AN EXAMINATION OF THE EFFECT OF MATCHED SAMPLING ON THE EFFICIENCY OF ESTIMATORS IN THE THEORY OF SUCCESSIVE SAMPLING

By B. D. Tikkiwal

Karnatak University

1. Introduction

The general theory of Successive Sampling has been studied for one character independently by Patterson[1] and the author[2] both arriving at similar results. Later the author has also given results[4,7] for the study of several characters on each occasion. The sampling scheme used in arriving at these results is one that of matched sampling, *i.e.*, some of the sample units on the current occasion are necessarily the units studied on the previous occasions giving rise to matching. It is shown in[4,7] that, on any one occasion after the first, each of the estimators for various characters, based on matched sampling and thus utilising the information collected on earlier occasions through matched units, has less variance and therefore greater efficiency than the estimator for the corresponding character when there is no matching on the occasion. The purpose of this paper is to examine the effect of matched sampling on the efficiency of the various estimators from occasion to occasion.

2. The Reduction of the Multivariate Case to the Univariate Case

The present examination can only be done when the variability of the character is the same on each occasion and the sample sizes for the character on different occasions are composed of the same proportions of new and old units. Further one has to use the same sub-sampling fractions for any one of the characters other than the first on each occasion when there are several characters. Under these conditions in notations[4,6,7] we have $n_h'' = n\lambda$, $n_h' = n(1-\lambda)$ and $\lambda'_{ht'} = \lambda''_{ht'} = \lambda t'$ for $t' = 2, \ldots, k$ and the variance of the best estimator for the k-th character $(k > 1)$ on the h-th occasion is given by the expression

$$\sum_{i=1}^{k} \left[\frac{1 - \rho'^{2}_{i-1, i})}{n\pi^{i}_{t'=1}\lambda_{t'}} \rho'^{2}_{1k} \right] \cdot \frac{\phi_h}{\lambda} \sigma^{2}_{xk} \qquad (2.1)$$

where σ^{2}_{xk} is the variance for the character k on each of the h occasions. For $k = 1$, the variance expression is

$$\frac{\phi_h}{n\lambda} \sigma^{2}_{X_1}. \qquad (2.2)$$

The weighting coefficient ϕ_h for $\rho_{h-1, h} = \rho$ for all $h \geqslant 2$ is itself given by

$$\phi_h = \frac{\frac{\lambda}{1-\lambda}(1-\rho^2) + \rho^2 \phi_{h-1}}{1 + \frac{\lambda}{1-\lambda}(1-\rho^2) + \rho^2 \phi_{h-1}}. \qquad (2.3)$$

In order to examine the behaviour of the efficiency of the estimator for any one character from one occasion to another, it is clear from (2.1) and (2.2) that we have to examine the behaviour of the weighting coefficient ϕ on these occasions irrespective of the value of k. Thus the problem for the multivariate case reduces to that for the univariate case.

3. The non-increasing behaviour of the weighting coefficient and hence the increase in efficiency due to matched sampling with increasing number of occasions.

From Equation (2.3), we have

$$\phi_h = 1 - \frac{1}{1 + a + \rho^2 \phi_{h-1}},$$

where

$$a = \frac{(1-\rho^2)\lambda}{\mu}$$

and

$$\lambda + \mu = 1;$$

or

$$\phi_h = 1 - \frac{1}{1 + a + \rho^2 - \frac{\rho^2}{1 + a + \rho^2 \phi_{h-2}}}.$$

Proceeding in this way, we have

$$\phi_h = 1 - \frac{1}{1 + \rho^2 + a -} \frac{\rho^2}{\rho^2 + a + 1 -} \frac{\rho^2}{\rho^2 + a + 1 -}$$

the last term in the continued fraction for ϕ_h being

$$\frac{\rho^2}{(1 + a + \lambda\rho^2)}$$

as $\phi_1 = \lambda$.

We immediately note that

$$\phi_2 \leqslant \phi_1$$

if,

$$\left(1 - \frac{1}{1 + a + \lambda\rho^2}\right) \leqslant \lambda. \tag{3.2}$$

That (3.2) is also the condition for $\phi_h \leqslant \phi_{h-1}$ for $h > 2$ is easily seen by noting, that, if the last term in the continued fraction for ϕ_h is combined with ρ^2 in the denominator of the second last term, then the resulting portion is the left-hand side expression in (3.2) when multiplied by ρ^2 and the corresponding portion, the only differing portion in the continued fraction for ϕ_{h-1} is $\lambda\rho^2$ and that if,

$$\phi_h = 1 - \frac{1}{1 + a + \rho^2 - f_h}$$

and,

$$\phi_{h-1} = 1 - \frac{1}{1 + a + \rho^2 - f_{h-1}} \tag{3.3}$$

then,

$$f_{h-1} \leqslant f_h < \rho^2. \tag{3.4}$$

The condition (3.2) is easily seen to be true and so for $h \geqslant 2$

$$\phi_h \leqslant \phi_{h-1}. \tag{3.5}$$

It may be noted from (3.2) that the equality sign holds in (3.5) for the trivial case when $\lambda = 1$ and so there is no matching. Thus we see that ϕ_h is non-increasing function of h and that whenever there is matching of units, the efficiency of the estimator for the same character increases with the occasions.

