Ann. Inst. Statist. Math.
33 (1981), Part A, 437-448

SOME FINITE SAMPLE RESULTS FOR THE SELECTION DIFFERENTIAL

H. N. NAGARAJA

(Received Nov. 17, 1980; revised Feb. 20, 1981)

1. Introduction

Let $X_1, X_2, \cdots, X_n$ be a random sample of size n from a continuous distribution with mean μ and variance σ^2 and distribution function (df) F. Let $X_{1:n} \leqq X_{2:n} \leqq \cdots \leqq X_{n:n}$ denote the order statistics of this sample. Suppose we select the top k X-values. Then $k^{-1} \sum_{i=n-k+1}^{n} (X_{i:n}-\mu)$ represents the average difference between the selected group and the population mean. This quantity expressed in standard deviation units is called the selection differential and may be written as

$$D_{k,n}=\frac{1}{k} \sum_{i=n-k+1}^{n} (X_{i:n}-\mu)/\sigma .$$

"Selection differential" has long been a familiar term to geneticists and breeders who often refer to it as "intensity of selection" (Falconer [3]). It represents a measure of improvement in the given trait due to selection. Hence it is useful in the construction of suitable breeding plans and in the comparison of different plans in plant as well as animal breeding. Of course it can be used in other kinds of selection problems as well. In this paper we obtain some finite sample results for $D_{k,n}$. An expression for the df of $D_{k,n}$ is given and several bounds on $\mathrm{E}\, D_{k,n}$ are presented. Numerical comparison of these bounds are made for the standard normal population when the sample size is 10. The last section considers the dependent sample case and develops bounds for $\mathrm{E}\, D_{k,n}$. In our discussion we usually assume that μ and σ are known and without loss of generality (WLOG) take $\mu=0$, $\sigma=1$. When μ and σ are replaced by $\bar{X}$ and S, the sample mean and sample standard deviation, the resulting quantity will be called the sample selection differential and is denoted by $\hat{D}_{k,n}$.

2. Distribution of $D_{k,n}$

$$\mathrm{P}\,(D_{k,n} \leqq x)=\mathrm{P}\,(X_{n-k+1:n}+\cdots+X_{n:n} \leqq kx)$$

$$=\int \mathrm{P}\,(X_{n-k+1:n}+\cdots+X_{n:n}\leqq kx\,|\,X_{n-k:n}=u)dF_{X_{n-k:n}}(u)$$

where $F_{X_{n-k:n}}$ is the df of $X_{n-k:n}$. It is known that (see David [2], p. 20) given $X_{n-k:n}=u$, $X_{n-k+1:n},\cdots,X_{n:n}$ form the order statistics from a random sample of size k from the df G_u given by

$$G_u(t)=\begin{cases} 0\,, & t<u \\ \dfrac{F(t)-F(u)}{1-F(u)}\,, & t\geqq u\,. \end{cases}$$

Hence,

$$\mathrm{P}\,(D_{k,n}\leqq x)=\int_{-\infty}^{x} G_u^{(k)}(kx)dF_{X_{n-k:n}}(u)\,, \tag{1}$$

where $G_u^{(k)}$ is the k-fold convolution of G_u; that is, the df of the sum of k independent identically distributed (iid) random variables (rvs) each with df G_u. As is evident from (1), there is no closed form expression for the df of $D_{k,n}$ in general. However, in the case of the exponential distribution, an expression for the probability density function (pdf) of $D_{k,n}$ can be given as discussed below.

Example. Let the parent distribution be Exp(1), that is exponential with mean unity. From the well known representation for exponential order statistics (see e.g., David [2], pp. 20–21) one obtains

$$M_{k,n}\equiv\frac{1}{k}\sum_{i=n-k+1}^{n} X_{i:n}\overset{d}{=}\frac{Z_1}{n}+\frac{Z_2}{n-1}+\cdots+\frac{Z_{n-k+1}}{k}\cdots+\frac{Z_n}{k} \tag{2}$$

where Z_i's are iid Exp(1) rvs. Hence,

$$M_{k,n}\overset{d}{=}Z_1^*+\cdots+Z_{n-k}^*+Z^*/k$$

where $Z_i^*\sim\mathrm{Exp}\,(\lambda_i^{-1})$, $\lambda_i=(n-i+1)$ and Z^*, the sum of k iid Exp(1) rvs, is Gamma $(k, 1)$, and are mutually independent. Consequently,

$$f_{M_{k,n}}(u)=\int_0^{\infty} f_{Z_1^*+\cdots+Z_{n-k}^*}(u-x)f_{Z^*/k}(x)dx\,.$$

