Hitotsubashi Journal of Economics 25 (1984) 61-69. © The Hitotsubashi Academy

FIRST, SECOND AND THIRD ORDER EFFICIENCIES OF THE ESTIMATORS FOR A COMMON MEAN

By TAKEAKI KARIYA, BIMAL K. SINHA* AND KASALA SUBRAMANYAM*

Based on the concepts of first, second and third order efficiencies developed by Kariya, Krishnaiah and Rao (1981) and Kariya (1981), this paper attempts to order the estimators proposed in the problem of estimating the common mean of K univariate normal populations. Only the Graybill and Deal (1959) estimator is shown to be third order efficient.

I. *Introduction*

Let $(\bar{x}_i, s_i^2)$ $(i=1, \ldots, K)$ be the K independent pairs of the sample mean and the unbiased sample variance from K univariate normal populations $N(\mu, \sigma_i^2)$, where the sample size for the i-th pair $(\bar{x}_i, s_i^2)$ is N_i. In this situation, the problem of estimating the common mean μ of the K populations has been extensively treated in the literature ([2], [3], [6], [7], [10], [14], [15], [16], [17]), and many estimators have been proposed. However, no unified comparison among those estimators has been yet attempted. A main difficulty in the comparison is caused by the incompleteness of the model which renders us to have ancillary statistics $\bar{x}_i - \bar{x}_j$ $(i \neq j)$. In order to make possible a comparison in such a model as above where it admits some ancillary statistics, Kariya, Krishnaiah and Rao (1981) and Kariya (1981) developed concepts of FOE (first order efficiency or often first order efficient), SOE (second order efficiency or often second order efficient) and higher order efficiencies for Fisher consistent estimators, where the MSE (mean squared error) criterion is adopted. These concepts are defined for each fixed sample size and different from those defined in such asymptotic manners as in Rao (1961, 1963), Ghosh and Subramanyam (1974), Ghosh, Sinha and Weiand (1980), Akahira and Takeuchi (1980), Pfanzagl (1980 etc. (see Kariya (1981) for some differences.) In this paper, applying these concepts to the problem of estimating the common mean μ, several estimators proposed so far by various authors are ordered.

More specifically, in section 2, when $K=2$, we obtain necessary and sufficient conditions for a Fisher consistent estimator to be FOE, for an FOE estimator to be SOE, and for an SOE estimator to be TOE (third order efficient (or sometimes efficiency)). The concepts and implications of these efficiencies are reviewed in terms of the present problem. In section III, these conditions are checked stepwise for the following estimators:

(1) Graybill-Deal (1959) type estimator of the form

* The authors were sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, under Contract F49620-79-C-0161.

(1.1) $\hat{\mu}_1(c_1,c_2)=[\sum_{i=1}^{2} c_i\bar{x}_i/s_i^2]/[\sum_{i=1}^{2} c_i/s_i^2]$,

where c_i's are constants. In particular, $\hat{\mu}_1{}^*=\hat{\mu}_1(N_1, N_2)$ will be called Graybill-Deal estimator.

(2) Zacks (1966) estimator.
(1.2) $\hat{\mu}_2=\bar{x}G+\hat{\mu}_1(1-G)$
where $\bar{x}=(N_1\bar{x}_1+N_2\bar{x}_2)/(N_1+N_2)$, and $G=I(a^{-1}<s_2^2/s_1^2<a)$, the indicator function of $\{a^{-1}<s_2^2/s_1^2<a\}$. This estimator is a preliminary test estimator, where the hypothesis is $\sigma_1^2=\sigma_2^2$.

(3) Gurland-Mehta (1969) estimator.
(1.3) $\hat{\mu}_3=\phi_3\bar{x}_1+(1-\phi_3)\bar{x}_2$ with $\phi_3=(cs_1^2+s_2^2)/((c+a)s_1^2+s_2^2)$.
In Gurland and Mehta (1969), $a=1$ and $c=0.4$ are proposed through a numerical comparison.

(4) Brown-Cohen (1974) type estimator.
(1.4) $\hat{\mu}_4(a,c_1,c_2,c_3)=\bar{x}_1+(\bar{x}_2-\bar{x}_1)\{av_1/[c_1v_1+c_2v_2+c_3(\bar{x}_2-\bar{x}_1)^2]\}$,
where $v_i=s_i^2/N_i$ $(i=1,2)$, a and c_i's are constants. In Brown and Cohen (1974), $\hat{\mu}_4(a_1,1,(N_2-1)/(N_2+2), 1/(N_2+2))$ for N_2 small and $\hat{\mu}_4(a_2,1,1,0)$ for N_2 large are proposed, where a_1 and a_2 are certain constants.

