PREDICTION THEORY IN FINITE POPULATION SAMPLING: THE HORVITZ-THOMPSON ESTIMATOR AND ESTIMATES OF ITS VARIANCE

William G. Cumberland, U.C.L.A.
Richard M. Royall, Johns Hopkins University

1. INTRODUCTION

Prediction models in finite population sampling theory can reveal relationships which are essential for making inferences but which are often concealed in probability sampling analyses. Here we examine the Horvitz-Thompson estimator under some linear prediction models and compare the results with those derived through probability sampling theory under a probability-proportional-to-size (πps) sampling plan. Various estimators of variance are compared theoretically, using prediction models, and empirically, using a πps sampling plan to draw samples of n=32 repeatedly from six real populations.

2. THEORY USING THE PROBABILITY SAMPLING DISTRIBUTION

With each of the N population units two numbers are associated: the variable of interest y and an auxiliary variable x whose value is known. A sample s of n units is chosen, leaving a set r of N-n nonsample units. The sample means are denoted by $\bar{x}_s$ and $\bar{y}_s$, the non-sample means by $\bar{x}_r$ and $\bar{y}_r$ and the population means by $\bar{x}$ and $\bar{y}$. The sampling fraction n/N is denoted by f.

A probability sampling plan p specifies for every possible sample s the probability p(s) that s will be selected. The probability that unit i will be in s is $\pi_i=\Sigma_{s\ni i}p(s)$. When $x_i>0$ is a measure of the size of unit i, any plan p for which π_i is proportional to x_i is called a probability-proportional-to-size, or πps, sampling plan. We consider only sampling without replacement with a fixed sample size n. In this case $\pi_i=fx_i/\bar{x}$ for all i=1, ..., N (possible only when max $x_i\leq\bar{x}/f$), and the Horvitz-Thompson estimator of the population total $T=N\bar{y}$ is $\hat{T}_{HT}=Nb\bar{x}$ where $b=\Sigma_s(y_i/nx_i)$. With respect to the πps sampling plan, this estimator is unbiased with variance

$$\mathrm{var}_p(\hat{T}_{HT})=\Sigma_1^N\Sigma_1^N(\pi_{ij}-\pi_i\pi_j)(y_i/\pi_i)(y_j/\pi_j) \qquad (2.1)$$

where π_{ij} is the probability that a sample containing both units i and j will be selected. Two variance estimators are commonly presented with $\hat{T}_{HT}$. One (Horvitz and Thompson 1952) is

$$v_{HT}=\Sigma_s\Sigma_s(1-\pi_i\pi_j/\pi_{ij})(y_i/\pi_i)(y_j/\pi_j)$$

and the other (Yates and Grundy 1953, Sen 1953) is

$$v_{YG}=\tfrac{1}{2}\Sigma_s\Sigma_s(\pi_i\pi_j/\pi_{ij}-1)(y_i/\pi_i-y_j/\pi_j)^2 \ .$$

If the probabilities π_{ij} are positive for all (i,j) pairs, then both v_{HT} and v_{YG} are unbiased estimators of $\mathrm{var}_p(\hat{T}_{HT})$.

Theoretical comparison of these two variance estimators has proved difficult. Although all πps sampling plans for a given population and sample size n have the same inclusion probabilities, the joint inclusion probabilities π_{ij} can be different for different plans. This means that $\mathrm{var}_p(\hat{T}_{HT})$, v_{HT}, and v_{YG} and the relationships among them can be different for different plans. In the special case of perfect proportionality $\hat{T}_{HT}=T$ for every sample, so $\mathrm{var}_p(\hat{T}_{HT})=0$. In this case $v_{YG}=0$, but v_{HT} need not equal zero. This implies that v_{HT} has greater sampling variance than v_{YG} in populations where y_i is very nearly proportional to x_i. It also shows that v_{HT} can be negative in such populations. It is less obvious but also well-known that v_{YG} too can take on negative values. On the other hand, v_{YG} but not v_{HT} has been shown to be non-negative for any sample chosen using some particular πps sampling plans (see Cochran 1977). The available empirical evidence favors v_{YG} over v_{HT}, but this is based primarily on samples of n=2 from small populations (Rao and Singh 1973).

A third variance estimator sometimes recommended for its simplicity is one derived under the assumption of sampling with replacement. If we define $d_i=y_i-bx_i$, then this estimator is $v_C=(N\bar{x}^2/f)\Sigma_s(d_i/x_i)^2/(n-1)$. Under a πps sampling plan v_C is biased. The bias is positive if the πps plan is more efficient than sampling independently with replacement and with probability $x_i/N\bar{x}$ of choosing unit i on each draw (Durbin 1953).

