The Annals of Statistics
1978, Vol. 6, No. 5, 1168–1169

STRATIFIED SAMPLING WITH EXCHANGEABLE PRIOR DISTRIBUTIONS

BY M. E. THOMPSON[1]

University of Waterloo

Suppose a finite population is stratified such that the variate values, appropriately scaled, are exchangeable within strata and independent from stratum to stratum. Consider designs with fixed sample sizes within strata, and unbiased estimators of the population total. Minimum expected variance is achieved with a stratified πps design and the usual (Horvitz-Thompson) unbiased estimator.

1. Introduction. Optimality theorems for estimator-design pairs have been proved by Godambe and Thompson (1973) when the population has an exchangeable prior distribution, and by Bellhouse et al. (1977) when the population is "multistage exchangeable." The corresponding result for stratified populations is straightforward; nevertheless it is important, since it provides a justification for the usual unbiased estimator used with stratified random sampling when the population itself is stratified a priori.

2. Notation. Suppose that $U = \{1, 2, \cdots, N\}$ is a finite population, partitioned into strata U_h, $h = 1, \cdots, k$. With each $i \in U$ is associated an unknown value y_i of the variable of interest, and a known value $a_i > 0$ of some auxiliary characteristic. The estimand is the population total $T(\mathbf{y}) = \sum_{i=1}^{N} y_i$.

A sample is a subset of the population U, and a sampling design is a probability function p on the set S of all samples s. If s is a sample, let $s_h = s \cap U_h$ for each h. A sampling design will be said to be of fixed size, $(\mathbf{n})$, where $\mathbf{n} = (n_1, \cdots, n_h, \cdots, n_k)$, if $p(s) > 0$ implies that the cardinality of s_h is n_h for each h.

3. The theorem. Let C be a class of "prior" distributions ξ of $\mathbf{y} \in R^N$ satisfying the following conditions: under ξ, the quantities $z_i = y_i/a_i$ within each stratum are exchangeably distributed; and these exchangeable sets of random variables are independent from stratum to stratum. For example, these conditions are satisfied if $\mathbf{z}$ is obtained from some fixed vector $\mathbf{w}$ by a randomly selected "stratified permutation" π, where π permutes the labels of the w-values within each stratum. Let $\mathscr{E}$ denote expectation with respect to ξ, let E denote sampling expectation, and let B_p be the class of all estimators e which are unbiased for $T(\mathbf{y})$ under the design p.

Received April 1977; revised November 1977.

[1] Research supported by the National Research Council of Canada.

AMS 1970 *subject classification.* Primary 62O05.

Key words and phrases. Optimality, finite populations, superpopulation, stratified sampling, exchangeable priors.

We assume that the 'possible' values of $\mathbf{z}$ are sufficiently many that the completeness of the "stratified order statistic" for sampled values of z can be proved, in analogy with Royall's (1968) proof for the ordinary order statistic.

THEOREM. *Suppose $n_1, \cdots, n_k$ are fixed positive numbers and consider the class of estimator-design pairs (e, p) where p has fixed size $(\mathbf{n})$ and $e \in B_p$. If $\xi \in C$, then the expected variance*

$$\mathscr{E}\,\mathrm{E}(e - T(\mathbf{y}))^2 \tag{3.1}$$

is minimized when the inclusion probabilities $\pi_i = \sum_{s \ni i} p(s)$ are proportional to the a_i within strata, and when e is given by

$$\sum_{i \in s} \frac{y_i}{\pi_i} = \sum_h \left(\sum_{i \in U_h} a_i / n_h\right) \sum_{i \in s_h} (y_i / a_i)\,. \tag{3.2}$$

REFERENCES

BELLHOUSE, D. R., THOMPSON, M. E. and GODAMBE, V. P. (1977). Two-stage sampling with exchangeable prior distributions. *Biometrika* **64** 97–103.

GODAMBE, V. P. and THOMPSON, M. E. (1973). Estimation in sampling theory with exchangeable prior distributions. *Ann. Statist.* **1** 1212–21.

ROYALL, R. (1968). An old approach to finite population sampling. *J. Amer. Statist. Assoc.* **63** 1269–79.

DEPARTMENT OF STATISTICS
UNIVERSITY OF WATERLOO
WATERLOO, ONTARIO
CANADA N2L 3G1