JOURNAL OF MULTIVARIATE ANALYSIS **57**, 156–174 (1996)
ARTICLE NO. 0027

Projection Method for Moment Bounds on Order Statistics from Restricted Families

I. Dependent Case*

LESŁAW GAJEK

Technical University of Łódź, Poland and Polish Academy of Sciences, Toruń, Poland

AND

TOMASZ RYCHLIK

Polish Academy of Sciences, Toruń, Poland

We present a method of projections onto convex cones for establishing the sharp bounds in terms of the first two moments for the expectations of L-estimates based on samples from restricted families. In this part, we consider the case of possibly dependent identically distributed parent random variables. For the classes of decreasing failure probability, DFR, and symmetric unimodal marginal distributions, we first determine parametric subclasses which contain the distributions attaining the extreme expectations for all L-estimates. Then we derive the bounds for single order statistics. The results provide some new characterizations of uniform and exponential distributions.

1. INTRODUCTION

Let $X_1, ..., X_n$ be possibly dependent identically distributed random variables. Let F, μ, μ_2, and $\sigma^2 = \mu_2 - \mu^2 < \infty$ stand for a common distribution function, the first and second moments, and variance, respectively. Define the quantile function by $Q_F(x) = \sup\{t\colon F(t) \leqslant x\}$. Let $X_{i:n}$ denote the ith order statistic. Rychlik [14] proved that for the expectation of an arbitrary L-estimate with coefficients $c_i \in \mathscr{R}$, $i = 1, ..., n$, the inequality

$$\mathrm{E}_F \sum_{i=1}^{n} c_i X_{i:n} \leqslant \int_0^1 Q_F(x) \sum_{i=1}^{n} d_i \delta_i(x)\, dx \tag{1}$$

Received December 9, 1994; revised December 1995.

AMS 1991 subject classifications: primary 62G30, secondary 62E10.

Key words and phrases: dependent identically distributed random variables, decreasing failure probability distribution, decreasing failure rate distribution, symmetric unimodal distribution, uniform distribution, exponential distribution, L-estimate, order statistic, convex cone, projection.

* Research supported by KBN Grant 2P 30101004.

0047-259X/96 $18.00

holds, where

$$d_i = n\left[C\left(\frac{i}{n}\right) - C\left(\frac{i-1}{n}\right)\right], \qquad i = 1, ..., n,$$

with C being the greatest convex function on $[0, 1]$ that satisfies

$$C\left(\frac{j}{n}\right) \leqslant \sum_{i=1}^{j} c_i, \qquad j = 0, 1, ..., n,$$

and δ_i denoting the indicator of $[(i-1)/n, i/n)$. Note that $\sum_{i=1}^n d_i\delta_i$ constitutes an n-dimensional family of jump functions and d_i, $i = 1, ..., n$, is the l_2-projection of c_i, $i = 1, ..., n$, onto nondecreasing sequences. E.g., for a single jth order statistic, $\sum_{i=1}^n d_i\delta_i = n/(n+1-j)$ on $[(j-1)/n, 1]$ and 0 elsewhere, and, as a consequence,

$$\mathrm{E}_F X_{j:n} \leqslant \frac{n}{n+1-j} \int_{(j-1)/n}^{1} Q_F(x)\, dx \tag{2}$$

(cf. also Caraux and Gascuel [5], Rychlik [13]). Inequality (1) is sharp; i.e., it becomes equality for some joint distribution with given common marginals F (for details, see [14]). For instance, (2) becomes equality iff

$$\Pr\left(X_{j-1:n} \leqslant Q_F\left(\frac{j-1}{n}\right) \leqslant X_{j:n} = X_{n:n}\right) = 1. \tag{3}$$

Applying (1) and the Schwarz inequality, Rychlik [16] obtained a general sharp bound for expected L-estimates, depending merely on the expectation and variance of parent variables:

$$\mathrm{E}_F \sum_{i=1}^{n} c_i(X_{i:n} - \mu) \leqslant \left[\sum_{i=1}^{n} d_i^2 - \left(\sum_{i=1}^{n} c_i\right)^2\right]^{1/2} \frac{\sigma}{\sqrt{n}}. \tag{4}$$

The equality in (4) holds if each X_i takes on n values (not necessarily different),

$$x_j = \mu + \frac{d_j - \sum_{i=1}^n c_i}{[\sum_{i=1}^n d_i^2 - (\sum_{i=1}^n c_i)^2]^{1/2}}\, \sigma, \qquad j = 1, ..., n,$$

with equal probabilities, for some specific joint distributions, e.g., for the random choice of a permutation of numbers $x_1, ..., x_n$. The cases of the sample maximum and range were solved earlier by Arnold [1]. If X_i are symmetrically distributed, an improvement of (4) was given by Rychlik [16] (for the case of sample maximum; see Arnold [1]).

Our purpose is to provide tighter bounds for expected L-estimates, when the marginals belong to restricted families of absolutely continuous distributions. In the case of independent variables, there are known bounds for some families of parent distribution functions: decreasing failure probability, monotone failure ratio, and symmetric unimodal ones (see Arnold and Balakrishnan [2, Section 3.4] and David [7, Section 4.4]). These bounds were established by means of the Jensen inequality. They depend on some quantiles of the parent distribution.

Below we present sharp bounds for the above-mentioned families of distributions in terms of the first two moments. The bounds are obtained by projecting quantile functions, or their modifications, on a convex cone generated by a considered class of distributions. To this end we shall apply the following basic lemma, describing projections onto convex cones in inner product spaces.

LEMMA 1. *Let $(\mathcal{X}, (\cdot, \cdot))$ be a real inner product space and let $\mathcal{C} \subset \mathcal{X}$ be a convex cone. If for a given $x \in \mathcal{X}$ there exists $P_{\mathcal{C}} x \in \mathcal{C}$ which satisfies*

$$\|x - P_{\mathcal{C}} x\| \leqslant \|x - y\| \qquad \text{for all} \quad y \in \mathcal{C},$$

then $P_{\mathcal{C}} x$ is unique and satisfies

$$(x, y) \leqslant (P_{\mathcal{C}} x, y) \qquad \text{for all} \quad y \in \mathcal{C}, \tag{5}$$

$$(x, P_{\mathcal{C}} x) = \|P_{\mathcal{C}} x\|^2. \tag{6}$$

If $\mathcal{X}$ is complete and $\mathcal{C}$ is closed, then $P_{\mathcal{C}} x$ does exist for every $x \in \mathcal{X}$, and it is called the metric projection of x onto $\mathcal{C}$. We omit the proof of Lemma 1 here, because it almost coincides with the proof of the statement for Hilbert spaces (see, e.g., Balakrishnan [3, Section 1.4]).