(5) Cohen-Sackrowitz (1974) estimator $(N_1=N_2)$.
(1.5) $\hat{\mu}_5=[1-c_oG(s_1^2,s_2^2)]\bar{x}_1+c_oG(s_1^2,s_2^2)\bar{x}_2$,
where $c_o=(N-4)/(N+2)$ for $N=N_1=N_2$ even, $c_o=(N-3)^2/(N+1)(N-1)$ for N odd, and $G(s_1^2,s_2^2)$ is the unique unbiased estimator of $\sigma_1^2/(\sigma_1^2+\sigma_2^2)$ based on (s_1^2, s_2^2).

(6) The likelihood equation estimator and its modified version.
Here the modification is made for the degrees of freedom of s_1^2 and s_2^2. The estimating equations of these estimators are given by cubic polynomials.

In the literature, no attention has been paid to the MLE (maximum likelihood estimator) because of its intractability. Apart from the estimators in (6), common features of the estimators (1)–(5) are that they are all unbiased and that they are written in the following form
(1.6) $\tilde{\mu}=\phi\bar{x}_1+(1-\phi)\bar{x}_2$.
It is noted that any estimator of this form is Fisher consistent (see section II for definition). In section III, the Graybill-Deal estimator, a Brown-Cohen estimator and the modified likelihood equation estimator are shown to be SOE as well as FOE, but later only the Graybill-Deal estimator is shown to be third order efficient (TOE). Consequently, from the viewpoint of a stepwise ordering based on FOE, SOE, and TOE, the Graybill-Deal estimator is preferred.

We remark that the argument is applicable to the case of K populations $(K\geq 3)$ without any difficulty but with some complication. Secondly, it is also remarked that many authors considered the problem of comparing $\mathrm{Var}(\tilde{\mu})$ with $\tilde{\mu}$ in (1.6) and $\mathrm{Var}(\bar{x}_i)$ $(i=1, \ldots, K)$, and have obtained necessary and sufficient conditions for which $\mathrm{Var}(\tilde{\mu})\leq\mathrm{Var}(\bar{x}_i)$ for some i or all i with a particular form of ϕ ([10], [14]). Sinha (1979) and Sinha and Mouqadem (1981) considered the admissibility problems of the Graybill-Deal type estimators and some other estimators.

II. *Conditions for FOE, SOE and TOE*

In this section, following the framework of KKR (Kariya, Krishnaiah and Rao) (1981) and Kariya (1981), necessary and sufficient conditions for FOE, SOE and TOE as applicable to the present problem are derived and the concepts of the FOE, SOE and TOE in general are reviewed. Assume $K=2$ and let

(2.1) $z\equiv(z_1, z_2, z_3, z_4)'=(\bar{x}_1, \bar{x}_2, s_1^2, s_2^2)'$

which is a sufficient statistic for

(2.2) $\eta\equiv(\eta_1, \eta_2, \eta_3)'=(\mu, \sigma_1^2, \sigma_2^2)'$.

Clearly the expected value of z is

(2.3) $\theta\equiv\theta(\eta)\equiv(\theta_1, \theta_2, \theta_3, \theta_4)'=(\mu, \mu, \sigma_1^2, \sigma_2^2)'$.

Here an estimator $h(z)$ of μ is called *Fisher consistent* if it satisfies

(2.4) $h(\theta(\eta))=\eta_1=\mu$ for all η.

Let C be the class of Fisher consistent estimators satisfying the following regularity condition:

(a) $\partial^2 h/\partial z_i\partial z_j(z)$ exists, continuous and of order $0(1)$ when h depends on N_1 and N_2, and $N_1\to\infty$, $N_2\to\infty$.

An estimator h in C can be expanded as

$$(2.5)\quad h(z)=\mu+\sum_{i=1}^{4}h^i(\theta)d_i+\frac{1}{2}\sum_i\sum_j h^{ij}\lambda(z+(1-\lambda)\theta)d_id_j,$$

where $d_i=z_i-\theta_i$, $h^i=\partial h/\partial z_i$, $h^{ij}=\partial^2 h/\partial z_i\partial z_j$, and $0\leq\lambda\equiv\lambda(z, \theta)\leq 1$. Hence, adopting the MSE criterion yields

(2.6) $E(h(z)-\mu)^2=x'Ax+R$, where

(2.7) $x\equiv x(\theta)=(h^1(\theta), \dots, h^4(\theta))'$,

(2.8) $A=(E(d_id_j))=\operatorname{diag}\{\sigma_1^2/N_1, \sigma_2^2/N_2, 2\sigma_1^4/n_1, 2\sigma_2^4/n_2\}$, $\quad n_i=N_i-1$,

and R is the remainder term. Since $z\to\theta$ a.s. as $N_1\to\infty$ and $N_2\to\infty$ and since h^{ij} is continuous, $R=0((N_1+N_2)^{-1})$ (see Remark 2.1 in [9]). The implication of the following definition is straightforward.

