A GENERALIZATION OF SEQUENTIAL SAMPLING STRUCTURE*

F. S. CHAUDHURY
H.A.U., Hissar
and
D. SINGH
B4/126, Paschim Vihar, New Delhi.

(Received : July 1980)

SUMMARY

It is proposed to develop a generalised sequential sampling procedure and evolve a technique of generating estimators for any sequential sampling system. The procedure suggested puts at our disposal a number of estimators. The problem of choice among different estimators needs further discussion and in most cases extensive empirical studies would be necessary to arrive at the best or the near best estimator. A particular general estimator given by the generating technique happens to include most of the estimators commonly used in practice so that these general estimators may be taken as an acceptable estimator and used on all such occasions where the efficiency of any particular estimator is doubtful. Necessary illustrations have been done to test the veracity of the method.

Introduction

Roy and Chakravarty [8] and Godambe [2] gave admissible estimators, Hanurav [3] applied admissibility concept to sampling theory. Murthy and Singh [5], Joshi [4], Prabhu-Ajgaonkar [7] and other contributed on best and admissible estimators with fixed-sample size only. Singh [9] has applied this concept with some modifications to sequential sampling, and Chaudhary and Singh [1], and Singh and Singh [10] have discussed some more generalized sequential estimators. A new line of acceptable sequential

*The work was supported partially by a grant from the University Grants Commission Grant No. SRI-25-3/7587/77.

estimators has been presented by Singh and Singh [11]. In the present study an attempt has been made to extend these ideas further to arrive at an exact generalization of sequential sampling structure and generate acceptable generalised estimators.

DEFINITION 1.1. Let a finite population $\mathcal{U}$ consist N distinguishable units U_i associated with a real variate y_i, $i = 1, 2, \ldots, N$. A parameter $\theta(= \underset{\sim}{\theta} = (y_1, y_2, \ldots, y_N)$ is a point in the Euclidean space or class of point sets (for brevity class A). Usually the problem is to estimate θ on the basis of individuals i sampled from the population $\mathcal{U}$ and the values y_i associated them, i.e., on the basis of $s(s, y_i \in s)$ where s is a subset of $\mathcal{U}$ drawn with a given sampling design p.

Sampling Design: is any function p on A, the set of all possible subsets of s of $\mathcal{U}$ such that $p(s) > 0$, $\Sigma\, p(s) = 1$, $s \in A$.

Probability Field: consider a non-negative function P defined for every combination $(y_{j1}, y_{j2}, \ldots, y_{jn})$ of s_j. A probability measure may be constructed in which the combination $(y_{j1}, y_{j2}, \ldots, y_{jn})$ will be sampled with probability proportionate to $P(y_{j1}, y_{j2}, \ldots, y_{jn})$ over the combination such that $\Sigma\, P = 1$ and it will be referred as a probability field (Ω).

DEFINITION 1.2. *Parametric Function*: Let us consider the parametric functions defined by Singh (1977) say, $\theta(= \theta\,(\underset{\sim}{y}))$ that can be expressed over the class A, i.e., a more general method to express a parametric function may be

$$\theta = \sum_{a_i \in A} \lambda_i\, \pi_{a_i}\, f(a_i) \tag{1.1}$$

where $f(a_i)$ is a single-valued set function defined over the class A, $\sum\limits_{a_i \in A}$ is the summation over all sets 'a_i' belonging to class A, π_{a_i} is a probability measure defined over a_i in the class A, and λ_i is some adjustment constant.

DEFINITION 1.3. *Sequential Estimator*: A 'statistic' t defined over the probability field Ω is a function over the sample s. A statistic used to estimate a parametric function θ is called an estimator of θ and a most general form of a linear estimator may be

$$t = \sum_{a_i \in s} f(a_i)\, \phi(a_i, s) \Big/ \sum_{a_i \in s} \phi(a_i, s) \tag{1.2}$$

where $\phi\,(a_i' s)$ is a probability measure defined over a point with a_i in s. In case $\sum\limits_{a_i \in s} \phi(a_i, s) = 1$, then the estimator t is called an unbiased estimator.

