In: Recent Developments in Ordered Random Variables ISBN 1-60021-301-4
Editors: M. Ahsanullah and M. Raqab, pp. 39-53

Chapter 3

ON RYCHLIK'S EXPECTATION BOUND FOR L-ESTIMATES BASED ON IDENTICALLY DISTRIBUTED VARIATES*

Nickos Papadatos†
Department of Mathematics, Section of Statistics and O.R.,
University of Athens, Panepistemiopolis, 157 84 Athens, Greece

Abstract

We provide an alternative proof for the best possible inequality on the expectation of any linear combination of order statistics based on dependent samples with identical marginals, originally proved by Rychlik (1993), *Statistics*, vol. **24**, pp. 1–7, and we present a generalization.

MSC: Primary 62G30.

Key words and phrases: Linear combination of order statistics; Expectation bounds; Identical Marginals; Rychlik's inequality.

1 Introduction

For a fixed univariate distribution function F, we denote by $\mathcal{M}_n(F)$ the space of all n-variate random vectors $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ such that $\mathbf{P}(X_i \le x) = F(x)$ for all $x \in \mathbf{R}$ and $i = 1, 2, \ldots, n$. The order statistics $X_{1:n} \le X_{2:n} \le \cdots \le X_{n:n}$ obtained from $\boldsymbol{X}$ will be denoted by $\boldsymbol{X}_{(\,)} := (X_{1:n}, X_{2:n}, \ldots, X_{n:n})$. Let $\mathbf{c} = (c_1, c_2, \ldots, c_n) \in \mathbf{R}^n$ be an arbitrary vector of scalars. Any linear combination on order statistics (i.e., any L-estimate) has the form

$$L(\mathbf{c}; \boldsymbol{X}) = \sum_{i=1}^{n} c_i X_{i:n},$$

*Work partially supported by the University of Athens' Research fund under Grant 70/4/5637.

†E-mail address: npapadat@math.uoa.gr

and Rychlik (1993a) was the first to obtain an explicit formula for the best possible expectation bound,

$$\max_{\boldsymbol{X}\in\mathcal{M}_n(F)} \mathbf{E}L(\boldsymbol{c};\boldsymbol{X}),$$

in terms of F and $\boldsymbol{c}$. Clearly, $\min_{\boldsymbol{X}} \mathbf{E}L(\boldsymbol{c};\boldsymbol{X}) = -\max_{\boldsymbol{X}} \mathbf{E}L(-\boldsymbol{c};\boldsymbol{X})$, so that the lower bounds can be immediately obtained from the upper ones; thus, Rychlik's work unified and extended several previous results; see, e.g., Arnold (1980, 1985, 1988).

In order to describe Rychlik's inequality, we first define some useful functions:

Definition 1.1 (i) For a univariate distribution function F, define its left-continuous inverse by

$$F^{-1}(u) = \min\{x : F(x) \ge u\}, \ \ 0 < u < 1.$$

(ii) Let $C(x)$, $x \in [0,1]$ be the piecewise linear function satisfying $C(0) = 0$, $C(j/n) = c_1 + \cdots + c_j$ for $j = 1,2,\ldots,n$, and $C(x)$ is linear in any interval of the form $[(i-1)/n, i/n]$, i.e,

$$C(x) = (nx - i + 1)c_i + \sum_{j=1}^{i-1} c_j, \ \ \frac{i-1}{n} \le x \le \frac{i}{n}, \ \ i = 1,2,\ldots,n.$$

(iii) Let $D(x)$, $0 \le x \le 1$ be the greatest convex minorant of $C(x)$, i.e., the greatest convex function such that $D(x) \le C(x)$; equivalently, D is the greatest convex function satisfying $D(0) = 0$ and $D(j/n) \le c_1 + \cdots + c_j$, $j = 1,2,\ldots,n$. Define $\boldsymbol{d} = (d_1, d_2, \ldots, d_n) \in \mathbf{R}^n$ by $d_j = D(j/n) - D((j-1)/n)$, $j = 1,2,\ldots,n$.

Clearly, the function $D(x)$ is continuous and piecewise linear in $[0,1]$, $d_1 \le d_2 \le \cdots \le d_n$ and $C(1) = \sum_{i=1}^n c_i = \sum_{i=1}^n d_i = D(1)$. It can be shown (see Remark 2.1, below) that $\boldsymbol{d}$ is the l^2-projection (with respect to the usual norm) of $\boldsymbol{c}$ onto the closed convex cone of componentwise nondecreasing vectors in $\mathbf{R}^n$, i.e.,

$$\min_{x_1 \le x_2 \le \cdots \le x_n} \sum_{i=1}^n (c_i - x_i)^2 = \sum_{i=1}^n (c_i - d_i)^2.$$

For construction purposes, it can be also shown (see e.g., Papadatos and Rychlik, 2004; Robertson et al., 1988, Section 1.2) that the vector $\boldsymbol{d}$ can be obtained with the help of the following procedure (Pool Adjacent Violators Algorithm–PAVA):

Step 1. Define $i_0 = 0$ and i_1 as the maximal index $j \in \{1,2,\ldots,n\}$ such that the arithmetic mean $(c_1 + \ldots + c_j)/j$ is minimal, i.e,

$$i_1 := \max\left\{ j \in \{1,2,\ldots,n\} : \frac{c_1 + \cdots + c_j}{j} \le \frac{c_1 + \cdots + c_k}{k} \text{ for all } k \in \{1,2,\ldots,n\} \right\}$$

and set

$$d_1 = \cdots = d_{i_1} := \frac{c_1 + \cdots + c_{i_1}}{i_1}.$$

Step 2. Suppose that the indices $i_1, i_2, \ldots, i_m$ have been defined, with $1 \le i_1 < i_2 < \cdots < i_m \le n$, and the corresponding d's are as follows:

$$d_1 = \cdots = d_{i_1} < d_{i_1+1} = \cdots = d_{i_2} < \cdots < d_{i_{m-1}+1} = \cdots = d_{i_m}.$$

If $i_m < n$, then define the index i_{m+1} as

$$i_{m+1} := \max\left\{ j \in \{i_m+1,\ldots,n\} : \frac{c_{i_m+1}+\cdots+c_j}{j-i_m} \le \frac{c_{i_m+1}+\cdots+c_k}{k-i_m} \text{ for all } k \in \{i_m+1,\ldots,n\} \right\},$$

and set

$$d_{i_m+1} = \cdots = d_{i_{m+1}} := \frac{c_{i_m+1}+\cdots+c_{i_{m+1}}}{i_{m+1}-i_m}.$$

Repeat **Step 2** until $i_m = n$.