Definition 2.1 (KKR (1981)). An estimator h in C is said to be FOE if for each (N_1, N_2), it minimizes $x'Ax$ in C.

To minimize $x'Ax$ in C, differentiate (2.4) with respect to η_j to have the side condition $Bx=c$, where

$$(2.9)\quad B=(\partial\theta_i/\partial\eta_j)=\begin{pmatrix}1&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}\quad\text{and}\quad c=\begin{pmatrix}1\\0\\0\end{pmatrix}$$

Hence, directly minimizing $x'Ax$ under $Bx=c$ or applying Theorem 2.1 in KKR (1981) yields

Theorem 2.1. An estimator h in C is FOE if and only if $x=x_o$, where

(2.10) $x_o = A^{-1}B'(BA^{-1}B')^{-1}c = \left(\frac{N_1\sigma_1^2}{N_1\sigma_2^2+N_2\sigma_1^2}, \frac{N_2\sigma_1^2}{N_1\sigma_2^2+N_2\sigma_1^2}, 0, 0\right)'$

The minimum value of $x'Ax$ in C is $J_1 = c'(BA^{-1}B')^{-1}c = \sigma_1^2\sigma_2^2/(N_1\sigma_2^2+N_2\sigma_1^2)$.

Any estimator of the form (1.6) satisfies (2.4) whatever $\phi=\phi(z)$ may be.

Corollary 2.1 An estimator h of the form (1.6) satisfying (a) is FOE if and only if
(2.11) $\phi(\theta(\eta)) = N_1\sigma_2^2/(N_1\sigma_2^2+N_2\sigma_1^2)$.

Proof Differentiating $h(z)=\phi(z)z_1+(1-\phi(z))z_2$ at $z=\theta$ and equating $x\equiv(h^1(\theta), \ldots, h^4(\theta))'$ with x_0 in (2.10) yields (2.11).

Next, to define a concept of SOE, we assume for FOE estimators
(b) $h^{ijk}(z)\equiv\partial^3h/\partial z_i\partial z_j\partial z_k(z)$ exists, continuous, and of order $0(1)$ as $N_1\to\infty$, $N_2\to\infty$ and let

$C_1 = \{h\in C \mid h \text{ satisfies (b)}, x=x_o\}$.

Then for $h\in C_1$, h can be written as

(2.12) $$h(z)=\mu+\sum_i x_{oi}d_i+\frac{1}{2}\sum_i\sum_j h^{ij}(\theta)d_id_j + \frac{1}{6}\sum_i\sum_j\sum_k h^{ijk}(\lambda z+(1-\lambda)\theta)d_id_jd_k$$

where $0\le\lambda\le1$, and $x_0=(x_{01}, \ldots, x_{04})'$ is given by (2.10). Based on Kariya (1981), we consider the measure

(2.13) $\nu_2 = E[h(z)-\mu-\sum_i x_{oi}d_i]^2 = \xi_2+R_2$,

where

(2.14) $$\xi_2=\frac{1}{4}\sum_i\sum_j\sum_k\sum_l h^{ij}(\theta)h^{kl}(\theta)c_{ijkl}$$

(2.15) $c_{ijkl}=E(d_id_jd_kd_l)$

and R_2 is the remainder term. It is noted that $c_{iiii}=E(d_i^4)=O(N_i^{-2})$ and $R_2=O((N_1+N_2)^{-2})$.

Definition 2.2 (Kariya (1981)). An estimator h in C_1 is said to be SOE if for each (N_1, N_2), it minimizes ξ_2 in C_1.

An intuitive rationale for this definition follows from (2.12) and (2.13) since ν_2 measures the degree of concentration of $h(z)$ toward $\mu+\sum_i x_{oi}d_i$ and since ξ_2 is the leading term of ν_2.

