BAYES UNBIASED ESTIMATION OF THE COMMON MEAN OF TWO NORMAL DISTRIBUTIONS BASED ON SMALL SAMPLES

by

RONALD LEE DILLON

B. S., Kansas State University, 1961

A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Statistics

KANSAS STATE UNIVERSITY

Manhattan, Kansas

1967

Approved by:

Shelemyahu Zacks

Major Professor

TABLE OF CONTENTS

INTRODUCTION

When estimating the common mean of two normal and independent distributions, $NID(\mu;\sigma_i^2)(i = 1,2)$ a well known procedure is to take independent simple random samples from both distributions, find the sample means $\bar{x}$ and $\bar{y}$, and determine a weighted mean where the weights are dependent on the ratio of variances with the restriction that they add to one; expressed parametrically the estimator is

$$\tilde{\bar{\mu}} = A\bar{x} + B\bar{y}, \quad A,B \geq 0, \ A + B = 1 \tag{1.1}$$

where A and B are the weighting functions. The problem is to find A and B to weight the estimators $\bar{x}$ and $\bar{y}$ to arrive at a combined estimator having desired properties.

When the variance ratio is known, the uniformly minimum variance unbiased estimator of μ is the maximum likelihood (M.L.) estimator

$$\hat{\mu}_0 = \phi(\rho)\bar{x} + \left(1 - \phi(\rho)\right)\bar{y} , \tag{1.2}$$

where $\phi(\rho) = \frac{n_1}{n_2}\rho/\left(1 + \frac{n_1}{n_2}\rho\right)$, $\rho = \sigma_2^2/\sigma_1^2$, and n_1, n_2 are the corresponding sample sizes. In applied statistics, however, ρ is generally unknown and other estimators for the common mean, i.e. estimators for the weighting functions A and B, must be found.

Several studies have been made using the classical approach to find an estimator when ρ is unknown, and are of two general classes which Zacks [9] expressed parametrically as;

<u>Class I</u>

$$\hat{\mu}(\rho^*) = I\left(\frac{s_2^2}{s_1^2}; \rho^*\right)\bar{\mu} + \left(1 - I\left(\frac{s_2^2}{s_1^2}; \rho^*\right)\right)\hat{\mu} \tag{1.3}$$

and

<u>Class II</u>

$$\tilde{\mu}(\rho^*) = I\left(\frac{s_2^2}{s_1^2};\ \rho^*\right)\bar{\mu} + J_1\left(\frac{s_2^2}{s_1^2};\ \rho^*\right)\bar{x} + J_2\left(\frac{s_2^2}{s_1^2};\ \rho^*\right)\bar{y}\ , \tag{1.4}$$

where:

$$\bar{\mu} = \frac{(n_1/n_2)\bar{x} + \bar{y}}{1 + n_1/n_2}\ , \tag{1.5}$$

$$\hat{\mu} = \frac{(n_1 s_2^2/n_2 s_1^2)\ \bar{x} + \bar{y}}{1 + n_1 s_2^2/n_2 s_1^2}\ , \tag{1.6}$$

$$I\left(\frac{s_2^2}{s_1^2};\ \rho^*\right) = \begin{cases} 1, & \text{if } 1/\rho^* \le s_2^2/s_1^2 \le \rho^* \\ 0, & \text{otherwise} \end{cases}\ , \tag{1.7}$$

$$J_1\left(\frac{s_2^2}{{}^{2}_{1}};\ \rho^*\right) = \begin{cases} 1, & \text{if } s_2^2/s_1^2 > \rho^* \\ 0, & \text{otherwise} \end{cases}\ , \tag{1.8}$$

and

$$J_2\left(\frac{s_2^2}{s_1^2};\ \rho^*\right) = \begin{cases} 1, & \text{if } s_2^2/s_1^2 < 1/\rho^* \\ 0, & \text{otherwise} \end{cases} \tag{1.9}$$

The $s_i^2 (i = 1,2)$ are the unbiased estimators for $\sigma_i^2 (i = 1,2)$. The values ρ^* in $\hat{\mu}(\rho^*)$ and $\tilde{\mu}(\rho^*)$ are critical values of the F-test of significance, according to which one decides to apply the estimators $\bar{\mu}$, $\hat{\mu}$, $\bar{x}$ or $\bar{y}$.

Graybill and Deal [3] have shown that $\hat{\mu}$ (eqn. 1.6) is uniformly better than $\bar{x}$ or $\bar{y}$ in estimating the common mean if and only if both n_1 and n_2 are greater than 10. Therefore with this information one wonders whether $\hat{\mu}(\rho^*)$ and $\tilde{\mu}(\rho^*)$ are equally as good an estimator for the common mean when samples are small. Both $\hat{\mu}(\rho^*)$ and $\tilde{\mu}(\rho^*)$ have a distinct disadvantage when based on small samples, since the values of their characteristic functions

$I(\cdot;\cdot)$, $J_1(\cdot;\cdot)$ and $J_2(\cdot;\cdot)$ are dependent upon sample variances. This disadvantage can easily be observed; since $E(s_i^2) = \sigma_i^2$, then $\mathrm{Var}\ (s_i^2) = 2\sigma_i^2/(n_i-1)$ attains near-maximum values when n_i is small. Therefore accuracy of the sample variances become a problem and the choice of $\bar{\mu}$, $\hat{\mu}$, $\bar{x}$ or $\bar{y}$ as estimators is somewhat dubious. Another possible disadvantage occuring in estimators $\tilde{\mu}(\rho^*)$ exists when $\rho=1$, and that is, all available information is not used since either $\bar{x}$ or $\bar{y}$ might be discarded, depending on the relative size of the sample variances. Therefore it is said that $\hat{\mu}(\rho^*)$ when based on small samples would be the best estimator under all circumstances, and this is verified in a study by Zacks [9]. Zacks studied the efficiency functions of $\hat{\mu}(\rho^*)$ and $\tilde{\mu}(\rho^*)$ when based on small samples of equal size and found that $\hat{\mu}(\rho^*)$ was a superior estimator for the common mean. By studying the general behavior of the efficiency functions and observing the explicit efficiency function for $\hat{\mu}(\rho^*)$, when $n=3$ and $\rho^* = 1, 3.4, 9, 19$ and ∞, Zacks recommended using $\hat{\mu}(\rho^*=9)$ as an estimator for the common mean, when ρ can assume any value $(\rho > 0)$. This recommendation was made because the efficiency function over the range of ρ has desired properties. (For further discussion see Zacks [9])

When prior information concerning the value of variance ratio ρ is available, Zacks [9] suggested that a Bayes approach might lead to a more efficient estimator of the common mean. It also seems reasonable that this estimator for the common mean will improve the use of the somewhat dubious reliability of $s_i^2 (i = 1,2)$ when based on small samples.

