Sankhyā : The Indian Journal of Statistics
1984, Volume 46, Series B, Pt. 2, pp. 217-227.

TWO EXISTENCE THEOREMS IN SURVEY SAMPLING OF CONTINUOUS POPULATIONS

By V. R. PADMAWAR
Indian Statistical Institute

SUMMARY. Interpreting the traditional survey sampling set-up in the continuous infinite population framework some optimality results w.r.t. a measure of uncertainty, under the well-known regression model, are obtained.

1 Introduction

Consider a population of infinitely many pairs $(y(x), x)$, $x \geqslant 0$, such that the joint distribution of $y(x)$, $x \geqslant 0$, is not known completely. For convenience we assume that $y(x)$, $x \geqslant 0$, 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 specified by

$$F(x) = \int_0^x f(u)du, \ x \geqslant 0.$$

In the continuous set up, the label of a population unit is a continuous index λ, where for convenience $\lambda \in [0, 1]$. A more specific ordering is imposed on λ by identifying it with λ-th quantile of the X-distribution. Having drawn and observed n units the data are recorded as $(y(x_i), x_i)$; $i = 1, 2, \ldots, n$; or equivalently $(y_s(\boldsymbol{x}), \boldsymbol{x})$; where $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$. The problem under consideration is to estimate the population mean for the variate Y, namely

$$m_y = E_f(Y) = \int_0^\infty y(x)f(x)dx.$$

This, incidentally, defines the operator, E_f.

Let $\mathcal{B}$ be the Borel σ-algebra of $\mathcal{R}_n^+ = \{\boldsymbol{x} : x_i \geqslant 0, \ i = 1, 2, \ldots, n\}$. Any continuous probability measure Q on $\mathcal{B}$ is called a sampling design. $Q(\boldsymbol{x})$ is the probability of drawing a sample such that the auxiliary variate value in the i-th draw does not exceed x_i, $i = 1, 2, \ldots, n$. Let $q(\boldsymbol{x}) = \frac{dQ(\boldsymbol{x})}{d\boldsymbol{x}}$; then $q(\boldsymbol{x})$ can be expressed as $q(\boldsymbol{x}) = p(\boldsymbol{x})f(\boldsymbol{x})$, where $f(\boldsymbol{x}) = \prod_{i=1}^n f(x_i)$. $p(\boldsymbol{x})$ is called a design function giving rise to the sampling design $Q(\boldsymbol{x})$.

AMS (1970) *subject classification* : 62D05.

Key words : Regression model ; Design function ; Measure of uncertainty ; Banach space ; Dual space ; Fréchet derivative ; Lagrangian ; Integral equation ; Taylor's expansion.

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

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

where for every fixed $x \geqslant 0$

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

and for every fixed $x \neq x'$, $x, x' \geqslant 0$

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

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

Any function T of the observed data $(y(x), x)$ is called an estimator of m_Y, whereas (p, T) an estimator T together with a design function p is called a strategy.

A strategy (p, T) is said to be p-unbiased (design-unbiased) if

$$E_p(T) = \int_{\mathcal{R}_n^+} T(y(x), x)p(x)f(x)dx = \int_{\mathcal{R}^+} y(x)f(x)dx$$

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

A strategy (p, T) is said to be ξ-unbiased (model-unbiased) if

$$E_\xi(T(Y(x), x) - m_Y) = 0 \text{ a.e. } [Q].$$

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

$$M(p, T) = E_\xi E_p(T - m_Y)^2. \qquad \dots \quad (1.2)$$

In actually computing (1.2) we assume that the population conforms to the model (1.1) with $g \epsilon [0, 2]$ known.

In this paper we consider the following :

(a) For a given design the problem of obtaining a best p as well as ξ-unbiased linear estimator, under the model (1.1) with g known, w.r.t. the measure of uncertainty (1.2).

(b) For a given design the problem of obtaining a best p-unbiased linear-estimator, under the model (1.1) with g known and the ratio $\frac{\sigma^2}{\beta^2}$ also known, w.r.t. the measure of uncertainty (1.2).

