An Innovative Additive Mathematical Model Using Auxiliary Information

Tanveer A. Tarray[1*], Javid Gani Dar[2], Ishfaq S. Ahmad[3]

[1,2,3]Department of Mathematical Science, Islamic University of Science and Technology, Jammu and Kashmir, India

Emails: tanveerstat@gmail.com ; javinfo.stat@yahoo.co.in ; peerishfaq007@gmail.com

Corresponding Author: tanveerstat@gmail.com

Abstract

This article proposes innovative ratio and regression estimators based on additive randomized response model. Expressions for the biases and mean squared errors of the recommended estimators are derived. It has been revealed that the advised groundbreaking ratio and regression estimators are improved than ratio and regression estimators under a very realistic condition. Numerical illustrations and simulation study are also given in support of the present study.

Keywords: Estimation; Mean Square error; Bias; Auxiliary variable; RRM.

AMS Subject Classification: 62D05.

1 Introduction

Captivating the inklings from different authors like [1-17], authors propose the unbiased estimator

$$\hat{\mu}_{(1)} = (1/n)\sum_{i=1}^{n} z_{ai} \qquad (1)$$

with

$$MSE(\hat{\mu}_{(1)}) = V(\hat{\mu}_{(1)}) = \frac{(1-f)}{n}(\sigma_y^2 + w\sigma_s^2), \qquad (2)$$

$$\sigma_y^2 = \frac{1}{(N-1)}\sum_{i=1}^{N}(y_i-\mu_y)^2 \text{ and } ,\sigma_s^2 = \frac{1}{(N-1)}\sum_{i=1}^{N}(s_i-\mu_s)^2$$

are respectively the variance.

Further we note that,

$E(Z_s) = E(Y\text{-}ST) = E(Y) = \mu_y$, since $E(S) = 0$.

Thus an unbiased estimator based on subtractive optional randomized response model Z_s = Y-ST is given by

$$\hat{\mu}_{(2)} = (1/n)\sum_{i=1}^{n} z_{si} \qquad (3)$$

whose variance is

DOI: https://doi.org/10.54216/AJBOR.060201
Received: January 12, 2022 Accepted: March 25, 2022

$$MSE(\hat{\mu}_{(2)}) = V(\hat{\mu}_{(2)}) = \frac{(1-f)}{n}(\sigma_y^2 + w\sigma_s^2), \quad (4)$$

From [9] and [2]

$$Z = Y + S, \quad (5)$$

an unbiased estimator

$$\hat{\mu}_{(3)} = (1/n)\sum_{i=1}^{n} z_i \quad (6)$$

with

$$MSE(\hat{\mu}_{(3)}) = V(\hat{\mu}_{(3)}) = \frac{(1-f)}{n}(\sigma_y^2 + \sigma_s^2). \quad (7)$$

From (2), (4) and (7), we have

$$MSE(\hat{\mu}_{(3)}) - MSE(\hat{\mu}_{(i),i=1,2}) = \frac{(1-f)}{n}\sigma_s^2(1-w), \quad (8)$$

since $0 \le w \le 1$, i.e.

$$MSE(\hat{\mu}_{(i),i=1,2}) < MSE(\hat{\mu}_{(3)}). \quad (9)$$

2. Improved Ratio and Regression Estimation of Mean

let $\bar{y} = (1/n)\sum_{i=1}^{n} y_i$, $\bar{x} = (1/n)\sum_{i=1}^{n} x_i$, $\bar{z}^* = (1/n)\sum_{i=1}^{n} z_i^*$ be the sample means and $\mu_y = E(Y)$, $\mu_x = E(X), \mu_{z^*} = E(Z^*)$ be the population mean we define

$$\bar{z}^* = \mu_{z^*}(1 + e_{z^*}), \ \bar{x} = \mu_x(1 + e_x),$$

with

$$E(e_{z^*}) = E(e_x) = 0$$

and

$$E(e_{z^*}^2) = \frac{(1-f)}{n}C_{z^*}^2, \ E(e_x^2) = \frac{(1-f)}{n}C_x^2, \ E(e_x e_z) = \frac{(1-f)}{n}\rho_{xz^*} C_x C_{z^*},$$

where

$$C_{z^*}^2 = \frac{\sigma_{z^*}^2}{\mu_{z^*}^2} = \frac{\sigma_y^2 + w\phi^2\sigma_s^2}{\mu_y^2} = (C_y^2 + \phi^2 w \frac{\sigma_s^2}{\mu_y^2}) = C_y^2(1 + \phi^2 r), \quad C_x^2 = \frac{\sigma_x^2}{\mu_x^2},$$

