Pak. J. Statist.
2009 Vol. 25(2), 165-194

SOME IMPROVED VARIANCE ESTIMATORS FROM A BIVARIATE NON-NORMAL POPULATION

Pulakesh Maiti and **T.P. Tripathi**
Indian Statistical Institute, Kolkata, India.
Email: pulakesh@isical.ac.in

ABSTRACT

Given paired observations, $\{(x_i, y_i); i = 1, 2, ..., n\}$ on two variables x and y for a random sample s, from some bivariate non-normal population like bivariate gamma, beta-stacy which are of much use in modelling data obtained in Physical, Social and Life-Sciences. This paper considers an improvement of the customary estimator of population variance. A mixture (i.e. a weighted combination) of the customary estimator of the variance and a suitably chosen statistic t is proposed. It is also indicated that under some conditions for a broad range of the values of the mixing constants, the improvement in the sense of having a smaller mean square error, over the traditional estimator is possible.

KEYWORDS AND PHRASES

Paired observations; estimation of variance; searles estimators; bivariate gamma population; beta-stacy population.

1. INTRODUCTION

In the statistical literature, it is well demonstrated that estimation of variance will be as important as estimating means, even may be more. Here, in this paper, improved estimators for population variance σ_y^2 of a variable y of bivariate gamma and beta-staey population have been considered, when we have information on both the variables y and x available in the form of paired observations $\{x_i, y_i\}$ only. The proposed class of estimators is as follows.

$$d = \lambda_1 s_y^2 + \lambda_2 t , \qquad (1.1)$$

where $s_y^2 = \sum_{i=1}^{n} \left(y_i - \bar{y}\right)^2 \Big/ \left(n-1\right)$ is an unbiased estimator of σ_y^2 and t is a suitably chosen statistic based on sample values of x alone or on sample values of both the variables x and y.

The motivation for the form of the proposed estimators in (1.1) arises from recognising the fact that in many real life situations, there may exist a functional relationship between the variables x and y. For example, in Rain-storm, according to

Etoh and Murota (1986), duration x, maximum intensity (y) and total amount z$\left(\alpha \frac{xy}{2}\right)$ have a gamma distribution. It is also observed that $y = \eta x^{\alpha}$.

We discuss some of the situations when a specific t would appear to be more appropriate than any other *t's*. Suppose the scatter diagram reveals approximately a linear relationship between y and x i.e.,

$$\text{y} = \alpha + \beta x,$$

then an estimator of σ_y^2, can be taken as

$$\hat{\sigma}_y^2 = k_1 s_x^2 \text{ (say)},$$

and with the choice of a t = $k_1 s_x^2$, the appropriate class would be

$$\text{d } \left(\lambda_1, \lambda_2; s_y^2, s_x^2\right) = \lambda_1 s_y^2 + \left(\lambda_2^* k_1\right) s_x^2 = \lambda_1 s_y^2 + \lambda_2 s_x^2.$$

The choice of a t is motivated by the relationship of the parameter $\theta = \sigma_y^2$ with another moment ξ which is related with the variance σ_y^2 through some relationship of the form $\sigma_y^2 = k\,\xi$. Motivating the choice of a t as $\hat{\xi}$, an estimate of ξ, estimation of σ_y^2 has been considered.

Similarly, if it is expected even distantly, $\mu_y \simeq \mu_x$ or $\sigma_y \simeq \sigma_x$ or $\sigma_y^2 \simeq \mu_x$ or $\sigma_x^2 \simeq \mu_x$ hold in some situations, then in such cases, following choices of t can be suggested.

$$t(y,x) = \begin{cases} \left(\bar{y} - \bar{x}\right), & \textit{if } \mu_y \simeq \mu_x \\ \left(s_y^2 - s_x^2\right), & \textit{if } \sigma_y \simeq \sigma_x \\ \left(s_y^2 - \bar{x}\right), & \textit{if } \sigma_y^2 \simeq \mu_x \\ \left(s_x^2 - \bar{x}\right), & \textit{if } \sigma_x^2 \simeq \mu_x \end{cases} \tag{1.2}$$

As an illustration one may observe that for the bivariate gamma population,

$$f(x,y) = \frac{a^{p+q}}{\Gamma p \, \Gamma q} x^{p-1} (y-x)^{q-1} e^{-ay}, \; 0<x<y<\propto, \; a>0, \; p>0, \; q>0. \tag{1.3}$$

If 0<p, q<1, p+q = 1 and p $\simeq$ 1 then one would have, (i) $\sigma_y \simeq \sigma_x$ and (ii) $\mu_y \simeq \mu_x$. Similarly for, a $\simeq$ 1, ap $\simeq$ 1 together with the condition in (1.3), we would have, (iii) $\sigma^2(x) = \mu_x$ and (iv) $\sigma^2(y) = \mu_y$ respectively.

The organization of the paper is as follows. In Section 2, some general results for estimation of any parameter θ have been provided followed by section 3, where in particular, the problem of estimation of σ_y^2 has been considered. In section 4, improvement over Searles- type estimator has been made through utilization of a pair of observations $\left\{\left(x_i, y_i\right); 1, 2, \ldots n\right\}$ for any bivariate population. In Section 5, we have considered observations as if drawn from a bivariate gamma and a beta-stacy population. In Section 6, we examine the superiority of the proposed estimators through a real life data which is supposed to be a realization from a bivariate gamma population.

2. SOME GENERAL RESULTS FOR ESTIMATION OF ANY PARAMETER θ

Let $\hat{\theta}$ be an unbiased estimator for the parameter θ. The generalised Searles estimator for θ may be defined as

$$T_1 = \lambda_1 \hat{\theta}, \tag{2.1}$$

where λ_1 is a suitably chosen constant. We have

$$M \left(T_1\right) = \theta^2 \left[\lambda_1^2 \left(1 + C^2\left(\hat{\theta}\right)\right) - 2\lambda_1 + 1\right]. \tag{2.2}$$

The natural question arises: For what choice of a λ_1, the estimator T_1 is better than $\hat{\theta}$ and what could be the best choice of λ_1? To answer this, we have the following.

Theorem 1:

For $T_1 = \lambda_1 \hat{\theta}$, the optimum choice of λ_1, which minimises mean square error of T_1 and the minimum mean square error are given respectively by

$$\lambda_{01} = 1 \Big/ \left[1 + C^2\left(\hat{\theta}\right)\right]$$

and

$$M_0\left(T_1\right) = \theta^2 . C^2\left(\hat{\theta}\right) \Big/ \left[1 + C^2\left(\hat{\theta}\right)\right]. \tag{2.3}$$

A sufficient condition for T_1 to be better than $\hat{\theta}$ can be obtained by taking a λ_1 such that

$$\left[1 + C_{(1)}^2\left(\hat{\theta}\right)\right]^{-1} \le \lambda_1 < 1, \tag{2.4}$$

where $C_{(1)}^2\left(\hat{\theta}\right) \left(\le C^2\left(\hat{\theta}\right)\right)$ is a value known apriori and $C^2\left(\hat{\theta}\right)$ is the square of coefficient of variation of $\hat{\theta}$.

Proof:

Minimising $M\left(T_1\right)$ in (2.2) with respect to λ_1, the results in (2.3) follow. Comparing $M\left(T_1\right)$ with $V\left(\hat{\theta}\right)$, it may be shown that T_1 would be better than $\hat{\theta}$ for all λ_1 satisfying

$$\left[1-C^2\left(\hat{\theta}\right)\right]\Big/\left[1+C^2\left(\hat{\theta}\right)\right]\le\lambda_1<1,$$

and hence a sufficient condition as in (2.4) follows.

It is interesting to note that use of $C_{(1)}^2\left(\hat{\theta}\right)$ for $C^2\left(\hat{\theta}\right)$ in λ_{01} still helps T_1 perform better than $\hat{\theta}$, but the use of $C_{(2)}^2\left(\hat{\theta}\right)$ $\left(\ge C^2\left(\hat{\theta}\right)\right)$ would not preserve this property of T_1.

Next we consider the problem of generating estimators better than $\hat{\theta}$ as well as T_1 through a class of weighted estimators defined by

$$d\ \left(\lambda_1,\lambda_2\right)=\left\{d: d=\lambda' v\right\}, \tag{2.5}$$

where $v'=\left(\hat{\theta},t\right)$ and $\lambda'=\left(\lambda_1,\lambda_2\right), \hat{\theta}$ being an unbiased estimator for θ and *t*, being a suitably chosen statistic such that σ_t^2 exists and λ_1,λ_2 being suitably chosen constants.

It may be shown that

$$M(d)=\lambda' G\lambda-2\theta\ \lambda'\psi+\theta^2 \tag{2.6}$$

where,

$$G=\begin{pmatrix} E\left(\hat{\theta}^2\right), & E\left(\hat{\theta}\ t\right) \\ E\left(\hat{\theta}\ t\right), & E\ \left(t^2\right)\end{pmatrix}; \qquad \psi'=\left(\theta, E\ (t)\right).$$

To find the estimators better than $\hat{\theta}$ as well as T_1, we have the following.

Theorem 2:

The optimum value of λ, say λ_0 which minimises *M(d)*, the mean square error of *d*, would be a solution of

$$G\lambda_0=\theta\psi \tag{2.7}$$

and min MSE , would be

$$M_0(d)=\theta^2\left[1-\psi'\ \left(\bar{G}\right)'\ \psi\right],$$

where, $\bar{G}$ is a g-inverse of the matrix G.

