REVSTAT – Statistical Journal
Volume 21, Number 2, April 2023, 235–266
https://doi.org/10.57805/revstat.v21i2.402

Conditional Evaluations of Sums of Sample Maxima and Records

Authors: Tomasz Rychlik
– Institute of Mathematics, Polish Academy of Sciences,
Śniadeckich 8, 00656 Warsaw, Poland
trychlik@impan.pl

Magdalena Szymkowiak ✉
– Institute of Automatic Control and Robotics, Poznan University of Technology,
Plac Marii Skłodowskiej-Curie 5, 60965 Poznań, Poland
magdalena.szymkowiak@put.poznan.pl

Received: October 2020 Revised: November 2021 Accepted: November 2021

Abstract:

- We consider sequences of independent and identically absolutely continuously distributed random variables assuming that they have finite expectation and variance. We determine sharp lower and upper bounds on the expectation of the sum of n first sample maxima and n first upper record values under the condition that the value of the j-th $(1 \leq j \leq n)$ sample maximum and record value, respectively, are known and equal to a given quantile of the parent distribution. The bounds are expressed in terms of the expectation and standard deviation of the parent distribution. Analogous evaluations are presented for the sum of record values in n observations, when the j-th sample maximum is known. The theoretical results are numerically compared.

Keywords:

- *sample maximum; upper record; conditional expectation; bound.*

AMS Subject Classification:

- 60E15, 62G32.

✉ Corresponding author.

1. INTRODUCTION

Let $X_1, ..., X_n, ...$ denote i.i.d. random variables with a common absolutely continuous distribution function F and density function f, say. We assume that they have a finite second moment. Let $M_n = \max\{X_1, ..., X_n\}$, $n = 1, 2, ...$, stand for the maximum of the first n observations. For fixed $1 \le j \le n$ and $0 < q < 1$, we determine tight lower and upper bounds for the standardized versions of the conditional expectations

$$\frac{1}{\sigma}\mathbb{E}\left(\sum_{k=1}^{n} M_k - n\mu \middle| M_j = F^{-1}(q)\right) \tag{1.1}$$

over all parent distribution functions F where μ and σ denote the respective mean and standard deviation. It is clear that manipulating with location and scale of the parent distribution function F, we may obtain arbitrarily large and small values of the conditional expectation in (1.1), and a proper standardization allows to get rid of the trivial extremes. We chose the mean and standard deviation of the parent distribution as the most classic location and scale parameters, respectively. Normalization (1.1) allows us to get rid of dependence on location and scale, and its variability depends only on the shape of the parent distribution. It is also intuitively clear that the conditional expectation depends on the location of M_j in the support of X_1, and its distribution over the support. This is well expressed by the order of respective quantile.

A similar problem is solved for the upper records. We define the first record time and value as $T_1 = 1$ and $R_1 = X_1$, respectively. The further record times and values are determined recursively $T_n = \min\{k > T_{n-1} : \; X_k > M_{k-1}\}$, and $R_n = X_{T_n} = M_{T_n}$. By definition, the sequence of upper records is the maximal increasing subsequence of the non-decreasing sequence of sample maxima, arisen by crossing out all the repetitions. The second problem we cope with here is evaluating

$$\frac{1}{\sigma}\mathbb{E}\left(\sum_{k=1}^{n} R_k - n\mu \middle| R_j = F^{-1}(q)\right), \qquad 1 \le j \le n, \; 0 < q < 1. \tag{1.2}$$

For describing our last problem, we introduce the record indicators $\eta_k = 1$ if $X_k > M_{k-1}$ and $\eta_k = 0$ otherwise. Our purpose is to evaluate

$$\frac{1}{\sigma}\mathbb{E}\left(\sum_{k=1}^{n} X_k\eta_k - n\mu \middle| M_j = F^{-1}(q)\right), \qquad 1 \le j \le n, \; 0 < q < 1. \tag{1.3}$$

Expression $\sum_{k=1}^{n} X_k\eta_k$ represents the sum of all the record values observed among the first n observations. Condition $M_j = F^{-1}(q)$ means that the actual record value after j observations amounts to $F^{-1}(q)$.

Exemplary applications of our problems are connected with sponsoring and rewarding sportsmen.

Example. Some sports disciplines consists in gaining the greatest possible results. The examples are here the track and field competitions in jumping and throwing. The sportsmen receive scholarships and rewards proportional to (or linearly dependent on) their achievements. Suppose that due to an agreement with a sponsor a person receives a scholarship in the period of n months based on sports level which is measured by his/her personal best result.

(In the meantime he/she can achieve worse results, but it is known that he/she is able to attain the results on the level of his/her personal best.) Therefore his/her joint earnings in n months are proportional to $\sum_{i=1}^{n} M_i$.

In the second case, the sponsoring company signs the agreement with the organizer of a competition series that it pays honoraria for n consecutive records during the competitions of the amounts linearly dependent on the values of records. The sum of payoffs is a linear function of $\sum_{i=1}^{n} R_i$ then.

Another variant of the agreement is that the company sponsors n track and field meetings so that it pays a random number of honoraria to the people who gain new records during these events. The total amount of the rewards is proportional to $\sum_{i=1}^{n} X_i \eta_i$, where X_i is the result of the winner in the i-th competition, and η_i is the respective record indicator.

We try to evaluate the total sums of payments in these three models on the basis of knowledge of j-th value of the payment, $1 \leq j \leq n$, which are M_j, R_j and M_j again, respectively, but we do not know a substantially random mechanism generating the results. However, such generally stated problems do not have nontrivial solutions. We should know at least approximate values of the location and scale parameters. Therefore we included the mean and standard deviation in the models, which are the most popular parameters of location and scale. Also, one other factor specific to a given sport discipline should be taken into account. For instance, it is obvious that one can expect more progress in the triple jump or hammer throw in the ladies competitions rather than among the men, because the women version of these sport competitions were introduced quite recently. Mathematically, the tendency of the given discipline for gaining new records is expressed the small value of the quantile order q of the parent distribution.

We solve our three problems using a similar approach. We represent expressions (1.1)–(1.3) in integral forms, depending on indices j and n, quantile order q, and parent distribution function F. Then for fixed j, n, and q, we determine the lower and upper bounds on the integrals representing (1.1)–(1.3), and distribution functions F which attain the bounds. These distributions have atoms, and formally do not satisfy the continuity assumptions. However, if we skilfully spread out (uniformly, for simplicity) the atom masses over their small neighborhoods preserving the parent mean and variance, we may attain values of conditional expectations arbitrarily close to the respective bounds by an absolutely continuous distributions. This means that our bounds are optimal: there are sequences of continuous distributions tending weakly to a discontinuous ones which approach respective bounds arbitrarily close. For brevity of presentation, we merely present these limiting discontinuous distributions, and imprecisely write that the bounds are attained by them.

The integral bounds are calculated with use of the method proposed in Moriguti [21] who used it for evaluating the expectations of order statistics from i.i.d. samples and their differences.

Lemma 1.1. *Let H be a non-decreasing right-continuous function on an interval $[a, b]$, and continuous at a and b. Let $\overline{H}$ and $\underline{H}$ be the smallest concave majorant, and the greatest convex minorant of H, respectively. Let $\overline{h}$ and $\underline{h}$ denote the the right-hand side derivatives of $\overline{H}$ and $\underline{H}$, respectively. Then for every non-decreasing function f on $[a, b]$ we have*

$$\int_a^b f(x)\overline{h}(x)dx \leq \int_a^b f(x)H(dx) \leq \int_a^b f(x)\underline{h}(x)dx \tag{1.4}$$

under the assumption that the integrals exist and are finite. The lower (upper) bound is attained iff f is constant in every interval of the open set $\{x \in [a,b] : \overline{H}(x) > H(x)\}$ ($\{x \in [a,b] : \underline{H}(x) < H(x)\}$, respectively), and $f(x)$ is left-continuous (right-continuous, resp.) at every discontinuity point (if any) of H.

Moriguti ([21], Theorem 1) determined the upper bound in (1.4) under a more general assumption that H has a bounded variation on $[a,b]$, and is continuous at the interval ends. The lower one is easily concluded from Theorem 1 of Moriguti [21]:

$$-\int_a^b f(x)H(dx) = \int_a^b f(x)(-H)(dx) \le -\int_a^b f(x)\overline{h}(x)dx,$$

because $-\overline{h}$ is the derivative of the greatest convex minorant of $-H$.

Order statistics, especially sample extremes, and records were the objects of extensive studies. Arnold *et al.* [2], and David and Nagaraja [9] are the most popular textbooks devoted to order statistics. Comprehensive studies of records were presented in Arnold *et al.* [3] and Nevzorov [23]. Gumbel [13] and Hartley and David [14] independently derived sharp upper mean-variance bounds on the maxima of i.i.d. random variables. Analogous estimates for the record values were presented in Nagaraja [22]. These bounds were determined with use of the Schwarz inequality. Applying the same tool one can establish analogous bounds on sums of maxima and records, but the respective analytic formulae are complicated.

Predictions of order statistics and record values were analyzed by Raqab and Balakrishnan [28], Ahmadi and Balakrishnan [1], MirMostafaee and Ahmadi [20], and Volterman *et al.* [31], among others. In particular, Rychlik [29] and Klimczak [17] determined bounds on conditional expectations of future order statistics and records. Balakrishnan *et al.* [5], Asgharzadeh *et al.* [4], Khatib and Ahmadi [15], and Khatib *et al.* [16] studied reconstructions of previous failure times and records in various models. Klimczak and Rychlik [18] presented evaluations of conditional expectations of previous order statistics and records.

Conditional expectations of (1.1), (1.2), and (1.3) are studied in survival analysis, the gambling, finance, and reliability theories. A problem of prediction of the sum of minima (dual to (1.1)) was treated by Nevzorov *et al.* [24]. Problem (1.3) is a modification of a classical secretary choice problem which consist in maximizing the probability of finding the maximal record value in a finite sequence of i.i.d. observation in an on-line decision procedure (see, e.g., Gilbert and Mosteller [11] or Chow *et al.* [8]). Various generalizations of the secretary problem can be found in Freeman [10] and Samuels [30]. Recent developments in the subject are presented in Ramsey [27], Kuchta [19], Woryna [32], and Grau Ribas [12]. Sums of records in fixed numbers of trials were treated in Bel'kov and Nevzorov [6]. For a fixed parent distribution, they maximized $\mathbb{E}\left(\sum_{k=j}^n X_k\eta_k|X_1,...X_j\right)$ with respect to $j = 1,...,n-1$. Nevzorov and Tovmasyan [26] analyzed a similar problem if the number of upper records was maximized instead of the sum of their values. Bel'kov and Nevzorov [7] maximized the joint sum of upper and lower records in the analogous model. Nevzorov and Stepanov [25] maximized the expected sum of maxima by choosing an optimal starting time.

Evaluations of (1.1), (1.2) and (1.3) are presented in Sections 2, 3, and 4, respectively. Section 5 is devoted to numerical comparisons.

2. SUMS OF SAMPLE MAXIMA

Lemma 2.1. *Let $X_1, ..., X_n$ be i.i.d. with an absolutely continuous distribution function F and density f. Then $\mathbb{E}\big(\frac{1}{n}\sum_{k=1}^{n} M_k \big| M_j = x\big)$ is identical with the expectation of the distribution function*

$$F_{j,n,F}(y|x) = \begin{cases} \dfrac{1}{jn}\displaystyle\sum_{k=1}^{j-1}(j-k)\dfrac{F^k(y)}{F^k(x)}, & y < x, \\ \dfrac{j}{n} + \dfrac{1}{n}\displaystyle\sum_{k=1}^{n-j} F^k(y), & y \geq x. \end{cases}$$

$$(2.1) \qquad = \begin{cases} \dfrac{1}{jn}\dfrac{F(x)F(y)}{[F(y)-F(x)]^2}\left[j - 1 - j\dfrac{F(y)}{F(x)} + \dfrac{F^j(y)}{F^j(x)}\right], & y < x, \\ \dfrac{j}{n} + \dfrac{1}{n}\dfrac{F(y)}{1-F(y)}[1 - F^{n-j}(y)], & y \geq x. \end{cases}$$

We adhere to the convention that $\sum_{k=i}^{j} a_i = 0$ for $j < i$.

