Jour. Ind. Soc. Ag. Statistics
53(2), 2000 : 161-171

Methodological Investigation on Post-harvest Losses

H.C. Gupta, Jagbir Singh and O.P. Kathuria[1]
Indian Agricultural Statistics Research Institute, New Delhi-110012
(Received : July, 1999)

SUMMARY

Often considerable losses of agricultural produce due to various causes occur at various Post-harvest stages. The reliable and objective estimates of such losses based on a statistical methodology were not available. In this context, I.A.S.R.I. developed a suitable statistical methodology for the estimation of losses at various post-harvest stages by conducting a pilot sample survey in Bulandshahr district of U.P. state (India) during 1985-88 but the estimators developed in this survey were not precise. In this paper, Minimum Variance Linear Unbiased estimators were developed by Projective Geometry approach and applied to estimate the foodgrain losses at harvest and threshing stages. Percentage losses estimated by Usual as well as by Projective Geometry approach reveal that though percentage losses at harvest as well as at threshing stage were more or less of the same order but the percentage standard errors estimated by second approach were significantly less than that estimated by the first approach at both stages. Fluctuations in percentage losses between years have also been estimated.

Key words : Parametric functions, Minimum Variance Linear Unbiased Estimator, Projective Geometry approach, Hadamard product of matrices.

1. Introduction

Often considerable losses of agricultural produce due to various causes occur at various Post-harvest stages. As the reliable and objective estimates of food grain losses based on a statistical methodology were not available. Indian Agricultural Statistics Research Institute (I.A.S.R.I.) developed a suitable statistical methodology for the estimation of food grain losses at various Post-harvest stages by conducting a pilot sample survey in Bulandshahr district of U.P. state (India) during 1985-88. The sampling design used in the survey was based on successive sampling technique. Lot of literature is available on this approach. Some of the important references are due to Jessen [4], Yates [16], Patterson [8], Tikkiwal ([13], [14], [15]) Eckler [1], Rao and Graham [9], Narain

1 Institute of Statistics and Applied Economics, Makere University, Kampala, Uganda (South Africa)

et al. [7], Srivastava *et al.* [12], Holt and Skinner [3], Singh and Yadav [10], Lamba and Singh [5], Nandram [6], Singh and Kathuria [11] etc. Gurney and Daly [2] generalized the theory of successive sampling by breaking down the repeated surveys into elementary components (sub-samples) of the overlap pattern and by defining elementary estimate as one which does not make use of the survey data for any time period except that period to which the estimate refers. With this elementary estimate as a conceptual base, using Projective Geometry approach, they derived two matrices viz. a matrix **C** (Say) of the best weights to be used on all the elementary estimators in order to yield the Minimum Variance Linear Unbiased Estimators (MVLUEs) of level parameters and a dispersion matrix **D** (say) of these MVLUEs. In this paper Minimum Variance Linear Unbiased Estimators (MVLUEs) have been developed by Projective Geometry approach for the overlap pattern and these estimators are applied to obtain the food grain losses at harvest and threshing stages in the pilot sample survey as mentioned above.

2. *Estimator of* $\mathbf{f}(\theta_{zy})$

The estimator of parametric functions $f(\theta_{zy})$ *i.e.*, levels, changes, average, total, etc. of the parameters θ_{zy} (Percentage food grain losses at harvesting and threshing stages), belonging to the y-th year/occasion in the z-th zone under any sampling design are developed $\forall$ z = 1, 2, 3, ... , 9 and y = 1, 2, 3 as follows :

Let,

m_{zy} = Elementary estimate of the percentage losses at the harvesting/threshing stage for the y-th year in the z-th zone based on a sub-sample of cultivators matched over all the years.

m'_{zy} = Elementary estimate of the percentage losses at the harvesting/threshing stage for the y-th year in the z-th zone based on the sub-sample of cultivators selected afresh every year.

