ESTIMATING ONE MEAN OF A BIVARIATE NORMAL DISTRIBUTION USING A PRELIMINARY TEST OF SIGNIFICANCE AND A TWO STAGE SAMPLING SCHEME

By

D. R. Brogan and J. Sedransk

Department of Biostatistics, U. N. C
and
Department of Statistics, University of Wisconsin

Institute of Statistics Mimeo Series No. 752

June 1971

Estimating One Mean of a Bivariate Normal Distribution Using a Preliminary Test of Significance and a Two Stage Sampling Scheme

by D. R. Brogan and J. Sedransk

ABSTRACT

Let (X,Y) have a bivariate normal distribution with unknown mean vector (μ_X,μ_Y) and known covariance matrix $\sum$. It is desired to estimate μ_Y. A first stage sample is obtained on (X,Y), X only, and Y only. A preliminary test of H_0: $\mu_Y=\mu_X$ is performed, the result of which specifies a second-stage sample. The estimator of μ_Y is either a regression estimator or a pooled estimator which pools estimators of μ_X and μ_Y. The bias and mean square error of this estimation procedure are derived, and a numerical example is discussed.

Estimating One Mean of a Bivariate Normal Distribution Using a Preliminary Test of Significance and a Two Stage Sampling Scheme

by D. R. Brogan and J. Sedransk*

1. Statement of the Problem and Examples

Let the random variable $(X,Y)'$ follow the bivariate normal distribution with mean vector $E(X,Y)' = (\mu_X,\mu_Y)'$ and known covariance matrix $\sum$, where

$$\sum = \text{Cov}\begin{pmatrix}X\\Y\end{pmatrix} = \begin{pmatrix}\sigma_X^2 & \rho\sigma_X\sigma_Y\\ \rho\sigma_X\sigma_Y & \sigma_Y^2\end{pmatrix} . \tag{1.1}$$

It is desired to estimate μ_Y when there is evidence that perhaps $\mu_Y=\mu_X$. There are available $n\geq 0$ bivariate observations on (X,Y), an additional $n_X\geq 0$ independent observations on X, and an additional $n_Y\geq 0$ independent observations on Y. Using these observations, the null hypothesis H_0: $\mu_Y=\mu_X$ versus the alternative hypothesis H_1: $\mu_Y\neq\mu_X$ is tested. After this preliminary test, a second stage of sampling is carried out on one, two or three of the following random variables: (X,Y) only, X only, and Y only. The estimator of μ_Y, using data from the first and second stages of sampling, depends upon the acceptance or rejection of H_0. This general estimator reduces in special cases to estimators proposed by other authors who have considered the "preliminary test" approach to

---

*D. R. Brogan is Assistant Professor, Department of Biostatistics, University of North Carolina School of Public Health. J. Sedransk is Associate Professor, Department of Statistics, University of Wisconsin. This research was partially supported by National Institutes of Health Biometry Training Grant 5T1GM34, National Institute of Mental Health Training Grant MH10373, and National Center for Educational Statistics (U.S. Office of Education) Contract Number OEC-3-002041-2041.

estimating μ_Y. In this paper, the bias and mean square error of the proposed estimator are derived.

A situation where such an estimation scheme would be appropriate is the estimation of the average systolic blood pressure (SBP) of a hospital population. SBP is available on each admitted patient's hospital record. However, it is well known that SBP varies within each person depending upon time of day, general level of excitement, and so on. Hence, theoretically, measurements of SBP should be taken on a random sample of admitted patients under some set of standard conditions. Then, one could easily estimate μ_Y, where Y is SBP measured under the set of standard conditions. However, since measurements on Y are difficult (and expensive) to obtain, it would be hoped that measurements on X, SBP measured under non-standard conditions, for a relatively large sample of patients might be used in conjunction with measurements on Y for a relatively small sample of patients in order to estimate μ_Y. In this example, it may be that $|\mu_Y-\mu_X|$ is small; almost certainly, though, σ_X^2 would be larger than σ_Y^2.

A random sample on X at the first stage could be collected from past hospital records where, it is assumed, SBP was measured under non-standard conditions. An independent bivariate random sample on (X,Y) could be obtained by measuring SBP under standard conditions on a sample of patients, as well as by using the usual SBP from the hospital records of the same patients. A further, independent, random sample on Y at the first stage (though not necessary) could possibly come from a research project done on some of the hospital population where SBP was purposely taken under standard conditions and the SBP under non-standard conditions is not available. Of course, it is necessary to take caution that these three samples do, indeed, come from the same population. Further sampling can be done at the second stage under several options (see

Section 3).

Letting C_X, C_Y, and C_{XY} be the respective per unit costs of measuring SBP under non-standard, standard, and both conditions, it is obvious in this example that $C_X<C_Y\leq C_{XY}\leq C_X+C_Y$. Hence, observations on X may be "preferred to" observations on Y if X can be used effectively to estimate μ_Y. This type of cost configuration is typical of many situations in which one might wish to apply the preliminary test approach discussed in this paper. That is, one wishes to estimate μ_Y, but an observation on X is less costly than an observation on Y and there is some prior evidence that $\mu_Y=\mu_X$.

Another example is that of estimating the average volume of trees in a forest. For any given tree it is possible to measure Y, the actual volume. This, however, is difficult and expensive. A much cheaper method which may provide a good estimate of the volume is to measure the height H of the tree and the diameter D at a specified height off the ground. Then the volume is estimated by $X = kD^2H$. If $|\mu_Y-\mu_X|$ is small, then measurements of X could be pooled with measurements of Y in order to estimate μ_Y.

Still another example is the post-enumeration surveys which the Census Bureau conducts to check on the adequacy of coverage and content in the decennial census of population and housing [Bailar, 8]. For each of several geographic areas, there could be available the population count or the attribute data from the post-enumeration survey (i.e. Y) and from the original census (i.e. X). These observations can possibly be combined in order to estimate the total population or the average value of some attribute for these areas. It would be hoped that the combined estimate would be more accurate than the census figures alone.

These three examples illustrate the following considerations: (1) a pooled estimator seems appropriate if $|\mu_Y-\mu_X|$ is small; (2) observations on X instead of Y are desirable from a cost viewpoint; and (3) a correlation between X and Y may allow utilization of information on X even though $|\mu_Y-\mu_X|$ is not small.

2. Previous Investigations of Related Problems

Some of the authors who have studied parametric estimation problems by using pooling procedures after a preliminary test of significance are Mosteller [17], Bennett [10, 11], Kitagawa [15], Bancroft [9], Asano [3, 4, 5], Kale and Bancroft [14], Han and Bancroft [12], Asano and Sugimura [7], and Huntsberger [13]. Asano and Sato [6] and Sato [19] have considered two bivariate populations and two multivariate populations, respectively. Tamura [20, 21] has considered non-parametric estimation after a preliminary test of significance.

This study differs from these other investigations in three respects. First, there is no published research where a two-stage sampling procedure has been considered in conjunction with a preliminary test of significance. The two-stage estimation procedures discussed, for example, by Yen [22] and Arnold and Al-Bayyati [2], use information from the first stage to determine the sampling plan at the second stage, but a preliminary test of significance is not used to make this determination. The two-stage sampling scheme is useful because, for a given budget (assuming $C_X<C_Y\leq C_{XY}$), it may be advantageous to do some additional sampling after the preliminary test is done. For example, if H_0: $\mu_Y=\mu_X$ is accepted, then a large sample where only X is measured is reasonable for the second stage. However if H_0: $\mu_Y=\mu_X$ is rejected, then a small sample where only Y is measured is probably more feasible for the second stage. In addition, the method proposed here allows a bivariate sample on (X,Y) at the second stage.

