**Science Journal of Applied Mathematics and Statistics**
2014; 2(6): 107-111
Published online November 18, 2014 (http://www.sciencepublishinggroup.com/j/sjams)
doi: 10.11648/j.sjams.20140206.11

# An alternative estimator for estimating the finite population mean in presence of measurement errors with the view to financial modelling

**Rajesh Singh[1], Sachin Malik[1], Mohd Khoshnevisan[2]**

[1]Department of Statistics, BHU, Varanasi, India
[2]Associate Professor of Finance, Ajman University of Science and Technology, Ajman, UAE

**Email address:**
rsinghstat@gmail.com (R. Singh), sachinkurava999@gmail.com (S. Malik), m.khoshnevisan@ajman.ac.ae (M. Khoshnevisan)

**To cite this article:**
Rajesh Singh, Sachin Malik, Mohd Khoshnevisan. An Alternative Estimator for Estimating the Finite Population Mean in Presence of Measurement Errors with the View to Financial Modelling. *Science Journal of Applied Mathematics and Statistics.* Vol. 2, No. 6, 2014, pp. 107-111. doi: 10.11648/j.sjams.20140206.11

**Abstract:** This article presents the problem of estimating the population mean using auxiliary information in the presence of measurement errors. We have compared the three proposed estimators being the exponential ratio-type estimator, Solanki et al. (2012) estimator, and the mean per unit estimator in the presence of measurement errors. Financial Model by Gujrati and Sangeetha (2007) has been employed in our empirical analysis. In that, our investigation has indicated that our proposed general class of estimator $t_4$ is the most suitable estimator with a smaller MSE relative to other estimators under measurement errors.

**Keywords:** Population Mean, Study Variate, Auxiliary Variates, Mean Squared Error, Measurement Errors, Efficiency, Financial Model

## 1. Introduction

In survey sampling, the properties of the estimators based on data usually presuppose that the observations are the correct measurements on characteristics being studied. Unfortunately, this ideal is not met in practice for a variety of reasons, such as non response errors, reporting errors, and computing errors, and sensitivity errors. When the measurement errors are negligible small, the statistical inferences based on observed data continue to remain valid. On the contrary, when they are not appreciably small and negligible, the inferences may not be simply invalid and inaccurate but may often lead to unexpected, undesirable and unfortunate consequences ( Shalabh, 2001). Some authors including Allen et al.(2003), Manisha and Singh (2001, 2002), Shalabh (1997), Bahl, S. and Tuteja, R. K. (1991), Koyuncu, N. and Kadilar, C. (2010),Singh and Karpe (2008, 2009), Kumar et al. (2011a,b) and Singh et al. (2011) have paid their attention towards the estimation of population mean $\mu_y$ of the study variable y using auxiliary information in the presence of measurement errors.

For a simple random sampling scheme, let ($x_i$, $y_i$) be observed values instead of the true values ($X_i$, $Y_i$) on two characteristics (x, y) respectively for the $i^{th}$ (i=1.2….n) unit in the sample of size n.

Let the measurement errors be

$$u_i = y_i - Y_i \tag{1.1}$$

$$v_i = x_i - X_i \tag{1.2}$$

which are stochastic in nature with mean zero and variances $\sigma_u^2$ and $\sigma_v^2$ respectively, and are independent. Further, let the population means of (x, y) be ( $\mu_x$ , $\mu_y$ ), population variances of (x, y) be ( $\sigma_x^2$ , $\sigma_y^2$ ) and $\sigma_{xy}$ and $\rho$ be the population covariance and the population correlation coefficient between x and y respectively (see Manisha and Singh (2002)).

