Tomasz Rychlik (Toruń)

# OPTIMAL MEAN-VARIANCE BOUNDS ON
# ORDER STATISTICS FROM FAMILIES
# DETERMINED BY STAR ORDERING

*Abstract.* We present optimal upper bounds for expectations of order statistics from i.i.d. samples with a common distribution function belonging to the restricted family of probability measures that either precede or follow a given one in the star ordering. The bounds for families with monotone failure density and rate on the average are specified. The results are obtained by projecting functions onto convex cones of Hilbert spaces.

**1. Introduction.** Consider a finite sequence $X_1, \ldots, X_n$ of independent identically distributed random variables, with common distribution function $F$, mean $\mu$ and variance $\sigma^2 < \infty$. The respective order statistics will be denoted by $X_{1:n}, \ldots, X_{n:n}$. Moriguti (1953) proved that for $2 \leq j \leq n-1$,

$$
(1.1) \qquad \mathrm{E}_F X_{j:n} - \mu = \int_0^1 F^{-1}(x)[f_{j:n}(x) - 1] \, dx
$$

$$
= \int_0^1 [F^{-1}(x) - \mu] \, d(F_{j:n} - U)(x)
$$

$$
\leq \int_0^1 [F^{-1}(x) - \mu] \, d(\overline{F}_{j:n} - U)(x)
$$

$$
\leq \Big\{ \int_0^1 [F^{-1}(x) - \mu]^2 \, dx \ \int_0^1 [\overline{f}_{j:n}(x) - 1]^2 \, dx \Big\}^{1/2}
$$

$$
= \|\overline{f}_{j:n} - 1\|_U \sigma,
$$

where
$$(1.2) \qquad f_{j:n}(x) = n\binom{n-1}{j-1}x^{j-1}(1-x)^{n-j}, \quad 1 \le j \le n,$$

is the density function of the $j$th order statistic of the standard uniform sample with distribution function $U(x) = x$, $0 \le x \le 1$, and

$$(1.3) \qquad F_{j:n}(x) = \int_0^x f_{j:n}(t)\,dt = \sum_{k=j}^n \binom{n}{k}x^k(1-x)^{n-k};$$

moreover

$$(1.4) \qquad \overline{F}_{j:n}(x) = \begin{cases} F_{j:n}(x) & \text{if } 0 \le x \le \delta, \\ f_{j:n}(\delta)(x-1)+1 & \text{if } \delta \le x \le 1, \end{cases}$$

with $\delta \in \left(0, \frac{j-1}{n-1}\right)$ uniquely determined by

$$(1.5) \qquad 1 - F_{j:n}(\delta) = (1-\delta)f_{j:n}(\delta),$$

is the greatest convex minorant of $F_{j:n}$, with derivative

$$(1.6) \qquad \overline{f}_{j:n}(x) = f_{j:n}(\min\{x,\delta\}),$$

and $\|\cdot\|_U$ denotes the $L^2$-norm on the unit interval with uniform weighting. Since $\int_0^1 \overline{f}_{j:n}(x)\,dx = \overline{F}_{j:n}(1) = F_{j:n}(1) = 1$, we have

$$(1.7) \quad \|\overline{f}_{j:n} - 1\|_U^2 = \int_0^1 \overline{f}_{j:n}^{\,2}(x)\,dx - 1 = \int_0^\delta f_{j:n}^2(x)\,dx + (1-\delta)f_{j:n}^2(\delta) - 1$$

$$= n\frac{\binom{2j-2}{j-1}\binom{2n-2j}{n-j}}{\binom{2n-1}{n}}F_{2j-1:2n-1}(\delta) + (1-\delta)f_{j:n}^2(\delta) - 1.$$

Explicit formulae for (1.7) in the special cases $j = 2, 3, n-2, n-1$ were presented in Balakrishnan (1993). The functions $\overline{f}_{j:n} - 1$ are the $L^2([0,1], dx)$-projections of the increasing-decreasing $f_{j:n} - 1$, $2 \le j \le n-1$, onto the convex cone of nondecreasing functions integrating to 0, which coincides with the class of all possible $F^{-1} - \mu$. Generally, the projection $Ph$ of an arbitrary $h$ onto a convex cone $\mathcal{C}$ of a real Hilbert space is characterized by two relations:

$$(1.8) \qquad\qquad \forall g \in \mathcal{C} \quad (g, h) \le (g, Ph),$$

$$(1.9) \qquad\qquad (Ph, h) = (Ph, Ph)$$

(see, e.g. Balakrishnan (1981, Section 1.4)). In our case, (1.8) is equivalent to (1.1), and (1.9) follows from

$$(1.10) \quad \int_0^1 [f_{j:n}(x) - 1][\overline{f}_{j:n}(x) - 1]\,dx$$

$$= \int_0^\delta [f_{j:n}(x) - 1]^2\,dx + [f_{j:n}(\delta) - 1]^2(1-\delta)$$

$$= \|\overline{f}_{j:n} - 1\|_U^2$$

(cf. (1.5)). By (1.10), the distribution function $F$ defined by

$$(1.11) \qquad F^{-1} - \mu = \frac{\overline{f}_{j:n} - 1}{\|\overline{f}_{j:n} - 1\|_U} \sigma,$$

i.e. the inverse of a polynomial with jump $1 - \delta$ at the right end-point of support is the unique marginal distribution with given first two moments that provides equality in (1.1).

Note that for the decreasing functions $f_{1:n} - 1$, the projection is constant $\overline{f}_{1:n} - 1 = 0$ (since $\overline{F}_{j:n} = U$), and the trivial bound $\mathrm{E}_F X_{1:n} \leq \mu$ holds, which becomes equality for the Dirac measure at $\mu$. For $j = n$, Gumbel (1954) and Hartley and David (1954) applied the Schwarz inequality to the second line of (1.1), and obtained

$$(1.12) \qquad \mathrm{E}_F X_{n:n} - \mu \leq \|f_{n:n} - 1\|_U \sigma = (n-1)(2n-1)^{-1/2} \sigma,$$

with equality attained for

$$(1.13) \qquad F\left(\frac{x - \mu}{\sigma}\right) = \left[\frac{n-1}{(2n-1)^{1/2} n} \cdot \frac{x - \mu}{\sigma} + \frac{1}{n}\right]^{1/(n-1)},$$

$$-\frac{(2n-1)^{1/2}}{n-1} \leq \frac{x - \mu}{\sigma} \leq (2n-1)^{1/2}$$

(cf. (1.11)). This is a consequence of the fact that $f_{n:n} - 1$ is actually increasing, and so it coincides with its projection. Accordingly, formula (1.1) expressed in terms of projections onto nondecreasing functions holds true for general $1 \leq j \leq n$, and (1.11) is a necessary and sufficient condition for equality.

The Moriguti projection method used together with the Hölder inequality provides more general bounds in terms of various central absolute moments of a single observation (see Rychlik (1998)). Raqab (1997) used the Moriguti method to find exact moment bounds on $k$th records. Gajek and Okolewski (2000) combined the greatest convex minorants with the Steffensen inequality and derived some general bounds for order and record statistics. We also refer to Robertson *et al.* (1988) for numerous applications of discrete versions of the Moriguti algorithm in order restricted statistical inference.