$$\mathbf{m} = (m_{11}\ m_{12}\ m_{13} \ldots m_{91}\ m_{92}\ m_{m_{93}})$$

$$\mathbf{u} = (m'_{11}\ m'_{12}\ m'_{13} \ldots m'_{91}\ m'_{92}\ m'_{93})$$

$$\mathbf{e} = (\mathbf{m}\ \ \mathbf{u})$$

$$s_{zy} = \sqrt{\hat{v}(m_{zy})}$$

$$s'_{zy} = \sqrt{\hat{v}(m'_{zy})}$$

$$\mathbf{s}_m = (s_{11}\ s_{12}\ s_{13} \ldots s_{91}\ s_{92}\ s_{93})$$

$$\mathbf{s}_u = (s'_{11}\ s'_{12}\ s'_{13} \ldots s'_{91}\ s'_{92}\ s'_{93})$$

$$\mathbf{s} = (\mathbf{s}_m\ \ \mathbf{s}_u)$$

$$\hat{\theta}_{zy} = m_{zy} \text{ and } m'_{zy} \text{ separately}$$

$$\theta = (\Sigma\, \theta_{zy}\, \mathbf{u}_{zy}),\ \forall\ z = 1, 2, \ldots, 9 \text{ and } y = 1, 2, 3$$

$$\mathbf{U} = [\mathbf{u}'_{11}\ \mathbf{u}'_{12}\ \mathbf{u}'_{13} \ldots \mathbf{u}'_{91}\ \mathbf{u}'_{92}\ \mathbf{u}'_{93}]'$$

$$= [\, I_{27\times 27}\ \ I_{27\times 27}\,]$$

$$\mathbf{R}_m = [\,\rho_{ij}\,] = \text{correlation matrix of m}'$$

$$\mathbf{R}_u = [\,\rho'_{ij}\,] = \text{correlation matrix of u}'$$

$$R = \begin{bmatrix} \mathbf{R}_m & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_u \end{bmatrix}$$

$$\mathbf{K} = \text{Dispersion matrix of } \mathbf{e}'$$

$$= \text{Hadamard product of } \mathbf{R} \text{ and s's}$$

$$= \begin{bmatrix} R_m\, \mathbf{s}'_m\, \mathbf{s}_m & 0 \\ 0 & R_u\, \mathbf{s}'_u\, \mathbf{s}_u \end{bmatrix}$$

$$= \begin{bmatrix} K_m & 0 \\ 0 & K_u \end{bmatrix} \text{(say)}$$

The matrix **C** of the best weights to be assigned to the elementary estimators **e′** and the dispersion matrix **D** of MVLUEs for θ′ are obtained as follows :

$$\mathbf{C} = \mathbf{U}'\,(\mathbf{U}\mathbf{K}^{-1}\,\mathbf{U}')^{-1}\,\mathbf{U}\mathbf{K}^{-1} \qquad \text{(Gurney and Daly [2])}$$

$$= [\,\mathbf{K}_m^{-1} + \mathbf{K}_u^{-1}\,]^{-1} \begin{bmatrix} \mathbf{I}_{27\times 27} & \mathbf{I}_{27\times 27} \\ \mathbf{I}_{27\times 27} & \mathbf{I}_{27\times 27} \end{bmatrix} \begin{bmatrix} \mathbf{K}_m^{-1} & 0 \\ 0 & \mathbf{K}_u^{-1} \end{bmatrix}$$

$\mathbf{I}_{27\times 27}$ being Identity matrix of order 27.

$$\mathbf{D} = \mathbf{U}'\ (\mathbf{U}\mathbf{K}^{-1}\,\mathbf{U}')^{-1}\,\mathbf{U} \qquad \text{(Gurney and Daly [2])}$$

$$= [\,\mathbf{K}_m^{-1} + \mathbf{K}_u^{-1}\,]^{-1} \begin{bmatrix} \mathbf{I}_{27\times 27} & \mathbf{I}_{27\times 27} \\ \mathbf{I}_{27\times 27} & \mathbf{I}_{27\times 27} \end{bmatrix}$$