The πps sampling plan we have used in this study is the well-known one investigated by Hartley and Rao (1962). The sampling procedure consists of a random permutation of the units followed by a random-start systematic sample with step size $\bar{x}/f$ through the interval $(0,N\bar{x})$. Unit i is selected if $\Sigma_1^{i-1}x_j\leq U<\Sigma_1^i x_j$ for one of the systematic sample points U. For this plan v_C has a positive sampling bias.

3. THEORY USING PREDICTION MODELS

The preceding theory treated the y's as fixed constants. All probabilistic calculations referred to the distribution created by the probability sampling plan p. For instance, the bias in $\hat{T}_{HT}$ was defined as $\Sigma^* p(s)[N\bar{x}\Sigma_s(y_i/nx_i)-\Sigma_1^N y_i]$ where Σ^* denotes the summation over all s. The only random quantity was s, the set of units chosen as the sample.

In many applications it is useful to regard the y's as realized values of random variables $(Y_1,Y_2,\ldots,Y_N)$. Knowledge about relationships among the Y's is represented in a probability model for their joint distribution. In such cases a different approach to inference becomes available, one based on the probability model instead of the sampling plan. For instance, the bias can be defined as $E_M(\hat{T}_{HT}-T)=E_M[N\bar{x}\Sigma_s(Y_i/nx_i)-\Sigma_1^N Y_i]$. Here the expectation E_M is with respect to the Y-distribution model and is conditioned on the sample s.

3.1 Performance of $\hat{T}_{HT}$. The estimator $\hat{T}_{HT}$ is often studied under a model which represents y_i as a realized value of a random variable Y_i generated according to the model

$$EY_i=\beta x_i,\ \mathrm{var}Y_i=\sigma^2x_i^2,\ \mathrm{cov}(Y_i,Y_j)=0 \quad i\neq j. \qquad (3.1)$$

Under this model $\hat{T}_{HT}$ is unbiased with variance

$$\mathrm{var}_M(\hat{T}_{HT}-T)=\sigma^2N\bar{x}(\bar{x}/f-2\bar{x}_s+c) \qquad (3.2)$$

where $c=N^{-1}\Sigma_1^N x_i^2/\bar{x}$ and the subscript M is a reminder that this calculation is made with respect to the model for a fixed sample s. This variance is a decreasing linear function of $\bar{x}_s$.

3.2 Performance of Variance Estimators. If inferences are to be made conditional on s, then under model (3.1) the variance to be estimated is (3.2), not (2.1). Standard techniques for estimating σ^2 from the weighted squared residuals give

$\hat{\sigma}^2=\Sigma_s(d_i/x_i)^2/(n-1)$, and substituting this estimate in (3.2) yields the least squares variance estimator v_L. This estimator is unbiased under model (3.1) and is closely related to v_C: $v_L=v_C[1-f(2\bar{x}_s-c)/\bar{x}]$.

Royall and Cumberland (1978) have studied procedures for generating bias-robust variance estimators for linear statistics. These techniques produce estimators of $\text{var}_M(\hat{T}_{HT}-T)$ which are unbiased under model (3.1) and which remain approximately unbiased when $\text{var}(Y_i)$ is not proportional to x_i^2. One of these was used in this study:

$$v_D=(N\bar{x}^2/f)\Sigma_s(d_i/x_i)^2/(n-1)-2N\bar{x}\Sigma_s(d_i^2/x_i)/(n-1)$$
$$+n(\Sigma_1^N x_i^2/\Sigma_s x_i^2)\Sigma_s d_i^2/(n-1) \ .$$

The statistics v_{HT}, v_{YG}, and v_C can also be viewed as estimators of (3.2) under model (3.1). For the sampling plan we have chosen Hartley and Rao (1962) showed that for $i\neq j$

$$n\pi_{ij}/\pi_i\pi_j=(n-1)[1+a_{ij}/N+b_{ij}/N^2+0(1/N^3)] \quad (3.3)$$

where $a_{ij}=(x_i+x_j-c)/\bar{x}$ and $b_{ij}=2(x_i^2+x_j^2+x_ix_j/\bar{x}^2)$ $-3c(x_i+x_j-c)\bar{x}^2-2N^{-1}\Sigma_1^N(x_i/\bar{x})^3$. Approximating the right hand side of (3.3) by $(n-1)(1+a_{ij}/N)$ and substituting in the formula for v_{YG} leads to the following approximation: $v_{YG}^*=(N\bar{x}^2/f)\Sigma_s(d_i/x_i)^2/(n-1)$ $-N\bar{x}\Sigma_s(d_i^2/x_i)/(n-1)-N\bar{x}(\bar{x}_s-c)\Sigma_s(d_i/x_i)^2/(n-1)$. A similar substitution leads to an approximation for v_{HT}: $v_{HT}^*=v_{YG}^*+N\bar{x}\,\Sigma_s[\bar{x}_s-c-(x_i-c)/n](y_i/x_i)^2/(n-1)$