The idea of using projections to determine sharp bounds for statistical functionals on restricted families was first presented in Gajek and Rychlik (1996). That paper contains applications of the method for determining bounds on expected order statistics of dependent identically distributed samples with decreasing density and failure rate, and symmetric unimodal distribution, when either the second moment or the mean and variance of a single observation are given. Analogous problems for independent samples were solved in Gajek and Rychlik (1998). Dependent samples with increasing

densities and failure rates and symmetric $U$-shaped densities were studied in Rychlik (1998), and those with monotone failure rates on the average in Rychlik (2001). An application of the projection method for improvements of records was presented in Rychlik (1997).

The purpose of this paper is to find sharp upper mean-variance bounds for the expectations of order statistics of independent identically distributed random variables from restricted families of distributions defined by means of star ordering. We consider an absolutely continuous marginal distribution function $F$ with density function $f$ concentrated on an interval $[a_F, d_F)$, $-\infty < a_F < d_F \leq \infty$. We say that $F$ precedes $G$ in the star ordering ($F \prec_\star G$) iff $G^{-1}F$ is starshaped, i.e. $G^{-1}F(x)/(x - a_F)$ is nondecreasing on the support of $F$. Observe that the relation is a location-scale invariant partial ordering so that $F \prec_\star G$ implies the same for any location-scale transformations of $F$ and $G$. We extend here the standard notion of star ordering defined for distribution functions with the left end-point $a_F = a_G = 0$, which is merely scale invariant. Properties of starshaped functions were studied in Barlow *et al.* (1969). Barlow and Proschan (1966) determined some inequalities for $L$-statistics imposed by the star ordering of the distributions of observations. Shaked (1982) discovered a close relation of the star and dispersive orderings.

We consider restricted classes of parent distributions $F$ which either precede or succeed a given distribution $W$ in the star ordering. A particular emphasis will be laid on the cases of uniform $W(x) = U(x) = x$, $0 \leq x \leq 1$, and exponential distributions $W(x) = E(x) = 1 - e^{-x}$, $0 \leq x < \infty$. For the former, $F \prec_\star U$ (resp. $F \succ_\star U$) means that $(x - a_F)^{-1} \int_{a_F}^x f(t)\,dt$ is nondecreasing (resp. nonincreasing). We can say that the density is *nondecreasing* (resp. *nonincreasing*) *on the average*: although $f$ can be multimodal, the larger values in $[a_F, d_F)$ are more (resp. less) probable than the smaller ones. In the latter case, the relations $F \prec_\star E$ and $F \succ_\star E$ with $a_F = 0$ define notions of increasing and decreasing failure rate on the average $\frac{1}{x} \int_0^x \frac{f(t)}{1 - F(t)}\,dt$, respectively, which are of importance in reliability theory.

In Section 2, we study the bounds for order statistics of i.i.d. samples for classes of parent distributions with fixed maximal and minimal element $W$ in the star ordering. Assumptions on $W$ will be chosen so as to cover the cases of interest. We state and solve auxiliary problems of projecting a given function onto convex cones in a Hilbert space. In particular, we show that for $F \prec_\star W$, the general bounds (1.1) hold. For $F \succ_\star W$, the solution, dependent on the sample size $n$ and the number $j$ of the order statistic, is given up to a pair of numerical parameters determined by two equations. In Section 3, we specify the results for families with monotone density and failure rate on the average. For each bound, we describe a (limiting) location-

scale family of parent distributions that attains it. Section 4 contains the proofs of the results presented in Section 2.

**2. Projection problem.** Changing variables in (1.1), we get

$$
(2.1) \qquad \mathrm{E}_F X_{j:n} - \mu = \int_a^d [F^{-1}W(x) - \mu][f_{j:n}W(x) - 1]w(x)\, dx
$$
$$
= (F^{-1}W - \mu, f_{j:n}W - 1)_W,
$$

where $W$ is an absolutely continuous distribution function with density $w$, support $[a, d)$, and a finite second moment, and $(\cdot, \cdot)_W$ denotes the inner product in the real Hilbert space $L^2([a, d), w(x)\, dx)$ of square integrable functions with weight $w$ on the interval $[a, d)$. Since $W$ will be fixed, we simplify the notation by dropping subscripts of its support end-points. The transformation $F^{-1}W - \mu$ for arbitrary $F$ with a finite second moment satisfying $F \succ_\star W$ (resp. $F \prec_\star W$) defines a family $\mathcal{C}_1$ (resp. $\mathcal{C}_2$) of functions in $L^2([a, d), w(x)\, dx)$ as

$$
\mathcal{C}_i = \Big\{ g \in \widehat{\mathcal{C}}_i : \int_a^d g(x)w(x)\, dx = 0 \Big\}, \qquad i = 1, 2,
$$

with

$$
(2.2) \qquad \widehat{\mathcal{C}}_1 = \{ g \in L^2([a, d), w(x)\, dx) : g(x) \text{ and } [g(x) - g(a)]/(x - a)
$$
$$
\text{are nondecreasing} \},
$$

$$
(2.3) \qquad \widehat{\mathcal{C}}_2 = \{ g \in L^2([a, d), w(x)\, dx) : g(x) \text{ is nondecreasing}
$$
$$
\text{and } [g(x) - g(a)]/(x - a) \text{ is nonincreasing} \}.
$$

Observe that all $\mathcal{C}_i, \widehat{\mathcal{C}}_i$, $i = 1, 2$, are convex cones, and the latter pair is translation invariant: $g \in \widehat{\mathcal{C}}_i$ implies that $g + c \in \widehat{\mathcal{C}}_i$, $i = 1, 2$, for all real $c$. By the arguments of Section 1, for determining the bounds for $F \succ_\star W$ and $F \prec_\star W$ it is crucial to find the projections $P_i(f_{j:n}W - 1)$ of $f_{j:n}W - 1$ onto $\mathcal{C}_i$, $i = 1, 2$. Indeed, combining (2.1) with (1.8), we obtain

$$
(2.4) \qquad \mathrm{E}_F X_{j:n} - \mu \le \| P_i(f_{j:n}W - 1) \|_W \sigma, \qquad i = 1, 2
$$

(cf. (1.1)). Similarly, for

$$
F^{-1}W - \mu = \frac{P_i(f_{j:n}W - 1)}{\| P_i(f_{j:n}W - 1) \|_W} \sigma \in \mathcal{C}_i, \qquad i = 1, 2,
$$

we obtain equality in (2.4), using (2.1) and (1.9) (cf. also (1.10)).

We also notice that we can replace the original projection problems by simpler ones of projecting $f_{j:n}W$ onto (2.2) and (2.3). We first recall Lemma 1 of Rychlik (2001).

LEMMA 1. *If $\widehat{\mathcal{C}}$ is a translation invariant convex cone in the space $L^2([a, d), w(x)\, dx)$ with $\int_a^d w(x)\, dx = 1$, and the projection $\widehat{P}h$ of some $h \in L^2([a, d), w(x)\, dx)$ onto $\widehat{\mathcal{C}}$ exists, then $\int_a^d \widehat{P}h(x)w(x)\, dx = \int_a^d h(x)w(x)\, dx$.*

Since $\int_a^d [f_{j:n}W(x) - 1]w(x)\,dx = \int_0^1 f_{j:n}(x)\,dx - 1 = 0$, by Lemma 1, $P_i(f_{j:n}W - 1) = \widehat{P}(f_{j:n}W - 1)$, $i = 1, 2$. Furthermore, $\widehat{P}(h + c) = \widehat{P}h + c$ for all real $c$ and every translation invariant convexcone of $\widehat{C}$ implies that $\widehat{P}(f_{j:n}W - 1) = \widehat{P}f_{j:n}W - 1$, $i = 1, 2$.

