RT-MAE-8916

EQUIVARIANT PREDICTION IN FINITE POPULATIONS

by

Heleno Bolfarine

Palavras CHaves: Equivariant predictors, location,
(Key words) scale and location-scale superpopulation
models, general linear models, Pitman predictors

Classificação AMS: 62D05
(AMS Classification)

- Agosto de 1989 -

EQUIVARIANT PREDICTION IN FINITE POPULATIONS

HELENO BOLFARINE
Instituto De Matemática E Estatística
Universidade De São Paulo
Caixa Postal 20570-Agência Iguatemi
São Paulo-Brasil

Key Words and Frases: *equivariant predictors; location,scale and location–scale superpopulation models; general linear models; Pitman predictors.*

ABSTRACT

A theory of equivariant prediction is developed for predicting the population total in finite populations. Minimum risk equivariant predictors (MREP) are derived under the location, scale and location–scale superpopulation models. Under the general linear model, it is shown that the best(linear) unbiased predictor (B(L)UP) is an MREP.

1. INTRODUCTION

The purpose of this paper is to develop a general theory of equivariance for predicting the population total in finite populations. Equivariant estimation in the infinite population context is considered in Lehmann (1985) and a more advanced treatment can be found in Zacks (1971). In the finite population context, very litlle has been done. Population variance prediction is considered in the important works of Zacks and Solomon (1981) and Zacks (1981), using a Bayesian equivariant approach.

The approach we follow to find the MREP, whenever it exists, is first to characterize the class of all equivariant predictors and then, within this class, find the one with minimum risk. For the simple location model, the task of finding the MREP using the squared error loss function is made easier by exploring some relationships with the B(L)UP.

The finite population is indexed by $\mathcal{P} = \{1, \ldots, N\}$, where N is known. Associated with unit k, there is an unknown quantity y_k and it is considered that $y = (y_1, \ldots, y_N)$ assumes its values on a space $\mathcal{Y}$ and is a realization of a vector Y ($Nx1$), of independent random variables such that

$$Y \sim F_\xi(y), \tag{1}$$

where $F_\xi \in \mathcal{F} = \{F_\xi, \xi \in \Pi_\Omega\}$, and Π_Ω is a subspace of R^N of dimension p. However, no distinction is made between y and Y.

To gain information about a function of y, that might be, for instance, the population total, $T(\mathbf{y}) = \sum_{i=1}^{N} y_i$, a sample s of size $n(> p)$ is selected from $\mathcal{P}$. Let $r = \mathcal{P} - s$ be the unobserved part of $\mathcal{P}$. After s has been selected, we may reorder (without loss of generality) $\mathbf{y} = (\mathbf{y}_s, \mathbf{y}_r)$, in such a way that $\mathbf{y}_s = (y_1, \dots, y_n)$ and $\mathbf{y}_r = (y_{n+1}, \dots, y_N)$ correspond, respectively, to the observed and unobserved parts of y. Then, we may write $T(\mathbf{y}) = \sum_{i \in s} y_i + \sum_{i \in r} y_i$, in such a way that predicting $T(\mathbf{y})$ is then equivalent to predicting $T_2(\mathbf{y}_r) = \sum_{i \in r} y_i = (N - n)\bar{y}_r$. Optimal (best unbiased) prediction of $T(\mathbf{y})$ is considered in Rodrigues et al. (1985). Bayesian prediction of $T(\mathbf{y})$ is considered in Bolfarine et al. (1987) and minimax prediction is considered in Bolfarine (1987).

Section 2 presents some important and general results for developing a theory of equivariant prediction in finite populations. Equivariant prediction under the location sampling model is considered in Section 3. The results of Section 3 are generalyzed in Section 4 to the general linear model, where it is shown that the ratio and regression estimators are MREP. In Sections 5 and 6, the scale and location–scale superpopulation models are considered. Pitman type predictors are introduced in Sections 3, and 5 under squared error loss functions.

2. A GENERAL FORMULATION FOR EQUIVARIANT PREDICTION IN FINITE POPULATIONS

As specified in the introduction, $\mathbf{y} = (y_1, \dots, y_N)$ assumes its values in a sample space $\mathcal{Y}$, and is distributed according to (1).

We shall say that a transformation g of $\mathcal{Y}$ onto itself leaves model (1) invariant if the distribution of

$$\mathbf{y}^* = g\mathbf{y} = (g\mathbf{y}_s, g\mathbf{y}_r) = (\mathbf{y}_s^*, \mathbf{y}_r^*), \tag{2}$$

is again a member $\mathcal{F}$, say $F_{\xi^*}(\mathbf{y})$,where

$$\xi^* = \bar{g}\xi, \tag{3}$$

and $\bar{g}$ is the one to one transformation of the parametric space Π_Ω induced by g. A class of transformation leaving a model invariant can be assumed to be a group (Zacks, 1971, Lehmann, 1985) and so, g and $\bar{g}$ generate group structures on $\mathcal{Y}$ and Π_Ω, respectively. As a consequence, for any integrable function Ψ,

$$E_\xi[\Psi(g\mathbf{y})] = E_{\xi^*}[\Psi(\mathbf{y})]. \tag{4}$$

Similar results hold for $\Psi(\mathbf{y}_s)$ and $\Psi(\mathbf{y}_r)$.

Since our main interest is on predicting the population total $T(\mathbf{y})$ (taking values on a set $\mathcal{H}$), applying g to $\mathbf{y}$ leads to

$$T(\mathbf{y}^*) = T(g\mathbf{y}) = \tilde{g}T(\mathbf{y}),$$

where $\tilde{g}$ generates a group of transformations of $\mathcal{H}$ into itself. So, predictor $\hat{T}(\mathbf{y}_s)$ of $T(\mathbf{y})$ is called equivariant if, for every transformation g and all $\mathbf{y}_s$,

$$\hat{T}(g\mathbf{y}_s) = \tilde{g}\hat{T}(\mathbf{y}_s). \tag{5}$$

Since the problem of predicting $T(\mathbf{y})$ in terms of $\mathbf{y}_s$ and $T(g\mathbf{y})$ in terms of $g\mathbf{y}_s$ represent the same physical situation, expressed in a new coordinate system, a loss fuction L should be invariant, that is,

$$L(\hat{T}(g\mathbf{y}_s); T(g\mathbf{y})) = L(\hat{T}(\mathbf{y}_s); T(\mathbf{y})). \tag{6}$$