Proof: For $j < k$ we have

$$\mathbb{P}(M_k = x | M_j = x) = \mathbb{P}(X_i \leq x,\ i = j+1, ..., k) = F^{k-j}(x),$$

and for $x < y$ yields

$$\mathbb{P}(M_j \leq x, M_k \leq y) = \mathbb{P}(X_i \leq x,\ i = 1, ..., j,\ X_i \leq y,\ i = j+1, ..., k) = F^j(x)F^{k-j}(y).$$

Therefore the joint density function of M_j and M_k is

$$(2.2) \qquad f_{M_j,M_k}(x,y) = j(k-j)F^{j-1}(x)F^{k-j-1}(y)f(x)f(y), \qquad x < y.$$

Since

$$(2.3) \qquad f_{M_j}(x) = jF^{j-1}(x)f(x),$$

the conditional density of M_k under condition $M_j = x$ has the form

$$f_{M_k|M_j}(y|x) = (k-j)F^{k-j-1}(y)f(y), \qquad y > x,$$

and the respective conditional distribution function is

$$(2.4) \qquad F_{M_k|M_j}(y|x) = \begin{cases} 0, & y < x, \\ F^{k-j}(y), & y \geq x. \end{cases}$$

Take now $k < j$. Using the exchangeability argument we conclude that $\mathbb{P}(M_k = x \,|\, M_j = x) = \frac{k}{j}$ for any x. Applying (2.2) and (2.3) we also obtain

$$f_{M_k|M_j}(y|x) = \frac{k(j-k)}{j}\frac{F^{k-1}(y)}{F^k(x)}, \qquad y < x.$$

It follows that the conditional distribution function of M_k with respect to $M_j = x$ is

$$F_{M_k|M_j}(y|x) = \begin{cases} \dfrac{j-k}{j} \dfrac{F^k(y)}{F^k(x)}, & y < x, \\ 1, & y \geq x. \end{cases} \tag{2.5}$$

Obviously, the distribution of M_j given $M_j = x$ is the degenerate measure concentrated at x. Combing this fact with (2.4) and (2.5), we get

$$\sum_{k=1}^{n} F_{M_k|M_j}(y|x) = \begin{cases} \dfrac{1}{j} \displaystyle\sum_{k=1}^{j-1} (j-k) \dfrac{F^k(y)}{F^k(x)}, & y < x, \\ j + \displaystyle\sum_{k=1}^{n-j} F^k(y), & y \geq x. \end{cases}$$

Finally,

$$\mathbb{E}\left(\frac{1}{n} \sum_{k=1}^{n} M_k \,\middle|\, M_j = x \right) = \int_{\mathbb{R}} y \sum_{k=1}^{n} F_{M_k|M_j}(dy|x) = \int_{\mathbb{R}} y \, F_{j,n,F}(dy|x). \qquad \square$$

Distribution function (2.1) in the standard uniform case has the form

$$\begin{aligned} F_{j,n}(u|q) &= \begin{cases} \dfrac{1}{jn} \displaystyle\sum_{k=1}^{j-1} (j-k) \dfrac{u^k}{q^k}, & 0 < u < q, \\ \dfrac{j}{n} + \dfrac{1}{n} \displaystyle\sum_{k=1}^{n-j} u^k, & q \leq u < 1, \end{cases} \\ &= \begin{cases} \dfrac{1}{jn} \dfrac{uq}{(u-q)^2} \left[j - 1 - j\dfrac{u}{q} + \dfrac{u^j}{q^j} \right], & u < q, \\ \dfrac{j}{n} + \dfrac{1}{n} \dfrac{u}{1-u} [1 - u^{n-j}], & u \geq q, \end{cases} \qquad 0 < q < 1. \end{aligned} \tag{2.6}$$

It has the density function

$$\begin{aligned} f_{j,n}(u|q) &= \begin{cases} \dfrac{1}{jnq} \displaystyle\sum_{k=0}^{j-2} (j-k-1)(k+1) \dfrac{u^k}{q^k}, & 0 < u < q, \\ \dfrac{1}{n} \displaystyle\sum_{k=0}^{n-j-1} (k+1) u^k, & q \leq u < 1, \end{cases} \\ &= \begin{cases} \dfrac{(j+1)q}{jn(q-u)^2} \left[1 - j\dfrac{u^{j-1}}{q^{j-1}} + (j-1)\dfrac{u^j}{q^j} \right] - \dfrac{q^2}{jn(q-u)^3} \\ \times \left[2 - j(j+1)\dfrac{u^{j-1}}{q^{j-1}} + 2(j-1)(j+1)\dfrac{u^j}{q^j} - j(j-1)\dfrac{u^{j+1}}{q^{j+1}} \right], & 0 < u < q, \\ \dfrac{1}{(1-u)^2} [1 - (n+1-j)u^{n-j} + (n-j)u^{n+1-j}], & q \leq u < 1, \end{cases} \end{aligned} \tag{2.7}$$

when $0 < q < 1$, and the jump of height $\frac{j+1}{2n} + \frac{1}{n} \sum_{k=1}^{n-j} q^k = \frac{j+1}{2n} + \frac{1}{n} \frac{q}{1-q}(1 - q^{n-j})$ at q.

We also note that

$$F_{j,n}(q-|q)=\frac{j-1}{2n},\quad F_{j,n}(q|q)=\frac{j}{n}+\frac{1}{n}\sum_{k=1}^{n-j}q^k=\frac{j}{n}+\frac{1}{n}\frac{q}{1-q}(1-q^{n-j}),\tag{2.8}$$

$$f_{j,n}(0|q)=\frac{j-1}{jnq},\quad f_{j,n}(1|q)=\frac{(n-j)(n-j+1)}{2n},\tag{2.9}$$

$$f_{j,n}(q-|q)=\frac{(j-1)(j+1)}{6nq},\quad f_{j,n}(q+|q)=\frac{1}{n}\sum_{k=0}^{n-j-1}(k+1)q^k$$
$$=\frac{1-(n+1-j)q^{n-j}+(n-j)q^{n+1-j}}{n(1-q)^2}.$$

Before we formulate the main results of this section, we define some auxiliary notions. Put

$$j_*=j_*(n)=\frac{2n+1-\sqrt{8n+1}}{2},\tag{2.10}$$

$$\begin{aligned} If_{j,n}(u|q) &= \int_0^u f_{j,n}^2(v|q)dv \\ &= \frac{1}{(jnq)^2}\int_0^u\left[\sum_{k=0}^{j-2}(j-k-1)(k+1)\frac{v^k}{q^k}\right]^2 dv \\ &= \frac{1}{(jnq)^2}\sum_{r=0}^{2j-4}\frac{1}{(r+1)q^r}\left[\sum_{k=\max\{1,r-j+3\}}^{\min\{r+1,j-1\}}k(j-k)(r-k+2)(j-r+k-2)\right]u^{r+1} \end{aligned}\tag{2.11}$$

for $0<u\le q$, and

$$\begin{aligned} Jf_{j,n}(u|q) &= \int_u^1 f_{j,n}^2(v|q)dv \\ &= \frac{1}{n^2}\int_u^1\left[\sum_{k=0}^{n-j-1}(k+1)u^k\right]^2 dv \\ &= \frac{1}{n^2}\sum_{r=0}^{2(n-j-1)}\frac{1}{(r+1)}\left[\sum_{k=\max\{1,r-n+j+2\}}^{\min\{r+1,n-j\}}k(r-k+2)\right](1-u^{r+1}) \end{aligned}\tag{2.12}$$

for $q\le u<1$. We first describe the upper bounds for $2\le j\le n-1$. The extreme cases $j=1$ and $j=n$ are treated separately.

Theorem 2.1. *Let $X_1,...,X_n$ be i.i.d. with some distribution and density functions F and f, mean μ and variance σ^2. Fix $2\le j\le n-1$, and $0<q<1$.*

(i) *If $j_*\le j\le n-1$ (see (2.10)), and $q\le\frac{j-1}{jn}$, then*

$$\frac{1}{\sigma}\mathbb{E}\left(\sum_{k=1}^{n}M_k-n\mu\middle|M_j=F^{-1}(q)\right)\le 0.\tag{2.13}$$

The bound is attained in limit by sequences of continuous distributions tending to degenerate ones.

(ii) *Assume that $q>\frac{j-1}{jn}$ and either of two conditions holds. One is $j_*\le j\le n-1$, and the other is $2\le j<j_*$ with the assumption that the equation*

$$f_{j,n}(1|q)(u-1)+1=F_{j,n}(u|q)\tag{2.14}$$

has a solution in $(0,q)$.

(a) *If moreover* $\frac{j-1}{jn} < q < \frac{(j-1)(j+1)}{6n+(j-1)(j-2)}$ *then the equation*

$$1 - F_{j,n}(u|q) = (1-u)f_{j,n}(u|q) \tag{2.15}$$

has a unique solution $0 < u_* < q$, *and then*

$$\frac{1}{\sigma}\mathbb{E}\left(\sum_{k=1}^{n} M_k - n\mu \middle| M_j = F^{-1}(q)\right) \le nA_{j,n}(q), \tag{2.16}$$

where

$$A_{j,n}^2(q) = If_{j,n}(u_*|q) + f_{j,n}^2(u_*|q)(1-u_*) - 1. \tag{2.17}$$

The equality in (2.16) is attained by the parent distribution with the quantile function

$$F^{-1}(u) = \mu + \frac{\sigma}{A_{j,n}(q)}[f_{j,n}(\min\{u, u_*\}|q) - 1]. \tag{2.18}$$

(b) *However, if* $q \ge \frac{(j-1)(j+1)}{6n+(j-1)(j-2)}$, *then (2.16) holds with*

$$A_{j,n}^2(q) = If_{j,n}(q|q) + \frac{[1 - F_{j,n}(q-|q)]^2}{1-q} - 1, \tag{2.19}$$

and attainability condition

$$F^{-1}(u) = \mu + \frac{\sigma}{A_{j,n}(q)} \times \begin{cases} f_{j,n}(u|q) - 1, & u < q, \\ \dfrac{1 - F_{j,n}(q-|q)}{1-q}, & u \ge q. \end{cases} \tag{2.20}$$

(iii) *Suppose that* $2 \le j < j_*$, *and either of two assumptions holds. The first is* $q \le \frac{j-1}{jn}$. *The other admits* $q > \frac{j-1}{jn}$, *but demands that the equation*

$$f_{j,n}(0|q)u = F_{j,n}(u|q) \tag{2.21}$$

has a solution in $(q, 1)$ *then. In consequence, the equation*

$$F_{j,n}(u|q) = uf_{j,n}(u|q) \tag{2.22}$$

has a unique solution $q < u_{**} < 1$, *and (2.16) holds with*

$$A_{j,n}^2(q) = f_{j,n}^2(u_{**}|q)u_{**} + Jf_{j,n}(u_{**}|q) - 1. \tag{2.23}$$

In this case the bound in (2.16) is attained by the distribution with the quantile function

$$F^{-1}(u) = \mu + \frac{\sigma}{A_{j,n}(q)}[f_{j,n}(\max\{u, u_{**}\}|q) - 1]. \tag{2.24}$$

(**iv**) *Finally, let $2 \le j < j_*$, and $q > \frac{j-1}{jn}$, and equations (2.15) and (2.22) do not have solutions in $(0,q)$ and $(q,1)$, respectively.*

(**a**) *If moreover the equation*

$$f_{j,n}(q-|q)(u-q) + F_{j,n}(q-|q) = F_{j,n}(u|q) \tag{2.25}$$

has a solution in $(q,1)$, though, then there exist unique $0 < u_* < q < u_{**} < 1$ satisfying the equations

$$f_{j,n}(u_*|q) = f_{j,n}(u_{**}|q) = \frac{F_{j,n}(u_{**}|q) - F_{j,n}(u_*|q)}{u_{**} - u_*}, \tag{2.26}$$

and (2.16) holds with

$$A^2_{j,n}(q) = I f_{j,n}(u_*|q) + f^2_{j,n}(u_*|q)(u_{**} - u_*) + J f_{j,n}(u_{**}|q) - 1. \tag{2.27}$$

The equality in (2.16) is attained then if

$$F^{-1}(u) = \mu + \frac{\sigma}{A_{j,n}(q)} \times \begin{cases} f_{j,n}(u_*|q) - 1, & u_* \le u \le u_{**}, \\ f_{j,n}(u|q) - 1, & \text{otherwise.} \end{cases} \tag{2.28}$$

(**b**) *If (2.25) does not have a solution in $(q,1)$, then there exists a unique $q < u_{**} < 1$ such that*

$$f_{j,n}(q-|q) < \frac{F_{j,n}(u_{**}|q) - F_{j,n}(q-|q)}{u_{**} - q} = f_{j,n}(u_{**}|q), \tag{2.29}$$

and (2.16) holds with

$$A^2_{j,n}(q) = I f_{j,n}(q|q) + \frac{[F_{j,n}(u_{**}|q) - F_{j,n}(q-|q)]^2}{u_{**} - q} + J f_{j,n}(u_{**}|q) - 1, \tag{2.30}$$

and the equality in (2.10) holds for

$$F^{-1}(u) = \mu + \frac{\sigma}{A_{j,n}(q)} \times \begin{cases} \dfrac{F_{j,n}(u_{**}|q) - F_{j,n}(q-|q)}{u_{**} - q} - 1, & q \le u < u_{**}, \\ f_{j,n}(u|q) - 1, & \text{otherwise.} \end{cases} \tag{2.31}$$

Proof: By Lemma 2.1,

$$\begin{aligned} n\mathbb{E}\left(\frac{1}{n}\sum_{k=1}^{n} M_k - \mu \middle| M_j = F^{-1}(q)\right) &= n\int_{\mathbb{R}} (y-\mu) F_{j,n,F}(dy|F^{-1}(q)) \\ &= n\int_0^1 [F^{-1}(u) - \mu] F_{j,n}(du|q). \end{aligned} \tag{2.32}$$

For using Lemma 1.1, we need to determine the greatest convex minorant of (2.6). This distribution function is convex on the intervals $[0,q)$ and $[q,1]$, and has a jump up at q. We easily notice that the greatest convex minorant may have four possible shapes. It is certainly linear near q and possibly identical with $F_{j,n}(u|q)$ at the ends of $[0,1]$. However, it may happen that the linear part reaches either of the end-points of the interval, or even the line may cover the whole interval.