1.
$$\sigma_{xz^*} = \sigma_{xy}, \rho_{xz^*} = \frac{\rho_{xy}}{\sqrt{(1 + w\phi^2 r^2)}}, \rho_{xy} = \frac{\sigma_{yx}}{\sigma_y \sigma_x}, \ \sigma_x^2 = \frac{1}{(N-1)}\sum_{i=1}^{N}(x_i - \mu_x)^2,$$

$$\sigma_y^2 = \frac{1}{(N-1)}\sum_{i=1}^{N}(y_i - \mu_y)^2, r = \sigma_s^2/\sigma_y^2, \text{ and } \sigma_{yx} = \frac{1}{(N-1)}\sum_{i=1}^{N}(y_i - \mu_y)^2(x_i - \mu_x)^2.$$

We foremost advise a difference estimator as

DOI: https://doi.org/10.54216/AJBOR.060201
Received: January 12, 2022 Accepted: March 25, 2022

$$t_d = \bar{z}^* + d(\mu_x - \bar{x}), \quad (10)$$

with

$$E(t_d) = \mu_{z^*} = \mu_y,$$

and

$$MSE(t_d) = V(\bar{z}^*) + d^2V(\bar{x}) - 2dCov(\bar{z}^*, \bar{x})$$

$$= \frac{(1-f)}{n}[\sigma^2_{z^*} + d^2\sigma^2_x - 2d\sigma_{xz^*}] \quad (11)$$

minimized for

$$d = \frac{\sigma_{z^*x}}{\sigma^2_x} = \beta_{z^*x} = \frac{\sigma_{xy}}{\sigma^2_x} = \beta_{yx}, \quad (12)$$

where

$$\sigma_{z^*x} = \frac{1}{(N-1)}\sum_{i=1}^{N}(z^*_i - \mu_{z^*})(x_i - \mu_x).$$

Substitution of (12) in (10), $\bar{Y}$ may be

$$t^* = \bar{z}^* + \beta_{z^*x}(\mu_x - \bar{x}) \quad . \quad (13)$$

with

$$MSE(t^*) = \frac{(1-f)}{n}\left[\sigma^2_{z^*} - \frac{\sigma^2_{z^*x}}{\sigma^2_x}\right] \quad (14)$$

$$= \frac{(1-f)}{n}\sigma^2_{z^*}\left[1 - \rho^2_{z^*x}\right]$$

2. $$= \frac{(1-f)}{n}\left[\sigma^2_{z^*} - \rho^2_{z^*x}\sigma^2_{z^*}\right]$$

$$= \frac{(1-f)}{n}\left[\sigma^2_y + \phi^2\sigma^2_s w - \rho^2_{yx}\sigma^2_y\right]$$

$$= \frac{(1-f)}{n}\sigma^2_y\left[1 + \phi^2 wr^2 - \rho^2_{yx}\right]$$

and

$$\hat{\beta}_{z^*x} = \frac{s_{z^*x}}{s^2_x}, \quad (15)$$

the resulting μ_y as

$$t^*_{lr} = \bar{z}^* + \hat{\beta}_{xz^*}(\mu_x - \bar{x}), \quad (16)$$

where $s_{z^*x} = \frac{1}{(n-1)}\sum_{i=1}^{n}(z^*_i - \bar{z}^*)(x_i - \bar{x})$ and $s^2_x = \frac{1}{(n-1)}\sum_{i=1}^{n}(x_i - \bar{x})^2$.