The estimator (1.2) may also be as

$$t = \sum_{a_i \epsilon s} f(a_i)\, p(a_i, s/E_{a_i}) / \sum_{a_i \epsilon s} p(a_i, s/E_{a_i}) \tag{1.3}$$

where E_{a_i} is an event depending on the occurrence of the set 'a_i' in the sample s and $p\ (a_i,\ s/E_{a_i})$ is a probability measure for a_i when E_{a_i} has occurred.

DEFINITION 1.4. *Sampling System*: It may be considered as the specification of all possible samples alongwith their probability fields over the combination of units in the sample with reference to θ, i.e., it is a combination of estimators of ordered sequence of samples s from $\mathcal{U}$ with probability field (Ω) symbolically, $F = F(t, \Omega)$.

DEFINITION 1.5. *Sequential Decision Rule*: Before deciding how the sampling process will terminate, some terms used within the text may be defined:

A sequential decision rule is a pair (ϕ, δ) in which ϕ is a stopping rule and δ is a terminal rule.

Stopping Rule: A stopping rule is a sequence of function

$$\phi(y) = (\phi_0, \phi_1(y_1), \phi_2(y_1, y_2), \ldots) \tag{1.5}$$

with $\phi_j(y_1, y_2, \ldots, y_j)$ such that $0 \leqslant \phi_j \leqslant 1$ for all j. Where ϕ_j stands for conditional probability that the experimenter will cease sampling, given that he has taken j observations.

Terminal Rule: A terminal rule is a sequence of functions

$$\delta(y) = (\theta_0, \delta_1\ (y_1), \delta_2(y_1, y_2), \ldots) \quad \text{for all } j. \tag{1.6}$$

δ_j is a sequential terminal rule for a statistical decision problem in the probability distribution $\sigma =$ field for which expected lose $E(\theta, (\phi, \delta))$ is finite.

Risk Function: The risk function of a sequential decision rule (ϕ, δ) is the expected value of the risk when θ is the true value of the parameter and will be denoted as $d = d\ (\theta, (\phi, \delta))$.

Sampling Structure: A sampling system F alongwith its risk function d defined over the probbility field (Ω) is called a sampling structure for estimation of θ, symbolically, $D = D(F, d) = (t, \Omega, d)$.

A sampling structure D is said to be unbiased if t is an unbiased estimator of θ.

A sampling structure D is said to be ultimate acceptable if t is a consistent and minimum risk unbiased estimator (MRUE). If the risk of D_1 is smaller than that of D_2, i.e., $d_1 \leqslant d_2$, then t_1 is said to be an acceptable estimator.

2. Sequential Algorithm

A sequential algorithm $\mathcal{R}$ can be defined to generate a sampling structure.

$$\mathcal{R} = \mathcal{R}\,(p(U_i),\ q(s_n),\ r(s_n,\ U_i)) \tag{2.1}$$

where p is a probability measure on U_i

such that $p(U_i) \geqslant 0$ for $1 \leqslant i \leqslant n$

and $\sum\limits_{i \epsilon s} p(U_i) = 1$

q defines a sequential sample s_n of size n lying in (0, 1) and r is defined as a probability measure on the population $\mathcal{U}$ for which

$$q\,(s_n) = 0 \text{ such that } r(s_n,\ U_i) \geqslant 0$$

and $\sum\limits_{i \epsilon s} r(s_n,\ U_i) = 1$

Using the algorithm $\mathcal{R}$ the sequential system runs as follows.