Remark 1.1 It is clear that the above procedure defines ν indices $1 \le i_1 < i_2 < \cdots < i_\nu = n$, so that the projection $\boldsymbol{d}$ of $\mathbf{c}$ is of the form

$$d_1 = \cdots = d_{i_1} < d_{i_1+1} = \cdots = d_{i_2} < \cdots < d_{i_{\nu-1}+1} = \cdots = d_{i_\nu}.$$

It should be also noted that the points $i_0/n = 0 < i_1/n < \cdots < i_{\nu-1}/n < i_\nu/n = 1$ are exactly the knots in the graph of the function $D(x)$, $x \in [0,1]$, given in Definition 1.1(iii). Specifically, the explicit form of the function D is

$$D(x) = nd_{i_j}x - \sum_{k=1}^{j-1} i_k(d_{i_{k+1}} - d_{i_k}), \text{ for } x \in \left[\frac{i_{j-1}}{n}, \frac{i_j}{n}\right], \ j = 1,\ldots,\nu,$$

and it is useful to observe that

$$\{i_0/n = 0, i_1/n, \ldots, i_{\nu-1}/n, i_\nu/n = 1\} \subseteq \{x \in [0,1] : D(x) = C(x)\}.$$

Rychlik's expectation bound is now stated in the following Theorem.

Theorem 1.1 (Rychlik, 1993a). (i) *If* $\boldsymbol{X} = (X_1, X_2, \ldots, X_n) \in \mathcal{M}_n(F)$, $\boldsymbol{c} \in \mathbf{R}^n$ *and* $\mathbf{E}|X_1| < \infty$, *then*

$$\mathbf{E}\left(\sum_{i=1}^n c_i X_{i:n}\right) \le n \sum_{i=1}^n d_i \int_{(i-1)/n}^{i/n} F^{-1}(u)du. \tag{1}$$

(ii) *For any fixed* $\mathbf{c} \in \mathbf{R}^n$ *and any fixed univariate distribution* F *with finite first moment, there exists some* $\boldsymbol{X} \in \mathcal{M}_n(F)$ *attaining the equality in* (1).

Note that the RHS of (1) can be rewritten as

$$\int_0^1 F^{-1}(u)dD(u) = n\sum_{j=1}^{\nu} d_{i_j} \int_{i_{j-1}/n}^{i_j/n} F^{-1}(u)du,$$

and it is finite because $\mathbf{E}|X_1| = \int_0^1 |F^{-1}(u)|du$ has been assumed finite.

The crucial step in the original proof of Theorem 1.1 is to show that for any fixed $x \in \mathbf{R}$,

$$\min_{\boldsymbol{X} \in \mathcal{M}_n(F)} \sum_{i=1}^n c_i \mathbf{P}(X_{i:n} \le x) = D(F(x)),$$

and then, to find a particular $\boldsymbol{X} \in \mathcal{M}_n(F)$ that attains the minimum uniformly in $x \in \mathbf{R}$. In the present article we provide a modified proof that makes use of a known deterministic inequality (Lemma 2.1), in combination with a stochastic ordering one (Lemma 2.3). The crucial inequality for the present approach is given in Lemma 2.4 (see (10), below), and some simple by-products of this approach are discussed in Section 4.

2 Some Auxiliary Facts

We first state some auxiliary facts; most are known, but we provide the proofs for completeness.

Lemma 2.1 (Rychlik, 1992a). *If $x_1 \le x_2 \le \cdots \le x_n$ then*

$$\sum_{i=1}^{n} c_i x_i \le \sum_{i=1}^{n} d_i x_i, \tag{2}$$

and the equality in (2) holds for any vector $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ with $x_1 \le x_2 \le \cdots \le x_n$ satisfying

$$x_{i_{j-1}+1} = \cdots = x_{i_j}, \text{ for } j = 1, \ldots \nu. \tag{3}$$

Proof: Let $y_1 = x_1 \in \mathbf{R}$, $y_2 = x_2 - x_1 \ge 0, \ldots, y_n = x_n - x_{n-1} \ge 0$. Since $C(1) = c_1 + \cdots + c_n = d_1 + \cdots + d_n = D(1)$ and $D(x) \le C(x)$, we get

$$\begin{aligned}
\sum_{i=1}^{n} c_i x_i &= \sum_{i=1}^{n} c_i \sum_{j=1}^{i} y_j \\
&= \sum_{j=1}^{n} y_j \sum_{i=j}^{n} c_i \\
&= y_1 C(1) + \sum_{j=2}^{n} y_j \left(C(1) - C\left(\frac{j-1}{n}\right)\right) \\
&\le y_1 D(1) + \sum_{j=2}^{n} y_j \left(D(1) - D\left(\frac{j-1}{n}\right)\right) \qquad (4) \\
&= y_1 D(1) + \sum_{j=2}^{n} y_j \sum_{i=j}^{n} d_i \\
&= \sum_{j=1}^{n} y_j \sum_{i=j}^{n} d_i \\
&= \sum_{i=1}^{n} d_i \sum_{j=1}^{i} y_j \\
&= \sum_{i=1}^{n} d_i x_i,
\end{aligned}$$

and inequality (2) holds. Regarding conditions for equality, observe that any vector of the form (3) satisfies the following: If $x_{j-1} < x_j$ for some $j \in \{2, \ldots, n\}$ then $(j-1)/n \in$

$(0, 1)$ is a knot in the graph of D, i.e., $j \in \{i_1+1, \ldots, i_{\nu-1}+1\}$. Obviously, the inequality (4) (equivalent to (2)) becomes equality if and only if $y_j = 0$ for all $j \in \{2, \ldots, n\}$ for which $D((j-1)/n) < C((j-1)/n)$. Since $D((j-1)/n) < C((j-1)/n)$ implies that $(j-1)/n$ is not a knot in the graph of D, it is clear that any vector of the form (3) attains the equality in (2).

