Permanents, Order Statistics, Outliers, and Robustness

N. BALAKRISHNAN

Department of Mathematics and Statistics
McMaster University
Hamilton, Ontario, Canada L8S 4K1
bala@mcmaster.ca

ABSTRACT

In this paper, we consider order statistics and outlier models, and focus primarily on multiple-outlier models and associated robustness issues. We first synthesise recent developments on order statistics arising from independent and non-identically distributed random variables based primarily on the theory of permanents. We then highlight various applications of these results in evaluating the robustness properties of several linear estimators when multiple outliers are possibly present in the sample.

Key words: order statistics, permanents, log-concavity, outliers, single-outlier model, multiple-outlier model, recurrence relations, robust estimators, sensitivity, bias, mean square error, location-outlier, scale-outlier, censoring, progressive Type-II censoring, ranked set sampling.

2000 Mathematics Subject Classification: 62E15, 62F10, 62F35, 62G30, 62G35, 62N01.

Introduction

Order statistics and their properties have been studied rather extensively since the early part of the last century. Yet, most of these studies focused only on the case when order statistics are from independent and identically distributed (IID) random variables. Motivated by robustness issues, studies of order statistics from outlier models began in early 70s. Though much of the early work in this direction concentrated only on the case when there is one outlier in the sample (single-outlier model), there has been a lot of work during the past fifteen years or so on multiple-outlier

models and more generally on order statistics from independent and non-identically distributed (INID) random variables. These results have also enabled useful and interesting discussions on the robustness of different estimators of parameters of a wide range of distributions.

These generalizations, of course, required the use of special methods and techniques. Since the book by Barnett and Lewis [43] has authoritatively covered the developments on the single-outlier model, we focus our attention here primarily on the multiple-outlier model which is quite often handled as a special case in the INID framework. We present many results on order statistics from multiple-outlier models and illustrate their use in robustness studies. We also point out some unresolved issues as open problems at a number of places which hopefully would perk the interest of some readers!

1. Order statistics from IID variables

Let $X_1, \ldots, X_n$ be IID random variables from a population with cumulative distribution function $F(x)$ and probability density function $f(x)$. Let $X_{1:n} < X_{2:n} < \cdots < X_{n:n}$ be the order statistics obtained by arranging the n X_i's in increasing order of magnitude. Then, the distribution function of $X_{r:n}$ $(1 \le r \le n)$ is

$$F_{r:n}(x) = \Pr(\text{at least } r \text{ of the } n \text{ } X\text{'s are at most } x)$$
$$= \sum_{i=r}^{n} \Pr(\text{exactly } i \text{ of the } n \text{ } X\text{'s are at most } x)$$
$$= \sum_{i=r}^{n} \binom{n}{i} \{F(x)\}^i \{1 - F(x)\}^{n-i}, \quad x \in \mathbb{R}. \tag{1}$$

Using the identity, obtained by repeated integration by parts,

$$\sum_{i=r}^{n} \binom{n}{i} \{F(x)\}^i \{1 - F(x)\}^{n-i} = \int_0^{F(x)} \frac{n!}{(r-1)!(n-r)!} t^{r-1} (1-t)^{n-r} \, dt,$$

we readily obtain from (1) the density function of $X_{r:n}$ $(1 \le r \le n)$ as

$$f_{r:n}(x) = \frac{n!}{(r-1)!(n-r)!} \{F(x)\}^{r-1} \{1 - F(x)\}^{n-r} f(x), \quad x \in \mathbb{R}. \tag{2}$$

The density function of $X_{r:n}$ $(1 \le r \le n)$ in (2) can also be derived using multinomial argument as follows. Consider the event $(x < X_{r:n} \le x + \Delta x)$. Then,

$$\Pr(x < X_{r:n} \le x + \Delta x)$$
$$= \frac{n!}{(r-1)!(n-r)!} \{F(x)\}^{r-1} \{F(x + \Delta x) - F(x)\} \{1 - F(x + \Delta x)\}^{n-r} + O((\Delta x)^2),$$

where $O((\Delta x)^2)$ denotes terms of higher order (corresponding to more than one of the X_i's falling in the interval $(x, x + \Delta x]$), and so

$$
\begin{aligned}
f_{r:n}(x) &= \lim_{\Delta x \downarrow 0} \frac{\Pr(x < X_{r:n} \leq x + \Delta x)}{\Delta x} \\
&= \frac{n!}{(r-1)!(n-r)!} \{F(x)\}^{r-1} \{1 - F(x)\}^{n-r} f(x).
\end{aligned}
$$

Proceeding similarly, we obtain the joint density of $X_{r:n}$ and $X_{s:n}$ $(1 \leq r < s \leq n)$ as

$$
\begin{aligned}
f_{r,s:n}(x, y) = \frac{n!}{(r-1)!(s-r-1)!(n-s)!} &\{F(x)\}^{r-1} \{F(y) - F(x)\}^{s-r-1} \\
&\times \{1 - F(y)\}^{n-s} f(x) f(y), \qquad -\infty < x < y < \infty. \quad (3)
\end{aligned}
$$

The single and product moments of order statistics can be obtained from (2) and (3) by integration. This computation has been carried out for numerous distributions and for a list of available tables, one may refer to the books in [57, 66].

