

Class of Estimators of Population Mode Using New Parametric Relationship for Mode

Mukesh K. Sharma, Sarbjit S. Brar*, Harinder Kaur

Department of Statistics, Punjabi University, Patiala, India

Abstract For estimating the population mode using the known information of population mean (\bar{X}) of the auxiliary variable, a class of estimators has been proposed by making use of new parametric relationship for mode developed by Sharma et.al. (2016). We have derived the asymptotic expression for the bias and MSE of any estimators of the proposed class and also obtain the minimum MSE in the proposed class. The results are also illustrated numerically.

Keywords Mode, Skewness, Auxiliary variable, Bias and Mean square error, Bivariate normal distribution

1. Introduction

In many practical situations, the more appropriate measure of location is mode rather than mean and median. For example, it is useful in finding the ideal size as in studies relating to marketing, trade and industry and also to find the ideal size such as in the manufacturing of shoes or ready-made garments, etc. In addition, the mode is also useful when dealing with skewed distributions such as income, drugs, AIDS etc. The impact of the proposed research is expected to be useful for scientists in Sociology, Psychology, Demography, Business and Economics, where the mode value is routinely being used in practice.

Sharma et al (2016) established the new parametric relationship for population mode (M_o) as

$$M_o = \bar{Y} - k \frac{\mu_{30}}{S_y^2}$$

where k is unknown constant to be determined for the given population. They proposed three estimators of M_o under the different situations as

$$\hat{M}_{o1} = \bar{y} - k_1 \frac{m_{30}}{s_y^2},$$

$$\hat{M}_{o2} = \bar{y} \frac{\bar{X}}{\bar{x}} - k_2 \frac{m_{30}}{s_y^2} \frac{\bar{X}}{\bar{x}},$$

and

$$\hat{M}_{o3} = \bar{y} \frac{\bar{x}}{\bar{X}} - k_3 \frac{m_{30}}{s_y^2} \frac{\bar{x}}{\bar{X}},$$

Where k_1, k_2 and k_3 are unknown constants, whose values are determined by minimizing the MSE's of

respective estimators $\hat{M}_{o1}, \hat{M}_{o2}$ and \hat{M}_{o3} . Here the estimator \hat{M}_{o1} uses no information on auxiliary variable x , which is highly correlated with y , whereas \hat{M}_{o2} and \hat{M}_{o3} uses the known information of \bar{X} , which are of ratio and product type estimators respectively.

In the present paper, we propose a class of estimators of population mode using the new parametric relationship for population mode when the population mean (\bar{X}) of the auxiliary variable is known. Asymptotic expression for the bias and MSE of any estimators of the proposed class and also their minimum values are obtained upto to terms of order n^{-1} . The results are also illustrated numerically.

2. Notations and Expectations

Let a simple random sample of size n is drawn from a population of size N by using simple random sampling without replacement (SRSWOR). Without loss of generality, we ignore the finite population correction (fpc) throughout the paper.

Let Y_i and X_i be the values of study and auxiliary variables for i^{th} unit in the population respectively and the corresponding small letters y_i and x_i denote the sample values.

Taking,

$$\bar{Y} = \frac{1}{N} \sum_{i=1}^N Y_i, \quad \bar{X} = \frac{1}{N} \sum_{i=1}^N X_i,$$

$$S_y^2 = \frac{1}{N-1} \sum_{i=1}^N (Y_i - \bar{Y})^2,$$

$$S_x^2 = \frac{1}{N-1} \sum_{i=1}^N (X_i - \bar{X})^2,$$

$$\mu_{rs} = \frac{1}{N} \sum_{i=1}^N (Y_i - \bar{Y})^r (X_i - \bar{X})^s,$$

* Corresponding author:

sarbjitbrar@pbi.ac.in (Sarbjit S. Brar)

Published online at <http://journal.sapub.org/ajms>

Copyright © 2016 Scientific & Academic Publishing. All Rights Reserved

$$\lambda_{rs} = \frac{\mu_{rs}}{\mu_{20}^{r/2} \mu_{02}^{s/2}},$$

$$m_{30} = \frac{n}{(n-1)(n-2)} \sum_{i=1}^n (y_i - \bar{y})^3.$$

Obviously

$\lambda_{11} = \rho_{xy} = \rho$ (Correlation between x and y),

$\lambda_{30} = \beta_{1y}$ (Coefficient of skewness of y),

$\lambda_{40} = \beta_{2y}$ (Coefficient of kurtosis of y).

