THE COORDINATE-FREE PREDICTION IN FINITE POPULATIONS WITH CORRELATED OBSERVATIONS

by

Silvia N. Elian
and
Josemar Rodrigues

THE COORDINATE-FREE PREDICTION IN FINITE POPULATIONS WITH CORRELATED OBSERVATIONS

Silvia N. Elian[1] and Josemar Rodrigues[2]

[1] Department of Statistics,
Institute of Mathematics and Statistics,
University of São Paulo,
C. P. 66281 – São Paulo, SP, Brazil
E-mail: selian@ime.usp.br

[2] Department of Statistics,
University Federal of São Carlos,
C. P. 676 – São Carlos, SP, Brazil
E-mail: vjosemar@power.ufscar.br

ABSTRACT

In this paper we got the best linear unbiased predictor of any linear function of the elements of a finite population under coordinate-free models. The optimal predictor of these quantities was obtained in an earlier work considering models with a known diagonal covariance matrix. We extended this result assuming any known covariance matrix.

It is shown that in the particular case of the coordinatized models, this general predictor coincides with the optimal predictor of the total population under a regression superpopulation model with correlated observations.

Key Words: Coordinate-Free Models, Best Linear Unbiased Predictor, Covariance Matrix, Orthogonal Projection.

1. INTRODUCTION

A coordinate-free approach in finite populations was introduced by Rodrigues (1989) as an alternative to the Gauss-Markov set up, used with the purpose of predicting linear functions. The Gauss-Markov approach is characterized by a dependence on a particular basis matrix, but in the coordinate-free language we need only to describe a parametric subspace of $\mathbb{R}^N$, where N is the size of the finite population. Coordinate-free models in the linear models context are discussed by Arnold (1980) and Drygas (1970).

In a finite population $P = \{1, 2, \ldots, N\}$, where N is the known population size, let y_i, $i = 1, 2, \ldots, N$ be the value of a random variable y associated to each population unit. Under the superpopulation approach, we will assume that Y is a random vector such that $Y \in Q$, where Q is an N-dimensional real vector space with the usual inner product.

The superpopulation model is expressed by

$$E(Y) = \mu \in \Omega$$

$$\text{Var}(Y) = \sigma^2 V, \tag{1.1}$$

where Ω is a p-dimensional subspace of Q, σ^2 is a unknown positive parameter and V is a known positive definite matrix.

The considered model is coordinate free, in the sense that no basis is defined for Ω, the parametric space of μ.

Our main objective is predicting $\ell'Y$, a linear combination of the elements of Y. With this purpose, a sample of n observations is drawn of the population and the values of y_i in Y become known for the sample elements. Let s and r be the sets of sample and non sample elements, respectively, such that $P = s \cup r$.

2

We will consider, without loss of generality that Y and V are reordered as

$$Y = \begin{bmatrix} Y_s \\ Y_r \end{bmatrix} \quad \text{and} \quad V = \begin{bmatrix} V_s & V_{sr} \\ V_{rs} & V_r \end{bmatrix},$$

with Y_s containing the n observed sample elements, Y_r, containing the unobserved elements, $V_s = \text{Var}(Y_s)$, $V_r = \text{Var}(Y_r)$ and $V_{sr} = \text{Cov}(Y_s, Y_r)$ is the covariance matrix.

Under a less general model, with $\text{Var}(Y) = \sigma^2 D$, D a known diagonal matrix, Rodrigues (1989) presented the optimal linear predictor of $\ell'Y$. In the next section, we extended the result, obtaining the best linear unbiased predictor of $\ell'Y$ in the model (1.1) and this is the main contribution of the paper. In Section 3, we show that under the coordinatized model, this predictor coincides with that given by Royall (1976). Finally, we conclude the paper with some examples in Section 4.

2. BEST LINEAR UNBIASED PREDICTOR OF LINEAR FUNCTIONS

The linear function $\theta = \ell'Y$ to be predicted may be written as

$$\theta = \ell'Y = \ell'I_s Y + \ell'(I - I_s)Y,$$

where $I_s = \text{diag}(i_1, i_2, \ldots, i_N)$ is a diagonal matrix with its kth diagonal element i_k, where $i_k = 1$ if $k \in s$ and $i_k = 0$ if $k \in r$, $s = \{1, 2, \ldots, n\}$, $r = \{n+1, n+2, \ldots, N\}$.

We note that with this notation, $\ell'I_s Y$ corresponds to the linear combination of the components of Y in the sample and $\ell'(I - I_s)Y$ is the combination of the unobserved elements.

Before stating the predicting results, it is necessary to introduce some definitions and preliminary results.