The area of order statistics has had a long and rich history. While the book in [6] provides an introduction to this area, the books in [53, 57] provide comprehensive reviews on various developments on order statistics. The books in [25, 66] describe various inferential methods based on order statistics. The two volumes in [35, 36] highlight many methodological and applied aspects of order statistics. Order statistics have especially found key applications in parametric inference, nonparametric inference and robust inference.

In this paper, we synthesise some recent advances on order statistics from INID random variables and pay special emphasis to results on order statistics from single-outlier and multiple-outlier models, and then illustrate their applications in the robust estimation of parameters of different distributions. It is important to mention here, however, that some developments on topics such as inequalities, stochastic orderings, and characterizations that are not directly relevant to the present discussion on outliers and robustness have not been stressed in this article.

2. Order statistics from a single-outlier model and robust estimation for normal distribution

2.1. Introduction

The distributions of order statistics presented in the last section, though simple in form, become quite complicated once the assumption of identical distribution of the random variables is lost. A well-known case in this scenario is the single-outlier model wherein $X_1, \ldots, X_n$ are independent random variables with $X_1, \ldots, X_{n-1}$ being from a population with cumulative distribution function $F(x)$ and probability density function $f(x)$ and X_n being an outlier from a different population with cumulative

distribution function $G(x)$ and probability density function $g(x)$. As before, let $X_{1:n} \leq \cdots \leq X_{n:n}$ denote the order statistics obtained from this single-outlier model.

2.2. Distributions of order statistics

By using multinomial arguments and accounting for the fact that the outlier X_n may fall in any of the three intervals $(-\infty, x]$, $(x, x + \Delta x]$ and $(x + \Delta x, \infty)$, the density function of $X_{r:n}$ $(1 \leq r \leq n)$ can be obtained as (see [5, 43, 58])

$$f_{r:n}(x) = \frac{(n-1)!}{(r-2)!(n-r)!} \{F(x)\}^{r-2} G(x) f(x) \{1 - F(x)\}^{n-r}$$
$$+ \frac{(n-1)!}{(r-1)!(n-r)!} \{F(x)\}^{r-1} g(x) \{1 - F(x)\}^{n-r}$$
$$+ \frac{(n-1)!}{(r-1)!(n-r-1)!} \{F(x)\}^{r-1} f(x) \{1 - F(x)\}^{n-r-1} \{1 - G(x)\},$$
$$x \in \mathbb{R}, \quad (4)$$

where the first and last terms vanish when $r = 1$ and $r = n$, respectively. Proceeding similarly, the joint density function of $X_{r:n}$ and $X_{s:n}$ $(1 \leq r < s \leq n)$ can be expressed as

$$f_{r,s:n}(x, y)$$
$$= \frac{(n-1)!}{(r-2)!(s-r-1)!(n-s)!} \{F(x)\}^{r-2} G(x) f(x) \{F(y) - F(x)\}^{s-r-1}$$
$$\times f(y) \{1 - F(y)\}^{n-s}$$
$$+ \frac{(n-1)!}{(r-1)!(s-r-1)!(n-s)!} \{F(x)\}^{r-1} g(x) \{F(y) - F(x)\}^{s-r-1}$$
$$\times f(y) \{1 - F(y)\}^{n-s}$$
$$+ \frac{(n-1)!}{(r-1)!(s-r-2)!(n-s)!} \{F(x)\}^{r-1} f(x) \{F(y) - F(x)\}^{s-r-2}$$
$$\times \{G(y) - G(x)\} f(y) \{1 - F(y)\}^{n-s}$$
$$+ \frac{(n-1)!}{(r-1)!(s-r-1)!(n-s)!} \{F(x)\}^{r-1} f(x) \{F(y) - F(x)\}^{s-r-1}$$
$$\times g(y) \{1 - F(y)\}^{n-s}$$
$$+ \frac{(n-1)!}{(r-1)!(s-r-1)!(n-s-1)!} \{F(x)\}^{r-1} f(x) \{F(y) - F(x)\}^{s-r-1}$$
$$\times f(y) \{1 - F(y)\}^{n-s-1} \{1 - G(y)\}, \qquad -\infty < x < y < \infty, \qquad (5)$$

where the first, middle and last terms vanish when $r = 1$, $s = r + 1$, and $s = n$, respectively.