Let the risk function of any predictor $\hat{T}(\mathbf{y}_s)$ be defined as

$$R[\xi; \hat{T}(\mathbf{y}_s)] = E[L(\hat{T}(\mathbf{y}_s); T(\mathbf{y}))].$$

When (6) holds, we shall say that the problem of predicting $T(\mathbf{y})$ is invariant on the basis of model (1), with respect to the transformations (2) and (3). By restricting attention to predictors satisfying (5) and loss functions satisfying (6), we have from (4),

Theorem 2.1. *If $\hat{T}(\mathbf{y}_s)$ is an equivariant predictor in a problem which is invariant with respect to a transformation g, then*

$$R[\bar{g}\xi; \hat{T}(\mathbf{y}_s)] = R[\xi; \hat{T}(\mathbf{y}_s)].$$

Infinite population versions of Theorem 2.1 can be found in Zacks (1971)and Lehmann (1983). Note also that Theorem 2.1 implies that $R[\xi; \hat{T}(\mathbf{y}_s)]$ is constant on the orbits of the group generated by the transformtions g.

A group of transformations is said to be transitive if for any two points there is a transformation taking the first point into the second.

Corollary 2.1. *Under the assumptions of Theorem 2.1, if the group of transformations generated by $\bar{g}$ is transitive over Π_Ω, then, for any predictor $\hat{T}(\mathbf{y}_s)$, $R[\xi; \hat{T}(\mathbf{y}_s)]$ is constant, that is, independent of ξ.*

The result of Corollary 2.1 simplifies considerably the task of finding an MREP, since it is sufficient to find the equivariant predictor with the smallest risk within the class of all equivariant predictors. In the following sections, the results of this section are applied to location, scale and location–scale superpopulation models.

3. THE SIMPLE LOCATION MODEL

In this section it is considered that $y_1, \ldots, y_N$ are independent and

$$y_i \sim F(y_i - \theta), \tag{7}$$

$i = 1, \ldots, N$, where $\theta \in \mathcal{R}$ and F has density f. Model (7) remains invariant under the transformations

$$y_i^* = y_i + c \text{ and } \theta^* = \theta + c, \tag{8}$$

$i = 1, \ldots, N$. The population total $T(\mathbf{y})$ when expressed in the new coordinate system, becomes $T(\mathbf{y}^*) = T(\mathbf{y}) + Nc$, and so, according to (5), a predictor $\hat{T}(\mathbf{y}_s)$, to be equivariant should satisfy

$$\hat{T}(\mathbf{y}_s^*) = \hat{T}(\mathbf{y}_s) + Nc. \tag{9}$$

A predictor satisfying (9) is called a location–equivariant predictor. According to (6), the loss function should satisfy

$$L(\hat{T}(\mathbf{y}_s); T(\mathbf{y})) = \Delta(\hat{T}(\mathbf{y}_s) - T(\mathbf{y})), \tag{10}$$

that is, be a function of the difference $\hat{T}(\mathbf{y}_s) - T(\mathbf{y})$ only. Note that the squared error loss function

$$L(\hat{T}(\mathbf{y}_s); T(\mathbf{y})) = (\hat{T}(\mathbf{y}_s) - T(\mathbf{y}))^2, \tag{11}$$

satisfies condition (10).

Lemma 3.1 *Under the transformations (8), the bias, variance and risk functions of any location–equivariant predictor are all constant, that is, independent of θ.*

Lemma 3.1 follows from Corollary 2.1, since the group generated by transformations (8) is transitive.

A predictor $\hat{T}^*(\mathbf{y}_s)$ of $T(\mathbf{y})$ is said to be the B(L)UP of $T(\mathbf{y})$ if it is unbiased, that is, (i)$E_\theta[\hat{T}^*(\mathbf{y}_s) - T(\mathbf{y})] = 0$, for all θ and $(ii)E_\theta[\hat{T}^*(\mathbf{y}_s) - T(\mathbf{y})]^2 \leq E_\theta[\hat{T}(\mathbf{y}_s) - T(\mathbf{y})]^2$, for all θ and any other (linear) unbiased predictor $\hat{T}(\mathbf{y}_s)$, where $E_\theta[.]$ is the expectation operator with respect to model (7). A general characterization of the BUP of $T(\mathbf{y})$ is considered in Rodrigues et al. (1985). The result that follows relates B(L)UP and MREP in finite populations.

Theorem 3.1. *Let the loss fuction be squared error. Then, under model (7),(i) an MREP of $T(\mathbf{y})$ is unbiased and (ii) if the B(L)UP exists and is equivariant it is an MREP.*

Application 3.1. Let $y_1, \ldots, y_N$, be independent and $y_i \sim N(\theta; \sigma^2)$, where σ^2 is known. Then it is well known that $\hat{T}_E(\mathbf{y}_s) = N\bar{y}_s$ (the usual expansion predictor), where $\bar{y}_s = \sum_{i \in s} y_i/n$, is the BUP of $T(\mathbf{y})$. Since $\hat{T}(\mathbf{y}_s)$ is location–equivariant, it follows from Theorem 3.1 that $\hat{T}_E(\mathbf{y}_s)$ is the MREP.

The result that follows establishes a least favorable property of the normal distribution in finite populations.

Theorem 3.2. *Let $y_1, \ldots, y_N$, be independent and distributed according to $F(y_i - \theta)$, which belongs to the class of all univariate distributions having a density, where $E[y_i] = \theta$, and $Var[y_i] = \sigma^2$ is known, i=1,...,N. Let r_F be the risk function of the MREP of $T(\mathbf{y})$, with respect to the distribution F. Then, r_F takes on its maximum value when F is $N(\theta; \sigma^2)$.*

Proof. According to Application 3.1, the MREP of T under the $N(\theta; \sigma^2)$ is $\hat{T}_E(\mathbf{y}_s)$, with risk $r_N = \sigma^2 N(N-n)/n$. Since the risk of $\hat{T}_E(\mathbf{y}_s)$ is r_N, whatever be F, the MREP for any F must have risk $r_F \leq r_N$, and the result follows.

