

## Research Article

# On a Class of Almost Unbiased Ratio Type Estimators

A.K.P.C. Swain<sup>1\*</sup>, Priyaranjan Dash<sup>2</sup>

<sup>1</sup>Former Professor of Statistics, Utkal University, Bhubaneswar, India

<sup>2</sup>Department of Statistics, Utkal University, Bhubaneswar, India

## ARTICLE INFO

### Article History

Received 04 Jul 2018

Accepted 15 oct 2019

### Keywords

Simple random sampling

Ratio estimator

Bias

Almost unbiased ratio type estimator

Variance

Efficiency

### 2000 Mathematics Subject

Classification: 62D05

## ABSTRACT

In sample surveys ratio estimator has found extensive applications to obtain more precise estimators of the population ratio, population mean, and population total of the study variable in the presence of auxiliary information, when the study variable is positively correlated with the auxiliary variable. The theory underlying the ratio method of estimation is same whether we estimate the population ratio or population mean/population total, excepting the fact that in the latter case we assume the advance knowledge of the population mean or total of the auxiliary variable in question. In this paper we use the term ratio estimator for both the purposes. However, in spite of its simplicity the ratio estimator is accompanied by an unwelcome bias, although the bias decreases with increase in sample size and is negligible for large sample sizes. In small samples the bias may be substantial so as to downgrade its utility by affecting the reliability of the estimate. As pointed out by L.A. Goodman, H.O. Hartley, J. Am. Stat. Assoc. 53 (1958), 491–508, in sample surveys where we draw very small samples from a large number of strata in stratified random sampling with the ratio method of estimation in each stratum, the combined bias from all the strata may assume serious proportions, affecting the reliability of the estimate. This calls for devising techniques either at estimation stage or in the sampling scheme at the selection stage to reduce the bias or completely eliminating it to make it usable in practice. This has motivated many research workers like E.M.L. Beale, Ind. Organ. 31 (1962), 27–28 and M. Tin, J. Am. Stat. Assoc. 60 (1965), 294–307 among others to construct estimators at the estimation stage removing the bias of  $O(1/n)$ , where  $n$  is the sample size, and thus reducing the bias to  $O(1/n^2)$ . Such estimators are termed as Almost Unbiased ratio-type estimators found in literature. In this paper we have proposed a class of almost ratio type estimators following the techniques of E.M.L. Beale, Ind. Organ. 31 (1962), 27–28 and M. Tin, J. Am. Stat. Assoc. 60 (1965), 294–307 and made comparison with regard to bias and efficiency.

© 2020 The Authors. Published by Atlantis Press SARL.

This is an open access article distributed under the CC BY-NC 4.0 license (<http://creativecommons.org/licenses/by-nc/4.0/>).

## 1. INTRODUCTION

Large scale sample surveys are often conducted in countries around the world to assess the present status of certain sectors of economy for future planning. In such surveys it is a general practice to adopt stratification to divide the heterogeneous population into homogeneous groups called strata. Sometimes besides observing the main variable under study, observations on certain auxiliary variables stipulated at planning stage or even during the course of investigation to improve the efficiency of the estimators of the parameters of the main variable under study. The simplest method of using auxiliary information in case of a single auxiliary variable when the main variable under study and auxiliary variable is positively correlated, is the ratio method of estimation, advocated by Cochran [3] among many [5] earlier workers in sample surveys. It is well known that the ratio estimator of the population mean/total/ratio is a biased estimator although the bias may be negligible for large sample sizes. Even for moderately large sample the bias may be substantial, more so in stratification where these biases accumulate over strata to make the overall estimate sometimes unacceptable to be used for the purpose for which it is to be used (Goodman and Hartley [7], Cochran [4]). This suggests to devise ways to construct estimators whose biases of  $O(1/n)$ ,  $n$  being the sample size, is removed and the reduced bias becomes of  $O(1/n^2)$ . Beale [1] and Tin [14] devised ways to adjust the estimator for the bias by the asymptotic series expansion of the ratio estimator under certain assumptions. These improved type of ratio estimators having first order bias being removed are known in sampling theory literature as Almost Unbiased Ratio Type Estimators. De-graft Johnson [6] and David [5] have made some extensive studies on ratio method of estimation.

Let there be a finite population  $U$  having  $N$  distinct and identifiable units  $\{U_1, U_2, \dots, U_N\}$  indexed by paired values of the study variable  $y$  and positively correlated auxiliary variable  $x$  such as  $(Y_1, X_1), (Y_2, X_2), \dots, (Y_N, X_N)$ . Assume that both  $y$  and  $x$  are positively measured.

Draw a simple random sample without replacement of size  $n$  from the finite population of  $N$  units and the paired values on the sample units are  $(y_1, x_1), (y_2, x_2), \dots, (y_n, x_n)$ .