2.3. Moments of order statistics

The single and product moments of order statistics in this case need to be obtained by integration from (4) and (5), respectively. Except in a few cases like the exponential distribution, the required integrations need to be done by numerical methods, and as is evident from the expressions in (4) and (5) this may be computationally very demanding. For example, in the case of the normal distribution, the required computations were carried out in [56] for the two cases:

(i) Location-outlier model:

$$X_1, \ldots, X_{n-1} \stackrel{d}{=} N(0,1) \qquad \text{and} \qquad X_n \stackrel{d}{=} N(\lambda, 1),$$

(ii) Scale-outlier model:

$$X_1, \ldots, X_{n-1} \stackrel{d}{=} N(0,1) \qquad \text{and} \qquad X_n \stackrel{d}{=} N(0, \tau^2).$$

The values of means, variances and covariances of order statistics for sample sizes up to 20 for different choices of λ and τ were all tabulated in [56].

2.4. Robust estimation for normal distribution

By using the tables in [56], detailed robustness examination has been carried out in [5, 58] on various linear estimators of the normal mean, which included

(i) Sample mean:

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_{i:n};$$

(ii) Trimmed means:

$$T_n(r) = \frac{1}{n-2r} \sum_{i=r+1}^{n-r} X_{i:n};$$

(iii) Winsorized means:

$$W_n(r) = \frac{1}{n} \left[\sum_{i=r+2}^{n-r-1} X_{i:n} + (r+1)[X_{r+1:n} + X_{n-r:n}] \right];$$

(iv) Modified maximum likelihood estimators:

$$M_n(r) = \frac{1}{m} \left[\sum_{i=r+2}^{n-r-1} X_{i:n} + (1+r\beta)[X_{r+1:n} + X_{n-r:n}] \right],$$

where $m = n - 2r + 2r\beta$;

| | | | | λ | | | |
Estimator	0.0	0.5	1.0	1.5	2.0	3.0	4.0	∞
$\bar{X}_{10}$	0.0	0.05000	0.10000	0.15000	0.20000	0.30000	0.40000	∞
$T_{10}(1)$	0.0	0.04912	0.09325	0.12870	0.15400	0.17871	0.18470	0.18563
$T_{10}(2)$	0.0	0.04869	0.09023	0.12041	0.13904	0.15311	0.15521	0.15538
Med_{10}	0.0	0.04832	0.08768	0.11381	0.12795	0.13642	0.13723	0.13726
$W_{10}(1)$	0.0	0.04938	0.09506	0.13368	0.16298	0.19407	0.20239	0.20377
$W_{10}(2)$	0.0	0.04889	0.09156	0.12389	0.14497	0.16217	0.16504	0.16530
$M_{10}(1)$	0.0	0.04934	0.09484	0.13311	0.16194	0.19229	0.20037	0.20169
$M_{10}(2)$	0.0	0.04886	0.09137	0.12342	0.14418	0.16091	0.16369	0.16394
$L_{10}(1)$	0.0	0.04869	0.09024	0.12056	0.13954	0.15459	0.15727	0.15758
$L_{10}(2)$	0.0	0.04850	0.08892	0.11700	0.13328	0.14436	0.14576	0.14585
G_{10}	0.0	0.04847	0.08873	0.11649	0.13237	0.14285	0.14407	0.14414

Table 1 – Bias of various estimators of μ for $n = 10$ when a single outlier is from $N(\mu + \lambda, 1)$ and the others from $N(\mu, 1)$

(v) Linearly weighted means:

$$L_n(r) = \frac{1}{2\left(\frac{n}{2} - r\right)^2} \sum_{i=1}^{\frac{n}{2}-r} (2i - 1)[X_{r+i:n} + X_{n-r-i+1:n}]$$

for even values of n;

(vi) Gastwirth mean:

$$G_n = 0.3(X_{[\frac{n}{3}]+1:n} + X_{n-[\frac{n}{3}]:n}) + 0.2(X_{\frac{n}{2}:n} + X_{\frac{n}{2}+1:n})$$

for even values of n, where $\left[\frac{n}{3}\right]$ denotes the integer part of $\frac{n}{3}$.

By making use of the tables of means, variances, and covariances of order statistics from a single location-outlier normal model presented in [56], bias and mean square error of all these estimators were computed and are presented in tables 1 and 2, respectively for $n = 10$. From these tables, we observe that though median gives the best protection against the presence of outlier in terms of bias, it comes at the cost of a higher mean square error than some other robust estimators. The trimmed mean, linearly weighted mean and the modified maximum likelihood estimator turn out to be quite robust and efficient in general.

In table 3, similar results are presented for a single scale-outlier normal model. In this case, since all the estimators considered are unbiased, comparisons are made only in terms of variance, and similar conclusions are reached.