3

Let

$$\Omega_s = \{\mu_s | \mu_s = I_s\mu, \mu \in \Omega\} \text{ and}$$

$$\Omega_r = \{\mu_r | \mu_r = (I - I_s)\mu, \mu \in \Omega\}.$$

Since after the sample is observed, I_sY will be known, we restrict our attention to linear predictors of $\ell'Y$ in the form

$$\hat{\theta} = \ell' I_s Y + b' I_s Y,$$

where b is a N-dimensional vector.

Definition. A linear predictor $\hat{\theta}$ of θ is unbiased if and only if

$$E_\mu(\hat{\theta} - \theta) = 0,$$

for every $\mu \in \Omega$.

The class of all linear unbiased predictors of $\ell'Y$ will be denoted by U_ℓ.

Finally, next definition states the concept of optimality of the linear predictor of θ.

Definition. The linear predictor $\hat{\theta}_0$ is the best linear unbiased predictor of θ or the optimal linear predictor of θ if $\hat{\theta}_0 \in U_\ell$ and

$$E_\mu(\hat{\theta}_0 - \theta)^2 \le E_\mu(\hat{\theta} - \theta)^2,$$

for every $\mu \in \Omega$ and every $\hat{\theta} \in U_\ell$.

The value of $E_\mu(\hat{\theta}_0 - \theta)^2$ corresponds to the mean-squared error of the predictor $\hat{\theta}_0$.

4

The optimal linear predictor of θ under the model

$$E(Y) = \mu \in \Omega$$

$$\text{Var}(Y) = \sigma^2 D,$$

where D is a known diagonal matrix and σ^2 is unknown was obtained by Rodrigues (1989). It was shown that if $\dim(\Omega) = \dim(\Omega_s)$, where $\dim(\Omega)$ is the dimension of the linear space Ω, then the best linear unbiased predictor of $\theta = \ell'Y$ is given by

$$\hat{\theta}_* = \ell' I_s Y + \ell' \hat{\mu}_*$$

where $\hat{\mu}_* = \begin{bmatrix} 0 \\ \hat{\mu}_r \end{bmatrix}$, 0 is a null vector of dimension n, $\hat{\mu}_r$ is such that

$$\hat{\mu}_r = (I - I_s) P_\Omega (Y_s + \hat{\mu}_r)$$

and P_Ω is the orthogonal projector onto Ω.

Returning to the model (1.1), with a non diagonal covariance matrix V, let us consider the decomposition $V = P \cdot P'$, with P a lower triangular matrix. As shown by Graybill (1976, Theorem 7.2.1) there is a unique lower triangular matrix P such that $V = PP'$. In addition, P is nonsingular. Then, we define the random vector $Z = P^{-1}Y$ and, as a consequence,

$$\text{Var}(Z) = P^{-1} V P^{-1'} \sigma^2 = P^{-1} P P' P^{-1'} \sigma^2 = P' P'^{-1} \sigma^2 = \sigma^2 I.$$

Next theorem presents the best linear unbiased predictor of $\ell'Y$ under model (1.1).

Theorem 1. In the model (1.1)

$$E(Y) = \mu \in \Omega$$

$$\text{Var}(Y) = \sigma^2 V,$$

V a known positive definite matrix, the optimal linear predictor of any linear function of Y, $h'Y$, is

$$h'I_s Y + h'\hat{\mu}_Y \tag{2.1}$$

where $\hat{\mu}_Y = \begin{bmatrix} 0 \\ \hat{Y}_r \end{bmatrix}$, 0 is the null vector of dimension n, $\hat{Y}_r$ is the solution in Y_r of
the system of linear equations

$$(I - I_s)P^{-1}Y = (I - I_s)P^{-1}P_\Omega Y,$$

and P_Ω is the orthogonal projection matrix onto Ω.

Proof. Let $Z = P^{-1}Y = \begin{bmatrix} Z_s \\ Z_r \end{bmatrix}$ with P the lower triangular matrix such that
$V = PP'$,

$$\Omega^* = \{\mu^* | \mu^* = P^{-1}\mu, \mu \in \Omega\},$$
$$\Gamma = h'PI_s Z + h'P\hat{\mu}_Z,$$

where $\hat{\mu}_Z = \begin{bmatrix} 0 \\ \hat{Z}_r \end{bmatrix}$, 0 is the null vector of dimension n and $\hat{Z}_r$ the solution in Z_r of
the system of linear equations

$$(I - I_s)Z = (I - I_s)P_{\Omega^*}(Z).$$

We note that Γ does not depend on unknown quantities because, as it will be
shown in the appendix, $h'PI_s Z$ and $h'P\hat{\mu}_Z$ do not depend on unknown quantities.