Defining,

$$\delta_0 = \frac{\bar{y}}{\bar{Y}} - 1, \quad \delta = \frac{s_y^2}{S_y^2} - 1$$

$$\epsilon = \frac{\bar{x}}{\bar{X}} - 1, \quad \eta_1 = \frac{m_{30}}{\mu_{30}} - 1$$

For the given SRSWOR, we have the following expectations,

$$E(\delta_0) = E(\delta) = E(\epsilon) = E(\eta_1) = 0$$

$$E(\delta_0^2) = \frac{1}{n} C_y^2, \quad E(\epsilon^2) = \frac{1}{n} C_x^2,$$

$$E(\delta_0 \epsilon) = \frac{1}{n} C_{yx}, \quad E(\epsilon \delta) = \frac{1}{n} \lambda_{21} C_x$$

$$E(\delta_0 \delta) = \frac{1}{n} \lambda_{30} C_y = \frac{1}{n} \beta_{1y} C_y$$

and up to terms of order n^{-1}

$$E(\delta^2) = \frac{1}{n} (\lambda_{40} - 1) = \frac{1}{n} (\beta_{2y} - 1),$$

$$E(\eta_1^2) = \frac{1}{n} \frac{(\lambda_{60} - 6\lambda_{40} - \lambda_{30}^2 + 9)}{\lambda_{30}^2}$$

$$= \frac{1}{n} \frac{(\lambda_{60} - 6\beta_{2y} - \beta_{1y}^2 + 9)}{\beta_{1y}^2},$$

$$E(\delta_0 \eta_1) = \frac{1}{n} \frac{(\lambda_{40} - 3)}{\lambda_{30}} C_y = \frac{1}{n} \frac{(\beta_{2y} - 3)}{\beta_{1y}} C_y,$$

$$E(\delta \eta_1) = \frac{1}{n} \frac{(\lambda_{50} - 3\lambda_{30})}{\lambda_{30}} = \frac{1}{n} \frac{(\lambda_{50} - 3\beta_{1y})}{\beta_{1y}},$$

$$E(\epsilon \eta_1) = \frac{1}{n} \frac{(\lambda_{31} - 3\rho)}{\lambda_{30}} C_x = \frac{1}{n} \frac{(\lambda_{31} - 3\rho)}{\beta_{1y}} C_x$$

3. Proposed Class of Estimators

In this section, we generalize the mean per unit, ratio-type and product-type estimators of population mode (M_o) given by Sharma *et al.* (2016).

Whatever be the sample chosen, let $u = \frac{\bar{x}}{\bar{X}}$ assume values in a bounded closed convex subset R of the one-dimensional real space containing the value '1'. Let $t(u)$ be a function of u such that

$$t(1) = 1$$

and it satisfies the following conditions:

- (i) The function $t(u)$ is continuous and bounded in R .
- (ii) The first and second order partial derivatives of $t(u)$ exist and are continuous and bounded in R .

We propose the class of estimator of the population mode (M_o) as

$$\tilde{M}_{og} = \hat{M}_o t(u) \quad (3.1)$$

To find the biases and MSE of estimators of class \tilde{M}_{og} , we expand the function $t(u)$ about the value '1' in second-order Taylor's series, we get

$$\tilde{M}_{og} = \hat{M}_o \left[1 + (u-1) t'(1) + (u-1)^2 \frac{t''(1)}{2} + \dots \right]$$

where $t'(1) = \left(\frac{\partial t}{\partial u} \right)_{u=1} = t_1(\text{say})$ and $t''(1) = \left(\frac{\partial^2 t}{\partial u^2} \right)_{u=1} = t_2(\text{say})$.

