*International Journal of Mathematics and Statistics Invention (IJMSI)*
*E-ISSN: 2321 – 4767 P-ISSN: 2321 - 4759*
*www.ijmsi.org Volume 3 Issue 8 || December. 2015 || PP-18-22*

# An Improved Regression Type Estimator of Finite Population Mean using Coefficient of variation in PPS sampling.

Anita[1], Shashi Bahl[2]
[1] *Research Scholar, Department of Statistics, M .D.U.Rohtak(Haryana) India.*
[2] *Professor, Department of Statistics, M.D.U.Rohtak(Haryana) India.*

**Abstract:** *Coefficient of variation is a unitless measure of dispersion and is widely used rather than standard deviation. In this paper for estimating population mean, auxiliary information is used in the form of coefficient of variation. The estimator depending on the estimated optimum value of k constant based on sample value under pps sampling scheme have been proposed. Its Bias, MSE and comparative study with the usual ratio, product and linear regression estimators is made. An empirical study is also included for the practical utility of the proposed estimator.*
***Key words :-*** *Coefficient of variation Regression ,Ratio ,Product Estimators Bias & Mean Square Error Efficiency*

## I. Introduction

It is well know that at large scale surveys use of multiple auxiliary characteristics improves the precision of the estimators. In many practical situations, the coefficient of variation Cx of auxiliary variable x may be available. Auxiliary variable x in the form of known population mean are used in estimators like ratio, product and regression in the literature. We can use auxiliary variable in the form of coefficient of variation for the efficient estimation of population mean of study variable y. In this paper two auxiliary characters have been used in different ways as one for the selection of sample and other for the purpose of es*timation to estimate the population mean*.

Notations

Let $y_i$, $x_{1i}$ and $x_{2i}$ be two variables of characters y under study and values of the two auxiliary characters $x_1$ and $x_2$ for the $i^{th}$ unit in the population $(i=1, 2, 3\ldots\ldots N)$ of size N. Let a sample of size n be drawn with ppswr sa*mpling* $(\text{based on } x_1)$ *and* $Y, X_1, X_2$ *be population total of* $y, x_1, x_2$ *respectively.*

$$u_i=\frac{y_i}{Np_{1i}}, v_i=\frac{x_{2i}}{Np_{1i}}, p_{1i}=\frac{x_{1i}}{X_1}$$

$$\bar{u}_n=\frac{1}{n}\sum_{i=1}^{n}u_i=\bar{y}_{1pps}$$

$$\bar{v}_n=\frac{1}{n}\sum_{i=1}^{n}v_i=\bar{x}_{2pps}$$

$$S_u^{\ 2}=\sum_{i=1}^{N}p_{1i}\left(u_i-\bar{y}\right)^2, s_u^{\ 2}=\sum_{i=1}^{N}p_{1i}\left(u_i-\bar{u}_n\right)^2$$

$$S_v^{\ 2}=\sum_{i=1}^{N}p_{1i}\left(v_i-\bar{X}_2\right)^2, s_v^{\ 2}=\sum_{i=1}^{N}p_{1i}\left(v_i-\bar{v}_n\right)^2$$

$$S_{uv}=\sum_{i=1}^{N}p_{1i}\left(u_i-\bar{y}\right)\left(v_i-\bar{X}_2\right), b=\frac{s_{uv}}{s_v^{\ 2}}$$

$$\mu_{rs}=\sum_{i=1}^{N}p_{1i}\left(u_i-\bar{y}\right)^r\left(v_i-\bar{X}_2\right)^s$$

$$C_u=\frac{\sqrt{\mu_{20}}}{\overline{Y}}, C_v=\frac{\sqrt{\mu_{02}}}{\overline{X}_2}, \beta=\frac{S_{uv}}{S_v^{\ 2}}=\rho\frac{S_u}{S_v}$$

$$\rho=\frac{\mu_{11}}{\sqrt{\mu_{20}\mu_{02}}}=\frac{S_{uv}}{S_uS_v}$$

$$\overline{y}_R=\frac{\overline{u}_n}{\overline{v}_n}\overline{X}_2 \text{-----}(1.1)$$

$$\overline{y}_p=\frac{\overline{u}_n\overline{v}_n}{\overline{X}_2}\text{------}(1.2)$$

$$\overline{y}_{lr}=\overline{y}+b(\overline{X}-\overline{x})\text{----}(1.3)$$

$$\beta_1(v)=\frac{\mu_{40}}{\mu_{20}^{\ 2}}=\frac{\mu_{40}}{\overline{y}^4C_u^4}$$

$$\gamma_1(v)=\frac{\mu_{03}}{\mu_{02}^{\ 3/2}}=\frac{\mu_{03}}{\overline{X}_2^{\ 3}C_v^3}$$

