SOME INVESTIGATIONS ON THE USE OF SUCCESSIVE SAMPLING IN PEST AND DISEASE SURVEYS

By

T.P. Abraham,* R.K. Khosla and O.P. Kathuria
Institute of Agricultural Research Statistics, New Delhi-12

(Received in April, 1968)

1. Introduction and Summary

In repeat sample surveys, the application of successive sampling technique with partial replacement of sampling units on the subsequent occasions has certain advantages. The successive sampling technique has been developed mainly by Jessen (1942), by Patterson (1950) and by Tikkiwal (1950, 53, 56, 64, 65, 67). Singh and Kathuria (1969) investigated the application of the technique to an enquiry when the sampling design adopted in a survey is multi-stage.

Surveys to estimate the incidence of pests and diseases on field crops have to be generally repeated due to the large variation in the incidence from year to year. It is, therefore, interesting to examine the utility of partial replacement of units in such repeat surveys, specially when taking some of the sampling units common from one year to the other is operationally convenient. In particular, we examine how far partial matching of the sampling units is helpful in obtaining a better estimate of (*i*) the incidence in the second year of the survey, (*ii*) the changes in their occurrence from one year to the other and (*iii*) overall mean incidence over the two years. For this purpose, we present, in Sections 2 and 3, the details of sampling technique and of data recording of a pilot sample survey. In Section 4, we give various estimators, based on matching and the variances of these estimators. We also discuss in this section, the relative efficiencies of these estimators with respect of the corresponding estimators based on

*At present at Soil Institute and Pilot Development Project, FAO Mission in Iran, Teheran.

no matching. In Section 5, the analysis of pertinent data from the pilot survey is given. From this analysis, we note that partial matching of units on the second occasion does provide improved estimators for incidence of certain pests and diseases.

2. Sampling Technique

A pilot sample survey was conducted by the Institute of Agricultural Research Statistics in Cuttack district of Orissa State from 1959-60 to 1961-62 to estimate the incidence of pests and diseases on rice crop. The sampling design adopted in the survey consisted of stratified multi-stage random sampling. The district was divided into 10 homogeneous zones. These zones constituted the strata. In each stratum, 6 villages as primary sampling units (psu) growing paddy crop were selected at random and in each selected village, 4 fields as secondary sampling units (ssu), growing paddy were taken at random during the first year. In the second year, 3 villages were retained in each stratum and 3 villages were selected afresh at random from among remaining villages and further in each retained village, 2 fields were retained in the second year and 2 fields were selected afresh at random from the remaining fields in the village. Since the number of matched psu's and ssu's (within psu) were not sufficient in each stratum, the investigations have been made ignoring stratification in the district. This may disturb the results of this investigation to some extent. For recording observations on incidence of pests and diseases, 4 plots each of size about 0.84 sq. metre, were located at random in each selected field (ssu).

3. Recording of Data

The major pests observed during the survey in Cuttack district were Stemborer (*Tryporyza incertulas*) and Gallfly (*Pachydiplosis oryzae*). The major disease was Helminthosporium (*Helminthosporium oryzae*). In each of the selected fields, periodical observations, on the various pests and diseases, were taken at an interval of about a month upto and including harvest. The first observation was taken at about a month after planting.

In each of the plots the number of clumps (plants) and the total number of tillers were counted. The number of dead hearts due to Stemborer and silver shoots affected by Gallfly incidence were noted. In case of Helminthosporium disease, four corner plants and the central plant were taken from each of the plots instead of all the plants in a plot, and on each of these plants the maximum infected leaf was taken. The intensity of Helminthosporium infection for the selected leaf was scored by comparing it with standard charts supplied by the Central Rice Research Institute, Cuttack. There are about 10 grades for scoring. At the time of harvest the number of clumps (plants), number of earheads and the number of white earheads due to Borers were counted. Field-wise average percentages of incidence due to major pests and diseases were worked out. The study was confined to only those monthly observations for different pests and diseases, when the respective incidences were at their peaks.

4. Different Linear Estimators and Their Efficiencies

Before we investigate the use of successive sampling technique in incidence of pests and diseases survey referred to in Section 2, we discuss various estimators and their efficiencies, in general, with particular reference to the corresponding estimators based on no matching.

Let $\bar{X}_{..}$ and $\bar{Y}_{..}$ denote the population means on the two occasions for a particular character under study. Let N and M denote the number of psu's and ssu's in the population and n and m denote the corresponding values in the sample. On the second occasion a fraction p of the n psu's and in each of np psu's a fraction r of the ssu's was retained. The remaining units, *viz.*, fraction $q\,(p+q=1)$ of the n psu's, all the ssu's in nq psu's and fraction $s\ (r+s=1)$ of the ssu's in np primary sampling units were replaced by the units selected afresh on the second occasion.

