Available at WWW.MATHEMATICSWEB.ORG
POWERED BY SCIENCE DIRECT

Journal of Statistical Planning and Inference 113 (2003) 569–588

journal of statistical planning and inference

www.elsevier.com/locate/jspi

Bounds on expectation of order statistics from a finite population

N. Balakrishnan[a,*], C. Charalambides[b], N. Papadatos[b,1]

[a]*Department of Mathematics and Statistics, McMaster University, 1280 Main Street West, Hamilton, Ont., Canada L8S 4K1*

[b]*Department of Mathematics, University of Athens, Panepistemiopolis, 157 84 Athens, Greece*

Received 18 January 2001; accepted 10 November 2001

Abstract

Consider a simple random sample $X_1, X_2, \ldots, X_n$, taken without replacement from a finite ordered population $\Pi = \{x_1 \leqslant x_2 \leqslant \cdots \leqslant x_N\}$ $(n \leqslant N)$, where each element of Π has equal probability to be chosen in the sample. Let $X_{1:n} \leqslant X_{2:n} \leqslant \cdots \leqslant X_{n:n}$ be the ordered sample. In the present paper, the best possible bounds for the expectations of the order statistics $X_{i:n}$ $(1 \leqslant i \leqslant n)$ and the sample range $R_n = X_{n:n} - X_{1:n}$ are derived in terms of the population mean and variance. Some results are also given for the covariance in the simplest case where $n = 2$. An interesting feature of the bounds derived here is that they reduce to some well-known classical results (for the i.i.d. case) as $N \to \infty$. Thus, the bounds established in this paper provide an insight into Hartley–David–Gumbel, Samuelson–Scott, Arnold–Groeneveld and some other bounds. © 2002 Elsevier Science B.V. All rights reserved.

MSC: primary 62G30; secondary 60E15

Keywords: Order statistics; Finite population; Without replacement scheme; Expectation bounds

1. Introduction

In real applications we frequently observe ordered populations of the form $\Pi = \{x_1 \leqslant x_2 \leqslant \cdots \leqslant x_N\}$, and the applied statistician has to consider a without replacement simple random sample $X_1, X_2, \ldots, X_n$ from Π. Of course $n \leqslant N$ (usually, n is much smaller than N) and the case $n = N$ leads to an exhaustive (trivial) sample from Π.

* Corresponding author. Tel.: +1-905-525-9140 ext. 23420; fax: +1-905-522-1676.
E-mail address: bala@mcmail.cis.mcmaster.ca (N. Balakrishnan).

[1] Part of this work was done when he was visiting McMaster University.

PII: S0378-3758(01)00321-4

Let us consider the ordered sample $X_{1:n} \leqslant X_{2:n} \leqslant \cdots \leqslant X_{n:n}$, obtained from the simple random sample $X_1, X_2, \ldots, X_n$. In the present paper, we derive the best possible bounds for $\mathbb{E}[X_{i:n}]$, $1 \leqslant i \leqslant n$ and for $\mathbb{E}[R_n]$, where $R_n = X_{n:n} - X_{1:n}$ is the sample range, when the population mean μ and the population variance σ^2 are fixed. We also discuss some results for the covariance of $X_{1:2}$ and $X_{2:2}$ in the particular case where $n = 2$.

It should be noted that all the preceding results are available in the literature (and are well-known) for the i.i.d. case. In particular, the bound for the expectation of the maximum $X_{n:n}$ is due to Hartley and David (1954) and Gumbel (1954), the bound for the expectation of a single order statistic $X_{i:n}$ $(1 < i < n)$ is due to Moriguti (1953). Also, the bound for the expectation of the sample range was firstly obtained by Plackett (1947). Moriguti (1951) also derived the upper bound for $\mathbb{E}[X_{n:n}]$ when the population is symmetric (the extremal population in this case is the one that also maximizes the expectation of the sample range). For a comprehensive review of the above results see David (1981, Chapter 4).

The bound for $\mathrm{Cov}[X_{1:2}, X_{2:2}]$ was obtained by Papathanasiou (1990), and the equality in this bound characterizes the rectangular (uniform over some interval) distributions. Although Papathanasiou's bound was motivated by Terrell's (1983) result for the correlation of an ordered pair from two i.i.d. r.v.'s, this result was already hidden in the old Hartley–David–Gumbel bound for $n = 2$, as it was shown by Balakrishnan and Balasubramanian (1993).

As it was already noted, all the above results are valid only when the ordered sample $X_{1:n} \leqslant X_{2:n} \leqslant \cdots \leqslant X_{n:n}$ corresponds to n i.i.d. r.v.'s. Some general bounds of this kind for arbitrary (neither independent nor identically distributed) r.v.'s was obtained by Arnold and Groeneveld (1979); see also Arnold and Balakrishnan (1989) for a complete review of related results. Some interesting generalizations can be found in Olkin (1992) and Rychlik (1998). However, in the case of a simple random sample from a finite population, few things are known with respect to moment bounds for order statistics. In this case, the basic r.v.'s are identically distributed and dependent; in fact, they are exchangeable.

Recently, the stochastic properties of order statistics from finite populations have been studied by many authors. The general distribution theory can be found in Arnold et al. (1992, Chapter 3); see also Hájek (1981). Boland et al. (1996), Kochar and Korwar (1997) and Takahasi and Futatsuya (1998) obtained very interesting results on the dependence structure of order statistics from a finite population. More specifically, one of the main results in Takahasi and Futatsuya (1998) is given by the fact that the joint distribution of $X_{i:n}$ and $X_{j:n}$ (for any $1 \leqslant i < j \leqslant n$) is positively likelihood ratio dependent, i.e., their joint probability mass function f satisfies

$$\begin{vmatrix} f(x,y) & f(x,y') \\ f(x',y) & f(x',y') \end{vmatrix} \geqslant 0$$

for all $x, y, x', y' \in \Pi$ with $x < x'$ and $y < y'$. This result was also shown by Boland et al. (1996). Kochar and Korwar (1997) proved that, when no multiplicities occur in Π, then for any $1 \leqslant i < j \leqslant n$, the r.v.'s $X_{i:n}$ and $X_{j:n}$ are likelihood ratio ordered, that is, their marginal probability mass functions $f_{i:n}$ and $f_{j:n}$ satisfy the condition

$$f_{i:n}(x)/f_{j:n}(x) \quad \text{is non-increasing in } x.$$

In addition, they considered a different scheme of sampling from a finite population (Midzuno sampling) and proved that similar results do not hold, apart from some particular cases.

So far, with one way or another, the above results imply that the i.i.d. dependence structure of order statistics remains valid in finite populations where independence is lost. Our approach is in a different context but follows the same spirit: One of the main features of the expectation bounds obtained in the present paper is the fact that, letting $N \to \infty$, we can reobtain all the known classical results for the i.i.d. case, and at the same time, the distribution of the maximizing population for large N approximates the optimal population for the i.i.d. case. Therefore, from one point of view, the bounds provided here may be regarded as 'profitable' for the general i.i.d. bounds.

It is worth pointing out that in contrast to the case of $N \to \infty$, where the results approximate the classical ones of the i.i.d. theory, the case $n = N$ gives the optimal bounds for arbitrarily dependent identically distributed samples; cf. Arnold (1985) for the sample maximum and range and Rychlik (1993) for other order statistics.

2. Upper bound for the sample maximum and the sample range

Every finite quantitative population can be put in the form $\Pi = \{x_1 \leqslant x_2 \leqslant \cdots \leqslant x_N\}$, and in the most natural situation, each element of Π has equal probability $1/N$ to be chosen (in the sense that if there exist $k > 1$ elements in Π with the same value x_j then $\mathbb{P}[X = x_j] = k/N$, where X is the r.v. considered as a random element of Π).

