A New Approach to Sampling from Finite Populations. II Distribution-free Sufficiency

By V. P. GODAMBE†

The Johns Hopkins University, Baltimore, U.S.A.

[Received June 1965. Revised November 1965]

SUMMARY

The idea that in some situations the prior knowledge about the unknown parameters could be formulated as a class of prior distributions, which could be used, not necessarily through Bayes posterior probability, for subsequent inference, is already present in Godambe (1955). In the present paper, the concept of distribution-free linear sufficiency or in short linear sufficiency, originally due to Barnard (1963b) but redefined by the present author (Part I), is extended by defining distribution-free sufficiency, removing the restriction of linearity. This extension again is based on the assumption that in some situations prior knowledge could be formulated as a class of prior distributions. A certain linear estimator of the population total, which in Part I was shown to satisfy the redefined criteria of linear sufficiency uniquely in the class of all linear estimators, is now shown to satisfy this extended criteria of distribution-free sufficiency in the entire class of estimators. Further the general relationship between the linear sufficiency of Part I and the distribution-free sufficiency introduced here is investigated. Broadly the result is that, if we restrict to linear estimators only, distribution-free sufficiency is identical with linear sufficiency. Finally, some remarks are offered by way of comparison between the result obtained by the author previously (Godambe, 1955) and the result here, about the utilization of the prior information. The approach of this paper clearly implies a generalization of Fisherian sufficiency suitable for the situations when prior knowledge consists of a class of prior distributions. An alternative generalization when the prior knowledge consists of a group structure is due to Barnard (1963a).

1. INTRODUCTION

WE use the same notation as in Part I. The population consists of N units denoted by the integers $i = 1, \ldots, N$. The variate value associated with the unit i is x_i $(i = 1, \ldots, N)$. It is assumed that x_i $(i = 1, \ldots, N)$ is a real variate and the space of all possible vectors $\mathbf{x} = (x_1, \ldots, x_N)$ is the Euclidean space R_N. The population total is a function T on R_N given by

$$T(\mathbf{x}) = \sum_1^N x_i. \qquad (1)$$

If s denotes a subset of integers i, where $i = 1, \ldots, N$ and S the totality of such subsets, then a sampling design p is a function p on S such that $\sum_S p(s) = 1$ and $p(s) \geqslant 0$ for $s \in S$. We shall call s a sample. As has been demonstrated several times by the author (Godambe, 1955, 1960, 1965), all the known survey designs (such as stratification, subsampling, p.p.s., etc.) are special cases of the sampling design defined above.

† On leave from the Institute of Science, Bombay.

Often in survey sampling the problem is to estimate the population total T on the basis of a sample s and the values x_i, $i \in s$ when the sample s is drawn at random with probability $p(s)$ specified by the given sampling design p. Now we have

Definition 1.1. Any real function $e(s, \mathbf{x})$, on $S \times R_N$, depending on

$$\mathbf{x} = (x_1, \ldots, x_i, \ldots, x_N),$$

only through those x_i for which $i \in s$ is called an estimator.

2. NON-EXISTENCE OF AN UNBIASED ESTIMATE WITH MINIMUM VARIANCE

Given a sampling design p, we say an estimator $e(s, \mathbf{x})$ is unbiased for the population total $T(\mathbf{x})$ if

$$\sum_S e(s, \mathbf{x}) p(s) = T(\mathbf{x}) \quad \text{for all} \quad \mathbf{x} \in R_N. \tag{2}$$

Next, the variance of an unbiased estimator e is given by

$$V(e, \mathbf{x}) = \sum_S \{e(s, \mathbf{x}) - T(\mathbf{x})\}^2 p(s). \tag{3}$$

If now for a given sampling design p, $\mathscr{B}$ denotes the class of all unbiased estimators e, i.e. each $e \in \mathscr{B}$ satisfies (2), then it has been proved by Godambe and Joshi (1965) that $\mathscr{B}$ does not contain an estimator $\tilde{e}$ for which

$$V(\tilde{e}, \mathbf{x}) \leqslant V(e, \mathbf{x}) \quad \text{for all} \quad e \in \mathscr{B} \quad \text{and} \quad \mathbf{x} \in R_N. \tag{4}$$

It is important to note the generality of the above "non-existence of an unbiased minimum variance estimator". It is valid for every sampling design admitting more than one unbiased estimator for the population total T, and for every interval of R_N. In view of this result, then restricted to linear estimators only, the author (1955), proposed a criterion for an estimator to be optimal with respect to a certain type of prior knowledge. This is discussed further in Section 8. In what follows, an alternative approach to the utilization of prior knowledge in a more general and realistic way is suggested.