Second, there are no published results where both independent and dependent sampling can be included at each stage. Thus, given a budget and cost function, in the procedure proposed here one can, at least theoretically, determine the optimal allocation of resources (1) between sampling at the first and second stages, and (2) among bivariate and univariate (both X and Y) sampling at each stage.

Third, every investigator except Kitagawa [15] and Mehta and Gurland [16] has considered the random variables X and Y to be independent. Mehta and Gurland [16], however, are concerned with testing hypotheses about a bivariate normal population, one of which is the null hypothesis that $\rho=0$. Kitagawa [15], on the other hand, considers $\rho\neq 0$, where ρ is unknown, although his investigation is specialized by having only a one-stage bivariate sample of size n with $\sigma_X^2=\sigma_Y^2=\sigma^2$.

In many prospective applications of the preliminary test approach (such as the SBP example) it is to be expected that $\sigma_X^2\neq\sigma_Y^2$. Further, it is important to extend the results available for "pooling means" to include the numerous applications where ρ is not necessarily zero, and where random samples of sizes $n\geq 0$, $n_X\geq 0$, and $n_Y\geq 0$ as described in Section 1 are selected. Thus, to avoid having to use approximate distribution theory, $\sum$ has been assumed to be known in this investigation. A small Monte Carlo study has been carried out to determine the effect on the bias and mean square error of the estimator of μ_Y from estimating the components of $\sum$ [Ruhl, 18].

Finally, it may also be noted that (1) those authors considering $\sigma_X^2\neq\sigma_Y^2$ [e.g. 10] assume that σ_X^2 and σ_Y^2 are known; and (2) some of the authors [14, 17] who take $\sigma_X^2=\sigma_Y^2=\sigma^2$ assume, for simplicity, that σ^2 is known.

Special cases of the sampling procedure considered in this investigation are the same as those studied by many of the above authors except for those investigations where $\sigma_X^2=\sigma_Y^2=\sigma^2$ and/or ρ are assumed to be unknown.

3. The Sampling Procedure, Some Notation, and Some Special Cases

A random sample $(X_1,Y_1),\ldots,(X_n,Y_n)$ is selected from the bivariate normal distribution. In addition, an independent random sample of n_X observations is taken on X, and a random sample of n_Y observations is taken on Y. In the first stage sample, thus, are $(n+n_X)$ observations on X and $(n+n_Y)$ observations on Y. Note that only X_i and Y_i are correlated, $(i=1,\ldots,n)$, where (X_i,Y_i) denotes the i-th element in the bivariate sample.

At this point a preliminary test of the null hypothesis H_0: $\mu_Y=\mu_X$ versus the alternative hypothesis H_1: $\mu_Y\neq\mu_X$ is done using the sample data from the first stage. On the basis of the preliminary test H_0 is either accepted or rejected. A second stage sample is then taken, again allowing a bivariate sample on (X,Y) and two independent samples, one each on X and Y. If H_0 is accepted, the size of the bivariate sample will be n_0, and the size of the independent samples on X and Y will be n_{0X} and n_{0Y}, respectively. Similarly, if H_0 is rejected, the sizes of the samples at the second stage will be n_1, n_{1X}, and n_{1Y}.

Thus, the notation for this sampling procedure can be summarized as follows. Let $\overline{X}_n$ and $\overline{Y}_n$ be the sample means on X and Y from the bivariate sample at the first stage. Likewise, let $\overline{X}_{n_X}$ and $\overline{Y}_{n_Y}$ be the sample means from the two independent samples at the first stage. Analogously, if H_0 is accepted, let the respective sample means be denoted by $\overline{X}_{n_0}$, $\overline{Y}_{n_0}$, $\overline{X}_{n_{0X}}$, and $\overline{Y}_{n_{0Y}}$. If H_0 is rejected, the sample means will be $\overline{X}_{n_1}$, $\overline{Y}_{n_1}$, $\overline{X}_{n_{1X}}$, and $\overline{Y}_{n_{1Y}}$. The notation and sampling scheme are illustrated in Figure 1.

Figure 1

Two Stage Sampling Procedure
and Resultant Estimator of μ_Y

FIRST STAGE SAMPLE MEANS

$(\overline{X}_n, \overline{Y}_n)$

$\overline{X}_{n_X}$ $\overline{Y}_{n_Y}$

PRELIMINARY TEST

$H_0: \mu_Y = \mu_X$ $H_1: \mu_Y \neq \mu_X$

SECOND STAGE SAMPLE MEANS

$(\overline{X}_{n_0}, \overline{Y}_{n_0})$ $(\overline{X}_{n_1}, \overline{Y}_{n_1})$

$\overline{X}_{n_{0X}}$ $\overline{Y}_{n_{0Y}}$ $\overline{X}_{n_{1X}}$ $\overline{Y}_{n_{1Y}}$

ESTIMATOR OF μ_Y ESTIMATOR OF μ_Y

$c_1\overline{X}_{n+n_0} + c_2\overline{Y}_{n+n_0} + c_3\overline{X}_{n_X+n_{0X}} + c_4\overline{Y}_{n_Y+n_{0Y}}$ $w_1\overline{Y}_{n_Y+n_{1Y}} + w_2[\overline{Y}_{n+n_1} + \beta(\overline{X}_{N_X} - \overline{X}_{n+n_1})]$

Note that only one of the two possible second stage samples is realized. However, in determining the bias and mean square error of the resultant estimator, the values of the sample sizes of both second stage possibilities must be considered. Thus, in using such an approach, values of n, n_X, n_Y, n_0, n_{0X}, n_{0Y}, n_1, n_{1X}, and n_{1Y} would be fixed in advance of sampling, but only one of the two sets (n_1, n_{1X}, n_{1Y}) and (n_0, n_{0X}, n_{0Y}) would be realized.

This sampling procedure includes several possibilities. By taking $n=\rho=0$, one has the one-stage sampling scheme considered by several authors [10, 14, 17]. For applications of the type illustrated by the SBP example, one would typically have $n>0$, $n_X>0$, and $n_Y=0$. At the second stage of sampling one might take $n_0=n_{0Y}=0$ and $n_{0X}>0$ if H_0 is accepted, whereas one might choose $n_1=n_{1X}=0$ and $n_{1Y}>0$ if H_0 is rejected. However, the sampling procedure is completely general in that it includes the possibility of both dependent and independent sampling at each stage while at the second stage a different procedure may be followed depending on whether H_0 is accepted or rejected.