(i) Then problem is most simple when $f_{j,n}(1|q) \le 1 \le f_{j,n}(0|q)$, i.e. when $j \ge j_*$ and $q \le \frac{j-1}{jn}$ (cf. (2.9)). Then the straight line $\ell(u) = u$, $0 \le u \le 1$, connects the points $(0, F_{j,n}(0|q)) = (0,0)$ and $(1, F_{j,n}(1|q)) = (1,1)$, and runs beneath $F_{j,n}(u|q)$ in between. It follows that the line is the greatest convex minorant of (2.6), and its derivative amounts to constant 1. Therefore

$$\mathbb{E}\left(\sum_{k=1}^{n} M_k - n\mu \middle| M_j = F^{-1}(q)\right) \le n \int_0^1 [F^{-1}(u) - \mu] du = 0.$$

This proves inequality (2.13). In order to prove its optimality, for simplicity we consider the family of two-point distributions with the quantile functions

$$F_\varepsilon^{-1}(u) = \mu + \sigma \times \begin{cases} -\sqrt{\dfrac{1-\varepsilon}{\varepsilon}}, & u < \varepsilon, \\ \sqrt{\dfrac{\varepsilon}{1-\varepsilon}}, & u \ge \varepsilon, \end{cases} \qquad 0 < \varepsilon < 1. \tag{2.33}$$

Applying the de l'Hospital rule and boundedness of $f_{j,n}(u|q)$ near 0, we obtain

$$\begin{aligned} &\lim_{\varepsilon \to 0} \int_0^1 [F_\varepsilon^{-1}(u) - \mu] F_{j,n}(du|q) \\ &= \sigma \lim_{\varepsilon \to 0} \left[-\sqrt{\varepsilon(1-\varepsilon)} \, \frac{F_{j,n}(\varepsilon|q)}{\varepsilon} + \sqrt{\frac{\varepsilon}{1-\varepsilon}} \, [1 - F_{j,n}(\varepsilon|q)] \right] = 0. \end{aligned}$$

This argument shows that the zero bound is optimal if the greatest convex minorant is linear. We shall not repeat it in the future proofs. Note that here the same conclusion could be derived if we locate the vanishing atom on the right.

Now we observe that each of equations (2.14) and (2.21) has at most two solutions in $(0,q)$ and $(q,1)$, respectively, because their left-hand sides are linear, and the right-hand sides are strictly convex. We also note that existence of solutions to (2.14) excludes that for (2.21) and vice-versa. Assume for instance that u_0 is the solution (the single one or the smaller of two) to (2.14). It follows that

$$F_{j,n}(u|q) > f_{j,n}(1|q)(u-1) + 1, \qquad q < u < 1. \tag{2.34}$$

The straight line $f_{j,n}(0|q)u$ runs below the point $(u_0, F_{j,n}(u_0|q))$, and line $f_{j,n}(1|q)(u-1)+1$ right to u_0. By (2.34), it cannot meet $F_{j,n}(u|q)$ in $(q,1)$. When (2.21) has a solution, we argue in a similar way to exclude that of (2.14).

(ii) Let $f_{j,n}(0|q) < 1$, i.e., $q > \frac{j-1}{jn}$. Assume moreover that either $f_{j,n}(1|q) \le 1$ $(j \ge j_*)$ or $f_{j,n}(1|q) > 1$ $(j < j_*)$ holds together with existence of solution to (2.14). It follows that then the greatest convex minorant of $F_{j,n}(u|q)$ coincides first with the function itself, and then with the straight line $f_{j,n}(1|q)(u-1)+1$ (at least on $[q,1)$). The change point u_* amounts to q if

$$f_{j,n}(q-|q) \le \frac{1 - F_{j,n}(q-|q)}{1-q}. \tag{2.35}$$

(a) If $\frac{j-1}{2n} < q < \frac{(j-1)(j+1)}{6n+(j-1)(j-2)}$, we have the reversed inequality in (2.35). Function

$$\underline{F}_{j,n}(u|q) = \begin{cases} F_{j,n}(u|q), & u \le q, \\ \dfrac{1 - F_{j,n}(q-|q)}{1-q}(u-1) + 1, & u \ge q, \end{cases} \tag{2.36}$$

is not convex then. However, there exist $u_* < q$ such that the line $\frac{1-F_{j,n}(u_*|q)}{1-u_*} \times (u-1)+1$ connecting the points of the graph of $F_{j,n}(u|q)$ at u_* and 1 runs below the graph, and is tangent to it at u_*. This provides the change point of the minorant, and is certainly determined by (2.15). When $\frac{j-1}{jn} < q \leq \frac{j-1}{2n}$, we have $F_{j,n}(q-|q) \geq q$ which again implies that the change point is u_* defined in (2.15). It follows that for $\frac{j-1}{jn} < q < \frac{(j-1)(j+1)}{6n+(j-1)(j-2)}$ the derivative of the greatest convex minorant of (2.6) has the form $\underline{f}_{j,n}(u|q) = f_{j,n}(\min\{u, u_*\}|q)$. Coming back to (2.32) we obtain

$$\begin{aligned}
\mathbb{E}\left(\sum_{k=1}^{n} M_k - n\mu \middle| M_j = F^{-1}(q)\right) &\leq n \int_0^1 [F^{-1}(u) - \mu][\underline{f}_{j,n}(u|q) - 1]du \\
&\leq n\left(\int_0^1 [F^{-1}(u) - \mu]^2 du \int_0^1 [\underline{f}_{j,n}(u|q) - 1]^2 du\right)^{\frac{1}{2}} \\
&= n\sigma\left(\int_0^1 \underline{f}_{j,n}^2(u|q)du - 1\right)^{\frac{1}{2}}. \qquad (2.37)
\end{aligned}$$

The last equality follows from the fact that $\underline{f}_{j,n}(u|q)$ integrates to $\underline{F}_{j,n}(1|q) = 1$ on the interval $[0,1]$. Using (2.11) we easily check that $\int_0^1 \underline{f}_{j,n}^2(u|q)du - 1 = A_{j,n}^2(q)$ defined in (2.17). The equality in the latter inequality of (2.37) holds when

$$F^{-1}(u) - \mu = \alpha[\underline{f}_{j,n}^2(u|q) - 1], \qquad 0 < u < 1, \qquad (2.38)$$

for some $\alpha > 0$. Note that the right-hand side is right-continuous and integrates to 0, which allows to preserve the expectation condition for the left-hand side. The variance assumption implies $\alpha = \frac{\sigma}{A_{j,n}(q)}$. Observe that condition (2.38) for the equality in the latter Schwarz inequality of (2.37) preserves constancy intervals of the derivative of the greatest convex minorant which is necessary for satisfying the first equality condition of Lemma 1.1 in the first upper Moriguti inequality of (2.37). The other is satisfied as well since we defined the right continuous version of the quantile function in (2.38). It follows that (2.18) actually defines the parent distribution for which the bound (2.16) with the right-hand side defined in (2.17) is attained.

This approach is used in our further investigations. Determination of the greatest convex minorant is the crucial step in the evaluation method. The upper bound coincides with the Hilbert norm of its derivative decreased by one, and a proper linear modification of this function defines the quantile function of the distribution which satisfies the moment conditions and attains the bound. For brevity, in our further studies we stop calculations once we define the greatest convex minorant of a suitable integrand, and tacitly refer to the procedure described in the previous paragraph.

(b) Using (2.8) and (2.9) we check that (2.35) is satisfied when $q \geq \frac{(j-1)(j+1)}{6n+(j-1)(j-2)}$. Note that (2.35) implies $F_{j,n}(q-|q) < q$. Indeed, relation $F_{j,n}(q-|q) \geq q$ forces $\frac{1-F_{j,n}(q-|q)}{1-q} \leq 1$, and $f_{j,n}(q-|q) > 1$. The latter is a consequence of the fact that $F_{j,n}(u|q)$ crosses then the line $\ell(u) = u$ from bottom to top in $(0,q)$. Its derivative is necessarily greater than 1 at the crossing point, and increases later on. Also, relation $q \geq \frac{(j-1)(j+1)}{6n+(j-1)(j-2)}$ implies that (2.36) is actually a convex function, and it forms the greatest convex minorant of $F_{j,n}(u|q)$. Its right-hand side derivative is

$$\underline{f}_{j,n}(u|q) = \begin{cases} f_{j,n}(u|q), & u < q, \\ \dfrac{1 - F_{j,n}(q - |q)}{1 - q} = \dfrac{2n + 1 - j}{2n(1 - q)}, & u \geq q. \end{cases}$$

Following the arguments presented above we conclude that in this case we obtain the bound in (2.16) defined by (2.19) and its attainability condition are described in (2.20).

(iii) Under the assumptions of this point, $F_{j,n}(u|q)$ runs below the line $\ell(u) = u$ in some left neighborhood of 1. If $q \leq \frac{j-1}{jn}$, the function is located above the line for all $0 < u < q$. Therefore the greatest convex minorant has to be first linear, and then identical with $F_{j,n}(u|q)$. When $q > \frac{j-1}{jn}$ and $f_{j,n}(0|q) < 1$ in consequence, but (2.21) holds for some $q < u < 1$, then $F_{j,n}(u|q)$, $0 < u < q$, lies above the line $f_{j,n}(0|q)u$, but this is not true for some $u \in (q, 1)$. Again we deduce that the minorant is first linear and eventually strictly convex. The change point belongs to $(q, 1)$, and q is impossible. This point is determined by (2.22) which means that the linear part of the greatest convex minorant is tangent to $F_{j,n}(u|q)$ at the change point u_{**}. The derivative of the convex minorant is then $\underline{f}_{j,n}(u|q) = f_{j,n}(\max\{u, u_{**}\}|q)$. Proceeding as in the previous part on the proof we determine the mean-variance bound for the conditional expectation and the condition of its attainability.

(iv) The assumptions mean that $F_{j,n}(u|q)$ goes below $\ell(u) = u$ in some neighborhoods of 0 and 1. Moreover, the lines tangent to $F_{j,n}(u|q)$ at 0 and 1 run below the graph of the function. This implies that the greatest convex minorant of $F_{j,n}(u|q)$ cannot be linear at vicinities of the end-points. So the linear part may appear only in the central part, and it contains q.

(a) If (2.25) has a solution then the derivative of the greatest convex minorant of $F_{j,n}(u|q)$ can be written as

$$\underline{f}_{j,n}(u|q) = \begin{cases} f_{j,n}(u_*|q), & u_* \leq u \leq u_{**}, \\ f_{j,n}(u|q), & \text{elsewhere}, \end{cases}$$

where $0 < u_* < q < u_{**} < 1$ are determined from the tangency conditions (2.26). In the standard way we establish the bound in (2.16) with the right-hand side described in (2.27), and the attainability condition (2.28).

(b) The lack of solution to (2.25) implies that all the lines $u \mapsto f_{j,n}(v|q)(u - v) + F_{j,n}(v|q)$, tangent to $F_{j,n}(u|q)$ at $v < q$ run below $F_{j,n}(u|q)$ for $u < v$. The only candidate for the change point of the minorant from $F_{j,n}(u|q)$ into a line is q. Consider the functions $\ell_\alpha(u) = \alpha(u - q) + F_{j,n}(q - |q)$, and increase the slopes α starting from $f_{j,n}(q - |q)$ until we touch any point of $F_{j,n}(u|q)$ for $u \geq q$. Obviously q cannot be the first meeting point, because the line connecting $F_{j,n}(q - |q)$ and $F_{j,n}(q|q)$ is vertical. It cannot be 1, either, because then $\alpha = \frac{1 - F_{j,n}(q-|q)}{1-q} \geq f_{j,n}(1|q)$ which contradicts the assumption that (2.14) does not have a solution in $(0,q)$. Consequently, our assumptions imply

$$\underline{f}_{j,n}(u|q) = \begin{cases} \dfrac{F_{j,n}(u_{**}|q) - F_{j,n}(q-|q)}{u_{**}-q}, & q \le u \le u_{**}, \\ f_{j,n}(u|q), & \text{elsewhere,} \end{cases}$$

where u_{**} is determined by solving (2.29). This allows us to conclude (2.16) with (2.30) and (2.31) assuring the equality in (2.16). □

We separately consider the extreme cases $j=1$ and $j=n$, for which the distribution function (2.6) does not have any mass on the left and right, respectively, of q. This allows us to simplify the arguments of the above proof in order to get desired conclusions. The details of the reasoning are left to the reader.

