A Generalized Difference-cum-Ratio Type Estimator for the Population Variance in Double Sampling

H. S. Jhajj and G.S. Walia

Abstract: For estimating the population variance of variable under study, a generalized difference-cum-ratio type estimator has been proposed. Expressions for bias and mean square error have been obtained using random sampling at both the phases. The expressions have also been derived using simple random sampling at both the phases as a special case. Then the comparison has been made with the regression-type estimator and sample variance. Results have also been illustrated numerically and graphically.

Keywords : Regression-type estimator, Sample variance, Mean square error, Auxiliary variable, Double Sampling, Efficiency.

I. INTRODUCTION

In survey sampling, the estimation of population variance of the variable under study has attracted the attention of large number of statisticians to know the variation in the population. It is well known that the use of the auxiliary information can increase the efficiency of the estimators of parameters of interest. Using the prior knowledge of population variance S_x^2 of auxiliary variable x, which is highly correlated with study variable y, several estimators have been defined by different authors such as Tripathi et al.(1978), Jhajj et al.(1980), Ahmed et al.(2003), Jhajj et al.(2005), Kadilar & Cingi (2006), Pradhan B.K.(2010) in the literature for estimating the unknown population variance of study variable y. In the situation when information on population variance S_x^2 is not known in advance then generally two phase (double) sampling design has been widely used. In the two-phase sampling design, a large preliminary random sample (called first phase sample) is drawn from the population and information on auxiliary variable is taken, which is used to estimate the value of unknown population variance S_x^2 of auxiliary variable x. Then second phase sample is drawn either from the first phase sample or independently from the population and observations on study and auxiliary variable are taken.

Manuscript received Nov. 30, 2010; revised April 18, 2011.

Harbans Singh Jhajj, Prof. & Head, Department of Statistics, Punjabi University, Patiala-147002, India Email : drhsjhajj@yahoo.co.in

Gurjeet Singh Walia, Department of Statistics, Punjabi University, Patiala-147002, India. Email : walia.pbi.uni@gmail.com

In the present paper, we propose a generalized difference-cum-ratio type estimator for the population variance under double sampling design. The expressions for bias and mean square error of the proposed estimator have been obtained. The comparison of the proposed estimator has been made with the regression type and sample variance. Effort has also been made to illustrate the results numerically and graphically.

II. NOTATIONS AND RESULTS

A preliminary large random sample (first phase sample) of size n' is drawn from a finite population of size N and both auxiliary variable x and study variable y are measured on it. The second phase random sample of size $n(<n')$ is drawn from the first phase sample.

Let Y_i and X_i denote the respective values of variables y and x on the i^{th} $(i=1,2,...,N)$ unit of the population and the corresponding small letters denote the values in the sample.

Denoting

$$s_y^2=\frac{1}{n-1}\sum_{i=1}^{n}(y_i-\bar{y})^2, \qquad s_y'^2=\frac{1}{n'-1}\sum_{i=1}^{n'}(y_i-\bar{y})^2,$$

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

$$s_x'^2=\frac{1}{n'-1}\sum_{i=1}^{n'}(x_i-\bar{x})^2, \qquad S_x^2=\frac{1}{N-1}\sum_{i=1}^{N}(X_i-\bar{X})^2$$

$$\mu_{rs}=\frac{1}{N}\sum_{i=1}^{N}(Y_i-\bar{Y})^r(X_i-\bar{X})^s \qquad \lambda_{rs}=\frac{\mu_{rs}}{\mu_{20}^{r/2}\mu_{02}^{s/2}},$$

$$\rho_v=\frac{(\lambda_{22}-1)}{\sqrt{(\lambda_{04}-1)(\lambda_{40}-1)}}$$

where s_x^2, $s_x'^2$ and s_y^2, $s_y'^2$ are the sample variances of variables x and y based on the sampling units of first and second phase samples of sizes n' and n respectively.