Remark 2.1 (i) The vector $\boldsymbol{d}$ itself satisfies (3) and hence attains the equality in (2), that is

$$\sum_{i=1}^{n} c_i d_i = \sum_{i=1}^{n} d_i^2.$$

According to Theorem 1 in Rychlik (2001), this equality combined with the inequality (2) shows that the vector $\boldsymbol{d}$ is the l^2-projection of the vector $\boldsymbol{c}$ on to the closed convex cone of componentwise nondecreasing vectors in $\mathbf{R}^n$.

(ii) The function D is linear if and only if $\nu = 1$. In this case we have $D(x) = (c_1 + c_2 + \cdots + c_n)x$, $x \in [0, 1]$, and the projection $\boldsymbol{d}$ is the constant vector $(\bar{c}, \bar{c}, \ldots, \bar{c})$, where $\bar{c} = (c_1 + c_2 + \cdots + c_n)/n$. On the other hand, $D(x) = C(x)$ for all $x \in [0, 1]$ if and only if $c_1 \le c_2 \le \cdots \le c_n$, in which case $\boldsymbol{d} = \boldsymbol{c}$.

(iii) In the general case, condition (3) is sufficient but not necessary for the equality in (2). Clearly, the necessary and sufficient condition for the equality in (2) is that

$$x_{j-1} = x_j \text{ for all } j \in \{2, \ldots, n\} \text{ such that } D((j-1)/n) < C((j-1)/n). \tag{5}$$

It can be shown (see Rychlik, 2005) that the following explicit construction, equivalent to the algorithm PAVA described in the previous section, can be used in order to specify all vectors $\boldsymbol{x}$ attaining equality in (2): Replace the maximum operator in Steps 1 and 2 by the minimum one, and define in a similar way the indices $1 \le \kappa_1 < \kappa_2 < \cdots < \kappa_m = n$. It then follows that $1 \le \nu \le m \le n$ and

$$\{i_1, i_2, \ldots, i_{\nu-1}\} \subseteq \{\kappa_1, \kappa_2, \ldots, \kappa_{m-1}\}.$$

Writing $\kappa_0 = 0$ for convenience, it is easy to see that the set

$$\{0 = \kappa_0, \kappa_1, \ldots, \kappa_{m-1}, \kappa_m = n\}$$

consists of those integers $j \in \{0, 1, \ldots, n\}$ for which $D(j/n) = C(j/n)$. It then follows that (5) is equivalent to

$$x_{\kappa_{j-1}+1} = \cdots = x_{\kappa_j}, \text{ for } j = 1, \ldots, m. \tag{6}$$

Lemma 2.2 (Rychlik, 1993a). *Assume that $\boldsymbol{X} \in \mathcal{M}_n(F)$ and let $F_{i:n}$ be the distribution function of $X_{i:n}$, $i = 1, 2, \ldots, n$. Then,*

$$F_{1:n} \ge F_{2:n} \ge \cdots \ge F_{n:n} \tag{7}$$

and

$$F_{1:n} + F_{2:n} + \cdots + F_{n:n} = nF. \tag{8}$$

Proof: Just take expectations to the following obvious relationships, holding for all $x \in \mathbf{R}$:

$$I(X_{1:n} \le x) \ge I(X_{2:n} \le x) \ge \cdots \ge I(X_{n:n} \le x)$$

and

$$\sum_{i=1}^{n} I(X_{i:n} \le x) = \sum_{i=1}^{n} I(X_i \le x).$$

Lemma 2.3 (i) *If $F_1, F_2, \ldots, F_n$ and F are univariate distribution functions satisfying $F_1 + F_2 + \cdots + F_n = nF$ then, for all $x \in \mathbf{R}$ and $j \in \{1, 2, \ldots, n\}$,*

$$\frac{F_j(x) + \cdots + F_n(x)}{n - j + 1} \ge \max\left\{0, \frac{nF(x) - j + 1}{n - j + 1}\right\}. \tag{9}$$

(ii) *Assume that $\boldsymbol{X} \in \mathcal{M}_n(F)$ and consider the order statistics $X_{i:n}$, $i = 1, 2, \ldots, n$, from $\boldsymbol{X}$. For a fixed $j \in \{1, 2, \ldots, n\}$, consider a random variable I, independent of $\boldsymbol{X}$, and uniformly distributed over the set of integers $\{j, \ldots, n\}$. Let also U be a random variable with standard uniform distribution over the interval $(0, 1)$. Define the random variables X and Y by*

$$X := X_{I:n}, \quad Y := F^{-1}\left(\frac{j-1}{n} + \left(1 - \frac{j-1}{n}\right) U\right).$$

Then, for all $x \in \mathbf{R}$,

$$\mathbf{P}(X \le x) \ge \mathbf{P}(Y \le x).$$

Proof: (i) We have

$$\begin{aligned} F_j(x) + \cdots + F_n(x) &= nF(x) - (F_1(x) + \cdots + F_{j-1}(x)) \\ &\ge nF(x) - (1 + \cdots + 1) \\ &= nF(x) - (j - 1), \end{aligned}$$

and (9) follows.

(ii) Let $F_i = F_{i:n}$ be the distribution function of $X_{i:n}$, $i = 1, 2, \ldots, n$. By Lemma 2.2, the distribution functions $F_{1:n}, F_{2:n}, \ldots, F_{n:n}$ and F satisfy (8), and hence, by part (i), they also satisfy (9). It suffices to observe that the distribution function of X equals to the LHS of (9), while that of Y equals to the RHS of (9).