To illustrate the projection method, suppose that $\mathcal{X}$ is the linear space of real functions on $[0, 1)$ which are right continuous, have left limits, and are square integrable, and the inner product is defined by

$$(f, g) = \int_0^1 f(x)\, g(x)\, dx, \qquad f, g \in \mathcal{X}.$$

Observe that the family $\mathcal{Q}_0$ of all quantile functions of random variables with a finite second moment is a convex cone in $\mathcal{X}$. Let $\mathcal{Q} \subseteq \mathcal{Q}_0$ be also a convex cone in $\mathcal{X}$. By (1), (5), and the Schwarz inequality

$$\begin{aligned} \mathrm{E}_F \sum_{i=1}^{n} c_i X_{i:n} &\leqslant \int_0^1 Q_F(x) \sum_{i=1}^{n} d_i \delta_i(x)\, dx \\ &\leqslant \int_0^1 Q_F(x)\, P_Q\left(\sum_{i=1}^{n} d_i \delta_i\right)(x)\, dx \\ &\leqslant \|Q_F\| \left\| P_Q\left(\sum_{i=1}^{n} d_i \delta_i\right)\right\| \\ &= \left\| P_Q\left(\sum_{i=1}^{n} d_i \delta_i\right)\right\| \sqrt{\mu_2}. \end{aligned} \tag{7}$$

For the joint distributions, described in [14], with common quantile function

$$Q_F = \frac{P_Q(\sum_{i=1}^{n} d_i \delta_i)}{\|P_Q(\sum_{i=1}^{n} d_i \delta_i)\|} \sqrt{\mu_2} \tag{8}$$

we apply (6) to get

$$\begin{aligned} \mathrm{E}_F \sum_{i=1}^{n} c_i X_{i:n} &= \int_0^1 Q_F(x) \sum_{i=1}^{n} d_i \delta_i(x)\, dx \\ &= \frac{\sqrt{\mu_2}}{\|P_Q(\sum_{i=1}^{n} d_i \delta_i)\|} \int_0^1 P_Q\left(\sum_{i=1}^{n} d_i \delta_i\right)(x) \sum_{i=1}^{n} d_i \delta_i(x)\, dx \\ &= \left\| P_Q\left(\sum_{i=1}^{n} d_i \delta_i\right)\right\| \sqrt{\mu_2}. \end{aligned}$$

If the shifted quantile functions $Q - \mu$, $Q \in \mathcal{Q}$, constitute a convex cone such that the projection $P_{\mathcal{Q}-\mu}(\sum_{i=1}^{n} d_i \delta_i)$ exists, we can similarly obtain another sharp inequality in terms of the expectation μ and the variance σ.

In Sections 2, 3, and 4 we consider specific families of decreasing failure probability, DFR, and symmetric unimodal parent distributions, respectively. In each section we first determine a parametric class of marginals containing the unique element with the extreme expectation of a given L-statistic. Then we present upper bounds for the expectation of a single-order statistic and the distributions for which the bounds are attained. The results provide some new characterizations of the uniform and exponential distributions (see Section 5 for details).

In the Appendix the problem of the best approximation of a jump function by a nondecreasing convex one is considered. Two lemmas proved therein provide the basic tools for deriving the results of Sections 2–4. The projection method allows us to establish analogous sharp bounds for the independent identically distributed random variables as well. In this case we need to project polynomials instead of jump functions. The respective calculations and resulting formulas are significantly more complicated. The independent case will be studied in Part II of the paper.

2. Life Distributions with Decreasing Failure Probability

We consider the class $\mathscr{F}_1$ of distribution functions such that $F \in \mathscr{F}_1$ iff $F = 0$ on $\mathscr{R}_-$, and $F(0) \geqslant 0$, and F is absolutely continuous on $\mathscr{R}_+$, and the density function is nonincreasing. This means that the failure (or death) probability decreases as the object is aging. The first condition means that the object is possibly damaged (or dead) from the beginning. These distributions can also describe one-sided (or absolute) measurement errors. Some reliability results for the class were presented in Dharmadhikari and Joag-Dev [8, Chap. 9]. The respective class of quantile functions is

$$\mathscr{C}_1 = \{Q: [0, 1] \mapsto \mathscr{R}: Q(0) = 0 \text{ and } Q \text{ is nondecreasing and convex}\}.$$

Observe that $\mathscr{C}_1$ is a convex cone of quantile functions. We apply the projection method and the results of the Appendix (with $[a, b) = [0, 1)$ and $w \equiv 1$) to establish sharp bounds for expected L-estimates which depend on the second moment.

Proposition 1. *Let $\mathscr{F}_1^* \subset \mathscr{F}_1$ be the family of finite mixtures of $[0, a_i]$-uniform random variables, $i = 1, ..., k$, $k \leqslant n$, including a possible pole at zero. Then, for every $(c_1, ..., c_n) \in \mathscr{R}^n$ there exists $F^* \in \mathscr{F}_1^*$ such that*

$$\frac{E_{F^*} \sum_{i=1}^n c_i X_{i:n}}{\sqrt{\mu_2}} = \sup_{F \in \mathscr{F}_1} \frac{E_F \sum_{i=1}^n c_i X_{i:n}}{\sqrt{\mu_2}}.$$

Proof. From (7) and (8), it follows that the right-hand side can be attained when a quantile function is proportional to the projection $P_1(\sum_{i=1}^n d_i \delta_i)$ on $\mathscr{C}_1$. From Lemma A1 of the Appendix we conclude that the projection exists and is piecewise linear. This corresponds to a mixture of uniform distributions. ∎

Proposition 2. *Suppose that X_i, $i = 1, ..., n$, are dependent with a common distribution function $F \in \mathscr{F}_1$:*

If $(j-1)/n \leqslant \frac{1}{3}$, then

$$E_F X_{j:n} \leqslant \frac{\sqrt{3}}{2}\left(1 + \frac{j-1}{n}\right)\sqrt{\mu_2}$$

and the equality holds iff X_i, $i = 1, ..., n$, are uniformly distributed on $[0, \sqrt{3\mu_2}]$.

