Sankhyā : The Indian Journal of Statistics
1998, Volume 60, Series B, Pt. 2, pp. 301-314

A NOTE ON COMPARISON OF ESTIMATION STRATEGIES IN SURVEY SAMPLING OF CONTINUOUS POPULATIONS

By V.R. PADMAWAR
Indian Statistical Institute, Bangalore

SUMMARY. Interpreting the traditional survey sampling set-up in the continuous infinite population framework, the performances of some design-unbiased sampling strategies for estimating the population mean with respect to measures of uncertainty are compared under a well-known regression model.

1. **Introduction**

A large number of sampling strategies for estimating the population mean have been considered in the literature of survey sampling of continuous populations (Cassel and Särndal (1972, 1974), Särndal (1980), Padmawar (1982, 1984, 1996), Cordy (1993)). Särndal (1980) studied certain strategies in the continuous set-up, which were later taken up by Padmawar (1982). Results regarding nonexistence (Padmawar (1982)) and some regarding existence (Padmawar (1984)), of optimal strategies in certain classes of $p-$unbiased strategies are known. Padmawar (1996) defined Rao-Hartley-Cochran strategy in the continuous set-up and studied its efficiency.

In the absence of an optimal $p-$unbiased strategy, we take up, in this note, the problem of comparing the performances of various strategies for estimating the population mean under a well-known regression model. At the end of section 1, we list the strategies to be studied. In section 2, we establish their interesting properties and compare them. In section 3, we study some strategies in the stratified continuous set-up.

We shall use, in this note, the same framework as that in Padmawar (1996).

Paper received. January 1996; revised March 1998.

AMS (1991) *subject classification.* 62D05.

Key words and phrases. Continuous population, regression model, unbiasedness, measure of uncertainty, stratification.

Consider a population of infinitely many pairs $(y(x),\ x)\ ;\ x \geq 0,$ such that the joint distribution of $y(x),\ x \geq 0,$ is known only partially. For convenience let us assume that $y(x),\ x \geq 0,$ are defined on some probability space $(\Omega,\ \mathcal{A},\ \xi\ )$. The distribution of $X,$ whose observed values are $x,$ assumed to be continuous and known is given by

$$F(x) = \int\limits_0^x f(u)du \quad ; \quad x \geq 0.$$

Here Y is the study variable while X is the auxiliary variable.

Any continuous probability measure Q is called a sampling design. $Q(\mathbf{x})$ is the probability of drawing a sample such that the auxiliary variate value does not exceed x_i in the ith draw, $1 \leq i \leq n$. Let $q(\mathbf{x}) = \frac{\partial^n Q(\mathbf{x})}{\partial x_1 \partial x_2 \cdots \partial x_n}$. Then $q(\mathbf{x})$ can be expressed as $q(\mathbf{x}) = p(\mathbf{x}) f(\mathbf{x}),$ where $f(\mathbf{x}) = \prod\limits_{i=1}^{n} f(x_i)$. We shall call $p(\mathbf{x}),$ the design function associated with the sampling design $Q(\mathbf{x})$.

Consider a sampling design $Q(\mathbf{x})$ and the corresponding design function $p(\mathbf{x})$. Having drawn and observed n units, the data is recorded as $(y(x_i),\ x_i),\ i = 1,\ 2,\ \cdots,\ n\ ;$ or equivalently as $(y(\mathbf{x}),\ \mathbf{x}),$ where $\mathbf{x} = (x_1,\ x_2,\ \cdots,\ x_n)$.

A function t of the observed data $(y(\mathbf{x}),\ \mathbf{x})$ is called an estimator of the population mean $m_{\mathrm{Y}},$ whereas $(\ p,\ t),$ an estimator together with a design function p is called a strategy. The problem under consideration is to get an efficient strategy $(\ p,\ t)$ to estimate the population mean for the variate $Y,$ namely

$$m_{\mathrm{Y}} = E_f(y) = \int\limits_0^\infty y(x) f(x) dx\ .$$

Here we consider a specific superpopulation model, namely the regression model, induced by the probability space $(\Omega,\ \mathcal{A},\ \xi\ ),$ given by

$$Y(x) = \beta x + Z(x), \quad x \geq 0$$

where for every fixed $x \geq 0$

$$E_\xi(Z(x)) = 0,\ E_\xi(Z^2(x)) = \sigma^2 x^g \qquad \ldots(1.1)$$

and for every $x \neq x'\ ;\ x, x' \geq 0$

$$E_\xi(Z(x)Z(x')) = 0$$

where $\sigma^2 > 0$ and β are unknown and $g \in [0,\ 2]$ may be known or unknown.

We assume that $Y(x)$ is square integrable with respect to the product probability $(F \times \xi)$. To judge the performance of a strategy $(p,\ t)$ we use the following measures of uncertainty

$$M_1(p,\ t) = E_\xi E_p(t - m_{\text{Y}})^2 \qquad \ldots(1.2)$$

$$M_2(p,\ t) = E_\xi E_p(t - \mu_{\text{Y}})^2 \qquad \ldots(1.3)$$

where $\mu_{\text{Y}} = E_\xi(m_{\text{Y}}) = E_\xi \int_0^\infty y(x)f(x)dx = \beta E_f(X) = \beta\mu$ (say).