\*Corresponding author. Email: [akpcs@rediffmail.com](mailto:akpcs@rediffmail.com)

Define the population means of  $y$  and  $x$  as  $\bar{Y} = \frac{1}{N} \sum_{i=1}^N Y_i$  and  $\bar{X} = \frac{1}{N} \sum_{i=1}^N X_i$  respectively and the population variances and covariance between  $y$  and  $x$  as  $S_y^2 = \frac{1}{N-1} \sum_{i=1}^N (Y_i - \bar{Y})^2$ ,  $S_x^2 = \frac{1}{N-1} \sum_{i=1}^N (X_i - \bar{X})^2$ , and  $S_{xy} = \frac{1}{N-1} \sum_{i=1}^N (X_i - \bar{X})(Y_i - \bar{Y})$  respectively. Define further  $C_x^2 = \frac{S_x^2}{\bar{X}^2}$  and  $C_y^2 = \frac{S_y^2}{\bar{Y}^2}$  as the population squared coefficients of variation of  $x$  and  $y$  respectively. Also, the population coefficient of co-variation  $C_{xy} = \frac{S_{xy}}{\bar{Y}\bar{X}} = \rho C_x C_y$ ,  $\rho$  being the coefficient of correlation between  $y$  and  $x$ . The population regression coefficient of  $y$  on  $x$   $\beta = \frac{S_{xy}}{S_x^2}$ . Further, the population ratio  $R = \frac{\bar{Y}}{\bar{X}}$ .

The sample means of  $y$  and  $x$  are respectively  $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$  and  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ . The sample variances of  $y$  and  $x$  are  $s_y^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2$  and  $s_x^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$  respectively and sample covariance is  $s_{xy} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$ .

Define  $c_x^2 = \frac{s_x^2}{\bar{x}^2}$ ,  $c_y^2 = \frac{s_y^2}{\bar{y}^2}$ , and  $c_{xy} = \frac{s_{xy}}{\bar{x}\bar{y}}$ .

$c_x^2$ ,  $c_y^2$ , and  $c_{xy}$  are consistent estimators of  $C_x^2$ ,  $C_y^2$ , and  $C_{xy}$  respectively.

Also, sample ratio  $r = \frac{\bar{y}}{\bar{x}}$ .

The ratio estimator of the population mean  $\bar{Y}$  is given by

$$\hat{\bar{Y}}_r = \frac{\bar{y}}{\bar{x}} \bar{X}, \quad (1)$$

where  $\bar{X}$  is known in advance.

Expanding (1) in power series (Sukhatme *et al.* [12])

and using results  $V(\bar{y}) = \theta C_y^2$ ,  $V(\bar{x}) = \theta C_x^2$ , and  $Cov(\bar{x}, \bar{y}) = \theta C_{xy}$ ,

we have to  $O(1/n)$ ,

$$E(\hat{\bar{Y}}_r) = \bar{Y} [1 + \theta (C_x^2 - C_{xy})], \text{ where} \quad (2)$$

$$\theta = \left( \frac{1}{n} - \frac{1}{N} \right).$$

Hence bias to  $O(1/n)$  is given by

$$\text{Bias}(\hat{\bar{Y}}_r) = \theta \bar{Y} (C_x^2 - C_{xy})$$

$\hat{\bar{Y}}_r$  is a biased but consistent estimator of  $\bar{Y}$  with bias being negligible in large samples.

Alternatively we may put

$$E(\hat{\bar{Y}}_r) = \bar{Y} + \bar{Y} \theta_1 (C_{20} - C_{11}), \quad (3)$$

where  $C_{ij} = \frac{\mu_{ij}}{\bar{X}^i \bar{Y}^j} = \frac{\frac{1}{N} \sum (X_i - \bar{X})^i (Y_i - \bar{Y})^j}{\bar{X}^i \bar{Y}^j}$ , and  $\theta_1 = \frac{N-n}{(N-1)n}$ .

For very large  $N$ ,  $\theta_1 \approx \theta$ .

The variance of  $\hat{\bar{Y}}_r$  to  $O(1/n)$  is given by (Sukhatme *et al.* [12])

$$V(\hat{\bar{Y}}_r) = \theta \bar{Y}^2 (C_y^2 + C_x^2 - 2C_{xy}), \quad (4)$$

Alternatively,

$$V(\hat{\bar{Y}}_r) = \theta \bar{Y}^2 (C_{02} + C_{20} - 2C_{11}) \quad (5)$$

$\hat{\bar{Y}}_r$  is more efficient than  $\bar{y}$  if

$$\rho > \frac{1}{2} \frac{C_x}{C_y} = \frac{1}{2} \sqrt{\frac{C_{20}}{C_{11}}}, \quad (6)$$

where  $\rho$  is the correlation coefficient between  $y$  and  $x$ .