Defining

$$\delta_0=\frac{s_x^2}{S_x^2}-1 \qquad \varepsilon_0=\frac{s_y^2}{S_y^2}-1$$

$$\delta_1=\frac{s_x'^2}{S_x^2}-1 \qquad \varepsilon_1=\frac{s_y'^2}{S_y^2}-1$$

(Advance online publication: 9 November 2011)

We assume that

$$E\left(\delta_0\right)=E\left(\delta_1\right)=E\left(\varepsilon_0\right)=E\left(\varepsilon_1\right)=0$$

$$\left.\begin{array}{ll}
E\left(\delta_0^2\right)=\dfrac{Var\left(s_x^2\right)}{S_x^4} & E\left(\delta_1^2\right)=\dfrac{Var\left(s_x'^2\right)}{S_x^4}\\
E\left(\varepsilon_0^2\right)=\dfrac{Var\left(s_y^2\right)}{S_y^4} & E\left(\varepsilon_1^2\right)=\dfrac{Var\left(s_y'^2\right)}{S_y^4}\\
E\left(\varepsilon_0\varepsilon_1\right)=\dfrac{Cov\left(s_y^2,s_y'^2\right)}{S_y^4} & E\left(\delta_0\delta_1\right)=\dfrac{Cov\left(s_x^2,s_x'^2\right)}{S_x^4}\\
E\left(\delta_0\varepsilon_0\right)=\dfrac{Cov\left(s_x^2,s_y^2\right)}{S_x^2S_y^2} & E\left(\delta_0\varepsilon_1\right)=\dfrac{Cov\left(s_x^2,s_y'^2\right)}{S_x^2S_y^2}\\
E\left(\delta_1\varepsilon_0\right)=\dfrac{Cov\left(s_x'^2,s_y^2\right)}{S_x^2S_y^2} & E\left(\delta_1\varepsilon_1\right)=\dfrac{Cov\left(s_x'^2,s_y'^2\right)}{S_x^2S_y^2}
\end{array}\right\} \quad (2.1)$$

III. THE PROPOSED ESTIMATOR AND ITS MEAN SQUARE ERROR

When information on population variance S_x^2 of the auxiliary variable x is not known, we propose an estimator of population variance S_y^2 of study variable y under the double sampling design defined in section 2 as

$$s_{hgd}^2=\left[s_y^2+\theta\left(s_y'^2-s_y^2\right)\right]\left[\frac{s_x'^2}{s_x^2+\theta\left(s_x'^2-s_x^2\right)}\right]^{\alpha} \qquad (3.1)$$

where α and θ are unknown constants.

To find the bias and mean square error of estimator s_{hgd}^2, we expand s_{hgd}^2 in terms of ε's and δ's and retaining terms up to second degree of approximation

$$\begin{aligned}
s_{hgd}^2=S_y^2\big[&1+\varepsilon_0+\theta\left(\varepsilon_1-\varepsilon_0\right)+\alpha(1-\theta)\left(\delta_1-\delta_0\right)-\alpha\delta_0\left(\delta_1-\delta_0\right)\\
&+\alpha\theta^2\left(\delta_1-\delta_0\right)^2+2\alpha\theta\delta_0\left(\delta_1-\delta_0\right)-\alpha\theta\delta_1\left(\delta_1-\delta_0\right)\\
&+\frac{\alpha(\alpha-1)}{2}(1-\theta)^2\left(\delta_1-\delta_0\right)^2+\alpha(1-\theta)\varepsilon_0\left(\delta_1-\delta_0\right)\\
&+\alpha\theta(1-\theta)\left(\varepsilon_1-\varepsilon_0\right)\left(\delta_1-\delta_0\right)\big]
\end{aligned} \qquad (3.2)$$

Taking expectation of (3.2), we obtain

$$E\left(s_{hgd}^2\right)=S_y^2+S_y^2(1-\theta)^2\left\{\frac{\alpha(\alpha+1)}{2}\left(\frac{V\left(s_x^2\right)-V\left(s_x'^2\right)}{S_x^4}\right)-\alpha\left(\frac{Cov\left(s_y^2,s_x^2\right)-Cov\left(s_y^2,s_x'^2\right)}{S_y^2S_x^2}\right)\right\} \qquad (3.3)$$