All the variance estimators v_C, v_D, v_L, v_{HT}^* and v_{YG}^* have the same leading term, $(N\bar{x}^2/f)\Sigma_s(d_i/x_i)^2/(n-1)$, and are thus asymptotically equivalent (as $n\to\infty$ and $f\to 0$). But when f is moderate there are important differences among the estimators. Under the model (3.1) v_D, v_L, and v_{YG}^* are unbiased estimators of the variance (3.2), while $E_M(v_C)=\sigma^2N\bar{x}^2/f$ and $E_M(v_{HT}^*)=\sigma^2N\bar{x}^2/f-\sigma^2N\bar{x}\,\bar{x}_s+\beta^2N\bar{x}(\bar{x}_s-c)$. Defining the relative bias of an estimator by

rel bias$(v)=[E_M(v)-\text{var}_M(\hat{T}_{HT}-T)]/\text{var}_M(\hat{T}_{HT}-T)$ we have the following approximations under model (3.1):

$$\text{rel bias}(v_C)\doteq f(2\bar{x}_s-c)/\bar{x}$$
$$\text{rel bias}(v_{HT}^*)\doteq(1+g)f(\bar{x}_s-c)/\bar{x} \qquad (3.4)$$

where $g=(\beta/\sigma)^2$. These expressions show that the relative bias in v_C is small, for moderate f, except when there is extreme variability among the x's and the sample units are chosen from the extreme x's. The only variance estimator whose expected value depends on β is v_{HT}^*, and this dependence ensures that v_{HT}^* will perform disastrously when g is large unless the sample is one where $f|\bar{x}_s-c|/\bar{x}$ is very small. Not only <u>can</u> v_{HT}^* take negative values, its expectation is negative when $\bar{x}_s<x^*$, where

$$x^*=(c-\bar{x}/fg)\left(\frac{g}{g-1}\right) \ . \qquad (3.5)$$

Since $E_M(v_{HT}^*)$ is a linear function of $\bar{x}_s$, the situation gets progressively worse with smaller $\bar{x}_s$. In the limit, when $\sigma^2=0, y_i=\beta x_i$, and $g=\infty$, we have $v_{HT}^*=\beta^2N\bar{x}(\bar{x}_s-c)$, a linear function of $\bar{x}_s$ with slope $\beta^2N\bar{x}$ while the actual variance (3.2) and the estimators v_C, v_D, v_L, and v_{YG} all equal zero for every sample. Note that c is the expected value of $\bar{x}_s$ under πps sampling: $E_p(\bar{x}_s)=c$.

3.3 <u>Effects of Model Failure</u>. Failure of the variance specification in (3.1) does not bias $\hat{T}_{HT}$ but does affect its variance as well as all of the variance estimators. Under the more general model

$$EY_i=\beta x_i,\ \text{var}Y_i=v_i,\ \text{cov}(Y_i,Y_j)=0,\quad i\neq j \quad (3.6)$$

the variance of $\hat{T}_{HT}$ is given by

$$\text{var}_M(\hat{T}_{HT}-T)=(N\bar{x}^2/f)\Sigma_s v_i/nx_i^2-2N\bar{x}\Sigma_s v_i/nx_i+\Sigma_1^N v_i \ .$$

Modifications of the arguments given in Royall and Cumberland (1978) show that the estimator v_D is bias-robust: as an estimator of $\text{var}_M(\hat{T}_{HT}-T)$, v_D is unbiased under the model (3.1) and is asymptotically unbiased under the model (3.6) in the sense that the rel bias$(v_D)\to 0$ as $n\to\infty$ and $f\to 0$. Since the other variance estimators have the same dominant term as v_D, all these estimators have the bias-robust property.

The flaw in v_{HT}^* is also robust, persisting under the general model (3.6). This is because the term in $E_M(v_{HT}^*)$ which depends on β, namely $\beta^2N\bar{x}(\bar{x}_s-c)$, which caused the problem under model (3.1), remains unchanged under the more general model (3.6).