The first unit is selected from $\mathcal{U}$ according to the probability measure p_i. If the sample thus obtained is $s_i = \{U_i\}$, s_1 is then imputed in q and the value is noted. Next a binomial trial is conducted with a probability of $q(s_1)$. If the trial is a failure, the drawing is terminated and s_1 is taken as the sample. Otherwise a second unit is drawn from U according to the measure $r(s_1,\ U_2)$. Let the sample of size 2 be $s_2 = \{U_1,\ U_2\}$. Again s_2 is imputed in q and value is noted. Another binomial trial is conducted with probability of success $q(s_2)$. If the trial results in a failure, the drawing is terminated and s_2 is taken as the sample, otherwise a third unit is drawn from U using the measure $r(s_2,\ U_2)$ and so on. The sequence of sampling continues for which $q(s_n) = 0$.

Actually we have defined a general form of sampling structure for which a sampling system has been prepared. Corresponding to probability field Ω the probability measure $p(U_i)$ has been created, similarly $q(s_n)$ defines the risk function d of D. The sequential estimator has been defined by $r(s_n, U_n)$. Thus a complete sampling structure has been established.

THEOREM 2.1. *The parametric function defined in the relation (1.1) can be estimated unbiasedly by the estimator (1.2) if, and only if, each set 'a_i' is contained in atleast one sample s in the sampling structure D.*

Proof: (*a*) The condition is necessary.

Suppose that set 'a_i' is not included in any sample s of D. As defined in (1.1), the parametric function is estimable by the estimator (1.2) where $\phi(a_i, s)$ is some function of s and a_i.

Because 'a_i' is not included in any sample and therefore $p(s, a_i)$ shall be Zero. Hence there is no estimator. Thus proved that the condition is necessary.

(*b*) The condition is sufficient.

If t is an unbiased estimator of θ, then

$$E(t) = \sum_{s \epsilon A} \sum_{a_i \epsilon s} f(a)\, p(s, a_i)$$

Where $\sum_{a_i \epsilon s} p(s, a_i) = 1$ i.e., a_i is included in the sample and the class A alongwith the probability field defined completely. Hence the condition is sufficient.

THEOREM 2.2. *An unbiased estimator of the variance of the estimator given by the relation (1.2) can be obtained if and only if each set $a_i \cup a_j$ is contained in at least one sample s in the sampling structure D.*

Proof: We know that the variance of the estimator (1.2) may be written as

$$V(t) = E(t^2) - \theta^2$$

Thus it is sufficient to estimate t^2 unbiasedly when an unbiased estimator of $V(t)$ is required.

We know that

$$t^2 = \sum_i \sum_j f(a_i) f(a_j)\, \phi(s, a_i)\, \phi(s, a_j) \tag{2.2}$$

Thus it follows directly from the theorem (1.1) that an unbiased estimator of t^2 is possible if and only if every set $(a_i \cup a_j)$ is included in at least one sample of D. Hence an unbiased estimator of $V(t)$ is given by

$$\hat{V}(t) = t^2 - \sum_{ij} f(a_i)\, f(a_j)\, \phi(s, a_i \cup a_j) / \sum_{i,j} \phi(s, a_i \cup a_j) \tag{2.3}$$

with the condition that

$$\sum_{i,j} \phi(s, a_i \cup a_j) = 1$$

3. Acceptable Generalized Estimators

Let s_{or} be a set of elements of y_i taking into account the 'order' and 'repetition' of the units. The subscripts 'o' and 'r' are used, to denote the order and repetitions of the units respectively. Let S_{or} be the class of all such sets S_{or}, i.e., it may be considered the sample space. This together with a probability measure $\pi(s_{or})$ gives rise to a sampling scheme. If $\tilde{y}$ be a statistic then the sampling system is defined over the sampling scheme. Sample units arranged in the ascending order of their unit indices form an 'order statistics' which may be written as $\tilde{y}_{or} = [y_{(1)}, y_{(2)}, \ldots, y_{(n)}, \ldots]$. This 'order statistic' is always given by a sequential sampling scheme. Another order statistic $\tilde{y}'_{or} = [y_{(1)}, y_{(2)}, \ldots, y_{(r)}, \ldots]$ forms a sufficient statistic if all the units are distinct.