The following result establishes the optimality of $\hat{T}_E(\mathbf{y}_s)$ under a more general set up than that of Application 3.1. We provide an alternative proof to this fact that easily generalizes to more general situations.

Lemma 3.2. *Let $y_1, \ldots, y_N$ be independent and suppose that $E[y_i] = \theta$, and $Var[y_i] = \sigma^2$, $i = 1, \ldots, N$, which we call the W-model. Then, $\hat{T}_E(\mathbf{y}_s)$ is the BLUP of $T(\mathbf{y})$ under the W-model.*

Proof. The model that follows by adding the normality assumption to the W–model, we call the N–model. Therefore, it is easily checked that, for any linear predictor, $\hat{T}(\mathbf{y}_s) = \sum_{i \in s} c_i y_i$, for some known constants c_i (note that $\hat{T}_E(\mathbf{y}_s)$ is a linear predictor), $E_W[\hat{T}(\mathbf{y}_s) - T(\mathbf{y})]^2 = E_N[\hat{T}(\mathbf{y}_s) - T(\mathbf{y})]^2$. Now the result follows from the optimality of $\hat{T}_E(\mathbf{y}_s)$ under the N–model. $E_W[.](E_N[.])$ denotes the expectation operator defined with respect to the $W(N)$–model.

Application 3.2. If the normality assumption is dropped from the Application 3.1, it follows from Lemma 3.2 and Theorem 3.1 that $\hat{T}_E(\mathbf{y}_s)$ is the MREP within the class of all linear predictors.

Application 3.3. Suppose now that $y_1, \ldots, y_N$ are independent and distributed according to the density $e^{-(y_i - \theta)}; y_i \geq \theta, i = 1, \ldots, N$. It follows from Theorem 1 in Rodrigues et al. (1985) that the BUP of $T(\mathbf{y})$, based on the observed sample $\mathbf{y}_s = (y_1, \ldots, y_n)$, is

$$\hat{T}^*(\mathbf{y}_s) = n\bar{y}_s + (N-n)(y_{s(1)} - 1/n),$$

where $y_{s(1)} = min(y_1,\dots,y_n)$. Now, since $\hat{T}^*(\mathbf{y}_s)$ is equivariant, it is, according to Theorem 3.1, the MREP of $T(\mathbf{y})$.

To derive an explicit and general expression for the MREP with respect to model (7), let's begin by finding the most general location equivariant predictor.

Lemma 3.3. *Let $\hat{T}_o(\mathbf{y}_s) = n\bar{y}_s + \hat{T}_{o2}(\mathbf{y}_s)$ be any equivariant predictor of $T(\mathbf{y})$ under model (7). Then, $\hat{T}(\mathbf{y}_s)$ is equivariant if and only if $\hat{T}(\mathbf{y}) = \hat{T}_o(\mathbf{y}_s) - w(\mathbf{z}_s)$, where $\mathbf{z}_s = (z_1,\dots,z_{n-1})$, and $z_i = y_i - y_n, i = 1,\dots,n-1$.*

The proof parallels that of Theorem 3.1.2 in Lehmann (1983), and so, is omited. The next theorem provides a general characterization for the MREP under model (7). The proof is also omited.

Theorem 3.3. *The MREP of $T(\mathbf{y})$ with respect to any loss function satisfying (10) is given by*

$$\hat{T}^*(\mathbf{y}_s) = \hat{T}_o(\mathbf{y}_s) - w^*(\mathbf{z}_s), \tag{12}$$

where $w^*(\mathbf{z}_s)$ minimizes

$$E_o\{\Delta[\hat{T}(\mathbf{y}_s) - w(\mathbf{z}_s) - T(\mathbf{y})]|\mathbf{z}_s\}, \tag{13}$$

and $\hat{T}_o(\mathbf{y}_s)$ is any location–equivariant predictor with finite risk function.

In (13), $E_o[.]$ means that the expectation operator is computed with respect to $\theta = 0$, according to Corollary 2.1.

Corollary 3.1. *Under the squared error loss fuction (10), the MREP of $T(\mathbf{y})$ is $\hat{T}^*(\mathbf{y}_s)$ of (12), with*

$$w^*(z_s) = E_o[\hat{T}_{o2}(\mathbf{y}_s) - T_2(\mathbf{y}_r)|\mathbf{z}_s], \tag{14}$$

where, $\hat{T}(\mathbf{y}) = n\bar{y}_s + \hat{T}_{o2}(\mathbf{y}_s), T(\mathbf{y}) = n\bar{y} + T_2(\mathbf{y}_r)$, and $T_2(\mathbf{y}_r) = \sum_{i\in r} y_i = (N-n)\bar{y}_r$.

From Corollary 3.1, $\hat{T}_E(\mathbf{y}_s)$ is the MREP under a distribution F if and only if $E_o[\bar{y}_s|\mathbf{z}_s] = 0$, a result that holds if and only if F is the normal distribution (see Kagan, Linik and Rao, 1965). Therefore, $\hat{T}_E(\mathbf{y}_s)$ is the MREP only under the normal model.

The next theorem presents a Pitman–location predictor for finite populations.

Theorem 3.4. *The MREP of $T(\mathbf{y})$ with respect to the squared squared error loss function (11) is*

$$\hat{T}^*(\mathbf{y}_s) = n\bar{y}_s + (N-n)\frac{\int_{-\infty}^{\infty} u f(y_1 - u,\dots,y_n - u)du}{\int_{-\infty}^{\infty} f(y_1 - u,\dots,y_n - u)du}. \tag{15}$$

Proof. It is easily seen that $\hat{T}_o(\mathbf{y}_s) = n\bar{y}_s + (N-n)y_n$ is equivariant. According to Corollary 3.1, under the squared error loss function (11), the MREP of $T(\mathbf{y})$ is $\hat{T}(\mathbf{y}_s) = \hat{T}_o(\mathbf{y}_s) - w^*(\mathbf{z}_s)$, where $w^*(\mathbf{z}_s) = (N-n)E_o[(y_n - \bar{y}_r)|\mathbf{z}_s] = E_o[y_n|\mathbf{z}_s]$, since $\bar{y}_r$ and $\mathbf{z}_s$ are independent and $E_o[\bar{y}_r] = 0$. Therefore, (15) follows from the fact that

$$E_o[y_n|\mathbf{z}_s] = y_n - \frac{\int_{-\infty}^{\infty} u f(y_1,\ldots,y_n)du}{\int_{-\infty}^{\infty} f(y_1-u,\ldots,y_n-u)du},$$

a result that follows from (7.2.9) in Zacks (1971).

