Journal of Statistical Planning and Inference 93 (2001) 17–27

journal of statistical planning and inference

www.elsevier.com/locate/jspi

Expectation bounds on linear estimators from dependent samples

Nickos Papadatos *

Department of Mathematics and Statistics, University of Cyprus, P.O. Box 20537, 1678 Nicosia, Cyprus

Received 6 January 2000; received in revised form 15 May 2000; accepted 19 June 2000

Abstract

Let $X_1, X_2, \ldots, X_n$ be a sample of arbitrary, possibly dependent, random variables, with possibly different marginal distributions, and denote by $X_{1:n} \leqslant X_{2:n} \leqslant \cdots \leqslant X_{n:n}$ the corresponding order statistics. Using the notation $\mu_i = \mathbb{E}X_i$ and $\sigma_i^2 = \operatorname{Var} X_i$, $i = 1, 2, \ldots, n$ (assumed finite), it is proved that for any real constants $\lambda_1, \lambda_2, \ldots, \lambda_n$,

$$\sum_{i=1}^{n} \lambda_i(\mathbb{E}X_{i:n} - \bar{\mu}) \leqslant \left(\sum_{i=1}^{n}(c_i - \bar{\lambda})^2\right)^{1/2} \left(\sum_{i=1}^{n}\{(\mu_i - \bar{\mu})^2 + \sigma_i^2\} - n \operatorname{Var}\bar{X}\right)^{1/2},$$

where $\bar{\mu} = n^{-1}\sum_{i=1}^{n}\mu_i$, $\bar{\lambda} = n^{-1}\sum_{i=1}^{n}\lambda_i$, $\bar{X} = n^{-1}\sum_{i=1}^{n}X_i$ and $(c_1, c_2, \ldots, c_n)'$ is the l^2-projection of the vector $(\lambda_1, \lambda_2, \ldots, \lambda_n)'$ onto the convex cone of componentwise nondecreasing vectors in $\mathbb{R}^n$ (in particular, $c_i = \lambda_i$ for all i if and only if λ_i is nondecreasing in i). A similar lower bound is also given. The bound is sharp when the X's are exchangeable; moreover, it provides an improvement over the known bounds given by (Arnold and Groeneveld, 1979 Ann. Statist. 7, 220–223, Aven, 1985 J. Appl. Probab. 22, 723–728 and Lefèvre, 1986 Stochastic Anal. Appl. 4, 351–356). © 2001 Elsevier Science B.V. All rights reserved.

MSC: Primary 62G30

Keywords: Expectation bounds; Dependent samples; Linear estimators

1. Introduction

Let $X_1, X_2, \ldots, X_n$ be an arbitrary set of random variables (possibly dependent and with possibly different marginals), with means $\mu_i = \mathbb{E}X_i$ and finite variances $\sigma_i^2 = \operatorname{Var} X_i$, $i = 1, 2, \ldots, n$, and denote their order statistics by $X_{1:n} \leqslant X_{2:n} \leqslant \cdots \leqslant X_{n:n}$. A pioneer result

* Correspondence address. University of Athens, Department of Mathematics, Section of Statistics and O.R., Panepistemiopolis, 157 84 Athens, Greece.

E-mail address: npapadat@cc.uoa.gr, npapadat@ucy.ac.cy (N. Papadatos).

PII: S0378-3758(00)00167-1

of Arnold and Groeneveld (1979) states that for any real constants $\lambda_1, \lambda_2, \dots, \lambda_n$,

$$\left|\sum_{i=1}^{n} \lambda_i(\mathbb{E}X_{i:n} - \bar{\mu})\right| \leqslant \left(\sum_{i=1}^{n}(\lambda_i - \bar{\lambda})^2\right)^{1/2} \left(\sum_{i=1}^{n}\{(\mu_i - \bar{\mu})^2 + \sigma_i^2\}\right)^{1/2}, \tag{1.1}$$

where $\bar{\mu} = n^{-1}\sum_{i=1}^{n}\mu_i$ and $\bar{\lambda} = n^{-1}\sum_{i=1}^{n}\lambda_i$.

If the covariances $\sigma_{ij} = \mathrm{Cov}[X_i, X_j]$, $i, j = 1, 2, \dots, n$, of the initial sample are known, some improvements of (1.1) were shown by Aven (1985) (who considers bounds on expectations of the extreme observations) and Lefèvre (1986) (who treats the general case), namely,

$$\left|\sum_{i=1}^{n} \lambda_i(\mathbb{E}X_{i:n} - \bar{\mu})\right| \leqslant \left(\sum_{i=1}^{n}(\lambda_i - \bar{\lambda})^2\right)^{1/2} \left(\min_{1\leqslant j\leqslant n} S_j^2 + \sum_{i=1}^{n}(\mu_i - \bar{\mu})^2\right)^{1/2}, \tag{1.2}$$

where $S_j^2 = \sum_{i=1}^{n}\mathrm{Var}[X_i - X_j]$, $j = 1, 2, \dots, n$. However, as pointed out by Lefèvre, (1.2) is not always sharper than (1.1), e.g., for exchangeable X's with $\mu_j = \mu$, $\sigma_j^2 = \sigma^2$ and $\sigma_{ij} = c$, (1.2) improves (1.1) if and only if $2(n-1)c \geqslant (n-2)\sigma^2$.