It is noted that $E(d_i)=E(z_i-\theta_i)=0$, $d_i\to0$ a.s. when N_1 and $N_2\to\infty$, and x_{oi}'s are common for all $h\in C_1$. Another rationale for Definition 2.2 is given by

(2.16) $[E(h(z)-\mu)^2]^{\frac{1}{2}}\le J_1^{\frac{1}{2}}+\nu_2^{\frac{1}{2}}$,

where J_1 is the minimum value of $x'Ax$ in Theorem 2.1 which is common for all $h\in C_1$. Hence minimizing the leading term ξ_2 of order $0((N_1+N_2)^{-2})$ in ν_2, we can control the MSE up to order $0((N_1+N_2)^{-2})$ via (2.16). To carry out the minimization, let

(2.17) $y=(h^{11}(\theta), \ldots, h^{14}(\theta), h^{22}(\theta), \ldots, h^{24}(\theta), h^{33}(\theta), h^{34}(\theta), h^{44}(\theta))'$

(2.18) $\alpha_3=-N_1N_2\sigma_2^2/(N_1\sigma_2^2+N_2\sigma_1^2)^2$ and $\alpha_4=N_1N_2\sigma_1^2/(N_1\sigma_2^2+N_2\sigma_1^2)^2$.

Theorem 2.2 An estimator h in C_1 is SOE if and only if $y=y_0$, where

(2.19) $y_0=(0, 0, \alpha_3, \alpha_4, 0, -\alpha_3, -\alpha_4, 0, 0, 0)'$

The minimum value of ξ_2 is given by

(2.20) $J_2=2\{\alpha_3^2\sigma_1^6/N_1^2+\alpha_3^2\sigma_2^2\sigma_1^4/N_1N_2+\alpha_4^2\sigma_1^2\sigma_2^4/N_1N_2+\alpha_4^2\sigma_2^4/N_2^2\}$.

Proof Since $h\in C_1$ satisfies $x=x_0$, differentiating $x=x_0$ with respect to η_j $(j=1, 2, 3)$ yields

(2.21) $\begin{cases} h^{11}+h^{12}=0,\ h^{13}=\alpha_3,\ h^{14}=\alpha_4,\ h^{21}+h^{22}=0,\ h^{23}=-\alpha_3 \\ h^{24}=-\alpha_4,\ h^{31}+h^{32}=0,\ h^{33}+h^{34}=0,\ h^{41}+h^{42}=0,\ h^{43}+h^{44}=0 \end{cases}$

Substituting (2.21) into ξ_2,

(2.22) $$\begin{aligned} 4\xi_2 &= \sum(h^{ii})^2E(d_i^4)+\sum_{i\neq j}\sum h^{ii}h^{jj}E(d_i^2)E(d_j^2)+2\sum_{i\neq j}\sum(h^{ij})^2E(d_i^2d_j^2), \\ &=\delta^2[\sum_{i=1}^{2}E(d_i^4)+6E(d_1^2d_2^2)]+4\alpha_3^2[E(d_1^2d_3^2)+E(d_2^2d_3^2)] \\ &\qquad +4\alpha_4^2[E(d_1^2d_4^2)+E(d_2^2d_4^2)], \end{aligned}$$

where $\delta=h^{11}=-h^{12}=h^{22}$. This is clearly minimized if and only if $\delta=0$, which together with (2.21) gives (2.17). The minimum value of ξ_2 is obtained by inserting $\delta=0$ and the expected values $E(d_i^2d_j^2)$ into (2·22). This completes the proof.

Corollary 2.2 An FOE estimator of the form (1.6) satisfying (b) is SOE if and only if $\phi^1(\theta)=\phi^2(\theta)=0$, $\phi^3(\theta)=\alpha_3$ and $\phi^4(\theta)=\alpha_4$, where $\phi^i(\theta)=\partial\phi/\partial z_i(\theta)$.

Proof Computing the derivatives of $\tilde{\mu}$ in (1.6) and equating them with y_o in (2.19) yields the result.

Finally, we consider a condition for TOE. Let C_2 be the class of SOE estimators satisfying

(c) $h^{ijkl}(z)=\partial^4h/\partial z_i\partial z_j\partial z_k\partial z_l(z)$ exists, continuous and of order $0(1)$; as $N_1\to\infty$, $N_2\to\infty$, i.e. $C_2=\{h\in C_1 \mid h \text{ satisfies (c)}, y=y_o\}$.