From [4], we have

$$Bias\ (t^*_{lr}) = -\frac{(1-f)}{n}\beta_{z^*x}\left[\frac{\mu_{12}}{\mu_{11}} - \frac{\mu_{03}}{\mu_{02}}\right] \quad (17)$$

and

$$MSE(t^*_{lr}) = \frac{(1-f)}{n}\sigma^2_{z^*}(1 - \rho^2_{z^*x})$$

$$= \frac{(1-f)}{n}(\sigma^2_y + \phi^2 w\sigma^2_s - \rho^2_{yx}\sigma^2_y)$$

DOI: https://doi.org/10.54216/AJBOR.060201
Received: January 12, 2022 Accepted: March 25, 2022

$$=\frac{(1-f)}{n}\sigma_y^2\{(1+\phi^2 wr^2)-\rho_{yx}^2)\} \quad (18)$$

and μ_y as

$$\hat{\mu}_{Regw}=\bar{z}_a+\hat{\beta}_{z_a x}(\mu_x-\bar{x}), \quad (19)$$

with $\beta_{z_a x}=\frac{\sigma_{z_a x}}{\sigma_x^2}=\frac{\sigma_{yx}}{\sigma_x^2}=\rho_{yx}\frac{\sigma_y}{\sigma_x}=\beta_{yx}$. (20)

having

$$\text{Bias }(\hat{\mu}_{Regw})=-\frac{(1-f)}{n}\beta_{z_a x}\left[\frac{\mu_{12}^*}{\mu_{11}^*}-\frac{\mu_{03}^*}{\mu_{02}^*}\right] \quad (21)$$

and

$$MSE(\hat{\mu}_{Regw})=\frac{(1-f)}{n}C_y^2\{1+wr^2-\rho_{yx}^2\}, \quad (22)$$

where $\mu_{pq}^*=\frac{1}{(N-1)}\sum_{i=1}^{N}(z_{ai}-\mu_{z_a})^p(x_i-\mu_x)^q$, (p,q) being non – negative integers.

Suggested the regression as

$$\hat{\mu}_{Reg}=\bar{z}+\hat{\beta}_{zx}(\mu_x-\bar{x}), \quad (23)$$

with

$$\text{Bias }(\hat{\mu}_{Reg})=-\frac{(1-f)}{n}\beta_{zx}\left[\frac{\mu_{12}^{**}}{\mu_{11}^{**}}-\frac{\mu_{03}^{**}}{\mu_{02}^{**}}\right] \quad (24)$$

and

$$MSE(\hat{\mu}_{Reg})=\frac{(1-f)}{n}\sigma_y^2\{1+r^2-\rho_{yx}^2\}, \quad (25)$$

where $\mu_{pq}^{**}=\frac{1}{(N-1)}\sum_{i=1}^{N}(z_i-\mu_z)^p(x_i-\mu_x)^q$, (p,q) being non – negative integers.

3 Efficiency Comparisons

Our proposed regression estimator t_{lr}^* with the estimators $\hat{\mu}_{(i),i=1to4}, t_{1(1)}, t_{(1)}, t_{(2)}, \hat{\mu}_{Regw}$ and $\hat{\mu}_{Reg}$.

From (2), (4) and (18), we have

$$MSE(\hat{\mu}_{(i),i=1,2})-MSE(t_{lr}^*)=\frac{(1-f)}{n}\sigma_y^2\left[wr^2(1-\phi^2)+\rho_{yx}^2\right]$$

which is always positive if

$$(1-\phi^2)>0$$

i.e. if

$$|\phi|<1 \quad (26)$$

From (7) and (18), we have

$$MSE(\hat{\mu}_{(3)})-MSE(t_{lr}^*)=\frac{(1-f)}{n}\sigma_y^2\left[r^2(1-\phi^2 w)+\rho_{yx}^2\right]$$

which is always positive if

DOI: https://doi.org/10.54216/AJBOR.060201
Received: January 12, 2022 Accepted: March 25, 2022

$$(1-w\phi^2)>0$$

i.e. if

$$|\phi|<(1/w)^{1/2} \quad (27)$$

and

$$MSE(\hat{\mu}_{(4)})-MSE(t^*_{lr})=\frac{(1-f)}{n}\sigma_y^2\rho_{yx}^2>0 \quad (28)$$

and

$$MSE(t_{1(1)})-MSE(t^*_{lr})=\frac{(1-f)}{n}[R^2\sigma_x^2-2R\rho_{yx}\sigma_y\sigma_x+(1+wr^2)\sigma_y^2$$
$$-(1+\phi^2wr^2)\sigma_y^2+\rho_{yx}^2\sigma_y^2]$$
$$=\frac{(1-f)}{n}[(R\sigma_x-\rho_{yx}\sigma_y)^2+wr^2\sigma_y^2(1-\phi^2)]$$

