<table>
<thead>
<tr>
<th>Title</th>
<th>First, Second and Third Order Efficienciencies of the Estimators for a Common Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Kariya, Takeaki; Sinha, Bimal K.; Subramanyam, Kasala</td>
</tr>
<tr>
<td>Citation</td>
<td>Hitotsubashi Journal of Economics, 25(1): 61-69</td>
</tr>
<tr>
<td>Issue Date</td>
<td>1984-06</td>
</tr>
<tr>
<td>Type</td>
<td>Departmental Bulletin Paper</td>
</tr>
<tr>
<td>Text Version</td>
<td>publisher</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://doi.org/10.15057/7909">http://doi.org/10.15057/7909</a></td>
</tr>
</tbody>
</table>
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 \( (x_i, s_i^2) (i=1, \ldots, K) \) 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 \( (x_i, s_i^2) \) is \( N_i \). In this situation, the problem of estimating the common mean \( \mu \) 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 \( x_i - x_j (i \neq j) \). 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 Weiland (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 \( \mu \), 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.
\( \hat{p}_1(c_1, c_2) = \left[ \sum_{i=1}^{2} \frac{c_i \overline{x}_i / s_i^2}{\sum_{i=1}^{2} c_i / s_i^2} \right], \)

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

(2) Zacks (1966) estimator.

\( \hat{p}_2 = xG + \hat{p}_1(1 - G) \)

where \( x = (N_1 \overline{x}_1 + N_2 \overline{x}_2) / (N_1 + N_2) \), 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 \( \sigma_1^2 = \sigma_2^2 \).


\( \hat{p}_3 = c_3 \overline{x}_1 + (1 - c_3) \overline{x}_2 \) with \( c_3 = (c_1 s_2^2 + s_2^2) / ((c + a) s_1^2 + s_2^2) \).

In Gurland and Mehta (1969), \( a = 1 \) and \( c = 0.4 \) are proposed through a numerical comparison.


\( \hat{p}_4(a, c_1, c_2, c_3) = \overline{x}_1 + (\overline{x}_2 - \overline{x}_1) \left\{ \frac{av_i / [c_1 v_1 + c_2 v_2 + c_3 (\overline{x}_2 - \overline{x}_1)^2]}{c_1 v_1 + c_2 v_2 + c_3 (\overline{x}_2 - \overline{x}_1)^2} \right\}, \)

where \( v_i = s_i^2 / N_i \) \((i = 1, 2)\), \( a \) and \( c_i \)'s are constants. In Brown and Cohen (1974), \( \hat{p}_4(a_1, 1, (N_2 - 1) / (N_2 + 2), 1 / (N_2 + 2)) \) for \( N_2 \) small and \( \hat{p}_4(a_2, 1, 1, 0) \) for \( N_2 \) large are proposed, where \( a_1 \) and \( a_2 \) are certain constants.

(5) Cohen-Sackrowitz (1974) estimator \((N_1 = N_2)\).

\( \hat{p}_5 = [1 - c_0 G(s_1^2, s_2^2)] \overline{x}_1 + c_0 G(s_1^2, s_2^2) \overline{x}_2, \)

where \( c_0 = (N - 4) / (N + 2) \) for \( N = N_1 = N_2 \) even, \( c_0 = (N - 3)^2 / (N + 1)(N - 1) \) 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

\( \hat{p} = \phi \overline{x}_1 + (1 - \phi) \overline{x}_2, \)

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 \( \text{Var}(\hat{p}) \) with \( \text{Var}(\overline{x}_i) \) \((i = 1, \ldots , K)\), and have obtained necessary and sufficient conditions for which \( \text{Var}(\hat{p}) \leq \text{Var}(\overline{x}_i) \) for some \( i \) or all \( i \) with a particular form of \( \phi \) ([10], [14]). Sinha (1979) and Sinha and Mouqadem (1981)
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

\[
(2.1) \quad z = (z_1, z_2, z_3, z_4)' = (x_1, x_2, s_1^2, s_2^2)'
\]

which is a sufficient statistic for

\[
(2.2) \quad \eta = (\eta_1, \eta_2, \eta_3)' = (\mu, \sigma_1^2, \sigma_2^2)'.
\]