3. THE CONCEPT OF INDEPENDENCE WITH RESPECT TO SOME KNOWLEDGE

We now suppose that the prior knowledge of the statistician about the vector $\mathbf{x} = (x_1, \ldots, x_N)$ is $\mathbf{K}$, where

Assumption 3.1. $\mathbf{K}$: Different co-ordinates of $\mathbf{x}$ are in no way mutually related. That is, the value that any particular co-ordinate takes is in no way dependent (in the present sense) on the values of the other co-ordinates. In other words whatever may be the statistician's knowledge about some co-ordinates of the vector $\mathbf{x}$, it can impart no knowledge about the remaining co-ordinates of $\mathbf{x}$. Next we have

Assumption 3.2. The prior knowledge $\mathbf{K}$ as given by Assumption 3.1 is equivalent to the class Ω of prior distributions ξ on R_N such that for all $\xi \in \Omega$, when $x_1, \ldots, x_N$ are distributed as ξ, they are probabilistically independent. Further we introduce the following notion of "independence with respect to the knowledge $\mathbf{K}$".

Definition 3.1 (see Appendix). Any two real functions f and g on R_N are said to be independent with respect to $\mathbf{K}$ if for all $\xi \in \Omega$,

$$E_\xi(fg) = E_\xi(f)\, E_\xi(g), \tag{5}$$

$E_\xi(.)$ denoting the expectation when ξ is the distribution.

Now from Definition 1.1 it is clear that an estimator $e(s, \mathbf{x})$ for every *fixed* s is a function on R_N. Hence we have

Definition 3.2. Any two estimators e_1 and e_2 are said to be independent with respect to $\mathbf{K}$ if for all $\xi \in \Omega$,

$$E_\xi(e_1 e_2 | s) = E_\xi(e_1 | s) E_\xi(e_2 | s) \quad \text{for all} \quad s \in S, \tag{6}$$

where $E_\xi(.\,|s)$ denotes the expectation when s is fixed and the distribution is ξ.

Definition 3.3. An estimator e and a function g on R_N are said to be independent with respect to $\mathbf{K}$ if for every $\xi \in \Omega$,

$$E_\xi(eg | s) = E_\xi(e | s) E_\xi(g) \quad \text{for all} \quad s \in S. \tag{7}$$

Remark 3.1. It will be clear from what follows that, for any given sampling design p, no estimator need be defined for samples s with probability $p(s) = 0$. Hence in Definitions 3.2 and 3.3, for any given sampling design p, (6) and (7) need not be satisfied for samples s with $p(s) = 0$.

Remark 3.2. In Definitions 3.1, 3.2 and 3.3, the words "e_1 and e_2 (f and g) are said to be independent with respect to $\mathbf{K}$" can be replaced by "$e_1(f)$ is said to be $\mathbf{K}$-independent of $e_2(g)$". Hereafter "independence" always means "independence with respect to $\mathbf{K}$" as contained in the above definitions.

4. DISTRIBUTION-FREE SUFFICIENCY

We now introduce the notion of distribution-free sufficiency with respect to $\mathbf{K}$ or, more briefly, DF-sufficiency, as follows.

Definition 4.1. An estimator e is said to be DF-sufficient for the population total T if any other estimator e_1 which is $\mathbf{K}$-independent of e (Definition 3.2) is also $\mathbf{K}$-independent of T (Definition 3.3).

Remark 4.1. Definition 4.1 can of course be generalized for any function f on R_N by just replacing T by f. But in this paper we shall restrict ourselves to DF-sufficiency for T.

Theorem 4.1. An estimator $\bar{e}(s, \mathbf{x})$ given by

$$\bar{e}(s, \mathbf{x}) = k(s) \sum_{i \in s} x_i \tag{8}$$

is DF-sufficient for the population total T, where k is a function on S, the set of all possible samples s.

