AD-A062 577 CASE WESTERN RESERVE UNIV CLEVELAND OHIO DEPT OF MA--ETC F/G 12/1
BAYES AND EQUIVARIANT ESTIMATORS OF THE VARIANCE OF A FINITE PO--ETC(U)
NOV 78 S ZACKS, H SOLOMON N00014-75-C-0529
UNCLASSIFIED TR-35 NL

| OF |
AD
A062577

END
DATE
FILMED
3-79
DDC

(12) (11) LEVEL II

(6) BAYES AND EQUIVARIANT ESTIMATORS OF THE VARIANCE OF A FINITE POPULATION.

by

(10) S. Zacks and H. Solomon

(14) TR-

(9) Technical Report No. 35

November 27, 1978

(11) 27 Nov 78

(12) 23 p.

(15) PREPARED UNDER CONTRACT
NR-00014-75-C-0529 / PROJECT NR 042-276
OFFICE OF NAVAL RESEARCH

Department of Mathematics and Statistics
CASE WESTERN RESERVE UNIVERSITY
Cleveland, Ohio

Bayes and Equivariant Estimators of the Variance of a Finite Population

By

S. Zacks and H. Solomon

0. Introduction.

Let $x_1,\ldots,x_N$ be the values of a variable x that measures a characteristic in a finite population of N elements. Let

$$(0.1) \qquad \mu = \frac{1}{N}\sum_{i=1}^{N} x_i \ , \quad \sigma_N^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i-\mu)^2 \ ,$$

be the population mean and variance of the measurements. In the present paper the problem of estimating σ_N^2 on the basis of a sample $X_1,\ldots,X_n$, $2 \leq n \leq N$, from the population is studied. The commonly used estimator is the sample variance $\hat{\hat{\sigma}}_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X}_n)^2$, where $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ is the sample mean. It is well known that $\hat{\hat{\sigma}}_n^2$ is an unbiased estimator of $\frac{N}{N-1}\sigma^2$, under simple random sampling.

In the present study the "unbiased" estimator $\hat{\hat{\sigma}}_n^2$ is replaced by $\hat{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X}_n)^2$, which is called the "classical" estimator. The two estimators are nearly equivalent if the sample size is not trivially small. The "classical" estimator does not utilize any prior information on σ^2 that may be often available. There are many examples of repetitive sampling surveys, in agricultural or industrial areas, in which good information is available on the distribution of the seasonal or yearly yield (production) of a certain commodity. Samples may be taken during the season to observe the distribution of related characteristics that may improve the forecasting of a population value. Estimates of the variance in the population could be adjusted adaptively in order to imporve the prediction (confidence) intervals for population parameters.

In this paper we show how such prior information on the mean and variance of the population can be utilized to adjust the "classical" estimator. Specifically, by utilizing the special structure of the sample survey theoretical model and of the likelihood function we derive the general form of Bayes and Bayes Equivariant estimators. It is shown that for any prior distribution, H, of independent identically distributed variables $x_1,\ldots,x_N$, having a prior mean μ_0 and prior variance σ_0^2, the Bayes estimator of σ^2 for squared-error loss is (approximately)

$$\hat{\sigma}_B^2 = \frac{n}{N}\hat{\sigma}_n^2 + (1 - \frac{n}{N})[\sigma_0^2 + \frac{n}{N}(\bar{X}_n - \mu_0)^2] \quad . \tag{0.2}$$

Estimator (0.2) does not depend on the sampling procedure. This Bayes estimator is a weighted average of the "classical" estimator, $\hat{\sigma}_n^2$, based on the observed sample and the Bayes estimator of the "within" variance in the unobserved portion of the population and the estimator of the variance between the means of the observed and unobserved portions of the population. The estimator (0.2) could well be found very meaningful and good also in a non-Bayesian sense by considering μ_0 and σ_0^2 as proper estimates (or quesses) of the mean and variance of the unobserved part of the population.

Equivariant estimators of the variance σ_N^2 are considered with respect to the group $\mathcal{G}$ of real affine transformations on the parameter space $\mathcal{X}^{(N)}$ of $(x_1,\ldots,x_N)$. It is shown that every equivariant estimator of σ_N^2 can be expressed in the general form $\hat{\sigma}_n^2\psi(\underset{\sim}{u}_n)$, where $\psi(\underset{\sim}{u}_n)$ is a proper function of the maximal invariant statistic, which is the vector of standardized sample values. Bayes equivariant estimators are studied, with respect to the quadratic loss function $L(\hat{\sigma}_n^2\psi(\underset{\sim}{u}_n)\sigma_N^2) = (\hat{\sigma}_n^2\psi(\underset{\sim}{u}_n)-\sigma_N^2)^2/\hat{\sigma}_n^4$. In contrast

to the case of determining Bayes estimators, the form of the Bayes equivariant estimator depends strongly on the particular prior distribution specified for $x_1,\ldots,x_N$. For example, it is shown that

$$(0.3) \qquad \hat{\sigma}^2_{BE} = \hat{\sigma}^2_n(1 - \frac{3}{N})(1 + \frac{3}{n-3}) = \hat{\sigma}^2_n \frac{n}{n-3} + O(\frac{1}{N}) ,$$

is the Bayes equivariant estimator for prior normal i.i.d. variables, regardless of their prior mean and variance. The above formula (0.3) is relatively simple. It depends only on $\hat{\sigma}^2_n$ and does not depend on $\underset{\sim}{u}_n$. This is not always the case, as shown by Zacks [16] in the case of exponentially distributed i.i.d. variates.

