BAYES PREDICTION OF POPULATION QUANTITIES IN A FINITE POPULATION

HU Gui-kai, XIONG Peng-fei, WANG Tong-xin

(School of Science, East China University of Technology, Nanchang 330013, China)

Abstract: In this paper, we investigate the prediction in a finite population with the normal inverse-Gamma prior under the squared error loss. First, we obtain Bayes prediction of linear quantities and quadratic quantities based on Bayesian theory, respectively. Second, we compare Bayes prediction with the best linear unbiased prediction of linear quantities according to statistical decision theory, which shows that Bayes prediction is better than the best linear unbiased prediction.

Keywords:Bayes prediction; linear quantities; quadratic quantities; finite populations2010 MR Subject Classification:62C10; 62M20Document code:AArticle ID: 0255-7797(2018)05-0793-11

1 Introduction

Let $\mathscr{P} = \{1, \dots, N\}$ denote a finite population of N identifiable units, where N is known. Associated with the *i*th unit of \mathscr{P} , there are p+1 quantities: $y_i, x_{i1}, \dots, x_{ip}$, where all but y_i are known, $i = 1, \dots, N$. Let $y = (y_1, \dots, y_N)'$ and $X = (X_1, \dots, X_N)'$, where $X_i = (x_{i1}, \dots, x_{ip})', i = 1, \dots, N$. Relating the two sets of variables, we consider the linear model

$$y = X\beta + \varepsilon, \varepsilon \sim N_N(0, \sigma^2 V), \tag{1.1}$$

where β is a $p \times 1$ unknown parameter vector, V is a known symmetric positive definite matrix, but the parameter $\sigma^2 > 0$ is unknown.

For the superpopulation model (1.1), it is interesting to study the optimal prediction of the population quantity $\theta(y)$ such as the population Total $T = \sum_{i=1}^{N} y_i$, the population variance $S_y^2 = \sum_{i=1}^{N} (y_i - \bar{y}_N)^2 / N$, where $\bar{y}_N = T/N$ is the population mean and the finite population regression coefficient $\beta_N = (X'V^{-1}X)^{-1}X'V^{-1}y$, and so on. In the literature, a lot of predictions for the population quantities were produced. For example, Bolfarine and Rodrigues [1] gave the simple projection predictor, and obtained necessary and sufficient conditions for

Received date: 2016-11-13 Accepted date: 2017-11-07

Foundation item: Supported by the National Natural Science Foundation(11661003); the Natural Science Foundation of Jiangxi Province (20161BAB201033); Science and Technology Project of Education Department of Jiangxi Province (GJJ150582; GJJ160559).

Biography: Hu Guikai (1977–), male, born at Ji'an, Jiangxi, professor, major in the estimate of parameters and optimal predictor.

it to be optimal. Bolfarine et al. [2] studied the best unbiased prediction of finite population regression coefficient under the generalized prediction mean squared error in different kinds of models. Xu et al. [3] obtained a kind of optimal prediction of linear predictable function, and got the necessary and sufficient conditions for any linear prediction to be optimal under matrix loss. Xu and Yu [4] further gave the admissible prediction in superpopulation models with random regression coefficients under matrix loss function. Hu and Peng [5] obtained some conditions for linear prediction to be admissible in superpopulation models with and without the assumption that the underlying distribution is normal, respectively. Furthermore, Hu et al. [6–7] discussed the linear minimax prediction in the multivariate normal populations and Gauss-Markov populations, respectively. Their results showed that linear minimax prediction for finite population regression coefficient is admissible in some conditions. Bolfarine and Zacks [8] studied Bayes and minimax prediction under square error loss function in a finite population with single parametric prior. Meanwhile, Bansal and Aggarwal [9–11] considered Bayes prediction of finite population regression coefficient using a balanced loss function under the same prior information. There are two characteristics in the above studies.

On the one hand, they obtained the optimal, linear admissible and minimax predictions based on statistical decision theory. It is well known that statistical decision theory only consider the sample information and loss function and do not consider the prior information. However, people usually have these information.

On the other hand, they discussed the Bayes prediction by considering the prior information of single parameter, and did not consider the situation of multi-parameters. In other words, they only made use of the prior information of regression coefficient, but not use the prior information of error variance in model (1.1). In fact, multi-parameter situations are often encountered in the practical problems. Therefore, in this paper, we will study Bayes prediction of linear and quadratic quantities in a finite population where regression coefficient and error variance have the normal inverse-Gamma prior.

Assume that the prior distribution of β and σ^2 is normal inverse-Gamma distribution, that is,

$$\beta \mid \sigma^2 \sim N_p(\mu, \frac{\sigma^2}{k} I_p), \sigma^2 \sim \Gamma^{-1}(\frac{\alpha}{2}, \frac{\lambda}{2}), \tag{1.2}$$

where μ is a known $p \times 1$ vector, α and λ are known constants, k^{-1} is a ratio between the prior variance of β and sample variance of model (1.1). We can suppose that k^{-1} is known by experience or professional knowledge. Therefore, the joint prior distribution of (β, σ^2) is

$$\pi(\beta, \sigma^2) = p_1(\beta | \sigma^2) p_2(\sigma^2) = M_1(\sigma^2)^{-(\frac{p+\alpha}{2}+1)} \exp\{-\frac{1}{2\sigma^2} [k(\beta-\mu)'(\beta-\mu)+\lambda]\},$$
(1.3)

