ESTIMATION OF THE COMMON MEAN OF TWO NORMAL DISTRIBUTIONS

by

Shi-Hwa Yuan

B. A., Fu-Jen Catholic University, 1977

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

1981

Approved by:

David P. Hasza
Major Professor

TABLE OF CONTENTS

CHAPTER 1.

INTRODUCTION AND LITERATURE REVIEW

The problem considered in this report is the estimation of the mean based on two independent samples from normal distributions with common mean and unknown variances. Let $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_n$ be independent i.i.d. samples from the $N(\mu,\sigma_x^2)$ and $N(\mu,\sigma_y^2)$ distributions respectively. The parameter $(\mu,\sigma_x^2,\sigma_y^2)\ \varepsilon(-\infty,\infty)X(0,\infty)X(0,\infty)$ is unknown. We shall consider the estimation of the common mean μ.

Let $\rho = \sigma_y^2/\sigma_x^2$ denote the ratio of the variances. If ρ is known, then the minimal variance unbiased estimator of μ is given by

$\hat{\mu}_o = \frac{\rho\bar{x} + \bar{y}}{1 + \rho}$, where $\bar{x} = n^{-1} \sum_{i=1}^{n} X_i$ and $\bar{y} = n^{-1} \sum_{i=1}^{n} Y_i$. $\hat{\mu}_o$ is also the maximum likelihood estimator of μ when ρ is known. In general, ρ is unknown.

When the ratio ρ of variances is unknown Graybill and Deal (1959), Seshadri (1963), Hogg (1960), Richter (1960) considered the estimation properties for medium-sized and large samples. Graybill and Deal (1959) suggested the use of the estimator

$$\hat{\mu}_{GD} = \frac{s_y^2\bar{x} + s_x^2\bar{y}}{s_x^2 + s_y^2}$$

where

$$s_x^2 = n^{-1} \sum_{i=1}^{n} (X_i - \bar{x})^2$$

$$s_y^2 = n^{-1} \sum_{i=1}^{n} (Y_i - \bar{y})^2 \quad .$$

They showed that the unbiased estimator $\hat{\mu}_{GD}$ has uniformly smaller variance than either sample mean, provided both sample sizes are greater than 10. The efficiency of $\hat{\mu}_{GD}$ has been studied by Graybill and Deal (1959).

S. Zacks (1966), Metha and Gurland (1969) presented estimators and considered small sample sizes. Zacks developed estimators for very small sample sizes. He considered two classes of randomized unbiased procedures. For both classes of estimators a two-sided F-test is performed. If $F = S_y^2/S_x^2$ falls in the interval $(1/\rho^*,\rho^*)$, where ρ^* is a constant , then the estimator $\hat{\mu}_z$ used is $\bar{\mu} = (\bar{x} + \bar{y})/2$. Otherwise, $\hat{\mu}_z$ is $\hat{\mu}_{GD} = \dfrac{S_y^2\bar{x} + S_x^2\bar{y}}{S_x^2 + S_y^2}$ for the first class, whereas for the second class estimator $\tilde{\mu}_z$ is estimated by $\bar{x}$ if $S_y^2/S_x^2 > \rho^*$ and by $\bar{y}$ if $S_y^2/S_x^2 < 1/\rho^*$. The value of ρ^*, in both classes, is the critical value of the F-test of significance according to which one decides whether to apply the estimators $\bar{\mu}$, $\hat{\mu}_{GD}$, $\bar{x}$ or $\bar{y}$.

The variances and the efficiency functions of these estimators are also studied in Zacks (1966) (2.3), (2.4), (3.4), (3.5), respectively. Explicit formulae for the efficiencies were given for the case of samples of size n=3 in Zacks (1966), (2.4) and (3.7). Further, as seen in Fig.(1), the efficiency of the estimator $\hat{\mu}_z$ is higher than that of $\tilde{\mu}_z$, over the range of $1/6 \leq \rho \leq 6$ for all values of ρ^*. Also Zacks concluded that the estimators $\hat{\mu}_z$ are superior to the estimators $\tilde{\mu}_z$.

Recently Cohen and Sackrowitz (1974) obtained a new unbiased estimator for the equal sample size case. They proved that the sample mean of the first population could be improved on provided the sample size is greater than 4; i.e., the estimator is uniformly better than the sample mean based on only one of the populations for $n \geq 5$. A particular estimator

which results from Cohen and Sackrowitz (1974) is

$$\hat{\mu}_{CS} = [1 - C_n G(S_x^2, S_y^2)]\bar{x} + C_n G(S_x^2, S_y^2)\bar{y}$$

where

$$C_n = \begin{cases} (n-3)^2 (n+1)^{-1} (n-1)^{-1} & \text{for } n \text{ odd} \\ (n-4)(n+2)^{-1} & \text{for } n \text{ even} \end{cases}$$

and

$$G(S_x^2, S_y^2) = \begin{cases} F(1, (3-n)/2, (n-1)/2, S_y^2/S_x^2 & 0 \leq S_y^2/S_x^2 \leq 1 \\ \frac{(n-3)}{(n-1)} \cdot \frac{S_x^2}{S_y^2} F(1, (5-n)/2, (n+1)/2, S_y^2/X_x^2) & S_y^2/S_x^2 \geq 1 \end{cases}$$

where F is the hypergeometric function. The estimator is unbiased and minimax for all $n \geq 5$. Although the given estimator is not based only on a sufficient statistic, it has sensible monotonicity properties and has desirable large sample properties.

Brown and Cohen (1974) showed that the sample mean of the first population can be improved on, provided the sample size in that population is greater than 2.

The estimator $\hat{\mu}_{BC} = \bar{x} + (\bar{y} - \bar{x})\{aS_{\bar{x}}^2/[S_{\bar{x}}^2 + (n-1)(S_{\bar{y}}^2/(n+2)) + (\bar{y} - \bar{x})^2/(n-2)]\}$ given by Brown and Cohen (1974) in Remark (2.1) is suggested when it is reasonable to feel that $\sigma_{\bar{y}}^2/\sigma_{\bar{x}}^2$ is large.

The plan of the present study is to derive the maximum likelihood estimator of the common mean μ, and to present some properties of this estimator. Further a Monte Carlo study is used to compare properties of the maximum likelihood estimator with properties of the other estimators presented.

ILLEGIBLE DOCUMENT

THE FOLLOWING DOCUMENT(S) IS OF POOR LEGIBILITY IN THE ORIGINAL

THIS IS THE BEST COPY AVAILABLE

CHAPTER 2.

THE MAXIMUM LIKELIHOOD ESTIMATOR

Let $X_1, \ldots, X_n$ be i.i.d. $N(\mu,\sigma_x^2)$ and let $Y_1, \ldots, Y_n$ be i.i.d. $N(\mu,\sigma_y^2)$ where the X_i's and Y_i's are mutually independent. The parameter $(\mu,\sigma_x^2,\sigma_y^2) \in (-\infty,\infty)X(0,\infty)X(0,\infty)$ is unknown. In this chapter we derive the maximum likelihood estimator of the common mean μ and examine properties of the estimator.

The likelihood function can be written as:

$$L(\mu,\sigma_x^2,\sigma_y^2) = (2\pi)^{-n} (\sigma_x^2)^{-\frac{n}{2}} (\sigma_y^2)^{-\frac{n}{2}} X$$

$$\exp\left[-\frac{1}{2\sigma_x^2}\sum_{i=1}^{n}(x_i-\mu)^2 - \frac{1}{2\sigma_y^2}\sum_{i=1}^{n}(y_i-\mu)^2\right] .$$

Then the logarithm of the likelihood function is given by:

$$L^*(\mu,\sigma_x^2,\sigma_y^2) = \log L(\mu,\sigma_x^2,\sigma_y^2)$$

$$= -n \log (2\pi) - \frac{n}{2}\log(\sigma_x^2) - \frac{n}{2}\log(\sigma_y^2)$$

$$- \frac{1}{2\sigma_x^2}\sum_{i=1}^{n}(x_i-\mu)^2 - \frac{1}{2\sigma_y^2}\sum_{i=1}^{n}(y_i-\mu)^2 .$$

To maximize $L^*(\mu,\sigma_x^2,\sigma_y^2)$ w.r.t. σ_x^2, σ_y^2, we set

$$\left.\frac{\partial L}{\partial \sigma_x^2}\right|_{\hat{\sigma}_x^2} = -\frac{n}{2}\frac{1}{\hat{\sigma}_x^2} + \frac{1}{2\hat{\sigma}_x^4}\sum_{i=1}^{n}(x_i-\mu)^2 = 0$$

and

$$\left.\frac{\partial L}{\partial \sigma_y^2}\right|_{\hat{\sigma}_y^2} = -\frac{n}{2}\frac{1}{\hat{\sigma}_y^2} + \frac{1}{2\hat{\sigma}_y^4}\sum_{i=1}^{n}(y_i-\mu)^2 = 0 .$$

