Ann. Inst. Statist. Math.
35 (1983), Part A, 25-30

A CLASS OF ADMISSIBLE ESTIMATORS OF A FINITE POPULATION TOTAL

K. W. TSUI

(Received Mar. 1, 1982)

Summary

In order to estimate the total value of an attribute of a finite population, Brewer [2] proposed an estimator which is asymptotically design-unbiased and which is optimal with respect to a certain superpopulation model. In this note, it is shown that a class of estimators which includes Brewer's estimator, as well as the usual ratio estimator, is admissible for any fixed population size. The proof of the result follows that of Joshi [4], [5].

1. Introduction

The problem of estimating the total value of an attribute of a finite population is frequently encountered. Some attention has been directed towards determining the admissibility of various particular total estimators (e.g., Joshi [4], [5], [6]). Brewer [2] proposed an estimator which is asymptotically design-unbiased and which is optimal with respect to a certain superpopulation model. However, the question of admissibility of his estimator was unresolved. In this note, it is shown that a general class of estimators which includes Brewer's estimator, as well as the usual ratio estimator, is admissible under squared error loss for any fixed population size. The proof of the result follows that of Joshi [4], [5]. Although Joshi's notation is different than that generally used currently, it will be employed here in order to facilitate comparison with his proofs.

Consider a finite population U of N distinct units labeled i, $i=1, \cdots, N$. With each unit i is associated a pair of quantities (y_i, x_i), $i=1, \cdots, N$. The x_i are the quantities of interest while the y_i are positive and known constants. Denote $x=(x_1, \cdots, x_N)$. Let $p(\cdot)$ be a sam-

AMS 1980 subject classification: Primary 62D05.
Key words and phrases: Admissible estimator, Cramér-Rao inequality, squared error loss, superpopulation model.

pling design and S be the set of all possible samples $s \subset U$ (i.e. $\sum_{s \in S} p(s) = 1$). If a sample s is taken, then values $\{(i, y_i, x_i): i \in s\}$ will be recorded. A problem of interest in many surveys is to estimate the total $T(x) = \sum_{i \in U} x_i$, based on the sample drawn under $p(\cdot)$. Let π_i be the inclusion probability of unit i in the sample, i.e., $\pi_i = \sum_{i \in s} p(s)$. Brewer [2] proposes the following estimator of $T(x)$:

$$(1) \qquad e_B(s, x) = \sum_{i \in s} x_i + (\sum_{i \notin s} y_i) (\sum_{i \in s} (\pi_i^{-1} - 1) x_i) / \sum_{i \in s} (\pi_i^{-1} - 1) y_i \ ,$$

where the π_i satisfy certain conditions to guarantee that $e_B(s, x)$ is asymptotically design-unbiased and has asymptotically smallest mean square error (for appropriate increases of N and the sample size $n(s)$) under the model

$$(2) \qquad x_i = \beta y_i + \varepsilon_i \ , \quad \mathrm{E}(\varepsilon_i) = 0 \quad \text{and} \quad \mathrm{E}(\varepsilon_i \varepsilon_j) = \begin{cases} \sigma_i^2 & i = j \ , \\ 0 & \text{otherwise} \ . \end{cases}$$

Although Brewer [2] justified the estimator e_B in a very strictly conditioned asymptotic framework, Robinson and Tsui [8] relaxed the conditions and provided a general asymptotic framework in which the same conclusion can be derived. Note that e_B is simply the usual ratio estimator when a simple random sampling design is used.

In this note the general class of estimates of the form

$$(3) \qquad \hat{e}(s, x) = \sum_{i \in s} x_i + (\sum_{i \notin s} y_i)(\sum_{i \in s} w_i x_i) / (\sum_{i \in s} w_i y_i) \ ,$$

where $w_i > 0$, $i = 1, \cdots, N$, is shown to be admissible (for fixed N) under squared error loss. That is, there is no other estimator $e'(s, x)$ such that

$$(4) \qquad \sum_{s \in S} p(s)(e'(s, x) - T(x))^2 \leqq \sum_{s \in S} p(s)(\hat{e}(s, x) - T(x))^2$$

for all x in R_N, the set of all N-dimensional real vectors, and with strict inequality holding for some x in R_N.