4. The Preliminary Test Statistic

Before defining the preliminary test statistic, some additional notation is introduced. First, define $N_X = n+n_1+n_X+n_{1X}$. Then sample means such as $\overline{X}_{n+n_0}$ and $\overline{X}_{N_X}$ are defined as

$$\overline{X}_{n+n_0} = (n+n_0)^{-1}(n\overline{X}_n+n_0\overline{X}_{n_0}) \tag{4.1}$$

and

$$\overline{X}_{N_X} = N_X^{-1}(n\overline{X}_n+n_1\overline{X}_{n_1}+n_X\overline{X}_{n_X}+n_{1X}\overline{X}_{n_{1X}}). \tag{4.2}$$

Given the sampling scheme discussed in section 3, a preliminary test of H_0: $\mu_Y=\mu_X$ versus H_1: $\mu_Y\neq\mu_X$ is made using the test statistic

$$Z = \bar{Y}_{n+n_Y} - \bar{X}_{n+n_X}. \tag{4.3}$$

Under H_0, Z is normally distributed with a mean of zero and variance σ_Z^2, where

$$\sigma_Z^2 = \frac{\sigma_Y^2}{n+n_Y} + \frac{\sigma_X^2}{n+n_X} - \frac{2n\rho\sigma_X\sigma_Y}{(n+n_X)(n+n_Y)} . \tag{4.4}$$

If the correlation ρ is near one, then better power on the preliminary test might be obtained by using the test statistics Z' instead of Z, where $Z' = \bar{Y}_n - \bar{X}_n$ and has variance $\sigma_{Z'}^2 = n^{-1}(\sigma_Y^2+\sigma_X^2-2\rho\sigma_X\sigma_Y)$. This is because $\sigma_Z^2 < \sigma_{Z'}^2$ if, and only if,

$$\rho < \frac{\sigma_Y^2 n_Y(n+n_X)+\sigma_X^2 n_X(n+n_Y)}{2\sigma_X\sigma_Y(nn_X+nn_Y+n_X n_Y)} . \tag{4.5}$$

Obviously, inequality (4.5) is satisfied if $\rho\leq 0$. In general, inequality (4.5) is satisfied unless ρ is close to one. Since, for most applications, the correlation will be zero or moderately positive, Z as defined in (4.3) will be used as the preliminary test statistic in the estimation of μ_Y.

Let ξ_α be the critical value with Type I error equal to α for the test of H_0: $\mu_Y=\mu_X$ versus H_1: $\mu_Y\neq\mu_X$ using the N(0,1) distribution with probability density function $\phi(t)$ and cumulative distribution function $\Phi(t)$. That is, $\Phi(\xi_\alpha) - \Phi(-\xi_\alpha) = 1-\alpha$. Hence, H_0 will be rejected whenever $|Z|>\xi_\alpha\sigma_Z$.

5. The General Estimator $\hat{\mu}_Y$ of μ_Y

Let case 1 be defined by the following conditions: $\mu_Y\neq\mu_X$, $n\geq 0$, $n_X\geq 0$,

$n_Y \geq 0$, $n_1 \geq 0$, $n_{1X} \geq 0$, $n_{1Y} \geq 0$, $n_0 = n_{0X} = n_{0Y} = 0$. Assuming these sample sizes to be predetermined and not dependent upon a preliminary test of significance, the maximum likelihood estimator of μ_Y under these conditions is

$$\hat{\mu}_1 = (g_1+g_2)^{-1}[g_1\bar{Y}_{n_Y+n_{1Y}} + g_2\{\bar{Y}_{n+n_1} + \frac{\rho\sigma_Y}{\sigma_X}(\bar{X}_{N_X} - \bar{X}_{n+n_1})\}] \tag{5.1}$$

where

$$g_1 = \frac{(n_Y+n_{1Y})}{\sigma_Y^2} \; ; \; g_2 = \frac{(n+n_1)}{\sigma_Y^2[1-\rho^2 k]} \tag{5.2}$$

and

$$k = N_X^{-1}(n_X+n_{1X}) . \tag{5.3}$$

Note that $\hat{\mu}_1$ is unbiased and is a weighted average of the mean $\bar{Y}_{n_Y+n_{1Y}}$ and a regression estimator of μ_Y. The regression estimator adjusts Y_{n+n_1} on the basis of the difference between $\bar{X}_{N_X}$ and $\bar{X}_{n+n_1}$. Note, also, that the variances of $\bar{Y}_{n_Y+n_{1Y}}$ and the regression estimator are g_1^{-1} and g_2^{-1}, respectively. Hence, $\hat{\mu}_1$ is a weighted average of two unbiased, statistically independent estimators of μ_Y, with each estimator weighted inversely proportional to its variance.

Let case 0 be defined by the conditions $\mu_Y = \mu_X$, $n \geq 0$, $n_X \geq 0$, $n_Y \geq 0$, $n_0 \geq 0$, $n_{0X} \geq 0$, $n_{0Y} \geq 0$, $n_1 = n_{1X} = n_{1Y} = 0$. Assuming, again, these sample sizes to be predetermined and not dependent upon a preliminary test of significance, the maximum likelihood estimator of μ_Y is

$$\hat{\mu}_0 = [\sum_{i=1}^{4} h_i]^{-1}[h_1\bar{X}_{n+n_0} + h_2\bar{Y}_{n+n_0} + h_3\bar{X}_{n_X+n_{0X}} + h_4\bar{Y}_{n_Y+n_{0Y}}] \tag{5.4}$$

where

$$h_1 = \frac{(n+n_0)}{(1-\rho^2)}\left[\frac{1}{\sigma_X^2} - \frac{\rho}{\sigma_X\sigma_Y}\right]; \quad h_3 = (n_X+n_{0X})/\sigma_X^2$$

$$h_2 = \frac{(n+n_0)}{(1-\rho^2)}\left[\frac{1}{\sigma_Y^2} - \frac{\rho}{\sigma_X\sigma_Y}\right]; \quad h_4 = (n_Y+n_{0Y})/\sigma_Y^2. \qquad (5.5)$$

Note that $\hat{\mu}_0$ is a weighted average of four unbiased estimators of μ_Y and hence unbiased. The weights h_i, $i=1,\ldots,4$, as defined in (5.5), minimize the variance of $\hat{\mu}_0$.

These two maximum likelihood estimators suggest the definition of the estimator of μ_Y under the possibilities of accepting or rejecting H_0. If H_0 is accepted, then the estimator $\hat{\mu}_Y$ of μ_Y is defined as $\hat{\mu}_{Y0}$ where

$$\hat{\mu}_{Y0} = c_1\overline{X}_{n+n_0}+c_2\overline{Y}_{n+n_0}+c_3\overline{X}_{n_X+n_{0X}}+c_4\overline{Y}_{n_Y+n_{0Y}}; \quad \sum_{i=1}^{4} c_i=1. \qquad (5.6)$$

If H_0 is rejected, then the estimator $\hat{\mu}_Y$ of μ_Y is defined as $\hat{\mu}_{Y1}$ where

$$\hat{\mu}_{Y1} = w_1\overline{Y}_{n_Y+n_{1Y}}+w_2[\overline{Y}_{n+n_1}+\beta(\overline{X}_{N_X}-\overline{X}_{n+n_1})]; \quad w_1+w_2=1. \qquad (5.7)$$

For the derivation of the bias and mean square error of $\hat{\mu}_Y$, w_i, $i=1,2$, and c_i, $i=1,2,3,4$, are assumed to be arbitrary known constants such that (5.6) and (5.7) are satisfied. Also, the regression coefficient β in (5.7) is assumed to be an arbitrary known constant.

Reasonable choices for w_i and c_i would be $w_i = g_i/(g_1+g_2)$ and $c_i = h_i/[\sum_{i=1}^{4} h_i]$. Likewise, a reasonable choice for β would be $\rho\sigma_Y/\sigma_X$ as suggested by (5.1). However, these choices will not necessarily minimize the variance or mean square error of $\hat{\mu}_Y$ even though they do minimize the variance of the maximum likelihood estimators from which $\hat{\mu}_Y$ was defined. Hence, for generality, the bias and mean square error are derived for w_i, c_i, and β being known constants.