Theorem 2.2. *Let the assumptions of Theorem 2.1 hold.*

(i) *Let $M_1 = F^{-1}(q)$ for some $0<q<1$.*

(a) *If $q < \frac{n-3}{n-1}$, then there exists $q < u_{**} < 1$ solving the equation*

$$\begin{aligned} \frac{u}{1-u}(1-u^{n-1}) &= nF_{1,n}(u|q) = n(u-q)f_{1,n}(u|q) \\ &= \frac{u-q}{(1-u)^2}[1-nu^{n-1}+(n-1)u^n], \end{aligned}$$

and then we have (2.16) with $j=1$ and

$$A^2_{1,n}(q) = f^2_{1,n}(u_{**}|q)(u_{**}-q) + Jf_{1,n}(u_{**}|q) - 1$$

(see (2.12)). The equality is attained if

$$F^{-1}(u) = \mu + \frac{\sigma}{A_{1,n}(q)} \times \begin{cases} -1, & u<q, \\ f_{1,n}(u_{**}|q)-1, & q \le u \le u_{**}, \\ f_{1,n}(u|q)-1, & u \ge u_{**}. \end{cases}$$

(b) *If $q \ge \frac{n-3}{n-1}$, then $A_{1,n}(q) = \sqrt{\frac{q}{1-q}}$, and the bound is attained by the two-point parent distribution on $\mu - \sigma\sqrt{\frac{1-q}{q}}$ and $\mu + \sigma\sqrt{\frac{q}{1-q}}$ with respective probabilities q and $1-q$.*

(ii) *Under the condition $M_n = F^{-1}(q)$, there are three possible cases.*

(a) *When $q \le \frac{n-1}{n^2}$ (cf. Theorem 2.1*(i)*), the optimal upper bound on the standardized expectation of the first n maxima is equal to 0.*

(b) *If $\frac{n-1}{n^2} < q < \frac{n-1}{n+2}$ then the statements of Theorem 2.1*(iia) *hold with j replaced by n.*

(c) *If $q \ge \frac{n-1}{n+2}$ then the statements of Theorem 2.1*(iib) *hold with j replaced by n.*

In the following theorem we describe the lower bounds on the conditional expectations of sample maxima for all $1 \le j \le n$.

Theorem 2.3. *Assume the conditions of Theorem 2.1.*

(i) *Let $0 < q_* \leq 1$ be the unique solution to*

$$j + \frac{q}{1-q}(1 - q^{n-j}) = nq. \tag{2.39}$$

If either $j_* \leq j \leq n$ (comp. (2.10)) or $1 \leq j < j_*$ and $q < q_*$ defined above, then

$$\frac{1}{\sigma}\mathbb{E}\left(\sum_{k=1}^{n} M_k - n\mu \middle| M_j = F^{-1}(q)\right) \geq -\frac{j + \frac{q}{1-q}(1-q^{n-j}) - nq}{\sqrt{q(1-q)}}. \tag{2.40}$$

The lower bound in (2.40) is attained by the two-point distribution supported on $\mu - \sigma\sqrt{\frac{1-q}{q}}$ and $\mu + \sigma\sqrt{\frac{q}{1-q}}$ with respective probabilities q and $1-q$.

(ii) *If $1 \leq j < j_*$ and $q \geq q_*$ then the optimal bound is*

$$\frac{1}{\sigma}\mathbb{E}\left(\sum_{k=1}^{n} M_k - n\mu \middle| M_j = F^{-1}(q)\right) \geq 0.$$

Note that $q_* = 1$ for $j = n$, and so (2.40) holds for all $0 < q < 1$ with the right hand-side simplified to $-n\sqrt{\frac{1-q}{q}}$.

Proof: We rewrite representation (2.32), and apply the lower estimate of Lemma 1.1. In consequence of the shape of the distribution function (2.6), the only possible shapes of its smallest concave majorant is the linear function $\ell(u) = u$ when $F_{j,n}(q|q) \leq q$ or a broken line with the break point q otherwise. Inequality $F_{j,n}(q|q) \leq q$ is equivalent to

$$j + \sum_{k=1}^{n-j} q^k - nq \leq 0. \tag{2.41}$$

The left-hand side function is strictly convex, positive at 0, and vanishing at 1. Its derivative at 1 amounts to $\frac{(n-j)(n-j+1)}{2} - n = \frac{1}{2}\left[j^2 - (2n+1)j + n(n-1)\right]$ which is non-positive for $j \geq j_*$ (see (2.10)), and positive otherwise. Accordingly, inequality (2.41) is false for all $0 < q < 1$ when $j \geq j_*$. If $j < j_*$, (2.41) holds only for sufficiently large q. Precisely, this is true for $q \geq q_*$ defined in (2.39).

(i) Assume so that either $j_* \leq j \leq n$ or $1 \leq j < j_*$ and $q < q_*$. Then the smallest concave majorant has the form

$$\overline{F}_{j,n}(u|q) = \begin{cases} \dfrac{F_{j,n}(q|q)}{q}\, u = \dfrac{j + \sum_{k=1}^{n-j} q^k}{nq}\, u, & u \leq q, \\[2ex] \dfrac{1 - F_{j,n}(q|q)}{1-q}(u-1) + 1 = \dfrac{n - j - \sum_{k=1}^{n-j} q^k}{n(1-q)}(u-1) + 1, & u \geq q. \end{cases}$$

We use

$$\overline{f}_{j,n}(u|q) - 1 = \begin{cases} \dfrac{j + \sum_{k=2}^{n-j} q^k - (n-1)q}{nq}, & u \leq q, \\[2ex] \dfrac{(n-1)q - j - \sum_{k=2}^{n-j} q^k}{n(1-q)} & u > q, \end{cases} \tag{2.42}$$

for establishing the following lower mean-variance bound

$$n\mathbb{E}\left(\frac{1}{n}\sum_{k=1}^{n} M_k \middle| M_j = F^{-1}(q)\right) \geq n\int_0^1 [F^{-1}(u)-\mu][\overline{f}_{j,n}(u|q)-1]du$$

$$\geq -n\left[\int_0^1 [F^{-1}(u)-\mu]^2 du\right]^{1/2}\left[\int_0^1 [\overline{f}_{j,n}(u|q)-1]^2 du\right]^{1/2} = -n\sigma a_{j,n}(q), \tag{2.43}$$

where

$$a_{j,n}^2(q) = \int_0^1 [\overline{f}_{j,n}(u|q)-1]^2 du = \frac{1}{n^2 q(1-q)}\left(j + \sum_{k=1}^{n-j} q^k - nq\right)^2.$$

Note that under the assumption the expression in the parentheses is positive. Now set

$$F^{-1}(u-) - \mu = -\frac{\sigma}{a_{j,n}(q)}\left[\overline{f}_{j,n}(u|q)-1\right] \tag{2.44}$$

which asserts the equalities in both the inequalities of (2.43). Note that the right-hand side of (2.44) is non-decreasing and left-continuous. Moreover, its integral over $[0,1]$ is equal to 0, and the integral of its square amounts to 1. This implies that the left-hand side determines the standardized lower quantile function of a distribution with mean μ and variance σ^2. Plugging (2.42) into (2.44) we obtain

$$F^{-1}(u-) = \mu + \sigma \times \begin{cases} -\sqrt{\dfrac{1-q}{q}}, & u \leq q, \\ \sqrt{\dfrac{q}{1-q}}, & u > q, \end{cases}$$

which describes the two-point distribution defined in the first part of Theorem 2.3.

(ii) Otherwise, if $1 \leq j < j_*$ and $q \geq q_*$, the derivative of the smallest concave majorant $\overline{F}_{j,n}(u|q) = u$, $0 \leq u \leq 1$, of $F_{j,n}(u|q)$ is equal to 1. Consequently,

$$n\mathbb{E}\left(\frac{1}{n}\sum_{k=1}^{n} M_k \middle| M_j = F^{-1}(q)\right) \geq n\int_0^1 [F^{-1}(u)-\mu]du = 0. \qquad \square$$

Seemingly, the conditions for attaining the upper bounds in Theorem 2.2(ib) and Theorem 2.3(i) pretend to be identical. There are subtle differences between them, though. In the first case, the strictly increasing quantile functions $F_n^{-1}(u)$ should tend to the right-continuous version of the two-valued extreme quantile function. In the other one, they should tend to the left-continuous lower quantile function. We omit presenting elementary constructions of such sequences.

3. SUMS OF UPPER RECORDS

Lemma 3.1. *Let $X_1, ..., X_n, ...$ be i.i.d. with an absolutely continuous distribution function F and density f, and let $R_1, ..., R_n$ denote the values of the first upper records in the sequence. Then $\mathbb{E}\big(\frac{1}{n}\sum_{k=1}^{n} R_k \big| R_j = x\big)$ for some $1 \le j \le n$ is identical with the expectation of the distribution function*

$$G_{j,n,F}(y|x) = \begin{cases} \dfrac{j-1}{n}\,\dfrac{-\ln[1-F(y)]}{-\ln[1-F(x)]}, & y < x, \\[2ex] 1 - \dfrac{1}{n}\,\dfrac{1-F(y)}{1-F(x)} \displaystyle\sum_{k=0}^{n-j-1} \frac{n-j-k}{k!}\left[-\ln\frac{1-F(y)}{1-F(x)}\right]^k, & y \ge x. \end{cases} \tag{3.1}$$

Proof: The density function of the single record value R_j, and the joint density of a pair (R_j, R_k), $j < k$, have the forms

$$f_{R_j}(x) = \frac{\{-\ln[1-F(x)]\}^{j-1}}{(j-1)!}\, f(x), \tag{3.2}$$

$$f_{R_j,R_k}(x,y) = \frac{\{-\ln[1-F(x)]\}^{j-1}}{(j-1)!}\, \frac{\left[-\ln\frac{1-F(y)}{1-F(x)}\right]^{k-j-1}}{(k-j-1)!}\, \frac{f(x)f(y)}{1-F(x)}, \quad x < y, \tag{3.3}$$

respectively (see, e.g., Arnold *et al.* [3], p. 11). It follows that for $j < k$

$$f_{R_k|R_j}(y|x) = \frac{\left[-\ln\frac{1-F(y)}{1-F(x)}\right]^{k-j-1}}{(k-j-1)!}\, \frac{f(y)}{1-F(x)}, \qquad y > x,$$

is the conditional density function of R_k under the condition that $R_j = x$. We see that the conditional distribution is identical with the unconditional distribution of the $(k-j)$-th record value from a sequence with the left-truncated parent distribution function $\frac{1-F(y)}{1-F(x)}$, $y > x$. The respective distribution function is

$$F_{R_k|R_j}(y|x) = 1 - \frac{1-F(y)}{1-F(x)} \sum_{i=0}^{k-j-1} \frac{[-\ln[\frac{1-F(y)}{1-F(x)}]^i}{i!}, \qquad x < y.$$

We also note that

$$\sum_{k=j+1}^{n} F_{R_k|R_j}(y|x) = n - j - \frac{1-F(y)}{1-F(x)} \sum_{k=0}^{n-j-1} \frac{n-j-k}{k!}\left[-\ln\frac{1-F(y)}{1-F(x)}\right]^k, \quad x < y. \tag{3.4}$$

Referring again to (3.2) and (3.3), we obtain

$$f_{R_k|R_j}(y|x) = \frac{(j-1)!}{(k-1)!(j-k-1)}\left[\frac{-\ln[1-F(y)]}{-\ln[1-F(x)]}\right]^{k-1} \times \left[1 - \frac{-\ln[1-F(y)]}{-\ln[1-F(x)]}\right]^{j-k-1} \frac{-f(y)}{[1-F(y)]\ln[1-F(x)]}$$

for $y < x$ and $k < j$. This coincides with the density function of the k-th order statistic from an i.i.d. sample of size $j-1$ from the right-truncated distribution function $\frac{-\ln[1-F(y)]}{-\ln[1-F(x)]}$, $y < x$.

Obviously, the sum of ordered variables is identical with that of the original unordered ones. Therefore

$$\mathbb{E}\left(\sum_{k=1}^{j-1} R_k \middle| R_j = x\right) = (j-1)\int_{-\infty}^{x} y\,\frac{-\ln[1-F(dy)]}{-\ln[1-F(x)]} \tag{3.5}$$

Combining (3.4), (3.5) with the trivial fact $\mathbb{E}(R_j|R_j = x) = x = \int_{\mathbb{R}} y\mathbf{1}_{[x,+\infty)}(dy)$, we conclude

$$\begin{aligned}
&\mathbb{E}\left(\frac{1}{n}\sum_{k=1}^{n} R_k \middle| R_j = x\right)\\
&= \frac{j-1}{n}\int_{-\infty}^{x} y\,\frac{-\ln[1-F(dy)]}{-\ln[1-F(x)]} + \frac{1}{n}\int_{\mathbb{R}} y\mathbf{1}_{[x,+\infty)}(dy) + \frac{1}{n}\int_{x}^{+\infty} y \sum_{k=j+1}^{n} F_{R_k|R_j}(dy|x)\\
&= \int_{\mathbb{R}} yF_{j,n,F}(dy|x),
\end{aligned}$$

which proves our statement. □

In the standard uniform case (3.1) takes on the form

$$G_{j,n}(u|q) = \begin{cases} \dfrac{j-1}{n}\dfrac{-\ln(1-u)}{-\ln(1-q)}, & 0 < u < q < 1,\\[2ex] 1 - \dfrac{1}{n}\dfrac{1-u}{1-q}\displaystyle\sum_{k=0}^{n-j-1}\frac{n-j-k}{k!}\left(-\ln\frac{1-u}{1-q}\right)^k, & 0 < q \le u < 1. \end{cases} \tag{3.6}$$