2. Existence theorems

A linear estimator is of the form

$$T(y(\boldsymbol{x}), \boldsymbol{x}) = \sum_{i=1}^{n} a_i(\boldsymbol{x})y(x_i) \qquad \dots \quad (2.1)$$

where $a_i(\boldsymbol{x})$, $i = 1, 2, \dots, n$ are $\mathcal{B}$-measurable functions. The condition of ξ-unbiasedness for the linear estimator (2.1) is given by

$$\sum_{i=1}^{n} a_i(\boldsymbol{x})x_i = \mu \;\forall\; \boldsymbol{x} \in \mathcal{R}_n^+ \qquad \dots \quad (2.2)$$

where $\mu = E_f(X) = \int_0^\infty xf(x)dx$.

The condition of p-unbiasedness for the strategy (p, T) is given by

$$\phi(\boldsymbol{a}, x) = 1 \;\forall\; x > 0 \qquad \dots \quad (2.3)$$

where $\boldsymbol{a} = (a_1, a_2, \dots, a_n)$, $a_i = a_i(\boldsymbol{x})$,

$$\phi_i(\boldsymbol{a}, x_i) = \int_{\mathcal{R}_{n-1}^+} a_i(\boldsymbol{x})p(\boldsymbol{x}) \prod_{j \neq i}^{n} f(x_j)dx_j \qquad \dots \quad (2.4)$$

and

$$\phi(\boldsymbol{a}, x) = \sum_{i=1}^{n} \phi_i(\boldsymbol{a}, x).$$

Now for a "p as well as ξ-unbiased" linear strategy (p, T) the measure of uncertainty (1.2) takes a simpler form; namely

$$M(p, T) = \sigma^2 \int_{\mathcal{R}_n^+} \sum_{i=1}^{n} a_i^2(\boldsymbol{x})x_i^g \cdot p(\boldsymbol{x})f(\boldsymbol{x})d\boldsymbol{x} + \beta^2\mu^2 - E_\xi m_Y^2. \qquad \dots \quad (2.5)$$

Thus for a given design function our problem is to minimize (2.5) subject to the conditions (2.2) and (2.3).

For a design function $p(\boldsymbol{x})$ define

$$q_i(x_i) = \int_{\mathcal{R}_{n-1}^+} p(\boldsymbol{x})f(\boldsymbol{x}) \prod_{j \neq i}^{n} dx_j. \qquad \dots \quad (2.6)$$

Let us assume that the given design function $p(\boldsymbol{x})$ satisfies the following conditions.

For some fixed $\nu > 1$ and for every $i = 1, 2, \dots, n$

$$|x|^{2\nu/(\nu-1)} \quad \text{is } q_i(x)\text{-integrable}$$

and

$$\left|\frac{q_i(x)}{f(x)}\right|^{(2\nu-1/(\nu-1)} \quad \text{is } f(x)\text{-integrable.} \qquad \dots \quad (2.7)$$

The number ν is chosen as close to 1 as possible so that the conditions (2.7) are still satisfied.

Let $\boldsymbol{a} = (a_1, a_2, \ldots, a_n)$, where $a_i(\boldsymbol{x})$ is $\mathcal{B}$-measurable, $i = 1, 2, \ldots, n$ and $q(\boldsymbol{x}) = p(\boldsymbol{x})f(\boldsymbol{x})$.

Define

$$U = \{\boldsymbol{a} : |a_i(\boldsymbol{x})|^{2s} \text{ is } q(\boldsymbol{x})\text{-integrable } \forall\, i = 1, 2, \ldots, n\}$$

$$V_1 = \{a : |a(\boldsymbol{x})|^2 \text{ is } q(\boldsymbol{x})\text{-integrable}\}$$

$$V_2 = \{b(x) : |b(x)|^2 \text{ is } f(x)\text{-integrable}\}.$$

Note that with usual 'L_p' norms U, V_1, V_2 are all Banach spaces. Let $V = V_1 \times V_2$ and V^* be the dual space (the space of all bounded linear functionals on V w.r.t. usual L_2-norm) of V. Clearly $V^* = V$.

