Ratio estimators of population mean using auxiliary information in median ranked set sampling

Nursel Koyuncu*
Hacettepe University, Ankara, Turkey – nkoyuncu@hacettepe.edu.tr

Abstract

In this study we have proposed some ratio estimators in median ranked set sampling design. We have derived mean square error (MSE) formula of proposed estimators. We have compared our estimators with the classical ratio estimator in simple random sampling. Also to see the performance of our proposed estimators we have conducted a simulation study. We have found that proposed estimators are highly efficient than existing estimator in the literature.

Keywords: ratio estimator**;** median ranked set sampling; efficiency; mean square error.

1.Simple random sampling design

Let $U=(\mathrm{U}_1,\mathrm{U}_2,\ldots,\mathrm{U}_N)$ be finite population of size N units. The variate of interest y and the auxiliary variate x assume high correlated variables on the unit U_i .Let (X_1,Y_1), (X_2,Y_2), …, (X_n,Y_n) be a bivariate random sample with pdf $f(x,y)$, means μ_x , μ_y variances σ_x^2 , σ_y^2 and correlation coefficient ρ_{xy} . From the population U a simple random sample o size n is drawn without replacement. Then the classical ratio estimator and MSE are given by respectively

$$\hat{\mu}_R=\frac{\bar{y}}{\bar{x}}\mu_x \tag{1.1}$$

$$MSE(\hat{\mu}_R)=\frac{1-f}{n}\left(\sigma_y^2+R^2\sigma_x^2-2R\sigma_{yx}\right) \tag{1.2}$$

$\bar{x}=\frac{\sum_{i=1}^{n}x_i}{n}$, $\bar{y}=\frac{\sum_{i=1}^{n}y_i}{n}$, $\sigma_y^2=\frac{\sum_{i=1}^{N}(y_i-\bar{Y})^2}{N-1}$, $\sigma_x^2=\frac{\sum_{i=1}^{N}(x_i-\bar{X})^2}{N-1}$, $\sigma_{xy}=\frac{\sum_{i=1}^{N}(y_i-\bar{Y})(x_i-\bar{X})}{N-1}$ and $f=\frac{n}{N}$.

2. Median ranked set sampling design

Al-Omari (2012) introduced median ranked set sampling as in the following steps:

1. Select n random samples each of size n bivariate units from the population of interest.
2. The units within each sample are ranked by visual inspection or any other cost free method with respect to a variable of interest.
3. If n is odd, select the $\left(\frac{n+1}{2}\right)$th-smallest ranked unit X together with the associated Y from each set, i.e., the median of each set. If n is even, from the first $\frac{n}{2}$ sets select the $\left(\frac{n}{2}\right)$ th ranked unit X

together with the associated Y and from the other $\frac{n}{2}$ sets select the $\left(\frac{n+2}{2}\right)$ th ranked unit X together with the associated Y.

4. The whole process can be repeated m times if needed to obtain a sample of size nm units.

Let $\left(X_{i(1)},Y_{i[1]}\right),\left(X_{i(2)},Y_{i[2]}\right),\ldots,\left(X_{i(n)},Y_{i[n]}\right)$ be the order statistics of $X_{i1},X_{i2},\ldots,X_{in}$ and the judgement order of $Y_{i1},Y_{i2},\ldots,Y_{in}$ $(i=1,2,\ldots,n)$, where $(\ )$ and $[\ ]$ indicate that the ranking of X is perfect and ranking of Y has errors. For odd and even sample sizes the units measured using MRSS are denoted by MRSSO and MRSSE, respectively.

For odd sample size let $\left(X_{1\left(\frac{n+1}{2}\right)},Y_{1\left[\frac{n+1}{2}\right]}\right),\left(X_{2\left(\frac{n+1}{2}\right)},Y_{2\left[\frac{n+1}{2}\right]}\right),\ldots,\left(X_{n\left(\frac{n+1}{2}\right)},Y_{n\left[\frac{n+1}{2}\right]}\right)$ denote the observed units by MRSSO. $\overline{x}_{MRSSO}=\frac{1}{n}\sum_{i=1}^{n}X_{i\left(\frac{n+1}{2}\right)}$ and $\overline{y}_{MRSSO}=\frac{1}{n}\sum_{i=1}^{n}Y_{i\left[\frac{n+1}{2}\right]}$ be the sample mean of X and Y respectively.