6. Bias of $\hat{\mu}_Y$

The expectation of $\hat{\mu}_Y$ is defined as

$$E(\hat{\mu}_Y) = E[\hat{\mu}_{Y1} \Big| |Z|>\xi_\alpha\sigma_Z]\Pr[|Z|>\xi_\alpha\sigma_Z] + E[\hat{\mu}_{Y0} \Big| |Z|\leq\xi_\alpha\sigma_Z]\Pr[|Z|\leq\xi_\alpha\sigma_Z] \qquad (6.1)$$

where all terms have been previously defined in (4.3), (4.4), (5.6), and (5.7). This can be written as

$$E(\hat{\mu}_Y) = \int_{|z|>\xi_\alpha\sigma_Z} E(\hat{\mu}_{Y1}|z)h(z)dz + \int_{|z|\leq\xi_\alpha\sigma_Z} E(\hat{\mu}_{Y0}|z)h(z)dz \qquad (6.2)$$

where h(z) is the density function of Z, and Z is normally distributed with mean $\Delta = \mu_Y - \mu_X$ and variance σ_Z^2 as defined in (4.4).

If now A and Z are bivariate normal random variables, where A has unconditional expectation μ_A, Anderson [1] shows that

$$E(A|Z=z) = \mu_A + \sigma_Z^{-2}(z-\Delta)\mathrm{Cov}(A,Z). \qquad (6.3)$$

Using (6.3) to obtain the conditional expectations in (6.2), $E(\hat{\mu}_Y)$ is obtained as

$$E(\hat{\mu}_Y) = \mu_Y + \int_{-\xi_\alpha-\delta}^{\xi_\alpha-\delta} [-\Delta(c_1+c_3)+(H_1-H_2)t/\sigma_Z]\phi(t)dt \qquad (6.4)$$

where

$$H_1 = c_1\mathrm{Cov}(\overline{X}_{n+n_0},Z) + c_2\mathrm{Cov}(\overline{Y}_{n+n_0},Z) + c_3\mathrm{Cov}(\overline{X}_{n_X+n_{0X}},Z) + c_4\mathrm{Cov}(\overline{Y}_{n_Y+n_{0Y}},Z) \quad (6.5)$$

and

$$H_2 = w_1\mathrm{Cov}(\overline{Y}_{n_Y+n_{1Y}},Z) + w_2\mathrm{Cov}(\overline{Y}_{n+n_1},Z) + w_2\beta k[\mathrm{Cov}(\overline{X}_{n_X+n_{1X}},Z) - \mathrm{Cov}(\overline{X}_{n+n_1},Z)] \quad (6.6)$$

and δ, a standardized measure of the difference between μ_Y and μ_X, is defined as

$$\delta = \Delta/\sigma_Z = (\mu_Y-\mu_X)/\sigma_Z. \quad (6.7)$$

The covariances which appear in H_1 and H_2 are easily derived as

$$\mathrm{Cov}(\overline{X}_{n+n_0},Z) = (n+n_0)^{-1}n[-\sigma_X^2/(n+n_X) + \rho\sigma_X\sigma_Y/(n+n_Y)] \quad (6.8)$$

$$\mathrm{Cov}(\overline{Y}_{n+n_0},Z) = (n+n_0)^{-1}n[\sigma_Y^2/(n+n_Y) - \rho\sigma_X\sigma_Y/(n+n_X)] \quad (6.9)$$

$$\mathrm{Cov}(\overline{X}_{n_X+n_{0X}},Z) = -(n_X+n_{0X})^{-1}n_X[\sigma_X^2/(n+n_X)] \quad (6.10)$$

$$\mathrm{Cov}(\overline{Y}_{n_Y+n_{0Y}},Z) = (n_Y+n_{0Y})^{-1}n_Y[\sigma_Y^2/(n+n_Y)] \quad (6.11)$$

$$\mathrm{Cov}(\overline{Y}_{n_Y+n_{1Y}},Z) = (n_Y+n_{1Y})^{-1}n_Y[\sigma_Y^2/(n+n_Y)] \tag{6.12}$$

$$\mathrm{Cov}(\overline{Y}_{n+n_1},Z) = (n+n_1)^{-1}n[\sigma_Y^2/(n+n_Y) - \rho\sigma_X\sigma_Y/(n+n_X)] \tag{6.13}$$

$$\mathrm{Cov}(\overline{X}_{n+n_1},Z) = (n+n_1)^{-1}n[-\sigma_X^2/(n+n_X) + \rho\sigma_X\sigma_Y/(n+n_Y)] \tag{6.14}$$

$$\mathrm{Cov}(\overline{X}_{n_X+n_{1X}},Z) = -(n_X+n_{1X})^{-1}n_X[\sigma_X^2/(n+n_X)] \tag{6.15}$$

The bias is immediately obtained from (6.4) and can be considered as a function of δ, since all other terms in (6.4) are known constants. Denoting the bias as $B(\delta)$, then

$$B(\delta) = -\delta\sigma_Z(c_1+c_3)[\Phi(\xi_\alpha-\delta)-\Phi(-\xi_\alpha-\delta)] + (H_1-H_2)[-\phi(\xi_\alpha-\delta)+\phi(-\xi_\alpha-\delta)]/\sigma_Z \tag{6.16}$$

where $\Phi(t)$ and $\phi(t)$ are defined in section 4.

Even in this very general form, it is possible to demonstrate some properties of $B(\delta)$. First, the bias is zero if $\delta=0$. Also, it is necessary only to consider the behavior of $B(\delta)$ for $\delta\geq 0$ since $B(-\delta) = -B(\delta)$. Furthermore, it can be shown by using l'Hospital's Rule that $\lim_{\delta\to\infty} B(\delta) = 0$. The expression for $B(\delta)$ will simplify for some choices of the weights w_i and c_i, if some of the nine possible sample sizes are taken to be zero, or if $\rho=0$.

7. Mean Square Error of $\hat{\mu}_Y$

The mean square error of $\hat{\mu}_Y$ is derived by first finding $E(\hat{\mu}_Y^2)$ and then using $MSE(\hat{\mu}_Y) = E(\hat{\mu}_Y^2)-2\mu_Y E(\hat{\mu}_Y)+\mu_Y^2$. Starting the derivation similarly to that of $E(\hat{\mu}_Y)$ yields

$$E(\hat{\mu}_Y^2) = \int_{|z|>\xi_\alpha\sigma_Z} E(\hat{\mu}_{Y1}^2|z)h(z)dz + \int_{|z|<\xi_\alpha\sigma_Z} E(\hat{\mu}_{Y0}^2|z)h(z)dz. \quad (7.1)$$

If A and Z follow a bivariate normal distribution with unconditional means μ_A and Δ, respectively, then it follows from standard multivariate normal theory that

$$E(A^2|Z=z) = \mu_A^2 + 2\mu_A \text{Cov}(A,Z)(z-\Delta)/\sigma_Z^2 + \text{Var}(A) + \frac{\text{Cov}^2(A,Z)}{\sigma_Z^2}\left[\frac{(z-\Delta)^2}{\sigma_Z^2} - 1\right]. \quad (7.2)$$

Letting $A=\hat{\mu}_{Y0}$ and then $A=\hat{\mu}_{Y1}$, and using (7.1) and (7.2), gives, after substantial algebra, the mean square error as a function of δ, i.e.