A strategy $(p,\ t)$ is said to be $p-$unbiased (design-unbiased) for m_{Y} if

$$E_p(t) = \int_{I\!R_n^+} t(y(\mathbf{x}),\ \mathbf{x})p(\mathbf{x})f(\mathbf{x})d\mathbf{x} = \int_0^\infty y(x)f(x)dx = m_{\text{Y}}$$

for every real valued F-integrable function $y(x)$. This defines the operator E_p.

A strategy $(p,\ t)$ is said to be $\xi-$unbiased (model-unbiased) for m_{Y} if

$$E_\xi[(t(y(\mathbf{x}),\ \mathbf{x}) - m_{\text{Y}}] = 0 \quad a.e.[Q].$$

A strategy $(p,\ t)$ is said to be $p\xi-$unbiased (model-design-unbiased) for m_{Y} if

$$E_pE_\xi[(t(y(\mathbf{x}),\ \mathbf{x})] - E_\xi[m_{\text{Y}}] = 0\ .$$

In this note we assume that the auxiliary variable X has Gamma distribution with parameter α. Clearly $\mu = E_f(X) = \alpha$. We also use the convention that unless otherwise specified $\sum$ would denote $\sum_{i=1}^{n}$.

We will consider strategies $(srs,\ \overline{y}\ )$, $(srs,\ t_{\text{R}})$, $(p_{\text{M}},\ t_{\text{R}})$, $(ppx,\ t_{\text{HT}})$, $(p_g,\ t_g)$, $(p_{\text{RHC}},\ t_{\text{RHC}})$ defined as follows :

a) sampling designs :

srs : simple random sampling for which $p(\mathbf{x}) \equiv 1$.

ppx^a : sampling design for which $p(\mathbf{x}) \propto \prod_{i=1}^{n} x_i^a$.

p_{M} : continuous analogue of the Midzuno-Sen sampling design for which $p(\mathbf{x}) = \frac{1}{n\mu}\sum x_i$, where $\mu = E_f(X) = \int_0^\infty xf(x)dx$.

p_g : sampling design with $p(\mathbf{x}) = k\prod_{i=1}^{n} x_i^{g-1}\sum x_i^{2-g}$, where $k = \frac{1}{n\mu}\left[\frac{\Gamma(\alpha)}{\Gamma(\alpha+g-1)}\right]^{n-1}$; $(\mu = \alpha)$.

p_{RHC} : continuous analogue of the Rao-Hartley-Cochran sampling design, vide (Padmawar (1996)).

b) estimators :

$\overline{y}$: sample mean $\frac{1}{n}\sum y(x_i)$.

$t_{\rm R}$: ratio estimator $\mu \frac{\sum y(x_i)}{\sum x_i}$.

$t_{\rm HT}$: Horvitz-Thompson estimator given by $\sum \frac{y(x_i) f(x_i)}{\pi(x_i)}$, based on $q(\mathbf{x})$, where $\pi(x_i) = \sum\limits_{j=1}^{n} q_j(x_i)$, $\pi(x) > 0$ for $x > 0$, and $q_i(x_i) = \int\limits_{\mathbb{R}_{n-1}^+} q(\mathbf{x}) \prod\limits_{j \neq i}^{n} dx_j$, $1 \leq i \leq n$, vide (Cordy (1993)).

t_g : estimator given by $\frac{\mu}{\sum x_i^{2-g}} \sum x_i^{1-g} y(x_i)$, $g \in [0,\ 2]$.

$t_{\rm RHC}$: continuous analogue of the Rao-Hartley-Cochran estimator, vide (Padmawar (1996)).

2. Comparison of Strategies

Comparison of sampling strategies, in the absence of an optimal one, under a superpopulation model with respect to an uncertainty measure has been one of the major problems of interest to survey statisticians. In this section we take up this problem in the continuous set-up for the strategies listed in the previous section. We first establish some interesting properties of these strategies.

It is known that for estimating the population mean $(srs,\ t_{\rm R})$ is not $p-$unbiased whereas $(srs,\ \overline{y})$, $(ppx,\ t_{\rm HT})$ and $(p_{\rm RHC},\ t_{\rm RHC})$ are $p-$unbiased, vide (Särndal (1980), Padmawar (1982, 1996), Cordy (1993)). It is easy to prove the following

Theorem 2.1. *The strategies* $(p_{\rm M},\ t_{\rm R})$ *and* $(p_g,\ t_g)$ *are* $p-$*unbiased for estimating the population mean* $m_{\rm Y}$.