Let
$$G(\boldsymbol{a}) = \int_{\mathcal{R}_n^+} \sum_{i=1}^{n} a_i^2(\boldsymbol{x}) x_i^2 p(\boldsymbol{x}) f(\boldsymbol{x}) d\boldsymbol{x}$$

$$H_1(\boldsymbol{a}) = I(x)\left(\sum_{i=1}^{n} a_i(\boldsymbol{x}) x_i - \mu \right)$$

$$H_2(\boldsymbol{a}) = I_1(x)(\phi(\boldsymbol{a}, x) - 1)$$

$$H(\boldsymbol{a}) = (H_1(\boldsymbol{a}),\ H_2(\boldsymbol{a}))$$

where
$$I(\boldsymbol{x}) = 1 \quad \text{if } \boldsymbol{x} \in \mathcal{R}_n^+$$
$$= 0 \quad \text{otherwise}$$

and
$$I_1(x) = 1 \quad \text{if } x \geqslant 0$$
$$= 0 \quad \text{otherwise.}$$

We are now in a position to formulate our problem as a familiar minimization problem on vector spaces :

$$\text{Minimize } G(\boldsymbol{a}) \text{ subject to } H(\boldsymbol{a}) = \theta \qquad \ldots \quad (2.8)$$

where θ is the zero vector of V.

It can be checked that G and H are infinitely Fréchet differentiable. To solve (2.8) we make use of the Lagrangian multipliers technique. The Lagrangian corresponding to (2.8) is given by

$$L(\boldsymbol{a}, v^*) = G(\boldsymbol{a}) + v^* H(\boldsymbol{a}) \qquad \ldots \quad (2.9)$$

where $v^* \epsilon V^* = V$

It is known that if $(\boldsymbol{a}_0, v_0^*)$ is the unconstrained minimum of (2.9) then $\boldsymbol{a}_0$ solves (2.8). Now (2.9) can be written as

$$L(\boldsymbol{a}, v^*) = \int_{\mathcal{R}_n^+} \Sigma a_i^2(\boldsymbol{x}) x_i^2 p(\boldsymbol{x}) f(\boldsymbol{x}) d\boldsymbol{x} - 2 \int_{\mathcal{R}_n^+} \lambda(\boldsymbol{x}) (\Sigma a_i(\boldsymbol{x}) x_i - \mu) p(\boldsymbol{x}) f(\boldsymbol{x}) d\boldsymbol{x}$$
$$-2 \int_{\mathcal{R}^+} b(x)(\phi(\boldsymbol{a}, x) - 1) f(x) dx$$

where $v^* = -2(\lambda(x), b(x))$, $\lambda(x) \in V_1$ and $b(x) \in V_2$.

Let $\delta L(a, v^*; h, w^*)$ be the derivative of $L(a, v^*)$ with the increment (h, w^*). Setting $\delta L(a, v^*; h, w^*) = 0 \; \forall \, (h, w^*) \in U \times V^*$ we get

$$H(a) = \theta$$

and $\forall \, h \in U$

$$\int_{\mathcal{R}_n^+} \sum_{i=1}^{n} a_i(x) h_i(x) x_i^g p(x) f(x) dx - \int_{\mathcal{R}_n^+} \lambda(x) \sum_{i=1}^{n} h_i(x) x_i p(x) f(x) dx$$

$$- \int_{\mathcal{R}^+} b(x) \phi(h, x) f(x) dx = 0. \qquad \dots \quad (2.10)$$

Note that

$$\int_{\mathcal{R}^+} b(x) \phi_i(h, x) f(x) dx = \int_{\mathcal{R}_n^+} b(x_i) h_i(x) p(x) f(x) dx.$$

Hence from (2.10), we get, $\forall \, h_i(x) \in U_1 = \{a : |a(x)|^{2\nu} \text{ is } q(x)\text{-integrable}\}$,

$$\int_{\mathcal{R}_n^+} h_i(x)[a_i(x) x_i^g - \lambda(x) x_i - b(x_i)] p(x) f(x) dx = 0; \; i = 1, 2, \dots, n.$$

Hence

$$a_i(x) x_i^g = \lambda(x) x_i + b(x_i). \qquad \dots \quad (2.11)$$

Now using the constraint (2.2) we get

$$\lambda(x) = \frac{\mu - \Sigma b(x_i) x_i^{1-g}}{\Sigma x_i^{2-g}}. \qquad \dots \quad (2.12)$$