Let $\tilde{Y}_{or}$ be an unbiased estimator which takes into account the order and the repetitions of the units, that it may be written as

$$\tilde{Y}_{or} = \Sigma\, \pi_{(or)}\, Y_{(or)} \tag{3.1}$$

Further if only the repetition is taken into account estimator is given by

$$\tilde{Y}_r\, \Sigma\, \pi'_{(or)}\, \tilde{y}'_{(or)} / \Sigma\, \pi'_{(or)} \tag{3.2}$$

We can proceed to improve it by ignoring the repetitions of the units and defining a new estimator which is based only on the distinct units in the sample.

$$\tilde{Y}_1 = \Sigma\, \tilde{Y}''_{(r)}\, \pi''_{(r)} / \Sigma\, \pi''_{(r)} \tag{3.3}$$

Another approach may be taken by considering the order and ignoring the repetitions of the units in which the estimator may be taken as $\tilde{Y}_0$ based on the sample s_0, then another estimator based on the ordered samples corresponding to s_0 is defined by

$$\tilde{Y}_2 = \Sigma \tilde{Y}_0\, \pi''_0 / \Sigma \pi'''_0 \tag{3.4}$$

For the sake of convenience, let us consider the case in general. Let s be the sample of n units with or without order of units or with or without replacement (repetitions) or both, as the case may be. Let S be the corresponding sample space and the probability of getting the sample s be $\pi_s > 0$. It may be noted that s is being used here to denote s_{or}, s_o, s_r or s, as the case may be, without any loss of generality, it may be assumed that each sample s consists of atleast one set 'a'. Now with the application of Theorems (2.1) and (2.2) it can be shown that these estimators are unbiased with measurable variance. Further if these estimators can be written in the form as given in (1.3), then sampling will terminate. It should be noted that for the estimators to be useful in practice the terminal event E_a should be so specified that it would be possible to calculate the risk function from the information available about the population, the sample and the sampling system.

The technique of generalised estimators given in (1.3) shows that, given any sampling scheme, one can derive a number of unbiased estimators by defining the event E_{ai} in various ways. We may choose one of them by taking into consideration cost and efficiency. In other words, E_a is another way of defining the risk function and thus it is one of the approaches to decide the structure D. This technique systematizes the problem of getting the unique sampling structure.

4. Numerical Illustration

Let the example taken by Murthy [6] (p. 106) be examined. He has drawn a random sample of size 10 from a population of 128 villages. The 128 villages are given running serial numbers from 1 to 128 so that each village is associated with one and only one of these numbers. A number of characteristics has been discussed and the estimates of the population mean have been obtained by simple random sampling with and without replacement with equal probability notions, which can be seen in the reference. With this equal probability notion, we now attempt to illustrate the idea of sequential sampling structure discussed in this paper. Here it is assumed that both the random methods are sequential methods with fixed stages. Corresponding to the sampling structure notion given in this paper, the probability space is equiprobable and is well defined at every stage of the sample. Using the sequential estimators discussed by Singh [9] we may define the risk function as

$$d(\bar{y}) = a + \sum_{i=1}^{n} c_i + \lambda\, L(n)$$

where a is the over-head cost,

c_i is the cost per unit,

λ is the constant, and

$L(n)$ is the loss function of the sampling structure D.

And thus the stopping and terminal rules of the structure are defined. The estimators taken for illustration are:

$\bar{y}_{or}$ = sequential sample mean with order and repetition,

$\bar{y}_0$ = sequential sample mean with order ignoring repetition,

$\bar{y}''_\nu$ = sequential sample mean with r distinct units.

$$= \frac{r\ \bar{y}_0}{E(r)} = \frac{\sum_{i=1}^{\nu} Y_1}{E(r)}$$

${}_r\bar{y}_n$ = random sample mean with replacement and with fixed size n,

$\bar{y}_n$ = Random sample mean without replacement and with fixed size n,

$s(\bar{y})$ = denotes the standard error of corresponding mean $\bar{y}$.