Beale [1] suggested an ingenious almost unbiased ratio type estimator given by

$$\hat{\bar{Y}}_{rB} = \hat{\bar{Y}}_r \left[ \frac{1 + \theta s_{xy} / \bar{xy}}{1 + \theta s_x^2 / \bar{x}^2} \right] \quad (7)$$

Tin [14] derived another almost unbiased estimator by subtracting the estimate of first order bias of  $O(1/n)$  from the estimator  $\hat{\bar{Y}}_r$  itself to get an estimator whose bias of  $O(1/n)$  is removed, so as to get his estimator having bias of  $O(1/n^2)$ . Thus Tin's estimator is given by

$$\hat{\bar{Y}}_{rT} = \hat{\bar{Y}}_r \left[ 1 + \theta \left( \frac{s_{xy}}{\bar{xy}} - \frac{s_x^2}{\bar{x}^2} \right) \right] \quad (8)$$

Beale's estimator is in ratio form, which reduces to Tin's form after its asymptotic expansion retaining terms up to  $O(1/n)$ . It also eliminates the bias of  $O(1/n)$  of the ratio estimator. Both Tin's and Beale's estimators use same information  $\frac{s_{xy}}{\bar{xy}}$  and  $\frac{s_x^2}{\bar{x}^2}$  in their formulations.

Considering terms up to  $O(1/n^2)$ , Tin [14] has shown that Beale's estimator is less biased and equally efficient compared to his estimator. A disadvantage with Tin's estimator is that it may take negative values for a positive population ratio, a situation pointed out by Beale in a private communication with Tin [14] with a sample of size two. Such discerning picture has also been seen by drawing a sample of size 8 from a bivariate normal population having  $\bar{X} = 5$ ,  $\bar{Y} = 15$ ,  $S_x^2 = 45$ ,  $S_y^2 = 500$ , and correlation coefficient  $\rho = 0.4$ . The computed estimators are  $r = 117.11$ ,  $r_b = 1.55$  and  $r_t = -64963.33$  (privately communicated by Mr. Xiafei Zhang, Iowa State University).

Again, writing Beale's estimator  $r_b$  for  $R$  as

$$r_b = r \left[ \frac{1 + \theta s_{yx} / \bar{yx}}{1 + \theta s_x^2 / \bar{x}^2} \right] = \frac{\bar{yx} + \theta s_{yx}}{\bar{x}^2 + \theta s_x^2},$$

we find that when  $\bar{x}$  is small or even  $\bar{x} = 0$ , Beale's estimator is dominated by  $s_{yx}/s_x^2$  and hence will not give extreme values and thus has a control to avoid extremes. As compared to Beale's estimator, Tin's estimator is dominated by  $s_x^2/\bar{x}^2$  and hence is extremely large when  $\bar{x}$  is small. The advantage with Beale's estimator is that it can deal with case when  $\bar{x} = 0$ . These thoughts were expressed by Professor W.A. Fuller in a private communication.

As noted by Tin [14], Beale's estimator seems to be a better estimator than his estimator as regards reducing the bias of the ratio estimator and also in large samples there is marginal loss of efficiency compared to his estimator.

Beale's estimator has been fruitfully applied in hydrological studies by Lee *et al.* [8] in load estimation using dense water quality data. Assuming positive correlation between flux and flow Richards and Holloway [9] and Richards [10] used Beale's estimator for flux estimation in the Great Lakes region and other parts of United States, generally applying more complex strata. They showed that Beale's estimator generally exhibited greater estimation accuracy and lower bias. Carriquiry *et al.* [2] have studied the estimation of usual intake distributions of intake of ratios of dietary components using Beale's estimator.

In this paper Srivastava's [11] class of estimators is considered to derive its Beale type and Tin type almost unbiased ratio type estimators and to compare them with regard to bias and efficiency. Further, as a special case Swain's [13] square root transformation estimator is discussed at length to compare Beale type and Tin type estimators with regard to bias and efficiency. As exact comparisons are not possible we have used asymptotic expansions and considered terms to  $O(1/n^2)$ .

## 2. A CLASS OF ALMOST UNBIASED RATIO TYPE ESTIMATORS

Srivastava [11] proposed a class of power transformation ratio estimators of the population mean  $\bar{Y}$ , with known population mean  $\bar{X}$  as

$$t_s = \bar{y} \left( \frac{\bar{X}}{\bar{x}} \right)^\alpha, \quad (9)$$

where  $\alpha$  is a real constant.

Define,

$$\bar{y} = \bar{Y}(1 + e_0), \bar{x} = \bar{X}(1 + e_1), s_{xy} = S_{xy}(1 + e_2), s_x^2 = S_x^2(1 + e_3) \\ E(e_i) = 0, i = 0, 1, 2, 3$$