In order to obtain the MVLUEs for the desired parametric functions ($\mathbf{B}\,\theta'$) of θ' for the Bulandshahr district as a whole, let us further define a matrix $\mathbf{B}$ (say) of the coefficients of θ'_{zy} in various parametric functions of θ' as follows :

$$\mathbf{B} = \begin{bmatrix} \mathbf{I} & \mathbf{I} & \mathbf{I} & \mathbf{I} & \mathbf{I} & \mathbf{I} & \mathbf{I} & \mathbf{I} & \mathbf{I} \\ \mathbf{E} & \mathbf{E} & \mathbf{E} & \mathbf{E} & \mathbf{E} & \mathbf{E} & \mathbf{E} & \mathbf{E} & \mathbf{E} \end{bmatrix}$$

where $\mathbf{I}$ = Identity matrix of order 3 and

$$\mathbf{E} = \begin{bmatrix} 0 & -1 & 1 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

Then the MVLUEs for the parametric functions $\mathbf{B}\theta'$ are given by

$$\mathbf{B}\hat{\theta}' = \mathbf{BCe'}$$

and the corresponding dispersion matrix by

$$V(\mathbf{B}\hat{\theta}') = \mathbf{BDB'} \qquad \text{(Singh and Kathuria [11])}$$

3. Estimation of Food Grain Losses

The estimators developed in Section 2 were utilised for estimating the losses of wheat crop at harvest and threshing stages in a survey "Pilot sample survey for estimation of post harvest food grain losses (wheat)" conducted by I.A.S.R.I., in Bulandshahr district (U.P.) during 1985-88. In this survey stratified two stage random sampling design was adopted. A sample of 4 villages as psu's was selected by SRSWOR in each of the 9 zones/strata of Bulandshahr district. Out of these 4 villages, a set of 2 villages was matched over all the three years of the survey and remaining 2 villages were selected afresh every year. The cultivators growing wheat crop in each selected village were classified into three categories (I, II, III) *i.e.*, cultivators having area under wheat crop upto 1 ha., having area > 1 ha and ≤ 2 ha. and having area > 2 ha. respectively. Then from each category, 2 cultivators were selected as ssu's by SRSWOR. Thus 24 cultivators in all constituting the sample were observed for collection of data on wheat crop losses at harvest and threshing stages. The estimates of losses obtained by using the estimators developed in Section 2 were also compared with the estimates obtained by using the estimators given in Section 3.1 & 3.2 in which information on previous occasion is not utilised.

3.1 Estimate of Losses at Harvest

Let P_{ijh} be the estimated yield per hectare based on crop cutting experiment conducted in the field of the j^{th} cultivator of h^{th} size category of the i^{th} village and a_{ijh} be the area in hectares under the wheat crop in the holding, the estimate of yield per hectare for the zone/stratum is given by

$$\hat{P} = \frac{\sum_{i=1}^{n} \hat{P}_i}{\sum_{i=1}^{n} A_i}$$

where $$\hat{P}_i = \sum_{h=1}^{3} A_{ih} \frac{\sum_{j=1}^{m_{ih}} P_{ijh}\, a_{ijh}}{\sum_{j=1}^{m_{ih}} a_{ijh}}$$

A_{ih} and m_{ih} being the area in hectares under wheat crop and number of cultivators selected in the h^{th} size category of i^{th} village respectively. It may be noted that in this case $n = 4$ and $m_{ih} = 2$.

Also let I_{ijh} be the estimated loss per hectare based on crop cutting experiment conducted in the field of the j^{th} cultivator of h^{th} size category of the i^{th} village, the estimate of loss per hectare can be obtained by replacing P_{ijh} by I_{ijh} in the above formula. The loss in the randomly selected plot was worked out after the harvested crop was removed from the plot and all grains shedded or missed were picked up carefully. The percentage loss at harvest for the stratum/zone is given by

$$\hat{L}\% = \frac{\hat{L}}{\hat{P}} \times 100$$

Estimate of variance of $(\hat{L}\%)$ will be given by

$$\hat{V}(\hat{L}\%) = \left(\frac{\hat{L}}{\hat{P}} \times 100\right)^2 \left\{\frac{\hat{V}(\hat{L})}{(\hat{L})^2} + \frac{\hat{V}(\hat{P})}{(\hat{P})^2}\right\}$$

ignoring the covariance term.