Solving the above equation for $\hat{\sigma}_x^2$ and $\hat{\sigma}_y^2$ we have

$$\hat{\sigma}_x^2 = n^{-1} \sum_{i=1}^{n} (x_i-\mu)^2 \quad ,$$

$$\hat{\sigma}_y^2 = n^{-1} \sum_{i=1}^{n} (y_i-\mu)^2 \quad .$$

Then

$$L^*(\mu,\hat{\sigma}_x^2,\hat{\sigma}_y^2) = -n \log(2\pi) - \frac{n}{2} \log(\hat{\sigma}_x^2) - \frac{n}{2} \log(\hat{\sigma}_y^2) - n$$

$$= -n \log(2\pi) - \frac{n}{2} \log[S_x^2 + (\mu-\bar{x})^2]$$

$$- \frac{n}{2} \log[S_y^2 + (\mu-\bar{y})^2] - n$$

where

$$S_x^2 = n^{-1} \sum_{i=1}^{n} (X_i-\bar{x})^2$$

$$S_y^2 = n^{-1} \sum_{i=1}^{n} (Y_i-\bar{y})^2 \quad .$$

We obtain the maximum likelihood estimator of $(\mu,\sigma_x^2, \sigma_y^2)$ given by $(\hat{\mu},\hat{\sigma}_x^2,\hat{\sigma}_y^2)$ where

$$\hat{\sigma}_x^2 = n^{-1} \sum_{i=1}^{n} (X_i-\hat{\mu})^2$$

$$\hat{\sigma}_y^2 = n^{-1} \sum_{i=1}^{n} (Y_i-\hat{\mu})^2 \quad ,$$

by selecting $\hat{\mu}$ to maximize $L^*(\mu,\hat{\sigma}_x^2,\hat{\sigma}_y^2)$ w.r.t. μ; i.e., $\hat{\mu}$ is chosen to minimize

$$\log[S_x^2 + (\mu-\bar{x})^2] + \log[S_y^2 + (\mu-\bar{y})^2]$$

$$= \log\{[S_x^2 + (\mu-\bar{x})^2] \cdot [S_y^2 + (\mu-\bar{y})^2]\}$$

or equivalently minimize

$$[S_x^2 + (\mu-\bar{x})^2] \cdot [S_y^2 + (\mu-\bar{y})^2]$$

$$= A + (\mu-\bar{x})^2 S_y^2 + (\mu-\bar{y})^2 S_x^2 + (\mu-\bar{x})^2(\mu-\bar{y})^2$$

where A is a constant which does not depend on μ.

Now let

$$g(\hat{\mu}) = (\hat{\mu}-\bar{x})^2(\hat{\mu}-\bar{y})^2 + S_x^2(\hat{\mu}-\bar{y}) + S_y^2(\hat{\mu}-\bar{x})^2 \quad .$$

Then we have

$$g'(\hat{\mu}) = 2S_x^2(\hat{\mu}-\bar{y}) + 2S_y^2(\hat{\mu}-\bar{x}) + 2(\hat{\mu}-\bar{x})(\hat{\mu}-\bar{y})(2\hat{\mu}-\bar{x}-\bar{y}) \quad ,$$

$$g''(\hat{\mu}) = 2(S_x^2+S_y^2) + 4(\hat{\mu}-\bar{x})(\hat{\mu}-\bar{y}) + 2(2\hat{\mu}-\bar{x}-\bar{y})^2 \quad ,$$

$$g'''(\hat{\mu}) = 12(2\hat{\mu}-\bar{x}-\bar{y}) \quad ,$$

and obtain

$$g'(\bar{x}) = 2S_x^2(\bar{x}-\bar{y}) \quad ,$$

$$g'(\bar{y}) = 2S_y^2(\bar{y}-\bar{x}) \quad ,$$

$$g'\left(\frac{\bar{x}+\bar{y}}{2}\right) = S_x^2(\bar{x}-\bar{y}) + S_y^2(\bar{y}-\bar{x}) = (\bar{x}-\bar{y})(S_x^2-S_y^2) \quad .$$

Further,

$$g''(\bar{x}) = 2(S_x^2+S_y^2) + 2(\bar{x}-\bar{y})^2 \quad ,$$

$$g''(\bar{y}) = 2(S_x^2+S_y^2) + 2(\bar{x}-\bar{y})^2 \quad ,$$

$$g''(\tfrac{\bar{x}+\bar{y}}{2}) = 2(S_x^2+S_y^2) - (\bar{x}-\bar{y})^2 \quad ,$$

and

$$g'''(\tfrac{\bar{x}+\bar{y}}{2}) = 0 \ .$$

If $S_x^2 \leq S_y^2$, by Lemma, then $g(\hat{\mu})$ attains its absolute minimum between $\bar{x}$ and $\frac{\bar{x}+\bar{y}}{2}$. Similarly if $S_x^2 \geq S_y^2$, $g(\hat{\mu})$ attains its absolute minimum between $\bar{y}$ and $\frac{\bar{x}+\bar{y}}{2}$.

Lemma: Let $g(\mu) = (\mu-\bar{x})^2S_y^2 + (\mu-\bar{y})^2S_x^2 + (\mu-\bar{x})^2(\mu-\bar{y})^2$

then result: $g(\mu)$ attains its absolute minimum

i) in $[\bar{x}, \frac{\bar{x}+\bar{y}}{2}]$ if $\bar{x} \leq \bar{y}$ and $S_x^2 \leq S_y^2$

ii) in $[\frac{\bar{x}+\bar{y}}{2}, \bar{x}]$ if $\bar{y} \leq \bar{x}$ and $S_x^2 \leq S_y^2$

iii) in $[\bar{y}, \frac{\bar{x}+\bar{y}}{2}]$ if $\bar{y} \leq \bar{x}$ and $S_y^2 \leq S_x^2$

iv) in $[\frac{\bar{x}+\bar{y}}{2}, \bar{y}]$ if $\bar{x} \leq \bar{y}$ and $S_y^2 \leq S_x^2$

Proof: Without loss of generality consider i) Suppose $\bar{x} \leq \bar{y}$ and $S_x^2 \leq S_y^2$, clearly $g(\mu)$ attains its absolutely minimum in $[\bar{x}, \bar{y}]$. For if $\mu_o \notin [\bar{x}, \bar{y}]$, there exists $\mu_1 \in [\bar{x}, \bar{y}]$ such that $(\mu_1-\bar{x})^2 \leq (\mu_o-\bar{x})^2$ and $(\mu_1-\bar{y})^2 \leq (\mu_o-\bar{y})^2$

Now let $\mu_o \in [\bar{y}, \frac{\bar{x}+\bar{y}}{2}]$

let $\mu_1 = \frac{\bar{x}+\bar{y}}{2} - [\mu_o - \frac{\bar{x}+\bar{y}}{2}]$

then $(\mu_1-\bar{x})^2 = (\mu_o-\bar{y})^2$

$(\mu_1-\bar{y})^2 = (\mu_o-\bar{x})^2$

Since $S_x^2 \leq S_y^2$ it follows $g(\mu_1) \leq g(\mu_o)$.

Without loss of generality we consider the case in which $\bar{x} < \bar{y}$ and $S_x^2 \leq S_y^2$. We shall show that $g(\mu)$ has one and only one relative minimum in $[\bar{x}, \frac{\bar{x}+\bar{y}}{2}]$. We shall verify that $g'(\mu)$ has one and only one zero in $[\bar{x}, \frac{\bar{x}+\bar{y}}{2}]$. This zero is then the maximum likelihood estimator of μ.

In the case $\bar{x} < \bar{y}$ and $S_x^2 \leq S_y^2$ we obtain $g'(x) = 2S_x^2(\bar{x}-\bar{y}) < 0$ and $g'(\frac{\bar{x}+\bar{y}}{2}) = (\bar{x}-\bar{y})(S_x^2-S_y^2) > 0$. Thus $g'(\mu)$ has at least one zero in $[\bar{x}, \frac{\bar{x}+\bar{y}}{2}]$.

