Z. Wahrscheinlichkeitstheorie verw. Geb. 19, 29–42 (1971)

Maximum Likelihood Estimators for Ranked Means

EDWARD J. DUDEWICZ*

1. Introduction

Suppose that observations from populations $\pi_1, \ldots, \pi_k$ $(k \geqq 2)$ are normally distributed with unknown means $\mu_1, \ldots, \mu_k$ (respectively) and a common known variance σ^2. Let $\mu_{[1]} \leqq \cdots \leqq \mu_{[k]}$ denote the ranked means. Several ranking and selection procedures take n independent observations from each population, denote the sample mean of the n observations from π_i by $\bar{X}_i$ $(i=1, \ldots, k)$, and utilize the ranked sample means $\bar{X}_{[1]} \leqq \cdots \leqq \bar{X}_{[k]}$. (See [6] for details.) We assume throughout that both the numerical values of $\mu_1, \ldots, \mu_k$ and the pairings of the $\mu_{[1]}, \ldots, \mu_{[k]}$ with the populations $\pi_1, \ldots, \pi_k$ are completely unknown and consider problems of estimation of $\mu_{[i]}$ $(1 \leqq i \leqq k)$ by likelihood methods.

2. Maximum Likelihood Estimators

In maximum likelihood estimation of $\mu_1, \ldots, \mu_k$, we seek the *maximum likelihood estimators* (MLE's), those functions $\hat{\mu}_1, \ldots, \hat{\mu}_k$ (if such exist) such that the density of the observed statistics (whatever they may be) is maximized by setting $\mu_1 = \hat{\mu}_1, \ldots, \mu_k = \hat{\mu}_k$.

Our observed statistics, as stated above, are X_{ij} $(i=1, \ldots, k; j=1, \ldots, n)$, but since $\bar{X}_1, \ldots, \bar{X}_k$ are sufficient statistics we may take them as fundamental. Then, letting $\phi(\cdot)$ denote the standard normal density,

$$f_{\bar{X}_1, \ldots, \bar{X}_k}(x_1, \ldots, x_k) = (\sqrt{n}/\sigma)^k \, \phi\left(\frac{x_1 - \mu_1}{\sigma/\sqrt{n}}\right) \cdots \phi\left(\frac{x_k - \mu_k}{\sigma/\sqrt{n}}\right) \tag{1}$$

and (if $\mu_i \neq \mu_j$; $i \neq j$; $i, j = 1, \ldots, k$) the MLE's of $\mu_1, \ldots, \mu_k$ based on $\bar{X}_1, \ldots, \bar{X}_k$ exist and are uniquely

$$\hat{\mu}_1 = \bar{X}_1, \ldots, \hat{\mu}_k = \bar{X}_k. \tag{2}$$

(The restriction to MLE's based on $\bar{X}_1, \ldots, \bar{X}_k$ is a consequence of the general result that MLE's are functions only of sufficient statistics for a problem.) The problem of possible equalities among $\mu_{[1]}, \ldots, \mu_{[k]}$ is discussed below; similar results hold for the case of equalities among $\mu_1, \ldots, \mu_k$.

For the problem of finding an MLE of a 1-1 function $u(\mu_1, \ldots, \mu_k)$, it is well-known that (assuming the MLE of $\mu_1, \ldots, \mu_k$ exists) $u(\hat{\mu}_1, \ldots, \hat{\mu}_k) = \hat{u}$ (say) furnishes

* This research was supported by ONR contracts N00014-68A-0091 and Nonr 401(53), and by NSF grants GP-8958 and GP-7798. "This research is supported in whole or in part by The Center for Naval Analyses of the University of Rochester. Such support does not imply endorsement of the content by The Navy."

a solution, essentially because forcing $u=\hat{u}$ implies $\mu_1=\hat{\mu}_1, \ldots, \mu_k=\hat{\mu}_k$. If $u(\mu_1, \ldots, \mu_k)$ is not 1-1, i.e. if it is many-to-one, points other than $\mu_1=\hat{\mu}_1, \ldots, \mu_k=\hat{\mu}_k$ may also be implied by $u=\hat{u}$. In this case Zehna (1966) was the first to state explicitly a reason for picking only the "right" point $\mu_1=\hat{\mu}_1, \ldots, \mu_k=\hat{\mu}_k$ for attention (and thus for calling $\hat{u}$ an MLE). Berk (1967) gives a different justification for calling $\hat{u}$ an MLE.

From the above it is clear that, based on $\bar{X}_1, \ldots, \bar{X}_k$,

$$\hat{\mu}_{[i]}=\{i\text{-th smallest of } \bar{X}_1, \ldots, \bar{X}_k\}=\bar{X}_{[i]} \quad (i=1, \ldots, k) \tag{3}$$

is the Berk-Zehna-MLE of $\mu_{[1]}, \ldots, \mu_{[k]}$. Below we discuss the problem of MLE-type estimators of $(\mu_{[1]}, \ldots, \mu_{[k]})$ from another point of view. This method, Iterated-MLE's, is discussed further in [5].

Blumenthal and Cohen (1968a), (1968b) (who provided the author with preliminaries of their papers) studied, for a translation parameter family, (1) estimation of the pair $(\mu_{[1]}, \mu_{[2]})$ for the sum of squared errors as loss function and (2) estimation of $\mu_{[2]}$ for a squared error loss function.

Other work on the case $k=2$, in another formulation, was done by Katz (1963), who proposed to estimate $(\mu_{[1]}, \mu_{[2]})$ when one *knows* that (e.g.) π_1 is associated with $\mu_{[1]}$ and π_2 is associated with $\mu_{[2]}$. This work was done for binomial probabilities and also for normal means, with (e.g.) sum of squared error losses. (The fact that $(\bar{X}_1, \bar{X}_2)$ is not a totally desirable estimator may be seen intuitively from the fact that, although $\mu_{[1]} \leqq \mu_{[2]}$, in general $\{\bar{X}_1 > \bar{X}_2\}$ can occur with positive probability.) In our work one does not know the association of the $\mu_{[i]}$ with the π_j $(i, j=1, \ldots, k)$; see Robertson and Waltman (1968) for the case where one does.

Blumenthal and Cohen (1968), *who utilize the MLE of $\mu_{[2]}$ found below*, desired their estimate to be symmetric in $\bar{X}_1, \bar{X}_2$; in order to force this they based their estimate on the maximal invariant $\bar{X}_{[1]}, \bar{X}_{[2]}$. Note, however, that in order to obtain symmetry in $\bar{X}_1, \bar{X}_2$ (and certain other invariance conditions) in one's estimator, one *need* not go to $\bar{X}_{[1]}, \bar{X}_{[2]}$ (at least for the normal case; see (3)). Note that (although the Berk-Zehna-MLE of $\mu_{[2]}$ based on $\bar{X}_1, \bar{X}_2$ is $\bar{X}_{[2]}$) the MLE of $\mu_{[2]}$ based on $\bar{X}_{[1]}, \bar{X}_{[2]}$ is not. In [5] we give additional justification for basing the MLE on $\bar{X}_{[1]}, \bar{X}_{[2]}$.