Lemma 2.4 *If $\mathbf{E}|X_1| < \infty$, where $\boldsymbol{X} \in \mathcal{M}_n(F)$ and $X_{i:n}$, $i = 1, 2 \ldots, n$ are as in Lemma 2.3(ii), then for all $j \in \{2, \ldots, n\}$,*

$$\mathbf{E}(X_{j:n} + \cdots + X_{n:n}) \le n \int_{(j-1)/n}^{1} F^{-1}(u) du, \tag{10}$$

while

$$\mathbf{E}(X_{1:n} + \cdots + X_{n:n}) = n \int_0^1 F^{-1}(u) du. \tag{11}$$

Proof: Let $F_{i:n}$ be the distribution function of $X_{i:n}$, $i = 1, 2, \ldots, n$. Since $\mathbf{E}|X_1| = \int_0^1 |F^{-1}(u)|du < \infty$ and $|X_{i:n}| \le |X_1| + |X_2| + \cdots + |X_n|$, it follows that for any $i \in \{1, 2, \ldots, n\}$, $\mathbf{E}|X_{i:n}| = \int_0^1 |F_{i:n}^{-1}(u)|du \le n\mathbf{E}|X_1| < \infty$; thus, all the quantities appearing in (10) and (11) are finite. Since (11) is obvious, it suffices to verify (10). Fix $j \in \{2, \ldots, n\}$ and let X and Y be the random variables defined in Lemma 2.3(ii). Obviously,

$$\mathbf{E}X = \mathbf{E}X_{I:n} = \mathbf{E}\,\mathbf{E}(X_{I:n}|I) = \sum_{i=j}^{n} \mathbf{P}(I = i)\mathbf{E}(X_{i:n}) = \frac{1}{n-j+1}\sum_{i=j}^{n} \mathbf{E}X_{i:n}.$$

Also,

$$\mathbf{E}Y = \int_0^1 F^{-1}\left(\frac{j-1}{n} + \left(1 - \frac{j-1}{n}\right)u\right)du = \frac{n}{n-j+1}\int_{(j-1)/n}^{1} F^{-1}(u)du,$$

and since, by Lemma 2.3(ii), Y is stochastically greater than X, inequality (10) follows from $\mathbf{E}X \le \mathbf{E}Y$.

Remark 2.2 It is well-known (see, e.g., Arnold, 1980; Caraux and Gascuel, 1992; Gascuel and Caraux, 1992; Rychlik, 1992b; Papadatos, 2001b) that under the assumptions of Lemma 2.4,

$$\mathbf{E}X_{j:n} \le \frac{n}{n-j+1}\int_{(j-1)/n}^{1} F^{-1}(u)du. \tag{12}$$

The assertion of Lemma 2.4, i.e., the inequality

$$\mathbf{E}\left(\frac{X_{j:n} + \cdots + X_{n:n}}{n-j+1}\right) \le \frac{n}{n-j+1}\int_{(j-1)/n}^{1} F^{-1}(u)du,$$

provides a strengthened version of (12), because

$$X_{j:n} \le \frac{X_{j:n} + \cdots + X_{n:n}}{n-j+1}.$$

3 A Proof of Theorem 1.1, and the Corresponding Mean-Variance Bounds

We split the proof of Theorem 1.1 in two parts: in the first one we verify the inequality (1) and in the second one we construct a random vector $\boldsymbol{X} \in \mathcal{M}_n(F)$ for which (1) holds as an equality.

Proof of Inequality (1): From the obvious relation

$$\mathbf{E}X_{1:n} \le \mathbf{E}X_{2:n} \le \cdots \le \mathbf{E}X_{n:n}$$

(note that all the expectations above are finite) and Lemma 2.1 we get

$$\mathbf{E}L(c; \boldsymbol{X}) = \sum_{i=1}^{n} c_i \mathbf{E}X_{i:n} \le \sum_{i=1}^{n} d_i \mathbf{E}X_{i:n} = \mathbf{E}L(d; \boldsymbol{X}). \tag{13}$$

For $i = 1, 2, \dots, n$ set $d_i = \gamma_1 + \cdots + \gamma_i$, so that $\gamma_2 \ge 0, \dots, \gamma_n \ge 0$. Then,

$$\begin{aligned} L(\boldsymbol{d}; \boldsymbol{X}) &= \sum_{i=1}^{n} d_i X_{i:n} \\ &= \sum_{i=1}^{n} X_{i:n} \sum_{j=1}^{i} \gamma_j \\ &= \sum_{j=1}^{n} \gamma_j \sum_{i=j}^{n} X_{i:n} \\ &= n\gamma_1 \overline{X} + \sum_{j=2}^{n} \gamma_j \sum_{i=j}^{n} X_{i:n}, \end{aligned}$$

where $\overline{X} = (X_1 + \cdots + X_n)/n$. Using (10) and the fact that $\mathbf{E}\overline{X} = \mathbf{E}X_1$, we get

$$\mathbf{E}L(\boldsymbol{d}, \boldsymbol{X}) \le n\gamma_1 \int_0^1 F^{-1}(u)du + \sum_{j=2}^{n} n\gamma_j \int_{(j-1)/n}^{1} F^{-1}(u)du. \tag{14}$$

Now (1) follows from (13) and (14) if we rewrite the RHS of (14) as

$$n \sum_{j=1}^{n} \gamma_j \int_{(j-1)/n}^{1} F^{-1}(u)du = n \sum_{j=1}^{n} \gamma_j \sum_{i=j}^{n} \int_{(i-1)/n}^{i/n} F^{-1}(u)du$$

and change the order of summation.