We refer to predictor (15) as the Pitman–location predictor of the population total $T(\mathbf{y})$, by its similarities with the well known Pitman estimator of location.

Application 3.4. let $y_1,\ldots,y_N$ be independent and distributed according to the density $f(y_i - \theta) = 1$; $-1/2 \leq y_i - \theta \leq 1/2$, $i = 1,\ldots,N$, and zero otherwise. Since, after some algebraic manipulations , $E_o[y_n|\mathbf{z}_s] = y_n - (y_{s(1)} + y_{s(n)})/2$, it follows from (15) that the MREP of $T(\mathbf{y})$ is

$$\hat{T}(\mathbf{y}_s) = n\bar{y}_s + (N-n)(y_{s(1)} + y_{s(2)})/2,$$

where $y_{s(1)} = min(y_1,\ldots,y_n)$, and $y_{s(n)} = max(y_1,\ldots,y_n)$.

The main results of this section extends easily to the case of stratified populations.

4. THE GENERAL LINEAR MODEL

In this section, it is considered that associated with unit i there is a random variable y_i, where

(16) $$y_i \sim F(y_i - \xi_i),$$

$i = 1,\ldots,N$, F has density f, and $\xi = (\xi_1,\ldots,\xi_N) \in \Pi_\Omega$, a p–dimensional subspace of $\mathcal{R}^N$. It follows that model (16) remains invariant under the transformations

(17) $$y_i^* = y_i + c_i \text{ and } \xi_i^* = \xi_i + c_i,$$

with $c_i \in \Pi_\Omega, i = 1,\ldots,N$. Therefore, when expressed in the new coordinate system a predictor $\hat{T}(\mathbf{y}_s)$ should satisfy

(18) $$\hat{T}(\mathbf{y}_s^*) = \hat{T}(\mathbf{y}_s) + \sum_{i=1}^{N} c_i,$$

which is the equivariance condition under (17). Therefore, the problem of predicting $T(\mathbf{y})$ remains invarint under the transformations (17) if (10) is valid, a condition satisfied by the squared error loss function (11). Lemma 3.1 continues to hold in the present situation, since (17) induces a transitive group on Π_Ω. Similarly for Theorem 3.1.

Let $\hat{\xi}$, be the weighted least squares estimate of ξ based on $\mathbf{y}_s$, that is, $\hat{\xi} = (\hat{\xi}_1, \ldots, \hat{\xi}_N)$ minimize $\sum_{i \in s}(y_i - \xi_i)^2/\sigma_i^2$, where $Var[y_i] = \sigma_i^2$ is known, subject to the restriction that $\xi \in \Pi_\Omega$. The following theorem is a direct extension of Theorem 3.1.

Theorem 4.1. *If the y_i are independent and $N(\xi_i; \sigma_i^2)$, $i = 1, \ldots, N$, the BUP of $T(\mathbf{y})$,*

$$\hat{T}^*(\mathbf{y}_s) = n\bar{y}_s + \sum_{i \in r} \hat{\xi}_i, \tag{19}$$

where $\hat{\xi}_i$ is the weighted least squares estimator of ξ_i, is the MREP of $T(\mathbf{y})$.

If the normality assumption is dropped, it follows, by extending Lemma 3.2 to the present situation, that $\hat{T}^*(\mathbf{y}_s)$ in (19) is the MREP of $T(\mathbf{y})$ in the class of all linear predictors.

Application 4.1. Let $y_1, \ldots, y_N$ be independent and where $y_i \sim N(\beta x_i; \sigma^2 x_i)$, $i = 1, \ldots, N$. It is well known that the BUP of $T(\mathbf{y})$ in the present situation ($\xi_i = x_i\beta, i = 1, \ldots, N$), is

$$\hat{T}^*(\mathbf{y}_s) = n\bar{y}_s + \sum_{i \in r} x_i \hat{\beta} = \sum_{i=1}^{N} x_i \frac{\sum_{i \in s} y_i}{\sum_{i \in s} x_i}, \tag{20}$$

the usual ratio estimator. In this case, the subspace Π_Ω has dimension 1 and is generated by the vector $\mathbf{x}' = (x_1, \ldots, x_N)$. Since $\hat{T}^*(\mathbf{y}_s)$ in (20) is equivariant under the transformations (17), it follows from Theorem 4.1 that $\hat{T}^*(\mathbf{y}_s)$ is the MREP. If the normality assumption is dropped, $\hat{T}^*(\mathbf{y}_s)$ in (20) is the MREP of $T(\mathbf{y}_s)$ within the class of all linear equivariant predictors.

Application 4.2. Let $y_1, \ldots, y_N$ be independent and where $y_i \sim N(\alpha + \beta x_i; \sigma^2)$, $i = 1, \ldots, N$. It is well known that the BUP of $T(\mathbf{y})$ in the present situation ($\xi_i = \alpha + \beta x_i, i = 1, \ldots, N$) is

$$\hat{T}^*(\mathbf{y}_s) = N\bar{y}_s + N(\bar{X} - \bar{X}_s)\hat{\beta}, \tag{21}$$

the usual ratio estimator, where $\hat{\beta} = \sum_{i \in s}(x_i - \bar{X}_s)y_i / \sum_{i \in s}(x_i - \bar{X}_s)^2$ is the least squares estimator of β, $\bar{X} = \sum_{i=1}^{N} x_i/N$, and $\bar{X}_s = \sum_{i \in s} x_i/n$. In this case, the subspace Π_Ω has dimension 2 and is generated by the vectors $(1, \ldots, 1)$ and $(x_1, \ldots, x_N)$. Since $\hat{T}^*(\mathbf{y}_s)$ is equivariant under the transformations (17), it is, according to Theorem 4.1, the MREP of $T(\mathbf{y})$. If the normality assumption is dropped, it is the MREP within the class of all linear equivariant predictors.