Let $X_1, X_2, \ldots, X_n$ $(n \leqslant N)$ be a simple random sample without replacement from Π. Since this simple random sample consists of dependent and identically distributed (in fact, exchangeable) r.v.'s, the corresponding order statistics $X_{1:n} \leqslant X_{2:n} \leqslant \cdots \leqslant X_{n:n}$ form an ordered sample which does not arise from the i.i.d. case. Therefore, the extensive literature for the expectation bounds, discussed in the introduction, does not apply in this case.

Observe that the population mean $\mu = \mathbb{E}[X]$ and the population variance $\sigma^2 = \mathrm{Var}[X]$ are given by

$$\mu = \frac{1}{N} \sum_{k=1}^{N} x_k \quad \text{and} \quad \sigma^2 = \frac{1}{N} \sum_{k=1}^{N} (x_k - \mu)^2. \tag{2.1}$$

For our purposes we shall use the following simple lemma.

Lemma 2.1. *Assume that $\Pi_0 = \{1, 2, \ldots, N\}$ and consider a simple random sample $U_1, U_2, \ldots, U_n$ $(n \leqslant N)$, drawn from Π_0 without replacement. Then, the ordered samples $X_{1:n} \leqslant X_{2:n} \leqslant \cdots \leqslant X_{n:n}$ from Π and $U_{1:n} \leqslant U_{2:n} \leqslant \cdots \leqslant U_{n:n}$ from Π_0 are related* *through*

$$(X_{1:n}, X_{2:n}, \ldots, X_{n:n}) \overset{\mathrm{d}}{=} (g(U_{1:n}), g(U_{2:n}), \ldots, g(U_{n:n})), \tag{2.2}$$

where $g: \Pi_0 \to \Pi$ *is given by* $g(k) = x_k$, $k = 1, 2, \ldots, N$.

Proof. The result is obvious since g is non-decreasing and

$$(X_1, X_2, \ldots, X_n) \overset{\mathrm{d}}{=} (g(U_1), g(U_2), \ldots, g(U_n)).$$

Using this lemma, we can easily establish the following:

Theorem 2.1. *Consider a without replacement sample* $X_1, X_2, \ldots, X_n$ $(n \leqslant N)$ *from* Π *as above, and let* $X_{1:n} \leqslant X_{2:n} \leqslant \cdots \leqslant X_{n:n}$ *be the corresponding ordered sample. Set*

$$S_N(n) = \sum_{k=n}^{N} \binom{k-1}{n-1}^2 \Big/ \binom{N}{n}^2.$$

If the population has mean μ *and variance* σ^2 *then*

$$\mathbb{E}[X_{n:n}] \leqslant \mu + \sigma\sqrt{NS_N(n) - 1} \tag{2.3}$$

and the equality holds only for the population $\Pi = \{x_1 \leqslant x_2 \leqslant \cdots \leqslant x_N\}$ *with*

$$x_k = \begin{cases} \mu - \sigma(NS_N(n) - 1)^{-1/2}, & k \leqslant n-1, \\ \mu + \sigma\left(N\binom{k-1}{n-1}\Big/\binom{N}{n} - 1\right)(NS_N(n) - 1)^{-1/2}, & k \geqslant n. \end{cases} \tag{2.4}$$

Before providing a proof of Theorem 2.1, let us see two particular examples.

Example 2.1. Take $n = 2$. Then Theorem 2.1 yields

$$\mathbb{E}[X_{2:2}] \leqslant \mu + \frac{\sigma}{\sqrt{3}}\sqrt{\frac{N+1}{N-1}}, \tag{2.5}$$

with equality iff (if and only if)

$$x_k = \mu + \sqrt{\frac{3(N-1)}{N+1}}\left(\frac{2(k-1)}{N-1} - 1\right)\sigma, \quad k = 1, 2, \ldots, N.$$

Observe that for $\sigma > 0$, the finite sequence x_k, $k = 1, 2, \ldots, N$, presents an arithmetic progress with mean μ and variance σ^2; therefore, this optimal population corresponds to the discrete uniform r.v. U_N with probability mass function $\mathbb{P}[U_N = x_k] = 1/N$, $k = 1, 2, \ldots, N$. Letting $N \to \infty$ we readily see that $U_N \to_w U\ (\mu - \sigma\sqrt{3}, \mu + \sigma\sqrt{3})$ (where $U(a, b)$ denotes an r.v. uniformly distributed over the interval (a, b)) and (2.5) leads to the classical Hartley–David–Gumbel bound for the i.i.d. case for $n=2$, namely $\mathbb{E}[X_{2:2}] \leqslant \mu+\sigma/\sqrt{3}$, in which the equality characterizes the $U(\mu-\sigma\sqrt{3}, \mu+\sigma\sqrt{3})$ r.v.

Example 2.2. Take $n=N$. Then, since $X_{N:N}=x_N$ with probability 1, we have $\mathbb{E}[X_{N:N}]=x_N$. Also, $S_N(N)=1$ and thus, Theorem 2.1 yields the Samuelson–Scott (deterministic) inequality (see Arnold and Balakrishnan, 1989, pp. 44–46) $x_N \leqslant \mu+\sigma\sqrt{N-1}$, where the equality holds iff $x_1=x_2=\cdots=x_{N-1}\leqslant x_N$.

Proof of Theorem 2.1. If $\sigma=0$, the result is obvious (and, obviously, the equality is always attained in this case since Π consists of N identical members equal to μ). Assuming that $\sigma>0$ (equivalently, $x_1<x_N$), it is enough to show the result only when $\mu=0$ and $\sigma=1$. Indeed, if the assertion is valid in this particular case, the general one will follow immediately by the observation that $Y_{n:n}=(X_{n:n}-\mu)/\sigma$ is the sample maximum from a without replacement simple random sample of size n drawn from the population $\Pi'=\{y_1\leqslant y_2\leqslant\cdots\leqslant y_N\}$, where $y_k=(x_k-\mu)/\sigma,\ k=1,2,\ldots,N$, with mean 0 and variance 1. Assume then that $\mu=0$ and $\sigma=1$. By Lemma 2.1, since

$$\mathbb{P}[U_{n:n}=k]=\binom{k-1}{n-1}\Big/\binom{N}{n}, \quad k\geqslant n$$

(see Arnold et al., 1992, p. 54), we have

$$\mathbb{E}[X_{n:n}]=\mathbb{E}[g(U_{n:n})]=\sum_{k=n}^{N} x_k \binom{k-1}{n-1}\Big/\binom{N}{n}=\sum_{k=1}^{N} x_k(\delta_k-1/N),$$

where

$$\delta_k=\binom{k-1}{n-1}\Big/\binom{N}{n}, \quad k=1,2,\ldots,N,$$

(with the convention $\binom{a}{b}=0$ for $a<b$), and we used the fact that $\sum_{k=1}^N x_k=0$ (because $\mu=0$). Applying the Cauchy–Schwarz inequality and taking into account the fact that $\sum_{k=1}^N x_k^2=N$ (since $\sigma=1$), we get

$$\mathbb{E}^2[X_{n:n}]\leqslant N\sum_{k=1}^{N}(\delta_k-1/N)^2=NS_N(n)-1$$

and (2.3) is proved. Equality occurs iff $x_k=c(\delta_k-1/N)$ for some constant c and for all k, and from the relation $x_1<x_N$ we conclude that $c>0$ and, therefore, the condition $\sigma=1$ implies that

$$x_k=\frac{N\delta_k-1}{\sqrt{NS_N(n)-1}}, \quad k=1,2,\ldots,N,$$

which completes the proof. □

The upper bound of Theorem 2.1 for $\mathbb{E}[X_{n:n}]$ can be easily transformed to a lower bound for $\mathbb{E}[X_{1:n}]$. In fact we have the following:

Corollary 2.1. *Under the conditions of Theorem* 2.1,

$$\mathbb{E}[X_{1:n}] \geqslant \mu - \sigma\sqrt{NS_N(n)-1}$$

and the equality holds only for the population $\Pi = \{x_1 \leqslant x_2 \leqslant \cdots \leqslant x_N\}$ *with*

$$x_k = \begin{cases} \mu - \sigma\left(N\binom{N-k}{n-1}\Big/\binom{N}{n} - 1\right)(NS_N(n)-1)^{-1/2}, & k \leqslant N-n+1, \\ \mu + \sigma(NS_N(n)-1)^{-1/2}, & k \geqslant N-n+2. \end{cases}$$

Proof. Consider the population $\Pi' = \{y_1 \leqslant y_2 \leqslant \cdots \leqslant y_N\}$ with $y_k = -x_{N+1-k}$ for all k. It follows that the mean μ' and the standard deviation σ' of Π' are simply $\mu' = -\mu$ and $\sigma' = \sigma$. Therefore, if $Y_{n:n}$ is the sample maximum from a simple random sample of size n drawn from Π', then $Y_{n:n} \overset{\mathrm{d}}{=} -X_{1:n}$, and the desired result is an immediate consequence of Theorem 2.1 applied to Π'. □

In order to show a similar result for the sample range $R_n = X_{n:n} - X_{1:n}$, we need the following combinatorial lemma.

Lemma 2.2. *For positive integers* $j \leqslant r \leqslant N$,

$$\sum_{k=j}^{N-r+j} \binom{k-1}{j-1}\binom{N-k}{r-j} = \binom{N}{r}$$

and, in particular, if $2n-1 \leqslant N$ *then*

$$\sum_{k=n}^{N+1-n} \binom{k-1}{n-1}\binom{N-k}{n-1} = \binom{N}{2n-1}.$$

Proof. Let $\Omega = \{1,2,\ldots,N\}$ and consider the set $\mathscr{A}$ of subsets of Ω that include r elements. Also, consider the set $\mathscr{A}_k \subset \mathscr{A}$ of subsets of Ω that include the integer k, $j-1$ elements less than k and $r-j$ elements greater than k, for $k = j, j+1, \ldots, j+N-r$. Then, $\{\mathscr{A}_j, \mathscr{A}_{j+1}, \ldots, \mathscr{A}_{j+N-r}\}$ is a partition of $\mathscr{A}$. Therefore, $|\mathscr{A}| = \sum_{k=j}^{N-r+j} |\mathscr{A}_k|$, and since

$$|\mathscr{A}| = \binom{N}{r} \quad \text{and} \quad |\mathscr{A}_k| = \binom{k-1}{j-1}\binom{N-k}{r-j},$$

the required expression is deduced. Setting $r = 2n-1$ and $j = n$ to the first expression we conclude the second one.

We can now prove the following result for the sample range.

Theorem 2.2. *Under the assumptions of Theorem* 2.1,

$$\mathbb{E}[R_n] \leqslant \sigma\sqrt{2N}\left(S_N(n) - \binom{N}{2n-1}\bigg/\binom{N}{n}^2\right)^{1/2}, \tag{2.6}$$

where $\binom{a}{b}$ *should be treated as* 0 *whenever* $a < b$. *The equality in* (2.6) *holds only for the population* $\Pi = \{x_1 \leqslant x_2 \leqslant \cdots \leqslant x_N\}$ *with*

$$x_k = \mu + \sigma\frac{\sqrt{N}\left[\binom{k-1}{n-1} - \binom{N-k}{n-1}\right]}{\sqrt{2}\binom{N}{n}\left(S_N(n) - \binom{N}{2n-1}\bigg/\binom{N}{n}^2\right)^{1/2}}, \quad k = 1,2,\ldots,N.$$

Proof. For $\sigma = 0$ the result is trivial. If $\sigma > 0$, we may assume that $\mu = 0$ and $\sigma = 1$ without any loss of generality. Then, by Lemma 2.1,

$$\mathbb{E}[R_n] = \mathbb{E}[g(U_{n:n})] - \mathbb{E}[g(U_{1:n})] = \binom{N}{n}^{-1}\sum_{k=1}^{N}\delta_k x_k,$$

where

$$\delta_k = \binom{k-1}{n-1} - \binom{N-k}{n-1}, \quad k = 1,2,\ldots,N$$

(with the convention $\binom{a}{b} = 0$ for $a < b$). Therefore, taking into account the fact that $\sum_{k=1}^{N} x_k^2 = N$ (because $\mu = 0$ and $\sigma = 1$), we get

$$\mathbb{E}^2[R_n] = \binom{N}{n}^{-2}\left(\sum_{k=1}^{N}\delta_k x_k\right)^2 \leqslant N\binom{N}{n}^{-2}\sum_{k=1}^{N}\delta_k^2$$

by the Cauchy–Schwarz inequality. A simple calculation, using Lemma 2.2, shows that

$$\sum_{k=1}^{N}\delta_k^2 = 2\binom{N}{n}^2 S_N(n) - 2\binom{N}{2n-1};$$

this proves (2.6). Equality holds iff $x_k = c\delta_k$ for some constant c and for all k. Since δ_k is non-decreasing in k, $\sum_{k=1}^{N}\delta_k = 0$ and $\delta_1 < 0 < \delta_N$, it follows that the only population (with mean 0 and variance 1) that attains the equality in (2.6) is given by

$$x_k = \delta_k\sqrt{N}\left(\sum_{j=1}^{N}\delta_j^2\right)^{-1/2}, \quad k = 1,2,\ldots,N,$$

and the proof is completed. □

Three particular cases are discussed in the following examples.

Example 2.3. For $n=2$ and $N \geqslant 3$ Theorem 2.1 yields

$$\mathbb{E}[R_2] \leqslant \frac{2\sigma}{\sqrt{3}}\sqrt{\frac{N+1}{N-1}} \to \frac{2\sigma}{\sqrt{3}} \quad \text{as } N \to \infty. \tag{2.7}$$

Also, for $n=3$ and $N \geqslant 5$ we get the inequality

$$\mathbb{E}[R_3] \leqslant \sigma\sqrt{3}\sqrt{\frac{N+1}{N-1}} \to \sigma\sqrt{3} \quad \text{as } N \to \infty. \tag{2.8}$$

Equality in both bounds is attained iff the finite sequence x_k, $k=1,2,\ldots,N$, presents an arithmetic progress with mean μ and variance σ^2 as in Example 2.1; therefore, the optimal population corresponds to the discrete uniform r.v. U_N which, for large N, approximates $U(\mu-\sigma\sqrt{3},\mu+\sigma\sqrt{3})$. Both (2.7) and (2.8) are the finite analogues of the classical (Plackett, 1947, Moriguti, 1951) bounds for the expectation of the sample range (for $n=2$ and 3) in the i.i.d. case.

Example 2.4. Take $n=N$. Then, since $R_N = x_N - x_1$ with probability 1 and $2n-1 = 2N-1 > N$ for $N \geqslant 2$, Theorem 2.2 reduces to the old (deterministic) Nair–Thomson inequality (see Arnold and Balakrishnan, 1989, p. 48)

$$x_N - x_1 \leqslant \sigma\sqrt{2N}. \tag{2.9}$$

Since $\delta_N = 1 = -\delta_1$ and $\delta_k = 0$ for all other k, equality is attained in (2.9) iff $x_1 \leqslant x_2 = \cdots = x_{N-1} \leqslant x_N$ and $x_N - x_{N-1} = x_2 - x_1$.

