UNA VISITA A LOS ESTIMADORES DE TIPO RAZÓN EN EL MUESTREO SIMPLE ALEATORIO

Carlos N. Bouza-Herrera* y Agustín Santiago Moreno**
*Universidad de LA Habana, Cuba
**Universidad Autónoma de Guerrero

ABSTRACT
In this paper we pretend to draw a panoramic view of the basic elements of ratio estimation. In it are presented approaches for extending the classic estimator, generalizing and determining classes of ratio type estimators

KEYWORDS: ratio, extensions of classic ratio estimator, generalized ratio estimators, classes of ratio type estimators.

MSC: 62D05

RESUMEN
En este trabajo se intenta dar un panorama de los elementos básicos de la estimación usando razones. En él se presentan los enfoques usados para extender el estimador clásico, generalizarle y de determinar clases de estimadores del tipo razón.

PALABRAS CLAVE: razón, extensiones del clásico estimador, estimador de tipo razón generalizado, clases de estimadores del tipo razón.

1. UNA INTRODUCCIÓN

El estimador de razón es muy popularmente utilizado en las aplicaciones de la estadistica. Naturalmente aparecen como razones diversas tasas e índices. Tal es el caso en el desarrollo de encuestas para hacer auditorias, en el estudio de áreas pequeñas, etc. Su uso puede trazarse en el trabajo de J. Graunt en 1662 y en la encuesta desarrollada por el célebre matemático P.S. Laplace. Ambos usaron la razón entre el número de habitantes y el de nacimientos y, a la luz del desarrollo actual, la estimaron. Laplace lo aplicó en la famosa encuesta para estimar la población de Francia en 1820 en que introdujo el uso de la selección aleatoria. Vea Cochran (1977). Una formulación general de la estimación de una razón es considerar que la variable de interés Y se relaciona con una variable auxiliar X, cuyo valor es conocido para todo elemento de la población, mediante el modelo de regresión lineal simple: Y=A+BX+e. Si $(a , b)^T$ es el estimador mínimo cuadrático de $(A , B)^T$ tenemos que $b=n(n^{-1} \sum_{i=1}^n Y_i-a)/\sum_{i=1}^n X_i$ bajo el supuesto de que $\sum_{i=1}^n e_i =0$.

Por lo que si A=0 se tiene que b es la razón entre las medias muestrales de Y y X.
Al considerar que los errores no tienen la misma varianza se utiliza para estimar B la suma ponderada de los errores

$$\sum_{i=1}^n W_i(Y_i -bX_i)^2$$

determinando el estimador

$$.b_W = \sum_{i=1}^n W_i Y_iX_i./\sum_{i=1}^n W_iX_i^2$$

$$\text{Si } V(e^2_i)=X_i^{2t} \sigma^2$$

el peso adecuado es $W_i =X_i^{-2t}$ por lo que para t=0,5 y se tiene que b_W es la razón entre los totales de Y y X.

El estudio de la estimación basado en una razón utiliza el desarrollo de Series de Taylor como herramienta. El marco de referencia parte de tener en cada individuo una medición de dos variables $(Y,X) \in]0,\infty[^2$ y el interés de estudiar el desarrollo correspondiente de $\frac{Y}{X}$.
En general lo que nos interesa es trabajar con una función g(Y,X).
Su desarrollo en Series de Taylor alrededor de $\vec{\mu} = (\mu_Y, \mu_X)$ lo podemos escribir en general como

$$g(Y,X) = g(\vec{\mu}) + g'_Y(\vec{\mu})(Y-\mu_Y) + g'_X(\vec{\mu})(X-\mu_X) + \varepsilon$$

ε es el resto de la serie.
Si usáramos un desarrollo de orden 1

$$E[g(Y,X)] = g(\vec{\mu}) + g'_Y(\vec{\mu})E(Y-\mu_Y) + g'_X(\vec{\mu})E(X-\mu_X) + E(\varepsilon\ )$$

pues

$$E[g(Y,X)] = g(\vec{\mu}) + E(\varepsilon\ \ )$$
Como el orden se fija a partir de que se considera, desde el punto de vista del Análisis Numérico, que $\varepsilon \cong 0$
$$E[g(Y,X)] \cong g(\vec{\mu})$$
Si se acepta que el orden es 2, lo más común en este contexto,
$$g(Y,X) = g(\vec{\mu}) + g'_Y(\vec{\mu})E(Y-\mu_Y) + g'_X(\vec{\mu})E(X-\mu_X)$$
$$+\frac{1}{2}\left[g''_{YY}(\vec{\mu})(Y-\mu_Y)^2 + g''^2_{XX}(\vec{\mu})(X-\mu_X)^2 + 2g'_Y(\vec{\mu})g'_X(\vec{\mu})(Y-\mu_Y)(X-\mu_X) + \varepsilon\ \ \right]$$
y
$$E[g(Y,X)] = g(\vec{\mu}) + \frac{1}{2}\left[g'^2_{YY}(\vec{\mu})E(Y-\mu_Y)^2 + g'^2_{XX}(\vec{\mu})E(X-\mu_X)^2 + 2g'_Y g'_X(\vec{\mu})E(Y-\mu_Y)(X-\mu_X)\right]$$
pues, como nuestra función es $\frac{Y}{X}$
$$g'_Y(X,Y) = \frac{1}{\mu_X}, g'_X(X,Y) = -\frac{\mu_Y}{\mu_X^2}$$
$$g'^2_Y(\vec{\mu}) = \frac{1}{\mu_X^2}\ , \qquad g'^2_X(\vec{\mu}) = \frac{\mu_Y^2}{\mu_X^4}, \qquad g'_Y(\vec{\mu})g'_X(\vec{\mu}) = -\frac{\mu_Y}{\mu_X^3}$$
$$E[g(Y,X)] = g(\vec{\mu}) + \frac{Var(X)\mu_Y}{\mu_X^3} - \frac{Cov(X,Y)}{\mu_X^2}$$
Si consideramos que $\mu_Z = E(Z), Z = X, Y$ este término es el sesgo aproximado. Nótese que si asumimos la aproximación de primer orden como aceptable, tendremos que la razón es aproximadamente insesgada.
Respecto a la varianza, en el caso de la aproximación de primer orden
$$Var\big(g(Y,X)\big) \cong E[g(Y,X) - g(\vec{\mu})]^2 = E[g'_Y(\vec{\mu})(Y-\mu_Y) + g'_X(\vec{\mu})(X-\mu_X)]^2$$
$$= \frac{Var(Y)}{\mu_X^2} + \frac{\mu_Y^2 Var(X)}{\mu_X^4} - \frac{2Cov(X,Y)}{\mu_Y\ \mu_X^3}$$
En el caso de usar la aproximación de orden 2 todos los términos de orden t>2 tiene esperanza aproximadamente cero por lo que, aunque
$$Var[g(Y,X)] \cong E[g'_Y(\vec{\mu})E(Y-\mu_Y) + g'_X(\vec{\mu})E(X-\mu_X)$$
$$+\frac{1}{2}[g''_{YY}(\vec{\mu})(Y-\mu_Y)^2 + g''_{XX}(\vec{\mu})(X-\mu_X)^2 + 2g'_Y(\vec{\mu})g'_X(\vec{\mu})(Y-\mu_Y)(X-\mu_X)]]^2$$
la varianza aproximada de ambos modelos como solo tienen un valor diferente de cero las potencias de los dos primeros términos coinciden. Esto es
$$Var\big(g(Y,X)\big) \cong \frac{Var(Y)}{\mu_X^2} + \frac{\mu_Y^2 Var(X)}{\mu_X^4} - \frac{2Cov(X,Y)}{\mu_Y\ \mu_X^3} = \frac{\mu_Y^2}{\mu_X^2}\left[\frac{Var(Y)}{\mu_Y^2} + \frac{Var(X)}{\mu_X^2} - \frac{2Cov(X,Y)}{\mu_X\ \mu_Y}\right]$$
usualmente se simplifica la notación tomando los coeficientes de variación $C_Z = \frac{\sqrt{Var(Z)}}{\mu_Z}, Z = Y, X$ y se denota
$$Var\big(g(Y,X)\big) \cong \frac{\mu_Y^2}{\mu_X^2}[C_Y^2 + C_X^2 - 2\rho C_X C_Y]$$
Al hacer estimaciones las varianzas y covarianzas tomaran las expresiones correspondientes.
$$\sum_{(i,j,k)\in C_3^n} \varphi_4\big(Z_i, Z_j, Z_k\big) + o\left(n^{-\frac{3}{2}}\right), q = t, s; Q = T, S;\ \varphi = \tau, \delta$$
Asumiendo la nulidad, al menos asintótica (cs) de
$$E[\varphi_1(Z_i)] = E[\varphi_2(Z_i)] = E[\varphi_3\big(Z_i, Z_j\big)|Z_i] =^{cs} E\big[\varphi_4\big(Z_i, Z_j, Z_k\big)\big|Z_i, Z_j\big] =^{cs} 0$$
y que $E[|\ \tau_1(Z_i)|^3 + |\ \delta_1(Z_i)|^3] < \infty$.
Nótese que en el caso de la aproximación de orden 2 al considerar el error cuadrático media se debe añadirle el sesgo a cuadrado.
Una representación más general se puede obtener usando las ideas que sostienen el uso de los estadísticos del tipo U.
Considere que se trabaja con funcionales evaluados en una sucesión de variables aleatorias independientes con la misma distribución y tomemos $q_n\ y\ Q_n, q = t, s; Q = T, S$. El parámetro Q es estimado por q con un sesgo φ_0 y lo podemos representar por