If $(j-1)/n > \frac{1}{3}$, then

$$E_F X_{j:n} \leqslant \tfrac{2}{3}\sqrt{2n\mu_2/(n+1-j)}$$

and the equality holds iff X_i are $[0, \sqrt{(2n\mu_2/(n+1-j)}]$-uniformly distributed with probability $\frac{3}{2}(1-(j-1)/n)$ and have one atom at zero with probability $\frac{3}{2}(j-1)/n-\frac{1}{2}$.

Proof. By Lemma A2, it suffices to find $\alpha \in [0, (j-1)/n]$ maximizing (24), which in our case can be rewritten as

$$A(\alpha)=\frac{3}{4}\left(1-\frac{j-1}{n}\right)^2 \frac{(1+(j-1)/n-2\alpha)^2}{(1-\alpha)^3}.$$

An easy computation shows that this is maximized by 0 if $(j-1)/n \leqslant \frac{1}{3}$ and $\frac{3}{2}(j-1)/n-\frac{1}{2}$ elsewhere. Plugging it into (25), we can write the projection as

$$P_1\left(\frac{n}{n+1-j}\sum_{i=j}^{n}\delta_i\right)(x)$$
$$=\begin{cases}\dfrac{3}{2}\left(1+\dfrac{j-1}{n}\right)x, & \text{if } \dfrac{j-1}{n}\leqslant\dfrac{1}{3},\\[2ex] \dfrac{4n^2}{9(n+1-j)^2}\left(2x+1-3\dfrac{j-1}{n}\right)_+, & \text{if } \dfrac{j-1}{n}>\dfrac{1}{3}.\end{cases}$$

By (7) and (8), we obtain the desired conclusion. ▮

3. Life Distributions with Decreasing Failure Rate

We now examine the case of the subclass $\mathscr{F}_2$ of life distributions (i.e., $F=0$ on $\mathscr{R}_-$) such that the respective failure rate functions

$$H_F(x)=-[\ln(1-F(x)]', \qquad F\in\mathscr{F}_2,$$

exist and decrease on $\mathscr{R}_+$. The failure rate represents the infinitesimal conditional probability that the object which survived time x will fail immediately afterwards. Properties and applications of the notion were extensively studied (see, e.g., Barlow and Proschan [4]). Some basic properties and examples of distributions with monotone failure rates were briefly presented in [11]. The decreasing failure rate means that the object is improving with age which means the conditional failure probability in a given period of time becomes smaller as the time passes (for theoretical justifications of DFR, we refer to Proschan [12]).

The DFR density function is decreasing on $\mathscr{R}_+$. We also assume here that $F(0)\geqslant 0$, $F\in\mathscr{F}_2$. Observe that the quantile function of DFR distribution satisfies

$$Q(0)=0, \qquad Q(1-e^{-x}) \text{ is convex on } \mathscr{R}_+.$$

Furthermore, by substitution, (1) can be rewritten as

$$\mathrm{E}_F \sum_{i=1}^{n} c_i X_{i:n} \leqslant \int_0^{\infty} Q(1-e^{-x}) \sum_{i=1}^{n} d_i \Delta_i(x)\, e^{-x}\, dx,$$

where Δ_i is the indicator of $[\ln(n/(n+1-i)), \ln(n/(n-i)))$. Accordingly, to obtain the bounds analogous to (7) for the case of DFR distributions we need to project $\sum_{i=1}^{n} d_i \Delta_i$ onto the convex cone of compositions of DFR quantiles with the standard exponential distribution

$$\mathscr{C}_2 = \{R\colon [0, \infty) \mapsto \mathscr{R}\colon R(0) = 0,\ R \text{ is nondecreasing and convex}\}$$

in the respective inner product function space with the weight function $w(x) = e^{-x}$, $x \geqslant 0$. We will denote by P_2 the projections onto $\mathscr{C}_2$. Lemma A1 immediately implies that all $P_2(\sum_{i=1}^{n} d_i \Delta_i)$ are piecewise linear. Below we describe the respective parametric class of marginal distributions. For $k \leqslant n$ take two arbitrary sequences $0 = x_0 < x_1 < \cdots < x_k = +\infty$ and $0 \leqslant \lambda_1 < \cdots < \lambda_k < +\infty$. If

$$Q_F(1-e^{-x}) = \lambda_j(x - x_{j-1}) + \sum_{i=1}^{j-1} \lambda_i(x_i - x_{i-1}) \qquad \text{for} \quad x \in [x_{j-1}, x_j],$$

then

$$F(x) = 1 - \exp\left(-\frac{1}{\lambda_j}\left[x - \sum_{i=1}^{j-1} \lambda_i(x_i - x_{i-1})\right] - x_j\right)$$
$$\text{for} \quad x \in \left[\sum_{i=1}^{j-1} \lambda_i(x_i - x_{i-1}), \sum_{i=1}^{j} \lambda_i(x_i - x_{i-1})\right]. \tag{9}$$

This means that the marginal distribution is a specific mixture of truncated exponential distributions with a possible pole such that $F(0) = 1 - e^{-x_1}$, if $\lambda_1 = 0$. Define

$$\mathscr{F}_2^* = \{F \in \mathscr{F}_2\colon F \text{ satisfies (9) for some } x_1, ..., x_{k-1}, \lambda_1, ..., \lambda_k, k \leqslant n\}.$$

We are thus led to the following.

PROPOSITION 3. *For every L-estimate, based on dependent variables with a common DFR life distribution there is an* $F^* \in \mathscr{F}_2^*$ *such that*

$$\frac{E_{F^*} \sum_{i=1}^{n} c_i X_{i:n}}{\sqrt{\mu_2}} = \sup_{F \in \mathscr{F}_2} \frac{E_F \sum_{i=1}^{n} c_i X_{i:n}}{\sqrt{\mu_2}}.$$

Since $\mathscr{C}_2 \subset \mathscr{C}_1$, for every L-estimate the above supremum is evidently less than the respective one in Proposition 1. We can compare explicit formulas for single order statistics.