This paper will exhibit an unbiased estimator of μ, in which the weight function $\psi(s_2^2/s_1^2)$ is a certain Bayes estimator of $\phi(\rho)$, and is more efficient than $\hat{\mu}(\rho^*=9)$ over the interval $1 \leq \rho \leq 6$. Explicit formulae for $\psi(s_2^2/s_1^2)$ are studied. The efficiency functions are plotted in Fig. 1. A table is

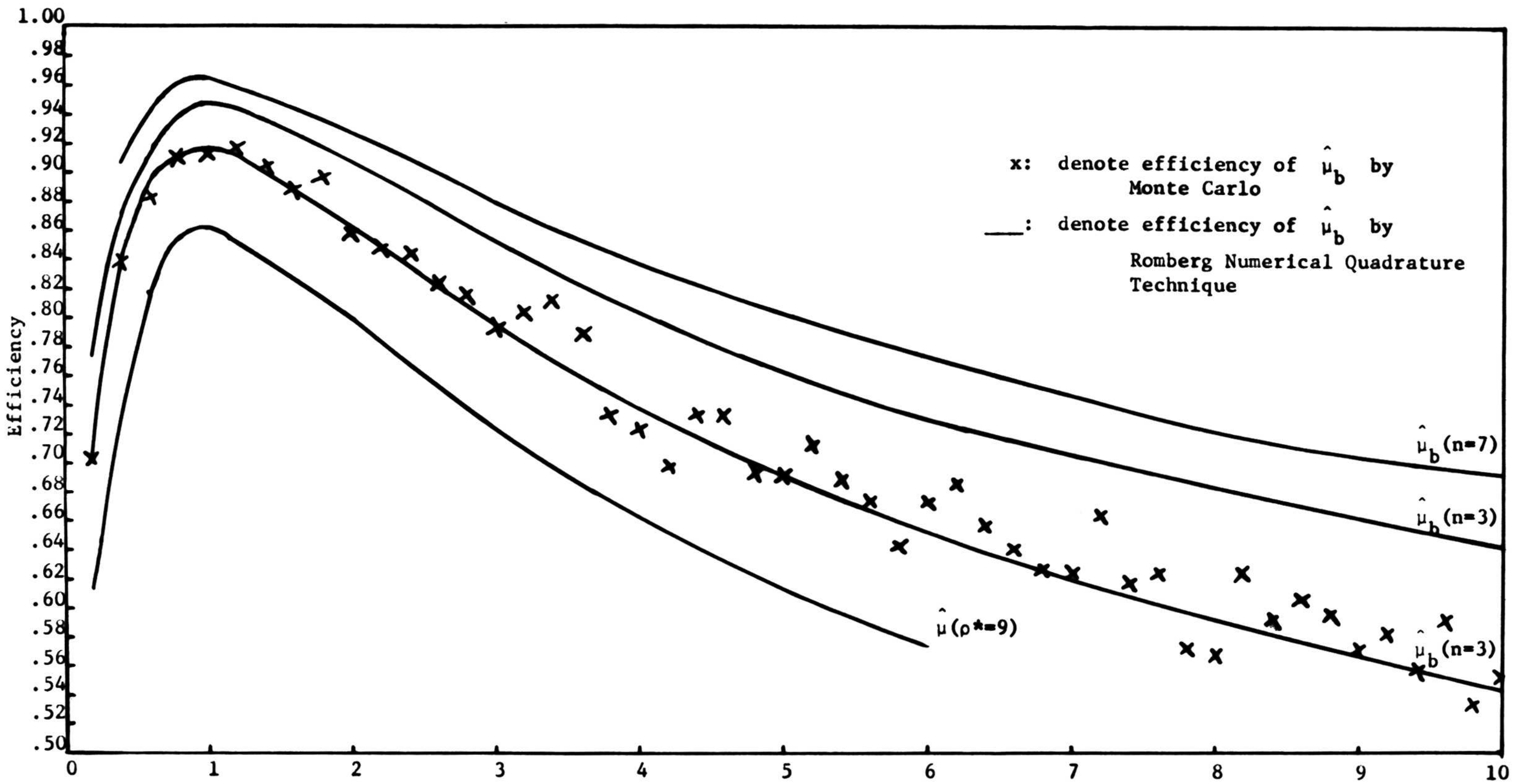

Figure 1. Efficiency curves of the unbiased estimators $\hat{\mu}_b$ for samples of equal size $n = 3, 5, 7$, and efficiency curve of $\hat{\mu}(\rho^*=9)$ for samples of equal size $n=3$.

given (Table I) which determines the value of the weighting function when $n = 3, 5, 7$ and $\rho = 0(.2)10$. Monte Carlo and numerical quadrature techniques for calculating the efficiency function are discussed and digital programs are given in Plates I and II.