where $M_1 = (\frac{k}{2\pi})^{\frac{p}{2}} (\frac{\lambda}{2})^{\frac{\alpha}{2}} [\Gamma(\frac{\alpha}{2})]^{-1}$. The Bayes model defined by (1.1) and (1.2) is designated by (1.4). In order to obtain Bayes prediction in the Bayes model (1.4), a sample \mathscr{S} of size n is selected from \mathscr{P} according to some specified sampling plan. Let $\mathscr{R} = \mathscr{P} - \mathscr{S}$ be the unobserved part of \mathscr{P} of size N - n. After the sample \mathscr{S} has been selected, we may reorder the elements of y such that we have the corresponding partitions of y, X and V, that is

$$y = \begin{pmatrix} y_s \\ y_r \end{pmatrix}, X = \begin{pmatrix} X_s \\ X_r \end{pmatrix}, V = \begin{pmatrix} V_s & V_{sr} \\ V_{rs} & V_r \end{pmatrix},$$

where X and X_s are known column full rank matrices.

The rest of this paper is organized as follows: in Section 2, we give Bayes predictor of population quantities in the Bayes model (1.4). Section 3 is devoted to discuss Bayesian prediction of linear quantities. In Section 4, we obtain Bayes prediction of quadratic quantities. Some examples are given in Section 5. Concluding remarks are placed in Section 6.

2 Bayes Prediction of Population Quantities

In this section, we will discuss the Bayes prediction of population quantities. Let $L(\hat{\theta}(y_s), \theta(y))$ be a loss function for predicting $\theta(y)$ by $\hat{\theta}(y_s)$. The corresponding Bayes prediction risk of $\hat{\theta}(y_s)$ in model (1.4) is defined as $\rho(\hat{\theta}(y_s), \theta(y)) = E_y[L(\hat{\theta}(y_s), \theta(y))]$, where the expectation operator E_y is performed with respect to the joint distribution of y and (β, σ^2) . The Bayes predictor is the one minimizing the Bayes prediction risk $\rho(\hat{\theta}(y_s), \theta(y))$. In particular, when we consider the squared error loss, then the Bayes prediction of $\theta(y)$ is

$$\hat{\theta}(y_s) = E_y[\theta(y)|y_s], \qquad (2.1)$$

and the Bayes prediction risk is

$$\varrho(\theta(y_s), \theta(y)) = E_{y_s} \{ \operatorname{Var}[\theta(y)|y_s] \}, \qquad (2.2)$$

where the expectation operator E_{y_s} is performed with respect to the joint distribution of y_s and (β, σ^2) . It is noted that $y_s | \beta, \sigma^2 \sim N_n(X_s\beta, \sigma^2 V_s)$ and

$$y_r - V_{rs}V_s^{-1}y_s | \beta, \sigma^2 \sim N_{N-n}((X_r - V_{rs}V_s^{-1}X_s)\beta, \sigma^2(V_r - V_{rs}V_s^{-1}V_{sr})).$$

This together with eq. (1.3) will yield the following results.

Theorem 2.1 Under the Bayes model (1.4), the following results hold.

(i) The joint posterior probability density of (β, σ^2) is

$$\pi(\beta, \sigma^2 | y_s) = M_2 |\Sigma|^{-\frac{1}{2}} (\sigma^2)^{-\frac{n+p+\alpha}{2}+1} \exp\{-\frac{1}{2\sigma^2} [c_0 + (\beta - \tilde{\beta}_s)' \Sigma^{-1} (\beta - \tilde{\beta}_s)]\}.$$
 (2.3)

(ii) The marginal posterior distribution of β is *p*-dimensional *t* distribution $MT_p(\hat{\beta}_s, \frac{c_0\Sigma}{n+\alpha}, n+\alpha)$ with probability density

$$\pi(\beta|y_s) = M_3 |\frac{c_0 \Sigma}{n+\alpha}|^{-\frac{1}{2}} [1 + \frac{1}{n+\alpha}(\beta - \tilde{\beta}_s)'(\frac{c_0 \Sigma}{n+\alpha})^{-1}(\beta - \tilde{\beta}_s)]^{-\frac{n+\alpha+p}{2}}.$$

(iii) The marginal posterior distribution of σ^2 is $\Gamma^{-1}(\frac{n+\alpha}{2}, \frac{c_0}{2})$ with probability density

$$\pi(\sigma^2 | y_s) = M_4(\sigma^2)^{-\frac{n+\alpha}{2}+1} \exp(-\frac{c_0}{2\sigma^2}).$$

(iv) Bayes prediction distribution of y_r given y_s is N - n dimensional t distribution $MT_{N-n}(\hat{y_r}, \frac{c_0 U}{n+\alpha}, n+\alpha)$ with probability density

$$\pi(y_r|y_s) = M_5 \left| \frac{c_0 U}{n+\alpha} \right|^{-\frac{1}{2}} \left[1 + \frac{1}{n+\alpha} (y_r - \hat{y_r})' (\frac{c_0 U}{n+\alpha})^{-1} (y_r - \hat{y_r}) \right]^{-\frac{N+r}{2}},$$