Since

$$E(Z) = P^{-1}E(Y) = P^{-1}\mu = \mu^* \in \Omega^*$$

and

$$\text{Var}(Z) = \sigma^2 I,$$

by Rodrigues's result, the optimal linear predictor of $\ell'Z$ is

$$\ell'I_s Z + \ell'\hat{\mu}_Z$$

6

with $\hat{\mu}_Z = \begin{bmatrix} 0 \\ \hat{Z}_r \end{bmatrix}$, where 0 is the null vector of dimension n and $\hat{Z}_r$ is the solution of the system of linear equations

$$(I - I_s)Z = (I - I_s)P_{\Omega^*}Z.$$

Taking $\ell' = h'P$, this predictor reduces to Γ and $\ell'Z = h'PP^{-1}Y = h'Y$. So, we have just proved that Γ is the optimal linear predictor of $h'Y$.

To finish the proof, it is enough to show that $\Gamma = h'I_sY + h'\hat{\mu}_Y$. For this purpose we write some of matrices already defined in the partitioned form as

$$P = \begin{bmatrix} P_1 & 0 \\ P_3 & P_4 \end{bmatrix}, \quad P^{-1} = \begin{bmatrix} C & 0 \\ B_1 & B_2 \end{bmatrix},$$

$$P_{\Omega^*} = \begin{bmatrix} H_1 & H_2 \\ A_1 & A_2 \end{bmatrix}, \quad I_s = \begin{bmatrix} I_n & 0 \\ 0 & 0 \end{bmatrix}, \quad I - I_s = \begin{bmatrix} 0 & 0 \\ 0 & I_{N-n} \end{bmatrix},$$

where the submatrix are of dimension $n \times n$, $n \times (N - n)$, $(N - n) \times n$ and $(N - n) \times (N - n)$ and 0 denotes the null matrix.

Since $Z = \begin{bmatrix} Z_s \\ Z_r \end{bmatrix}$,

$$(I - I_s)Z = (I - I_s)P_{\Omega^*}Z$$

implies that

$$Z_r = A_1Z_s + A_2Z_r \text{ and } \hat{Z}_r = (I - A_2)^{-1}A_1Z_s.$$

Further, $P_\Omega = PP_{\Omega^*}P^{-1}$ (Rao, 1973), then $P_{\Omega^*} = P^{-1}P_\Omega P$ and after some calculations we have

$$(I - I_s)P^{-1}Y = \begin{bmatrix} 0 \\ B_1Y_s + B_2Y_r \end{bmatrix}$$

and

$$(I - I_s)P^{-1}P_\Omega Y = (I - I_s)P^{-1}P_\Omega PP^{-1}Y$$
$$= (I - I_s)P_{\Omega^*}P^{-1}Y = \begin{bmatrix} 0 \\ (A_1C + A_2B_1)Y_s + A_2B_2Y_r \end{bmatrix}.$$

Thus, if $\hat{Y}_r$ is the solution in Y_r of

$$(I - I_s)P^{-1}Y = (I - I_s)P^{-1}P_\Omega Y,$$

it follows that

$$\hat{Y}_r = (B_2 - A_2 B_2)^{-1}(A_1 C + A_2 B_1 - B_1)Y_s.$$

Now, with this notation,

$$Z = P^{-1}Y = \begin{bmatrix} CY_s \\ B_1 Y_s + B_2 Y_r \end{bmatrix}$$

which implies that

$$Z_s = CY_s.$$

So,

$$
\begin{aligned}
B_1 Y_s + B_2 \hat{Y}_r &= B_1 Y_s + B_2(B_2 - A_2 B_2)^{-1}(A_1 C + A_2 B_1 - B_1)Y_s \\
&= \{B_1 + B_2 B_2^{-1}(I - A_2)^{-1}(A_1 C + A_2 B_1 - B_1)\}Y_s \\
&= \{B_1 + (I - A_2)^{-1}A_1 C + (I - A_2)^{-1}(A_2 - I)B_1\}Y_s \\
&= (I - A_2)^{-1}A_1 CY_s = \hat{Z}_r.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\Gamma &= h'PI_s Z + h'P\hat{\mu}_Z = h'P \begin{bmatrix} Z_s \\ 0 \end{bmatrix} + h'P \begin{bmatrix} 0 \\ \hat{Z}_r \end{bmatrix} \\
&= h' \begin{bmatrix} P_1 & 0 \\ P_3 & P_4 \end{bmatrix} \begin{bmatrix} CY_s \\ B_1 Y_s + B_2 \hat{Y}_r \end{bmatrix} \\
&= h' \begin{bmatrix} P_1 CY_s \\ P_3 CY_s + P_4 B_1 Y_s + P_4 B_2 \hat{Y}_r \end{bmatrix}
\end{aligned}
$$

and because

$$P.P^{-1} = \begin{bmatrix} P_1 C & 0 \\ P_3 C + P_4 B_1 & P_4 B_2 \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ 0 & I_{N-n} \end{bmatrix},$$

8

then

$$\Gamma = h' \begin{bmatrix} Y_s \\ \hat{Y}_r \end{bmatrix} = h' \begin{bmatrix} Y_s \\ 0 \end{bmatrix} + h' \begin{bmatrix} 0 \\ \hat{Y}_r \end{bmatrix} = h' I_s Y + h' \hat{\mu}_Y.$$

■

It is important to observe that, since P_Ω is a projection matrix,

$$(Y - P_\Omega Y)' V^{-1} (Y - P_\Omega Y) \leq (Y - v)' V^{-1} (Y - v),$$

for every vector $v \in \Omega$.