Proof. From Definition 4.1 it follows that $\bar{e}$ in (8) is DF-sufficient if, for all other estimators e_1,

$$\int \bar{e} e_1 \, d\xi = \int \bar{e} \, d\xi \int e_1 \, d\xi \tag{9}$$

implies

$$\int T e_1 \, d\xi = \int T \, d\xi \int e_1 \, d\xi \quad \text{for all} \quad s \in S, \xi \in \Omega. \tag{10}$$

In (9) and (10) $\bar{e}$ stands for $\bar{e}(s, \mathbf{x})$, e_1 for $e_1(s, \mathbf{x})$ and the integrals are taken over R_N for a fixed s. Now,

$$\int T e_1 \, d\xi = \int \left(\sum_{i \in s} x_i \right) e_1 \, d\xi + \int \left(T - \sum_{i \in s} x_i \right) e_1 \, d\xi. \tag{11}$$

Further, since by Assumption 3.2, $x_1, \ldots, x_N$ when distributed as ξ are probabilistically independent, in (11) we have from Definition 1.1,

$$\int \left(T - \sum_{i \in s} x_i\right) e_1 \, d\xi = \int \left(T - \sum_{i \in s} x_i\right) d\xi \int e_1 \, d\xi. \tag{12}$$

Thus from (11) and (12) we have

$$\int T e_1 \, d\xi = \int \left(\sum_{i \in s} x_i\right) e_1 \, d\xi + \int \left(T - \sum_{i \in s} x_i\right) d\xi \int e_1 \, d\xi. \tag{13}$$

Again from (8) and (9) we have in (13)

$$\int \left(\sum_{i \in s} x_i\right) e_1 \, d\xi = \int \left(\sum_{i \in s} x_i\right) d\xi \int e_1 \, d\xi. \tag{14}$$

Thus from (13) and (14) we get

$$\begin{aligned} \int T e_1 \, d\xi &= \int \left(\sum_{i \in s} x_i\right) d\xi \int e_1 \, d\xi + \int \left(T - \sum_{i \in s} x_i\right) d\xi \int e_1 \, d\xi \\ &= \int T \, d\xi \int e_1 \, d\xi, \end{aligned}$$

which proves Theorem 4.1. In this connection also note the remark in Section 7.

We repeat the definition of a fixed sample size design from Part I as

Definition 4.2. If $n(s)$ denotes the size of the sample s, i.e. the total number of units i such that $i \in s$, then p is said to be a fixed sample size design if for all $s \in S$

$$n(s) \neq \text{constant} \rightarrow p(s) = 0. \tag{15}$$

Theorem 4.2. For a sampling design p of fixed sample size n an unbiased DF-sufficient estimator e^*, for the population total T, is given by

$$e^*(s, \mathbf{x}) = \frac{1}{{}^{N-1}C_{n-1} \, p(s)} \sum_{i \in s} x_i, \tag{16}$$

assuming $p(s) > 0$ for all the ${}^N C_n$ samples s.

Theorem 4.2 is an immediate consequence of Theorem 4.1 and equation (2).

In the next section we shall present another approach to DF-sufficiency incorporating unbiased estimation more naturally. The advantage of Definition 4.1 is that it does not require the additional concept of unbiased estimation.

5. Another Approach to Distribution-free Sufficiency

In this section we replace Definition 4.1 by the following

Definition 5.1. For a given sampling design p, an estimator e is said to be DF-sufficient for the population total T if

$$\sum_{S} e(s, \mathbf{x}) p(s) = T(\mathbf{x}) \quad \text{for all} \quad \mathbf{x} \in R_N \tag{17}$$

and for every other estimator e_1 which is **K**-independent of e (Definition 3.2)

$$\sum_{S} e_1(s, \mathbf{x}) p(s) = g(\mathbf{x}) \quad \text{for all} \quad \mathbf{x} \in R_N, \tag{18}$$

where the function g in (18) is **K**-independent of T (Definition 3.1).

Remark 5.1. Here we dispense with Definition 3.3, defining the **K**-independence of an estimator e and a function g on R_N.

Theorem 5.1. For a sampling design p of fixed sample size n a DF-sufficient estimator e^* for the population total T is given by (16), assuming $p(s) > 0$ for all the ${}^N C_n$ samples s.

Proof. Since for e^* in (16) the condition (17) is obviously satisfied, according to Definition 5.1, e^* is DF-sufficient if for all other estimators e_1

$$\int e^* e_1 \, d\xi = \int e^* \, d\xi \int e_1 \, d\xi \quad \text{for all} \quad s \in S, \xi \in \Omega \tag{19}$$

(all integrals being taken on R_N) implies

$$\int T \left\{ \sum_S e_1(s, \mathbf{x}) p(s) \right\} d\xi = \int T d\xi \int \left\{ \sum_S e_1(s, \mathbf{x}) p(s) \right\} d\xi, \tag{20}$$

for all $\xi \in \Omega$, and (20) can be obtained by multiplying (10) by $p(s)$ and summing it over all $s \in S$. This proves Theorem 5.1 as "(19) implies (10)" due to Theorem 4.1.