Let

$$k_1 = \bar{y} - \mu_y = \frac{1}{\sqrt{n}}\left(w_y - w_u\right)$$

and,

$$k_2 = \bar{x} - \mu_x = \frac{1}{\sqrt{n}}\left(w_x - w_v\right), \mathrm{var}\left(\bar{y}\right) = \frac{\sigma_y^2}{n}\left[1 + \frac{\sigma_u^2}{\sigma_y^2}\right]$$

and

$$E(k_1)=E(k_2)=0$$

$$E\left(k_1^2\right)=\frac{\sigma_y^2}{n}\left(1+\frac{\sigma_u^2}{\sigma_y^2}\right)=V_{ym}$$

$$E\left(k_2^2\right)=\frac{\sigma_x^2}{n}\left(1+\frac{\sigma_v^2}{\sigma_x^2}\right)=V_{xm}$$

$$E(k_1k_2)=\frac{\rho\sigma_y\sigma_x}{n}=V_{yxm}$$

In this paper, we have studied the behaviour of some estimators in presence of measurement error.

## 2. Estimators in Literature

Singh et al. (2011) suggested an exponential ratio type and a difference type estimator under measurement error for estimating $\bar{y}$ as

$$t_1=\bar{y}\exp\left(\frac{\mu_x-\bar{x}}{\mu_x+\bar{x}}\right) \tag{2.1}$$

$$t_2=\omega_1\bar{y}+\omega_2(\mu_x-\bar{x}) \tag{2.2}$$

The biases and MSE's of the estimators are respectively given by

$$Bias(t_1)=\frac{1}{\mu_x}\left(\frac{3}{8}R_mV_{xm}-\frac{1}{2}V_{yxm}\right) \tag{2.3}$$

$$Bias(t_2)=\mu_y(\omega_1-1) \tag{2.4}$$

$$\mathrm{MSE}(\mathrm{t}_1)=\frac{\sigma_y^2}{\mathrm{n}}\left[1-\frac{C_x}{C_y}\left(\rho-\frac{C_x}{4C_y}\right)\right]+\frac{1}{n}\left[\frac{\mu_y^2}{4\mu_x^2}\sigma_v^2+\sigma_u^2\right] \tag{2.5}$$

$$MSE(t_2)=(\omega_1-1)^2\mu_y^2+\omega_1^2(a_1)+\omega_2^2a_2+2\omega_1\omega_2(-a_3) \tag{2.6}$$

where,

$$a_1=(V_{ym}),\ a_2=(V_{xm}),\ \text{and}\ a_3=(V_{yxm}).$$

Now, optimising MSE of the estimator $t_2$ with respect to $\omega_1$ and $\omega_2$, we get

$$\omega_1^*=\frac{b_3b_4}{b_1b_3-b_2^2}\quad and\quad \omega_2^*=-\frac{b_2b_4}{b_1b_3-b_2^2} \tag{2.7}$$

where,

$$b_1=\mu_y^2+a_1,\ b_2=-a_3,\ b_3=a_2\ \text{and}\ b_4=\mu_y^2.$$

Using theses optimum values of $\omega_1^*$ and $\omega_2^*$ from equation (2.7) into equation (2.6), we get the minimum MSE of the estimator $t_2$ as

$$\mathrm{MSE}(t_2)_{\min}=\left[\mu_y^2-\frac{b_3b_4^2}{b_1b_3-b_2^2}\right] \tag{2.8}$$

## 3. Proposed Estimators

Solanki et al. (2012) estimator under measurement error is given by

$$t_3=\bar{y}\left\{2\text{-}\left(\frac{\bar{x}}{\mu_x}\right)^{\alpha}\exp\left[\frac{\beta(\bar{x}-\mu_x)}{(\bar{x}+\mu_x)}\right]\right\} \tag{3.1}$$

where α and β are suitably chosen scalars.