Furthermore, P_Ω has $\frac{N(N+1)}{2}$ unknown elements and it may be difficult to calculate by the above definition. But it can be obtained as $P_\Omega = A(A'V^{-1}A)^{-1}A'V^{-1}$, when A is a basis matrix for Ω.

Some applications of the result in Theorem 1 will be presented in the examples.

3. BEST LINEAR UNBIASED PREDICTOR IN THE COORDINATIZED MODEL

We now consider a coordinatized version of the model (1.1), given by

$$E(Y) = X\beta, \quad \beta \in \mathbb{R}^p$$

$$\text{Var}(Y) = \sigma^2 V, \tag{3.1}$$

$\sigma^2 > 0$, with V a known positive definite matrix and X a basis matrix of Ω.

Under this formulation, X is a $N \times p$ matrix of full rank p and there exists a unique $\beta \in \mathbb{R}^p$ such that $\mu = X\beta$. Regression models are included in the class of models defined in (3.1).

Royall (1976) derived the best linear unbiased predictor of the population total $T = \sum_{i=1}^{N} y_i$. This predictor, adapted to the notation introduced here and to predict any linear combination of Y is given by

$$\hat{T} = h'Y_s + h'(I - I_s) \begin{bmatrix} 0 \\ X_r\hat{\beta} + V_{rs}V_s^{-1}(Y_s - X_s\hat{\beta}) \end{bmatrix}, \tag{3.2}$$

where $\hat{\beta} = (X_s'V_s^{-1}X_s)^{-1}X_s'V_s^{-1}Y_s$ and $X = \begin{bmatrix} X_s \\ X_r \end{bmatrix}$.

Next theorem shows that in the coordinatized model (3.1), the optimal linear predictor obtained in Theorem 1 reduces to the Royall's predictor defined in (3.2).

Theorem 2. Under model (3.1), the optimal linear predictor $h'I_sY + h'\hat{\mu}_Y$ given in (2.1) is equal to $\hat{T}$.

Proof. We must show that $\hat{Y}_r$ in (2.1) is equal to $X_r\hat{\beta} + V_{rs}V_s^{-1}(Y_s - X_s\hat{\beta})$.

As proved in Theorem 1

$$\hat{Y}_r = (B_2 - A_2B_2)^{-1}(A_1C + A_2B_1 - B_1)Y_s$$

which is equivalent to

$$\hat{Y}_r = B_2^{-1}[(I - A_2)^{-1}A_1C - B_1]Y_s.$$

Applying (A.3), (A.1) and (A.2) of the appendix, it follows that

$$
\begin{aligned}
\hat{Y}_r &= B_2^{-1}[(I - A_2)^{-1}(B_1X_s + B_2X_r)(X'V^{-1}X)^{-1}X_s'C'C - B_1]Y_s \\
&= B_2^{-1}[(I - A_2)^{-1}(B_1X_s + B_2X_r)(X'V^{-1}X)^{-1}X_s'V_s^{-1} - B_1]Y_s \\
&= V_{rs}V_s^{-1}Y_s + B_2^{-1}(I - A_2)^{-1}(B_1X_s + B_2X_r)(X'V^{-1}X)^{-1}X_s'V_s^{-1}Y_s.
\end{aligned}
$$

Now, it is enough showing that

$$B_2^{-1}(I - A_2)^{-1}(B_1X_s + B_2X_r)(X'V^{-1}X)^{-1} = [X_r - V_{rs}V_s^{-1}X_s](X_s'V_s^{-1}X_s)^{-1}.$$

By (A.6),

$$
\begin{aligned}
&B_2^{-1}(I - A_2)^{-1}(B_1X_s + B_2X_r)(X'V^{-1}X)^{-1} \\
&= B_2^{-1}[I + (B_1X_s + B_2X_r)(X_s'V_s^{-1}X_s)^{-1}(X_s'B_1' + X_r'B_2')] \\
&\quad \times (B_1X_s + B_2X_r)(X'V^{-1}X)^{-1}
\end{aligned}
$$