Proof.

$$\begin{aligned}
E_P(\,p_{\rm M},\ t_{\rm R}) &= \int\limits_{\mathbb{R}_n^+} \mu \frac{\sum y(x_i)}{\sum x_i} \frac{\sum x_i}{n\mu} f(\mathbf{x}) d\mathbf{x} \\
&= \int\limits_{\mathbb{R}_n^+} \frac{1}{n} \sum y(x_i) f(\mathbf{x}) d\mathbf{x} \\
&= \frac{1}{n} \sum_{i=1}^{n} \int\limits_{0}^{\infty} y(x_i) \left[\prod_{j \neq i}^{n} \int\limits_{0}^{\infty} f(x_j) dx_j \right] f(x_i) dx_i \\
&= \int\limits_{0}^{\infty} y(x) f(x) dx \\
&= m_{\rm Y} \,.
\end{aligned}$$

Thus the strategy $(p_{\mathrm{M}},\ t_{\mathrm{R}})$ is $p-$unbiased for the population mean. Similarly,

$$\begin{aligned} E_p(\,p_g,\ t_g) &= \int\limits_{I\!R_n^+} \frac{\mu}{\sum x_i^{2-g}} \sum x_i^{1-g} y(x_i) \frac{1}{n\mu} \left[\frac{\Gamma(\alpha)}{\Gamma(\alpha+g-1)}\right]^{n-1} \times \\ & \quad \prod_{i=1}^{n} x_i^{g-1} \sum x_i^{2-g} f(\mathbf{x}) d\mathbf{x} \\ &= \frac{1}{n}\left[\frac{\Gamma(\alpha)}{\Gamma(\alpha+g-1)}\right]^{n-1} \times \\ & \quad \sum_{i=1}^{n} \int\limits_0^\infty y(x_i) \left[\prod_{j\neq 1}^{n} \int\limits_0^\infty x_j^{g-1} f(x_j) dx_j\right] f(x_i) dx_i \\ &= \int\limits_0^\infty y(x) f(x) dx \\ &= m_{\mathrm{Y}}\ . \end{aligned}$$

Hence $(\,p_g,\ t_g)$ is also $p-$unbiased.
We now prove an important property of the estimator t_g.

THEOREM 2.2. *For any design p the estimator t_g is the best linear $\xi-$unbiased estimator for m_{Y} in the sense of minimum $M_1(\,p,\ t)$.*

PROOF. A linear estimator is of the type

$$t = t(y(\mathbf{x}),\ \mathbf{x}) = \sum_{i=1}^{n} a_i(\mathbf{x}) y(x_i) \qquad \ldots (2.1)$$

where $a_1,\ a_2,\ \cdots,\ a_n$ are known measurable functions.
The condition of $\xi-$unbiasedness under the model (1.1) for the estimator (2.1) is

$$\sum_{i=1}^{n} a_i(\mathbf{x}) x_i = \mu \quad a.e.[Q]\ . \qquad \ldots (2.2)$$

For a given design p we want to minimize $M_1(\,p,\ t)$ subject to the condition (2.2). Note that

$$\begin{aligned} M_1(\,p,\ t) &= E_\xi E_p (t - m_{\mathrm{Y}})^2 \\ &= E_\xi E_f p(\mathbf{x}) (t - m_{\mathrm{Y}})^2 \\ &= E_f p(\mathbf{x}) E_\xi (t - m_{\mathrm{Y}})^2\ . \end{aligned}$$

Hence it suffices to minimize $E_\xi(t-m_{\mathrm{Y}})^2$ subject to (2.2). Now

$$E_\xi(t - m_{\mathrm{Y}})^2 = E_\xi t^2 + E_\xi m_{\mathrm{Y}}^2 - 2E_\xi t m_{\mathrm{Y}}$$

and

$$\begin{aligned} E_\xi t m_{\text{Y}} &= E_\xi \sum a_i(\mathbf{x}) Y(x_i) E_f Y(x) \\ &= \sum_{i=1}^{n} a_i(\mathbf{x}) E_\xi E_f Y(x_i) Y(x) \\ &= \sum_{i=1}^{n} a_i(\mathbf{x}) E_f E_\xi Y(x_i) Y(x)\ . \end{aligned}$$

But

$$E_\xi Y(x_i) Y(x) = \beta^2 x_i x \ \text{ a.e. } \ [\text{F}] \ \ \forall\ i = 1,\ 2,\ \cdots,\ n\ .$$

Hence

$$\begin{aligned} E_\xi t m_{\text{Y}} &= \beta^2 \sum_{i=1}^{n} a_i(\mathbf{x}) x_i E_f X \\ &= \beta^2 \mu^2 \quad \text{using (2.2)}\ . \end{aligned}$$

Therefore it suffices to minimize $E_\xi t^2$ subject to (2.2), i.e., to

$$\begin{aligned} &\textit{minimize} && \sum_{i=1}^{n} a_i^2(\mathbf{x}) x_i^g \\ &\textit{subject to} && \sum_{i=1}^{n} a_i(\mathbf{x}) x_i = \mu. \end{aligned}$$

This immediately admits the following solution

$$a_i(\mathbf{x}) = \frac{\mu x_i^{1-g}}{\sum x_i^{2-g}}\ , \quad 1 \le i \le n.$$

Hence, $t_g = \frac{\mu}{\sum x_i^{2-g}} \sum x_i^{1-g} y(x_i)$ is the best linear $\xi-$unbiased estimator.

REMARK 2.1. Although t_g possesses the above optimal property for any p we consider the strategy $(p_g,\ t_g)$ as it is $p-$unbiased and hence even if the model breaks down it remains at least $p\xi-$unbiased. Moreover, in this note we would like to compare the performances of various $p-$unbiased strategies.