Substituting the value of $\lambda(x)$ from (2.12) in (2.11) we get

$$a_i(x) = b(x_i) x_i^{-g} + \frac{\mu - \Sigma b(x_j) x_j^{1-g}}{\Sigma x_j^{2-g}} \cdot x_i^{1-g}. \qquad \dots \quad (2.13)$$

From (2.13) we get

$$\int_{\mathcal{R}_{n-1}^+} a_i(x) p(x) \prod_{j \neq i}^{n} f(x_j) dx_j = b(x_i)[x_i^{-g} r_i(x_i) - x_i^{2-2g} c_i(x_i)]$$

$$+ \mu x_i^{1-g} c_i(x_i) - x_i^{1-g} \sum_{j \neq i}^{n} \int_{\mathcal{R}^+} x_j^{1-g} b(x_j) c_{ij}(x_i, x_j) f(x_j) dx_j \; \dots \quad (2.14)$$

where $\quad r_i(x_i) = q_i(x_i)/f(x_i)$, $c_i(x_i) = \displaystyle\int_{\mathcal{R}_{n-1}^+} \frac{1}{\Sigma x_j^{2-g}} \, p(x) \prod_{j \neq i}^{n} f(x_j) dx_j$

and $\quad c_{ij}(x_i, x_j) = \displaystyle\int_{\mathcal{R}_{n-2}^+} \frac{1}{\Sigma x_k^{2-g}} \, p(x) \prod_{k \neq ij}^{n} f(x_k) dx_k.$

Now observe that

$$\sum_{j \neq i}^{n} \int_{\mathcal{R}^+} x_j^{1-g} b(x_j) c_{ij}(x_i, x_j) f(x_j) dx_j = \sum_{j \neq i}^{n} \int_{\mathcal{R}^+} x^{1-g} b(x) c_{ij}(x_i, x) f(x) dx$$

$$= \int_{\mathcal{R}^+} x^{1-g} b(x) D_i(x_i, x) f(x) dx,$$

where $$D_i(x_i, x) = \sum_{j\neq i}^{n} c_{ij}(x_i, x).$$

Substituting this in (2.14) we get,

$$\phi_i(\boldsymbol{a}, x) = b(x)(x^{-\theta} r_i(x) - x^{2-2\theta} c_i(x)) + \mu x^{1-\theta} c_i(x) - x^{1-\theta} \int_{\mathcal{R}^+} b(w) w^{1-\theta} D_i(x, w) f(w) dw.$$

Now using the constraint (2.3) we get

$$1 = b(x) \left[x^{-\theta} \sum_{i=1}^{n} r_i(x) - x^{2-2\theta} \sum_{i=1}^{n} c_i(x) \right] + \mu x^{1-\theta} \sum_{i=1}^{n} c_i(x) - x^{1-\theta} \int_{\mathcal{R}^+} b(w) w^{1-\theta} D(x, w) f(w) dw. \quad \ldots \quad (2.15)$$

where $$D(x, w) = \sum_{i=1}^{n} D_i(x, w).$$

Observe that

$$r_i(x) - x^{2-\theta} c_i(x) = \int_{\mathcal{R}^+_{n-1}} \left[1 - \frac{x^{2-\theta}}{x^{2-\theta} + \sum_{j\neq i}^{n} x_j^{2-\theta}} \right] p(x) \prod_{j\neq i}^{n} f(x_j) dx_j > 0.$$

Hence (2.15) can be compressed to the familiar Fredholm equation, [vide Hochstadt, 1973]

$$m(x) = b(x) - \int_{\mathcal{R}^+} K(x, w) b(w) f(w) dw \quad \ldots \quad (2.16)$$

where $$m(x) = \frac{1 - \mu x^{1-\theta} \Sigma c_i(x)}{x^{-\theta} \Sigma r_i(x) - x^{2-2\theta} \Sigma c_i(x)}$$

and $$K(x, w) = \frac{x w^{1-\theta} D(x, w)}{\Sigma r_i(x) - x^{2-\theta} \Sigma c_i(x)}.$$