Several applications of the above results are given by these authors: Arnold and Groeneveld derived sharp bounds for the expectations of the ith-order statistic, of the trimmed mean, of the difference of two order statistics and of Downton's (1966) unbiased estimator of the normal standard deviation; Lefèvre obtained simple explicit upper bounds for the mean completion time in Pert networks.

Moreover, Nagaraja (1981) and Arnold and Balakrishnan (1989, Theorem 3.18) presented a method for obtaining (1.1) by using deterministic bounds involving n real numbers $x_{1:n} \leqslant x_{2:n} \leqslant \cdots \leqslant x_{n:n}$. More specifically, Nagaraja derived an improved versions of (1.1) by using a technique consisting of the following two steps: He first considered the quantities $\bar{X} = n^{-1}\sum_{i=1}^{n}X_i$, the sample mean, and $S = (S^2)^{1/2}$, where $S^2 = \sum_{i=1}^{n}[X_i - \bar{X}]^2$, a scalar multiple of the sample standard deviation, showing that

$$\mathbb{E}^2 S \leqslant \mathbb{E}S^2 \leqslant \textstyle\sum_{i=1}^{n}\{(\mu_i - \bar{\mu})^2 + \sigma_i^2\}. \tag{1.3}$$

He then applied expectations in the deterministic inequality

$$\left|\sum_{i=1}^{n} \lambda_i(X_{i:n} - \bar{X})\right| \leqslant \left(\sum_{i=1}^{n}(\lambda_i - \bar{\lambda})^2\right)^{1/2} S, \tag{1.4}$$

yielding

$$\left|\sum_{i=1}^{n} \lambda_i(\mathbb{E}[X_{i:n}] - \bar{\mu})\right| \leqslant \left(\sum_{i=1}^{n}(\lambda_i - \bar{\lambda})^2\right)^{1/2} \mathbb{E}S. \tag{1.5}$$

More recently, Rychlik (1993a,b, 1994, 1995) obtained the sharp expectation bounds for Linear Estimators, particularly applicable in the case where the X's are identically distributed (possibly dependent); in particular, he considered the l^2–projection, $(c_1, c_2, \dots, c_n)'$, of the vector $(\lambda_1, \lambda_2, \dots, \lambda_n)'$, onto the convex cone of nondecreasing functions $x : \{1, 2, \dots, n\} \to \mathbb{R}$, yielding the correct sharp version of (1.1) in terms of the coefficients c_i rather than λ_i. This projection method was extended by Gajek and Rychlik (1996, 1998), in order to determine sharper expectation bounds when the X's arise from restricted families of distributions (for a comprehensive review of all the above results (and more) see Rychlik (1998)).

In the present note we derive two improved versions of (1.1) in terms of the means, variances and covariances of the X's (Theorems 2.1 and 2.2). Some applications of the bounds are discussed in Section 3. It turns out that (2.8) and (2.9) improve (2.2), and that (2.2) is strictly better than both (1.1) and (1.2). It should be noted, however, that the result of Theorem 2.1 is, in fact, implied by Nagaraja's results (1.5) and (1.3) (see Remark 2.2, below). Nevertheless, it seems that expression (2.2) is very well hidden into these results, in such a way that many subsequent papers do not use it, although it provides better bounds. This fact can be clearly seen in the applications of Section 3 (see also the discussion after Eq. (22) in Nagaraja's paper).

2. Expectation bounds for Linear Estimators in terms of the first two moments

We first prove the following Lemma.

Lemma 2.1. *Let* $\mathbf{X}=(X_1,X_2,\ldots,X_n)'$, $n\geqslant 2$, *be an exchangeable random vector with* $\mathbb{E}X_1=\mu$, $\operatorname{Var}X_1=\sigma^2$ (*assumed finite*) *and* $\operatorname{Cov}[X_1,X_2]=c$. *Then, for any real constants* $\lambda_1,\lambda_2,\ldots,\lambda_n$,

$$\left|\sum_{i=1}^{n}\lambda_i(\mathbb{E}X_{i:n}-\mu)\right| \leqslant \left(\sum_{i=1}^{n}(\lambda_i-\bar{\lambda})^2\right)^{1/2}((n-1)(\sigma^2-c))^{1/2}, \tag{2.1}$$

where $\bar{\lambda}=n^{-1}\sum_{i=1}^{n}\lambda_i$ *and* $X_{1:n}\leqslant X_{2:n}\leqslant\cdots\leqslant X_{n:n}$ *are the order statistics corresponding to* $\mathbf{X}$. *Moreover, bound* (2.1) *is best possible* (*for any given values of* μ, σ^2 *and* c), *provided that the finite sequence* $\lambda_1,\lambda_2,\ldots,\lambda_n$ *is monotone.*

Proof. First note that $-\sigma^2/(n-1)\leqslant c\leqslant\sigma^2$, so that the upper bound in (2.1) is well defined. Next observe that $\mathbb{E}\bar{X}=\mu$, and thus

$$\begin{aligned}\sum_{i=1}^{n}\mathbb{E}^2[X_{i:n}-\mu] &= \sum_{i=1}^{n}\mathbb{E}^2[X_{i:n}-\bar{X}]\\ &\leqslant \sum_{i=1}^{n}\mathbb{E}[X_{i:n}-\bar{X}]^2\\ &= \mathbb{E}\sum_{i=1}^{n}[X_i-\bar{X}]^2\\ &= (n-1)(\sigma^2-c).\end{aligned}$$