It has the density function

$$g_{j,n}(u|q) = \begin{cases} \dfrac{j-1}{n}\dfrac{1}{-\ln(1-q)}\dfrac{1}{1-u}, & 0 < u < q < 1,\\[2ex] \dfrac{1}{n(1-q)}\displaystyle\sum_{k=0}^{n-j-1}\frac{1}{k!}\left(-\ln\frac{1-u}{1-q}\right)^k, & 0 < q \le u < 1, \end{cases}$$

and an atom with the weight $\frac{1}{n}$ at q. In particular we have

$$G_{j,n}(q-|q) = \frac{j-1}{n}, \quad G_{j,n}(q|q) = \frac{j}{n},$$

$$g_{j,n}(0|q) = \frac{j-1}{-n\ln(1-q)}, \quad g_{j,n}(1-|q) = \begin{cases} +\infty, & j < n-1,\\ \dfrac{1}{n(1-q)}, & j = n-1,\\ 0, & j = n, \end{cases} \tag{3.7}$$

$$g_{j,n}(q-|q) = \frac{j-1}{-n(1-q)\ln(1-q)}, \quad g_{j,n}(q+|q) = \frac{1}{n(1-q)}.$$

We also define

$$Ig_{j,n}(u|q) = \int_0^u g_{j,n}^2(v|q)dv = \left[\frac{j-1}{-n\ln(1-q)}\right]^2\int_0^u \frac{1}{(1-v)^2}dv = \left[\frac{j-1}{-n\ln(1-q)}\right]^2\frac{u}{1-u}, \tag{3.8}$$

for $0 < u \leq q$, and

$$Jg_{j,n}(u|q) = \int_u^1 g_{j,n}^2(v|q)dv = \frac{1}{n^2(1-q)^2}\int_u^1 \left[\sum_{k=0}^{n-j-1}\frac{1}{k!}\left(-\ln\frac{1-v}{1-q}\right)^k\right]^2 dv$$

$$= \frac{1}{n^2(1-q)^2}\sum_{r=0}^{2(n-j-1)}\left[\sum_{k=\max\{0,r-n+j+1\}}^{\min\{r,n-j-1\}}\binom{r}{k}\right]\frac{1}{r!}\int_u^1\left(-\ln\frac{1-u}{1-q}\right)^r dv$$

$$= \frac{1-u}{n^2(1-q)^3}\sum_{r=0}^{2(n-j-1)}\left[\sum_{k=\max\{0,r-n+j+1\}}^{\min\{r,n-j-1\}}\binom{r}{k}\right]\left[\sum_{k=0}^{r}\frac{1}{k!}\left(-\ln\frac{1-u}{1-q}\right)^k\right] \tag{3.9}$$

for $q \leq u < 1$. Note that for $r = 0, \dots, n-j-1$, the sum of binomial coefficients in the first square brackets of (3.9) amounts to 2^r.

Theorem 3.1. *Let $X_1, \dots, X_n, \dots$ be i.i.d. with some distribution and density functions F and f, mean μ and variance σ^2. Fix $2 \leq j \leq n-2$, and $0 < q < 1$.*

(i) *Suppose that either of two assumptions holds. One is $q \leq 1 - \exp(-\frac{j-1}{n})$. The other is $q > 1 - \exp(-\frac{j-1}{n})$ and the equation*

$$g_{j,n}(0|q)u = G_{j,n}(u|q) \tag{3.10}$$

has a solution in $(q, 1)$. Then the equation

$$G_{j,n}(u|q) = ug_{j,n}(u|q)$$

has a unique solution $q < u_{**} < 1$, and we have

$$\frac{1}{\sigma}\mathbb{E}\left(\sum_{k=1}^{n} R_k - n\mu \,\middle|\, R_j = F^{-1}(q)\right) \leq nB_{j,n}(q) \tag{3.11}$$

with

$$B_{j,n}^2(q) = g_{j,n}^2(u_{**}|q)u_{**} + Jg_{j,n}(u_{**}|q) - 1. \tag{3.12}$$

In this case the bound in (3.11) is attained by the distribution with the quantile function

$$F^{-1}(u) = \mu + \frac{\sigma}{B_{j,n}(q)}[g_{j,n}(\max\{u, u_{**}\}|q) - 1]. \tag{3.13}$$

(ii) *Assume that $q > 1 - \exp(-\frac{j-1}{n})$ and the equation (3.10) does not have a solution in $(q, 1)$.*

(a) *If moreover there exists in $(q, 1)$ a solution to the equation*

$$g_{j,n}(q-|q)(u-q) + G_{j,n}(q-|q) = G_{j,n}(u|q) \tag{3.14}$$

then there is a unique pair $0 < u_* < q < u_{**} < 1$ satisfying the equations

$$g_{j,n}(u_*|q) = g_{j,n}(u_{**}|q) = \frac{G_{j,n}(u_{**}|q) - G_{j,n}(u_*|q)}{u_{**} - u_*}, \tag{3.15}$$

and (3.11) holds with

$$B_{j,n}^2(q) = Ig_{j,n}(u_*|q) + g_{j,n}^2(u_*|q)(u_{**} - u_*) + Jg_{j,n}(u_{**}|q) - 1.$$

The equality in (3.11) is attained then if

$$F^{-1}(u) = \mu + \frac{\sigma}{B_{j,n}(q)} \times \begin{cases} g_{j,n}(u_*|q) - 1, & u_* \le u \le u_{**}, \\ g_{j,n}(u|q) - 1, & \text{otherwise.} \end{cases}$$

(b) *If (3.14) does not have a solution in* $(q,1)$*, then there is a unique* $q < u_{**} < 1$ *such that*

$$g_{j,n}(q-|q) < \frac{G_{j,n}(u_{**}|q) - G_{j,n}(q-|q)}{u_{**} - q} = g_{j,n}(u_{**}|q), \tag{3.16}$$

and (3.11) holds with

$$B_{j,n}^2(q) = Ig_{j,n}(q|q) + \frac{[G_{j,n}(u_{**}|q) - G_{j,n}(q-|q)]^2}{u_{**} - q} + Jg_{j,n}(u_{**}|q) - 1, \tag{3.17}$$

whereas the equality in (3.11) is attained for

$$F^{-1}(u) = \mu + \frac{\sigma}{B_{j,n}(q)} \times \begin{cases} \dfrac{G_{j,n}(u_{**}|q) - G_{j,n}(q-|q)}{u_{**} - q} - 1, & q \le u < u_{**}, \\ g_{j,n}(u|q) - 1, & \text{otherwise.} \end{cases} \tag{3.18}$$

The idea of proof of Theorem 3.1 as well as the following results is similar to that of Theorem 2.1. Therefore we sketch only the main points focusing merely on the differences.

Proof: Since $g_{j,n}(1-|q) = +\infty$ for $j \le n-2$, we can exclude the possibilities that the greatest convex minorant of (3.6) is linear at the neighborhood of 1.

(i) Suppose that either $g_{j,n}(0|q) \ge 1$ (i.e., $q \le 1 - \exp\left(-\frac{j-1}{n}\right)$, comp. (3.7)) or $g_{j,n}(0|q) < 1$ but the line tangent to $G_{j,n}(u|q)$ at 0 meets $G_{j,n}(u|q)$ somewhere in $(q,1)$. This implies that there is a line located below it in the positive half-axis which runs through $(0,0)$ and is tangent to $G_{j,n}(u|q)$ at some u_{**} in $(q,1)$. Its segment joining $(0,0)$ with $(u_{**}, G_{j,n}(u_{**}|q))$ extended by $G_{j,n}(u|q)$ itself on the right composes the greatest convex minorant of $G_{j,n}(u|q)$. This observation allows us to determine the bound (3.12) in (3.11), and the condition of its attainability (3.13) (comp. (2.16), (2.23) and (2.24)).

(ii) If $q > 1 - \exp\left(-\frac{j-1}{n}\right)$ and $g_{j,n}(0|q)u$ runs below $G_{j,n}(u|q)$ on $(q,1)$, then the convex minorant should coincide with the original function on a right neighborhood of 0 as well as that of 1, and be linear in between. There are two possible subcases.

  (a) If the line tangent to $G_{j,n}(u|q)$ at $q-$ runs above $G_{j,n}(u|q)$ on the whole $(q,1)$ (i.e., (3.14) does hold), the point where the minorant transforms into a line has to be less than q. The linear part should be tangent to the graph of $G_{j,n}(u|q)$ at the both its ends. Therefore the end points $u_* < q < u_{**}$ are determined by equations (3.15). Once we fix the convex minorant we are in a position to calculate the sharp upper bound on the conditional expectation, and the parent distribution which attains it.

(b) In the opposite case, the linear part starts at q, and its right end u_{**} is determined by the tangency condition (3.16). This provides the bound defined in (3.17) and its attainability condition described in (3.18). □

Below we present without a proof the upper bounds for conditional expectations of the sum of first upper records under condition $R_j = F^{-1}(q)$ for remaining $j = 1$, $n-1$, and n.

Theorem 3.2. *Suppose that $X_1, ..., X_n, ...$ satisfy the assumptions of Theorem 3.1.*

(i) *There exists $q < u_{**} < 1$ solving the equation*

$$g_{1,n}(u|q)(u-q) = G_{1,n}(u|q)$$

such that

$$\frac{1}{\sigma}\mathbb{E}\left(\sum_{k=1}^{n} R_k - n\mu \middle| R_1 = F^{-1}(q)\right) \leq nB_{1,n}(q) \tag{3.19}$$

where

$$B^2_{1,n}(q) = g^2_{1,n}(u_{**}|q)(u_{**}-q) + Jg_{1,n}(u_{**}|q) - 1.$$

The equality in (3.19) holds for the distribution function F satisfying

$$F^{-1}(u) = \mu + \frac{\sigma}{B_{1,n}(q)} \times \begin{cases} -1, & u < q, \\ g_{1,n}(u_{**}|q) - 1, & q \leq u < u_{**}, \\ g_{j,n}(u|q) - 1, & u \geq u_{**}. \end{cases}$$

(ii) *For $j = n-1$ and $j = n$ we have the following.*

(a) *If $q \leq 1 - \exp\left(-\frac{j-1}{n}\right)$, then the optimal bound is*

$$\frac{1}{\sigma}\mathbb{E}\left(\sum_{k=1}^{n} R_k - n\mu \middle| R_j = F^{-1}(q)\right) \leq 0.$$

(b) *If $1 - \exp\left(-\frac{j-1}{n}\right) < q < 1 - \exp\left(-\frac{j-1}{n+1-j}\right)$, then (3.11) holds with*

$$B^2_{j,n}(q) = Ig_{j,n}(u_*|q) + g^2_{j,n}(u_*|q)(1-u_*) - 1$$

for $0 < u_* < q$ satisfying the equation

$$g_{j,n}(u|q)(1-u) = 1 - G_{j,n}(u|q).$$

The condition for getting the equality in (3.11) is

$$F^{-1}(u) - \mu + \frac{\sigma}{B_{j,n}(q)}[g_{j,n}(\min\{u, u_*\}|q)].$$

(c) *Finally, if $q \geq 1 - \exp\left(-\frac{j-1}{n+1-j}\right)$, then (3.11) holds with*

$$B^2_{j,n}(q) = Ig_{j,n}(q|q) + \frac{[1 - G_{j,n}(q-|q)]^2}{1-q} - 1.$$

The equality in (3.11) holds then if

$$F^{-1}(u) = \mu + \frac{\sigma}{B_{j,n}(q)} \times \begin{cases} g_{j,n}(u|q) - 1, & u < q, \\ \dfrac{1 - G_{j,n}(q-|q)}{1-q} - 1, & u \geq q. \end{cases}$$

The lower bounds on the conditional expectations of the sums of consecutive record values are presented below. The proof mimics the proof of Theorem 2.3, and it is omitted.

Theorem 3.3. *Assume the conditions of Theorem 3.1. For any $1 \le j \le n$, we have two cases.*

(i) *If $q < \frac{j}{n}$, then*

$$\frac{1}{\sigma}\mathbb{E}\left(\sum_{k=1}^{n} R_k - n\mu \,\middle|\, R_j = F^{-1}(q)\right) \ge -\frac{j-nq}{\sqrt{q(1-q)}}. \tag{3.20}$$

The equality in (3.20) is attained by the two-point distribution supported on the points $\mu - \sigma\sqrt{\frac{1-q}{q}}$ and $\mu + \sigma\sqrt{\frac{q}{1-q}}$ with probabilities q and $1-q$, respectively.

(ii) *If $q \ge \frac{j}{n}$ then*

$$\frac{1}{\sigma}\mathbb{E}\left(\sum_{k=1}^{n} R_k - n\mu \,\middle|\, R_j = F^{-1}(q)\right) \ge 0.$$

is optimal.

A more precise description of the attainability conditions in case (i) is presented in the comment below Theorem 2.3. Note that for $j = n$ only this case occurs.

4. SUMS OF RECORDS IN FINITE SEQUENCES

The problem of maximizing the conditional expectation of $\sum_{k=1}^{n} X_k\eta_k$ makes sense if X_k are positive.