## II. Proposed Estimator

$$\overline{y}_{I,pps}=\overline{y}+b(\overline{X}-\overline{x})+k(\overline{x}-\frac{s_x}{c_x})\text{-----}(2.1)$$

Bias and Mean Square Error of $\overline{y}_{I,pps}$

Let

$\overline{y}=\overline{Y}+e_0, \overline{x}=\overline{X}+e_1$

$s_{uv}=S_{uv}+e_2, s_v^2=S_v^2+e_3$

So that

$E(e_0)=E(e_1)=E(e_2)=E(e_3)=0$

and $E(e_0^2)=\frac{S_u^2}{n}, E(e_1^2)=\frac{S_v^2}{n}, E(e_3^2)=\beta_1(v)\text{-}1\frac{S_v^4}{n}$,

$E(e_0e_1)=\frac{S_{uv}}{n}=\frac{\beta S_v^2}{n}, E(e_0e_3)=\frac{\mu_{12}}{n}, E(e_1e_2)=\frac{\mu_{12}}{n}, E(e_1e_3)=\frac{\mu_{03}}{n}$

From (2.1)we have

$$\overline{y}_{I,pps}=(\overline{Y}+e_0)+(-e_1)(\frac{S_{uv}+e_2}{S_v^2+e_3})+k[\overline{X}+e_1-\overline{X}_2\frac{(S_v^2+e_3)^{1/2}}{S_v}]$$

$$(\overline{y}_{I,pps}-\overline{Y})=e_0-\beta e_1(1+\frac{e_2}{S_{uv}})(1+\frac{e_3}{S_v^2})^{-1}+k[\overline{X}+e_1-\overline{X}_2(1+\frac{e_3}{S_v^2})^{1/2}$$

Expanding $(1+\frac{e_3}{S_v^2})^{-1}, (1+\frac{e_3}{S_v^2})^{1/2}$ and multiplying out and neglecting the terms of $e_i^{'s}$ of higher degree we obtain.

$$(\overline{y}_{I,pps}-\overline{Y})=(e_0-\beta e_1)+k(\overline{X}+e_1)-k\overline{X}_2(1+\frac{e_3}{2S_v^2}-\frac{e_3^{\ 2}}{8S_v^4})+\beta(\frac{e_1e_3}{S_v^2}-\frac{e_1e_2}{S_{uv}})$$

$$(\overline{y}_{I,pps}-\overline{Y})=(e_0-\beta e_1)+k(\overline{X}-\overline{X}(_2)+ke_1-k\overline{X}_2(\frac{e_3}{2S_v^2})+\beta(\frac{e_1e_3}{S_v^2}-\frac{e_1e_2}{S_{uv}})+\frac{k\overline{X}_2e_3^2}{8S_v^4}+\text{.......}(2.2)$$

taking expecation on both sides of (2.2) we have the bias of $(\overline{y}_{I,pps})$upto higher order

$$Bias(\overline{y}_{I,pps})=\frac{\beta}{n}(\frac{\mu_{03}}{S_v^2}-\frac{\mu_{12}}{S_{uv}})+\frac{k\overline{X}_2}{8n}(\beta_1(v)\text{-}1)$$

$$=\frac{1}{n}[(\beta\frac{\mu_{03}}{S_v^2}-\frac{\mu_{12}}{S_{uv}})+\frac{k\overline{X}_2}{8}(\beta_1(v)\text{-}1)]\text{-----}(2.3)$$