Remark 2.1. It is clear from (4) and (5) that analysis of multiple-outlier models in this direct approach would become extremely difficult if not impossible! For example, if we allow two outliers in the sample, the marginal density of $X_{r:n}$ will have 5 terms while

| | λ | | | | | | | |
Estimator	0.0	0.5	1.0	1.5	2.0	3.0	4.0	∞
$\bar{X}_{10}$	0.10000	0.10250	0.11000	0.12250	0.14000	0.19000	0.26000	∞
$T_{10}(1)$	0.10534	0.10791	0.11471	0.12387	0.13285	0.14475	0.14865	0.14942
$T_{10}(2)$	0.11331	0.11603	0.12297	0.13132	0.13848	0.14580	0.14730	0.14745
Med_{10}	0.13833	0.14161	0.14964	0.15852	0.16524	0.17072	0.17146	0.17150
$W_{10}(1)$	0.10437	0.10693	0.11403	0.12405	0.13469	0.15039	0.15627	0.15755
$W_{10}(2)$	0.11133	0.11402	0.12106	0.12995	0.13805	0.14713	0.14926	0.14950
$M_{10}(1)$	0.10432	0.10688	0.11396	0.12385	0.13430	0.14950	0.15513	0.15581
$M_{10}(2)$	0.11125	0.11395	0.12097	0.12974	0.13770	0.14649	0.14853	0.14876
$L_{10}(1)$	0.11371	0.11644	0.12337	0.13169	0.13882	0.14626	0.14797	0.14820
$L_{10}(2)$	0.12097	0.12386	0.13105	0.13933	0.14598	0.15206	0.15310	0.15318
G_{10}	0.12256	0.12549	0.13276	0.14111	0.14777	0.15376	0.15472	0.15479

Table 2 – Mean square error of various estimators of μ for $n = 10$ when a single outlier is from $N(\mu + \lambda, 1)$ and the others from $N(\mu, 1)$

| | τ | | | | | |
Estimator	0.5	1.0	2.0	3.0	4.0	∞
$\bar{X}_{10}$	0.09250	0.10000	0.13000	0.18000	0.25000	∞
$T_{10}(1)$	0.09491	0.10534	0.12133	0.12955	0.13417	0.14942
$T_{10}(2)$	0.09953	0.11331	0.12773	0.13389	0.13717	0.14745
Med_{10}	0.11728	0.13833	0.15375	0.15953	0.16249	0.17150
$W_{10}(1)$	0.09571	0.10437	0.12215	0.13221	0.13801	0.15754
$W_{10}(2)$	0.09972	0.11133	0.12664	0.13365	0.13745	0.14950
$M_{10}(1)$	0.09548	0.10432	0.12187	0.13171	0.13735	0.15581
$M_{10}(2)$	0.09940	0.11125	0.12638	0.13328	0.13699	0.14876
$L_{10}(1)$	0.09934	0.11371	0.12815	0.13436	0.13769	0.14820
$L_{10}(2)$	0.10432	0.12097	0.13531	0.14101	0.14398	0.15318
G_{10}	0.10573	0.12256	0.13703	0.14270	0.14565	0.15479

Table 3 – Variance of various estimators of μ for $n = 10$ when a single outlier is from $N(\mu, \tau^2)$ and the others from $N(\mu, 1)$

the joint density of $(X_{r:n}, X_{s:n})$ will have 13 terms. For this reason, majority of such work in the outlier literature have dealt with only the single-outlier model case; see [43]. Therefore, special tools and techniques are needed to deal with multiple-outlier models as will be demonstrated in subsequent sections.

3. Permanents

3.1. Introduction

The permanent function was introduced by Binet and Cauchy (independently) as early as in 1812, more or less simultaneously with the determinant function. The famous conjecture posed by van der Waerden [91] concerning the minimum permanent over the set of doubly stochastic matrices was primarily responsible for attracting the attention of numerous mathematicians towards the theory of permanents. van der Waerden's conjecture was finally solved by Egorychev, and independently by Falikman, around 1980. This resulted in an increased activity in this area as it is clearly evident from the expository book on permanents by Minč [78] and the two subsequent survey papers [79, 80]. These works will make excellent sources of reference for any reader interested in the theory of permanents.

Suppose $\boldsymbol{A} = ((a_{i,j}))$ is a square matrix of order n. Then, the permanent of the matrix $\boldsymbol{A}$ is defined to be

$$\operatorname{Per} \boldsymbol{A} = \sum_{P} \prod_{j=1}^{n} a_{j,i_j}, \tag{6}$$

where $\sum_{P}$ denotes the sum over all $n!$ permutations $(i_1, i_2, \ldots, i_n)$ of $(1, 2, \ldots, n)$. The definition of the permanent in (6) is thus similar to that of the determinant except that it does not have the alternating sign (depending on whether the permutation is of even or odd order). Consequently, it is not surprising to see the following basic properties of permanents.