6. A Relationship between Distribution-free Sufficiency and Linear Sufficiency

It is of interest to note that the estimators given by (8) and (16) are the same as those which were proved to be unbiased and linearly sufficient in Part I.

In the following, we shall prove that if we restrict ourselves to a sub-class $\bar{\Omega} \subset \Omega$ of prior distributions in Assumption 3.2, given by

$$\bar{\Omega} = [\xi \in \Omega \colon E_\xi\{x_i - E_\xi(x_i)\}^2 = \sigma^2, i = 1, \ldots, N], \tag{21}$$

then the linear sufficiency of Part I follows as a special case of DF-sufficiency. To do this, we very briefly recapitulate some notation of Part I. A linear estimator $e_b(s, \mathbf{x})$ is a function on $S \times R_N$ given by

$$e_b(s, \mathbf{x}) = \sum_{i=1}^{N} b(s, i) \, x_i, \tag{22}$$

where b is any real function on $S \times U$ (U is the set of integers 1 to N), subject to the condition $b(s, i) = 0$ for all (s, i) such that $i \notin s$. If $\mathbf{b}(s)$ denotes the vector

$$\mathbf{b}(s) = \{b(s, 1), \ldots, b(s, i), \ldots, b(s, N)\}, \tag{23}$$

then (22) can be written as the scalar product of two vectors, viz.

$$e_b(s, \mathbf{x}) = \mathbf{b}(s) \, \mathbf{x}, \tag{24}$$

where $\mathbf{x} = (x_1, \ldots, x_i, \ldots, x_N)$. Similarly any linear function f on R_N is given by

$$f(\mathbf{x}) = \sum_{i=1}^{N} f_i \, x_i = \mathbf{f}\mathbf{x}, \tag{25}$$

where $\mathbf{f} = (f_1, \ldots, f_i, \ldots, f_N)$, f_i $(i = 1, \ldots, N)$ are any real constants.

Theorem 6.1. If in Definitions 3.1, 3.2 and 3.3, Ω is replaced by $\bar{\Omega}$ given by (21), the estimators restricted to linear estimators defined by (22) and functions on R_N to linear functions defined by (25), then equations (5), (6) and (7) are equivalent to the following equations (27), (28) and (29) respectively.

Using the notation in (25), if the two functions f and g in (6) are given by

$$f(\mathbf{x}) = \sum_1^N f_i x_i = \mathbf{fx} \quad \text{and} \quad g(\mathbf{x}) = \sum_1^N g_i x_i = \mathbf{gx},$$

then for all $\xi \in \bar{\Omega}$ in (21), we have

$$E_\xi(fg) = \sigma^2 \sum_1^N f_i g_i + E_\xi(f)\, E_\xi(g). \tag{26}$$

Thus (5) and (26) together imply the scalar product

$$\mathbf{fg} = \sum_1^N f_i g_i = 0. \tag{27}$$

Similarly, we can prove that if in (6) the two estimators e_1 and e_2 are given by

$$e_1(s, \mathbf{x}) = \sum_1^N b^1(s,i)\, x_i = \mathbf{b}^1(s)\,\mathbf{x} \quad \text{and} \quad e_2(s, \mathbf{x}) = \sum_1^N b^2(s,i)\, x_i = \mathbf{b}^2(s)\,\mathbf{x},$$

for all $\xi \in \bar{\Omega}$, then equation (6) is equivalent to

$$\mathbf{b}^1(s)\,\mathbf{b}^2(s) = \sum_1^N b^1(s,i)\, b^2(s,i) = 0 \quad \text{for all} \quad s \in S. \tag{28}$$

Again for the same reason, if in (7) we have

$$e(s, \mathbf{x}) = \sum_1^N b(s,i), x_i = \mathbf{b}(s)\,\mathbf{x} \quad \text{and} \quad g(\mathbf{x}) = \sum_1^N g_i x_i = \mathbf{gx},$$

for all $\xi \in \bar{\Omega}$ then equation (7) is equivalent to

$$\mathbf{b}(s)\,\mathbf{g} = \sum_1^N b(s,i)\, g_i = 0 \quad \text{for all} \quad s \in S, \tag{29}$$

proving Theorem 6.1.