In Section 1 we introduce the sample survey model and discuss sample statistics and likelihood functions. The Bayesian approach extending the sample survey model is discussed in Section 2. Sections 3 and 4 define and analyze equivariant estimators, Bayesian concepts and loss functions. Bayesian measures of relative efficiency are introduced in Section 5. We provide a numerical example in which fifty populations of size N = 100 were simulated from an exponential distribution. From each such population a sample of size n = 10 was drawn and the estimators $\hat{\sigma}^2_n$ and $\hat{\sigma}^2_B$ were computed. It is interesting to observe the extent to which the Bayes estimator $\hat{\sigma}^2_B$ is more effective than the classical estimator $\hat{\sigma}^2_n$ in small samples. Estimate of their Bayes relative efficiency is provided in that example. General efficiency analysis is provided for prior normal distributions. It is shown that the classical estimator is considerably less efficient than the Bayes estimator. Some sensitivity analysis is performed to study the effects of erroneous prior parameters on the relative efficiency.

There are only a few published papers on the Bayesian estimation of the variance of a finite population. Liu [9] considered unbiased estimators of σ_N^2 under various possible sampling designs. Since the sampling variance of these estimators depends on the population values $x_1,\ldots,x_N$, Liu considered the Bayes risk of these estimators. He derived a lower bound to the Bayes risk function and thus showed the optimality of the Horwitz-Thompson type estimator, under certain conditions. We remark that Liu's study is not really a Bayesian study, since proper Bayes estimators are independent of the sampling design and are generally not unbiased. Royall [10], [11] and Royall and Cumberland [12] studied the problem of developing confidence intervals of the population variance by regression estimates. We cannot compare their results with ours since the problems are different and so are the approaches.

2. Foundations.

Consider a finite population of size N whose units have values (real finite) $x_1,\ldots,x_N$. According to the modern theory of sampling surveys (see Godambe [5,6,7], Basu [2] and others) the population vector $\underset{\sim}{x}_N = (x_1,\ldots,x_N)$ is considered a parametric point in a parameter space $\mathfrak{x}^{(N)}$, which belongs to the Euclidean N-space. In the present paper a sample, $\underset{\sim}{s}$, of size n, $1 \leq n \leq N$, designates a subvector of $\underset{\sim}{x}_N$ consisting of n components $\underset{\sim}{s} = < x_{i_1},\ldots,x_{i_n} >$, where $i_j \in \{1,2,\ldots,N\}$ for all $j = 1,\ldots,N$. A sampling procedure is a plan according to which the components of $\underset{\sim}{x}_N$ are chosen. In a non-Bayesian theory of sampling surveys one has to introduce probability functions

$P(\underset{\sim}{s})$ on the sample space, $\mathcal{S}$, of all possible samples, in order to discuss random samples. In a Bayesian theory the parametric vector $\underset{\sim}{x}_N$ is considered a random vector having a prior joint distribution $H(\underset{\sim}{x}_N)$ on $\mathcal{X}^{(N)}$. According to this approach, the population vector, $\underset{\sim}{x}_N$, is a realization of a sample from a "superpopulation", generated (like in a Monte Carlo procedure) according to $H(\underset{\sim}{x}_N)$. According to this approach, given any sample $\underset{\sim}{s} = < x_{i_1},\ldots,x_{i_n} >$, the joint prior distribution of $\underset{\sim}{s}$ can be derived from $H(\underset{\sim}{x}_N)$ and the posterior joint distribution of $\underset{\sim}{x}^*_{N-n} = < x_\nu;\ \nu \notin \underset{\sim}{s} >$ and is <u>independent</u> of the sampling probability function $P(\underset{\sim}{s})$, which is immaterial for a Bayesian analysis (see Solomon and Zacks (1970)). For this reason we will assume in what follows, without loss of generality, that the sample consists of the subvector $\underset{\sim}{x}_n = (x_1,\ldots,x_n)$ and $\underset{\sim}{x}^*_{N-n} = (x_{n+1},\ldots,x_N)$. If $x_1,\ldots,x_N$ are assumed to be priorly independent and identically distributed then any sample $\underset{\sim}{s}$ can be considered a simple random sample from H, as in the classical model of inference.

The estimation problem is that of estimating a specified parametric function $\theta(\underset{\sim}{x}_N)$ of the population vector (e.g. the population mean, variance, etc.).

3. <u>Estimators of the Population Variance.</u>

3.1 <u>General Structure.</u>

Let $\underset{\sim}{x}_n = (x_1,\ldots,x_n)$ be an observed sample. Designate by $\bar{x}_n$, $\hat{\sigma}^2_n$ the sample mean and the sample variance, respectively; where

$$\bar{x}_n = \frac{1}{n}\sum_{i=1}^{n} x_i \quad \text{and} \quad \hat{\sigma}^2_n = \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x}_n)^2 .$$

If $\bar{x}^*_{N-n}$ and τ^2_{N-n} designate the mean and variance of the N-n units not in the sample then the population variance is

$$(3.1) \qquad \sigma^2(\underset{\sim}{x}_N) = \frac{n}{N} \cdot \hat{\sigma}^2_n + (1-\frac{n}{N})\tau^2_{N-n} + \frac{n}{N}(1-\frac{n}{N})(\bar{x}_n - \bar{x}^*_{N-n})^2 .$$

Formula (3.1) can be verified since total variance may be written as the average of the conditional variances plus the variance of the conditional expectations.

Estimators of the population variance are sample statistics with range in $(0,\infty)$. The most common estimators in use are the sample variance $\hat{\sigma}^2_n$ or the "unbiased estimator" $\hat{\hat{\sigma}}^2_n = \frac{n}{n-1}\hat{\sigma}^2_n$. Formula (3.1) shows that, regardless of the sampling procedure, a proper estimator of $\sigma^2(\underset{\sim}{x}_N)$ can be obtained by substituting estimators of τ^2_{N-n} and $\bar{x}^*_{N-n}$ in (3.1). The "unbiased" estimator $\hat{\hat{\sigma}}^2_n$ can be obtained from (3.1) by substituting $\bar{x}^*_{N-n} = \bar{x}_n$ and $\tau^2_{N-n} = \frac{n}{n-1} \cdot \frac{N-n+1}{N-1} \cdot \hat{\sigma}^2_n$. As will be shown in Section 4, Bayes estimators of $\sigma^2(\underset{\sim}{x}_N)$ can be obtained by substituting corresponding Bayes estimators for τ^2_{N-n} and $\bar{x}^*_{N-n}$ in (3.1).