Let

$$w_u=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}u_i,\ w_y=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(y_i-\mu_y)$$

$$w_v=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}v_i,\ w_x=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(x_i-\mu_x)$$

$$C_x=\frac{\sigma_x}{\mu_x}\ \text{and}\ C_y=\frac{\sigma_y}{\mu_y}$$

Expanding equation (3.1) and subtracting $\mu_y$ from both sides, we get

$$(t_3-\mu_y)=(\mu_y+k_1)\left[1-\frac{k_2}{\mu_x}\left(\alpha+\frac{\beta}{2}\right)-\frac{k_2^2}{\mu_x^2}\left(\alpha(\alpha-1)+\frac{\beta(\beta-2)}{8}+\frac{\alpha\beta}{2}\right)\right] \tag{3.2}$$

Taking expectation of both sides of (3.2), we get the bias of the estimator $t_3$ to the order $O(n^{-1})$ as

$$Bias(t_3)=\mu_y\left\{\frac{V_{xm}A}{\mu_x^2}\right\}-\left\{\frac{B}{\mu_x}V_{xm}\right\} \tag{3.3}$$

where,

$$A=\left[\alpha(\alpha-1)+\frac{\beta(\beta-2)}{8}+\frac{\alpha\beta}{2}\right],\ \text{and}\ \mathrm{B}=\left(\alpha+\frac{\beta}{2}\right).$$

Squaring both sides of (3.2) and taking expectations, the

MSE of $t_3$ to the order $O(n^{-1})$ is given by

$$MSE(t_3) = E(t_3 - \mu_y)^2 = V_{yxm} + V_{xm}R_m^2B^2 - 2R_mV_{yxm}B \quad (3.4)$$

where $R_m = \frac{\mu_y}{\mu_x}$.

Following Solanki et al. (2012), we propose a general class of estimator $t_4$ as

$$t_4 = \left[m_1\bar{y} + m_2(\mu_x - \bar{x})\right]\left\{2 - \left(\frac{\bar{x}}{\mu_x}\right)^{\alpha} \exp\left[\frac{\beta(\bar{x}-\mu_x)}{(\bar{x}+\mu_x)}\right]\right\} \quad (3.5)$$

Expanding equation (3.4) and subtracting $\mu_y$ from both sides, we get

$$(t_4 - \mu_y) = \left\{(m_1 - 1)\mu_y - m_1\mu_y\left\{B\frac{k_2}{\mu_x} + \frac{k_2^2A}{\mu_x^2}\right\} + m_1k_1\left\{1 - \frac{Bk_2}{\mu_x}\right\} - m_2k_2\left\{1 - \frac{Bk_2}{\mu_x}\right\}\right\} \quad (3.6)$$

On taking expectation of both sides of (3.5), we get the bias of the estimator $t_4$ to the order $O(n^{-1})$ as

$$Bias(t_4) = (m_1 - 1)\mu_y - m_1\mu_y\left\{\frac{V_{xm}A}{\mu_x^2}\right\} - m_1\left\{\frac{B}{\mu_x}V_{yxm}\right\} + m_2\left\{\frac{B}{\mu_x}V_{xm}\right\} \quad (3.7)$$

Squaring both sides of (3.5) and taking expectations, the MSE of $t_4$ to the order $O(n^{-1})$ is

$$MSE(t_4) = E(t_4 - \mu_y)^2 = (m_1 - 1)^2\mu_y^2 + m_1^2R_m^2B^2V_{xm} + m_1^2V_{yxm} - 2m_1^2R_mBV_{yxm} - 2m_1(m_1 - 1)R_m^2AV_{xm} + m_2^2V_{xm}$$

$$+2\left\{m_2(m_1 - 1)BR_mV_{xm} + m_1m_2R_mV_{xm} - m_1m_2V_{yxm} + m_1(m_1 - 1)BR_mV_{yxm}\right\} \quad (3.8)$$