From Feller [4], p. 40, problem 12, it follows that

$$f_{Z_1^*+\cdots+Z_{n-k}^*}(u-x)=\lambda_1\lambda_2\cdots\lambda_{n-k}\left[\sum_{i=1}^{n-k}\Psi_{i,n-k}e^{-\lambda_i(u-x)}\right]\,,\qquad u-x>0$$

where

$$\Psi_{i,n-k}^{-1}=(\lambda_1-\lambda_i)\cdots(\lambda_{i-1}-\lambda_i)(\lambda_{i+1}-\lambda_i)\cdots(\lambda_{n-k}-\lambda_i)\,.$$

Therefore,

$$f_{M_{k,n}}(u)=\lambda_1\lambda_2\cdots\lambda_{n-k}\sum_{i=1}^{n-k}\Psi_{i,n-k}\int_0^u e^{-\lambda_i(u-x)}\frac{k^k}{(k-1)!}e^{-kx}x^{k-1}dx$$
$$=\frac{k^k}{(k-1)!}\lambda_1\lambda_2\cdots\lambda_{n-k}\sum_{i=1}^{n-k}\Psi_{i,n-k}e^{-\lambda_i u}\int_0^u e^{x(\lambda_i-k)}x^{k-1}dx\ ,\qquad u>0\ .$$

For a given k the integral can be evaluated explicitly and hence an explicit expression for the pdf of $D_{k,n}$ is available, since $D_{k,n}=(M_{k,n}-1)$.

3. Bounds on the expected value of the selection differential

3.1. *Bounds on* $\hat{D}_{k,n}$

Let $x_{1:n}\leqq x_{2:n}\leqq\cdots\leqq x_{n:n}$ be the order statistics from an observed sample $x_1, x_2,\cdots, x_n$. Mallows and Richter [5] have established sharp bounds for $v_k=k^{-1}\sum_{i=n-k+1}^{n} x_{i:n}$, which is the sample selection differential except for a change of location and scale. Their Corollary 6.1 (p. 1931) states that

$$\bar{x}+\frac{n-k}{t}\frac{s}{\sqrt{n-1}}\leqq v_k\leqq\bar{x}+\sqrt{\frac{n-k}{k}}s \tag{3}$$

where $t=\max(k, n-k)$ and $s^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$ so that $S^2=\frac{n}{n-1}s^2$. Assuming that $S\neq 0$ (i.e., x_i's are not all equal), we obtain

$$\frac{n-k}{t}\frac{1}{\sqrt{n}}\leqq\frac{v_k-\bar{x}}{S}\equiv\hat{D}_{k,n}\leqq\sqrt{\frac{n-k}{k}}\sqrt{\frac{n-1}{n}}\ .$$

These bounds are sharp.

3.2. *Bounds on* $\mathrm{E}\,D_{k,n}$*—Cauchy-Schwarz technique*

Since $\mu=0$, $\sigma=1$, $\int_0^1 F^{-1}(u)du=0$ and $\int_0^1[F^{-1}(u)]^2du=1$.

$$\mathrm{E}\,D_{k,n}=\frac{1}{k}\sum_{i=n-k+1}^{n}\mathrm{E}\,X_{i:n}=\int_0^1\Big[\sum_{i=n-k+1}^{n}\frac{1}{k}\frac{n!}{(i-1)!(n-i)!}\cdot u^{i-1}(1-u)^{n-i}-1\Big]F^{-1}(u)du$$
$$\leqq\left\{\int_0^1\Big[\sum_{i=n-k+1}^{n}\frac{n}{k}\binom{n-1}{i-1}u^{i-1}(1-u)^{n-i}-1\Big]^2du\right\}^{1/2}\cdot\left\{\int_0^1[F^{-1}(u)]^2du\right\}^{1/2},$$

by the Cauchy-Schwarz inequality. Hence,

$$\text{(4)} \qquad \mathrm{E}\, D_{k,n} \leqq \left\{ \left(\frac{n}{k}\right)^2 \frac{1}{2n-1} \sum_{i,j=n-k+1}^{n} \frac{\binom{n-1}{i-1}\binom{n-1}{j-1}}{\binom{2n-2}{i+j-2}} - 1 \right\}^{1/2} .$$

Of course, $\mathrm{E}\, D_{k,n} \geqq 0$. Equality in (4) is attained if and only if (iff), for some constant c,

$$\text{(5)} \qquad F^{-1}(u) = c\left[\frac{n}{k} \sum_{i=n-k+1}^{n} \binom{n-1}{i-1} u^{i-1}(1-u)^{n-i} - 1\right] .$$