Failure of the assumption that $E(Y_i)=\beta x_i$ in (3.1) can seriously bias the estimator $\hat{T}_{HT}$. For example, under the general J^{th} order polynomial regression model

$$EY_i=\Sigma_{j=0}^J\beta_jx_i^j,\ \text{var}(Y_i)=v_i,\ \text{cov}(Y_i,Y_j)=0\quad i\neq j \quad (3.7)$$

the bias is

$$E_M(\hat{T}_{HT}-T)=N\Sigma_{j=0}^J\beta_j\Delta_j(s) \qquad (3.8)$$

where $\Delta_j(s)=\bar{x}\Sigma_s x_i^{j-1}/n-N^{-1}\Sigma_1^N x_i^j$. The bias vanishes when $\Delta_j(s)=0$ for $j=0,1,\ldots,J$, and we describe a sample satisfying these conditions as "π-balanced". Note that $\Delta_1(s)\equiv 0$ and that the π-balance conditions are met in expectation under a πps sampling plan: $E_p[\Delta_j(s)]=0$.

The π-balance conditions are analogous to the balance conditions studied by Royall and Herson (1973) for the ratio estimator. Scott, Brewer, and Ho (1978) discussed different balance conditions appropriate for the BLU estimator under the model (3.1).

When the regression function contains an omitted or unknown regressor z, say $EY_i=\beta x_i+\delta z_i$, $\hat{T}_{HT}$ incurs a bias $E_M(\hat{T}_{HT}-T)=[N\bar{x}\Sigma_s(z_i/nx_i)-\Sigma_1^N z_i]\delta$. The bias is zero when $(N\bar{x})\Sigma_s(z_i/nx_i)=\Sigma_1^N z_i$, and this is again a condition which is met in expectation under a πps sampling plan. Here πps sampling plays the same role in producing approximately π-balanced samples (on both x and z) as simple random sampling plays in producing approximately balanced samples. For a discussion of this role see Royall and Herson (1973) Section 6.

Failure of the condition $EY_i=\beta x_i$ biases the variance estimates but does not affect the variance of $\hat{T}_{HT}$. At π-balance the effect is to make v_C, v_D, and v_L conservative as estimates of $E_M(\hat{T}_{HT}-T)^2$ in that their biases are positive when EY_i is a polynomial.

4. EMPIRICAL STUDY

Six real populations described in Table 1 were used in an empirical study of the preceding theoretical results. These are the same populations used in a previous study of the ratio estimator except that the four largest units have been removed from the original Sales population so that max $x_i<\bar{x}/f$. These are populations where a straight line through the origin regression model with variance proportional to either x or x^2 might be a reasonable first approximation. More detailed descriptions of the populations, including scatter plots, appear in Royall and Cumberland (1980).

From each population we drew 1000 πps samples of n=32 using the sampling plan described in Section 2.1. For each sample we calculated the Horvitz-Thompson estimator, the actual error, and the five variance estimates. The estimators v_{HT} and v_{YG} were calculated from the Hartley-Rao formula (3.3), including the terms of order $1/N^2$. The average values are presented in Table 2. As expected, v_C tends to overestimate the variance. The averages for the other estimators are consistent with their being unbiased under πps sampling.

The prediction theory outlined in Section 3 suggests that the performance of the variance estimates will depend strongly on $\bar{x}_s$. The bias incurred by $\hat{T}_{HT}$ when $E_M(\hat{T}_{HT}-T)$ contains a quadratic term also depends on $\bar{x}_s$. To examine performance as a function of $\bar{x}_s$, we arranged the 1000 samples from each population in order of increasing value of $\bar{x}_s$. We then grouped the samples in 20 sets of 50, so that the first group contains the 50 samples with the smallest $\bar{x}_s$, the next group the samples with the next 50 smallest $\bar{x}_s$, etc. For each group we calculated the average values of $\bar{x}_s$, the average error, the mean square error (mse), and the averages of each of the five variance estimates, $\bar{v}_C$, $\bar{v}_D$, $\bar{v}_L$, $\bar{v}_{HT}$ and $\bar{v}_{YG}$. We then plotted the average errors and the values of $(mse)^{\frac{1}{2}}$, $\bar{v}_C^{\frac{1}{2}}$, $\bar{v}_D^{\frac{1}{2}}$, $\bar{v}_L^{\frac{1}{2}}$ and $\bar{v}_{YG}^{\frac{1}{2}}$ against the average values of $\bar{x}_s$. The trajectory showing the average error plotted against average value of $\bar{x}_s$ is labelled error. Those showing $\bar{v}_C^{\frac{1}{2}}$, $\bar{v}_D^{\frac{1}{2}}$, etc., are labelled C, D, etc. Figures 1-6 show the results. The performance of v_{HT} was so dramatically different from the others that we have made separate figures, 7-12, for comparing v_{HT} with v_{YG}.