$$q_n = Q_n + \frac{\varphi_0}{n} + \frac{1}{n^2}\sum_{i=1}^{n}\varphi_1(Z_i) + \frac{1}{n}\sum_{i=1}^{n}\varphi_2(Z_i) + \frac{1}{n^2}\sum_{(i,j)\in C_2^n}\varphi_3(Z_i,Z_j) + \frac{1}{n^3}$$
$$E\left[|\varphi_2(Z_i)|^{4+\varepsilon} + \left|\varphi_3(Z_i,Z_j)\right|^{4+\varepsilon} + \left|\varphi_4(Z_i,Z_j,Z_k)\right|^{4+\varepsilon}\right] < \infty$$
es posible garantizar , bajo una serie de condiciones suaves (ver Maesono, 2005) que la razón es asintóticamente un estadístico U.

$$\frac{t_n}{s_n} = U_n + o\left(n^{-\frac{3}{2}}\right)$$

por tanto su estudio se puede llevar a cabo en el marco de la estadística funcional.
En esta monografía veremos que existen clases de estimadores suficientemente amplias como para que diferentes estimadores de razón desarrollados en la literatura pertenezcan a ella como casos particulares. Esto permite un estudio unificado de ellos.

2. EL ESTIMADOR CLÁSICO

Tomemos a $U = \left(U_1, U_2, ...U_N\right)$ como una población finita y que su tamaño es $|U| = N$

Para cada $U_i, \left(i = 1,2,...,N\right)$ se tienen un par de variables: X, es una auxiliar correlacionada con Y, la de interés. En la población tenemos las medias de ellas y su razón:

$$\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} y_i, \quad \bar{X} = \frac{1}{N}\sum_{i=1}^{N} x_i\ ,\ R = \frac{\bar{Y}}{\bar{X}}\ .$$

Así como sus varianzas y la covarianza

$$S_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(y_i - \bar{Y}\right)^2, S_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(x_i - \bar{X}\right)^2,\ S_{yx} = \frac{1}{N-1}\sum_{i=1}^{N}\left(y_i - \bar{Y}\right)\left(x_i - \bar{X}\right)$$

El coeficiente de correlación es $\rho_{yx} = \rho = \frac{S_{yx}}{S_y S_x}$

También son de interés los coeficientes de variación

$$C_y = \frac{S_y}{\bar{Y}}, C_x = \frac{S_x}{\bar{X}},\ \&\ C_{yx} = \frac{S_{yx}}{\overline{YX}} = \rho_{yx} C_y C_x\ ,$$

Si se selecciona una muestra aleatoria $s \subset U, |s| = n$ si se hace y $\left(y_i, x_i\right), i = 1,2,...,n$ son observados el estimador del total d e Y es

$$\bar{y}_R = \frac{\bar{y}}{\bar{x}}\bar{X}$$

donde $\bar{z} = \frac{1}{n}\sum_{i=1}^{n} z_i,\ z = x, y$. Para ello debe ser conocida $\bar{X}$.
El estudio de este estimador parte del desarrollo en Series de Taylor de este y considerar que es válido tomar O(1/n). Para ello un método es considerar que$|e_1| < 1$ y que podemos escribir ahora

$$\bar{y} = \bar{Y}\left(1+e_0\right),\ \bar{x} = \bar{X}\left(1+e_1\right),\ E\left(e_0\right) = E\left(e_1\right) = 0$$
$$V\left(e_0\right) = \theta C_y^2, \quad V\left(e_1\right) = \theta C_x^2,\ COV(e_0 e_1) = \theta C_{yx}$$

Entonces el sesgo (B) y el error cuadrático medio (MSE) son aproximados por

$$B\left(\bar{y}_R\right) = \theta\bar{Y}\left(C_x^2 - C_{yx}\right),\ MSE\left(\bar{y}_R\right) = \theta\bar{Y}^2\left(C_y^2 + C_x^2 - 2C_{yx}\right), \theta = \left(\frac{1}{n} - \frac{1}{N}\right).$$

Un resultado básico es que $\bar{y}_R$ es más eficiente que $\bar{y}$ si $K = \rho\frac{C_y}{C_x} > \frac{1}{2}$..

Si el orden aproximación fuere O(n^{-2}) se tiene que al definir

$$C_{th} = \frac{E(\bar{y}-\bar{Y})^t(\bar{x}-\bar{X})^h}{E(\bar{y})^t E(\bar{x})^h}$$

$$B(\bar{y}_R) = \bar{Y}\frac{N-n}{N-1}\frac{1}{n}(C_{02}-C_{11}) + \bar{Y}\frac{N-n}{N-1}\frac{1}{n}\left[A\frac{1}{n}(C_{12}-C_{03}) + B\frac{1}{n^2}(C_{04}-C_{13}) + 3D\frac{1}{n^2}(C_{02}^2 - C_{02}C_{11})\right]$$

$$MSE(\bar{y}_R) = \bar{Y}^2\frac{N-n}{N-1}\frac{1}{n}\left[(C_{20}+C_{02}-2C_{11}) + \frac{2}{n}A(2C_{12}-C_{21}-C_{03})\right]$$

$$+\bar{Y}^2\frac{N-n}{N-1}\frac{1}{n}\left[\frac{3}{n^2}B(C_{04}+C_{22}-2C_{13}) + \frac{3}{n^2}D(C_{20}C_{02}+2C_{11}^2-6C_{11}C_{02}+3C_{02}^2)\right]$$

Si ambas variables fueran normales

$$MSE(\bar{y}_R) = \bar{Y}^2\frac{C_{02}}{n}\left[2(1-\rho) + \frac{C_{02}}{n}(12-18\rho+6\rho^2)\right]$$

A lo largo de los años de han propuesto varias modificaciones del llamado estimador de razón clásico. Entre ellos podemos señalar Murthy (1967), Cochran (1977), Presad (1989), Sen (1993), Singh-Tailor (2003, 2005), Singh et al. (2004), Kadilar-Cingi (2004, 2006), Koyuncu-Kadilar (2009), Alomari et al. (2009), Yan -Tian (2010) , entre otros. Estos toman una forma del tipo

$$\hat{\mu}_{\text{SRS}} = \frac{\bar{y}+\beta(\mu_x-\bar{x})}{(\alpha\bar{x}+\gamma)}(\alpha\mu_x+\gamma),\ \beta = {\sigma_{xy}}/{\sigma_x^2}$$

Por ejemplo una extensión fue propuesta por Sisodid-Dwivedi (1981) al suponer conocido el coeficiente de variación C_x de la variable auxiliar X. El estimador desarrollado fue:

$$\bar{y}_{SD} = \bar{y}\left(\frac{\bar{X}+C_x}{\bar{x}+C_x}\right)$$

Tomando $a_1 = \frac{\bar{Y}}{\bar{X}+C_x}$ el sesgo y el MSE aproximados obtenidos son

$$B(\bar{y}_{SD}) = \frac{\theta a_1}{\bar{Y}}\left(Ra_1S_x^2 - S_{yx}\right)$$

$$MSE(\bar{y}_{SD}) = \theta\left(S_y^2 + a_1S_x^2 - 2a_1S_{yx}\right)$$

Bahl- Tuteja(1991) propusieron por su parte el estimador

$$\bar{y}_{BTR} = \bar{y}\exp\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right)$$

Este es más eficiente que $\bar{y}$ si $\frac{1}{4} < K < \frac{3}{4}$ i). Por su parte Upadhyaya, et al. (2011) derivaron que si el orden de convergencia es O(n^{-2}) también es más eficiente que $\bar{y}_R$ bajo la normalidad conjunta de (y, x) , los coeficientes de variación son iguales y $0 < \rho < \frac{3}{4}$.