REMARK 2.2. It is interesting to note that for $g = 1$ the strategy $(p_g,\ t_g)$ coincides with the strategy $(p_{\text{M}},\ t_{\text{R}})$ and for $g = 2$ it coincides with the strategy $(ppx,\ t_{\text{HT}})$.

Särndal (1980) studied the strategy $(srs,\ t_{\text{R}})$. He observed that the strategy $(srs,\ t_{\text{R}})$ is not $p-$unbiased and if the model (1.1) breaks down then it is not even $p\xi-$unbiased. The strategy $(p_{\text{M}},\ t_{\text{R}})$ is $\xi-$unbiased and, since we have just proved that it is $p-$unbiased, it would remain $p\xi-$unbiased even if the model (1.1) breaks down. We prove here that the strategy $(p_{\text{M}},\ t_{\text{R}})$, apart from possessing the above advantage over the strategy $(srs,\ t_{\text{R}})$, is, in fact, superior to $(srs,\ t_{\text{R}})$. Let us, however, first prove the following lemma :

Lemma 2.1. *For* $\delta \in I\!R$ *and* $n\alpha + g - \delta > 0$,

$$\mathbf{J} = \int\limits_{I\!R_n^+} \frac{x_1^g}{[\sum x_i]^\delta} e^{-\sum x_i} \prod_{i=1}^n x_i^{\alpha-1} dx_i = \frac{\Gamma(n\alpha+g-\delta)\Gamma(g+\alpha)(\Gamma(\alpha))^{n-1}}{\Gamma(n\alpha+g)} .$$

Proof. Consider the following transformation
$x_1 = u_1(1-u_2)$, $x_2 = u_1u_2(1-u_3)$, $x_3 = u_1u_2u_3(1-u_4)$, $\cdots$,
$x_{n-1} = u_1u_2\cdots u_{n-1}(1-u_n)$ and $x_n = u_1u_2\cdots u_{n-1}u_n$.

For this transformation, $0 \le u_1 < \infty$ and $0 \le u_i \le 1 \quad \forall\, i = 2,\ 3,\ \cdots,\ n$. The Jacobian of the transformation is $\prod\limits_{1=1}^n u_i^{n-i}$. Thus

$$\begin{aligned}\mathbf{J} &= \int_0^\infty \int_0^1 \cdots \int_0^1 \frac{[u_1(1-u_2)]^g}{u_1^\delta} e^{-u_1} \Big[u_1(1-u_2)u_1u_2(1-u_3)\cdots \\ &\quad (u_1u_2\cdots u_n)\Big]^{\alpha-1} \prod_{i=1}^n u_i^{n-i} du_i \\ &= \int_0^\infty e^{-u_1} u_1^{n\alpha+g-\delta-1} du_1 \int_0^1 u_2^{(n-1)\alpha-1}(1-u_2)^{g+\alpha-1} du_2 \\ &\quad \times \int_0^1 u_3^{(n-2)\alpha-1}(1-u_3)^{\alpha-1} du_3 \cdots \int_0^1 u_{n-1}^{2\alpha-1}(1-u_{n-1})^{\alpha-1} du_{n-1} \\ &\quad \times \int_0^1 u_n^{\alpha-1}(1-u_n)^{\alpha-1} du_n .\end{aligned}$$

Thus
$$\begin{aligned}\mathbf{J} &= \Gamma(n\alpha+g-\delta) \frac{\Gamma\{(n-1)\alpha\}\ \Gamma(g+\alpha)}{\Gamma(n\alpha+g)} \frac{\Gamma\{(n-2)\alpha\}\ \Gamma(\alpha)}{\Gamma\{(n-1)\alpha\}} \cdots \\ &\quad \cdots \frac{\Gamma(2\alpha)\Gamma(\alpha)}{\Gamma(3\alpha)} \frac{\Gamma(\alpha)\Gamma(\alpha)}{\Gamma(2\alpha)} \text{ , or,} \\ \mathbf{J} &= \frac{\Gamma(n\alpha+g-\delta)\Gamma(g+\alpha)}{\Gamma(n\alpha+g)} [\Gamma(\alpha)]^{n-1} .\end{aligned}$$

This completes the proof of the lemma that would be used to prove the following

Theorem 2.3. *Under the model (1.1) the strategy* $(p_{\rm M},\ t_{\rm R})$ *is superior to the strategy* $(srs,\ t_{\rm R})$ *with respect to the measure of uncertainty* $M_2(p,\ t)$ *if* $n\alpha + g - 2 > 0$.

Proof. For design function p,

$$\begin{aligned} M_2(p,\ t_{\rm R}) &= E_\xi E_p(t_{\rm R} - \beta\mu)^2 \\ &= E_p E_\xi(t_{\rm R} - \beta\mu)^2 \\ &= E_p V_\xi(t_{\rm R}) \\ &= \sigma^2\mu^2 E_p \frac{\sum x_i^g}{[\sum x_i]^2} .\end{aligned}$$