where

$$\begin{split} \hat{\beta}_{s} &= (X_{s}'V_{s}^{-1}X_{s})^{-1}X_{s}'V_{s}^{-1}y_{s}, \hat{\sigma}^{2} = \frac{(y_{s} - X_{s}\hat{\beta}_{s})'V_{s}^{-1}(y_{s} - X_{s}\hat{\beta}_{s})}{n - p}, \\ c_{0} &= (\hat{\beta}_{s} - \mu)'A(\hat{\beta}_{s} - \mu) + (n - p)\hat{\sigma}^{2} + \lambda, \tilde{\beta}_{s} = \Sigma[k\mu + (X_{s}'V_{s}^{-1}X_{s})\hat{\beta}_{s}], \\ A &= [(X_{s}'V_{s}^{-1}X_{s})^{-1} + k^{-1}I_{p}]^{-1}, \\ \Sigma &= (X_{s}'V_{s}^{-1}X_{s} + kI_{p})^{-1}, M_{2} = (2\pi)^{-\frac{p}{2}}M_{4}, M_{3} = \frac{\Gamma(\frac{n + \alpha + p}{2})}{\Gamma(\frac{n + \alpha}{2})}[\pi(n + \alpha)]^{-\frac{p}{2}}, \\ M_{4} &= \frac{\left(\frac{c_{0}}{2}\right)^{n + \alpha}}{\Gamma(\frac{n + \alpha}{2})}, M_{5} = \frac{\Gamma(\frac{N + \alpha}{2})}{\Gamma(\frac{n + \alpha}{2})}[\pi(n + \alpha)]^{-\frac{N + \alpha}{2}}, \\ \tilde{y}_{r} &= X_{r}\tilde{\beta}_{s} + V_{rs}V_{s}^{-1}(y_{s} - X_{s}\tilde{\beta}_{s}), \\ U &= V_{r} - V_{rs}V_{s}^{-1}V_{sr} + (X_{r} - V_{rs}V_{s}^{-1}X_{s})\Sigma(X_{r} - V_{rs}V_{s}^{-1}X_{s})'. \end{split}$$

Proof The proof of (i): since

$$(y_s - X_s \beta)' V_s^{-1} (y_s - X_s \beta)$$

= $(y_s - X_s \hat{\beta}_s)' V_s^{-1} (y_s - X_s \hat{\beta}_s) + (X_s \hat{\beta}_s - X_s \beta)' V_s^{-1} (X_s \hat{\beta}_s - X_s \beta)$
= $(n - p)\hat{\sigma}^2 + (\hat{\beta}_s - \beta)' X_s' V_s^{-1} X_s (\hat{\beta}_s - \beta)$

and $y_s|\beta, \sigma^2 \sim N_n(X_s\beta, \sigma^2 V_s)$, the conditional probability density of y_s given (β, σ^2) is

$$p_3(y_s|\beta,\sigma^2) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\{-\frac{(n-p)\hat{\sigma}^2 + (\hat{\beta}_s - \beta)'X_s'V_s^{-1}X_s(\hat{\beta}_s - \beta)}{2\sigma^2}\}.$$

This together with eq. (1.3) will yield that the joint posterior probability density of (β, σ^2) is

$$\begin{aligned} \pi(\beta, \sigma^{2}|y_{s}) &= \frac{p_{3}(y_{s}|\beta, \sigma^{2})\pi(\beta, \sigma^{2})}{m(y_{s})} \\ &\propto p_{3}(y_{s}|\beta, \sigma^{2})\pi(\beta, \sigma^{2}) \\ &\propto (\sigma^{2})^{-(\frac{n+\alpha+p}{2}+1)}\exp\{-\frac{1}{2\sigma^{2}}[(n-p)\hat{\sigma}^{2} + (\beta - \hat{\beta}_{s})'X_{s}'V_{s}^{-1}X_{s}(\beta - \hat{\beta}_{s})]\} \\ &\exp\{-\frac{1}{2\sigma^{2}}[k(\beta - \mu)'(\beta - \mu) + \lambda]\} \\ &\propto (\sigma^{2})^{-(\frac{n+\alpha+p}{2}+1)}\exp\{-\frac{1}{2\sigma^{2}}[c_{0} + (\beta - \tilde{\beta}_{s})'\Sigma^{-1}(\beta - \tilde{\beta}_{s})]\}, \end{aligned}$$

where $m(y_s)$ is the marginal probability density of y_s , symbol \propto denotes proportional to. By adding the regularization constant $M_2|\Sigma|^{-\frac{1}{2}}$ to eq. (2.3), we obtain result (i). The proof of (ii): by the integral of eq. (2.2) about σ^2 , we have

$$\begin{aligned} \pi(\beta|y_s) &= \int_0^{+\infty} \pi(\beta, \sigma^2|y_s) d\sigma^2 \\ &= M_2 |\Sigma|^{-\frac{1}{2}} \int_0^{+\infty} (\sigma^2)^{-(\frac{n+p+\alpha}{2}+1)} \exp\{-\frac{1}{2\sigma^2} [c_0 + (\beta - \tilde{\beta})'\Sigma^{-1}(\beta - \tilde{\beta})]\} d\sigma^2 \\ &= M_3 |\frac{c_0 \Sigma}{n+\alpha}|^{-\frac{1}{2}} [1 + \frac{1}{n+\alpha} (\beta - \tilde{\beta}_s)'(\frac{c_0 \Sigma}{n+\alpha})^{-1} (\beta - \tilde{\beta}_s)]^{-\frac{n+\alpha+p}{2}}, \end{aligned}$$

which implies that the marginal posterior distribution of β is *p*-dimensional *t* distribution with mean vector $\tilde{\beta}$, correlation matrix $\frac{c_0 \Sigma}{n+\alpha}$ and degrees of freedom $n + \alpha$.