We will now consider the general case in which it is desired to find the MLE's of $\mu_{[1]}, \ldots, \mu_{[k]}$ based on $\bar{X}_{[1]}, \ldots, \bar{X}_{[k]}$. The likelihood function is given in (A.1), and (due to its symmetry in $\mu_{[1]}, \ldots, \mu_{[k]}$) if $\hat{\mu}_{[1]}, \ldots, \hat{\mu}_{[k]}$ is an MLE then so is any permutation of it (so that it is not necessarily the case that $\hat{\mu}_{[1]} \leqq \cdots \leqq \hat{\mu}_{[k]}$). In order to eliminate such undesirable occurrences, we require a consistency condition.

(4) *Consistency Criterion.* Among the (at most $k!$) permutation MLE's which any $\hat{\mu}_{[1]}, \ldots, \hat{\mu}_{[k]}$ which maximizes (A.1) provides, only the one with $\hat{\mu}_{[1]} \leqq \cdots \leqq \hat{\mu}_{[k]}$ will be called an MLE.

From (A.1) and the form of $\phi(\cdot)$, it is clear that we may restrict our search for the maximum to $\mu_{[1]}, \ldots, \mu_{[k]}$ such that $x_1 \leqq \{\mu_{[1]}, \ldots, \mu_{[k]}\} \leqq x_k$. By (4) we need only consider the case $\mu_{[1]} \leqq \cdots \leqq \mu_{[k]}$, and not all $k!$ (fewer if there are any equalities) orderings. It is well-known that in such a case the maximum must occur at $\mu_{[1]}, \ldots, \mu_{[k]}$ such that

$$\frac{\partial f_{\bar{X}_{[1]},\ldots,\bar{X}_{[k]}}(x_1,\ldots,x_k)}{\partial \mu_{[i]}}=0 \qquad (i=1,\ldots,k); \tag{5}$$

any point $\mu_{[1]}, \ldots, \mu_{[k]}$ (which depends on the values of $x_1, \ldots, x_k$) where (5) holds is called a *critical point*.

In taking the derivatives (5), the results depend on how many, and which, of the $k-1$ inequalities $\mu_{[1]} \leqq \cdots \leqq \mu_{[k]}$ are equalities. There are thus 2^{k-1} mutually exclusive and exhaustive cases, say

$$\Omega_0 \equiv \{(\mu_1, \ldots, \mu_k)\colon \mu_i \in \Re\,(i=1,\ldots,k)\} = \Omega_{(1)} + \Omega_{(2)} + \cdots + \Omega_{(2^{k-1})} \tag{6}$$

where the $\Omega_{(i)}$ are disjoint, $\Omega_{(1)} = \Omega(\neq) \equiv \{\mu\colon \mu_{[1]} \neq \mu_{[2]} \neq \cdots \neq \mu_{[k]}\}$, and the $\Omega_{(i)}$ $(i=2,\ldots,2^{k-1})$ are the other $2^{k-1}-1$ cases in some order. Fix any i $(2 \leqq i \leqq 2^{k-1})$ and suppose that some $\mu^* \in \Omega_{(i)}$ solves the system (5) (i.e., is a critical point *when* the derivatives are taken for $\mu \in \Omega_{(i)}$). *Then* it is easy to verify (using (A.1)) that μ^* is a critical point of system (5) when derivatives are taken for $\mu \in \Omega_{(1)}$. We thus have the

(7) **Theorem.** *Any critical point for our problem is a solution of system* (5) *with derivatives taken for* $\mu \in \Omega(\neq)$, *provided only that we allow boundary points (i.e., points of* $\Omega_{(2)} + \cdots + \Omega_{(2^{k-1})}$*) to be considered solutions.*

To completely justify calling the boundary points included in Theorem (7) critical points, one should show that any such point is a solution of system (5) when derivatives are taken for μ in its $\Omega_{(i)}$; this is clear from the proof of Theorem (7).

Now (taking derivatives when $\mu_{[1]} < \cdots < \mu_{[k]}$) system (5) is

$$\sum_{\beta \in S_k} (\sqrt{n}/\sigma)^k\, \phi\left(\frac{x_{\beta(1)}-\mu_{[1]}}{\sigma/\sqrt{n}}\right) \cdots \phi\left(\frac{x_{\beta(k)}-\mu_{[k]}}{\sigma/\sqrt{n}}\right) \frac{x_{\beta(i)}-\mu_{[i]}}{\sigma/\sqrt{n}}\,(\sqrt{n}/\sigma)=0 \qquad (i=1,\ldots,k), \tag{8}$$

where S_k is the symmetric group on k elements, or

$$\mu_{[i]} = \frac{\sum_{\beta \in S_k} x_{\beta(i)}\, \phi\left(\frac{x_{\beta(1)}-\mu_{[1]}}{\sigma/\sqrt{n}}\right) \cdots \phi\left(\frac{x_{\beta(k)}-\mu_{[k]}}{\sigma/\sqrt{n}}\right)}{\sum_{\beta \in S_k} \phi\left(\frac{x_{\beta(1)}-\mu_{[1]}}{\sigma/\sqrt{n}}\right) \cdots \phi\left(\frac{x_{\beta(k)}-\mu_{[k]}}{\sigma/\sqrt{n}}\right)} \qquad (i=1,\ldots,k) \tag{9}$$

or

$$\frac{\mu_{[j]}}{\mu_{[i]}} = \frac{\sum_{\beta \in S_k} x_{\beta(j)}\, \phi\left(\frac{x_{\beta(1)}-\mu_{[1]}}{\sigma/\sqrt{n}}\right) \cdots \phi\left(\frac{x_{\beta(k)}-\mu_{[k]}}{\sigma/\sqrt{n}}\right)}{\sum_{\beta \in S_k} x_{\beta(i)}\, \phi\left(\frac{x_{\beta(1)}-\mu_{[1]}}{\sigma/\sqrt{n}}\right) \cdots \phi\left(\frac{x_{\beta(k)}-\mu_{[k]}}{\sigma/\sqrt{n}}\right)} \qquad (i,j=1,\ldots,k;\ i<j). \tag{10}$$

(11) **Theorem.** $(\hat{\mu}_{[1]}, \ldots, \hat{\mu}_{[k]}) = (\bar{x}, \ldots, \bar{x})$ *with* $\bar{x} = \frac{x_1 + \cdots + x_k}{k}$ *is a critical point.*

Proof. It is clear that this is so from system (9).

