Scientific
Research
Publishing

The Coordinate-Free Prediction in Finite Populations with Correlated Observations

Silvia N. Elian

Department of Statistics, Institute of Mathematics and Statistics, University of São Paulo, São Paulo, Brazil
Email: selian@ime.usp.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 super population model with correlated observations.

Keywords

Coordinate-Free Models, Best Linear Unbiased Predictor, Covariance Matrix, Orthogonal Projection

1. Introduction

A coordinate-free approach in finite populations was introduced by [1] 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 IR^N, where N is the size of the finite population. Coordinate-free models in the linear models context are discussed by [2] and [3].

In a finite population $P = \{1, 2, \cdots, N\}$, where N is the known population size, let y_i, $i = 1, 2, \cdots, N$ 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 $P = s \cup r$.

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 $V_{sr} = \text{Cov}(Y_s, Y_r)$ are the covariance matrix.

Under a less general model, with $\text{Var}(Y) = \sigma^2 D$, D a known diagonal matrix, [1] 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 was the main contribution of the paper. In Section 3, we show that under the coordinatized model, this predictor coincides with that given by [4]. 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 $I_s = \text{diag}(i_1, i_2, \cdots, i_N)$ is a diagonal matrix with its k-th diagonal element i_k, where $i_k = 1$ if $k \in s$ and $i_k = 0$ if $k \in r$, $s = \{1, 2, \cdots, n\}$, $r = \{n+1, n+2, \cdots, N\}$.

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.

Let

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

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

$$Y_s = I_s Y, \quad Y_r = (I - I_s)Y, \quad N \times 1 \text{ matrices.}$$

Since after the sample is observed, $I_s Y$ 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

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

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

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

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

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

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 [1]. 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, $\hat{\mu}_*$ is such that

$$\hat{\mu}_* = \left(I - I_s \right) P_\Omega \left(Y_s + \hat{\mu}_* \right) \tag{1.2}$$

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 [5] (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, by multivariate properties of covariance matrix of random vectors and matrix results,

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

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

$$h'I_sY + h'\hat{\mu}_Y \qquad (2.1)$$

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 $V = PP'$,

$$\Omega^* = \left\{ \mu^* \middle| \mu^* = P^{-1}\mu, \mu \in \Omega \right\},$$

$$\Gamma = h'PI_sZ + h'P\hat{\mu}_Z,$$

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

$$\hat{\mu}_Z = (I - I_s)P_{\Omega^*}(I_sZ + \hat{\mu}_Z).$$

We note that Γ does not depend on unknown quantities because, as it will be shown in the appendix, $h'PI_sZ$ 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

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

by [1] results, the optimal linear predictor of $\ell'Z$ is

$$\ell'I_sZ + \ell'\hat{\mu}_Z$$

with $\hat{\mu}_Z = \begin{bmatrix} 0 \\ \hat{Z}_r \end{bmatrix}$, where 0 is the null vector of dimension n and $\hat{Z}_r$ obtained by (1.2) is the solution of the system of linear equations

$$\hat{\mu}_Z = (I - I_s)P_{\Omega^*}(I_sZ + \hat{\mu}_Z).$$

Taking $\ell' = h'P$, this predictor reduces to Γ and $\ell'Z = h'PP^{-1}Y = h'Y$. So, by (1.2), 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 $Z = \begin{bmatrix} Z_s \\ Z_r \end{bmatrix}$,

$$\hat{\mu}_Z = (I - I_s) P_{\Omega^*} (I_s Z + \hat{\mu}_Z)$$

implies that

$$\hat{Z}_r = A_1 Z_s + A_2 \hat{Z}_r \quad \text{and} \quad \hat{Z}_r = (I - A_2)^{-1} A_1 Z_s.$$

Further, $P_\Omega = P P_{\Omega^*} P^{-1}$ [6], then $P_{\Omega^*} = P^{-1} P_\Omega P$ and after some calculations we have

$$(I - I_s) P^{-1} Y = \begin{bmatrix} 0 \\ B_1 Y_s + B_2 Y_r \end{bmatrix}$$

and

$$(I - I_s) P^{-1} P_\Omega Y = (I - I_s) P^{-1} P_\Omega P P^{-1} Y$$
$$= (I - I_s) P_{\Omega^*} P^{-1} Y = \begin{bmatrix} 0 \\ (A_1 C + A_2 B_1) Y_s + A_2 B_2 Y_r \end{bmatrix}.$$

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,

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

which implies that

$$Z_s = C Y_s.$$

So,

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

Hence,

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

and because

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

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 P_Ω has $\dfrac{N(N+1)}{2}$ 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

$$\begin{aligned} E(Y) &= X\beta, \ \ \beta \in \mathrm{IR}^p \\ \mathrm{Var}(Y) &= \sigma^2 V. \end{aligned} \tag{3.1}$$

$\sigma^2 > 0$, 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 $\beta \in \mathrm{IR}^p$ such that $\mu = X\beta$. Regression models are included in the class of models defined in (3.1).