5. EQUIVARIANT PREDICTION UNDER SCALE MODELS

In this section, it is considered that $y_1, \dots, y_N$ are independent and

$$y_i \sim \frac{1}{\tau} f\left(\frac{y_i}{\tau}\right), \tag{22}$$

$i = 1, \dots, N$. Model (22) remains invariant under the transformations

$$y_i^* = by_i \text{ and } \tau^* = b\tau, b \geq 0, \tag{23}$$

$i = 1, \dots, N$. Under the transformations (23), it follows that $T(\mathbf{y}) = bT(\mathbf{y})$. Therefore, predictor $\hat{T}(\mathbf{y}_s)$ is equivariant under the transformations (23), provided

$$\hat{T}(\mathbf{y}^*) = b\hat{T}(\mathbf{y}_s). \tag{24}$$

A predictor satisfying (24), is called a scale–equivariant predictor. Then, the problem of predicting $T(\mathbf{y})$ remains invariant under the transformations (23) provided

$$L(\hat{T}(\mathbf{y}_s); T(\mathbf{y})) = \Delta\left[\frac{\hat{T}(\mathbf{y}_s) - T(\mathbf{y})}{\tau}\right].$$

Note that the squared error loss function

$$L(\hat{T}(\mathbf{y}_s); T(\mathbf{y})) = \frac{(\hat{T}(\mathbf{y}_s) - T(\mathbf{y}))^2}{\tau^2}, \tag{25}$$

satisfies the required condition.

Now, since the group of transformations induced by (23) is transitive, it follows from Corollary 2.1 that the risk function of any equivariant predictor is constant. So, one might expect an MREP to exist. As before, let $T(\mathbf{y}) = n\bar{y}_s + T_2(\mathbf{y})$, be the population total, where $T_2(\mathbf{y}) = (N - n)\bar{y}_r$. The result that follows parallels that of Theorems 3.3 for the location model. The proof is omited.

Theorem 5.1. *Let $\hat{T}_o(\mathbf{y}) = n\bar{y}_s + \hat{T}_{o2}(\mathbf{y}_s)$ be any equivariant predictor. Then, a necessary and sufficient condition for $\hat{T}(\mathbf{y}_s) = n\bar{y}_s + \hat{T}_2(\mathbf{y}_s)$ to be scale–equivariant is that $\hat{T}_2(\mathbf{y}_s) =$*

$\hat{T}_{o2}(\mathbf{y}_s)/w(\mathbf{z}_s)$, where $\mathbf{z}_s = (y_1/y_n, \ldots, y_{n-1}/y_n, |y_n|/y_n)$. If there is a number $w^*(\mathbf{z}_s)$ that minimizes

$$E_1\left\{\Delta\left[\frac{\hat{T}_{o2}(\mathbf{y}_s)}{w(\mathbf{z}_s)} - T_2(\mathbf{y}_s)\right]|\mathbf{z}_s\right\}, \tag{26}$$

then

$$\hat{T}^*(\mathbf{y}_s) = n\bar{y}_s + \frac{\hat{T}_{o2}(\mathbf{y}_s)}{w^*(\mathbf{z}_s)} \tag{27}$$

is an MREP of $T(\mathbf{y})$.

In (26) above, $E_1[.]$ indicates that the expectation operator is taken with respect to $r = 1$, according to Corollary 2.1.

Corollary 5.1. *Under the squared error loss function (25), the MREP of $T(\mathbf{y})$ is $\hat{T}^*(\mathbf{y}_s)$ of (27), with*

$$w^*(\mathbf{z}_s) = \frac{E_1[\hat{T}_{o2}^2(\mathbf{y}_s)|\mathbf{z}_s]}{E_1[\hat{T}_{o2}(\mathbf{y}_s)T_2(\mathbf{y}_r)|\mathbf{z}_s]}.$$

Application 5.1. Let $y_1, \ldots, y_N$ be independent and where $y_i \sim \frac{1}{\tau}e^{-y_i/\tau}, i = 1, \ldots, N$. It is easily checked that $\hat{T}_o = n\bar{y}_s + (N-n)\bar{y}_s$ is scale equivariant. Now, $\bar{y}_r$, $\bar{y}_s$ and $\mathbf{z}_s$ are independent (Basu's Theorem). Therefore, it follows from Corollary 5.1, with respect to the squared error loss function (25) that

$$w^*(\mathbf{z}_s) = \frac{E_1[\bar{y}_s^2|\mathbf{z}_s]}{E_1[\bar{y}_s\bar{y}_r|\mathbf{z}_s]} = \frac{n+1}{n},$$

since $E_1[\bar{y}_s^2] = (n+1)/n$, and $E_1[\bar{y}_s] = E_1[\bar{y}_r] = 1$. From (27), it follows that the MREP of $T(\mathbf{y})$ is given by

$$\hat{T}^*(\mathbf{y}_s) = n\bar{y}_s + (N-n)\frac{\sum_{i\in s} y_i}{n+1}.$$

Application 5.2. Suppose now that $y_1, \ldots, y_N$ are independent and such that $y_i \sim U(0,\theta), i = 1, \ldots, N$, and U stands for the uniform distribution. It is easy to see that predictor $\hat{T}_o(\mathbf{y}_s) = n\bar{y}_s + (N-n)y_{s(n)}$, where $y_{s(n)} = max(y_1, \ldots, y_n)$, is scale-equivariant. Since $y_{s(n)}$ is independent of $\mathbf{z}_s$, it follows from Corollary 5.1, under squared error loss function, that

$$w^*(\mathbf{z}) = \frac{E_1[y_{s(n)}^2]}{E_1[y_{s(n)}]E_1[\bar{y}_r]} = \frac{2(n+1)}{n+2}.$$

From (27), it follows that the MREP of $T(\mathbf{y})$ is

$$\hat{T}^*(\mathbf{y}_s) = n\bar{y}_s + (N-n)\frac{n+2}{2(n+1)}y_{s(n)}.$$

Under squared error loss function and after some algebraic manipulations, it can be shown that the general form of the MREP of $T(\mathbf{y})$ is

$$\hat{T}^*(\mathbf{y}_s) = n\bar{y}_s + (N-n)\frac{c\int_{-\infty}^{\infty} u^n f(uy_1,\ldots,uy_n)du}{\int_{-\infty}^{\infty} u^{n+1} f(uy_1,\ldots,uy_n)du}$$

where $c = E_1[y_i]$ and we call this predictor, the Pitman–scale predictor of $T(\mathbf{y})$.