Clearly the expected value of \( z \) is

\[
(2.3) \quad \theta = \theta(\eta) = (\theta_1, \theta_2, \theta_3, \theta_4)' = (\mu, \mu, \sigma_1^2, \sigma_2^2)'.
\]

Here an estimator \( h(z) \) of \( \mu \) is called Fisher consistent if it satisfies

\[
(2.4) \quad h(\theta(\eta)) = \eta_k = \mu \text{ for all } \eta.
\]

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 \( O(1) \) 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

\[
(2.5) \quad h(z) = \mu + \sum_{i=1}^{4} h'(\theta) d_i + \frac{1}{2} \sum_{i,j} h''(\zeta + (1 - \lambda)\theta) d_i d_j,
\]

where \( d_i = z_i - \theta_i, \quad h' = \partial h/\partial z_i, \quad h'' = \partial^2 h/\partial z_i \partial z_j, \) and \( 0 \leq \lambda \equiv \zeta(z, \theta) \leq 1 \). Hence, adopting the MSE criterion yields

\[
(2.6) \quad E(h(z) - \mu)^2 = x'Ax + R,
\]

where

\[
(2.7) \quad x = x(\theta) = (h'(\theta), \ldots, h'(\theta))',
\]

\[
(2.8) \quad A = (E(d_i d_j)) = \text{diag} (\sigma_1^2/N_1, \sigma_2^2/N_2, 2\sigma_1^4/n_1, 2\sigma_2^4/n_2), \quad n_i = N_i - 1,
\]

and \( R \) is the remainder term. Since \( z \to \theta \) a.s. as \( N_1 \to \infty \) and \( N_2 \to \infty \) and since \( h'' \) is continuous, \( R = O((N_1 + N_2)^{-1}) \) (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 \( \eta_j \) to have the side condition \( Bx = c \), where

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

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 \( x = x_0 \), where
The minimum value of $x'Ax$ in $C$ is $J_4 = c'(BA^{-1}B')^{-1}c = a_1^2a_2^2/(N_1\sigma_2^2 + N_2\sigma_1^2)$.

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

\[ \phi(\theta(z)) = \frac{N_1\sigma_2^2}{(N_1\sigma_2^2 + N_2\sigma_1^2)} \]

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

Next, to define a concept of SOE, we assume for FOE estimators

(b) $h^{i,k}(z) = \partial^2 h/\partial z_i \partial z_j \partial z_k(z)$ exists, continuous, and of order $O(1)$ as $N_1 \to \infty$, $N_2 \to \infty$ and let

$C_1 = \{ h \in C \mid h \text{ satisfies (b), } x = x_0 \}$.

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

\[ h(z) = \mu + \sum_i x_{0i}d_i + \frac{1}{2} \sum_i \sum_j h^{i,j}(\theta)d_id_j + \frac{1}{6} \sum_i \sum_j \sum_k h^{i,j,k}(\lambda z + (1 - \lambda)\theta)d_id_jd_k \]

where $0 \leq \lambda \leq 1$, and $x_0 = (x_{01}, \ldots, x_{0k})'$ is given by (2.10). Based on Kariya (1981), we consider the measure

\[ \nu_2 = E[h(z) - \mu - \sum_i x_{0i}d_i]^2 = \bar{\xi}_2 + R_2, \]

where

\[ \bar{\xi}_2 = \frac{1}{4} \sum_j \sum_i \sum_k h^{i,j}(\theta)h^{j,i}(\theta)c_{ij,k}, \]

\[ c_{ij,k} = E(d_id_jd_k) \]

(2.13) and $R_2$ is the remainder term. It is noted that $c_{ij,k} = E(d_id_jd_k) = O(N_1^{-2})$ 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 $\bar{\xi}_2$ in $C_1$.

An intuitive rationale for this definition follows from (2.12) and (2.13) since $\nu_2$ measures the degree of concentration of $h(z)$ toward $\mu + \sum_i x_{0i}d_i$ and since $\bar{\xi}_2$ is the leading term of $\nu_2$.