Property 3.1. Per $\boldsymbol{A}$ *is unchanged if the rows or columns of* $\boldsymbol{A}$ *are permuted.*

Property 3.2. *If* $\boldsymbol{A}(i,j)$ *denotes the sub-matrix of order* $n-1$ *obtained from* $\boldsymbol{A}$ *by deleting the* i*-th row and the* j*-th column, then*

$$\operatorname{Per} \boldsymbol{A} = \sum_{i=1}^{n} a_{i,j} \operatorname{Per} \boldsymbol{A}(i,j), \qquad j = 1, 2, \ldots, n$$

$$= \sum_{j=1}^{n} a_{i,j} \operatorname{Per} \boldsymbol{A}(i,j), \qquad i = 1, 2, \ldots, n.$$

That is, the permanent of a matrix can be expanded by any row or column.

Property 3.3. *If* $\boldsymbol{A}^*$ *denotes the matrix obtained from* $\boldsymbol{A}$ *simply by replacing the elements in the i-th row by* $c\,a_{i,j}, \ j = 1, 2, \ldots, n,$ *then*

$$\operatorname{Per} \boldsymbol{A}^* = c \operatorname{Per} \boldsymbol{A}.$$

Property 3.4. *If* $\boldsymbol{A}^{**}$ *denotes the matrix obtained from* $\boldsymbol{A}$ *by replacing the elements in the i-th row by* $a_{i,j} + b_{i,j}$ $(j = 1, 2, \ldots, n)$ *and* $\boldsymbol{A}^*$ *the matrix obtained from* $\boldsymbol{A}$ *by replacing the elements in the i-th row by* $b_{i,j}$ $(j = 1, 2, \ldots, n)$, *then*

$$\operatorname{Per} \boldsymbol{A}^{**} = \operatorname{Per} \boldsymbol{A} + \operatorname{Per} \boldsymbol{A}^*.$$

Due to the absence of the alternating sign in (6), the permanent of a matrix in which two or more rows (or columns) are repeated *need not be zero* (unlike in the case of a determinant). Let us use

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \cdot & \cdot & \cdots & \cdot \end{pmatrix} \begin{matrix} \} \, i_1 \\ \} \, i_2 \\ \end{matrix}$$

to denote a matrix in which the first row is repeated i_1 times, the second row is repeated i_2 times, and so on.

3.2. Log-concavity

An interesting and important result in the theory of permanents of non-negative matrices is the Alexandroff inequality. This result, as illustrated in [40], is useful in establishing the log-concavity of distribution functions of order statistics.

For the benefit of readers, we present below a brief introduction to log-concavity and some related properties. A sequence of non-negative numbers $\alpha_1, \alpha_2, \ldots, \alpha_n$ is said to be log-concave if $\alpha_i^2 \geq \alpha_{i-1}\,\alpha_{i+1}$ $(i = 2, 3, \ldots, n-1)$. The following lemma presents a number of elementary properties of such log-concave sequences.

Lemma 3.5. *Let* $\alpha_1, \alpha_2, \ldots, \alpha_n$ *and* $\beta_1, \beta_2, \ldots, \beta_n$ *be two log-concave sequences. Then the following statements hold:*

(i) *If* $\alpha_i > 0$ *for* $i = 1, 2, \ldots, n,$ *then*

$$\frac{\alpha_i}{\alpha_{i-1}} \geq \frac{\alpha_{i+1}}{\alpha_i}, \qquad i = 2, \ldots, n-1;$$

   that is, α_i/α_{i-1} *is non-increasing in* i.

(ii) *If* $\alpha_i > 0$ *for* $i = 1, 2, \ldots, n,$ *then* $\alpha_1, \alpha_2, \ldots, \alpha_n$ *is unimodal; that is,*

$$\alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_k \geq \alpha_{k+1} \geq \cdots \geq \alpha_n$$

   for some k *($1 \leq k \leq n$)*.

(iii) *The sequence* $\alpha_1 \beta_1, \alpha_2 \beta_2, \ldots, \alpha_n \beta_n$ *is log-concave.*

(iv) *The sequence* $\gamma_1, \gamma_2, \ldots, \gamma_n$ *is log-concave, where*

$$\gamma_k = \sum_{i=1}^{k} \alpha_i \, \beta_{k+1-i}, \qquad k = 1, 2, \ldots, n.$$

(v) *The sequences* $\alpha_1, \alpha_1 + \alpha_2, \ldots, \sum_{i=1}^{n} \alpha_i$ *and* $\alpha_n, \alpha_{n-1} + \alpha_n, \ldots, \sum_{i=1}^{n} \alpha_i$ *are both log-concave.*

(vi) *The sequence of combinatorial coefficients* $\binom{n}{i}$, $i = 0, 1, \ldots, n$, *is log-concave.*

Proof. (i), (iii), and (vi) are easily verified.