The proof of (iii): by the integral of eq. (2.2) about β , we can obtain the result. Here it is omitted.

The proof of (iv): by $y_s|\beta, \sigma^2 \sim N_n(X_s\beta, \sigma^2V_s), y_r|\beta, \sigma^2, y_s \sim N_{N-n}(X_r\beta + V_{rs}V_s^{-1}(y_s - X_s\beta), \sigma^2(V_r - V_{rs}V_s^{-1}V_{sr}))$, and eq. (2.2), we know that

$$\begin{aligned} &\pi(y_r,\beta,\sigma^2|y_s) \propto p_3(y_s|\beta,\sigma^2)\pi(\beta,\sigma^2)p_4(y_r|\beta,\sigma^2,y_s) \\ &\propto \quad (\sigma^2)^{-(\frac{N+p+\alpha}{2}+1)} \exp\{-\frac{1}{2\sigma^2}[c_0+(\beta-\tilde{\beta})'\Sigma^{-1}(\beta-\tilde{\beta})]\} \\ &\times \quad \exp\{-\frac{1}{2\sigma^2}[y_r-X_r\beta-V_{rs}V_s^{-1}(y_s-X_s\beta)]'(V_r-V_{rs}V_s^{-1}V_{sr})^{-1} \\ & [y_r-X_r\beta-V_{rs}V_s^{-1}(y_s-X_s\beta)]\} \\ &\propto \quad (\sigma^2)^{-(\frac{N+p+\alpha}{2}+1)} \exp\{-\frac{1}{2\sigma^2}[(\beta-D^{-1}\beta^*)'D(\beta-D^{-1}\beta^*)+c_0+(y_r-\tilde{y_r})'U(y_r-\tilde{y_r})]\}, \end{aligned}$$

where

$$\beta^* = \Sigma^{-1} \tilde{\beta}_s + (X_r - V_{rs} V_s^{-1} X_s)' (V_r - V_{rs} V_s^{-1} V_{sr})^{-1} (y_r - V_{rs} V_s^{-1} y_s),$$

$$D = \Sigma^{-1} + (X_r - V_{rs} V_s^{-1} X_s)' (V_r - V_{rs} V_s^{-1} V_{sr})^{-1} (X_r - V_{rs} V_s^{-1} X_s).$$

Adding the regularization constant to eq. (2.3) and integrating it by β and σ^2 , respectively, we can obtain the result.

3 Bayes Prediction of Linear Quantities

In order to obtain Bayes prediction of $\theta(y)$, we consider the squared error loss

$$L(\hat{\theta}(y_s), \theta(y)) = [\hat{\theta}(y_s) - \theta(y)]^2, \qquad (3.1)$$

then Bayes prediciton of $\theta(y)$ is

$$\hat{\theta}(y_s) = E_y[\theta(y)|y_s], \qquad (3.2)$$

and Bayes prediction risk is

$$\varrho(\theta(y_s), \theta(y)) = E_y[\theta(y_s) - \theta(y)]^2 = E_{y_s}\{\operatorname{Var}[\theta(y)|y_s]\},$$
(3.3)

where the expectation operator E_{y_s} is performed with respect to the joint distribution of y_s and (β, σ^2) . By result (iv) of Theorem 2.1, we know

$$E_y(y_r|y_s) = X_r \tilde{\beta}_s + V_{rs} V_s^{-1}(y_s - X_s \tilde{\beta}_s)$$
(3.4)

and

$$\operatorname{Var}(y_r|y_s) = \frac{c_0}{n+\alpha-2}U.$$
(3.5)

Now, let $\theta(y) = Qy$ be any linear quantity, where $Q = (Q'_s, Q'_r)$ is a known $1 \times N$ vector. According to Theorem 2.1, eqs. (3.4) and (3.5), we have the following conclusions.

Theorem 3.1 Under model (1.4) and squared error loss function, Bayes predictor of linear quantity Qy is $\tilde{\theta}(y_s) = Q'_s y_s + Q'_r \tilde{y}_r$, and Bayes predictor risk is $\frac{E_{ys}(c_0)}{n+\alpha-2}Q'_r UQ_r$.

As we know, the best linear unbiased prediction of Qy under the squared error loss is $\hat{\theta}(y_s)$, where $\hat{\theta}(y_s) = Q'_s y_s + Q'_r \hat{y}_r$, and $\hat{y}_r = X_r \hat{\beta}_s + V_{rs} V_s^{-1} (y_s - X_s \hat{\beta}_s)$. In the following, we will discuss the superiority between Bayes prediction and the best linear unbiased prediction under the predicative mean squared error (PMSE), which is defined by $PMSE(d(y_s), Qy) = E[(d(y_s) - Qy)^2]$.

Theorem 3.2 Under model (1.4), Bayes prediction $\tilde{\theta}(y_s)$ of Qy is better than the best linear unbiased prediction $\hat{\theta}(y_s)$ under the predicative mean squared error.