Expanding  $t_s = \bar{y} \left( \frac{\bar{X}}{\bar{x}} \right)^\alpha$  in power series, assuming  $|e_i| < 1$  for all possible samples,  $i = 0, 1, 2, 3$  and retaining terms up to degree four, we have

$$\begin{aligned} t_s &= \bar{Y}(1 + e_0)(1 + e_1)^{-\alpha} \\ &= \bar{Y}(1 - \alpha e_1 + \frac{\alpha(\alpha + 1)}{2} e_1^2 - \frac{\alpha(\alpha + 1)(\alpha + 2)}{6} e_1^3 + \frac{\alpha(\alpha + 1)(\alpha + 2)}{24} e_1^4 \\ &\quad + e_0 - \alpha e_1 e_0 + \frac{\alpha(\alpha + 1)}{2} e_1^2 e_0 - \frac{\alpha(\alpha + 1)(\alpha + 2)}{6} e_1^3 e_0 + \dots) \\ &= \bar{Y}(1 - \lambda_1 e_1 + \lambda_2 e_1^2 - \lambda_3 e_1^3 + \lambda_4 e_1^4 + e_0 - \lambda_1 e_1 e_0 + \lambda_2 e_1^2 e_0 - \lambda_3 e_1^3 e_0 + \dots), \end{aligned} \quad (10)$$

$$\text{where } \lambda_1 = \alpha, \lambda_2 = \frac{\alpha(\alpha + 1)}{2}, \lambda_3 = \frac{\alpha(\alpha + 1)(\alpha + 2)}{6}, \lambda_4 = \frac{\alpha(\alpha + 1)(\alpha + 2)(\alpha + 3)}{24}$$

After some lengthy derivations using traditional techniques adopted for asymptotic expansion of the ratio estimator (see Sukhatme *et al.* [12]), we have to  $O(1/n^2)$

$$E(t_s) = \bar{Y} + \bar{Y} [\theta (\lambda_2 C_{20} - \lambda_1 C_{11}) + \theta^2 (-\lambda_3 C_{30} + 3\lambda_4 C_{20}^2 + \lambda_2 C_{21} - 3\lambda_3 C_{20} C_{11})], \quad (11)$$

Now,

$$\begin{aligned} V(t_s) &= \bar{Y}^2 \theta (\lambda_1^2 C_{20} - 2\lambda_1 C_{11} + C_{02}) \\ &\quad + \bar{Y}^2 \theta^2 \{C_{20}^2 (2\lambda_2^2 + 6\lambda_1 \lambda_3) + C_{11}^2 (\lambda_1^2 + 4\lambda_2) + C_{20} C_{02} (\lambda_1^2 + 2\lambda_2) + C_{30} (-2\lambda_1 \lambda_2)\} \\ &\quad + \bar{Y}^2 \theta^2 \{C_{21} (2\lambda_1^2 + 2\lambda_2) + C_{20} C_{11} (-6\lambda_3 - 10\lambda_1 \lambda_2) - 2\lambda_1 C_{12}\} \end{aligned} \quad (12)$$

When  $\lambda_i = 1$  for  $i = 1, 2, 3, 4$ ,

$$E(\hat{\bar{Y}}_r) = \bar{Y} + \bar{Y} [\theta (C_{20} - C_{11}) + \theta^2 (-C_{30} + 3C_{20}^2 + C_{21} - 3C_{20} C_{11})] \quad (13)$$

$$\begin{aligned} V(\hat{\bar{Y}}_r) &= \bar{Y}^2 \theta \{C_{20} - 2C_{11} + C_{02}\} \\ &\quad + \bar{Y}^2 \theta^2 \{8C_{20}^2 + 5C_{11}^2 + 3C_{20} C_{02} - 2C_{30} + 4C_{21} - 16C_{20} C_{11} - 2C_{12}\} \end{aligned} \quad (14)$$

The expressions for  $E(\hat{\bar{Y}}_r)$  and  $V(\hat{\bar{Y}}_r)$  are the same as those derived by Tin [14] and De-Graft Johnson [6].

Following Beale [1] we write an almost unbiased ratio estimator of  $\bar{Y}$  using  $t_s$  given by

$$t_{sB} = \bar{y} \left( \frac{\bar{X}}{\bar{x}} \right)^\alpha \left[ \frac{1 + \lambda_1 \theta \frac{s_{xy}}{xy}}{1 + \lambda_2 \theta \frac{s_x^2}{x^2}} \right] \quad (15)$$

Now,

$$\begin{aligned} 1 + \lambda_1 \theta \frac{s_{xy}}{xy} &= 1 + \lambda_1 \theta C_{11} (1 + e_0)^{-1} (1 + e_1)^{-1} (1 + e_2) \\ &= 1 + \lambda_1 \theta C_{11} (1 - e_0 + e_0^2 - e_0^3 + e_0^4 + \dots) (1 - e_1 + e_1^2 - e_1^3 + e_1^4 + \dots) (1 + e_2) \\ &= B, \text{ say.} \end{aligned}$$

Further,

$$\begin{aligned} \left( 1 + \lambda_2 \theta \frac{s_x^2}{x^2} \right)^{-1} &= [1 + \lambda_2 \theta C_{20} (1 + e_3) (1 + e_1)^{-2}]^{-1} \\ &= 1 - \lambda_2 \theta C_{20} (1 - 2e_1 + 3e_1^2 - 4e_1^3 + e_3 - 2e_1 e_3) + \lambda_2^2 \theta^2 C_{20}^2 \\ &= A, \text{ say} \end{aligned}$$