We will now investigate the nature of this critical point. For $i, j = 1, \ldots, k$, for $x_1 \leqq \cdots \leqq x_k$,

$$\frac{\partial^2}{\partial \mu_{[i]} \partial \mu_{[j]}} f_{\bar{X}_{[1]}, \ldots, \bar{X}_{[k]}}(x_1, \ldots, x_k)$$

$$= \begin{cases} \sum_{\beta \in S_k} \left(\frac{\sqrt{n}}{\sigma}\right)^{k+2} \phi\left(\frac{x_{\beta(1)} - \mu_{[1]}}{\sigma/\sqrt{n}}\right) \cdots \phi\left(\frac{x_{\beta(k)} - \mu_{[k]}}{\sigma/\sqrt{n}}\right) \\ \qquad \cdot \frac{x_{\beta(i)} - \mu_{[i]}}{\sigma/\sqrt{n}} \frac{x_{\beta(j)} - \mu_{[j]}}{\sigma/\sqrt{n}}, & i \neq j \\ \sum_{\beta \in S_k} \left(\frac{\sqrt{n}}{\sigma}\right)^{k+2} \phi\left(\frac{x_{\beta(1)} - \mu_{[1]}}{\sigma/\sqrt{n}}\right) \cdots \phi\left(\frac{x_{\beta(k)} - \mu_{[k]}}{\sigma/\sqrt{n}}\right) \\ \qquad \cdot \left[\left(\frac{x_{\beta(i)} - \mu_{[i]}}{\sigma/\sqrt{n}}\right)^2 - 1\right], & i = j. \end{cases} \tag{12}$$

Thus, for the matrix $Q = (d_{ij})$ of evaluations of (12) at $(\bar{x}, \ldots, \bar{x})$ we find

$$d_{ij} = \left(\frac{\sqrt{n}}{\sigma}\right)^{k+4} \left[\prod_{l=1}^{k} \phi\left(\frac{x_l - \bar{x}}{\sigma/\sqrt{n}}\right)\right] \cdot \begin{cases} \sum_{\beta \in S_k} (x_{\beta(i)} - \bar{x})(x_{\beta(j)} - \bar{x}), & i \neq j \\ \sum_{\beta \in S_k} \left[(x_{\beta(i)} - \bar{x})^2 - \frac{\sigma^2}{n}\right], & i = j \end{cases}$$

$$= (k-2)! \left(\frac{\sqrt{n}}{\sigma}\right)^{k+4} \left[\prod_{l=1}^{k} \phi\left(\frac{x_l - \bar{x}}{\sigma/\sqrt{n}}\right)\right] \cdot \begin{cases} \sum_{\substack{i,j=1 \\ i \neq j}}^{k} (x_i - \bar{x})(x_j - \bar{x}), & i \neq j \\ (k-1) \sum_{i=1}^{k} (x_i - \bar{x})^2 - k(k-1)\frac{\sigma^2}{n}, & i = j \end{cases}$$

$$= (k-2)! (\sqrt{n}/\sigma)^{k+4} \left[\prod_{l=1}^{k} \phi\left(\frac{x_l - \bar{x}}{\sigma/\sqrt{n}}\right)\right] \cdot \begin{cases} \operatorname{cov}(R, S)\, k(k-1), & i \neq j \\ k(k-1) \operatorname{var}(R) - k(k-1)(\sigma^2/n), & i = j \end{cases} \tag{13}$$

$$= k! (\sqrt{n}/\sigma)^{k+4} \left[\prod_{l=1}^{k} \phi\left(\frac{x_l - \bar{x}}{\sigma/\sqrt{n}}\right)\right] \cdot \begin{cases} \operatorname{cov}(R, S), & i \neq j \\ \operatorname{var}(R) - \sigma^2/n, & i = j, \end{cases}$$

where R and S are numbers selected at random (without replacement) from $\{x_1, \ldots, x_k\}$. If we let

$$\begin{aligned} c &= c(x_1, \ldots, x_k) = k! (\sqrt{n}/\sigma)^{k+4} \left[\prod_{l=1}^{k} \phi\left(\frac{x_l - \bar{x}}{\sigma/\sqrt{n}}\right)\right] \\ d_1 &= \operatorname{cov}(R, S) \cdot c \\ d_0 &= (\operatorname{var}(R) - \sigma^2/n) \cdot c, \end{aligned} \tag{14}$$

then $d_{ij}=d_1$ $(i\neq j)$ and $d_{ij}=d_0$ $(i=j)$. Now, if we find the eigenvalues of Q we can determine the nature of the critical point $(\bar{x},\ldots,\bar{x})$. Now

(15)
$$|Q-\lambda I|=\det\begin{bmatrix} d_0-\lambda & d_1 & d_1\ldots & d_1 & d_1 \\ d_1 & d_0-\lambda & d_1\ldots & d_1 & d_1 \\ \vdots & & \ddots & & \\ d_1 & d_1 & d_1\ldots & d_0-\lambda & d_1 \\ d_1 & d_1 & d_1\ldots & d_1 & d_0-\lambda \end{bmatrix}$$
$$=(d_0-\lambda-d_1)^{k-1}\left(d_0-\lambda+(k-1)\,d_1\right)$$

where we have subtracted the last column from all others, added all rows to the last row, and taken minors. Thus, the k eigenvalues of Q are

(16)
$$\lambda_1=\cdots=\lambda_{k-1}=d_0-d_1$$
$$\lambda_k=d_0+(k-1)\,d_1$$

and we have the

(17) **Theorem.** *The nature of the critical point* $(\bar{x},\ldots,\bar{x})$ *is:*

(i) *relative minimum if* $-\dfrac{d_0}{k-1}<d_1<d_0$;

(ii) *relative maximum if* $d_0<d_1<-\dfrac{d_0}{k-1}$;

(iii) *undecided if either*: (a) $-\dfrac{d_0}{k-1}\leqq d_1=d_0$ *or* $-\dfrac{d_0}{k-1}=d_1\leqq d_0$

or: (b) $d_0=d_1\leqq\dfrac{-d_0}{k-1}$ *or* $d_0\leqq d_1=\dfrac{-d_0}{k-1}$;

(iv) *saddle point if* $d_1<\min\left(d_0,-\dfrac{d_0}{k-1}\right)$ *or if* $d_1>\max\left(d_0,-\dfrac{d_0}{k-1}\right)$.