Construction of a random vector $\boldsymbol{X} \in \mathcal{M}_n(F)$ that attains the equality in (1). Let $(U_1, U_2, \dots, U_\nu)$ be any random vector with standard uniform marginals, and assume that $(\pi(1), \pi(2), \dots, \pi(n))$ is independent of $(U_1, U_2, \dots, U_\nu)$ and has uniform distribution over the $n!$ permutations of $\{1, 2, \dots, n\}$. Define the random vector $(V_1, \dots, V_\nu)$ by $V_j := i_{j-1}/n + (i_j - i_{j-1})U_j/n$, $j = 1, \dots, \nu$, the random vector $(Y_1, Y_2, \dots, Y_n)$ by

$$Y_i := F^{-1}(V_j), \quad i = i_{j-1}+1, \dots, i_j, \quad j = 1, 2, \dots, \nu,$$

and set

$$\boldsymbol{X} := (Y_{\pi(1)}, Y_{\pi(2)}, \dots, Y_{\pi(n)}).$$

Since $Y_1 \le Y_2 \le \cdots \le Y_n$, it follows that $X_{i:n} = Y_i$, and thus, $\mathbf{E}X_{i:n} = \mathbf{E}Y_i$, $i = 1, 2, \dots, n$. On the other hand, $Y_{i_{j-1}+1} = \cdots = Y_{i_j}$, and therefore, (13) becomes an equality because of (3). Also, it is easy to see that

$$\mathbf{P}(F^{-1}(V_j) \le x) = \mathbf{P}(V_j \le F(x)) = \begin{cases} 1, & \text{for } j < k, \\ \dfrac{nF(x) - i_{k-1}}{i_k - i_{k-1}}, & \text{for } j = k, \\ 0, & \text{for } j > k, \end{cases}$$

where $k \in \{1, \dots, n\}$ is the unique integer satisfying $i_{k-1}/n < F(x) \le i_k/n$. Therefore, for any index $s \in \{1, 2, \dots, n\}$ and any $x \in \mathbf{R}$,

$$\mathbf{P}(X_s \le x) = \mathbf{P}(Y_{\pi(s)} \le x)$$

$$\begin{aligned}
&= \sum_{i=1}^{n} \mathbf{P}(\pi(s)=i)\mathbf{P}(Y_{\pi(s)} \le x | \pi(s)=i) \\
&= \frac{1}{n}\sum_{i=1}^{n} \mathbf{P}(Y_i \le x) \\
&= \frac{1}{n}\sum_{j=1}^{\nu} (i_j - i_{j-1})\mathbf{P}(F^{-1}(V_j) \le x) \\
&= \frac{1}{n}\sum_{j<k} (i_j - i_{j-1}) + \frac{1}{n}(i_k - i_{k-1})\mathbf{P}(F^{-1}(V_k) \le x) \\
&= \frac{i_{k-1}}{n} + \frac{1}{n}(i_k - i_{k-1})\frac{nF(x) - i_{k-1}}{i_k - i_{k-1}} \\
&= F(x),
\end{aligned}$$

which shows that $\boldsymbol{X} \in \mathcal{M}_n(F)$. Finally, it remains to prove that $\boldsymbol{X}$ attains the equality in (14), i.e., it suffices to show that for all $j = 1, 2, \ldots, n$,

$$\gamma_j \mathbf{E}(X_{j:n} + \cdots + X_{n:n}) = n\gamma_j \int_{(j-1)/n}^{1} F^{-1}(u)du. \tag{15}$$

Observe that (15) is satisfied if either $j = 1$ (see (11)) or $j \in \{2, \ldots, n\}$ and $\gamma_j = 0$. In the remaining case where $\gamma_j > 0$ for some $j \ge 2$ (that is, $d_j > d_{j-1}$), the point $(j-1)/n$ must be a knot in the graph of D. It then follows that $\nu \ge 2$ and $j = i_k + 1$ for some $k \in \{1, \ldots, \nu - 1\}$; thus,

$$\begin{aligned}
\mathbf{E}(X_{j:n} + \cdots + X_{n:n}) &= \mathbf{E}(Y_j + \cdots + Y_n) \\
&= \mathbf{E}(Y_{i_k+1} + \cdots + Y_{i_\nu}) \\
&= \sum_{s=k+1}^{\nu} (i_s - i_{s-1})\mathbf{E}F^{-1}(V_s) \\
&= \sum_{s=k+1}^{\nu} (i_s - i_{s-1}) \int_0^1 F^{-1}\left(\frac{i_{s-1}}{n} + \frac{i_s - i_{s-1}}{n}u\right) du \\
&= n \sum_{s=k+1}^{\nu} \int_{i_{s-1}/n}^{i_s/n} F^{-1}(u)du \\
&= n \int_{i_k/n}^{1} F^{-1}(u)du,
\end{aligned}$$

and since $i_k = j - 1$, (15) is again satisfied. This completes the proof.

Definition 3.1 Any random vector $\boldsymbol{X} \in \mathcal{M}_n(F)$ attaining the equality in (1) will be called extremal, and the collection of extremal vectors will be denoted by $\mathcal{E}_n(\boldsymbol{c}; F)$.

Using the notation given in Definition 3.1, the above construction shows that for all $\boldsymbol{c} \in \mathbf{R}^n$ and for all univariate distributions F with finite first moment, the set $\mathcal{E}_n(\boldsymbol{c}; F)$ is a nonempty subset of $\mathcal{M}_n(F)$.

Theorem 3.1 (Rychlik, 1993b). *If* $\boldsymbol{X} \in \mathcal{M}_n(F)$ *with* $\mathbf{E}X_1^2 < \infty$, *and if* $\mathbf{E}X_1 = \mu$ *and* $\mathrm{Var}\,X_1 = \sigma^2$, *then*

$$\mathbf{E}\left(\sum_{i=1}^{n} c_i X_{i:n}\right) \le \mu \sum_{i=1}^{n} c_i + \sigma \left(n \sum_{i=1}^{n} d_i^2 - \left(\sum_{i=1}^{n} c_i\right)^2\right)^{1/2}. \tag{16}$$