Thus we write

$$t_{sB} = \bar{y} \left( \frac{\bar{X}}{\bar{x}} \right)^\alpha (BA)$$

After some mathematical simplifications we write the expressions for the expected value and variance of  $t_{sB}$  to  $O(1/n^2)$  as

$$E(t_{sB}) = \bar{Y} + \bar{Y}\theta^2 \{C_{30}(3\lambda_2 - 1) + C_{20}^2(3 - 6\lambda_2 + \lambda_2^2) + C_{20}C_{11}(3\lambda_2 - \lambda_1\lambda_2) + C_{21}(1 - 2\lambda_1 - \lambda_2)\} \quad (16)$$

$$\begin{aligned} V(t_{sB}) = & \bar{Y}^2 \theta (\lambda_1^2 C_{20} - 2\lambda_1 C_{11} + C_{02}) + \bar{Y}^2 \theta^2 \{C_{11}^2 (4\lambda_2 - 2\lambda_1 - \lambda_1^2) + \lambda_1^2 C_{20} C_{02}\} \\ & + \bar{Y}^2 \theta^2 \{C_{20}^2 (6\lambda_4 + 6\lambda_1\lambda_3 + 6\lambda_2 - 8\lambda_1\lambda_2 - 2\lambda_1^2\lambda_2 - 6) + C_{20}C_{11} (2\lambda_1^3 + 4\lambda_1^2\lambda_2 + 2\lambda_1 - 2\lambda_2 + 2\lambda_1\lambda_2 - 12\lambda_3)\} \end{aligned} \quad (17)$$

Substituting,

$$\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 1,$$

$$E(\hat{Y}_{rB}) = \bar{Y} + \bar{Y}\theta^2 \{2C_{30} - 2C_{20}^2 + 2C_{20}C_{11} - 2C_{21}\} \quad (18)$$

$$V(\hat{Y}_{rB}) = \bar{Y}^2 [\theta(C_{20} - 2C_{11} + C_{02}) + \theta^2(C_{11}^2 + C_{20}C_{02} + 2C_{20}^2 - 4C_{20}C_{11})] \quad (19)$$

Further, following Tin [14] we have another almost unbiased ratio-type estimator given by

$$\begin{aligned} t_{sT} = & \bar{y} \left( \frac{\bar{X}}{\bar{x}} \right)^\alpha \left( 1 + \lambda_1 \theta \frac{s_{xy}}{xy} - \lambda_2 \theta \frac{s_x^2}{x^2} \right) \\ = & \bar{y} \left( \frac{\bar{X}}{\bar{x}} \right)^\alpha \left[ 1 + \lambda_1 \theta C_{11} (1 + e_0)^{-1} (1 + e_1)^{-1} (1 + e_2) - \lambda_2 \theta C_{20} (1 + e_3) (1 + e_1)^{-2} \right] \\ = & \bar{y} \left( \frac{\bar{X}}{\bar{x}} \right)^\alpha \left[ 1 + \lambda_1 \theta C_{11} (1 - e_1 + e_1^2 - e_0 + e_1 e_0 + e_0^2 + e_2 - e_1 e_2 - e_0 e_2) - \lambda_2 \theta C_{20} (1 - 2e_1 + 3e_1^2 - 4e_1^3 + e_3 - 2e_1 e_3) \right] \end{aligned} \quad (20)$$

Retaining terms up to  $O(1/n^2)$

$$\begin{aligned} E(t_{sT}) = & \bar{Y} + \bar{Y}\theta^2 \{C_{30}(-\lambda_3 + 2\lambda_2 + \lambda_1\lambda_2) + C_{20}^2(3\lambda_4 - 4\lambda_2 - 2\lambda_1\lambda_2)\} \\ & + \bar{Y}\theta^2 \{C_{20}C_{11}(-\lambda_3 + \lambda_1\lambda_2 + 2\lambda_2 + \lambda_1) + C_{21}(-\lambda_1^2 - \lambda_1)\} \end{aligned} \quad (21)$$

$$\begin{aligned} V(t_{sT}) = & \bar{Y}^2 \theta (\lambda_1^2 C_{20} - 2\lambda_1 C_{11} + C_{02}) + \bar{Y}^2 \theta^2 \{C_{11}^2 (4\lambda_2 - 2\lambda_1 - \lambda_1^2) + \lambda_1^2 C_{20} C_{02}\} \\ & + \bar{Y}^2 \theta^2 \{C_{20}^2 (6\lambda_1\lambda_3 + 2\lambda_2 - 4\lambda_1\lambda_2 - 2\lambda_1^2\lambda_2) + C_{20}C_{11} (2\lambda_1^3 + 4\lambda_1^2 + 4\lambda_2 - 4\lambda_1\lambda_2 - 10\lambda_3)\} \end{aligned} \quad (22)$$