DERIVATION OF A BAYES ESTIMATOR OF THE WEIGHT FUNCTION $\phi(\rho)$

The Bayes estimator for the common mean of two normal distributions when the variance ratio is unknown is derived in this report by finding a Bayes estimator for the weighting function $\phi(\rho)$. Let $\psi(z)$ be an estimator for $\phi(\rho)$, where z, a random variate, is a function of the two independent simple random samples from a density function $g(z \mid \rho)$, where ρ is defined as before. Also assume that ρ has a priori density function $h(\rho)$, and an associated loss function $L\big(\psi(z);\ \phi(\rho)\big) \geq 0$. Then it is said that the estimator $\psi(z)$ that minimizes the loss function is a good estimator, and further, an estimator $\psi(z)$ that minimizes the a priori risk, $E_\rho[R\big(\psi(z),\ \phi(\rho)\big)]$, where $R\big(\psi(z),\ \phi(\rho)\big) = E[L\big(\psi(z);\ \phi(\rho)\big)]$, is a Bayes estimator (Wilks [8]). It is easily shown that to minimize the a priori risk is equivalent to minimizing the a posteriori risk, $E_\rho[L\big(\psi(z);\ \phi(\rho)\big) \mid Z]$ (Mood and Graybill [5]). By letting the loss function be the squared-error, $\big(\psi(z) - \phi(\rho)\big)^2$, the Bayes estimator is found by setting the first derivative of the average a posteriori risk, with respect to $\psi(z)$, equal to zero, which gives

$$\frac{d\left\{E_\rho[L\big(\psi(z);\ \phi(\rho)\big)]\right\}}{d\ \psi(z)} = 0 \quad ,$$

or equivalently

$$\int_0^\infty \frac{d\left\{L\big(\psi(z);\ \phi(\rho)\big)\right\}}{d\ \psi(z)}\ h(\rho \mid z)\ d\rho = 0.$$

After substitution of the squared-error loss and taking the derivative we arrive at the Bayes estimator

$$\psi(z) = E_{\rho}[\phi(\rho) \mid z] = \int_0^{\infty} \phi(\rho)h(\rho \mid z)\, d\rho \,, \tag{2.1}$$

where $h(\rho \mid z)$ is the a posteriori density function. By Bayes theorem, the a posteriori density function of ρ, given z is:

$$h(\rho \mid z) = \frac{g(z \mid \rho)\, h(\rho)}{k(z)} \,, \quad 0 \le \rho \le \infty,\ 0 < z < \infty.$$

$k(z)$ is the marginal density of z, averaged with respect to the priori density of ρ, i.e.

$$k(z) = \int_0^{\infty} g(z \mid \rho)\, h(\rho)\, d\rho \,.$$

To find the Bayes estimator $\psi(z)$, the a posteriori density function must first be determined. Let $z = s_2^2/s_1^2$, a function of the two independent random samples, then since s_1^2 and s_1^2 are independent, $z \sim \rho F[\gamma_2,\gamma_1]$; where $F[\gamma_2,\gamma_1]$ is a central F-statistic with $\gamma_i = n_i - 1$ $(i = 1,2)$ degrees of freedom. The density function of $F[\gamma_2,\gamma_1]$ at the point F is given by:

$$f(F) = \frac{1}{B\left(\frac{\gamma_1}{2},\frac{\gamma_2}{2}\right)} \left(\frac{\gamma_2}{\gamma_1}\right)^{\frac{\gamma_2}{2}} \frac{F^{\frac{\gamma_2}{2}-1}}{\left(1 + \frac{\gamma_2}{\gamma_1} F\right)^{\frac{\gamma_1+\gamma_2}{2}}} \,, \quad 0 \le F \le \infty \,.$$

Making the transformation $z = \rho F$, the density function of $z = s_2^2/s_1^2$ is found to be:

$$g(z \mid \rho) = \frac{1}{B\left(\frac{\gamma_1}{2},\frac{\gamma_2}{B}\right)} \left(\frac{\gamma_1}{\gamma_2}\right)^{\frac{\gamma_1}{2}} \left(\frac{1}{z}\right)^{\frac{\gamma_1}{2}+1} \frac{\rho^{\frac{\gamma_1}{2}}}{\left[1 + \frac{\gamma_1}{\gamma_2}\frac{\rho}{z}\right]^{\frac{\gamma_1+\gamma_2}{2}}} \,,$$

$$0 < Z < \infty, \quad 0 \le \rho \le \infty. \tag{2.2}$$

Since ρ is a ratio of variances, the a priori density function is chosen to be

$$h(\rho) \propto \frac{\rho^{\frac{\gamma_2}{2} - 1}}{\left(1 + \frac{\gamma_2}{\gamma_1}\rho\right)^{\frac{\gamma_1+\gamma_2}{2}}} \qquad 0 \leq \rho \leq \infty \ . \tag{2.3}$$

From equations (2.2) and (2.3) the a posteriori density function, $h(\rho \mid z)$, is:

$$h(\rho \mid z) = \frac{\left(\frac{\gamma_1}{\gamma_2}\right)^{\frac{\gamma_1}{2}} \left(\frac{1}{z}\right)^{\frac{\gamma_1}{2}+1} \frac{1}{\rho}}{B\left(\frac{\gamma_1}{2},\frac{\gamma_2}{2}\right)\ k(z)} \left[\frac{\rho}{\left(1 + \frac{\gamma_1}{\gamma_2}\frac{\rho}{z}\right)\left(1 + \frac{\gamma_1}{\gamma_2}\rho\right)}\right]^{\frac{\gamma_1+\gamma_2}{2}} , \tag{2.4}$$

where

$$k(z) = \frac{\left(\frac{\gamma_1}{\gamma_2}\right)^{\frac{\gamma_1}{2}} \left(\frac{1}{z}\right)^{\frac{\gamma_1}{2}+1}}{B\left(\frac{\gamma_1}{2},\frac{\gamma_2}{2}\right)} \int_0^\infty \frac{1}{\rho} \left[\frac{\rho}{\left(1 + \frac{\gamma_1}{\gamma_2}\frac{\rho}{z}\right)\left(1 + \frac{\gamma_2}{\gamma_1}\rho\right)}\right]^{\frac{\gamma_1+\gamma_2}{2}} d\rho \tag{2.5}$$

Under the condition that ρ is known, the best estimator for the common mean is the M.L. estimator (eqn. 1.2) where

$$\phi(\rho) = \frac{(n_1/n_2)\rho}{1 + (n_1/n_2)\rho}$$

By substitution of equations (2.4) and $\phi(\rho)$ into equation (2.1), the Bayes estimator of $\phi(\rho)$ given $z = s_2^2/s_1^2$ is:

$$\psi(z) = E_\rho[\phi(\rho) \mid z] = \int_0^\infty \frac{(n_1/n_2)\rho}{\left(1 + (n_1/n_2)\rho\right)} h(\rho \mid z)\, d\rho$$

$$= \frac{\frac{n_1}{n_2}\left(\frac{\gamma_1}{\gamma_2}\right)^{\frac{\gamma_1}{2}} \left(\frac{1}{z}\right)^{\frac{\gamma_1}{2}+1}}{B\left(\frac{\gamma_1}{2},\frac{\gamma_2}{2}\right)\ k(z)} \int_0^\infty \frac{1}{1 + (n_1/n_2)\rho} \left[\frac{\rho\left(1 + (\gamma_2/\gamma_1)\rho\right)^{-1}}{\left(1 + \frac{\gamma_1}{\gamma_2}\frac{\rho}{z}\right)}\right]^{\frac{\gamma_1+\gamma_2}{2}} d\rho \tag{2.6}$$

where $k(z)$ is defined in equation (2.5).