>0 if $|\phi|<1$ (29)

and

$$MSE(t_{(2)})-MSE(t^*_{lr})=\frac{(1-f)}{n}[\sigma_y^2(1+r^2)+R^2\sigma_x^2-2R\rho_{yx}\sigma_y\sigma_x$$
$$-\sigma_y^2(1+wr^2)+\rho_{yx}^2\sigma_y^2]$$
$$=\frac{(1-f)}{n}[r^2(1-w\phi^2)+(R\sigma_x-\rho_{yx}\sigma_y)^2]$$

>0 if $|\phi|<(1/w)^{1/2}$ (30)

and

$$MSE(t_{(1)})-MSE(t^*_{lr})=\frac{(1-f)}{n}[R^2\sigma_x^2-2R\rho_{yx}\sigma_y\sigma_x+(1+\phi^2wr^2)\sigma_y^2$$
$$-(1+\phi^2wr^2)\sigma_y^2+\rho_{yx}^2\sigma_y^2]$$
$$=\frac{(1-f)}{n}[(R\sigma_x-\rho_{yx}\sigma_y)^2]>0 \text{ provides } R\neq\beta \quad (31)$$

and

$$MSE(\hat{\mu}_{Regw})-MSE(t^*_{lr})=\frac{(1-f)}{n}[1+wr^2-\rho_{yx}^2-1-\phi^2wr^2+\rho_{yx}^2]\sigma_y^2$$
$$=\frac{(1-f)}{n}\sigma_y^2(1-\phi^2)wr^2$$

>0 if $|\phi|$ <1 (32)

and

$$MSE(\hat{\mu}_{Reg})-MSE(t^*_{lr})=\frac{(1-f)}{n}[1+r^2-\rho_{yx}^2-1-\phi^2wr^2+\rho_{yx}^2]\sigma_y^2$$
$$=\frac{(1-f)}{n}\sigma_y^2(1+\phi^2w)r^2$$

DOI: https://doi.org/10.54216/AJBOR.060201
Received: January 12, 2022 Accepted: March 25, 2022

$$> 0 \text{ if } |\phi| < \left(1/w\right)^{1/2} \tag{33}$$

4. Numerical efficiency comparison

To magistrate the qualities we have

$$PRE(t_{lr}^*, \hat{\mu}_{(3)}) = \frac{MSE(\hat{\mu}_{(3)})}{MSE(t_{lr}^*)} \times 100$$

$$= \frac{(\sigma_y^2 + \sigma_s^2)}{[\sigma_y^2 + \phi^2 w \sigma_s^2 - \sigma_y^2 \rho_{yx}^2]} \times 100$$

and findings are displayed in Table 1.

Table 1. PRE's of $\hat{\mu}_{(lr^*)}$ with respect to $\hat{\mu}_{(3)}$ with $\rho_{xy} = 0.80, \mu_x = 6, \mu_y = 6, \sigma_x = 2, \sigma_y = 3$ and $\sigma_s = 3$

ϕ \ w	± 1.0 $PRE(t_{Regw}, \hat{\mu}_{(3)})$	± 0.9	± 0.8	± 0.5	± 0.4
0.1	434.78	453.51	471.70	519.48	531.91
0.2	357.14	383.14	409.84	487.80	510.20
0.3	303.03	331.67	362.32	459.77	490.20
0.4	263.16	292.40	324.68	434.78	471.70
0.5	232.56	261	294	412	454.55

DOI: https://doi.org/10.54216/AJBOR.060201
Received: January 12, 2022 Accepted: March 25, 2022