When

$$\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 1$$

$$E(\hat{Y}_{rT}) = \bar{Y} + \bar{Y}\theta^2 (2C_{30} - 3C_{20}^2 + 3C_{20}C_{11} - 2C_{21}) \quad (23)$$

$$V(\hat{Y}_{rT}) = V(\hat{Y}_{rB}), \quad (24)$$

Note: The expressions for  $E(\hat{Y}_{rB})$ ,  $E(\hat{Y}_{rT})$ ,  $V(\hat{Y}_{rB})$ , and  $V(\hat{Y}_{rT})$  are derived by Tin [14] using bivariate cumulants and by De-Graft Johnson [6] using bivariate moments. Some of the higher order bivariate moments neglecting finite population correction factor for large finite population mentioned by De-Graft Johnson [6] are given below:

$$\begin{aligned} E(e_1^2 e_0^2) = & \frac{1}{n^2} (2C_{11}^2 + C_{20}C_{02}), \quad E(e_1^4) = \frac{3C_{20}^2}{n^2}, \quad E(e_1^3 e_0) = \frac{1}{n^2} 3C_{20}C_{11} \\ E(e_1 e_2) = & \frac{1}{n} \frac{C_{21}}{C_{11}}, \quad E(e_1 e_3) = \frac{1}{n} \frac{C_{30}}{C_{20}}, \quad E(e_0 e_2) = \frac{1}{n} \frac{C_{12}}{C_{11}}. \end{aligned}$$

## 2.1. Comparison of Bias and Variance of the Class of Beale Type and Tin Type Almost Unbiased Ratio-Type Estimators

Consider the situation when  $y$  and  $x$  follow a bivariate symmetric distribution with odd order moments being zero.

We have  $E(t_{sB}) = \bar{Y} + \bar{Y}\theta^2 [(3 - 6\lambda_2 + \lambda_2^2) C_{20}^2 + (3\lambda_2 - \lambda_1\lambda_2) C_{20}C_{11}] = \bar{Y} + \bar{Y}\theta^2 (B_1 C_{20}^2 + B_2 C_{20}C_{11})$ ,

Where  $B_1 = (3 - 6\lambda_2 + \lambda_2^2)$  and  $B_2 = (3\lambda_2 - \lambda_1\lambda_2)$ ,

and

$$\begin{aligned} E(t_{sT}) &= \bar{Y} + \bar{Y}\theta^2 [(3\lambda_4 - 4\lambda_2 + -2\lambda_1\lambda_2) C_{20}^2 + (-\lambda_3 + \lambda_1\lambda_2 + 2\lambda_2 + 2\lambda_1) C_{20}C_{11}] \\ &= \bar{Y} + \bar{Y}\theta^2 (T_1 C_{20}^2 + T_2 C_{20}C_{11}), \end{aligned}$$

Where  $T_1 = (3\lambda_4 - 4\lambda_2 + -2\lambda_1\lambda_2)$  and  $T_2 = (-\lambda_3 + \lambda_1\lambda_2 + 2\lambda_2 + 2\lambda_1)$

Hence,  $t_{sB}$  will be less biased than  $t_{sT}$

$$\text{if } \left| B_1 + B_2 \frac{\beta}{R} \right| < \left| T_1 + T_2 \frac{\beta}{R} \right| \quad (25)$$

Further,

$$V(t_{sB}) < V(t_{sT})$$

$$\text{if } \frac{\beta}{R} > \frac{6\lambda_4 + 4\lambda_2 - 4\lambda_1\lambda_2 - 6}{6\lambda_2 + 2\lambda_3 - 6\lambda_1\lambda_2 - 2\lambda_1 + 4\lambda_1^2 - 4\lambda_1^2\lambda_2}. \quad (26)$$

that is, if  $\rho > \frac{1}{2} \frac{C_x}{C_y} \frac{(6\lambda_4 + 4\lambda_2 - 4\lambda_1\lambda_2 - 6)}{(6\lambda_2 + 2\lambda_3 - 6\lambda_1\lambda_2 - 2\lambda_1 + 4\lambda_1^2 - 4\lambda_1^2\lambda_2)}$ , provided the denominator does not vanish.

## 2.2. A Special Case of Class of Almost Unbiased Ratio-Type Estimators

Consider a special case considered by Swain [13] of  $t_s$  when  $\alpha = 1/2$  of the form

$$t_{sqr} = \bar{y} \left( \frac{\bar{X}}{\bar{x}} \right)^{1/2} \quad (27)$$

This estimator is termed as square root transformation estimator by Swain [13].