Inequality (2.1) follows immediately from the above inequality and the following one (see Arnold and Groeneveld, 1979)

$$\begin{aligned}\left|\sum_{i=1}^{n}\lambda_i(\mathbb{E}X_{i:n}-\mu)\right| &= \left|\sum_{i=1}^{n}(\lambda_i-\bar{\lambda})(\mathbb{E}X_{i:n}-\mu)\right|\\ &\leqslant \left(\sum_{i=1}^{n}(\lambda_i-\bar{\lambda})^2\right)^{1/2}\left(\sum_{i=1}^{n}\mathbb{E}^2[X_{i:n}-\mu]\right)^{1/2}.\end{aligned}$$

We next study the case of equality. If the λ's are all equal, (2.1) takes the trivial form $0=0$. If the λ's form a nondecreasing sequence with $\lambda_1<\lambda_n$, then we consider an arbitrary random variable Z with $\mathbb{E}\,Z=\mu$ and $\operatorname{Var} Z=(\sigma^2+(n-1)c)/n$, and we set $Y_j=Z+A(\lambda_j-\bar{\lambda})$, $j=1,2,\ldots,n$, where

$$A=\left(\frac{(n-1)(\sigma^2-c)}{\sum_{i=1}^n(\lambda_i-\bar{\lambda})^2}\right)^{1/2}.$$

Then, it is easy to check that the random vector

$$(X_1,X_2,\ldots,X_n)'=(Y_{\pi(1)},Y_{\pi(2)},\ldots,Y_{\pi(n)})',$$

where the vector $(\pi(1),\pi(2),\ldots,\pi(n))'$ is stochastically independent of Z and uniformly distributed over the $n!$ permutations of $\{1,2,\ldots,n\}$, satisfies all the conditions of the lemma and, moreover, attains the equality in (2.1). This shows that (2.1) is optimal. The case where the λ's form a nonincreasing sequence can be treated similarly. $\square$

The general case for an arbitrary sample can be derived from this lemma as follows.

Theorem 2.1. *For an arbitrary random vector $\mathbf{X}=(X_1,X_2,\ldots,X_n)'$ with mean vector $(\mu_1,\mu_2,\ldots,\mu_n)'$ and variance–covariance matrix (σ_{ij}), and for arbitrary real constants $\lambda_1,\lambda_2,\ldots,\lambda_n$, we have*

$$\left|\sum_{i=1}^n\lambda_i(\mathbb{E}\,X_{i:n}-\bar{\mu})\right|\leqslant\left(\sum_{i=1}^n(\lambda_i-\bar{\lambda})^2\right)^{1/2}\left(\sum_{i=1}^n\{(\mu_i-\bar{\mu})^2+\sigma_i^2\}-n\operatorname{Var}\bar{X}\right)^{1/2}, \tag{2.2}$$

where $X_{1:n}\leqslant X_{2:n}\leqslant\cdots\leqslant X_{n:n}$ are the order statistics corresponding to $\mathbf{X}$, $\sigma_i^2=\sigma_{ii}$, $i=1,2,\ldots,n$, $\bar{\mu}=n^{-1}\sum_{i=1}^n\mu_i$, $\bar{\lambda}=n^{-1}\sum_{i=1}^n\lambda_i$ and $\bar{X}=n^{-1}\sum_{i=1}^n X_i$.

Proof. If $n=1$ the result is obvious; for $n\geqslant 2$, set

$$(Y_1,Y_2,\ldots,Y_n)'=(X_{\pi(1)},X_{\pi(2)},\ldots,X_{\pi(n)})',$$

where $(\pi(1),\pi(2),\ldots,\pi(n))'$ is independent of $\mathbf{X}$ and distributed as in Lemma 2.1. It is easy to check that the Y's are exchangeable with

$$\mathbb{E}\,Y_1=\bar{\mu},\quad \operatorname{Var} Y_1=\frac{1}{n}\sum_{i=1}^n\{(\mu_i-\bar{\mu})^2+\sigma_i^2\},$$

and

$$\begin{aligned}\operatorname{Cov}[Y_1,Y_2]&=\frac{1}{n(n-1)}\left(2\sum_{1\leqslant i<j\leqslant n}\sigma_{ij}-\sum_{i=1}^n(\mu_i-\bar{\mu})^2\right)\\&=\frac{n}{n-1}\operatorname{Var}\bar{X}-\frac{1}{n-1}\operatorname{Var} Y_1.\end{aligned}$$