10

and employing (A.2), last expression reduces to

$$[B_2^{-1} - V_{rs}V_s^{-1}X_s(X_s'V_s^{-1}X_s)^{-1}(X_s'B_1' + X_r'B_2') + X_r(X_s'V_s^{-1}X_s)^{-1}(X_s'B_1'$$

$$+X_r'B_2')](B_1X_s + B_2X_r)(X'V^{-1}X)^{-1}$$

$$= \{-V_{rs}V_s^{-1}X_s + X_r + [X_r - V_{rs}V_s^{-1}X_s](X_s'V_s^{-1}X_s)^{-1}(X_s'B_1' + X_r'B_2')$$

$$\times (B_1X_s + B_2X_r)\}(X'V^{-1}X)^{-1}.$$

Finally, using (A.5), we get

$$B_2^{-1}(I - A_2)^{-1}(B_1X_s + B_2X_r)(X'V^{-1}X)^{-1}$$

$$= (X_r - V_{rs}V_s^{-1}X_s)[I + (X_s'V_s^{-1}X_s)^{-1}(X'V^{-1}X - X_s'V_s^{-1}X_s)](X'V^{-1}X)^{-1}$$

$$= (X_r - V_{rs}V_s^{-1}X_s)(X_s'V_s^{-1}X_s)^{-1}. \qquad \blacksquare$$

4. EXAMPLES

In this section we present two examples to illustrate the optimal predictors that were obtained in the theorems.

In the first one, we consider a coordinate free model and the predictor is derived applying Theorem 1. Second example shows an application of Theorem 2 in a particular coordinatized model.

Example 1. Our objective is to predict the population total $T = \sum_{i=1}^{N} y_i$ in the model

$$E(Y) = \mu \quad \text{and}$$

$$\text{Var}(Y) = \frac{\sigma^2}{1 - \rho^2} \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{N-1} \\ & 1 & \rho & \cdots & \rho^{N-2} \\ & & 1 & \cdots & \rho^{N-3} \\ & & & \vdots & \\ & & & & 1 \end{bmatrix},$$

11

with ρ a known parameter and $\sigma^2 > 0$.

Because of the great quantity of calculations, without loss of generality, we restrict the attention to the situation where $N = 4$, $n = 3$, such that

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}, \quad Y_s = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}, \quad Y_r = [y_4] \text{ and }$$

$$V = \frac{1}{1-\rho^2} \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{bmatrix}, \quad |\rho| \neq 1, \ \rho \text{ known.}$$

In this case,

$$V^{-1} = \begin{bmatrix} -1 & -\rho & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & 0 \\ 0 & -\rho & 1+\rho^2 & -\rho \\ 0 & 0 & -\rho & 1 \end{bmatrix} \text{ and } P^{-1} = \begin{bmatrix} \sqrt{1-\rho^2} & 0 & 0 & 0 \\ -\rho & 1 & 0 & 0 \\ 0 & -\rho & 1 & 0 \\ 0 & 0 & -\rho & 1 \end{bmatrix}.$$

Since

$$\Omega = \{ v \in \mathbb{R}^4 \mid v = (\mu, \mu, \mu, \mu)', \ \mu \in \mathbb{R} \},$$

a base for Ω is given by $A = [1 \ 1 \ 1 \ 1]'$.

Then, it is easy to see that

$$\begin{aligned} P_\Omega &= A(A'V^{-1}A)^{-1}A'V^{-1} \\ &= \frac{1}{4 - 6\rho + 2\rho^2} \begin{bmatrix} 1-\rho & 1+\rho^2 - 2\rho & 1+\rho^2 - 2\rho & 1-\rho \\ 1-\rho & 1+\rho^2 - 2\rho & 1+\rho^2 - 2\rho & 1-\rho \\ 1-\rho & 1+\rho^2 - 2\rho & 1+\rho^2 - 2\rho & 1-\rho \\ 1-\rho & 1+\rho^2 - 2\rho & 1+\rho^2 - 2\rho & 1-\rho \end{bmatrix}. \end{aligned}$$

Also

$$(I - I_s)P^{-1}Y = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -\rho y_3 + y_4 \end{bmatrix}.$$

and

$$(I - I_s)P^{-1}P_\Omega Y =$$

$$\frac{1}{4 - 6\rho + 2\rho^2} \begin{bmatrix} 0 \\ 0 \\ 0 \\ (1 - 2\rho + \rho^2)(y_1 + y_4) + (1 - 3\rho + 3\rho^2 - \rho^3)(y_2 + y_3) \end{bmatrix}.$$