Let $\bar{x}'$ = mean per ssu on the first occasion based on $npmr$ units which are common to the two occasions,

$\bar{y}'$ = mean per ssu based on these units on the second occasion,

$\bar{x}''$ = mean per ssu on the first occasion based on $npms$ fresh units taken from common psu on both occasions,

$\bar{y}''$ = mean per ssu on the second occasion based on these units,

$\bar{x}'''$ = mean per ssu on the first occasion based on nqm units which are in the sample on the first occasion only,

$\bar{y}'''$ = mean per ssu on the second occasion based on nqm units selected afresh on the second occasion.

(*a*) Estimator I : For estimating the mean on the second occasion

Let the first unbiased estimator $\bar{y}_2$, of $\bar{Y}_{..}$, be given by

$$(4.1)\quad \bar{y}_2 = a\bar{x}' + b\bar{x}'' - (a+b)\,\bar{x}''' + d\bar{y}' + e\bar{y}'' + (1-d-e)\,\bar{y}'''.$$

where a, b, d and e are some constants to be chosen in such a way that $V(\bar{y}_2)$ is minimum. We assume that S_b^2, the mean square between the psu's, and S_w^2, the mean square between the ssu's are the same on each of the two occasions. Let,

$$\rho_1 S_b^2 = \frac{1}{(N-1)}\left[\sum_{i=1}^{N}\left(\bar{Y}_{i.} - \bar{Y}_{..}\right)\left(\bar{X}_{i.} - \bar{X}_{..}\right)\right]$$

be the covariance between psu's in the population on both the occasions and

$$\rho_2 S_w^2 = \frac{1}{N(M-1)}\sum_{i=1}^{N}\sum_{j=1}^{M}\left(y_{ij} - \bar{Y}_{i.}\right)\left(x_{ij} - \bar{X}_{i.}\right)$$

be the covariance between ssu's within psu's on both occasions. The values of a, b, d, and e for which $V(\bar{y}_2)$ is minimum are given by

$$a = -\Delta pr\left[\left(\rho_1 q + \rho_2 s\right)S_b^2 + \left(q + ps\right)\frac{S_w^2}{m}\right],$$

$$b = \Delta ps\left[\left(r\rho_2 - \rho_1 q + \rho_1\rho_2^2 qs\right)S_b^2 + \rho_2 pr\,\frac{S_w^2}{m}\right],$$

$$d = \Delta pr\left[\left(1 + \rho_1\rho_2 qs\right)S_b^2 + \frac{S_w^2}{m}\right],$$

and

$$e = \Delta ps\left[\left(1 - \rho_1\rho_2 rq - \rho_2^2 s\right)S_b^2 + \left\{1 - \left(q + ps\right)\rho_2^2\right\}\frac{S_w^2}{m}\right]$$

where

$$\Delta = \frac{\left(S_b^2 + \frac{S_w^2}{m}\right)}{\left[\left\{(1 + \rho_1\rho_2 qs)S_b^2 + \frac{S_w^2}{m}\right\}^2 - \left\{\left(\rho_1 q + \rho_2 s\right)S_b^2 + \left(q + ps\right)\frac{S_w^2}{m}\right\}^2\right]}$$

The minimum variance of the estimator $\bar{y}_2$, after ignoring finite correction factors, is given by

$$(4.2)\ V(\bar{y}_2)=\frac{1}{np}\left(a^2+d^2\right)\left(S_b^2+\frac{S_w^2}{mr}\right)+\frac{1}{np}\left(b^2+e^2\right)\left(S_b^2+\frac{S_w^2}{ms}\right)+$$

$$\frac{1}{nq}\left\{\left(a+b\right)^2+\left(1-d-e\right)^2\right\}\left(S_b^2+\frac{S_w^2}{m}\right)+$$

$$\frac{2}{np}\left(ab+de+ae\rho_1+bd\rho_1+be\rho_1\right)S_b^2+2\frac{ad}{np}\left(\rho_1 S_b^2+\rho_2\frac{S_w^2}{mr}\right).$$

We shall be ignoring finite population correction factors in subsequent calculations of the variances of different estimators also.

It is not possible to get a simple expression of the variance of the estimator although numerical values of the estimator and its variance can be obtained by taking particular values of the parameters involved. It is, therefore, necessary to adopt a technique with a simple expression of the estimator and its variance. The method consists in obtaining the estimator in two stages by taking appropriate weights at each stage.

(*b*) Estimator II : For estimating the mean on the second occasion

Let us first consider *np* primary sampling units which are common to two occasions. An estimator $\bar{y}_c$ (Singh and Kathuria, 1969) of the mean $\bar{Y}_{..}$ based on the common psu's is given by

$$(4.3)\qquad \bar{y}_c=k\,[\bar{y}'+\rho_2(\bar{x}_c-\bar{x}')]+(1-k)\,\bar{y}''$$

where $\bar{x}_c$ is the mean, per ssu on the first occasion, based on *np* psu's common to the two occasions.