Probability sampling theory assures us that $\hat{T}_{HT}$ is unbiased in πps sampling. Table 2 confirms this. However, the error curves in Figures 1-6 show that the bias defined with respect to the prediction model and conditioned on s is more appropriate for inference from a given sample. In four populations (Cities, Counties 60, Counties 70 and Hospitals) there is a clear bias in $\hat{T}_{HT}$ as a function of $\bar{x}_s$. The positive and negative biases at the extremes cancel each other when the results are averaged over all samples. Thus averaging over the sampling distribution conceals an important property of the estimator, its bias in samples not achieving good π-balance. The bias seen in $\hat{T}_{HT}$ as a function of $\bar{x}_s$ can be explained in terms of curvature of the true regression function. The term in the bias of $\hat{T}_{HT}$ found in (3.8) resulting from quadratic curvature is $N\bar{x}\beta_2(\bar{x}_s-c)$. Hence for populations like Counties 70 with a convex regression (slope increasing with x), the bias should be an increasing function of $\bar{x}_s$. For populations like Counties 60 and Hospitals where the regression is concave, the bias should be decreasing linear function of $\bar{x}_s$. This bias will be near zero when $\bar{x}_s=c$. This is precisely the behavior suggested by the error curves in the figures.

Note that πps sampling has not provided adequate π-balance to protect $\hat{T}_{HT}$ against bias in these populations. This is particularly true of Counties 60, where the 10-20% of samples in which $|\bar{x}_s-c|$ was greatest produced biases which are large relative to the $(mse)^{\frac{1}{2}}$. This illustrates an important point: the fact the πps sampling produces π-balance "on the average" does not imply that we can make inferences as if each sample were approximately π-balanced.

The variance estimates behave as predicted under model (3.1). The curves for v_D, v_L and v_{YG} all tend to decrease as $\bar{x}_s$ increases, and all three estimators track the mse curve reasonably well except in cases when a large bias influences the mse. The difference between v_C and the others increases as $\bar{x}_s$ increases in all the populations.

To examine the performance of v_{HT} we plotted the average values of v_{HT} and v_{YG} against the average values of $\bar{x}_s$ in Figures 7-12. Although v_{YG} and v_{HT} are unbiased in expectation with respect to a πps sampling plan, prediction theory reveals that v_{HT} can have a serious bias when $\bar{x}_s \neq c$. Under the model (3.1) this bias is a linear increasing function of $\bar{x}_s$. When $\bar{x}_s < x^*$, given in (3.5), the bias will be so large that the expected value under (3.1) of v_{HT}^* is negative. This will also be true of v_{HT}, if v_{HT}^* is an adequate approximation. The theory sketched in Section 3.3 suggests that the larger β^2/σ^2, the more severe the bias in v_{HT}. For each population the least-squares estimates of β and σ^2 under (3.1) based on the entire population were used to estimate β^2/σ^2 (see Table 1). These values were used to estimate x^* (3.5) where $E_M(v_{HT}^*)=0$. The points x^* are on the x-axes in Figures 7-12.

The extent to which prediction theory has correctly summarized the performance of v_{HT} in these populations is remarkable. In every population the v_{HT} line increases linearly with $\bar{x}_s$. In the populations in which the v_{HT} line becomes negative (Cities, Counties 1960, Counties 1970, and Sales), it crosses the axis near x^*. In those populations with the largest values of β^2/σ^2 (Counties 1960 and Sales) performance of v_{HT} is worst. Cancer and Hospitals, having the smallest values of β^2/σ^2, are the only two populations for which the v_{HT} line is non-negative. In both of these populations, x^* is smaller than the smallest average $\bar{x}_s$ value plotted. As theory predicts, the lines for v_{YG} and v_{HT} cross when $\bar{x}_s \doteq c$ in every population.

5. CONCLUSIONS

The variance estimator v_{YG} is generally recommended over v_{HT}. The present study provides strong theoretical and empirical support for this advice. It also suggests that v_D and v_L deserve consideration as practical alternatives. These have two important advantages over v_{YG}--they are always nonnegative and they do not require that the joint inclusion probabilities π_{ij} be available.

We believe these results show again that finite population inferences should be based on prediction models, not on the probability sampling distribution. They help to clarify how probability sampling can contribute to robust inference with $\hat{T}_{HT}$ by producing samples which are approximately π-balanced. But they illustrate the danger in making inferences which are not conditioned on the sample s actually observed.

REFERENCES

[1]Cochran, W. G. (1977), Sampling Techniques, 3rd Edition. New York: John Wiley and Sons.

[2]Durbin, J. (1953). "Some Results in Sampling Theory When The Units Are Selected With Unequal Probabilities", Journal of the Royal Statistical Society B, 15, 262-9.

[3]Hartley, H. O. and Rao, J. N. K. (1962). "Sampling With Unequal Proababilities and Without Replacement", Annals of Mathematical Statistics, 33, 350-74.

[4]Horvitz, D. G. and Thompson, D. J. (1952), "A Generalization of Sampling Without Replacement from A Finite Universe", Journal of the American Statistical Association, 47, 663-85.