Graphically,

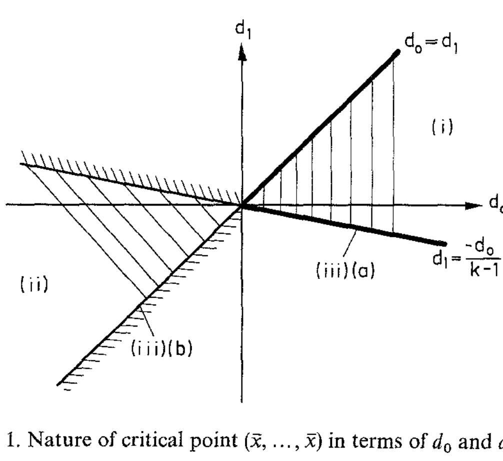

Fig. 1. Nature of critical point $(\bar{x},\ldots,\bar{x})$ in terms of d_0 and d_1

The method of determinants can also be used to prove Theorem (17) (since the required determinants can be evaluated as in (15)), but is cumbersome.

We now wish to investigate the nature (asymptotic as $n \to \infty$ as well as small sample) of the critical point $(\bar{x}, \ldots, \bar{x})$. Let $\chi_a^2(b)$ denote a non-central chi-square r.v. with "a" degrees of freedom and noncentrality "b".

(18) **Theorem.** I. $P_\mu[(\bar{X}, \ldots, \bar{X})$ *is a relative minimum, or undecided*$] = 0$.

II. $P_\mu[(\bar{X}, \ldots, \bar{X})$ *is a saddle point*$] = P_\mu\left[\chi_{k-1}^2\left(\frac{1}{2}\frac{k\,n}{\sigma^2}\operatorname{Var}(M)\right) > k-1\right]$; *otherwise* $(\bar{X}, \ldots, \bar{X})$ *is a relative maximum. This probability does not depend on n if* $\mu_{[1]} = \cdots = \mu_{[k]}$.

III. *As* $n \to \infty$, $P_\mu[(\bar{X}, \ldots, \bar{X})$ *is a saddle point*$] \to 1$ *unless* $\mu_{[1]} = \cdots = \mu_{[k]}$ (*in which case it is constant as given in* II).

Proof. I. Case (i) or case (iii)(a) of Theorem (17) holds iff (see (13))

$$-\frac{d_0}{k-1} \leqq d_1 \leqq d_0,$$

i.e. iff

$$-\frac{1}{k-1}(\operatorname{Var}(R) - \sigma^2/n) \leqq \operatorname{cov}(R, S) \leqq \operatorname{Var}(R) - \sigma^2/n,$$

i.e. iff (since $\operatorname{Var}(R) > 0$ w.p. 1)

$$-\frac{1}{k-1} + \frac{\sigma^2/n}{(k-1)\operatorname{Var}(R)} \leqq \rho(R, S) \leqq 1 - \frac{\sigma^2/n}{\operatorname{Var}(R)}. \tag{19}$$

Since (w.p. 1) $\rho(R, S) = \frac{-1}{k-1}$, w.p. 1 Eq. (19) fails to hold. W.p. 1 case (iii)(b) fails to hold since (for it to hold) at least one of the inequalities in (19) must be an equality; this occurs w.p. 0.

II. As in I, it can be seen that case (ii) holds iff

$$1 - \frac{\sigma^2/n}{\operatorname{Var}(R)} < \rho(R, S) < \frac{-1}{k-1} + \frac{\sigma^2/n}{(k-1)\operatorname{Var}(R)}. \tag{20}$$

Since the r.h.s. of (20) holds w.p. 1, case (ii) holds iff

$$1 - \frac{\sigma^2/n}{\operatorname{Var}(R)} < \frac{-1}{k-1}, \tag{21}$$

i.e. iff $\operatorname{Var}(R)\frac{k}{k-1} < \sigma^2/n$; otherwise (by I) case (iv) must hold. Now from Graybill (1961), p. 88 (Theorem 4.20), p. 91 (Problem 4.24), $\operatorname{Var}(R) = (1/k)\sum_{i=1}^{k}(\bar{X}_i - \bar{X})^2$ is $(\sigma^2/(n\,k))\,\chi_{k-1}^2(\lambda)$ with

$$\begin{aligned}\lambda &= \frac{1}{2}\frac{k\,n}{\sigma^2}\left(\frac{\sum \mu_i^2}{k} - \frac{(\sum \mu_i)^2}{k^2}\right)\\ &= \frac{1}{2}\frac{k\,n}{\sigma^2}\operatorname{Var}(M),\end{aligned} \tag{22}$$

where M is a number selected at random from $\{\mu_1, \ldots, \mu_k\}$. Thus,

$$(23)\quad \begin{aligned} P_\mu[(\bar{X}, \ldots, \bar{X}) \text{ is a relative maximum}] &= P_\mu\left[\operatorname{Var}(R) > \frac{k-1}{k}\,\frac{\sigma^2}{n}\right] \\ &= P_\mu\left[\frac{\sigma^2}{n\,k}\,\chi^2_{k-1}(\lambda) > \frac{k-1}{k}\,\frac{\sigma^2}{n}\right] = P_\mu\left[\chi^2_{k-1}\left(\frac{1}{2}\,\frac{k\,n}{\sigma^2}\operatorname{Var}(M)\right) > k-1\right]. \end{aligned}$$

III. This follows from II.

Note that even when $(\bar{X}, \ldots, \bar{X})$ is a relative maximum it is not necessarily an absolute one (which it would be if, e.g., the system had no other solution). Below we will find reason to believe that the maximum is "near" $(\hat{\mu}_{[1]}, \ldots, \hat{\mu}_{[k]}) = (\bar{X}_{[1]}, \ldots, \bar{X}_{[k]})$.

For *the case* $k=2$, Theorem (17) shows (after some reduction) that $(\bar{x}, \bar{x})$ is

$$(24)\quad \begin{array}{ll} \text{a relative maximum} & \text{iff } (x_1-x_2)^2 < 2\sigma^2/n \\ \text{undecided (negative semi-definite)} & \text{iff } (x_1-x_2)^2 = 2\sigma^2/n \\ \text{a saddle point} & \text{iff } (x_1-x_2)^2 > 2\sigma^2/n. \end{array}$$

The limiting results of Theorem (18) can, for the case $k=2$, be obtained using (24).