Since the ordered sample of the Y's coincides with that of the X's and

$$(n-1)(\operatorname{Var} Y_1 - \operatorname{Cov}[Y_1, Y_2]) = n \operatorname{Var} Y_1 - n \operatorname{Var} \bar{X}$$
$$= \sum_{i=1}^{n} \{(\mu_i - \bar{\mu})^2 + \sigma_i^2\} - n \operatorname{Var} \bar{X},$$

the desired result follows from (2.1). □

Remark 2.1. Clearly, (2.2) is sharper than (1.1), and the two bounds coincide only in the trivial case where $\bar{X}$ is a.s. constant. Moreover, (2.2) is always better than (1.2), as can be seen from the relation (see (1.2))

$$\left[\min_{1\leqslant j\leqslant n} S_j^2 + \sum_{i=1}^{n}(\mu_i - \bar{\mu})^2\right] - \left[\sum_{i=1}^{n}\{(\mu_i - \bar{\mu})^2 + \sigma_i^2\} - n \operatorname{Var} \bar{X}\right]$$
$$= n \min_{1\leqslant j\leqslant n} \operatorname{Var}[\bar{X} - X_j],$$

which shows that bounds (1.2) and (2.2) coincide only in the trivial case where $\bar{X} - X_j$ is a.s. constant for some $j \in \{1, 2, \ldots, n\}$.

Remark 2.2. Using Nagaraja's (1981) results, one can give another proof of (2.2) as follows: Writing $S^2 = \sum_{i=1}^{n} [X_i - \bar{X}]^2$ as

$$S^2 = \sum_{i=1}^{n} [X_i - \bar{\mu}]^2 - n[\bar{X} - \bar{\mu}]^2$$
$$= \sum_{i=1}^{n} [X_i - \mu_i]^2 + \sum_{i=1}^{n}(\mu_i - \bar{\mu})^2 + 2\sum_{i=1}^{n}(X_i - \mu_i)(\mu_i - \bar{\mu}) - n[\bar{X} - \bar{\mu}]^2,$$

and taking expectations to the last expression, we have

$$\mathbb{E}\, S^2 = \sum_{i=1}^{n} \{(\mu_i - \bar{\mu})^2 + \sigma_i^2\} - n \operatorname{Var} \bar{X}, \tag{2.3}$$

which, combined with (1.5) and (1.3), yields (2.2).

The bound (2.1) is attainable only when the λ's form a monotone finite sequence. Therefore, in order to obtain the sharp result for all sequences, we need the following considerations due to Rychlik (see, for example, Rychlik, 1998):

For any real constants λ_j, $j = 1, 2, \ldots, n$, define $C(x)$, $0 \leqslant x \leqslant 1$, to be the greatest convex function such that $C(0) = 0$, and $C(j/n) \leqslant \sum_{i=1}^{j} \lambda_i$, and set $c_j = C(j/n) - C((j-1)/n)$, $j = 1, 2, \ldots, n$. Similarly, define $D(x)$, $0 \leqslant x \leqslant 1$, to be the smallest concave function satisfying $D(0) = 0$, and $D(j/n) \geqslant \sum_{i=1}^{j} \lambda_i$, and set $d_j = D(j/n) - D((j-1)/n)$, $j = 1, 2, \ldots, n$. By construction, $C(1) = D(1) = n\bar{\lambda}$, showing that (using an obvious notation) $\bar{c} = \bar{d} = \bar{\lambda}$. It can be shown that $(c_1, c_2, \ldots, c_n)'$ is the l^2-projection of the vector $(\lambda_1, \lambda_2, \ldots, \lambda_n)'$ onto the convex cone of componentwise nondecreasing vectors in $\mathbb{R}^n$, while $(d_1, d_2, \ldots, d_n)'$ is the corresponding projection onto the convex cone of componentwise nonincreasing vectors. Thus, the c's form a monotone nondecreasing finite sequence while the d's form a nonincreasing one; both c's and d's are determined from the λ's and, finally, the c's (d's) coincide with the λ's if and only if the finite sequence of λ's is nondecreasing (nonincreasing).

Using the above notation we can prove the following lemma.

Lemma 2.2. *Under the conditions of Lemma* 2.1,

$$\sum_{i=1}^{n} \lambda_i(\mathbb{E}X_{i:n} - \mu) \leqslant \left(\sum_{i=1}^{n}(c_i - \bar{\lambda})^2\right)^{1/2} ((n-1)(\sigma^2 - c))^{1/2}, \tag{2.4}$$

and

$$\sum_{i=1}^{n} \lambda_i(\mathbb{E}X_{i:n} - \mu) \geqslant - \left(\sum_{i=1}^{n}(d_i - \bar{\lambda})^2\right)^{1/2} \left((n-1)(\sigma^2 - c)\right)^{1/2}. \tag{2.5}$$

Both bounds are best possible.

Proof. We prove only (2.4), since the other part is similar.