Incluyendo los términos de orden 4 en la aproximación se tiene que

$$\bar{y}_{BTR} - \bar{Y} = \bar{Y}(e_0 - \frac{e_1}{2} + \frac{3}{8}e_1^2 - \frac{13}{48}e_1^3 + \frac{73}{384}e_1^4 - \frac{e_0e_1}{2} + \frac{3}{8}e_0e_1^2 - \frac{13}{48}e_0e_1^3)$$

$$(\bar{y}_{BTR} - \bar{Y})^2 = \bar{Y}^2(e_0^2 + \frac{1}{4}e_1^2 - e_0e_1 + \frac{5}{4}e_0e_1^2 - e_0^2e_1 - \frac{3}{8}e_1^3 + \frac{79}{192}e_1^4 + e_0^2e_1^2 - \frac{31}{24}e_0e_1^3)$$

por lo que

$$B(\bar{y}_{BTR}) = \bar{Y}\frac{N-n}{N-1}\frac{1}{n}\left[(\frac{3}{8}C_{02} - \frac{1}{2}C_{11})\right]$$

$$+\bar{Y}\frac{N-n}{N-1}\frac{1}{n}\Big[\frac{A}{n}(\frac{3}{8}C_{12}-\frac{13}{48}C_{03})+\frac{B}{n^2}(\frac{73}{384}C_{04}-\frac{13}{48}C_{13})$$
$$+\frac{D}{n^2}(\frac{73}{128}C_{02}^2-\frac{13}{16}C_{02}C_{11})\ \Big]$$

$$MSE(\bar{y}_{BTR})=E(\bar{y}_{BTR}-\bar{Y})^2=\bar{Y}^2E(e_0^2+\tfrac{1}{4}e_1^2-e_0e_1)$$
$$+\bar{Y}^2E(\frac{5}{4}e_0e_1^2-e_0^2e_1-\frac{3}{8}e_1^3+\frac{79}{192}e_1^4+e_0^2e_1^2-\frac{31}{24}e_0e_1^3)\ =\bar{Y}^2\frac{N-n}{N-1}\frac{1}{n}\Big[(C_{20}+\frac{1}{4}C_{02}-C_{11})+\frac{1}{n}A(\frac{5}{4}C_{12}-C_{21}-\frac{3}{8}C_{03})$$
$$+$$
$$\frac{1}{n^2}B(\frac{79}{192}C_{04}+C_{22}-\frac{31}{24}C_{13})+\frac{1}{n^2}D(C_{20}C_{02}+2C_{11}^2-\frac{31}{8}C_{11}C_{02}+\frac{79}{64}C_{02}^2)\ \Big]$$

si trabajamos con una distribución normal bivariante

$$MSE(\bar{y}_{BTR})=\bar{Y}^2\frac{C_{02}}{n}\left[(\frac{5}{4}-\rho)+\frac{C_{02}}{n}(\frac{143}{64}-\frac{31}{8}\rho+2\rho^2)\right]$$

Note que

$$MSE(\bar{y}_R)-MSE(\bar{y}_{SQR})=\theta\bar{Y}^2\left(\frac{3C_{02}}{4}-C_{11}\right)$$

3. ALGUNAS CLASES DE ESTIMADORES DEL TIPO RAZÓN

Una de las primeras clases del tipo exponencial fue la que propuso Srivastava (1967). Digamos que esta es

$$S=\left\{\bar{y}_{SR}=\bar{y}\left(\frac{\bar{X}}{\bar{x}}\right)^{\alpha}|\alpha\in\Re\right\},$$

Un caso particular es el estimador

$$\bar{y}_{SQR}=\bar{y}\left(\frac{\bar{X}}{\bar{x}}\right)^{\frac{1}{2}}$$

debido a Swain (2014) cuya expansión hasta el grado cuarto de e_0 y e_1 es

$$\bar{y}_{SQR}=\bar{Y}+\bar{Y}(e_0-\frac{1}{2}e_1+\frac{3}{8}e_1^2-\frac{15}{48}e_1^3+\frac{105}{384}e_1^4-\frac{1}{2}e_0e_1+\frac{3}{8}e_0e_1^2-\frac{15}{48}e_0e_1^3)$$

por lo que

$$(\bar{y}_{SQR}-\bar{Y})^2=\bar{Y}^2[\ (e_0^2+\frac{1}{4}e_1^2-e_0e_1)\ +e_0^2e_1^2+\frac{5}{4}e_0e_1^2-e_0^2e_1-\frac{33}{24}e_0e_1^3-\frac{3}{8}e_1^3+\frac{87}{192}e_1^4\ ]$$

obteniendo para el sesgo la aproximación ,

$$B(\bar{y}_{SQR})=\bar{Y}[\ E(e_0)-\frac{1}{2}E(e_1)+\frac{3}{8}E(e_1^2)-\frac{15}{48}E(e_1^3)+\frac{105}{384}E(e_1^4)$$
$$-\frac{1}{2}E(e_0e_1)+\frac{3}{8}E(e_0e_1^2)-\frac{15}{48}E(e_0e_1^3)\ ]$$

y para el MSE

$$MSE(\bar{y}_{SQR})=E(\bar{y}_{SQR}-\bar{Y})^2=\bar{Y}^2\left[E(e_0^2)+\frac{1}{4}E(e_1^2)-E(e_0e_1)\right]$$

$$+\bar{Y}^2\left[E(e_0^2e_1^2)+\frac{5}{4}e_0e_1^2-E(e_0^2e_1)-\frac{33}{24}E(e_0e_1^3)-\frac{3}{8}E(e_1^3)+\frac{87}{192}E(e_1^4)\right]$$

Tomemos

$$C_{tu}=\frac{E(y\quad-\bar{Y})^t(x-\bar{X})^u}{\bar{Y}^t\bar{X}^u}$$

En este caso se necesita el cálculo de

$$E(e_1^4)=\frac{N-n}{n^3}\left[\frac{N^2+N-6nN+6n^2}{(N-1)(N-2)(N-3)}C_{04}+\frac{3N(N-n-1)(n-1)}{(N-1)(N-2)(N-3)}C_{02}^2\right]$$

$$E(e_0e_1^3)=\frac{N-n}{n^3}\left[\frac{N^2+N-6nN+6n^2}{(N-1)(N-2)(N-3)}C_{13}+\frac{3N(N-n-1)(n-1)}{(N-1)(N-2)(N-3)}C_{11}C_{02}\right]$$

$$E(e_0^2e_1^2)=\frac{N-n}{n^3}\left[\frac{N^2+N-6nN+6n^2}{(N-1)(N-2)(N-3)}C_{22}+\frac{N(N-n-1)(n-1)}{(N-1)(N-2)(N-3)}(C_{20}C_{02}+2C_{11}^2\right]$$

$$E(e_1^3)=\frac{(N-n)(N-2n)}{(N-1)(N-2)n^2}C_{03}$$

$$E(e_0e_1^2)=\frac{(N-n)(N-2n)}{(N-1)(N-2)n^2}C_{12}\qquad E(e_0^2e_1)=\frac{(N-n)(N-2n)}{(N-1)(N-2)n^2}C_{21}$$