First we note that for $k<n$, $\sum_{i=n-k+1}^{n} \binom{n-1}{i-1} u^{i-1}(1-u)^{n-i}$ represents the df of $(n-k)$th order statistic from a random sample of size $(n-1)$ from $\mathcal{U}(0, 1)$ distribution, that is, uniform distribution over $(0, 1)$. Hence, the right-hand side in (5) is increasing if $c>0$ and consequently there exists an F satisfying (5). For this F, $\mathrm{E}\, D_{k,n}$ is the bound given in (4). However, a closed form expression for such an F is not possible. But, since $\int_0^1 [F^{-1}(u)]^2 du = 1$,

$$c = \left\{ \left(\frac{n}{k}\right)^2 \frac{1}{2n-1} \sum_{i,j=n-k+1}^{n} \frac{\binom{n-1}{i-1}\binom{n-1}{j-1}}{\binom{2n-2}{i+j-2}} - 1 \right\}^{-1/2} .$$

Also, $F^{-1}(0) = -c$ and $F^{-1}(1) = c(n-k)/k$. Hence, this extremal F has bounded support, and is nonsymmetric.

Remarks. 1. The above technique has been employed for finding bounds for $\mathrm{E}\, X_{j:n}$, $1 \leqq j \leqq n$ in David ([2], p. 63) where it is noted that the bounds are attained only when $j=n$. But, in the case of the selection differential, or equivalently in the case of the average of $X_{n-k+1:n}, \cdots, X_{n:n}$ the bound is attainable for all k.

2. Let $h(X)$ and $g(X)$ be two functions of a rv X where $\mathrm{E}\,[h(X)]^2$ and $\mathrm{E}\,[g(X)]^2$ are finite. Let $\mathrm{E}\, h(X)=0$. Then sharper bounds can be obtained for $\mathrm{E}\, h(X)g(X)$ by using the Cauchy-Schwarz inequality for $\mathrm{E}\,(h(X)-\mathrm{E}\, h(X))(g(X)-\mathrm{E}\, g(X))$ instead of the given expectation even though the two integrals are essentially the same. This procedure would yield a tighter bound than the one obtained by direct application of the Cauchy-Schwarz inequality.

3. We can obtain sharper upper bounds for $\mathrm{E}\, D_{k,n}$ assuming a symmetric parent distribution and using similar techniques. The Cauchy-Schwarz inequality applied to some orthonormal systems can be used to obtain tighter bounds and approximations for $\mathrm{E}\, D_{k,n}$. These would closely follow Section 4.3 of David ([2], pp. 66–70) and are omitted. But, some nontrivial extensions of his Section 4.4 are possible and we

pursue these in the next section.

3.3. *c-Comparison and s-comparison*

Let $\mathcal{F}$ be the class of all dfs which have positive continuous derivatives on their supports. If F and F^* are in $\mathcal{F}$ then we say that $F \underset{c}{<} F^*$ iff $F^{*-1}F$ is convex on I, the support of F, and in such a case F is said to c-precede F^*. Van Zwet [7] has shown that if $F \underset{c}{<} F^*$, then

$$(6) \qquad F(\mathrm{E}\, X_{r:n}) \leqq F^*(\mathrm{E}\, X^*_{r:n})$$

for all $r=1, 2, \cdots, n$, and for all n for which $\mathrm{E}\, X_{r:n}$ and $\mathrm{E}\, X^*_{r:n}$ exist (see David [2], p. 73). We assume that both F and F^* have finite variances. Since c-ordering is independent of location and scale, WLOG we take both F and F^* to be standardized dfs.

From (6) we have

$$g(\mathrm{E}\, X_{r:n}) \leqq \mathrm{E}\, X^*_{r:n}\,, \qquad r=1, 2, \cdots, n$$

where $g=F^{*-1}F$ is a convex function on I. Hence,

$$(7) \qquad \frac{1}{k} \sum_{i=n-k+1}^{n} g(\mathrm{E}\, X_{i:n}) \leqq \frac{1}{k} \sum_{i=n-k+1}^{n} \mathrm{E}\, X^*_{i:n}\,.$$