Then for $h\in C_2$, h can be expanded as

(2.23) $$\begin{aligned} h(z)&=\mu+\sum_i x_{oi}d_i+\frac{1}{2}\sum_i\sum_j y_{oij}d_id_j+\frac{1}{6}\sum_i\sum_j\sum_k h^{ijk}(\theta)d_id_jd_k \\ &\quad+\frac{1}{24}\sum_i\sum_j\sum_k\sum_l h^{ijkl}(\lambda z+(1-\lambda)\theta)d_id_jd_kd_l, \end{aligned}$$

where $0\leq\lambda\leq1$, and $y_o=(y_{o11}, \ldots, y_{o44})$ is given by (2.19). Similar to the case of the definition of SOE, we consider the measure

(2.24) $\nu_3=E[h(z)-\mu-\sum_i x_{oi}d_i-\frac{1}{2}\sum_i\sum_j y_{oij}d_id_j]^2=\xi_3+R_3$,

where

(2.25) $\xi_3=\frac{1}{36}\Sigma\Sigma\Sigma\Sigma\Sigma\Sigma h^{ijk}h^{lmn}e(i, j, k, l, m, n)$,

(2.26) $e(i, j, k, l, m, n)=E[d_id_jd_kd_ld_md_n]$

and R_3 is the remainder term. It is noted that $e(i, \ldots, i)=0(N_i^{-3})$ and $R_3=o((N_1+N_2)^{-3})$.

Definition 2.3 An estimator h in C_2 is said to be TOE if for each (N_1,N_2), it minimizes ξ_3 in C_2.

The rationale and implications of this definition are similar to those of Definition 2.2. Especially, like (2.16), we have

(2.27) $[E(h(z)-\mu)^2]^{\frac{1}{2}} \leq J_1^{\frac{1}{2}}+J_2^{\frac{1}{2}}+\nu_3^{\frac{1}{2}}$,
where J_2 is given by (2.20). Hence minimizing the leading term ξ_3 of ν_3, we can control the MSE up to $0((N_1+N_2)^{-3})$ via (2.27). To minimize ξ_3, we first derive the side conditions implied by $y=y_0$, where y and y_0 are given by (2.17) and (2.19) respectively. Differentiating $h^{ij}=y_{0ij}$ with respect to μ, σ_1^2 and σ_2^2 gives the following set of conditions on h^{ijk}.

$$
(2.28)\quad
\begin{array}{lll}
h^{111}+h^{112}=0, & h^{113}=0, & h^{114}=0 \\
h^{121}+h^{122}=0, & h^{123}=0, & h^{124}=0 \\
h^{131}+h^{132}=0, & h^{133}=\beta_1, & h^{134}=\beta_2 \\
h^{141}+h^{142}=0, & h^{143}=\gamma_1, & h^{144}=\gamma_2 \\
h^{221}+h^{222}=0, & h^{223}=0, & h^{224}=0 \\
h^{231}+h^{232}=0, & h^{233}=-\beta_1, & h^{234}=-\beta_2 \\
h^{241}+h^{242}=0, & h^{243}=-\gamma_1, & h^{244}=-\gamma_2 \\
h^{331}+h^{332}=0, & h^{333}=0, & h^{334}=0 \\
h^{341}+h^{342}=0, & h^{343}=0, & h^{344}=0 \\
h^{441}+h^{442}=0, & h^{443}=0, & h^{444}=0
\end{array}
$$

where α_3 and α_4 are as in (2.18) and
(2.29) $\beta_1=\partial\alpha_3/\partial\sigma_1^2$, $\beta_2=\partial\alpha_3/\partial\sigma_2^2=\partial\alpha_4/\partial\sigma_1^2=\gamma_1$, $\gamma_2=\partial\alpha_4/\partial\sigma_2^2$.

Theorem 2.3 An estimator h in C_2 is TOE if and only if $h^{133}=\beta_1$, $h^{134}=\beta_2$, $h^{144}=\gamma_2$, $h^{233}=-\beta_1$, $h^{234}=-\beta_2$, $h^{244}=-\gamma_2$, and all the other h^{ijk}'s are zero.
Proof Let $h^{111}=x$. From (2.28), it is easily shown that $h^{112}=-x$, $h^{122}=x$, $h^{222}=-x$, $h^{133}=\beta_1$, $h^{134}=\beta_2$, $h^{144}=\gamma_2$, $h^{233}=-\beta_1$, $h^{234}=-\beta_2$, $h^{244}=-\gamma_2$ and all the other h^{ijk}'s are zero. Substituting these values into ξ_3, using the independence of d_i's and minimizing ξ_3 with respect to x yielxs $x=0$. Therefore the result follows.