The estimate of variance of $\hat{P}$ and $\hat{L}$ are obtained as

$$\hat{V}(\hat{\bar{X}}) = \frac{1}{n(n-1)\bar{A}^2} \sum_{i=1}^{n} \sum_{h=1}^{3} A_{ih}^2 (\hat{\bar{X}}_{ih} - \hat{\bar{X}}_h)^2$$

where X is the variable (loss or yield) and $\hat{\bar{X}}_{ih}$ stands for the estimate of average X for h^{th} size category of the i^{th} village, $\bar{A}$ and $\hat{\bar{X}}_h$ being the average area per village and the estimate of average X for h^{th} size category respectively in the stratum/zone. The finite population correction factor is ignored because the sampling fractions are expected to be small.

The estimate of average over strata/zones for the district was obtained as a weighted average of stratum wise estimates, the estimates of production in the respective strata serving as weights.

The variance of the weighted average was calculated from stratum wise estimates of variance.

3.2 Estimate of Losses at Threshing

At this stage, the loss can be occurred due to three reasons *i.e.*, (i) Loss due to grains remaining in the left-over straw, (ii) Loss due to grains fallen on the threshing floor and (iii) Loss due to grains remaining unthreshed in the left-over material. If x_{ijh} is the percentage loss due to any of the three reasons for the j^{th} cultivator of h^{th} size category of the i^{th} village, the estimate of average for the stratum/zone is given by

$$\hat{\bar{X}} = \frac{\sum_{i=1}^{n} \hat{X}_i}{\sum_{i=1}^{n} P_i}$$

$$\hat{X}_i = \sum_{h=1}^{3} P_{ih} \frac{\sum_{j=1}^{m_{ih}} P_{ih} X_{ih}}{\sum_{j=1}^{m_{ih}} P_{ijh}}$$

where P_{ijh}, P_{ih} and m_{ih} being the yield of crop of the j^{th} cultivator of h^{th} size category of the i^{th} village, the estimated yield of grains of cultivators belonging to h^{th} size category of the i^{th} village and the number of cultivators selected from h^{th} size category of i^{th} selected village respectively.

Estimated variance of $\hat{\bar{X}}$ will be

$$\hat{V}(\hat{\bar{X}}) = \frac{1}{n(n-1)\bar{P}'^2} \sum_{i=1}^{n} \sum_{h=1}^{3} P_{ih}^2 (\hat{\bar{X}}_{ih} - \hat{\bar{X}}_h)^2$$

where $\hat{\bar{X}}_{ih}$ is the estimate of average percentage of loss for the cultivators belonging to h^{th} size category in the i^{th} village and $\bar{P}'$ is the average yield per village in the stratum/zone.

4. *Results and Discussion*

4.1 Losses at Harvest

(i) Year wise % losses estimated by Usual as well as Projective Geometry approach are given in Table 1.

Table 1. Percentage loss at harvest

Year	Usual Method		Projective Geometry approach		Gain (%) of II method over I method	No. of Observations
	Loss (%)	S.E. (%)	Loss (%)	S.E. (%)		
1985	0.48	14.58	0.54	12.78	14.08	216
1986	0.31	9.68	0.34	8.03	20.55	214
1987	0.72	11.11	0.70	9.02	23.17	214
Over all	0.50	8.00	0.53	6.50	23.08	644

Table 1 indicates that though the percentage losses estimated by the two methods are more or less of the same order but the percentage S.E.s estimated by second method are less than that estimated by the first method in all the three years as well as pooled over the years. Results also indicate that percentage gain in efficiency of II method over the I method ranged from 14.08 to 23.08 indicating that the Projective Geometry approach is superior than Usual method for estimating loss at harvesting.