Making the transformation $u = (1 + \frac{\gamma_1}{\gamma_2}\frac{\rho}{z})^{-1}$ to obtain bounded integration limits, the estimator is

$$\psi(z) = \frac{\frac{n_1}{n_2}\left(\frac{\gamma_2}{\gamma_1}\right)^{\frac{\gamma_2}{2}+1} z^{\frac{\gamma_2}{2}}}{B\left(\frac{\gamma_1}{2},\frac{\gamma_2}{2}\right)\; k(z)} \int_0^1 \frac{(1-u)^{\frac{\gamma_1+\gamma_2}{2}}\, u^{\frac{\gamma_1+\gamma_2}{2}-1}\, du}{\left(u+\frac{n_1}{n_2}\frac{\gamma_1}{\gamma_2} z(1-u)\right)\left(u+\left(\frac{\gamma_2}{\gamma_1}\right)^2 z(1-u)\right)^{\frac{\gamma_1+\gamma_2}{2}}}\,, \tag{2.7}$$

where

$$k(z) = \frac{\left(\frac{\gamma_2}{\gamma_1}\right)^{\frac{\gamma_2}{2}} z^{\frac{\gamma_2}{2}-1}}{B\left(\frac{\gamma_1}{2},\frac{\gamma_2}{2}\right)} \int_0^1 \frac{(1-u)^{\frac{\gamma_1+\gamma_2}{2}-1}\, u^{\frac{\gamma_1+\gamma_2}{2}-1}\, du}{\left(u+\left(\frac{\gamma_2}{\gamma_1}\right)^2 z(1-u)\right)^{\frac{\gamma_1+\gamma_2}{2}}} \tag{2.8}$$

In investigating equations (2.7) and (2.8) for unequal sample sizes, it was found that solutions required laborious calculations, therefore only estimators of equal sample sizes were considered. Explicit formulae for the Bayes estimator $\psi(z)$ when the equal sample sizes are $n = 3$, 5 and 7, were found by making the transformation $t = u + \left(\frac{\gamma_2}{\gamma_1}\right)^2 \left(z(1-u)\right)$, and integrating by direct procedures. The obtained Bayes estimators $\psi_n(z)$ are:

$$\psi_3(z) = \frac{1-4z-5z^2+(4z+2z^2)\ln_e z}{2(z^2-1)\ln_e z - 4(1-z)^2} \tag{2.9}$$

$$\psi_5(z) = \frac{z\left(\frac{32}{3}+9z+16z^2-\frac{47}{12}z^3+\frac{1}{4z}+(4+18z+12z^2+z^3)\ln_e z\right)}{(1-z)\left(-\frac{11}{3}-9z+9z^2+\frac{11}{3}z^3-(z^3+9z^2+9z+1)\ln_e z\right)} \tag{2.10}$$

$$\psi_7(z) = \frac{z\left(\frac{107}{5}+125z+\frac{200}{3}z^2-\frac{275}{2}z^3-71z^4-\frac{71}{15}z^5+\frac{1}{6z}+T_1(z)\right)}{(1-z)\left(-\frac{137}{30}(1-z^5)-\frac{325}{6}(z-z^4)-\frac{200}{3}(z^2-z^3)+T_2(z)\right)}\,, \tag{2.11}$$

where

$$T_1(z) = (6+75z+200z^2+150z^3+30z^4+z^5)\ln_e z$$

$$T_2(z) = (1+25z+100z^2+100z^3+25z^4+z^5)\ln_e z \ . \qquad (2.12)$$

By using l'Hospitals rule one can show that the above Bayes estimators have the expected property:

$$\operatorname{Lim} \psi_i(z) = \begin{cases} 0, & \text{when } z \to 0 \\ \frac{1}{2}, & \text{when } z \to 1 \\ 1, & \text{when } z \to \infty \end{cases} \qquad \text{for } i = 3, 5, 7 \ .$$

These limiting values are the same as those of $\phi(\rho)$ when $n_1 = n_2$. For aiding the experimenter, tables for $\psi_i(z) (i = 3, 5, 7)$ are given (Table I) which determine the value of the weighting function when $\rho = .2(.2)10$.

EFFICIENCY OF THE BAYES UNBIASED ESTIMATOR

The Bayes estimator of the common mean can be written as:

$$\hat{\mu}_b = \psi(z)\bar{x} + \big(1 - \psi(z)\big)\bar{y} \qquad (3.1)$$

where $\psi(z)$ is the Bayes estimator for $\phi(\rho)$, a function of sample variances, and applying the well known property that the sample mean and variance are independent in normal distributions (Mood and Graybill [5]), it can readily be shown that $\hat{\mu}_b$ is an unbiased estimator of the common mean μ. The variance of $\hat{\mu}_b$ is

$$\begin{aligned} \operatorname{Var}[\hat{\mu}_b] &= E_z[\operatorname{Var}(\hat{\mu}_b \mid z)] + \operatorname{Var}_z[E(\hat{\mu}_b \mid z)] \\ &= \frac{\sigma_1^2}{n} E_z[\psi^2(z)] + \frac{\sigma_2^2}{n} E_z[\big(1 - \psi(z)\big)^2] \\ &= \frac{\sigma_1^2}{n}\Big\{E_z[\psi^2(z)] + \rho E_z[\big(1 - \psi(z)\big)^2]\Big\} \end{aligned} \qquad (3.2)$$

All formulae in the present section are restricted to cases of equal sample size.

