Journal of Statistical Planning and Inference 41 (1994) 151-161
North-Holland

Prediction in finite population under error-in-variables superpopulation models

P. Mukhopadhyay

Computer Science Unit, Indian Statistical Institute, 203 Barrackpore Trunk Road, Calcutta 700 035, India

Received 25 January 1993; revised manuscript received 21 October 1993

Abstract

In this paper we consider the optimal prediction of finite population total and variance under location model with measurement errors. Bayes predictors of population total and variance under a class of priors have been derived and a minimax predictor for population total has been obtained. Under regression superpopulation model with measurement errors, an optimal predictor of population total has been derived. A Bayesian approach for this model shows the linear regression predictor to be a Bayes predictor.

AMS Subject Classification: 62D05.

Key words: Superpopulation; prediction; measurement error; Bayesian; posterior distribution; risk function; minimax.

1. Introduction

In practical sample survey situations the true values of the variables are rarely observed but values mixed with measurement errors. We consider, in this paper, problems of estimation of finite population total and variance when the survey data contain observations mixed with random measurement errors.

Consider a finite population $\mathscr{U}$ of a known number N of identifiable units labelled $1, \ldots, i, \ldots, N$. Associated with i is a value y_i of a study variable y. We assume that y_i cannot be observed correctly but a different value Y_i which is mixed with measurement errors is observed.

We also assume that the true value y_i in the finite population is actually a realisation of a random variable $\mathscr{Y}_i$, the vector $\underset{\sim}{\mathscr{Y}} = (\mathscr{Y}_1, \ldots, \mathscr{Y}_N)$ having a joint distribution model ξ. However, both y_i and $\mathscr{Y}_i$ are not observable and we cannot make any distinction between them. Our problem is to predict the population total $T = \sum_{i=1}^N y_i$ or the population variance $S_y^2 = 1/(N-1) \sum_{i=1}^N (y_i - \bar{y})^2$, $\bar{y} = T/N$ by drawing a sample

Correspondence to: Dr. P. Mukhopadhyay, Computer Science Unit, Indian Statistical Institute, 203 Barrackpore Trunk Road, Calcutta 700 035, India.

s according to a sampling design p with selection probability $p(s)$ and observing the data $Y_s=(i, Y_i; i\in s)$. A combination of a sampling design and a predictor is called a sampling strategy. We shall often confine to the class ρ_n of sampling designs with fixed size n, $\rho_n=\{p: p(s)>0 \Rightarrow n(s)=n\}$, $n(s)$ denoting the size of the sample s.

In Section 2 we obtain the optimal strategies for predicting T or $\bar{y}$ and S_y^2 under the simple location model with measurement errors. Bayes predictor of T under a class of priors along with their Bayes risks is obtained. A minimax predictor of T has been derived. A Bayes predictor of S_y^2 has also been obtained.

In Section 3 we obtain an optimal strategy for T under simple regression model with measurement errors in both variables. We have obtained the Bayes predictor of T under this model when only the observations on the regressor variable are mixed with errors.

A general treatment for inference problem under measurement error models has been considered in Fuller (1987, 1989). Prediction in finite population under measurement error models has been considered in Bolfarine (1991).

2. The location model with measurement error

Consider the simple location model with measurement errors

$$y_i=\mu+e_i, \quad E(e_i)=0, \quad E(e_i^2)=\sigma_e^2, \quad E(e_i e_i')=0, \qquad (i\neq i'), \tag{1a}$$

$$Y_i=y_i+u_i, \quad E(u_i)=0, \quad E(u_i^2)=\sigma_u^2, \quad E(u_i u_i')=0, \quad E(e_i u_j)=0$$
$$i,j=1,\ldots,N, \tag{1b}$$

where $\mu, \sigma_e^2\ (>0), \sigma_u^2\ (>0)$ are constants. Note that e_i's are random variables due to superpopulation distribution of $\underset{\sim}{\mathscr{Y}}$, whereas u_i's are due to measurement errors. We shall use the same symbols E, V and C to denote expectation, variance and covariance respectively, with respect to joint or marginal distribution of e_i and u_i.

The models (1a) and (1b) are the simple location error-in-variable superpopulation models and have been considered by Bolfarine (1991). In Section 2.1 we consider the optimal strategies for predicting $\bar{y}$ and S_y^2 under this model. Section 2.2 derives Bayes predictor for T and S_y^2 under a class of priors. A minimax predictor of T is also obtained.