Squaring both sides of (2.2) and taking expectation upto terms of higher order is

$$MSE(\bar{y}_{I,pps})=(e_0-\beta e_1)^2+k^2(e_1-\frac{\bar{X}_2e_3}{2S_v^2})^2+\beta^2(\frac{e_1e_3}{S_v^2}-\frac{e_1e_2}{S_{uv}})^2+\frac{k^2\bar{X}_2^2e_3^4}{64S_v^{16}}+2k(e_0-\beta e_1)(e_1-\frac{\bar{X}_2e_3}{2S_v^2})$$

$$MSE(\bar{y}_{I,pps})=E(e_0)^2+\beta^2E(e_1)^2-2\beta E(e_0e_1)+k^2[E(e_1)^2+E(\frac{\bar{X}_2^2e_3^2}{4S_v^4})-\frac{\bar{X}_2E(e_1e_3)}{S_v^2}]+2k(e_0-\beta e_1)(e_1-\frac{\bar{X}_2e_3}{2S_v^2})$$

$$=E(e_0)^2+\beta^2E(e_1)^2-2\beta E(e_0e_1)+k^2[E(e_1)^2+E(\frac{\bar{X}_2^2e_3^2}{4S_v^4})-\frac{\bar{X}_2E(e_1e_3)}{S_v^2}]+2kE[e_0e_1-\frac{\bar{X}_2(e_0e_3)}{2S_v^2}+\frac{\beta\bar{X}_2(e_0e_3)}{2S_v^2}-\beta(e_1^2)]$$

$$=\frac{\bar{Y}^2C_u^2(1-\rho^2)}{n}+k^2[\frac{S_v^2}{n}+\frac{\bar{X}_2^2}{4n}\beta_1(v)-1-\frac{\bar{X}_2}{S_v^2}\frac{\mu_{03}}{n}]+2k[\frac{S_{uv}}{n}-\frac{\mu_{12}}{2n\bar{X}_2C_v^2}+\frac{\beta\bar{X}_2\mu_{03}}{2nS_v^2}-\beta\frac{S_v^2}{n}]$$

$$=\frac{\bar{Y}^2C_u^2(1-\rho^2)}{n}+\frac{k^2\bar{X}_2^2}{4n}[4C_v^2+\beta_1(v)-1-4\gamma_1(v)C_v]+k[\frac{S_{uv}\gamma_1(v)}{nC_v}-\frac{\mu_{12}}{n\bar{X}_2C_v^2}]$$

$$=\frac{\bar{Y}^2C_u^2(1-\rho^2)}{n}+\frac{k^2\bar{X}_2^2}{4n}[4C_v^2+\beta_1(v)-1-4\gamma_1(v)C_v]+k[\frac{\bar{X}_2C_v\gamma_1(v)S_{uv}-\mu_{12}}{n\bar{X}_2C_v^2}] \text{----(2.4)}$$

The optimum value of k&MinimumMean Square Error of ($\bar{y}_{I,pps}$) in (2.4) is given by

$$k_{opt}=\frac{2\bar{X}_2C_v\gamma_1(v)S_{uv}-\mu_{12}}{\bar{X}_2^3C_v^2[4C_v^2+\beta_1(v)-1-4\gamma C_v}\text{------(2.5)}$$

$$\text{MinimumMSE}(\bar{y}_{I,pps})=\frac{\bar{Y}^2C_u^2(1-\rho^2)}{n}-\frac{[\bar{X}_2C_v\gamma(v)S_{uv}-\mu_{12}]^2}{n\bar{X}_2^4C_v^4[4C_v^2+\beta_1(v)-1-4\gamma(v)C_v]}------(2.6)$$

## III. Estimator Based on Estimated optimum K

The value of k may not be known for us practically so we replace by its estimated $S_{uv,}$ $\mu_{12,}$ ,$\gamma_1(v)$ ,$\beta_1(v)$, $\mu_{03}$ &$\mu_{04}$ is involved in optimum K is

$$\hat{S}_{uv}=s_{uv}=\sum_{i=1}^{n}p_{1i}(u_i-\bar{u}_n)(v_i-\bar{v}_n)$$

$$\hat{\mu}_{12}=\sum_{i=1}^{n}p_{1i}(u_i-\bar{u}_n)(v_i-\bar{v}_n)^2$$

$$\hat{\gamma}_1(v)=\frac{\hat{\mu}_{03}}{\bar{X}_2^3C_v^3},\hat{\beta}_1(v)=\frac{\hat{\mu}_{04}}{\bar{X}_2^4C_v^4}$$