PROPOSITION 4. *Suppose that* $F \in \mathscr{F}_2$. *If* $(j-1)/n \leqslant 1-e^{-1}$ *then*

$$E_F X_{j:n} \leqslant \left[\ln \frac{n}{n+1-j} + 1 \right] \sqrt{\mu_2/2},$$

which is attainable for the exponential marginal distribution with the scale parameter $\sqrt{\mu_2/2}$.

Otherwise,

$$E_F X_{j:n} \leqslant \sqrt{2n\mu_2/[(n+1-j)\, e]}.$$

This becomes equality for the mixture of the exponential distribution with the scale $\sqrt{(n/(n+1-j))(\mu_2/2e)}$ *and zero, with probabilities* $(n+1-j)\, e/n$ *and* $1-(n+1-j)\, e/n$, *respectively.*

Proof. Since

$$\sup_{F \in \mathscr{F}_2} \frac{\mathrm{E}_F X_{j:n}}{\sqrt{\mu_2}} = \inf_{R \in \mathscr{C}_2,\ \|R\| = \sqrt{\mu_2}} \int_0^\infty R(x) \frac{n}{n+1-j} \sum_{i=j}^n \Delta_i(x)\, e^{-x}\, dx,$$

our problem resolves to finding $P_2(n/(n+1-j) \sum_{i=j}^n \Delta_i)$ and calculating its norm. By Lemma A2, the projection is the positive part of a linear function. To determine the root of it, we should maximize (24), which in this case amounts to

$$A(\alpha) = \frac{(n+1-j)^2}{2n^2} \left(\ln \frac{n}{n+1-j} + 1 - \alpha \right)^2 e^{\alpha},$$

for $\alpha \in [0, \ln(n/(n+1-j)))$. An elementary algebra yields $\alpha^* = (\ln(n/(n+1-j)) - 1)_+$.

Suppose first that $\alpha^* = 0$, i.e., $(j-1)/n < 1-e^{-1}$. Then the optimal slope is

$$\lambda^*(0) = \frac{1}{2} \left(\ln \frac{n}{n+1-j} + 1 \right),$$

due to (25) and, further,

$$P_2 \left(\frac{n}{n+1-j} \sum_{i=j}^n \Delta_i \right)(x) = \left(\ln \frac{n}{n+1-j} + 1 \right) \frac{x}{2},$$

$$\left\| P_2 \left(\frac{n}{n+1-j} \sum_{i=j}^n \Delta_i \right) \right\| = \frac{1}{\sqrt{2}} \left(\ln \frac{n}{n+1-j} + 1 \right),$$

and, finally, the quantile function satisfying

$$Q_{F^*}(1-e^{-x})=\sqrt{\mu_2/2}\,x, \qquad x>0,$$

gives the sharp bound for the expectation, as we claimed.

Turning to the other case, we verify that

$$P_2\left(\frac{n}{n+1-j}\sum_{i=j}^{n}\Delta_i\right)(x)=\frac{n}{(n+1-j)\,e}\left(x+1-\ln\frac{n}{n+1-j}\right)_+$$

has the norm

$$\left\|P_2\left(\frac{n}{n+1-j}\sum_{i=j}^{n}\Delta_i\right)\right\|=\sqrt{2n/[(n+1-j)e]},$$

which yields

$$Q_{F^*}(1-e^{-x})=\sqrt{n\mu_2/[2e(n+1-j)]}\left(x+1-\ln\frac{n}{n+1-j}\right)_+, \qquad x>0.$$

This completes the proof. ∎

4. Symmetric Unimodal Distributions

We denote by $\mathscr{F}_3$ the class of symmetric unimodal distributions which, by definition, satisfy $F(\mu+x)=1-F(\mu-x)$, $x\in\mathscr{R}$, and F is concave on $[\mu, +\infty)$. The class contains a number of popular parametric families, with the normal one as the most eminent representative. Unimodality properties and applications in statistics were described in [8]. The respective quantiles are convex on $[\frac{1}{2}, 1]$, and satisfy

$$Q_F(x)-\mu=\mu-Q_F(1-x). \tag{10}$$

It is convenient to concentrate on the modifications of quantiles

$$S_F=(Q_F-\mu)_{|[1/2,\,1]}, \qquad F\in\mathscr{F}_3, \tag{11}$$

which constitute a convex cone

$$\mathscr{C}_3=\{S\colon [\tfrac{1}{2}, 1]\mapsto\mathscr{R}\colon S(\tfrac{1}{2})=0 \text{ and } S \text{ is nondecreasing and convex}\}.$$

Then, by (1), (7), (8), and (11), we have

$$\sup_{F \in \mathscr{F}_3} \frac{\mathrm{E}_F \sum_{i=1}^n c_i (X_{i:n} - \mu)}{\sigma}$$

$$= \sup_{S \in \mathscr{C}_3} \frac{\int_{1/2}^1 S(x) \sum_{i=1}^{[n/2]} (d_{n+1-i} - d_i)\, \delta_{n+1-i}(x)\, dx}{\sigma}$$

$$= \frac{1}{\sqrt{2}} \left\| P_3 \left(\sum_{i=1}^{[n/2]} (d_{n+1-i} - d_i)\, \delta_{n+1-i} \right) \right\|,$$

where P_3 denotes the projection onto $\mathscr{C}_3$. The supremum on the left-hand side is attained iff

$$S_F = \frac{P_3(\sum_{i=1}^{[n/2]} (d_{n+1-i} - d_i)\, \delta_{n+1-i})}{\| P_3(\sum_{i=1}^{[n/2]} (d_{n+1-i} - d_i)\, \delta_{n+1-i}) \|} \frac{\sigma}{\sqrt{2}}. \qquad (12)$$

Note that the projected function is a jump function, which takes on $[(n+1)/2]$ values at most. Therefore the projection of it is a piecewise linear function with $k \leqslant [(n+1)/2]$ pieces. Precisely,

$$P_3 \left(\sum_{i=1}^{[n/2]} (d_{n+1-i} - d_i)\, \delta_{n+1-i} \right)(x)$$

$$= \alpha_j (x - x_{j-1}) + \sum_{i=1}^{j-1} \alpha_i (x_i - x_{i-1}) \qquad \text{for} \quad x \in [x_{j-1}, x_j], \quad (13)$$

for two nondecreasing sequences x_i, α_i, $i = 1, ..., k$, where $x_0 = \frac{1}{2}$, $x_k = 1$, and $\alpha_1 \geqslant 0$. By (11) and (12), the respective quantile function is an affine transformation of (13) on $[\frac{1}{2}, 1]$, defined by (10) on the remaining part of its domain. Finally, we can write the following.