The MSE of the estimator $t_4$ can also be written as

$$MSE(t_4) = (m_1 - 1)^2\mu_y^2 + m_1^2P_1 + m_2^2P_2 + 2m_1m_2P_3 - 2m_1P_3 - 2mPA_5 \quad (3.9)$$

where,

$$P_1 = \left(V_{ym} + B^2R_m^2V_{xm} - 2R_m^2AV_{xm}\right),\ P_2 = (V_{xm}),$$

$$P_3 = (2BR_mV_{xm} - V_{yxm}),\ P_4 = (R_mBV_{yxm} - AV_{xm}R_m^2),$$

$$P_5 = (BR_mV_{xm}).$$

Now, optimising MSE $t_4$ with respect to, $m_1$ and $m_2$, we get the optimum values as -

$$m_1^* = \frac{B_1P_2 - P_3P_5}{B_2P_2 - P_3^2} \quad and \quad m_2^* = \frac{B_2P_5 - B_1P_3}{B_2P_2 - P_3^2}$$

where,

$$B_1 = \mu_y^2 + P_4, \qquad B_2 = \mu_y^2 + P_1.$$

# 4. Theoretical Efficiency Comparisons

The MSE of the proposed estimator $t_4$ proposed in (3.4) will be smaller than usual estimator under measurement error case if the following condition is satisfied by the data set

$$\frac{\sigma_y^2}{n}\left[1 - \frac{C_x}{C_y}\left(\rho - \frac{C_x}{4C_y}\right)\right] + \frac{1}{n}\left[\frac{\mu_y^2}{4\mu_x^2}\sigma_v^2 + \sigma_u^2\right] \le \frac{\sigma_y^2}{n}\left(1 + \frac{\sigma_u^2}{\sigma_y^2}\right)$$

or

$$R_m^2\frac{V_{xm}}{V_{yxm}} \le 4 \quad (4.1)$$

As we know that the estimators $t_3$ defined in (3.1) is the particular member of the generalised estimator $t_4$ so, if the above condition is satisfied for different values of $\alpha, \beta$, and $m_1, m_2$ the estimator $t_3$ will be better than usual estimator under measurement errors.

Also,

$$MSE(t_4)_{min} \le V(\bar{y}_m) \quad (4.2)$$

# 5. Empirical Study

Data statistics: The data used for empirical study has been taken from Gujrati and Sangeetha (2007) .

Where, Yi= True consumption expenditure,

Xi= True income,

yi= Measured consumption expenditure,

xi = Measured income.

From the data given, we get the following parameter values

| n | $\mu_y$ | $\mu_x$ | $\sigma_y^2$ | $\sigma_x^2$ | $\rho$ | $\sigma_u^2$ | $\sigma_v^2$ |
|---|---|---|---|---|---|---|---|
| 10 | 127 | 170 | 1278 | 3300 | 0.964 | 36.00 | 36.00 |

***Table 5.1.*** *Showing the Percent Relative Efficiencies' PRE's of estimators with respect to* $\overline{y}_m$ .

| Estimators | Values of $\alpha$ and $\beta$ | PRE |
|---|---|---|
| $\overline{y}_m$ | - | 100.00 |
| $t_1$ | - | 437.59 |
| $t_{regm}$ | $\alpha=1, \beta_m=\frac{V_{yxm}}{\sqrt{V_{xm}V_{ym}}}$ | 946.54 |
| $t_{2\min}$ | $\alpha=0, \beta=0$ | 944.94 |
| $t_{3\min}$ | $\alpha=1, \beta=1$ | 123.23 |
| | $\alpha=1, \beta=0$ | 603.01 |
| | $\alpha=0, \beta=1$ | 437.27 |
| | $\alpha=1, \beta=-1$ | 437.27 |
| $t_{4\min}$ | $\alpha=1, \beta=0$ | 1012.77 |
| | $\alpha=0, \beta=1$ | 1031.13 |
| | $\alpha=1, \beta=1$ | 948.35 |
| | $\alpha=1, \beta=-1$ | 1031.11 |