After writing the above function in terms of ϵ 's, η 's, δ 's and then taking the expectations given in section 2, upto terms of order n^{-1} , we get,

$$B(\tilde{M}_{og}) = \frac{1}{n} \frac{\bar{y}}{2} \left[\{C_x^2 t_2 + 2C_{yx} t_1 - k C_y \{ \beta_{1y} C_x^2 t_2 - 2B_1 C_x t_1 + 2\beta_{1y} (\beta_{2y} + 3) - 2\lambda_{50} \} \} \right] \quad (3.2)$$

$$MSE(M_{og}) = \frac{1}{n} \frac{\bar{y}^2}{2} \left[\left\{ C_y^2 + C_x^2 t_1^2 + 2C_{yx} t_1 \right\} + k^2 C_y^2 \left\{ (\lambda_{60} - 6\beta_y + \beta_{0y}) + \beta_{1y}^2 C_x^2 t_1^2 - 2\beta_{1y} B_1 C_x t_1 \right\} \right. \\ \left. - 2k C_y \left\{ \beta_y C_y + \beta_{1y} C_x^2 t_1^2 + (\beta_{1y} \rho C_y + B_1) C_x t_1 \right\} \right] \quad (3.3)$$

where

$$\beta_y = \beta_{2y} - \beta_{1y}^2 - 3,$$

$$\beta_{0y} = \beta_{1y}^2 \beta_{2y} - 2\beta_{1y} \lambda_{50} - 9,$$

$$B_1 = \beta_{1y} \lambda_{21} - \lambda_{31} + 3\rho,$$

$$B = B_1 C_x.$$

Here k and t_1 are unknown constants whose values are determined by minimizing $MSE(\tilde{M}_{og})$.

To obtain the minimum value of $MSE(\tilde{M}_{og})$ we differentiate (3.3) w.r.t. k and t_1 , then equating to zero, we get,

$$(kC_y\beta_{1y} - 1)\beta_{1y}C_x^2t_1^2 - (2kC_y\beta_{1y}B_1 + \beta_{1y}\rho C_y - B_1)C_x t_1 + kC_y(\lambda_{60} - 6\beta_y + \beta_{0y}) - \beta_y C_y = 0$$

and

$$(1 + k^2C_y^2\beta_{1y}^2 - 2kC_y\beta_{1y})C_x t_1 + \{\rho - k^2C_y\beta_{1y}B_1 - k(\beta_{1y}\rho C_y - B_1)\}C_y = 0.$$

Solving above quadratic equations for k and t_1 , we get two pairs of solutions (k_1, t_{11}) and (k_2, t_{12}) where,

$$k_1 = \frac{(B_1\rho + \beta_y)}{\{(\lambda_{60} - 6\beta_y + \beta_{0y}) - B_1^2\}},$$

$$t_{11} = \frac{C_y\{(B_1\rho + \beta_y)B_1 + (\lambda_{60} - 6\beta_y + \beta_{0y})\rho - B_1^2\rho\}}{C_x\{(B_1\rho + \beta_y)\beta_{1y}C_y - (\lambda_{60} - 6\beta_y + \beta_{0y}) + B_1^2\}}$$

$$k_2 = \frac{1}{\beta_{1y}C_y},$$

$$t_{12} = \infty.$$

Clearly, the only feasible pair of values is (k_1, t_{11}) .

Substituting the values of pair (k_1, t_{11}) in (3.3), we get,

$$MSE_{min}(\tilde{M}_{og}) = \frac{1}{n}\bar{Y}^2C_y^2\left[1 - \rho^2 - \frac{(B_1\rho + \beta_y)^2}{(\lambda_{60} - 6\beta_y + \beta_{0y}) - B_1^2}\right]$$