Proof:

It follows from (2.6). It can be shown that (2.7) is always consistent i.e., it always yields a solution λ_0 for λ such that

$$M(d) = M_0(d) + (\lambda - \lambda_0)' G(\lambda - \lambda_0)) \geq M_0(d)\,.$$

Since the matrix *G*, in general, is a non-negative definite matrix and would be non-singular, it follows from (2.6) and (2.7) that

$$\lambda_0 = \theta\, G^{-1}\psi \text{ and } M_0(d) = \theta^2 \left[1 - \psi'\,(G^{-1})'\psi\right].$$

In case G is a positive definite matrix, $\lambda_0 = \left(\lambda_{01}, \lambda_{02}\right)'$ and $M_0(d)$ would be given by

$$\lambda_{01} = \theta \left[E\left(\hat{\theta}\right).E\left(t^2\right) - E(t).E\left(t\hat{\theta}\right)\right] \Big/ D\left(\hat{\theta}, t\right);$$

$$\lambda_{02} = \theta \left[E(t).E\left(\hat{\theta}^2\right). - E\left(\hat{\theta}\right).E\left(\hat{\theta}t\right)\right] \Big/ D\left(\hat{\theta}, t\right) \tag{2.8}$$

and

$$M_0(d) = \theta^2 \left[1 - \left\{N\left(\hat{\theta}, t\right) \Big/ D\left(\hat{\theta}, t\right)\right\}\right], \tag{2.9}$$

where,

$$D\left(\hat{\theta}, t\right) = E\left(\hat{\theta}^2\right) E\left(t^2\right) - \left(E\left(\hat{\theta}t\right)\right)^2$$

$$= \theta^2 \left(E(t)\right)^2 \left[\left(1 - \rho^2_{\hat{\theta},t}\right) C^2\left(\hat{\theta}\right) C^2(t) + C^2\left(\hat{\theta}\right) + C^2(t) - 2\rho_{\hat{\theta},t} C\left(\hat{\theta}\right) C(t)\right]; \tag{2.10}$$

$$N\left(\hat{\theta}t\right) = \theta^2 \left(E(t)\right)^2 \left[C^2\left(\hat{\theta}\right) - 2\rho_{\hat{\theta},t} C\left(\hat{\theta}\right) C(t) + C^2(t)\right];$$

$C(t)$ = Coefficient of Variation of *t*, and

$\rho_{\hat{\theta},t}$ = Correlation Coefficient between $\hat{\theta}$ and t .

However, in practice, λ_0 would not be known, as it may depend upon a number of parameters, including sometimes, even the parameter θ itself. Therefore in the absence of exact knowledge of λ_0, our approach is to improve $\hat{\theta}$ through a T_1 and then T_1 through an estimator of the type $d = T_1 + \lambda_2 t$ by appropriate choice of a constant depending on λ_1 and a specific *t* with $V(t) < \infty$. This has been possible because of the following representation of *M (d),* i.e., *M (d)* being possible to split itself into M (T_1) as

$$M(d) = M\left(T_1\right) + \lambda_2^2\, E\left(t^2\right) - 2\lambda_2\theta\, E(t)\left\{\left(1 - \lambda_1\right) - \lambda_1 \rho_{\hat{\theta},t} C\left(\hat{\theta}\right).C(t)\right\}. \tag{2.11}$$

The above idea of improving $\hat{\theta}$ through T_1 and then T_1 through a $d = T_1 + \lambda_2 t$ can be implemented through the following.

Theorem 3:

For a given λ_1 as in T_1, an estimator $d\left(\lambda_1, \lambda_2\right)$ would be better than T_1,

$$\text{iff } \lambda_2 \text{ lies between 0 and } 2\lambda_{02}^{*}, \tag{2.12}$$

where,

$$\lambda_{02}^{*} = \left[\left(1-\lambda_1\right)\theta\, E\,\left(t\right) - \lambda_1 Cov\,\left(\hat{\theta}, t\right)\right] \Big/ E\left(t^2\right)$$

$$= \theta\, E\left(t\right)\left[\left(1-\lambda_1\right) - \lambda_1 \rho_{\hat{\theta}t,t}\, C\left(t\right).C\left(\hat{\theta}\right)\right] \Big/ E\left(t^2\right),$$

is the optimum choice of λ_2 for a specific λ_1 in T_1.

Proof:

The result follows from (2.11). Obviously, the resulting MSE of *d*, in this case, would be

$$M_0(d) = M\left(T_1\right) - \lambda_{02}^{*2}\, E\left(t^2\right).$$

Therefore, it is noted that $\hat{\theta}$ may be improved through T_1 which, in turn, could be improved through *d*, even if $\hat{\theta}$ is uncorrelated with *t*.

It may be observed that for a given λ_1.

$$\lambda_{02}^{*} = \theta\, E\left(t\right)\left[1 - \lambda_1\left(1 + \rho_{\hat{\theta},t} .C\left(\hat{\theta}\right) C\left(t\right)\right)\right] \Big/ E\left(t^2\right).$$

To ensure the non-negatively of the estimator $d\left(\lambda_1, \lambda_2\right)$ for the non-negative parameter θ, we should avoid taking that *t* for which

$$\theta\, E\,(t) > 0 \text{ and } \lambda_1\left(1 + \rho_{\hat{\theta},t} .C\left(\hat{\theta}\right).C(t)\right) > 1,$$

as, in this case λ_{02}^{*} in (2.12) would be negative and $d\left(\lambda_1, \lambda_2\right)$ may also sometimes turn out to be negative.

A subclass of *d*, say $d\left(\lambda_1 = 1, \lambda_2 = \lambda_2'\right) = \hat{\theta} + \lambda_2' t$ may be quite interesting in some situations to generate estimators better than $\hat{\theta}$. Comparing $M\left[d\left(\lambda_1 = 1, \lambda_2 = \lambda_2'\right)\right]$ with $V(\hat{\theta})$, a sufficient condition for $d\left(\lambda_1 = 1, \lambda_2 = \lambda_2'\right)$ to be better than $\hat{\theta}$ would be either

$$0 < \lambda_2' < 2\lambda_{02}' \text{ in case } \rho_{\hat{\theta},t} < 0$$

or

$$2\lambda'_{02} < \lambda'_2 < 0 \text{ in case } \rho_{\hat{\theta},t} > 0 , \qquad (2.13)$$

where, $\lambda'_{02} = -Cov\left(\hat{\theta}, t\right) \big/ E\left(t^2\right)$. Therefore, for the situation $\lambda_1 = 1$, one should never choose a *t* to be uncorrelated with $\hat{\theta}$, as for such choices of *t*, $\hat{\theta}$ would be uniformly better than $d\left(\lambda_1 = 1, \lambda_2 = \lambda'_2\right)$. In practice, in the absence of exact knowledge of λ'_{02}, a set of sufficient conditions for $d\left(\lambda_1 = 1, \lambda_2 = \lambda'_2\right)$ to be better than $\hat{\theta}$ would be

$$0 < \lambda'_2 < 2\rho^* K_{(1)} \left(C^2_{(1)}(t) \Big/ \left(1 + C^2_{(2)}(t)\right)\right), \text{ in case } \rho_{\hat{\theta},t} < 0$$

or

$$-2\rho^* K_{(1)} \left(C^2_{(1)}(t) \Big/ \left(1 + C^2_{(2)}(t)\right)\right) < \lambda'_2 < 0 , \text{ in case } \rho_{\hat{\theta},t} > 0 . \qquad (2.14)$$

3. SOME GENERAL RESULTS FOR ESTIMATION OF σ_y^2

Let $\hat{\sigma}_y^2$ be an unbaised estimator for σ_y^2 based on any sampling design and let a Searles-type estimator be defined as

$$\mathrm{T}_1 = \lambda_1 \hat{\sigma}_y^2 , \qquad (3.1)$$

where λ_1 is a suitably chosen constant minimising mean square error of T_1. This optimal estimator T_{01} is an improvement over $\hat{\sigma}_y^2$, since

$$\mathrm{V}\left(\hat{\sigma}_y^2\right) - M_0\left(\mathrm{T}_1\right) = \sigma_y^4 \left[C^2\left(\hat{\sigma}_y^2\right) - \frac{C^2\left(\hat{\sigma}_y^2\right)}{\left(1 + C^2\left(\hat{\sigma}_y^2\right)\right)} \right] \geq 0 , \qquad (3.2)$$

where $C^2\left(\hat{\sigma}_y^2\right)$ is the square of coefficient of variation of $\hat{\sigma}_y^2$.

But the optimum estimator T_{01} can never be used in practice, unless the optimum choice of λ_1 in T_1, namely,

$$\lambda_{01} = 1 \Big/ \left[1 + C^2\left(\hat{\sigma}_y^2\right) \right] , \qquad (3.3)$$

is known exactly. However, in the absence of exact knowledge of $\lambda_{o1}, \hat{\sigma}_y^2$ can still be improved through the estimators of the type(3.1) by choosing a λ_1 such that

$$\left[1 + C^2_{(1)}\left(\hat{\sigma}_y^2\right)^{-1} \right] \leq \lambda_1 < 1, \qquad (3.4)$$

where, $C^2_{(1)}\left(\hat{\sigma}^2_y\right)\left(\le C^2\left(\hat{\sigma}^2_y\right)\right)$ is a value known apriori.

T_1 in (3.1) with a λ_1 satisfying (3.4) can further be improved through an estimator of the type

$$d = \mathrm{T}_1 + \lambda_2 t\,, \tag{3.5}$$

where t is a suitably chosen statistic and λ_2 is a value lying between

$$0 \text{ and } 2\lambda^*_{02}\,, \tag{3.6}$$

λ^*_{02} being the optimum choice of λ_2 and given by

$$\lambda^*_{02} = \left[\left(1-\lambda_1\right)\sigma^2_y E\left(t\right) - \lambda_1 Cov\left(\hat{\sigma}^2_y, t\right)\right] \Big/ E\left(t^2\right). \tag{3.7}$$

This follows immediately by observing that

$$M(d) = M\left(T_1\right) + \lambda^2_2 E\left(t^2\right) - 2\lambda_2\left\{\left(1-\lambda_1\right)\sigma^2_y E\left(t\right) - \lambda_1 Cov\left(\hat{\sigma}^2_y, t\right)\right\},$$

and

$$M_0\left(d\right) = M\left(d\right)\Big|_{\lambda_2=\lambda^*_{02}} = M\left(\mathrm{T}_1\right) - \lambda^{*2}_{o2} E\left(t^2\right). \tag{3.8}$$

Therefore, the procedure to improve $\hat{\sigma}^2_y$, the usual unbiased estimator for σ^2_y would be as follows:

(i) $\hat{\sigma}^2_y$ is first improved through a T_1 or through the T_{01}, and then (ii) T_1 or T_{01} is improved through an estimator of the type in (3.5).