Let Y be a rv which takes values $\mathrm{E}\, X_{n-k+1:n}, \cdots, \mathrm{E}\, X_{n:n}$ with probability $1/k$ each. Since g is a convex function on I and these expectations belong to I, we have, by Jensen's inequality

$$g\left(\frac{1}{k} \sum_{i=n-k+1}^{n} \mathrm{E}\, X_{i:n}\right) = g(\mathrm{E}\, Y) \leqq \mathrm{E}\, g(Y) = \frac{1}{k} \sum_{i=n-k+1}^{n} g(\mathrm{E}\, X_{i:n})\,.$$

Hence, we have

$$(8) \qquad g\left(\mathrm{E}\left(\frac{1}{k} \sum_{i=n-k+1}^{n} X_{i:n}\right)\right) \leqq \frac{1}{k} \sum_{i=n-k+1}^{n} g(\mathrm{E}\, X_{i:n}) \leqq \mathrm{E}\left(\frac{1}{k} \sum_{i=n-k+1}^{n} X^*_{i:n}\right)\,.$$

Recalling that F and F^* are standardized dfs it follows that

$$g(\mathrm{E}\, D_{k,n}) \leqq \frac{1}{k} \sum_{i=n-k+1}^{n} g(\mathrm{E}\, X_{i:n}) \leqq \mathrm{E}\,(D^*_{k,n})\,.$$

That is,

$$(9) \qquad F(\mathrm{E}\, D_{k,n}) \leqq F^*\left(\frac{1}{k} \sum_{i=n-k+1}^{n} g(\mathrm{E}\, X_{i:n})\right) \leqq F^*(\mathrm{E}\, D^*_{k,n})\,.$$

Again, from (6), we have

$$\mathrm{E}\, X_{r:n} \leqq g^{-1}(\mathrm{E}\, X^*_{r:n})\,.$$

Hence, proceeding on similar lines as above, and using the fact that g^{-1} is concave, one obtains,

$$F(\mathrm{E}\,D_{k,n}) \leqq F\left(\frac{1}{k}\sum_{i=n-k+1}^{n} g^{-1}(\mathrm{E}\,X^*_{i:n})\right) \leqq F^*(\mathrm{E}\,D^*_{k,n}) \,. \tag{10}$$

(9) can be used to give lower bounds for $\mathrm{E}\,D^*_{k,n}$ whereas (10) is handy if we are interested in an upper bound for $\mathrm{E}\,D_{k,n}$. However, note that the intermediate bounds are not easy to compute. If any of F and F^* is not standardized, the corresponding selection differential has to be replaced by the average of the top k order statistics. In that case, one does not even need the finiteness of the mean, just the existence of expectations of order statistics appearing in (9) or (10).

Applications. (i) c-comparison with the $\mathcal{U}(0, 1)$ df gives, for any (standardized) convex F,

$$F(\mathrm{E}\,D_{k,n}) \leqq F\left(\frac{1}{k}\sum_{i=n-k+1}^{n} (i/(n+1))\right) \leqq (2n-k+1)/2(n+1) \,;$$

for any concave F, the inequalities are reversed.

(ii) For a standardized df F having increasing failure rate, that is for which $F'(x)/(1-F(x))$ is non-decreasing, we get

$$F(\mathrm{E}\,D_{k,n}) \leqq 1-\exp(-\mathrm{E}\,M_{k,n})$$

on c-comparison with Exp (1) distribution. Here $M_{k,n}$ is as given by (2) and hence

$$\mathrm{E}\,M_{k,n} = \sum_{i=k+1}^{n} \frac{1}{i} + 1 \leqq \int_{k+1/2}^{n+1/2} x^{-1}dx + 1 = \log\frac{2n+1}{2k+1} + 1 \,.$$

Consequently, $F(\mathrm{E}\,D_{k,n}) \leqq 1-(2k+1)/e(2n+1)$.

(iii) For the standard normal parent with df $\Phi(x)$, $1/\Phi(x)$ is convex. Hence, with $F(x) = -1/x$, $x < -1$ and $F^*(x) = \Phi(x)$, $F^{-1}F^*$ is concave. Consequently, $g = F^{*-1}F$ is convex since g is increasing and its inverse function is concave. Also, note that F does not have a mean but $\mathrm{E}\,D_{k,n}$ exists for $k<n$. $\mathrm{E}\,X_{r:n} = -n/(r-1)$, $r>1$ (David [2], p. 74) and hence from (9) we have

$$\Phi(\mathrm{E}\,D^*_{k,n}) \geqq \Phi\left(\frac{1}{k}\sum_{i=n-k+1}^{n} \Phi^{-1}((i-1)/n)\right) \geqq F\left(-\frac{n}{k}\sum_{i=n-k}^{n-1}\frac{1}{i}\right) \,.$$

That is,

$$\mathrm{E}\,D^*_{k,n} \geqq \frac{1}{k}\sum_{i=n-k+1}^{n} \Phi^{-1}\left(\frac{i-1}{n}\right) \geqq \Phi^{-1}\left(\frac{k}{n\sum_{i=n-k}^{n-1} i^{-1}}\right) \,, \qquad k<n \,. \tag{11}$$

s-Comparison.