Proof By the definition of PMSE and $\tilde{\beta}_s = \hat{\beta}_s - k\Sigma(\hat{\beta}_s - \mu)$, we have

$$\begin{split} & \operatorname{PMSE}(\tilde{\theta}(y_{s}), Qy) \\ = & E[(\tilde{\theta}(y_{s}) - Qy)^{2}] \\ = & E[Q_{r}'(\tilde{y}_{r} - y_{r})(\tilde{y}_{r} - y_{r})'Q_{r}] \\ = & E[Q_{r}'(X_{r}\tilde{\beta}_{s} + V_{rs}V_{s}^{-1}(y_{s} - X_{s}\tilde{\beta}_{s}) - y_{r})(X_{r}\tilde{\beta}_{s} + V_{rs}V_{s}^{-1}(y_{s} - X_{s}\tilde{\beta}_{s}) - y_{r})'Q_{r}] \\ = & Q_{r}'E_{y_{s}}[(X_{r} - V_{rs}V_{s}^{-1}X_{s})(\tilde{\beta}_{s} - \beta)(\tilde{\beta}_{s} - \beta)'(X_{r} - V_{rs}V_{s}^{-1}X_{s})' + \sigma^{2}(V_{r} - V_{rs}V_{s}^{-1}V_{sr})]Q_{r} \\ = & Q_{r}'(X_{r} - V_{rs}V_{s}^{-1}X_{s})E_{y_{s}}[(\hat{\beta}_{s} - \beta) - k\Sigma(\hat{\beta}_{s} - \mu)][(\hat{\beta}_{s} - \beta) - k\Sigma(\hat{\beta}_{s} - \mu)]' \\ & (X_{r} - V_{rs}V_{s}^{-1}X_{s})'Q_{r} + \frac{\lambda}{\alpha - 2}Q_{r}'(V_{r} - V_{rs}V_{s}^{-1}V_{sr})Q_{r} \\ = & \operatorname{PMSE}(\hat{\theta}(y_{s}), Qy) \\ & - \frac{k^{2}\lambda}{\alpha - 2}Q_{r}'(X_{r} - V_{rs}V_{s}^{-1}X_{s})\Sigma(k^{-1}I_{p} + (X_{s}'V_{s}^{-1}X_{s})^{-1})\Sigma'(X_{r} - V_{rs}V_{s}^{-1}X_{s})'Q_{r}. \end{split}$$

That is, $\text{PMSE}(\hat{\theta}(y_s), Qy) - \text{PMSE}(\tilde{\theta}(y_s), Qy) > 0$. Therefore, $\tilde{\theta}(y_s)$ is better than $\hat{\theta}(y_s)$ under the predicative mean squared error.

Corollary 3.1 Bayes predictor of the population total T under model (1.4) and the loss function (3.1) is $\tilde{T}(y_s) = 1'_n y_s + 1'_{N-n} [X_r \tilde{\beta}_s + V_{rs} V_s^{-1} (y_s - X_s \tilde{\beta}_s)]$, and Bayes risk of this predictor is $\frac{E_{y_s}(c_0)}{n+\alpha-2} 1'_{N-n} U 1_{N-n}$. Moreover, $\hat{T}(y_s)$ is dominated by $\tilde{T}(y_s)$ under the predicative mean squared error, where $\hat{T}(y_s) = 1'_n y_s + 1'_{N-n} \hat{y}_r$.

For the finite population regression coefficient $\beta_N = (X'V^{-1}X)^{-1}X'V^{-1}y$, following

Bolfarine et al. [2], we can write it as

$$\begin{aligned} \beta_N &= (X'V^{-1}X)^{-1}X'V^{-1}y \\ &= \left[(X'_s \ X'_r) \left(\begin{array}{cc} V_s & V_{sr} \\ V_{rs} & V_r \end{array} \right)^{-1} \left(\begin{array}{cc} X_s \\ X_r \end{array} \right) \right]^{-1} (X'_s \ X'_r) \left(\begin{array}{cc} V_s & V_{sr} \\ V_{rs} & V_r \end{array} \right)^{-1} \left(\begin{array}{cc} y_s \\ y_r \end{array} \right) \\ &= K_s y_s + K_r y_r, \end{aligned}$$

where

$$\begin{split} K_s &= G^{-1}JC^{-1}, \ K_r = G^{-1}FE^{-1}, \ J = X'_s - X'_rV_r^{-1}V_{rs}, \\ C &= V_s - V_{sr}V_r^{-1}V_{rs}, \ F = X'_r - X'_sV_s^{-1}V_{sr}, \ E = V_r - V_{rs}V_s^{-1}V_{sr}, \end{split}$$

and

$$G = JC^{-1}X_s + FE^{-1}X_r.$$

Then by Theorem 3.1, we have the following corollary.

Corollary 3.2 Bayes predictor of the population total β_N under model (1.4) and the loss function (3.1) is $\tilde{\beta}_N(y_s) = K_s y_s + K_r E(y_r|y_s)$, and Bayes risk of this predictor is $\frac{E_{ys}(c_0)}{n+\alpha-2}K_r UK'_r$. Moreover, it is better than $\hat{\beta}_N(y_s)$ under the predicative mean squared error, where $\hat{\beta}_N(y_s) = K_s y_s + K_r \hat{y}_r$.

In order to illustrate our results, we give the following example.