2.1.

Consider the class of linear predictors

$$e(s, Y_s)=b_s+\sum_{k\in s} b_{ks} Y_k, \tag{2}$$

where b_s, b_{ks} are constants not depending on Y-values.

A predictor g is said to be a design-model (pm-)unbiased predictor of $\theta(\boldsymbol{y})$ (or pm-unbiased estimator of $E(\theta(\boldsymbol{y}))$ where $\boldsymbol{y}=(y_1,\dots,y_N)$) if

$$E_pE(g(s,\boldsymbol{Y}_s))=E(\theta(\boldsymbol{y})) \tag{3}$$

when E_p and V_p denote, respectively, expectation and variance with respect to the sampling design p. Hence, $e(s,\boldsymbol{Y}_s)$ will be pm-unbiased for $\bar{y}$ iff

$$E_pE\left(b_s+\sum_{k\in s} b_{ks}Y_k\right)=E(\bar{y})=\mu,$$

i.e. iff

$$E_p(b_s)=0, \tag{4a}$$

$$E_p\left(\sum_{k\in s} b_{ks}\right)=1. \tag{4b}$$

Following the usual variance-minimisation criterion, a predictor g^* will be said to be optimal in a class of predictors G for predicting $\theta(\boldsymbol{y})$ for a fixed p, if

$$E_pE(g^*-\theta)^2\leqslant E_pE(g-\theta)^2 \tag{5}$$

for all $g\in G$.

To find an optimal pm-unbiased predictor of $\bar{y}$, we consider the following theorem on UMVU-estimation (Rao, 1973). Let C denote a class of pm-unbiased estimators of τ and C_0 the corresponding class of pm-unbiased estimators of zero.

Theorem 1. *A predictor g^* in C is optimal for τ iff for any f in C_0, $E_pE(g^*f)=0$.*

From the above theorem, Theorem 2 readily follows.

Theorem 2. *Under models* (1a) *and* (1b), *optimal pm-unbiased predictor of $\bar{y}$ in the class of all linear pm-unbiased predictors, where $p\in\rho_n$, is given by $\bar{Y}_s$. Again any $p\in\rho_n$ is optimal for using $\bar{Y}_s$.*

If $\mathscr{V}$ denotes the variance operator with respect to models (1a) and (1b) and sampling design p,

$$\begin{aligned}\mathscr{V}(\bar{Y}_s-\bar{y})&=E_pV(\bar{Y}_s-\bar{y})+V_pE(\bar{Y}_s-\bar{y})\\&=E_pV(\bar{Y}_s-\bar{y})\\&=E_p[V(\bar{Y}_s)+V(\bar{y})-2C(\bar{Y}_s-\bar{y})]\\&=\left(\frac{1}{n}-\frac{1}{N}\right)\sigma_e^2+\frac{\sigma_u^2}{n}.\end{aligned} \tag{6}$$

Equation (6) actually gives the value of $E(\bar{Y}_s-\bar{y})^2$ which is the risk corresponding to $\bar{Y}_s$, assuming a squared-error loss function.

Theorem 2 states that any $FS(n-)$ design including a purposive sampling design ($p(s^*)=1$ for s^*, a specified set of n units) is optimal for predicting $\bar{y}$. However, for purpose of robustness under model failures one should consider (as shown in a different context by Godambe and Thompson (1977)) a probability sampling design $p \in \rho_n$ along with $\bar{Y}_s$.

We now consider optimal prediction of S_y^2. For this we shall confine to the class of pm-unbiased quadratic predictors

$$e_q(s, Y_s) = b_s + \sum_{k \in s} b_{ks} Y_k^2 + \sum_{k \neq k' \in s} b_{kk's} Y_k Y_{k'},$$

where b_s, $b_{ks'}$ $b_{kk's}$ are suitable constants that do not depend on Y_s.

Assumption A. We assume that y_i's are identically distributed with finite fourth-order moments and that the conditional distributions of Y_i, given y_i, are identical with finite moments of fourth order.

Under these assumptions, all moments of the joint distribution of Y_i, Y_j, Y_k, etc., are constants, depending only on the order of moments.

By virtue of Theorem 1 the following result can easily be verified.