The efficiency of $\hat{\mu}_b$ when compared to the M.L. estimator $\hat{\mu}_0$ (eqn. 1.2) as a function of ρ is:

$$\mathrm{Eff}[\hat{\mu}_b \mid \rho, n] = \frac{\mathrm{Var}[\hat{\mu}_0]}{\mathrm{Var}[\hat{\mu}_0]} = \frac{\sigma_1^2 \rho}{n(1+\rho)\mathrm{Var}[\hat{\mu}_b]}$$

$$= \frac{\rho/(1+\rho)}{E_z[\psi^2(z)] + \rho E_z[(1-\psi(z))^2]} \tag{3.3}$$

The efficiency functions of the Bayes estimators were calculated for samples of equal size $n = 3$, 5 and 7. The graphs appear in Fig. 1, where $\rho = .2(.2)10$. In the previous study of Zacks [9] the efficiency function of $\hat{\mu}(\rho^*)$, when $n=3$ and $\rho^*=9$, was calculated similiarly with respect to the M.L. estimator $\hat{\mu}_0$. This efficiency function is presented in Fig. 1. We see in Fig. 1 that $\hat{\mu}_b(n=3)$ has a higher efficiency than $\hat{\mu}(\rho^*=9)$ for all values of ρ in the interval $.2 \leq \rho \leq 6$.

NUMERICAL TECHNIQUES

To find the function $\mathrm{Eff}[\hat{\mu}_b \mid \rho, n_1, n_2]$, the moments $E_z[\psi(z)]$ and $E_z[\psi^2(z)]$ should be determined. It is observed that neither moments can be found by exact integration methods because $\psi(z)$ is too complicated. To overcome this difficulty, two approximating techniques were used; one, a Monte Carlo procedure, which uses the mean estimate

$$\overline{\psi^i} = \frac{1}{k_i} \sum_{j=1}^{k_i} \psi^i(z_j) \qquad (i = 1,2) \tag{4.1}$$

to approximate $E_z[\psi^i(z)]$ $(i = 1,2)$; and two, a Romberg numerical quadrature procedure which is a recursive calculation based on the trapezoidal rule, and is an extension (but more than a reformulation) of the Newton Cotes formula. (Bauer, et. al. [1])

The Monte Carlo procedure was adapted for use on the IBM 1410 Computer and the FORTRAN program (Plate I) uses the following steps to generate independent random z_i variates:

(1) Generate independent psuedo-random uniformly distributed $(U(0,1))$ variates, u_i, by a subroutine RECTAN. A multiplicative congruential procedure developed by D. H. Lehmer in 1951 is used, utilizing the relation,

$$u_{i+1} = 23u_i \text{ (Modulus } 10^8 + 1) \qquad (i = 0, 1, 2, \cdots), \quad (4.2)$$

where u_0 is the starting value (any 8 digit number chosen from a random number table) and the $u_i (i = 1, 2, \cdots)$ are the resulting 8 digit psuedo-random numbers that are split into two 4 digit numbers and used as two $U(0,1)$ variates. The 8 digit u_i's were tested by Taussky and Todd [7] and it was found that the method is a suitable generator with recycle period 5882352.

(2) Generate $\chi^2[\gamma_i]$ variates. Let $u_i (i = 1, 2, \cdots)$ be independent psuedo-random numbers from $U(0,1)$ distribution, then the inverse transformation relation (Naylor, et. al. [6]),

$$x_i = -2 \ln_e(u_i) \qquad i = 1, 2, \cdots \qquad (4.3)$$

yields $x_i \sim \chi^2[2]$ independent psuedo-random variates.

Since the generating function of $\chi^2[\gamma_i]$ is a convolution of the generating function of $\chi^2[2]$ (Feller [4]) when γ_i is even,

$$t_{\gamma_i} = \sum_{j=1}^{\gamma_i/2} x_j \sim \chi^2[\gamma_i] \quad , \tag{4.4}$$

where the $\gamma_i/2$ values of x_j are generated independently. When γ_i is odd we use the formula

$$t_{\gamma_i} = \sum_{j=1}^{\frac{\gamma_i-1}{2}} x_j + v^2 \sim \chi^2[\gamma_i] \quad , \tag{4.5}$$

where v is independent of x_j and $v \sim N(0,1)$, then it is well known that $v^2 \sim \chi^2[1]$. To generate v, we generate two additional independent u_1 and u_2 and use the inverse transformation relation (Box and Muller [2])

$$\begin{aligned} v_1 &= (-2 \ln_e u_1)^{1/2} \sin 2\pi u_2 \\ v_2 &= (-2 \ln_e u_1)^{1/2} \cos 2\pi u_2 \end{aligned} \tag{4.6}$$

Either v_1 or v_2 is then used. Since in this report $\gamma_i (i = 1,2)$ are confined to even numbers, only relation (4.4) is used in the computer program.

(3) Generate $F[\gamma_2,\gamma_1]$ variates. This is done by using the well known relation

$$F = \frac{\gamma_1}{\gamma_2} \frac{t_{\gamma_2}}{t_{\gamma_1}} \sim F[\gamma_2,\gamma_1] \quad , \tag{4.7}$$

where the x_i's in t_{γ_1} and t_{γ_2} are independently generated for all i.