From Eq. (40) in Rychlik (1998), it follows that for any real numbers $x_{1:n} \leqslant x_{2:n} \leqslant \cdots \leqslant x_{n:n}$,

$$\sum_{i=1}^{n} \lambda_i(x_{i:n} - \bar{x}) \leqslant \left(\sum_{i=1}^{n}(c_i - \bar{\lambda})^2\right)^{1/2} (s^2)^{1/2}, \tag{2.6}$$

where $\bar{x} = n^{-1}\sum_{i=1}^{n} x_{i:n}$, $s^2 = \sum_{i=1}^{n} [x_{i:n} - \bar{x}]^2$ (this is the discrete version of a result of Moriguti (1953); note also that a denominator of $n^{1/q}$ is missing in the RHS of Eqs. (24) and (40) of Rychlik (1998)). Moreover, equality in (2.6) occurs if and only if there exists some constant $A \geqslant 0$, such that

$$x_{i:n} - \bar{x} = A(c_i - \bar{\lambda}), \quad i = 1, 2, \ldots, n. \tag{2.7}$$

Therefore, with S as in (1.3), we have

$$\sum_{i=1}^{n} \lambda_i(X_{i:n} - \bar{X}) \leqslant \left(\sum_{i=1}^{n}(c_i - \bar{\lambda})^2\right)^{1/2} S.$$

Taking expectations in the last expression and using the first inequality in (1.3) and the fact that $\mathbb{E}S^2 = (n-1)(\sigma^2 - c)$ (see (2.3)), we conclude (2.4).

Regarding the case of equality, this is trivial if all the λ's are equal. Also, this is trivial if $\sigma^2 = c$ (which implies that the X's are all equal a.s.). Observe that if $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$, $\lambda_1 > \lambda_n$ and $\sigma^2 > c$, then $c_1 = c_2 = \cdots = c_n = \bar{\lambda}$, and the equality never holds in (2.4). However, when this is the case, the best possible upper bound for the LHS of (2.4) is 0; this can be easily seen from the following construction: Let $\varepsilon > 0$ be sufficiently small so that $p = n(n+1)\varepsilon^2/(12(\sigma^2 - c)) < 1$, and set $A = 12(\sigma^2 - c)/(\varepsilon n(n+1)) > 0$. Define $Y_j = Z + A(j - (n+1)/2)W$, $j = 1, 2, \ldots, n$, where Z is a random variable as in Lemma 2.1, and W is a Bernoulli 0–1 random variable with success probability p, independent of Z. Define $X_1, X_2, \ldots, X_n$ to be a random permutation of $Y_1, Y_2, \ldots, Y_n$, and observe that the X's are exchangeable and satisfy $\mathbb{E}X_1 = \mu$, $\mathrm{Var}\, X_1 = \sigma^2$, $\mathrm{Cov}[X_1, X_2] = c$, and $X_{i:n} - \bar{X} = A(i - (n+1)/2)W$ for all i;

thus, $\mathbb{E}X_{i:n} - \mu = A(i-(n+1)/2)p = (i-(n+1)/2)\varepsilon$, which implies that

$$\left|\sum_{i=1}^{n} \lambda_i(\mathbb{E}X_{i:n} - \mu)\right| = \varepsilon \left|\sum_{i=1}^{n} \lambda_i \left(i - \frac{n+1}{2}\right)\right| = \varepsilon M,$$

where M is a positive constant depending only on the λ's. Since ε is arbitrary, the LHS of (2.4) can be arbitrarily close to 0, and hence, (2.4) is optimal. In any other case we can proceed as in Lemma 2.1, taking $Y_j = Z + A(c_j - \bar{\lambda})$, $j = 1,2,\ldots,n$, where Z is as above, and

$$A = \left(\frac{(n-1)(\sigma^2 - c)}{\sum_{i=1}^{n}(c_i - \bar{\lambda})^2}\right)^{1/2}.$$

Choosing the X's to be a random permutation of the Y's, it is easy to check that $\mathbb{E}X_1 = \mu$, $\operatorname{Var} X_1 = \sigma^2$, $\operatorname{Cov}[X_1, X_2] = c$, $X_{i:n} - \bar{X} = A(c_i - \bar{\lambda})$ (nonrandom), and $S^2 = (n-1)(\sigma^2 - c)$ (nonrandom), and thus, (2.4) becomes an equality. $\square$

Remark 2.3. Since for the finite sequence $x_{i:n} = c_i - \bar{\lambda}$ the equality occurs in (2.6) (see (2.7)), an application of the ordinary Cauchy–Schwarz inequality yields

$$\begin{aligned}
\sum_{i=1}^{n}(c_i - \bar{\lambda})^2 &= \sum_{i=1}^{n} \lambda_i(c_i - \bar{\lambda}) \\
&= \sum_{i=1}^{n}(\lambda_i - \bar{\lambda})\,(c_i - \bar{\lambda}) \\
&\leqslant \left(\sum_{i=1}^{n}(\lambda_i - \bar{\lambda})^2\right)^{1/2} \left(\sum_{i=1}^{n}(c_i - \bar{\lambda})^2\right)^{1/2},
\end{aligned}$$

showing that $\sum_{i=1}^{n}(c_i - \bar{\lambda})^2 \leqslant \sum_{i=1}^{n}(\lambda_i - \bar{\lambda})^2$, and the inequality is strict unless $\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n$. It follows that (2.4) is uniformly better than the upper bound of (2.1). Similarly, (2.5) is uniformly better than the lower bound (with the $-$ sign) of (2.1).

Our most general result is stated in the following theorem. The proof follows the same arguments as in Theorem 2.1, the only difference being that one has to use Lemma 2.2 rather than Lemma 2.1; the details are left to the reader.