Throughout this section we consider general weight functions $w : [a, d) \to [0, \infty)$ such that

$$(2.5) \qquad \int_a^d w(x)\,dx = 1, \quad \int_a^d x^2 w(x)\,dx < \infty,$$

and $0 \le h \in L^2([a, d), w(x)\,dx)$ satisfying

$$(2.6) \qquad \int_a^d h(x)w(x)\,dx = 1.$$

We also assume that $h$ is bounded continuously differentiable, and $h(a) = 0$, and there are $a \le b < c \le d$ such that $h$ is increasing and strictly convex in $(a, b)$, and increasing and strictly concave in $(b, c)$, and decreasing in $(c, d)$. Either of $(a, b)$ and $(c, d)$ may be empty. Here $h = f_{j:n}W$ for brevity of notation. Conditions (2.5) mean that $w$ is a density function on $[a, d)$ with a finite variance. Writing $\int_a^d f_{j:n}W(x)w(x)\,dx = \int_0^1 f_{j:n}(x)\,dx = 1$, we trivially check (2.6). The functions $f_{j:n}$, $2 \le j \le n - 1$, start with zero value, then increase, and eventually decrease, and so do the compositions $f_{j:n}W$ by strict increase of $W$. If $W = U$, then the assumptions of consecutive convexity and concavity are satisfied for $2 \le j \le n - 1$ with $a = b$, when $j = 2$. If $W = E$, the shape assumptions hold for $2 \le j \le n$ with $a = b$ and $c = d$ for $j = 2$ and $j = n$, respectively. Since $f_{1:n}W$ is decreasing, $P_i(f_{1:n}W - 1)$, $i = 1, 2$, are identical with the projections $\widehat{P}_i(f_{1:n}W - 1) = 0$ onto the class of all nondecreasing functions. This gives the zero bound for the sample minimum. We exclude this trivial case from the further study.

We now formulate the main results of this section. For $a \le p \le c$ and $p \le q \le c$, $q \ne a$, set

$$(2.7) \qquad G(p, q) = \int_q^d \left[ \frac{h(q) - h(p)}{q - a}(x - a) + h(p) - h(x) \right](x - a)w(x)\,dx,$$

$$(2.8) \qquad H(p, q) = \int_a^p [h(p) - h(x)]w(x)\,dx$$

$$+ \int_q^d \left[ \frac{h(q) - h(p)}{q - a}(x - a) + h(p) - h(x) \right]w(x)\,dx.$$

Then for $a < p \le c$,

$$(2.9) \qquad \overline{G}(p) = G(p,p) = \int_p^d [h(p) - h(x)](x - a)w(x)\,dx,$$

$$(2.10) \qquad \overline{H}(p) = H(p,p) = \int_a^d [h(p) - h(x)]w(x)\,dx = h(p) - 1.$$

THEOREM 1. *Let $p_0$ be the unique zero of* (2.10) *in* $(a,c)$. *If*

$$(2.11) \qquad \overline{G}(p_0) = \int_{p_0}^d [1 - h(x)](x - a)w(x)\,dx \ge 0,$$

*then $P_1 h = 1$. Otherwise there exists a unique pair $(p^\star, q^\star)$, $a \le p^\star < p_1 = \overline{G}^{-1}(0) < p_0$, $\max\{p^\star, b\} < q^\star < c$, determined by the equations*

$$(2.12) \qquad\qquad\qquad G(p,q) = 0,$$
$$(2.13) \qquad\qquad\qquad H(p,q) = 0,$$

*such that $P_1 h = h_{p^\star q^\star}$ for*

$$(2.14) \qquad h_{pq}(x) = \begin{cases} h(p) & \text{if } a \le x \le p, \\ h(x) & \text{if } p \le x \le q, \\ \dfrac{h(q) - h(p)}{q - a}(x - a) + h(p) & \text{if } q \le x < d. \end{cases}$$

Since $h$ is strictly increasing from $h(a) = 0$ to $h(c) = \sup h > 1$ (cf. (2.5) and (2.6)), $p_0$ is indeed well defined. If $a \le p \le q \le c$, $q \ne a$, then (2.14) is nondecreasing, and constant for $a < p = q \le c$ in particular. If $q^\star < \overline{q}(p^\star) \in (\max\{p^\star, b\}, c)$ defined in Lemma 2 of Section 4, then $h_{p^\star q^\star}$ is also starshaped. Existence and uniqueness of $(p^\star, q^\star)$ will follow from Lemma 3. Taking $h = f_{j:n}W$, we can deduce the following sharp bounds for the expectations of order statistics.

THEOREM 2. *Let $X_1, \ldots, X_n$ be i.i.d. random variables with finite mean $\mu$ and variance $\sigma^2$, and a distribution function $F$ such that $F \succ_\star W$. Suppose that the density $w$ of $W$ and $h = f_{j:n}W$ for some $2 \le j \le n$ satisfy the assumptions presented above (see (2.5)–(2.6)). If for $p_0 = W^{-1}f_{j:n}^{-1}(1) \in (a,c) = \left(W^{-1}(0), W^{-1}\left(\frac{j-1}{n-1}\right)\right)$ we have*

$$(2.15) \qquad \overline{G}_W(p_0) = \int_{p_0}^d [1 - f_{j:n}W(x)](x - a)w(x)\,dx \ge 0,$$

*then*
$$(2.16) \qquad\qquad\qquad \mathrm{E}_F X_{j:n} \le \mu,$$

*with equality when $F$ is concentrated at $\mu$. Otherwise there exists a pair*
$(p^\star, q^\star)$, $a \le p^\star < c$, $a < q^\star \in [p^\star, c)$, *determined by*

$$(2.17) \quad G_W(p,q) = \frac{Q_W - P_W}{q-a} \int\limits_q^d (x-a)^2 w(x)\,dx + P_W \int\limits_q^d (x-a)w(x)\,dx$$

$$- \int\limits_q^d f_{j:n}W(x)(x-a)w(x)\,dx = 0,$$

$$(2.18) \quad H_W(p,q) = P_W[W(p) + 1 - W(q)] - F_{j:n}W(p) - 1 + F_{j:n}W(q)$$

$$+ \frac{Q_W - P_W}{q-a} \int\limits_q^d (x-a)w(x)\,dx = 0,$$

*with* $P_W, Q_W$ *denoting* $f_{j:n}W(p), f_{j:n}W(q)$, *respectively. Then*

$$(2.19) \qquad\qquad \mathrm{E}_F X_{j:n} - \mu \le B(j,n)\sigma$$

*for*

$$(2.20) \quad B^2 = B^2(j,n) = P_W^2[W(p) + 1 - W(q)]$$

$$+ 2P_W \frac{Q_W - P_W}{q-a} \int\limits_q^d (x-a)w(x)\,dx$$

$$+ \left(\frac{Q_W - P_W}{q-a}\right)^2 \int\limits_q^d (x-a)^2 w(x)\,dx - 1$$

$$+ n \frac{\binom{2j-2}{j-1}\binom{2n-2j}{n-j}}{\binom{2n-1}{n}} [F_{2j-1:2n-1}W(q) - F_{2j-1:2n-1}W(p)].$$