We now consider two different cases:

Case I: Assume $g''(\frac{\bar{x}+\bar{y}}{2}) = 2(S_x^2+S_y^2)(\bar{x}-\bar{y})^2 > 0$. We have $g'''(\bar{x}) < 0$, $g'''(\bar{y}) > 0$, $g'''(\frac{\bar{x}+\bar{y}}{2}) = 0$. Thus $g''(\hat{\mu})$ attains its absolute minimum at $\frac{\bar{x}+\bar{y}}{2}$. By assumption $g''(\frac{\bar{x}+\bar{y}}{2}) > 0$ implying $g''(\hat{\mu}) > 0$ on $[\bar{x}, \bar{y}]$. Therefore $g''(\hat{\mu})$ is increasing on $[\bar{x}, \bar{y}]$, and $g'(\hat{\mu})$ is increasing on $[\bar{x}, \frac{\bar{x}+\bar{y}}{2}]$. We have $g'(\bar{x}) < 0$, $g'(\frac{\bar{x}+\bar{y}}{2}) > 0$. Thus $g'(\hat{\mu})$ has one and only one zero in $[\bar{x}, \frac{\bar{x}+\bar{y}}{2}]$.

Case II: Assume $g''(\frac{\bar{x}+\bar{y}}{2}) < 0$. Then $g'(\hat{\mu})$ is decreasing at $\frac{\bar{x}+\bar{y}}{2}$. Since $g'(\hat{\mu})$ is a polynomial in $\hat{\mu}$ of degree 3, $g'(\hat{\mu})$ has either one zero or three zeroes in $[\bar{x}, \frac{\bar{x}+\bar{y}}{2}]$. If $g'(\hat{\mu})$ has three zeros in $[\bar{x}, \frac{\bar{x}+\bar{y}}{2}]$, then because $g'(\bar{x}) < 0$, $g'(\frac{\bar{x}+\bar{y}}{2}) > 0$ and $g''(\frac{\bar{x}+\bar{y}}{2}) < 0$, $g'(\hat{\mu})$ would have three critical values in $[\bar{x}, \frac{\bar{x}+\bar{y}}{2}]$, which would imply $g''(\hat{\mu})$ has three zeros in $[\bar{x}, \frac{\bar{x}+\bar{y}}{2}]$, which is in contradiction to $g''(\hat{\mu})$ having degree 2.

Thus $g'(\hat{\mu})$ has only one zero in $[\bar{x}, \frac{\bar{x}+\bar{y}}{2}]$. It follows that $g'(\hat{\mu})$ has one and only one zero in $[\bar{x}, \frac{\bar{x}+\bar{y}}{2}]$. Now let $\hat{\theta} = \hat{\mu} - \frac{\bar{x}+\bar{y}}{2}$ and substitute $\hat{\theta}$ into $g'(\hat{\mu})$ to obtain

$$g'(\hat{\mu}) = 2s_x^2[\hat{\theta} - \frac{\bar{y}-\bar{x}}{2}] + 2s_y^2[\hat{\theta} - \frac{\bar{x}-\bar{y}}{2}] + 4\hat{\theta}[\hat{\theta} - \frac{\bar{y}-\bar{x}}{2}][\hat{\theta} - \frac{\bar{x}-\bar{y}}{2}]$$

$$= 4\hat{\theta}^3 + 4\hat{\theta}[\frac{1}{2}(s_x^2+s_y^2) - \frac{1}{4}(\bar{x}-\bar{y})^2] + (\bar{x}-\bar{y})(s_x^2-s_y^2) \quad .$$

Then

$$\frac{1}{4}g'(\hat{\mu}) = \hat{\theta}^3 + \hat{\theta}[\frac{1}{2}(s_x^2+s_y^2) - \frac{1}{4}(\bar{x}-\bar{y})^2] + \frac{1}{4}(\bar{x}-\bar{y})(s_x^2-s_y^2)$$

$$\overset{\text{Say}}{=} f'(\hat{\theta}) \quad .$$

The zero of $g'(\hat{\mu})$ in $[\bar{x}, \frac{\bar{x}+\bar{y}}{2}]$ is the zero of $f(\hat{\theta})$ lying in $[\frac{\bar{x}-\bar{y}}{2}, 0]$. Then the zero of $f(\hat{\theta})$ in $[\frac{\bar{x}-\bar{y}}{2}, 0]$ is given by

$$\hat{\theta} = \left\{\frac{(\bar{x}-\bar{y})(s_y^2-s_x^2)}{8} + \left[\frac{(\bar{x}-\bar{y})^2(s_y^2-s_x^2)^2}{64} + \frac{[(s_x^2+s_y^2) - \frac{1}{2}(\bar{x}-\bar{y})^2]^3}{216}\right]^{\frac{1}{2}}\right\}^{\frac{1}{3}}$$

$$+ \left\{\frac{(\bar{x}-\bar{y})(s_y^2-s_x^2)}{8} - \left[\frac{(\bar{x}-\bar{y})^2(s_y^2-s_x^2)^2}{64} + \frac{[(s_x^2+s_y^2) - \frac{1}{2}(\bar{x}-\bar{y})^2]^3}{216}\right]^{\frac{1}{2}}\right\}^{\frac{1}{3}} \quad .$$

If the quantities inside the large brackets are imaginary the cube root is taken so that it is analytic of the negative real axis. Therefore $\Psi(\bar{x}, \bar{y}, s_x^2, s_y^2)$ the maximum likelihood estimator of μ is given by $\hat{\mu} = \frac{\bar{x}+\bar{y}}{2} + \hat{\theta}$. Therefore

$$\hat{\mu}_{MLE} = \frac{\bar{x}+\bar{y}}{2} + \left\{\frac{(\bar{x}-\bar{y})(s_y^2-s_x^2)}{8} + \left[\frac{(\bar{x}-\bar{y})^2(s_y^2-s_x^2)^2}{64} + \frac{[(s_x^2+s_y^2) - \frac{1}{2}(\bar{x}-\bar{y})^2]^3}{216}\right]^{\frac{1}{2}}\right\}^{\frac{1}{3}}$$

$$+ \left\{\frac{(\bar{x}-\bar{y})(s_y^2-s_x^2)}{8} - \left[\frac{(\bar{x}-\bar{y})^2(s_y^2-s_x^2)^2}{64} + \frac{[(s_x^2+s_y^2) - \frac{1}{2}(\bar{x}-\bar{y})^2]^3}{216}\right]^{\frac{1}{2}}\right\}^{\frac{1}{3}} \quad .$$

We now present several results concerning $\hat{\mu}_{MLE}$. In this report, we concentrate on the equal sample size problem. When the sample sizes are

n_1 and n_2, the maximum likelihood estimator is

$$\hat{\mu}_{MLE} = \bar{x} + (1-\alpha)\bar{y} + -\left[\frac{C_2}{2} + \left[\frac{C_2^2}{2} + \frac{C_1^3}{3}\right]^{\frac{1}{2}}\right]^{\frac{1}{3}}$$

$$+ -\left[\frac{C_2}{2} - \left[\frac{C_2^2}{2} + \frac{C_1^3}{3}\right]^{\frac{1}{2}}\right]^{\frac{1}{3}}$$

where

$$\alpha = \frac{n_1 + 2n_2}{3(n_1+n_2)}$$

$$C_1 = \frac{n_1 s_y^2 + n_2 s_x^2}{n_1 + n_2} - \frac{1}{3}(x-y)^2\left[1 - \frac{n_1 n_2}{(n_1+n_2)^2}\right]$$

$$C_2 = (\bar{x}-\bar{y})\left[\frac{\alpha n_2 S_x^2-(1-\alpha)n_1 S_y^2}{n_1+n_2}\right] + \frac{(\bar{x}-\bar{y})^3}{3}\left[\frac{\alpha(1-\alpha)(n_2-n_1)}{n_1+n_2}\right] .$$

Theorem 1. The maximum likelihood estimator of μ, $\hat{\mu}$, has a symmetric distribution about the common mean μ.