Theorem 2.2. *Under the assumptions of Theorem* 2.1 *and the notation of Lemma* 2.2,

$$\sum_{i=1}^{n} \lambda_i(\mathbb{E}X_{i:n} - \bar{\mu}) \leqslant \left(\sum_{i=1}^{n}(c_i - \bar{\lambda})^2\right)^{1/2} \left(\sum_{i=1}^{n}\left\{(\mu_i - \bar{\mu})^2 + \sigma_i^2\right\} - n\operatorname{Var}\bar{X}\right)^{1/2}, \tag{2.8}$$

and

$$\sum_{i=1}^{n} \lambda_i(\mathbb{E}X_{i:n} - \bar{\mu}) \geqslant -\left(\sum_{i=1}^{n}(d_i - \bar{\lambda})^2\right)^{1/2} \left(\sum_{i=1}^{n}\left\{(\mu_i - \bar{\mu})^2 + \sigma_i^2\right\} - n\operatorname{Var}\bar{X}\right)^{1/2}. \tag{2.9}$$

It should be also noted that the above result improves the refinement of Lefèvre's inequality, described in Rychlik (1998, p. 114).

3. Applications

For the purposes of the present section, we consider an arbitrary random vector $\mathbf{X} = (X_1, X_2, \ldots, X_n)'$, with means, variances and covariances as in Theorem 2.1, and we set

$$\Delta = \Delta(\mu_i, \sigma_{ij}; i, j = 1, 2, \ldots, n) = \left(\sum_{i=1}^{n} \{(\mu_i - \bar{\mu})^2 + \sigma_i^2\} - n \operatorname{Var} \bar{X} \right)^{1/2},$$

which, under the homogeneity assumptions $\mu_i = \mu$, $\sigma_{ii} = \sigma^2$, $\sigma_{ij} = c$ $(i \neq j)$, $i, j = 1, 2, \ldots, n$, simplifies to

$$\Delta_0 = \Delta_0(\sigma^2, c) = \sqrt{(n-1)(\sigma^2 - c)} \leqslant \sigma \sqrt{n}.$$

It is useful to note that the inequality $\Delta \leqslant \sigma\sqrt{n}$ remains valid in the quasi-homogeneous case, in which we merely assume that $\mu_i = \mu$, $\sigma_i^2 = \sigma^2$ (i.e., σ_{ij} is allowed to vary with i, j).

(a) *Bounds for a single-order statistic and for the difference between two-order statistics*. In this case, Theorem 2.2 yields the bounds

$$\bar{\mu} - \Delta \sqrt{\frac{n-i}{ni}} \leqslant \mathbb{E} X_{i:n} \leqslant \bar{\mu} + \Delta \sqrt{\frac{i-1}{n(n-i+1)}}$$

(which improves Eq. (9) in Gascuel and Caraux (1992)), and for $i < j$,

$$0 \leqslant \mathbb{E}[X_{j:n} - X_{i:n}] \leqslant \Delta \sqrt{\frac{n+1-(j-i)}{i(n+1-j)}},$$

which improves Eq. (6) in Arnold and Groeneveld (1979).

(b) *Gupta's simple least-squares estimator for the location parameter of the uniform location-scale family*. If a random sample of size n follows the uniform $U(\mu - \sigma\sqrt{3}, \mu + \sigma\sqrt{3})$ model, where μ and $\sigma > 0$ are unknown, then the simple least-squares estimator of Gupta's (1952) for σ takes the form (see Eq. (6.2.14) in David (1981)):

$$\hat{\sigma}_n = \frac{2\sqrt{3}}{n(n-1)} \sum_{i=1}^{n} \left(\mathrm{i} - \frac{n+1}{2} \right) X_{i:n};$$

the quantity $\hat{\sigma}_n$ is a scalar multiple of Gini's mean difference and of Downton's (1966) unbiased estimator of the normal standard deviation (see Arnold and Groeneveld (1979)). It is easy to see that, under the homogeneity assumptions, Theorem 2.1 yields the bound

$$\mathbb{E}\,\hat{\sigma}_n \leqslant \sqrt{\frac{n+1}{n(n-1)}} \Delta_0 = \sqrt{\left(1 + \frac{1}{n}\right)(\sigma^2 - c)} \leqslant \sigma \sqrt{\frac{n+1}{n-1}}.$$

This shows that $\hat{\sigma}_n$ does not overestimate σ by a large amount, even if the X's are not independent, not identically distributed and not even uniformly distributed. The only assumption required is $\mu_i=\mu$, $\sigma_i^2=\sigma^2$ and $\sigma_{ij}=c$. Furthermore, even in the quasi-homogeneous case (in which σ_{ij} is allowed to vary with i,j), Theorem 2.1 (in fact, (1.1)) yields the estimate $\mathbb{E}\,\hat{\sigma}_n \leqslant \sigma((n+1)/(n-1))^{1/2}$, as in the corresponding example (c) of Arnold and Groeneveld (1979).