*Equality in (2.19) is attained by*

$$(2.21) \quad F_W\left(\frac{x-\mu}{\sigma}\right)$$

$$= \begin{cases} 0 & \text{if } \dfrac{x-\mu}{\sigma} < -\dfrac{1-P_W}{B}, \\[2mm] f_{j:n}^{-1}\left(B\dfrac{x-\mu}{\sigma} + 1\right) & \text{if } -\dfrac{1-P_W}{B} \le \dfrac{x-\mu}{\sigma} < -\dfrac{1-Q_W}{B}, \\[2mm] W\left(a + \dfrac{(q-a)\left(B\frac{x-\mu}{\sigma} + 1 - P_W\right)}{Q_W - P_W}\right) & \text{if } \dfrac{x-\mu}{\sigma} \ge -\dfrac{1-Q_W}{B}. \end{cases}$$

The proof is straightforward once we apply the conclusions of Theorem 1 for determining $P_1(f_{j:n}W - 1) = \widehat{P}_1 h - 1$. Formula (2.11) can be specified as (2.15), and (2.16), apparently true for small order statistics, follows from the first statement. Rewriting (2.12)–(2.13) as (2.17)–(2.18), we calculate $B^2(j,n) = \|h_{p^\star q^\star} - 1\|_W^2 = \|h_{p^\star q^\star}\|_W^2 - 1$. Observe that (2.21) has a jump of height $W(p)$ at the left end-point, and, under linear transformations of

its argument, coincides then with the inverse of a Bernstein polynomial and ultimately with $W$ generating the class of parent distributions.

Formally, (2.21) does not belong to the class of absolutely continuous distributions following $W$ in the star ordering. However, it is easy to find sequences of absolutely continuous $F_k \succ_\star W$, $k \to \infty$, which attain the bounds asymptotically. It suffices that

$$\|F_k^{-1}W - F_W^{-1}W\|_W^2 = \int_0^1 [F_k^{-1}(x) - F_W^{-1}(x)]^2 \, dx \to 0.$$

For instance, if $P_1 h = h_{p^\star q^\star}$ for some $a \leq p^\star < q^\star$, then we can replace the constant part in (2.14) by

$$h_k(x) = h(p) - \frac{p - x}{k(p - a)} \sim F_k^{-1}W(x) - \mu, \quad a \leq x \leq p,$$

for sufficiently large $k$. Consequently, $\|F_k^{-1}W - F_W^{-1}W\|_W^2 \leq \text{const } W(p)/k^2 \to 0$, and $F_k^{-1}W - \mu \in \widehat{\mathcal{C}}_1$. Since $(F_k^{-1}W - F_W^{-1}W, 1)_W = (F_k^{-1}W - \mu, 1)_W \to 0$, we have $\mathcal{C}_1 \ni F_k^{-1}W - (F_k^{-1}W, 1)_W \to F_W^{-1}W - \mu$. Such $F_k$ have increasing densities proportional to $w$ in shrinking left neighborhoods of the left-end support point of $F_W$. Details of the construction of approximations for other forms of projections are left to the reader.

Theorem 3 asserts that the best bounds for order statistics of i.i.d. samples with distributions $F \prec_\star W$ are identical with general ones described in (1.1)–(1.7). Since the distributions attaining the bounds (see (1.11)) have jumps at their right-end support points, except for (1.13) in case $j = n$, they should also be approximated by absolutely continuous sequences.

THEOREM 3. *If $X_1, \ldots, X_n$ are i.i.d. with distribution function $F \prec_\star W$, finite mean and variance, and the assumptions on $w = dW/dx$ and $h = f_{j:n}W$ are satisfied, then (1.1) with (1.7) provide the best possible bounds for $E_F(X_{j:n} - \mu)/\sigma$. They are attained in limit by sequences $F_k \prec_\star W$ whose quantile functions tend in $L^2([0,1], dx)$ to one defined in (1.11). In particular, the general bounds cannot be improved in the classes of distributions with increasing density function and failure rate on the average.*

**3. Distributions with decreasing density and failure rate on the average.** To shorten notation, we write $P_U = f_{j:n}(p)$ and $Q_U = f_{j:n}(q)$ for some fixed $2 \leq j \leq n-1$ and some $0 \leq p < q < \frac{j-1}{n-1}$ defined in Proposition 1 below. For $W = U$, (2.17) and (2.18) have the forms

(3.1) $\quad G_U(p, q) = \dfrac{1 - q^3}{3q} Q_U - \dfrac{(1-q)^2(2+q)}{6q} P_U - \dfrac{j}{n+1} [1 - F_{j+1:n+1}(q)],$

(3.2) $\quad H_U(p, q) = \left[ p - \dfrac{(1-q)^2}{2q} \right] P_U + \dfrac{1 - q^2}{2q} Q_U - 1 - F_{j:n}(p) + F_{j:n}(q).$

PROPOSITION 1. *Let $X_1, \ldots, X_n$ be i.i.d. with distribution function $F$, and $\mu, \sigma^2$ and $a$ stand for the finite mean, variance and left-end support point. Assume that $F(x)/(x-a)$ is nonincreasing for $x > a$. For given $2 \le j \le n-1$, let $p_U \in \left(0, \frac{j-1}{n-1}\right)$ be the unique point such that $f_{j:n}(p_U) = 1$. If*

$$(3.3) \qquad\qquad 1 - p_U^2 \ge \frac{2j}{n+1}\,[1 - F_{j+1:n+1}(p_U)],$$

*then (2.16) holds. Otherwise there are unique $0 < p < p_U$, $p < q < \frac{j-1}{n-1}$ at which (3.1) and (3.2) vanish, and*

$$(3.4) \qquad\qquad \mathrm{E}_F X_{j:n} - \mu \le B_U(j,n)\sigma,$$

*where*

$$(3.5) \quad B_U^2 = B_U^2(j,n)$$
$$= \left[p + \frac{(1-q)^3}{3q^2}\right] P_U^2 - \frac{(1-q)^2(2+q)}{3q^2}\,P_U Q_U + \frac{1-q^3}{3q^2}\,Q_U^2$$
$$+ n\,\frac{\binom{2j-2}{j-1}\binom{2n-2j}{n-j}}{\binom{2n-1}{n}}\,[F_{2j-1:2n-1}(q) - F_{2j-1:2n-1}(p)] - 1.$$