Proposition 5. *For every L-estimate there exists a distribution function $F^* \in \mathscr{F}_3$ which is a mixture $k \leqslant [(n+1)/2]$ uniform, symmetric about μ distributions, possibly including the degenerate μ-valued distribution such that*

$$\frac{E_{F^*} \sum_{i=1}^n c_i (X_{i:n} - \mu)}{\sigma} = \sup_{F \in \mathscr{F}_3} \frac{E_F \sum_{i=1}^n c_i (X_{i:n} - \mu)}{\sigma}.$$

The remainder of this section will be devoted to the case of single order statistics.

Proposition 6. *For the sample minimum we have*

$$E_F X_{1:n} \leqslant \mu, \qquad F \in \mathscr{F}_3.$$

The equality holds iff $X_1 = \cdots = X_n$ *have any distribution* $F \in \mathscr{F}_3$*:*

If $0 < (j-1)/n < \frac{1}{3}$*, then*

$$E_F X_{j:n} \leqslant \mu + \frac{2n}{3(n+1-j)} \sqrt{(j-1)/n}\, \sigma, \qquad F \in \mathscr{F}_3.$$

This becomes an equality iff each $X_i = \mu$ *with probability* $1 - 3(j-1)/n$ *and is uniformly distributed on* $[\mu - \sqrt{n/(j-1)}\,\sigma, \mu + \sqrt{n/(j-1)}\,\sigma]$ *with probability* $3(j-1)/n$.

If $\frac{1}{3} \leqslant (j-1)/n \leqslant \frac{2}{3}$*, then*

$$E_F X_{j:n} \leqslant \mu + \sqrt{3}\, \frac{j-1}{n}\, \sigma, \qquad F \in \mathscr{F}_3,$$

with the equality attainable only for $[\mu - \sqrt{3}\,\sigma, \mu + \sqrt{3}\,\sigma]$*-uniform marginal distribution.*

If $\frac{2}{3} < (j-1)/n$*, then*

$$E_F X_{j:n} \leqslant \mu + \tfrac{2}{3} \sqrt{n/(n+1-j)}\, \sigma, \qquad F \in \mathscr{F}_3.$$

Here the equality holds iff F *is the mixture of the jump distribution at* μ *and the* $[\mu - \sqrt{n/(n+1-j)}\,\sigma, \mu + \sqrt{n/(n+1-j)}\,\sigma]$*-uniform distribution with probabilities* $3(j-1)/n - 2$ *and* $3(1-(j-1)/n)$*, respectively.*

Proof. The first bound is trivial. We have $\mathrm{E}_F X_{1:n} \leqslant \mathrm{E}_F X_1$, whatever the interdependence and the common marginal F are. Obviously, the equality holds iff the random variables are identical.

Let us observe that for the jth-order statistic

$$\sum_{i=1}^{[n/2]} (d_{n+1-i} - d_i)\, \delta_{n+1-i} = \frac{n}{n+1-j} \sum_{i=r}^{n} \delta_i,$$

where $r = \max\{j-1, n+1-j\}$. We can apply Lemma A2 to assert that

$$P_3\left(\frac{n}{n+1-j} \sum_{i=r}^{n} \delta_i\right)(x) = \lambda^*(x - \alpha^*)_+.$$

Maximizing (24), we obtain $\alpha^* = \max\{\frac{1}{2}, (3r-n)/2n\}$.

If $\frac{1}{3} \leqslant (j-1)/n \leqslant \frac{2}{3}$, then $\alpha^* = \frac{1}{2}$, and so

$$P_3\left(\frac{n}{n+1-j} \sum_{i=r}^{n} \delta_i\right)(x) = 12\, \frac{(n-r)\, r}{(n+1-j)\, n} \left(x - \frac{1}{2}\right) = 12\, \frac{j-1}{n} \left(x - \frac{1}{2}\right)$$

has the norm

$$\left\| P_3\left(\frac{n}{n+1-j} \sum_{i=r}^{n} \delta_i\right) \right\| = \sqrt{6}\, \frac{j-1}{n}.$$

Therefore, by (11) and symmetry,

$$Q_{F^*}(x) = \mu + 2\sqrt{3}\,(x - \tfrac{1}{2})\,\sigma,$$

which is the quantile function of $[\mu - \sqrt{3}\,\sigma, \mu + \sqrt{3}\,\sigma]$-uniform distribution.

Otherwise $\alpha^* = (3r - n)/2n$,

$$P_3\left(\frac{n}{n+1-j}\sum_{i=r}^{n}\delta_i\right)(x) = \frac{8n^2}{9(n-r)(n+1-j)}\left(x - \frac{3r-n}{2n}\right)_+,$$

$$\left\| P_3\left(\frac{n}{n+1-j}\sum_{i=r}^{n}\delta_i\right)\right\| = \frac{2\sqrt{2}}{3}\,\frac{\sqrt{(n-r)\,n}}{n+1-j},$$

and for $x \geqslant \frac{1}{2}$,

$$Q_{F^*}(x) = \mu + \frac{2}{3}\left(\frac{n}{n-r}\right)^{3/2}\left(x - \frac{3r-n}{2n}\right)_+ \sigma.$$

Therefore F^* is the combination of the jump distribution at μ and the uniform distribution on $[\mu - \sqrt{n/(n-r)}\,\sigma, \mu + \sqrt{n/(n-r)}\,\sigma]$ with coefficients $3r/n - 2$ and $3(1 - r/n)$, respectively. Taking $r = n+1-j$ and $j-1$, we obtain the second and fourth cases of Proposition 7, respectively. ∎

It is worth noting that $\sup_{F \in \mathscr{F}_3} \mathrm{E}_F X_{j:n} > \mu$, for all $j \geqslant 2$. This is impossible for independent parent samples.