The value of *k* for which variance of the estimator $\bar{y}_c$ is minimum, is $r(1-\rho_2^2 s^2)^{-1}$. The minimum variance of $\bar{y}_c$ is given by

$$(4.4)\quad V(\bar{y}_c)=\frac{S_b^2}{np}+\frac{(1-\rho_2^2 s)}{(1-\rho_2^2 s^2)}\cdot\frac{S_w^2}{npm}.$$

Let the second estimator $\bar{y}_w$ based on all the sampling units on the second occasion be given by

$$(4,5)\qquad \bar{y}_w=K\bar{y}_c+(1-K)\,\bar{y}'''+K'\,(\bar{x}_c-\bar{x}''')$$

where K and K' are some constants chosen in a way so as to minimise $V(\bar{y}_w)$. This gives

$$K=\frac{p\left(S_b^2+\frac{S_w^2}{m}\right)^2}{\left[\left(S_b^2+\frac{S_w^2}{m}\right)\left\{S_b^2+\frac{1-(q+ps)\rho_2^2 s}{(1-\rho_2^2 s^2)}\frac{S_w^2}{m}\right\}-q^2\left(\rho_1 S_b^2+\frac{\rho_2 r}{(1-\rho_2^2 s^2)}\frac{S^2_w}{m}\right)^2\right]}$$

and

$$K'=\frac{-q\left(\rho_1 S_b^2+\rho_2\frac{S_w^2}{m}\right)}{\left(S_b^2+\frac{S_w^2}{m}\right)}K$$

The minimum variance of the estimator $\bar{y}_w$ is given by

(4.6) $V(\bar{y}_w)$

$$=\frac{\frac{1}{n}\left(S^2_b+\frac{S_w^2}{m}\right)\left[\left(S_b^2+\frac{S_w^2}{m}\right)\left\{S_b^2+\frac{(1-\rho_2^2 s)}{(1-\rho_2^2 s^2)}\frac{S_w^2}{m}\right\}-q\left(\rho_1 S_b^2+\frac{\rho_2 r}{(1-\rho_2^2 s^2)}\frac{S_w^2}{m}\right)^2\right]}{\left[\left(S_b^2+\frac{S_w^2}{m}\right)\left\{S_b^2+\frac{1-(q+ps)\rho_2^2 s}{(1-\rho_2^2 s^2)}\frac{S_w^2}{m}\right\}-q^2\left(\rho_1 S_b^2+\frac{\rho_2 r\, S_w^2}{(1-\rho_2^2 s^2)m}\right)^2\right]}$$

The efficiencies of the estimators I and II over that of the sample mean per ssu on the second occasion, based on all the nm units for $m=4$ and for different values of ρ_1, ρ_2, p, r and $\phi=S_w^2/S_b^2$ are given in Tables 1 and 2.

We observe the following regarding the efficiencies of the two estimators for different values of ϕ.

The efficiency of the estimator I : (1) When $\phi=.1$ and therefore, $S_b^2 > S_w^2$.

(*i*) efficiency increases appreciably with increase in ρ_1 but not with increase in ρ_2, (*ii*) when ρ_1 is large, it is better to retain 50 per cent or less than 50 per cent of the psu's, (*iii*) different values of r do not make appreciable change in the efficiency.

(2) When $\phi=1$ and therefore, $S_w^2=S_b^2$.

(*i*) efficiency increases with increase in ρ_1 and ρ_2, (*ii*) same as (*ii*) above for $\phi=.1$, (*iii*) efficiency increases with the increase in the values of r.

(3) When $\phi=10$ and therefore, $S_w^2 > S_b^2$.

(*i*) efficiency increases as ρ_1 and ρ_2 increase, but the increase of efficiency with the increase of ρ_2 is more rapid.

(*ii*) the efficiency is higher, for $\rho_1=.8$, $\rho_2=.9$, $p=.5$ and $r=.5$ and .75, than for other values of the corresponding factors.

TABLE 1

The efficiency of the Estimator I over that of the sample mean per ssu on the second occasion, based on all nm units for m = 4 and for different values of ρ_1, ρ_2, p, r and ϕ.