Entonces se deriva que bajo esta convergencia

$$B(\bar{y}_{SQR})=\bar{Y}\frac{N-n}{N-1}\frac{1}{n}\left[(\frac{3}{8}C_{02}-\frac{1}{2}C_{11})\right]$$

$$+\bar{Y}\frac{N-n}{N-1}\frac{1}{n}\left[\frac{A}{n}(\frac{3}{8}C_{12}-\frac{15}{48}C_{03})+\frac{B}{n^2}(\frac{105}{384}C_{04}-\frac{15}{48}C_{13})\right.$$

$$\left.+\frac{3D}{n^2}(\frac{105}{384}C_{02}^2-\frac{15}{48}C_{11}C_{02})\quad\right]$$

$$MSE(\bar{y}_{SQR})=\bar{Y}^2\frac{N-n}{N-1}\frac{1}{n}\left[(C_{20}+\frac{1}{4}C_{02}-C_{11})+\frac{1}{n}A(\frac{5}{4}C_{12}-C_{21}-\frac{3}{8}C_{03})\right]$$

$$+\bar{Y}^2\frac{N-n}{N-1}\frac{1}{n}\left[\frac{1}{n^2}B(\frac{87}{192}C_{04}+C_{22}-\frac{33}{24}C_{13})+\frac{1}{n^2}D(C_{20}C_{02}+2C_{11}^2-\frac{33}{8}C_{11}C_{02}+\frac{87}{64}C_{02}^2)\right]$$

Tomando

$$A=\frac{N-2n}{N-2}\quad,B=\frac{N^2+N-6nN+6n^2}{(N-2)(N-3)}\quad,D=\frac{N(N-n-1)(n-1)}{(N-2)(N-3)}$$

Si asumimos la normalidad de X y Y el MSE se transforma en

$$MSE(\bar{y}_{SQR})=\bar{Y}^2\frac{C_{02}}{n}\left[(\frac{5}{4}-\rho)+\frac{C_{02}}{n}(\frac{151}{64}-\frac{33}{8}\rho+2\rho^2)\right]$$

Así que :

1. En el caso O(n^{-1})

$\bar{y}_{BTR}$ y $\bar{y}_{SQR}$

- Tienen la misma eficiencia
- Son más eficientes que $\bar{y}_R$ si $\frac{C_{11}}{C_{02}} < \frac{3}{4}$ o $K > \frac{1}{4}$
- Más eficientes que $\bar{y}$ si $K > \frac{1}{4}$
- Más eficientes que $\bar{y}_R$ y $\bar{y}$ si $\frac{1}{4} < K > \frac{3}{4}$
- Tienen el mismo sesgo que $\bar{y}_R$ si $\frac{5}{4} < K$ o $K < \frac{11}{12}$

2. En el caso O(n^{-2}) para N grande se obtiene otro panorama. De la comparación de los MSE`s se tiene la siguiente expresión:

$$MSE(\bar{y}_{SQR}) - MSE(\bar{y}_{BTR}) = \frac{\bar{Y}^2}{4n^2}\left(\frac{{C_{02}}^2}{2} - C_{11}C_{02}\right)$$

Se tiene que $\bar{y}_{SQR}$ es más eficiente que $\bar{y}_{BTR}$ si $\frac{{C_{02}}^2}{2} - C_{11}C_{02} < 0$, $C_{11}C_{02} > \frac{1}{2}$ o $\rho > \frac{1}{2}$ y $C_{20} = C_{02}$
Bajo la normalidad si $C_{20} = C_{02}$

$$MSE(\bar{y}_{BTR}) - MSE(\bar{y}_R) = \frac{1}{n}\bar{Y}^2 C_{02}\left[(\rho - \frac{3}{4}) - \frac{C_{02}}{n}(\frac{625}{64} - \frac{113}{8}\rho + 4\rho^2)\right]$$

$$MSE(\bar{y}_{SQR}) - MSE(\bar{y}_R) = \frac{1}{n}\bar{Y}^2 C_{02}\left[(\rho - \frac{3}{4}) - \frac{C_{02}}{n}(\frac{617}{64} - \frac{111}{8}\rho + 4\rho^2)\right]$$

De ser $\rho < \frac{3}{4}$, son negativas las diferencias $MSE(\bar{y}_{BTR}) - MSE(\bar{y}_R)$ y $MSE(\bar{y}_{SQR}) - MSE(\bar{y}_R)$. Entonces bajo el orden $O(1/n^2)$ una condición de suficiencia para que $\bar{y}_{BTR}$ y $\bar{y}_{SQR}$ sean más eficientes que $\bar{y}_R$ es que $0 < \rho < \frac{3}{4}$.

Además

- $\bar{y}_{SQR}$ es más eficiente que $\bar{y}_{BTR}$ si $\frac{1}{2} < \rho$
- $\bar{y}_{SQR}$ es mas eficiente que $\bar{y}_{BTR}$ y $\bar{y}_R$ si $\frac{1}{2} < \rho < \frac{3}{4}$

Swain (2014) elaboró una tabla para evaluar la preferencia de estos estimadores

Tabla- Comparación de Estimadores de tipo razón

$\bar{y}_{SQR}$	Preferir el estimador
0<ρ<1/4	$\bar{y}$
1/4<ρ<1/2	$\bar{y}_{SBR}$
1/2<ρ<3/4	$\bar{y}_{SQR}$
3/4<ρ<1	$\bar{y}_R$

Vishwakarma et al. (2011) revisitaron la propuesta de Bahl-Tuteja (1991) y les denotó

$$t_1 = \bar{y}\exp\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right), \quad t_2 = \bar{y}\exp\left(\frac{\bar{x}-\bar{X}}{\bar{x}+\bar{X}}\right).$$

Su propuesta fue usar las funciones lineales de estos estimadores y la media muestral

$$t_3 = \alpha\,\bar{y} + (1-\alpha)t_1$$

$$t_4 = \beta\,\bar{y} + (1-\beta)t_2$$

como es usual α y β son constantes a determinar. La minimización de los *MSE´s* prima en su búsqueda . Nótese que podemos determinar la existencia de una clase de estimadores del tipo razón que depende de estas constantes.

$$G = \{t \ \ (\vartheta) | \vartheta = \alpha, \beta, \vartheta \in \Re \ \}$$

Tenemos en esta clase a

$$\bar{y} = t(1|\vartheta = \alpha, \beta)$$

Los estimadores de Bahl.Tuteja (1991), están dados por fijar $t_1 = t(0|\vartheta = \alpha)$ y $t_2 = t(0|\vartheta = \beta)$
Como

$$\bar{y} = \bar{Y}(1+e_0), \bar{x} = \bar{X}(1+e_1)$$

$$t_3 - \bar{Y} = \bar{Y}\left[e_0 - \left(\frac{e_1}{2} - \frac{3}{8}e_1^2 + \frac{e_0 e_1}{2}\right) + \alpha\left(\frac{e_1}{2} - \frac{3}{8}e_1^2 + \frac{e_0 e_1}{2}\right)\right]$$

$$t_4 - \bar{Y} = \bar{Y}\left[e_0 + \left(\frac{e_1}{2} - \frac{e_1^2}{8} + \frac{e_0 e_1}{2}\right) - \beta\left(\frac{e_1}{2} - \frac{e_1^2}{8} + \frac{e_0 e_1}{2}\right)\right]$$

para $O(n^{-1})$

$$B(t_3) = \theta\bar{Y}C_x^2\left[(1-\alpha)\left(\frac{3}{8} - \frac{k_{yx}}{2}\right)\right], B(t_4) = \theta\bar{Y}C_x^2\left[(1-\beta)\left(\frac{k_{yx}}{2} - \frac{1}{8}\right)\right], k_{yx} = \rho\frac{C_y}{C_x}$$

y respecto a los MSE´s se derivaron

$$MSE(t_3) = \theta\bar{Y}^2\left[C_y^2 + \frac{C_x^2}{4}(1-\alpha)^2 - \rho C_y\ C_x\ (1-\alpha)\ \right]$$

$$MSE(t_4) = \theta\bar{Y}^2\left[C_y^2 + \frac{C_x^2}{4}(1-\beta)^2 - \rho C_y\ C_x\ (1-\alpha)\ \right]$$

La minimización de los MSE´s es obtenida igualmente al resolver las ecuaciones respectivas

$$\frac{\partial}{\partial\alpha}MSE(t_3) = 0, \ \frac{\partial}{\partial\beta}MSE(t_4) = 0$$

que llevan a que los valores óptimos de los parámetros sean

$$\alpha^* = 1 - 2k_{yx}, \ \beta^* = 1 + 2k_{yx}$$

Al utilizar los valores óptimos tenemos como expresión particular de los estimadores

$$t_3^* = (1-2k_{yx})\bar{y} + 2k_{yx}t_1, t_4^* = (1+2k_{yx})\bar{y} - 2k_{yx}t_2$$