3. Expectation bounds for a single order statistic: Moriguti's method

Consider again a simple random sample of size n drawn from a finite population $\Pi = \{x_1 \leqslant x_2 \leqslant \cdots \leqslant x_N\}$ with mean μ and variance σ^2. Suppose that $3 \leqslant n \leqslant N$ and fix an integer i with $2 \leqslant i \leqslant n-1$. From Lemma 2.1 we have

$$\mathbb{E}[X_{i:n}] = \mathbb{E}[g(U_{i:n})] = \binom{N}{n}^{-1} \sum_{k=1}^{N} \binom{k-1}{i-1}\binom{N-k}{n-i} x_k = \sum_{k=1}^{N} p_k x_k. \tag{3.1}$$

Therefore, assuming that $\mu=0$ and $\sigma=1$ and proceeding exactly as in Theorem 2.1, one can easily establish the inequality

$$\mathbb{E}^2[X_{i:n}] \leqslant N S_N(n,i) - 1, \tag{3.2}$$

where

$$S_N(n,i) = \sum_{k=i}^{N-n+i} \binom{k-1}{i-1}^2 \binom{N-k}{n-i}^2 \Big/ \binom{N}{n}^2$$

(note that $S_N(n,n) = S_N(n,1) = S_N(n)$, where $S_N(n)$ is defined in Theorem 2.1). However, for $1 < i < n$, there does not exist finite populations with mean 0 and variance 1 attaining the equality in (3.2); this is so because the finite sequence p_k, $k=1,2,\ldots,N$, in (3.1), fails to be monotonic and, thus, the Cauchy–Schwarz inequality in (3.2) yields a greater upper bound than it is really needed. Based on convexity arguments, Moriguti (1953) proved an ingenious 'Cauchy–Schwarz inequality for increasing functions', in order to determine a sharp upper bound for $\mathbb{E}[X_{i:n}]$ in the i.i.d. case. The purpose of the present section is in finding the corresponding sharp bounds for the case of a finite population. The following lemma will be used in the sequel.

Lemma 3.1. *Let*

$$p_k = \binom{N}{n}^{-1}\binom{k-1}{i-1}\binom{N-k}{n-i}, \quad k = 1,2,\ldots,N$$

be as in (3.1) *and set* $P_k = \sum_{j=1}^{k} p_j$ *for* $k = 1,2,\ldots,N$ (*thus*, $P_1 = 0$, $P_N = 1$). *Then*,

(i) *There exists a unique integer* k_0 (*depending on* i, n *and* N) *such that*

$$(N-k_0)p_{k_0} \leqslant 1 - P_{k_0} < (N-k_0)p_{k_0+1}. \tag{3.3}$$

Moreover, $i-1 \leqslant k_0 < (i-1)N/(n-1)$.

(ii) *Let*

$$\delta_k = \begin{cases} p_k, & k = 1,2,\ldots,k_0, \\ \dfrac{1-P_{k_0}}{N-k_0}, & k = k_0+1,\ldots,N, \end{cases}$$

and set $\Delta_k = \sum_{j=1}^{k} \delta_j$ *for* $k = 1,2,\ldots,N$ (*thus*, $\Delta_1 = 0$). *Then*,

(a) $0 \leqslant \delta_1 \leqslant \delta_2 \leqslant \cdots \leqslant \delta_N$,

(b) $\Delta_N = 1$,

(c) $\Delta_k \leqslant P_k$ *for all* k, *and*

(d) $\sum_{k=1}^{N} p_k\delta_k = \sum_{k=1}^{N} \delta_k^2$.

We omit the proof of Lemma 3.1, because the assertion of existence and uniqueness of real numbers $\delta_1, \delta_2, \ldots, \delta_N$ satisfying (a), (b), (c) and (d) of the lemma is a standard fact in functional analysis; in fact, the vector $(\delta_1, \delta_2, \ldots, \delta_N)$ is the l^2-projection of the vector $(p_1, p_2, \ldots, p_N)$ onto the convex cone of componentwise non-decreasing vectors of $\mathbb{R}^N$ (see Balakrishnan, 1981, Section 1.4). The other details of (i) and (ii) of Lemma 3.1 can be verified directly.

Remark 3.1. The unique integer k_0 of Lemma 3.1 plays exactly the same role as ρ_1, the unique solution of the equation $1 - G(x) = (1-x)g(x)$, $0 < x < 1$, where

$$g(x) = \frac{1}{B(i, n+1-i)} x^{i-1}(1-x)^{n-i}, \quad G(x) = \int_0^x g(t)\,\mathrm{d}t = I_x(i, n+1-i)$$

and (3.3) is the discrete analogue of the above (continuous) equation. This root ρ_1 has been shown to play a vital role in the determination of sharp bounds for moments of order statistics; see Moriguti (1953), Balakrishnan (1993) and Papadatos (1997).

Corollary 3.1. *For any set of real numbers* $x_1 \leqslant x_2 \leqslant \cdots \leqslant x_N$,

$$\sum_{k=1}^{N} p_k x_k \leqslant \sum_{k=1}^{N} \delta_k x_k, \tag{3.4}$$

where p_k, δ_k, $k = 1,2,\ldots,N$, *are as in Lemma* 3.1. *A sufficient condition for the equality to hold in* (3.4) *is given by*

$$x_{k_0+1} = x_{k_0+2} = \cdots = x_N, \tag{3.5}$$

where k_0 *is defined in Lemma* 3.1.

Proof. Since $x_1 \leqslant x_2 \leqslant \cdots \leqslant x_N$, we may write $x_k = y_1 + \cdots + y_k$, $k = 1,2,\ldots,N$, where $y_j \geqslant 0$ for $j \geqslant 2$. Then, by Lemma 3.1,

$$\begin{aligned}\sum_{k=1}^{N} p_k x_k &= \sum_{k=1}^{N} p_k \left(\sum_{j=1}^{k} y_j\right) = \sum_{j=1}^{N} y_j \left(\sum_{k=j}^{N} p_k\right) = y_1 + \sum_{j=2}^{N} y_j (1 - P_{j-1}) \\ &\leqslant y_1 + \sum_{j=2}^{N} y_j (1 - \Delta_{j-1}) = \sum_{j=1}^{N} y_j \left(\sum_{k=j}^{N} \delta_k\right) \\ &= \sum_{k=1}^{N} \delta_k \left(\sum_{j=1}^{k} y_j\right) = \sum_{k=1}^{N} \delta_k x_k,\end{aligned}$$

which proves (3.4). Since (3.5) implies that $y_{k_0+2} = \cdots = y_N = 0$, we have

$$\begin{aligned}y_1 + \sum_{j=2}^{N} y_j (1 - P_{j-1}) &= y_1 + \sum_{j=2}^{k_0+1} y_j (1 - P_{j-1}) \\ &= y_1 + \sum_{j=2}^{k_0+1} y_j (1 - \Delta_{j-1}) = y_1 + \sum_{j=2}^{N} y_j (1 - \Delta_{j-1})\end{aligned}$$

(because $P_k = \Delta_k$ for $k \leqslant k_0$), which shows that the equality holds in (3.4). □

We can now prove the main result of this section, stated in the following:

Theorem 3.1. *Under the conditions of Theorem* 2.1, *for any fixed* i, $2 \leqslant i \leqslant n-1$, *we have the inequality*

$$\mathbb{E}[X_{i:n}] \leqslant \mu + \sigma \sqrt{N \sum_{k=1}^{N} \delta_k^2 - 1},$$

with equality iff

$$x_k = \mu + \sigma \frac{N\delta_k - 1}{\sqrt{N\sum_{j=1}^{N} \delta_j^2 - 1}}, \quad k = 1,2,\ldots,N,$$

where δ_k *is given by Lemma* 3.1.