6. EQUIVARIANT PREDICTION UNDER LOCATION–SCALE MODELS

Suppose now that $y_1,\ldots,y_N$ are independent and where

$$y_i \sim \frac{1}{\tau} f\left(\frac{y_i-\theta}{\tau}\right), \tag{28}$$

$i = 1,\ldots,N$, with both parameters unknown. Model (28) remains invariant under the transformations

$$y_i^* = by_i + a, \theta^* = b\theta + a \text{ and } \tau^* = b\tau, \tag{29}$$

where $b \geq 0$. Under the transformations (29), it follows that $T(\mathbf{y}^*) = a + bT(\mathbf{y})$. Therefore, it follows that a predictor $\hat{T}(\mathbf{y}_s)$ is equivariant under (29) provided

$$\hat{T}(\mathbf{y}_s^*) = a + b\hat{T}(\mathbf{y}_s). \tag{30}$$

A predictor satisfying (30) is called a location–scale equivariant predictor. Also, the problem of predicting $T(\mathbf{y})$ remains invariant under transformations (29) if

$$L(\hat{T}(\mathbf{y}_s); T(\mathbf{y})) = \Delta\left[\frac{\hat{T}(\mathbf{y}_s) - T(\mathbf{y})}{\tau}\right]. \tag{31}$$

Note that the loss function (25) satisfies condition (31). Note also that the transformations (29) induces a transitive group. Therefore, according to Corollary 2.1 the risk function of any location scale equivariant predictor is constant.

Lemma 6.1. *Suppose there exists an MREP $\hat{T}^*(\mathbf{y}_s)$ of $T(\mathbf{y})$ under model (28), for each fixed (known) τ, with respect to the transformations (8) and that $\hat{T}^*(\mathbf{y}_s)$ is independent of τ and satisfies (30). Then, $\hat{T}^*(\mathbf{y}_s)$ is an MREP of $T(\mathbf{y})$ among all predictors satifying (30) when τ is unknown.*

Application 6.1. Let $y_1, \ldots, y_N$ be independent and $y_i \sim N(\theta; \sigma^2)$, $i = 1, \ldots, N$, both parameters being unknown. It follows from Application 3.1 and Lemma 6.1 that $\hat{T}_E(\mathbf{y}_s)$ is the MREP of $T(\mathbf{y})$ under the transformations (29).

The theorem that follows provides an explicit and general expression for the MREP under the location–scale model (28) by first finding the most general equivariant predictor.

Theorem 6.1. *Let $\hat{T}_o(\mathbf{y}_s) = n\bar{y}_s + \hat{T}_{o2}(\mathbf{y}_s)$ be any location–scale equivariant predictor. Then, $\hat{T}(\mathbf{y}_s) = n\bar{y}_s + \hat{T}_2(\mathbf{y}_s)$ is location equivariant if and only if $\hat{T}_2(\mathbf{y}_s) = \hat{T}_{o2}(\mathbf{y}_s) - w(\mathbf{z}_s)\delta_1(\mathbf{y}_s)$, where $\delta_1(b\mathbf{y}_s) = b\delta_1(\mathbf{y}_s)$ and*

$$\mathbf{z}_s = \left(\frac{y_1 - y_n}{y_{n-1} - y_n}, \ldots, \frac{|y_{n-1} - y_n|}{y_{n-1} - y_n}\right).$$

The MREP is given by

$$\hat{T}^*(\mathbf{y}_s) = n\hat{y}_s + \hat{T}_{o2}(\mathbf{y}_s) - w^*(\mathbf{z}_s)\delta_1(\mathbf{y}_s), \tag{32}$$

where $w^*(\mathbf{z}_s)$ minimizes

$$E_{o,1}\{\Delta[\hat{T}_{o2}(\mathbf{y}_s) - w^*(\mathbf{z}_s)\delta_1(\mathbf{y}_s) - T_2(\mathbf{y}_r)] | \mathbf{z}_s\}. \tag{33}$$

As before, $T(\mathbf{y}) = n\bar{y}_s + T_2(\mathbf{y}_s)$, where $T_2(\mathbf{y}_s) = (N - n)\bar{y}_r$ corresponds to the sum of the units not in the sample. In (33), $E_{o,1}[.]$ is the expectation operator taken with respect to $\theta = 0$ and $\tau = 1$.

Corollary 6.1. *Under the squared error loss function (25), it follows that the MREP of $T(\mathbf{y})$ is $\hat{T}(\mathbf{y}_s)$ in (32), with*

$$w^*(\mathbf{z}_s) = \frac{E_{o,1}\{[\hat{T}_{o2}(\mathbf{y}_s) - T_2(\mathbf{y}_r)]\delta_1(\mathbf{y}_s) | \mathbf{z}_s\}}{E_{o,1}[\delta_1^2(\mathbf{y}_s) | \mathbf{z}_s]}$$

Application 6.2. Let $y_1, \ldots, y_N$ be independent and

$$y_i \sim \frac{1}{\tau} e^{-(y_i - \theta)/\tau}; y_i \geq \theta,$$

$i = 1, \ldots, N$, where θ and τ are unknown. It is clear that $\hat{T}(\mathbf{y}_s) = n\bar{y}_s + (N - n)y_{s(1)}(\hat{T}_{o2}(\mathbf{y}_s) = (N - n)y_{s(1)})$, where $y_{s(1)} = min(y_1, \ldots, y_N)$, is location equivariant. Also, let $\delta_1(\mathbf{y}_s) = \sum_{i \in s}(y_i - y_{s(1)})$. Note that $\delta_1(\mathbf{y}_s)$ satisfies $\delta_1(b\mathbf{y}_s) = b\delta_1(\mathbf{y}_s)$. By using Basu's Theorem, it follows that $\hat{T}_{o2}(\mathbf{y}_s)$ and $\delta_1(\mathbf{y}_s)$ are independent of each other and also independent of $\mathbf{z}_s$. Therefore, from Corollary 6.1, it follows that $w^*(\mathbf{z}_s) =$