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

\[ [E(h(z) - \mu)^2]^\frac{1}{2} \leq J_4^\frac{1}{2} + \nu_2^\frac{1}{2}, \]

where $J_4$ 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 $\xi_2$ of order $O((N_1 + N_2)^{-2})$ in $\nu_2$, we can control the MSE up to order $O((N_1 + N_2)^{-2})$ via (2.16). To carry out the minimization, let

\[ y = (h^{11}(\theta), \ldots, h^{22}(\theta), \ldots, h^{34}(\theta), h^{24}(\theta), h^{34}(\theta), h^{44}(\theta))', \]

\[ \alpha_3 = -N_1N_2\sigma_2^2/(N_1\sigma_2^2 + N_2\sigma_1^2)^2 \text{ and } \alpha_4 = N_1N_2\sigma_1^2/(N_1\sigma_2^2 + N_2\sigma_1^2)^2. \]
Theorem 2.2 An estimator $h$ in $C_1$ is SOE if and only if $y = y_0$, where
\begin{equation}
(2.19) \quad y_0 = (0, 0, \alpha_3, \alpha_4, 0, -\alpha_3, -\alpha_4, 0, 0, 0)^T.
\end{equation}
The minimum value of $\xi_2$ is given by
\begin{equation}
(2.20) \quad J_2 = 2 \left( \alpha_3^2 \sigma_1^2 N_1 + \alpha_4^2 \sigma_2^2 \sigma_1^2 N_1 N_2 + \alpha_4^2 \sigma_1^2 \sigma_2^2 N_1 N_2 + \alpha_4^2 \sigma_2^2 N_2^2 \right).
\end{equation}

Proof Since $h \in C_1$ satisfies $x = x_0$, differentiating $x = x_0$ with respect to $\gamma_j$ ($j = 1, 2, 3$) yields
\begin{equation}
(2.21) \quad \begin{cases}
\gamma_1^1 + \gamma_1^2 = 0, \\
\gamma_1^3 = \alpha_3, \\
\gamma_1^4 = \alpha_4, \\
\gamma_2^1 + \gamma_2^2 = 0, \\
\gamma_2^3 = -\alpha_3, \\
\gamma_2^4 = -\alpha_4, \\
\gamma_3^1 + \gamma_3^2 = 0, \\
\gamma_3^3 + \gamma_3^4 = 0, \\
\gamma_3^4 + \gamma_3^4 = 0.
\end{cases}
\end{equation}
Substituting (2.21) into $\xi_2$,
\begin{equation}
(2.22) \quad 4\xi_2 = \sum_{i, j} h^{ij}(\delta)^2 E(d_{ij}^2) + \sum_{i, j} h^{ij}(\delta)^2 E(d_{ij}^2) E(d_{ij}^2) + 2 \left( \sum_{i, j} h^{ij}(\delta)^2 E(d_{ij}^2) E(d_{ij}^2) \right)
\end{equation}
where $\delta = \gamma_1^1 = -\gamma_1^2 = \gamma_2^2$. This is clearly minimized if and only if $\delta = 0$, which together with (2.21) gives (2.17). The minimum value of $\xi_2$ is obtained by inserting $\delta = 0$ and the expected values $E(d_{ij}^2 d_{ij}^2)$ 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 $\phi'(\theta) = \phi'(\theta) = 0, \phi'(\theta) = \alpha_3$ and $\phi'(\theta) = \alpha_4$, where $\phi'(\theta) = \delta \phi'(\theta)$.

Proof Computing the derivatives of $\mu$ in (1.6) and equating them with $y_0$ in (2.19) yields the result.