4. IMPROVED ESTIMATORS OF VARIANCES OF BIVARIATE POPULATION

Let a random sample of size n yield the paired observations $\left\{(y_i, x_i);\ i = 1, 2, \ldots.n\right\}$ and let

$$\mathrm{T}_1 = \lambda_1 s^2_y,$$

be the Searles - type estimator for σ^2_y. It is found that optimum choice of λ_1 minimising $\mathrm{MSE}\left(\mathrm{T}_1\right)$ is

$$\lambda_{01} = n\big/\Delta\,,$$

with $\Delta = \beta_2\left(y\right) + \dfrac{n^2-2n+3}{\left(n-1\right)}$, $\beta_2\left(y\right)$ being the coefficient of Kurtosis of y.

Let $\lambda_1^* = \left(n / \Delta^*\right)$, where Δ^* is such that $n \le \Delta^* \le \Delta$ and $\mathrm{T}_1^* = \lambda_1^* s_y^2$ be the shrinkage type estimator for σ_y^2. We now define several estimators for σ_y^2 as follows.

$$d_1 = d\left(\lambda_1^*, \lambda_2; s_y^2, \bar{x}\right) = \mathrm{T}_1^* + \lambda_2 \bar{x} \tag{4.1}$$

$$d_2 = d\left(\lambda_1^*, \lambda_2; s_y^2, s_x^2\right) = \mathrm{T}_1^* + \lambda_2 s_x^2 \tag{4.2}$$

$$d_3 = d\left(\lambda_1^*, \lambda_2; s_y^2, \left(s_y^2 - s_x^2\right)\right) = \mathrm{T}_1^* + \lambda_2 \left(s_y^2 - s_x^2\right) \tag{4.3}$$

$$d_4 = d\left(\lambda_1^* = 1, \lambda_2; s_y^2, \bar{x}\right) = s_y^2 + \lambda_2 \bar{x} \tag{4.4}$$

$$d_5 = d\left(\lambda_1^* = 1, \lambda_2; s_y^2, s_x^2\right) = s_y^2 + \lambda_2 s_x^2 \tag{4.5}$$

$$d_6 = d\left(\lambda_1^* = 1, \lambda_2; s_y^2, \left(s_y^2 - s_x^2\right)\right) = s_y^2 + \lambda_2 \left(s_y^2 - s_x^2\right) \tag{4.6}$$

It may be observed that the estimators for the variance

$$d_i = \lambda_i s_y^2,\ i = 1, 2, \ldots, 6 \tag{4.7}$$

due to the Das and Tripathi (1978) are based on the availability of the knowledge on mean, variance and co-efficient of variation of an auxiliary variate x and λ_i's have been made to be dependent on the sample estimates as well as on the above parameters. In our case, in the absence of the exact knowledge on the parameter involved in λ_0, the optimum choice of λ, a useable λ has been obtained using the knowledge on some bounds of the parameter involved in λ_0 and an improved estimator over s_y^2 has been obtained.

The following proposition exhibits the range of values of λ_2, such that for a given value of λ_1, the estimators in (4.1) to (4.3) will be improvements over T_1^* and hence over s_y^2 also.

Proposition 4.1:

A set of necessary and sufficient conditions for the estimators in (4.1) to (4.3) to be better than T_1^* would be that λ_2 lies between 0 and $2\lambda_{o2}^*$, where,

$$\lambda_{o2}^* = \begin{cases} \left[\left(1-\lambda_1^*\right)\sigma_y^2 E\left(\bar{x}\right) - \lambda_1^* Cov\left(s_y^2, \bar{x}\right)\right] \Big/ E\left(\bar{x}^2\right),\ t = \bar{x} \\ \left[\left(1-\lambda_1^*\right)\sigma_y^2 E\left(s_x^2\right) - \lambda_1^* Cov\left(s_y^2, s_x^2\right)\right] \Big/ E\left(s_x^2\right)^2,\ t = s_x^2 \\ \left[\left(1-\lambda_1^*\right)\sigma_y^2 E\left(s_y^2 - s_x^2\right) - \lambda_1^* \left\{V\left(s_y^2\right) - Cov\left(s_y^2, s_x^2\right)\right\}\right] \Big/ E\left(s_y^2 - s_x^2\right)^2,\ t = s_y^2 - s_x^2 \end{cases} \tag{4.8}$$

with $E\left(s_x^2\right)^2 = \frac{\sigma_x^4}{n}\left[\beta_2 + \frac{\left(n^2-2n+3\right)}{(n-1)}\right]$

$$V\left(s_x^2\right) = \frac{\sigma_x^4}{n}\left[\beta_2(x) - \frac{(n-3)}{(n-1)}\right];\ V\left(s_y^2\right) = \frac{\sigma_y^4}{n}\left[\beta_2(y) - \frac{(n-3)}{(n-1)}\right] \tag{4.9}$$

$$Cov(\bar{x},\bar{y}) = p\sigma_x\sigma_y/n\ ;\ Cov\left(s_y^2,\bar{y}\right) = \mu_3(y)/n\ ;$$

$$\begin{aligned} Cov\left(s_y^2,\bar{x}\right) = \frac{1}{n}\Big[&\mu'_{1,2}(x,y) + \mu'_{1,0}(x,y)\Big\{(n-1)\mu'_{0,2}(x,y) \\ &-(n-2)\mu'^2_{0,1}(x,y) - 2\mu'_{1,1}(x,y)\mu'_{01}(x,y) - n\sigma_y^2\mu_x\Big\}\Big] \end{aligned}$$

$$= \frac{1}{n}\left[\mu'_{1,2}(x,y) + \mu_x\left\{(n-1)\mu'_2(y) - (n-2)\mu_y^2\right\} - 2\mu'_{1,1}(x,y)\mu_y - n\sigma_y^2\mu_x\right]$$

$$\begin{aligned} Cov\left(s_y^2,s_x^2\right) = \frac{1}{n}\Bigg[&\mu'_{2,2}(x,y) + (n-1)\mu'_2(x)\mu'_2(y) - 2\mu'_{1,2}(x,y)\mu_x \\ &-(n-2)\mu_x^2\mu'_2(y) - 2\mu'_{2,1}(x,y)\mu'_y - (n-2)\mu'_2(x)\mu_y^2 \\ &+ 4\left(\frac{n-2}{n-1}\right)\mu'_{1,1}(x,y)\mu_x\mu_y + \frac{2}{(n-1)}\mu'^2_{1,1}(x,y) \\ &+ \frac{(n-2)(n-3)}{(n-1)}\mu_x^2\mu_y^2 - n\mu_2(x)\mu_2(y)\Bigg] \end{aligned}$$

and $\mu'_{r,s} = E\left(x^r y^s\right)$

Proof:

It may be noted from (3.8) that M(d) was decomposed as

$$\text{M(d)} = M\left(\text{T}_1^*\right) + \lambda_2^2 E\left(t^2\right) - 2\lambda_2\left\{\left(1-\lambda_1^*\right)\sigma_y^2 E(t) - \lambda_1^* Cov\left(\hat{\sigma}_y E(t)\right)\right\}$$

and hence, the set of necessary and sufficient condition follows.

After routine calculations, one will obtain the above expressions in (4.9); [Please see Appendix A1.1 to A1.6]

Corollary 4.2:

A set of necessary and sufficient conditions for the estimators in (4.4) to (4.6) to be better than s_y^2 would be that, the corresponding λ_2 lies between 0 and $2\lambda_{02}^*$ of (4.8) with λ_1^* being replaced by 1 i.e., corresponding λ_2 should lie between 0 and $2\lambda_{02}^*$, where,

$$\lambda_{02}^{*} = \begin{cases} -Cov\left(s_y^2, \bar{x}\right) \Big/ E\left(\bar{x}^2\right) \\ -Cov\left(s_y^2, s_x^2\right) \Big/ E\left(s_x^2\right)^2 \\ -\left\{V\left(s_y^2\right) - Cov\left(s_y^2, s_x^2\right)\right\} \Big/ E\left(s_y^2 - s_x^2\right)^2 \end{cases} \tag{4.10}$$

This follows from the class defined in (2.13) and also from the expression in (4.8).

Remark 4.3:

i) Neither the optimum estimators in (4.1) to (4.6), nor the intervals of preference i.e., the interval between 0 and $2\lambda_{02}^{*}$ can be of any practical use, unless the exact values of the parameters involved in λ_{02}^{*} may be known exactly.

ii) In such situations, depending on some prior information on the bounds of the parameters involved in λ_{02}^{*}, if available, the interval of preference i.e., the interval between 0 and $2\lambda_{02}^{*}$ can be shrinked to the interval 0 and $2\lambda_{02}^{*(1)}$, where $\lambda_{02}^{*(1)}\left(\le \lambda_{02}^{*}\right)$ would be a quantity known apriori.

Therefore in the absence of exact knowledge of λ_{02}^{*}, a sufficient condition for an estimator d of the types in (4.1) to (4.6) to be better than T_1^{*} would be that corresponding λ_2 should satisfy

$0 < \lambda_2 \le 2\lambda_{02}^{*(1)}$, if $0 < \lambda_{02}^{*(1)} \le \lambda_{02}^{*}$

or $2\lambda_{02}^{*(1)} \le \lambda_2 < 0$, if $\lambda_{02}^{*(1)} \le \lambda_{02}^{*} < 0$.