Since from (i)

$$\frac{\alpha_2}{\alpha_1} \geq \frac{\alpha_3}{\alpha_2} \geq \cdots \geq \frac{\alpha_n}{\alpha_{n-1}}$$

and that there must exist some k $(1 \leq k \leq n)$ such that

$$\frac{\alpha_2}{\alpha_1} \geq \cdots \geq \frac{\alpha_k}{\alpha_{k-1}} \geq 1 \geq \frac{\alpha_{k+1}}{\alpha_k} \geq \cdots \geq \frac{\alpha_n}{\alpha_{n-1}},$$

(ii) follows.

(iv) may be proved directly by showing $\gamma_i^2 \geq \gamma_{i-1}\gamma_{i+1}$ after carefully pairing terms on both sides of the inequality.

The first part of (v) follows from (iv) simply by taking $\beta_i = 1$ for $i = 1, 2, \ldots, n$. Since $\alpha_n, \alpha_{n-1}, \ldots, \alpha_1$ is log-concave, the second part of (v) follows immediately. $\square$

Interested readers may refer to the classic book on inequalities in [65] for an elaborate treatment on log-concavity.

Now, we shall simply state the Alexandroff inequality for permanents of non-negative matrices and refer the readers to [92] for an elegant proof.

Theorem 3.6. *Let*

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{a}_1 \\ \vdots \\ \boldsymbol{a}_n \end{pmatrix}$$

be a non-negative square matrix of order n. Then,

$$(\operatorname{Per} \boldsymbol{A})^2 \geq \operatorname{Per} \left. \begin{pmatrix} \boldsymbol{a}_1 \\ \vdots \\ \boldsymbol{a}_{n-2} \\ \boldsymbol{a}_{n-1} \end{pmatrix} \right\} 2 \quad \operatorname{Per} \left. \begin{pmatrix} \boldsymbol{a}_1 \\ \vdots \\ \boldsymbol{a}_{n-2} \\ \boldsymbol{a}_n \end{pmatrix} \right\} 2$$

Remark 3.7. The above given Alexandroff's inequality was proved in [1] for a general function called the "mixed discriminant" and, as a matter of fact, there is no mention of permanents even in the paper. After almost forty years, Egorychev realized that the result, when specialized to a permanental inequality, is what is needed to prove the van der Waerden conjecture. As mentioned earlier, this inequality will be used later on to establish the log-concavity of distribution functions of order statistics arising from INID random variables.

Theorem 3.8 (Newton's theorem). *If $b_1, b_2, \ldots, b_n$ are all real and if*

$$\prod_{i=1}^{n}(x + b_i) = \sum_{r=0}^{n} \binom{n}{r} \alpha_r \, x^r,$$

then $\alpha_r^2 \geq \alpha_{r-1}\alpha_{r+1}$ for $1 \leq r \leq n - 1$; that is, $\alpha_0, \alpha_1, \ldots, \alpha_n$ form a log-concave sequence.

For a proof, interested readers may refer to [65, pp. 51–52].

4. Order statistics from INID variables

4.1. Distributions and joint distributions

Let $X_1, X_2, \ldots, X_n$ be independent random variables with X_i having cumulative distribution function $F_i(x)$ and probability density function $f_i(x)$. Let $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ be the order statistics obtained from the above n variables. Then, for deriving the density function of $X_{r:n}$, let us consider

$$\Pr(x < X_{r:n} \leq x + \Delta x)$$
$$= \frac{1}{(r-1)!(n-r)!} \sum_{P} F_{i_1}(x) \cdots F_{i_{r-1}}(x)\{F_{i_r}(x + \Delta x) - F_{i_r}(x)\}$$
$$\times \{1 - F_{i_{r+1}}(x + \Delta x)\} \cdots \{1 - F_{i_n}(x + \Delta x)\} + O((\Delta x)^2), \quad (7)$$

where $\sum_P$ denotes the sum over all $n!$ permutations $(i_1, i_2, \ldots, i_n)$ of $(1, 2, \ldots, n)$. Dividing both sides of (7) by Δx and then letting Δx tend to zero, we obtain the density function of $X_{r:n}$ $(1 \leq r \leq n)$ as

$$f_{r:n}(x) = \frac{1}{(r-1)!(n-r)!} \sum_{P} F_{i_1}(x) \cdots F_{i_{r-1}}(x) \, f_{i_r}(x)$$
$$\times \{1 - F_{i_{r+1}}(x)\} \cdots \{1 - F_{i_n}(x)\}, \qquad x \in \mathbb{R}. \quad (8)$$