We will now seek the MLE (for $k \geqq 2$): We may (without loss) choose our estimator to be of the form

$$(25)\quad \begin{aligned} \hat{\mu}_{[1]} &= x_1 + a_1(x_1, \ldots, x_k) \\ &\vdots \\ \hat{\mu}_{[k]} &= x_k + a_k(x_1, \ldots, x_k). \end{aligned}$$

As noted following (4), we may restrict ourselves without loss to

$$x_1 \leqq \{\hat{\mu}_{[1]}, \ldots, \hat{\mu}_{[k]}\} \leqq x_k,$$

from which it follows that we have

$$(26)\quad \begin{aligned} &0 \leqq a_1 \\ -(x_i - x_1) \leqq\ &a_i \leqq (x_k - x_i) \qquad (i=1, \ldots, k) \\ &a_k \leqq 0. \end{aligned}$$

Let (for $1 \leqq l \leqq k$; $i=1, \ldots, k$)

$$(27)\quad \begin{aligned} A_l(i) &= \sum_{\substack{\beta \in S_k \\ \beta(i)=l}} \phi\left(\frac{x_{\beta(1)} - x_1 - a_1}{\sigma/\sqrt{n}}\right) \ldots \phi\left(\frac{x_{\beta(k)} - x_k - a_k}{\sigma/\sqrt{n}}\right) \\ A &= \sum_{\beta \in S_k} \phi\left(\frac{x_{\beta(1)} - x_1 - a_1}{\sigma/\sqrt{n}}\right) \ldots \phi\left(\frac{x_{\beta(k)} - x_k - a_k}{\sigma/\sqrt{n}}\right). \end{aligned}$$

Then (note that, for any $1 \leqq i \leqq k$, $A = A_1(i) + \cdots + A_k(i)$) from system (9) we find that $a_1, \ldots, a_k$ must satisfy the system

$$(28)\quad (x_i + a_i)\,A = x_1\,A_1(i) + \cdots + x_k\,A_k(i) \qquad (i=1, \ldots, k).$$

If we add the terms of (28) over $i=1, \ldots, k$, we obtain (since $A=A_l(1)+\cdots+A_l(k)$ for $l=1, \ldots, k$)

$$A(x_1+\cdots+x_k)+(a_1+\cdots+a_k)A=A(x_1+\cdots+x_k),$$

or (since $A>0$) $a_1+\cdots+a_k=0$. Thus, we have the

(29) **Theorem.** *For $k\geqq 2$, the MLE is given by*

$$\hat{\mu}_{[1]}=\bar{X}_{[1]}+a_1(\bar{X}_{[1]}, \ldots, \bar{X}_{[k]}), \ldots, \hat{\mu}_{[k]}=\bar{X}_{[k]}+a_k(\bar{X}_{[1]}, \ldots, \bar{X}_{[k]}),$$

where $a_1, \ldots, a_k$ are some solution of system (28) *and must satisfy*

$$-(x_i-x_1)\leqq a_i\leqq(x_k-x_i) \qquad (i=1, \ldots, k)$$

and

$$a_1+\cdots+a_k=0.$$

(30) **Theorem.** *For $i, j=1, \ldots, k$, if $a_j\neq 0$ then*

$$a_i=a_j\frac{d_{1i}A_1(i)+\cdots+d_{ki}A_k(i)}{d_{1j}A_1(j)+\cdots+d_{kj}A_k(j)}$$

where $d_{ij}=x_i-x_j=-d_{ji}$ $(i, j=1, \ldots, k)$.

Proof. System (28) is equivalent to the system

$$a_i\sum_{\beta\in S_k}\phi\left(\frac{x_{\beta(1)}-x_1-a_1}{\sigma/\sqrt{n}}\right)\cdots\phi\left(\frac{x_{\beta(k)}-x_k-a_k}{\sigma/\sqrt{n}}\right)$$
$$=\sum_{\beta\in S_k}(x_{\beta(i)}-x_i)\,\phi\left(\frac{x_{\beta(1)}-x_1-a_1}{\sigma/\sqrt{n}}\right)\cdots\phi\left(\frac{x_{\beta(k)}-x_k-a_k}{\sigma/\sqrt{n}}\right) \qquad (i=1, \ldots, k),$$

or (substituting the d_{ij}'s)

$$a_i(A_1(i)+\cdots+A_k(i))=d_{1i}A_1(i)+\cdots+d_{ki}A_k(i) \qquad (i=1, \ldots, k).$$

Thus, the theorem follows. (Note that the denominator $d_{1j}A_1(j)+\cdots+d_{kj}A_k(j)$ is zero iff $a_j=0$.)

(31) **Lemma.** *For the case $k=2$, $a_1=-a_2$. Also, $0\leqq a_1\leqq x_2-x_1$.*

Proof. From Theorem (30),

$$a_1=a_2\frac{d_{11}A_1(1)+d_{21}A_2(1)}{d_{12}A_1(2)+d_{22}A_2(2)}=a_2\frac{d_{21}A_2(1)}{d_{12}A_1(2)}=-a_2\frac{A_2(1)}{A_1(2)}=a_2.$$

The theorem follows from Theorem (29).

(32) **Lemma.** *Let $d=x_2-x_1\geqq 0$. Then the MLE for $k=2$ is given by*

$$\hat{\mu}_{[1]}=\bar{X}_{[1]}+a_1(\bar{X}_{[1]}, \bar{X}_{[2]}), \qquad \hat{\mu}_{[2]}=\bar{X}_{[2]}-a_1(\bar{X}_{[1]}, \bar{X}_{[2]})$$

where a_1 *is some root of*

$$d = a_1\left(1 + e^{\frac{d^2 - 2a_1 d}{\sigma^2/n}}\right)$$

and $0 \leqq a_1 \leqq d$.

Proof. By Lemma (31) we must have $0 \leqq a_1 = -a_2 \leqq d$. Then by Theorem (29), the MLE must be of the form given where a_1 is some root of the system (28):

$$\begin{cases}(x_1 + a_1)A = x_1 A_1(1) + x_2 A_2(1)\\ (x_2 - a_1)A = x_1 A_1(2) + x_2 A_2(2),\end{cases}$$

$$\begin{cases}x_1 A_2(1) + a_1 A = x_2 A_2(1)\\ x_2 A_1(2) - a_1 A = x_1 A_1(2),\end{cases}$$

$$a_1 A = d A_2(1) = d A_1(2),$$

$$a_1 A = d A_1(2),$$

$$a_1 = d\frac{A_1(2)}{A_1(2) + A_2(2)},$$

$$a_1 = \frac{d}{1 + \dfrac{A_2(2)}{A_1(2)}}.$$

Now

$$A_1(2) = \sum_{\substack{\beta \in S_2\\ \beta(2)=1}} \phi\left(\frac{x_{\beta(1)} - x_1 - a_1}{\sigma/\sqrt{n}}\right)\phi\left(\frac{x_{\beta(2)} - x_2 - a_2}{\sigma/\sqrt{n}}\right) = \phi\left(\frac{d - a_1}{\sigma/\sqrt{n}}\right)\phi\left(\frac{a_1 - d}{\sigma/\sqrt{n}}\right)$$