Estos son equivalentes pues

$$MSE(t_3)_{min} = MSE(t_4)_{min} = \theta\bar{Y}^2[C_y^2(1-\rho^2)]$$

Al comparar estos estimadores con los restantes se tiene que

- $MSE(t_3) < V(\bar{y})$ si $\alpha > 1 - 4k_{yx}$

- $MSE(t_4) < V(\bar{y})$ si $\beta > 1 + 4k_{yx}$
- $MSE(t_3) < MSE(\bar{y}_R)$ si $\alpha < 3 - 4k_{yx}$
- $MSE(t_4) < MSE(\bar{y}_P)$ si $\beta < 3 + 4k_{yx}$
- $MSE(t_3) < MSE(t_1)$ si $\alpha < 2 - 4k_{yx}$
- $MSE(t_4) < MSE(t_2)$ si $\beta < 2 + 4k_{yx}$

Swain (2012) propuso la clase generalizada de estimadores del tipo razón

$$t_g = \bar{y}\left[\alpha\left(\frac{\bar{X}}{\bar{x}}\right)^g + (1-\alpha)\left(\frac{\bar{x}}{\bar{X}}\right)^h\right]^\delta,$$

En ella deben ser determinadas las constantes $\zeta \in \Re, \zeta)\alpha, \beta, \delta, h, g$.
Usando la aproximación debida tenemos que

$$t_g = \bar{Y}\left[1 + \delta\{h - \alpha(g+h)\}e_1 + \delta\left\{\frac{h(h-1)}{2} + \left(\frac{g(g+1)}{2} - \frac{h(h-1)}{2}\right)e_1^2\right\} + \frac{\delta(\delta-1)}{2}\right.$$

$$\left.\{h - \alpha(g+h)\}^2 e_1^2 + e_0 + \delta\{h - \alpha(g+h)\}e_0 e_1\right]$$

Su sesgo y MSE son aproximadamente

$$B(t_g) = \theta\bar{Y}\left[\delta\{h - \alpha(g+h\}C_{yx} + \delta\left\{\frac{h(h-1)}{2} + \alpha\left(\frac{g(g+1)}{2} - \frac{h(h-1)}{2}\right)\right\}C_x^2\right.$$

$$\left.+\frac{\delta(\delta-1)}{2}\{h - \alpha(g+h)\}^2 C_x^2\right]$$

$$MSE(t_g) = \theta\bar{Y}^2\left[C_y^2 + \delta^2\{h - \alpha(g+h)\}^2 C_x^2 + 2\delta\{h - \alpha(g+h)\}C_{yx}\right]$$

El Mínimo el MSE es obtenido cuando usamos

$$\alpha_{opt} = \frac{K + \delta h}{\delta(g+h)}$$

Al usarle obtenemos que

$$MSE(t_g)_{opt} = \theta\bar{Y}^2 C_y^2(1 - \rho^2)$$

$$B(t_g)_{opt} = \theta\bar{Y}C_x^2\left[-K^2 + \frac{\delta}{2}\big(h(h-1) + (K + \delta h)(g - h + 1)\big) + \frac{(\delta-1)K^2}{2\delta}\right]$$

Swain(2014) derivó aceptando que, bajo $O(1/n)$, $\bar{y}_{SQR} = \bar{y}(\frac{\bar{X}}{\bar{x}})^{1/2}$ y $\bar{y}_{BTR} = \bar{y}\exp(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}})$
poseen el mismo sesgo y que son eficientes en el mismo rango
Swain(2013) desarrolló lo que denominó clases generalizadas del tipo razón. La llamada exponencial del tipo razón fue

$$\bar{y}_{SW} = \bar{y}e^{\alpha\left(\frac{\bar{X} - \bar{x}}{\bar{X}}\right)}$$

α es también un constante a determinar por el muestrista y se obtuvo que

$$B(\bar{y}_{SW}) = \theta\bar{Y}(\frac{\alpha^2}{2}C_x^2 - \alpha C_{yx}) \quad , MSE(\bar{y}_{SW}) = \theta\bar{Y}^2(C_y^2 + \alpha^2 C_x^2 - 2\alpha C_{yx})$$

Así se tiene de este estudio que:

- $\bar{y}_{SR}$ y $\bar{y}_{SW}$ aparecen como equivalentes en términos de eficiencia.
- $\bar{y}_{SR}$ posee menor sesgo que $\bar{y}_{SW}$ cuando $K > \frac{2\alpha+1}{4}$.
- De usarse el valor óptimo de α , $\bar{y}_{SR}$ posee menor sesgo que $\bar{y}_{SW}$ cuando$K > \frac{1}{2}$.

Una clase mas amplia fue propuesta por Swain (2014) donde la clase es caracterizada por la estructura ya propuesta

$$t_g^* = \bar{y}\left[\alpha e^{g\left(\frac{\bar{X}-\bar{x}}{\bar{X}}\right)} + (1-\alpha)e^{h\left(\frac{\bar{x}-\bar{X}}{\bar{X}}\right)}\right]^{\delta}$$

α, g, h **y** δ siguen definidas como anteriormente. Este consideró que fijando . g, h and δ se puede minimizar el MSE respecto a la constante.
Veáse que la expansión de este estimador en Series de Taylor es

$$t_g^* = \bar{Y}(1+e_0)\left[\alpha\left(1 - ge_1 + \frac{g^2}{2}e_1^2 + ...\right) + (1-\alpha)\left(1 + he_1 + \frac{h^2 e_1^2}{2} + ...\right)\right]^{\delta}$$

$$= \bar{Y}\left[1 + \delta\{h - \alpha(g+h)\}e_1 + \delta\left\{\frac{h^2}{2} + \alpha\left(\frac{g^2 - h^2}{2}\right)\right\}e_1^2\right.$$

$$\left. + \frac{\delta(\delta-1)}{2}\{h - \alpha(g+h)\}^2 e_1^2 + e_0 + \delta\{h - \alpha(g+h)\}e_0 e_1\right]$$

resultando que el sesgo es aproximadamente

$$B(t_g^*) = \theta\bar{Y}\left[\delta\{h - \alpha(g+h)\}C_{yx}\right.$$

$$\left. + \delta\left\{\frac{h^2}{2} + \alpha\left(\frac{g^2 - h^2}{2}\right)\right\}C_x^2 + \frac{\delta(\delta-1)}{2}\{h - \alpha(g+h)\}^2 C_x^2\right]$$

y el MSE está dado aproximadamente por

$$MSE(t_g^*) = \theta\bar{Y}^2\left[C_y^2 + \delta^2\{h - \alpha(g+h)\}^2 C_x^2 + 2\delta\{h - \alpha(g+h)\}C_{yx}\right]$$

cuyo mínimo se obtiene para

$$\alpha_{opt} = \frac{K + \delta h}{\delta(g+h)}$$

Siendo el mínimo del MSE:

$$MSE(t_g^*)_{opt} = \theta\bar{Y}^2 C_y^2(1-\rho^2) \qquad ,$$

Este es igual el del estimador de regresión lineal

$$\bar{y}_{1r} = \bar{y} + b(\bar{X} - \bar{x})$$

Como se sabe b es el coeficiente de regresión muestral.
El sesgo obtenido es

$$B\left(t_g^*\right)_{opt} = \theta\bar{Y}C_x^2\left[-K^2 + \frac{1}{2}(K(g-h)+\delta gh)^2 + \frac{(\delta-1)K^2}{2\delta}\right]$$

Swain (2014) presenta una caracterización de varios estimadores de tipo razón que pertenecen a la clase propuesta y ahí su estudio numérico de las clases. En la tabla siguiente .

Tabla Comparación desarrollada por Swain sin considera el factor $\theta\bar{Y}$

Sl. No.	δ	g	h	α_{opt}	$B(t_g)$	$B(t_g^*)$
1	1	1	1	$\frac{K+1}{2}$	$(\frac{K+1}{2}-K^2)C_x^2$	$-(K^2-\frac{1}{2})C_x^2$
2	1	0	1	$K+1$	$-K^2C_x^2$	$-K(K+\frac{1}{2})C_x^2$
3.	1	1	0	K	$K(1-K)C_x^2$	$-K(K-\frac{1}{2})C_x^2$
4.	-1	1	1	$\frac{1-K}{2}$	$\frac{(K-1)}{2}C_x^2$	$-\frac{1}{2}C_x^2$
5.	-1	0	1	$1-K$	0	$-\frac{K}{2}C_x^2$
6.	-1	1	0	$-K$	$-KC_x^2$	$\frac{K}{2}C_x^2$

4. RAZONES Y SU DUAL

Al definirse un estimador del tipo razón podemos considerar este como un "primal" e igual que en la Programación Matemática en ocasiones es preferible trabajar con su dual.