Now, we consider a subclass $\mathcal{S}$ of symmetric distributions in $\mathcal{F}$. Let $F(x_0-x)+F(x_0+x)=1$ for some x_0 and all x if $F\in\mathcal{S}$. If F and F^* are in $\mathcal{S}$, then $F\underset{s}{<}F^*$ iff $g=F^{*-1}F$ is convex for $x>x_0$, $x\in I$, the support of F. From van Zwet [7], we have, whenever $F\underset{s}{<}F^*$,

$$g(\mathrm{E}\,X_{r:n})\leqq \mathrm{E}\,X^*_{r:n} \tag{12}$$

for all $(n+1)/2\leqq r\leqq n$ and all n for which $\mathrm{E}\,X^*_{r:n}$ exists (see David [2], p. 76). We assume that both F and F^* are standardized. Consequently, $x_0=0$ and $g(0)=0$. Now, noting that $\mathrm{E}\,X_{r:n}>0$ for $r>(n+1)/2$ and that g is convex for $x>0$, we get, on using arguments similar to those leading to (8),

$$g(\mathrm{E}\,D_{k,n})\leqq\frac{1}{k}\sum_{i=n-k+1}^{n} g(\mathrm{E}\,X_{i:n})\leqq \mathrm{E}\,D^*_{k,n}\,,\qquad k\leqq(n+1)/2\,. \tag{13}$$

We now show that (13) is true even when $k>(n+1)/2$. Since F is symmetric about zero, for $k>(n+1)/2$,

$$\mathrm{E}\,D_{k,n}=\frac{1}{k}\,\mathrm{E}\left(\sum_{i=n-k+1}^{n} X_{i:n}\right)=\frac{2k-n}{k}\cdot 0+\frac{1}{k}\sum_{i=k+1}^{n}\mathrm{E}\,X_{i:n}\,. \tag{14}$$

From (12), since $k>(n+1)/2$,

$$\frac{1}{k}\sum_{i=k+1}^{n} g(\mathrm{E}\,X_{i:n})\leqq\frac{1}{k}\sum_{i=k+1}^{n}\mathrm{E}\,X^*_{i:n}\,. \tag{15}$$

Define a rv Z which takes values 0, $\mathrm{E}\,X_{k+1:n},\cdots,\mathrm{E}\,X_{n:n}$ with probabilities $(2k-n)/k$, $1/k,\cdots,1/k$, respectively. Since g is convex on the support of Z, by Jensen's inequality, it follows that

$$\begin{aligned} g\left(\frac{2k-n}{k}\cdot 0+\frac{1}{k}\sum_{i=k+1}^{n}\mathrm{E}\,X_{i:n}\right)&\leqq \mathrm{E}\,g(Z)\\ &=g(0)\,\frac{2k-n}{k}+\frac{1}{k}\sum_{i=k+1}^{n} g(\mathrm{E}\,X_{i:n})\\ &=\frac{1}{k}\sum_{i=k+1}^{n} g(\mathrm{E}\,X_{i:n})\\ &=\frac{1}{k}\sum_{i=n-k+1}^{n} g(\mathrm{E}\,X_{i:n}) \end{aligned}$$

since g is antisymmetric about 0. Now recalling (14) and (15) we conclude that (13) is true for $k>(n+1)/2$ also. This is recorded as a theorem below:

THEOREM. *If F and F^* are standardized dfs in $\mathcal{S}$, and $F\underset{s}{<}F^*$,*

$$F(\mathrm{E}\,D_{k,n}) \leqq F^*\Big(\frac{1}{k}\sum_{i=t+1}^{n} g(\mathrm{E}\,X_{i:n})\Big) \leqq F^*(\mathrm{E}\,D_{k,n}^*), \text{ where } t=\max(k, n-k).$$

One can also show that

$$(16) \qquad F(\mathrm{E}\,D_{k,n}) \leqq F\Big(\frac{1}{k}\sum_{i=t+1}^{n} g^{-1}(\mathrm{E}\,X_{i:n}^*)\Big) \leqq F^*(\mathrm{E}\,D_{k,n}^*) .$$