ρ_1		$\phi = .1$									$\phi = 1.0$									$\phi = 10.0$								
		$\rho_2 = .5$			.7			.9			.5			.7			.9			.5			.7			.9		
	$r =$ / p	.25	.50	.75	.25	.50	.75	.25	.50	.75	.25	.50	.75	.25	.50	.75	.25	.50	.75	.25	.50	.75	.25	.50	.75	.25	.50	.75
	.25	1.03	1.03	1.03	1.03	1.04	1.04	1.04	1.04	1.04	1.03	1.03	1.04	1.04	1.04	1.05	1.06	1.06	1.06	1.02	1.03	1.04	1.05	1.07	1.09	1.13	1.18	1.19
.4	.50	1.04	1.04	1.04	1.04	1.05	1.05	1.05	1.05	1.05	1.04	1.05	1.05	1.05	1.06	1.06	1.08	1.09	1.08	1.03	1.05	1.06	1.08	1.11	1.12	1.21	1.24	1.23
	.75	1.03	1.03	1.03	1.03	1.03	1.03	1.04	1.04	1.04	1.03	1.04	1.04	1.05	1.05	1.05	1.08	1.08	1.07	1.04	1.06	1.06	1.09	1.12	1.11	1.24	1.25	1.20
	.25	1.09	1.09	1.09	1.09	1.09	1.09	1.09	1.10	1.10	1.07	1.07	1.08	1.08	1.09	1.10	1.11	1.12	1.12	1.03	1.04	1.06	1.06	1.09	1.11	1.16	1.22	1.24
.6	.50	1.11	1.11	1.11	1.11	1.11	1.11	1.11	1.11	1.11	1.08	1.09	1.11	1.10	1.11	1.12	1.14	1.15	1.15	1.04	1.06	1.08	1.09	1.13	1.15	1.24	1.28	1.28
	.75	1.07	1.07	1.07	1.07	1.08	1.08	1.08	1.08	1.08	1.06	1.07	1.07	1.08	1.09	1.09	1.12	1.12	1.11	1.05	1.07	1.07	1.10	1.14	1.13	1.26	1.27	1.23
	.25	1.22	1.22	1.22	1.22	1.22	1.23	1.23	1.23	1.23	1.13	1.15	1.16	1.16	1.18	1.20	1.22	1.24	1.25	1.04	1.06	1.07	1.07	1.11	1.14	1.20	1.28	1.32
.8	.50	1.22	1.23	1.23	1.23	1.23	1.23	1.24	1.24	1.24	1.15	1.17	1.18	1.18	1.20	1.22	1.24	1.26	1.26	1.05	1.08	1.10	1.11	1.16	1.18	1.27	1.33	1.33
	.75	1.14	1.14	1.14	1.14	1.14	1.14	1.15	1.15	1.15	1.11	1.12	1.12	1.13	1.14	1.14	1.17	1.18	1.17	1.05	1.08	1.08	1.11	1.15	1.15	1.28	1.30	1.26

TABLE 2

The efficiency of the Estimator II over that of the sample mean per ssu on the second occasion, based on all nm units for m = 4 and for different values of ρ_1, ρ_2, p, r and φ.

ρ_1		$\varphi = .1$									$\varphi = 1.0$									$\varphi = 10.0$								
		$\rho_2 = .5$			.7			.9			.5			.7			.9			.5			.7			.9		
	$r = p$	.25	.50	.75	.25	.50	.75	.25	.50	.75	.25	.50	.75	.25	.50	.75	.25	.50	.75	.25	.50	.75	.25	.50	.75	.25	.50	.75
.4	.25	1.03	1.04	1.03	1.04	1.04	1.04	1.04	1.04	1.04	1.04	1.04	1.04	1.06	1.06	1.05	1.08	1.08	1.07	1.06	1.07	1.06	1.15	1.16	1.14	1.42	1.39	1.30
	.50	1.04	1.05	1.04	1.05	1.05	1.05	1.05	1.05	1.05	1.06	1.06	1.05	1.08	1.08	1.07	1.11	1.11	1.09	1.09	1.09	1.08	1.19	1.20	1.17	1.47	1.44	1.33
	.75	1.03	1.03	1.03	1.03	1.04	1.03	1.04	1.04	1.04	1.04	1.05	1.04	1.06	1.07	1.06	1.10	1.10	1.08	1.08	1.09	1.07	1.17	1.18	1.15	1.39	1.36	1.26
.6	.25	1.09	1.09	1.09	1.09	1.10	1.09	1.10	1.10	1.10	1.09	1.09	1.09	1.11	1.12	1.11	1.15	1.15	1.14	1.08	1.09	1.08	1.19	1.19	1.17	1.57	1.53	1.39
	.50	1.11	1.11	1.11	1.11	1.11	1.11	1.12	1.12	1.12	1.11	1.11	1.11	1.14	1.14	1.13	1.19	1.18	1.17	1.11	1.11	1.11	1.23	1.24	1.21	1.57	1.53	1.40
	.75	1.07	1.07	1.07	1.08	1.08	1.08	1.08	1.08	1.08	1.08	1.08	1.08	1.10	1.11	1.10	1.15	1.14	1.12	1.09	1.10	1.09	1.19	1.20	1.17	1.43	1.40	1.29
.8	.25	1.22	1.22	1.22	1.23	1.23	1.23	1.24	1.24	1.24	1.18	1.18	1.18	1.23	1.23	1.23	1.32	1.31	1.29	1.11	1.11	1.10	1.24	1.25	1.22	1.83	1.75	1.54
	.50	1.23	1.23	1.23	1.24	1.24	1.24	1.24	1.24	1.24	1.20	1.20	1.20	1.25	1.25	1.24	1.32	1.32	1.29	1.13	1.14	1.13	1.28	1.29	1.25	1.69	1.64	1.49
	.75	1.14	1.14	1.14	1.15	1.15	1.14	1.15	1.15	1.15	1.13	1.13	1.13	1.16	1.16	1.15	1.22	1.21	1.19	1.11	1.12	1.10	1.22	1.23	1.19	1.48	1.44	1.33