Up to first order of approximation, mean square error (MSE) of the estimator s_{hgd}^2 is obtained by using (3.2) as

$$\begin{aligned}
MSE\left(s_{hgd}^2\right)&=E\left(s_{hgd}^2-S_y^2\right)^2\\
&=S_y^4E\left[\varepsilon_0+\theta\left(\varepsilon_1-\varepsilon_0\right)+\alpha(1-\theta)\left(\delta_1-\delta_0\right)\right]^2\\
&=V\left(s_y^2\right)+\left(\theta^2-2\theta\right)\left\{V\left(s_y^2\right)-V\left(s_y'^2\right)\right\}\\
&\quad+\alpha^2(1-\theta)^2\frac{S_y^4}{S_x^4}\left\{V\left(s_x^2\right)-V\left(s_x'^2\right)\right\}\\
&\quad-2\alpha(1-\theta)^2\frac{S_y^2}{S_x^2}\left\{Cov\left(s_y^2,s_x^2\right)-Cov\left(s_y^2,s_x'^2\right)\right\}
\end{aligned} \qquad (3.4)$$

Differentiating (3.4) w.r.t. α and equating to zero, we get

$$2\alpha(1-\theta)^2\frac{S_y^4}{S_x^4}\left\{V\left(s_x^2\right)-V\left(s_x'^2\right)\right\}-2(1-\theta)^2\frac{S_y^2}{S_x^2}\left\{Cov\left(s_y^2,s_x^2\right)-Cov\left(s_y^2,s_x'^2\right)\right\}=0$$

After solving, we get the optimum value of α as

$$(\alpha)_{opt}=\frac{S_x^2\left\{Cov\left(s_y^2,s_x^2\right)-Cov\left(s_y^2,s_x'^2\right)\right\}}{S_y^2\left\{V\left(s_x^2\right)-V\left(s_x'^2\right)\right\}} \qquad (3.5)$$

Substituting the optimum value of α from (3.5) in (3.4), we get minimum mean square error $MSE_{\min}\left(s_{hgd}^2\right)$ as

$$\begin{aligned}
MSE_{\min}\left(s_{hgd}^2\right)=&V\left(s_y^2\right)+\left(\theta^2-2\theta\right)\left\{V\left(s_y^2\right)-V\left(s_y'^2\right)\right\}\\
&-(1-\theta)^2\frac{\left\{Cov\left(s_y^2,s_x^2\right)-Cov\left(s_y^2,s_x'^2\right)\right\}^2}{\left\{V\left(s_x^2\right)-V\left(s_x'^2\right)\right\}}
\end{aligned} \qquad (3.6)$$

Theorem 1: Up to first order of approximation, the bias of estimator s_{hgd}^2 is

$$Bias\left(s_{hgd}^2\right)=S_y^2(1-\theta)^2\left\{\frac{\alpha(\alpha+1)}{2}\left(\frac{V\left(s_x^2\right)-V\left(s_x'^2\right)}{S_x^4}\right)-\alpha\left(\frac{Cov\left(s_y^2,s_x^2\right)-Cov\left(s_y^2,s_x'^2\right)}{S_y^2S_x^2}\right)\right\}$$

and its Mean Square Error is

$$\begin{aligned}
MSE\left(s_{hgd}^2\right)=&V\left(s_y^2\right)+\left(\theta^2-2\theta\right)\left\{V\left(s_y^2\right)-V\left(s_y'^2\right)\right\}\\
&+\alpha^2(1-\theta)^2\frac{S_y^4}{S_x^4}\left\{V\left(s_x^2\right)-V\left(s_x'^2\right)\right\}\\
&-2\alpha(1-\theta)^2\frac{S_y^2}{S_x^2}\left\{Cov\left(s_y^2,s_x^2\right)-Cov\left(s_y^2,s_x'^2\right)\right\}
\end{aligned}$$

Theorem 2: Up to first order of approximation, the MSE of s_{hgd}^2 is minimized for

$$(\alpha)_{opt}=\frac{S_x^2\left\{Cov\left(s_y^2,s_x^2\right)-Cov\left(s_y^2,s_x'^2\right)\right\}}{S_y^2\left\{V\left(s_x^2\right)-V\left(s_x'^2\right)\right\}}$$

and its minimum value is given by

$$MSE_{\min}\left(s_{hgd}^2\right)=V\left(s_y^2\right)+\left(\theta^2-2\theta\right)\left\{V\left(s_y^2\right)-V\left(s_y'^2\right)\right\}$$
$$-\left(1-\theta\right)^2\frac{\left\{Cov\left(s_y^2,s_x^2\right)-Cov\left(s_y^2,s_x'^2\right)\right\}^2}{\left\{V\left(s_x^2\right)-V\left(s_x'^2\right)\right\}}$$