4. THE CHECKS FOR THE COMPUTATION OF ϕ_h AND THE BEHAVIOUR OF ITS LIMITING VALUE

A lower limit of the value of ϕ_h for all h is obtained by putting $f_h = \rho^2$ in (3.3).

Thus,

$$\phi_h > \frac{\lambda(1-\rho^2)}{1-\lambda\rho^2} \tag{4.1}$$

for all h. This provides a computational check for all ϕ's. Further, we have seen in Section 3 that whenever there is matched sampling, ϕ_h is in fact a decreasing function of h and since ϕ_h is always positive and so it tends to a limit which is different from that given in (4.1). Thus it provides another computational check for ϕ's. The limiting value of ϕ_h itself can be obtained easily by putting $\phi_h = \phi_{h-1}$ in (2.3). But a more instructive method is the following one.

Let the $n\,(\leqslant h-3)$-th convergent to the continued fraction

$$\frac{-\rho^2}{\rho^2+a+1-}\;\frac{\rho^2}{\rho^2+a+1-}\cdots, \tag{4.2}$$

which is a part of the continued fraction for ϕ_h in (3.1), be denoted by p_n/q_n, then

$$\frac{p_n}{q_n} = \frac{-1}{\left[(a+\beta) - a\,\dfrac{\left\{1-\left(\frac{a}{\beta}\right)^{n-1}\right\}}{\left\{1-\left(\frac{a}{\beta}\right)^{n}\right\}}\right]} \tag{4.3}$$

where,

$$\rho^2(\beta-x)(a-x) = 1-(\rho^2+a+1)x+\rho^2x^2 \tag{4.4}$$

Therefore,

$$\lim_{n\to\infty}\frac{p_n}{q_n} = \frac{-1}{\beta}$$

and hence,

$$\phi_\infty = \lim_{h\to\infty}\phi_h = 1 - \frac{1}{1+a+\rho^2-\dfrac{1}{\beta}}$$

$$= 1 - \frac{1}{\rho^2 (\alpha + \beta) - \frac{1}{\beta}}$$

$$= 1 - \alpha.$$

$$= \frac{-(1 - \rho^2) + \sqrt{(1 - \rho^2)\{1 - \rho^2 (1 - 4\lambda\mu)\}}}{2\mu\rho^2}$$

the same as given by Yates.[8] Further, it can be noted that the absolute difference between the n-th and $(n - 1)$-th convergents is less than ρ^{2n} and so ϕ_{h-1} will tend to ϕ_h if ρ^h tends to zero and in fact more rapidly than indicated by the latter. When $|\rho| = 1$, ϕ_{h-1} will still tend to ϕ_h but slowest of all other values of ρ.

We know ϕ_h is a function of absolute value of ρ and so its actual behaviour is shown in the table given in the appendix reproduced from[2] for positive values of ρ and $\lambda = \cdot 1, \cdots, \cdot 5$. It is noted that ϕ_h tends to the limiting value ϕ more slowly for increasing value of ρ as indicated by the mathematical discussion. This is clear even otherwise, as the occasion, on which the maximum efficiency due to matched sampling is reached should occur later in the case when the correlation between the successive occasions is higher.

SUMMARY

We have shown in this paper that the effect, of using matched sampling to estimate the time-conditioned characters of a population from occasion to occasion, is to improve the efficiency of the various estimators, based on matched sampling, of the characters with increasing occasions reaching a limiting value. This is in addition that it gives more efficient estimators of the various parameters on an occasion as compared to the corresponding simple estimators based on no-matching on that occasion. The two checks are given for the computation of ϕ_h, the weighting coefficient used in the estimators and the behaviour of its limiting value is discussed. It is shown that the rapidity of convergence of the weighting coefficient decreases with increasing absolute value of the correlation-coefficient between two consecutive occasions. This is expected as the occasion, on which the limiting efficiency due to matched sampling is reached, should occur later in the case when the correlation between successive occasions is higher.

REFERENCES

1. Patterson, H. D. .. "Sampling on successive occasions with partial replacement of units," *J.R.S.S.*, 1950, **12**, 241–55.

2. Tikkiwal, B. D. .. "Theory of successive sampling," A thesis for diploma, I.C.A.R., New Delhi, 1951, 1–17.

3. ———— .. "Optimum allocation in successive sampling," *Jour. Ind. Soc. Agric. Stat.*, 1953, **5**, 100–02.