For $p(\mathbf{x}) \equiv 1$, using Lemma 2.1 with $\delta = 2$, we get

$$\begin{aligned} E_p \frac{\sum x_i^g}{[\sum x_i]^2} &= \frac{n}{[\Gamma(\alpha)]^n} \int\limits_{I\!R_n^+} \frac{x_1^g}{[\sum x_i]^2} e^{-\sum x_i} \prod_{i=1}^n x_i^{\alpha-1} dx_i \\ &= \frac{n}{[\Gamma(\alpha)]^n} \frac{\Gamma(n\alpha+g-2)\Gamma(g+\alpha)}{\Gamma(n\alpha+g)} [\Gamma(\alpha)]^{n-1} \ . \end{aligned}$$

Thus $$M_2(srs,\ t_{\rm R}) = \sigma^2\mu^2 \frac{n\Gamma(g+\alpha)/\Gamma(\alpha)}{(g+n\alpha-1)(g+n\alpha-2)} \ .$$

Similarly, for $p(\mathbf{x}) = \frac{\sum x_i}{n\mu}$, using Lemma 2.1 with $\delta = 1$, we get

$$M_2(\,p_{\rm M},\ t_{\rm R}) = \sigma^2\mu^2 \frac{\Gamma(g+\alpha)/\Gamma(\alpha+1)}{(g+n\alpha-1)} \ . \qquad \ldots(2.3)$$

Therefore, $$\frac{M_2(\,p_{\rm M},\ t_{\rm R})}{M_2(srs,\ t_{\rm R})} = \frac{n\alpha+g-2}{n\alpha} \ .$$

Since $g \in [0,\ 2]$, the strategy $(\,p_{\rm M},\ t_{\rm R})$ is always superior to $(srs,\ t_{\rm R})$.
In our next theorem we compare the strategies $(\,p_{\rm M},\ t_{\rm R})$ and $(\,ppx,\ t_{\rm HT})$.

Theorem 2.4. *Under the model (1.1) for* $n \geq 2$ *and* $g + n\alpha - 1 > 0$ *, we have,*

$$M_2(\,p_{\rm M},\ t_{\rm R}) \ \underset{¿}{\overset{¡}{=}} \ M_2(\,ppx,\ t_{\rm HT}) \quad \textit{according as} \quad g \ \underset{¿}{\overset{¡}{=}} \ 1.$$

Proof. We know from Särndal (1980) that

$$M_2(\,ppx,\ t_{\rm HT}) = \frac{\sigma^2\mu^2}{n} \frac{\Gamma(\alpha+g-1)}{\Gamma(\alpha+1)} \ . \qquad \ldots(2.4)$$

Using (2.3) and (2.4) we get,

$$\frac{M_2(\,p_{\rm M},\ t_{\rm R})}{M_2(\,ppx,\ t_{\rm HT})} = \frac{n(\alpha+g-1)}{(g+n\alpha-1)} = 1 + \frac{(n-1)(g-1)}{(g+n\alpha-1)} \ .$$

Clearly for $n \geq 2$ and $g + n\alpha - 1 > 0$, we have,

$$M_2(\,p_{\rm M},\ t_{\rm R}) \ \underset{¿}{\overset{¡}{=}} \ M_2(\,ppx,\ t_{\rm HT}) \quad \text{according as} \quad \text{g} \ \underset{¿}{\overset{¡}{=}} \ 1.$$

Hence the theorem.

It is interesting to note that for any strategy $(\,p,\ t)$ that is $p-$unbiased as well as $\xi-$unbiased, the measures of uncertainty $M_1(\,p,\ t)$ and $M_2(\,p,\ t)$ differ by a quantity that is independent of $(\,p,\ t)$. Since both the strategies in the above theorem are $p-$unbiased as well as $\xi-$unbiased we immediately have the following

THEOREM 2.5. *Under the model (1.1) for* $n \geq 2$ *and* $g + n\alpha - 1 > 0$, *we have,*

$$M_1(\, p_{\mathrm{M}},\ t_{\mathrm{R}}) \ \overset{\text{¡}}{\underset{\text{¿}}{=}}\ M_1(\, ppx,\ t_{\mathrm{HT}}) \quad \textit{according as} \quad g \ \overset{\text{¡}}{\underset{\text{¿}}{=}}\ 1.$$

REMARK 2.3. It is clear from the above results that if the parameter g of the model (1.1) is not known and if the sampler has to choose between the above two strategies then there is a clear demarcation of the range of the parameter g. If there are reasons to believe that the parameter g is less than unity then the sampler should go for the strategy $(\, p_{\mathrm{M}},\ t_{\mathrm{R}})$. On the other hand, if the sampler speculates g to be greater than unity then the strategy $(\, ppx,\ t_{\mathrm{HT}})$ is to be preferred.

REMARK 2.4. Theorem 2.5 agrees with the result due to Rao (1967) in which the same two strategies are compared in the finite set-up.
We now compare the strategies $(\, ppx,\ t_{\mathrm{HT}})$ and $(\, p_g,\ t_g)$.