[4] 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 $h' I_s Y + h' \hat{\mu}_Y$ 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

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

which is equivalent to

$$\hat{Y}_r = B_2^{-1}\left[\left(I - A_2\right)^{-1} A_1 C - B_1\right] Y_s.$$

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

$$\hat{Y}_r = B_2^{-1}\left[\left(I - A_2\right)^{-1}\left(B_1 X_s + B_2 X_r\right)\left(X'V^{-1}X\right)^{-1} X_s' C' C - B_1\right] Y_s$$

$$= B_2^{-1}\left[\left(I - A_2\right)^{-1}\left(B_1 X_s + B_2 X_r\right)\left(X'V^{-1}X\right)^{-1} X_s' V_s^{-1} - B_1\right] Y_s$$

$$= V_{rs} V_s^{-1} Y_s + B_2^{-1}\left(I - A_2\right)^{-1}\left(B_1 X_s + B_2 X_r\right)\left(X'V^{-1}X\right)^{-1} X_s' V_s^{-1} Y_s.$$

Now, it is enough showing that

$$B_2^{-1}\left(I - A_2\right)^{-1}\left(B_1 X_s + B_2 X_r\right)\left(X'V^{-1}X\right)^{-1} = \left[X_r - V_{rs} V_s^{-1} X_s\right]\left(X_s' V_s^{-1} X_s\right)^{-1}.$$

By (A.6),

$$B_2^{-1}\left(I - A_2\right)^{-1}\left(B_1 X_s + B_2 X_r\right)\left(X'V^{-1}X\right)^{-1}$$

$$= B_2^{-1}\left[I + \left(B_1 X_s + B_2 X_r\right)\left(X_s' V_s^{-1} X_s\right)^{-1}\left(X_s' B_1' + X_r' B_2'\right)\right]$$

$$\times\left(B_1 X_s + B_2 X_r\right)\left(X'V^{-1}X\right)^{-1}$$

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

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

$$\left. + X_r\left(X_s' V_s^{-1} X_s\right)^{-1}\left(X_s' B_1' + X_r' B_2'\right)\right]\left(B_1 X_s + B_2 X_r\right)\left(X'V^{-1}X\right)^{-1}$$

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

$$\left. \times\left(B_1 X_s + B_2 X_r\right)\right\}\left(X'V^{-1}X\right)^{-1}.$$

Finally, using (A.5), we get

$$B_2^{-1}\left(I - A_2\right)^{-1}\left(B_1 X_s + B_2 X_r\right)\left(X'V^{-1}X\right)^{-1}$$

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

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

4. Examples

In this section, we present two examples to illustrate the optimal predictors that are 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

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

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

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

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

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

In this case,

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

Since

$$\Omega=\left\{ v \in \mathrm{IR}^4 \,\middle|\, v=(\mu,\mu,\mu,\mu)', \mu \in \mathrm{IR} \right\},$$

a base for Ω is given by $A=\begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}'$.

Then, it is easy to see that

$$P_\Omega = A\left(A'V^{-1}A\right)^{-1}A'V^{-1}$$

$$=\frac{1}{4-6\rho+2\rho^2}\begin{bmatrix} 1-\rho & 1+\rho^2-2\rho & 1+\rho^2-2\rho & 1-\rho \\ 1-\rho & 1+\rho^2-2\rho & 1+\rho^2-2\rho & 1-\rho \\ 1-\rho & 1+\rho^2-2\rho & 1+\rho^2-2\rho & 1-\rho \\ 1-\rho & 1+\rho^2-2\rho & 1+\rho^2-2\rho & 1-\rho \end{bmatrix}.$$

Also

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

and

$$\left(I - I_s\right)P^{-1}P_\Omega Y = \frac{1}{4 - 6\rho + 2\rho^2} \begin{bmatrix} 0 \\ 0 \\ 0 \\ \left(1 - 2\rho + \rho^2\right)\left(y_1 + y_4\right) + \left(1 - 3\rho + 3\rho^2 - \rho^3\right)\left(y_2 + y_3\right) \end{bmatrix}.$$

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

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

After calculations, we get

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

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 [1] under the model $E(Y) = \mu$ and $\mathrm{Var}(Y) = \sigma^2 I$.