By Theorem 1, the optimal linear predictor of T is $\hat{T} = \sum_{i=1}^{3} y_i + \hat{y}_4$, where $\hat{y}_4$ is the solution in y_4 of the equation

$$(I - I_s)P^{-1}Y = (I - I_s)P^{-1}P_\Omega Y.$$

After calculations, we get

$$\hat{y}_4 = \frac{(1 - 2\rho + \rho^2)y_1 + (1 - 3\rho + 3\rho^2 - \rho^3)y_2 + (1 + \rho - 3\rho^2 + \rho^3)y_3}{3 - 4\rho + \rho^2}$$

and

$$\hat{T} = a_1 y_1 + a_2 y_2 + a_3 y_3,$$

where $a_1 = \dfrac{4 - 6\rho + 2\rho^2}{3 - 4\rho + \rho^2}$, $a_2 = \dfrac{4 - 7\rho + 4\rho^2 - \rho^3}{3 - 4\rho + \rho^2}$ and $a_3 = \dfrac{4 - 3\rho - 2\rho^2 + \rho^3}{3 - 4\rho + \rho^2}$.

It is interesting to note that, if $\rho = 0$, such that $V = I$ and y_i and y_j are uncorrelated, $i \neq j$, then $\hat{T} = 4\bar{y}_s$, where $\bar{y}_s$ is the sample mean. In this case, $\hat{T}$ is the expansion predictor which was found by Rodrigues (1989) under the model $E(Y) = \mu$ and $\text{Var}(Y) = \sigma^2 I$.

Example 2. Let us consider the superpopulation model

$$y_i = \beta x_i + \epsilon_i, \quad i = 1, 2, \ldots, N,$$

with $E(\epsilon_i) = 0$, $\text{Var}(\epsilon_i) = 1$, $\text{Cov}(\epsilon_i, \epsilon_j) = \rho$ for $i \neq j$, $i, j = 1, 2, \ldots, N$, and ρ a known parameter, $|\rho| \neq 1$.

Our objective is to calculate the best linear unbiased predictor of the population total $T = \sum_{i=1}^{N} y_i$.

In this situation, the model is coordinatized, and by Theorem 2, it is enough to obtain the value

$$\hat{Y}_r = X_r \hat{\beta} + V_{rs} V_s^{-1} (Y_s - X_s \hat{\beta}).$$

Let V_s and V_{rs} be written as

$$V_s = (1 - \rho) I_n + \rho J_n$$

$$V_{rs} = \rho J_{N-n,n} ,$$

where J_n and $J_{N-n,n}$ are respectively the $n \times n$ and $(N - n) \times n$ matrix of ones.

Thus, it is easy to see that

$$\hat{\beta} = \frac{\sum_{i=1}^{n} x_i y_i - \frac{\rho}{1+(n-1)\rho} \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i^2 - \frac{\rho}{1+(n-1)\rho} \left(\sum_{i=1}^{n} x_i \right)^2}$$

and

$$\hat{Y}_r = \begin{bmatrix} a_{n+1}\hat{\beta} + b \\ a_{n+2}\hat{\beta} + b \\ \vdots \\ a_N \hat{\beta} + b \end{bmatrix},$$

where

$$a_{n+j} = \left(x_{n+j} - \frac{\rho}{1 + (n-1)\rho} \sum_{i=1}^{n} x_i \right) \hat{\beta}, \quad j = 1, 2, \ldots, N - n,$$

and

$$b = \frac{\rho \sum_{i=1}^{n} y_i}{1 + (n - 1)\rho}.$$

APPENDIX

First, we show that $\Gamma = h'PI_*Z + h'P\hat{\mu}_Z$ defined in the proof of Theorem 1 does not depend on unknown quantities.

Since P is a lower triangular matrix, P^{-1} is lower triangular also, then

$$P^{-1}Y = \begin{bmatrix} \delta_{11} & 0 & \cdots & 0 \\ \delta_{21} & \delta_{22} & \cdots & 0 \\ \vdots & & & \\ \delta_{N1} & \delta_{N2} & \cdots & \delta_{NN} \end{bmatrix} Y = \begin{bmatrix} \delta_{11}y_1 \\ \delta_{21}y_1 + \delta_{22}y_2 \\ \vdots \\ \delta_{n1}y_1 + \delta_{n2}y_2 + \cdots + \delta_{nn}y_n \\ \delta_{n+11}y_1 + \delta_{n+12}y_2 + \cdots + \delta_{n+1\,n+1}y_{n+1} \\ \vdots \\ \delta_{N1}y_1 + \delta_{N2}y_2 + \cdots + \delta_{NN}y_N \end{bmatrix}$$

and

$$I_*Z = \begin{bmatrix} I_n & 0 \\ 0 & 0 \end{bmatrix} P^{-1}Y = \begin{bmatrix} \delta_{11}y_1 \\ y_{21}y_1 + \delta_{22}y_2 \\ \vdots \\ \delta_{n1}y_1 + \delta_{n2}y_2 + \cdots + \delta_{nn}y_n \\ 0 \\ \vdots \\ 0 \end{bmatrix} .$$

So, it is shown that $h'PI_*Z$ does not depend on unknown quantities. By the proof of Theorem 1, we can see that $\hat{Z}_r = (I - A_2)^{-1}A_1CY_*$ and thus, $h'P\hat{\mu}_Z$ also does not depend on unknown quantities. Then Γ is a predictor of $h'Y$.