$$MSE(\delta) = V_1 + \int_{-\xi_\alpha-\delta}^{\xi_\alpha-\delta} [V_0-V_1+\delta^2\sigma_Z^2(c_1+c_3)^2-2\delta(c_1+c_3)tH_1+(t^2-1)(H_1^2-H_2^2)/\sigma_2^2]\phi(t)dt, \quad (7.3)$$

where H_1 and H_2 are defined in (6.5) and (6.6),

$$V_1 = \frac{w_1^2\sigma_Y^2}{(n_Y+n_{1Y})} + \frac{w_2^2\sigma_Y^2}{(n+n_1)} + w_2^2\beta^2k^2\left[\frac{\sigma_X^2}{(n_X+n_{1X})} + \frac{\sigma_X^2}{(n+n_1)}\right] - \frac{2w_2^2\beta k\rho\sigma_X\sigma_Y}{(n+n_1)}, \quad (7.4)$$

and

$$V_0 = \frac{c_1^2\sigma_X^2}{(n+n_0)} + \frac{c_2^2\sigma_Y^2}{(n+n_0)} + \frac{c_3^2\sigma_X^2}{(n_X+n_{0X})} + \frac{c_4^2\sigma_Y^2}{(n_Y+n_{0Y})} + \frac{2c_1c_2\rho\sigma_X\sigma_Y}{(n+n_0)}. \quad (7.5)$$

Note that V_0 and V_1 are the unconditional variances of $\hat{\mu}_{Y0}$ and $\hat{\mu}_{Y1}$, respectively.

A few properties of MSE(δ) can be ascertained in this general form. First, mean square error is a symmetric function of δ, i.e. MSE($-\delta$) = MSE(δ). Hence, it is necessary only to investigate the behavior of MSE(δ) for $\delta \geq 0$. Second, it can be shown that $\lim_{\delta \to \infty}$ MSE(δ) = V_1, the unconditional variance of $\hat{\mu}_{Y1}$. Third, MSE(δ) is equal to the unconditional variance of $\hat{\mu}_{Y1}$, i.e. V_1, plus another term which can either be positive or negative.

The expression for MSE(δ) in (7.3) can be integrated and written as in (6.16) as a function of $\phi(x)$ and $\Phi(x)$, where x takes the value $(\xi_\alpha - \delta)$ or $(-\xi_\alpha - \delta)$. One of us has written a computer program to evaluate the bias and mean square error as given in equations (7.3) and (6.16). What is typically done in studies of preliminary test procedures is to evaluate the bias and mean square error for various values of α and Δ or δ, for some given sample sizes and variance-covariance matrix. This paper presents the additional problem, however, of determining values for the weights w_i and c_i and the regression coefficient β.

8. Choice of the Weights w_i and c_i and the Regression Coefficient β

It is theoretically possible to choose the w_i, c_i, and β so that the mean square error as given in (7.3) is minimized. If this is pursued, however, three things become evident. First, the solution for w_1, w_2, and β is independent of the solution for the c_i, which simplifies matters considerably. Secondly, however, the solution for w_2 and β involves two simultaneous equations with terms of order three such as $\beta^2 w_2$, βw_2^2, etc. Third, the solutions for w_i, c_i, and β will all be functions of δ, which, of course, is unknown. Hence, it does not appear feasible to choose the weights w_i, c_i, and β so that mean square error as given in (7.3) is minimized with respect to these parameters.

A logical choice for the regression coefficient β is $\beta = \rho\sigma_Y/\sigma_X$ as suggested by the maximum likelihood estimator in (5.1). If this is done, V_1 of equation (7.4) simplifies to

$$V_1 = \frac{w_1^2\sigma_Y^2}{(n_Y+n_{1Y})} + \frac{w_2^2\sigma_Y^2(1-\rho^2k)}{(n+n_1)}; \ \beta = \rho\sigma_Y/\sigma_X, \tag{8.1}$$

where k is defined in (5.3). Also, H_2 in equation (6.6) reduces to

$$H_2 = \frac{w_2 n\sigma_Y^2(1-\rho^2k)}{(n+n_1)(n+n_Y)} + \frac{(1-w_2)\sigma_Y^2 n_Y}{(n_Y+n_{1Y})(n+n_Y)} - \frac{w_2\rho\sigma_X\sigma_Y}{N_X}; \ \beta = \frac{\rho\sigma_Y}{\sigma_X}. \tag{8.2}$$

Now, with β defined as $\beta = \rho\sigma_Y/\sigma_X$ and with V_1 and H_2 defined as in (8.1) and (8.2), MSE(δ) in (7.3) can be minimized with respect to the w_i and c_i. This will produce one linear equation in w_2. Using the method of LaGrange multipliers to find the c_i which minimize MSE(δ) subject to the restriction $\sum_{i=1}^{4} c_i=1$ leads to five simultaneous linear equations in five unknowns, i.e. c_1, c_2, c_3, c_4, and λ, where λ is the LaGrange multiplier. These equations for w_i and c_i can be solved in the usual manner. However, the solutions for both the w_i and c_i are still a function of δ, which is unknown. It is possible, of course, to estimate δ by $\hat{\delta}$ from the sample data and, hence, have the c_i and w_i be functions of $\hat{\delta}$. However, the formulas given in this paper for bias and mean square error of $\hat{\mu}_Y$ would no longer be appropriate.

As a solution to this dilemma, the logical choices for the w_i and c_i are $w_i = g_i(g_1+g_2)^{-1}$, i=1,2, and $c_i = h_i[\sum_{i=1}^{4} h_i]^{-1}$, i=1,...,4, where g_i and h_i are defined in (5.2) and (5.5). Recall that these weights minimize the unconditional variances of $\hat{\mu}_{Y1}$ and $\hat{\mu}_{Y0}$, respectively. Using these weights yields $H_1 = 0$, and

V_0 reduces to $V_0^* = [\sum_{i=1}^{4} h_i]^{-1}$. If, in addition, $\beta = \rho\sigma_Y/\sigma_X$, then V_1 reduces to $V_1^* = (g_1+g_2)^{-1}$ and H_2 reduces to

$$H_2^* = \frac{\sigma_Y^2[1 - \frac{\rho\sigma_X(1-k)}{\sigma_Y(1-k\rho^2)}]}{[(n_Y+n_{1Y})+(n+n_1)/(1-k\rho^2)]} . \qquad (8.3)$$

Using these values for β and for the w_i and c_i yields the bias and mean square error as

$$B^*(\delta) = \int_{-\xi_\alpha-\delta}^{\xi_\alpha-\delta} [-\delta\sigma_Z(h_1+h_3)(\sum_{i=1}^{4} h_i)^{-1}-H_2^* t/\sigma_Z]\phi(t)dt \qquad (8.4)$$

and

$$MSE^*(\delta) = (g_1+g_2)^{-1} + \int_{-\xi_\alpha-\delta}^{\xi_\alpha-\delta} \{[\sum_{i=1}^{4} h_i]^{-1}-(g_1+g_2)^{-1} + \delta^2\sigma_Z^2(h_1+h_3)^2[\sum_{i=1}^{4} h_i]^{-2}$$

$$-(H_2^*)^2(t^2-1)/\sigma_Z^2\} \phi(t)dt. \qquad (8.5)$$

Even with these simplifications, however, it is still difficult to tell how the bias and mean square error will behave for various values of α, ρ, etc. Hence, numerical evaluations are necessary for any further analysis of any of the particular sampling plans which this estimation procedure encompasses.