Proof. Without any loss of generality, assume that $\mu = 0$ and $\sigma = 1$. Then, since $\sum_{k=1}^{N} x_k = 0$, it follows from Corollary 3.1 that

$$\mathbb{E}[X_{i:n}] = \sum_{k=1}^{N} p_k x_k \leqslant \sum_{k=1}^{N} \delta_k x_k = \sum_{k=1}^{N} (\delta_k - 1/N) x_k.$$

By using the relation $\sum_{k=1}^{N} x_k^2 = N$, an application of the Cauchy–Schwarz inequality to the last sum yields $\sum_{k=1}^{N} (\delta_k - 1/N) x_k \leqslant \sqrt{N \sum_{k=1}^{N} (\delta_k - 1/N)^2}$, and the desired inequality follows from the fact that $\sum_{k=1}^{N} (\delta_k - 1/N)^2 = \sum_{k=1}^{N} \delta_k^2 - 1/N$. The last inequality becomes equality iff there exists a constant c such that $x_k = c(\delta_k - 1/N)$ for all k. Therefore, since $\mu = 0$, $\sigma = 1$, and the finite sequence x_k, $k = 1,2,\ldots,N$, is non-decreasing with $x_1 < x_N$, it follows that the last inequality becomes equality iff

$$x_k = \frac{N\delta_k - 1}{\sqrt{N\sum_{j=1}^{N} \delta_j^2 - 1}}, \quad k = 1,2,\ldots,N.$$

On the other hand, since $\delta_{k_0+1} = \cdots = \delta_N$, we conclude that the x_k above satisfy $x_{k_0+1} = \cdots = x_N$, and by Corollary 3.1, the first inequality also becomes equality. Therefore, this is the unique population attaining equality in the upper bound (when $\mu = 0$ and $\sigma = 1$), and the proof is complete. □

Regarding the lower bounds, one can easily show the following result.

Corollary 3.2. *Under the conditions of Theorem* 2.1, *for any fixed* i, $2 \leqslant i \leqslant n - 1$, *we have the inequality*

$$\mathbb{E}[X_{n+1-i:n}] \geqslant \mu - \sigma \sqrt{N \sum_{k=1}^{N} \delta_k^2 - 1},$$

with equality iff

$$x_k = \mu - \sigma \frac{N\delta_{N+1-k} - 1}{\sqrt{N\sum_{j=1}^{N} \delta_j^2 - 1}}, \quad k = 1,2,\ldots,N,$$

where δ_k *is given by Lemma* 3.1.

We omit the simple proof; this can be done by using exactly the same arguments as in the proof of Corollary 2.1. We only consider some numerical examples.

Example 3.1. Let $i=2$ and $n=3$. Then, if $N\in\{3,4,5,6\}$, we have that $k_0=1$, and Theorem 3.1 yields the inequality $\mathbb{E}[X_{2:3}]\leqslant\mu+\sigma/\sqrt{N-1}$, with equality iff $x_1\leqslant x_2=\cdots=x_N$. For $N=7$, however, Theorem 3.1 leads to the inequality $\mathbb{E}[X_{2:3}]\leqslant\mu+\sigma\sqrt{6/35}$, with equality iff $\Pi=\{\mu-\sigma\sqrt{35/6},\mu,\mu+\sigma\sqrt{7/30},\ldots,\mu+\sigma\sqrt{7/30}\}$.

Example 3.2. Take $n=N$ and $1<i<N$. Then, $X_{i:N}=x_i$ with probability 1. Since $p_k\neq 0$ only when $k=i$, it follows that $k_0=i-1$ is the unique integer satisfying (3.3); thus $\delta_k=0$ for $k<i$, and $\delta_k=(N-i+1)^{-1}$ for $k\geqslant i$, yielding

$$N\sum_{k=1}^{N}\delta_k^2-1=\frac{i-1}{N-i+1}.$$

Applying Theorem 3.1 and Corollary 3.2, we get the bounds

$$\mu-\sigma\sqrt{\frac{N-i}{i}}\leqslant x_i\leqslant\mu+\sigma\sqrt{\frac{i-1}{N-i+1}}.$$

Furthermore, equality occurs in the lower bound iff $x_1=\cdots=x_i\leqslant x_{i+1}=\cdots=x_N$, and in the upper bound iff $x_1=\cdots=x_{i-1}\leqslant x_i=\cdots=x_N$. This inequality has been shown by several authors using different methods; see Arnold and Balakrishnan (1989, p. 48, Theorem 3.8), Scott (1936), Mallows and Richter (1969), Hawkins (1971), Boyd (1971), Arnold and Groeneveld (1979), and Wolkowicz and Styan (1979).

4. Covariance bounds and related results

In the two previous sections, the best possible upper and lower bounds for the expectations of the order statistics $X_{i:n}$, $1\leqslant i\leqslant n$, from a finite population, were obtained in terms of the population mean and variance; the only exceptions being the lower bound for the sample maximum (similarly, the upper bound for the sample minimum) and the lower bound for the sample range. In this section, we solve the exceptional cases, expanding the relation between these bounds and the corresponding covariance bound for $n=2$. First of all, we state two lemmas.

Lemma 4.1. *In any finite population Π of size $N\geqslant 2$ with mean μ and variance σ^2,*

$$\mathrm{Cov}[X_{1:2},X_{2:2}]=(\mathbb{E}[X_{2:2}]-\mu)^2-\sigma^2/(N-1), \tag{4.1}$$

where $X_{1:2}\leqslant X_{2:2}$ is an ordered sample of size 2 drawn from Π without replacement.

Proof. If $\sigma=0$ the result is obvious; otherwise we may assume that $\mu=0$ and $\sigma=1$. Then, from Lemma 2.1 and the fact that $g(U_1)g(U_2)=g(U_{1:2})g(U_{2:2})$ with probability 1, it follows that

$$\mathbb{E}[X_{1:2}X_{2:2}]=\mathbb{E}[g(U_{1:2})g(U_{2:2})]=\mathbb{E}[g(U_1)g(U_2)]$$

$$= \frac{1}{N(N-1)} \sum_{k=1}^{N} x_k \left(\sum_{s \neq k} x_s \right) = \frac{1}{N(N-1)} \sum_{k=1}^{N} x_k \left(-x_k + \sum_{s=1}^{N} x_s \right)$$

$$= -\frac{1}{N-1}.$$

On the other hand, since $\mathbb{E}[X_{2:2}] = -\mathbb{E}[X_{1:2}]$, we have

$$\mathrm{Cov}[X_{1:2}, X_{2:2}] = -\frac{1}{N-1} - (-\mathbb{E}[X_{2:2}])\mathbb{E}[X_{2:2}] = \mathbb{E}^2[X_{2:2}] - \frac{1}{N-1},$$

which is (4.1).

Lemma 4.2. *The sample maxima from a finite population of size $N \geqslant 2$ are stochastically ordered according to their sample size, that is,*

$$\mathbb{P}[X_{n:n} \leqslant x] \geqslant \mathbb{P}[X_{n+1:n+1} \leqslant x], \quad -\infty < x < \infty,$$

for any $n = 1, 2, \ldots, N-1$.