5. Concluding Remarks

1. According to (7), the projection method allows us to determine bounds for various restricted families of distributions for which quantile functions constitute convex cones. Applying auxiliary results of Appendix, we can establish, with a moderate effort, bounds for the counterparts of families considered in Sections 2–4: life distributions with increasing failure probability and rate, and symmetric U-shaped. In all these cases, for the single order statistics, the marginals attaining the bounds are certain transformations of projection of the form $\min\{\lambda^* x, M\}$.

2. It is easy to verify that for arbitrary possibly asymmetric trimmed means $T_{j,k:n} = (k+1-j)^1 \sum_{i=j}^{k} X_{i:n}$, as well as Winsorized means,

$$W_{j,k:n} = \frac{1}{n}\left[\, jX_{j:n} + \sum_{i=j+1}^{k-1} X_{i:n} + (n+1-k)\,X_{k:n}\right], \qquad 1 \leqslant j < k \leqslant n,$$

we have

$$\sum_{i=1}^{n} d_i \delta_i = \frac{n}{n+1-j} \sum_{i=j}^{n} \delta_i,$$

just like for the jth order statistic (see Rychlik [15]). As a consequence, we can replace $X_{j:n}$ by $T_{j,k:n}$ or $W_{j,k:n}$ in Propositions 2, 4, and 6, and the assertions will be still true. Likewise, for the jth range $R_{j:n} = X_{j:n} - X_{n+1-j:n}$, $j > n/2$, yields

$$\sum_{i=1}^{[n/2]} (d_{n+1-j} - d_i)\, \delta_i = \frac{2n}{n+1-j} \sum_{i=j}^{n} \delta_i$$

(cf. Rychlik [14]). Therefore the bounds for $\mathrm{E}_F R_{j:n}$, $F \in \mathscr{F}_3$, are twice as much as the bounds for the corresponding jth-order statistic in Proposition 6, with the identical conditions of being attained.

3. Propositions 2, 4, and 6 provide new characterizations of uniform and exponential distributions:

For possibly dependent identically distributed random variables with decreasing failure probability

$$\mathrm{E}_F X_{j:n} = \frac{\sqrt{3}}{2}\left(1 + \frac{j-1}{n}\right)\sqrt{\mu_2}, \qquad \frac{j-1}{n} \leqslant \frac{1}{3},$$

may hold iff the marginal is uniform.

The only common marginal DFR distribution F such that

$$\mathrm{E}_F X_{j:n} = \left[\ln \frac{n}{n+1-j} + 1\right] \sqrt{\mu_2/2}, \qquad \frac{j-1}{n} \leqslant 1 - e^{-1},$$

for a dependent sample is the exponential distribution.

$$\mathrm{E}_F X_{j:n} = \mu + \sqrt{3}\, \frac{j-1}{n}\, \sigma, \qquad \frac{1}{3} \leqslant \frac{j-1}{n} \leqslant \frac{2}{3},$$

holds for a sample of dependent identically distributed symmetric and unimodal random variables iff the distribution F is uniform.

Appendix

It will be convenient to consider some modifications of quantile functions. Therefore we introduce an inner product function space which is slightly more general than the one considered in the Introduction. Let $\mathscr{X}$ be the family of functions $f: [a, b) \mapsto \mathscr{R}$, $-\infty < a < b \leqslant +\infty$, which are right continuous and have left limits, and

$$\int_a^b f^2(x)\, w(x)\, dx < \infty$$

for a strictly positive weight function w such that all linear functions belong to $\mathscr{X}$. The inner product will be defined by

$$(f, g) = \int_a^b f(x)\, g(x)\, w(x)\, dx, \qquad f, g \in \mathscr{X}.$$

Let $\mathscr{C} \subset \mathscr{X}$ be the subfamily of nondecreasing, convex functions which vanish at a. Obviously, $\mathscr{C}$ is a convex cone. Now fix $f \in \mathscr{X} \backslash \mathscr{C}$. For applications to the bounds on L-statistics in Sections 2–4 we assume that f is a jump function. In the case of single order statistics the respective jump functions take on only two values. Our objective is to find $P_{\mathscr{C}} f \in \mathscr{C}$ (if it exists) which minimizes the weighted mean square distance

$$D_f(g) = \int_a^b [f(x) - g(x)]^2\, w(x)\, dx$$

over all $g \in \mathscr{C}$.

We now consider the case of f being stepwise with m steps.

LEMMA A1. *If* $f(x) = M_i$, *when* $x \in [x_{i-1}, x_i)$ *for some* $a = x_0 < x_1 < \cdots < x_m = b$ *and* $M_{i-1} \neq M_i$, $i = 2, ..., m$, *then* $P_{\mathscr{C}} f$ *exists and is piecewise linear function with at most* m *pieces.*

Proof. Let $\mathscr{C}_m \subset \mathscr{C}$ be the family of piecewise linear functions which have no more than m pieces. We first show that for every $g \in \mathscr{C}$ there exists $h \in \mathscr{C}_{m+1}$ such that

$$D_f(h) \leqslant D_f(g). \tag{14}$$

So let g be an arbitrary element of $\mathscr{C}$. We define below a partition of the domain into intervals with ends $a = y_0 \leqslant y_1 < \cdots < y_s = b$. The first, possibly degenerate, interval $[y_0, y_1]$ is the one where g vanishes. Observe that g is strictly increasing on (y_1, b), and, in consequence, g can intersect f at separate points of (y_1, b) at most once in each of (y_1, x_j), (x_{i-1}, x_i) for all $i > j$ such that $x_{j-1} \leqslant y_1 < x_j$. For a given interval, three cases are possible: either g intersects f in it, or g lies above or beneath f.

We define y_{i+1}, $i \geqslant 1$, recursively. Suppose that $x_{k-1} \leqslant y_i < x_k$. If g intersects or lies above f on $[y_i, x_k)$, then $y_{i+1} = x_k$. Otherwise, we define y_{i+1} as the greatest x_l such that g lies beneath f on $[y_i, x_l)$. Obviously, $\{y_2, ..., y_s\} \subset \{x_j, ..., x_m\}$ and so $s \leqslant m+1$.