Finally, we consider a condition for TOE. Let $C_2$ be the class of SOE estimators satisfying
\begin{equation}
(2.23) \quad h(z) = \mu + \sum_i x_i d_i + \frac{1}{2} \sum_i \sum_j y_{ij} d_i d_j + \frac{1}{6} \sum_i \sum_j \sum_k h^{ijk}(\theta) d_i d_j d_k
\end{equation}
where $0 \leq \lambda \leq 1$, and $y_0 = (y_{01}, \ldots, y_{04})$ is given by (2.19). Similar to the case of the definition of SOE, we consider the measure
\begin{equation}
(2.24) \quad \nu_3 = E[h(z) - \mu - \sum_i x_i d_i - \frac{1}{2} \sum_i \sum_j y_{ij} d_i d_j]^2 = \xi_3 + R_3,
\end{equation}
where
\begin{equation}
(2.25) \quad \xi_3 = \frac{1}{36} \sum_i \sum_j \sum_k \sum_l h^{ijk}(\lambda z + (1 - \lambda)\theta) d_i d_j d_k.
\end{equation}
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 $\xi_3$ 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) \[ E(h(z) - \mu)^2 \leq J_1 + J_2 + v_3, \]
where \( J_2 \) is given by (2.20). Hence minimizing the leading term \( \xi_3 \) of \( \nu_2 \), we can control the MSE up to \( O(N_1+N_2)^{-2} \) via (2.27). To minimize \( \xi_3 \), 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 \( h^{ij} = y_{ij} \) with respect to \( \mu, \sigma_1^2 \) and \( \sigma_2^2 \) gives the following set of conditions on \( h_{ijk} \)

\[
\begin{align*}
  h_{111} + h_{112} &= 0, & h_{113} &= 0, & h_{114} &= 0 \\
  h_{121} + h_{122} &= 0, & h_{123} &= 0, & h_{124} &= 0 \\
  h_{131} + h_{132} &= 0, & h_{133} &= \beta_1, & h_{134} &= \beta_2 \\
  h_{141} + h_{142} &= 0, & h_{143} &= \gamma_1, & h_{144} &= \gamma_2 \\
  h_{211} + h_{212} &= 0, & h_{213} &= 0, & h_{214} &= 0 \\
  h_{221} + h_{222} &= 0, & h_{223} &= 0, & h_{224} &= 0 \\
  h_{231} + h_{232} &= 0, & h_{233} &= -\beta_1, & h_{234} &= -\beta_2 \\
  h_{241} + h_{242} &= 0, & h_{243} &= -\gamma_1, & h_{244} &= -\gamma_2 \\
  h_{311} + h_{312} &= 0, & h_{313} &= 0, & h_{314} &= 0 \\
  h_{321} + h_{322} &= 0, & h_{323} &= 0, & h_{324} &= 0 \\
  h_{331} + h_{332} &= 0, & h_{333} &= 0, & h_{334} &= 0 \\
  h_{341} + h_{342} &= 0, & h_{343} &= 0, & h_{344} &= 0 \\
  h_{411} + h_{412} &= 0, & h_{413} &= 0, & h_{414} &= 0 \\
  h_{421} + h_{422} &= 0, & h_{423} &= 0, & h_{424} &= 0 \\
  h_{431} + h_{432} &= 0, & h_{433} &= 0, & h_{434} &= 0 \\
  h_{441} + h_{442} &= 0, & h_{443} &= 0, & h_{444} &= 0 \\
\end{align*}
\]

where \( \alpha_3 \) and \( \alpha_4 \) are as in (2.18) and

\[
(2.29) \quad \beta_1 = a_3 \sigma_1^2, \quad \beta_2 = a_3 \sigma_2^2, \quad \sigma_1 = \sigma_2^2, \quad \gamma_1 = \gamma_2.
\]

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 \( \xi_3 \), using the independence of \( d_i \)'s and minimizing \( \xi_3 \) with respect to \( x \) yields \( x = 0 \). 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 \( \phi_i(\theta) = N_1 \sigma_1^2 / (N_1 \sigma_1^2 + N_2 \sigma_2^2) \equiv \gamma \), and \( \phi_i \) is continuously twice differentiable. First, the Graybill-Deal type estimator \( \hat{\mu}_1 \) in (1.1) has \( \phi_1(z) = c_1 z_4 / (c_1 z_4 + c_2 z_3) \), 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

\[
(3.1) \quad \phi_2(z) = [N_1 / (N_1 + N_2)] G_2 + [N_2 / (N_1 + N_2)] (1 - G_2)
\]