Moreover, if the projection $\mathbf{d}$ *of* $\mathbf{c}$ *is non-constant* (equivalently, if $\nu \ge 2$), *then the equality in* (16) *is attained if and only if* $\boldsymbol{X} \in \mathcal{E}_n(\mathbf{c}; F_0)$, *with* F_0 *being the distribution function of the random variable* X *with probability function given by*

$$\mathbf{P}\left(X = \mu + \sigma \frac{n d_{i_j} - \sum_{i=1}^{n} c_i}{\left(n \sum_{i=1}^{n} d_i^2 - \left(\sum_{i=1}^{n} c_i\right)^2\right)^{1/2}}\right) = \frac{i_j - i_{j-1}}{n}, \ j = 1, 2, \ldots, \nu. \tag{17}$$

Proof: We follow the arguments given in Rychlik (2001). If $\sigma^2 = 0$ the result is trivial. Assuming that $\sigma^2 > 0$, there is no loss of generality if we further assume that $\mu = 0$ and $\sigma^2 = 1$; equivalently, $\int_0^1 F^{-1}(u)du = 0$, $\int_0^1 (F^{-1}(u))^2 du = 1$. Inequality (1) of Theorem 1.1 can be rewritten as

$$\mathbf{E}\left(\sum_{i=1}^{n} c_i X_{i:n}\right) \le n \int_0^1 F^{-1}(u) g(u) du = n \int_0^1 F^{-1}(u)(g(u) - \bar{c}) du, \tag{18}$$

where $\bar{c} = (c_1 + c_2 + \cdots + c_n)/n = (d_1 + d_2 + \cdots + d_n)/n = \int_0^1 g(u)du$ and for $u \in (0,1)$, the left-continuous nondecreasing function g is given by

$$g(u) = \sum_{i=1}^{n} d_i I\left(\frac{i-1}{n} < u \le \frac{i}{n}\right).$$

Applying the Cauchy-Schwarz inequality to the RHS of (18) we get

$$n \int_0^1 F^{-1}(u)(g(u) - \bar{c}) du \le n \left(\int_0^1 (F^{-1}(u))^2 du\right)^{1/2} \left(\int_0^1 (g(u) - \bar{c})^2 du\right)^{1/2}, \tag{19}$$

and since

$$\int_0^1 (F^{-1}(u))^2 du = 1 \text{ and } \int_0^1 (g(u) - \bar{c})^2 du = \int_0^1 g^2(u) du - (\bar{c})^2 = \frac{1}{n} \sum_{i=1}^{n} d_i^2 - (\bar{c})^2,$$

inequality (16) follows. It is obvious that equality in (16) is attained if and only if we have equalities into both (18) and (19). Observe that equality in (18) means that $\boldsymbol{X} \in \mathcal{E}_n(c; F)$, while equality in (19) means that there exists a constant a such that

$$F^{-1}(u) = a(g(u) - \bar{c}) \text{ for almost all } u \in (0,1).$$

Now, since the projection $\boldsymbol{d}$ of $\boldsymbol{c}$ is non-constant, we have

$$n \sum_{i=1}^{n} d_i^2 > \left(\sum_{i=1}^{n} d_i\right)^2 = \left(\sum_{i=1}^{n} c_i\right)^2,$$

and the condition $\int_0^1 (F^{-1}(u))^2 du = 1$, together with the fact that F^{-1} has to be nondecreasing, yield

$$a = \frac{n}{\left(n\sum_{i=1}^n d_i^2 - (\sum_{i=1}^n c_i)^2\right)^{1/2}}.$$

Thus, equality in (19) implies that

$$F^{-1}(u) = \frac{n\sum_{i=1}^n d_i I\left(\frac{i-1}{n} < u \le \frac{i}{n}\right) - \sum_{i=1}^n c_i}{\left(n\sum_{i=1}^n d_i^2 - (\sum_{i=1}^n c_i)^2\right)^{1/2}};$$

and it is easily seen that this quantile function corresponds to the distribution function F_0 with probability function given by (17) with $\mu = 0, \sigma^2 = 1$. Since $\mathcal{E}_n(\boldsymbol{c}; F_0)$ is nonempty, the proof is complete.

4 A Generalization

It became clear from the proof of inequality (1) given in the previous section, that (10) and (11) are the crucial steps. In the present section we try to formulate the same approach to more general situations. To this end, we give the following definition.

Definition 4.1 Let $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_n)$ be an arbitrary random vector such that $\mathbf{E}|Y_j| < \infty, j = 1, 2, \ldots, n$. Also, assume that

$$\mathbf{E}Y_1 \le \mathbf{E}Y_2 \le \cdots \le \mathbf{E}Y_n.$$

Any function $G \in \mathcal{L}^1(0,1)$ satisfying

$$\frac{1}{n}\mathbf{E}(Y_j + \cdots + Y_n) \le \int_{(j-1)/n}^1 G(u)du, \ \ j = 2, \ldots, n, \tag{20}$$

and

$$\frac{1}{n}\mathbf{E}(Y_1 + \cdots + Y_n) = \int_0^1 G(u)du, \tag{21}$$

will be called as *mean-dominant of* $\boldsymbol{Y}$.

It is obvious that we can always find a (trivial) mean-dominant function, if we know the expectations $\mathbf{E}Y_j, j = 1, 2 \ldots, n$; e.g., put

$$G(u) = \sum_{i=1}^n I\left(\frac{i-1}{n} < u \le \frac{i}{n}\right)\mathbf{E}Y_i.$$

The interesting case, however, appears when a mean-dominant function is known for $\boldsymbol{Y}$, while $\mathbf{E}\boldsymbol{Y}$ is not. The following result shows that the mean-dominant functions are useful for establishing inequalities for the expectation of any linear combination in $\boldsymbol{Y}$.