Proof: We shall show that $F_{\hat{\mu}-\mu} = F_{\mu-\hat{\mu}}$; i.e., that $\hat{\mu} - \mu$ and $\mu - \hat{\mu}$ are identically distributed. Now

$$\hat{\mu} - \mu = \frac{(\bar{x}-\mu)-(\bar{y}-\mu)}{2} + \left\{\frac{((\bar{x}-\mu)-(\bar{y}-\mu))(S_y^2-S_x^2)}{8} + \right.$$

$$\left. + \left[\frac{((\bar{x}-\mu)-(\bar{y}-\mu))^2(S_y^2-S_x^2)^2}{64} + \frac{[(S_x^2+S_y^2)-\frac{1}{2}((\bar{x}-\mu)-(\bar{y}-\mu))^2]^3}{216}\right]^{\frac{1}{2}}\right\}^{\frac{1}{3}}$$

$$+\left\{\frac{((\bar{x}-\mu)-(\bar{y}-\mu))(S_y^2-S_x^2)}{8} - \left[\frac{((\bar{x}-\mu)-(\bar{y}-\mu))^2(S_y^2-S_x^2)^2}{64}\right.\right.$$

$$\left.\left. + \frac{[(S_x^2+S_y^2)-\frac{1}{2}((\bar{x}-\mu)-(\bar{y}-\mu))^2]^3}{216}\right]^{\frac{1}{2}}\right\}^{\frac{1}{3}} .$$

Therefore $\hat{\mu} - \mu$ is a function of $\bar{x} - \mu$, $\bar{y} - \mu$, S_x^2, S_y^2 which we denote by $\hat{\mu} - \mu = g(\bar{x}-\mu, \bar{y}-\mu, S_x^2, S_y^2)$. Further,

$$\mu - \hat{\mu} = \frac{(\mu-\bar{x})-(\mu-\bar{y})}{2} + \left\{\frac{[(\mu-\bar{x})-(\mu-\bar{y})]}{8}\right.$$

$$\left. - \left[\frac{[(\mu-\bar{x})-(\mu-\bar{y})]^2(S_y^2-S_x^2)}{64} - \frac{[(S_x^2+S_y^2)-\frac{1}{2}((\mu-\bar{x})-(\mu-\bar{y}))^2]^3}{216}\right]^{\frac{1}{2}}\right\}^{\frac{1}{3}}$$

$$+\left\{\frac{[(\mu-\bar{x})-(\mu-\bar{y})]}{8} + \left[\frac{[(\mu-\bar{x})-(\mu-\bar{y})]^2(S_y^2-S_x^2)^2}{64}\right.\right.$$

$$\left.\left. - \frac{[(S_x^2+S_y^2)-\frac{1}{2}((\mu-x)-(\mu-y))^2]^3}{216}\right]^{\frac{1}{2}}\right\}^{\frac{1}{3}}$$

and $\mu - \hat{\mu} = g(\mu-\bar{x}, \mu-\bar{y}, S_x^2, S_y^2)$.

The result then follows because $(\mu-\bar{x}, \mu-\bar{y}, S_x^2, S_y^2)$ and $(\bar{x}-\mu, \bar{y}-\mu, S_x^2, S_y^2)$ are identically distributed.

Theorem 2. *All the moments of* $\hat{\mu}$ exist and are finite.

Proof: Because $\hat{\mu} \in [\min\{\bar{x},\bar{y}\} , \max\{\bar{x},\bar{y}\}]$

$$|\hat{\mu}| \leq |\bar{x}| + |\bar{y}|$$

$$|\hat{\mu}|^k \leq |\max\{\bar{x},\bar{y}\}|^k$$

$$\leq |\bar{x}|^k + |\bar{y}|^k$$

Then we obtain $E[|\hat{\mu}|^k] \leq E[|\bar{x}|^k+|\bar{y}|^k] = E[|\bar{x}|^k] + E[|\bar{y}|^k]$.

Since $\bar{x}$, $\bar{y}$ are normally distributed, $\bar{x} \sim N(\mu,\sigma_1^2/n)$, $\bar{y} \sim N(\mu,\sigma_2^2/n)$, $E[|\bar{x}|^r] < \infty$, and $E[|\bar{y}|^r] < \infty$, $\forall$ r. Then $E[|\bar{x}|^k] + E[|\bar{y}|^k] < \infty \; \forall$ k. Hence all moments of $\hat{\mu}$ exist and all finite.

Theorem 3: The estimator $\hat{\mu}$ is unbiased for μ.

Proof: By Theorem 2, $E[|\hat{\mu}|^k]$ exists and is finite for $\forall$ k. Let k = 1 then $E[|\hat{\mu}|]$ exists and is finite. Also by Theorem 1, $\hat{\mu}$ has a symmetric distribution about μ and therefore $E[\hat{\mu}] = \mu$.

Theorem 4: $\hat{\mu}$ is invariant with respect to location, scale, and the naming of the two samples.

Proof: To show $\hat{\mu}$ is location invariant we have

$$\hat{\mu}(\bar{x}, \bar{y}, S_x^2, S_y^2) \overset{\text{say}}{=} t(x_1,\ldots,x_n, y_1,\ldots,y_n).$$

Also

$$t(x_1+C,\ldots,x_n+C,\ y_1+C,\ldots,y_n+C) = \frac{\bar{x}+\bar{y}}{2} + C$$

$$+ \left\{ \frac{(\bar{x}-\bar{y})(s_y^2-s_x^2)}{8} + \left[\frac{(\bar{x}-\bar{y})^2(s_y^2-s_x^2)^2}{64} + \frac{(s_x^2+s_y^2-\frac{1}{2}(\bar{x}-\bar{y})^2)^3}{216} \right]^{\frac{1}{2}} \right\}^{\frac{1}{3}}$$

$$+ \left\{ \frac{(\bar{x}-\bar{y})(s_y^2-s_x^2)}{8} - \left[\frac{(\bar{x}-\bar{y})^2(s_y^2-s_x^2)^2}{64} + \frac{(s_x^2+s_y^2-\frac{1}{2}(\bar{x}-\bar{y})^2)^3}{216} \right]^{\frac{1}{2}} \right\}^{\frac{1}{3}}$$

$$= t(x_1,\ldots,x_n,\ y_1,\ldots,y_n) + C$$

$$= \hat{\mu}(\bar{x},\ \bar{y},\ s_x^2,\ s_y^2) + C.$$

To show $\hat{\mu}$ is scalar invariant we have

$$t(Cx_1,\ldots,Cx_n,\ Cy_1,\ldots,Cy_n) = \frac{C(\bar{x}+\bar{y})}{2}$$

$$+ \left\{ \frac{C^3(\bar{x}-\bar{y})(s_y^2-s_x^2)}{8} + \left[\frac{C^2(\bar{x}-\bar{y})^2C^4(s_y^2-s_x^2)^2}{64} + \frac{(C^2(s_x^2+s_y^2-\frac{1}{2}(\bar{x}-\bar{y})^2)^3}{216} \right]^{\frac{1}{2}} \right\}^{\frac{1}{3}}$$

$$+ \left\{ \frac{C^3(\bar{x}-\bar{y})(s_y^2-s_x^2)}{8} - \left[\frac{C^2(\bar{x}-\bar{y})^2C^4(s_y^2-s_x^2)^2}{64} \right.\right.$$

$$\left.\left. + \frac{(C^2(s_x^2+s_y^2-\frac{1}{2}(\bar{x}-\bar{y})^2)^3}{216} \right]^{\frac{1}{2}} \right\}^{\frac{1}{3}}$$

$$= Ct(x_1,\ldots,x_n,\ y_1,\ldots,y_n) = C\hat{\mu}(\bar{x},\ \bar{y},\ s_x^2,\ s_y^2)\ .$$

To show that $\hat{\mu}$ is invariant with respect to the naming of the variables, we shall show that

$$\hat{\mu}(\bar{x},\ \bar{y},\ s_x^2,\ s_y^2) = \hat{\mu}(\bar{y},\ \bar{x},\ s_y^2,\ s_x^2).$$

Because

$$\frac{(\bar{x}-\bar{y})(s_y^2-s_x^2)}{8} + \left[\frac{(\bar{x}-\bar{y})(s_y^2-s_x^2)^2}{64} + \frac{[s_x^2+s_y^2-\frac{1}{2}(\bar{x}-\bar{y})^2]^3}{216}\right]^{\frac{1}{2}}$$

$$= \frac{(-(\bar{x}-\bar{y}))(-(s_y^2-s_x^2))}{8} + \left[\frac{(-(\bar{x}-\bar{y}))^2(-(s_y^2-s_x^2))^2}{64}\right.$$

$$\left. + \frac{[s_y^2+s_x^2-\frac{1}{2}(-(\bar{x}-\bar{y}))^2]^3}{216}\right]^{\frac{1}{2}}$$

$$= \frac{(\bar{y}-\bar{x})(s_x^2-s_y^2)}{8} + \left[\frac{(\bar{y}-\bar{x})^2(s_x^2-s_y^2)^2}{64} + \frac{[s_y^2+s_x^2-\frac{1}{2}(\bar{y}-\bar{x})^2]^3}{216}\right]^{\frac{1}{2}} ,$$

$$\hat{\mu}(\bar{x}, \bar{y}, s_x^2, s_y^2) = \frac{\bar{y}+\bar{x}}{2} +$$

$$\left\{\frac{(\bar{y}-\bar{x})(s_x^2-s_y^2)}{8} + \frac{(\bar{y}-\bar{x})^2(s_x^2-s_y^2)^2}{64} + \frac{[s_y^2+s_x^2-\frac{1}{2}(\bar{y}-\bar{x})^2]^3}{216}\right.^{\frac{1}{2}}\Bigg\}^{\frac{1}{3}}$$

$$+\left\{\frac{(\bar{y}-\bar{x})(s_x^2-s_y^2)}{8} - \frac{(\bar{y}-\bar{x})^2(s_x^2-s_y^2)^2}{64} + \frac{[s_y^2+s_x^2-\frac{1}{2}(\bar{y}-\bar{x})^2]^3}{216}\right.^{\frac{1}{2}}\Bigg\}^{\frac{1}{3}}$$

$$= \hat{\mu}(\bar{y}, \bar{x}, s_y^2, s_x^2) \quad .$$

<u>Theorem</u>: $\hat{\mu}_{MLE} = \hat{\mu}_{GD} + O_p(n^{-1})$ and $\sqrt{n}(\hat{\mu}_{MLE} - \mu) \xrightarrow{L} N(0, \sigma_x^2\sigma_y^2/\sigma_x^2+\sigma_y^2)$.