$$\hat{\mu}_{03}=\sum_{i=1}^{n}p_{1i}(v_i-\bar{v}_n)^3$$

$$\hat{\mu}_{04}=\sum_{i=1}^{n}p_{1i}(v_i-\bar{v}_n)^4$$

we get the estimated optimum $\hat{k}$ as

$$\hat{k}=-\frac{2[\bar{X}_2C_v\hat{\gamma}_1(v)s_{uv}-\hat{\mu}_{12}]}{\bar{X}_2^3C_v^2[4C_v^2+\hat{\beta}_1(v)-1-4\hat{\gamma}_1(v)C_v]}$$

$$\hat{k}=-\frac{2[\frac{\bar{X}_2C_v\hat{\mu}_{03}s_{uv}}{\bar{X}_2^3C_v^3}-\hat{\mu}_{12}]}{\bar{X}_2^3C_v^2[4C_v^2+\frac{\hat{\mu}_{04}}{\bar{X}_2^4C_v^4}-1-4\frac{\hat{\mu}_{03}C_v}{\bar{X}_2^3C_v^3}]}\text{------(3.1)}$$

Estimator of poppulation Mean $\bar{Y}$ based on estimated optimum $\hat{k}$ is

$$\bar{y}_{I,pps}=\bar{y}+b(\bar{X}-\bar{x})+\hat{k}(\bar{x}-\frac{s_x}{c_x})\text{-----(3.2)}$$

let $\hat{\mu}_{03}=\mu_{03}+e_4,\hat{\mu}_{12}=\mu_{12}+e_5,\hat{\mu}_{04}=\mu_{04}+e_6$

so that from (3.1),we get

$$\hat{k}=-\frac{\frac{2[\bar{X}_2C_v(\mu_{03}+e_4)(s_{uv}+e_2)}{\bar{X}_2^3C_v^3}-\mu_{12}+e_5]}{\bar{X}_2^3C_v^2[4C_v^2+\frac{\mu_{04}+e_6}{\bar{X}_2^4C_v^4}-1-\frac{4(\mu_{03}+e_5)C_v}{\bar{X}_2^3C_v^3}]}$$

$$\hat{k}=-\frac{2[\frac{\bar{X}_2C_v\mu_{03}s_{uv}}{\bar{X}_2^3C_v^3}+\frac{\mu_{03}e_2+s_{uv}e_4+e_2e_4}{\bar{X}_2^2C_v^2}-\mu_{12}-e_5]}{\bar{X}_2^3C_v^2[4C_v^2+\frac{\mu_{04}}{\bar{X}_2^4C_v^4}+\frac{e_6}{\bar{X}_2^4C_v^4}-1-\frac{4\mu_{03}C_v}{\bar{X}_2^3C_v^3}-\frac{4e_5}{\bar{X}_2^3C_v^2}]}$$

$$=-\frac{2[\bar{X}_2C_v\gamma_1(v)s_{uv}-\mu_{12}+\frac{\mu_{03}e_2+s_{uv}e_4+e_2e_4-\bar{X}_2^2C_v^2e_5}{\bar{X}_2^2C_v^2}}{\bar{X}_2^3C_v^2[4C_v^2+\beta_1(v)-1-4\gamma_1(v)C_v+\frac{e_6-4\bar{X}_2C_v^2e_4}{\bar{X}_2^4C_v^4}]}$$

$$=-\frac{2[\bar{X}_2C_v\gamma_1(v)s_{uv}-\mu_{12}]}{\bar{X}_2^3C_v^2[4C_v^2+\beta_1(v)-1-4\gamma_1(v)C_v]}[1+\frac{\mu_{03}e_2+s_{uv}e_4+e_2e_4-\bar{X}_2^2C_v^2e_5}{\bar{X}_2^2C_v^2[\bar{X}_2C_v\gamma_1(v)s_{uv}-\mu_{12}]}][1+\frac{e_6-4\bar{X}_2C_v^2e_4}{\bar{X}_2^4C_v^4[4C_v^2+\beta_1(v)-1-4\gamma_1(v)C_v]}]^{-1}$$