By referring to random numbers there, the first 10 three-digited numbers, which form the sample, are 112, 059, 112, 116, 124, 090, 037, 078, 092, 062 and with slight modification we have taken them as cost per sampling unit i.e. 11.20, 5.90, 11.20, 11.60, 12.40, 9.00, 3.70, 7.80, 9.20, 6.20 respectively. It may be noted that in the sample selected above, which gives sequential sample with replacement, the village with serial number 112 has figured twice. For sampling without replacement, the third draw which is a repetition, has been dropped and to make a sample of size 10, a further draw has been made by taking the next random number in the sequence, which comes out to be 077. Let us take in the risk function, $a = 100$, $\lambda = .10$ (say), c_i is already assumed and $L(n)$ be the standard error of the estimator or the coefficient of variation as the case may be. The results of different characteristics are presented in appendix I (Tables 4.1 to 4.4).

Here the stopping rule of the sampling structure is: Stop sampling when the c.v. or the variance (or standard error) is the least value among all the available results. The terminal rule may be defined as: Terminate sampling when it further increases or the risk function (or the budget) exceeds the given limit or the required number of distinct units are obtained in the sample.

From the Table 4.1, we find that the standard error of the mean for the value $n = 8$ is the least of all the values and that it increases as n increases. Hence we suggest to terminate further sampling for the system $F(\bar{y}, \Omega)$ at the sample size 8. Hence $\bar{y}$ is an acceptable estimator for the sampling structure D. A similar inference may be drawn for the Table 4.2.

We define differently the stopping and terminal rules for the sampling structures in Tables 4.3 and 4.4. The stopping rule is: Stop sampling when (i) the coefficient of variation is the least value among all the available results sequentially. The terminal rule is:

(i) if the budget exceeds, or (ii) the value c.v. or risk function increases further as the sample size increases. From the results in Table 4.3, we suggest to terminate sampling when the c.v. is .45, i.e., $n = 7$ gives the optimal sample size and the estimates given for this value are equally acceptable. In Table 4.4 we recommend to terminate sampling at the sample size $n = 9$, as a further increase in sample will entail to increase the risk function of the sampling structure D. With the help of these results we conclude that the sequential sampling structure always gives acceptable estimators and they are more serviceable than the estimators given by the random sampling scheme.

REFERENCES

[1] Chaudhary, F. S. and Singh, J. B. (1979): On the generalized sequential estimators, *J. Ind. Soc. Agri. Statist.*, **31**; 72.

[2] Godambe, V. P. (1960): An admissible estimate for any sampling design, *Sankhya*, **22**; 285-88.

[3] Hanurav, T. V. (1968): Hyper-admissibility and optimum estimators for Sampling finits populations. *Ann. Math. Statist.*, **39**; 621-42.

[4] Joshi, V. M. (1968): Admissibility of the sample mean as an estimate of the mean of a finite population. *Ann. Math. Statist.*, **39**; 606-20.

[5] Murthy, M. N. and Singh, M. P. (1969): On the concept of best admissible estimators in sampling theory. *Sankhya*, **31**; 343-54.

[6] Murthy, M. N. (1967): *Sampling theory and Methods*. Statist. Pub. Soc., Calcutta.

[7] Prabhu-Ajgaonkar, S. G. (1969): On the non-existence of a best estimator for the entire class of linear estimators. *Sankhya*, **31**; 455-62.

[8] Roy, J. and Chakravarty, I. M. (1960): Estimating the mean of a finite population. *Ann. Math. Statist.*, **31**; 392-98.

[9] Singh, F. (1977): A sequential approach to sample surveys. Ph. D. *Thesis*, Meerut University.

[10] Singh, F. and Singh, D. (1980): On the generalized sequential estimator. *Jour. of Ind. Soc. of Agri. Statistics*, **32**; 1-5.

[11] Singh, F. and Singh, D. (1981): Acceptable Sequential Estimators of Population Mean, J. Ind. Soc. Agri. Statistics, **33**; 29-37.