Now we consolidate all the above discussion in this section in

Theorem 6.2. If we restrict ourselves to linear estimators (defined by (22)), linear functions (defined by (25)) and the class of prior distributions $\bar{\Omega}$ given by (21), then Definitions 4.1 and 5.1 of DF-sufficiency here are identical with Definitions 4.4 and 6.1, respectively, of linear sufficiency in Part I. Similarly, Theorems 4.1 and 5.1 are here identical with Theorems 4.1 and 6.3, respectively, of Part I.

7. Distribution-Free Sufficiency and Bayes Estimation

A Bayes estimator $e^*(s, \mathbf{x})$ with respect to the prior distribution ξ on R_N and squared error as loss for the population total T is defined by the inequality,

$$\int \left[\sum_S p(s)\{e^*(s,\mathbf{x}) - T(\mathbf{x})\}^2\right] d\xi \leqslant \int \left[\sum_S p(s)\{e(s,\mathbf{x}) - T(\mathbf{x})\}^2\right] d\xi, \tag{30}$$

where e is any other estimator. Obviously an estimator e^* satisfying (30) is given by

$$e^*(s,\mathbf{x}) = \frac{\int_{R_N(x_i, i\in s)} T d\xi}{\int_{R_N(x_i, i\in s)} d\xi} \quad \text{for all} \quad s\in S, \mathbf{x}\in R_N, \tag{31}$$

where $R_N(x_i, i\in s)\subset R_N$ such that for every $\mathbf{x}\in R_N(x_i, i\in s)$ the ith co-ordinate has a specified value, namely x_i, $i\in s$. It is important to note that the Bayes estimator e^* in (31) does not depend on the sampling design. Now (31) also implies an informative posterior distribution for T, in contrast to the non-informative likelihood (cf. remark in Section 3 of Part I). A similar observation in another connection is due to Cornfield (1965).

Now consider a sub-class $\Omega_y\subset\Omega$ (Ω as in Assumption 3.2) such that, for some specified numbers y_i ($i=1,\ldots,N$) for every $\xi\in\Omega_y$,

$$\int x_i\, d\xi = y_i \quad (i=1,\ldots,N). \tag{32}$$

It is easy to see that for every $\xi\in\Omega_y$ the Bayes estimator e^* in (31) is given by

$$e^*(s,\mathbf{x}) = \sum_{i\in s} x_i + \left(\sum_1^N y_i - \sum_{i\in s} y_i\right) \tag{33}$$

for all $s\in S$ and $\mathbf{x}\in R_N$.

It is interesting to note that the estimator e^* in (33) is DF-sufficient.

Remark. With slight modification of the proof of Theorem 4.1, we can show that any estimator given by

$$\bar{e}(s,\mathbf{x}) = k_1(s)\sum_{i\in s} x_i + k_2(s), \quad s\in S,\ \mathbf{x}\in R_N \tag{34}$$

(where k_1 and k_2 are constants *not* depending on x_i, $i\in s$, but depending on s) is DF-sufficient. Clearly the estimator e^* in (33) is a special case of $\bar{e}$ in (34).

This again suggests that the approach through DF-sufficiency is possibly more general than the Bayes approach which depends on a stronger prior knowledge (characterized by the class of prior distributions Ω_y) than $\mathbf{K}$ in Assumptions 3.1 and 3.2. Actually the estimator e^* in (33) has so far never been recommended in practice. On the other hand, ratio-type and regression estimates which are special cases of (34) are already in use. However, in a subsequent publication we shall discuss more fully the Bayes approach to this problem.

8. DISTRIBUTION-FREE SUFFICIENCY AND UNBIASED ESTIMATION WITH MINIMUM EXPECTED VARIANCE

Now with Assumption 3.2 and Theorem 5.1, following the argument in Section 9 of Part I, if the statistician happens to know the values y_i ($i=1,\ldots,N$) of a variate y, which is correlated with x, we recommend the use of a sampling design p with a fixed sample size n, given by

$$p(s) = \frac{\sum_{i\in s} y_i}{\sum_1^N y_i} \Big/ {}^{N-1}C_{n-1} \tag{35}$$

for all samples s containing n units i and the use of the ratio-type estimate e^* obtained by substituting (35) in (16).