Example 3.1 Let $X = (x_1, x_2, \dots, x_N)'$, $V = \text{diag}(x_1, x_2, \dots, x_N)$ in the Bayesian model (1.4), where $x_i \neq 0, i = 1, 2, \dots, N$. If $X_s = (x_1, x_2, \dots, x_n)', y_s = (y_1, y_2, \dots, y_n)'$, we have $\tilde{\beta}_s = \frac{1}{\sum\limits_{i=1}^n x_i + k} (k\mu + 1'_n y_s), \ \hat{\beta}_s = \frac{1}{\sum\limits_{i=1}^n x_i} 1'_n y_s$. According to Theorem 3.1, we have the following conclusions.

(i)
$$\tilde{T}(y_s) = 1'_n y_s + \frac{\sum\limits_{i=n+1}^N x_i}{\sum\limits_{i=1}^n x_i+k} (k\mu + 1'_n y_s)$$
. Its Bayes prediction risk is $\frac{\lambda(\sum\limits_{i=1}^N x_i+k)}{(\alpha-2)(\sum\limits_{i=1}^n x_i+k)}$

Moreover, $T(y_s)$ is better than $T(y_s)$.

(ii)
$$\tilde{\beta}_N(y_s) = \frac{1}{\sum\limits_{i=1}^N x_i} \hat{T}(y_s)$$
, and its Bayes prediction risk is $\frac{\lambda(\sum\limits_{i=1}^N x_i+k)\sum\limits_{i=n+1}^N x_i}{(\alpha-2)(\sum\limits_{i=1}^n x_i+k)(\sum\limits_{i=1}^N x_i)^2}$. Moreover,

 $\beta_N(y_s)$ is better than $\beta_N(y_s)$.

In the following, we continue to give the simulation study to explain our results according to the following steps, which are executed on a personal computer using Version 7.9 (R2009b) Matlab software.

(i) Generating randomly a $N \times p$ full column rank matrix X and a *p*-dimensional vector μ ;

(ii) The number σ^2 and random error ε are generated from distribution $\Gamma^{-1}(\frac{\alpha}{2}, \frac{\lambda}{2})$ and $N(0, \sigma^2 V)$, respectively;

- (iii) Generating a *p*-dimensional vector β by the distribution $N(\mu, \frac{\sigma^2}{h}I_p)$;
- (iv) Obtaining the dependent variable y by the model $y = X\beta + \varepsilon$.

Vol. 38

(v) Generating randomly N-dimensional vector Q, then Bayes prediction and the best linear unbiased prediction of Qy are derived by Theorem 3.1, respectively.

(vi) Finally, we compare the PMSE between Bayes prediction and best linear unbiased prediction.

Now, we assume that $N = 10, n = 6, p = 3, \alpha = 8, \lambda = 12, k = 10$, and obtain the above data. The simulation study shows that Bayes prediction is better than the best linear unbiased prediction, which is consistent to our theoretical conclusions. Here, we give the above data in one experiment as following.

$$\varepsilon = \begin{pmatrix} 0.7079 & -0.6014 & -2.3252 \\ 1.9574 & 0.5512 & -1.2316 \\ 0.5045 & -1.0998 & 1.0556 \\ 1.8645 & 0.0860 & -0.1132 \\ -0.3398 & -2.0046 & 0.3792 \\ -1.1398 & -0.4931 & 0.9442 \\ -0.2111 & 0.4620 & -2.1204 \\ 1.1902 & -0.3210 & -0.6447 \\ -1.1162 & 1.2366 & -0.7043 \\ 0.6353 & -0.6313 & -1.0181 \end{pmatrix}, \quad \beta = \begin{pmatrix} -1.3868 \\ 0.3785 \\ 1.7166 \end{pmatrix},$$

$$\varepsilon = \begin{pmatrix} 0.6693 \\ 0.3681 \\ 0.6319 \\ 0.2148 \\ 0.6315 \\ 0.2148 \\ 0.3615 \\ 0.1250 \\ 0.6272 \\ 0.3941 \\ 0.4814 \\ 0.6497 \end{pmatrix}, \quad y = \begin{pmatrix} -4.5316 \\ -4.2521 \\ 1.3281 \\ -2.5328 \\ 0.7251 \\ 3.1399 \\ -2.5452 \\ -2.4847 \\ 1.2884 \\ -2.2180 \end{pmatrix}.$$

At this time, we get randomly

Q = (0.3139, 0.6382, 0.9866, 0.5029, 0.9477, 0.8280, 0.9176, 0.1131, 0.8121, 0.9083)'.

By direct computation, we have Qy = -4.3971. By Theorem 2.1, we know $\hat{\theta}(y_s) = -4.8497, \hat{\theta}(y_s) = -5.7928$, and $\text{PMSE}(\hat{\theta}(y_s)) - \text{PMSE}(\tilde{\theta}(y_s)) = 0.0844 > 0$. Therefore, Bayes prediction of Qy is better than the best linear unbiased predictor.

4 Bayes Prediction of Quadratic Quantities

In this section, we will discuss Bayes prediction of quadratic quantities f(H) = y'Hy, where H is a known symmetric matrix. Assume that $H = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix}$ with $H_{12} = H'_{21}$,

$$f(H) = (y'_s, y'_r) \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} \begin{pmatrix} y_s \\ y_r \end{pmatrix} = y'_s H_{11} y_s + y'_s H_{12} y_r + y'_r H_{21} y_s + y'_r H_{22} y_r$$

By Theorem 2.1 and eq. (3.2), we have the following results.