(ii) Size category wise percentage losses estimated by Usual method as well as Projective Geometry approach are given in Table 2.

Table 2. Percentage loss at harvest (Size category wise)

Size Category	Year	Usual Method		Projective Geometry approach		Gain (%) of II method over I method	No. of Observations
		Loss (%)	S.E. (%)	Loss (%)	S.E. (%)		
I	1985	0.48	20.83	0.46	16.73	24.50	77
	1986	0.28	17.86	0.27	16.06	11.21	79
	1987	0.74	29.73	0.80	25.84	15.05	78
	Over all	0.50	18.00	0.50	12.78	40.85	234
II	1985	0.45	31.11	0.48	27.06	14.97	72
	1986	0.30	23.33	0.33	21.18	10.15	71
	1987	0.81	34.56	0.77	23.42	47.57	70
	Over all	0.53	22.64	0.53	15.81	43.20	213
III	1985	0.47	40.42	0.63	30.12	34.20	67
	1986	0.32	25.00	0.37	18.91	32.21	64
	1987	0.70	34.28	0.65	23.62	45.13	66
	Over all	0.50	22.00	0.55	17.77	23.80	197

Table 2 indicates that percentage losses estimated by first as well as second method do not differ significantly between size categories but differ significantly between years within size category. It was also observed that percentage S.E.'s estimated by second method are significantly less than the corresponding percentage S.E.'s estimated by first method. In this case also the percentage gain in efficiency indicates the supremacy of II method of estimation over the I method.

4.2 Fluctuation at Harvest

Estimate of fluctuation in loss (%) at harvest is given in Table 3.

Estimate of fluctuations in percentages losses is maximum (0.36) between 1987 and 1986 whereas it is minimum (0.15) between 1987 and 1985.

Table 3. Estimate of fluctuation (%) at harvest

Between Years	Loss (%)	S.E. (%)	No. of Observations
1986 and1985	0.21	32.85	430
1987 and 1986	0.36	16.16	428
1987 and 1985	0.15	54.06	430

4.3 Losses at Threshing

(i) Table 4 represents the year wise percentage total losses estimated by Usual method as well as Projective Geometry approach.

Table 4. Percentage total loss at threshing

Year	Usual Method		Projective Geometry approach		Gain (%) of II method over I method	No. of Observa-tions
	Loss (%)	S.E. (%)	Loss (%)	S.E. (%)		
1985	0.23	21.74	0.18	2.34	*	216
1986	0.41	7.32	0.41	4.62	58.44	209
1987	0.32	12.50	0.30	7.42	68.46	212
Over all	0.34	5.88	0.30	2.90	102.76	637

Note : * indicate that value is exceptionally high.

Table 4 reveals that in case of threshing also the percentage losses estimated by two approaches were more or less of the same order but percentage S.E.s estimated by the second approach were significantly less than the percentage S.E.s estimated by the first approach. The percentage gain also indicates that in case of threshing also the II method is much superior than the I method.

(ii) Table 5 given below represents the size category wise percentage losses estimated by Usual method as well as Projective Geometry approach.

Table 5 indicates that the percentage S.E.s estimated by Projective Geometry approach are significantly less than that estimated by Usual method. The percentage gain given in the table also indicate that Projective Geometry approach is much better than the usual method.

Table 5. Percentage total loss at threshing (Size category wise)

Size Category	Year	Usual Method		Projective Geometry approach		Gain (%) of II method over I method	No. of Observations
		Loss (%)	S.E. (%)	Loss (%)	S.E. (%)		
I	1985	0.23	47.83	0.19	26.23	82.35	77
	1986	0.40	12.50	0.44	10.28	21.60	77
	1987	0.37	18.92	0.34	13.01	45.43	76
	Over all	0.36	11.11	0.32	7.33	51.57	230
II	1985	0.23	21.74	0.14	9.64	125.52	72
	1986	0.44	27.27	0.51	10.52	159.22	69
	1987	0.40	25.00	0.39	13.03	91.86	69
	Over all	0.39	17.95	0.35	5.99	199.67	210
III	1985	0.20	45.00	0.22	23.31	93.05	67
	1986	0.41	9.76	0.40	6.81	43.32	63
	1987	0.29	31.03	0.26	16.24	91.07	67
	Over all	0.32	12.50	0.29	7.22	73.13	197

4.4 Fluctuation at Threshing

Estimate of fluctuation (%) at threshing is given in Table 6.