[5]Rao, J. N. K. and Singh, M. P. (1973), "On The Choice of Estimator in Survey Sampling", Australian Journal of Statistics, 15, 95-104.

[6]Royall, R. M. and Herson, J. (1973), "Robust Estimation in Finite Populations I.", Journal of the American Statistical Association, 68, 880-9.

[7]Royall, R. M. and Cumberland, W. G. (1978),

"Variance Estimation in Finite Population Sampling", Journal of the American Statistical Association, 73, 351-8.
[8]Royall, R. M. and Cumberland, W. G. (1980), "An Empirical Study of The Ratio Estimator and Estimators of Its Variance", Journal of the American Statistical Association, 75, to appear.
[9]Scott, A. J., Brewer, K. R. W. and Ho, E. W. H. (1978), "Finite Population Sampling and Robust Estimation", Journal of the American Statistical Association, 73, 359-61.
[10]Sen, A. R. (1953), "On The Estimate of Variance in Sampling With Varying Probabilities", Journal of the Indian Society for Agricultural Statistics, 5, 119-27.
[11]Yates, F. and Grundy, P. M. (1953), "Selection Without Replacement from Within Strata With Probability Proportional To Size", Journal of the Royal Statistical Society B, 15, 253-61.

TABLE 1: STUDY POPULATIONS

	Population	x	y	T	b^2/s^2 †
Cancer	301 counties in N.C., S.C. and Ga.	adult white female population 1960	breast cancer mortality 1950-69	11994	9.5
Cities	125 U.S. cities	population 1960	population 1970	35.691×10^6	16.0
Counties 60	304 counties in N.C., S.C. and Ga.	households 1960	population 1960	10.007×10^6	210.6
Counties 70	304 counties in N.C., S.C. and Ga.	households 1960	population 1970	11.243×10^6	39.6
Hospitals	393 hospitals from national sample	number of beds	patients discharged	320159	11.9
Sales	327 U.S. Corporations	gross sales 1974	gross sales 1975	667.88×10^9	50.9

†b and s^2 are weighted least-squares estimates of β and σ^2 under the model (3.1).

TABLE 2: RESULTS FOR 1000 πps SAMPLES OF $n = 32$ (HORVITZ-THOMPSON ESTIMATOR)

		(Average in 1000 samples)$^{\frac{1}{2}}$					
Population	Average Error	$(\hat{T}_{HT}-T)^2$	v_C	v_D	v_L	v_{YG}	v_{HT}
Cancer	4.2	502	539	479	464	489	487
Cities (millions)	.02	1.26	1.46	1.18	1.13	1.23	1.24
Counties 60 (thousands)	.155	110	124	108	105	109	107
Counties 70 (thousands)	15.4	267	318	271	269	266	269
Hospitals (thousands)	-.42	14.0	14.9	14.1	13.9	14.0	14.0
Sales (billions)	.10	13.0	14.1	12.7	12.3	12.9	13.1