Como tenemos toda la información sobre la variable auxiliar , consideremos que esta es positiva tomando como su mínimo y su máximo a X_m, y X_M Entonces podemos definir las variables transformadas

$$V_i = \frac{X_i + X_m}{X_M + X_m} \in]0,1]$$

$$W_i = \frac{X_i + X_M}{X_M + X_m} \in [1,2[$$

Note que V(C)≤Max(C)-Min(C)=1, C=V, W . En nuestro caso s$\rho_{x,y} = \rho_{v,y} = \rho_{w,y}$.

Mohanty-Sahoo (1995) propusieron como estimadores a

$$\bar{y}_{r(MSt)} = \frac{\bar{y}}{\bar{Z}_{est}}\bar{Z},\ Z = \begin{cases} V\ if\ t=1 \\ W\ if\ t=2 \end{cases}, \bar{Z}_{est} = \sum_{i=1}^{n}\frac{z_i}{n}$$

Srivenkataramana (1980) por su parte propuso como dual del clásico estimador de razón a

$$\bar{y}_s = \bar{y}\left(\frac{N\bar{Z} - n\bar{Z}_{est}}{(N-n)\bar{Z}}\right), Z = X$$

Bouza (2013) propuso la clase de estimadores de razón a:

$$F* = \left\{\bar{y}_\theta = \frac{\bar{y} + \alpha}{B\bar{Z}_{est} + \lambda}\left(B\bar{Y} + \lambda\right); \theta = (a, B, \lambda)^T \in A \times B \times L\right\},$$

siendo

$$A = \left\{0, \quad b(\bar{X} - \bar{x}), \quad \sigma_x\right\} = \{a_1, a_2, a_3\}$$
$$B = \{1, \quad B_2(x), \quad C_x, \quad \rho\} = \{B_1, B_2, B_3.B_4\}$$
$$L = \{0, \quad \rho, \quad B_2(x), \quad C_x\} = \{\lambda_1, \lambda_2, \lambda_3.\lambda_4\}$$

posteriormente propuso otra clase de estimadores del tipo razón

$$F = \left\{ \bar{y}_\theta = \frac{\bar{Y}_{est} + \alpha}{B\bar{Z}_{est}}(\lambda); \theta = (a, B, \lambda)^T \in A \times B \times L \right\} \tag{2.2}$$

$\bar{A}_{est}$ estima a $\bar{A}$. Se deriva la pertenencia a ella de diversos estimadores. Por ejemplo clásicos esta dado por usar Z=X a, θ=(0,1,0)=(α_1 ,B_1 ,λ_1) pues

$$\bar{y}_r = \bar{y}_{(0,1,\bar{Z})} = \frac{\bar{y}}{\bar{Z}_{est}}\bar{Z}, \qquad Z = X$$

A ella también pertenece el estimador de Mohanty-Sahoo (1995) pues si $\theta = (0,1,\bar{Z})$

$$\bar{y}_{r(MSt)} = \frac{\bar{y}}{\bar{Z}_{est}}\bar{Z}, \qquad Z = \begin{cases} V \; if \; t = 1 \\ W \; if \; t = 2 \end{cases}$$

Usando el desarrollo en Series de Taylor se deriva la aproximación del MSE pero

$$M(\bar{y}_{(0,1,\bar{Z})}) = \frac{\sigma_y^2 + R_Z^2\sigma_Z^2 w(w - 2\gamma)}{n}$$

para $\bar{y}_r$

$$w = 1$$

$$\gamma = \rho_{z,y}\frac{C_y}{C_z} = \rho_{z,y}\left(\frac{\frac{\sigma_y}{\bar{Y}}}{\frac{\sigma_z}{\bar{Z}}}\right)$$

En Bouza (2013), al anlizar F*, se obtuvo que

$$M(\bar{y}_{(0,1,\bar{Z})}) = \frac{\sigma_y^2 + R_Z^2\sigma_Z^2 - 2\rho_{z,y}R_Z\sigma_y\sigma_z}{n}$$

Haciendo Z=X, se tiene la pertenencia a la clase del el estimador dual de Srivenkataramana (1980)

$$\bar{y}_{\left(0,1,\frac{(N\bar{X} - n\bar{X}_{est})\bar{X}_{est}}{(N-n)\bar{X}}\right)} = \bar{y}_s = \frac{\bar{y}}{\bar{X}_{est}}\left(\frac{(N\bar{X} - n\bar{X}_{est})\bar{X}_{est}}{(N-n)\bar{X}}\right)$$

La propuesta de Jhajj et al. (2006) determina subclase de F determinada por

$$\bar{y}_{\left(0,1,\frac{(N\bar{Z} - n\bar{z}_{est})\bar{Z}_{est}}{(N-n)\bar{Z}}\right)t} = \bar{y}_{Dt} = \frac{\bar{y}}{\bar{Z}_{est}}\left(\frac{(N\bar{Z} - n\bar{Z}_{est})\bar{Z}_{est}}{(N-n)\bar{Z}}\right), Z = \begin{cases} V \; if \; t = 1 \\ W \; if \; t = 2 \\ X \; if \; t = 3 \end{cases}$$

El sesgo y la varianza de esta clase es aproximada por

$$B\left[\bar{y}_{\left(0,1,\frac{(N\bar{Z} - n\bar{z}_{est})\bar{Z}_{est}}{(N-n)\bar{Z}}\right)t}\right] = -\bar{Y}\rho_{z,y}C_y\left(\frac{\bar{Z}C_z}{a + \bar{Z}}\right),$$

$$Z=\begin{cases} V\ if\ t=1 \\ W\ if\ t=2 \\ X\ if\ t=3 \end{cases},\ a=\begin{cases} X_m\ if\ Z=V \\ X_M\ if\ Z=Z \\ 0\ if\ Z=X \end{cases}$$

$$V\left[\bar{y}_{\left(0,1,\frac{(N\bar{Z}-n\bar{z}_{est})\bar{Z}_{est}}{(N-n)\bar{Z}}\right)t}\right]=\frac{\bar{Y}^2}{n}\left(C_y^2+\left(\frac{n}{N-n}\right)^2\left(\frac{\bar{Z}^2C_z^2}{(a+\bar{Z})^2}\right)-2\rho_{z,y}\left(\frac{n}{N-n}\right)\left(\frac{\bar{Z}C_yC_z}{a+\bar{Z}}\right)\right),$$

$$Z=\begin{cases} V\ if\ t=1 \\ W\ if\ t=2 \\ X\ if\ t=3 \end{cases},\ a=\begin{cases} X_m\ if\ Z=V \\ X_M\ if\ Z=Z \\ 0\ if\ Z=X \end{cases}$$

entonces el MSE es aproximado por

$$M\left[\bar{y}_{\left(0,1,\frac{(N\bar{Z}-n\bar{z}_{est})\bar{Z}_{est}}{(N-n)\bar{Z}}\right)t}\right]=V\left[\bar{y}_{\left(0,1,\frac{(N\bar{Z}-n\bar{z}_{est})\bar{Z}_{est}}{(N-n)\bar{Z}}\right)t}\right]+B\left[\bar{y}_{\left(0,1,\frac{(N\bar{Z}-n\bar{z}_{est})\bar{Z}_{est}}{(N-n)\bar{Z}}\right)t}\right]^2$$

se desarrolla una comparación entre las clases y sus duales para determinar en que caso es mejor trabajar con el estimador o su dual.. Bouza determinó que era preferible los miembros de la clase tales que $\left(0,1,\frac{(N\bar{X}-n\bar{X}_{est})\bar{X}_{est}}{(N-n)\bar{X}}\right)$ a su dual evaluando

$$D=M[0,1,\bar{Z}]-M\left[\bar{y}_{\left(0,1,\frac{(N\bar{Z}-n\bar{z}_{est})\bar{Z}_{est}}{(N-n)\bar{Z}}\right)t}\right]=\left(-2\rho_{z,y}\sigma_y\sigma_z R_Z+\left(\frac{n}{N-n}\right)\left(\frac{\bar{Y}}{a+\bar{Z}}\right)\right)+\sigma_{\bar{Z}}^2\left(R_{\bar{Z}}^2-\left(\frac{n}{N-n}\right)\left(\frac{\bar{Y}^2}{(a+\bar{Z})^2}\right)\right)-\left[\rho_{z,y}\sigma_y\left(\frac{\sigma_z}{a+\bar{Z}}\right)\right]^2$$