Table 1. Description of populations

Sr. No.	Source	y	x	\bar{Y}	k_1	μ_{20}	β_{1y}	ρ
1	Murthy(1967) P.91 (1-35)	Cultivated area (acres)	Holding size (acres)	2.3650	-0.2217	1.5818	0.9119	0.3685
2	Murthy(1967) P.398	No. of absentees	No. of workers	9.6512	0.0442	42.1341	1.5575	0.6608
3	Murthy(1967) P.399	Area under wheat in 1964	Cultivated area in 1961	199.4412	-0.0220	21900.8936	1.1295	0.9043
4	Chakravarti et al.(1967) P-183	Length(cm) measured by 1 st person	Length (cm) measured by 2 nd person	4.9737	-0.0437	0.1346	-0.0546	0.9317
5	Chakravarti et al.(1967) P-207	Weight (kg) of male	Height (cm) of male	29.2625	-0.0240	6.5836	0.3670	0.7709
6	Chakravarti et al.(1967) P-207	Weight (kg) of female	Height (cm) of female	28.5313	-0.3896	1.8109	0.1099	0.2306
7	Chakravarti et al.(1967) P-185 (1-35)	Weight (lb) of Kayastha males	Stature (cm) of Kayastha males	82.2000	-0.2012	191.7029	0.0439	0.8578
8	Chakravarti et al.(1967) P-185 (1-76)	Weight (lb) of Kayastha males	Stature (cm) of Kayastha males	89.4211	0.0516	278.4806	0.6068	0.4361
9	Chochran (1999) P-325	Total number of persons	Average persons per room	101.1000	-0.3015	214.6900	0.3248	0.6515
10	Maddala p-96	Consumption per capital of Lamb	Deflated prices of Lamb	4.5188	-0.0281	0.2103	-0.6578	-0.7517
11	Guajrati p-27,(1-50)	Price per dozen (cent) in 1990	Egg production in 1991 (million)	78.2880	0.0111	445.3787	0.9959	-0.3096
12	http://content.hc.cfl.edu	Highway fuel efficiency of vehicles (in miles)	Weight of vehicles (in 1000 lbs.)	30.6154	-0.2045	15.6213	0.0549	-0.8978

Table 2. MSE's of \tilde{M}_{01} , \tilde{M}_{02} , \tilde{M}_{03} , $\hat{\bar{Y}}_R$, $\hat{\bar{Y}}_P$, \tilde{M}_{og} and $\hat{\bar{Y}}_{lr}$ up to terms of n^{-1}

Pop. No.	\tilde{M}_{01}	\tilde{M}_{02}	$1/n \times \text{MSE of } \tilde{M}_{03}$	$\hat{\bar{Y}}_R$	$\hat{\bar{Y}}_P$	\tilde{M}_{og}	$\hat{\bar{Y}}_{lr}$
1	1.2908	5.2902	14.1623	-	-	1.2181	1.3670
2	38.9929	22.4016	84.6724	23.7459	-	22.3983	23.7380
3	20163.1975	4081.6692	66594.1526	4286.4483	-	3955.1587	3992.7274
4	0.1085	0.0201	0.4480	0.0201	-	0.0173	0.0178
5	6.4593	3.8957	10.4341	3.9590	-	2.6614	2.6713
6	1.2534	1.7332	2.2410	-	-	1.2315	1.7146
7	102.2873	58.5715	172.6350	105.5227	-	41.0129	50.6533
8	263.7385	220.9858	532.8689	237.2253	-	210.0375	225.5076
9	146.2998	117.9843	494.0659	135.1725	-	92.0838	123.5609
10	0.1926	0.6273	0.1022	-	0.1023	0.0910	0.0915
11	445.3281	9649.3442	6697.4891	-	-	402.1633	402.7018
12	6.7198	39.3146	6.1189	-	6.7647	2.4786	3.0308

Where $\hat{\bar{Y}}_R$, $\hat{\bar{Y}}_P$ and $\hat{\bar{Y}}_{lr}$ denotes the usual ratio, product and regression estimators for population mean.

We can construct the large number of estimators belonging to the proposed class \tilde{M}_{og} . Here it should be noted that the efficient use of estimators of the proposed class \tilde{M}_{og} requires the optimum values of constants k and t_1 , which are further functions of unknown population parameters. However, if it is possible to guess accurately the values of such parameters either through past experience or through a pilot sample survey and the value of optimum constants so obtained by using these guessed values of parameters are close enough to the optimum values of constants, then the resulting estimators will be better than the convention estimators. Srivastava and Jhaji (1983) have shown that upto the first order of approximation, the estimators of the class with estimated values of optimum parameters obtained by their respective consistent estimators, attain the same optimum MSE of class based on optimum values upto the first order of approximation. Therefore, the presence of unknown population parameters in optimum values of constants will not create any problem for practical use of the proposed class \tilde{M}_{og} .