From (8) and (6), we then readily see that the density function of $X_{r:n}$ $(1 \leq r \leq n)$ can be written as

$$f_{r:n}(x) = \frac{1}{(r-1)!(n-r)!} \operatorname{Per} \boldsymbol{A}_1, \quad x \in \mathbb{R}, \quad (9)$$

where

$$
\boldsymbol{A}_1 = \begin{pmatrix} F_1(x) & F_2(x) & \cdots & F_n(x) \\ f_1(x) & f_2(x) & \cdots & f_n(x) \\ 1 - F_1(x) & 1 - F_2(x) & \cdots & 1 - F_n(x) \end{pmatrix} \begin{matrix} \} \, r-1 \\ \} \, 1 \\ \} \, n-r \end{matrix} . \tag{10}
$$

The permanent representation of $f_{r:n}(x)$ in (9) is originally due to [93].

Similarly, for deriving the joint density function of $X_{r:n}$ and $X_{s:n}$ $(1 \le r < s \le n)$, let us consider

$$
\Pr(x < X_{r:n} \le x + \Delta x, \ y < X_{s:n} \le y + \Delta y)
$$
$$
= \frac{1}{(r-1)!(s-r-1)!(n-s)!} \sum_P F_{i_1}(x) \cdots F_{i_{r-1}}(x)\{F_{i_r}(x + \Delta x) - F_{i_r}(x)\}
$$
$$
\times \{F_{i_{r+1}}(y) - F_{i_{r+1}}(x + \Delta x)\} \cdots \{F_{i_{s-1}}(y) - F_{i_{s-1}}(x + \Delta x)\}
$$
$$
\times \{F_{i_s}(y + \Delta y) - F_{i_s}(y)\}\{1 - F_{i_{s+1}}(y + \Delta y)\} \cdots \{1 - F_{i_n}(y + \Delta y)\}
$$
$$
+ O((\Delta x)^2 \, \Delta y) + O(\Delta x \, (\Delta y)^2), \tag{11}
$$

where $O((\Delta x)^2 \, \Delta y)$ denotes terms of higher order corresponding to more than one of the X_i's falling in $(x, x + \Delta x]$ and exactly one in $(y, y + \Delta y]$, and $O(\Delta x \, (\Delta y)^2)$ corresponding to exactly one of the X_i's falling in $(x, x + \Delta x]$ and more than one in $(y, y + \Delta y]$. Dividing both sides of (11) by $\Delta x \Delta y$ and then letting both Δx and Δy tend to zero, we obtain the joint density function of $X_{r:n}$ and $X_{s:n}$ $(1 \le r < s \le n)$ as

$$
f_{r,s:n}(x,y) = \frac{1}{(r-1)!(s-r-1)!(n-s)!} \sum_P F_{i_1}(x) \cdots F_{i_{r-1}}(x) f_{i_r}(x)
$$
$$
\times \{F_{i_{r+1}}(y) - F_{i_{r+1}}(x)\} \cdots \{F_{i_{s-1}}(y) - F_{i_{s-1}}(x)\}
$$
$$
\times \ f_{i_s}(y)\{1 - F_{i_{s+1}}(y)\} \cdots \{1 - F_{i_n}(y)\}, \qquad -\infty < x < y < \infty. \tag{12}
$$

From (12) and (6), we readily see that the joint density function of $X_{r:n}$ and $X_{s:n}$ $(1 \le r < s \le n)$ can be written as

$$
f_{r,s:n}(x,y) = \frac{1}{(r-1)!(s-r-1)!(n-s)!} \ \text{Per} \, \boldsymbol{A}_2, \qquad -\infty < x < y < \infty, \tag{13}
$$

where

$$
\boldsymbol{A}_2 = \begin{pmatrix} F_1(x) & F_2(x) & \cdots & F_n(x) \\ f_1(x) & f_2(x) & \cdots & f_n(x) \\ F_1(y) - F_1(x) & F_2(y) - F_2(x) & \cdots & F_n(y) - F_n(x) \\ f_1(y) & f_2(y) & \cdots & f_n(y) \\ 1 - F_1(y) & 1 - F_2(y) & \cdots & 1 - F_n(y) \end{pmatrix} \begin{matrix} \} \, r-1 \\ \} \, 1 \\ \} \, s-r-1 \\ \} \, 1 \\ \} \, n-s \end{matrix} .
$$