3.2 Equivariant Estimators.

Following Fraser [4] we will denote by $[\alpha,\beta]$, with $-\infty < \alpha < \infty$ and $\beta \neq 0$, a real affine transformation, i.e. $[\alpha,\beta]x = \alpha+\beta x$. Let $\mathcal{H}$ denote the group of all such transformations. We define $[\alpha,\beta]\underset{\sim}{x}_N = (y_1,\ldots,y_N)$, where $y_i = [\alpha,\beta]\underset{\sim}{x}_i$, $i=1,\ldots,N$. Every element of $\mathcal{H}$ transforms $\mathcal{X}^{(N)}$ into $\mathcal{X}^{(N)}$ in a 1:1 fashion. Let $\mathcal{G}$ be the group of transformations on the parameter space of $\sigma^2(\underset{\sim}{x}_N)$ induced by the elements of $\mathcal{H}$. That is, if

$[\alpha,\beta]\underset{\sim}{x}_N = \underset{\sim}{y}_N$ then $\sigma^2(\underset{\sim}{y}_N) = \beta^2\sigma^2(\underset{\sim}{x}_N)$ where β^2 is the element of $\bar{\mathcal{H}}$ corresponding to $[\alpha,\beta]$ of G. An estimator $\sigma^2(\underset{\sim}{x}_n)$ is called equivariant with respect to G if, for every $[\alpha,\beta] \in \mathcal{H}$

$$\hat{\sigma}^2([\alpha,\beta]\underset{\sim}{x}_n) = \beta^2\hat{\sigma}^2(\underset{\sim}{x}_n) , \qquad \underset{\sim}{x}_n \in \mathcal{X}^{(n)} . \tag{3.2}$$

The sample variance $\hat{\sigma}_n^2$ is equivariant with respect to $\mathcal{H}$. The statistic

$$\underset{\sim}{u}_n = [-\frac{\bar{x}_n}{\hat{\sigma}_n} , \frac{1}{\hat{\sigma}_n}]\underset{\sim}{x}_n$$

is maximal invariant with respect to $\mathcal{H}$. Thus, every equivariant estimator of $\sigma^2(\underset{\sim}{x}_N)$ can be expressed in the form

$$\hat{\sigma}_\psi^2(\underset{\sim}{x}_n) = \hat{\sigma}_n^2\psi(\underset{\sim}{u}_n) , \tag{3.3}$$

where $\psi(\underset{\sim}{u}_n)$ is a proper positive function of the maximal invariant statistic $\underset{\sim}{u}_n$. For further reading on invariance structures for sampling from finite populations see Chaudhuri [3].

4. Bayes and Bayes Equivariant Estimators.

4.1 Bayes Estimators.

Let $H(\underset{\sim}{x}_N)$ be a prior distribution in a specified family $\mathcal{H}$. Let $L(\hat{\sigma}^2,\sigma^2)$ denote a loss function associated with estimating $\sigma^2(\underset{\sim}{x}_N)$ by $\hat{\sigma}^2(\underset{\sim}{x}_n)$. An estimator $\hat{\sigma}_H^2(\underset{\sim}{x}_n)$ is Bayes with respect to $H(\underset{\sim}{x}_N)$ and $L(\hat{\sigma}^2,\sigma^2)$ if it minimizes the prior risk function

$$R(\hat{\sigma}^2,H) = \int_{\mathcal{X}^{(N)}} L(\hat{\sigma}^2(\underset{\sim}{x}_n),\sigma^2(\underset{\sim}{x}_N))dH(\underset{\sim}{x}_N) . \tag{4.1}$$

The following is a general result for the squared-error loss function:

If $x_1,\ldots,x_N$ are i.i.d. random variables having <u>any</u> prior distribution $H(x)$ with a finite prior variance, σ_0^2, then the Bayes estimator of $\sigma^2(\underset{\sim}{x}_N)$ is

$$\hat{\sigma}_B^2 = \frac{n}{N}\hat{\sigma}_n^2 + (1 - \frac{n}{N})[\sigma_*^2 + \frac{n}{N}(\bar{x}_n - \mu_0)^2] \quad , \tag{4.2}$$

where $\sigma_*^2 = \sigma_0^2(1 - \frac{1}{N})$ and μ_0 is the prior expectation.

The proof of (4.2) proceeds as follows. The Bayes estimator of $\sigma^2(\underset{\sim}{x}_N)$, given $\underset{\sim}{x}_n$, for the squared-error loss function is the posterior expectation of (3.1). Furthermore, since the components of $\underset{\sim}{x}_N$ are priorly independent

$$E_H\{\tau_{N-n}^2 | \underset{\sim}{x}_n\} = E_H\{\tau_{N-n}^2\} = \frac{N-n-1}{N-n}\sigma_0^2 \quad , \tag{4.3}$$

for any prior distribution H, having variance σ_0^2. Moreover,

$$E_H\{(\bar{x}_n - \bar{x}_{N-n}^*)^2 | \underset{\sim}{x}_n\} = (\bar{x}_n - \mu_0)^2 + \frac{\sigma_0^2}{N-n} \quad . \tag{4.4}$$

Substituting these expressions in (3.1) one obtains (4.2). In many situations it is not unreasonable to assume that $x_1,\ldots,x_N$ are priorly i.i.d. Hence, formula (4.2) is a very general formula, since it does not depend on the form of $H(x)$, but only on the prior mean and variance. These values may be known from previous experience.