We now construct $h \in \mathscr{C}_{m+1}$ satisfying (14). Consider an interval (y_{i-1}, y_i), $i \geqslant 2$, where g and f intersect at some z_i. Let h_i be the linear function determined by two points $(y_{i-1}, g(y_{i-1}))$ and $(z_i, g(z_i))$. Further, suppose that $g > f$ on some (y_{i-1}, y_i). We define h_i choosing any linear function tangent to g at y_{i-1}. Take finally any (y_{i-1}, y_i), where f has values $M_k, ..., M_l$ and $g < f$. Let h_i be the greatest convex function on $[y_{i-1}, y_i]$, not greater than f on (y_{i-1}, y_i), and equal to g at y_{i-1} and y_i and extended linearly to the whole domain. This is a piecewise linear function determined by $l+1-k$ conditions $h_i(y_r) = g(y_r)$, $r = i-1, i$, and $h_i(x_r) \leqslant M_r$, $r = k, ..., l-1$. It has $l-k$ pieces at most.

Observe that for all $i = 1, ..., s-1$

$$0 \leqslant h_i(y_i) \leqslant g(y_i) = h_{i+1}(y_i), \tag{15}$$

$$0 \leqslant h_i'(y_i) \leqslant g'(y_i -) \leqslant h_{i+1}'(y_i). \tag{16}$$

We now examine the behaviour of $\max\{h_i, h_{i+1}\}$, $i = 2, ..., s$. Suppose first that g and f intersect on (y_{i-1}, y_i). Then, by definition of z_i and (15)–(16),

$$h_{i+1}(z_i) \leqslant h_i(z_i) = f(z_i) \leqslant h_i(y_i) \leqslant h_{i+1}(y_i) = g(y_i).$$

Therefore there exists $u_i \in [z_i, y_i]$ such that $\max\{h_i, h_{i+1}\} = h_i$ on $[a, u_i]$ and h_{i+1} elsewhere. Moreover,

$$g(x) \leqslant \max\{h_i, h_{i+1}\}(x) \leqslant f(x) \qquad \text{for} \quad x \in [y_{i-1}, z_i], \tag{17}$$

$$f(x) \leqslant \max\{h_i, h_{i+1}\}(x) \leqslant g(x) \qquad \text{for} \quad x \in [z_i, y_i). \tag{18}$$

If $g > f$ on (y_{i-1}, y_i), then, by (15)–(16),

$$h_{i+1}(y_{i-1}) \leqslant h_i(y_{i-1}) = g(y_{i-1}) \leqslant h_i(y_i) \leqslant h_{i+1}(y_i) = g(y_i),$$

and so $\max\{h_i, h_{i+1}\}$ is equal to h_i and h_{i+1} on the left and right to some $u_i \in [y_{i-1}, y_i]$, respectively. Also,

$$f \leqslant \max\{h_i, h_{i+1}\} \leqslant g \qquad \text{on } [y_{i-1}, y_i]. \tag{19}$$

If $g < f$ on (y_{i-1}, y_i), then h_i and h_{i+1} are equal at y_i and we have

$$g \leqslant \max\{h_i, h_{i+1}\} = h_i \leqslant f \qquad \text{on } (y_{i-1}, y_i). \tag{20}$$

Define $h_1 \equiv 0$ and $h = \max_{1 \leqslant i \leqslant s} h_i$. Then $h \in \mathscr{C}_{m+1}$ and

$$h(x) = \begin{cases} 0, & \text{if} \quad x \in [a, y_1], \\ \max\{h_i, h_{i+1}\}(x), & \text{if} \quad x \in [y_{i-1}, y_i], \quad i = 2, ..., s-1, \\ h_s(x), & \text{if} \quad x \in [y_{s-1}, b). \end{cases}$$

Moreover, $h = g$ on $[a, y_1]$ and, by (17)–(20), it lies between g and f, and so (14) holds.

The next step of the proof is showing that

$$\inf_{h^* \in \mathscr{C}_m} D_f(h^*) = \inf_{g \in \mathscr{C}} D_f(g). \tag{21}$$

To this end, it suffices to modify h so that we obtain $h^* \in \mathscr{C}_m$ less distant from f in the case $s = m+1$. This implies $y_1 < x_1$. Suppose first that $h_2(z_2) = h_3(z_2)$ for a $z_2 \in (y_1, x_1)$. Construction of h shows that $M_1 \leqslant h_2(z_2)$. If $M_1 \leqslant 0$, then $h^* = \max\{h_1, h_3, ..., h_s\}$ satisfies $D_f(h^*) \leqslant D_f(h)$. If $M_1 > 0$, then we can replace $\max\{h_1, h_2\}$ by the linear $\tilde{h}_1$, which vanishes at a and intersects h_3 at the level M_1. Finally suppose that h is linear on $(y_1, x_1]$, which forces $h < M_1$ there. However, since $s = m+1$ holds, h must break at x_1. Applying $\tilde{h}_1(x) = h_3(x_1)(x-a)/(x_1 - a)$, as in the previous case, we improve the approximation of f, and, in consequence, (21) holds.

The proof is completed by proving that the left-hand side infimum is attainable. We notice that for $h^* \in \mathscr{C}_m$ given by

$$h^*(x) = \lambda_j (x - z_{j-1}) + \sum_{i=1}^{j-1} \lambda_i (z_i - z_{i-1}), \qquad x \in [z_{j-1}, z_j],$$

$D_f(h^*)$ is a continuous function of $2m-1$ parameters $0 \leqslant \lambda_1 \leqslant \cdots \leqslant \lambda_m$, $z_i \in [x_{i-1}, x_i]$, $i = 1, ..., m-1$. Our purpose is to prove that the infimum is attained on a compact subset of $\mathscr{R}^{2m-1}$. We only need to find an upper estimate on λ_m.