The efficiency of the estimator II for different values of ϕ behaves in a similar manner except with some difference such as the following. For $\phi=1$ and $\rho_2=.9$, the efficiency decreases with the increase in r. For $\phi=10$, $\rho_1=.8$ and $\rho_2=.9$, the behavior of the efficiency is erratic.

We note that the technique of matching does provide better estimators. We further note that the Estimator II is equal or more efficient than the Estimator I for different values of ρ_1, ρ_2, p, r and ϕ and is simple in form. Therefore, the estimator II should be preferred over the Estimator I.

(*c*) An estimator of the change :

An unbiased estimator of change $(=\bar{Y}_{..}-\bar{X}_{..})$ may be taken as

(4.7) $D_w=a(\bar{y}'-\bar{x}')+b(\bar{y}''-\bar{x}'')+(1-a-b)(\bar{y}'''-\bar{x}''')$.

By minimising the variance of D_w, we get

$$a=\frac{rp\left(S_b^2+\frac{S_w^2}{m}\right)}{\left[\left(1-\rho_1 q\right)\left(1-\rho_2 s\right)S_b^2+\left\{1-\left(q+ps\right)\rho_2\right\}\frac{S_w^2}{m}\right]}$$

and $$b=\frac{sp\left(1-\rho_2\right)\left(S_b^2+\frac{S_w^2}{m}\right)}{\left[\left(1-\rho_1 q\right)\left(1-\rho_2 s\right)S_b^2+\left\{1-\left(q+ps\right)\rho_2\right\}\frac{S_w^2}{m}\right]}$$

The minimum variance of the estimator of the change D_w is given by

(4.8) $V(D_w)$

$$=\frac{\frac{2}{n}\left(S_b^2+\frac{S_w^2}{m}\right)\left[\left(1-\rho_1\right)\left(1-\rho_2 s\right)S_b^2+\left(1-\rho_2\right)\frac{S_w^2}{m}\right]}{\left[\left(1-\rho_1 q\right)\left(1-\rho_2 s\right)S_b^2+\left\{1-\left(q+ps\right)\rho_2\right\}\frac{S_w^2}{m}\right]}$$

It may be seen from Table 3 on page 33 that the estimator D_w is more efficient than the estimator obtained by taking the difference of the two sample means per ssu, on the two occasions, based on all nm units. The increase in efficiency is more rapid for higher values of ρ_1, ρ_2, p and r for different values of ϕ.

(*d*) An over-all estimator of the mean over the two occasions.

An over-all estimator of the over-all mean, $\frac{1}{2}(\bar{X}_{..}+\bar{Y}_{..})$, for

TABLE 3

The efficiency of the estimator Dw of the change as compared to the estimator obtained by taking the difference of the two sample means per ssu, on the two occasions, based on all nm units, for $m = 4$ and for different values of ρ_1, ρ_2, p r and ϕ.

ρ_1		$\phi = .1$									$\phi = 1.0$									$\phi = 10.0$								
		$\rho_2 = .5$			.7			.9			.5			.7			.9			.5			.7			.9		
	$r = p$	.25	.50	.75	.25	.50	.75	.25	.50	.75	.25	.50	.75	.25	.50	.75	.25	.50	.75	.25	.50	.75	.25	.50	.75	.25	.50	.75
.4	.25	1.16	1.17	1.17	1.17	1.17	1.17	1.17	1.17	1.17	1.14	1.16	1.17	1.16	1.19	1.20	1.21	1.23	1.24	1.09	1.14	1.18	1.15	1.25	1.33	1.39	1.58	1.70
	.50	1.33	1.33	1.33	1.33	1.34	1.34	1.34	1.35	1.35	1.28	1.32	1.34	1.32	1.37	1.40	1.42	1.47	1.49	1.17	1.27	1.36	1.30	1.50	1.66	1.78	2.16	2.40
	.75	1.49	1.50	1.50	1.50	1.51	1.51	1.51	1.52	1.52	1.42	1.47	1.51	1.49	1.56	1.61	1.63	1.70	1.73	1.26	1.41	1.54	1.45	1.75	1.99	2.17	2.74	3.10
.6	.25	1.36	1.36	1.37	1.37	1.37	1.38	1.38	1.38	1.39	1.27	1.30	1.33	1.31	1.36	1.39	1.41	1.45	1.47	1.11	1.17	1.23	1.19	1.31	1.42	1.50	1.77	1.96
	.50	1.72	1.73	1.74	1.73	1.75	1.75	1.76	1.77	1.77	1.54	1.60	1.65	1.62	1.71	1.77	1.81	1.90	1.95	1.23	1.35	1.46	1.38	1.63	1.84	2.00	2.55	2.92
	.75	2.08	2.10	2.11	2.10	2.12	2.13	2.14	2.15	2.16	1.81	1.90	1.98	1.93	2.07	2.16	2.22	2.35	2.42	1.34	1.52	1.69	1.58	1.94	2.26	2.50	3.32	3.88
.8	.25	1.91	1.93	1.95	1.94	1.96	1.98	1.98	2.00	2.01	1.53	1.60	1.66	1.62	1.74	1.82	1.88	2.02	2.10	1.15	1.22	1.29	1.24	1.40	1.54	1.65	2.09	2.42
	.50	2.83	2.87	2.89	2.88	2.92	2.95	2.97	3.01	3.02	2.06	2.20	2.32	2.25	2.48	2.65	2.76	3.05	3.19	1.30	1.44	1.57	1.48	1.79	2.08	2.31	3.17	3.85
	.75	3.74	3.80	3.84	3.81	3.88	3.93	3.95	4.01	4.03	2.59	2.81	2.98	2.87	3.22	3.47	3.64	4.07	4.29	1.44	1.66	1.86	1.73	2.19	2.62	2.96	4.26	5.27