For even sample size let $\left(X_{1\left(\frac{n}{2}\right)},Y_{1\left[\frac{n}{2}\right]}\right)$, $\left(X_{2\left(\frac{n}{2}\right)},Y_{2\left[\frac{n}{2}\right]}\right),\ldots,\left(X_{\frac{n}{2}\left(\frac{n}{2}\right)},Y_{\frac{n}{2}\left[\frac{n}{2}\right]}\right),\left(X_{\frac{n+2}{2}\left(\frac{n+2}{2}\right)},Y_{\frac{n+2}{2}\left[\frac{n+2}{2}\right]}\right)$ $\left(X_{\frac{n+4}{2}\left(\frac{n+2}{2}\right)},Y_{\frac{n+4}{2}\left[\frac{n+2}{2}\right]}\right)$ $,\ldots,$ $\left(X_{n\left(\frac{n+2}{2}\right)},Y_{n\left[\frac{n+2}{2}\right]}\right)$ denote the observed units by MRSSE. $\overline{x}_{MRSSE}=\frac{1}{n}\left(\sum_{i=1}^{\frac{n}{2}}X_{i\left(\frac{n}{2}\right)}+\sum_{i=\frac{n+2}{2}}^{n}X_{i\left(\frac{n+2}{2}\right)}\right)$ and $\overline{y}_{MRSSE}=\frac{1}{n}\left(\sum_{i=1}^{\frac{n}{2}}Y_{i\left[\frac{n}{2}\right]}+\sum_{i=\frac{n+2}{2}}^{n}Y_{i\left[\frac{n+2}{2}\right]}\right)$ be the sample mean of X and Y respectively.

3. Suggested estimators in median ranked set sampling

In this section, following Upadhyaya and Sing(1999) and Al-Omari (2012) and we have proposed ratio estimators of population mean using auxiliary information as follows:

$$\hat{\mu}_{N1}=\overline{y}_{MRSS(j)}\left(\frac{\mu_x+C_x}{\overline{x}_{MRSS(j)}+C_x}\right), \tag{3.1}$$

$$\hat{\mu}_{N2}=\overline{y}_{MRSS(j)}\left(\frac{\mu_x+\beta_2(x)}{\overline{x}_{MRSS(j)}+\beta_2(x)}\right), \tag{3.2}$$

$$\hat{\mu}_{N3}=\overline{y}_{MRSS(j)}\left(\frac{\mu_x C_x+\beta_2(x)}{\overline{x}_{MRSS(j)}C_x+\beta_2(x)}\right), \tag{3.3}$$

$$\hat{\mu}_{N4}=\overline{y}_{MRSS(j)}\left(\frac{\mu_x\beta_2(x)+C_x}{\overline{x}_{MRSS(j)}\beta_2(x)+C_x}\right) \tag{3.4}$$

where $j=(E,O)$ denote the sample size even or odd; coefficient of variation C_x and kurtosis $\beta_2(x)$. For odd and even sample sizes to the first degree of approximation the *MSE* of $\hat{\mu}_{Nk}$ $(k=1,2,3,4)$ are respectively given by

$$MSE\left(\hat{\mu}_{N1(O)}\right)\cong\frac{1}{n}\left(\frac{\mu_y^2}{(\mu_x+C_x)^2}\sigma^2_{x\left(\frac{n+1}{2}\right)}+\sigma^2_{y\left[\frac{n+1}{2}\right]}-2\frac{\mu_y}{(\mu_x+C_x)}\sigma_{xy\left[\frac{n+1}{2}\right]}\right)$$

$$MSE\left(\hat{\mu}_{N1(E)}\right)\cong\frac{1}{2n}\left(\frac{\mu_y^2}{(\mu_x+C_x)^2}\left(\sigma^2_{x\left(\frac{n}{2}\right)}+\sigma^2_{x\left(\frac{n+2}{2}\right)}\right)+\left(\sigma^2_{y\left[\frac{n}{2}\right]}+\sigma^2_{y\left[\frac{n+2}{2}\right]}\right)-2\frac{\mu_y}{\mu_x+C_x}\left(\sigma_{yx\left[\frac{n}{2}\right]}+\sigma_{yx\left[\frac{n+2}{2}\right]}\right)\right)$$