Theorem 4.1 Under model (1.4) and the loss function (1.3), the Bayes prediction of f(H) is

$$\hat{f}(H) = y'_{s}H_{11}y_{s} + y'_{s}H_{12}E(y_{r}|y_{s}) + [E(y_{r}|y_{s})]'H_{21}y_{s} + [E(y_{r}|y_{s})]'H_{22}[E(y_{r}|y_{s})] + tr[H_{22}Var(y_{r}|y_{s})].$$

For the population variance S_y^2 , we know that

$$S_{y}^{2} = y' \frac{1}{N} (I_{N} - \frac{1}{N} \mathbf{1}_{N} \mathbf{1}_{N}') y = \frac{1}{N} (y'_{s}, y'_{r}) \begin{pmatrix} I_{n} - \frac{1}{N} \mathbf{1}_{n} \mathbf{1}_{n}' & -\frac{1}{N} \mathbf{1}_{n} \mathbf{1}_{N-n}' \\ -\frac{1}{N} \mathbf{1}_{N-n} \mathbf{1}_{n}' & I_{N-n} - \frac{1}{N} \mathbf{1}_{N-n} \mathbf{1}_{N-n}' \end{pmatrix} \begin{pmatrix} y_{s} \\ y_{r} \end{pmatrix},$$

where 1_n denotes *n* dimensional vector with elements 1. Then by Theorem 4.1, we can obtain the following corollary.

Corollary 4.1 The Bayes prediction of the population variance S_y^2 under model (1.4) and the loss function (3.1) is

$$\hat{S}_{y}^{2} = \frac{1}{N} (y'_{s}, E(y_{r}|y_{s})') \begin{pmatrix} I_{n} - \frac{1}{N} \mathbf{1}_{n} \mathbf{1}'_{n} & -\frac{1}{N} \mathbf{1}_{n} \mathbf{1}'_{N-n} \\ -\frac{1}{N} \mathbf{1}_{N-n} \mathbf{1}'_{n} & I_{N-n} - \frac{1}{N} \mathbf{1}_{N-n} \mathbf{1}'_{N-n} \end{pmatrix} \begin{pmatrix} y_{s} \\ E(y_{r}|y_{s}) \end{pmatrix} \\ + \frac{1}{N} \operatorname{tr}[(I_{N-n} - \frac{1}{N} \mathbf{1}_{N-n} \mathbf{1}'_{N-n}) \operatorname{Var}(y_{r}|y_{s})].$$

It is noted that $S_y^2 = \frac{n}{N}S_{y_s}^2 + (1 - \frac{n}{N})[S_{y_r}^2 + \frac{n}{N}(\bar{y}_s - \bar{y}_r)^2]$, where \bar{y}_s and $S_{y_s}^2$ are the mean and variance of y_s , \bar{y}_r and $S_{y_r}^2$ are the mean and variance of y_r . Therefore, the Bayes prediction of the population variance can also be expressed as follows.

Remark 4.1 The Bayes prediction of the population variance S_y^2 under model (1.4) and the loss function (3.1) is

$$\hat{S}_{y}^{2} = \frac{n}{N} S_{y_{s}}^{2} + (1 - \frac{n}{N}) \left\{ \operatorname{tr}\left[\frac{1}{N - n} (I_{N - n} - \frac{1}{N - n} \mathbf{1}_{N - n} \mathbf{1}'_{N - n}) \operatorname{Var}(y_{r} | y_{s})\right] \\ + E(y_{r} | y_{s})' \frac{1}{N - n} (I_{N - n} - \frac{1}{N - n} \mathbf{1}_{N - n} \mathbf{1}'_{N - n}) E(y_{r} | y_{s}) \\ + \frac{n}{N} [(\bar{y}_{s} - \frac{1}{N - n} \mathbf{1}'_{N - n} E(y_{r} | y_{s}))^{2} + \frac{1}{(N - n)^{2}} \mathbf{1}'_{N - n} \operatorname{Var}(y_{r} | y_{s}) \mathbf{1}_{N - n}] \right\}$$

Proof Since $S_y^2 = \frac{n}{N}S_{y_s}^2 + (1-\frac{n}{N})[S_{y_r}^2 + \frac{n}{N}(\bar{y}_s - \bar{y}_r)^2]$, we only derive the Bayes prediction of $S_{y_r}^2 + \frac{n}{N}(\bar{y}_s - \bar{y}_r)^2$. Moreover, we know that $S_{y_r}^2 = \frac{1}{N-n}y'_r(I_{N-n} - \frac{1}{N-n}1_{N-n}1'_{N-n})y_r$, and $\bar{y}_r = \frac{1}{N-n}1'_{N-n}y_r$. Therefore, the Bayes prediction of $S_{y_r}^2$ is

$$E(S_{y_r}^2|y_s) = \frac{1}{N-n} \operatorname{tr}[(I_{N-n} - \frac{1}{N-n} \mathbf{1}_{N-n} \mathbf{1}'_{N-n}) \operatorname{Var}(y_r|y_s)] + E(y_r|y_s)' \frac{1}{N-n} (I_{N-n} - \frac{1}{N-n} \mathbf{1}_{N-n} \mathbf{1}'_{N-n}) E(y_r|y_s).$$
(4.1)

And, the Bayes prediction of $(\bar{y}_s - \bar{y}_r)^2$ is

$$E[(\bar{y}_s - \bar{y}_r)^2 | y_s] = (\bar{y}_s - \frac{1}{N-n} \mathbf{1}'_{N-n} E(y_r | y_s))^2 + \frac{1}{(N-n)^2} \mathbf{1}'_{N-n} \operatorname{Var}(y_r | y_s) \mathbf{1}_{N-n}.$$
 (4.2)

According to eqs.(4.1)–(4.2) and the expression of S_y^2 , we can derive the result of this remark. It is easy to verify that the result of this remark is consistent to Corollary 4.1.