(c) *Lefèvre's bounds for the mean completion time in Pert networks*. Lefèvre (1986) applied (1.2) to obtain bounds on mean completion time for Pert networks. If the completion times of the arcs (activities) are the independent random variables $Y_{k,l}$, this is equivalent to find an expectation bound for the stochastic completion time $T=X_{n:n}$ of the network, when the random lengths are defined by

$$X_i = \sum_{(k,l)\in p_i} Y_{k,l}, \quad i=1,2,\dots,n,$$

where p_i, $1\leqslant i\leqslant n$, are the n network paths leading from the source to the sink (see Lefèvre for more details). Therefore, the means, variances and covariances of X's can be calculated from the means $m_{k,l}$ and the variances $s_{k,l}^2$ of Y's (assumed known), and thus, one may easily apply the results of the present work. As an example, we consider the network with the three paths $p_1=\{(1,2),(2,4)\}$, $p_2=\{(1,3),(3,4)\}$, $p_3=\{(1,2),(2,3),(3,4)\}$ and the five arcs $(1,2)$, $(1,3)$, $(2,3)$, $(2,4)$ and $(3,4)$, where 1 is the source and 4 is the sink node of the network, given in Table 1 of Lefèvre (in fact, two networks A and B were given, with identical sets of means and different sets of variances). With the given sets of values (see Lefèvre, 1986, p. 355), one finds that

$$\mu_A=\mu_B=\begin{pmatrix}5\\2\\5\end{pmatrix},\quad \Sigma_A=\begin{pmatrix}2.21&0&1\\0&0.72&0.36\\1&0.36&2\end{pmatrix},\quad \Sigma_B=\begin{pmatrix}1.28&0&0.64\\0&0.72&0.36\\0.64&0.36&1.36\end{pmatrix}.$$

If T_A and T_B denote the respective completion times of the systems, Lefèvre showed (with the help of (1.1), (1.2) and a slight variation of (1.2)) that $\mathbb{E}\,T_A\leqslant 6.609$ and $\mathbb{E}\,T_B\leqslant 6.347$; however, one can see that (2.2) yields the bounds $\mathbb{E}\,T_A\leqslant 6.364$ and $\mathbb{E}\,T_B\leqslant 6.093$.

(d) *Papathanasiou's bound on the covariance of an ordered pair*. Papathanasiou (1990) obtained an upper bound for the covariance of $X_{1:2}$ and $X_{2:2}$, when X_1 and X_2 are independent identically distributed with (common) variance σ^2, namely $\mathrm{Cov}[X_{1:2},X_{2:2}]\leqslant\sigma^2/3$. Balakrishnan and Balasubramanian (1993) obtained a generalization of this bound, namely

$$\begin{aligned}\mathrm{Cov}[X_{1:2},X_{2:2}] &= (\mathbb{E}X_{2:2}-\mu)^2+c &(3.1)\\ &\leqslant \sigma^2+c, &(3.2)\end{aligned}$$

provided that X_1 and X_2 are identically distributed (possibly dependent) with $\mathbb{E}X_1=\mu$, $\mathrm{Var}\,X_1=\sigma^2$ and $\mathrm{Cov}[X_1,X_2]=c$. Since (3.1) holds in the most general homogeneous case, Theorem 2.2 yields

$$c\leqslant \mathrm{Cov}[X_{1:2},X_{2:2}]\leqslant(\sigma^2+c)/2, \tag{3.3}$$

and the upper bound in (3.3) is the one half of that in (3.2). Moreover, we observe that in the uncorrelated case ($c=0$), $\sigma^2/3$ is simply replaced by $\sigma^2/2$ in Papathanasiou's bound, if neither independent nor identical distributions can be assumed for X_1 and X_2.

(e) *A criterion for infinite nonexchangeability*. Consider a set of $n \geqslant 2$ exchangeable random variables $X_1, X_2, \ldots, X_n$, with $\mathbb{E}X_1 = \mu$, $\operatorname{Var} X_1 = \sigma^2$ and $\operatorname{Cov}[X_1, X_2] = c$. It is well known that if $c < 0$, then the X's cannot be extended to an infinite sequence of exchangeable random variables (this happens because $\operatorname{Var}\bar{X} = (\sigma^2 + (n-1)c)/n$ has to be nonnegative for all n). If $c \geqslant 0$, however, the problem of extendibility (or not) becomes nontrivial. Our results can give a negative answer in some particular cases where $c \geqslant 0$. Indeed, from Lemma 2.1 we have the inequality

$$\mathbb{E}X_{n:n} \leqslant \mu + \frac{n-1}{\sqrt{n}} \sqrt{\sigma^2 - c}. \tag{3.4}$$

On the other hand, if we assume that the X's can be extended to an infinite exchangeable sequence, then de Finetti theorem (see, for example, Galambos, 1978, Theorem 3.6.1, or Arnold et al., 1992, Eq. (9.7.1)) asserts that there exists a random variable V such that for all $x_1, x_2, \ldots, x_n$,

$$\mathbb{P}[X_1 \leqslant x_1, X_2 \leqslant x_2, \ldots, X_n \leqslant x_n \mid V] = \prod_{i=1}^{n} \mathbb{P}[X_i \leqslant x_i \mid V] \quad \text{a.s.}, \tag{3.5}$$

that is, given V, the X's are independent and identically distributed. Therefore, the classical Hartley–David–Gumbel bound (see, for example, David, 1981, Chapter 4) yields

$$\mathbb{E}[X_{n:n} \mid V] \leqslant \mathbb{E}[X_1 \mid V] + \frac{n-1}{\sqrt{2n-1}} \sqrt{\operatorname{Var}[X_1 \mid V]} \quad \text{a.s.}$$