Lemma 4.1. *Let $X_1, ..., X_n$ be positive i.i.d. with an absolutely continuous distribution function F, and finite expectation. Then $\mathbb{E}\big(\frac{1}{n}\sum_{k=1}^{n} X_k\eta_k \big| M_j = x\big)$ for some $1 \le j \le n$ is identical with the expectation of the distribution function*

$$H_{j,n,F}(y|x) = \begin{cases} 0, & y < 0, \\ \dfrac{n-1}{n} - \displaystyle\sum_{k=1}^{n-j} \frac{1-F^k(x)}{nk} - \left(\sum_{k=2}^{j} \frac{1}{nk}\right)\left[1 - \frac{-\ln[1-F(y)]}{-\ln[1-F(x)]}\right], & 0 \le y < x, \\ 1 - \displaystyle\sum_{k=1}^{n-j} \frac{1-F^k(y)}{nk}, & y \ge x. \end{cases} \tag{4.1}$$

Proof: Since we may observe at most j records among $X_1, ..., X_j$, we have

$$\begin{aligned} \mathbb{E}\left(\sum_{k=1}^{j} X_k\eta_k \,\middle|\, M_j = x\right) &= \sum_{k=1}^{j} \mathbb{E}\left(\sum_{i=1}^{k} R_i \,\middle|\, M_j = R_k = x, \sum_{i=1}^{j} \eta_i = k\right) \mathbb{P}\left(\sum_{i=1}^{j} \eta_i = k\right) \\ &= \sum_{k=1}^{j} \mathbb{E}\left(\sum_{i=1}^{k-1} R_i + x \,\middle|\, R_k = x, \sum_{i=1}^{j} \eta_i = k\right) \mathbb{P}\left(\sum_{i=1}^{j} \eta_i = k\right) \\ &= x + \sum_{k=2}^{j} \mathbb{E}\left(\sum_{i=1}^{k-1} R_i \,\middle|\, R_k = x, \sum_{i=1}^{j} \eta_i = k\right) \mathbb{P}\left(\sum_{i=1}^{j} \eta_i = k\right). \end{aligned}$$

Let $Y_1(x), ..., Y_{j-1}(x)$ denote i.i.d. random variables with a common distribution function $F_x(y) = \frac{-\ln[1-F(y)]}{-\ln[1-F(x)]}$, $y < x$. By arguments of the proof of Lemma 3.1 we notice that

$$\begin{aligned}
\mathbb{E}\left(\sum_{k=1}^{j} X_k\eta_k \middle| M_j = x\right) &= x + \sum_{k=2}^{j} \mathbb{E}\left(\sum_{i=1}^{k-1} Y_i(x)\right) \mathbb{P}\left(\sum_{i=1}^{j} \eta_i = k\right) \\
&= x + \mathbb{E}Y_1(x) \sum_{k=2}^{j} (k-1)\, \mathbb{P}\left(\sum_{i=2}^{j} \eta_i = k-1\right) \\
&= x + \mathbb{E}Y_1(x)\, \mathbb{E}\left(\sum_{i=2}^{j} \eta_i\right) \\
&= x + \mathbb{E}Y_1(x) \sum_{k=2}^{j} \frac{1}{k},
\end{aligned}$$

because $\mathbb{P}(\eta_k = 1) = \frac{1}{k} = 1 - P(\eta_k = 0)$. Note that under the condition $M_j = x$, just one among $X_k\eta_k$, $k = 1, ..., j$, has value x for sure. The other ones take on either some values in $(0, x)$ as the order statistics from the sample with the distribution function F_x, or they amount to 0. The first ones appear with probabilities $\frac{1}{k}$, and the others with probabilities $1 - \frac{1}{k}$, $k = 2, ..., j$. Therefore we can write

$$\begin{aligned}
\mathbb{E}\left(\sum_{k=1}^{j} X_k\eta_k \middle| M_j = x\right) &= \int_{\mathbb{R}} y\, \mathbf{1}_{[x,\infty)}(dy) + \sum_{k=2}^{j} \frac{1}{k} \int_0^x y\, \frac{-\ln[1-F(dy)]}{-\ln[1-F(x)]} \\
&\quad + \left(j - \sum_{k=1}^{j} \frac{1}{k}\right) \int_{\mathbb{R}} y\, \mathbf{1}_{[0,\infty)}(dy). \qquad (4.2)
\end{aligned}$$

For $k > j$, the conditional distribution of $X_k\eta_k$ has an atom at 0 with probability

$$\begin{aligned}
\mathbb{P}(X_k\eta_k = 0 | M_j = x) &= \mathbb{P}(X_k \le x) + \mathbb{P}(x < X_k \le \max\{X_{j+1}, ..., X_{k-1}\}) \\
&= F(x) + \int_x^\infty \mathbb{P}(\max\{X_{j+1}, ..., X_{k-1}\} \ge y) f(y)dy \\
&= F(x) + \int_x^\infty [1 - F^{k-j-1}(y)] f(y)dy = 1 - \frac{1 - F^{k-j}(x)}{k-j}.
\end{aligned}$$

Moreover, for $y > x$ we have

$$\begin{aligned}
\mathbb{P}(X_k\eta_k > y | M_j = x) &= \mathbb{P}(X_k > \max\{y, X_{j+1}, ..., X_{k-1}\}) \\
&= \int_y^\infty \mathbb{P}(\max\{X_{j+1}, ..., X_{k-1}\} < t) f(t)dt \\
&= \int_y^\infty F^{k-j-1}(t) f(t)dt = \frac{1 - F^{k-j}(y)}{k-j}.
\end{aligned}$$

Summing up, we obtain

$$\mathbb{P}(X_k\eta_k \le y | M_j = x) = \begin{cases} 0, & y < 0, \\ 1 - \dfrac{1 - F^{k-j}(x)}{k-j}, & 0 \le y \le x, \\ 1 - \dfrac{1 - F^{k-j}(y)}{k-j}, & y \ge x. \end{cases} \qquad (4.3)$$

Combining (4.2) and (4.3) yields

$$\mathbb{E}\left(\frac{1}{n}\sum_{k=1}^{n} X_k\eta_k \middle| M_j = x\right) = \frac{1}{n}\left[j - \sum_{k=1}^{j}\frac{1}{k} + n - j - \sum_{k=j+1}^{n}\frac{1-F^{k-j}(x)}{k-j}\right]\int_{\mathbb{R}} y\,\mathbf{1}_{[0,\infty)}(dy)$$
$$+\sum_{k=2}^{j}\frac{1}{nk}\int_0^x y\,\frac{-\ln[1-F(dy)]}{-\ln[1-F(x)]} + \frac{1}{n}\int_{\mathbb{R}} y\,\mathbf{1}_{[x,\infty)}(dy)$$
$$(4.4)\qquad +\frac{1}{n}\int_x^\infty y\left[n-j+\sum_{k=j+1}^{n}\frac{1-F^{k-j}(dy)}{k-j}\right] = \int_{\mathbb{R}} yH_{j,n,F}(dy|x).$$

This completes the proof. □

If $X_1, ..., X_n$ are standard uniform, (4.1) simplifies to

$$(4.5)\qquad H_{j,n}(u|q) = \begin{cases} \dfrac{n-1}{n} - \displaystyle\sum_{k=1}^{n-j}\frac{1-q^k}{nk} - \left(\sum_{k=2}^{j}\frac{1}{nk}\right)\left[1 - \frac{-\ln(1-u)}{-\ln(1-q)}\right], & 0 \le u < q, \\ 1 - \displaystyle\sum_{k=1}^{n-j}\frac{1-u^k}{nk}, & q \le u \le 1. \end{cases}$$

It has two atoms at 0 and q with respective probabilities $1 - \sum_{k=1}^{j}\frac{1}{nk} - \sum_{k=1}^{n-j}\frac{1-q^k}{nk}$ and $\frac{1}{n}$, and the density function

$$(4.6)\qquad h_{j,n}(u|q) = \begin{cases} \left(\displaystyle\sum_{k=2}^{j}\frac{1}{nk}\right)\dfrac{1}{-(1-u)\ln(1-q)}, & 0 < u < q, \\ \dfrac{1}{n}\displaystyle\sum_{k=0}^{n-j-1} u^k = \frac{1-u^{n-j}}{n(1-u)}, & q < u < 1. \end{cases}$$

Below we use the following values

$$(4.7)\qquad H_{j,n}(q-|q) = \frac{n-1}{n} - \sum_{k=1}^{n-j}\frac{1-q^k}{nk}, \qquad h_{j,n}(q-|q) = \frac{\sum_{k=2}^{j}\frac{1}{k}}{-n\ln(1-q)(1-q)},$$

and the the following function (comp. (3.8))

$$(4.8)\qquad Ih_{j,n}(u|q) = \int_0^u h_{j,n}^2(v|q)dv = \left[\frac{\sum_{k=2}^{j}\frac{1}{k}}{-n\ln(1-q)}\right]^2\int_0^u \frac{1}{(1-v)^2}dv = \left[\frac{\sum_{k=2}^{j}\frac{1}{k}}{-n\ln(1-q)}\right]^2\frac{u}{1-u}.$$

Theorem 4.1. *Let $X_1, ..., X_n$ be positive i.i.d. with an absolutely continuous distribution function F, and finite variance σ^2. Let μ stand for the respective expectation.*

(i) *Let $0 < q_* < 1$ be the unique solution to the equation*

$$(4.9)\qquad 1 + \sum_{k=1}^{n-1}\frac{1-q^k}{k} - n(1-q) = 0.$$

(a) *If $q \le q_*$, then the bound*

$$\frac{1}{\sigma}\mathbb{E}\left(\sum_{k=1}^{n} X_k\eta_k - n\mu \middle| M_1 = F^{-1}(q)\right) \le 0 \tag{4.10}$$

is sharp and attained in limit by the degenerate distribution.

(b) *If $q > q_*$, then*

$$\frac{1}{\sigma}\mathbb{E}\left(\sum_{k=1}^{n} X_k\eta_k - n\mu \middle| M_1 = F^{-1}(q)\right) \le nC_{j,n}(q) = \frac{1+\sum_{k=1}^{n-1}\frac{1-q^k}{k} - n(1-q)}{\sqrt{q(1-q)}}, \tag{4.11}$$

and the equality is attained by the parent distribution assigning the masses q and $1-q$ to the points 0 and $\frac{\sigma}{\sqrt{q(1-q)}}$, respectively.

(ii) *Assume $2 \le j \le n$.*

(a) *If*

$$H_{j,n}(u|q) \ge u, \tag{4.12}$$

(comp. (4.5)) for all $0 < u < q$, then the optimal inequality is

$$\frac{1}{\sigma}\mathbb{E}\left(\sum_{k=1}^{n} X_k\eta_k - n\mu \middle| M_j = F^{-1}(q)\right) \le 0.$$

Otherwise we have three possibilities.

(b) *If*

$$h_{j,n}(q-|q) \le \frac{H_{j,n}(q-|q)}{q} < 1, \tag{4.13}$$

then (4.11) holds with M_1 and $\sum_{k=1}^{n-1}\frac{1-q^k}{k}$ replaced by M_j and $\sum_{k=1}^{n-j}\frac{1-q^k}{k}$, respectively, and identical conditions of attainability.

(c) *If*

$$1 > \frac{H_{j,n}(q-|q)}{q} < h_{j,n}(q-|q) \le \frac{1-H_{j,n}(q-|q)}{1-q}, \tag{4.14}$$

then there exists a unique $0 < u_* < q$ solving the equation

$$H_{j,n}(u|q) = uh_{j,n}(u|q)$$

(see (4.5) and (4.6)), and then

$$\frac{1}{\sigma}\mathbb{E}\left(\sum_{k=1}^{n} X_k\eta_k - n\mu \middle| M_j = F^{-1}(q)\right) \le nC_{j,n}(q), \tag{4.15}$$

where

$$C_{j,n}^2(q) = h_{j,n}^2(u_*)u_* + Ih_{j,n}(q|q) - Ih_{j,n}(u_*|q) + \frac{[1-H_{j,n}(q-|q)]^2}{1-q} - 1.$$

(see also (4.8)). The equality in (4.15) holds for F with the quantile function

$$F^{-1}(u) = \frac{\sigma}{C_{j,n}(q)} \times \begin{cases} 0, & 0 < u < u_*, \\ h_{j,n}(u|q) - h_{j,n}(u_*|q), & u_* \le u < q, \\ \dfrac{1-H_{j,n}(q-|q)}{1-q} - h_{j,n}(u_*|q), & q \le u < 1. \end{cases}$$

(**d**) *Finally, if*

$$h_{j,n}(q-|q) > \min\left\{\frac{H_{j,n}(q-|q)}{q}, \frac{1-H_{j,n}(q-|q)}{1-q}\right\}$$

then there also exists $u_* < u_{**} < q$ *satisfying the equation*

$$(1-q)h_{j,n}(u|q) = 1 - H_{j,n}(u|q),$$

and then (4.15) holds with

$$C_{j,n}^2(q) = h_{j,n}^2(u_*)u_* + Ih_{j,n}(u_{**}|q) - Ih_{j,n}(u_*|q) + h_{j,n}^2(u_{**})(1-u_{**}) - 1,$$

and the equality condition

$$F^{-1}(u) = \frac{\sigma}{C_{j,n}(q)} \times \begin{cases} 0, & 0 < u < u_*, \\ h_{j,n}(u|q) - h_{j,n}(u_*|q), & u_* \leq u < u_{**}, \\ h_{j,n}(u_{**}|q) - h_{j,n}(u_*|q), & u_{**} \leq u < 1. \end{cases}$$

Proof: We first notice that in contrast to the maximization problems studied in Sections 2 and 3, one treated here is not location-scale invariant. Indeed, if we translate the parent distribution by $c > 0$ to the right, we obtain

$$\begin{aligned} &\frac{1}{\sigma}\mathbb{E}\left(\sum_{k=1}^{n}(X_k+c)\eta_k - n(\mu+c)\middle| M_j = x+c\right) \\ &= \frac{1}{\sigma}\mathbb{E}\left(\sum_{k=1}^{n}X_k\eta_k - n\mu\middle| M_j = x\right) - c\left(n - \sum_{k=1}^{j}\frac{1}{k} - \sum_{k=1}^{n-j}\frac{1-F^k(x)}{k}\right) \\ &< \frac{1}{\sigma}\mathbb{E}\left(\sum_{k=1}^{n}X_k\eta_k - n\mu\middle| M_j = x\right) \end{aligned}$$

(see (4.1)). Accordingly, it suffices to restrict our investigations to the distributions whose supports start from 0. Alternatively, we can consider the problem modification where the lack of record gives the gain equal to the minimal value of the distribution support, and then remove the solutions for which $F^{-1}(0) \neq 0$.