Proof:

$$\hat{\mu}_{MLE} = \frac{(\bar{x}-\bar{y})}{2} + \left\{\frac{(\bar{x}-\bar{y})(s_y^2-s_x^2)}{8} + \left[\frac{(\bar{x}-\bar{y})^2(s_y^2-s_x^2)^2}{64} + \frac{[(s_x^2+s_y^2)-\frac{1}{2}(\bar{x}-\bar{y})^2]^3}{216}\right]^{\frac{1}{2}}\right\}^{\frac{1}{3}} + \left\{\frac{(\bar{x}-\bar{y})(s_y^2-s_x^2)}{8} - \left[\frac{(\bar{x}-\bar{y})^2(s_y^2-s_x^2)^2}{64} + \frac{[(s_x^2+s_y^2)-\frac{1}{2}(\bar{x}-\bar{y})^2]^3}{216}\right]^{\frac{1}{2}}\right\}^{\frac{1}{3}} .$$

Let

$$A = \frac{(\bar{x}-\bar{y})(s_y^2-s_x^2)}{8}$$

$$B = \frac{s_x^2+s_y^2-\frac{1}{2}(\bar{x}-\bar{y})^2}{6} .$$

Therefore

$$\hat{\mu}_{MLE} = \frac{\bar{x}+\bar{y}}{2} + [A + \sqrt{A^2 + B^3}\ ]^{\frac{1}{3}} + [A - \sqrt{A^2 + B^3}\ ]^{\frac{1}{3}}$$

$$= \frac{\bar{x}+\bar{y}}{2} + [A + B^{\frac{3}{2}}\sqrt{1 + A^2/B^3}\ ]^{\frac{1}{3}} + [A - B^{\frac{3}{2}}\sqrt{1 + A^2/B^3}\ ]^{\frac{1}{3}} .$$

Since $\bar{x}-\bar{y} \sim N(0, \frac{\sigma_x^2+\sigma_y^2}{n})$, by Corollary (5.1.1.1.) in Fuller (1976), we have

$$\bar{x}-\bar{y} = O_p(n^{-\frac{1}{2}}) .$$

Further

$$\frac{\frac{S_y^2-S_x^2}{8}}{\left\{\frac{(S_x^2+S_y^2-\frac{1}{2}(\bar{x}-\bar{y})^2)^2}{6}\right\}^{\frac{3}{2}}} \xrightarrow{P} \frac{6^{\frac{3}{2}}}{8} \cdot \frac{\sigma_y^2-\sigma_x^2}{(\sigma_y^2+\sigma_x^2)^{3/2}}$$

implying

$$\frac{A}{B^{\frac{3}{2}}} = O_p(n^{-\frac{1}{2}}) \quad \text{and}$$

$$\frac{A^2}{B^3} = O_p(n^{-1}) \quad .$$

By Corollary (5.1.5.) in Fuller (1976) with

$$X_n = 1 + \frac{A^2}{B^3} \ , \ g(x) = \sqrt{x}$$

we have

$$\sqrt{1 + A^2/B^3} = 1 + O_p(n^{-1})$$

Now

$$\hat{\mu}_{MLE} = \frac{\bar{x}+\bar{y}}{2} + [A + B^{\frac{3}{2}}(1 + O_p(n^{-1}))]^{\frac{1}{3}} + [A - B^{\frac{3}{2}}(1+O_p(n^{-1}))]^{\frac{1}{3}}$$

$$= \frac{\bar{x}+\bar{y}}{2} + B^{\frac{1}{2}}\left[1 - \frac{A}{B^{\frac{3}{2}}} + O_p(n^{-1})\right]^{\frac{1}{3}} \quad .$$

By Corollary (5.1.1.1) in Fuller (1976)

$$[1 + \frac{A}{B^{\frac{3}{2}}} + O_p(n^{-1})]^{\frac{1}{3}} = 1 + \frac{A}{3B^{\frac{3}{2}}} + O_p(n^{-1}) \quad .$$

Now

$$\hat{\mu}_{MLE} = \frac{\bar{x}+\bar{y}}{2} + B^{\frac{1}{2}}(1 + \frac{A}{3B^{\frac{3}{2}}}) - B^{\frac{1}{2}}(1 - \frac{A}{3B^{\frac{3}{2}}}) + O_p(n^{-1})$$

$$= \frac{\bar{x}+\bar{y}}{2} + \frac{2}{3}(\frac{A}{B^{\frac{3}{2}}}) + O_p(n^{-1})$$

$$= \frac{\bar{x}+\bar{y}}{2} + \frac{2}{3} \cdot \frac{A}{(S_x^2+S_y^2)/6} + O_p(n^{-1})$$

$$= \frac{\bar{x}+\bar{y}}{2} + \frac{1}{2} \cdot \frac{(\bar{x}-\bar{y})(S_y^2-S_x^2)}{S_x^2+S_y^2} + O_p(n^{-1})$$

$$= \frac{1}{2} \cdot \frac{2(\bar{x}S_y^2+\bar{y}S_x^2)}{S_x^2+S_y^2} + O_p(n^{-1})$$

$$= \frac{\bar{x}S_y^2+\bar{y}S_x^2}{S_x^2+S_y^2} + O_p(n^{-1})$$

$$= \hat{\mu}_{GD} + O_p(n^{-1}).$$

Therefore $\hat{\mu}_{MLE} = \hat{\mu}_{GD} + O_p(n^{-1})$.

Then $\hat{\mu}_{MLE}$ has the same asymptotic distribution as $\hat{\mu}_{GD}$. Because

$$S_x^2 \xrightarrow{P} \sigma_x^2 \quad \text{and} \quad S_y^2 \xrightarrow{P} \sigma_y^2 \quad ,$$

$$\sqrt{n}\left[\frac{\bar{x}S_y^2+\bar{y}S_x^2}{S_x^2+S_y^2} - \mu\right]$$

has the same asymptotic distribution as

$$\sqrt{n}\left[\frac{\bar{x}\sigma_y^2+\bar{y}\sigma_x^2}{\sigma_x^2+\sigma_y^2}-\mu\right]$$

$$=\sqrt{n}\left[\frac{(\bar{x}-\mu)\sigma_y^2+(\bar{y}-\mu)\sigma_x^2}{\sigma_x^2+\sigma_y^2}\right]$$

which converges in law to the $N(0,\frac{\sigma_x^2\sigma_y^2}{\sigma_x^2+\sigma_y^2})$ distribution. Therefore

$$\sqrt{n}\,(\hat{\mu}_{MLE}-\mu)\xrightarrow{L}N(0,\frac{\sigma_x^2\sigma_y^2}{\sigma_x^2+\sigma_y^2}).$$

CHAPTER 3.

MONTE CARLO STUDY

In this chapter the mean square error for $\hat{\mu}_{GD}$, $\hat{\mu}_{MLE}$, $\hat{\mu}_{CS}$, $\hat{\mu}_{BC}$, and $\hat{\mu}_{Z}$ are examined for various values of ρ and n where n denotes the degrees of freedom for S_x^2 and S_y^2. In Table 1 the variance of $\hat{\mu}_{GD}$ is given for the various values of ρ and n. In every case it is n times the variance which is tabulated. These entries are obtained using the infinite series formula (2.1) given in Nair (1980). For the other estimators exact formulae for the variances are not available and a Monte Carlo study was undertaken.

To generate independent standard normal deviates the method described in Marsaglia, Ananthanrayanan, and Paul (1976) was used. The values of n used were 2, 4, 6, 8, 10, 20, ∞ and the values of ρ used were .001, .01, .05, .1, .2, .3, .4, .5, .6, .7, .8, .9, and 1.0. In estimating the mean square error at each n, combination 25,000 samples were used to obtain the sample variances for n = 2, 4, 6, 8, 10. For n = 20 a sample of size 10,000 was used. For n = ∞ the asymptotic distribution is used to obtain the tabulated value.