Theorem 3. *Under models* (1*a*) *and* (1*b*) *and under Assumption A*

$$s_Y^2 = \frac{1}{(n-1)} \sum_{k \in s} (Y_k - \bar{Y}_s)^2$$

is an optimal predictor of S_y^2 *in the class of all pm-unbiased quadratic predictors for any given* $p \in \rho_n$. *Again any* $p \in \rho_n$ *is optimal for using* s_Y^2.

2.2.

We now consider Bayes prediction of T under different priors for (μ, σ_e^2). E and V will denote throughout expectation and variance with respect to different prior distributions, posterior distributions, etc.

2.2.1.

Assume that σ_e^2, σ_u^2 are known. As the distribution of a large number of variables including socioeconomic variables is (at least approximately) normal in large sample, we consider a normal prior $N(0, \theta^2)$ for μ. On the assumption that the errors e_i, u_i are independent normal variables, the likelihood of Y_s, given μ, is

$$L(Y_s | \mu) \propto \exp\left[-\frac{n}{2\sigma^2}(\bar{Y}_s - \mu)^2\right], \tag{7}$$

where

$$\sigma^2 = \sigma_e^2 + \sigma_u^2. \tag{8}$$

The posterior distribution of μ is, therefore,

$$N\left(\frac{nY_s\theta^2}{n\theta^2+\sigma^2}, \frac{\sigma^2\theta^2}{n\theta^2+\sigma^2}\right). \tag{9}$$

Again, likelihood of Y_s, given $y_s=(y_i,\ i\in s)$, is

$$L(Y_s|y_s)\propto\exp\left[-\frac{1}{2\sigma_u^2}\sum_{i\in s}(Y_i-y_i)^2\right]. \tag{10}$$

Hence, posterior distributions of y_i's, given (Y_s,μ), are independent:

$$y_i\sim N\left(\frac{Y_i\sigma_e^2+\mu\sigma_u^2}{\sigma^2}, \frac{\sigma_e^2\sigma_u^2}{\sigma^2}\right) \quad (i\notin s), \tag{11}$$

$$y_i\sim N(\mu,\sigma_e^2) \quad (i\notin s).$$

We shall assume throughout a squared-error loss function. Therefore, Bayes predictor of T is

$$\hat{T}_B=E\left(\sum_{i=1}^{N} y_i|\,Y_s\right)=E\left\{E\left(\sum_{i=1}^{N} y_i\,\middle|\,Y_s,\mu\right)\middle|\,Y_s\right\}$$

$$=n\bar{Y}_s\left[\frac{\sigma_e^2}{\sigma^2}+\frac{(N-n)\sigma_e^2+N\sigma_u^2}{\sigma^2(n\theta^2+\sigma^2)}\theta^2\right] \quad \text{(using (9) and (11))}. \tag{12}$$

The variance of the posterior distribution of T is independent of Y_s. Hence, Bayes risk of $\hat{T}_B$ is the posterior variance of T and is given by

$$V(T\,|\,Y_s)=E\{V(T\,|\,\mu,Y_s)\}+V\{E(T\,|\,\mu,Y_s)|\,Y_s\}$$

$$=\frac{n\sigma_e^2\sigma_u^2}{\sigma^2}+(N-n)\sigma_e^2+\left(N-\frac{n\sigma_e^2}{\sigma^2}\right)^2\frac{\sigma^2\theta^2}{n\theta^2+\sigma^2}$$

$$=r_\theta(\hat{T}_B) \quad \text{(say)}. \tag{13}$$

As $\theta\to\infty$,

$$r_\theta(\hat{T}_B)\to\frac{N(N-n)}{n}\sigma_e^2+\frac{N^2}{n}\sigma_u^2$$

$$=r_0 \quad \text{(say)}. \tag{14}$$

Now we consider two theorems connecting Bayes estimate and minimax risk estimate.

Theorem 4 (Lehmann, 1950). *If $\{\lambda_n\}$ is a sequence of a priori probability distributions, $\{r_n\}$ the sequence of associated Bayes risks and if $r_n\to r$ as $n\to\infty$, and if there exists some predictor δ for which the risk $R(\delta,T)=E(\delta-T)^2\leqslant r$, whatever be the value of T, then δ is a minimax predictor.*

Theorem 5 (Aggarwal, 1959). *If δ and r are a minimax procedure and minimax risk, respectively, assuming that the observations Y_s follow any probability distributions*

$w \in \Omega^*$ and if $\Omega^* \subset \Omega$ is a space of distributions for which the risk associated with δ does not exceed r, then δ is a minimax procedure and r the minimax risk for all distributions of Y_s in Ω.