Special Case : When simple random sampling is used for selection of samples in given double sampling design, then we have

$$\left.\begin{array}{ll} Var\left(s_y^2\right)=\frac{1}{n}S_y^4\left(\lambda_{40}-1\right) & Var\left(s_y'^2\right)=\frac{1}{n'}S_y^4\left(\lambda_{40}-1\right)\\ Var\left(s_x^2\right)=\frac{1}{n}S_x^4\left(\lambda_{04}-1\right) & Var\left(s_x'^2\right)=\frac{1}{n'}S_x^4\left(\lambda_{04}-1\right)\\ Cov\left(s_x^2,s_y^2\right)=\frac{1}{n}S_x^2S_y^2\left(\lambda_{22}-1\right) & Cov\left(s_x'^2,s_y'^2\right)=\frac{1}{n'}S_x^2S_y^2\left(\lambda_{22}-1\right)\end{array}\right\}\quad(3.7)$$

Substituting results of (3.7) in (3.3) and (3.4) respectively, we have

$$Bias\left(s_{hgd}^2\right)=\left(\frac{1}{n}-\frac{1}{n'}\right)S_y^2\left(1-\theta\right)^2\left\{\frac{\alpha\left(\alpha+1\right)}{2}\left(\lambda_{04}-1\right)-\alpha\left(\lambda_{22}-1\right)\right\}\quad(3.8)$$

$$MSE\left(s_{hgd}^2\right)=\frac{1}{n}S_y^4\left(\lambda_{40}-1\right)+\left(\frac{1}{n}-\frac{1}{n'}\right)S_y^4\left\{\left(\theta^2-2\theta\right)\left(\lambda_{40}-1\right)+\alpha^2\left(1-\theta\right)^2\left(\lambda_{04}-1\right)-2\alpha\left(1-\theta\right)^2\left(\lambda_{22}-1\right)\right\}\quad(3.9)$$

For minimizing $MSE\left(s_{hgd}^2\right)$, we differentiate (3.9) w.r.t. α and equating to zero and after some simplification, we get

$$\left(\alpha\right)_{opt}=\frac{\left(\lambda_{22}-1\right)}{\left(\lambda_{04}-1\right)}\quad(3.10)$$

Substituting the optimum value of α from (3.10) in (3.9), we obtain

$$MSE_{\min}\left(s_{hgd}^2\right)=\frac{1}{n}S_y^4\left(\lambda_{40}-1\right)+\left(\frac{1}{n}-\frac{1}{n'}\right)S_y^4\left(\lambda_{40}-1\right)\left[\left(\theta^2-2\theta\right)-\left(1-\theta\right)^2\rho_v^2\right]\quad(3.11)$$

Cor 1.1: Up to first order of approximation $\left(n^{-1}\right)$, under double sampling design in which simple random sampling is used at both phases, the bias of estimator s_{hgd}^2 is

$$Bias\left(s_{hgd}^2\right)=\left(\frac{1}{n}-\frac{1}{n'}\right)S_y^2\left(1-\theta\right)^2\left\{\frac{\alpha\left(\alpha+1\right)}{2}\left(\lambda_{04}-1\right)-\alpha\left(\lambda_{22}-1\right)\right\}$$

and its MSE is

$$MSE\left(s_{hgd}^2\right)=\frac{1}{n}S_y^4\left(\lambda_{40}-1\right)+\left(\frac{1}{n}-\frac{1}{n'}\right)S_y^4\left\{\left(\theta^2-2\theta\right)\left(\lambda_{40}-1\right)+\alpha^2\left(1-\theta\right)^2\left(\lambda_{04}-1\right)-2\alpha\left(1-\theta\right)^2\left(\lambda_{22}-1\right)\right\}$$