4. ———— .. "Multiphase sampling on successive occasions," *A Thesis submitted to the Faculty of North Carolina State College, Raleigh, N.C., U.S.A., in partial fulfilment of the requirements for the Degree of Doctor of Philosophy in the Department of Experimental Statistics*, 1955, 1–90.

5. ———— .. "A further contribution to the theory of univariate sampling on successive occasions," *Jour. Ind. Soc. Agric. Stat.*, 1956, **8**, 84–90.

6. ———— .. "An application of the theory of multiphase sampling on successive occasions to surveys of livestock marketing," *Jour. of the Karnatak University*, 1956, **1** (1), 120–30.

7. ———— .. "Theory of multiphase sampling from a finite or an infinite population on successive occasions," A paper under publication in *Ann. Math. Stat.*

8. Yates, F. .. *Sampling Methods for Census and Surveys*, Charles Griffin & Company Ltd., London, 1949.

APPENDIX

TABLE*

	$\rho=$	·1	·2	·3	·4	·5	·6	·7	·8	·9	1·0
$\lambda=\cdot 1$ ..	ϕ_2	·0999	·0996	·0992	·0986	·0977	·0968	·0956	·0942	·0921	·0909
	ϕ_3	·0999	·0996	·0991	·0984	·0973	·0958	·0938	·0922	·0873	·0833
	ϕ_4	·0999	·0996	·0991	·0983	·0972	·0955	·0931	·0901	·0841	·0769
	ϕ_5	·0999	·0996	·0991	·0983	·0972	·0954	·0928	·0890	·0819	·0714
	ϕ_6	·0999	·0996	·0991	·0983	·0972	·0954	·0927	·0384	·0804	·0667
	ϕ_∞	·0999	·0996	·0991	·0983	·0972	·0954	·0925	·0877	·0772	·0000
$\lambda=\cdot 2$ ..	ϕ_2	·1997	·1987	·1971	·1949	·1919	·1883	·1840	·1790	·1732	·1667
	ϕ_3	·1997	·1987	·1969	·1943	·1906	·1855	·1788	·1698	·1581	·1429
	ϕ_4	·1997	·1987	·1969	·1943	·1904	·1849	·1770	·1658	·1493	·1250
	ϕ_5	·1997	·1987	·1969	·1943	·1903	·1847	·1764	·1639	·1442	·1111
	ϕ_6	·1997	·1987	·1969	·1943	·1903	·1847	·1762	·1631	·1411	·1000
	ϕ_∞	·1997	·1987	·1969	·1943	·1903	·1847	·1762	·1625	·1365	·0000
$\lambda=\cdot 3$ ..	ϕ_2	·2994	·2975	·2943	·2898	·2839	·2766	·2677	·2572	·2450	·2308
	ϕ_3	·2994	·2974	·2940	·2889	·2818	·2721	·2591	·2418	·2187	·1875
	ϕ_4	·2994	·2974	·2940	·2889	·2816	·2713	·2568	·2361	·2054	·1579
	ϕ_5	·2994	·2974	·2940	·2889	·2815	·2711	·2562	·2339	·1986	·1364
	ϕ_6	·2994	·2974	·2940	·2889	·2815	·2711	·2560	·2331	·1950	·1200
	ϕ_∞	·2994	·2974	·2940	·2889	·2815	·2711	·2560	·2326	·1911	·0000
$\lambda=\cdot 4$ ..	ϕ_2	·3990	·3961	·3912	·3842	·3750	·3633	·3490	·3316	·3107	·2791
	ϕ_3	·3990	·3961	·3909	·3833	·3726	·3579	·3382	·3114	·2745	·2182
	ϕ_4	·3990	·3961	·3909	·3832	·3723	·3571	·3359	·3052	·2587	·1791
	ϕ_5	·3990	·3961	·3909	·3832	·3723	·3570	·3354	·3031	·2516	·1519
	ϕ_6	·3990	·3961	·3903	·3832	·3723	·3570	·3353	·3027	·2484	·1319
	ϕ_∞	·3990	·3961	·3909	·3832	·3723	·3570	·3352	·3024	·2457	·0000
$\lambda=\cdot 5$ ..	ϕ_2	·4988	·4950	·4885	·4792	·4667	·4505	·4302	·4048	·3730	·3333
	ϕ_3	·4987	·4949	·4882	·4783	·4643	·4451	·4189	·3824	·3298	·2499
	ϕ_4	·4987	·4949	·4882	·4782	·4641	·4445	·4170	·3768	·3137	·2000
	ϕ_5	·4987	·4949	·4882	·4782	·4641	·4444	·4167	·3755	·3075	·1667
	ϕ_6	·4987	·4949	·4882	·4782	·4641	·4414	·4166	·3751	·3051	·1429
	ϕ_∞	·4987	·4949	·4882	·4782	·4641	·4444	·4166	·3750	·3036	·0000

* The weighting coefficieut ϕ_1 is not tabulated above because for any given value of λ, $\phi_1 = \lambda$ for different values of ρ.