III. *FOE, SOE and TOE of $\hat{\mu}_i$*

In this section we check whether or not the estimators $\hat{\mu}_i$ $(i=1, \ldots, 5)$ listed in (1.1)-(1.5) and the likelihood equation estimator are FOE, SOE and TOE. Since all $\hat{\mu}_i$ are of the form (1.6) and since any estimator of the form (1.6) is Fisher consistent in the sense of (2.5), by Corollary 2.1, $\hat{\mu}_i(z)=\phi_i(z)z_1+(1-\phi_i(z))z_2$ is FOE if and only if $\phi_i(\theta)=N_1\sigma_2^2/(N_1\sigma_2^2+N_2\sigma_1^2)\equiv\gamma$, and ϕ_i is continuously twice differentiable. First, the Graybill-Deal type estimator $\hat{\mu}_1$ in (1.1) has $\phi_1(z)=c_1z_4/(c_1z_4+c_2z_3)$, and so it is FOE if and only if $c_1=N_1$ and $c_2=N_2$. Consequently the Graybill-Deal estimator $\hat{\mu}_1^*=\hat{\mu}_1(N_1,N_2)$ is FOE. Secondly, write the Zacks estimator $\hat{\mu}_2$ in (1.2) in the form of (1.6) with $\phi=\phi_2$, where
(3.1) $\phi_2(z)=[N_1/(N_1+N_2)]G_2+[N_1z_4/(N_1z_4+N_2z_3)](1-G_2)$
where G_2 is the indicator function of $a^{-1}<z_4/z_3<a$. Here G_2 is not differentiable. But G_2 can be approximated by a continuously twice differentiable function G_2^* such that G_2^* agrees with G_2 except on the intervals $[a^{-1}-\varepsilon, a^{-1}+\varepsilon]$ and $[a-\varepsilon, a+\varepsilon]$, where $\varepsilon>0$ is arbitrarily small. Then ϕ_2^* with G_2^* for G_2 in (3.1) does not satisfy $\phi_2^*(\theta)=\gamma$, and so the estimator $\hat{\mu}_2$ with ϕ_2^* is not FOE for any $\varepsilon>0$. Hence, the Zacks estimator is not approximately FOE. Thirdly, in the Gurland-Mehta estimator $\hat{\mu}_3$, ϕ_3 in (1.3) satisfies $\phi_3(\theta)=\gamma$ if and only if $c=0$ and $a=N_2/N_1$, in which case $\hat{\mu}_3=\hat{\mu}_1^*$. Therefore, the Gurland-Mehta estimator with $c>0$ is not FOE. Fourthly, the Brown-Cohen type estimator $\hat{\mu}_4$ in (1.4) has

(3.2) $\phi_4(z)=[(c_1-a)N_1^{-1}z_3+c_2N_2^{-1}z_4+c_3(z_1-z_2)^2]/[c_1N_1^{-1}z_3+c_2N_2^{-1}z_4+c_3(z_1-z_2)^2$

This ϕ_4 satisfies $\phi_4(\theta)=\gamma$ if and only if $c=c_1=c_2\neq 0$. Hence the estimator $\hat{\mu}_4{}^*(b)=\phi_4{}^*(z)z_1+(1-\phi_4{}^*(z))z_2$ with

(3.3) $\phi_4{}^*(z)=[N_1z_4+b(z_1-z_2)^2]/[N_2z_3+N_1z_4+b(z_1-z_2)^2]$

is FOE, where $b=c_3/a\geq 0$ is arbitrary. It is noted that $\hat{\mu}_4{}^*(\theta)+\hat{\mu}_1{}^*$ and that the Brown-Cohen estimator proposed for N_2 small is not FOE. Fifthly, it is easy to see that the Cohen-Sackrowitz estimator $\hat{\mu}_5$ in (1.5) is not FOE.

Finally we consider the likelihood equation estimator. As is easily shown, maximizing the log likelihood function with respect to $\sigma_i{}^2$ $(i=1,2)$ yields the log likelihood function of μ;

(3.4) $l(\mu: z)=-(N_1/2)\log[(z_1-\mu)^2+n_1N_1^{-1}z_3]-(N_2/2)\log[(z_2-\mu)^2+n_2N_2^{-1}z_4]$

and so the likelihood equation of μ is given by the cubic polynomial

(3.5) $N_1n_2N_2^{-1}z_4(z_1-\hat{\mu})+N_2n_1N_1^{-1}z_3(z_2-\hat{\mu})+N_1(z_1-\hat{\mu})(z_2-\hat{\mu})^2+N_2(z_2-\hat{\mu})(z_1-\hat{\mu})^2=0$

In addition to an analytical difficulty in handling this equation, the equation sometimes gives 3 real roots, say $m_i=m_i(z)$ $(i=1,2,3)$. Hence the MLE is defined as $\hat{\mu}_6$ satisfying

(3.6) $l(\hat{\mu}_6(z): z)=\max\{l(m_i(z): z)\mid i=1,2,3\}$.