Now we derive the results (A.1) through (A.6) which are necessary to prove Theorem 2.

Let P^{-1} partitioned as in the proof of Theorem 1, $P^{-1} = \begin{bmatrix} C & 0 \\ B_1 & B_2 \end{bmatrix}$ which implies that

$$P = \begin{bmatrix} C^{-1} & 0 \\ -B_2^{-1}B_1C^{-1} & B_2^{-1} \end{bmatrix} .$$

Then using the equality $V = PP'$ and after some algebraic manipulations, it follows that

$$C^{-1}C^{-1'} = V_s,$$

and so,

$$C'C = V_s^{-1}. \qquad (A.1)$$

Furthermore,

$$-C^{-1}C^{-1'}B_1'B_2^{-1} = -V_s B_1'B_2^{-1'} = V_{sr}$$

and hence

$$-B_2^{-1}B_1 = V_{rs}V_s^{-1}. \qquad (A.2)$$

In the coordinatized model with $\mu = X\beta$ and covariance matrix V, it is well known (Rao, 1973), that

$$P_\Omega = X(X'V^{-1}X)^{-1}X'V^{-1}$$

and thus

$$P_{\Omega^*} = P^{-1}P_\Omega P = P^{-1}X(X'V^{-1}X)^{-1}X'P^{-1'}.$$

In the partitioned form, this matrix can be written as

$$P_{\Omega^*} = \begin{bmatrix} H_1 & H_2 \\ A_1 & A_2 \end{bmatrix} = \begin{bmatrix} C & 0 \\ B_1 & B_2 \end{bmatrix} \begin{bmatrix} X_s \\ X_r \end{bmatrix} (X'V^{-1}X)^{-1}[X_s' \ \ X_r'] \begin{bmatrix} C' & B_1' \\ 0 & B_2' \end{bmatrix}$$

$$= \begin{bmatrix} CX_s(X'V^{-1}X)^{-1} \\ (B_1X_s + B_2X_r)(X'V^{-1}X)^{-1} \end{bmatrix} [X_s'C' \ \ X_s'B_1' + X_r'B_2']$$

16

$$= \begin{bmatrix} CX_s(X'V^{-1}X)^{-1}X_s'C' & CX_s(X'V^{-1}X)^{-1}(X_s'B_1' + X_r'B_2') \\ (B_1X_s + B_2X_r)(X'V^{-1}X)^{-1}X_s'C' & (B_1X_s + B_2X_r)(X'V^{-1}X)^{-1}(X_s'B_1' + X_r'B_2') \end{bmatrix},$$

then

$$A_1 = (B_1X_s + B_2X_r)(X'V^{-1}X)^{-1}X_s'C', \tag{A.3}$$

and

$$A_2 = (B_1X_s + B_2X_r)(X'V^{-1}X)^{-1}(X_s'B_1' + X_r'B_2'). \tag{A.4}$$

Using the fact that $V^{-1} = P^{-1'}P^{-1}$, it follows that

$$V^{-1} = \begin{bmatrix} C' & B_1' \\ 0 & B_2' \end{bmatrix} \begin{bmatrix} C & 0 \\ B_1 & B_2 \end{bmatrix} = \begin{bmatrix} C'C + B_1'B_1 & B_1'B_2 \\ B_2'B_1 & B_2'B_2 \end{bmatrix}$$

and

$$X'V^{-1}X = X_s'C'CX_s + X_s'B_1'B_1X_s + X_r'B_2'B_1X_s + X_s'B_1'B_2X_r + X_r'B_2'B_2X_r.$$

Applying (A.1),

$$X'V^{-1}X = X_s'V_s^{-1}X_s + X_s'B_1'B_1X_s + X_r'B_2'B_1X_s + X_s'B_1'B_2X_r + X_r'B_2'B_2X_r. \tag{A.5}$$

Application of a result of inverse matrix in conjunction with (A.4) and (A.5) yields

$$\begin{aligned} (I - A_2)^{-1} &= [I - (B_1X_s + B_2X_r)(X'V^{-1}X)^{-1}(X_s'B_1' + X_r'B_2')]^{-1} \\ &= I - (B_1X_s + B_2X_r)[(X_r'B_2' + X_s B_1')(B_1X_s \\ &\quad + B_2X_r) - X'V^{-1}X]^{-1}(X_s'B_1' + X_r'B_2') \\ &= I + (B_1X_s + B_2X_r)(X_s'V_s^{-1}X_s)^{-1}(X_s'B_1' + X_r'B_2'). \end{aligned} \tag{A.6}$$