*Equality in (3.4) holds for the location-scale family of distributions*

$$(3.6) \quad F_U\!\left(\frac{x-\mu}{\sigma}\right)$$

$$= \begin{cases} 0 & \text{if } \dfrac{x-\mu}{\sigma} < -\dfrac{1-P_U}{B_U}, \\[2ex] f_{j:n}^{-1}\!\left(B_U\,\dfrac{x-\mu}{\sigma}+1\right) & \text{if } -\dfrac{1-P_U}{B_U} \le \dfrac{x-\mu}{\sigma} < -\dfrac{1-Q_U}{B_U}, \\[2ex] \dfrac{q\left(B_U\frac{x-\mu}{\sigma}+1-P_U\right)}{Q_U - P_U} & \text{if } -\dfrac{1-Q_U}{B_U} \le \dfrac{x-\mu}{\sigma} < \dfrac{(Q_U-P_U)/q - 1 + P_U}{B_U}, \\[2ex] 1 & \text{if } \dfrac{x-\mu}{\sigma} \ge \dfrac{(Q_U-P_U)/q - 1 + P_U}{B_U}. \end{cases}$$

It is clear that condition (3.3) and the trivial bound (2.16) hold for small order statistics. For greater ones, we have (3.4)–(3.5), which are specified from general results of Theorem 2. The extreme distribution function (3.6) has a jump at the left end, then coincides with the inverse of a Bernstein polynomial of degree $n-1$, and is uniform in the right part. Observe that (1.13) has a decreasing density. Therefore it satisfies the assumptions of Proposition 1, and provides the accurate bound (1.12) for the maximum of a sample with decreasing density on the average.

With the notation $P_E = f_{j:n}(1 - e^{-p}) < Q_E = f_{j:n}(1 - e^{-q})$, $2 \le j \le n$, we now specify

(3.7)    $G_E(p, q) = (q + 2 + 2/q)e^{-q}Q_E - (1 + 2/q)e^{-q}P_E$

$$- \sum_{k=1}^{j} \frac{1}{n + 1 - k} F_{n+1-k:n}(e^{-q}) - qF_{n+1-j:n}(e^{-q}),$$

(3.8)    $H_E(p, q) = (1 - e^{-p} - e^{-q}/q)P_E + (1 + 1/q)e^{-q}Q_E$
$$- 1 + F_{n+1-j:n}(e^{-p}) - F_{n+1-j:n}(e^{-q}).$$

We derive (3.7)–(3.8) from (2.17)–(2.18), respectively, combining elementary calculus formulae with

$$\int_{q}^{\infty} xf_{j:n}(1 - e^{-x})e^{-x}\, dx = \sum_{k=1}^{j} \frac{1}{n + 1 - k} F_{n+1-k:n}(e^{-q}) + qF_{n+1-j:n}(e^{-q}).$$

PROPOSITION 2. *In the notation of Proposition 1, set* $p_E = -\ln(1 - p_U)$ *and assume that* $-\ln[1 - F(x)]/(x - a)$ *is nonincreasing for* $x > a$. *If for* $2 \le j \le n$ *and* $p = p_E$, *we have*

$$\overline{G}_E(p) = (1 + p)e^{-p} - \sum_{k=1}^{j} \frac{1}{n + 1 - k} F_{n+1-k:n}(e^{-p}) - pF_{n+1-j:n}(e^{-p}) \ge 0,$$

*then* (2.16) *holds. Otherwise there exist* $0 < p < p_E$, $p < q < \ln\frac{n-1}{n-j}$ *for which both* (3.7) *and* (3.8) *are equal to* 0, *and*

(3.9)                       $\mathrm{E}_F X_{j:n} - \mu \le B_E(j, n)\sigma,$

*where*

(3.10)    $B_E^2 = B_E^2(j, n) = (1 - e^{-p} + 2q^{-2}e^{-q})P_E^2 - 2(2 + q)q^{-2}e^{-q}P_EQ_E$
$$+ (2 + 2q + q^2)q^{-2}e^{-q}Q_E^2 - 1$$
$$+ n\frac{\binom{2j-2}{j-1}\binom{2n-2j}{n-j}}{\binom{2n-1}{n}}[F_{2n+1-2j:2n-1}(e^{-p}) - F_{2n+1-2j:2n-1}(e^{-q})].$$

*Equality holds in* (3.9) *for*

(3.11)    $F_E\left(\frac{x - \mu}{\sigma}\right)$

$$= \begin{cases} 0 & \text{if } \dfrac{x - \mu}{\sigma} < -\dfrac{1 - P_E}{B_E}, \\[2ex] f_{j:n}^{-1}\left(B_E\dfrac{x - \mu}{\sigma} + 1\right) & \text{if } -\dfrac{1 - P_E}{B_E} \le \dfrac{x - \mu}{\sigma} < -\dfrac{1 - Q_E}{B_E}, \\[2ex] 1 - \exp\left(-\dfrac{q\left(B_E\frac{x-\mu}{\sigma} + 1 - P_E\right)}{Q_E - P_E}\right) & \text{if } \dfrac{x - \mu}{\sigma} \ge -\dfrac{1 - Q_E}{B_E}. \end{cases}$$

The shapes of the distribution functions (3.6) and (3.11) differ only in the right tails: the former is uniform, and the latter is exponential. If $\mu = (1 - P_E)\sigma/B_E$, then (3.11) is a DFRA life distribution starting at 0.

In Table 1, we present a numerical comparison of the bounds described in Propositions 1 and 2 with general ones for order statistics from samples of size 15. By Theorem 3, the bounds for distributions with increasing density and failure rate on the average are identical with the general ones. Zero bounds indicate that the expectation of the order statistic cannot exceed $\mu$ for any distribution of a given family. Observe that

$$(3.12) \quad \mathrm{E}_U X_{7:15} < \sup_{F \succ_\star U} \mathrm{E}_F X_{7:15} \leq \mu = \mathrm{E}_U X_{8:15} < \sup_{F \succ_\star U} \mathrm{E}_F X_{8:15},$$

$$(3.13) \quad \mathrm{E}_E X_{9:15} < \sup_{F \succ_\star E} \mathrm{E}_F X_{9:15} \leq \mu < \mathrm{E}_E X_{10:15} < \sup_{F \succ_\star U} \mathrm{E}_F X_{10:15}.$$

The last inequalities in (3.12) and (3.13) are obvious, because for any $W$ we have $W \succ_\star W$, and (2.19) becomes equality for $F_W \neq W$ (see (2.21)). It is not apparent why $\mathrm{E}_W X_{j:n} < \mu$ implies $B_W(j, n) = 0$ for $W = U, E$ for a number of examples of sample sizes, different from 15 as well. Table 1 also contains values of parameters (see (1.5), (3.1)–(3.2), and (3.7)–(3.8)) which enable us to determine the bounds and respective extreme distributions (see (1.11) with (1.6), and (3.6), and (3.11)).