It is now just a matter of generating the $F[\gamma_2,\gamma_1]$ variates for different fixed ρ, n_1, n_2 in order to obtain a $\rho F[\gamma_2,\gamma_1]$ distribution, and subsequently to determine estimates for $E_z[\psi^i(z)]$ $(i = 1,2)$. It was found that when $k_i = 200$ $(i = 1,2)$ in equation (4.1), the values of $Eff[\hat{\mu}_b]$ when $.2 \leq \rho \leq 10$, $n_1 = n_2 = 3$, gave a reasonable estimate of a smooth curve. (see Fig. 1)

The Romberg quadrature method was chosen in preference to other quadrature methods because it is numerically stable and allows for a recursive calculation procedure for higher orders to be easily adapted to computer programming. The FORTRAN program for the IBM 1410 was written by J. O. Mingle, Kansas State University, Department of Nuclear Engineering, and is given in a modified form in Plate II. By definition,

$$E_z[\psi^i(z)] = \int_0^\infty \psi^i(z)\; g(z \mid \rho)\; dz \qquad (i = 1,2) \tag{4.8}$$

where $g(z \mid \rho)$ is given by equation (2.2). The limits of integration can not be handled easily by computer methods, therefore the transformation $u = (1 + z/\rho)^{-1}$ when $\gamma_1 = \gamma_2$ was used, giving

$$E_z[\psi^i(z)] = \int_0^1 \psi^i\left(\frac{\rho(1-u)}{u}\right) g\left(\frac{\rho(1-u)}{u}\right) d\left(\frac{\rho(1-u)}{u}\right) , \tag{4.9}$$

where the limits of integration can be easily handled.

The FORTRAN programs which are given are for $n=3$ and can be easily adapted for other sizes. The two methods were used as a procedural check and to determine which had a faster calculation time. It was found that the Romberg procedure gave best results in the shortest time although the graphs of the efficiency function of $\hat{\mu}_b(n=3)$ for the two methods were not significantly different. (see Fig. 1)

SUMMARY AND CONCLUSION

An unbiased estimator $\hat{\mu}_b$ for the common mean of two normal distributions was derived, in which a weight function $\psi(z)$ is a certain Bayes estimator for $\phi(\rho)$. Attention was focused on the efficiency of this estimator when samples from each distribution are very small. In particular, explicit formulae of the Bayes estimator $\psi(z)$ were derived for samples of equal size $n = 3, 5, 7$ and the efficiencies for the estimators of the common mean determined by these $\psi(z)$ were studied. In investigating the Bayes estimator for $\phi(\rho)$ for unequal sample size, it was discovered that solutions required laborious calculations, therefore they were not considered.

It was found that the efficiency functions for $\hat{\mu}_b(n = 3, 5, 7)$ over the interval $.2 \leq \rho \leq 10$, are uniformly greater than 0.54. Moreover, when the efficiency of $\hat{\mu}_b$ was compared to $\hat{\mu}(\rho^*=9)$ for $n=3$, it was found that $\hat{\mu}_b$ is uniformly more efficient in the interval $1 \leq \rho \leq 6$; in fact, $\hat{\mu}_b$ is uniformly 6% more efficient than $\hat{\mu}(\rho^*=9)$.

It is therefore concluded that this Bayes unbiased estimator for the common mean of two normal distributions does offer an improvement over existing procedures when samples are very small.

ACKNOWLEDGEMENTS

I wish to express my sincere appreciation to my major professor, Dr. S. Zacks, for his suggestion of this topic and for his advice and assistance during the preparation of this report. Thanks are due also to my professors for their invaluable instruction in course work. Finally, I thank my wife for her encouragement to return to college.

REFERENCES

1. Bauer, F. L., Rutishauser, H., and Stiefel, E., "New Aspects in Numerical Quadrature," Proceedings of Symposia in Applied Mathematics, Vol. XV, American Mathematical Society, Providence, R. I., 1936, pp. 199-218.

2. Box, G. E. P. and Muller, M. E., "A Note on the Generation of Random Normal Deviates," Annals of Mathematical Statistics, **29**(1958), 610-11.

3. Graybill, F. A. and Deal, R. B., "Combining Unbiased Estimators," Journal of the Biometric Society, **15**(1959), 543-50.

4. Feller, W., An Introduction to Probability Theory and its Applications, Vol. I (2nd Ed.), John Wiley & Sons, Inc., New York, 1966.

5. Mood, A. M., and Graybill, F. A., Introduction to the Theory of Statistics, McGraw-Hill, New York, 1963.

6. Naylor, T. H., Balintfy, J. L., Burdick, D. S., and Kong Chu, Computer Simulation Techniques, John Wiley & Sons, Inc., New York, 1966

7. Taussky, O., and Todd, J., "Generation and Testing of Psuedo-Random Numbers," Symposium on Monte Carlo Methods, John Wiley & Sons, Inc., New York, 1956.

8. Wilks, S. S., Mathematical Statistics, John Wiley & Sons, Inc., New York, 1962.

9. Zacks, S., "Unbiased Estimation of the Common Mean of Two Normal Distributions Based on Small Samples," Journal of the American Statistical Association, **61**(1966), 467-76.

APPENDIX

TABLE I

VALUES OF WEIGHTING FUNCTION $\psi_n(z)$ $(n = 3, 5, 7)$ FOR $z = .2(.2)10$

$z = s_2^2/s_1^2$	$\psi_3(z)$	$\psi_5(z)$	$\psi_7(z)$
.2	.346	.330	.324
.4	.410	.400	.396
.6	.449	.444	.441
.8	.478	.475	.474
1.0	.500	.500	.500
1.2	.518	.520	.517
1.4	.534	.537	.539
1.6	.547	.552	.554
1.8	.558	.565	.567
2.0	.569	.576	.579
2.2	.578	.586	.590
2.4	.586	.596	.599
2.6	.594	.604	.608
2.8	.601	.612	.616
3.0	.607	.619	.624
3.2	.613	.626	.631
3.4	.619	.632	.637
3.6	.624	.638	.643
3.8	.629	.643	.649
4.0	.634	.648	.654
4.2	.638	.653	.659
4.4	.642	.658	.664
4.6	.646	.662	.668
4.8	.650	.666	.672
5.0	.653	.670	.676