Example 4.1 Let $X = 1_N, V = (1 - \rho)I_N + \rho 1_N 1'_N$ in the Bayesian model (1.4), where $\rho \in (0, 1)$ is known. It can be checked that $X'_s V_s^{-1} X_s = \frac{n}{1 + (n-1)\rho}$, and

$$\tilde{\beta}_s = \left(\frac{n}{1+(n-1)\rho} + k\right)^{-1} \left[k\mu + \frac{n}{1+(n-1)\rho}\bar{y}_s\right],\\c_0 = \frac{nk}{k+k\rho(n-1)+n} \left(\frac{1}{n}1'_n y_s - \mu\right)^2 + \lambda + \frac{1}{1-\rho}y'_s (I_n - \frac{1}{n}1_n 1'_n)y_s.$$

Then,

$$E(y_r|y_s) = a1_{N-n}$$
 and $\operatorname{Var}(y_r|y_s) = \frac{c_0}{n+\alpha-2}[(1-\rho)I_{N-n} + b1_{N-n}1'_{N-n}]$

where $a = \frac{1-\rho}{1+(n-1)\rho}\tilde{\beta}_s + \frac{n\rho}{1+(n-1)\rho}\bar{y}_s$, $b = \frac{1-\rho}{1+(n-1)\rho}[\rho + \frac{1-\rho}{n+k+(n-1)k\rho}]$. According to Remark 4.1, we know that

$$\hat{S}_{y}^{2} = \frac{n}{N}S_{y_{s}}^{2} + \frac{(1-\rho)(N-n-1)c_{0}}{N(n+\alpha-2)} + \frac{(N-n)n}{N^{2}}[(\bar{y}_{s}-a)^{2} + \frac{c_{0}}{n+\alpha-2}(b+\frac{1-\rho}{N-n})].$$

5 Concluding Remarks

In this paper, we obtain Bayes prediction of linear and quadratic quantities in the finite population with normal inverse-Gamma prior information. In our studies, on the one hand, the distribution of the superpopulation model is need to be normal. However, in many occasions, the distribution of the model is usually unknown in addition to the mean vector and covariance matrix. At this time, how to deal with the Bayes prediction? On the other hand, if the prior distribution is hierarchical and improper, how to obtain the generalized Bayes prediction and discuss its optimal properties? Such as these problems are deserved to discuss in the future.

References

- Bolfarine H, Rodrigues J. On the simple projection predictor in finite populations[J]. Aust. J. Stat., 1988, 30: 338–341.
- [2] Bolfarine H, Zacks S, Elian S N, Rodrigues J. Optimal prediction of the finite population regression coefficient[J]. Sankhyā. Ser. B., 1994, 56: 1–10.
- [3] Xu L W, Lu M, Jiang C F. Optimal prediction in finite populations under matrix loss[J]. J. Stat. Plan. Infer., 2011, 141(8): 2503–2512.

- [4] Xu L W, Yu S H. Admissible prediction in superpopulation models with random regression coefficients under matrix loss function[J]. J. Multiv. Anal., 2012, 103(1): 68–76.
- [5] Hu G K, Peng P. Linear admissible predictor of finite population regression coefficient under a balanced loss function [J]. J. Math., 2014, 34: 820–828.
- [6] Hu G K, Li Q G, Yu S H. Optimal and minimax prediction in multivariate normal populations under a balanced loss function[J]. J. Multiv. Anal., 2014, 128: 154–164.
- [7] Hu G K, Li Q G, Yu S H. Linear minimax prediction of finite population regression coefficient under a balanced loss function[J]. Commun. Stat-The. M., 2016, 45(24): 7197–7209.
- [8] Bolfarine H, Zacks S. Bayes and minimax prediction in finite populations[J]. J. Stat. Plan. Infer., 1991, 28: 139–151.
- Bansal A K, Aggarwal P. Bayes prediction for a heteroscedastic regression surperpopulation model using balanced loss function[J]. Commun. Stat-The. M., 2007,36: 1565–1575.
- [10] Bansal A K, Aggarwal P. Bayes prediction of the regression coefficient in a finite population using balanced loss function [J]. Metron., 2009, 67: 1–16.
- [11] Bansal A K, Aggarwal P. Bayes prediction for a stratified regression superpopulation model using balanced loss function[J]. Commun. Stat-The. M., 2010, 39: 2789–2799.

有限总体中总体数量的贝叶斯预测

胡桂开, 熊鹏飞, 王同心

(东华理工大学理学院,江西南昌 330013)

摘要: 本文在误差平方损失下研究了具有正态逆伽马先验信息有限总体中的预测问题.首先,基于贝 叶斯思想分别获得了线性数量和二次型数量的贝叶斯预测;其次,利用统计决策理论对线性数量的贝叶斯预 测和最佳线性无偏预测进行了比较.结果表明在预测均方误差下线性数量的贝叶斯预测一致优于最佳线性 无偏预测.

关键词: 贝叶斯预测;线性数量;二次型数量;有限总体 MR(2010)主题分类号: 62C10;62M20 中图分类号: O212.8