Cor2.1: Up to first order of approximation $\left(n^{-1}\right)$, under double sampling design in which simple random sampling is used at both phases, the MSE of s_{hgd}^2 is minimized for

$$\left(\alpha\right)_{opt}=\frac{\left(\lambda_{22}-1\right)}{\left(\lambda_{04}-1\right)}$$

and its minimum value is given by

$$MSE_{\min}\left(s_{hgd}^2\right)=\frac{1}{n}S_y^4\left(\lambda_{40}-1\right)+\left(\frac{1}{n}-\frac{1}{n'}\right)S_y^4\left(\lambda_{40}-1\right)\left[\left(\theta^2-2\theta\right)-\left(1-\theta\right)^2\rho_v^2\right]$$

IV. COMPARISON

For comparing the proposed estimator with the existing ones, we first write the expressions of their mean square errors. The MSE of the linear regression-type estimator s_{lrd}^2 under double sampling design is given by

$$MSE\left(s_{lrd}^2\right)=\frac{1}{n}S_y^4\left(\lambda_{40}-1\right)-\left(\frac{1}{n}-\frac{1}{n'}\right)S_y^4\left(\lambda_{40}-1\right)\rho_v^2\quad(4.1)$$

and the MSE of the sample variance s_y^2 is given by

$$MSE\left(s_y^2\right)=\frac{1}{n}S_y^4\left(\lambda_{40}-1\right)\quad(4.2)$$

Using (3.11) and (4.1), we obtain

$$MSE\left(s_{lrd}^2\right)-MSE_{\min}\left(s_{hgd}^2\right)=\left(\frac{1}{n}-\frac{1}{n'}\right)S_y^4\left(\lambda_{40}-1\right)\left(\theta^2-2\theta\right)\left(\rho_v^2-1\right)\geq 0$$

$$\Rightarrow MSE_{\min}\left(s_{hgd}^2\right)\leq MSE\left(s_{lrd}^2\right)\qquad if\qquad 0<\theta<2\quad(4.3)$$

Using (3.11) and (4.2), we obtain

$$MSE\left(s_y^2\right)-MSE_{\min}\left(s_{hgd}^2\right)=\left(\frac{1}{n}-\frac{1}{n'}\right)S_y^4\left(\lambda_{40}-1\right)\left[\rho_v^2-\left(\theta^2-2\theta\right)\left(1-\rho_v^2\right)\right]\geq 0$$

$$\Rightarrow MSE_{\min}\left(s_{hgd}^2\right)\leq MSE\left(s_y^2\right)\qquad if\qquad 1-\frac{1}{\sqrt{1-\rho_v^2}}<\theta<1+\frac{1}{\sqrt{1-\rho_v^2}}\quad(4.4)$$

V. NUMERICAL ILLUSTRATION

To have a rough idea about the gain in efficiency of the proposed estimator (s^2_{hgd}) over the Regression-type estimator $\left(s^2_{lrd}\right)$ in double sampling, we take the empirical population considered in the literature (Source: Sukhatme & Sukhatme, 1970, p-256). The values of the population parameters obtained are given in Table 1. The mean square error and relative efficiency of the proposed estimator (s^2_{hgd}) w.r.t. Regression-type estimator $\left(s^2_{lrd}\right)$ and sample variance (s^2_y) are given for some different values of θ in the table 2.

Table 1: Value of Population Parameters

N	n'	n	S_y	ρ_v
89	45	23	716.65	0.81

Table 2: Mean Square Errors and Relative Efficiency Proposed Estimator Vs. Regression-type Estimator and Sample variance

θ	$MSE\left(s^2_y\right)$	$MSE\left(s^2_{lrd}\right)$	$MSE\left(s^2_{hgd}\right)$	Efficiency		
				s^2_y	s^2_{lrd}	s^2_{hgd}
0	83539.758	56329.12	56329.12	100	148.30	148.30
0.5	83539.758	56329.12	46105.85	100	148.30	181.19
1	83539.758	56329.12	42698.09	100	148.30	195.65
1.5	83539.758	56329.12	46105.85	100	148.30	181.19
2	83539.758	56329.12	56329.12	100	148.30	148.30