Thus determining the Lagrangian multiplier $b(x)$ is equivalent to solving the equation (2.16). If $b(x)$ is a solution to (2.16) then by substituting in (2.13) we get n functions $\bar{a}_1, \bar{a}_2, \ldots, \bar{a}_n$. We now show that this $\bar{\boldsymbol{a}}$ is indeed a solution to (2.8). Treating $L(\boldsymbol{a}, v^*)$ as a functional in $\boldsymbol{a}$ we note that the second derivative with the increment $\boldsymbol{h}$, $\boldsymbol{h}$, $\delta^2 L(\boldsymbol{a}, v^*; \boldsymbol{h}, \boldsymbol{h}) > 0 \ \forall\ \boldsymbol{h} \in U$ and the higher order derivatives are uniformly zero. Hence using Taylor's expansion, namely

$$L(\boldsymbol{a}+\boldsymbol{h}, v^*) = L(\boldsymbol{a}, v^*) + \delta L(\boldsymbol{a}, v^*; \boldsymbol{h}) + \frac{\delta^2 L(\boldsymbol{a}, v^*; \boldsymbol{h}, \boldsymbol{h})}{2!} + \sum_{m \geqslant 3} \frac{1}{m!} \delta^m L(\boldsymbol{a}, v^*; \boldsymbol{h}, \boldsymbol{h}, \ldots, \boldsymbol{h}) \quad \ldots \quad (2.17)$$

we get for $\bar{\boldsymbol{a}}$, $L(\bar{\boldsymbol{a}}) \leqslant L(\bar{\boldsymbol{a}}+\boldsymbol{h}) \ \forall\ \boldsymbol{h} \in U$.

As a matter of fact, depending on solutions to (2.16), even if $\bar{\boldsymbol{a}}$ is not unique, again using (2.17) it is clear that the value $L(\boldsymbol{a}, v^*)$ is same for all of them, i.e., we may get different vectors $\bar{\boldsymbol{a}}$ leading to the same value of the functional $L(\boldsymbol{a}, v^*)$.

We now state a theorem which can be used to solve (2.16) :

Theorem 2.1 : *If* $m(x) \in V_2$ *and* $\int_{\mathcal{R}_2^+} K^2(x, w)f(x)f(w)dxdw < \infty$

then

$$b(x) - \lambda \int_{\mathcal{R}^+} K(x, w)b(w)f(w)dw = m(x) \qquad \ldots \quad (2.18)$$

has a unique solution if and only if

$$b(x) - \lambda \int_{\mathcal{R}} K(x, w)b(w)f(w)dw = 0 \qquad \ldots \quad (2.19)$$

has only the trivial solution $b(x) = 0$.

If (2.19) has at least one nontrivial solution then (2.18) will have a solution if

$$\int_{\mathcal{R}^+} m(x)l(x)f(x)dx = 0$$

for every $l(x)$ satisfying the equation

$$l(x) - \lambda \int_{\mathcal{R}^+} K(w, x)l(w)f(w)dw = 0.$$

We are now in a position to state our first existence theorem :

Theorem 2.2 : *For any design function* $p(x)$ *satisfying* (2.7) *and for which* (2.16) *has a solution there exists a best* p *as well as* ξ*-unbiased linear estimator under the model* (1.1) *with* g *known w.r.t. the measure of uncertainty* (1.2).

Example 2.1 : Let us consider an example so as to get the idea about the above result.

Let
$$p(x) = A \sum_{i=1}^{n} x_i^{2-g} \prod_{i=1}^{n} p(x_i) \qquad \ldots \quad (2.20)$$

where A is the normalizing constant and $p(x)$ is such that (2.7) is satisfied and $x^{1-g}p(x)$ and $x^g/p(x)$ belong to V_2. For the $p(x)$ given by (2.20), we have

$$\Sigma r_i(x) = nAp(x)[(n-1)\lambda_1\lambda_2 + x^{2-g}\lambda_3]$$

$$\Sigma c_i(x) = Ap(x)\lambda_3$$

$$D(x, w) = n(n-1)\lambda_2 A p(x)p(w)$$

$$K(x, w) = \frac{xw^{1-g}p(w)}{\lambda_1}$$

$$m(x) = \frac{\lambda_4 x^g}{p(x)} - x\lambda_5$$

where $\lambda_1 = \int_{\mathcal{R}^+} x^{2-g}\, p(x)f(x)dx, \qquad \lambda_2 = \left(\int_{\mathcal{R}^+} p(x)f(x)dx \right)^{n-2}$

$$\lambda_3 = \lambda_2 \int_{\mathcal{R}^+} p(x)f(x)dx, \quad \lambda_4 = \frac{1}{n(n-1)\lambda_1\lambda_2 A}, \; \lambda_5 = \frac{\mu\lambda_3}{(n-1)\lambda_1\lambda_2}$$