THEOREM 2.6. *Under the model (1.1) the strategies* $(\, ppx,\ t_{\mathrm{HT}})$ *and* $(\, p_g,\ t_g)$ *are equally efficient with respect to either measure of uncertainty.*

PROOF. Since both the strategies $(ppx,\ t_{\mathrm{HT}})$ and $(p_g,\ t_g)$ are $p-$unbiased as well as $\xi-$unbiased it is enough to consider the measure of uncertainty M_2 . Let us first evaluate $M_2(\, p_g,\ t_g)$.

$$\begin{aligned}
M_2(\, p_g,\ t_g) &= E_{p_g} E_\xi t_g^2 - \beta^2\mu^2 \\
&= E_{p_g} \frac{\sigma^2\mu^2}{[\sum x_i^{2-g}]} \\
&= \int\limits_{\mathbb{R}_n^+} \frac{\sigma^2\mu^2}{\sum x_i^{2-g}} \frac{1}{n\alpha} \left[\frac{\Gamma(\alpha)}{\Gamma(\alpha+g-1)}\right]^{n-1} \prod_{i=1}^n x_i^{g-1} \sum x_i^{2-g} f(\mathbf{x}) d\mathbf{x} \\
&= \frac{\sigma^2\mu^2}{n\alpha} \left[\frac{\Gamma(\alpha)}{\Gamma(\alpha+g-1)}\right]^{n-1} \prod_{i=1}^n \int_0^\infty x_i^{g-1} f(x_i) dx_i \\
&= \frac{\sigma^2\mu^2}{n\alpha} \left[\frac{\Gamma(\alpha)}{\Gamma(\alpha+g-1)}\right]^{n-1} \left[\frac{\Gamma(\alpha+g-1)}{\Gamma(\alpha)}\right]^n
\end{aligned}$$

Thus
$$M_2(\, p_g,\ t_g) = \frac{\sigma^2\mu^2}{n} \frac{\Gamma(\alpha+g-1)}{\Gamma(\alpha+1)} \qquad \ldots(2.5)$$

which is same as $M_2(ppx,\ t_{\mathrm{HT}})$. Thus the strategies $(ppx,\ t_{\mathrm{HT}})$ and $(p_g,\ t_g)$ are equally efficient.
It is easy to prove the following

COROLLARY 2.1. *Under the model (1.1) with* $g = 1$ *the strategies* $(p_{\mathrm{M}},\ t_{\mathrm{R}})$, $(ppx,\ t_{\mathrm{HT}})$ *and* $(p_g,\ t_g)$ *are equally efficient with respect to either measure of uncertainty.*

Padmawar (1996) defined Rao-Hartley-Cochran strategy, $(p_{\text{RHC}},\ t_{\text{RHC}})$, in the continuous set-up. It is proved there that, this strategy is $p-$unbiased as well as $\xi-$unbiased and that in the limiting sense, the value of $M_2(p_{\text{RHC}},\ t_{\text{RHC}})$ is given by

$$M_2(p_{\text{RHC}},\ t_{\text{RHC}}) = \frac{\sigma^2\mu^2}{n}\frac{\Gamma(\alpha+g-1)}{\Gamma(\alpha+1)}\ . \qquad \ldots(2.6)$$

In view of (2.6) and Theorem 2.6 we conclude this section with the following

THEOREM 2.7. *Under the model (1.1) the strategy* $(p_{\text{RHC}},\ t_{\text{RHC}})$ *is as efficient, in the limiting sense, as the strategies* $(ppx,\ t_{\text{HT}})$ *and* $(p_g,\ t_g)$ *with respect to either measure of uncertainty.*

REMARK 2.5. It was, however, observed in Padmawar (1996) that from the practical point of view the strategy $(ppx,\ t_{\text{HT}})$ is better than the other two competing strategies $(p_g,\ t_g)$ and $(p_{\text{RHC}},\ t_{\text{RHC}})$ as $(p_g,\ t_g)$ depends on the parameter g of the model (1.1) that may not always be known and $(p_{\text{RHC}},\ t_{\text{RHC}})$ is equally efficient only in the limiting sense.

In the next section we compare some more strategies in the stratified set-up.

3. Stratified Sampling

In this section we consider the stratified sampling set-up having L strata. Let $0 = z_0 < z_1 < z_2 < \cdots < z_L = \infty$ be the given stratification points. A unit is said to belong to the hth stratum if its $x-$value belongs to $[z_{h-1},\ z_h)$, $1 \leq h \leq L$. For the stratified sampling we have to modify our basic set-up suitably. Define, for the hth stratum, $f_h(x)$, the analogue of $f(x)$ on $I\!R^+$, as

$$\begin{aligned} f_h(x) &= \tfrac{f(x)}{W_h} \qquad \text{if } x \in [z_{h-1},\ z_h) \\ &= 0 \qquad \text{otherwise} \end{aligned}$$

where $W_h = F(z_h) - F(z_{h-1}) = \int\limits_{z_{h-1}}^{z_h} f(x)dx$.

Let n_h be the number of units to be sampled from the hth stratum, $1 \leq h \leq L$, then the total sample size n is given by $n = \sum\limits_{h=1}^{L} n_h$.

We can now think of a design function $p_h(\mathbf{x}_h)$ for the hth stratum where $\mathbf{x}_h = (x_{h1},\ x_{h2},\ \cdots,\ x_{hn_h})$ now is a vector with n_h coordinates, i.e., x_{hi} denotes the ith unit from the hth stratum , $1 \leq i \leq n_h$, $1 \leq h \leq L$. The sampling design for the hth stratum, $1 \leq h \leq L$, is defined as

$$q_h(\mathbf{x}_h) = p_h(\mathbf{x}_h)f_h(\mathbf{x}_h) \quad \text{where} \quad f_h(\mathbf{x}_h) = \prod_{i=1}^{n_h} f_h(x_{hi})\ .$$

The overall stratified sampling design is now given by $\prod_{h=1}^{L} q_h(\mathbf{x}_h)$.