5. RESULTS FOR SOME BIVARIATE NON-NORMAL POPULATION

5.1 Bivariate Gamma Population:

We study the properties of the proposed estimators for the following bivariate gamma population due to MC-Kay,

$$f\left(x, y\right) = \frac{a^{p+q}}{\Gamma p\,\Gamma q} x^{p-1}\left(y-x\right)^{q-1} e^{-ay}, 0 < x < y < \infty , (a, p, q>0) \tag{5.1.1}$$

we obtain, from Appendix A2.1

$$\mu'_{r,s} = E\left(x^r y^s\right) = \frac{\beta\left(r+p,q\right)\Gamma(s+q+r+p)}{a^{s+r}\,\Gamma p\,\Gamma q} .$$

Hence, $\mu_x = p/a, \mu_2\left(x\right) = p/a^2; \mu_3\left(x\right) = 2p/a^3, \mu_4\left(x\right) = 3p\left(p+2\right)/a^4;$

$$\mu_y = \left(p+q\right)/a, \mu_2\left(y\right) = \left(p+q\right)/a^2,$$

$$\mu_3(y) = 2(p+q)/a^3, \mu_4(y) = 3(p+q)(p+q+2)/a^4;$$

$$E\left(\bar{x}^2\right) = p(1+p)/a^2; Cov(\bar{x},\bar{y}) = p/na^2; Cov\left(s_y^2,\bar{x}\right) = 2p/na^3 \qquad (5.1.2)$$

$$Cov\left(s_y^2,\bar{x}\right) = 2(p+q)/na^3; \beta_2(x) = 3(p+2)/p; \beta_2(y) = 3(p+q+2)/(p+q)$$

For n such that $\frac{n}{n-1} \simeq 1$ and $\frac{n+1}{n-1} \simeq 1$, we have,

$$Cov\left(s_y^2, s_x^2\right) = \frac{1}{na^4}\left[2p^2\left(\frac{n}{n-1}\right)+6p\right] = \frac{2p(p+3)}{na^4};$$

$$V\left(s_y^2\right) = \frac{2(p+q)}{na^4}\left[(p+q)\left(\frac{n}{n-1}\right)+3\right] = \frac{2(p+q)(p+q+3)}{na^4};$$

$$V\left(s_x^2\right) = \frac{2p}{na^4}\left[p\left(\frac{n}{n-1}\right)+3\right] = \frac{2p(p+3)}{na^4} \qquad (5.1.3)$$

$$E\left(s_x^2\right)^2 = \frac{p}{na^4}\left[\frac{(n+1)}{(n-1)}np+6\right] = \frac{p}{na^4}[np+6];$$

$$E\left(s_y^2\right)^2 = \frac{(p+q)}{na^4}\left[\frac{(n+1)}{(n-1)}n(p+q)+6\right] = \frac{p+q}{na^4}\left[n(p+q)+6\right].$$

Proposition 5.1.1.

For bivariate gamma population as in 5.1.1, $0 < p,q < 1, \frac{n}{n-1} \simeq 1$, $\frac{n+1}{n-1} \simeq 1$, a set of necessary and sufficient conditions for the estimators in (4.1) to (4.3) to be better than T_1^* for a specified λ_1^* and hence, than s_y^2 also would be that the corresponding λ_2 lies between 0 and $2\lambda_{02}^*$, where,

$$\lambda_{02}^* = \begin{cases} \dfrac{\left[\left(1-\lambda_1^*\right)(p+q)-\left(2\lambda_1^*/n\right)\right]}{a\left(p+\dfrac{1}{n}\right)}, & for\ d_1 \qquad (5.1.4) \\ \dfrac{1}{\left(p+\dfrac{6}{n}\right)}\left[(p+q)-\lambda_1^*\left\{(p+q)+\dfrac{2(p+3)}{n}\right\}\right], & for\ d_2 \qquad (5.1.5) \\ \left[(p+q)-\lambda_1^*\left\{(p+q)+\theta\right\}\right]/(q+\theta), & for\ d_3 \qquad (5.1.6) \end{cases}$$

where $\theta = 2(2p+q+3)/n$

Proof:

From (4.8), (4.9), (5.1.2), (5.1.3) and Appendix [A2.3], the result follows.

Corollary 5.1.2.

From (5.1.4), (5.1.5) and (5.1.6), it follows that a set of necessary and sufficient conditions for the estimators in (4.4) to (4.6) to be better than T_1^* as well as s_y^2 would be that corresponding λ_2 should lie between 0 and $2\lambda_{02}^*$, where

$$\lambda_{02}^* = \begin{cases} -2/a(np+1), & \text{for } d_4 \\ -2(p+3)/(np+6), & \text{for } d_5 \\ -\theta/(q+\theta), & \text{for } d_6 \end{cases} \tag{5.1.7}$$

Thus, in the absence of exact knowledge on the parameters p,q, but depending on some bounds $\left(p \le p^{(2)}, q \le q^{(2)}\right)$, a set of sufficient conditions for estimators in (4.4) to (4.6) to be better than T_1^* can be obtained.

Corollary 5.1.3.

One can observe that λ_{02}^*'s in (5.1.4), (5.1.5) and (5.1.6) will be greater than 0 if respective values of n is such that $n > \frac{2}{(p+q)} \cdot \frac{\lambda_1^*}{\left(1-\lambda_1^*\right)}$, $n > \frac{2(p+3)}{(p+q)} \cdot \frac{\lambda_1^*}{1-\lambda_1^*}$ and $n > \frac{2(2p+q+3)}{(p+q)} \cdot \frac{\lambda_1^*}{\left(1-\lambda_1^*\right)}$ respectively.

Thus a set of sufficient conditions for the estimators in (4.1) to (4.3) to be better than T_1^* would be that corresponding λ_2 should lie between $0 < \lambda_2 \le 2\lambda_{02}^{*(1)}$, where $\lambda_{02}^{*(1)}$ can be obtained from λ_{02}^* on the basis of some bounds on the parameters p,q,a in the absence of them knowing exactly.

Corollary 5.1.4:

For some bivariate gamma populations with $(p+q)=1, (0<p,q<1)$, a set of necessary and sufficient conditions for the estimators in (4.1) to (4.3) to be better than T_1^* would be that corresponding λ_2 should lie between 0 and $2\lambda_{02}^*$, where,

$$\lambda_{02}^{*}=\begin{cases}\left[1-\lambda_1^{*}\left(1+\dfrac{2}{n}\right)\right]\Big/\left[\left(p+\dfrac{1}{n}\right)a\right], & \text{for } d_1\\ \dfrac{1}{\left(p+\dfrac{6}{n}\right)}\left[1-\lambda_1^{*}\left(1+\dfrac{2\left(p+3\right)}{n}\right)\right], & \text{for } d_2\\ \left[1-\lambda_1^{*}\left\{1+\dfrac{2\left(2p+q+3\right)}{n}\right\}\right]\Big/\left[\dfrac{2\left(2p+q+3\right)}{n}+q\right], & \text{for } d_3\end{cases} \tag{5.1.8}$$

Proposition 5.1.5:

Let the terms n/(n-1) and (n+1)/(n-1) be replaced by unity and the bivariate gamma population as in (5.1.1) be considered. If p+q=1, and $0< p,q <1$, then we have,

$$M_0\left(d_1\right)= M_0\left(d_2\right)=M_0\left(d_3\right).$$

Proof:

From (3.8) and (4.8) and under the conditions of the proposition, we obtain

$$\left[M\left(\mathrm{T}_1^{*}\right)-M_0\left(d\right)\right]=\begin{cases}\left(\lambda_{02}^{*}\right)^2.E\left(\overline{x}\right)^2=\dfrac{\left(1-\lambda_1^{*}\right)^2}{a^4\left(p+1\right)^2}.E\left(\overline{x}\right)^2=\dfrac{\left(1-\lambda_1^{*}\right)^2}{a^4}., & \text{for } d_1\\ \left(\lambda_{02}^{*}\right)^2.E\left(s_x^2\right)^2=\dfrac{\left(1-\lambda_1^{*}\right)^2}{p^2}.E\left(s_x^2\right)^2=\dfrac{\left(1-\lambda_1^{*}\right)^2}{a^4}, & \text{for } d_2\\ \left(\lambda_{02}^{*}\right).E\left(s_y^2-s_x^2\right)^2=\dfrac{\left(1-\lambda_1^{*}\right)^2}{\left(1-p\right)^2}.E\left(s_y^2-s_x^2\right)^2=\dfrac{\left(1-\lambda_1^{*}\right)^2}{a^4}, & \text{for } d_3\end{cases}$$

Hence the result follows.

5.2 Beta-Stacy population:

We consider the observations $\{(x_i,y_i);i=1,2,....,n\}$ as if drawn from beta-stacy population and study the properties of the proposed estimators of σ_y^2 for this population;

$$f\left(x,y\right)=\gamma x^{p-1}\left(y-x\right)^{q-1}y^{k-\gamma-\left(p+q\right)}\exp\left\{-\left(\frac{y}{\beta}\right)^{\gamma}\right\}\Big/\beta^{yk}\,\Gamma k.\beta\left(p,q\right), \tag{5.2.1}$$

$$0<x<y<\propto,\beta>0,\gamma>0,k>0,p>0,q>0$$

We obtain, from Appendix (A2.4),

$$\mu'_{r,s}(x,y) = E\left(x^r y^s\right) = \frac{\beta(r+p,q)}{\beta(p,q)} \frac{\beta^{r+s}\,\Gamma(r+s+\gamma k)/\gamma}{\Gamma k}. \qquad (5.2.2)$$

Now, for Beta-stacy population with $\gamma = 1, \frac{k+2}{k} \simeq 1, \frac{k+1}{k} \simeq 1$, we have the followings:

$$\mu_y = k\beta; \quad \mu_x = \left(\frac{p}{p+q}\right) k\beta,$$

$$\mu_2(y) = k\beta^2;\ \mu_2(x) = \frac{\beta^2 k(k+1)p}{(p+q+1)}\left[\frac{p+1}{p+q+1} - \left(\frac{k}{k+1}\right).\frac{p}{p+q}\right] \simeq \frac{\beta^2 k^2 pq}{(p+q)^2(p+q+1)},$$

$$\mu_3(y) = 2k\beta^3;\ \mu_3(x) \simeq \frac{2pq(q-p)k^3\beta^3}{(q+p)^3(p+q+1)(p+q+2)},$$

$$\mu_4(y) = 3k(k+2)\beta^4.$$

Now for the population with p+q=1, γ = 1, $\frac{k+1}{k} \simeq 1$ and $\frac{k+2}{k} \simeq 1$, we have

$$\begin{aligned}\mu_4(x) &= \frac{k^4\beta^4 pq}{8}(2-5pq)\\ &\geq \frac{k^4\beta^4 pq}{8}\left(2-\frac{5}{4}\right)\left[\because pq = \left(\frac{p+q}{2}\right)^2 - \left(\frac{p-q}{2}\right)^2 \leq \frac{1}{4} \therefore -5pq \geq -\frac{5}{4}\right] \qquad (5.2.3)\\ &> 0\end{aligned}$$

$$\beta_2(y) = 3.\frac{k(k+2)}{k^2} \simeq 3,$$

$$\beta_2(x) = \frac{(2-5pq)}{2pq} = \left(\frac{1}{pq} - \frac{5}{2}\right),$$

$$\begin{aligned}E\left(\bar{x}^2\right) &= \frac{\beta^2 k^2 p\left[q + np(p+q) + np\right]}{n(p+q)^2(p+q+1)}\\ &= \frac{\beta^2 k^2 p}{2n}\left[(p+q) + (2n-1)p\right] = \frac{\beta^2 k^2 p\left[1 + (2n-1)p\right]}{2n},\end{aligned}$$