It is easy to check that $A^{-1} = n\lambda_1\lambda_3$.

The equation (2.16) reduces to

$$b(x) - \frac{x}{\lambda_1} \int_{\mathcal{R}^+} w^{1-g} p(w)b(w)f(w)dw = \frac{\lambda_4 x^g}{p(x)} - \lambda_5 x. \qquad \dots \quad (2.21)$$

Let
$$b_1 = \int_{\mathcal{R}^+} w^{1-g} p(w)b(w)f(w)dw;$$

then
$$b(x) - \frac{xb_1}{\lambda_1} = \frac{\lambda_4 x^g}{p(x)} - \lambda_5 x. \qquad \dots \quad (2.22)$$

Multiplying both sides of (2.22) by $x^{1-g}p(x)$ and integrating we get

$$b_1 \cdot 0 = \lambda_4 \mu - \lambda_5 \lambda_1. \qquad \dots \quad (2.23)$$

This shows that (2.22) has a solution for any real value of b_1 if and only if (2.23) is satisfied. But note that

$$\begin{aligned} \lambda_4\mu - \lambda_5\lambda_1 &= \frac{\mu}{n(n-1)\lambda_1\lambda_2 A} - \frac{\mu\lambda_3}{(n-1)\lambda_2} \\ &= \frac{\mu}{n(n-1)\lambda_2 A\lambda_1}(1 - nA\lambda_1\lambda_3) \\ &= 0, \end{aligned}$$

hence $b(x) = \dfrac{x\eta}{\lambda_1} + \dfrac{\lambda_4 x^g}{p(x)} - \lambda_5 x$ is a solution to (2.21), where η is any real number.

Substituting $b(x)$ in (2.13) we get a unique set of n functions $\bar{a}_1, \bar{a}_2, \dots, \bar{a}_n$, i.e., they do not depend on any particular choice of η. Thus for $p(x)$ in (2.20) the best p as well as ξ-unbiased estimator is given by

$$\sum_{i=1}^{n} \bar{a}_i(x)y(x_i)$$

where
$$\bar{a}_i(x) = \frac{\lambda_4}{p(x_i)} + \frac{\mu - \lambda_4 \Sigma x_i/p(x_i)}{\Sigma x_i^{2-g}}\, x_i^{1-g}, \; i = 1, 2, \dots, n. \qquad \dots \quad (2.24)$$

Remark 2.1 : In particular for $p(x_i) = x_i^{g-1}$ we get $\bar{a}_i(x) = \dfrac{\mu x_i^{1-g}}{\Sigma x_j^{2-g}}$, and further if $g = 2$ we get $\bar{a}_i(x) = \dfrac{\mu}{nx_i}$.

We now proceed to our next existence result.

For a p-unbiased linear strategy the measure of uncertainty (1.2) takes the form

$$M(p, T) = \sigma^2 \int_{\mathcal{R}_n^+} \sum_{i=1}^{n} a_i^2(x)x_i^g p(x)f(x)dx$$

$$+\beta^2 \int_{\mathcal{R}_n^+} \left(\sum_{i=1}^{n} a_i(x)x_i \right)^2 p(x)f(x)dx - E_\xi m_Y^2.$$

Let $\sigma^2/\beta^2 = k$. Our attempt is to find a best linear p-unbiased estimator for a given design function when k is known. We assume that $p(x)$ satisfies (2.7).