REFERENCES

1. Rodrigues, J. The Coordinate-Free Prediction in Finite Populations. Pakistan Journal of Statistics 1989, 5, 119–129.

2. Arnold, S. F. *The Theory of Linear Models and Multivariate Analysis*, John Wiley-New York, **1980**.

3. Drygas, H. *The Coordinate-Free Approach to Gauss-Markov Estimation*, Springer-Verlag, **1970**.

4. Royall, R. M. The Linear Least-Squares Prediction Approach to Two-Stage Sampling, Journal of the American Statistical Association **1976**, 71 (355), 657–664.

5. Graybill, F. A. *Theory and Application of the Linear Model*, North Scituate, MA: Duxbury Press, **1976**.

6. Rao, C. R. *Linear Statistical Inference and Its Applications*, 2nd Ed., New York. Wiley, **1973**.

ÚLTIMOS RELATÓRIOS TÉCNICOS PUBLICADOS

2002-01 - **BUENO, V.C.** Asymptotic reliability for a coherent system with an infinite number of components. 2002. 10p. (RT-MAE-2002-01)

2002-02 - **TAVARES, H.R., ANDRADE, D.F.** Item Response Theory for Longitudinal Data: Item and Population Ability Parameters Estimation. 2002. 20p. (RT-MAE-2002-02)

2002-03 - **BOLFARINE, H., VALENÇA, D.M.** Score tests for Weibull-regression models with random effect. 2002. 16P. (RT-MAE-2002-03)

2002-04 - **BUENO, V.C.** Using a second order stochastic dominance for active redundancy allocation in a K-out-of-n system. 2002. 7p. (RT-MAE-2002-4)

2002-05 - **FERRARI, S.L.P., LUCAMBIO, F., CRIBARI-NETO, F** Adjusted and Bartlett-Adjusted profile likelihood ratio tests. 2002. 19p. (RT-MAE-2002-05).

2002-06 - **PEREIRA, C.A.B., STERN, J.M.** Full Bayesian Significance Test: Invariant Formulation. 2002. 10p. (RT-MAE-2002-06).

2002-07 - **BUENO, V.C.** Optimal arrangement of components for systems of dependent components. 2002. 13p. (RT-MAE-2002-07)

2002-08- **KOLEV, N., PAIVA, D.** Multinomial Latent Model for Random Sums. 2002. 9p. (RT-MAE-2002-08)

2002-09 - **BUENO, V.C.** A Note on a Hazard Processes Ordering: Application on a *K-out-of-N System*. 2002. 9P. (RT-MAE-2002-09)

2002-10 - **FERRARI, S.L.P., DAVID, J.S.E., ANDRÉ, A.P., PEREIRA, L.A.A.** Use of overdispersed regression models in analysing the association between air pollution and human health. 2002. 15p. (RT-MAE-2002-10)

2002-11 - **PERES, A.A.Q., BOTTER, D.A.B.** Association between air pollution and human health: filling in the gaps of the pollutant series. 2002. 13P. (RT-MAE-2002-11)

2002-12 - **CYSNEIROS, F.J.A., PAULA, A.G.** One-Sided Tests in Linear Models with Multivariate t-Distribution. 2002. 33p. (RT-MAE-2002-12)

2002-13 - **CRIBARI-NETO, F., FERRARI, S.L.P., OLIVEIRA, W.A.S.C.** A numerical evaluation of tests based on different heteroskedasticity-consistent covariance matrix estimators. 2002. 19p. (RT-MAE-2002-13)

2002-14 - **BARROSO, L.P., CORDEIRO, G.M.** Bartlett Corrections in Heteroskedastic t Regression Models. 2002. 12p. (RT-MAE-2002-14)

2002-15 - **KOLEV, N.** Volodya, I Miss You (Two Correlated Collective Risk Models) 2002. 20p. (RT-MAE-200215)

2002-16 - **BUENO, V.C.** Burn-in a coherent system in a random environment. 2002. 7p. (RT-MAE-2002-16)

2002-17 - **FERRARI, S.L.P., CYSNEIROS, A.H.M.A., CRIBARI-NETO, F.** An improved test for heteroskedasticity using adjusted modified profile likelihood inference. 2002. 15p. (RT-MAE-2002-17)

2002-18 - **CYSNEIROS, F.J.A., PAULA, G.A.** One-sided tests in univariate elliptical linear regression models. 2002. 19p. (RT-MAE-2002-18)

The complete list of "Relatórios do Departamento de Estatística", IME-USP, will be sent upon request.