Särndal (1980) considered the strategy $(srst,\ \overline{y}_{st})$ that consists of $srst,$ the stratified simple random sampling and the estimator $\overline{y}_{st}$, given by

$$\overline{y}_{st} = \sum_{h=1}^{L} W_h \overline{y}_h \qquad \text{where} \qquad \overline{y}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} y(x_{hi}), \quad 1 \le h \le L \ .$$

Let us consider the strategy $(ppxst,\ t^*_{\text{HT}})$ that consists of $ppxst,$ the stratified ppx sampling and the estimator t^*_{HT} given by

$$t^*_{\text{HT}} = \sum_{h=1}^{L} W_h \frac{\mu_h}{n_h} \sum_{i=1}^{n_h} \frac{y(x_{hi})}{x_{hi}} \qquad \text{where} \qquad \mu_h = \frac{1}{W_h} \int_{z_{h-1}}^{z_h} x f(x) dx \ .$$

It is easy to see that the strategy $(\,ppxst,\ t^*_{\text{HT}})$ is $p-$unbiased as well as $\xi-$unbiased. For this strategy let us evaluate M_2 .

$$\begin{aligned} M_2(\,ppxst,\ t^*_{\text{HT}}) &= E_p V_\xi t^*_{\text{HT}} \\ &= E_p \sum_{h=1}^{L} W_h^2 \frac{\mu_h^2}{n_h^2} \sum_{i=1}^{n_h} \sigma^2 x_{hi}^{g-2} \\ &= \sigma^2 \sum_{h=1}^{L} W_h \frac{\mu_h}{n_h} \int_{z_{h-1}}^{z_h} x^{g-1} f(x) dx \ . \end{aligned}$$

For the allocation $$n_h = \frac{nW_h\mu_h}{\mu}, \quad 1 \le h \le L, \qquad \ldots(3.1)$$

$$M_2(\,ppxst,\ t^*_{\text{HT}}) = \frac{\sigma^2\mu^2}{n} \frac{\Gamma(\alpha+g-1)}{\Gamma(\alpha+1)} \ . \qquad \ldots(3.2)$$

For the optimal allocation

$$n_h \propto \left(W_h \mu_h \int_{z_{h-1}}^{z_h} x^{g-1} f(x) dx \right)^{\frac{1}{2}} , \quad 1 \le h \le L \ , \qquad \ldots(3.3)$$

$$M_2(\,ppxst,\ t^*_{\text{HT}}) = \frac{\sigma^2}{n} \left[\sum_{h=1}^{L} \left\{ W_h \mu_h \int_{z_{h-1}}^{z_h} x^{g-1} f(x) dx \right\}^{1/2} \right]^2 . \qquad \ldots(3.4)$$

We now state the following

Theorem 3.1. *Under the model (1.1) we have, for the allocation (3.1)*

$$M_2(\,ppxst,\ t^*_{\text{HT}}) = M_2(\,ppx,\ t_{\text{HT}})$$

and for the optimal allocation (3.3)

$$M_2(\,ppxst,\ t^*_{\text{HT}}) \le M_2(\,ppx,\ t_{\text{HT}}),$$

and the equality holds if and only if $g = 2$.

Remark 3.1. Thus for any allocation that is better than the allocation (3.1) there would be *gain due to stratification,* in the sense that the stratified strategy $(\,ppxst,\ t^*_{\text{HT}})$ would perform better than its unstratified counterpart $(\,ppx,\ t_{\text{HT}})$. The above result for the optimal allocation agrees with the result due to Rao (1968) in which the same two strategies are compared in the finite set-up. It may be mentioned here that in the finite stratified set-up the problem of comparing different allocations in terms M_1 was first considered by Hanurav (1965), followed by Rao (1968, 1977).

We now proceed to comment on a result due to Särndal (1980). Let us first evaluate

$$\begin{aligned} M_2(srst,\ \overline{y}_{st}) &= E_p[V_\xi(\overline{y}_{st}) + E_\xi(\overline{y}_{st})^2] - \beta^2\mu^2 \\ &= E_p\left(\sigma^2 \sum_{h=1}^{L} \tfrac{W_h^2}{n_h^2} \sum_{i=1}^{n_h} x^g_{hi} + \beta^2\left[\sum_{h=1}^{L} W_h\overline{x}_h\right]^2\right) - \beta^2\mu^2 \end{aligned}$$

where $\overline{x}_h = \frac{1}{n_h}\sum\limits_{i=1}^{n_h} x_{hi}$

$$\begin{aligned} &= \sigma^2 \sum_{h=1}^{L} \tfrac{W_h}{n_h} \int_{z_{h-1}}^{z_h} x^g f(x)dx \\ &\quad + \beta^2 \sum_{h=1}^{L} \tfrac{1}{n_h}\left[W_h \int_{z_{h-1}}^{z_h} x^2 f(x)dx - \left(\int_{z_{h-1}}^{z_h} xf(x)dfx\right)^2\right]. \end{aligned} \quad \ldots(3.5)$$