It is seen from (6) that the risk of the predictor $N\bar{Y}_s$ is given by r_0. Hence, $N\bar{Y}_s$ is a minimax predictor of T under the assumption of normality of the distribution of e_i's and u_i's as considered above. Again, since expression (6) was obtained without any assumption about the form of the distribution, the predictor $N\bar{Y}_s$ is minimax in the general class of distributions (not necessarily normal) which satisfy models (1a) and (1b). Hence, we have the following theorem.

Theorem 6. *The predictor $N\bar{Y}_s$ is a minimax predictor of T under the general class of prior distributions of errors (e_i's and u_i's) which satisfy models (1a) and (1b).*

2.2.2.

We now assume $\tau = 1/\sigma_e^2$ is unknown. Also assume that $1/\sigma_u^2 = k\tau$, where k is a known positive constant.

If we assume a normal–gamma prior for (μ, τ) (e.g. Broemeling, 1985) with parameters (v, α, β),

$$P(\mu, \tau) \propto \tau^{\alpha - 1/2} \exp\left\{-\frac{\tau}{2}[(\mu - v)^2 + 2\beta]\right\},$$

$$\mu \in R_1,\ \tau > 0,\ \alpha > 0,\ \beta > 0,\ v \in R_1, \tag{15}$$

marginal posterior distribution of μ is a Student's t-distribution with $(n+2\alpha)$ d.f. and posterior mean and variance given, respectively,

$$E(\mu/Y_s) = \frac{v + nq\bar{Y}_s}{1 + nq}, \tag{16a}$$

$$V(\mu/Y_s) = \frac{2\beta + q\sum_{i \in s} Y_i^2 + v^2 - \dfrac{(v + nq\bar{Y}_s)^2}{(1 + nq)}}{(1 + nq)(n + 2\alpha)}, \tag{16b}$$

where $q = k/(k+1)$. It is assumed that $n > 1$.

Marginal posterior distribution of τ is a gamma with parameters

$$\alpha^* = \frac{n + 2\alpha}{2},$$

$$\beta^* = \beta + \frac{v^2}{2} + \frac{q}{2}\sum_{i \in s} Y_i^2 - \frac{(v + nq\bar{Y}_s)^2}{2(1 + nq)}. \tag{17}$$

Here Bayes predictor of T is

$$\hat{T}_B^{(1)} = E\left\{E\left(\sum_{i\in s} y_i + \sum_{i\in \bar{s}} y_i/\mu, \tau, \mathbf{Y}_s\right)/\mathbf{Y}_s\right\}$$

$$= \frac{kn(N+1)}{k(n+1)+1}\bar{Y}_s + \frac{N+k(N-n)}{k(n+1)+1}\theta. \tag{18}$$

To use $\hat{T}_B^{(1)}$ one needs to know only the value of $k=\sigma_e^2/\sigma_u^2$.

We note that when $n=N$, $\hat{T}_B^{(1)} \neq Y$ and hence $\hat{T}_B^{(1)}$ is not a consistent estimator in Cochran's sense (Cochran, 1977, p. 21). This is, however, not surprising since $\hat{T}_B^{(1)}$ was derived under a prior distribution for (μ, τ).

2.2.3.

We now consider Jeffrey's (1959) noninformative prior

$$P(\mu,\tau) \propto \frac{1}{\tau}, \quad \mu \in R_1, \ \tau > 0. \tag{19}$$

Bayes predictor of T is

$$\hat{T}_B'' = E(T/\mathbf{Y}_s) = E\left\{\frac{n\mu + kn\bar{Y}_s}{k+1} + (N-n)\mu/\mathbf{Y}_s\right\}$$

$$= \bar{Y}_s. \tag{20}$$

Posterior variance of T is

$$V(T/\mathbf{Y}_s) = E\{V(T/\mathbf{Y}_s, \mu, \tau)/\mathbf{Y}_s\} + V\{E(T/\mathbf{Y}_s, \mu, \tau)/\mathbf{Y}_s\}$$

$$= E\left[\frac{N(k+1)-nk}{k+1}\cdot\frac{1}{\tau}\Big/\mathbf{Y}_s\right] + V\left[\left(N - \frac{nk}{k+1}\right)\mu\Big/\mathbf{Y}_s\right]$$