The estimator $\hat{e}(s, x)$ can be motivated by the prediction theory approach (cf. Royall [9]) as follows: with respect to the model given by (2), the Best Linear Unbiased Estimator (BLUE) for β is $\hat{\beta} = (\sum_{i \in s} x_i \cdot y_i \sigma_i^{-2}) / (\sum_{i \in s} y_i^2 \sigma_i^{-2})$, which becomes $(\sum_{i \in s} w_i x_i) / (\sum_{i \in s} w_i y_i)$ if $w_i = y_i \sigma_i^{-2}$, $i = 1, \cdots, N$. Based on the sample s, the estimator of the population total $T(x)$ based on the prediction theory approach is $\sum_{i \in s} x_i + \hat{\beta} \sum_{i \notin s} y_i$, which is exactly the estimator $\hat{e}(s, x)$. Our admissibility result in the next section shows that even though (3) can be motivated under a superpopulation model, it is actually admissible from the randomization theory approach. With appropriate choices of the w_i's, the estimator $\hat{e}(s, x)$ encompasses the following special cases: (a) the usual ratio estimator, (b) Brewer's estimator given by (1), and (c) the estimator $\sum_{i \in s} x_i + (\sum_{i \notin s} y_i)(\sum_{i \in s} x_i/y_i)/n(s)$, suggested by Basu [1], where $n(s)$ is the number of units in the sample s. Meeden and Ghosh [7] provide an alternative proof of the admissibility of Basu's estimator. Their proof, however, pertains only to the admissibility of Basu's estimator, while the admissibility proof in this paper covers a broad class of estimators.

2. The admissibility result

The proof is similar to the one given by Joshi [5], who proved that the usual ratio estimator is admissible. Modifications were required in the definitions of $B(s)$ and $\bar{x}(s)$ below, in the assumptions about the prior variances σ_i^2, and in the calculation of the constant h.

Suppose there exists an estimate $e'(s, x)$ satisfying (4) with strict inequality holding for some $x \in R_N$. Denote

$$A(s) = \sum_{i \notin s} y_i ,$$

$$(5) \qquad g(s, x) = [A(s)]^{-1}(e'(s, x) - \sum_{i \in s} x_i) ,$$

$$\bar{x}(s) = \sum_{i \in s} w_i x_i / \sum_{i \in s} w_i y_i .$$

Suppose the prior distribution of the x_i is such that the x_i are all distributed independently with mean $\mathrm{E}(x_i) = \theta y_i$. We have

$$(6) \qquad \begin{aligned} \mathrm{E}\,[e'(s, x) - T(x)]^2 &= \mathrm{E}\,[e'(s, x) - \sum_{i \in s} x_i - \theta A(s) - \sum_{i \notin s} (X_i - \theta y_i)]^2 \\ &= A^2(s)\,\mathrm{E}\,[g(s, x) - \theta]^2 + \sum_{i \notin s} \mathrm{E}\,(x_i - \theta y_i)^2 , \end{aligned}$$

where the expectation E is taken with respect to the prior described. Moreover,

$$(7) \qquad \begin{aligned} &\mathrm{E}\,[\hat{e}(s, x) - T(x)]^2 \\ &= \mathrm{E}\,\{[\sum_{i \in s} w_i x_i / \sum_{i \in s} w_i y_i](\sum_{i \notin s} y_i) - \theta A(s) - \sum_{i \notin s} (x_i - \theta y_i)\}^2 \\ &= A(s)^2\,\mathrm{E}\,(\bar{x}(s) - \theta)^2 + \sum_{i \notin s} \mathrm{E}\,(x_i - \theta y_i)^2 . \end{aligned}$$

Thus, taking expectations of both sides of (4) with respect to the prior and cancelling out common terms, we have

$$(8) \qquad \sum_{s \in \bar{S}} p(s) A^2(s)\,\mathrm{E}[g(s, x) - \theta]^2 \leqq \sum_{s \in \bar{S}} p(s) A^2(s)\,\mathrm{E}\,[\bar{x}(s) - \theta]^2$$

where $\bar{S}=\{s \in S: p(s) \neq 0\}$. If we further assume that each x_i is distributed normally with variance $\sigma_i^2 = ky_i/w_i$ where $k>0$, then

$$\text{(9)} \qquad \mathrm{E}\,[\bar{x}(s)-\theta]^2 = k/B(s)$$

where

$$\text{(10)} \qquad B(s)=\sum_{i \in s} w_i y_i \, .$$

The joint distribution of the x_i for $i \in s$ is

$$L=(2\pi)^{-n(s)/2} \prod_{i \in s} \sigma_i^{-1} \exp\left(-\frac{1}{2} \sum_{i \in s} (x_i-\theta y_i)^2/\sigma_i^2\right) .$$