Example 2. Let us consider the superpopulation model

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

with $E(\epsilon_i) = 0$, $\mathrm{Var}(\epsilon_i) = 1$, $\mathrm{Cov}(\epsilon_i, \epsilon_j) = \rho$ for $i \neq j$, $i, j = 1, 2, \cdots, N$, 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

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

Let V_s and V_{rs} be written as

$$V_s = \left(1 - \rho\right)I_n + \rho J_n$$
$$V_{rs} = \rho J_{N-n,n},$$

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

Thus, it is easy to see that

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

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

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

and

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

References

[1] Rodrigues, J. (1989) The Coordinate-Free Prediction in Finite Populations. *Pakistan Journal of Statistics*, **5**, 119-129.

[2] Arnold, S.F. (1980) The Theory of Linear Models and Multivariate Analysis. John Wiley, New York.

[3] Drygas, H. (1970) The Coordinate-Free Approach to Gauss-Markov Estimation. Springer-Verlag, Berlin. https://doi.org/10.1007/978-3-642-65148-9

[4] Royall, R.M. (1976) The Linear Least-Squares Prediction Approach to Two-Stage Sampling. *Journal of the American Statistical Association*, **71**, 657-664. https://doi.org/10.1080/01621459.1976.10481542

[5] Graybill, F.A. (1976) Theory and Application of the Linear Model. Duxbury Press, North Scituate, MA.

[6] Rao, C.R. (1973) Linear Statistical Inference and Its Applications. 2nd Edition, Wiley, New York. https://doi.org/10.1002/9780470316436

Appendix

First, we show that $\Gamma = h'PI_sZ + h'P\hat{\mu}_Z$ 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

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

and

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

So, it is shown that $h'PI_sZ$ does not depend on unknown quantities. By the proof of Theorem 1, we can see that $\hat{Z}_r = (I - A_2)^{-1} A_1CY_s$ 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,

$$C'C = V_s^{-1}. \tag{A.1}$$

Furthermore,

$$-C^{-1}C^{-1'}B_1'B_2^{-1} = -V_sB_1'B_2^{-1'} = V_{sr}$$

and hence

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

In the coordinatized model with $\mu = X\beta$ and covariance matrix V, it is

well known [6], that

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

and thus

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

In the partitioned form, this matrix can be written as

$$
\begin{aligned}
P_{\Omega^*} &= \begin{bmatrix} H_1 & H_2 \\ A_1 & A_2 \end{bmatrix} = \begin{bmatrix} C & 0 \\ B_1 & B_2 \end{bmatrix} \begin{bmatrix} X_s \\ X_r \end{bmatrix} \left(X'V^{-1}X \right)^{-1} \begin{bmatrix} X_s' & X_r' \end{bmatrix} \begin{bmatrix} C' & B_1' \\ 0 & B_2' \end{bmatrix} \\
&= \begin{bmatrix} CX_s \left(X'V^{-1}X \right)^{-1} \\ \left(B_1 X_s + B_2 X_r \right) \left(X'V^{-1}X \right)^{-1} \end{bmatrix} \begin{bmatrix} X_s'C' & X_s'B_1' + X_r'B_2' \end{bmatrix} \\
&= \begin{bmatrix} CX_s \left(X'V^{-1}X \right)^{-1} X_s'C' & CX_s \left(X'V^{-1}X \right)^{-1} \left(X_s'B_1' + X_r'B_2' \right) \\ \left(B_1 X_s + B_2 X_r \right) \left(X'V^{-1}X \right)^{-1} X_s'C' & \left(B_1 X_s + B_2 X_r \right) \left(X'V^{-1}X \right)^{-1} \left(X_s'B_1' + X_r'B_2' \right) \end{bmatrix},
\end{aligned}
$$

then

$$A_1 = \left(B_1 X_s + B_2 X_r \right) \left(X'V^{-1}X \right)^{-1} X_s'C', \tag{A.3}$$

and

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

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

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

Applying (A.1),

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

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

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

Scientific Research Publishing

Submit or recommend next manuscript to SCIRP and we will provide best service for you:

Accepting pre-submission inquiries through Email, Facebook, LinkedIn, Twitter, etc.
A wide selection of journals (inclusive of 9 subjects, more than 200 journals)
Providing 24-hour high-quality service
User-friendly online submission system
Fair and swift peer-review system
Efficient typesetting and proofreading procedure
Display of the result of downloads and visits, as well as the number of cited articles
Maximum dissemination of your research work

Submit your manuscript at: http://papersubmission.scirp.org/
Or contact ojs@scirp.org