However, this does not mean that one of the roots is the MLE. That is, depending on z, sometimes m_i maximizes $l(\mu;z)$ and sometimes m_2 or m_3 maximizes it. Moreover, setting $z=\theta$ in (3.5), one gets

(3.7) $(\hat{\mu}-\mu)[N_1n_2N_2^{-1}\sigma_2{}^2+N_2n_1N_1^{-1}\sigma_1{}^2+N_1(\hat{\mu}-\mu)^2+N_2(\hat{\mu}-\mu)^2]=0.$

From this, it is observed that only one of the m_i's is Fisher consistent ($\hat{\mu}(\theta)=\mu$), and the other roots of (3.7) are complex. However, the authors have been unable to identify which root of (3.5) is Fisher consistent since analytical solutions of (3.5) are intractable. Without identifying it, we shall treat the Fisher consistent root of (3.5) and call it $\hat{\mu}_6{}^*$. Regarding (3.5) as an implicit function $F(z,\hat{\mu}(z))=0$ and differentiating it with respect to z_j's at $z=\theta$ produces the gradient vector of $\hat{\mu}$ at $z=\theta$:

(3.8) $(\hat{\mu}^1(\theta),\ldots,\hat{\mu}^4(\theta))=(N_1n_2N_2^{-1}\sigma_2{}^2/D,\ N_2n_1N_1^{-1}\sigma_1{}^2/D,\ O,\ O)$, where $D=N_1n_2N_2^{-1}\sigma_2{}^2+N_2n_1N_1^{-1}\sigma_1{}^2$ and $\hat{\mu}^i=\partial\hat{\mu}/\partial z_i$.

It is noted that this gradient is common for $\hat{\mu}=m_i$, the roots of (3.5). Therefore by Theorem 2.1, $\hat{\mu}_6{}^*$ is not FOE. On the other hand, if $n_iN_i^{-1}$'s are replaced by 1 in (3.8), the gradient vector agrees with x_0 in Theorem 2.1. This suggests the following modification. Replace both $n_2N_2^{-1}$ and $n_1N_1^{-1}$ by 1 in the likelihood equation (3.5) and define $\hat{\mu}_7$ as the Fisher consistent root of the modified equation. We shall call this estimator the modified likelihood equation estimator. Then, in a similar manner, $\hat{\mu}_7$ is easily shown to be FOE. The estimator $\hat{\mu}_7$ is not yet identified, but since in Theorem 3.3, $\hat{\mu}_7$ is shown to be *not* TOE, this identification is really not necessary and a TOE estimator is preferred to $\hat{\mu}_7$. The above results are summerized as

Theorem 3.1 Among the estimators $\hat{\mu}_i$ $(i=1,\ldots,7)$, the Graybill-Deal estimator $\hat{\mu}_1{}^*=\hat{\mu}_1(N_1,N_2)$, the Brown-Cohen type estimator $\hat{\mu}_4{}^*$(b) with $\phi_4{}^*$ in (3.3) and the modified likelihood equation estimator $\hat{\mu}_7$ are FOE, and the others are not.

To see whether these estimators are SOE, it is noted that $\hat{\mu}_1{}^*=\hat{\mu}_4{}^*(\theta)$, and $\hat{\mu}_4{}^*$(b) and $\hat{\mu}_7$ belong to the class C_1 defined in section 2. Applying Corollary 2.2 and computing the partial derivatives of $\phi_4{}^*$ at $z=\theta$ verifies that for any $b\geq 0$, $\hat{\mu}_4{}^*$(b) is SOE. Similarly, from the equation (3.5) with both $n_1N_1^{-1}$ and $n_2N_2^{-1}$ replaced by 1, the vector of second partial derivatives of $\hat{\mu}_7$ is computed and it is shown to be equal to y_0 in Theorem 2.2. Hence we obtain

Theorem 3.2 The Graybill-Deal estimator $\hat{\mu}_1^*=\hat{\mu}_4^*(0)$, the Brown-Cohen type estimator $\hat{\mu}_4^*(b)$ and the modified likelihood equation estimator $\hat{\mu}_7$ are all SOE.

Consequently, based on the SOE, these estimators are not discriminated. However, the concept of TOE does discriminate them.

Theorem 3.3 Among the three estimators $\hat{\mu}_1^*$, $\hat{\mu}_4^*$ and $\hat{\mu}_7$, only the Graybill-Deal estimator $\hat{\mu}_1^*$ is TOE.