TABLE I CONTINUED

$z = s_2^2/s_1^2$	$\psi_3(z)$	$\psi_5(z)$	$\psi_7(z)$
5.2	.657	.674	.680
5.4	.660	.677	.684
5.6	.663	.681	.688
5.8	.666	.684	.691
6.0	.669	.687	.694
6.2	.672	.690	.697
6.4	.674	.693	.700
6.6	.677	.696	.703
6.8	.679	.698	.706
7.0	.682	.701	.709
7.2	.684	.704	.711
7.4	.686	.706	.714
7.6	.688	.708	.716
7.8	.690	.711	.718
8.0	.692	.713	.721
8.2	.694	.715	.723
8.4	.696	.717	.725
8.6	.698	.719	.727
8.8	.700	.721	.729
9.0	.702	.723	.731
9.2	.703	.725	.733
9.4	.705	.726	.735
9.6	.707	.728	.737
9.8	.708	.730	.738
10.0	.710	.732	.740

```
C*****MONTE CARLO TECHNIQUE FOR EQUAL SAMPLE SIZES AND EVEN NUMBER DEG
C      OF FREEDOM.  INTEGER CONSTANT WORD SIZE MINIMUM OF 8.  FLOATING
C      POINT CONSTANT SHOULD BE AT MACHINE MAXIMUM IN ORDER TO CALCULATE
C      LOG-BASE(E) IN FUNCTION F(X).  IF LOG-BASE(E) ACCURACY ERROR OC-
C      CURS, THE PROGRAM OMITS THAT ITERATION.
C      DIMENSION X(4),B(2*(N-1))
       DIMENSIONX(10),B(12)
00001  FORMAT(6HLRHO= ,F7.3,5X,6HEFF = ,F10.4)
00002  FORMAT(6H START,3X,I5,2X,I5)
00003  FORMAT(1H ,2E16.9)
00004  FORMAT(5H RAND,2X,F12.8,2X,F12.8)
00005  FORMAT(6X,3HIS=,I8)
C**    G(Z)=BAYES ESTIMATOR FOR WEIGHT FUNCTION
       G(Z)=((1.+4.*Z-5.*Z*Z+(4.*Z+2.*Z*Z)*ALOG(Z))/(2.*((Z*Z-1.)*ALOG(Z)
      1-2.*(1.-Z)*(1.-Z))))
C**    IS = 8 DIGIT RANDOM NUMBER START
       IS=20938802
C**    N = SIZE OF SAMPLE
       N=3
C**    KRN=NUMBER OF ITERATIONS
       KRN=200
       KDF=N-1
       KDFD2=KDF/2
       P=0.0
       DO52I=20,30,10
       R=FLOAT(I)/10.
       PSQ=0.0
       TN=0.0
       TN1=0.0
       DO44JJ=1,KRN
       DO20KK=1,4
00020  X(KK)=0.0
C**    GENERATE U(0,1) PSUEDORANDOM NUMBERS
       DO26K=1,KDF
       CALLRECTAN(IS,U1,U2,CHECK)
       IF(CHECK.EQ.0.0)STOP
       B(K)=ABS(U1)
       KDFK=K+KDF
00026  B(KDFK)=ABS(U2)
       DO31M=1,KDFD2
       K=4*(M-1)
C**    CHI SQUARE TRANSFORMATION(N-1 DEGREES OF FREEDOM)
       DO31J=1,4
       J1=J+K
00031  X(J)=X(J)+(-2.*ALOG(1.-B(J1)))
C**    F-DISTRIBUTION TRANSFORMATION
       F1=X(1)/X(2)
       F2=X(3)/X(4)
       Z=R*F1
       ZP=R*F2
       T=G(Z)
       IF(T.GT.1.0)GOTO39
       TN=TN+1.0
00039  T1=G(ZP)
       IF(T1.GT.1.0)GOTO45
       P=P+T1
       TN1=TN1+1.0
       PSQ=PSQ+T*T
00044  CONTINUE
00045  CONTINUE
       WRITE(3,3)TN,TN1
       BP=P/TN
       BPSQ=PSQ/TN1
C**    NUMBER OF ITERATIONS USED TO CALCULATE F(X) AND F(X)**2
       WRITE(3,3)BP,BPSQ
C**    CALCULATION OF EFFICIENCY FUNCTION
       EFF=(R/(1.+R))/(BP*(1.-2.*R)+R*(1.+BPSQ))
       WRITE(3,1)R,EFF
00052  CONTINUE
       WRITE(2,5)IS
       STOP
       END
```

```
      SUBROUTINERECTAN(IS,U1,U2,CHECK)                                  04157001
00001 FORMAT(1H1,61HRANDOM NUMBER GENERATOR HAS OBTAINED TWO ZEROS SIMUL04157002
     6TANEOUSLY)                                                        04157003
C         THE NEXT FOUR STATEMENTS PERFORM WHAT IS KNOWN AS THE RESIDUE  04157004
C         CLASS METHOD OF GENERATING RANDOM DIGITS.  THIS METHOD WAS     04157005
C         DEVELOPED BY D.H.LEHMER IN 1951.  THE PROCEEDURE IS TO         04157006
C         GENERATE RANDOM DIGITS BY USING THE RELATION, X(N+1)=K*X(N) IN 04157007
C         MOD M WHICH ACTUALLY MEANS TO DIVIDE K*X(N) THROUGH BY M AND   04157008
C         TO SET X(N+1) EQUAL TO THE REMAINDER.  THE FOLLOWING WAS SET AT04157009
C         THE BEGINNING, X(0)=XXXXXXXX, K=23, AND M=10**8.    ACCORDING  04157010
C         TO TAUSSKY AND TODD, THIS SEQUENCE HAS A PERIOD OF 5,882,352   04157011
C         DIGITS WHICH IS FAR MORE THAN THIS SUBPROGRAM ACTUALLY NEEDS.  04157012
00002 N1=IS*23                                                          04157013
      N2=N1/100000000                                                   04157014
      N3=N2*100000000                                                   04157015
      N4=N1-N2-N3                                                       04157016
      IF(N4.NE.0)GOTO10                                                 04157017
      WRITE(3,1)                                                        04157018
      CHECK=0.0                                                         04157019
      GOTO22                                                            04157020
00010 CHECK=1.0                                                         04157021
      IRA=N4/10000                                                      04157022
      N6=IRA*10000                                                      04157023
      IRB=N4-N6                                                         04157024
      DRA=IRA                                                           04157025
      DRB=IRB                                                           04157026
      U1=DRA/10000.0                                                    04157027
      IF(U1.NE.0.0)GOTO20                                               04157028
      IS=N4                                                             04157029
      GOTO2                                                             04157030
00020 U2=DRB/10000.0                                                    04157031
      IS=N4                                                             04157032
00022 RETURN                                                            04157033
      END                                                               04157034
```

```
C*****ROMBERG NUMERICAL INTEGRATION PROGAM FOR IBM 1410 WRITTEN BY J.O.
C      MINGLE NUCLEAR ENGINEERING, K.S.U. 8-16-65, AND MODIFIED BY R. L.
C      DILLON WHERE,
C      A = LOWER BOUNDARY LIMIT OF INTEGRATION
C      B = UPPER BOUNDARY LIMIT OF INTEGRATION
C      EP = CRITERIUM FOR THE LARGEST FRACTIONAL CHANGE OF THE ANSWER BY
C           AN ADDITIONAL CALCULATION
C      ERROR = ACTUAL FRACTIONAL CHANGE IN LAST TWO SUCCESIVE CALCULATION
C*****MAX = NUMBER OF PARTS THAT INTERVAL (B-A) IS TO BE DIVIDED INTO
C**    DIMENSION T1(MAX,MAX),T2(MAX,MAX)
       DIMENSIONT1(10,10),T2(10,10)
00001  FORMAT(7HLRHO = ,F5.1,3X,8HE(PSI)= ,F10.5,3X,11HE(PSI SQ)= ,F10.5,
      15X,2F10.5)
00002  FORMAT(5X,5HROW =,I3,4X,10F10.4/)
00003  FORMAT(7HLRHO = ,F5.1,3X,6HEFF = ,F10.5)
       G(Z)=((1.+4.*Z-5.*Z*Z+(4.*Z+2.*Z*Z)*ALOG(Z))/(2.*((Z*Z-1.)*ALOG(Z)
      1-2.*(1.-Z)*(1.-Z))))
       F(Z)=((1.+4.*Z-5.*Z*Z+(4.*Z+2.*Z*Z)*ALOG(Z))/(2.*((Z*Z-1.)*ALOG(Z)
      1-2.*(1.-Z)*(1.-Z))))**2
       BETA=1.
       EP=0.01
       A=0.0
       B=1.0
       DO51I=2,100,2
       R=FLOAT(I)/10.
       RHO=R
C**    CALCULATE INTEGRALS ( G(Z) AND F(Z) )
       MAX=10
       DO44K=1,MAX
       N=2**K-1
       H=(B-A)/2.**K
       S1=.25*H
       S2=.125*H
       DO33J=1,N
       Q=A+FLOAT(J)*H
       P=R*(1.-Q)/Q
       IF(P.EQ.1.0)GOTO24
       GOTO27
C**    NEXT TWO STEPS ARE CORRECTION FOR DISCONTINUITY OF PSI AT Z=1
00024  S1=S1+.5*H
       S2=S2+.25*H
       GOTO33
00027  P1=P*P
       P2=P*P*P
       P3=P2*P2
       P4=P2*P1
       S1=S1+H*G(P)
       S2=S2+H*F(P)
00033  CONTINUE
       T1(1,K)=S1
       T2(1,K)=S2
       IF(K.EQ.1)GOTO44
       DO39M=2,K
       T1(M,K)=(4.**M*T1(M-1,K)-T1(M-1,K-1))/(4.**M-1.)
00039  T2(M,K)=(4.**M*T2(M-1,K)-T2(M-1,K-1))/(4.**M-1.)
       ERR1=ABS(T1(K-1,K-1)/T1(K,K)-1.)
       ERR2=ABS(T2(K-1,K-1)/T2(K,K)-1.)
       IF(ERR1.GE.EP)GOTO44
       IF(ERR2.LT.EP)GOTO45
00044  CONTINUE
00045  IF(K.GT.MAX)K=MAX
C**    AREA UNDER INTEGRALS (ANS1=E(PSI)),(ANS2=E(PSI**2))
       ANS1=T1(K,K)*BETA
       ANS2=T2(K,K)*BETA
       WRITE(3,1)RHO,ANS1,ANS2,ERR1,ERR2
C**    CALCULATE EFFICIENCY
       EFF=(RHO/(1.0+RHO))/(RHO+(1.0+RHO)*ANS2-2.0*RHO*ANS1)
       WRITE(3,3)RHO,EFF
00051  CONTINUE
       STOP
       END
```