4.2 Bayes Equivariant Estimators.

Consider the structure of Bayes equivariant estimators. We have

$$\underset{\sim}{u}^*_{N-n} = [-\frac{\bar{x}_n}{\hat{\sigma}_n}, \frac{1}{\hat{\sigma}_n}]\underset{\sim}{x}^*_{N-n},$$

where $\underset{\sim}{x}^*_{N-n} = (x_{n+1},\ldots,x_N)$ and $\underset{\sim}{u}^*_{N-n}$ is maximal invariant with respect to $\mathcal{H}$. Let v_{N-n} and w^2_{N-n} be the mean and the variance of $\underset{\sim}{u}^*_{N-n}$. One can express the population variance in these terms in the form

$$\sigma^2(\underset{\sim}{x}_N) = \hat{\sigma}^2_n[\frac{n}{N} + (1 - \frac{n}{N})(w^2_{N-n} + \frac{n}{N} v^2_{N-n})] . \tag{4.5}$$

Thus, comparing (3.3) and (4.5), the ψ-function of an equivariant estimator should be chosen to estimate the function

$$D(w^2_{N-n}, v^2_{N-n}) = \frac{n}{N} + (1 - \frac{n}{N})(w^2_{N-n} + \frac{n}{N} v^2_{N-n}) . \tag{4.6}$$

Let $L(\psi,D)$ be a loss function for the estimation of $D(w^2_{N-n}, v^2_{N-n})$ by $\psi(\underset{\sim}{u}_n)$. $L(\psi,D)$ is invariant with respect to $\mathcal{H}$. Let $G(\underset{\sim}{u}_n, \underset{\sim}{u}^*_{N-n})$ be a prior distribution induced by $H(\underset{\sim}{x}_N)$. The prior risk associated with ψ and G is

$$R(\psi,G) = \int L(\psi(\underset{\sim}{u}_n), D(w^2_{N-n}, v^2_{N-n}))dG(\underset{\sim}{u}_n, \underset{\sim}{u}^*_{N-n}) . \tag{4.7}$$

An estimator $\hat{\sigma}^2_n\psi_G(\underset{\sim}{u}_n)$ is called Bayes equivariant if ψ_G minimizes (4.7).

Notice that the criterion of minimizing (4.7) is the same as minimizing the Bayes risk for the quadratic loss $L(\hat{\theta},\sigma^2_N) = (\hat{\theta}-\sigma^2_N)^2/\hat{\sigma}^4_n$, where $\hat{\theta} = \hat{\sigma}^2_n\psi(\underset{\sim}{u}_n)$.

In many applications it would be reasonable to assume that the family $\mathcal{H}$ of prior distributions is a family with location and scale parameters. In other words, assume that all the prior distributions of $\mathcal{H}$ are of the form

$$H\left(\frac{x_1-\mu_0}{\sigma_0}, \frac{x_2-\mu_0}{\sigma_0}, \ldots, \frac{x_N-\mu_0}{\sigma_0}\right),$$

where $-\infty < \mu_0 < \infty$ and $0 < \sigma_0 < \infty$. In this case the Bayes equivariant estimator depends only on the general form of $H(x_1,\ldots,x_N)$. Indeed, the distribution $G(\underset{\sim}{u}_n, \underset{\sim}{u}^*_{N-n})$ is the same for all μ_0 and σ_0 of distributions in $\mathcal{H}$.

Equivariant estimators in the <u>strict</u> sense were defined as those of the form $\hat{\sigma}^2_n \psi(\underset{\sim}{u}_n)$. The Bayes estimator $\hat{\sigma}^2_B$ (4.2) is thus not strictly equivariant. However, if $\underset{\sim}{x}_N$ is transformed to $[\alpha,\beta]\underset{\sim}{x}_N$ the prior parameters (μ_0,σ_0) should be transformed to $[\alpha,\beta](\mu_0,\sigma_0) = (\alpha+\beta\mu_0, |\beta|\sigma_0)$. Let $\hat{\sigma}^2_{(\mu_0,\sigma_0)}(\underset{\sim}{x}_n)$ denote the Bayes estimator $\hat{\sigma}^2_B$ with the prior parameters μ_0 and σ_0, respectively. Then $\hat{\sigma}^2_{(\mu_0,\sigma_0)}(\underset{\sim}{x}_n)$ is <u>generalized equivariant</u> in the sense that

$$\hat{\sigma}^2_{[\alpha,\beta](\mu_0,\sigma_0)}([\alpha,\beta]\underset{\sim}{x}_n) = \beta^2 \hat{\sigma}^2_{(\mu_0,\sigma_0)}(\underset{\sim}{x}_n), \tag{4.8}$$

for all $-\infty < \alpha < \infty$, $0 < \beta < \infty$; and all $\underset{\sim}{x}_n$. Furthermore, the Bayes estimator $\hat{\sigma}^2_{(\mu_0,\sigma_0)}(\underset{\sim}{x}_n)$ is also Bayes in the class of all generalized equivariant estimators with respect to the quadratic loss $(\hat{\theta}-\sigma^2_N)^2/\hat{\sigma}^4_n$.