Proof. By Lemma 2.1, $X_{n:n} \stackrel{\mathrm{d}}{=} g(U_{n:n})$, and since g is non-decreasing, it suffices to prove the lemma for the population Π_0. However, in this case, the result is evident because

$$\mathbb{P}[U_{n:n} \leqslant x] = \binom{\min\{[x], N\}}{n} \bigg/ \binom{N}{n}, \quad 0 \leqslant x < \infty.$$

In particular, it follows from Lemma 4.2 that

$$\mu = \mathbb{E}[X_{1:1}] \leqslant \mathbb{E}[X_{2:2}] \leqslant \cdots \leqslant \mathbb{E}[X_{N:N}] = x_N. \tag{4.2}$$

Using the above lemmas and a result of Takahasi and Futatsuya (1998), one can easily obtain the following:

Theorem 4.1. *Let $X_{n:n}$ (resp. $X_{1:n}$) be the sample maximum (resp. minimum) from a random sample of size $n \geqslant 2$ drawn without replacement from a finite population Π of size $N \geqslant n$ with mean μ and variance σ^2. Then,*

$$\mathbb{E}[X_{n:n}] \geqslant \mu + \sigma/\sqrt{N-1} \tag{4.3}$$

and similarly,

$$\mathbb{E}[X_{1:n}] \leqslant \mu - \sigma/\sqrt{N-1}. \tag{4.4}$$

Equality in (4.3) holds iff any one of the following conditions is satisfied: either (a) $\Pi = \{x_1 \leqslant x_2 = \cdots = x_N\}$, or (b) $n = 2$ and $\Pi = \{x_1 = \cdots = x_{N-1} < x_N\}$. Similarly, the equality in (4.4) occurs iff any one of the following conditions holds: either (a′) $\Pi = \{x_1 = \cdots = x_{N-1} \leqslant x_N\}$, or (b′) $n = 2$ and $\Pi = \{x_1 < x_2 = \cdots = x_N\}$.

Proof. From (4.2), $\mathbb{E}[X_{n:n}] \geqslant \mathbb{E}[X_{2:2}] \geqslant \mu$. On the other hand, Takahasi and Futatsuya (1998) showed that $\mathrm{Cov}[X_{1:2}, X_{2:2}] \geqslant 0$, with equality iff either $\Pi = \{x_1 \leqslant x_2 = \cdots = x_N\}$

or $\Pi = \{x_1 = \cdots = x_{N-1} < x_N\}$. Therefore, from (4.1) and (4.2) we have

$$\mathbb{E}[X_{n:n}] \geqslant \mathbb{E}[X_{2:2}] = \mu + \sqrt{\mathrm{Cov}[X_{1:2}, X_{2:2}] + \sigma^2/(N-1)} \geqslant \mu + \sigma/\sqrt{N-1},$$

which is (4.3). For $n=2$, equality holds in (4.3) iff $\mathrm{Cov}[X_{1:2}, X_{2:2}] = 0$; that is, iff (a) or (b) holds. For $n \geqslant 3$, however, the equality is attained iff $\mathrm{Cov}[X_{1:2}, X_{2:2}] = 0$ and $\mathbb{E}[X_{n:n}] = \mathbb{E}[X_{2:2}]$; therefore, the populations of the form $\Pi = \{x_1 = \cdots = x_{N-1} < x_N\}$ do not yield equality in (4.3) (since $\mathbb{E}[X_{n:n}] \geqslant \mathbb{E}[X_{3:3}]$, and it is easy to check that $\mathbb{E}[X_{3:3}] > \mathbb{E}[X_{2:2}]$ in this case), and thus, the only case of equality, when $n \geqslant 3$, is given by (a). The case for $\mathbb{E}[X_{1:n}]$ can be treated similarly. □

For $n = N$, Theorem 4.1 yields the Hawkins (deterministic) bounds (see Arnold and Balakrishnan, 1989, p. 49; Hawkins, 1971; Boyd, 1971), namely

$$x_N \geqslant \mu + \sigma/\sqrt{N-1}, \quad x_1 \leqslant \mu - \sigma/\sqrt{N-1},$$

where the equality holds in the former (latter) bound iff the $N-1$ greatest (lowest) elements of Π are equal.

Corollary 4.1. *Under the notation of Lemma* 4.1, *for the sample range* $R_2 = X_{2:2} - X_{1:2}$ *we have the inequality* $\mathbb{E}[R_2] \geqslant 2\sigma/\sqrt{N-1}$, *with equality iff either* $\Pi = \{x_1 \leqslant x_2 = \cdots = x_N\}$ *or* $\Pi = \{x_1 = \cdots = x_{N-1} < x_N\}$.

Proof. It is evident because $\mathbb{E}[R_2] = 2(\mathbb{E}[X_{2:2}] - \mu)$. □

Finally, for the covariance of an ordered sample of size 2, we have the following:

Corollary 4.2. *Under the notation of Corollary* 4.1,

$$\mathrm{Cov}[X_{1:2}, X_{2:2}] \leqslant \frac{N-2}{N-1}\left(\frac{1}{3}\sigma^2\right), \tag{4.5}$$

with equality iff Π *is as in Example* 2.1.

Proof. It is evident from Lemma 4.1 and Theorem 2.1. □

The bound (4.5) is the discrete analogue of the result given by Papathanasiou (1990) for the i.i.d. case, namely $\mathrm{Cov}[X_{1:2}, X_{2:2}] \leqslant \sigma^2/3$, where the equality characterizes the $U(\mu - \sigma\sqrt{3}, \mu + \sigma\sqrt{3})$ r.v. Note that the bound (4.5) approximates Papathanasiou's bound for large N (and the optimal population approximates the corresponding uniform r.v.). In fact, Balakrishnan and Balasubramanian (1993) have shown that Papathanasiou's bound and the associated uniform characterization are equivalent to those based on Hartley–David–Gumbel bound (see Section 2).

It should be noted that we have not been able to find a general sharp lower bound for $\mathbb{E}[R_n]$, and the sharp (upper and lower) bounds for $\mathrm{Cov}[X_{1:n}, X_{n:n}]$, when $n \geqslant 3$. It seems, however, that the techniques needed for derivation of this kind of bounds are completely different than those used here (see, for example, Papadatos, 1999, for the upper bound of $\mathrm{Cov}[X_{1:3}, X_{3:3}]$ in the i.i.d. case, where it is proved that the maximum is attained by the *hyperbolic sine* density); note also that the best lower bound for $\mathbb{E}[R_N] = x_N - x_1$, when $N \geqslant 2$, is given by (see Arnold and Balakrishnan, 1989, p. 50; Fahmy and Proschan, 1981; Thomson, 1955)

$$x_N - x_1 \geqslant \begin{cases} 2\sigma & \text{if } N \text{ is even}, \\ \dfrac{2N}{\sqrt{N^2-1}}\sigma & \text{if } N \text{ is odd}, \end{cases}$$

where the equality is attained iff there exist some constants a, b, such that $N/2$ elements of Π are equal to a, and the remaining $N/2$ elements are equal to b, when N is even, or $(N-1)/2$ elements are equal to a, and the remaining $(N+1)/2$ elements are equal to b, when N is odd).

5. Limiting behavior of the bounds for large populations

In this section, we shall show that the results discussed in Section 2 approximate the well known classical ones for the i.i.d. case, as $N \to \infty$. For this reason, we will make use of the following lemma.

Lemma 5.1. *Let $S_N(n)$ be as in Theorems* 2.1 *and* 2.2. *Then,*

$$\lim_{N\to\infty} NS_N(n) = \frac{n^2}{2n-1}. \tag{5.1}$$

Proof. We have

$$\begin{aligned} NS_N(n) &= N \sum_{k=n}^{N} \left(\frac{[(k-1)!]^2}{[(n-1)!]^2\,[(k-n)!]^2} \right) \left(\frac{(n!)^2\,[(N-n)!]^2}{(N!)^2} \right) \\ &= \frac{1}{N} \sum_{k=n}^{N} n^2 \left\{ \frac{(k-1)(k-2)\cdots(k-n+1)}{(N-1)(N-2)\cdots(N-n+1)} \right\}^2 \\ &\to \int_0^1 n^2 (u^{n-1})^2\,\mathrm{d}u \quad (\text{as } N \to \infty), \end{aligned}$$

by the Riemann integral, which completes the proof.