Let

$$G_1(a) = k \int_{\mathcal{R}_n^+} \sum_{i=1}^{n} a_i^2(x)x_i^g p(x)f(x)dx + \int_{\mathcal{R}_n^+} \left(\sum_{i=1}^{n} a_i(x)x_i \right)^2 p(x)f(x)dx.$$

Thus our problem is to

$$\text{Minimize } G_1(a) \text{ subject to } H_2(a) = \theta, \qquad \dots \quad (2.25)$$

where $H_2(a)$ is same as in the previous problem and θ is the zero vector of V_2.

The Lagrangian corresponding to (2.25) is given by

$$L_1(a,\ v^*) = G_1(a) + v^* H_2(a) \qquad \dots \quad (2.26)$$

where $v^* \in V_2^* = V_2$.

It is easy to check that G_1 is infinitely Fréchet differentiable. Proceeding on the lines similar to that used in solving the previous problem we get

$$a_i(x) = k^{-1}\left[b(x_i)x_i^{-g} - x_i^{1-g} \frac{\Sigma b(x_j)x_j^{1-g}}{k+\Sigma x_j^{2-g}} \right] \qquad \dots \quad (2.27)$$

where $-2b(x)$, $b(x) \in V_2$, is the Lagrangian multiplier.

To determine $b(x)$ we make use of the constraint (2.3). Now

$$\phi_i(a, x_i) = \int_{\mathcal{R}_{n-1}^+} a_i(x)p(x) \prod_{j \neq i}^{n} f(x_j)dx_j$$

$$= k^{-1} \int_{\mathcal{R}_{n-1}^+} x_i^{-g} \left[b(x_i) - x_i \frac{\Sigma b(x_j)x_j^{1-g}}{k+\Sigma x_j^{2-g}} \right] p(x) \prod_{j \neq i}^{n} f(x_j)dx_j$$

$$= k^{-1} \Bigg[x_i^{-g} b(x_i) \int_{\mathcal{R}_{n-1}^+} p(x) \prod_{j \neq i}^{n} f(x_j)dx_j$$

$$- x_i^{2-2g} b(x_i) \int_{\mathcal{R}_{n-1}^+} \frac{p(x)}{k+\Sigma x_j^{2-g}} \prod_{j \neq i}^{n} f(x_j)dx_j$$

$$- x_i^{1-g} \sum_{j \neq i}^{n} \int_{\mathcal{R}^+} b(x_j)x_j^{1-g} \left\{ \int_{\mathcal{R}_{n-2}^+} \frac{p(x)}{k+\Sigma x_l^{2-g}} \prod_{l \neq i, j}^{n} f(x_l)dx_l \right\} f(x_j)dx_j \Bigg]$$

$$= k^{-1}\left[b(x_i)x_i^{-g}(r_i(x_i) - x_i^{2-g}c_i'(x_i)) - x_i^{1-g} \int_{\mathcal{R}^+} b(w)w^{1-g} D_i'(x_i, w)f(w)dw \right]$$

where $$r_i(x_i) = q_i(x_i)/f(x_i);$$

$$c'_{ij}(x_i, x_j) = \int_{\mathcal{R}^+_{n-2}} \frac{p(x)}{k+\Sigma x_l^{2-g}} \prod_{l \neq i, j}^{n} f(x_l)dx_l;$$

$$c'_i(x_i) = \int_{\mathcal{R}^+} c'_{ij}(x_i, x_j)f(x_j)dx_j.$$

$$D'_i(x_i, x) = \sum_{j \neq i}^{n} c'_{ij}(x_i, x).$$

Now using $\sum_{i=1}^{n} \phi_i(a, x) = \phi(a, x) = 1$ we get

$$b(x) - \int_{\mathcal{R}^+} K'(x, w)b(w)f(w)dw = m'(x) \qquad \dots \quad (2.28)$$

where $K'(x, w) = \dfrac{xw^{1-g}D'(x, w)}{r(x)-x^{2-g}B(x)}$ and $m'(x) = \dfrac{kx^g}{r(x)-x^{2-g}B(x)}$

with $D'(x, w) = \sum_{i=1}^{n} D'_i(x, w)$, $r(x) = \sum_{i=1}^{n} r_i(x)$ and $B(x) = \sum_{i=1}^{n} c'_i(x)$.