Let k be the smallest subscript for which $\lambda_k = \lambda_m$. We claim that we can confine to the case when λ_m either equals 0 or is the optimal slope for approximating f on $[z_{k-1}, b)$ by linear functions starting from $(z_{k-1}, h^*(z_{k-1}))$. Indeed,

$$\int_{z_{k-1}}^{b} [f(x) - h^*(z_{k-1}) - \lambda_m (x - z_{k-1})]^2 \, dx$$

is a quadratic function of λ_m, minimized by

$$\lambda^*(z_{k-1}) = \frac{\int_{z_{k-1}}^{b} (x - z_{k-1})(f(x) - h^*(z_{k-1}))\, w(x)\, dx}{\int_{z_{k-1}}^{b} (x - z_{k-1})^2\, w(x)\, dx}. \tag{22}$$

If $0 < \lambda_m \neq \lambda^*(z_{k-1})$, we could decrease $D_f(h^*)$ slightly moving λ_m towards $\lambda^*(z_{k-1})$. For $M = \max_{1 \leqslant i \leqslant m} M_i$ and a positive ε, we have

$$\sup_{a \leqslant z \leqslant x_{m-1}} \lambda^*(z) \leqslant \sup_{a \leqslant z \leqslant x_{m-1}} \frac{M \int_z^b (x-z)\, w(x)\, dx}{\varepsilon^2 \int_{z+\varepsilon}^b w(x)\, dx}$$
$$\leqslant \frac{M \int_a^b (x-a)\, w(x)\, dx}{\varepsilon^2 \int_{x_{m-1}+\varepsilon}^b w(x)\, dx} < \infty,$$

which ends the proof. ∎

It is clear that $P_{\mathscr{C}} f$ has exactly m pieces if $M_1 \geqslant 0$ and $M_i - M_{i-1}$, $i = 2, ..., m$, increase sufficiently fast.

Lemma A2. *Assume that $f(x) = 0$ for $x < \beta$ and $f(x) = M > 0$ for $x \geqslant \beta$, where $\beta \in [a, b]$. Then*

$$P_{\mathscr{C}} f(x) = \lambda^*(x - \alpha^*)_+, \tag{23}$$

where α^* maximizes

$$A(\alpha) = \frac{[\int_\beta^b (x-\alpha)\, w(x)\, dx]^2}{\int_\alpha^b (x-\alpha)^2\, w(x)\, dx} \qquad \text{for} \quad \alpha \in [a, \beta) \tag{24}$$

and

$$\lambda^* = M \frac{\int_\beta^b (x-\alpha^*)\, w(x)\, dx}{\int_{\alpha^*}^b (x-\alpha^*)^2\, w(x)\, dx} > 0. \tag{25}$$

Proof. From the proof of Lemma A1 we deduce that $P_{\mathscr{C}} f$ has no more than two pieces. If it has exactly two pieces, then they meet at an $\alpha \in [a, \beta)$ and therefore the first one must equal 0. The second one must have optimal slope (22), which can be rewritten here as

$$\lambda^*(\alpha) = M \frac{\int_\beta^b (x-\alpha)\, w(x)\, dx}{\int_\alpha^b (x-\alpha)^2\, w(x)\, dx}.$$

If $P_{\mathscr{C}} f$ is linear, then either $P_{\mathscr{C}} f \equiv 0$ or $P_{\mathscr{C}} f(x) = \lambda^*(a)(x-a)$. Since $\lambda^*(a) > 0$, the former case is excluded. Accordingly, $P_{\mathscr{C}} f$ belongs to the class of functions $h_\alpha(x) = \lambda^*(\alpha)(x-\alpha)$, $a \leqslant \alpha \leqslant \beta$. Since

$$D_f(h_\alpha) = -M^2 \frac{[\int_\beta^b (x-\alpha)\, w(x)\, dx]^2}{\int_\alpha^b (x-\alpha)^2\, w(x)\, dx} + M^2 \int_\beta^b w(x)\, dx,$$

it remains to determine $\alpha^* \in [a, \beta]$ that maximizes (24). It is worth noting that $\alpha^* \neq \beta$, because differentiating $D_f(h_\alpha)$ in α gives

$$D'_f(h_\alpha) = -2M^2 \frac{\int_\beta^b (x-\alpha)\, w(x)\, dx}{[\int_\alpha^b (x-\alpha)^2\, w(x)\, dx]^2}$$

$$\times \left[\int_\alpha^b (x-\alpha)\, w(x)\, dx \int_\beta^b (x-\alpha)\, w(x)\, dx \right.$$

$$\left. - \int_\alpha^b (x-\alpha)^2\, w(x)\, dx \int_\beta^b w(x)\, dx \right]$$

and so $D'_f(h_\beta) > 0$ by the Schwarz inequality. ▌

Acknowledgment

The authors are grateful to the referees for their comments which improved presentation of the paper.

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., and Balakrishnan, N. (1989). *Relations, Bounds and Approximations for Order Statistics.* Lecture Notes in Statistics, Vol. 53, Springer-Verlag, New York.

[3] Balakrishnan, A. V. (1981). *Applied Functional Analysis*, 2nd ed., Springer-Verlag, New York.

[4] Barlow, R. E., and Proschan, F. (1975). *Statistical Analysis of Reliability and Life Testing.* Holt, Rinehart, and Winston, New York.

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

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

[7] David, H. A. (1981). *Order Statistics*, 2nd ed. Wiley, New York.

[8] Dharmadhikari, S., and Joag-Dev, K. (1988). *Unimodality, Convexity, and Applications.* Academic Press, Boston.

[9] Lai, T. L., and Robbins, H. (1976). Maximally dependent random variables. *Proc. Nat. Acad. Sci. U.S.A.* **73** 286–288.

[10] Mallows, C. L. (1969). Extrema of expectations of uniform order statistics. *SIAM Rev.* **11** 410–411.

[11] Patel, J. K. (1983). Hazard rate and other classifications of distributions. In *Encyclopedia of Statistical Sciences* (S. Kotz, N. L. Johnson and C. B. Read, Eds.), pp. 590–594. Wiley, New York.

[12] Proschan, F. (1963). Theoretical explanations of observed decreasing failure rate. *Technometrics* **5** 375–384.

[13] RYCHLIK, T. (1992). Stochastically extremal distributions of order statistics for dependent samples. *Statist. Probab. Lett.* **13** 337–341.
[14] RYCHLIK, T. (1993). Bounds for expectation of *L*-estimates for dependent samples. *Statistics* **24** 1–7.
[15] RYCHLIK, T. (1993). Bias-robustness of *L*-estimates of location against dependence. *Statistics* **24** 9–15.
[16] RYCHLIK, T. (1993). Sharp bounds on *L*-estimates and their expectations for dependent samples. *Commun. Statist. Theory Methods* **22** 1053–1068.

Printed in Belgium