Tables 2 to 4 give the estimated mean square errors from $\hat{\mu}_{MLE}$, $\hat{\mu}_{CS}$, and $\hat{\mu}_{BC}$. Table 5A to 5D gives the estimated mean square errors for $\hat{\mu}_{Z}$ with associated significance levels α = .01, .05, .10 and .25 respectively.

Tables 6 to 8 give the efficiencies of $\hat{\mu}_{MLE}$, $\hat{\mu}_{CS}$, and $\hat{\mu}_{BC}$, relative to $\hat{\mu}_{GD}$. The efficiencies are given relative to $\hat{\mu}_{GD}$ because $\hat{\mu}_{GD}$ is the estimator commonly used. The efficiencies for $\hat{\mu}_{MLE}$ for each n were smoothed using a regression function of the form $a + b \exp\{-c_1\rho^{-1/2} - c_2\rho\}$. Tables 9A to 9D give the efficiencies of $\hat{\mu}_{Z}$ with α = .01, .05, .10, and .25 relative to $\hat{\mu}_{GD}$.

In every case the actual tabulated value is the mean square error of the estimator multiplied by the degrees of freedom associated with both S_x^2 and S_y^2 and the values for $n = \infty$ are obtained from the asymptotic distribution of the estimator.

TABLE 1. VARIANCE FOR GRAYBILL-DEAL ESTIMATOR

ρ \ n	2	4	6	8	10	20	∞
.001	.00196	.00116	.00099	.00101	.00101	.00099	.00100
.010	.01853	.01086	.01026	.00994	.00987	.00987	.00990
.050	.08187	.05615	.05106	.04945	.04946	.04816	.04760
.100	.14381	.11009	.10117	.09623	.09440	.09203	.09090
.200	.24705	.20297	.18876	.18220	.17854	.17198	.16667
.300	.32819	.28018	.26287	.25428	.24923	.23955	.23080
.400	.39618	.34561	.32614	.31604	.30992	.29767	.28570
.500	.45482	.40202	.28080	.36948	.36248	.34808	.33330
.600	.50643	.45131	.42852	.41614	.40840	.39216	.37500
.700	.55250	.49484	.47057	.45724	.44883	.43099	.41180
.800	.59410	.53364	.50793	.49371	.48470	.46545	.44440
.900	.63225	.56850	.54135	.52629	.51671	.49620	.47370
1.000	.66667	.60000	.57140	.55556	.54555	.52380	.50000

NOTE: n denotes the degrees of freedom for S_x^2 and S_y^2. ρ denotes σ_x^2/σ_y^2. The entries in the table give $nV(\hat{\mu}_{GD})$ when $\sigma_y^2 = 1$.

TABLE 2 . VARIANCE FOR MAXIMUM LIKELIHOOD ESTIMATOR

ρ \ n	2	4	6	8	10	20	∞
.001	.00216	.00113	.00098	.00099	.00100	.00099	.00100
.010	.02042	.01055	.01011	.00977	.00974	.00986	.00990
.050	.09072	.05489	.05031	.04963	.04883	.04810	.04760
.100	.16094	.11012	.10020	.09474	.09330	.09191	.09090
.200	.27997	.21008	.19031	.18103	.17761	.17175	.16667
.300	.37461	.29554	.26900	.25547	.24967	.23927	.23080
.400	.45437	.36880	.33737	.32055	.31223	.29773	.28570
.500	.52341	.43233	.39704	.37763	.36680	.35122	.33333
.600	.58433	.48802	.44950	.42794	.41470	.39274	.37500
.700	.63882	.53732	.49592	.47256	.45703	.43233	.41180
.800	.68808	.58132	.53731	.51236	.49468	.46775	.44444
.900	.73332	.62089	.57443	.54805	.52834	.49964	.47370
1.000	.77417	.65667	.60786	.58023	.55872	.52852	.50000

NOTE: n denotes the degrees of freedom for S_x^2 and S_y^2. ρ denotes σ_x^2/σ_y^2. The entires in the table give $nV(\hat{\mu}_{MLE})$ when $\sigma_y^2 = 1$.

TABLE 3. VARIANCE FOR COHEN-SACKROWITZ ESTIMATOR

ρ \ n	2	4	6	8	10	20	∞
.001	.00099	.00099	.00098	.00101	.00100	.00099	.00100
.010	.00989	.00995	.00998	.00982	.00982	.00986	.00990
.050	.04957	.04931	.04825	.04828	.04846	.04802	.04760
.100	.09900	.09737	.09575	.09324	.09301	.09154	.09090
.200	.19614	.19004	.18431	.17856	.17410	.17300	.16667
.300	.29563	.27753	.26349	.25542	.25062	.23828	.23080
.400	.39490	.36166	.34078	.32417	.31926	.29995	.28570
.500	.49608	.44732	.41316	.39064	.37987	.35080	.33333
.600	.59122	.52384	.47806	.45321	.43553	.39614	.37500
.700	.69971	.61097	.54536	.51218	.48573	.44286	.41180
.800	.79301	.68752	.60792	.56147	.53483	.48437	.44444
.900	.88983	.76543	.66920	.61807	.57890	.51674	.47370
1.000	1.00255	.84503	.73065	.66694	.62373	.55071	.50000

NOTE: n denotes the degrees of freedom for S_x^2 and S_y^2. ρ denotes σ_x^2/σ_y^2. The entries in the table give $nV(\hat{\mu}_{CS})$ when $\sigma_y^2 = 1$.

TABLE 4. VARIANCE FOR BROWN-COHEN ESTIMATOR

ρ \ n	2	4	6	8	10	20	∞
.001	.00099	.00099	.00098	.00101	.00100	.00099	.00100
.010	.00987	.00993	.00996	.00983	.00981	.00986	.00990
.050	.04908	.04875	.04788	.04801	.04828	.04796	.04760
.100	.09718	.09536	.09437	.09226	.09231	.09135	.09090
.200	.19000	.18370	.18002	.17561	.17203	.17232	.16667
.300	.28377	.26591	.25620	.25065	.24682	.23715	.23080
.400	.37617	.34469	.33064	.31769	.31471	.29864	.28570
.500	.47004	.42392	.40073	.38340	.37592	.34997	.33333
.600	.55748	.49499	.46449	.44620	.43161	.39624	.37500
.700	.65725	.57640	.52961	.50595	.48348	.44374	.41180
.800	.74230	.64735	.59301	.55613	.53397	.48691	.44444
.900	.83070	.71925	.65373	.61535	.58143	.52175	.47370
1.000	.93338	.79368	.71664	.66571	.62943	.55798	.50000

NOTE n denotes the degrees of freedom for S_x^2 and S_y^2. ρ denotes σ_x^2/σ_y^2. The entries in the table give $nV(\hat{\mu}_{BC})$ when $\sigma_y^2 = 1$.

TABLE 5A. VARIANCE FOR ZACK'S ESTIMATOR (α=.01)

ρ \ n	2	4	6	8	10	20	∞
.001	.02335	.00122	.00099	.00101	.00100	.00099	.00100
.010	.12963	.02245	.01115	.01009	.00987	.00987	.00990
.050	.22596	.13704	.08157	.05917	.05277	.04817	.04760
.100	.25613	.21425	.17060	.13573	.11516	.09253	.09090
.200	.29332	.28082	.26958	.24906	.22976	.18615	.16667
.300	.32309	.31874	.31823	.30810	.30398	.26898	.23080
.400	.34974	.34737	.34517	.34460	.34542	.33106	.28570
.500	.37579	.37314	.37666	.37702	.37723	.37089	.33333
.600	.40042	.39718	.40388	.40412	.40327	.39693	.37500
.700	.43022	.42661	.42875	.42960	.42620	.42610	.41180
.800	.45304	.45437	.45884	.45122	.45075	.45761	.44444
.900	.47845	.47578	.47882	.48136	.47540	.47885	.47340
1.000	.51117	.50245	.50462	.50778	.50413	.50582	.50000

NOTE: n denotes the degrees of freedom for S_x^2 and S_y^2. ρ denotes σ_x^2/σ_y^2. The entires in the table give $nV(\hat{\mu}_Z)$ when $\sigma_y^2 = 1$ and $\alpha = .01$.