the two years may similarly be written as

$$(4\cdot9)\quad T_w = \tfrac{1}{2}a(\bar{y}'+\bar{x}') + \tfrac{1}{2}b(\bar{y}''+\bar{x}'') + \tfrac{1}{2}(1-a-b)(\bar{y}'''+\bar{x}''')$$

By minimising the variance of T_w, we get

$$a = \frac{pr\left(S_b^2 + \frac{S_w^2}{m}\right)}{\left[\left(1+\rho_1 q\right)\left(1+\rho_2 s\right)S_b^2 + \left\{1+\left(q+ps\right)\rho_2\right\}\frac{S_w^2}{m}\right]}$$

and

$$b = \frac{ps\left(1+\rho_2\right)\left(S_b^2 + \frac{S_w^2}{m}\right)}{\left[\left(1+\rho_1 q\right)\left(1+\rho_2 s\right)S_b^2 + \left\{1+\left(q+ps\right)\rho_2\right\}\frac{S_w^2}{m}\right]}$$

The minimum variance of the estimator T_w is given by

$$(4\cdot10)\quad V(T_w) = \frac{\frac{1}{2n}\left(S_b^2 + \frac{S_w^2}{m}\right)\left[\left(1+\rho_1\right)\left(1+\rho_2 s\right)S_b^2 + \left(1+\rho_2\right)\frac{S_w^2}{m}\right]}{\left[\left(1+\rho_1 q\right)\left(1+\rho_2 s\right)S_b^2 + \left\{1+\left(q+ps\right)\rho_2\right\}\frac{S_w^2}{m}\right]}$$

It may be noted from Table 4 on page 35 that the efficiency of the estimator T_w, as compared to the estimator obtained by taking the simple average of the two sample means per ssu, on two occasions, based on all nm units decreases as ρ_1, ρ_2, p and r increase for different values of ϕ. Thus, it appears that for estimating the overall mean, it is better to take fresh units at both the stages on the second occasion rather than resort to matching when ρ_1 and ρ_2 are positive. With negative values of ρ_1 and ρ_2, T_w can be seen to be more efficient than the estimator based on no matching for different values of ϕ.

5. The suitability of successive sampling technique in surveys for measuring the incidence of pests and diseases on rice crop.

The various efficiency comparisons suggest the utility of successive sampling technique in general. As mentioned earlier, the purpose of the present study was to investigate the use of this technique in incidence of pests and diseases survey for (*i*) improving the estimate of mean incidence of pests and diseases on the second occasion, (*ii*) estimating the change in incidence of pests and diseases from one year to another year and (*iii*) finding the effect on efficiency of overall estimator of mean incidence over two years.

TABLE 4

The efficiency of the over-all estimator over that of the estimator obtained by taking the simple average of the two sample means per ssu, on the two occasions, based on all nm units, for $m = 4$ and for different values of ρ_1, ρ_2, p, r and φ.