FIGURES 1 - 6

HORVITZ-THOMPSON ESTIMATOR

FIGURE 1. CANCER POPULATION
1000 πPS SAMPLES OF n=32
HORVITZ-THOMPSON ESTIMATOR

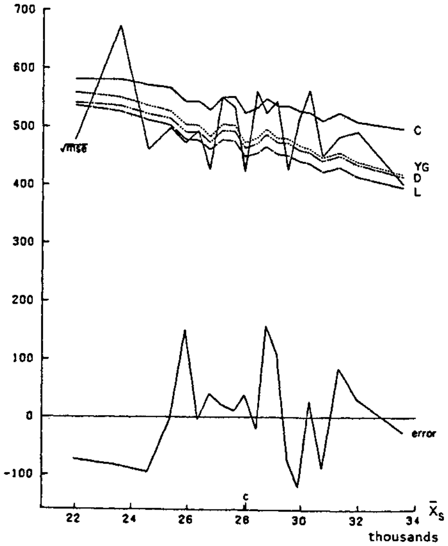

FIGURE 2. CITIES POPULATION
1000 πPS SAMPLES OF n=32
HORVITZ-THOMPSON ESTIMATOR

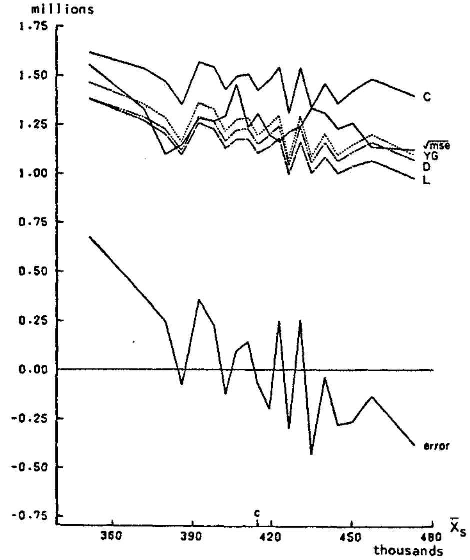

FIGURE 3. COUNTIES 60 POPULATION
1000 πPS SAMPLES OF n=32
HORVITZ-THOMPSON ESTIMATOR

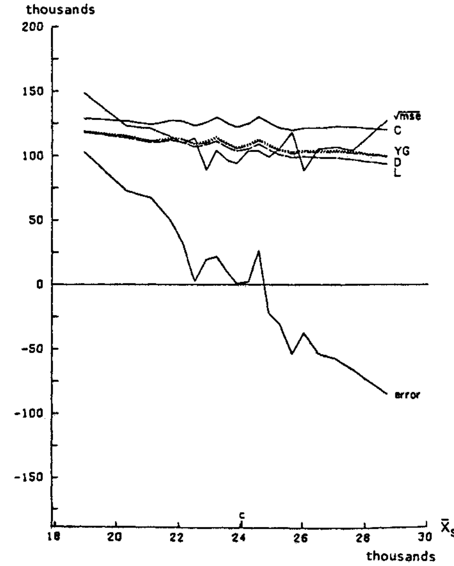

FIGURE 4. COUNTIES 70 POPULATION
1000 πPS SAMPLES OF n=32
HORVITZ-THOMPSON ESTIMATOR

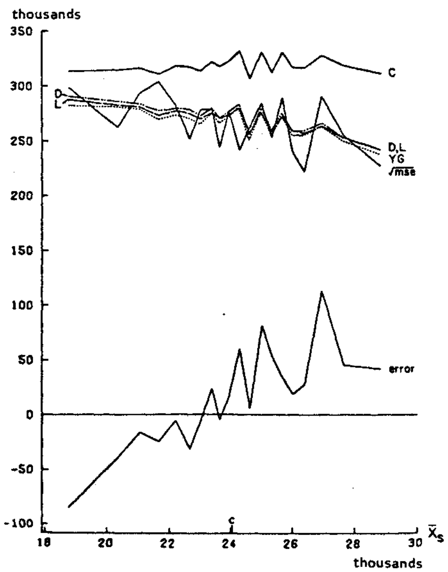

FIGURE 5. HOSPITALS POPULATION
1000 πPS SAMPLES OF n=32
HORVITZ-THOMPSON ESTIMATOR

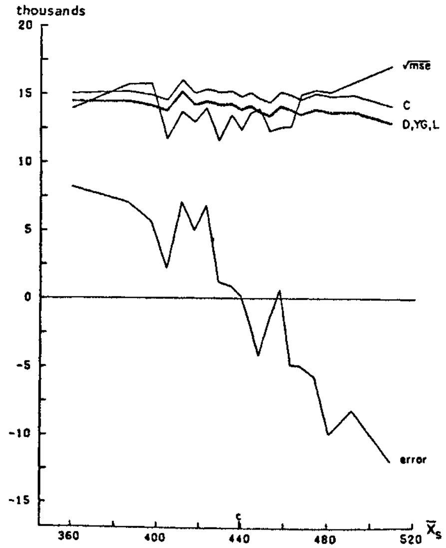