$$MSE\left(\hat{\mu}_{N2(O)}\right)\cong\frac{1}{n}\left(\frac{\mu_y^2}{(\mu_x+\beta_2(x))^2}\sigma^2_{x\left(\frac{n+1}{2}\right)}+\sigma^2_{y\left[\frac{n+1}{2}\right]}-2\frac{\mu_y}{(\mu_x+\beta_2(x))}\sigma_{xy\left[\frac{n+1}{2}\right]}\right)$$

$$MSE\left(\hat{\mu}_{N2(E)}\right)\cong\frac{1}{2n}\left(\frac{\mu_y^2}{(\mu_x+\beta_2(x))^2}\left(\sigma^2_{x\left(\frac{n}{2}\right)}+\sigma^2_{x\left(\frac{n+2}{2}\right)}\right)+\left(\sigma^2_{y\left[\frac{n}{2}\right]}+\sigma^2_{y\left[\frac{n+2}{2}\right]}\right)-2\frac{\mu_y}{\mu_x+\beta_2(x)}\left(\sigma_{yx\left[\frac{n}{2}\right]}+\sigma_{yx\left[\frac{n+2}{2}\right]}\right)\right)$$

$$MSE\left(\hat{\mu}_{N3(O)}\right)\cong\frac{1}{n}\left(\frac{\mu_y^2C_x^2}{(\mu_xC_x+\beta_2(x))^2}\sigma^2_{x\left(\frac{n+1}{2}\right)}+\sigma^2_{y\left[\frac{n+1}{2}\right]}-2\frac{\mu_yC_x}{(\mu_xC_x+\beta_2(x))}\sigma_{xy\left[\frac{n+1}{2}\right]}\right)$$

$$MSE\left(\hat{\mu}_{N3(E)}\right)\cong\frac{1}{2n}\left(\frac{\mu_y^2C_x^2}{(\mu_xC_x+\beta_2(x))^2}\left(\sigma^2_{x\left(\frac{n}{2}\right)}+\sigma^2_{x\left(\frac{n+2}{2}\right)}\right)+\left(\sigma^2_{y\left[\frac{n}{2}\right]}+\sigma^2_{y\left[\frac{n+2}{2}\right]}\right)-2\frac{\mu_yC_x}{\mu_xC_x+\beta_2(x)}\left(\sigma_{yx\left[\frac{n}{2}\right]}+\sigma_{yx\left[\frac{n+2}{2}\right]}\right)\right)$$

$$MSE\left(\hat{\mu}_{N4(O)}\right)\cong\frac{1}{n}\left(\frac{\mu_y^2\beta_2(x)^2}{(\mu_x\beta_2(x)+C_x)^2}\sigma^2_{x\left(\frac{n+1}{2}\right)}+\sigma^2_{y\left[\frac{n+1}{2}\right]}-2\frac{\mu_y\beta_2(x)}{(\mu_x\beta_2(x)+C_x)}\sigma_{xy\left[\frac{n+1}{2}\right]}\right)$$

$$MSE\left(\hat{\mu}_{N4(E)}\right)\cong\frac{1}{2n}\left(\frac{\mu_y^2\beta_2(x)^2}{(\mu_x\beta_2(x)+C_x)^2}\left(\sigma^2_{x\left(\frac{n}{2}\right)}+\sigma^2_{x\left(\frac{n+2}{2}\right)}\right)+\left(\sigma^2_{y\left[\frac{n}{2}\right]}+\sigma^2_{y\left[\frac{n+2}{2}\right]}\right)-2\frac{\mu_y\beta_2(x)}{\mu_x\beta_2(x)+C_x}\left(\sigma_{yx\left[\frac{n}{2}\right]}+\sigma_{yx\left[\frac{n+2}{2}\right]}\right)\right)$$

4. Simulation study

In this section, we conducted a simulation study to investigate the properties of proposed estimators. In the simulation study, we consider finite populations of size $N=10000$ generated from a bivariate normal distribution $N\left(\mu_x,\mu_y,\sigma_x^2,\sigma_y^2,\rho_{xy}\right)$. The samples were generated from a bivariate normal distribution using mvrnorm function in R programme. In the simulation, we considered $\mu_x=2$, $\mu_y=4$, $\sigma_x^2=\sigma_y^2=1$ and different values of ρ_{xy}. We have computed mean square errors (MSEs) and percent relative efficiencies (PREs) of estimators with respect to $\hat{\mu}_R$ for $n=2,3,4,5,6,7$ on the basis of 1000 replications and displayed in Table1 and Table2.