$$= \frac{1}{2\pi\sigma^2/n} e^{-\frac{(d-a_1)^2}{\sigma^2/n}};$$

$$A_2(2) = \sum_{\substack{\beta \in S_2\\ \beta(2)=2}} \phi\left(\frac{x_{\beta(1)} - x_1 - a_1}{\sigma/\sqrt{n}}\right)\phi\left(\frac{x_{\beta(2)} - x_2 - a_2}{\sigma/\sqrt{n}}\right) = \phi\left(\frac{-a_1}{\sigma/\sqrt{n}}\right)\phi\left(\frac{a_1}{\sigma/\sqrt{n}}\right)$$

$$= \frac{1}{2\pi\sigma^2/n} e^{-\frac{a_1^2}{\sigma^2/n}}.$$

Thus,

$$\frac{A_2(2)}{A_1(2)} = e^{-\frac{a_1^2}{\sigma^2/n} + \frac{(d-a_1)^2}{\sigma^2/n}} = e^{\frac{d^2 - 2a_1 d}{\sigma^2/n}}$$

and the lemma follows.

(33) **Lemma.** *For fixed* d *and* $0 \leqq a_1 \leqq d$, *the roots of*

(34) $$d = a_1\left(1 + e^{\frac{d^2 - 2a_1 d}{\sigma^2/n}}\right)$$

are (1) $a_1 = d/2$, *and* (2) $a_1 = \frac{d}{2} + \frac{\varepsilon_0}{2d}\sigma^2/n$ *if* $d > \sqrt{2}\,\sigma/\sqrt{n}$. *Here* ε_0 *is either of the two solutions of*

(35) $$d^2 n/\sigma^2 = \varepsilon \coth(\varepsilon/2).$$

Proof. First, $a_1 = d/2$ is seen to satisfy (34). Now, suppose there is another solution of (34), say (without loss of generality)

$$a_1 = d/2 + \frac{\varepsilon}{2d}\sigma^2/n$$

with $-d^2 n/\sigma^2 \leqq \varepsilon \leqq d^2 n/\sigma^2$ (since $0 \leqq a_1 \leqq d$), $\varepsilon \neq 0$. Substituting in (34), we find ε must satisfy

$$(36)\quad \begin{aligned} d &= \left(\frac{d}{2} + \frac{\varepsilon}{2d}\frac{\sigma^2}{n}\right)\left(1 + e^{\frac{d^2 - 2\left(\frac{d}{2} + \frac{\varepsilon}{2d}\frac{\sigma^2}{n}\right)d}{\sigma^2/n}}\right) = \left(\frac{d}{2} + \frac{\varepsilon}{2d}\frac{\sigma^2}{n}\right)(1 + e^{-\varepsilon}) \\ &= \frac{1}{2}\left\{d + de^{-\varepsilon} + \frac{\varepsilon}{d}\frac{\sigma^2}{n} + \frac{\varepsilon}{d}\frac{\sigma^2}{n}e^{-\varepsilon}\right\}, \end{aligned}$$

or

$$d^2 = d^2 e^{-\varepsilon} + \varepsilon\frac{\sigma^2}{n} + \varepsilon\frac{\sigma^2}{n}e^{-\varepsilon},$$

or (since $\varepsilon \neq 0 \Rightarrow 1 - e^{-\varepsilon} \neq 0$)

$$d^2 = \varepsilon\frac{\sigma^2}{n}\frac{1 + e^{-\varepsilon}}{1 - e^{-\varepsilon}} = \varepsilon\frac{\sigma^2}{n}\frac{e^{\varepsilon/2} + e^{-\varepsilon/2}}{e^{\varepsilon/2} - e^{-\varepsilon/2}} = \varepsilon\frac{\sigma^2}{n}\coth(\varepsilon/2).$$

Since $\coth(-z) = -\coth(z)$, $\varepsilon\coth(\varepsilon/2)$ is an even function. Now,

$$\lim_{\varepsilon\to 0}\varepsilon\coth(\varepsilon/2) = \lim_{\varepsilon\to 0}(1 + e^{-\varepsilon})\cdot\lim_{\varepsilon\to 0}\frac{\varepsilon}{1 - e^{-\varepsilon}} = 2\lim_{\varepsilon\to 0}\frac{1}{e^{-\varepsilon}} = 2.$$

Since

$$\begin{aligned} \frac{\partial}{\partial\varepsilon}[\varepsilon\coth(\varepsilon/2)] &= \coth(\varepsilon/2) - (\varepsilon/2)\operatorname{csch}^2(\varepsilon/2) \\ &= \frac{\cosh(\varepsilon/2)}{\sinh(\varepsilon/2)} - (\varepsilon/2)\frac{1}{\sinh^2(\varepsilon/2)} \\ &= \frac{1}{\sinh(\varepsilon/2)}\left\{\cosh(\varepsilon/2) - \frac{\varepsilon/2}{\sinh(\varepsilon/2)}\right\}, \end{aligned}$$

the facts $\sinh(\varepsilon/2) > 0$ if $\varepsilon > 0$ and

$$\begin{aligned} \cosh(\varepsilon/2) - \frac{\varepsilon/2}{\sinh(\varepsilon/2)} &= \frac{1}{\sinh(\varepsilon/2)}[\sinh(\varepsilon/2)\cosh(\varepsilon/2) - \varepsilon/2] \\ &= \frac{1}{\sinh(\varepsilon/2)}\left[\frac{\sinh(\varepsilon)}{2} - \varepsilon/2\right] \\ &= \frac{1}{2\sinh(\varepsilon/2)}\left[\varepsilon + \frac{\varepsilon^3}{3!} + \frac{\varepsilon^5}{5!} + \cdots - \varepsilon\right] \\ &= \frac{1}{2\sinh(\varepsilon/2)}\left[\frac{\varepsilon^3}{3!} + \frac{\varepsilon^5}{5!} + \frac{\varepsilon^7}{7!} + \cdots\right] > 0 \end{aligned}$$

imply that $\frac{\partial}{\partial \varepsilon}[\varepsilon \coth(\varepsilon/2)] > 0$. Combining the above information, we may plot Fig. 2.