where \( G_2 \) is the indicator function of \( a < z_4 / z_3 < a \). 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 \([a^- - \epsilon, a^- + \epsilon] \) and \([a^- - \epsilon, a^- + \epsilon] \), where \( \epsilon > 0 \) is arbitrarily small. Then \( \phi_2^* \) with \( G_2^* \) for \( G_2 \) in (3.1) does not satisfy \( \phi_2^*(\theta) = \gamma \), and so the estimator \( \hat{\mu}_2 \) with \( \phi_2^* \) is not FOE for any \( \epsilon > 0 \). Hence, the Zacks estimator is not approximately FOE. Thirdly, in the Gurland-Mehta estimator \( \hat{\mu}_3, \hat{\mu}_3 \) in (1.3) satisfies \( \phi_3(\theta) = \gamma \) if and only if \( c = 0 \) and \( a = N_2 / N_1 \), in which case \( \hat{\mu}_3 = \hat{\mu}_3^* \). Therefore, the Gurland-Mehta estimator with \( c > 0 \) is not FOE. Fourthly, the Brown-Cohen type estimator \( \hat{\mu}_4 \) in (1.4) has
\[ \phi_4(z) = [c_1 - a]N_1^{-1}z_3 + c_2N_2^{-1}z_4 + c_3(z_1 - z_2)^2] / [c_1N_1^{-1}z_3 + c_2N_2^{-1}z_4 + c_3(z_1 - z_2)^2] \]

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

\[ \phi_4^*(z) = \{N_1z_2 + b(z_1 - z_2)^2\} / [N_2z_2 + N_1z_4 + b(z_1 - z_2)^2] \]

is FOE, where \( b = c_3/a = 0 \) is arbitrary. It is noted that \( \hat{\mu}_4^*(\theta) + \hat{\mu}_4^* \) 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 \( \mu \):

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

and so the likelihood equation of \( \mu \) is given by the cubic polynomial

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

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 \( \hat{\mu}_8 \) satisfying

\[ l(\hat{\mu}_8(z); z) = \max \{ l(m_i(z); z) \mid i = 1, 2, 3 \} \]

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 \( m_i \) or \( m_3 \) maximizes it. Moreover, setting \( z = \theta \) in (3.5), one gets

\[ \left( \hat{\mu} - \mu \right)[N_1n_2N_2^{-1}z_4^2 + N_2n_1N_1^{-1}z_3^2 + N_1(\hat{\mu} - \mu)^2 + N_2(\hat{\mu} - \mu)^2] = 0 \]

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 \( \hat{\mu}_8^* \). Regarding (3.5) as an implicit function \( F(z, \hat{\mu}(z)) = 0 \) and differentiating it with respect to \( z \)'s at \( z = \theta \) produces the gradient vector of \( \hat{\mu} \) at \( z = \theta \):

\[ (3.8) \quad (\hat{\mu}^1(\theta), \ldots, \hat{\mu}^4(\theta)) = (N_1n_2N_2^{-1}z_4^2/D, N_2n_1N_1^{-1}z_3^2/D, O, O) \]

where

\[ D = N_1n_2N_2^{-1}z_4^2 + N_2n_1N_1^{-1}z_3^2 \quad \text{and} \quad \hat{\mu}^i = \partial \hat{\mu} / \partial z_i. \]

It is noted that this gradient is common for \( \hat{\mu} = m_c \), the roots of (3.5). Therefore by Theorem 2.1, \( \hat{\mu}_8^* \) 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 summarized as

**Theorem 3.1** Among the estimators \( \hat{\mu}_i \) \( (i = 1, \ldots, 7) \), the Graybill-Deal estimator \( \hat{\mu}_4^* = \hat{\mu}_4(N_1, N_2) \), 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 \( \hat{\mu}_4^* = \hat{\mu}_4^*(\theta) \), 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 \( \partial^3 \hat{\mu}_4^* / \partial z^3 \neq 0 \) and \( \partial^3 \hat{\mu}_7 / \partial z^3 \neq 0 \). 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 \( h^{111}[N_1\sigma_1^2 + N_2\sigma_2^2] + (h^1)^3[6(N_1 + N_2)] + (h^1)^2[-6(2N_2 + N_1)] + h^1(6N_2) = 0 \), from which we obtain \( h^{111} \neq 0 \) where \( h = \hat{\mu}_7 \). Similarly \( \partial^3 \hat{\mu}_4^* / \partial z^3 \neq 0 \) 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.

**REFERENCES**

137–184.