$$= \frac{N(k+1)-nk}{(k+1)^2} s_Y^2 \left[\frac{2k(n-1)}{n+3} + \frac{N(k+1)-nk}{n}\right]$$

$$\simeq \frac{N^2(k+1)^2 - n^2k^2}{n(k+1)^2} s_Y^2 \quad \left(\text{assuming } \frac{n-1}{n-3} \simeq 1\right)$$

$$= r(\hat{T}_B'') \quad \text{(say)}. \tag{21}$$

In particular, for $k=1$,

$$r(\hat{T}_B'') = \frac{4N^2 - n^2}{4n} s_Y^2. \tag{22}$$

Result (20) was also obtained by Bolfarine (1991) for non-informative prior distribution of μ. We, however, considered here a non-informative prior joint distribution for (μ, τ) and as a result the expression for posterior variance of T in (21) differed from his expression. His result is the same as expression (6) in the present paper.

2.2.4.

Under Jeffrey's prior, (19), Bayes predictor of S_y^2 is

$$\hat{S}_{yB}^2 = E(S_y^2/Y_s)$$

$$= \frac{k^2}{(k+1)^2}\frac{n(N+n-2)}{N(n-1)}\bar{Y}_s^2 + \frac{ks_Y^2}{(k+1)^2}\left[k + \frac{k(N-n)}{N(N-1)} + \frac{2(n-1)}{N(n-3)}((N-n)(k+1)+1)\right]. \quad (23)$$

Under non-informative prior distribution of μ, Bolfairne obtained a different Bayes predictor of S_y^2 (equation (5) of his (Bolfairne (1991)) paper). However, we considered joint non-informative prior for (μ, τ) and as such the expression (23) depends only on the value of k and does not require the knowledge of both σ_e^2 and σ_u^2.

3. Regression model with measurement errors

We assume that, associated with each i, there is a true value x_i of an auxiliary variable x closely related to the main variable y. The values x_i's, however, cannot be measured without error and instead some other values X_i's are observed. It is assumed that $X_1, \ldots, X_N$ are known fixed quantities. We assume, further, that the unknown true value y_i of the study variable y is a realisation of a random variable y_i obeying a superpopulation model such that

$$y_i = \beta_0 + \beta_1 x_i + e_i, \quad E(e_i) = 0, \quad E(e_i^2) = \sigma_e^2, \quad E(e_i \cdot e_{i'}) = 0, \ i \neq i'. \quad (24a)$$

Again

$$X_i = x_i + v_i, \quad E(X_i/x_i) = x_i, \quad V(X_i/x_i) = \sigma_v^2, \quad E(X_i X_{i'}/x_i, x_{i'}) = x_i x_{i'}, \ i \neq i', \quad (24b)$$

$$Y_i = y_i + u_i, \quad E(Y_i/y_i) = y_i, \quad V(Y_i/y_i) = \sigma_u^2, \quad E(Y_i Y_{i'}/y_i, y_{i'}) = y_i y_{i'}, \ i \neq i', \quad (24c)$$

where e_i, v_i, u_i are assumed to be mutually independent and Y_i's are as defined in Section 1.

Here $\beta_0, \beta_1, \sigma_e^2, \sigma_v^2, \sigma_u^2$ are constants. Distributions of e_i's are due to superpopulation model ξ, whereas conditional distributions of $Y_i(X_i)$, given $y_i(x_i)$, are due to measurement error. The operators E, V and C will denote expectation, variance and covariance, respectively, with respect to joint or marginal distribution of y_i, X_i, Y_i.

3.1.

We shall first consider optimal prediction of $\bar{y}$ under the models (24a)–(24c). The following result follows from Theorem 1.

Theorem 7. *Under models* (24*a*)–(24*c*) *the best linear optimal pm-unbiased predictor of* $\bar{y}$ *for any given* $p \in \rho_n$ *is*

$$e_1^* = \left(\sum_{k=1}^{N} \frac{\pi_k}{Z_k} \right)^{-1} \sum_{k \in s} \frac{Y_k}{Z_k}, \tag{25}$$

where

$$Z_k = \beta_0 + \beta_1 x_k.$$

Again,

$$\begin{aligned} &E_p E(e_1^* - \beta_0 - \beta_1 \bar{x})^2 \\ &= \frac{1}{\left(\sum_{k=1}^{N} \frac{\pi_k}{Z_k} \right)^2} \left[n^2 + (\beta^2 \sigma_v^2 + \sigma_u^2 + \sigma_e^2) \sum_{k=1}^{N} \frac{\pi_k}{Z_k^2} \right] - \bar{Z}^2 \\ &= \delta_n \quad (say), \end{aligned}$$

a constant dependent only on n. Hence, any $p \in \rho_n$ *is an optimal sampling design for using the optimal predictor* e_1^*.