TABLE 5B. VARIANCE FOR ZACK'S ESTIMATOR (α=.05)

ρ \ n	2	4	6	8	10	20	∞
.001	.00553	.00107	.00099	.00101	.00100	.00099	.00100
.010	.04831	.01281	.01034	.00994	.00987	.00987	.00990
.050	.15463	.08009	.05751	.05093	.05000	.04816	.04760
.100	.21382	.15436	.12315	.10624	.09832	.09207	.09090
.200	.28140	.24921	.23012	.21052	.19522	.17541	.16667
.300	.32688	.30773	.29641	.28360	.27686	.24825	.23080
.400	.36206	.34828	.34100	.33536	.33103	.31343	.28570
.500	.39484	.38272	.38140	.37630	.37323	.36112	.33333
.600	.42527	.41228	.41205	.41042	.40721	.39452	.37500
.700	.46036	.44529	.44122	.44015	.43412	.43001	.41180
.800	.48457	.47561	.47433	.46393	.46220	.46287	.44444
.900	.51215	.49906	.49621	.49628	.48703	.48508	.47370
1.000	.54617	.52645	.52363	.52426	.51733	.51231	.50000

NOTE: n denotes the degrees of freedom for S_x^2 and S_y^2. ρ denotes σ_x^2/σ_y^2. The entries in the table give $nV(\hat{\mu}_Z)$ when $\sigma_y^2 = 1$ and $\alpha = .05$.

TABLE 5C. VARIANCE FOR ZACK'S ESTIMATOR (α=.10)

ρ \ n	2	4	6	8	10	20	∞
.001	.00338	.00102	.00099	.00101	.00100	.00099	.00100
.010	.03021	.01153	.01028	.00994	.00987	.00987	.00990
.050	.11915	.06617	.05341	.05000	.04958	.04816	.04760
.100	.18491	.13146	.11001	.10003	.09575	.09205	.09090
.200	.26894	.22938	.21033	.19491	.18431	.17377	.16667
.300	.32731	.29705	.28232	.27046	.26387	.24259	.23080
.400	.37328	.34701	.33508	.32678	.32274	.30541	.28570
.500	.41261	.38857	.38239	.37438	.36981	.35450	.33333
.600	.44845	.42314	.41820	.41261	.40836	.39186	.37500
.700	.48884	.46155	.45106	.44746	.43805	.43022	.41180
.800	.51735	.49326	.48654	.47367	.46980	.46551	.44444
.900	.54541	.51935	.51200	.50659	.49616	.48912	.47370
1.000	.58088	.54850	.53900	.53579	.52631	.51723	.50000

NOTE: n denotes the degrees of freedom for S_x^2 and S_y^2. ρ denotes σ_x^2/σ_y^2. The entries in the table give $nV(\hat{\mu}_Z)$ when $\sigma_y^2 = 1$ and $\alpha = .10$.

TABLE 5D. VARIANCE FOR ZACK'S ESTIMATOR (α=.25)

ρ \ n	2	4	6	8	10	20	∞
.001	.00210	.00101	.00099	.00101	.00100	.00099	.00100
.010	.02003	.01094	.01025	.000994	.00987	.00987	.00990
.050	.08726	.05732	.05131	.04950	.04946	.04816	.04760
.100	.15115	.11319	.10214	.09674	.09453	.09203	.09090
.200	.24929	.20645	.19302	.18300	.17683	.17306	.16667
.300	.32731	.28191	.26668	.25604	.25189	.23847	.23070
.400	.38787	.34230	.32466	.31542	.31241	.29846	.28570
.500	.44142	.39430	.38105	.36974	.36403	.34811	.33333
.600	.48717	.43839	.42521	.41455	.40821	.38847	.37500
.700	.53785	.48295	.46466	.45630	.44259	.43015	.41180
.800	.57377	.52194	.50427	.48653	.47940	.46812	.44444
.900	.60462	.54997	.53304	.52246	.50823	.49468	.47370
1.000	.64467	.58294	.56113	.55268	.54055	.52464	.50000

NOTE: n denotes the degrees of freedom for S_x^2 and S_y^2. ρ denotes σ_x^2/σ_y^2. The entries in the table give $nV(\hat{\mu}_Z)$ when $\sigma_y^2 = 1$ and $\alpha = .25$.

TABLE 6. EFFICIENCY OF MAXIMUM LIKELIHOOD ESTIMATOR

ρ \ n	2	4	6	8	10	20	∞
.001	.90754	1.02967	1.01510	1.01692	1.01285	1.00133	1.00000
.010	.90753	1.02967	1.01510	1.01692	1.01285	1.00133	1.00000
.050	.90245	1.02302	1.01481	1.01691	1.01283	1.00133	1.00000
.100	.89355	.99974	1.00973	1.01573	1.01178	1.00133	1.00000
.200	.88242	.96615	.99185	1.00645	1.00525	1.00132	1.00000
.300	.87608	.94804	.97721	.99534	.99824	1.00118	1.00000
.400	.87193	.93713	.96672	.98593	.99261	.99981	1.00000
.500	.86895	.92990	.95909	.97842	.98823	.99107	1.00000
.600	.86668	.92477	.95334	.97242	.98480	.99852	1.00000
.700	.86488	.92094	.94888	.96758	.98205	.99691	1.00000
.800	.86341	.91798	.94532	.96360	.97982	.99509	1.00000
.900	.86218	.91562	.94242	.96029	.97798	.99312	1.00000
1.000	.86114	.93170	.94002	.95749	.97643	.99107	1.00000

TABLE 7. EFFICIENCY OF COHEN-SACKROWITZ ESTIMATOR

ρ \ n	2	4	6	8	10	20	∞
.001	1.9798	1.1717	1.0102	1.0000	1.0100	1.0000	1.0000
.010	1.8736	1.0915	1.0281	1.0122	1.0051	1.0010	1.0000
.050	1.6516	1.1387	1.0582	1.0242	1.0206	1.0029	1.0000
.100	1.4526	1.1306	1.0566	1.0321	1.0149	1.0054	1.0000
.200	1.2596	1.0680	1.0241	1.0204	1.0255	.9941	1.0000
.300	1.1101	1.0095	.9976	.9955	.9945	1.0053	1.0000
.400	1.0032	.9556	.9570	.9749	.9707	.9924	1.0000
.500	.9168	.8987	.9217	.9458	.9542	.9922	1.0000
.600	.8566	.8615	.8964	.9182	.9277	.9900	1.0000
.700	.7896	.8099	.8629	.8927	.9240	.9732	1.0000
.800	.7492	.7762	.8355	.8793	.9063	.9609	1.0000
.900	.7105	.7427	.8090	.8515	.8926	.9603	1.0000
1.000	.6650	.7100	.7820	.8330	.8747	.9511	1.0000

TABLE 8. EFFICIENCY OF BROWN-COHEN ESTIMATOR

ρ \ n	2	4	6	8	10	20	∞
.001	1.9800	1.1717	1.0102	1.0000	1.0100	1.0000	1.0000
.010	1.8774	1.0937	1.0301	1.0112	1.0061	1.0010	1.0000
.050	1.6681	1.1518	1.0664	1.0300	1.0244	1.0042	1.0000
.100	1.4798	1.1545	1.0721	1.0430	1.0226	1.0074	1.0000
.200	1.3003	1.1049	1.0486	1.0375	1.0378	.9980	1.0000
.300	1.1565	1.0537	1.0260	1.0145	1.0098	1.0101	1.0000
.400	1.0532	1.0027	.9864	.9948	.9848	.9968	1.0000
.500	.9676	.9483	.9503	.9637	.9642	.9946	1.0000
.600	.9084	.9118	.9226	.9326	.9462	.9897	1.0000
.700	.8406	.8585	.8885	.9037	.9283	.9713	1.0000
.800	.8004	.8243	.8565	.8878	.9077	.9560	1.0000
.900	.7611	.7904	.8281	.8553	.8887	.9510	1.0000
1.0000	.7143	.7560	.7973	.8345	.8667	.9387	1.0000

TABLE 9A. EFFICIENCY OF ZACK'S ESTIMATOR (α=.01)

ρ \ n	2	4	6	8	10	20	∞
.001	.0839	.9508	1.0000	1.0000	1.0100	1.0000	1.0000
.010	.1429	.4837	.9202	.9851	1.0000	1.0000	1.0000
.050	.3623	.4097	.6260	.8357	.9373	.9998	1.0000
.100	.5615	.5138	.5930	.7090	.8197	.9946	1.0000
.200	.8423	.7228	.7002	.7316	.7771	.9239	1.0000
.300	1.0158	.8790	.8260	.8253	.8199	.8906	1.0000
.400	1.1328	.9949	.9449	.9171	.8972	.8991	1.0000
.500	1.2103	1.0774	1.0110	.9800	.9609	.9385	1.0000
.600	1.2647	1.1363	1.0610	1.0297	1.0127	.9880	1.0000
.700	1.2842	1.1600	1.0975	1.0643	1.0531	1.0115	1.0000
.800	1.3114	1.1745	1.1070	1.0942	1.0753	1.0171	1.0000
.900	1.3215	1.1949	1.1306	1.0933	1.0869	1.0362	1.0000
1.000	1.3042	1.1941	1.1323	1.0941	1.0822	1.0355	1.0000