9. A Numerical Example

As an indication of the effect of α, δ, and ρ upon the bias and the mean square error of the procedure, a simple numerical example is given in this section. This example assumes no second-stage sampling, i.e. $n_0=n_{0X}=n_{0Y}=n_1=n_{1X}=n_{1Y}=0$ while $n>0$, $n_X>0$, and $n_Y>0$.

Using the weights w_i, $i=1,2$ and c_i, $i=1,\ldots,4$ as suggested in section 8, the pooled estimator which is used whenever H_0: $\mu_Y=\mu_X$ is accepted becomes

$$\hat{\mu}_0 = [\sum_{i=1}^{4} h_i]^{-1}[h_1\overline{X}_n+h_2\overline{Y}_n+h_3\overline{X}_{n_X}+h_4\overline{Y}_{n_Y}] \tag{9.1}$$

where

$$h_1 = \frac{n}{(1-\rho^2)}\left[\frac{1}{\sigma_X^2} - \frac{\rho}{\sigma_X\sigma_Y}\right]; \quad h_3 = n_X/\sigma_X^2$$
$$h_2 = \frac{n}{(1-\rho^2)}\left[\frac{1}{\sigma_Y^2} - \frac{\rho}{\sigma_X\sigma_Y}\right]; \quad h_4 = n_Y/\sigma_Y^2. \tag{9.2}$$

If, in addition, $\rho=0$, then the pooled estimator in (9.1) reduces to a simple weighted average of $\overline{X}_{n+n_X}$ and $\overline{Y}_{n+n_Y}$, i.e.

$$\hat{\mu}_0^* = \frac{\dfrac{(n+n_X)\overline{X}_{n+n_X}}{\sigma_X^2} + \dfrac{(n+n_Y)\overline{Y}_{n+n_Y}}{\sigma_Y^2}}{(n+n_X)/\sigma_X^2 + (n+n_Y)/\sigma_Y^2}, \quad \rho=0. \tag{9.3}$$

For this example, the regression estimator, which is used whenever H_0: $\mu_Y=\mu_X$ is rejected, becomes

$$\hat{\mu}_1 = (g_1+g_2)^{-1}[g_1\overline{Y}_{n_Y}+g_2\{\overline{Y}_n + \frac{\rho\sigma_Y}{\sigma_X}(\overline{X}_{n+n_X}-\overline{X}_n)\}], \tag{9.4}$$

where

$$g_1 = n_Y/\sigma_Y^2;\ g_2 = n/[\sigma_Y^2(1-\rho^2 k)];\ k = n_X/(n+n_X). \qquad (9.5)$$

If, in addition, $\rho=0$, then the regression estimator simply reduces to $\overline{Y}_{n+n_Y}$, i.e. the unpooled estimator of μ_Y which uses none of the available information regarding μ_X.

Furthermore, this example is for the following specified parameters: $\sigma_X^2=25$, $\sigma_Y^2=16$, $n=15$, $n_X=30$, and $n_Y=10$. The values of δ, $B(\delta)$, and $MSE(\delta)$ have been computed for 3 values of $\alpha(.50, .25, .10)$, 7 values of $\rho(-0.5, -.25, 0, .25, .33, .50, .67)$, and 8 values of $\Delta(0, .8, 1.6, 2.4, 3.2, 4.0, 4.8, 5.6)$. These results are in Tables 1 through 3.

To give a reference point for mean square error when reading Tables 1 through 3, Table 4 gives, for 7 values of ρ, the variance of the regression estimator in (9.4), the variance of the pooled estimator in (9.1), and the variance of the unpooled estimator $\overline{Y}_{n+n_Y}$. Recall that the unpooled estimator and the regression estimator are unbiased for all values of $\Delta = \mu_Y-\mu_X$, whereas the pooled estimator is unbiased if, and only if, $\Delta=0$. Table 4 shows that the pooled estimator has the smallest variance, and the variance increases as ρ increases. This happens because the additional information on μ_X provided by the bivariate sample becomes less useful for estimating μ_Y as ρ increases. The variance of the regression estimator is maximum when $\rho=0$, and decreases as $|\rho|$ increases. The unpooled estimator has the largest variance, and is the same as the regression estimator when $\rho=0$.

If it is known that $\Delta=0$, then the obvious estimator for μ_Y is the unpooled estimator. If one does not know the value of Δ, but knows that $\Delta\neq0$, then the regression estimator should be used to estimate μ_Y. By making a preliminary test of significance, one expects to use the pooled estimator whenever $\Delta \doteq 0$ or $\delta \doteq 0$ and hence reduce the mean square error below the value of the variance (also mean square error) of the regression estimator.

Consider first the effect of δ on $-B(\delta)$ for all α and for all ρ. Tables 1 through 3 show that for $\delta=0$, $B(\delta)=0$ as noted previously in section 6. As δ increases beyond 0, $-B(\delta)$ increases monotonically to a maximum at a value of δ around 1.3 or 1.4. For δ increasing beyond 1.3 or 1.4, $-B(\delta)$ approaches zero asymptotically as stated in section 6.

The effect of α upon $-B(\delta)$ can be seen by noting that the maximum value of $-B(\delta)$ increases as α gets smaller. E.g., for $\rho=.25$, the maximum values of $-B(\delta)$ for $\alpha=.50$, $\alpha=.25$, and $\alpha=.10$ are, approximately, .03, .10, and .25, respectively. This relationship holds for all values of ρ. In addition, as α increases, $-B(\delta)$ attains its asymptotic value of zero for smaller values of δ, i.e. it approaches zero more rapidly. These properties occur because, as α gets smaller, the probability of making a Type II error increases, and the Type II error then results in a biased estimator of μ_Y. These effects of δ and α upon $B(\delta)$ are the same as those found by other investigators who considered only the special case when $\rho=0$.

The effect of ρ upon $-B(\delta)$ is more difficult to indicate directly because ρ affects $-B(\delta)$ in at least two ways. First, equations (4.4) and (6.7) show that as ρ increases, δ increases. Hence, an increase in ρ produces the same effects as an increase in δ. Second, equations (9.2) and (9.5) show that ρ, as a known parameter, is a component of the weights used in the weighted average estimators in (9.1) and (9.4). Hence, ρ is a component in the expressions for $B(\delta)$ and $MSE(\delta)$ other than via δ. Table 1 illustrates the interaction of these two factors. For $\Delta=.8$, as ρ increases from $-.50$ to $.67$, $-B(\delta)$ begins at .03, decreases to

.02, and then increases to .05. Now, if ρ were having no effect over and above its effect through δ, then -B(δ) should steadily increase as ρ increases because δ is increasing from .66 to .87. This isn't true, however, since -B(δ) decreases until ρ attains some point in the interval [0,¼). For Δ=1.6 in Table 1, one would expect -B(δ) to decrease as ρ increases if ρ was having its only effect through δ. However, -B(δ) decreases until ρ reaches a point in [0,¼), and then it begins to increase. The same general behavior is seen for Δ=2.4, although the minimum value of -B(δ) appears to occur for ρ ε [¼,⅓]. For Δ≥3.2 in Table 1, an increase in ρ produces a decrease in -B(δ), most likely primarily through the influence of δ. Similar patterns are seen in Tables 2 and 3, except that the value of Δ for which an increase in ρ always produces a decrease in -B(δ) gets smaller as α gets smaller (i.e. Δ=3.2, 2.4, and 1.6 in Tables 1, 2, and 3, respectively). In general, it appears from this example that an increase in ρ will cause the same effects as an increase in δ, with the following exceptions:

1) For δ in the range of, approximately 0 to .7, an increase in ρ produces a <u>decrease</u> in -B(δ) rather than an <u>increase</u>.