***Table 5.2.*** *Showing the MSE's of the estimators with and without measurement errors.*

| Estimators | | MSE without measurement error | Contribution of measurement error in MSE | MSE with measurement error |
|---|---|---|---|---|
| $\overline{y}_m$ | | 127.800 | 3.600 | 131.400 |
| $t_1$ | | 25.925 | 4.102 | 30.028 |
| $t_{reg}$ | | 9.000 | 4.896 | 13.882 |
| $t_{2\min}$ | $(\alpha=0, \beta=0)$ | 8.995 | 4.910 | 13.905 |
| $t_{3\min}$ | $(\alpha=1, \beta=0)$ | 17.203 | 4.587 | 21.790 |
| | $(\alpha=0, \beta=1)$ | 25.798 | 4.252 | 30.050 |
| | $(\alpha=1, \beta=1)$ | 101.874 | 4.747 | 106.621 |
| | $(\alpha=1, \beta=-1)$ | 25.772 | 4.278 | 30.050 |
| $t_{4\min}$ | $(\alpha=1, \beta=0)$ | 8.397 | 4.577 | 12.974 |
| | $(\alpha=0, \beta=1)$ | 8.536 | 4.207 | 12.743 |
| | $(\alpha=1, \beta=1)$ | 8.990 | 4.865 | 13.855 |
| | $(\alpha=1, \beta=-1)$ | 7.868 | 4.874 | 12.742 |

# 6. Conclusion

We observe that our proposed estimator $t_4$ is the most appropriate estimator given the set of optimality conditions depicted in Table 5.1. That is, the MSE of our proposed estimator is lower than the MSE of estimators that have been studied in this paper. Furthermore, it shall be noted that the class of usual estimator is the least that is impacted by the measurement error, and unequivocally it has maintained its topological stability. Our result that is illustrated in Table 5.2. confirms that it is imperative to consider observational errors in order to obtain true variances, and to minimize the topological overshooting and undershooting of measurement errors. Our Future research could potentially include the fuzzy efficiency of our estimator $t_4$ with a single Fuzzy Logic Controller. This de novo investigation perhaps enables us to test the sensitivity and specificity of various data structures more decisively under fuzzy measurement errors.

# References

[1] Allen, J., Singh, H. P. and Smarandache, F. (2003): A family of estimators of population mean using multiauxiliary information in presence of measurement errors. International Journal of Social Economics 30(7), 837–849.

[2] Bahl, S. and Tuteja, R. K. (1991): Ratio and product type exponential estimator. Information and optimization sciences12 (1), 159-163.

[3] Gujarati, D. N. and Sangeetha (2007): Basic econometrics. Tata McGraw – Hill.

[4] Koyuncu, N. and Kadilar, C. (2010): On the family of estimators of population mean in stratified sampling. Pakistan Journal of Statistics (to be published).

[5] Kumar, M., Singh, R., Sawan, N. and Chauhan, P. (2011a): Exponential ratio method of estimators in the presence of measurement errors. Int. J. Agricult. Stat. Sci. 7(2): 457-461.

[6] Kumar, M., Singh, R., Singh, A.K. and Smarandache, F. (2011b): Some ratio type estimators under measurement errors. WASJ 14(2) :272-276.

[7] Manisha and Singh, R. K. (2001): An estimation of population mean in the presence of measurement errors. Journal of Indian Society of Agricultural Statistics 54(1), 13–18.

[8] Manisha and Singh, R. K. (2002): Role of regression estimator involving measurement errors. Brazilian journal of probability Statistics 16, 39- 46.

[9] Shalabh (1997): Ratio method of estimation in the presence of measurement errors. Journal of Indian Society of Agricultural Statistics 50(2):150– 155.

[10] Singh, H. P. and Karpe, N. (2008): Ratio-product estimator for population mean in presence of measurement errors. Journal of Applied Statistical Sciences 16, 49–64.

[11] Singh, H. P. and Karpe, N. (2009): On the estimation of ratio and product of two populations means using supplementary information in presence of measurement errors. Department of Statistics, University of Bologna, 69(1), 27-47.

[12] Singh, R., Kumar, M., Chaudhary, M.K. (2011): Improved Family of Estimators of Population Mean in Simple Random Sampling. WASJ 13(10), 2131-2136.

[13] Solanki, R.S., Singh, H. P. and Rathour, A. (2012): An alternative estimator for estimating the finite population mean using auxiliary information in sample surveys. ISRN Probability and Statistics doi:10.5402/2012/657682