Note 1: In deriving e_1^* it is assumed that β_0, β_1 are all known. However, in practice, the parameters β_0, β_1 will remain unknown and require to be estimated. For this we take recourse to the following procedure.

In addition to assumptions (24a)–(24c) we assume that e_i, v_i, u_i are independent normal $N(0, \sigma^2), N(0, \sigma_v^2), N(0, \sigma_u^2)$, respectively. Further, x_i's are assumed to be independently normally distributed (μ_x, σ_x^2) and independent of e_j, v_l, u_t $(i, j, l, t = 1, \ldots, N)$. Under these assumptions (Y_i, X_i) has a bivariate normal distribution with mean vector

$$\begin{bmatrix} \mu_Y \\ \mu_X \end{bmatrix} = \begin{bmatrix} \beta_0 + \beta_1 \mu_x \\ \mu_x \end{bmatrix}$$

and dispersion matrix

$$\begin{bmatrix} \beta_1^2 \sigma_x^2 + \sigma_u^2 + \sigma_e^2 & \beta_1 \sigma_x^2 \\ \beta_1 \sigma_x^2 & \sigma_x^2 + \sigma_v^2 \end{bmatrix}.$$

We denote by m_{XX}, m_{YY}, m_{XY} the sample variances of X and Y and covariances of (X, Y), respectively. β_0, β_1 are estimated in the following situations.

Case a. The ratio $\sigma_x^2 / \sigma_X^2 = k_{xx}$, *called the reliability ratio, is known*:

Here

$$\hat{\beta}_1 = \frac{\gamma_e}{k_{xx}},$$

where

$$\gamma_e = \sum_{i=1}^{n} (X_i - \bar{X}_s)(\bar{Y}_i - \bar{Y}_s) \Big/ \sum_{i=1}^{n} (X_i - \bar{X}_s)^2, \quad \bar{X}_s = \frac{1}{n} \sum_{i \in s} X_i.$$

Case b. The measurement error variance σ_v^2 is known:

$$\hat{\beta}=\frac{m_{XY}}{m_{XX}-\sigma_v^2}.$$

Case c. The ratio $(\sigma_u^2+\sigma_e^2)/\sigma_v^2=\delta$ is known:

$$\hat{\beta}=\frac{1}{2m_{XY}}\left[m_{YY}-\delta m_{XX}+\sqrt{(m_{YY}-\delta m_{XX})^2+4\delta m_{XY}^2}\right].$$

In all the cases $\hat{\beta}_0=\bar{Y}_s-\hat{\beta}\bar{X}_s$. The above derivations follow from Fuller (1987).

Note 2: In case, X_k's are known only for $k\in s$, e_1^* may be replaced by

$$e_1'^*=\frac{\sum_{k\in s}\dfrac{Y_k}{Z_k}}{\sum_{k\in \bar{s}}\dfrac{1}{Z_k}}, \tag{26}$$

which is a Ha'jek (1959) type predictor. The predictor $e_1'^*$ is pm-biased.

3.2.

We now consider Bayes prediction of T. For simplicity, we assume that x's can be measured without error so that models (24a) and (24b) only are relevant. We also assume that e_i's, u_i's are independently normally distributed with the parameters as stated and σ_e^2, σ_u^2 are known. Suppose also that x_k, $k=1,\ldots,N$, are all known quantities.