APPENDIX

TABLE 4.1—SHOWING THE RESULTS OF THE CHARACTERISTIC—NUMBER OF PERSONS IN 1951 CENSUS

$N = 128, \bar{Y} = 3243, S = 1953; {}_r\bar{y}_{10} = 1703, s({}_r\bar{y}_{10}) = 342, d({}_r\bar{y}_{10}) = 222.40, \bar{y}_{10} = 1878, s(\bar{y}_{10}) = 317, d(\bar{y}_{10}) = 216.40$

n	r	c_n	y_n	$\bar{y}_{or}$	$S(\bar{y}_{or})$	$d(\bar{y}_{or}')$	$\bar{y}_o$	$s(\bar{y}_o)$	$d(\bar{y}_o)$	$\bar{y}_r''$	$s(\bar{y}_r'')$	$d(\bar{y}_r'')$
1	1	11.20	695	695	—	—	695	—	—	695	—	—
2	2	5.90	1693	1169	547.3	171.83	1169	1386.3	255.73	1173	1391.0	256.20
3	2	11.20	695	1011	656.0	193.96	1169	1134.1	241.71	785	1147.9	242.69
4	3	11.60	2639	1448	466.1	186.51	1659	984.1	238.31	1258	1997.5	239.65
5	4	12.40	1577	1449	366.3	188.93	1638	881.9	240.49	1331	898.3	242.13
6	5	9.00	1146	1399	366.6	197.96	1540	806.6	241.96	1308	825.6	243.86
7	6	3.70	2654	1578	320.5	197.05	1727	748.2	239.82	1514	769.4	241.94
8	7	7.80	1010	1507	286.2	201.42	1623	701.2	242.92	1459	724.4	245.23
9	8	9.20	953	1445	382.8	220.28	1539	662.4	248.24	1411	687.2	250.72
10	9	6.20	4020	1703	325.6	220.76	1815	629.6	251.16	1691	656.0	253.80
11	10	7.70	2444									

TABLE 4.2—SHOWING THE RESULTS OF THE CHARACTERISTIC—NUMBER OF PERSONS IN 1961 CENSUS

$N = 128$ $\overline{Y} = 3463$, $S = 2065$, ${}_r\overline{y}_{10} = 1797$, $s({}_r\overline{y}_{10}) = 301$, $d({}_r\overline{y}_{10}) = 218.30$, $\overline{y}_{10} = 1942$, $s(\overline{y}_{10}) = 277$, $d(\overline{y}_{10}) = 212.40$

n	r	c_n	y_n	$\overline{y_{or}}$	$s(\overline{y_{or}})$	$d(\overline{y_o})$	$\overline{y}_o$	$s(\overline{y}_o)$	$d(\overline{y}_o)$	$\overline{y}''_r$	$s(\overline{y}''_r)$	$d(\overline{y}''_r)$
1	1	11.30	925	—	—	—	925	—	—	925	—	—
2	2	5.90	1808	1366	509.8	168.08	1366	1465.8	263.68	1371	1470.9	264.19
3	2	11.20	925	1219	612.9	189.59	1306	1199.2	248.22	981	1209.8	249.27
4	3	11.60	2740	1599	435.3	183.43	1824	1040.5	243.95	1384	1055.1	245.41
5	4	12.40	1757	1631	350.9	187.39	1807	937.5	245.55	1468	950.3	247.32
6	5	9.00	1154	1551	352.7	196.57	1676	352.9	246.59	1424	873.4	248.63
7	6	3.70	2768	1725	314.9	196.49	1858	791.2	244.11	1630	814.0	246.40
8	7	7.80	1021	1637	277.7	200.57	1739	741.5	246.95	1563	766.4	249.43
9	8	9.20	1223	1591	336.2	215.62	1674	700.4	252.04	1535	727.2	254.71
10	9	6.20	3652	1797	386.2	218.82	1894	665.8	254.77	1765	694.2	275.62
11	10	7.70	2373									