$$=-\frac{2[\bar{X}_2C_v\gamma_1(v)s_{uv}-\mu_{12}]}{\bar{X}_2^3C_v^2[4C_v^2+\beta_1(v)-1-4\gamma_1(v)C_v]}[1+\frac{\mu_{03}e_2+s_{uv}e_4+e_2e_4-\bar{X}_2^2C_v^2e_5}{\bar{X}_2^2C_v^2[\bar{X}_2C_v\gamma_1(v)s_{uv}-\mu_{12}]}-\frac{e_6-4\bar{X}_2C_v^2e_4}{\bar{X}_2^4C_v^4[4C_v^2+\beta_1(v)-1-4\gamma_1(v)C_v]}+...]\text{----(3.3)}$$

from (3.2) we have

$$(\bar{y}_{Ipps}-\bar{Y})=(e_0-\beta e_1)+\hat{k}(\bar{X}-\bar{X}_2)+\hat{k}(e_1-\bar{X}_2\frac{e_3}{2S_v^2})+\beta(\frac{e_1e_3}{S_v^2}-\frac{e_1e_2}{S_{uv}})+\frac{kX_2e_3^2}{8S_v^4}+\ldots\text{-----}(3.4)$$

Substituting $\hat{k}$ from (3.3) in (3.4), Squaring both sides and taking expectation, the MSE of $\bar{y}_{Ipps}$ to the terms of higher order is

$$MSE(\bar{y}_{Ipps})=E[(e_0-\beta e_1)-\frac{2[\bar{X}_2C_v\gamma_1(v)s_{uv}-\mu_{12}]}{\bar{X}_2^3C_v^2[4C_v^2+\beta_1(v)-1-4\gamma_1(v)C_v]}(e_1-\bar{X}_2\frac{e_3}{2S_v^2})]^2$$

$$=\frac{\bar{Y}^2C_u^2(1-\rho^2)}{n}+\frac{[\bar{X}_2C_v\gamma(v)S_{uv}-\mu_{12}]^2}{n\bar{X}_2^4C_v^4[4C_v^2+\beta_1(v)-1-4\gamma(v)C_v]}-\frac{2[\bar{X}_2C_v\gamma(v)S_{uv}-\mu_{12}]^2}{n\bar{X}_2^4C_v^4[4C_v^2+\beta_1(v)-1-4\gamma(v)C_v]}$$

$$=\frac{\bar{Y}^2C_u^2(1-\rho^2)}{n}-\frac{[\bar{X}_2C_v\gamma(v)S_{uv}-\mu_{12}]^2}{n\bar{X}_2^4C_v^4[4C_v^2+\beta_1(v)-1-4\gamma(v)C_v]}\text{-----}(3.5)$$

Which shows that the mean square error of the estimator $\bar{y}_{Ipps}$ in (3.2) based on the estimated optimum tothe terms of higher order is same as $\bar{y}_{Ipps}$ *in* (2.6).

## IV. Comparison of the proposed Estimator

4.11 $MSE(\bar{y}_{Ipps})<MSE(\bar{y})$

$$\frac{\bar{Y}^2C_u^2(1-\rho^2)}{n}-\frac{[\bar{X}_2C_v\gamma_1(v)S_{uv}-\mu_{12}]^2}{n\bar{X}_2^4[4C_v^2+\beta_1(v)-1-4\gamma_1(v)C_v]}<\frac{\bar{Y}^2C_u^2}{n}$$

$$-\frac{\rho^2\bar{Y}^2C_u^2}{n}-\frac{[\bar{X}_2C_v\gamma_1(v)S_{uv}-\mu_{12}]^2}{n\bar{X}_2^4[4C_v^2+\beta_1(v)-1-4\gamma_1(v)C_v]}<0$$