We assume that the prior distribution of $\boldsymbol{\beta}=(\beta_0,\beta_1)^{\mathrm{T}}$ is bivariate normal with mean $\boldsymbol{b}^0=(b_0^0,b_1^0)^{\mathrm{T}}$ and precision matrix $q\boldsymbol{S}^0$ $(q=k/(k+1))$, where $\boldsymbol{S}^0$ is a 2×2 positive semidefinite matrix. The posterior distribution of $\boldsymbol{\beta}$ given $\boldsymbol{Y}_s, \boldsymbol{X}_s$ where $\boldsymbol{X}_s=[1,x_k; k\in s]_{n\times 2}$ is normal (Raiffa and Schlaifer, 1961) with

$$E(\dot{\boldsymbol{\beta}}/\boldsymbol{Y}_s,\boldsymbol{X}_s)=\boldsymbol{b}^{00}=\boldsymbol{S}^{00^{-1}}(\boldsymbol{S}\boldsymbol{b}+\boldsymbol{S}^0\boldsymbol{b}^0) \tag{27a}$$

and dispersion matrix

$$D(\boldsymbol{\beta}/\boldsymbol{Y}_s,\boldsymbol{X}_s)=S_{00}^{-1}=\frac{1}{q\tau}(\boldsymbol{S}^0+\boldsymbol{S})^{-1}, \tag{27b}$$

where

$$\boldsymbol{b}=[\bar{Y}_s-b\bar{x}_s, b]^{\mathrm{T}},$$

$$b=\sum_s(Y_i-\bar{Y}_s)(x_i-\bar{x}_s)/\sum_s(x_i-\bar{x}_s)^2,$$

$$\boldsymbol{S}=\begin{pmatrix} n & \sum_s x_k \\ \sum_s x_k & \sum_s x_k^2\end{pmatrix}. \tag{28}$$

Again, assuming that the model errors e_i's are independent normal, Bayes predictor of T is

$$\hat{T}_B''' = \left(N - \frac{nk}{k+1}\right) b_0^{00} + \left(\frac{n\bar{x}_s}{k+1} + (N-n)\bar{x}_{\bar{s}}\right) b_1^{00} + \frac{nk\bar{Y}_s}{k+1},$$

where

$$\bar{x}_{\bar{s}} = \sum_{k \in \bar{s}} x_k/(N-n), \quad \bar{s} = u - s.$$

In particular, if we assume a natural conjugate prior of $\boldsymbol{\beta}$ so that $\boldsymbol{b}^0 = \boldsymbol{b}$, $\boldsymbol{S}^0 = \boldsymbol{S}$, then $\boldsymbol{b}^{00} = \boldsymbol{b}$, and $\hat{T}_B'''$ reduces to

$$\hat{T}_B''' = N\bar{Y}_s - b(N-n)(\bar{x}_s - \bar{x}_{\bar{s}})$$
$$= N[\bar{Y}_s + b(\bar{X} - \bar{x}_s)],$$

the linear regression predictor of T.

Acknowledgement

The author is thankful to the referee for many valuable suggestions leading to the overall improvement of the paper.

References

Aggarwal, O.P. (1959). Bayes and minimax procedure in sampling from finite and infinite populations. *Ann. Math. Statist.* **30**, 206–218.

Bolfarine, H. (1991). Finite population prediction under error-in-variables superpopulation models. *Can. J. Statist.* **19**(2), 191–207.

Broemeling, L.D. (1985). *Bayesian Analysis of Linear Models.* Marcel Drekker, New York.

Cochran, W.G. (1977). *Sampling Techniques*, 3rd ed. Wiley Eastern, New Delhi.

Fuller, W.A. (1987). *Measurement Error Models.* Wiley, New York.

Fuller, W.A. (1989). Prediction of true values for the measurement error models. In: P.J. Brown and W.A. Fuller, Eds., *Statistical Analysis of Measurement Error Models and Applications.* Am. Mathematical Soc., Providence, RI.

Godambe, V.P. and M.E. Thompson (1977). Robust near optimal estimation in survey practice. *Bull. Int. Statist. Instt.* **47**(3), 129–146.

H'ajek, J. (1959). Optimum strategy and other problems in probability sampling. *Casopis. Pest. Math.* **84**, 387–423.

Lehmann, E.L. (1950). Mimeographed notes on the theory of estimation, Chap. IV. University of California, Berkeley, CA.

Mukhopadhyay, P. (1992). On prediction in finite population under error-in-variables superpopulation models. *Tech. Rep. ASC/92/11*, Ind. Stat. Instt.

Rao, C.R. (1973). *Linear Statistical Inference and its Applications*, 2nd ed., Wiley Eastern, New Delhi.

Raiffa, H. and R. Schlaifer (1961). *Applied Statistical Decision Theory.* Harvard University, Boston.