5. Conclusions

From the Table1 and Table2 we can see that our proposed estimators are higly efficient than classical ratio estimator.

References

Al-Omari, A. I. (2012) Ratio estimation of the popu

Koyuncu, N. & Kadilar, C. (2009) Family of estimators of population mean using two auxiliary variables in stratified random sampling, Communications in Statistics: Theory and Methods 38, 14, 2398-2417.

Koyuncu, N. (2015), New difference-cum-ratio and exponential type estimators in median ranked set sampling, Hacettepe Journal of Mathematics and Statistics (accepted)

Upadhyaya, L.N. & Singh, H.P. (1999), Use of transformed auxiliary variable in estimating the finite population mean, Biometrical Journal, 41, 5, 627-636.

Table1: Mean Square Error (MSE) and the Percent Relative Efficiency (PRE) of estimators with respect to $\hat{\mu}_R$ for $n=2,4,6$

		MSE			**PRE**		
Estimator		$n=2$	$n=4$	$n=6$	$n=2$	$n=4$	$n=6$
$\hat{\mu}_R$	$\rho_{yx}=0.7$	3.09	0.98	0.48	100.00	100.00	100.00
$\hat{\mu}_{N1}$		0.40	0.13	0.07	762.57	765.85	688.08
$\hat{\mu}_{N2}$		0.73	0.19	0.10	422.40	524.51	504.18
$\hat{\mu}_{N3}$		0.63	0.17	0.09	493.19	578.58	547.08
$\hat{\mu}_{N4}$		**0.13***	**0.06***	**0.04***	**2306.25***	**1542.63***	**1169.90***
$\hat{\mu}_R$	$\rho_{yx}=0.9$	2.09	0.63	0.32	100.00	100.00	100.00
$\hat{\mu}_{N1}$		0.22	0.06	0.03	955.31	979.30	970.21
$\hat{\mu}_{N2}$		0.47	0.11	0.05	450.14	586.60	613.28
$\hat{\mu}_{N3}$		0.37	0.09	0.05	571.12	688.54	709.22
$\hat{\mu}_{N4}$		**0.07***	**0.03***	**0.02***	**2988.62***	**2169.85***	**1759.97***
***represent most efficent estimator**							

Table 2: : Mean Square Error (MSE) and the Percent Relative Efficiency (PRE) of estimators with respect to $\hat{\mu}_R$ for $n=3,5,7$

		MSE			**PRE**		
Estimator		$n=3$	$n=5$	$n=7$	$n=3$	$n=5$	$n=7$
$\hat{\mu}_R$	$\rho_{yx}=0.7$	1.63	0.72	0.40	100.00	100.00	100.00
$\hat{\mu}_{N1}$		0.32	0.15	0.10	510.48	488.04	409.40
$\hat{\mu}_{N2}$		0.47	0.19	0.12	347.91	378.20	336.00
$\hat{\mu}_{N3}$		0.42	0.18	0.11	385.72	405.06	354.55
$\hat{\mu}_{N4}$		**0.17***	**0.10***	**0.07***	**961.17***	**730.19***	**554.70***
$\hat{\mu}_R$	$\rho_{yx}=0.9$	1.08	0.44	0.26	100.00	100.00	100.00
$\hat{\mu}_{N1}$		0.15	0.07	0.04	707.39	634.52	637.25
$\hat{\mu}_{N2}$		0.25	0.10	0.06	423.70	438.17	466.32
$\hat{\mu}_{N3}$		0.22	0.09	0.05	499.95	494.33	517.61
$\hat{\mu}_{N4}$		**0.08***	**0.04***	**0.03***	**1429.56***	**1051.54***	**831.49***