TABLE 9B. EFFICIENCY OF ZACK'S ESTIMATOR (α=.05)

ρ \ n	2	4	6	8	10	20	∞
.001	.3544	1.0841	1.0000	1.0000	1.0100	1.0000	1.0000
.010	.3836	.8478	.9923	1.0000	1.0000	1.0000	1.0000
.050	.5295	.7011	.8878	.9709	.9892	1.0000	1.0000
.100	.6726	.7133	.8215	.9058	.9601	.9996	1.0000
.200	.8779	.8145	.8203	.8655	.9146	.9804	1.0000
.300	1.0040	.9105	.8868	.8966	.9002	.9650	1.0000
.400	1.0942	.9923	.9564	.9424	.9362	.9497	1.0000
.500	1.1520	1.0504	.9984	.9819	.9712	.9639	1.0000
.600	1.1908	1.0947	1.0400	1.0140	1.0029	.9940	1.0000
.700	1.2001	1.1113	1.0665	1.0388	1.0339	1.0023	1.0000
.800	1.2260	1.1220	1.0708	1.0642	1.0487	1.0056	1.0000
.900	1.2345	1.1391	1.0910	1.0605	1.0609	1.0230	1.0000
1.000	1.2206	1.1397	1.0912	1.0597	1.0545	1.0224	1.0000

TABLE 9C. EFFICIENCY OF ZACK'S ESTIMATOR (α=.10)

ρ \ n	2	4	6	8	10	20	∞
.001	.5799	1.1373	1.0000	1.0000	1.0100	1.0000	1.0000
.010	.6134	.9419	.9981	1.0000	1.0000	1.0000	1.0000
.050	.6871	.8486	.9560	.9890	.9976	1.0000	1.0000
.100	.7777	.8374	.9196	.9620	.9860	.9998	1.0000
.200	.9186	.8849	.8974	.9348	.9687	.9897	1.0000
.300	1.0027	.9960	.9311	.9402	.9445	.9875	1.0000
.400	1.0613	.9432	.9733	.9671	.9603	.9747	1.0000
.500	1.1023	.9960	.9958	.9870	.9802	.9819	1.0000
.600	1.1293	1.0346	1.0247	1.0086	1.0001	1.0008	1.0000
.700	1.1302	1.0666	1.0433	1.0219	1.0246	1.0018	1.0000
.800	1.1484	1.0721	1.0440	1.0423	1.0317	.9999	1.0000
.900	1.1592	1.0819	1.0573	1.0389	1.0414	1.0145	1.0000
1.000	1.1477	1.0939	1.0607	1.0369	1.0366	1.0127	1.0000

TABLE 9D. EFFICIENCY OF ZACK'S ESTIMATOR (α=.25)

ρ \ n	2	4	6	8	10	20	∞
.001	.9333	1.1485	1.0000	1.0000	1.0100	1.0000	1.0000
.010	.9251	.9927	1.0010	1.0000	1.0000	1.0000	1.0000
.050	.9382	.9796	.9951	.9990	1.0000	1.0000	1.0000
.100	.9514	.9726	.9905	.9947	.9986	1.0000	1.0000
.200	.9910	.9831	.9779	.9956	1.0097	.9938	1.0000
.300	1.0027	.9939	.9857	.9931	.9894	1.0045	1.0000
.400	1.0214	1.0097	1.0046	1.0020	.9920	.9974	1.0000
.500	1.0304	1.0196	.9993	.9993	.9957	.9999	1.0000
.600	1.0395	1.0295	1.0078	1.0038	1.0005	1.0095	1.0000
.700	1.0272	1.0246	1.0127	1.0021	1.0141	1.0020	1.0000
.800	1.0354	1.0224	1.0073	1.0148	1.0111	.9943	1.0000
.900	1.0467	1.0337	1.0156	1.0073	1.0167	1.0031	1.0000
1.000	1.0341	1.0293	1.0183	1.0052	1.0092	.9984	1.0000

While no single estimator of the common mean dominates any of the others for all values of n and ρ, certain conclusions may be made. Considering Table 6, we see that the efficiency of $\hat{\mu}_{MLE}$ with respect to $\hat{\mu}_{GD}$ increases to 1 as n increases when $\rho = \sigma_x^2/\sigma_y^2$ is fixed and near unity. Also for fixed n, the efficiency decreases as ρ increases toward 1. For moderate sample sizes and small values of ρ the efficiencies are larger than 1. It thus appears that $\hat{\mu}_{GD}$ is preferable to $\hat{\mu}_{MLE}$ when the sample size is very small and when ρ is near 1. $\hat{\mu}_{MLE}$ is more efficient than $\hat{\mu}_{GD}$ when the sample size is moderate and ρ is small. $\hat{\mu}_{GD}$ is slightly more efficient than $\hat{\mu}_{MLE}$ when the sample size is large and ρ is near 1.

Tables 7 and 8 show that the efficiencies of $\hat{\mu}_{CS}$, $\hat{\mu}_{BC}$ are larger than 1 for small values of ρ. The efficiencies decrease as ρ increases for small sample sizes. For fixed ρ, the efficiencies decrease as n increases when ρ is near 1.

Tables 9A to 9D indicate the efficiencies of $\hat{\mu}_Z$ with α = .01, .05, .10, .25 are larger than $\hat{\mu}_{GD}$ for large values of ρ. For fixed ρ near unity the efficiencies decrease toward one as n increases. For small sample sizes and ρ close to 1 the efficiencies increase as ρ increases. Therefore we conclude for small sample sizes and ρ near 1 the estimator $\hat{\mu}_Z$ is preferable to $\hat{\mu}_{GD}$.

REFERENCES

Bement, T. E. and Williams, J. S. (1969). Variance of weighted regression estimators when sampling errors are independent and heteroscedastic. Journal of the American Statistical Association, 64, 1369-1382.

Bennet, B. M. and Allshouse, M. P. (1976). On the use of the log likelihood ratio in combining several means. Biometrics Zeit., 18, 101-104.

Bhattacharya, C. G. (1980). Estimation of a common mean and recovery of interblock information. Annals of Statistics, 8, 205-211.

Brown, L. D. and Cohen, A. (1974). Point and confidence estimation of a common mean and recovery of interblock information. Annals of Statistics, 2, 963-976.

Cohen, A. and Sackrowitz, H. B. (1974). On estimating the common mean of two normal distributions. Annals of Statistics, 2, 1274-1282.

Fuller, W. A. (1976). Introduction to Statistical Time Series. John Wiley and Sons.

Graybill, F. A. and Deal, R. B. (1959). Combining unbiased estimators. Biometrics, 15, 543-550.

Hogg, R. V. (1960). On conditional expectations of location statistics. Journal of the American Statistical Association, 55, 714-717.

Levy, P. S. and Mantal, N. (1974). Combining unbiased estimates -- a further examination of some old estimators. Journal of Statistical Computation and Simulation, 3, 147.

Mehta, J. S. and Gurland, J. (1969). Combinations of unbiased estimators of the mean which consider inequality of unknown variables. Journal of the American Statistical Association, 64, 1042-1055.

Nair, K. A. (1980). Distribution of an estimator of the common mean of two normal populations. Annals of Statistics, 8, 212-216.

Olkin, I. and Pratt, J. (1958). Unbiased estimation of certain correlation coefficients. Annals of Mathematical Statistics, 29, 201-211.

Richter, D. (1960). Two-stage experiments for estimating a common mean. Annals of Mathematical Statistics, 31, 1164-1173.

Sehadri, V. (1963). Constructing uniformly better estimators. Journal of the American Statistical Association, 58, 172-175.

Zacks, S. (1966). Unbiased estimation of the common mean of two normal distributions based on small samples of equal size. Journal of the American Statistical Association, 61, 467-476.

Marsaglia, G., Ananthanarayanan, K. and Paul, N. J. (1976). Improvements on fast methods for generating normal random variables. Information Processing Letters, 5, 2, 27-30.

ACKNOWLEDGMENTS

The author is sincerely grateful to her major professor, Dr. David Hasza, for his suggestions during this study, his patience and valuable guidance in the preparation of this report.

Appreciation is also extended to Dr. S. K. Perng and Dr. W. Yang for being on her graduate committee.

Thanks also to her parents for their encouragement and support.

ESTIMATION OF THE COMMON MEAN OF TWO NORMAL DISTRIBUTIONS

by

Shi-Hwa Yuan

B. A., Fu-Jen Catholic University, 1977

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

1981

ABSTRACT

Several estimators are considered for the problem of estimation of the common mean of two independent normal distributions when the variances are unknown. The maximum likelihood estimator of the common mean is derived. For the equal sample size case, the maximum likelihood estimator together with estimators previously proposed by various authors are compared in a Monte Carlo study.