4.3 Examples of Bayes Equivariant Estimators.

4.3.1 Normal Priors.

Suppose that $x_1,\ldots,x_N$ are priorly independent and identically distributed (i.i.d.) normal variables with prior mean μ_0 and prior variance σ_0^2. $\mathcal{H}$ consists of all such distributions with $-\infty < \mu_0 < \infty$ and $0 < \sigma_0^2 < \infty$. Under this model the sample statistics $\bar{x}_n$ and $\hat{\sigma}_n^2$ are independent of τ_{N-n}^2 and $\bar{x}_{N-n}^*$. The Bayes equivariant estimation is

$$(4.9) \qquad \hat{\sigma}_{BE}^2(\underset{\sim}{x}_n) = \hat{\sigma}_n^2\{\frac{n}{N} + (1-\frac{n}{N})[E\{w_{N-n}^2|\underset{\sim}{u}_n\} + \frac{n}{N}E\{v_{N-n}^2|\underset{\sim}{u}_n\}]\}.$$

We now show that w_{N-n}^2 and v_{N-n} are independent of $\underset{\sim}{u}_n$. Indeed, by the Bayes model $\underset{\sim}{x}_n$ and $\underset{\sim}{x}_{N-n}^*$ are independent. Hence $(\bar{x}_n,\hat{\sigma}_n,\underset{\sim}{u}_n)$ is independent of $\underset{\sim}{x}_{N-n}^*$. Furthermore, $(\bar{x}_n,\hat{\sigma}_n)$ is a complete sufficient statistic for the subfamily of prior distributions of $\underset{\sim}{x}_n$. Hence, from Basu's theorem (Basu, [1]) $\underset{\sim}{u}_n$ is independent of $(\underset{\sim}{x}_{N-n}^*,\bar{x}_n,\hat{\sigma}_n)$. Finally, since $\underset{\sim}{u}_{N-n}^*$ is a function of $(\underset{\sim}{x}_{N-n}^*,\bar{x}_n,\hat{\sigma}_n)$, $\underset{\sim}{u}_n$ and $\underset{\sim}{u}_{N-n}^*$ are independent. Hence,

$$(4.10) \qquad E\{w_{N-n}^2|\underset{\sim}{u}_n\} = E\{w_{N-n}^2\} = \frac{n}{N-n}\cdot\frac{N-n-1}{n-1}E\{F[N-n-1,n-1]\} = \frac{n}{n-3}\cdot\frac{N-n-1}{N-n}.$$

Similarly,

$$(4.11) \qquad E\{v^2_{N-n}|\underset{\sim}{u}_n\} = E\{v^2_{N-n}\} = \frac{N}{(N-n)(n-3)} .$$

Substituting these results in (4.9) we obtain as the Bayes equivariant estimator

$$(4.12) \qquad \hat{\sigma}^2_{BE}(\underset{\sim}{x}_n) = \hat{\sigma}^2_n(1 - \frac{3}{N})(1 + \frac{3}{n-3}) ,$$

which, in large populations is close to $\Sigma^n_{i=1} (x_i-\bar{x})^2/(n-3)$.

It is well known (see Zacks [17; pp. 346]) that the minimum mean-squared-error equivariant estimator of σ^2_0 in the i.i.d. case is $\Sigma^n_{i=1} (x_i-\bar{x})^2/(n+1)$. On the other hand, if the loss function is the quadratic loss $(\hat{\theta}-\sigma^2_0)^2/\hat{\sigma}^4_n$ the best equivariant estimator is $\Sigma^n_{i=1} (x_i-\bar{x})^2/(n-3)$. This confirms the above result.

4.3.2 Exponential Priors

Suppose that $x_1,\ldots,x_N$ are priorly i.i.d., with a common exponential distribution, with mean μ_0 (the prior variance is $\sigma^2_0 = \mu^2_0$). It is shown in [16] that the Bayes equivariant estimator is

$$(4.13) \qquad \hat{\sigma}^2_{BE} = \hat{\sigma}^2_n\{\frac{n}{N} + (1 - \frac{n}{N}) \frac{f^2(\underset{\sim}{u}_n)n^2}{(n-2)(n-3)} \} + O(\frac{1}{N}) ,$$

where $f(\underset{\sim}{u}_n) = -\min(u_1,\ldots,u_n)$. Thus, in contrast to the normal case, in the exponential priors model the Bayes equivariant estimator depends on $\underset{\sim}{u}_n$ too.

5. Bayes Risk Efficiency.

In the present section we introduce an index of *Bayesian efficiency* of estimators of σ^2_N. Given a prior distribution, $H(\underset{\sim}{x}_N)$, we denote by $R(\hat{\sigma}^2, H)$ the prior risk function of an estimator $\hat{\sigma}^2$ under H. In the present section we consider squared-error loss, $(\hat{\sigma}^2 - \sigma^2_N)^2$, only. Generalization to quadratic loss functions can be readily attained. Since the minimal prior risk is attained by the Bayes estimator $\hat{\sigma}^2_B$, with proper prior parameters μ_0, σ^2_0, we define the prior relative efficiency of an estimator $\hat{\sigma}^2$ as

$$RE(\hat{\sigma}^2, H) = \frac{R(\hat{\sigma}^2_B, H)}{R(\hat{\sigma}^2, H)} \quad . \tag{5.1}$$

For any estimator, $0 \leq RE(\hat{\sigma}^2, H) \leq 1$.

5.1 Exponential Priors-Example.

We provide now a numerical example of estimating the variance of a small population, $N = 100$, when the population variates are priorly i.i.d. exponential random variables with expectation $\mu_0 = 10$. The sample size is $n = 10$. In Table 1 we present the values of the classical and the Bayes estimators determined by 50 independent simulation runs. In each case we give also the value of σ^2_N. We see that generally the Bayes estimator is closer to the population variance. The prior relative efficiency of $\hat{\sigma}^2_n$ against $\hat{\sigma}^2_B$ is estimated to be $RE = .122$. Note $\hat{R}(\hat{\sigma}^2_n)$ and $\hat{R}(\hat{\sigma}^2_B)$ are the sample estimates of the prior mean-squared-errors

$E_H\{(\hat{\sigma}_n^2-\sigma_N^2)^2\}$ and $E_H\{(\hat{\sigma}_B^2-\sigma_N^2)^2\}$, respectively. We see in this example that the classical sample variance is very inefficient compared to the Bayes estimator. In the following example we show some analytical comparisons for the normal case.