Substituting  $\lambda_1 = \frac{1}{2}$ ,  $\lambda_2 = \frac{3}{8}$ ,  $\lambda_3 = \frac{5}{16}$ , and  $\lambda_4 = \frac{35}{128}$  in the expressions in (15)

Beale's almost unbiased ratio estimator is given by

$$t_{sqrB} = \bar{y} \left( \frac{\bar{X}}{\bar{x}} \right)^{1/2} \left[ \frac{1 + \frac{1}{2} \theta \frac{s_{xy}}{xy}}{1 + \frac{3}{8} \theta \frac{s_x^2}{\bar{x}^2}} \right] \quad (28)$$

and Tin's almost unbiased ratio estimator is written as

$$t_{sqrT} = \bar{y} \left( \frac{\bar{X}}{\bar{x}} \right)^{1/2} \left( 1 + \frac{1}{2} \theta \frac{s_{xy}}{xy} - \frac{3}{8} \theta \frac{s_x^2}{\bar{x}^2} \right) \quad (29)$$

Substituting  $\lambda_1 = \frac{1}{2}$ ,  $\lambda_2 = \frac{3}{8}$ ,  $\lambda_3 = \frac{5}{16}$ , and  $\lambda_4 = \frac{35}{128}$  in the expressions in (11), (16), (21), (12), (17), and (22) respectively,

we have

$$E(t_{sqr}) = \bar{Y} + \bar{Y} \left[ \theta \left( \frac{3}{8} C_{20} - \frac{1}{2} C_{11} \right) + \theta^2 \left( -\frac{5}{16} C_{30} + \frac{105}{128} C_{20}^2 + \frac{3}{8} C_{21} - \frac{15}{16} C_{20}C_{11} \right) \right] \quad (30)$$

$$E(t_{sqrB}) = \bar{Y} + \bar{Y}\theta^2 \left[ \left( \frac{1}{8} C_{30} - \frac{3}{8} C_{21} \right) + \left( \frac{57}{64} C_{20}^2 + \frac{15}{16} C_{20}C_{11} \right) \right] \quad (31)$$

$$E(t_{sqRT}) = \bar{Y} + \bar{Y}\theta^2 \left[ \left( \frac{5}{8}C_{30} - \frac{3}{4}C_{21} \right) + \left( -\frac{135}{128}C_{20}^2 + \frac{9}{8}C_{20}C_{11} \right) \right] \quad (32)$$

$$V(t_{sqR}) = \bar{Y}^2\theta \left( \frac{1}{4}C_{20} - C_{11} + C_{02} \right) + \bar{Y}^2\theta^2 \left( \frac{39}{32}C_{20}^2 + \frac{7}{4}C_{11}^2 + C_{20}C_{02} - \frac{3}{8}C_{30} - \frac{15}{4}C_{20}C_{11} + \frac{5}{4}C_{21} - C_{12} \right) \quad (33)$$

$$V(t_{sqRB}) = \bar{Y}^2\theta \left( \frac{1}{4}C_{20} - C_{11} + C_{02} \right) + \bar{Y}^2\theta^2 \left( -\frac{183}{64}C_{20}^2 + \frac{1}{4}C_{11}^2 + \frac{1}{4}C_{20}C_{02} - \frac{5}{2}C_{20}C_{11} \right) \quad (34)$$

$$V(t_{sqRT}) = \bar{Y}^2\theta \left( \frac{1}{4}C_{20} - C_{11} + C_{02} \right) + \bar{Y}^2\theta^2 \left( \frac{3}{4}C_{20}^2 + \frac{1}{4}C_{11}^2 + \frac{1}{4}C_{20}C_{02} - \frac{9}{8}C_{20}C_{11} \right) \quad (35)$$

**Comparison of biases and variances of almost unbiased  $t_{sqR} = \bar{y} \left( \frac{\bar{X}}{\bar{x}} \right)^{1/2}$**

$t_{sqRB}$  will be less biased than  $t_{sqRT}$  to  $O(1/n^2)$  if

$$\left| \frac{57}{64} + \left( \frac{15}{16} \right) \frac{\beta}{R} \right| < \left| -\frac{135}{128} + \left( \frac{9}{8} \right) \frac{\beta}{R} \right| \quad (36)$$

$t_{sqRB}$  will be more efficient than  $t_{sqRT}$  to  $O(1/n^2)$  if

$$\frac{231}{64} + \frac{11}{8} \left( \frac{\beta}{R} \right) > 0, \quad (37)$$

which is always true since both  $\beta$  and  $R$  are positive.

### 3. NUMERICAL ILLUSTRATION

To compare Beale type and Tin type square root transformation estimators, we have considered four natural populations-1, 2, 3, and 5, having size  $N = 3164$ , whose parameters in terms of product moments are mentioned in De-Graft Johnson [6]. The biases and variances, ignoring finite population correction factor, to  $O(1/n^2)$  for different sample sizes are given in Table 1.

**Table 1** | Comparison biases and variances of Beale type and Tin type estimators for the square root transformation estimator, omitting constant multipliers.