			.44	.12	.37	
	0.6	208.33	236.41	268.82	392.16	438.60
5.	0.7	188.68	215.75	247.52	373.83	423.73
	0.8	172.41	198.41	229.36	357.14	409.84
	0.9	158.73	183.65	213.68	341.88	396.83
	1.0	$PRE(t_{Reg}, \hat{\mu}_{(\beta)})$ 147.06	170.94	200.00	327.87	384.62

Conclusion

This paper premeditated the problem based on additive randomized response model. Properties of the proposed model have been studied along with recommendations. Efficiency comparison is worked out to investigate the performance of the suggested procedures. It is interesting to mention that the proposed procedure is superior to the one recently envisaged estimator.

Funding Statement: The author received no specific funding for this study.

Conflicts of Interest: The author declare that they have no conflicts of interest to report regarding the present study.

References

[1] Eichhorn B.H. and Hayre .LS. (1983): Scrambled randomized response methods for obtaining sensitive quantitative dada. Jour. Statist. Plann. Inf. 7,307-316.

[2] Himmelfarb S. and Edgell S.E. (1980): Additive constant model: a randomized response technique for eliminating evasiveness to quantitative response questions. Psychol. Bull. 87, 525-530.

[3] Pandey B.N. (1979): On shrinkage estimation of normal population variance. Commun. Statist. Theo. and Metho., 8, 359-365.

DOI: https://doi.org/10.54216/AJBOR.060201
Received: January 12, 2022 Accepted: March 25, 2022

[4] Panday and Singh (1977): Estimation of variance of normal population using prior information. Jour. Ind. Statist. Assoc.,15,141-150.

[5] Pollock K.H. and Bek Y. (1976): A comparison of three randomized response models for quantitative data. Jour. Amer. Statist. Assoc., 71, 884-886.

[6] Singh H. P. and Tarray T. A. (2012): A Stratified Unknown repeated trials in randomized response sampling. Common. Korean Statist. Soc., 19, (6), 751-759.

[7] Singh H.P. and Tarray T.A. (2013): An alternative to Kim and Warde's mixed randomized response model. Statist. Oper. Res. Trans.(SORT), 37 (2), 189-210.

[8] Singh H.P. and Tarray T.A. (2014): An Improvement Over Kim and Elam Stratified Unrelated Question Randomized Response Model Using Neyman Allocation. Sankhya – B, The Ind. Jour. Statist.,DOI 10.1007/s13571-014-0088-5.

[9] Singh H.P., Jong M. Kim and Tarray T.A. (2015): A family of estimators of population variance in two occasion rotation patterns. Comm. Statist. Theo. Meth.., 45(14) 4106 - 4116 .

[10] Singh HP and Tarray TA (2017): A stratified unrelated question randomized response model using Neyman allocation. Comm. Statist. Theo.- Metho. **DOI:** 10.1080/03610926.2014.983612.

[11] Tarray T.A. and Singh H.P. (2015): A general procedure for estimating the mean of a sensitive variable using auxiliary information. Investigacion Operacionel. 36(3),249-262.

[12] Tarray T.A. and Singh H.P. (2015): A randomized response model for estimating a rare sensitive attribute in stratified sampling using Poisson distribution. Model Assist. Statist. Appl., 10, 345-360.

[13] Tarray TA and Singh HP (2016): New procedures of estimating proportion and sensitivity using randomized response in a dichotomous finite population. Jour. Mod. Appli. Statist. Meth. 15(1), 635-669.

[14] Tarray TA (2017): Scrutinize on Stratified Randomized Response Technique, Munich, GRIN Verlag, ISBN: 9783668554580.

[15] Tarray T.A. and Singh H.P. (2018): Missing data in clinical trials: stratified Singh and Grewal's randomized response model using geometric distribution. Trends in Bioinformatics, 11(1),44-55.

[16] Tarray T.A., Sigh H.P. and Saadia Masood (2019): An Endowed Randomized Response Model for Estimating a Rare Sensitive Attribute Using Poisson Distribution. Trends in Applied Sciences Research, 12, 1-6.

DOI: https://doi.org/10.54216/AJBOR.060201
Received: January 12, 2022 Accepted: March 25, 2022