2) For δ in the range of, approximately, 1.6 to 1.8, an increase in ρ produces an <u>increase</u> in -B(δ) rather than a decrease.

Consider now the effect of δ on MSE(δ) for any given value of ρ and α. Looking down any column of Tables 1, 2, or 3, it can be seen that the minimum MSE(δ) occurs at δ=0. This minimum value of MSE(δ) is less than the variance of the regression estimator, but greater than the variance of the pooled estimator. As δ increases beyond 0, MSE(δ) increases monotonically until it reaches the value of the variance of the regression estimator. This occurs approximately around δ=.8. As δ increases beyond .8, MSE(δ) increases monotonically until it attains its maximum value for δ approximately equal to 2. As δ increases beyond 2, MSE(δ) decreases monotonically toward a limiting value which is the variance of the regression estimator. Hence, the pooling procedure yields maximum mean square error around δ=2, with MSE(δ) approaching from above the variance of the regression estimator for δ>2 and MSE(δ) less than the variance of the regression estimator for δ<.8 (approximately).

The effect of α upon MSE(δ) can be seen by noting that, for fixed ρ, the minimum value of MSE(δ) decreases as α decreases. E.g., for ρ=.25, the minimum value of MSE(δ) is .60, .54, and .45 for α=.50, .25, and .10, respectively. Also, the maximum value of MSE(δ) increases as α decreases. E.g., for ρ=.25, the maximum values of MSE(δ) are (approximately) .65, .74, and .90 for α=.50, .25, and .10, respectively. In addition, it can be noted from Tables 1 through 3 that, as α decreases, MSE(δ) approaches its asymptotic value more slowly. Hence, a smaller value of α will result in a larger reduction in MSE(δ) if δ is small (approximately less than .8), but, on the other hand, will result in a larger increase in MSE(δ) if δ is moderate (approximately equal to 2). These effects of δ and α upon MSE(δ) are the same as those found by other investigators for the special cases where ρ=0.

ρ can affect MSE(δ) either via δ or through its influence on the weights in the weighted estimators. The general effect upon MSE(δ) of increasing ρ from -.50 to .67, as illustrated in Tables 1 through 3, is to first increase MSE(δ) to some maximum value and then to decrease MSE(δ). For any given value of Δ, Table 1 shows that the maximum value for MSE(δ) is attained for ρ approximately equal to zero, whereas in Tables 2 and 3 the maximum value for MSE(δ) is attained for some value of ρ in the interval (-.25,.25). Hence, in general, it appears that an increase in the absolute value of ρ will decrease MSE(δ), although this relationship between ρ and MSE(δ) is definitely not symmetric about the point ρ=0.

10. Some General Conclusions

The bias and mean square error of a two-stage sampling scheme which involves a preliminary test of significance have been derived. Numerical investigation of some one-stage examples only indicate that in this procedure α and δ have an effect on $B(\delta)$ and $MSE(\delta)$ which is similar to that reported by other authors who have considered similar procedures. In addition, it appears that an increase in the absolute value of ρ will generally decrease $MSE(\delta)$. The relationship between ρ and $B(\delta)$ is not obvious, but it appears that ρ influences $B(\delta)$ primarily through the effect of ρ upon δ.

Table 1

Value of δ, $-B(\delta)$, and MSE(δ) for Various ρ and Δ where $\sigma_Y^2=16$, $\sigma_X^2=25$, n=15, $n_X=30$, $n_Y=10$, $\alpha=.50$

	ρ						
Δ	-.50	-.25	.00	.25	.33	.50	.67
0	0.0000[1]	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0000[2]	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.5379[3]	0.5955	0.6156	0.5997	0.5860	0.5444	0.4832
.8	0.6616	0.6940	0.7317	0.7762	0.7929	0.8301	0.8730
	0.0309	0.0263	0.0239	0.0263	0.0287	0.0372	0.0512
	0.5674	0.6220	0.6404	0.6267	0.6149	0.5792	0.5244
1.6	1.3232	1.3880	1.4633	1.5524	1.5858	1.6601	1.7459
	0.0342	0.0273	0.0230	0.0232	0.0246	0.0297	0.0375
	0.6083	0.6556	0.6685	0.6532	0.6418	0.6078	0.5527
2.4	1.9847	2.0819	2.1950	2.3286	2.3787	2.4902	2.6189
	0.0192	0.0137	0.0101	0.0089	0.0089	0.0096	0.0104
	0.6096	0.6527	0.6623	0.6432	0.6301	0.5912	0.5300
3.2	2.6463	2.7759	2.9266	3.1049	3.1717	3.3202	3.4919
	0.0064	0.0039	0.0024	0.0017	0.0016	0.0015	0.0013
	0.5895	0.6354	0.6475	0.6290	0.6156	0.5756	0.5140
4.0	3.3079	3.4699	3.6583	3.8811	3.9646	4.1503	4.3648
	0.0013	0.0007	0.0003	0.0002	0.0002	0.0001	0.0001
	0.5763	0.6263	0.6413	0.6244	0.6113	0.5718	0.5110
4.8	3.9695	4.1639	4.3899	4.6573	4.7575	4.9803	5.2378
	0.0002	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000
	0.5722	0.6241	0.6401	0.6238	0.6108	0.5714	0.5108
5.6	4.6311	4.8578	5.1216	5.4335	5.5504	5.8104	6.1107
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.5715	0.6238	0.6400	0.6237	0.6107	0.5714	0.5108

[1] δ

[2] $B(\delta)$

[3] MSE(δ)

Table 2

Value of δ, $-B(\delta)$, and MSE(δ) for Various ρ and Δ where $\sigma_Y^2=16$, $\sigma_X^2=25$, $n=15$, $n_X=30$, $n_Y=10$, $\alpha=.25$

	ρ						
Δ	-.5	-.25	0	.25	.33	.50	.67
0	0.0000[1]	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0000[2]	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.4620[3]	0.5205	0.5454	0.5385	0.5288	0.4974	0.4501
.8	0.6616	0.6940	0.7317	0.7762	0.7929	0.8301	0.8730
	0.1041	0.0996	0.0966	0.0978	0.0999	0.1076	0.1202
	0.5515	0.6110	0.6355	0.6285	0.6190	0.5877	0.5368
1.6	1.3232	1.3880	1.4633	1.5524	1.5858	1.6601	1.7459
	0.1273	0.1150	0.1044	0.0984	0.0980	0.1002	0.1058
	0.6942	0.7445	0.7566	0.7370	0.7234	0.6835	0.6194
2.4	1.9847	2.0819	2.1950	2.3286	2.3787	2.4902	2.6189
	0.0829	0.0680	0.0552	0.0459	0.0437	0.0406	0.0384
	0.7235	0.7566	0.7520	0.7172	0.6989	0.6497	0.5771
3.2	2.6463	2.7759	2.9266	3.1049	3.1717	3.3202	3.4919
	0.0335	0.0240	0.0166	0.0115	0.0102	0.0082	0.0065
	0.6602	0.6910	0.6880	0.6568	0.6399	0.5936	0.5265
4.0	3.3079	3.4699	3.6583	3.8811	3.9646	4.1503	4.3648
	0.0088	0.0053	0.0029	0.0016	0.0013	0.0008	0.0005
	0.6016	0.6428	0.6510	0.6296	0.6155	0.5744	0.5124
4.8	3.9695	4.1639	4.3899	4.6573	4.7575	4.9803	5.2378
	0.0015	0.0007	0.0003	0.0001	0.0001	0.0000	0.0000
	0.5777	0.6269	0.6414	0.6243	0.6111	0.5716	0.5109
5.6	4.6311	4.8578	5.1216	5.4335	5.5504	5.8104	6.1107
	0.0002	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000
	0.5723	0.6241	0.6401	0.6238	0.6108	0.5714	0.5108