Distribution functions (4.5) contain atoms at their left-end points of the supports, and hence they do not satisfy the assumptions of Lemma 1.1. We show that we get the sharp right-hand inequality in (1.4) if we replace the greatest convex minorant of $H_{j,n}(u|q)$ by the greatest convex minorant $\underline{H}_{j,n,0}(u|q)$ of $H_{j,n}(u|q)$ and the point $(0,0)$. Take $\varepsilon > 0$ sufficiently small so that $\underline{H}_{j,n,0}(u|q)$ is also the greatest convex minorant of $H_{j,n,\varepsilon}(u|q) = \min\{H_{j,n}(u|q), \frac{u}{\varepsilon}\} \leq H_{j,n}(u|q)$. For every non-deceasing function f yields

$$\int_0^1 f(u)H_{j,n}(du|q) \leq \int_0^1 f(u)H_{j,n,\varepsilon}(du|q) \leq \int_0^1 f(u)\underline{h}_{j,n,0}(u|q)\,du,$$

where $\underline{h}_{j,n,0}(u|q)$ denotes the right derivative of $\underline{H}_{j,n,0}(u|q)$. Let f_0 satisfy the equality conditions in the latter inequality: f_0 is constant on each interval of $\{H_{j,n,\varepsilon}(u|q) < \underline{H}_{j,n,0}(u|q)\}$ and right-continuous. In particular, it is constant on $\{H_{j,n,\varepsilon}(u|q) < H_{j,n}(u|q)\}$, and 0 can be attached to this interval by right-continuity of f_0. For brevity, denote the extended interval by $[0,\delta)$. Therefore

$$\int_0^\delta f_0(u)H_{j,n}(du|q) = \int_0^\delta f_0(u)H_{j,n,\varepsilon}(du|q) = f_0(0)H_{j,n,\varepsilon}(\delta-) = f_0(0)H_{j,n}(\delta-).$$

The respective integrals over $[\delta,1)$ are identical, because $H_{j,n}(u|q)$ and $H_{j,n,\varepsilon}(u|q)$ are identical there. Consequently,

$$\int_0^1 f_0(u)H_{j,n}(du|q) = \int_0^1 f_0(u)H_{j,n,\varepsilon}(du|q) = \int_0^1 f_0(u)\underline{h}_{j,n,0}(u|q)\,du,$$

which proves sharpness of the upper bound.

It follows that for proving our bounds, we need to determine the greatest convex minorants of the functions $H_{j,n,0}(u|q)$, which amount to 0 at $u=0$, and coincide with $H_{j,n}(u|q)$ otherwise. Note that each $H_{j,n}(u|q)$ is convex non-decreasing in $(q,1)$, and its derivative satisfies $h_{j,n}(1-|q) = 1-\frac{j}{n} < 1$. So this part of the function runs above the line $\ell(u)=u$, and does not affects the convex minorant.

(i) Function $H_{1,n}(u|q)$ is constant on the interval $(0,q)$. Therefore the greatest convex minorant of $H_{1,n,0}(u|q)$ is either the straight line $\underline{H}_{1,n,0}(u|q)=u$, $0<u<1$, when $H_{1,n}(q-|q)\ge q$, or the broken line

$$\underline{H}_{1,n,0}(u|q) = \begin{cases} \dfrac{H_{1,n}(q-|q)}{q}\,u, & u<q, \\ \dfrac{1-H_{1,n}(q-|q)}{1-q}(u-1)+1, & u\ge q, \end{cases} \tag{4.16}$$

otherwise. Function

$$H_{1,n}(q-|q) - q = \frac{n-1}{n} - \sum_{k=1}^{n-1}\frac{1-q^k}{nk} - q, \qquad 0<q<1,$$

amounts to $\frac{n-1}{n} - \sum_{k=1}^{n-1}\frac{1}{nk} > 0$ at 0, and to $-\frac{1}{n} < 0$ at 1. Moreover, its derivative $\frac{1}{n}\sum_{k=0}^{n-2} q^k - 1$ is negative for all $0<q<1$. Therefore $\underline{H}_{1,n.0}(u|q) = u$ for $q\le q_*$ defined in (4.9), and has the form (4.16) for $q>q_*$.
Repeating the reasoning of the previous proofs we determine the sharp bounds (4.10) and (4.11). In the modified location-scale invariant problem, the former is attained by (2.33) with $\varepsilon\to 0$. In order to obey the restriction $F_\varepsilon^{-1}(0)=0$ we put $\mu = \sigma\sqrt{\frac{1-\varepsilon}{\varepsilon}}$. In the latter, the modified problem has solution (2.33) with ε replaced by q. Again, the support requirement narrows the attainability condition to the last statement of Theorem 4.1(i).

(ii) Relation (4.12) implies that $H_{j,n,0}(u|q)\ge u = \underline{H}_{j,n,0}(u|q)$, $0<u<1$. It follows that zero provides the optimal bound for (1.3). Otherwise we obtain non-trivial evaluations. Under condition (4.13) the line $\frac{H_{j,n}(q-|q)}{q}u$ runs beneath function $H_{j,n,0}(u|q)$ on $(0,q)$, and connects its end-points. Another linear function $\frac{1-H_{j,n}(q-|q)}{1-q}(u-1)+1$ minorizes $H_{j,n,0}(u|q)$ in $[q,1]$. Gluing together the lines we obtain the greatest convex minorant of $H_{j,n,0}(u|q)$ (note that the inequalities $\frac{H_{j,n}(q-|q)}{q} < 1 < \frac{1-H_{j,n}(q-|q)}{1-q}$ guarantee convexity and compare with (4.16)).

Mimicking the arguments of the previous proofs we calculate the upper bounds and determine the location-scale family of two-point distributions attaining the bounds in the location-scale invariant problem. Under the support restriction, we distinguish the scale family of distributions with the left support end-point equal to 0. If $\frac{H_{j,n}(q-|q)}{q} < h_{j,n}(q-|q)$ (see (4.7) and (4.14)), then the right part of the line $\frac{H_{j,n}(q-|q)}{q}u$, $0<u<q$, lies above $H_{j,n}(u|q)$, and cannot constitute a part of the minorant. It should be replaced by a line with a smaller slope $h_{j,n}(u_*|q) = \frac{H_{j,n}(u_*|q)}{u_*}$, tangent to $H_{j,n}(u|q)$ at some $0<u_*<q$, and $H_{j,n}(u|q)$ on the right which ultimately transforms into a line. If moreover $h_{j,n}(q-|q) \le \frac{1-H_{j,n}(q-|q)}{1-q}$, then $\underline{H}_{j,n,0}(u|q) = H_{j,n}(u|q)$ for all $u_* \le u < q$. The last part of the minorant is the line connecting $(q, H_{j,n}(q-|q))$ with $(1,1)$. Otherwise $\underline{H}_{j,n,0}(u|q)$ should transform into a line at some $u_* < u_{**} < q$ determined by the tangency condition $h_{j,n}(u_{**}|q)(1-u_{u**}) = 1 - H_{j,n}(u_{**}|q)$. Note that in the last case it is admitted that $H_{j,n}(q-|q) \ge q$ but necessarily $H_{j,n}(u_{**}|q) < u_{**}$.

Once we determine the greatest convex minorants, we further proceed in a standard way. The bound amounts to n multiplied by the square root of the integral of the squared derivative of the minorant decreased by 1. The standardized quantile function of the distribution attaining the bound is proportional to the greatest convex minorant derivative decreased by 1. The last step of the proof consists in removing the distributions whose left-end support points differ from 0. Detailed calculations are left to the reader. □

Establishing lower bounds for (1.3) does not make sense, because when we consider random variables $X_1, ..., X_n$ taking on very large values, and we may get $X_k\eta_k = 0$ as the results of not reaching records in some trials, would make (1.3) negative and arbitrarily small. We illustrate the phenomenon in the following example.

Example. Suppose that X_k, $k = 1, ..., n$, are uniformly distributed on the interval $[m, m+1]$. They have the distribution function $F(x) = x - m$, $m < x < m+1$, and quantile function $F^{-1}(q) = m + q$, $0 < q < 1$. Applying (4.4) we calculate

$$\begin{aligned}
\mathbb{E}\left(\sum_{k=1}^{n} X_k\eta_k \middle| M_j = m+q\right) &= \frac{\sum_{k=2}^{j}\frac{1}{k}}{-\ln(1-q)}\int_m^{m+q}\frac{y\,dy}{1-y+m} + q + m + \int_{m+q}^{m+1} y \sum_{k=0}^{n-j-1}(y-m)^k\,dy \\
&= \frac{\sum_{k=2}^{j}\frac{1}{k}}{-\ln(1-q)}[-(m+1)\ln(1-q) - q] + q + m + \sum_{k=2}^{n-j+1}\frac{1-q^k}{k} + m\sum_{k=1}^{n-j}\frac{1-q^k}{k} \\
&= m\left[1 + \sum_{k=2}^{j}\frac{1}{k} + \sum_{k=1}^{n-j}\frac{1-q^k}{k}\right] + q + \sum_{k=2}^{j}\left[1 - \frac{q}{-\ln(1-q)}\right] + \sum_{k=2}^{n-j+1}\frac{1-q^k}{k}.
\end{aligned}$$

Since $\mathbb{E}X_1 = m + \frac{1}{2}$ and $\operatorname{Var} X_1 = \frac{1}{12}$, we have

$$\begin{aligned}
\frac{1}{\sigma}\mathbb{E}\left(\sum_{k=1}^{n} X_k\eta_k - n\mu \middle| M_j = m+q\right) &= 12m\left[1 + \sum_{k=2}^{j}\frac{1}{k} + \sum_{k=1}^{n-j}\frac{1-q^k}{k} - n\right] \\
&\quad + 12\left[q + \sum_{k=2}^{j}\left[1 - \frac{q}{-\ln(1-q)}\right] + \sum_{k=2}^{n-j+1}\frac{1-q^k}{k} - \frac{n}{2}\right].
\end{aligned} \tag{4.17}$$

Putting $m = 0$, we obtain the standardized conditional expectation for the standard uniform variables. However, when m increases to $+\infty$, then (4.17) tends to $-\infty$, because the factor at m is strictly negative.

We would avoid obtaining trivial lower bounds in (1.3) if we replaced $X_k\eta_k = 0$ by a quantity connected with the distribution of random variables, e.g. by the mean or a quantile of F of a positive order.

5. NUMERICAL EVALUATIONS

Here we present numerical upper bounds on the conditional expectations of sums of sample maxima and records described analytically in Sections 2–4 for $n = 10$, $j = 1, 5, 8$ and 10, and quantile orders $q = 0.1..., (0.1), ..., 0.9$. We also include two extreme cases $q = 0.05$ and 0.99. For comparison, we present numerical results for $n = 20$ and $j = 10$ as well. We do not evaluate numerically respective lower bounds because they have simple analytic forms.

The bounds strongly depend on the number n of summands. Therefore instead of bounds $nA_{j,n}(q)$, $nB_{j,n}(q)$, and $nC_{j,n}(q)$ on the expectations of the total sums, we present in Tables 1–5 the average bounds $A_{j,n}(q)$, $B_{j,n}(q)$, and $C_{j,n}(q)$ determined per each particular summand. Each numerical bound is accompanied by the reference to a particular part of the theorem which provides the tools for calculating it. This allows the reader to realize the shape of the parent distribution which attains the corresponding bound. For instance, the average upper bounds $A_{5,10}(q)$ are determined with use of Theorem 2.1. For $q = 0.05$ the quantile function is first a constant (which generates a jump of height $u_{**} > q$), and than is a curve linearly transforming $f_{5,10}(u|0.05)$ (see (2.7)). For $q = 0.1, ..., 0.4$ the extreme quantile function is first increasing, then constant, and again increasing. When $q = 0.1, 0.2, 0.3$ the transition from the curve to the horizontal line occurs at q (see Theorem 2.1(iva)), but for $q = 0.4$ it happens at some $u_* < q$ (see Theorem 2.1(ivb)). For $q \geq 0.5$ the conditions of Theorem 2.1(iib) hold which implies that the distribution functions attaining the respective bounds are continuous on some intervals, and have jumps of size $1 - q$ at their right-end points.

All the average bounds presented in Tables 1–5 are increasing with respect to q. This is easily justifiable: the greater is the extreme j-th variable, the greater is the expectation of the sum of n analogous observations. When we fix n and q, we observe that the bounds decrease when j increases. It has a clear explanation as well. E.g., when we assume that $M_{j_1} = x$ we may suspect that $\sum_{k=1}^n M_k$ is greater than in the case $M_{j_2} = x$ for some $j_2 > j_1$, because in the latter case the maximum equal to x appears later than in the former one. We note that $C_{j,n}(q) = 0$ except for about 10% upper quantile orders q. Trivial zero bounds $A_{j,n}(q)$ and $B_{j,n}(q)$ for the sums of maxima and records, respectively, appear only for relatively small q and large j.