POP-1	B( $t_{sqR}$ )	B( $t_{sqRB}$ )	B( $t_{sqRT}$ )	V( $t_{sqR}$ )	V( $t_{sqRB}$ )	V( $t_{sqRT}$ )
$\theta = 1/10$	-0.02490112	0.00045152	-0.00179427	0.15693527	0.14491935	0.14988051
$\theta = 1/20$	-0.01255322	0.00011288	-0.00044857	0.07643719	0.07343321	0.07467350
$\theta = 1/50$	-0.00504592	0.00001806	-0.00007177	0.03008757	0.02960693	0.02980538
$\theta = 1/100$	-0.00252707	0.00000452	-0.00001794	0.01496257	0.01484241	0.01489202
<b>POP-2</b>						
$\theta = 1/10$	-0.01953039	0.00154964	-0.00048017	0.16621535	0.15482812	0.16030793
$\theta = 1/20$	-0.00972207	0.00038741	-0.00012004	0.08135202	0.07850522	0.07987517
$\theta = 1/50$	-0.00387848	0.00006199	-0.00001921	0.03211945	0.03166396	0.03188316
$\theta = 1/100$	-0.00193751	0.00001550	-0.00000480	0.01598950	0.01587563	0.01593043
<b>POP-3</b>						
$\theta = 1/10$	-0.03381434	0.01896277	-0.00493000	0.10236741	0.01425384	0.08892160
$\theta = 1/20$	-0.01682702	0.00474069	-0.00123250	0.04729623	0.02526784	0.04393478
$\theta = 1/50$	-0.00671157	0.00075851	-0.00019720	0.01798550	0.01446095	0.01744766
$\theta = 1/100$	-0.00335258	0.00018963	-0.00004930	0.00883725	0.00795611	0.00870279
<b>POP-5</b>						
$\theta = 1/10$	-0.01751400	0.00049767	-0.00041688	0.17321951	0.16626384	0.16834761
$\theta = 1/20$	-0.00874797	0.00012442	-0.00010422	0.08530856	0.08356965	0.08401559
$\theta = 1/50$	-0.00349702	0.00001991	-0.00001668	0.03381114	0.03353291	0.03360626
$\theta = 1/100$	-0.00174815	0.00000498	-0.00000417	0.01685352	0.01678397	0.01680480

#### Comments:

For all populations, Beale's estimator is more efficient than Tin's estimator for all sample sizes under consideration. For population 1 Beale's estimator is less biased than Tin's estimator, but for populations 2, 3, and 5 Tin's estimator is marginally less biased than Beale's estimator.

Thus, Beale's estimator appears to be a preferred estimator over Tin's estimator for the square root transformation estimator with regard to bias and efficiency.

## 4. CONCLUSIONS

Almost unbiased Tin type and Beale type estimators for Srivastava's [11] class of estimators ( $t_s$ ) are derived and compared with regard to bias and efficiency. As a special case Beale's and Tin's almost unbiased estimators for Swain's [13] square root transformation estimator are formulated and compared. It is seen that Beale's estimator is conditionally less biased than Tin's estimator, but interestingly is more efficient than Tin's estimator to  $O(1/n^2)$ . Numerical illustrations show that Beale's estimator have better performances with regard to bias and efficiency.

## CONFLICT OF INTEREST

There is no conflict of interest involved and the research was carried out with authors's own contribution without any outside funding.

## REFERENCES

1. E.M.L. Beale, *Ind. Organ.* 31 (1962), 27–28.
2. A.L. Carriquiry, W.A. Fuller, J.J. Goneyeche, K.W. Dodd, Estimation of the Usual Intake of Distributions of Ratios of Dietary Supplements, Research Report – foe Agricultural Research Service, Department of Agriculture, Department of Statistics, Iowa State University, Iowa, USA, 1995.
3. W.G. Cochran, *J. Agric. Sci.* 30 (1940), 262–275.
4. W.G. Cochran, *Sampling Techniques*, John Wiley and Sons, New York, USA, 1977.
5. I.P. David, Contribution to Ratio Method of Estimation, Ph.D. Dissertation, Iowa State University, Ames, IA, USA, 1971.
6. K.T. De-Graft Johnson, Some Contributions to the Theory of Two Phase Sampling, a Dissertation Submitted for the Degree of Doctor of Philosophy, Iowa State University, Ames, IA, USA, 1969.
7. L.A. Goodman, H.O. Hartley, *J. Am. Stat. Assoc.* 53 (1958), 491–508.
8. C.J. Lee, R.M. Hirsch, G.E. Schwarz, D.J. Holtschlag, S.D. Preston, C.G. Crawford, A.V. Vecchia, *J. Hydrol.* 542 (2016), 185–203.
9. R.P. Richards, J. Holloway, *Water Resour. Res.* 23 (1987), 1339–1348.
10. R.P. Richards, Estimation of Pollutant Loads in Rivers and Streams: A Guidance Document for NPS Programs, Water Quality Laboratory, Heidelberg College, Project Report Prepared under Grant from US Environment Protection Agency, OH, USA, 1998.
11. S.K. Srivastava, *Calcutta Stat. Assoc. Bull.* 16 (1967), 121–132.
12. P.V. Sukhatme, B.V. Sukhatme, C. Asok, *Sampling Theory in Surveys with Applications*, Iowa State University Press, Ames, IA, USA, 1984.
13. A.K.P.C. Swain, *Revista Investigacion Operacional.* 35 (2014), 49–57.
14. M. Tin, *J. Am. Stat. Assoc.* 60 (1965), 294–307.