With an assumption of a stronger prior knowledge than **K** in Assumption 3.1, we can restrict ourselves to a sub-class $\bar{\Omega}$, of the class Ω (in Assumption 3.2), of prior distributions on R_N. If, for any specified numbers y_i ($i = 1, \ldots, N$) and any constant c,

$$\bar{\Omega} = \{\xi \in \Omega : E_\xi(x_i) = y_i, E_\xi(x_i - y_i)^2 = cy_i^2, i = 1, \ldots, N\} \quad (36)$$

it was earlier proved and further confirmed (Godambe, 1955, 1965) that for any sampling design p with fixed sample size (n given), the average of the variance in (3) given by

$$E_\xi\{V(e, \mathbf{x})\} = \int_{R_N} V(e, \mathbf{x})\, d\xi \quad (37)$$

has a lower bound, not depending upon the sampling design, for the class of all linear unbiased estimators e and for all $\xi \in \bar{\Omega}$ in (36). Further it was shown that this lower bound is attained for a sampling design p for which

$$\sum_{s \ni i} p(s) = \frac{ny_i}{\sum_1^N y_i} \quad (i = 1, \ldots, N) \quad (38)$$

($s \ni i$ denoting all samples s which include the unit i) and for the estimator

$$\tilde{e}(s, \mathbf{x}) = \frac{\sum_1^N y_i}{n} \cdot \sum_{i \in s} \frac{x_i}{y_i}. \quad (39)$$

There is an implicit assumption that y_i ($i = 1, \ldots, N$) are such that the right-hand side of (38), for $i = 1, \ldots, N$, lies between 0 and 1. Recently this result has been generalized by Godambe and Joshi (1965), removing the restriction of linearity.

Thus for a given class of prior distributions $\bar{\Omega}$ in (36), (38) and (39) provide an answer to the problem, left open due to non-existence of an unbiased estimator satisfying (4), by way of recommending the use of a fixed sample size sampling design p, satisfying (38) and the estimator (39). Incidentally the estimator $\tilde{e}$ in (39) is the usual Horvitz–Thompson estimator for the population total T.

Now it will be clear that our recommendation in the first paragraph of this section to use the fixed sample size sampling design p given by (35) and the ratio-type estimator e^* given by (16) and (35) together is based on the assumption of a much *weaker* (and more realistic) prior knowledge **K** characterized by the class of prior distributions Ω (Assumption 3.2) which is far broader than $\bar{\Omega}$ in (36). Further, the approach through the concept of distribution-free sufficiency as put forward in Parts I and II of this paper appears to be more basic and fundamental than the conventional approach.

ACKNOWLEDGEMENTS

The author acknowledges with pleasure some useful discussions with Professors L. J. Savage, Allan Birnbaum and J. Durbin. Of course they are not responsible for the views expressed in the paper. He also has to thank Mrs Carole Mullen and Mrs Sara Donato for doing the necessary typing . A part of this work was carried out at the Dominion Bureau of Statistics, Ottawa, and part was sponsored by the Air Force Office of Scientific Research, Contract No. AF49(638)-1302.

REFERENCES

BARNARD, G. A. (1963a), "Some logical aspects of the fiducial argument", *J. R. statist. Soc.* B, **25**, 111–114.

—— (1963b), "The logic of least squares", *J. R. statist. Soc.* B, **25**, 124–127.

CORNFIELD, J. (1965), "A note on the likelihood function generated by randomization over a finite set". Read before the 35th Session of the International Statistical Institute at Belgrade.

GODAMBE, V. P. (1955), "A unified theory of sampling from finite populations", *J. R. statist. Soc.* B, **17**, 269–278.

—— (1960), "An admissible estimate for any sampling design", *Sankhyā*, **22**, 285–288.

—— (1965), "A review of the contributions towards a unified theory of sampling from finite populations", *Int. statist. Inst. Rev.*, **33**, 242–258.

GODAMBE, V. P. and JOSHI, V. M. (1965), "Admissibility and Bayes estimation in sampling from finite populations: I", *Ann. math. Statist.*, **36**, 1707–1722.

APPENDIX

It has been proved by Joshi (1965, unpublished) that if two real functions f and g on R_N are "independent with respect to the knowledge **K**", as in Definition 3.1, then f and g are necessarily probabilistically independent when $x_1, ..., x_N$ are distributed as ξ for all $\xi \in \Omega$, Ω being the class defined in Assumption 3.2.

The above result suggests an evident modification of Definitions 3.1, 3.2, 3.3 and consequently of Definition 4.1, replacing in them the "uncorrelatedness" by "probabilistic independence". This modification helps to clarify further the intuitive relationship between DF-sufficiency (Section 4) and the conventional sufficiency. This will be discussed in a separate publication by Joshi.

Joshi's (1965, unpublished) work also strengthens Theorem 4.1 as: A DF-sufficient estimator of the population total T must necessarily be of the form

$$k_1(s) \sum_{i \in s} x_i + k_2(s),$$

where k_1 and k_2 are any functions defined on S.