$-(N-n)(n-1)/n^2$, so that the MREP of $T(\mathbf{y}_s)$, using the squared error loss function (25) is

$$\hat{T}^*(\mathbf{y}_s) = n\bar{y}_s + (N-n)\left(y_{s(1)} + \frac{(n-1)}{n^2}\sum_{i\in s}(y_i - y_{s(1)})\right).$$

Application 6.3. Let $y_1, \ldots, y_N$ be independent and such that $f(y_i) = \tau, \theta - \tau/2 \le y_i \le \theta + \tau/2$, and zero otherwise, $i = 1, \ldots, N$. Therefore, it follows from Lemma 6.1 and Application 3.4 that

$$\hat{T}^*(\mathbf{y}_s) = n\bar{y}_s + (N-n)(y_{s(1)} + y_{s(2)})/2$$

is the MREP of $T(\mathbf{y})$ with respect to the squared error loss function (25).

ACKNOWLEDGEMENT

Special thanks are due to professor S. Zacks for making valuable comments and suggestions on an early version of this paper.

References

Basu, D. (1955) *On Statistics Independent of a Complete and Sufficient Statistic.* Sankhya, A, 15, 377-380.

Bolfarine, H., Pereira, C.A.B., Rodrigues, J. (1987) *Robust Linear Prediction in Finite Populations-A Bayesian Perspective.* Sankhya, B, 49, 23–35

Bolfarine, H. (1987) *Minimax Prediction in Finite Populations.* Commun. in Statist., A, 3683–3700

Kagan, A., Linnik, Y.V., Rao, C.R. (1965) *On a Characterization of the Normal Law Based on a Property of the Sample Average.* Sankhya, A, 405–406.

Lhemann, E. L. (1983) *Theory of Point Estimation.* Wiley

Rodrigues, J., Bolfarine, H., Rogakto, A. (1985) *A General Theory of Prediction in Finite Populations.* Intern. Statistic. Rev., 53, 239–254.

Zacks, S. (1971) *The Theory of Statistical Inference.* Wiley.

Zacs S. and Solomon, H. (1981) *Bayes and Equivariant Estimators of the Variance of a Finite Population. Simple Random Sampling.* Commun. in Statist., A, 10, 407–426.

Zacks, S. (1981) *Bayes Equivariant Estimators of the Variance of a Finite Population for Exponential Priors..* Commun. in Statistic., A, 10, 427-437.

RELATÓRIO TÉCNICO
DO
DEPARTAMENTO DE ESTATÍSTICA

TÍTULOS PUBLICADOS EM 1987

8701 - ACHCAR, J.A. & BOLFARINE, H.; Constant Hazard Against a Change-Point Alternative: A Bayesian Approach with Censored Data, São Paulo, IME-USP, 1987, 20p.

8702 - RODRIGUES, J.; Some Results on Restricted Bayes Least Squares Predictors for Finite Populations, São Paulo, IME-USP, 1987, 16p.

8703 - LEITE, J.G., BOLFARINE, H. & RODRIGUES, J.; Exact Expression for the Posterior Mode of a Finite Population Size: Capture-Recapture Sequential Sampling, São Paulo, IME-USP, 1987, 14p.

8704 - RODRIGUES, J., BOLFARINE, H. & LEITE, J.G.; A Bayesian Analysis in Closed Animal Populations from Capture Recapture Experiments with Trap Response, São Paulo, IME-USP, 1987, 21p.

8705 - PAULINO, C.D.M.; Analysis of Categorical Data with Full and Partial Classification: A Survey of the Conditional Maximum Likelihood and Weighted Least Squares Approaches, São Paulo, IME-USP, 1987, 52p.

8706 - CORDEIRO, G.M. & BOLFARINE, H.; Prediction in a Finite Population under a Generalized Linear Model, São Paulo, IME-USP, 1987, 21p.

8707 - RODRIGUES, J. & BOLFARINE, H.; Nonlinear Bayesian Least-Squares Theory and the Inverse Linear Regression, São Paulo, IME-USP, 1987, 15p.

8708 - RODRIGUES, J. & BOLFARINE, H.; A Note on Bayesian Least-Squares Estimators of Time-Varying Regression Coefficients, São Paulo, IME-USP, 1987, 11p.

8709 - ACHCAR, J.A., BOLFARINE, H. & RODRIGUES, J.; Inverse Gaussian Distribution: A Bayesian Approach, São Paulo, IME-USP, 1987, 20p.

8710 - CORDEIRO, G.M. & PAULA, G.A.; Improved Likelihood Ratio Statistics for Exponential Family Nonlinear Models, São Paulo, IME-USP, 1987, 26p.

8711 - SINGER, J.M.; PERES, C.A.& HARLE, C.E.; On the Hardy -Weinberg Equilibrium in Generalized ABO Systems, São Paulo, IME-USP, 1987, 16p.

8712 - BOLFARINE, H. & RODRIGUES, J.; A Review and Some Extensions on Distribution Free Bayesian Approaches for Estimation and Prediction, São Paulo, IME-USP, 1987, 19p.

8713 - RODRIGUES, J.; BOLFARINE, H. & LEITE, J.G.; A Simple Nonparametric Bayes Solution to the Estimation of the Size of a Closed Animal Population, São Paulo, IME-USP, 1987, 11p.

8714 - BUENO, V.C.; Generalizing Importance of Components for Multistate Monotone Systems, São Paulo, IME-USP, 1987, 12p.

8801 - PEREIRA, C.A.B. & WECHSLER, S.; On the Concept of P-value, São Paulo, IME-USP, 1988, 22p.

8802 - ZACKS, S., PEREIRA, C.A.B. & LEITE, J.G.; Bayes Sequential Estimation of the Size of a Finite Population, São Paulo,IME-USP, 1988, 23p.

8803 - BOLFARINE, H.; Finite Population Prediction Under Dynamic Generalized Linear Models, São Paulo, IME-USP, 1988, 21p.

8804 - BOLFARINE, H.; Minimax Prediction in Finite Populations, São Paulo, IME-USP, 1988, 18p.