FIGURE 6. SALES POPULATION
1000 πPS SAMPLES OF n=32
HORVITZ-THOMPSON ESTIMATOR

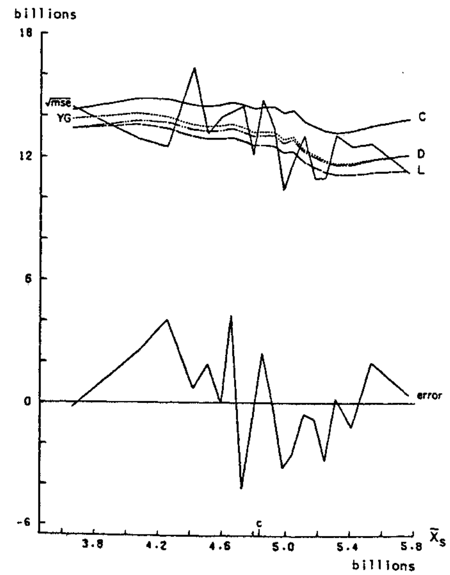

FIGURE 7. CANCER POPULATION
1000 πPS SAMPLES OF n=32

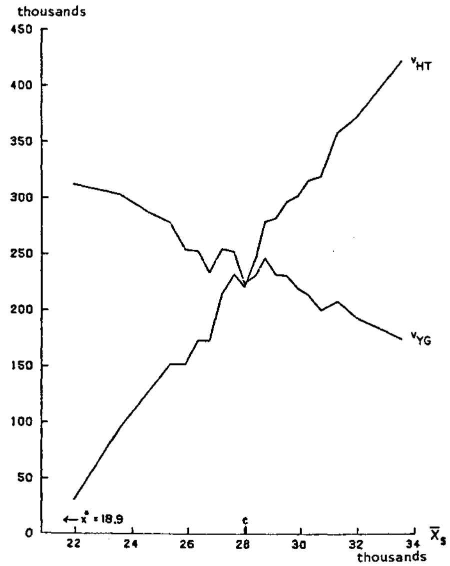

FIGURE 8. CITIES POPULATION
1000 πPS SAMPLES OF n=32

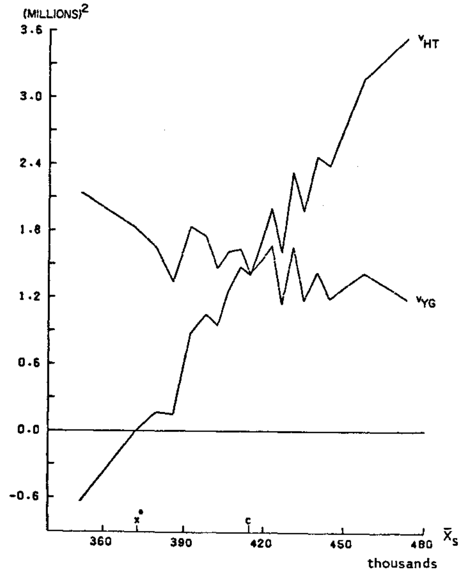

FIGURE 9. COUNTIES 60 POPULATION
1000 πPS SAMPLES OF n=32

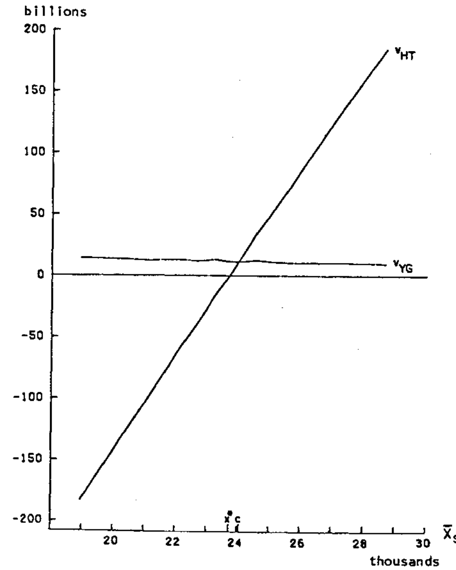

FIGURE 10. COUNTIES 70 POPULATION
1000 πPS SAMPLES OF n=32

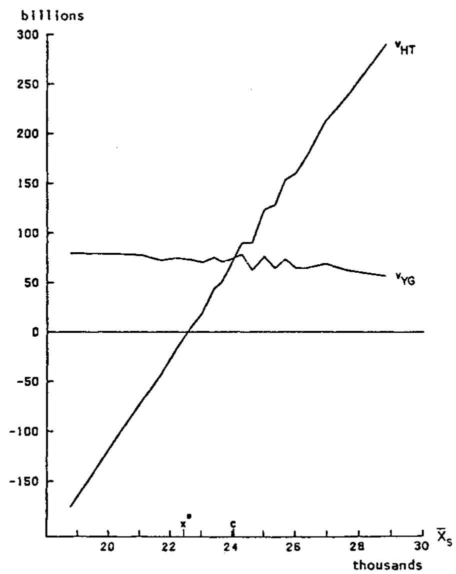

FIGURE 11. HOSPITALS POPULATION
1000 πPS SAMPLES OF n=32

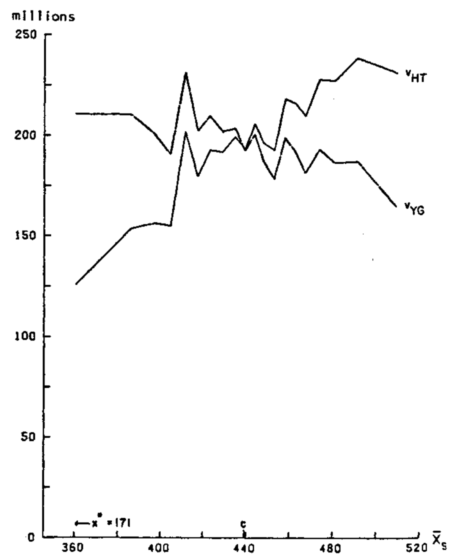

FIGURE 12. SALES POPULATION
1000 πPS SAMPLES OF n=32