**Table 1.** Bounds on expectations of $X_{j:15}$, $1 \leq j \leq 15$, for i.i.d. samples with general distribution (increasing density and failure rate on the average), and decreasing density and failure rate on the average

| | general distribution | | decreasing density on average | | | decreasing failure rate on average | | |
|---|---|---|---|---|---|---|---|---|
| $j$ | $\delta$ | $\|\bar{f}_{j:15} - 1\|_U$ | $p$ | $q$ | $B_U(j, 15)$ | $p$ | $q$ | $B_E(j, 15)$ |
| 1 | — | **0** | — | — | **0** | — | — | **0** |
| 2 | 0.00510 | **0.04056** | — | — | **0** | — | — | **0** |
| 3 | 0.03358 | **0.12918** | — | — | **0** | — | — | **0** |
| 4 | 0.07789 | **0.21656** | — | — | **0** | — | — | **0** |
| 5 | 0.13206 | **0.30127** | — | — | **0** | — | — | **0** |
| 6 | 0.19339 | **0.38562** | — | — | **0** | — | — | **0** |
| 7 | 0.26050 | **0.47223** | — | — | **0** | — | — | **0** |
| 8 | 0.33264 | **0.56384** | 0.29572 | 0.30248 | **0.09398** | — | — | **0** |
| 9 | 0.40945 | **0.66376** | 0.33460 | 0.36546 | **0.31211** | — | — | **0** |
| 10 | 0.49085 | **0.77663** | 0.37258 | 0.44111 | **0.52643** | 0.56252 | 0.57152 | **0.10876** |
| 11 | 0.57705 | **0.90967** | 0.40975 | 0.53016 | **0.74334** | 0.66193 | 0.70031 | **0.32040** |
| 12 | 0.66863 | **1.07586** | 0.44577 | 0.63093 | **0.97623** | 0.77152 | 0.87277 | **0.57884** |
| 13 | 0.76682 | **1.30263** | 0.47863 | 0.74153 | **1.25356** | 0.89202 | 1.13102 | **0.91738** |
| 14 | 0.87441 | **1.66662** | 0.49832 | 0.86228 | **1.65242** | 1.01481 | 1.59609 | **1.42872** |
| 15 | — | **2.59973** | — | — | **2.59973** | 0.95203 | 3.25991 | **2.58621** |

**4. Proofs.** The proof of Theorem 1 is preceded by two lemmas. For fixed $p \in [a, c)$, let

$$(4.1) \qquad s_p(q) = [h(q) - h(p)]/(q - a), \qquad a \neq q \in [p, c),$$

denote the slopes of the lines $l_{AQ}(x) = s_p(q)(x - a) + h(p)$ passing through $A = (a, h(p))$ and the points $Q = (q, h(q))$, $p \leq q < c$, of the graph of $h$ (cf. the last line of (2.14)). Write $s_a(a) = h'(a) \geq 0$ for completeness.

LEMMA 2. *For every $p \in [a, c)$ there exists $\overline{q}(p) \in \mathcal{B} = [\max\{p, b\}, c)$ such that $s_p$ increases on $(p, \overline{q}(p))$ and decreases on $(\overline{q}(p), c)$.*

*Proof.* The statement is trivial for $p = a = b$ with $\overline{q}(p) = a$. Indeed, by concavity of $h$ on $(a, c)$, for any $a < q_1 < q_2 < c$ the point $Q_1$ lies above $l_{AQ_2}$ and so $s_a(q_1) > s_a(q_2)$.

We now claim that $s_p$ is increasing-decreasing on $\mathcal{B}$ if either $p > a = b$ or $p \geq a \neq b$. The proof consists in checking that the following is false: for every fixed $q \in \mathcal{B}$ there exist $q_1 < q < q_2$, $q_i \in \mathcal{B}$, such that $s_p(q_i) \geq s_p(q)$, $i = 1, 2$. If $l_{AQ}$ is tangent to $h$ at $q$, it runs above all graph points $Q'$, $q \neq q' \in \mathcal{B}$. Accordingly, $s_p(q') < s_p(q)$. If $l_{AQ}$ is a secant line to the curve, then these cross each other in $\mathcal{B}$ once more at most. If $h - l_{AQ}$ changes its sign from $-$ to $+$ at $q$, then $h(q_1) < l_{AQ}(q_1)$ and $s_p(q_1) < s_p(q)$ for all $\mathcal{B} \ni q_1 < q$. If the sign changes from $+$ to $-$, then the analogous relations hold for $\mathcal{B} \ni q_2 > q$. Therefore our claim is indeed true.

By continuity of $s_p$, it remains to show that (4.1) is increasing in $(p, b)$ when $a \leq p < b$. Note that $\max\{h(p), h(x)\}$ is convex in $(a, b)$ and strictly convex in $(p, b)$. For every $p \leq q_1 < q_2 \leq b$ the point $Q_1$ lies below $l_{AQ_2}$, and hence $s_p(q_1) < s_p(q_2)$. This completes the proof. ∎

It is clear that $\overline{q}(p)$ is the point where $l_{A\overline{Q}(p)}$ is tangent to $h$ and therefore can be determined by

(4.2) $$h(q) - h(p) = h'(q)(q - a).$$

We now take an arbitrary $g \in \widehat{\mathcal{C}}_1$ and modify it to improve the approximation of $h$. If $g(a) \geq h(c) = \max_{a \leq x \leq d} h(x)$, then $h \neq g \geq h$, and, by Lemma 1, subtracting a constant $c > g(a) - h(c)$ from $g$ gives a function closer to $h$ than the original one. If $g(a) < 0$, then $\max\{g, 0\} \in \widehat{\mathcal{C}}_1$ lies between $g$ and $h$ by nonnegativity of $h$. Therefore it suffices to study the functions $g \in \widehat{\mathcal{C}}_1$ such that $h(a) = 0 \leq g(a) < h(c)$.

LEMMA 3. *For every $p \in [a, c)$ and $g \in \widehat{\mathcal{C}}_1$ such that $g(a) = h(p)$ there exists $a \neq q \in [p, \overline{q}(p))$ (see (4.2)) such that for $h_{pq}$ defined by (2.14) we have*

(4.3) $$\|h_{pq} - h\|_W \leq \|g - h\|_W.$$

*Proof.* If $g(\overline{q}(p)) \geq h(\overline{q}(p))$, then we can take $h_{p\overline{q}(p)}$. Since $h(x) < h(p)$ on $[a, p)$, we have $h < h_{p\overline{q}(p)} \leq g$ there, because the constant $h(p)$ is the smallest nondecreasing function starting from level $h(p)$. Also, $h_{p\overline{q}(p)} = h$ is the optimal approximation of $h$ on $[p, \overline{q}(p)]$. Finally, for $x > \overline{q}(p)$ we have

$$h(x) \leq \frac{h(\overline{q}(p)) - h(p)}{\overline{q}(p) - a} (x - a) + h(p) \leq g(x).$$

The former inequality holds because the middle term defines the line that is tangent to $h$ at $\overline{q}(p)$ and majorizes $h$. The latter is a consequence of the fact that the line is the smallest starshaped function in $[\overline{q}(p), d)$ that passes through $A$ and $\overline{Q}(p) = (\overline{q}(p), h(\overline{q}(p)))$. The above arguments show that $h_{p\overline{q}(p)}$ indeed satisfies (4.3).