$$Cov\left(s_y^2, \bar{y}\right) = \frac{2\beta^3 k}{n} > 0,$$

$$Cov\left(s_y^2,\bar{x}\right)=-\left(\frac{p}{p+q}\right)k^2\beta^3=-pk^2\beta^3<0\ ,$$

$$Cov\left(s_y^2,s_x^2\right)=-\frac{\beta^4k^3pq}{\left(p+q\right)^2\left(p+q+1\right)}$$

$$=-\frac{\beta^4k^3}{2}.\left\{\left(\frac{p+q}{2}\right)^2-\left(\frac{p-q}{2}\right)^2\right\}=-\frac{\beta^4k^3}{8}\left(1-\left(p-q\right)^2\right),$$

$$Cov\left(\bar{x},\bar{y}\right)=\frac{1}{n}.\left(\frac{p}{p+q}\right).\beta^2k=\frac{p}{n}\beta^2k\ ,$$

$$V\left(s_y^2\right)=\frac{2\beta^4k^2}{\left(n-1\right)}\ ;\ V\left(s_x^2\right)=\frac{\beta^4k^4\left(pq\right)^2}{4n}\left[\frac{2-5pq}{2pq}-\frac{n-3}{n-1}\right],$$

$$E\left(s_y^2\right)^2=\beta^2k^2.\frac{n+1}{n-1}\sim\beta^2k^2;\ \ E\left(s_x^2\right)^2=\frac{\beta^4k^4\left(pq\right)^2}{4n}\left[\frac{2-5pq}{2pq}+\frac{n^2-2n+3}{n-1}\right].$$

Remarks:

It may be noted that in the bivariate gamma population, '$\mu_2(y)$' depends on the parameters of the '*x*' distribution, so improvement may be possible. However, for the bivariate beta-stacy distribution '$\mu_2(y)$' does not depend on the parameters of the '*x*' distribution, so a minimum variance unbiased estimators of σ_y^2 is available in this case (Johnson and Kotz (1972). However, if the roles of *x* and *y* are changed then it would make sense.

The following proposition provides the range of values of λ_2 for bivariate beta-stacy population which, when imputed, to the proposed estimators makes them improved over T_1^* as well as over the usual estimator s_y^2.

Proposition 5.2.1:

Let the terms $\left(\frac{k+1}{k}\right)$ and $\left(\frac{k+2}{k}\right)$ be replaced by unity and the bivariate beta-staey population as in (5.2.1) be considered. Let $\delta=kpq$. If $p+q=1$, then for large n, a set of necessary and sufficient conditions for estimators in (4.1) to (4.4) to be better than T_1^* and hence than s_y^2 would be that the corresponding λ_2 should lie between 0 and $2\lambda_{02}^*$, where,

$$\lambda_{02}^{*} = \begin{cases} \beta / p, \; for \;\; t = \overline{x} \\ 2/\delta, \; for \;\; t = s_x^2 \\ \left[\left(1-\lambda_1^{*}\right)-\delta\right] / \left(1+\delta^2\right), \; for \; t = \left(s_y^2 - s_x^2\right) \end{cases}$$

Proof:

The proposition follows from (4.8), (4.9), (5.2.3) and Appendix A2.6

6. NUMERICAL ILLUSTRATION

Table 24.53 from page 354 of the Book by Hutchinson and Lai (1990) have been used. X in the above table is our Y and their Y is our X. Assuming the bivariate data arising from bivariate gamma population, relative efficiencies of T_1 and $d_i\,'s\,(i=1,2,3)$ in 5.1.4 to 5.1.6 over s_y^2 have been computed and presented in Table 6.2. Rain volumes from clouds which were seeded (y) and matched clouds (x) which were not, are given in the following table.

Table 6.1:
Rain volumes from seeded (Y) and unseeded clouds (X).

Trials 1-5		**Trials 6-10**		**Trials 11-16**	
x	**Y**	**X**	**y**	**X**	**Y**
26.1	129.6	0.0	302.8	68.5	200.7
26.3	31.4	17.3	119.0	81.2	274.7
87.0	2349.6	24.4	4.1	97.3	261.7
95.0	489.1	11.5	92.4	28.6	7.7
372.4	430.0	321.2	17.5	830.1	1606.0
				345.5	978.0

On computation, we have,

$\mu_y = 455.64,\; \mu_x = 152.03,\; \sigma_y^2 = 548321.99,\; \sigma_x^2 = 48056.05$

$a = \mu_x \big/ \sigma_x^2 = .00316,\; p+q = \mu_y a = 1.44,$

$p = a.\mu_x = .480399,\; q = .9596601,$

$\beta_2\left(y\right) = 7.1666,\; c^2\left(s_y^2\right) = 0.63889,\; \Delta = 14.71205,$

Let $n = 10$, and let $\lambda_1^{*} = \lambda_0 = \dfrac{n}{\Delta} = 0.6797$. Now, we have,

$M_0\left(T_1^{*}\right) =$ 1.172051201 x 10^{11}

$V\left(s_y^2\right) =$ 1.920867537 x 10^{11}

$M_0\left(d_1\right) =$ 1.163236471 x 10^{11}

$M_0\left(d_2\right) =$ 1.172043839 x 10^{11}

$M_0\left(d_3\right) =$ 1.169916504 x 10^{11}.

Table 6.2:
Percentage Relative Efficiency of T_1, $d_i(i = 1,2,3)$ over s_y^2 .

$\left[V\left(s_y^2\right)/M_0\left(T_1^*\right)\right]x100$	$\left[V\left(s_y^2\right)/M_0\left(d_1\right)\right]x100$	$\left[V\left(s_y^2\right)/M_0\left(d_2\right)\right]x100$	$\left[V\left(s_y^2\right)/M_0\left(d_3\right)\right]x100$
163.88%	165.13%	163.89%	164.18%

ACKNOWLEDGEMENT

The first author acknowledges the referee and Editor-in-Chief for their valuable suggestion which have improved the earlier version and mourns at the sudden demise of his co-author late Professor T.P. Tripathi.

APPENDIX A1:

A.1.1: Let $\mu_x = \mu_1'$ and for simplicity, $\mu_r'(x)\,(r = 1,2,3,4)$ be written as μ_r' ;

$$E(s_x^2)^2 = \frac{1}{n}\left[\mu_4' - 4\mu_3'\mu_1' - \frac{2(n-2)(n-3)}{(n-1)}\mu_2'\mu_1'^2 + \frac{(n-2)(n-3)}{(n-1)}\mu'^4 + \frac{n^2-2n+3}{(n-1)}\mu_2'^2\right]$$

$$= \frac{1}{n}\left[\mu_4 - 6\mu_2'\mu_1'^2 + 3\mu_1'^4 - \frac{2(n-2)(n-3)}{(n-1)}\mu_2'\mu_1'^2 + \frac{(n-2)(n-3)}{(n-1)}\mu_1'^4 + \frac{n^2-2n+3}{(n-1)}\mu_2'^2\right]$$

$$= \frac{\sigma_x^4}{n}\left[\beta_2 + \frac{2(n^2-2n+3)}{(n-1)}\right].$$

A.1.2: $V(s_x^2) = E(s_x^2)^2 - (E(s_x^2))^2 = \dfrac{\sigma_x^4}{n}\left[\beta_2 + \dfrac{n^2-2n+3}{n-1}\right] - \sigma_x^4$

$$= \frac{\sigma_x^4}{n}\left[\beta_2 - \frac{(n-3)}{(n-1)}\right]$$

A.1.3: $Cov(\bar{x},\bar{y}) = E(\bar{x},\bar{y}) - E(\bar{x}).E(\bar{y}) = \dfrac{1}{n^2}E\left[\sum x_i y_i + \sum_{i\neq j}\sum x_i y_j\right] - \mu_x\mu_y$

$$= \frac{1}{n^2}.\left[n.\mu_{11}' + n(n-1)\mu_x\mu_y\right] - \mu_x\mu_y$$

$$= \frac{\rho.\sigma_y\sigma_x}{n}$$

A.1.4: $Cov(s_y^2, \bar{y}) = E(s_y^2.\bar{y}) - E(s_y^2).E(\bar{y})$

Now, $E(s_y^2.\bar{y}) = E\left[\frac{1}{(n-1)}\sum_{i=1}^{n}(y_i - \bar{y})^2\right]\left(\frac{1}{n}\sum_{i=1}^{n} y_i\right)$

$$= \frac{1}{(n-1)}E\left[\sum_{i=1}^{n} y_i^2 - \frac{\sum_{i=1}^{n} y_i^2 + \sum_{i\neq j}^{n}\sum y_i y_j}{n}\right]\left(\sum_{i=1}^{n} y_i\right)$$

$$= \frac{1}{n}\left[\mu_3' + (n-3)\mu_2'\mu_1' - (n-2)\mu_1'^3\right]$$

Therefore,

$$Cov(s_y^2, \bar{y}) = \frac{1}{n}\left[\mu_3' + (n-3)\mu_2'\mu_1' - (n-2)\mu_1'^3\right] - \sigma_y^2\mu_y$$

$$= \frac{1}{n}\left[\mu_3 + n\mu_2\mu_1'\right] - \mu_2\mu_1' = \frac{\mu_3}{n}\ .$$

A.1.5: $Cov(s_y^2, \bar{x}) = E(s_y^2.\bar{x}) - E(s_y^2).E(\bar{x})$

Now, $E(s_y^2.\bar{x}) = E\left[\frac{1}{n(n-1)}\sum_{i=1}^{n}(y_i - \bar{y})^2\left(\sum_{i=1}^{n} x_i\right)\right]$

$$= \frac{1}{n(n-1)}E\left[\left\{\frac{(n-1)}{n}\sum_{i=1}^{n} y_i^2 - \frac{\sum_{i\neq j} y_i y_j}{n}\right\}\sum_{i=1}^{n} x_i\right]$$

$$= \frac{1}{n}\left[\mu_{12}'(x,y) + \mu_{10}'(x,y)\left\{(n-1)\mu_{02}'(x,y) - (n-2)\mu_{01}'^2(x,y)\right\} - 2\mu_{11}'(x,y).\mu_{01}'(x,y)\right]$$