3.1. Special Case of Bivariate Normal Population

Let $(Y, X) \sim N(\mu_y, \mu_x, \sigma_y^2, \sigma_x^2, \rho)$, then we have $\lambda_{60} = 15$, $\lambda_{40} = 3$, $\lambda_{31} = 3\rho$, $\lambda_{22} = 1 + 2\rho^2$, $\lambda_{r,s} = 0$ if $r + s$ is odd. Also, $X \sim N(\mu_x, \sigma_x^2)$ and $Y \sim N(\mu_y, \sigma_y^2)$.

Using these values, we get,

$$MSE_{min}(\tilde{M}_{og}) = MSE(\hat{\bar{Y}}_{lr}) = \frac{1}{n} S_y^2 (1 - \rho^2).$$

4. Numerical Illustrations

To illustrate the result numerically, we have made computations for 12 populations taken from literature by using Microsoft Excel 2010.

The source of the populations, the nature of the variables, the values of \bar{Y} , k_1 , μ_{20} , β_{1y} and ρ are listed in Table 1.

The efficiencies of proposed estimators are given in Table 2.

From Table 2, it has been clarified that upto terms of order $1/n$, $MSE_{min}(\tilde{M}_{og})$ is less than $MSE(\tilde{M}_{01})$, $MSE(\tilde{M}_{02})$ and $MSE(\tilde{M}_{03})$ which implies that the proposed estimator is more efficient than the existing ones. Further, $MSE_{min}(\tilde{M}_{og})$ is also smaller than $MSE(\hat{\bar{Y}}_R)$, $MSE(\hat{\bar{Y}}_P)$, $MSE(\hat{\bar{Y}}_{lr})$ which shows that at optimum value of (k, t_1) , the proposed class estimators is more stable than the $\hat{\bar{Y}}_R$, $\hat{\bar{Y}}_P$ and $\hat{\bar{Y}}_{lr}$.

5. Conclusions

A general class of estimators of population mode has been proposed and the lower bound for the MSE has been obtained for the class of estimators. Further, the empirical study shows that $MSE_{min}(\tilde{M}_{og})$ is less than $MSE(\hat{\bar{Y}}_R)$, $MSE(\hat{\bar{Y}}_P)$, $MSE(\hat{\bar{Y}}_{lr})$ which implies that the optimum estimator is more stable than $\hat{\bar{Y}}_R$, $\hat{\bar{Y}}_P$ and $\hat{\bar{Y}}_{lr}$.

REFERENCES

- [1] Chakravarty, I. M., Laha, R. G., & Roy, J. (1967). *Handbook of Methods of Applied Statistics: Techniques of Computation, Descriptive Methods, and Statistical Inference*. John Wiley & Sons.
- [2] Cochran, W. G. (1999). *Sampling Techniques* (Vol.3). John Wiley & Sons.
- [3] Gujarati, D. N. (2004). *Basic Econometrics*. Mc. Graw Hills Pub. Co, New York.

- [4] Kuk, A. Y., & Mak, T. K. (1989). Median estimation in the presence of auxiliary information. *Journal of the Royal Statistical Society. Series B (Methodological)*, 261-269.
- [5] Maddala, G. S., & Lahiri, K. (1992). *Introduction to econometrics* (Vol. 2). New York: Macmillan.
- [6] Sedory, S. A., & Singh, S. (2014). Estimation of mode using auxiliary information. *Communications in Statistics-Simulation and Computation*, 43(10), 2390-2402.
- [7] Sharma, M.K., Brar, S.S. & Kaur, H. (2016). Estimators of population mode using new parametric relationship mode. *International Journal of Probability and Statistics*, 5(1), 18-24.
- [8] Srivastava, S.K. & Jhaji, H.S. (1983). A class of estimators of the population mean using multi-auxiliary information. *Cal. Stat. Assoc. Bull.*, 32, 47-56.
- [9] Murthy, M. N. (1967). Sampling theory and methods. *Sampling theory and methods*.