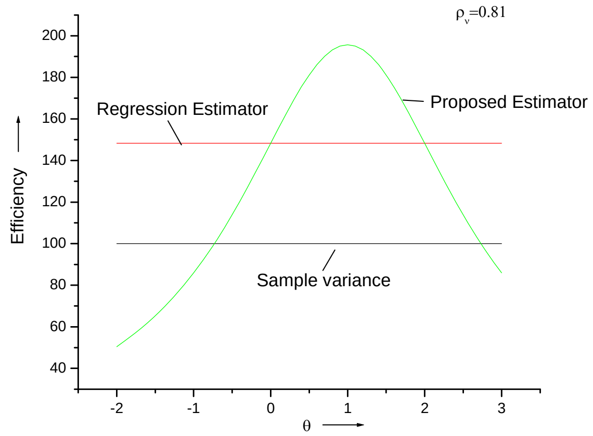

Figure 1: Comparison of Proposed estimator with Regression-type estimator and sample variance

VI. CONCLUSION

From table 2, we can see that there is a significant gain in efficiency of the proposed estimator (s^2_{hgd}) over the Regression-type estimator $\left(s^2_{lrd}\right)$ in double sampling for $0<\theta<2$ and sample variance (s^2_y) for $1-\frac{1}{\sqrt{1-\rho_v^2}}<\theta<1+\frac{1}{\sqrt{1-\rho_v^2}}$. From the graph we also see that for $0<\theta<2$, the efficiency of proposed estimator is more than the regression-type estimator and it is also more efficient than sample variance for wider range of θ at moderate value of correlation coefficient. Hence we conclude that proposed estimator will always be better than the existing regression-type estimator under double sampling for $0<\theta<2$ and sample variance for $1-\frac{1}{\sqrt{1-\rho_v^2}}<\theta<1+\frac{1}{\sqrt{1-\rho_v^2}}$.

ACKNOWLEDGEMENT

Authors are thankful to the Referees for their valuable suggestions in improving the paper.

REFERENCES

[1] S.N. Nath, "On Product Moments from a Finite Universe", Journal *of the American Statistical Association*, 63(322), 535-541, 1968.

[2] S.N. Nath, "More Results on Product Moments from a Finite Universe", Journal *of the American Statistical Association*, 64(327), 864-869, 1969.

[3] A.K. Das and B. Tripathi, "Use of auxiliary information in estimating the finite population variance", *Sankhya, C* 40**,** 139–148, 1978.

[4] H.S. Jhajj and S.K. Srivastava, "A class of estimators using auxiliary information for estimating finite population variance", *Sankhya, C* 42(1–2), 87–96, 1980.

[5] M.S. Ahmed, A.D. Walid and A.O.H. Ahmed, "Some estimators for finite population variance under two-phase sampling", *Statistics in Transition*, **6**(1), 143–150, 2003.

[6] H.S. Jhajj, M.K. Sharma and L.K. Grover, "An Efficient class of chain estimators of population variance under sub-sampling scheme", *Jour Japan Statistical Society,* **35**(2), 273–286, 2005.

[7] C. Kadilar and H. Cingi, "Ratio estimators for the population variance in simple and stratified random sampling", *Applied Mathematics and Computation*, 173, 1047–1059, 2006.

[8] R. Fernandez-Alcala, J. Navarro-Moreno, J. C. Ruiz-Molina and A. Oya, "LMMSE Estimation based on Counting Observations", IAENG Int. J. Appl. Math, 37:2, 2007.

[9] B.K. Pradhan, "Regression Type Estimators of Finite Population Variance Under Multiphase Sampling", *Bull. Malays. Math. Sci. Soc.*, (2) 33(2), 345–353, 2010.

[10] S.A. Al-hadhrami, "Estimation of the population variance using Ranked set sampling with auxiliary variable", *Int. J. Contemp. Math. Sciences*, Vol. 5, no.52, 2567-2576, 2010.

[11] H.P. Singh, R. Tailor, S. Singh and Jong-Min Kim, "Estimation of population variance in successive sampling", *Qual. Quant*, 45:477–494, 2011.

APPENDIX

Appendix: I. For getting the expected value of $\delta_1\varepsilon_1$, we proceed as

$$E\left(\delta_1\varepsilon_1\right)=E\left[\left(\frac{s_x'^2}{S_x^2}-1\right)\left(\frac{s_y'^2}{S_y^2}-1\right)\right]$$

$$=\frac{1}{S_x^2S_y^2}E\left[\left(s_x'^2-S_x^2\right)\left(s_y'^2-S_y^2\right)\right]$$

$$=\frac{Cov\left(s_x'^2,s_y'^2\right)}{S_x^2S_y^2} \tag{1}$$

On the similar line, other results involved in (2.1) can be derived.