Taking expectations in the last expression with respect to V, and using the concavity of the square root and the fact that $\mathbb{E}\operatorname{Var}[X_1 \mid V] = \sigma^2 - \operatorname{Var}\mathbb{E}[X_1 \mid V] = \sigma^2 - c$, we obtain

$$\mathbb{E}X_{n:n} \leqslant \mu + \frac{n-1}{\sqrt{2n-1}} \sqrt{\sigma^2 - c}. \tag{3.6}$$

The upper bound in (3.4) is strictly larger than that of (3.6), because for the latter we assumed infinite exchangeability; therefore, all the constructions attaining equalities in Lemmas 2.1 and 2.2 lead to exchangeable random vectors that are not infinitely exchangeable. In other words, if $\mathbb{E}X_{n:n}$ is greater than the upper bound in (3.6), then the X's cannot be infinitely exchangeable; this provides a very simple criterion. As an example, consider the random variables X, Y and Z supported in $\{0,1\}^3$ with probability function

$$\mathbb{P}[X = x, Y = y, Z = z] = \begin{cases} 1/4 & \text{if } x+y+z \text{ is odd}, \\ 0 & \text{if } x+y+z \text{ is even}, \end{cases} \qquad x, y, z \in \{0,1\}.$$

It is easy to see that X, Y and Z are exchangeable and, moreover, X and Y are independent Bernoulli with $p = \frac{1}{2}$. Thus, $\mu = \frac{1}{2}$, $\sigma^2 = \frac{1}{4}$ and $c = 0$. The bound (3.4) is $\frac{1}{2} + 1/\sqrt{3}$, while (3.6) is $\frac{1}{2} + 1/\sqrt{5}$. Since $\mathbb{E}\max\{X, Y, Z\} = 1$ and $\frac{1}{2} + 1/\sqrt{5} < 1 < \frac{1}{2} + 1/\sqrt{3}$, it follows that the X, Y and Z are not infinitely exchangeable.

References

Arnold, B.C., Balakrishnan, N., 1989. Relations, Bounds and Approximations for Order Statistics. Lecture Notes in Statistics, Vol. 53, Springer, New York.

Arnold, B.C., Balakrishnan, N., Nagaraja, H.N., 1992. A First Course in Order Statistics. Wiley, New York.

Arnold, B.C., Groeneveld, R.A., 1979. Bounds on expectations of linear systematic statistics based on dependent samples. Ann. Statist. 7, 220–223; Correction 8 1401.

Aven, T., 1985. Upper (lower) bounds on the mean of the maximum (minimum) of a number of random variables. J. Appl. Probab. 22, 723–728.

Balakrishnan, N., Balasubramanian, K., 1993. Equivalence of Hartley–David–Gumbel and Papathanasiou bounds and some further remarks. Statist. Probab. Lett. 16, 39–41.

David, H.A., 1981. Order Statistics, 2nd edition. Wiley, New York.

Downton, F., 1966. Linear estimates with polynomial coefficients. Biometrika 53, 129–141.

Gajek, L., Rychlik, T., 1996. Projection method for moment bounds on order statistics from restricted families. I. Dependent case. J. Multivariate Anal. 57, 156–174.

Gajek, L., Rychlik, T., 1998. Projection method for moment bounds on order statistics from restricted families. II. Independent case. J. Multivariate Anal. 64, 156–182.

Galambos, J., 1978. The Asymptotic Theory of Extreme Order Statistics. Wiley, New York.

Gascuel, O., Caraux, G., 1992. Bounds on expectations of order statistics via extremal dependences. Statist. Probab. Lett. 15, 143–148.

Gupta, A.K., 1952. Estimation of the mean and standard deviation of a normal population from a censored sample. Biometrika 39, 260–273.

Lefèvre, C., 1986. Bounds on the expectations of linear combinations of order statistics with applications to Pert networks. Stochastic Anal. Appl. 4, 351–356.

Moriguti, S., 1953. A modification of Schwarz's inequality with applications to distributions. Ann. Math. Statist. 24, 107–113.

Nagaraja, H.N., 1981. Some finite sample results for the selection differential. Ann. Inst. Statist. Math. 33, 437–448.

Papathanasiou, V., 1990. Some characterizations of distributions based on order statistics. Statist. Probab. Lett. 9, 145–147.

Rychlik, T., 1993a. Bounds for expectations of L-estimates for dependent samples. Statistics 24, 1–7.

Rychlik, T., 1993b. Sharp bounds on L-estimates and their expectations for dependent samples. Commun. Statist.–Theory Meth. 22, 1053–1068; Erratum 23, 305–306.

Rychlik, T., 1994. Distributions and expectations of order statistics for possibly dependent random variables. J. Multivariate Anal. 48, 31–42.

Rychlik, T., 1995. Bounds for order statistics based on dependent variables with given nonidentical distributions. Statist. Probab. Lett. 23, 351–358.

Rychlik, T., 1998. Bounds for expectations of L-estimates. In: Balakrishnan, N., Rao, C.R. (Eds.), Order Statistics: Theory & Methods, Handbook of Statistics–16. Elsevier, North-Holland, Amsterdam, pp. 105–145.