By definition $\sum_{k=1}^n X_k\eta_k \leq \sum_{k=1}^n M_k \leq \sum_{k=1}^n R_k$ for any random sequence $X_1, ..., X_n, ...$. The corresponding relations for the bounds $C_{j,n}(q) < A_{j,n}(q) < B_{j,n}(q)$ are preserved and their values are significantly different when j is small with respect to n. When j is equal or close to n, the latter inequality is reversed, though. For $j = n$ it is justified by the following arguments.

Table 1: Average upper bounds for $n = 10$, and $j = 1$.

q	T2	$A_{j,n}(q)$	T5	$B_{j,n}(q)$	T7	$C_{j,n}(q)$
0.05	(ia)	0.99901	(i)	23.19053	(ia)	0
0.1	(ia)	1.00446	(i)	24.48122	(ia)	0
0.2	(ia)	1.01964	(i)	27.54620	(ia)	0
0.3	(ia)	1.04398	(i)	31.48623	(ia)	0
0.4	(ia)	1.08488	(i)	36.73885	(ia)	0
0.5	(ia)	1.15691	(i)	44.09161	(ia)	0
0.6	(ia)	1.28895	(i)	55.11962	(ia)	0
0.7	(ia)	1.53655	(i)	73.49811	(ia)	0
0.8	(ib)	2.00000	(i)	110.25284	(ia)	0
0.9	(ib)	3.00000	(i)	220.51248	(ib)	0.24836
0.99	(ib)	9.94987	(i)	2205.14721	(ib)	0.99321

Table 2: Average upper bounds for $n = 10$, and $j = 5$.

q	T1	$A_{j,n}(q)$	T4	$B_{j,n}(q)$	T7	$C_{j,n}(q)$
0.05	(iii)	0.11265	(i)	1.53182	(iia)	0
0.1	(iva)	0.11312	(i)	1.62854	(iia)	0
0.2	(iva)	0.16859	(i)	1.85718	(iia)	0
0.3	(iva)	0.28233	(i)	2.14943	(iia)	0
0.4	(ivb)	0.43379	(i)	2.53691	(iia)	0
0.5	(iib)	0.61390	(iia)	3.09867	(iia)	0
0.6	(iib)	0.82506	(iia)	3.92617	(iia)	0
0.7	(iib)	1.09660	(iia)	5.29056	(iia)	0
0.8	(iib)	1.50351	(iib)	7.86104	(iia)	0
0.9	(iib)	2.33534	(iib)	15.77310	(iib)	0.15107
0.99	(iib)	7.94034	(iib)	157.76816	(iic)	0.94803

Table 3: Average upper bounds for $n = 20$, and $j = 10$.

q	T1	$A_{j,n}(q)$	T4	$B_{j,n}(q)$	T7	$C_{j,n}(q)$
0.05	(iii)	0.38885	(i)	22.61002	(iia)	0
0.1	(iva)	0.39153	(i)	23.86739	(iia)	0
0.2	(iva)	0.41974	(i)	26.85346	(iia)	0
0.3	(iva)	0.47203	(i)	30.69239	(iia)	0
0.4	(iva)	0.54734	(i)	35.81061	(iia)	0
0.5	(iva)	0.65396	(i)	42.97568	(iia)	0
0.6	(ivb)	0.81552	(iia)	53.72374	(iia)	0
0.7	(ivb)	1.06348	(iia)	71.63557	(iia)	0
0.8	(iib)	1.45459	(iia)	107.45727	(iia)	0
0.9	(iib)	2.25974	(iia)	214.91881	(iia)	0
0.99	(iib)	7.69114	(iib)	2149.144709	(iic)	0.43245

There are no future maxima and records after the j-th one. Conditionally on $R_n = x$, the previous record values $R_1, ..., R_{n-1}$ are distributed as ordered i.i.d. random variables from the right-truncated at x parent distribution (cf. Lemma 3.1). The distributions of M_k, $k = 1, ..., n-1$, under the condition $M_n = x$ are the mixtures of maxima from k independent observations from the right-truncated baseline distribution and an atom at x (see Lemma 2.1). This implies $\mathbb{E}\big(\sum_{k=1}^n R_k \big| R_n = x\big) < \mathbb{E}\big(\sum_{k=1}^n M_k \big| M_n = x\big)$ for any parent distribution.

Numerical calculations show that the reversed inequality $B_{j,n}(q) < A_{j,n}(q)$ holds for $j = n-1$ and all q as well. Then the distribution $\mathcal{L}(R_n|R_{n-1} = x)$ is just the left-truncated parent distribution at x, and this does not affect much the whole sum. Table 4 shows that for $n = 10$, $j = n-2 = 8$ the reversed inequalities $B_{8,10}(q) < A_{8,10}(q)$ are satisfied merely for some central q.

Table 4: Average upper bounds for $n = 10$, and $j = 8$.

q	T1	$A_{j,n}(q)$	T4	$B_{j,n}(q)$	T7	$C_{j,n}(q)$
0.05	(i)	0	(i)	0.01450	(iia)	0
0.1	(iia)	0.00644	(i)	0.01874	(iia)	0
0.2	(iia)	0.10956	(i)	0.03141	(iia)	0
0.3	(iia)	0.21969	(i)	0.05288	(iia)	0
0.4	(iia)	0.33303	(i)	0.08963	(iia)	0
0.5	(iia)	0.45780	(i)	0.15357	(iia)	0
0.6	(iia)	0.60936	(iia)	0.59290	(iia)	0
0.7	(iib)	0.82368	(iia)	0.96188	(iia)	0
0.8	(iib)	1.16140	(iia)	1.56008	(iia)	0
0.9	(iib)	1.85340	(iia)	3.22165	(iib)	0.06500
0.99	(iib)	6.43747	(iib)	16.06139	(iic)	0.92619

Table 5: Average upper bounds for $n = 10$, and $j = 10$.

q	T2	$A_{j,n}(q)$	T5	$B_{j,n}(q)$	T7	$C_{j,n}(q)$
0.05	(iia)	0	(iia)	0	(iia)	0
0.1	(iib)	0.00448	(iia)	0	(iia)	0
0.2	(iib)	0.10267	(iia)	0	(iia)	0
0.3	(iib)	0.20796	(iia)	0	(iia)	0
0.4	(iib)	0.31432	(iia)	0	(iia)	0
0.5	(iib)	0.42761	(iia)	0	(iia)	0
0.6	(iib)	0.55702	(iib)	0.00138	(iia)	0
0.7	(iib)	0.72122	(iib)	0.08846	(iia)	0
0.8	(iic)	0.98150	(iib)	0.25087	(iia)	0
0.9	(iic)	1.55748	(iib)	0.54681	(iia)	0
0.99	(iic)	5.44190	(iib)	1.91034	(iic)	0.91287

We finally focus on Tables 2 and 3 which contain results for $(j,n) = (5,10)$ and $(10,20)$. One could expect that the average bounds are similar if the proportion j/n is preserved. This actually happens in the case of sums of maxima. We see that $A_{5,10}(q) < A_{10,20}(q)$ for small q, and the opposite holds for large q. In the former case, when $M_j = F^{-1}(q)$ is relatively small, it is a great chance that the total sum of maxima shall increase more when we observe 10 future i.i.d. observations rather than 5. This chance decreases when M_j is close to the right end-point of the support. Much the same observation concerns $C_{j,n}(q)$, but the difference is not visible for small q, for which $C_{5,10}(q) = C_{10,20}(q) = 0$. The average bounds $B_{j,n}(q)$ for the sums of record values behave quite differently: $B_{5,10}(q)$ are much less than $B_{10,20}(q)$, and the latter are rather close to $B_{1,10}(q)$ (see Table 1). This shows that the average bounds $B_{j,n}(q)$ depend rather on the differences $n-j$, i.e., on the number of future records.

ACKNOWLEDGMENTS

The research of the second author has been partially supported by PUT under grant 0211/SBAD/0121. The authors thank the associate editor for valuable comments which allowed them to improve the presentation of the paper.

REFERENCES

[1] Ahmadi, J. and Balakrishnan, N. (2010). Prediction of order statistics and record values from two independent sequences, *Statistics*, **44**, 417–430.

[2] Arnold, B.C.; Balakrishnan, N. and Nagaraja, H.N. (1992). *A First Course in Order Statistics*, Wiley, New York.

[3] Arnold, B.C.; Balakrishnan, N. and Nagaraja, H.N. (1998). *Records*, Wiley, New York.

[4] Asgharzadeh, A.; Ahmadi, J.; Ganji, Z.M. and Valiollahi, R. (2012). Reconstruction of the past failure times for the proportional reversed hazard rate model, *J. Stat. Comput. Simul.*, **82**, 475–489.

[5] Balakrishnan, N.; Doostparast, M. and Ahmadi, J. (2009). Reconstruction of past records, *Metrika*, **70**, 89–109.

[6] Bel'kov, I.V. and Nevzorov, V.B. (2018). On a problem on the optimal choice of record values, *Vestnik St. Petersburg Univ. Math.*, **51**, 107–113.

[7] Bel'kov, I.V. and Nevzorov, V.B. (2020). On a problem of the optimal choice of record values, *J. Math. Sci. (N.Y.)*, **244**, 718–722.

[8] Chow, Y.S.; Robbins, H. and Siegmund, D. (1971). *Great Expectations: The Theory of Optimal Stopping*, Houghton Mifflin Co., Boston, Mass.

[9] David, H.A. and Nagaraja, H.N. (2003). *Order Statistics*, 3rd ed., Wiley, Hoboken, NJ.

[10] Freeman, P.R. (1983). The secretary problem and its extensions: a review, *Int. Stat. Rev.*, **51**, 189–206.

[11] Gilbert, J.P. and Mosteller, F. (1966). Recognizing the maximum of a sequence, *J. Am. Stat. Assoc.*, **61**, 35–73.

[12] Grau Ribas, J.M. (2019). A new look at the returning secretary problem, *J. Comb. Optim.*, **37**, 1216–1236.

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

[14] Hartley, H.O. and David, H.A. (1954). Universal bounds for mean range and extreme observation, *Ann. Math. Statist.*, **25**, 85–99.

[15] Khatib, B. and Ahmadi, J. (2014). Best linear unbiased and invariant reconstructors for the past records, *Bull. Malays. Math. Sci. Soc.*, **37**(2), 1017–1028.

[16] Khatib, B.; Ahmadi, J. and Razmkhah, M. (2014). Reconstruction of the past lower record values in a proportional reversed hazard rate model, *Statistics*, **48**, 421–435.

[17] Klimczak, M. (2006). Prediction of k-th records, *Statist. Probab. Lett.*, **76**, 117–127.

[18] Klimczak, M. and Rychlik, T. (2005). Reconstruction of previous failure times and records, *Metrika*, **61**, 277–290.

[19] Kuchta, M. (2017). Iterated full information secretary problem, *Math. Methods Oper. Res.*, **86**, 277–292.

[20] MirMostafaee, S.M.T.K. and Ahmadi, J. (2011). Point prediction of future order statistics from an exponential distribution, *Statist. Probab. Lett.*, **81**, 360–370.

[21] Moriguti, S. (1953). A modification of Schwarz's inequality with applications to distributions, *Ann. Math. Statist.*, **24**, 107–113.

[22] Nagaraja, H.N. (1978). On the expected values of record values, *Austral. J. Statist.*, **20**, 176–182.

[23] Nevzorov, V.B. (2001). *Records: Mathematical Theory*, Translations of Mathematical Monographs, **194**, AMS, Providence, RI.

[24] Nevzorov, V.B.; Savinova, V. and Stepanov, A. (2022). Linear prediction of sequential minima, *Comm. Statist. Theory Methods*, **51**, 5446–5454.

[25] Nevzorov, V.B. and Stepanov, A. (2021). On maximum of expected sums of sequential minima, *Comm. Statist. Theory Methods*, **50**, 1362–1369.

[26] Nevzorov, V.B. and Tovmasyan, S.A. (2014). On the maximal value of the expectation of record numbers, *Vestnik St. Petersburg Univ. Math.*, **47**(2), 64–67.

[27] Ramsey, D.M. (2016). A secretary problem with missing observations, *Math. Appl. (Warsaw)*, **44**, 149–165.

[28] Raqab, M.Z. and Balakrishnan, N. (2008). Prediction intervals for future records, *Statist. Probab. Lett.*, **78**, 1955–1963.

[29] Rychlik, T. (2002). Predictions of increments of order and record statistics in nonparametric families of distributions, *J. Statist. Theory Appl.*, **1**, 43–56.

[30] Samuels, S.M. (1991). *Secretary problems*. In "Handbook of Sequential Analysis" (B.K. Ghosh and P.K. Sen, Eds.), Marcel Dekker, New York, pp. 381–405.

[31] Volterman, W.; Davies, K.F.; Balakrishnan, N. and Ahmadi, J. (2014). Nonparametric prediction of future order statistics, *J. Stat. Comput. Simul.*, **84**, 683–695.

[32] Woryna, A. (2017). The solution of a generalized secretary problem via analytic expressions, *J. Comb. Optim.*, **33**, 1469–1491.