5.2 Normal Priors.

The prior relative efficiency index (5.1) can be expressed also in the form

$$(5.2) \qquad RE(\hat{\sigma}^2,H) = \left[1 + \frac{E_H\{(\hat{\sigma}^2-\hat{\sigma}_B^2)^2\}}{E_H\{PVR\}}\right]^{-1}$$

where PVR is the posterior variance of the Bayes estimator $\hat{\sigma}_B^2$. In the case of prior i.i.d. normal (μ_0, σ_0^2) variables, one obtains

$$(5.3) \qquad E\{PVR(H,\underset{\sim}{x}_n)\} = \frac{2\sigma_0^4}{N}(1-f)\left(1-\frac{1}{N}(1-f)\right),$$

where $f = n/N$. Consider the sample variance $\hat{\sigma}_n^2$. Due to the prior independence of $\bar{x}_n$ and $\hat{\sigma}_n^2$ we obtain

$$(5.4) \qquad E\{[\hat{\sigma}_n^2 - \frac{n}{N}\hat{\sigma}_n^2 - (1-\frac{1}{n})[\sigma_0^2(1-\frac{1}{N}) + \frac{n}{N}(\bar{x}_n-\mu_0)^2]^2\}$$

$$= (1-f)^2\sigma_0^4 E\{[\frac{1}{n}\chi_1^2[n-1] - (1-\frac{1}{N}) - \frac{1}{N}\chi_2^2[1]]^2\},$$

where $\chi_1^2[n-1]$ and $\chi_2^2[1]$ designate independent chi-squared r.v.'s. From (5.3) and (5.4), the prior relative efficiency of σ_n^2, relative

Table 1. 50 Independent Simulation Runs of Exponential Populations of Size N = 100. Variance Estimates are Based on Sample of Size n = 10.

i	$\hat{\sigma}^2_n(i)$	$\hat{\sigma}^2_B(i)$	$\sigma^2_N(i)$
1	11.0279	91.3822	59.3401
2	25.9047	93.5383	99.7939
3	163.1974	108.4712	117.9033
4	114.8876	102.4455	112.7694
5	48.5473	96.1006	140.4537
6	24.1673	92.6005	95.3784
7	233.9413	115.0390	110.2580
8	24.2149	92.8096	73.1228
9	36.0928	93.7894	88.0842
10	58.1589	95.8768	92.6123
11	82.4155	98.2541	100.8301
12	110.1057	101.4133	82.7390
13	30.7288	94.2891	82.7665
14	92.7353	99.3392	66.7146
15	14.1656	92.9971	100.5613
16	30.6368	93.1324	87.0248
17	75.4294	97.7225	74.1037
18	123.1242	102.9664	104.8201
19	34.9740	95.1125	86.4942
20	226.5559	116.6118	118.4538
21	16.9147	92.2562	134.9088
22	98.5459	100.9711	72.2479
23	48.8585	94.9472	134.9985
24	54.3939	95.4909	114.8043
25	52.1598	95.2457	158.8236
26	55.5302	95.6501	79.0303
27	171.1030	107.1958	87.0796
28	105.0818	100.6807	128.0554
29	17.5409	92.9367	95.6511
30	23.1610	93.2519	94.2407
31	116.6341	101.8029	118.2461
32	27.0976	94.1436	74.1948
33	69.4181	98.1458	101.3439
34	20.7325	93.0247	98.1385
35	191.2939	112.5688	121.3687
36	199.9331	110.6755	105.3757
37	83.3183	98.5574	67.8614
38	17.2250	95.0543	77.5798
39	218.7134	112.1939	108.5281
40	30.0034	93.1121	83.3298
41	34.1185	93.4987	80.7266
42	32.7168	93.6423	106.1845
43	50.9459	95.3254	77.0860
44	41.4210	94.8028	64.2112
45	36.6681	94.1386	75.0202
46	22.9323	93.6654	88.6936
47	41.4523	94.1452	64.5216
48	149.5766	106.0161	99.7121
49	19.2172	101.9416	79.8605
50	64.9260	96.6517	60.5816

$\hat{R}(\hat{\sigma}^2_n) = 3887.67$

$\hat{R}(\hat{\sigma}^2_B) = 474.90$

$RE(\hat{\sigma}^2_n, H) = .12216$

to the prior normal distributions is:

$$(5.5) \qquad RE(\hat{\sigma}_n^2, H) = \left[1 + \frac{N(1-f)\left[\frac{3}{N^2} + \left(1-\frac{1}{N}\right)^2 - \left(1-\frac{1}{Nf}\right)^2\right]}{2\left(1-\frac{1}{N}(1-f)\right)} \right]^{-1}.$$

The relative efficiency function is independent of the prior parameters, since σ_0 is a scale parameter of the distribution. We therefore provide in the following table some relative efficiency values as functions of the sample fraction f and the population size, N.

Table 2. The Prior Relative Efficiency of $\hat{\sigma}_n^2$.

N\f	0.10	0.25	0.50	0.75
100.	0.114	0.310	0.666	0.921
200.	0.112	0.309	0.666	0.922
300.	0.111	0.308	0.666	0.922
400.	0.111	0.308	0.666	0.922
500.	0.111	0.308	0.666	0.923
600.	0.111	0.308	0.666	0.923
700.	0.111	0.308	0.667	0.923
800.	0.110	0.308	0.667	0.923
900.	0.110	0.308	0.667	0.923
1000.	0.110	0.308	0.667	0.923

From Table 2 the prior relative efficiency of $\hat{\sigma}_n^2$ is almost independent of the population size N and is somewhat greater than the sample fraction, f. These numerical results show the extent of possible improvement in estimation if good information is available on the prior distribution.