Entonces es mejor usar W cuando la variable auxiliar tiene un valor alto de X_M, pues

$$D<R_{\bar{Z}}^2-\rho_{ZY}\sigma_y\left(2R_Z+\frac{\rho_{z,y}\sigma_y}{(a+\bar{Z}\ \ )^2}\right)+\left(\frac{n}{N-n}\right)\frac{\bar{Y}}{a+\bar{Z}}\left(1-\frac{\bar{Y}}{(a+\bar{Z})}\right)<0$$

5. **CLASES DE ESTIMADORES GENERALIZADOS DE RAZÓN.**

La literatura es plagada de papers donde se habla del desarrollo de un clase de estimadores generalizados de razón. Cada autor al hacer una generalización utiliza frases similares. Esto crea a veces confusión al pensarse que todos hacen lo mismo. Hay muchas formas de hacer una generalización. Lo que en general se hace es fijar una clase dependiente de parámetros a fijar. Tal es el caso de

$$F=\left\{\bar{y}_\theta=\frac{\bar{Y}_{est}+\alpha}{B\bar{X}_{est}+\lambda}\left(B\bar{X}+\lambda\right);\theta=(\alpha,B,\lambda)^T\in A\times B\times L\right\}$$

donde $\bar{Z}_{est}$, Z=X, Y, es la media muestral. Los parámetros que la caracterizan son tales que

$$A=\left\{0,\qquad b(\bar{X}-\bar{x}),\qquad \sigma_x\right\}=\{\alpha_1,\alpha_2,\alpha_3\}$$
$$B=\{1,\qquad B_2(x),\qquad C_x,\qquad \rho\}=\{B_1,B_2,B_3.B_4\}.$$
$$L=\{0,\qquad \rho,\qquad B_2(x),\qquad C_x\}=\{\lambda_1,\lambda_2,\lambda_3.\lambda_4\}$$

Tomando

- $b=s_{xy}/s_x^2$ como el coeficiente de regresión lineal simple muestral,
- $B_2(x)$ como el coeficiente de curtosis de la distribución de X,
- C_x como el coeficiente de variación de X
- ρ como el coeficiente de correlación lineal entre X y Y.

A esta clase pertenece el estimador clásico en el que $\theta=(0,1,0)=(\alpha_1, B_1, \lambda_1)$ pues

$$\bar{y}_r = \bar{y}_{(0,1,0)} = \frac{\bar{y}}{\bar{x}}\bar{X}$$

El error cuadrático medio de la subclase de *F* a la que pertenece este es caracterizado por

$$M(\bar{y}_{razn}) = \frac{\sigma_y^2 + R^2\sigma_x^2 w(w-2\gamma)}{n}$$

En el caso del estimador clásico

$$w = 1$$

$$\gamma = \rho\frac{C_y}{C_x} = \rho\frac{\frac{\sigma_y}{\bar{Y}}}{\frac{\sigma_x}{\bar{X}}}$$

y se deduce con facilidad que

$$M(\bar{y}_n) = \frac{\sigma_y^2 + R^2\sigma_x^2 - 2\rho R\sigma_y\sigma_x}{n}$$

Singh-Taylor (2003) desarrollaron un estimador alternativo que utilizaba información poblacional sobre el coeficiente de correlación ρ. Este es de la misma familia pues al fijar que θ=(0, 1, ρ), se obtiene:

$$\bar{y}_{ST} = \frac{\bar{y}}{\bar{x}+\rho}(\bar{X}+\rho)$$

Su error está dado por

$$M(\bar{y}_{ST}) = \left(\bar{y}_{razon}\middle| w = \frac{\bar{X}}{\bar{X}+\rho}, \qquad \gamma = 2\rho C_y / C_x\right) = \frac{\sigma_y^2 + R^2\sigma_x^2\left(\frac{\bar{X}}{\bar{X}+\rho}\right)^2 - 2\frac{\bar{Y}\rho\sigma_y\sigma_x}{\bar{X}+\rho}}{n}$$

La comparación de los errores permite hacer una ordenación. Así evaluamos cuando

$M(\bar{y}_n) - M(\bar{y}_{ST}) < 0$. Una condición suficiente es que $\rho<R\sigma_x (w+1)/\sigma_y$

Los estimadores de Kadilar-Cingi {Kadilar-Cingi , 2003, 2004 y 2005} son clasificables en *F* al denotarles en función de θ mediante $\bar{y}\ (\theta_t)$ donde

- $\theta_1 =(\alpha_2 , B_1 , \lambda_1)$
- $\theta_2 =(\alpha_2 , B_1 , \lambda_3)$
- $\theta_3 =(\alpha_2 , B_1 , \lambda_4)$
- $\theta_4 =(\alpha_2 , B_2 , \lambda_1)$
- $\theta_5 =(\alpha_2 , B_3, \lambda_3)$
- $\theta_6 =(\alpha_2 , B_1 , \lambda_2)$
- $\theta_7 =(\alpha_2 , B_3\ \lambda_2)$
- $\theta_8 =(\alpha_2 , B_4 ,\lambda_4)$
- $\theta_9 =(\alpha_2 , B_2 ,\lambda_2)$
- $\theta_{10} =(\alpha_2 , B_4, \lambda_3)$

En forma explícita estos están dados por

$$\bar{y}(\theta_1) = \frac{\bar{y} + b(\bar{X} - \bar{x})}{\bar{x}} \bar{X}, \qquad \bar{y}(\theta_2) = \frac{\bar{y} + b(\bar{X} - \bar{x})}{\bar{x} + B_2(x)} (\bar{X} + B_2(x))$$

$$\bar{y}(\theta_3) = \frac{\bar{y} + b(\bar{X} - \bar{x})}{\bar{x} + C_x} (\bar{X} + C_x), \qquad \bar{y}(\theta_4) = \frac{\bar{y} + b(\bar{X} - \bar{x})}{\bar{x} B_2(x) + C_x} (\bar{X}\, B_2(x) + C_x)$$

$$\bar{y}(\theta_5) = \frac{\bar{y} + b(\bar{X} - \bar{x})}{\bar{x} C_x + B_2(x)} (\bar{X}\, C_x + B_2(x)), \qquad \bar{y}(\theta_6) = \frac{\bar{y} + b(\bar{X} - \bar{x})}{\bar{x} + \rho} (\bar{X} + \rho)$$

$$\bar{y}(\theta_7) = \frac{\bar{y} + b(\bar{X} - \bar{x})}{\bar{x} C_x + \rho} (\bar{X}\, C_x + \rho), \qquad \bar{y}(\theta_8) = \frac{\bar{y} + b(\bar{X} - \bar{x})}{\bar{x}\, \rho + C_x} (\bar{X}\, \rho + C_x)$$

$$\bar{y}(\theta_9) = \frac{\bar{y} + b(\bar{X} - \bar{x})}{\bar{x} B_2(x) + \rho} (\bar{X}\, B_2(x) + \rho), \qquad \bar{y}(\theta_{10}) = \frac{\bar{y} + b(\bar{X} - \bar{x})}{\bar{x}\, \rho + B_2(x)} (\bar{X}\rho + B_2(x))$$

Así que ellos determinan una subclase en la que el error es:

$$M(\bar{y}(\theta_t)) = \frac{R^2(\theta_t)\sigma_x^2 + \sigma_y^2(1 - \rho^2)}{n}, \qquad t = 1,\ldots,10$$