Theorem 4.1 *Let* $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)$ *be an arbitrary random vector such that* $\mathbf{E}|Y_j| < \infty$, $j = 1, 2, \ldots, n$. *Also, assume that*

$$\mathbf{E}Y_1 \leq \mathbf{E}Y_2 \leq \cdots \leq \mathbf{E}Y_n. \tag{22}$$

If G *is any mean-dominant function of* $\mathbf{Y}$, *then*

$$\mathbf{E}\left(\sum_{i=1}^{n} c_i Y_i\right) \leq n \sum_{i=1}^{n} d_i \int_{(i-1)/n}^{i/n} G(u)du. \tag{23}$$

Proof: From (22) and Lemma 2.1 we get

$$\mathbf{E}\left(\sum_{i=1}^{n} c_i Y_i\right) = \sum_{i=1}^{n} c_i \mathbf{E}Y_i \leq \sum_{i=1}^{n} d_i \mathbf{E}Y_i. \tag{24}$$

As in the proof of inequality (1), for $i = 1, 2, \ldots, n$ set $d_i = \gamma_1 + \cdots + \gamma_i$, so that $\gamma_2 \geq 0, \ldots, \gamma_n \geq 0$. Then,

$$\begin{aligned} \sum_{i=1}^{n} d_i \mathbf{E}Y_i &= \sum_{j=1}^{n} \gamma_j \mathbf{E}\left(\sum_{i=j}^{n} Y_i\right) \\ &= n\gamma_1 \mathbf{E}\overline{Y} + \sum_{j=2}^{n} \gamma_j \mathbf{E}\left(\sum_{i=j}^{n} Y_i\right), \end{aligned}$$

where $\overline{Y} = (Y_1 + \cdots + Y_n)/n$. Since, by (21), $\mathbf{E}\overline{Y} = \int_0^1 G(u)du$, an application of (20) yields

$$\sum_{i=1}^{n} d_i \mathbf{E}Y_i \leq n\gamma_1 \int_0^1 G(u)du + \sum_{j=2}^{n} n\gamma_j \int_{(j-1)/n}^{1} G(u)du. \tag{25}$$

The desired inequality (23) follows from (24) and (25) if we rewrite the RHS of (25) as

$$n \sum_{j=1}^{n} \gamma_j \int_{(j-1)/n}^{1} G(u)du = n \sum_{j=1}^{n} \gamma_j \sum_{i=j}^{n} \int_{(i-1)/n}^{i/n} G(u)du$$

and change the order of summation.

It is now obvious that the basic inequality (1) is an immediate Corollary of Theorem 4.1 and the following Lemma, which is just a restatement of Lemma 2.4.

Lemma 4.1 *If* $\mathbf{X} \in \mathcal{M}_n(F)$ *and* $\mathbf{E}|X_1| < \infty$, *then* $G = F^{-1}$ *is a mean-dominant of the vector of order statistics* $\mathbf{X}_{(\,)} = (X_{1:n}, X_{2:n}, \ldots, X_{n:n})$.

Moreover, it is easy to extend inequality (1) for any random vector $\mathbf{X}$, without assuming identical marginals, as follows (see, also, Gascuel and Caraux, 1992; Rychlik, 1998):

Lemma 4.2 *Let* $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ *be any random vector with* $\mathbf{E}|X_i| < \infty$, $i = 1, 2, \ldots, n$, *and let* F_i *be the distribution function of* X_i, $i = 1, 2, \ldots, n$. *Let* I *be a random variable, independent of* $\mathbf{X}$, *and uniformly distributed over* $\{1, 2, \ldots, n\}$, *and consider the distribution function* $\overline{F} = (F_1 + \cdots + F_n)/n$, *which is the distribution function of the mixture random variable* X_I. *Then,* $(\overline{F})^{-1}$ *is a mean-dominant of the vector of order statistics* $\mathbf{X}_{(\,)} = (X_{1:n}, X_{2:n}, \ldots, X_{n:n})$.

Proof: Define the random vector $\boldsymbol{Y} = (Y_1, Y_2, \dots, Y_n) := (X_{\pi(1)}, X_{\pi(2)}, \dots, X_{\pi(n)})$, where $(\pi(1), \pi(2), \dots, \pi(n))$ is independent of $\boldsymbol{X}$ and uniformly distributed over the $n!$ permutations of $\{1, 2, \dots, n\}$, and let $\boldsymbol{Y}_{(\,)} := (Y_{1:n}, Y_{2:n} \dots, Y_{n:n})$ be the vector of order statistics of $\boldsymbol{Y}$. Clearly, $\boldsymbol{X}_{(\,)} = \boldsymbol{Y}_{(\,)}$, because the random vector $\boldsymbol{Y}$ is just a permutation of the random vector $\boldsymbol{X}$. Since $\pi(i) \stackrel{d}{=} I$ for all $i = 1, 2, \dots, n$, it follows that $Y_i \stackrel{d}{=} X_I$ for all $i = 1, 2, \dots, n$, and thus $\boldsymbol{Y} \in \mathcal{M}_n(\overline{F})$. Therefore, $(\overline{F})^{-1}$ is a mean-dominant of $\boldsymbol{Y}_{(\,)}$, by Lemma 4.1. Since $\boldsymbol{X}_{(\,)} = \boldsymbol{Y}_{(\,)}$, it is also a mean-dominant of $\boldsymbol{X}_{(\,)}$, and the proof is complete.

Combining Lemma 4.2 and Theorem 4.1, we get the following Corollary.

Corollary 4.1 *Let $\boldsymbol{X} = (X_1, X_2, \dots, X_n)$ be any random vector with $\mathbf{E}|X_i| < \infty$, $i = 1, 2, \dots, n$, and let F_i be the distribution function of X_i, $i = 1, 2, \dots, n$. Then,*

$$\mathbf{E}\left(\sum_{i=1}^{n} c_i X_{i:n}\right) \le n \sum_{i=1}^{n} d_i \int_{(i-1)/n}^{i/n} (\overline{F})^{-1}(u)du, \tag{26}$$

where $\overline{F} = (F_1 + \cdots + F_n)/n$ and $(X_{1:n}, X_{2:n}, \dots, X_{n:n})$ is the vector of order statistics of $\boldsymbol{X}$.

It is worth pointing out that bound (26) is not attained for arbitrary non-identical marginal distribution functions F_i; e.g. necessary and sufficient conditions on F_i, $i = 1, 2, \dots, n$ for the attainability of the bound (26) in the case of a single order statistic can be found in Rychlik (1995).

Finally, working as in Theorem 3.1 or in Lemma 4.2, we can easily show the following result.