In order to analyze the extent of errors in the prior assumptions concerning the values of μ_0 and σ_0 we derive, on the basis of (5.2), the prior relative efficiencies of $\hat{\sigma}^2_{\mu_1,\sigma_1}(\underset{\sim}{x}_n)$, under (μ_0,σ_0). It is a straightforward matter to show that the prior relative efficiency of $\hat{\sigma}^2_{\mu_1,\sigma_1}(\underset{\sim}{x}_n)$ is

$$(5.6) \qquad RE(\hat{\sigma}^2_{H'}, H) = \left[1 + \frac{N(1-f)\left[(\rho-1+f\delta^2)^2 + 4f^2\frac{\delta^2}{n}\right]}{2\left(1-\frac{1}{N}(1-f)\right)}\right]^{-1},$$

where $\rho = \sigma_1^2/\sigma_0^2$ and $\delta = (\mu_1-\mu_0)/\sigma_0$.

In Table 3 we present the prior relative efficiency of the Bayes estimator $\hat{\sigma}^2_{H'}(\underset{\sim}{x}_n)$ as a function of f, δ and $\lambda=\rho-1$, where H' is the $N(\mu_1,\sigma_1^2)$ distribution. We see that the magnitude of δ is not so important, but deviations from σ_0 larger in magnitude than 10 percent reduce the prior relative efficiency below that of $\hat{\sigma}^2_n$. In Table 4 we provide these prior relative efficiency values for values of λ between -7.5% to 7.5% . We see that in this range the Bayes estimator is considerably more efficient than the classical sample variance.

Table 3. The Prior Relative Efficiency of $\hat{\sigma}^2_{H'}(\underset{\sim}{x}_n)$, N = 1,000

Sample Fraction = .10							
$\varepsilon \backslash \lambda$	-.30	-.20	-.10	0.00	0.10	0.20	0.30
-.50	0.028	0.067	0.279	0.754	0.124	0.042	0.021
-.40	0.027	0.061	0.258	0.874	0.141	0.045	0.022
-.30	0.026	0.057	0.211	0.950	0.157	0.048	0.023
-.20	0.025	0.055	0.194	0.986	0.170	0.051	0.023
-.10	0.024	0.053	0.185	0.998	0.179	0.052	0.024
.00	0.024	0.053	0.182	1.000	0.182	0.053	0.024
0.10	0.024	0.053	0.135	0.998	0.179	0.052	0.024
0.20	0.025	0.055	0.194	0.986	0.170	0.051	0.023
0.30	0.026	0.057	0.211	0.950	0.157	0.048	0.023
0.40	0.027	0.061	0.258	0.874	0.141	0.045	0.022
0.50	0.028	0.067	0.279	0.754	0.124	0.042	0.021
Sample Fraction = .25							
$\varepsilon \backslash \lambda$	-.30	-.20	-.10	0.00	0.10	0.20	0.30
-.50	0.045	0.122	0.617	0.391	0.091	0.037	0.020
-.40	0.038	0.094	0.415	0.602	0.119	0.044	0.023
-.30	0.033	0.078	0.304	0.817	0.150	0.051	0.025
-.20	0.031	0.069	0.247	0.950	0.180	0.057	0.027
-.10	0.029	0.064	0.219	0.994	0.202	0.061	0.028
.00	0.029	0.062	0.210	1.000	0.210	0.062	0.029
0.10	0.029	0.064	0.219	0.994	0.202	0.061	0.028
0.20	0.031	0.069	0.247	0.950	0.180	0.057	0.027
0.30	0.033	0.078	0.304	0.817	0.150	0.051	0.025
0.40	0.038	0.094	0.415	0.602	0.119	0.044	0.023
0.50	0.045	0.122	0.617	0.391	0.091	0.037	0.020
Sample Fraction = .50							
$\varepsilon \backslash \lambda$	-.30	-.20	-.10	0.00	0.10	0.20	0.30
-.50	0.114	0.395	0.780	0.199	0.073	0.036	0.022
-.40	0.076	0.214	0.847	0.373	0.109	0.048	0.027
-.30	0.058	0.142	0.555	0.645	0.159	0.062	0.032
-.20	0.048	0.110	0.382	0.893	0.216	0.076	0.038
-.10	0.044	0.095	0.307	0.989	0.266	0.087	0.041
.00	0.043	0.091	0.286	1.000	0.286	0.091	0.043
0.10	0.044	0.095	0.307	0.989	0.266	0.087	0.041
0.20	0.048	0.110	0.382	0.893	0.216	0.076	0.038
0.30	0.058	0.142	0.555	0.645	0.159	0.062	0.032
0.40	0.076	0.214	0.847	0.373	0.109	0.048	0.027
0.50	0.114	0.395	0.780	0.199	0.073	0.036	0.022

Table 4. Prior Relative Efficiency of $\hat{\sigma}^2_{H'}(\underset{\sim}{x}_n)$, N = 1,000.