For nonstandardized dfs, the selection differential has to be replaced by $M_{k,n}$, the average of the top k order statistics.

s-Comparison of the standard normal df (F) with the logistic distribution (F^*), where $F^*(x)=(1+\exp(-x))^{-1}$, $-\infty<x<\infty$ shows that $F \underset{s}{<} F^*$ (see David [2], p. 77) and hence from (16) we have

$$(17) \qquad \mathrm{E}\,D_{k,n} \leqq \frac{1}{k}\sum_{i=t+1}^{n} \Phi^{-1}(F^*(\mathrm{E}\,X_{i:n}^*)) \leqq \Phi^{-1}(F^*(\mathrm{E}\,M_{k,n}^*)) .$$

It is known from David ([2], p. 78), $\mathrm{E}\,X_{r:n}^* = \sum_{i=n-r+1}^{r-1} i^{-1}$ for $r \geqq (n+1)/2$ and hence

$$(18) \qquad \mathrm{E}\,M_{k,n}^* = \frac{1}{k}\sum_{i=t+1}^{n} \mathrm{E}\,X_{i:n}^*$$

$$= \begin{cases} \displaystyle\sum_{i=1}^{n-1} \frac{1}{i}\,, & k=1 \\[2ex] \displaystyle\frac{n}{k}\sum_{i=n-k+1}^{n-1} \frac{1}{i} + \sum_{i=k}^{n-k} \frac{1}{i}\,, & 1<k\leqq \dfrac{n}{2} \\[2ex] \displaystyle\frac{n}{k}\sum_{i=n-k}^{n-1} \frac{1}{i} - \sum_{i=n-k}^{k} \frac{1}{i}\,, & \dfrac{n}{2}\leqq k<n \end{cases}$$

on simplification.

Now we compare some of the bounds discussed so far when the parent distribution is standard normal and the sample size is 10. For this define the following:

$\mathrm{UB1} = \frac{1}{k}\sum_{i=t+1}^{n} \Phi^{-1}(F^*(\mathrm{E}\,X_{i:n}^*))$ of (17),

$\mathrm{UB2} = \Phi^{-1}(F^*(\mathrm{E}\,M_{k,n}^*))$ of (17) where $\mathrm{E}\,M_{k,n}^*$ is given by (18),

UB3 = Bound given by (4) using the Cauchy-Schwarz technique,

$\mathrm{LB} = \frac{1}{k}\sum_{i=t+1}^{n} \Phi^{-1}\Big(\frac{i-1}{n}\Big)$, an improved version of the intermediate bound of (11) which exploits the symmetry of the normal distribution.

$\mathrm{E}\,D_{k,n}$ was computed using the table of expected values given by Teichroew [6]. Column (3) in Table 4.4 of David [2] was used to com-

Bounds for $\mathrm{E}\, D_{k,n}$ for $n=10$

k	$\mathrm{E}\, D_{k,n}$	UB1	UB2	UB3	LB
1	1.539	1.591	1.591	2.065	1.282
2	1.270	1.309	1.321	1.526	1.062
3	1.065	1.096	1.115	1.211	0.883
4	0.893	0.918	0.942	0.987	0.725
5	0.739	0.760	0.787	0.810	0.580
6	0.595	0.612	0.641	0.658	0.483
7	0.457	0.470	0.499	0.519	0.378
8	0.318	0.328	0.354	0.381	0.265
9	0.171	0.177	0.196	0.229	0.142

pute UB1. All these bounds and $\mathrm{E}\, D_{k,n}$ are given for $k=1(1)9$, $n=10$ in the table given above. Of the upper bounds, the ones obtained using s-comparison perform well in comparison with the one which uses the Cauchy-Schwarz technique. The lower bound is too low to be useful.

3.4. *Dependent sample case*

In this section we first consider bounds on the expectation of any linear function of order statistics when the variables are dependent and possibly nonidentically distributed. While doing so, we improve a result due to Arnold and Groeneveld [1]. Then, we discuss the case of the selection differential.