Now to solve the equation (2.28) we make use of Theorem 2.1. Thus if $b(x)$ is a solution to (2.28) then by substituting it in (2.27) we get a vector $\bar{a}'$. As shown in first problem, we can indeed prove that this $\bar{a}'$ is a solution to (2.25) and if there is more than one choice for $\bar{a}'$ the corresponding value of the functional $L_1(a, v^*)$ is same for all of them. Thus we have our second existence theorem as follows.

Theorem 2.3 : *For any design function $p(x)$ satisfying* (2.7) *and for which* (2.28) *has a solution there exists a best p-unbiased linear estimator under the model* (1.1) *with g and $\sigma^2/\beta^2 = k$ known; w.r.t. the measure of uncertainty* (1.2).

Let us consider an example so as to get an idea about the above result.

Example 2.2 : Let

$$p(x) = A\left(k + \sum_{i=1}^{n} x_i^{2-g}\right) \prod_{i=1}^{n} p(x_i)$$

where A is the normalizing constant. Let $p(x)$ be such that (2.7) is satisfied and $x^{1-g}p(x)$, $x^g/p(x)$ belong to V_2. It can be checked that

$$m'(x) = \frac{kx^g}{nAp(x)(k\lambda_1+(n-1)\lambda_2)\lambda_1^{n-2}} = \lambda_3 \frac{x^g}{p(x)} \quad \text{(say)}$$

and $$K'(x, w) = \frac{(n-1)xw^{1-g}p(w)}{(k\lambda_1+(n-1)\lambda_2)} = \lambda_4 xw^{1-g}p(w) \quad \text{(say)}$$

where $$\lambda_1 = \int_{\mathcal{R}^+} p(x)f(x)dx \quad \text{and} \quad \lambda_2 = \int_{\mathcal{R}^+} x^{2-g}p(x)f(x)dx.$$

Thus the equation (2.29) reduces to

$$b(x)-\lambda_4 x \int_{R^+} w^{1-g}p(w)b(w)f(w)dw = \lambda_3 \frac{x^g}{p(x)}. \qquad \ldots \quad (2.29)$$

Letting $b_1 = \int_{R^+} w^{1-g}p(w)b(w)f(w)dw$, multiplying both sides of (2.29) by $x^{1-g}p(x)$ and integrating we get

$$b_1(1-\lambda_4\lambda_2) = \lambda_3\mu$$

note that $\lambda_4\lambda_2 < 1$.

Hence
$$b_1 = \frac{\lambda_3\mu}{1-\lambda_4\lambda_2}.$$

Thus
$$b(x) = \frac{\lambda_3 x^g}{p(x)} + \frac{\lambda_3\lambda_4\mu x}{1-\lambda_4\lambda_2}$$

$$= \frac{\lambda_3 x^g}{p(x)} + \lambda_5 x \quad \text{(say)}.$$

Substituting in (2.27) we get

$$a_i(\boldsymbol{x}) = \frac{\lambda_3}{kp(x_i)} + \frac{x_i^{1-g}}{k+\Sigma x_j^{2-g}}\left[\lambda_5 - \frac{\lambda_3}{k}\Sigma\frac{x_j}{p(x_j)}\right]; \; i = 1, 2, \ldots, n. \qquad \ldots \quad (2.30)$$

Thus for the above $p(\boldsymbol{x})$ the best linear p-unbiased estimator is given by

$$\sum_{i=1}^{n} \bar{a}_i'(\boldsymbol{x})y(x_i)$$

where $\bar{a}_i'(\boldsymbol{x})$, $i = 1, 2, \ldots, n$ are given by (2.30).

Remark 2.2 : In particular for $g = 2$ and $p(x_i) = \frac{x_i}{\mu}$ we get $\bar{a}_i'(\boldsymbol{x})$ independent of k, namely $\bar{a}_i'(\boldsymbol{x}) = \frac{\mu}{nx_i}$.

Acknowledgement. The author is thankful to a referee for his suggestions leading to a better presentation of the results.

References

Hochstadt, H. (1973) : *Integral Equations.* Wiley Interscience, New York, 108-109.

Särndal, C. E. (1980) : A method of assessing efficiency and bias of estimation strategies in survey sampling. *S. African J. Statist.*, **24**, 17–30.

Paper received : September, 1981.
Revised : April, 1983.