Theorem 4.2 *Let $\boldsymbol{X} = (X_1, X_2, \dots, X_n)$ be any random vector with $\mathbf{E}X_i^2 < \infty$, $i = 1, 2, \dots, n$, and let $\mu_i = \mathbf{E}X_i$ and $\sigma_i^2 = \operatorname{Var} X_i$, $i = 1, 2, \dots, n$. Then*

$$\mathbf{E}\left(\sum_{i=1}^{n} c_i X_{i:n}\right) \le \bar{\mu} \sum_{i=1}^{n} c_i + \bar{\sigma}\left(n \sum_{i=1}^{n} d_i^2 - \left(\sum_{i=1}^{n} c_i\right)^2\right)^{1/2}, \tag{27}$$

where $\bar{\mu} := (\mu_1 + \ldots + \mu_n)/n$ and $\bar{\sigma}^2 := (1/n)\sum_{i=1}^{n} \{(\mu_i - \bar{\mu})^2 + \sigma_i^2\}$.

Proof: Using the notation as in the proof of Lemma 4.2, it follows that

$$\sum_{i=1}^{n} c_i X_{i:n} = \sum_{i=1}^{n} c_i Y_{i:n},$$

where $\boldsymbol{Y} \in \mathcal{M}_n(\overline{F})$. Therefore, since $\overline{F}$ is the distribution of X_I with mean $\bar{\mu}$ and variance $\bar{\sigma}^2$, the desired result follows from (16).

It should be noted that inequality (27) is an improved version of the main result given in Arnold and Groeneveld (1979); a further improvement of (27) can be found in Papadatos (2001a).

As a final observation, we mention the following: If $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ is an i.i.d. sample from F with finite first moment then, trivially, $\boldsymbol{X} \in \mathcal{M}_n(F)$. Thus, Lemma 4.1 applies to $\boldsymbol{X}$, showing that F^{-1} is a mean-dominant of $\boldsymbol{X}_{(\,)}$. Since $X_{i:n} \stackrel{\mathrm{d}}{=} F^{-1}(U_{i:n})$ with $U_{i:n}$ being a Beta$(i, n+1-i)$ random variable, (20) yields the inequality

$$\int_0^1 F^{-1}(u) \left(\sum_{i=j}^{n} \frac{n!}{(i-1)!(n-i)!} u^{i-1}(1-u)^{n-i} \right) du \le n \int_{(j-1)/n}^{1} F^{-1}(u)du,$$

satisfied by any $j \in \{1, 2, \ldots, n\}$ and any nondecreasing function $F^{-1} \in \mathcal{L}^1(0,1)$. Of course, due to the binomial formula, the above inequality becomes an identity when $j = 1$ (see, also, (21)), but it does not seem to be obvious for $j \ge 2$. Note that, using the well-known relationship between binomial and Beta distribution functions, this inequality can be rewritten, for $j = 2, \ldots, n$, as

$$\frac{(n-1)!}{(j-2)!(n-j)!} \int\int_{0<u<v<1} u^{j-2}(1-u)^{n-j} F^{-1}(v) dv du \le \int_{(j-1)/n}^{1} F^{-1}(u)du.$$

References

[1] Arnold, B.C. (1980). Distribution-free bounds on the mean of the maximum of a dependent sample. *SIAM J. Appl. Math.* **38**, 163–167.

[2] Arnold, B.C. (1985). p-Norm bounds on the expectation of the maximum of possibly dependent sample. *J. Mulivariate Anal.* **17**, 316–332.

[3] Arnold, B.C. (1988). Bounds on the expected maximum. *Commun. Statist.–Theory Meth.* **17**, 2135–2150.

[4] Arnold, B.C. and Groeneveld, R.A. (1979). Bounds on expectations of linear systematic statistics based on dependent samples. *Ann. Statist.* **7**, 220–223. Correction: 8, 1401.

[5] Caraux, G. and Gascuel, O. (1992). Bounds on distribution functions of order statistics for dependent variates. *Statist. Probab. Lett.* **14**, 103–105.

[6] Gascuel, O. and Caraux, G. (1992). Bounds on expectations of order statistics via extremal dependences. *Statist. Probab. Lett.* **15**, 143–148.

[7] Papadatos, N. (2001a). Expectation bounds on linear estimators from dependent samples. *J. Statist. Plann. Inference* **93**, 17–27.

[8] Papadatos, N. (2001b). Distribution and expectation bounds on order statistics from possibly dependent variates. *Statist. Probab. Lett.* **54**, 21–31.

[9] Papadatos, N. and Rychlik, T. (2004). Bounds on expectations of L-statistics from without replacement samples. *J. Statist. Plann. Inference* **124**, 317–336.

[10] Robertson, T., Wright, F.T. and Dykstra, R.L. (1988). *Order Restricted Statistical Inference*. Wiley, Chichester.

[11] Rychlik, T. (1992a). Sharp inequalities for linear combinations of elements of monotone sequences. *Bull. Polish Acad. Sci. Math.* **40**, 247–254.

[12] Rychlik, T. (1992b). Stochastically extremal distributions of order statistics for dependent samples. *Statist. Probab. Lett.* **13**, 337–341.

[13] Rychlik, T. (1993a). Bounds for expectations of L-estimates for dependent samples. *Statistics* **24**, 9–15.

[14] Rychlik, T. (1993b). Sharp bounds on L-estimates and their expectations for dependent samples. *Commun. Statist.–Theory Meth.* **22**, 1053–1068. Correction: **23**, 305–306.

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

[16] Rychlik, T. (1998). Bounds on expectations of L-estimates. In: *Order Statistics: Theory and Methods* (N. Balakrishnan and C.R. Rao, eds.), *Handbook of Statistics*, vol. 16, North-Holland, Amsterdam, 105–145.

[17] Rychlik, T. (2001). *Projecting statistical functionals. Lecture Notes in Statistics*, **160**, Springer-Verlag, New York.

[18] Rychlik, T. (2005). Optimal deterministic bounds on L-statistics. Submitted for publication.