We can now prove the main result of this section, stated in the following:

Theorem 5.1. *For fixed n, the upper bound of Theorem* 2.1 *approximates the Hartley–David–Gumbel bound for the i.i.d. case, as $N \to \infty$, that is*

$$\lim_{N\to\infty} \left(\mu + \sigma\sqrt{NS_N(n) - 1}\right) = \mu + \frac{n-1}{\sqrt{2n-1}}\sigma. \tag{5.2}$$

Moreover, if X_N is the r.v. corresponding to the optimal population Π given by (2.4) *and* $\sigma > 0$, *then*

$$X_N \to_w X \quad \text{as } N \to \infty,$$

where X is the r.v. with d.f.

$$F_X(x) = \left(\frac{1}{n} + \frac{(n-1)(x-\mu)}{\sigma n\sqrt{2n-1}}\right)^{1/(n-1)}, \quad \mu - \sigma\frac{\sqrt{2n-1}}{n-1} \leqslant x \leqslant \mu + \sigma\sqrt{2n-1}.$$

Proof. From (5.1), it follows immediately that $\lim_{N\to\infty}(NS_N(n)-1)^{1/2} = (n-1) \times (2n-1)^{-1/2}$, which proves (5.2). Fix now a number $u \in (0,1)$, and define $F_N^{-1}(u) = \inf\{x\colon \mathbb{P}[X_N \leqslant x] \geqslant u\}$, and

$$F_X^{-1}(u) = \inf\{x\colon F_X(x) \geqslant u\} = \mu + \sigma\sqrt{2n-1}(nu^{n-1}-1)/(n-1)$$

(note that, by definition, the maximizing population $\Pi = \Pi_N$ varies with N, in such a way that both $\mathbb{E}[X_N] = \mu$ and $\mathrm{Var}[X_N] = \sigma^2$ remain constant for all $N \geqslant n$). For $N > \max\{n/u, 2/(1-u)\}$, let $k_N(u)$ be the unique integer in $\{n,\ldots,N-2\}$ satisfying $k_N(u)/N < u \leqslant (k_N(u)+1)/N$. Since $k_N(u) \geqslant n$, it follows from (2.4) that

$$\begin{aligned} F_N^{-1}(u) &= F_N^{-1}((k_N(u)+1)/N) \\ &= \mu + \sigma\left(N\binom{k_N(u)}{n-1}\Big/\binom{N}{n} - 1\right)(NS_N(n)-1)^{-1/2}. \end{aligned}$$

Obviously, $\lim_{N\to\infty}(k_N(u)-j)/N = u$, for any fixed j (because $[Nu]-1 \leqslant k_N(u) \leqslant [Nu]+1$), and therefore,

$$\begin{aligned} N\binom{k_N(u)}{n-1}\Big/\binom{N}{n} &= n\frac{(k_N(u)/N(k_N(u)/N-1/N))\cdots(k_N(u)/N-(n-2)/N)}{(1-1/N)(1-2/N)\cdots(1-(n-1)/N)} \\ &\to nu^{n-1} \end{aligned}$$

as $N \to \infty$. Thus, $\lim_{N\to\infty} F_N^{-1}(u) = F_X^{-1}(u)$ for all $u \in (0,1)$, and the proof is complete. □

Using exactly the same arguments, one can easily establish a similar result for the sample range. We give without proof the following:

Theorem 5.2. *For fixed $n \geqslant 2$, the upper bound of Theorem* 2.2 *approximates the Plackett–Moriguti bound for the i.i.d. case, as $N \to \infty$, that is*

$$\begin{aligned} &\lim_{N\to\infty} \sigma\sqrt{2N}\left(S_N(n) - \binom{N}{2n-1}\Big/\binom{N}{n}^2\right)^{1/2} \\ &= \frac{n\sqrt{2}}{\sqrt{2n-1}}\left(1 - \binom{2n-2}{n-1}^{-1}\right)^{1/2}\sigma. \end{aligned}$$

Moreover, if X_N is the r.v. corresponding to the optimal population Π given by Theorem 2.2 *and* $\sigma > 0$, *then* $X_N \to_w X$, *as* $N \to \infty$, *where X is the r.v. with inverse* *d.f. given by* $F_X^{-1}(u) = \mu + a_n\sigma(u^{n-1} - (1-u)^{n-1})$, $0 < u < 1$, *with*

$$a_n = \frac{\sqrt{2n-1}}{\sqrt{2}}\left(1 - \binom{2n-2}{n-1}^{-1}\right)^{-1/2}.$$

It should be noted that similar results could also be valid for the case of a single order statistic $X_{i:n}$, $1 < i < n$, and the upper bound of Theorem 3.1, for large N, presumably approximates the corresponding Moriguti's bound; we do not treat this case, however, because these bounds are not available in a closed form.

6. Concluding remarks

Since in most real applications, the applied statistician usually considers finite populations, the results of this paper can be viewed as 'direct', in contrast to the i.i.d. case where the bounds are, from this point of view, 'limiting cases'. It has been shown in Section 5 that, almost all bounds discussed in this article, approximate (when the population is large) the limiting ones. This fact is fairly expected, however, in view of the following lemma, which may be of some independent interest. (We note that the assertion of this lemma is probably known; since, however, we have not been able to trace a proof, we provide one for the completeness of the presentation.)

Lemma 6.1. *Let* $X_1, X_2, \ldots, X_n$ *be an i.i.d. sample from a d.f.* F *with mean* μ *and finite variance* σ^2. *Then, there exist a sequence of finite populations* Π_N *of size* N *with mean* μ *and variance* σ^2 *such that*

$$(X_{1,N}, X_{2,N}, \ldots, X_{n,N}) \to_w (X_1, X_2, \ldots, X_n), \quad \text{as } N \to \infty,$$

where $X_{1,N}, X_{2,N}, \ldots, X_{n,N}$ *is a sample taken without replacement from* Π_N.

Proof. Since the case $\sigma^2 = 0$ is trivial, assume that $\sigma^2 > 0$. Let $U_1, U_2, \ldots, U_n$ be an i.i.d. sample from $U(0,1)$ and denote by $U_{1,N}, U_{2,N}, \ldots, U_{n,N}$ the without replacement sample from $\Pi_0^N = \{1/(N+1), 2/(N+1), \ldots, N/(N+1)\}$, with $N \geqslant n$. Let $h : \mathbb{R} \to \mathbb{R}$ be a continuous bounded function and consider n real constants $c_1, c_2, \ldots, c_n$. It follows that

$$\begin{aligned}
\mathbb{E}\left[h\left(\sum_{j=1}^{n} c_j U_{j,N}\right)\right] &= \frac{1}{(N)_n}\sum_{(k_1,k_2,\ldots,k_n)} h\left(\sum_{j=1}^{n} c_j k_j/(N+1)\right) \\
&\to \int_0^1\int_0^1 \cdots \int_0^1 h\left(\sum_{j=1}^{n} c_j u_j\right) \mathrm{d}u_n\,\mathrm{d}u_{n-1}\cdots \mathrm{d}u_1 \\
&\quad (\text{as } N \to \infty) \\
&= \mathbb{E}\left[h\left(\sum_{j=1}^{n} c_j U_j\right)\right],
\end{aligned}$$

where the sum extends over the $(N)_n$ n-permutations $(k_1,k_2,\dots,k_n)$ of $\{1,2,\dots,N\}$ and the limit is easily verified by the Riemann integral. This implies that $\sum_{j=1}^n c_j U_{j,N} \to_w \sum_{j=1}^n c_j U_j$, which in turn yields

$$(U_{1,N},U_{2,N},\dots,U_{n,N}) \to_w (U_1,U_2,\dots,U_n) \tag{6.1}$$

by the Cramér–Wold device. For $u\in(0,1)$, let $F^{-1}(u) = \inf\{x\colon F(x) \geqslant u\}$ be the left-continuous inverse of F. Since $F^{-1}(U_j)$ is distributed like X_j, it follows that F^{-1} belongs to $L^2(0,1)$ and the Lebesgue integrals are

$$\mu = \int_0^1 F^{-1}(u)\,\mathrm{d}u \quad \text{and} \quad \mu^2+\sigma^2 = \int_0^1 (F^{-1}(u))^2\,\mathrm{d}u.$$