Sample Fraction = .10							
$\delta\backslash\lambda$	-.075	-.050	-.025	0.000	0.025	0.050	0.075
-.50	0.461	0.754	0.957	0.754	0.461	0.279	0.180
-.40	0.385	0.645	0.939	0.874	0.560	0.334	0.210
-.30	0.336	0.564	0.884	0.950	0.651	0.387	0.238
-.20	0.305	0.510	0.829	0.986	0.722	0.431	0.262
-.10	0.288	0.480	0.793	0.998	0.766	0.460	0.278
-.00	0.283	0.470	0.780	1.000	0.780	0.470	0.283
0.10	0.288	0.480	0.793	0.998	0.766	0.460	0.278
0.20	0.305	0.510	0.829	0.986	0.722	0.431	0.262
0.30	0.336	0.564	0.884	0.950	0.651	0.387	0.238
0.40	0.385	0.645	0.939	0.874	0.560	0.334	0.210
0.50	0.461	0.754	0.957	0.754	0.461	0.279	0.180
Sample Fraction = .25							
$\delta\backslash\lambda$	-.075	-.050	-.025	0.000	0.025	0.050	0.075
-.50	0.868	0.868	0.617	0.391	0.252	0.171	0.122
-.40	0.658	0.911	0.874	0.602	0.378	0.244	0.166
-.30	0.484	0.759	0.965	0.817	0.532	0.333	0.217
-.20	0.385	0.619	0.910	0.950	0.678	0.423	0.268
-.10	0.336	0.540	0.838	0.994	0.777	0.491	0.307
-.00	0.321	0.516	0.810	1.000	0.810	0.516	0.321
0.10	0.336	0.540	0.888	0.994	0.777	0.491	0.307
0.20	0.385	0.619	0.910	0.950	0.678	0.423	0.268
0.30	0.484	0.759	0.965	0.817	0.523	0.333	0.217
0.40	0.658	0.911	0.874	0.602	0.378	0.244	0.166
0.50	0.868	0.868	0.617	0.391	0.252	0.171	0.122
Sample Fraction = .50							
$\delta\backslash\lambda$	-.075	-.050	0.025	0.000	0.025	0.050	0.075
-.50	0.571	0.395	0.276	0.199	0.148	0.114	0.090
-.40	0.921	0.766	0.544	0.373	0.261	0.188	0.141
-.30	0.787	0.951	0.873	0.645	0.440	0.303	0.215
-.20	0.573	0.803	0.974	0.893	0.655	0.445	0.305
-.10	0.448	0.662	0.905	0.989	0.813	0.568	0.384
-.00	0.415	0.615	0.865	1.000	0.863	0.615	0.415
0.10	0.448	0.662	0.905	0.989	0.813	0.568	0.384
0.20	0.563	0.803	0.974	0.893	0.655	0.445	0.305
0.30	0.787	0.951	0.873	0.645	0.440	0.303	0.214
0.40	0.921	0.766	0.544	0.373	0.261	0.188	0.141
0.50	0.571	0.395	0.276	0.199	0.148	0.114	0.090

References

[1] Basu, D. (1955), "On statistics independent of complete sufficient statistic."

[2] Basu, D. (1969), "The role of sufficiency and the likelihood principles in sample survey theory." Sankhya, A, 31: 441-454.

[3] Chaudhuri, A. (1977), "Some applications of the principles of Bayesian sufficiency and invariance to inference problems with finite populations," Sankhya, Sample Surveys; Theory and Practice, 39, Series C: 140-149.

[4] Fraser, D.A.S. (1968), The Structure of Statistical Inference, John Wiley and Sons, New York.

[5] Godambe, V.P. (1955), "A unified theory of sampling from finite populations," J.R. Statist. Soc., B, 17: 268-278.

[6] Godambe, V.P. (1965), " Contributions to the unified theory of sampling from finite populations, Int. Stat. Rev., 33: 242-258.

[7] Godambe, V.P. (1966), "A new approach to sampling from finite populations," J.R. Stat. Soc., B, 28: 310-328.

[8] Lindley, D.V. (1972), "Bayes estimates for the linear model," J. Roy. Statist. Soc., B, 34: 1-42.

[9] Liu, T.P. (1974), "Bayes estimation for the variance of a finite population," Metrika, 21: 127-132.

[10] Royall, R.M. (1971), "Linear regression models in finite population sampling," Foundations Of Statistical Inference (V.P. Godambe and D.A. Sprott, Ed.), Holt Rinehart and Winston of Canada, Toronto.

[11] Royall, R.M. (1976), "The linear least-squares prediction approach to two-stage sampling," Jour. Amer. Statist. Assoc., 71: 657-664.

[12] Royall, R.M. and Cumberland, W.G. (1978), "Variance estimation in finite population sampling," Jour. Amer. Statist. Assoc., 73: 351-358.

[13] Solomon, H. and Zacks, S. (1970), "Optimal designs of sampling from finite populations: A critical review and indication of new research areas," Jour. Amer. Statist. Assoc., 65: 653-677.

[14] Smith, A.F.M. (1973), "A general Bayesian linear model," J. Roy. Statist. Assoc., B, 35: 67-75.

[15] Zacks, S. (1969), "Bayes sequential designs of fixed size samples from finite populations," Jour. Amer. Statist. Assoc., 64: 1342-1349.

[16] Zacks, S. (1978), "Bayes equivariant estimators of the variance of a finite population for exponential priors," Tech. Report No. 33, ONR Project NR 042-276, Department of Math. and Statist., Case Western Reserve University.

[17] Zacks, S. (1971), The Theory Of Statistical Inference, John Wiley, New York.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER TECHNICAL REPORT No. 35	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Bayes and Equivariant Estimators of the Variance of a Finite Population		5. TYPE OF REPORT & PERIOD COVERED 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) S. Zacks and H. Solomon		8. CONTRACT OR GRANT NUMBER(s) NR 00014-75-C-0529 PROJECT NR 042-276
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Mathematics and Statistics CASE WESTERN RESERVE UNIVERSITY Cleveland, Ohio 44106		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS OFFICE OF NAVAL RESEARCH ARLINGTON, VIRGINIA 22217		12. REPORT DATE November 27, 1978 13. NUMBER OF PAGES 22
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

DISTRIBUTION OF DOCUMENT UNLIMITED

DISTRIBUTION STATEMENT A
Approved for public release; Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Sampling surveys, Bayes estimators of variance, equivariant estimators, Bayes-equivariant estimators, prior relative efficiency.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The problem of estimating the variance of a finite population is studied in a Bayesian framework. On the basis of the modern theoretical approach to sampling from finite populations and the special structure of the likelihood functions Bayes estimators of the population variance are derived. The structure of equivariant estimators is analyzed and Bayes equivariant estimators in the strict and the generalized sense are derived. Posterior and prior efficiency of the estimators is discussed.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)