Proof To show that $\hat{\mu}_4^*$ and $\hat{\mu}_7$ are not TOE, from Theorem 2.3, it suffices to show that $\partial^3\hat{\mu}_4^*/\partial z_1^3(\theta)\neq 0$ and $\partial^3\hat{\mu}_7/\partial z_1^3(\theta)\neq 0$. For example, for $\hat{\mu}_7$, regarding (3.5) as $F(z,\hat{\mu}_7(z))=0$, differentiating it three times with respect to z_1 and evaluating it at $z=\theta$ yields $h^{111}[N_1\sigma_2^2+N_2\sigma_1^2]+(h^1)^3[6(N_1+N_2)]+(h^1)^2[-6(2N_2+N_1)]+h^1(6N_2)=0$, from which we obtain $h^{111}\neq 0$ where $h=\hat{\mu}_7$. Similarly $\partial^3\hat{\mu}_4^*/\partial z_1^3(\theta)\neq 0$ is easily shown. To show that $\hat{\mu}_1^*$ is TOE, we need to verify that all the third derivatives of $\hat{\mu}_1^*$ at $z=\theta$ are equal to the ones in Theorem 2.3. This is directly checked. Therefore, the proof is completed.

By this theorem, from the viewpoint of the present paper, the Graybill-Deal estimator $\hat{\mu}_1$ is most preferred.

HITOTSUBASHI UNIVERSITY, UNIVERSITY OF PITTSBURGH, AND UNIVERSITY OF PITTSBURGH

REFERENCES

[1] Akahira, M. and Takeuchi, K., 1980, The Concept of Asymptotic Efficiency and Higher Order Asymptotic Efficiency in Statistical Estimation Theory, University of Electro-Comm., Japan.

[2] Brown, L.D. and Cohen, A., 1974, Point and Confidence Estimation of a Common Mean and Recovery of Inter-block Information, *Annals* of *Statistics* 2, 963–976.

[3] Cohen, A. and Sackrowitz, H.B., 1974, On Estimating the Common Mean of Two Normal Populations, *Annals* of *Statistics* 2, 1274–1282.

[4] Ghosh, J.K. and Subramanyam, K., 1974, Second Order Efficiency of Maximum Likelihood Estimators, *Sankhyā* (A) 36, 325–358.

[5] Ghosh, J.K., Sinha, B.K. and Wieand, H.S., 1980, Second Order Efficiency of the MLE with Respect to Any Bounded Bowl-shaped Loss Function, *Annals* of *Statistics* 8, 506–521.

[6] Graybill, F.A. and Deal, R.B., 1959, Combining Unbiased Estimators, *Biometrics* 15, 543–550.

[7] Gurland, J. and Mehta, J.S., 1969, Combination of Unbiased Estimates of the Mean Which Consider Inequality of Unknown Variances, *Journal of American Statistical Association* 64, 1042–1055.

[8] Kariya, T., Krishnaiah, P.R. and Rao, C.R., 1981, Inference on Parameters of Milti-variate Normal Population When Some Data is Missing, *Development in Statistics* 4,

137–184.

[9] Kariya, T., 1981, A Concept of Second Order Efficiencies and Its Application to a Missing Data Problem, To appear from *Proceedings of Pacific Area Statistical Conference.*

[10] Norwood, T.E. and Hinkelmann, K., 1977, Estimating the Common Mean of Several Normal Populations, *Annals* of *Statistics* 5, 1047–1050.

[11] Pfanzagl, J., 1980, Asymptotic Expansions in Parametric Statistical Theory, *Development in Statistics* (P.R. Krishnaiah, Editor), Vol. III.

[12] Rao, C.R., 1961, Asymptotic Efficiency and Limiting Information, *Proceedings Fourth Berkeley Symposium Mathematical Statistics Probability* 1, 531–546.

[13] Rao, C.R., 1963, Criteria of Estimation in Large Samples, *Sankhyā Ser. A* 25, 189–206.

[14] Shinozaki, N., 1978, A Note on Estimating the Common Mean of K Normal Distributions and the Stein Problem, *Communications in Statistics,* A7(15), 1421–1432.

[15] Sinha, B.K., 1979, Is the MLE of the Common Mean of Several Normal Populations Admissible? *Sankhyā (B).*

[16] Sinha, B.K. and Mouqadem, O., 1981, Estimation of the Common Mean of Two Univariate Normal Populations, *Technical Report* No. 81–07, Institute for Statistics and Applications, University of Pittsburgh.

[17] Zacks, S., 1966, Unbiased Estimation of the Common Mean Normal Distributions Based on Small Samples of Equal Size, *Journal of American Ststistical Association* 61, 467–476.