TABLE 4.3—SHOWING THE RESULTS OF THE CHARACTERISTIC—NUMBER OF WORKERS IN INDUSTRY IN 1961

$N = 128, \overline{y} = 70, S = 94; {}_r\overline{y} = 33.1, s({}_r\overline{y}_{10}) = 12.9, d({}_r\overline{y}_{10}) = 189.49; \overline{y}_{10} = 39.2, s(\overline{y}_{10}) = 12.1, d(\overline{y}_{10}) = 185.99.$

n	r	c_n	y_n	$\overline{y}_{0r}$	$s(\overline{y}_{0r})$	$d(\overline{y}_{0r})$	$\overline{y}_0$	$s(\overline{y}_0)$	$d(\overline{y}_0)$	$\overline{y}''_r$	$s(\overline{y}''_r)$	$d(\overline{y}''_r)$
1	1	11.20	1	18.15	20.2	—	1	—	—	1	—	—
2	2	5.90	36	18.5	20.2	119.12	18.5	66.7	123.77	18.57	66.6	123.76
3	2	11.20	1	12.6	11.9	129.4	18.5	54.5	133.75	12.4	54.5	133.75
4	3	11.60	6	11.0	24.4	142.3	14.3	47.3	144.63	10.8	47.4	144.64
5	4	12.40	100	28.8	18.9	154.1	35.7	42.4	156.54	29.0	42.5	156.55
6	5	9.00	24	28.0	20.4	163.3	33.4	38.8	165.18	28.0	38.9	165.15
7	6	3.70	107	39.3	17.9	166.7	45.7	36.0	168.60	28.3	36.1	168.61
8	7	7.80	5	35.0	16.1	174.4	39.8	33.7	176.17	40.0	33.9	176.19
9	8	9.20	0	41.1	14.4	183.4	34.9	37.8	185.18	35.8	32.0	185.20
10	9	6.20	51	33.1	13.1	189.4	36.7	30.3	191.23	41.9	30.5	191.25
11	10	7.70	62									

TABLE 4.4—SHOWING THE RESULTS OF THE CHARACTERISTIC—CULTIVATED AREA IN 1961 CENSUS

$N = 1258$ $\bar{y} = 1943$, $S = 1107$; $\bar{y}_{10} = 1132$, $s(\bar{y}_{10}) = 221$, $d({}_r\bar{y}_{10}) = 210{,}30$; $\bar{y}_{10} = 1230$, $s(\bar{y}_{10}) = 199$, $d(\bar{y}_{10}) = 208.10$

n	r	C_n	y_n	$\bar{y}_{or}$	$s(\bar{y}_{or})$	$d(\bar{y}_{or})$	$\bar{y}_o$	$s(\bar{y}_o)$	$d(\bar{y}_o)$	$\bar{y}''_r$	$s(\bar{y}''_r)$	$d(\bar{y}''_r)$
1	1	11.20	428	428	—	—	428	—	—	428	—	—
2	2	5.90	1314	871	511.5	160.65	871	785.8	195.68	874	788.9	195.99
3	2	11.20	428	723	364.5	164.75	571	642.8	192.58	585	649.1	193.22
4	3	11.60	1328	874	259.1	165.81	1023	557.8	195.68	776	566.4	196.54
5	4	12.40	772	854	212.7	173.57	960	499.0	202.29	780	510.4	203.54
6	5	9.00	509	796	354.0	196.70	870	457.2	207.01	739	469.3	208.23
7	6	3.70	2622	1057	299.4	194.94	1161	424.1	207.41	1019	437.5	208.75
8	7	7.80	980	1047	279.7	200.77	1136	397.5	212.55	1021	412.1	214.01
9	8	9.20	1881	1040	246.9	203.69	1229	375.4	219.54	1127	391.2	221.12
10	9	6.20	1053	1132	201.4	208.34	1209	356.9	223.69	1127	373.6	225.56
11	10	7.70	1414									