We have, by (9) and the special form of the σ_i^2:

$$\text{(11)} \qquad \mathrm{E}\,(\partial \log L/\partial\theta)^2 = \mathrm{E}\,[\sum_{i \in s} (y_i/\sigma_i^2)(x_i-\theta y_i)]^2 = B(s)/k \, .$$

Let $\mathrm{E}\,(g(s, x)) = \theta + b(s, \theta)$, where $b(s, \theta)$ is the bias of the estimate. Using the Cramér-Rao inequality, we have

$$\mathrm{E}\,(g(s, x)-\theta)^2 \geqq [k/B(s)](1+b'(s, \theta))^2 + b^2(s, \theta) \, .$$

Inequality (8) now becomes:

$$\sum_{s \in \bar{S}} p(s)A^2(s)b^2(s, \theta) + k \sum_{s \in \bar{S}} p(s)[A^2(s)/B(s)](1+b'(s, \theta))^2 \leqq k \sum_{s \in \bar{S}} p(s)A^2(s)/B(s) \, .$$

Proceeding as Joshi did ([5], pp. 1661–1662 and [4], pp. 1733–1734), we see that $b(s, \theta)=0$ for all $s \in \bar{S}$. Therefore, $g(s, x)=\bar{x}(s)$ for almost all $x \in R_N$, which in turn implies that $e'(s, x)=\hat{e}(s, x)$ for almost all $x \in R_N$. In other words, $\hat{e}(s, x)$ is weakly admissible in the class of all measurable estimates.

To strengthen the result so that $e'(s, x)=\hat{e}(s, x)$ for *all* $x \in R_N$, we next show that Theorem 4.1 of Joshi [4] holds for $\hat{e}(s, x)$. Following Joshi's [5] argument again, let Q_{N-k}^{α} and $Q_{N-k}^{\alpha'}$ be hyperplanes in R_N such that the last k coordinates are fixed to be $\alpha=(\alpha_{N-k+1}, \cdots, \alpha_N)$ and $\alpha'=(\alpha'_{N-k+1}, \cdots, \alpha'_N)$, respectively. Let $\bar{S}_k=\{s \in \bar{S}: i \in s,\ i=N-k+1, \cdots, N\}$, i.e., $\bar{S}_k$ consists of all the samples s containing the last k units. To establish a $1-1$ correspondence between the points of Q_{N-k}^{α} and $Q_{N-k}^{\alpha'}$, we put

$$x_i' = x_i + hy_i \, , \qquad i=1, \cdots, N-k \, .$$

The constant h is found by equating

$$\text{(12)} \qquad \hat{e}(s, x') - T(x') = \hat{e}(s, x) - T(x) \, ,$$

which as shown in the appendix, is given by

$$h=\sum_{i=N-k+1}^{N} w_i(\alpha_i'-\alpha_i)\Big/\sum_{i=N-k+1}^{N} w_i y_i \;.$$

As Joshi [4] argues, the estimate $e'(s, x)$ can be extended to every hyperplane $Q_{N-k}^{\alpha'}$ by setting

$$e'(s, x')-T(x')=e'(s, x)-T(x) \;.$$

Applying the remaining argument in the proof of Theorem 4.1 of Joshi [4], we see that his Theorem 4.1 holds for our $\hat{e}$. For reference, we state the result below.

THEOREM. *If an estimate $e'(s, x)$ satisfies*

$$\sum_{s\in \bar{S}_k} p(s)(e'(s, x)-T(x))^2 \leqq \sum_{s\in \bar{S}_k} p(s)(\hat{e}(s, x)-T(x))^2$$

for almost all x in Q_{N-k}^{α}, then

$$e'(s, x)=\hat{e}(s, x) \qquad \textit{for almost all } x \textit{ in } Q_{N-k}^{\alpha} \;.$$

Furthermore, Theorem 5.1 in Joshi [4] also applies and $e'(s, x)-\hat{e}(s, x)$ cannot be different from zero for any $x \in R_N$. This shows that (4) cannot be a strict inequality for any $x \in R_N$. In other words, $\hat{e}(s, x)$ is admissible.