Suppose $X_1, X_2, \cdots, X_n$ are possibly dependent rvs with $\mathrm{E}\, X_i=\mu_i$ and $\mathrm{Var}\,(X_i)=\sigma_i^2$. Let $X_{1:n}\leqq X_{2:n}\leqq \cdots \leqq X_{n:n}$ be the order statistics with $\mu_{i:n}=\mathrm{E}\, X_{i:n}$. Let $\bar{X}$ be the sample mean and $s^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2$.

$$\begin{aligned}\mathrm{E}\, s^2&=\frac{1}{n}\sum_{i=1}^{n}\mathrm{E}\, X_i^2-\mathrm{E}\, \bar{X}^2\leqq\frac{1}{n}\sum_{i=1}^{n}\mathrm{E}\, X_i^2-(\mathrm{E}\, \bar{X})^2\,, \qquad \text{since } \mathrm{Var}\,(\bar{X})\geqq 0\\ &=\frac{1}{n}\sum_{i=1}^{n}(\mu_i^2+\sigma_i^2)-\bar{\mu}^2\,, \qquad \text{where } \bar{\mu}=\frac{1}{n}\sum\mu_i=\frac{1}{n}\sum\mu_{i:n}\\ &=\frac{1}{n}\sum_{i=1}^{n}[\sigma_i^2+(\mu_i-\bar{\mu})^2]\end{aligned}$$

and the equality holds iff $\bar{X}$=constant almost surely (a.s.). Also,

$$(\mu_{i:n}-\bar{\mu})^2=[\mathrm{E}\,(X_{i:n}-\bar{X})]^2\leqq\mathrm{E}\,(X_{i:n}-\bar{X})^2$$

and hence

$$\sum_i(\mu_{i:n}-\bar{\mu})^2\leqq\sum_i\mathrm{E}\,(X_{i:n}-\bar{X})^2=\mathrm{E}\,(\sum_i(X_i-\bar{X})^2)=n\,\mathrm{E}\, s^2$$

where the equality holds iff $X_{i:n}-\bar{X}=c_i$ a.s. with $\sum c_i=0$. Hence, we have the following:

$$\sum_{i=1}^{n}(\mu_{i:n}-\bar{\mu})^2 \leqq n\,\mathrm{E}\,s^2 \leqq \sum_{i=1}^{n}[\sigma_i^2+(\mu_i-\bar{\mu})^2]\ . \tag{19}$$

Arnold and Groeneveld ([1], pp. 220–221) have shown that:

$$\sum_{i=1}^{n}(\mu_{i:n}-\bar{\mu})^2 \leqq \sum_{i=1}^{n}[\sigma_i^2+(\mu_i-\bar{\mu})^2]$$

and hence, for constants λ_i, $1 \leqq i \leqq n$, that

$$\begin{aligned}|\textstyle\sum \lambda_i(\mu_{i:n}-\bar{\mu})| &\leqq [\textstyle\sum (\lambda_i-\bar{\lambda})^2]^{1/2}[\sum (\mu_{i:n}-\bar{\mu})^2]^{1/2} \\ &\leqq [\textstyle\sum (\lambda_i-\bar{\lambda})^2]^{1/2}[\sum (\sigma_i^2+(\mu_i-\bar{\mu})^2)]^{1/2}\ .\end{aligned} \tag{20}$$

However, using the first inequality in (19), we obtain

$$|\textstyle\sum \lambda_i(\mu_{i:n}-\bar{\mu})| \leqq \sqrt{n}\,[\sum (\lambda_i-\bar{\lambda})^2]^{1/2}(\mathrm{E}\,s^2)^{1/2} \tag{21}$$

which is strictly better than (20) unless the sample mean is a constant a.s. Also, if we start with $\sum \lambda_i(X_{i:n}-\bar{X})$ instead of $\sum \lambda_i(\mu_{i:n}-\bar{\mu})$, use the Cauchy-Schwarz inequality, and take expectations at the end, we end up with still better bounds. To be precise, consider

$$\begin{aligned}|\textstyle\sum \lambda_i(X_{i:n}-\bar{X})| &= |\textstyle\sum (\lambda_i-\bar{\lambda})(X_{i:n}-\bar{X})| \\ &\leqq [\textstyle\sum (\lambda_i-\bar{\lambda})^2]^{1/2}[\sum (X_{i:n}-\bar{X})^2]^{1/2} \\ &= \sqrt{n}\,[\textstyle\sum (\lambda_i-\bar{\lambda})^2]^{1/2} s\ .\end{aligned}$$

Therefore,

$$\begin{aligned}|\textstyle\sum \lambda_i(\mu_{i:n}-\bar{\mu})| &= |\mathrm{E}\textstyle\sum \lambda_i(X_{i:n}-\bar{X})| \leqq \mathrm{E}\,|\sum \lambda_i(X_{i:n}-\bar{X})| \\ &\leqq \sqrt{n}\,[\textstyle\sum (\lambda_i-\bar{\lambda})^2]^{1/2}\,\mathrm{E}\,s\ .\end{aligned}$$