II. Under Simple Random Sampling, covariance between sample variance of x and y can be obtained as

$$Cov\left(s_x^2,s_y^2\right)=E\left[\left(s_x^2-S_x^2\right)\left(s_y^2-S_y^2\right)\right]$$
$$=E\left(s_x^2s_y^2\right)-S_x^2S_y^2$$
$$=E\left[\left(\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2\right)\left(\frac{1}{n-1}\sum_{i=1}^{n}\left(y_i-\bar{y}\right)^2\right)\right]-S_x^2S_y^2$$
$$=\frac{1}{\left(n-1\right)^2}E\left\{\left[\sum_{i=1}^{n}\left(x_i-\bar{X}\right)^2-n\left(\bar{x}-\bar{X}\right)^2\right]\left[\sum_{i=1}^{n}\left(y_i-\bar{Y}\right)^2-n\left(\bar{y}-\bar{Y}\right)^2\right]\right\}-S_x^2S_y^2$$

$$=\frac{1}{\left(n-1\right)^2}E\left\{\sum_{i=1}^{n}\left(x_i-\bar{X}\right)^2\sum_{i=1}^{n}\left(y_i-\bar{Y}\right)^2-n\left(\bar{x}-\bar{X}\right)^2\sum_{i=1}^{n}\left(y_i-\bar{Y}\right)^2\right.$$
$$\left.-n\left(\bar{y}-\bar{Y}\right)^2\sum_{i=1}^{n}\left(x_i-\bar{X}\right)^2+n^2\left(\bar{x}-\bar{X}\right)^2\left(\bar{y}-\bar{Y}\right)^2\right\}-S_x^2S_y^2 \tag{2}$$

The following results can be obtained by usual procedure

$$E\left[\sum_{i=1}^{n}\left(x_i-\bar{X}\right)^2\sum_{i=1}^{n}\left(y_i-\bar{Y}\right)^2\right]$$
$$=\frac{\left(N-n\right)n}{\left(N-1\right)}\mu_{22}+\frac{nN\left(n-1\right)}{\left(N-1\right)}\mu_{20}\mu_{02} \tag{3}$$

$$E\left[\left(\bar{x}-\bar{X}\right)^2\sum_{i=1}^{n}\left(y_i-\bar{Y}\right)^2\right]$$
$$=\frac{\left(N-n\right)\left(N-2n\right)}{\left(N-1\right)\left(N-2\right)n}\mu_{22}+\frac{N\left(N-n\right)\left(n-1\right)}{n\left(N-1\right)\left(N-2\right)}\mu_{20}\mu_{02} \tag{4}$$

$$E\left[\left(\bar{y}-\bar{Y}\right)^2\sum_{i=1}^{n}\left(x_i-\bar{X}\right)^2\right]$$
$$=\frac{\left(N-n\right)\left(N-2n\right)}{\left(N-1\right)\left(N-2\right)n}\mu_{22}+\frac{N\left(N-n\right)\left(n-1\right)}{n\left(N-1\right)\left(N-2\right)}\mu_{20}\mu_{02} \tag{5}$$

$$E\left[\left(\bar{x}-\bar{X}\right)^2\left(\bar{y}-\bar{Y}\right)^2\right]=\frac{\left(N-n\right)\left[N^2-\left(6n-1\right)N+6n^2\right]}{\left(N-1\right)\left(N-2\right)\left(N-3\right)n^3}\mu_{22}$$
$$+\frac{N\left(N-n\right)\left(N-n-1\right)\left(n-1\right)}{\left(N-1\right)\left(N-2\right)\left(N-3\right)n^3}\mu_{20}\mu_{02}$$
$$+\frac{2N\left(N-n\right)\left(N-n-1\right)\left(n-1\right)}{\left(N-1\right)\left(N-2\right)\left(N-3\right)n^3}\mu_{11}^2 \tag{6}$$

Using the results of (3), (4), (5) and (6) in (2), and retaining terms up to first order of approximation, we have

$$Cov\left(s_x^2,s_y^2\right)=\frac{1}{n}S_x^2S_y^2\left(\lambda_{22}-1\right)$$

On the similar line, other results involved in (3.8) can be derived.