8805 - SINGER, J.M. & ANDRADE, D.F.; On the Choice of Appropriate Error Terms for Testing the General Linear Hypothesis in Profile Analysis, São Paulo, IME-USP, 1988, 23p.

8806 - DACHS, J.N.W. & PAULA, G.A.; Testing for Ordered Rate Rations in Follow-up Studies with Incidence Density Data, São Paulo, IME-USP, 1988, 18p.

8807 - CORDEIRO, G.M. & PAULA, G.A.; Estimation, Significance Tests and Diagnostic Methods for the Non-Exponential Family Nonlinear Models, São Paulo, IME-USP, 1988, 29p.

8808 - RODRIGUES, J. & ELIAN, S.N.; The Coordinate - Free Estimation in Finite Population Sampling, São Paulo, IME-USP, 1988, 5p.

8809 - BUENO, V.C. & CUADRADO, R.Z.B.; On the Importance of Components for Continuous Structures, São Paulo, IME-USP, 1988, 14p.

8810 - ACHCAR, J.A., BOLFARINE, H & PERICCHI, L.R.; Some Applications of Bayesian Methods in Analysis of Life Data, São Paulo, IME-USP, 1988, 30p.

8811 - RODRIGUES, J.; A Bayesian Analysis of Capture-Recapture Experiments for a Closed Animal Population, São Paulo, IME-USP, 1988, 10p.

8812 - FERRARI, P.A.; Ergodicity for Spin Systems, São Paulo, IME-USP, 1988, 25p.

8813 - FERRARI, P.A. & MAURO, E.S.R.; A Method to Combine Pseudo-Random Number Generators Using Xor, São Paulo, IME-USP, 1988, 10p.

8814 - BOLFARINE, H. & RODRIGUES, J.; Finite Population Prediction Under a Linear Functional Superpopulation Model a Bayesian Perspective, São Paulo, IME-USP, 1988, 22p.

8815 - RODRIGUES, J. & BOLFARINE, H.; A Note on Asymptotically Unbiased Designs in Survey Sampling, São Paulo, IME-USP, 1988, 6p.

8816 - BUENO, V.C.; Bounds for the a Availabilities in a Fixed Time Interval for Continuous Structures Functions, São Paulo, IME-USP, 1988, 22 p.

8817 - TOLOI, C.M.C. & MORETTIN, P.A.; Spectral Estimation for Time Series with Amplitude Modulated Observations: A Review, São Paulo, IME-USP, 1988, 16p.

8818 - CHAYES, J.T.; CHAYES, L.; GRIMMETT, G.R.; KESTEN, H. & SCHONMANN, R.H.; The Correlation Length for the High Density Phase of Bernoulli Percolation, São Paulo, IME-USP, 1988, 46p.

8819 - DURRETT, R.; SCHONMANN, R.H. & TANAKA, N.I.: The Contact Process on a Finite Set, III: The Critical Case, São Paulo, IME-USP, 1988, 31p.

8820 - DURRET, R,; SCHONMANN, R.H. & TANAKA, N.I.; Correlation Lengths for Oriented Percolation, São Paulo, IME-USP, 1988, 18p.

8821 - BRICMONT, J.; KESTEN, H,; LEBOWITZ, J.L. & SCHONMANN, R,H,; A Note on the Ising Model in High Dimensions, São Paulo, IME-USP, 1988, 21p.

8822 - KESTEN, H, & SCHONMANN, R,H,; Behavior in Large Dimensions of the Potts and Heisenberg Models, São Paulo, IME-USP, 1988, 61p.

8823 - DURRET, R, & TANAKA, N.I.; Scaling Inequalities for Oriented Percolation, São Paulo, IME-USP, 1988, 21p.

3901 - RODRIGUES, J.; Asymptotically Design - Unbiased Predictors to Two-Stage Sampling, São Paulo, IME-USP, 1989, 9p.

8902 - TOLOI, C.M.C. & MORETTIN, P.A.; Spectral Analysis for Amplitude Modulated Time Series, São Paulo, IME-USP, 1989, 24p.

8903 - PAULA, G.A.; Influence Measures for Generalized Linear Models with Restrictions in Parameters, São Paulo, IME-USP, 1989, 18p.

8904 - MARTIN, M.C. & BUSSAB, W.O.; An Investigation of the Properties of Raking Ratio Estimators for Cell Frequencies with Simple Random Sampling, São Paulo, IME-USP, 1989, 11p.

8905 - WECHSLER, S.; Yet Another Refutation of Allais' Paradox, São Paulo, IME-USP, 1989, 6p.

8906 - BARLOW, R.E. & PEREIRA, C.A.B.; Conditional Independence and Probabilistic Influence Diagrams, São Paulo, IME-USP, 1989, 21p.

8907 - BARLOW, R.E.; PEREIRA, C.A.B. & WECHSLER, S.; The Bayesian Approach to Ess, São Paulo, IME-USP, 1989, 20p.

8908 - PEREIRA, C.A.B. & BARLOW, R.E.; Medical Diagnosis Using Influence Diagrams, São Paulo, IME-USP, 1989, 13p.

8909 - BOLFARINE, H.; A Note on Finite Population Prediction Under Asymmetric Loss Functions, São Paulo, IME-USP, 1989, 8p.

8910 - BOLFARINE, H.; Bayesian Modelling in Finite Populations, São Paulo, IME-USP, 1989, 8p.

8911 - NEVES, M.M.C & MORETTIN, P.A.; A Generalized Cochrane-Orcutt-Type Estimator for Time Series Regression Models, São Paulo, IME-USP, 1989, 29p.

8912 - MORETTIN, P.A., TOLOI, C.M.C., GAIT, N. & MESQUITA, A.R.: Analysis of the Relationships Between Some Natural Phenomena: Atmospheric Precipitation, Mean Sea Level and Sunspots, São Paulo, IME-USP, 1989, 35p.

8913 - BOLFARINE, H.; Population Variance Prediction Under Normal Dynamic Superpopulation Models, São Paulo, IME-USP, 1989, 7p.

8914 - BOLFARINE, H.; Maximum Likelihood Prediction in Two Stage Sampling, São Paulo, IME-USP, 1989, 4p.

8915 - WECHSLER, S.; Exchangeability and Predictivism, São Paulo, IME-USP, 1989, 10p.