ρ_1		$\varphi = .1$									$\varphi = 1.0$									$\varphi = 10.0$								
		$\rho_2 = .5$			.7			.9			.5			.7			.9			.5			.7			.9		
	$r = p$	.25	.50	.75	.25	.50	.75	.25	.50	.75	.25	.50	.75	.25	.50	.75	.25	.50	.75	.25	.50	.76	.25	.50	.75	.25	.50	.75
	.25	.93	.93	.93	.93	.93	.93	.93	.93	.93	.94	.93	.93	.94	.93	.93	.94	.93	.92	.96	.95	.93	.96	.94	.92	.96	.94	.92
.4	.50	.86	.86	.86	.86	.86	.86	.86	.86	.86	.87	.87	.86	.87	.86	.85	.87	.86	.85	.92	.90	.87	.92	.88	.85	.91	.87	.83
	.75	.79	.79	.79	.79	.79	.79	.79	.79	.78	.81	.80	.79	.81	.80	.78	.81	.79	.77	.89	.85	.80	.88	.83	.77	.87	.81	.75
	.25	.91	.91	.91	.91	.91	.91	.91	.91	.91	.92	.91	.91	.92	.91	.91	.92	.91	.91	.95	.94	.93	.95	.93	.92	.95	.93	.91
.6	:50	.81	.81	.81	.81	.81	.81	.81	.81	.81	.83	.83	.82	.83	.82	.82	.83	.82	.81	.90	.88	.85	.90	.87	.84	.89	.86	.82
	.75	.72	.72	.72	.72	.72	.72	.72	.72	.72	.75	.74	.73	.75	.74	.73	.75	.74	.72	.86	.82	.78	.85	.80	.75	.84	.79	.73
	.25	.89	.89	.89	.89	.89	.89	.89	.89	.89	.90	.90	.90	.90	.90	.89	.90	.90	.89	.94	.93	.92	.94	.93	.91	.94	.92	.90
.8	.50	.78	.78	.78	.78	.78	.78	.78	.78	.78	.80	.80	.79	.80	.80	.79	.80	.79	.79	.89	.86	.84	.88	.85	.82	.88	.84	.81
	.75	.67	.67	.67	.67	.67	.67	.67	.67	.67	.70	.70	.69	.70	.69	.68	.70	.69	.68	.83	.80	.76	.82	.78	.73	.82	.77	.71

The present study is confined to the analysis of incidence data collected during the months of March (*Rabi* crop) and October (*Kharif* crop), for Stemborer, Gallfly and Helminthosporium during the years 1959-60, 1960-61, and 1961-62. The data for the months of March and October were examined as the respective incidences were at their peaks during these months in the *Rabi* and *Kharif* seasons respectively. The data on the incidence of White earheads was also examined. The results are presented in Table 5.

The various means in column 1 are based on the observations in a given month in each of the two years in a given column from amongst the columns 2-10. For example, in column 2, the data in month of March, 1960 is taken as the data on the first occasion and the data in the month of March, 1961 is taken as the data on the second occasion. The quantities ρ_1 and ρ_2 and S_b^2 and S_w^2 for the second occasion are calculated from the sample data. Assuming that the matched units and the sample size n on each occasion are large enough, these estimated quantities may be taken as population values for calculating the various estimators and their efficiencies from the various formulae given in Section 4. This step is justified in view of the results due to Tikkiwal (1956) regarding the estimation procedure in univariate sampling on successive occasions. The various estimators, based on matching, are calculated using the information on previous year only. The efficiency of various estimators, based on matching, are obtained by dividing the variances of these estimators with those of corresponding estimators based on no matching.

We draw the following salient conclusions from the analysis given in the above table. The estimates of mean incidences, due to Stemborer in March and October, 1961 and March, 1962 and due to Gallfly in October, 1960 and October, 1961 were found to be more efficient than those based on no matching. The gain in efficiency is more when either ρ_1 or ρ_2 is large. There is no gain in efficiency in case of Helminthosporium and White earheads where ρ_1 and ρ_2 are very low.

The estimates of change in incidence of Stemborer and Gallfly in the months of March and October during *Rabi* and *Kharif* seasons were found more efficient than those based on no matching. The gain in efficiency of estimates of change in incidence during *Rabi* season was however much less as compared to the gain in efficiency of estimator of change in incidence during *Kharif* season.

TABLE 5

Estimate of mean incidence on second occasion, estimate of change in incidence from year to year and over-all estimate of incidence over two years in Cuttack district of Orissa State.