For the proportional allocation, $n_h = nW_h,\ \ 1 \le h \le L,$ we have

$$\begin{aligned} M_2(srst,\ \overline{y}_{st}) &= \tfrac{\sigma^2\mu^2}{n}\tfrac{\Gamma(\alpha+g)}{\alpha\Gamma(\alpha+1)} \\ &\quad + \beta^2 \sum_{h=1}^{L} \tfrac{1}{nW_h}\left[W_h \int_{z_{h-1}}^{z_h} x^2 f(x)dx - \left(\int_{z_{h-1}}^{z_h} xf(x)dx\right)^2\right]. \end{aligned} \quad \ldots(3.6)$$

Särndal (1980) stated that for the proportionally allocated stratified random sample, $(n_h = nW_h)$, and under 'maximum benefit from stratification' (many strata with optimally located boundaries),

$$M_2(srst,\ \overline{y}_{st}) = \frac{\sigma^2\mu^2}{n}\frac{\Gamma(\alpha+g)}{\alpha\Gamma(\alpha+1)}\ . \quad \ldots(3.7)$$

However, it follows easily from the Cauchy-Schwartz inequality that the coefficient of β^2 in (3.5) and that in (3.6) are positive. Therefore even under 'maximum benefit from stratification', the right hand side expression in (3.7) can only be a lower bound for $M_2(srst,\ \overline{y}_{st})$. The comparison between the strategies $(srs,\ t_{\mathrm{R}})$ and $(srst,\ \overline{y}_{st})$ carried out by Särndal (1980) is valid for large values of n and L. However, for a given sample size n, we cannot arbitrarily increase the total number of strata, as it is necessary to sample at least two units from each stratum. Further, in (3.6), the value of β^2 may be large as compared to that of σ^2, both of which are unknown parameters of the model (1.1). In the following theorem we, therefore, carry out some exact comparisons involving the strategy $(srst,\ \overline{y}_{st})$.

Theorem 3.2. *Under the model (1.1) with* $g \geq 1$ *for the proportional allocation the strategy* $(srst,\ \overline{y}_{st})$ *is inferior to the strategies* $(p_{\mathrm{M}},\ t_{\mathrm{R}})$, $(ppx,\ t_{\mathrm{HT}})$ *and* $(p_g,\ t_g)$.

Proof. Using (2.3) and (3.6) we get

$$\frac{M_2(srst,\ \overline{y}_{st})}{M_2(p_{\mathrm{M}},\ t_{\mathrm{R}})} \geq 1 + \frac{g-1}{n\alpha}\ .$$

Thus for $g \geq 1$, $(p_{\mathrm{M}},\ t_{\mathrm{R}})$ and hence $(ppx,\ t_{\mathrm{HT}})$ and $(p_g,\ t_g)$ are all more efficient than the strategy $(srst,\ \overline{y}_{st})$.

Remark 3.2. If the parameter g of the model (1.1) is greater than unity, then, even the stratified version $(srst,\ \overline{y}_{st})$ of the strategy $(srs,\ \overline{y})$ loses out to the strategies that depend on x.

References

Cassel, C.M. and Särndal, C.E. (1972). A model for studying robustness of estimators and informativeness of labels in sampling with varying probabilities. *J.R. Statist. Soc.* B, **34**, 279-289.

– – –– (1974). Evaluation of some sampling strategies for finite populations using a continuous variable framework. *Commun. Statist.,* **3**, 373-390.

Cordy, C.B. (1993) An extension of the Horvitz-Thompson theorem to point sampling from a continuous universe. *Statistics and Probability Letters.*, **18**, 353-362.

Hanurav, T.V. (1965). Optimum sampling strategies and some related problems. *Ph.D. Thesis*, Indian Statistical Institute.

Padmawar, V.R. (1981). A note on the comparison of certain sampling strategies *J.R. Statist. Soc.* B, **43**, 321-326.

– – –– (1982). Optimal strategies under superpopulation models. *Ph.D. Thesis*, Indian Statistical Institute.

– – –– (1984). Two existence theorems in survey sampling of continuous populations. *Sankhyā* B, **46**, 217-227.

–––– (1996). Rao-Hartley-Cochran strategy in survey sampling of continuous populations. *Sankhyā* B, **57**, 90-104.

RAO, T.J. (1967). On the choice of a strategy for the ratio method of estimation. *J.R. Statist. Soc.* B, **29**, 392-397.

– – –– (1968). On the allocation of sample size in stratified sampling. *Ann. Inst. Statist. Math.*, **20**, 159-166.

– – –– (1977). Optimum allocation of sample size and prior distributions : a review. *Inter. Statist. Review.*, **45**, 173-179.

SÄRNDAL, C.E. (1980). A method for assessing efficiency and bias of estimation strategies in survey sampling. *South African Statist. J.*, **14**, 17-30.

V.R. PADMAWAR
STAT−MATH DIVISION
INDIAN STATISTICAL INSTITUTE
8TH MILE, MYSORE ROAD
BANGALORE, 560 059
INDIA.
e-mail : vrp@isibang.ac.in