Therefore,

$$Cov(s_y^2, \bar{x}) = \frac{1}{n}\left[\begin{array}{l}\mu_{12}'(x,y) + \mu_{10}'(x,y)\left\{(n-1)\mu_{02}'(x,y) - (n-2)\mu_{01}'^2(x,y)\right\}\\ -2\mu_{11}'(x,y).\mu_{01}'(x,y) - \mu_2(y)\mu_1'(x)\end{array}\right]$$

$$Cov\left(s_y^2, \bar{x} - \bar{y}\right) = Cov\left(s_y^2, \bar{x}\right) - Cov\left(s_y^2, \bar{y}\right)$$

$$Cov\left(s_y^2, s_x^2 - s_y^2\right) = Cov\left(s_y^2, s_x^2\right) - V\left(s_y^2\right)$$

$$V(s_y^2) = \frac{\mu_2^2(y)}{n}\left[\beta_2 - \frac{(n-3)}{(n-1)}\right] = \frac{\sigma_y^4}{n}\left[\beta_2 - \frac{n-3}{n-1}\right]$$

$$E\left(s_x^2\right)^2=\frac{1}{n}\left[\begin{array}{l}\mu'_{40}(x,y)+\dfrac{\left(n^2-2n+3\right)}{(n-1)}\mu'^{2}_{20}(x,y)-4\mu'_{30}(x,y)\,\mu'_{10}(x,y)\\ -\dfrac{2(n-2)(n-3)}{(n-1)}\mu'_{20}(x,y).\mu'^{2}_{01}(x,y)+\dfrac{(n-2)(n-3)}{(n-1)}\mu'^{4}_{10}(x,y)\end{array}\right]$$

A.1.6: $Cov(s_y^2,s_x^2)=E(s_y^2.s_x^2)-E(s_y^2).E(s_x^2)$

$$=E\left\{\frac{1}{(n-1)}\sum_{i=1}^{n}(y_i-\overline{y})^2\times\frac{1}{(n-1)}\sum_{i=1}^{n}(x_i-\overline{x})^2\right\}-\mu_2(y)\,\mu_2(x)$$

$$+\frac{1}{n^2}\left\{4\sum_{i\neq j\neq k}y_i\,x_i\,y_jx_k+2\sum_{i\neq j}x_i\,y_ix_jy_j+\sum y_iy_jx_kx_l\right\}\bigg]-\mu_2(y)\,\mu_2(x)$$

$$=\frac{1}{n^2}\left\{n\,\mu'_{22}(x,y)+n\,(n-1)\,\mu'_{20}(x,y)\mu'_{02}(x,y)\right\}$$

$$-\frac{1}{(n-1)}\left\{2\,\mu'_{12}(x,y)\,n(n-1)\,\mu'_{10}(x,y)\right\}$$

$$+n(n-1)(n-2)\mu'_{02}(x,y)\mu'^{2}_{10}(x,y)\bigg\}-\frac{1}{n^2\,(n-1)}\bigg\{2\,\mu'_{21}(x,y)\,\mu'_{01}(x,y)\,n(n-1)$$

$$+n(n-1)(n-2)\mu'_{20}(x,y)\mu'^{2}_{0,1}(x,y)\bigg\}$$

$$+\frac{1}{(n(n-1))^{2}}\bigg\{4n(n-1)(n-2)\,\mu'_{11}(x,y)\,\mu'_{10}(x,y)\,\mu'_{01}(x,y)$$

$$+2n(n-1)\mu'^{2}_{11}(x,y)+n(n-1)\,(n-2)\,(n-3)\mu'^{2}_{01}(x,y)\,\mu'^{2}_{10}(x,y)\bigg\}-\mu_2(y)\,\mu_2(x)$$

$$=\frac{1}{n}\Big[\mu'_{22}(x,y)+(n-1)\,\mu'_{20}(x,y)\,\mu'_{02}(x,y)-2\mu'_{12}(x,y)\,\mu'_{10}(x,y)$$

$$-(n-2)\,\mu'^{2}_{10}(x,y)\,\mu'_{02}(x,y)-2\mu'_{21}(x,y)\,\mu'_{01}(x,y)$$

$$-(n-2)\,\mu'_{20}(x,y)\,\mu'^{2}_{20}(x,y)\,\mu'^{2}_{01}(x,y)$$

$$+4\left(\frac{n-2}{n-1}\right)\mu'_{11}(x,y)\,\mu'_{10}(x,y)\,\mu'_{01}(x,y)+\frac{2}{(n-1)}\mu'^{2}_{11}(x,y)$$

$$+\frac{(n-2)\,(n-3)}{(n-1)}\,\mu'^{2}_{10}(x,y)\times\mu'^{2}_{0,1}(x,y)\Big]-\sigma_y^2\sigma_x^2$$

APPENDIX A2.1:

Calculation of moments of different order of bivariate gamma distribution. We have,

$$\mu'_{r,s}(x,y)=\frac{a^{p+q}}{\Gamma p\Gamma q}\iint x^r y^s x^{p-1}y^{q-1}\left(1-\frac{x}{y}\right)^{q-1} e^{-ay}dxdy$$

$$=\frac{a^{-(s+r)}\,\Gamma(r+p)\;\Gamma(s+q+r+p)}{\Gamma p\;\Gamma(p+q+r)};$$

Hence, relating $\mu'_{r,s}(x,y)$ with $\mu_{r,s}(x,y)$, we have,

$$\mu_1(x)=\mu_{1,0}(x,y)=p/a\;;\;\mu_1(y)=\mu_{0,1}(x,y)=(p+q)/a$$

$$\mu_2(x)=\mu_{2,0}(x,y)=p/a^2\;;\;\mu_2(y)=\mu_{0,2}(x,y)=(p+q)/a^2$$

$$\mu_3(x)=\mu_{3,0}(x,y)=2p/a^3\;;\;\mu_3(y)=\mu_{0,3}(x,y)=2(p+q)/a^3$$

$$\mu_4(x)=\mu_{4,0}(x,y)=\frac{3p(p+2)}{a^4};\;\mu_4(y)=\mu_{0,4}(x,y)=\frac{3(p+q)(p+q+2)}{a^4}$$

$$\mu'_{1,1}(x,y)=p(p+q+1)/a^2$$

$$\mu'_{2,1}(x,y)=(p+q+2)(p+1)p/a^3$$

$$\mu'_{1,2}(x,y)=(p+q+2)(p+q+1)p/a^3$$

$$\mu'_{3,1}(x,y)=(p+q+3)(p+2)(p+1)p/a^4$$

$$\mu'_{1,3}(x,y)=(p+q+3)(p+q+2)(p+q+1)p/a^4$$

$$\mu'_{2,2}(x,y)=(p+1)p(p+q+3)(p+q+2)/a^4$$

APPENDIX A2.2:

Calculations of $Cov\left(s_y^2,t\right)$ for $t=s_x^2,\bar{x},\bar{y}$ under the assumption of

$$\frac{n+1}{n-1}\simeq 1,\;\frac{n}{n-1}\simeq 1.$$

(i) $$Cov\left(s_y^2,s_x^2\right)=\frac{1}{na^4}\Big[p(p+1)(p+q+2)(p+q+3)+(n-1)p(p+1).$$

$$\text{x }(p+q)(p+q+1)-2p^2(p+q+1)(p+q+2)$$

$$-(n-2)p^2(p+q)(p+q+1)-2p(p+1)(p+q)(p+q+2)$$

$$-(n-2)p(p+1)(p+q)^2+4\left(\frac{n-2}{n-1}\right)p^2(p+q)(p+q+1)$$

$$+\frac{2}{(n-1)}p^2(p+q+1)^2+\frac{(n-2)(n-3)}{(n-1)}p^2(p+q)^2-np(p+q)\;\Big]$$

$$= \frac{1}{na^4}\left\{2p^2.\left(\frac{n}{n-1}\right)+6p\right\} \simeq \frac{1}{na^4}2p\left(p+3\right);$$

(ii) $Cov\left(s_y^2,\bar{x}\right)=\frac{2p}{na^3}$;

(iii) $Cov\left(s_y^2,\bar{y}\right)=\frac{2\left(p+q\right)}{na^3}$;

(iv) $Cov\left(\bar{x},\bar{y}\right)=\frac{p}{na^2}$;

(v) $\beta_2\left(x\right)=\frac{3\left(p+2\right)}{p}$;

(vi) $\beta_2\left(y\right)=\frac{3\left(p+q+2\right)}{\left(p+q\right)}$;

(vii) $V\left(s_x^2\right)=\frac{p}{na^4}\left[2p\left(\frac{n}{n-1}\right)+6\right] \simeq \frac{2p\left(p+3\right)}{na^4}$;

(viii) $V\left(s_y^2\right)=\frac{2\left(p+q\right)}{na^4}\left[\left(p+q\right)\left(\frac{n}{n-1}\right)+3\right] \simeq \frac{2\left(p+q\right)\left(p+q+3\right)}{na^4}$;

(ix) $E\left(s_y^2\right)^2=\frac{p}{na^4}\left[\frac{\left(n+1\right)}{\left(n-1\right)}.np+6\right] \simeq \frac{p}{na^4}\left(np+6\right)$;

APPENDIX A2.3:

Calculation of λ_{02}^*'s for different t's $\left(t=\bar{x},s_x^2,s_y^2-s_x^2\right)$.