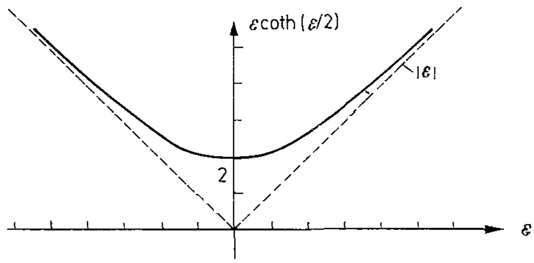

Fig. 2. The function ε coth (ε/2)

Since $\coth(x) > 1$ for $x > 0$, the range of $\varepsilon \coth(\varepsilon/2)$ will be $\left[2, d^2 \frac{n^+}{\sigma^2}\right]$ when ε is in $[-d^2 n/\sigma^2, d^2 n/\sigma^2]$. Thus, there will be two additional solutions if $d^2 n/\sigma^2 > 2$ and none if $d^2 n/\sigma^2 \leqq 2$.

Note that $a_1 = 0$ corresponds to the estimator (x_1, x_2); $a_1 = d/2$ corresponds to $(\bar{x}, \bar{x})$; and $a_1 = d$ corresponds to (x_2, x_1). Consistency Criterion (4) rules out values $a_1 > d/2$; thus, in seeking the MLE we only consider ε_0 which is the negative solution of (35) in Theorem (33) (or, what is the same, $-\varepsilon_0$ where ε_0 is the positive solution).

(37) **Theorem.** *If* $0 \leqq d \leqq \sqrt{2}\sigma/\sqrt{n}$, $(\bar{x}, \bar{x})$ *is the only critical point and is the MLE.*

If $d > \sqrt{2}\sigma/\sqrt{n}$, *there are two critical points. One yields* $(\bar{x}, \bar{x})$ *and is a saddle point. The other yields the MLE*

(38) $$\left(\bar{x} - \frac{\varepsilon_0}{2d}\sigma^2/n, \bar{x} + \frac{\varepsilon_0}{2d}\sigma^2/n\right),$$

where ε_0 *is the positive solution of*

(39) $$d^2 n/\sigma^2 = \varepsilon \coth(\varepsilon/2).$$

Theorem (37) follows from previous results, notably Lemma (32) for the form of the MLE, Lemma (33) for the solutions of a certain equation, and (24) for the nature of $(\bar{x}, \bar{x})$. In obtaining the form of (38), relations such as

$$\begin{aligned}\hat{\mu}_{[1]} &= x_1 + a_1 = x_1 + \frac{d}{2} - \frac{\varepsilon_0}{2d}\sigma^2/n \\ &= \bar{x} - \frac{\varepsilon_0}{2d}\sigma^2/n\end{aligned}$$

are used. Note that, for $d^2 n/\sigma^2$ "large", $\varepsilon_0 \simeq d^2 n/\sigma^2$, so that (38) is "close" to (x_1, x_2). The following lemma studies the approach of ε_0 to $d^2 \frac{n}{\sigma^2}$.

(40) **Lemma.** *If* ε_0 *is the positive solution of* (39), *then* (*with* $o(n) \geqq 0$)

$$\varepsilon_0 = \frac{d^2 n}{\sigma^2} - o(n).$$

Proof. If we write $a = d^2/\sigma^2$, then we are interested in the positive solution of $\varepsilon \coth(\varepsilon/2) = a \cdot n$. Let us set this solution as $\varepsilon_0 = a \cdot n - c_n$ and investigate the order of c_n. Substituting in the equation,

$$(a \cdot n - c_n) \coth\left(\frac{a \cdot n - c_n}{2}\right) = a \cdot n$$

or

(41) $$\left(1 - \frac{c_n}{a \cdot n}\right) \coth\left(\frac{a \cdot n - c_n}{2}\right) = 1.$$

From Fig. 2 we see that $\varepsilon_0 \to \infty$ as $n \to \infty$, and since $\varepsilon_0 > 0$ we have $c_n < a \cdot n$ or $c_n/n < a$. Since $\coth(x) > 1$ if $x > 0$, and since (41) must be satisfied, $c_n/n > 0$. Now, taking the limit of (41) as $n \to \infty$, we find that

$$(1 - b/a) \cdot 1 = 1$$

where $0 \leqq b = \lim_{n \to \infty} \frac{c_n}{n} \leqq a$. This is a contradiction unless $\lim_{n \to \infty} \frac{c_n}{n} = 0$, so that $c_n = o(n)$.

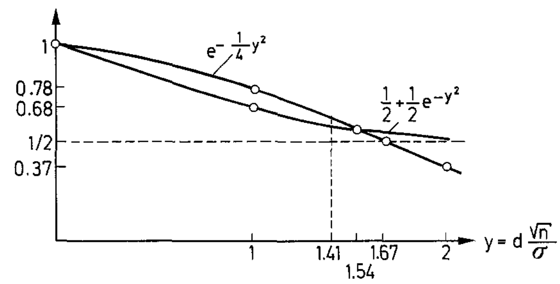

Fig. 3. Likelihoods of $(\bar{X}, \bar{X})$ (MLE if $0 \leqq y \leqq \sqrt{2}$) ... $e^{-y^2/4}$; and of $(\bar{X}_{[1]}, \bar{X}_{[2]})$ ... $\frac{1}{2} + \frac{1}{2}e^{-y^2}$

It is of interest to compare (for the case $k = 2$) the likelihoods of the three estimators $(\bar{X}, \bar{X})$, $(\bar{X}_{[1]}, \bar{X}_{[2]})$, and the MLE. With $d = x_2 - x_1$, we find (see (A.1))

$$\pi \frac{\sigma^2}{n} f_{\bar{X}_{[1]}, \bar{X}_{[2]}}(x_1, x_2)$$

$$= \pi \phi\left(\frac{x_1 - \mu_{[1]}}{\sigma/\sqrt{n}}\right) \phi\left(\frac{x_2 - \mu_{[2]}}{\sigma/\sqrt{n}}\right) + \pi \phi\left(\frac{x_2 - \mu_{[1]}}{\sigma/\sqrt{n}}\right) \phi\left(\frac{x_1 - \mu_{[2]}}{\sigma/\sqrt{n}}\right)$$

(42) $$= \begin{cases} e^{-\frac{d^2}{4\sigma^2/n}} & \text{if } (\mu_{[1]}, \mu_{[2]}) = (\bar{x}, \bar{x}), \text{ the MLE for } 0 \leqq d^2 \leqq 2\sigma^2/n \\ \frac{1}{2} + \frac{1}{2} e^{-\frac{d^2}{\sigma^2/n}} & \text{if } (\mu_{[1]}, \mu_{[2]}) = (x_1, x_2) \\ \frac{1}{2} e^{-\frac{1}{4\sigma^2/n}\left(\frac{\varepsilon_0}{d}\frac{\sigma^2}{n} - d\right)^2} + \frac{1}{2} e^{-\frac{1}{4\sigma^2/n}\left(\frac{\varepsilon_0}{d}\frac{\sigma^2}{n} + d\right)^2} & \\ & \text{if } (\mu_{[1]}, \mu_{[2]}) = \left(\bar{x} - \frac{\varepsilon_0}{2d}\frac{\sigma^2}{n}, \bar{x} + \frac{\varepsilon_0}{2d}\frac{\sigma^2}{n}\right), \\ & \text{the MLE for } d^2 > 2\sigma^2/n. \end{cases}$$