Las razones indexadas por θ_t llevan a que:

$$R(\theta_1) = R = \frac{\bar{Y}}{\bar{X}}, \qquad R(\theta_2) = \frac{\bar{Y}}{\bar{X} + B_2(x)}, \qquad R(\theta_3) = \frac{\bar{Y}}{\bar{X} + C_x}, \qquad R(\theta_4) = \frac{\bar{Y}\, B_2(x)}{\bar{X}\, B_2(x) + C_x}$$

$$R(\theta_5) = \frac{\bar{Y}\, C_x}{\bar{X}\, C_x + B_2(x)}, \qquad R(\theta_6) = \frac{\bar{Y}}{\bar{X} + \rho}, \qquad R(\theta_7) = \frac{\bar{Y}\, C_x}{\bar{X}\, C_x + \rho}, \qquad R(\theta_8) = \frac{\bar{Y}\, \rho}{\bar{X}\, \rho + C_x},$$

$$R(\theta_9) = \frac{\bar{Y}\, B_2(x)}{\bar{X}\, B_2(x) + \rho}, \qquad R(\theta_{10}) = \frac{\bar{Y}\, B_2(x)}{\bar{X}\, B_2(x) + \rho},$$

Por lo que se prefieren los estimadores de la subclase Kadilar-Cingi cuando

$$\rho > \frac{\sigma_x^2\left(R^2 w^2 - R(\theta_t)\right)}{\sigma_y\left(\sigma_y + 2\sigma_x\right)}$$

El valor absoluto del término de la derecha es con mucha frecuencia mayor que 1 . Respecto a la clase preferimos los estimadores anteriores cuando

$$\rho > \frac{\sigma_x^2\left(R^2(\theta_t) - R^2\right)}{\sigma_y\left(\sigma_y + 2R\sigma_x\right)}$$

con lo que podríamos hacer una valoración similar.

En el volumen creciente de papers en este tema se enmarca la mayor parte de las investigaciones actuales.

Acknowledgments : Uno de los autores agradece el suporte de los Projectos "A CUBAN-FLEMISH TRAINING AND RE- SEARCH PROGRAM IN DATA SCIENCE AND BIG DATA ANALYSIS" and RED AGRO-BIGDATA Y "DECISION SUPPORT SYSTEMS" (DSS) PARA UN SECTOR AGROPECUARIO SOSTENIBLE BIGDSSAGRO que permitieron el desarrrollo de la investigacion durante estancias de investigación.

RECEIVED: MARCH, 2019.
REVISED: DECEMBER, 2019.

REFERENCIAS

[1] AL-OMARI, A.I., A. A. JEMAIN and I. KAMARULZAMAN (2009): New ratio estimators of the mean using simple random sampling and ranked set sampling methods. **Revista Investigación Operacional** . 30, 97-108.

[2] BAHL,S. and TUTEJA,R.K.(1991): Ratio and product type exponential estimator. **Information and Optimization Sciences**,12,159-175.

[3] BOUZA, C. (2016): Una clase de estimadores basados en una razón: muestreo simple aleatorio y muestreo por conjuntos ordenados. **Revista Investigacion Operacional**, 37, 1, 59-70.

[4] CHAUDHURI, A. and J. MITRA (1996): Setting confidence intervals by ratio estimator. **Comm. In Stat. Theory and Methods** , 25, 1135-1148.

[5] COCHRAN, W. G. (1977): **Sampling Techniques**. Wiley and Sons, N. York.

[6] COCHRAN , W. G. (1978): Laplace's ratio estimator. En **"Contributions to Survey Sampling and Applied Statistics"**, H.A. David (editor). Academic Press, N. York.

[7] HOEFFDING, W. (1961). The strong law of large numbers for U-statistics, Univ. of North Carolina Institute of statistics. Mimeo Series. No. 302.

[8] JEELANI, M.I., MAQBOOL, S., MIR, S.A., KHAN, I., NAZIR, N. and JEELANI, F. (2013). Modified ratio estimators of population mean using linear combination of co-efficient of kurtosis and quartile deviation. **International Journal of Modern Mathematical Sciences**, 8, 149-153.

[9] JHAJJ, H. S. M M. K. SHARMA and L. K. GROVER (2006): Dual of ratio estimators of finite population mean obtained on using linear transformation to auxiliary variable. **J. Japan Statist. Soc**. 36, 107–119.

[10] KADILAR, C. and H. ÇINGI (2003): A study on the chain ratio-type estimator. **Hacettepe Journal of Mathematics and Statistics,** 32, 105-108

[11] KADILAR, C. and H. ÇINGI (2004): Ratio estimators in simple random sampling. **Applied Math. And Computation**, 151, 893-902.

[12] KADILAR, C. and H. ÇINGI (2004): Ratio estimators in simple random sampling, Applied Mathematics and Computation 151, 893-902.

[13] KADILAR, C. and H. ÇINGI (2005): A new ratio estimator in stratified random sampling. Comm. in Stat:. Theory and Methods, 34, 597-602.

[14] KOYUNCU, N. and KADILAR, C. (2009). Efficient Estimators for the Population mean. **Hacettepe Journal of Mathematics and Statistics**, 38, 217-225.

[15] MAESONO, Y. (1998). An asymptotic representation of a ratio of two statistics and its applications. **Communications in Statistics-Theory and Methods**, A27, 305–327.

[16] MAESONO, Y. (2005): Asymptotic representation of ratio statistics and their mean squared errors . **J. Japan Statist. Soc.** 35, 73–97

[17] MOHANTY, S. and SAHOO, J. (1995):. A note on improving the ratio method of estimators through linear transformation using certain known population parameters, **Sankhya B,** 57, 93–102.

[18] PRASAD, B. (1989): Some improved ratio type estimators of population mean and ratio in finite population sample surveys. **Communications in Statistics: Theory and M ethods**, 18, 379–392.

[19] RAO, T. J. (1991): On certain methods of improving ratio and regression estimators. **Communications in Statistics-Theory and Methods** 20, 3325–3340.

[20] SINGH , R., P. CHAUHAN, N. SAWAN and F. SMARANDACHE (2008): Ratio Estimators in Simple Random Sampling Using Information on Auxiliary Attribute. **Pakistan Journal of Statistics and Operation Research** 4, 47-53.

[21] SINGH, H.P. and TAILOR, R. (2003). Use of known correlation coefficient in estimating the finite population means. **Statistics in Transition** 6 , 555-560.

[22] SINGH, H.P. TAILOR, R. and KAKRAN, M.S. (2004). An Improved Estimator of population mean using Power transformation. **Journal of the Indian Society of Agricultural Statistics** 58, 223-230.

[23] SISODIA, B. V. S. and DWIVEDI, V. K. (1981): A modified ratio estimator using coefficient of variation of auxiliary variable. **Journal of the Indian Society of Agricultural Statistics** 33, 13-18.

[24] SWAIN, A.K.P.C.(2003): **Finite Population Sampling: Theory and Methods**. South Asia Publishers, New Delhi.

[25] SWAIN, A.K.P.C.((2012): On Classes of modified ratio type and regression-cum-Ratio type estimators in sample surveys using two auxiliary variables. **Statistics in transition – New series**, 13, 473-494.

[26] SWAIN, A.K.P.C.((2014): On improved ratio type estimator of finite population mean in sample surveys. **Revista Investigación Operacional**, 35, 49-57.

[27] SRIVASTAVA, S.K. (1967): An estimator using auxiliary information in sample surveys. **Cal. Stat. Assoc. Bull.**, 16, 121-132.

[28] SRIVASTAVA, S.K.(1971): A generalized estimator for the mean of the finite population using multi-auxiliary information. **J. Amer. Stat.Assoc**.,66,404-407.

[29] SRIVENKATARAMANA, T. (1980). A dual to ratio estimator in sample surveys, **Biometrika,** 67, 199–204.

[30] SOHAIL, M. U., J. SHABBIR and S. AHMED (2018): A class of ratio type estimators for imputing the missing values under rank set sampling . **Journal of Statistical Theory and Practice**. DOI: 10.1080/15598608.2018.1460886

[31] UPADHYAYA, L.N., SINGH, H.P., CHATTERJEE, S. and YADAV, R.(2011): Improved Ratio and Product Exponential Type Estimator. **Journ. Stat.Theo. and Pract.** 5, 285-293.

[32] VISHWAKARMA, G. , R. SINGH, P.C. GUPTA and S. PAREEK (2016): Improved ratio and product type estimators of finite population mean in simple random sampling. **Revista Investigacion Operacional** , 36, 70-76.

[33] YAN, Z. and TIAN, B. (2010). Ratio Method to the Mean Estimation Using Coefficient of Skewness of Auxiliary Variable. **ICICA 2010, Part II, CCIS 106**, 103–110.