353	
Total	$\bar{Y}_{1.}$ = 14471	$\bar{Y}_{2.}$ = 15774	$\bar{Y}_{3.}$ = 18026	$\bar{Y}_{..} = 16090$

Cuadro 2. Varianzas estimadas del estimador regular  $\hat{Y}_{ij}$  y desvíos medios cuadráticos estimados del estimador aditivo  $Y_{ij}^{\circ}$ , (10)<sup>6</sup>.

Región	A Ñ O S			Total
	1	2	3	
1. $\hat{Var}(\hat{Y}_{i1})$	31	47	40	118
$\hat{DMC}(Y_{i1}^{\circ})$	9.5	14.1	12.2	35.8
2. $\hat{Var}(\hat{Y}_{i2})$	32	82	62	176
$\hat{DMC}(Y_{i2}^{\circ})$	10.2	24.7	19.8	54.7
3. $\hat{Var}(\hat{Y}_{i3})$	9	8	12	29
$\hat{DMC}(Y_{i3}^{\circ})$	5.9	2.5	8.0	16.4
4. $\hat{Var}(\hat{Y}_{i4})$	2	2	2	6
$\hat{DMC}(Y_{i4}^{\circ})$	2.8	.7	2.9	6.4

Cuadro 3. Eficiencias estimadas del estimador regular  $\hat{Y}_{ij}$  con respecto al estimador aditivo  $Y_{ij}^{\circ}$  utilizando el cociente de las medias.

Región	A Ñ O S			Eficiencia estimada
	1	2	3	
1	.31	.30	.31	.31
2	.32	.30	.32	.31
3	.66	.31	.67	.57
4	1.40	.35	1.45	1.07

8.2 La eficiencia estimada de  $\hat{Y}_{ij}$  con respecto a  $Y_{ij}^o$  de todo el esquema es .57 aplicando la media de los cocientes y es .34 aplicando el cociente de las medias.

9.1 Resumen. El estimador aditivo ha sido investigado sobre la base de mejorar los estimadores regulares que se usan en las estimaciones de cruzamiento de variables que se presentan en clasificaciones a doble entrada.

Es un estimador compuesto, definido como la suma de las medias marginales correspondientes a una celda, menos la media total del esquema.

El estimador aditivo ajusta los datos respetando los totales marginales del esquema. Es de hacer notar que los valores ajustados para una determinada celda pueden aparecer negativos. Esto ocurre cuando la suma de las medias marginales de esa celda en particular es menor a la media total del esquema. En estos casos esos valores negativos deben ser reemplazados por ceros. Este estimador aditivo resulta más eficiente que el estimador regular para una celda, cuando la varianza del estimador regular de esa celda en particular es mayor que la media de las varianzas, calculada sobre todos los estratos para ese periodo, dividida por el número de estratos.

En el ejemplo presentado se observa, Cuadro 3, que el estimador aditivo trabajó con mayor eficiencia en 10 de las 12 celdas y en 3 de las 4 regiones.

Bibliografía

1. Cochran, W.G., Técnicas de Muestreo. Primera Edición en Español. Compañía Editorial Continental, S.A. México, 1971.
2. Kleane, S.C., Introduction to Metamathematics. American Elsevier Publishing Company, Inc., New York, 1974.
3. Ranjan Kumar Som, A Manual of Sampling Techniques. Heinemann Educational Books Ltd., London, 1973.
4. The Currents Population Survey. A Report on Methodology. Technical Paper N°7. Bureau of the Census, 1963.
5. Yates, F., Sampling Methods for Censuses and Surveys. Charles Griffin & Company Limited, London, 1949.

CELADE - SISTEMA DOCPAL  
DOCUMENTACION  
SOBRE POBLACION EN  
AMERICA LATINA