If $0 \leqq d\sqrt{n}/\sigma \leqq \sqrt{2}$, $(\bar{X}, \bar{X})$ is the MLE, and the curve of $(\bar{X}, \bar{X})$ has ordinate $\frac{1}{2}$ when $d\sqrt{n}/\sigma = 2\sqrt{\ln 2} \simeq 1.67$. The curves of $(\bar{X}, \bar{X})$ and $(\bar{X}_{[1]}, \bar{X}_{[2]})$ cross at $d\sqrt{n}/\sigma \simeq 1.54$. At $d\sqrt{n}/\sigma = 2$, for $(\bar{X}_{[1]}, \bar{X}_{[2]})$ we find $\frac{1}{2} + \frac{1}{2}e^{-y^2} = \frac{1}{2} + \frac{1}{2}(0.01831) = 0.5092$, while for the MLE, a solution of $4 = \varepsilon \coth(\varepsilon/2)$ is approximately $\varepsilon_0 = 3.8$ (thus $\varepsilon_0/4 = 0.95$) and $1 - \varepsilon_0/4 = 0.05$. Thus, for the MLE we find

$$\frac{1}{2}e^{-\frac{y^2}{4}\left(\frac{\varepsilon_0}{y^2}-1\right)^2} + \frac{1}{2}e^{-\frac{y^2}{4}\left(\frac{\varepsilon_0}{y^2}+1\right)^2} = \frac{1}{2}\{e^{-0.0025} + e^{-3.8025}\}$$
$$\geqq \frac{1}{2}\{e^{-0.003} + e^{-3.81671}\} \simeq \frac{1}{2}(1.019) = 0.5095.$$

Note that Theorem (2.1.33) of [5] indicates the reasonableness of an estimator which compensates, as does the MLE $= (x_1 + a, x_2 - b)$, for under and over estimation with regard to expectation; the likelihood approach bears this out.

The above results indicate a weakness of taking a function of MLE's to estimate that function of the parameters for a problem (as discussed at (3)): namely, other methods yield *different* estimators with *higher* likelihoods. $\left(\text{In fact, with the other method the likelihood could never exceed } \frac{1}{2\pi} n/\sigma^2\text{; with our method it can never be less than } \frac{1}{2\pi} n/\sigma^2.\right)$

Acknowledgment. Thanks are especially due to Professor Lionel Weiss for his suggestions for and guidance of [5], which contains these results.

Appendix A. Joint Distribution of $\bar{X}_{[1]}, \dots, \bar{X}_{[k]}$

The joint density of $\bar{X}_1, \dots, \bar{X}_k$ is

$$f_{\bar{X}_1, \dots, \bar{X}_k}(y_1, \dots, y_k) = f_{\bar{X}_1}(y_1) \dots f_{\bar{X}_k}(y_k) \qquad (y_i \in \Re;\ i = 1, \dots, k)$$

where $f_{\bar{X}_i}(\cdot)$ is the $N(\mu_i, \sigma^2/n)$ density function $(i = 1, \dots, k)$.

It is well-known that then the joint density of the ordered $\bar{X}_i$ $(i = 1, \dots, k)$, i.e. of $\bar{X}_{[1]} \leqq \dots \leqq \bar{X}_{[k]}$, is

$$f_{\bar{X}_{[1]}, \dots, \bar{X}_{[k]}}(x_1, \dots, x_k)$$

$$(\text{A.1}) \qquad = \begin{cases} \sum_{\beta \in S_k} f_{\bar{X}_1, \dots, \bar{X}_k}(x_{\beta(1)}, \dots, x_{\beta(k)}), & x_1 \leqq \dots \leqq x_k \\ 0, & \text{otherwise} \end{cases}$$

$$= \begin{cases} \sum_{\beta \in S_k} (\sqrt{n}/\sigma)^k \, \phi\left(\frac{x_{\beta(1)} - \mu_1}{\sigma/\sqrt{n}}\right) \dots \phi\left(\frac{x_{\beta(k)} - \mu_k}{\sigma/\sqrt{n}}\right), & x_1 \leqq \dots \leqq x_k \\ 0, & \text{otherwise} \end{cases}$$

$$= \begin{cases} \sum_{\beta \in S_k} (\sqrt{n}/\sigma)^k \, \phi\left(\frac{x_{\beta(1)} - \mu_{[1]}}{\sigma/\sqrt{n}}\right) \dots \phi\left(\frac{x_{\beta(k)} - \mu_{[k]}}{\sigma/\sqrt{n}}\right), & x_1 \leqq \dots \leqq x_k \\ 0, & \text{otherwise.} \end{cases}$$

References

1. Berk, R.H.: Zehna, P.W. Invariance of maximum likelihood estimators, Review #1922. Mathematical Reviews **33**, 342–343 (1967).
2. Blumenthal, S., Cohen, A.: Estimation of two ordered translation parameters. Ann. math. Statistics **39**, 517–530 (1968a).
3. — — Estimation of the larger translation parameter. Ann. math. Statistics **39**, 502–516 (1968b).
4. — — Estimation of the larger of two normal means. Amer. statist. Assoc. **63**, 861–876 (1968c).

5. Dudewicz, E. J.: Estimation of ordered parameters. Technical Report No. 60, Department of Operations Research, Cornell University, Ithaca, New York (1969).
6. – Confidence intervals for ranked means. Naval Res. Logist. Quart. **17**, 69–78 (1970).
7. Graybill, F. A.: An introduction to linear statistical models, Volume I. New York: McGraw-Hill Book Company, Inc. 1961.
8. Katz, M. W.: Estimating ordered parameters. Ann. math. Statistics **34**, 967–972 (1963).
9. Robertson, T., Waltman, P.: On estimating monotone parameters. Ann. math. Statistics **39**, 1030–1039 (1968).
10. Zehna, P. W.: Invariance of maximum likelihood estimators. Ann. math. Statistics **37**, 744 (1966).

E. J. Dudewicz
The University of Rochester
Department of Statistics
Rochester, New York 14627
USA

(Received August 4, 1969, December 23, 1970)