Table 6. Estimate of fluctuation (%) at threshing

Between Years	Loss (%)	S.E. (%)	No. of Observations
1986 and1985	0.22	8.43	1275
1987 and 1986	0.10	27.66	1263
1987 and 1985	0.11	19.16	1284

Estimate of fluctuation in percentage total loss was observed maximum (0.22) between 1986 and 1985 and minimum (0.10) between 1987 and 1986.

5. Conclusion

The study reveals that though the food grain losses at harvest and threshing stages can be estimated by both Usual method and Projective Geometry approach, Projective Geometry approach provides the linear unbiased estimates with minimum variance indicating that the later approach is superior to Usual method. This fact is also proved by the application of both methods in estimating the percentage losses along with their percentage S.E.s and percentage gain in efficiency of Projective Geometry approach over the Usual method.

REFERENCES

[1] Eckler, A.R. (1955). Rotation sampling. *A.M.S.*, **26**, 664-685.

[2] Gurney, M. and Daly, J.F. (1965). A multivariate approach to estimation in periodic sample surveys. *Proc. Soc. Stat. Sec. Amer. Stat. Assoc.*, 242-257.

[3] Holt, D. and Skinner, C.J. (1989). Components of change in surveys. *Intt. Stat .Rev.*, **57(1)**, 1-18.

[4] Jessen, R.J. (1942). Statistical investigation of a sample survey for obtaining farm facts. *lowa Agri. Expt. Station Res. Bull.*, **304.**

[5] Lamba, I.M.S. and Singh, P. (1993). Estimation of prevalence and relative risk in repetitive survey. *Biom. J.*, **35**, 877-892.

[6] Nandram, Balgobin (1994). Bayesian Predictive Inference for multivariate sample surveys. *J. Off. Stat.*, **10(2)**, 167-179.

[7] Narain, P., Kathuria, O.P. and Srivastava, A.K. (1987). Measuring change in Agricultural Surveys. Invited paper, 46th Session of the International Statistical Institute, Tokyo. 1-15.

[8] Patterson, H.D. (1950). Sampling on successive occasions with partial replacement of units. *J. Roy. Stat. Soc.*, **Sr B 12**, 241-255.

[9] Rao, J.N.K. and Graham, J.E. (1964). Rotation design for sampling on repeated occasons. *Jour. Amer. Stat. Assoc.*, **59**, 492-509.

[10] Singh, Padam and Yadav, R.J. (1992). Generalized estimation under successive sampling. *J. Ind. Soc. Ag. Stat.*, **1**, 27-36.

[11] Singh, Jagbir and Kathuria, O.P. (1997). Estimation of parametic functions in repeat surveys. *Jour. Ind. Soc. Ag. Stat.*, **50(1)**, 18-28.

[12] Srivastava, A.K., Khosla, R.K., Narain, P., Gupta, H.C., Sone Lal, Srivastava, J.L., Nagar, A.K. and Gupta, O.P. (1988). Pilot sample survey for estimation of post harvest foodgrain losses (wheat). Project report-an IASRI publication.

[13] Tikkiwal, B.D. (1951). Theory of successive sampling. Unpublished thesis for Diploma, I.C.A.R., New Delhi.

[14] Tikkiwal, B.D. (1960). On the theory of classical regression and double sampling estimation. *J. K. Statist. Soc.*, **B 22**, 131-138.

[15] Tikkiwal, B.D. (1979). Successive sampling-a review. Proc. 42nd session ISI. **48, II**, 367-384.

[16] Yates, F. (1949). *Sampling Methods For Censuses and Surveys.* Charles Griffin and Company Ltd., London.