[1] δ

[2] $B(\delta)$

[3] MSE(δ)

Table 3

Value of δ, $-B(\delta)$, and $MSE(\delta)$ for Various ρ and Δ where $\sigma_Y^2=16$, $\sigma_X^2=25$, $n=15$, $n_X=30$, $n_Y=10$, $\alpha=.10$

	ρ						
Δ	-.5	-.25	0	.25	.33	.50	.67
0	0.0000[1]	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0000[2]	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.3601[3]	0.4173	0.4479	0.4547	0.4517	0.4373	0.4119
.8	0.6616	0.6940	0.7317	0.7762	0.7929	0.8301	0.8730
	0.2104	0.2093	0.2072	0.2063	0.2067	0.2090	0.2118
	0.5146	0.5794	0.6125	0.6165	0.6110	0.5879	0.5443
1.6	1.3232	1.3880	1.4633	1.5524	1.5858	1.6601	1.7459
	0.2917	0.2778	0.2614	0.2461	0.2419	0.2353	0.2298
	0.8108	0.8728	0.8911	0.8702	0.8540	0.8049	0.7237
2.4	1.9847	2.0819	2.1950	2.3286	2.3787	2.4902	2.6189
	0.2294	0.2020	0.1734	0.1470	0.1392	0.1251	0.1120
	0.9488	0.9775	0.9565	0.8950	0.8660	0.7924	0.6902
3.2	2.6463	2.7759	2.9266	3.1049	3.1717	3.3202	3.4919
	0.1171	0.0917	0.0686	0.0494	0.0441	0.0348	0.0268
	0.8554	0.8591	0.8221	0.7555	0.7272	0.6593	0.5716
4.0	3.3079	3.4699	3.6583	3.8811	3.9646	4.1503	4.3648
	0.0397	0.0265	0.0164	0.0095	0.0078	0.0051	0.0031
	0.6984	0.7133	0.6973	0.6567	0.6375	0.5881	0.5200
4.8	3.9695	4.1639	4.3899	4.6573	4.7575	4.9803	5.2378
	0.0089	0.0049	0.0024	0.0010	0.0008	0.0004	0.0002
	0.6069	0.6442	0.6503	0.6282	0.6140	0.5730	0.5114
5.6	4.6311	4.8578	5.1216	5.4335	5.5504	5.8104	6.1107
	0.0013	0.0006	0.0002	0.0001	0.0000	0.0000	0.0000
	0.5777	0.6266	0.6411	0.6241	0.6110	0.5715	0.5108

[1] δ

[2] $B(\delta)$

[3] $MSE(\delta)$

Table 4

Variances of Three Estimators of μ_Y, for Various Values of ρ where $\sigma_Y^2=16$, $\sigma_X^2=25$, $n=15$, $n_X=30$, $n_Y=10$

	Estimator		
ρ	Regression (9.4)	Pooled (9.1)	Unpooled $\overline{Y}_{n+n_y}$
-.50	.5714	.2051	.6400
-.25	.6237	.2587	.6400
.00	.6400	.2974	.6400
.25	.6237	.3263	.6400
.33	.6107	.3342	.6400
.50	.5714	.3478	.6400
.67	.5108	.3581	.6400

REFERENCES

[1] Anderson, T. W., An Introduction to Multivariate Statistical Analysis, New York: John Wiley and Sons, Inc., 1958.

[2] Arnold, J. C. and H. A. Al-Bayyati, "On Double-Stage Estimation of the Mean Using Prior Knowledge," Biometrics, 26 (1970) 787-800.

[3] Asano, Chooichiro, "A Note on Modified Two-Sample-Theoretical Estimation of Biological Assay," Bulletin of Mathematical Statistics, 9 (1960) 41-56.

[4] Asano, Chooichiro, "Estimations after Preliminary Test of Significance and Their Applications to Biometrical Researches," Bulletin of Mathematical Statistics, 9 (1960) 1-23.

[5] Asano, Chooichiro, "Some Considerations on the Combination of Estimates from Different Biological Assays," Bulletin of Mathematical Statistics, 10 (1961) 17-32.

[6] Asano, Chooichiro and Sokuro Sato, "A Bivariate Analogue of Pooling of Data," Bulletin of Mathematical Statistics, 10 (1962) 39-59.

[7] Asano, Chooichiro and Masahiko Sugimura, "Some Considerations on Estimation of Population Variance Due to the Use of Pooling Data," Bulletin of Mathematical Statistics, 10 (1961) 33-44.

[8] Bailar, Barbara, "Recent Research in Reinterview Procedures," Journal of the American Statistical Association, 63 (1968) 41-63.

[9] Bancroft, T. A., "Analysis and Inference for Incompletely Specified Models Involving the Use of Preliminary Test(s) of Significance," Biometrics, 20 (1964) 427-442.

[10] Bennett, B. M., "Estimation of Means on the Basis of Preliminary Tests of Significance," Institute of Statistical Mathematics Annals, 4 (1952) 31-43.

[11] Bennett, B. M., "On the Use of Preliminary Tests in Certain Statistical Procedures," Institute of Statistical Mathematics Annals, 8 (1956) 45-52.

[12] Han, C. P. and T. A. Bancroft, "On Pooling Means When Variance is Unknown," Journal of the American Statistical Association, 63 (1968) 1333-1342.

[13] Huntsberger, D. V., "A Generalization of a Preliminary Testing Procedure for Pooling Data," Annals of Mathematical Statistics, 26 (1955) 734-743.

[14] Kale, B. K. and T. A. Bancroft, "Inference for Some Incompletely Specified Models Involving Normal Approximations to Discrete Data," Biometrics, 23 (1967) 335-348.

[15] Kitagawa, T., "Estimation After Preliminary Tests of Significance," University of California Publications in Statistics, 3 (1963) 147-186.

[16] Mehta, J. S. and John Gurland, "Testing Equality of Means in the Presence of Correlation," Biometrika, 56 (1969) 119-126.

[17] Mosteller, F., "On Pooling Data," Journal of the American Statistical Association, 43 (1948) 231-242.

[18] Ruhl, D. J. B., Preliminary Test Procedures and Bayesian Procedures for Pooling Correlated Data, unpublished Ph.D Dissertation, Ames, Iowa: Iowa State University, 1967.

[19] Sato, Sokuro, "A Multivariate Analogue of Pooling of Data," Bulletin of Mathematical Statistics, 10 (1962) 61-76.

[20] Tamura, Ryoji, "Nonparametric Inferences with a Preliminary Test," Bulletin of Mathematical Statistics, 11 (1965) 39-61.

[21] Tamura, Ryoji, "Some Estimate Procedures with a Nonparametric Preliminary Test I," Bulletin of Mathematical Statistics, 11 (1965) 63-71.

[22] Yen, Elizabeth, "On Two State Non-Parametric Estimation," Annals of Mathematical Statistics, 35 (1964) 1099-1114.