However, $h \neq h_{p\overline{q}(p)} \geq h$ and, due to Lemma 1, the approximation can be further improved by a downward translation of $h_{p\overline{q}(p)}$. So we consider the case $g(\overline{q}(p)) < h(\overline{q}(p))$. Define $r = \inf\{x > \overline{q}(p) : g(x) \geq h(x)\}$, putting $r = d$ if $g < h$ in $(\overline{q}(p), d)$. Consider

$$l_{AR}(x) = \frac{h(r) - h(p)}{r - a}(x - a) + h(p).$$

We first prove that $l_{AR} < h$ in $(\overline{q}(p), r)$. If $(c, r) \neq \emptyset$, then $l_{AR} - h$ is strictly increasing there, and $l_{AR}(r) - h(r) \leq 0$. For $x \in (\overline{q}(p), \min\{c, r\})$ the inequality follows from the strict convexity of $l_{AR} - h$ in the interval, and its nonpositivity at the end-points. Now we check that $l_{AR} - h$ changes sign in $(a, \overline{q}(p))$ once at most. Assume that $q$ is the largest point of sign change in $(a, \overline{q}(p))$. Evidently $l_{AR} - h$ is positive and negative to the left and right of $q$, respectively. If $q \in [b, \overline{q}(p))$, then convexity of $l_{AR} - h$ implies its positivity in $(b, q)$. This is also positive in $(a, b)$ by its concavity there, and nonnegativity at the end-points. If $q \in (a, b)$, it suffices to repeat the above argument with $b$ replaced by $q$. Assume that a unique $q$ exists. We can write

$$l_{AR}(x) = \frac{h(q) - h(p)}{q - a}(x - a) + h(p).$$

Since $l_{AR} \geq h(p)$, we get $q \in [p, \overline{q}(p))$. Note that in the class of nondecreasing starshaped functions passing through $A = (a, h(p))$ and $(r, g(r))$, the function $l_{AR}$ is maximal in $[a, r]$ and minimal in $[r, d)$. Therefore $g \leq l_{AR} \leq h$ in $[q, r]$ and $h \leq l_{AR} \leq g$ in $[r, d)$. Also, $h \leq h(p) \leq g$ in $[a, p]$. Consequently, $h_{pq}$ defined to equal $h(p)$, $h$, and $l_{AR}$ in $[a, p]$, $[p, q]$, and $[q, d)$, respectively, lies closer to $h$ than $g$ does. Since $p \leq q < \overline{q}(p)$, we see that $h_{pq} \in \widehat{\mathcal{C}}_1$.

If $p = a$, it may happen that $l_{AR}$ has no sign changes in $[a, r]$. Then $l_{AR}(a) = h(a) = 0$ and $l_{AR} < h$ in $(a, r)$, while $l_{AR} > h$ in $(r, d)$. The last relation is a consequence of concavity and ultimate decrease of $h$. If $l_{AR} = 0 \leq h$, referring to Lemma 1 we decrease the $L^2$-distance to $h$ by adding a positive constant $l_{AR}^{+}$. If $l_{AR}$ has a positive slope, we can take a line $l_{AR}^{-}$ running through $(r, g(r))$ with a slightly smaller slope. Observe that $l_{AR}^{-}$ lies closer to $h$ than $l_{AR}$ does in $[a, r]$ and $[r, d)$. Both the modifications lead to linear functions which cross $h$ once at some $q < \overline{q}(p)$. We are thus in a position to apply the constuction of the previous paragraph, which ends the proof. ∎

*Proof of Theorem 1.* By Lemma 3, we should minimize

$$(4.4) \quad D(p,q) = \int\limits_a^p [h(p) - h(x)]^2 w(x)\, dx + \int\limits_q^d [l_{AQ}(x) - h(x)]^2 w(x)\, dx$$

with respect to two parameters $p \in [a, c)$ and $a \neq q \in [\max\{p, b\}, \overline{q}(p)) \subset (a, c)$. Fixing $p$ and differentiating (4.4) with respect to $q$, we obtain

$$(4.5) \qquad \frac{\partial D(p,q)}{\partial q} = \frac{2}{q - a} [h'(q) - l'_{AQ}] G(p,q)$$

(cf. (2.7)). If $p < q < \overline{q}(p)$, then $h - l_{AQ}$ changes sign from $-$ to $+$ at $q$ and hence the expression in brackets is positive. This vanishes at $\overline{q}(p)$, but $\overline{q}(p)$ cannot be optimal, because $h_{p\overline{q}(p)} \geq h$. Analyzing the sign of (4.5), it suffices to concentrate on (2.7). We have

$$\frac{\partial G(p,q)}{\partial q} = \frac{h'(q) - l'_{AQ}}{q - a} \int\limits_q^d (x - a)^2 w(x)\, dx > 0,$$

which implies that (4.5) is the product of a positive function and increasing $G(p, \cdot)$. Since the integrand is positive for $q = \overline{q}(p)$, we have $G(p, \overline{q}(p)) > 0$. If $q = p > a$, then $l_{AQ} = h(p)$. It follows that $\overline{G}(p) < 0$ as $p \searrow a$ and $\overline{G}(p) > 0$ for $p \nearrow c$.

We can summarize the behaviour of (4.4) as follows. If $p$ is small enough then (4.5) is negative for $q$ close to $p$, and changes its sign at a $q^\star(p) \in (p, \overline{q}(p))$ where (2.7) vanishes and the unique minimum of $D(p, \cdot)$ is attained. If $p \geq p_1$ satisfying $\overline{G}(p_1) = 0$ then (4.5) is positive for all $q > p$. Then $D(p, \cdot)$ is minimized at $q^\star(p) = p$, which gives a constant approximation $h_{pp} = h(p)$. It remains to choose $p \in [a, c)$ such that $(p, q^\star(p))$ minimizes (4.4), where $q^\star(p) > p$ satisfies $G(p, q^\star(p)) = 0$ for $p < p_1$ and $q^\star(p) = p$ for $p \geq p_1$.

By Lemma 1, a necessary condition for that is

$$(4.6) \qquad H(p, q^\star(p)) = \int\limits_a^d [h_{pq^\star(p)}(x) - h(x)] w(x)\, dx = 0$$

(cf. (2.8)). It is clear that (2.10) strictly increases from negative $\overline{H}(a)$ to positive $\overline{H}(c) = H(c, q^\star(c))$. We have $\overline{H}(p_0) = 0$ if $h_{p_0 p_0} = h(p_0) = \int_a^d h(x) w(x)\, dx = 1$. We show that $H(p, q^\star(p))$ is also increasing when $q^\star(p) > p$ is determined by (2.12). Consider

$$(4.7) \quad \frac{dH(p, q^\star(p))}{dp} = h'(p) \left[ W(p) + 1 - W(q^\star(p)) - \int\limits_{q^\star(p)}^d \frac{(x - a)w(x)}{q^\star(p) - a}\, dx \right]$$

$$+ \frac{dq^\star(p)}{dp} [h'(q^\star(p)) - l'_{AQ^\star(p)}] \int\limits_{q^\star(p)}^d \frac{(x - a)w(x)}{q^\star(p) - a}\, dx.$$

Plugging

$$\frac{dq^\star(p)}{dp} = \frac{h'(p)\int_{q^\star(p)}^d [x - q^\star(p)](x-a)w(x)\,dx}{[h'(q^\star(p)) - l'_{AQ^\star(p)}]\int_{q^\star(p)}^d (x-a)^2 w(x)\,dx} > 0$$

into (4.7), we obtain

$$(4.8) \quad \frac{dH(p, q^\star(p))}{dp} = \frac{h'(p)}{\int_{q^\star(p)}^d (x-a)^2 w(x)\,dx}\left\{ W(p)\int_{q^\star(p)}^d (x-a)^2 w(x)\,dx \right.$$

$$\left. + [1 - W(q^\star(p))]\int_{q^\star(p)}^d (x-a)^2 w(x)\,dx - \left[\int_{q^\star(p)}^d (x-a)w(x)\,dx\right]^2 \right\}.$$

The last line is positive by the Schwarz inequality and the same holds for (4.8).