Case 1: $t=\bar{x}$:

$$\lambda_{02}^*=\left[\left(1-\lambda_1^*\right)\sigma_y^2E\left(t\right)-\lambda_1^*Cov\left(s_y^2,\bar{x}\right)\right] \Big/ E\left(\bar{x}^2\right)$$

$$= \frac{p\left[\left(1-\lambda_1^*\right)\left(p+q\right)-\frac{2\lambda_1^*}{n}\right]}{a^3} x \frac{a^2}{p\left(\frac{1}{n}+p\right)}$$

$$= \frac{\left(1-\lambda_1^*\right)\left(p+q\right)-\left(2\lambda_1^*/n\right)}{ax\left(\frac{1}{n}+p\right)}$$

Thus, for the bivariate gamma population with $p+q=1$, we have,

$$\lambda_{02}^* = \frac{1}{a\left(p+1\right)}\left[\frac{1}{a}-\frac{\lambda_1^*}{na}\left(n+2a\right)\right]$$

$$= \frac{\left[1-\lambda_1^*\left(1+\frac{2a}{n}\right)\right]}{a^2\left(p+\frac{1}{n}\right)};$$

and for large n $\left(n\to\alpha\right)$, we have,

$$\lambda_{02}^* = \left(1-\lambda_1^*\right)\big/ap \qquad \text{(A 2.3.1)}$$

It may be noted that,

$$\lambda_{02}^* > 0,\ \text{or} < 0\ \text{according as}\ n > \frac{2}{p+q}.\frac{\lambda_1^*}{\left(1-\lambda_1^*\right)},\ \text{or}\ n < \frac{2}{p+q}.\frac{\lambda_1^*}{\left(1-\lambda_1^*\right)}.$$

Hence, a set of sufficient conditions for d_1 to be better than λ_1^* for all bivariate gamma population with $p+q=1$, would be that

$$0<\lambda_2 \le 2\lambda_{02}^{*(1)} < 2\lambda_{02}^*\ \text{for,}\ n > \frac{2}{p+q}.\frac{\lambda_1^*}{\left(1-\lambda_1^*\right)}$$

or $2\lambda_{02}^{*(1)} \le \lambda_2 < 0$, for, $n < \dfrac{2}{p+q}.\dfrac{\lambda_1^*}{\left(1-\lambda_1^*\right)}$

where, $\lambda_{02}^{*(1)} = \left(1-\lambda_1^*\right)\big/2a_{(2)}^2$, $a_{(2)} \le a$ $\left[\because p+1<2\right]$

<u>Case 2</u>: $t = s_x^2$;

$$\lambda_{02}^* = \left[\left(1-\lambda_1^*\right)\sigma_y^2\sigma_x^2 - \lambda_1^* Cov\left(s_y^2, s_x^2\right)\right]\Big/E\left(s_x^2\right)^2$$

$$= \frac{1}{\left(np+6\right)}\left[n\left(p+q\right)-\lambda_1^*\left\{n\left(p+q\right)+2\left(p+3\right)\right\}\right]$$

Thus, for bivariate gamma population with $p+q=1$, we have,

$$\lambda_{02}^* = \left[n-\lambda_1^*\left(n+2\left(p+3\right)\right)\right]\Big/n\left(p+\frac{6}{n}\right),$$

$$= n\left[1-\lambda_1^*\left(1+\frac{2\left(p+3\right)}{n}\right)\right]\Big/n\left(p+\frac{6}{n}\right)$$

which, for large n i.e. $n \to \alpha$, becomes,

$$\lambda_{02}^{*} = \left(1-\lambda_1^{*}\right) \Big/ p > 0 \ . \qquad \text{(A 2.3.2)}$$

Hence, a set of sufficient conditions for d_2 to be better than λ_1^{*} would be,

$$0 < \lambda_2 \le 2\lambda_{02}^{*(1)} ;$$

Case 3: $t = \left(s_y^2 - s_x^2\right)$;

$$\lambda_{02}^{*} = \frac{\left[\left(1-\lambda_1^{*}\right)\sigma_y^2\left(\sigma_y^2 - \sigma_x^2\right) - \lambda_1^{*} Cov\left\{s_y^2, \left(s_y^2 - s_x^2\right)\right\}\right]}{E\left(s_y^2 - s_x^2\right)^2}$$

Thus, $$\lambda_{02}^{*} = \frac{\left[(p+q) - \lambda_1^{*}\left\{(p+q) + \dfrac{2(2p+q+3)}{n}\right\}\right]}{\left[q + \dfrac{2(2p+q+3)}{n}\right]}$$

Thus, for the bivariate gamma populations with $p+q=1$ and for large n , $(n \to \alpha)$, we have,

$$\lambda_{02}^{*} = \frac{\left[1-\lambda_1^{*}\left(1+\dfrac{2(2p+q+3)}{n}\right)\right]}{\left[q+\dfrac{2(2p+q+3)}{n}\right]} = \left(1-\lambda_1^{*}\right)/q \qquad \text{(A 2.3.3)}$$

Clearly, $\lambda_{02}^{*} > 0$, if $n > \dfrac{\lambda_1^{*}}{\left(1-\lambda_1^{*}\right)} . \dfrac{2(2p+q+3)}{n}$ and hence a sufficient condition for d_3 to be better than T_1^{*} would be

$$0 < \lambda_2 \le 2\lambda_{02}^{*(1)} ,$$

where, $\lambda_{02}^{*(1)}$ is any known quantity depending on some bounds on the parameters involved.

It may be observed that, for $n \to \propto$,

$$\lambda_{02}^{*2} . E\left(s_x^2\right)^2 = \left(1-\lambda_1^{*}\right)^2 \Big/ a^4$$

$$\lambda_{02}^{*2}.E\left(s_y^2 - s_x^2\right)^2 = \left(1-\lambda_1^*\right)^2 \Big/ a^4$$

and $\lambda_{02}^{*2}.E\left(\overline{x}^2\right) = \left(1-\lambda_1^*\right)^2 \Big/ a^4$

APPENDIX A2.4:

Calculation of moments of different orders of Beta-Stacy population, (5.2.1). We have,

$$\mu'_{r,s}\left(x,y\right) = \frac{\beta\left(r+p,q\right)}{\beta\left(p,q\right)} E\left(y^{r+s}\right),$$

where,

$$E\left(y^{r+s}\right) = \int_0^{\infty} y^{r+s} f\left(y\right) dy\ ;$$

and ,

$$f(y) = \frac{\int\limits_{0<x<y} \gamma.\left(\frac{x}{y}\right)^{p-1} y^{\left(p-1\right)}.y^{(q-1)}\left(1-\frac{x}{y}\right)^{q-1}.y^{k\gamma-p-q}.\exp\left\{-\left(\frac{y}{\beta}\right)^{\gamma}\right\} dx}{\beta^{yk}\Gamma k\beta\left(p,q\right)}$$

$$= \int\limits_{0<u<1} \gamma.y^{\gamma k-1} k^{p-1}\left(1-u\right)^{q-1} \exp\left\{-\left(\frac{y}{\beta}\right)^{\gamma}\right\} dy \Big/ \beta^{\gamma k}\,\Gamma k\beta\left(p,q\right)$$

$$= \gamma.y^{\gamma k-1} e^{-\left(\frac{y}{\beta}\right)^{\gamma}} \Big/ \beta^{\gamma k}\,\Gamma k, \qquad y > 0$$

Therefore,

$$E\left(y^{r+s}\right) = \frac{1}{\beta^{\gamma k}\Gamma k}\int_0^{\alpha} \gamma y^{r+s+\gamma k-1} e^{-\left(\frac{y}{\beta}\right)^{\gamma}} dy$$

$$= \frac{\beta^{r+s+\gamma k}\sqrt{\left(r+s+\gamma k\right)/\gamma}}{\beta^{\gamma k}\Gamma k}$$

Hence,

$$\mu'_{r,s}\left(x,y\right) = \frac{\beta\left(r+p,q\right)}{\beta\left(p,q\right)}.\frac{\beta^{r+s}\,\Gamma(\left(r+s+\gamma k\right)/\gamma)}{\Gamma k}$$

Now, for beta-staey distribution with $\gamma = 1$, we have, finally,

$$\mu'_{r,s}\left(x,y\right) = \frac{\beta\left(r+p,q\right)}{\beta\left(p,q\right)}\frac{\beta^{r+s}\ r+s+k}{\Gamma k} \tag{A.2.4.1}$$

APPENDIX A2.5:

Computation of $Cov\left(s_y^2, t\right)$ for different choices of t.

From A.2.4.1, we have,

$$\mu'_{1,2}\left(x,y\right)=\frac{p}{\left(p+q\right)}\beta^3 k\left(k+1\right)\left(k+2\right)$$

$$\mu'_{2,1}\left(x,y\right)=\frac{p\left(p+1\right)}{\left(p+q\right)\left(p+q+1\right)}\beta^3 k\left(k+1\right)\left(k+2\right) \qquad \text{(A.2.5.1)}$$

$$\mu'_{1,1}\left(x,y\right)=\frac{p}{\left(p+q\right)}\beta^2 k\left(k+1\right)$$

$$\mu'_{2,2}\left(x,y\right)=\frac{p\left(p+1\right)}{\left(p+q\right)\left(p+q+1\right)}\beta^4 k\left(k+1\right)\left(k+2\right)\left(k+3\right).$$

(i) $$Cov\left(s_y^2,\bar{x}\right)=\frac{1}{n}\Big[\mu'_{1,2}\left(x,y\right)+\mu_x\left\{\left(n-1\right)\mu'_2\left(y\right)-\left(n-2\right)\mu_y^2\right\}$$
$$-2\mu'_{1,1}\left(x,y\right)\mu_y-n\sigma_y^2\mu_x\Big]$$

$$=\frac{1}{n}\left[\left(\frac{p}{p+q}\right)k\left(k+1\right)\left(k+2\right)\beta^3+\left(\frac{p}{p+q}\right)k\beta\left\{\left(n-1\right)\beta^2 k\left(k+1\right)-\left(n-2\right)k^2\beta^2\right\}\right.$$
$$\left.-\ 2\left(\frac{p}{p+q}\right)k\left(k+1\right)\beta^2 k\beta-n\beta^2 k\left(\frac{p}{p+q}\right).k\beta\ \right]$$

$$=\frac{1}{n}.\left(\frac{p}{p+q}\right).k^3\beta^3\left[\frac{k\left(k+1\right)\left(k+2\right)}{k^3}+\left\{\left(n-1\right)\left(\frac{k+1}{k}\right)-\left(n-2\right)\right\}-2\left(\frac{k+1}{k}\right)-\frac{n}{k}\right]$$

$$=\frac{1}{n}\left(\frac{p}{p+q}\right).\left(-\frac{n}{k}\right).k^3\beta^3\ \left[\frac{k+1}{k}\simeq 1 \text{ and } \frac{k+2}{k}\simeq 1\right]$$

$$=-\left(\frac{p}{p+q}\right)k^2\beta^3 \qquad \text{(A.2.5.2)}$$

Let the terms $\frac{\left(k+1\right)}{k}$, $\frac{\left(k+2\right)}{k}$ and $\frac{\left(k+3\right)}{k}$ be replaced by unity, then from A.2.5.1, we have the followings:

(a) $$\mu'_{2,2}\left(x,y\right)=\beta^4 k^4\frac{p\left(p+1\right)}{\left(p+q\right)\left(p+q+1\right)}$$

(b) $\mu_2'(x).\mu_2'(y) = \dfrac{\beta^4 k^4 p(p+1)}{(p+q)(p+q+1)}$

(c) $\mu_{1,2}'(x,y)\mu_x = \beta^4 k^4 . \left(\dfrac{p}{p+q}\right)^2$

(d) $\mu_{2,1}'(x,y)\mu_y = \dfrac{\beta^4 k^4 p(p+1)}{(p+q)(p+q+1)}$ (A.2.5.3)

(e) $\mu_2'(y)\mu_x^2 = \beta^4 k^4 \left(\dfrac{p}{p+q}\right)^2$

(f) $\mu_2'(x)\mu_y^2 = \dfrac{\beta^4 k^4 p(p+1)}{(p+q)(p+q+1)}$

(g) $\mu_{1,1}'(x,y).\mu_x\mu_y = \beta^4 k^4 . \left(\dfrac{p}{p+q}\right)^2$

(h) $\mu_{1,1}'^2(x,y) = \beta^4 k^4 \left(\dfrac{p}{p+q}\right)^2$

(i) $\mu_x^2.\mu_y^2 = \beta^4 k^4 . \left(\dfrac{p}{p+q}\right)^2$

(j) $\mu_2(x).\mu_2(y) = \beta^4 k^4 \left[\dfrac{pq}{k(p+q)^2(p+q+1)}\right]$

(ii)
$$\begin{aligned} Cov\left(s_y^2, s_x^2\right) = \frac{1}{n}\Bigg[& \mu_{2,2}'(x,y) + (n-1)\mu_2'(x).\mu_2'(y) \\ & - 2\mu_{1,2}'(x,y)\mu_x - (n-2)\mu_x^2.\mu_2'(y) - 2\mu_{2,1}'(x,y)\mu_y \\ & - (n-2)\mu_2'(x)\mu_y^2 + 4\frac{(n-2)}{(n-1)}\mu_{1,1}'(x,y).\mu_x.\mu_y \\ & + \left(\frac{2}{n-1}\right)\mu_{1,1}'^2(x,y) + \frac{(n-2)(n-3)}{(n-1)}\mu_x^2\mu_y^2 - n\mu_2(x).\mu_2(y) \Bigg] \end{aligned}$$

Substituting the values from (A.2.5.3), we have,

$$Cov\left(s_y^2, s_x^2\right) = -\frac{\beta^4 k^3 pq}{(p+q)^2(p+q+1)} \qquad \text{(A.2.5.4)}$$

(iii) $Cov\left(s_y^2, \bar{y}\right) = \dfrac{\mu_3(y)}{n} = \dfrac{2\beta^3 k}{n} > 0$

(iv) $Cov\left(\overline{x},\overline{y}\right)=\frac{\mu_{1,1}\left(x,y\right)}{n}=\frac{1}{n}\left[\mu'_{1,1}\left(x,y\right)-\mu_x\mu_y\right]$

$$=\frac{1}{n}\left[\left(\frac{p}{p+q}\right)\beta^2k\left(k+1\right)-\left(\frac{p}{p+q}\right)k\beta.k\beta\right]$$

$$=\frac{1}{n}.\left(\frac{p}{p+q}\right).\beta^2k.$$

APPENDIX A2.6:

Computation of λ_{02}^* for different choices of t

<u>Case 1</u>: $t=\overline{x}$;

$$\lambda_{02}^*=\left[\left(1-\lambda_1^*\right)\sigma_y^2E\left(\overline{x}\right)-\lambda_1^*Cov\left(s_y^2,\overline{x}\right)\right]\Big/E\left(\overline{x}^2\right)$$

$$=\frac{k^2\beta^3p\big/\left(p+q\right)}{k^2\beta^2p\left\{\frac{q}{n\left(p+q+1\right)}+p\right\}\Big/\left(p+q\right)^2}$$

when $p+q=1$, we have

$$\lambda_{02}^*=2\beta\Big/\left[\frac{1}{n}+\left(2-\frac{1}{n}\right)p\right],$$

which, for large n, i.e. for $n\to\propto$ becomes,

$$\lambda_{02}^*=\left(\beta/p\right)>0$$

<u>Case 2</u>: $t=s_x^2$;

$$\lambda_{02}^*=\left[\left(1-\lambda_1^*\right)\sigma_y^2\sigma_x^2-\lambda_1^*Cov\left(s_y^2,s_x^2\right)\right]\Big/E\left(s_x^2\right)^2$$

$$=\frac{\beta^4k^3pq\big/\left(p+q\right)^2\left(p+q+1\right)}{\frac{\beta^4k^4\left(pq\right)^2}{4}\left\{\frac{2-5pq}{2npq}+\frac{\left(1-\frac{2}{n}\right)}{\left(1-\frac{1}{n}\right)}\right\}},$$

which, for $p+q=1$ and for large n, i.e. for $n\to\propto$ we have

$$\lambda_{02}^*=\frac{2}{\delta}>0\text{, where }\delta=kpq.$$

Case 3: $t = \left(s_y^2 - s_x^2\right)$;

$$\lambda_{02}^* = \frac{\left[\left(1-\lambda_1^*\right)\sigma_y^2\left(\sigma_y^2-\sigma_x^2\right)-\lambda_1^* Cov\left\{s_y^2,\left(s_y^2-s_x^2\right)\right\}\right]}{E\left(s_y^2-s_x^2\right)^2}$$

Numerator = $\left(1-\lambda_1^*\right)\left(k\beta^2\right)^2\left(1-\frac{kpq}{(p+q)^2(p+q+1)}\right)-\ \lambda_1^*\left\{\frac{2\beta^4k^2}{(n-1)}+\frac{\beta^4k^3pq}{(p+q)^2(p+q+1)}\right\}$

$$= \ k^2\beta^4\left[\left(1-\lambda_1^*\right)-\frac{kpq}{(p+q)^2(p+q+1)}\right]$$

Denominator = $V\left(s_y^2\right)+V\left(s_x^2\right)-2Cov\left(s_y^2,s_x^2\right)+\left(\sigma_y^2-\sigma_x^2\right)^2$

$$= \ \frac{2\beta^4k^2}{(n-1)}+\frac{\beta^4k^4pq}{4n}\left(\frac{2-5pq}{2pq}-\left(\frac{n-3}{n-1}\right)\right)+\ \frac{\beta^4k^3}{4}\left\{1-(p-q)^2\right\}+\beta^4k^4\left(1-\frac{kpq}{2}\right)^2$$

Assuming $\frac{n-3}{n-1} \simeq 1$, we have,

$$E\left(s_y^2-s_x^2\right)^2 = k^2\beta^4 \ \text{ x } \left[\text{terms containing }\left(\frac{1}{n}\right)+kpq+\left(1-\frac{kpq}{2}\right)^2\right]$$

and for large n, i.e. for $n \to \propto$ $E\left(s_y^2-s_x^2\right)^2$ becomes

$$E\left(s_y^2-s_x^2\right) = k^2\beta^4\left[1+\frac{k^2p^2q^2}{4}\right]$$

Thus, for $p+q=1$ and for large n, we have,

$$\lambda_{02}^* = \frac{k^2\beta^4\left[\left(1-\lambda_1^*\right)-\frac{kpq}{2}\right]}{k^2\beta^4\left[1+\left(\frac{kpq}{2}\right)^2\right]}$$

$$= \ \frac{\left(1-\lambda_1^*\right)-\delta}{\left(1+\delta^2\right)}$$

REFERENCES

1. Das, A.K. and Tripathi, T.P. (1978). Use of auxiliary information in estimating the finite population variance. *Sankhya*, Series C, 40,139-148.
2. Etoh, T. and Murota, A. (1986). Probabilistic model of rainfall of a single storm. *J. Hydroscience and Hydraulic Engineering*, 4, 65-67.
3. Hutchinson, T.P. and Lai, C.D. (1990). *Continuous bivariate distributions emphasizing applications*. Runsby Scientific Publishing, Remdle Mall, Adelaide, South Australia 5000.
4. Johnson Norman L. and Kotz Samuel (1972). Comparison among estimators of a scale parameter of the Beta-Stacy Distribution. *J. Amer. Statist. Assoc.* 67, 203-205.
5. Kotz, Samuel, N. Balakrishnan and Norman L. Johnson (2000). *Continuous Multivariate Distributions*, Volume 1: Models and Applications, 2nd edition, New York: John Wiley & Sons.
6. Liu, P.C. (1967). Estimating long term wave statistics from long term wind statistics. In B.L. Edge (Editor). *Twentieth Coastal Engineering Conference Proceedings*, 1, 512-521; New York.
7. Maiti, P. and Tripathi, T.P. (1981). A class of estimators for population variance using some apriori information. Stat-Math. Tech report No. 22/81, ISI, Calcutta, Presented at the *Golden Jubilee Conference on Statistics: Applications and new directions*, held at ISI, Calcutta.
8. Maiti, P. (1983). *Some contributions to the sampling theory using apriori and aposteriori information*. Ph.D. Thesis submitted to the Indian Statistical Institute, Calcutta.
9. Maiti, P. (2007). *A class of shrinkage type of estimators for variance of some univariate populations*. To appear in *JISPS*.
10. Pieph H-P. and McCulloch, C.E. (2002). Can the sample variance estimator be improved by using a covariate? *J. Agri. Bio. Envir. Statist.*, 7, 157-175.
11. Searles, D.T. (1964). The utilization of a known coefficient of variation in estimation procedure. *J. Amer. Statist. Assoc.*, 59, 1225-1226.
12. Shinozaki (1995). Some modification of improved estimators of a normal variance. *Ann. Inst. Statist. Math.*, 47, 273-286.
13. Singh, J., Pandey B.N. and Hirano, K. (1973). On the utilisation of a known coefficient of kurtosis in the estimation procedure of variance. *Ann. Inst. Statist. Math.*, 25, 51-55.
14. Tripathi, T.P. and Chaubey, Y.P. (1992). Improved estimation of a finite population mean based on paired observations. *Comm. Statist. Theory and Methods*, 21(11), 3329-3333.