That is,

$$|\textstyle\sum \lambda_i(\mu_{i:n}-\bar{\mu})| \leqq \sqrt{n}\,[\sum (\lambda_i-\bar{\lambda})^2]^{1/2}\,\mathrm{E}\,s\ . \tag{22}$$

Noting that $\mathrm{E}\,s^2 \geqq [\mathrm{E}\,(s)]^2$ we see that (22) gives a sharper bound than (21), with equality of bounds occurring only when s^2 is a constant a.s. The only shortcoming of (21) or (22) is that we need to know $\mathrm{E}\,s^2$ or $\mathrm{E}\,s$ in order to compute the bounds. But, at the same time, one can dispense with the knowledge of σ_i^2's which are needed in (20).

Finally, we consider a special case of dependence where X_i's are uncorrelated. Then, it can be shown that

$$n\,\mathrm{E}\,(s^2) = \sum_i \left[(\mu_i-\bar{\mu})^2+\sigma_i^2\left(\frac{n-1}{n}\right)\right]$$

and hence (21) reduces to

$$|\sum \lambda_i(\mu_{i:n}-\bar{\mu})|\leqq[\sum (\lambda_i-\bar{\lambda})^2]^{1/2}\left[\sum \left((\mu_i-\bar{\mu})^2+\left(1-\frac{1}{n}\right)\sigma_i^2\right)\right]^{1/2}$$

indicating clearly the improvement over (20). The above inequality is dealt with in Exercise 4.5.1. of David [2].

Now we can assume that X_i's have the same mean μ and the same variance σ^2 and turn our attention to the selection differential. Here, sharp bounds can be obtained by dealing with (3), rather than appealing to any of the inequalities derived above. Taking expectations in (3), we get

$$\mu+\frac{n-k}{t}\frac{\mathrm{E}\, s}{\sqrt{n-1}}\leqq\mathrm{E}\left(\frac{1}{k}\sum_{i=n-k+1}^{n} X_{i:n}\right)\leqq\mu+\sqrt{\frac{n-k}{k}}\,\mathrm{E}\, s$$

where $t=\max(k, n-k)$. Therefore,

$$\frac{n-k}{\max(k, n-k)}\frac{1}{\sqrt{n-1}}\frac{\mathrm{E}\, s}{\sigma}\leqq\mathrm{E}\, D_{k,n}\leqq\sqrt{\frac{n-k}{k}}\frac{\mathrm{E}\, s}{\sigma}\,. \tag{23}$$

Since the bounds in (3) are sharp, these bounds are also sharp. (A necessary condition is that s is constant a.s.). If $\mathrm{E}\, s$ is unknown, the fact that $\mathrm{E}\, s\leqq\sigma$ can be used to replace the upper bound in (23) by $\sqrt{(n-k)/k}$. In addition, if X_i's are uncorrelated,

$$\mathrm{E}\, s\leqq\sqrt{\mathrm{E}\, s^2}=\sigma\sqrt{(n-1)/n}$$

gives a slightly better upper bound, namely $\sqrt{(n-k)(n-1)/kn}$. But, a good lower bound for $\mathrm{E}\, s/\sigma$ is not possible without additional conditions on the parent distribution.

Acknowledgements

This paper is based on a portion of my Ph.D. dissertation written under the direction of Dr. H. A. David at Iowa State University. The research was supported by a grant from U.S. Army Research Office.

THE OHIO STATE UNIVERSITY

REFERENCES

[1] Arnold, B. C. and Groeneveld, R. A. (1979). Bounds on expectations of linear systematic statistics based on dependent samples, *Ann. Statist.*, **7**, 220-223.

[2] David, H. A. (1981). *Order Statistics*, 2nd ed., John Wiley & Sons, Inc., New York.

[3] Falconer, D. S. (1960). *Introduction to Quantitative Genetics*, The Ronald Press Co., New York.

[4] Feller, W. (1966). *An Introduction to Probability Theory and its Applications*, Vol. II, John Wiley & Sons, Inc., New York.

[5] Mallows, C. L. and Richter, D. (1969). Inequalities of Chebychev type involving conditional expectations, *Ann. Math. Statist.*, **40**, 1922-1932.
[6] Teichroew, D. (1956). Tables of expected values of order statistics and products of order statistics for samples of size 20 and less from the normal distribution, *Ann. Math. Statist.*, **27**, 410-426.
[7] van Zwet, W. R. (1964). *Convex Transformations of Random Variables*, Mathematical Centre Tracts 7 Mathematisch Centrum, Amsterdam.