We are thus led to the following conclusions. If $p_0 \geq p_1$ then $(p^\star, q^\star(p^\star)) = (p_0, p_0)$ is the unique pair satisfying the necessary condition (4.6) for minimizing (4.4). This gives the first statement of Theorem 1. To see that $p_0 \geq p_1$ coincides with (2.11), we note that (2.9) satisfies $\lim_{p \searrow a} \overline{G}(p) < 0 < \lim_{p \nearrow c} \overline{G}(p)$ and $\overline{G}'(p) = h'(p)\int_p^d (x-a)w(x)\,dx > 0$. Therefore $\overline{G}$ is increasing and has a single zero at $p_1$. The same holds for $\overline{H}$ which vanishes at $p_0$. Hence the conditions $p_0 \geq p_1$ and (2.11) are equivalent. If $p_0 < p_1$, then $\overline{H}(p_1) = \lim_{p \nearrow p_1} H(p, q^\star(p)) > 0$, and we can make $H(p, q^\star(p))$ smaller by decreasing $p$. There is a $p^\star \in [a, p_1)$ that satisfies (4.6), because the opposite contradicts the existence of solution. Taking $q^\star = q^\star(p^\star)$, we see that (4.6) and (2.13) are identical, and (2.12) holds by the definition of $q^\star(p^\star)$. ∎

*Proof of Theorem 3.* The proof consists in showing that $\bar{h} = \bar{f}_{j:n}W$ is a limit point of $\widehat{\mathcal{C}}_2$. Then $\int_a^d \bar{h}(x)w(x)\,dx = \int_0^1 \bar{f}_{j:n}(x)\,dx = 1$ implies that $\bar{h} - 1$ belongs to the border of $\mathcal{C}_2$. On the other hand, by a change of variables, (1.1) can be rewritten as $E_F X_{j:n} - \mu \leq \|\bar{h} - 1\|_W \sigma$. By construction, $\bar{h}(x) = f_{j:n}W(\min\{x, \gamma\})$ for $a = W^{-1}(0) < \gamma = W^{-1}(\delta) < c = W^{-1}\left(\frac{j-1}{n-1}\right)$ (cf. (1.6)). We assumed that $h$ is increasing on $(a, c)$, strictly convex on some $(a, b)$ and strictly concave on $(b, c) \neq \emptyset$. Therefore $\bar{h}$, identical with $h$ in $[a, \gamma]$, is increasing, and either concave or convex-concave or convex there. It follows that $\sup_{a \leq x \leq \gamma} h'(x)$ is attained at $\beta = \min\{b, \gamma\}$. Define straight lines $l_t$ secant to the graph of $h$ at $t \in [a, d)$ by $l_t(x) = h'(t)(x - t) + h(t)$, and consider $s_\beta(t) = [h(t) - l_\beta(a)]/(t - a)$.

We show that $s_\beta$ is nonincreasing for $t > a$ by checking that

$$(4.9) \qquad s'_\beta(t) = [h'(t)(t-a) + l_\beta(a) - h(t)]/(t-a)^2 \leq 0.$$

If $t \in (a, \beta)$, then $l_\beta(t) < l_t(t) = \bar{h}(t)$, and $0 \leq \bar{h}'(t) = l'_t < \bar{h}'(\beta) = l'_\beta$ by convexity of $h$. It follows that $l_t(x) > l_\beta(x)$ for all $x < t$, and $l_t(a) > l_\beta(a)$ in

particular. Since the tangent lines lie above the concave curve $\bar{h}(t)$, $t > \beta$, we have $l_\beta(\beta) = \bar{h}(\beta) < l_t(\beta)$ and $l_t(t) = \bar{h}(t) < l_\beta(t)$. This implies that $l_t$ and $l_\beta$ cross each other somewhere in $(\beta, t)$, and so $l_t(t) > l_\beta(t)$ for all $t$. Therefore $h(t) = l_t(t) = h'(t)(t-a) + l_t(a) \geq h'(t)(t-a) + l_\beta(a)$ and (4.9) holds.

Consider now

$$\bar{h}_k(t) = \begin{cases} s_\beta(a + 1/k)(t - a) + l_\beta(a) & \text{if } a \leq t \leq a + 1/k, \\ \bar{h}(t) & \text{if } a + 1/k \leq t < d, \end{cases}$$

for integers $k > 1/(\beta - a)$. By definition, $\bar{h}_k(a) = l_\beta(a)$, $\bar{h}_k$ is continuous nondecreasing and $[\bar{h}_k(t) - \bar{h}_k(a)]/(t-a) = s_\beta(\max\{a + 1/k, t\})$ is continuous nonincreasing. Since $\|\bar{h} - \bar{h}_k\|_W^2 \leq [\bar{h}(a + 1/k) - l_\beta(a)]W(a + 1/k) \to 0$, we have $\mathcal{C}_2 \ni \bar{h}_k - \int_a^d \bar{h}_k(x)w(x)\,dx \to \bar{h}$ in $L^2([a,d), w(x)\,dx)$ as well. ∎

## References

A. V. Balakrishnan (1981), *Applied Functional Analysis*, 2nd ed., Springer, New York.

N. Balakrishnan (1993), *A simple application of binomial-negative relationship in the derivation of sharp bounds for moments of order statistics based on greatest convex minorants*, Statist. Probab. Lett. 18, 301–305.

R. E. Barlow, A. W. Marshall and F. Proschan (1969), *Some inequalities for starshaped and convex functions*, Pacific J. Math. 29, 19–42.

R. E. Barlow and F. Proschan (1966), *Inequalities for linear combinations of order statistics from restricted families*, Ann. Math. Statist. 37, 1574–1591.

L. Gajek and A. Okolewski (2000), *Sharp bound on moments of generalized order statistics*, Metrika 52, 27–43.

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

L. Gajek and T. Rychlik (1998), *Projection method for moment bounds on order statistics from restricted families. II. Independent case*, ibid. 64, 156–182.

E. J. Gumbel (1954), *The maxima of the mean largest value and of the range*, Ann. Math. Statist. 25, 76–84.

H. O. Hartley and H. A. David (1954), *Universal bounds for mean range and extreme observation*, ibid. 25, 85–99.

S. Moriguti (1953), *A modification of Schwarz's inequality with applications to distributions*, ibid. 24, 107–113.

M. Z. Raqab (1997), *Bounds based on greatest convex minorants for moments of record values*, Statist. Probab. Lett. 36, 35–41.

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

T. Rychlik (1997), *Evaluating improvements of records*, Appl. Math. (Warsaw) 24, 315–324.

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

T. Rychlik (2001), *Mean-variance bounds for order statistics from dependent DFR*, *IFR*, *DFRA and IFRA samples*, J. Statist. Plann. Inference 92, 21–38.

M. Shaked (1982), *Dispersive ordering of distributions*, J. Appl. Probab. 19, 310–320.

Institute of Mathematics
Polish Academy of Sciences
Chopina 12
87-100 Toruń, Poland
E-mail: trychlik@impan.gov.pl