BAYES UNBIASED ESTIMATION OF THE COMMON MEAN OF TWO NORMAL DISTRIBUTIONS BASED ON SMALL SAMPLES

by

RONALD LEE DILLON

B. S., Kansas State University, 1961

AN ABSTRACT OF A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Statistics

KANSAS STATE UNIVERSITY

Manhattan, Kansas

1967

Given two independent simple random samples from two normal distributions $N(\mu,\sigma_i^2)$ $(i = 1,2)$, the problem is to estimate the common mean μ, $-\infty \leq \mu \leq \infty$, when the variance ratio $\rho = \sigma_2^2/\sigma_1^2$ is unknown.

When ρ is known, the uniformly minimum variance unbiased estimator of μ is the maximum likelihood estimator: $\hat{\mu}_0 = \phi(\rho)\bar{x} + (1 - \phi(\rho))\bar{y}$, where $\phi(\rho) = (n_1/n_2)\rho/(1 + (n_1/n_2)\rho)$ and $(\bar{x}, \bar{y}, n_1, n_2)$ are the sample means and sizes respectively.

This report derives an unbiased estimator for the common mean when ρ is unknown, in which the weight function $\psi(s_2^2/s_1^2)$ is a certain Bayes estimator for $\phi(\rho)$ where $s_i^2(i = 1,2)$ are unbiased estimators for $\sigma_i^2(i = 1,2)$. Explicit formulae for the Bayes estimator $\psi(s_2^2/s_1^2)$ are derived for samples of equal size $n = 3, 5, 7$ and the efficiency functions of the unbiased estimator of μ, determined by there $\psi(s_2^2/s_1^2)$ are studied. For $n=3$, the efficiency of the Bayes unbiased estimator is compared to the efficiency of an unbiased estimator of classical form and is found to be superior.

It is concluded that the Bayes unbiased estimator for the common mean of two normal distributions does offer an improvement over existing procedures when samples are very small.