Pest/Disease	*Stemborer*			*Gallfly*		*Helminthosporium*		*White earheads*	
Time of recording observation with season	*March (Rabi)*	*October (Kharif)*	*March (Rabi)*	*October (Kharif)*	*October (Kharif)*	*March (Rabi)*	*March (Rabi)*	*At harvest (Rabi)*	*At harvest (Rabi)*
Agricultural Year	(1) to (2)*	(2) to (3)	(2) to (3)	(1) to (2)	(2) to (3)	(1) to (2)	(2) to (3)	(1) to (2)	(2) to (3)
n	40	54	40	48	54	37	40	46	47
p	0.35	0.50	0.38	0.46	0.50	0.35	0.38	0.37	0.36
q	0.65	0.50	0.62	0.54	0.50	0.65	0.62	0.63	0.64
r	0.45	0.49	0.47	0.50	0.49	0.44	0.47	0.48	0.47
s	0.55	0.51	0.53	0.50	0.51	0.56	0.53	0.52	0.53
m	4.00	3.70	4.00	3.50	3.70	4.00	4.00	3.80	4.00
Est : ρ_1	0.52	0.86	0.46	0.48	0.24	−0.11	0.17	0.15	−0.03
,, ρ_2	0.86	−0.09	−0.17	0.13	0.97	0 44	0.08	0.02	0.11
,, $\bar{x}'$	3.74%	4.75%	2.61%	1.89%	2.39%	1.64 Score	1.64 Score	2.07%	1.56%
,, $\bar{x}''$	4.59%	3.69%	0.37%	2.55%	1.98%	1.85 ,,	1.87 ,,	2.29%	1.42%
,, $\bar{x}'''$	4.80%	1.80%	12.03%	1.94%	1.02%	2.15 ,,	1.59 ,,	2.02%	1.42%
,, $\bar{y}'$	5.29%	2.58%	1.73%	1.99%	1.47%	2.16 ,,	1.18 ,,	1.58%	1.40%
,, $\bar{y}''$	8.62%	1.79%	0.83%	2.90%	1.54%	2.03 ,,	1.30 ,,	1.50%	1.71%
,, $\bar{y}'''$	8.73%	1.25%	3.15%	0.69%	1.91%	1.58 ,,	1.50 ,,	1.42%	1.39%
,, S_b^2	139.73	34.27	131.11	10.63	6.52	1.68	1.48	1.54	2.18
,, S_w^2	26.96	3.01	17.82	13.56	12.79	0.28	0.21	0.98	0.95
,, $\bar{y}_w$	7.99 (3.37)**	1.20 (0.51)	3.56 (3.22)	1.51 (0.29)	1.51 (0.16)	1.75 (0.05)	1.40 (0.04)	1.46 (0.04)	1.45 (0.05)
,, D_w	2.85 (5.03)	−4.29 (0.35)	−4.29 (5.05)	−0.41 (0.36)	−0.29 (0.17)	−0.26 (0.09)	−0.27 (0.07)	−0.62 (0.06)	0.01 (0.09)
,, T_w	6.56 (2.08)	2.12 (0.42)	5.74 (1.92)	1.72 (0.17)	1.62 (0.10)	1.89 (0.025)	1.53 (0.02)	1.77 (0.02)	1.45 (0.025)
Variance of the sample mean incidence $= (S_b^2+S_w^2/m)/n$	3.66	0.65	3.39	0.30	0.18	0.05	0.04	0.04	0.05
Efficiency of $\bar{y}_w$	1.09	1.27	1.05	1.03	1.13	1 00	1.00	1.00	1.00
Efficiency of D_w	1.46	3.71	1.34	1.67	2.12	1.11	**1.14**	1.33	1.11
Efficiency of T_w	.88	.77	.88	.87	.88	1.00	1.00	1.00	1.00

N.B. : *(1)=1959-60, (2)=1960-61, and (3)=1961-62. **The figures in the brackets in the body of the table are the variances of the estimates.

The over-all estimator for two years was found to be equally efficient as the one based on no matching in case of white earheads and Helminthosporium because of low values of ρ_1 and ρ_2. In the case of Stemborer and Gallfly incidence the over-all estimator was, however, 12 to 23 per cent less efficient as compared to no matching.

ACKNOWLEDGEMENTS

The authors are thankful to Dr. G.R. Seth, Statistical Adviser, for his valuable suggestions and to the referee for his useful comments leading to the improvement of the paper.

REFERENCES

1. Jessen, R J. (1942) Statistical Investigations of a sample survey for obtaining Farm Facts. Iowa Agri. Exp. Station Res. Bull. 304.

2. Patterson, H.D. (1950) Sampling on successive occasions with partial replacement of units. J.R.S.S., Series B, 12, 241-255.

3. Singh, D. and Kathuria, O.P. (1969) On Two-stage successive sampling. Australian Journal of Statistics, Vol. 11, No. 2.

4. Tikkiwal B.D. (1950) Estimation by Successive Sampling. Paper presented at 4th Annual Meeting of Ind. Soc Agri. Stat., Details in Diploma Thesis, 1951. The Institute of Agricultural Research Statistics, New Delhi.

5. ——————————(1953) Optimum allocation in Successive Sampling. J. Ind. Soc. Agri. Stat., 5, 101-102.

6. ——————————(1956) A further contribution to the theory of Univariate Sampling on Successive Occasions. J. Ind. Soc. Agri. Stat., 8, 84-90.

7. ——————————(1964) A note on Two-Stage Sampling on Successive Occasions. Sankhya, Series A, 26, Part I, 97-100.

8. ——————————(1965) The Theory of Two-Stage Sampling on Successive Occasions. Jour. Ind. Stat. Assoc. Vol. 3, No. 2 and 3, April-July, 125-135.

9. ——————————(1967) Theory of Multiphase Sampling from a Finite or an Infinite Population on Successive Occasions. Review Int. Stat. Instt., Vol. 35.