Appendix

The constant h that satisfies equation (12)

$$
\begin{aligned}
&[\hat{e}(s, x')-T(x')]-[\hat{e}(s, x)-T(x)]=0\\
&\iff \sum_{\substack{i\in s\\ i\leqq N-k}} (x_i+hy_i)+\sum_{i=N-k+1}^{N} \alpha_i'\\
&\quad +\Big(\sum_{i\notin s} y_i\Big)\frac{\sum_{\substack{i\in s\\ i\leqq N-k}} (w_i x_i+h w_i y_i)+\sum_{i=N-k+1}^{N} w_i\alpha_i'}{\sum_{i\in s} w_i y_i}\\
&\quad -\sum_{i\leqq N-k} (x_i+hy_i)-\sum_{i=N-k+1}^{N} \alpha_i'-\sum_{\substack{i\in s\\ i\leqq N-k}} x_i-\sum_{i=N-k+1}^{N} \alpha_i\\
&\quad -\frac{\sum_{\substack{i\in s\\ i\leqq N-k}} w_i x_i+\sum_{i=N-k+1}^{N} w_i\alpha_i}{\sum_{i\in s} w_i y_i}\sum_{i\notin s} y_i\\
&\quad +\sum_{i\leqq N-k} x_i+\sum_{i=N-k+1}^{N} \alpha_i=0
\end{aligned}
$$

$$\iff \sum_{\substack{i\in s\\ i\leq N-k}} hy_i + \frac{\sum_{\substack{i\in s\\ i\leq N-k}} hw_iy_i + \sum_{i=N-k+1}^{N} w_i(\alpha_i'-\alpha_i)}{\sum_{i\in s} w_iy_i} \sum_{i\notin s} y_i$$

$$- \sum_{i\leq N-k} hy_i = 0$$

$$\iff -h\sum_{i\notin s} y_i + \frac{h\sum_{i\in s} w_iy_i - h\sum_{i=N-k+1}^{N} w_iy_i + \sum_{i=N-k+1}^{N} w_i(\alpha_i'-\alpha_i)}{\sum_{i\in s} w_iy_i}$$

$$\times \sum_{i\notin s} y_i = 0$$

$$\iff h = \sum_{i=N-k+1}^{N} w_i(\alpha_i'-\alpha_i) \Big/ \sum_{i=N-k+1}^{N} w_iy_i .$$

UNIVERSITY OF WISCONSIN-MADISON

REFERENCES

[1] Basu, D. (1971). An essay on the logical foundations of survey sampling, part one, *Foundations of Statistical Inference* (eds. V. P. Godambe and D. A. Sprott), Holt, Rinehart and Winston, Toronto, 203-233.

[2] Brewer, K. R. W. (1979). A class of robust sampling designs for large scale surveys, *J. Amer. Statist. Ass.*, **74**, 911-915.

[3] Hodges, J. L., Jr. and Lehmann, E. L. (1951). Some applications of the Cramér-Rao inequality, *Proc. Second Berkeley Symp. Math. Statist. Prob.*, University of California Press, 13-22.

[4] Joshi, V. M. (1965). Admissibility and Bayes estimation in sampling finite populations, III, *Ann. Math. Statist.*, **36**, 1723-1729.

[5] Joshi, V. M. (1966). Admissibility and Bayes estimation in sampling finite populations, IV, *Ann. Math. Statist.*, **37**, 1658-1670.

[6] Joshi, V. M. (1979). Joint admissibility of the sample means as estimators of the means of finite populations, *Ann. Statist.*, **7**, 995-1002.

[7] Meeden, G. and Ghosh, M. (1983). Choosing between experiments: applications to finite population sampling, *Ann. Statist.*, **11**, to appear.

[8] Robinson, P. M. and Tsui, K. W. (1981). Optimal asymptotically design-unbiased estimation of a population total, *Technical Report* No. 634, Dept. of Statistics, Univ. of Wisconsin-Madison.

[9] Royall, R. M. (1970). On finite population sampling theory under certain linear regression models, *Biometrika*, **64**, 499-513.