4.12 $MSE(\bar{y}_{Ipps})<MSE(\bar{y}_R)$

$$\frac{\bar{Y}^2C_u^2(1-\rho^2)}{n}-\frac{[\bar{X}_2C_v\gamma_1(v)S_{uv}-\mu_{12}]^2}{n\bar{X}_2^4[4C_v^2+\beta_1(v)-1-4\gamma_1(v)C_v]}<\frac{\bar{Y}^2}{n}[C_u^2+C_v^2-2\rho C_uC_v]$$

$$-\frac{\bar{Y}^2[\rho C_u-C_v]^2}{n}-\frac{[\bar{X}_2C_v\gamma_1(v)S_{uv}-\mu_{12}]^2}{n\bar{X}_2^4[4C_v^2+\beta_1(v)-1-4\gamma_1(v)C_v]}<0$$

4.13 $MSE(\bar{y}_{Ipps})<MSE(\bar{y}_p)$

$$\frac{\bar{Y}^2C_u^2(1-\rho^2)}{n}-\frac{[\bar{X}_2C_v\gamma_1(v)S_{uv}-\mu_{12}]^2}{n\bar{X}_2^4[4C_v^2+\beta_1(v)-1-4\gamma_1(v)C_v]}<\frac{\bar{Y}^2}{n}[C_u^2+C_v^2+2\rho C_uC_v]$$

$$-\frac{\bar{Y}^2[\rho C_u+C_v]^2}{n}-\frac{[\bar{X}_2C_v\gamma_1(v)S_{uv}-\mu_{12}]^2}{n\bar{X}_2^4[4C_v^2+\beta_1(v)-1-4\gamma_1(v)C_v]}<0$$

4.14 $MSE(\bar{y}_{Ipps})<MSE(\bar{y}_{lr})$

$$\frac{\bar{Y}^2C_u^2(1-\rho^2)}{n}-\frac{[\bar{X}_2C_v\gamma_1(v)S_{uv}-\mu_{12}]^2}{n\bar{X}_2^4[4C_v^2+\beta_1(v)-1-4\gamma_1(v)C_v]}<\frac{\bar{Y}^2C_u^2(1-\rho^2)}{n}$$

$$-\frac{[\bar{X}_2C_v\gamma_1(v)S_{uv}-\mu_{12}]^2}{n\bar{X}_2^4[4C_v^2+\beta_1(v)-1-4\gamma_1(v)C_v]}<0$$

## V. Empirical Study

To compare proposed Estimator numerically with other estimator.The description of population are given below

POPULATION(source:Singh.D and Chaudhary F.S.Sample Survey Designs ,ppt-132

$N=8,\bar{X}_1=36.25,\bar{X}_2=33.182,\bar{Y}=38.875$

$\bar{X}_2^2=1101.045124,\bar{X}_2^3=36534.879305,\bar{X}_2^4=1212300.3651$

$\bar{Y}^2=1511.265,\bar{Y}^3=58750.4269,\bar{Y}^4=2283922.85$

$\sigma_u^2=1.463,\sigma_v^2=392.222,C_u^2=0.00096,C_v^2=0.00556,\rho_{uv}=-0.2683$

$\gamma_1(v)=4394.23077,\beta_1(v)=564.919931,\mu_{12}=218.372735$

Mean Square Errors of $\bar{y}$ ,$\bar{y}_r$ ,$\bar{y}_p$ ,$\bar{y}_{lr}$ ,$\bar{y}_{Ipps}$

$M(\bar{y})=0.181351,M(\bar{y}_r)=1.462148,M(\bar{y}_p)=1.001213$

$M(\bar{y}_{lr})=0.168297,M(\bar{y}_{Ipps})=0.163741$

Relative Efficiencies of $\bar{y}$ ,$\bar{y}_r$ ,$\bar{y}_p$ ,$\bar{y}_{lr}$

$\bar{y}=110.754,\bar{y}_r=892.963$

$\bar{y}_p=611.461,\bar{y}_{lr}=102.782$

## VI. Conclusion

Thus the mean square error of our proposed estimator is less than the usual estimators Ratio, Product & Regression

## References

1.Cochran W.G. Sampling Techniq*ues* $3^{rd}$ *ed*. (1977) *wiley Eastern, New Delhi.*

2.*Daroga singh, F.S Chaudhary* (1986) *Theory & Analysis of sample survey Degins .New Delhi.*

3.*S.Sampath* (*Sampling theory & methods*)