On the other hand, both F^{-1} and $(F^{-1})^2$ are almost everywhere continuous in $(0,1)$, and therefore, Riemann integrable; thus, as $N\to\infty$,

$$\mu_N{:=}\mathbb{E}[F^{-1}(U_{j,N})] = \frac{1}{N}\sum_{k=1}^N F^{-1}(k/(N+1)) \to \mu, \quad \text{and}$$

$$\sigma_N^2+\mu_N^2{:=}\mathbb{E}[(F^{-1}(U_{j,N}))^2] = \frac{1}{N}\sum_{k=1}^N (F^{-1}(k/(N+1))^2 \to \mu^2+\sigma^2.$$

By Skorohod's Theorem in $\mathbb{R}^n$ (see, for example, Billingsley, 1986, Theorem 29.6) and (6.1) it follows that

$$(X_{1,N},X_{2,N},\dots,X_{n,N}) \to_w (F^{-1}(U_1),F^{-1}(U_2),\dots,F^{-1}(U_n)) \stackrel{\mathrm{d}}{=} (X_1,X_2,\dots,X_n),$$

where $X_{j,N}{:=}(\sigma/\sigma_N)F^{-1}(U_{j,N})+(\sigma/\sigma_N)(\mu-\mu_N)+\mu(1-\sigma/\sigma_N)$, $j=1,2,\dots,n$. Obviously, $\mathbb{E}[X_{j,N}]=\mu$ and $\mathrm{Var}[X_{j,N}]=\sigma^2$. Moreover, it is easily verified that the random vector $(X_{j,N},\ j=1,2,\dots,n)$ has the same distribution as a without replacement sample of size n from $\Pi_N=\{\mu+\sigma(F^{-1}(k/(N+1))-\mu_N)/\sigma_N,\ k=1,2,\dots,N\}$, completing the proof.

From another point of view, Arnold and Groeneveld (1979) obtained expectation bounds for the order statistics in the completely general case, where the n-variate d.f. of the random vector $(X_1,X_2,\dots,X_n)$, is allowed to be completely arbitrary. They proved, for example, that, under the assumptions $\mathbb{E}[X_j]=\mu$ and $\mathrm{Var}[X_j]=\sigma^2$ for all $j\in\{1,2,\dots,n\}$,

$$\left|\sum_{j=1}^n \lambda_j\left(\mathbb{E}[X_{j:n}]-\mu\right)\right| \leqslant \sigma\sqrt{n}\left(\sum_{j=1}^n\left(\lambda_j-\lambda\right)^2\right)^{1/2},$$

where $\lambda_1,\lambda_2,\dots,\lambda_n$ are arbitrary scalars and $\lambda=n^{-1}\sum_{j=1}^n\lambda_j$. In particular, using the obvious inequalities $(X_{1:n}+\cdots+X_{i:n})/i \leqslant X_{i:n} \leqslant (X_{i:n}+\cdots+X_{n:n})/(n-i+1)$ it follows that (compare with Example 3.2)

$$\mu-\sigma\sqrt{\frac{n-i}{i}} \leqslant \mathbb{E}[X_{i:n}] \leqslant \mu+\sigma\sqrt{\frac{i-1}{n-i+1}}, \quad 1\leqslant i\leqslant n.$$

Because of its generality, Arnold and Groeneveld's bound has many important applications in robustness (for example, the application to Downton–Gini estimator of the normal standard deviation, given by the above authors). Despite its generality, however, this result is derived by taking expectations to a corresponding deterministic inequality (as in Example 3.2), and therefore, the case of equality corresponds to a trivial (exhaustive) sample from a finite population. For the bounds discussed here, however, this is not true in general; the optimal random samples (that attain the equalities in the bounds) correspond to order statistics with non-zero variances, except in some very particular cases (for example, when $n = N$).

References

Arnold, B.C., 1985. p-Norm bounds on the expectation of the maximum of possibly dependent sample. J. Multivariate Anal. 17, 316–332.

Arnold, B.C., Balakrishnan, N., 1989. Relations, bounds and approximations for order statistics. Lecture Notes in Statistics, Vol. 53. Springer, New York.

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

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

Balakrishnan, A.V., 1981. Applied Functional Analysis, 2nd Edition. Springer, New York.

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

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

Billingsley, P., 1986. Probability and Measure, 2nd Edition. Wiley, New York.

Boland, P.J., Hollander, M., Joag-Dev, K., Kochar, S., 1996. Bivariate dependence properties of order statistics. J. Multivariate Anal. 56, 75–89.

Boyd, A.V., 1971. Bounds on order statistics. Publications of the Electrotechnical Faculty of Belgrade University, Mathematics and Physics Series, Vol. 365, pp. 31–32.

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

Fahmy, S., Proschan, F., 1981. Bounds on differences of order statistics. Amer. Statist. 35, 46–47.

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

Hájek, J., 1981. Sampling from a Finite Population. Marcel Dekker, New York.

Hartley, H.O., David, H.A., 1954. Universal bounds for mean range and extreme observations. Ann. Math. Statist. 25, 85–99.

Hawkins, D.M., 1971. On the bounds of the range of order statistics. J. Amer. Statist. Assoc. 66, 644–645.

Kochar, S.C., Korwar, R., 1997. Stochastic properties of order statistics from finite populations. Sankhyā Ser. A 59, 102–116.

Mallows, C.L., Richter, D., 1969. Inequalities of Chebychev type involving conditional expectations. Ann. Math. Statist. 40, 1922–1932.

Moriguti, S., 1951. Extremal properties of extreme value distributions. Ann. Math. Statist. 22, 523–536.

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

Olkin, I., 1992. A matrix formulation on how deviant an observation can be. Amer. Statist. 46, 205–209.

Papadatos, N., 1997. Exact bounds for the expectations of order statistics from non-negative populations. Ann. Inst. Statist. Math. 49, 727–736.

Papadatos, N., 1999. Upper bound for the covariance of extreme order statistics from a sample of size three. Sankhyā Ser. A 61, 229–240.

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

Plackett, R.L., 1947. Limits of the ratio of mean range to standard deviation. Biometrika 34, 120–122.
Rychlik, T., 1993. Sharp bounds on *L*-estimates and their expectations for dependent samples. Commun. Statist. Theory Methods 22, 1053–1068 (Erratum 23, 305–306).
Rychlik, T., 1998. Bounds for expectations of *L*-estimates. In: Balakrishnan, N., Rao, C.R. (Eds.), Handbook of Statistics—16, Order Statistics: Theory and Methods. North-Holland, Amsterdam, pp. 105–145.
Scott, J.M.C., 1936. Appendix to paper by Pearson and Chandra Sekar. Biometrika 28, 319–320.
Takahasi, K., Futatsuya, M., 1998. Dependence between order statistics in samples from a finite population and its application to ranked set sampling. Ann. Inst. Statist. Math. 50, 49–70.
Terrell, G.R., 1983. A characterization of rectangular distributions. Ann. Probab. 11, 823–826.
Thomson, G.W., 1955. Bounds for the ratio of range to standard deviation. Biometrika 42, 268–269.
Wolkowicz, H., Styan, G.P.H., 1979. Extensions of Samuelson's inequality. Amer. Statist. 33, 143–144.