Proceeding similarly, we can show that the joint density function of $X_{r_1:n}, X_{r_2:n}, \ldots, X_{r_k:n}$ $(1 \le r_1 < r_2 < \cdots < r_k \le n)$ can be written as

$$f_{r_1, r_2, \ldots, r_k}(x_1, x_2, \ldots, x_k)$$
$$= \frac{1}{(r_1 - 1)!(r_2 - r_1 - 1)! \cdots (r_k - r_{k-1} - 1)!(n - r_k)!} \text{ Per } \boldsymbol{A}_k,$$
$$-\infty < x_1 < x_2 < \cdots < x_k < \infty,$$

where

$$\boldsymbol{A}_k = \begin{pmatrix} F_1(x_1) & \cdots & F_n(x_1) \\ f_1(x_1) & \cdots & f_n(x_1) \\ F_1(x_2) - F_1(x_1) & \cdots & F_n(x_2) - F_n(x_1) \\ f_1(x_2) & \cdots & f_n(x_2) \\ \cdot & \cdots & \cdot \\ F_1(x_k) - F_1(x_{k-1}) & \cdots & F_n(x_k) - F_n(x_{k-1}) \\ f_1(x_k) & \cdots & f_n(x_k) \\ 1 - F_1(x_k) & \cdots & 1 - F_n(x_k) \end{pmatrix} \begin{matrix} \} \, r_1 - 1 \\ \} \, 1 \\ \} \, r_2 - r_1 - 1 \\ \} \, 1 \\ \\ \} \, r_k - r_{k-1} - 1 \\ \} \, 1 \\ \} \, n - r_k \end{matrix} \; .$$

Permanent expressions may also be presented for cumulative distribution functions of order statistics. For example, let us consider

$$F_{r:n}(x) = \Pr(X_{r:n} \le x)$$
$$= \sum_{i=r}^{n} \Pr(\text{exactly } i \text{ of } X\text{'s are } \le x)$$
$$= \sum_{i=r}^{n} \frac{1}{i!(n-i)!} \sum_{P} F_{j_1}(x) \cdots F_{j_i}(x) \{1 - F_{j_{i+1}}(x)\} \cdots \{1 - F_{j_n}(x)\}, \quad (14)$$

where $\sum_P$ denotes the sum over all $n!$ permutations $(j_1, j_2, \ldots, j_n)$ of $(1, 2, \ldots, n)$. From (14) and (6), we see that the cumulative distribution function of $X_{r:n}$ $(1 \le r \le n)$ can be written as

$$F_{r:n}(x) = \sum_{i=r}^{n} \frac{1}{i!(n-i)!} \text{ Per } \boldsymbol{B}_1, \qquad x \in \mathbb{R}, \qquad (15)$$

where

$$\boldsymbol{B}_1 = \begin{pmatrix} F_1(x) & F_2(x) & \cdots & F_n(x) \\ 1 - F_1(x) & 1 - F_2(x) & \cdots & 1 - F_n(x) \end{pmatrix} \begin{matrix} \} \, i \\ \} \, n - i \end{matrix}$$

The permanent form of $F_{r:n}(x)$ in (15) is due to [40]. It should be mentioned here that an equivalent expression for the cumulative distribution function of $X_{r:n}$ is (see [53, p. 22])

$$F_{r:n}(x) = \sum_{i=r}^{n} \sum_{P_i} \left[\prod_{\ell=1}^{i} F_{j_\ell}(x) \right] \left[\prod_{\ell=i+1}^{n} \{1 - F_{j_\ell}(x)\} \right], \qquad (16)$$

where $\sum_{P_i}$ denotes the sum over all permutations $(j_1, j_2, \ldots, j_n)$ of $(1, 2, \ldots, n)$ for which $j_1 < j_2 < \cdots < j_i$ and $j_{i+1} < j_{i+2} < \cdots < j_n$. Realize that $\sum_{P_i}$ includes $\binom{n}{i}$ terms in (16), while $\sum_P$ in (14) includes $n!$ terms; see [57].

Proceeding similarly, the joint cumulative distribution function of $X_{r_1:n}, X_{r_2:n}, \ldots, X_{r_k:n}$ $(1 \le r_1 < r_2 < \cdots < r_k \le n)$ can be written as

$$F_{r_1, r_2, \ldots, r_k:n}(x_1, x_2, \ldots, x_k) = \Pr(X_{r_1:n} \le x_1, X_{r_2:n} \le x_2, \ldots, X_{r_k:n} \le x_k)$$
$$= \sum \frac{1}{j_1!\, j_2! \cdots j_{k+1}!}\ \mathrm{Per}\ \boldsymbol{B}_k,$$
$$-\infty < x_1 < x_2 < \cdots < x_k < \infty, \quad (17)$$

where

$$\boldsymbol{B}_k = \begin{pmatrix} F_1(x_1) & \cdots & F_n(x_1) \\ F_1(x_2) - F_1(x_1) & \cdots & F_n(x_2) - F_n(x_1) \\ . & \cdots & . \\ F_1(x_k) - F_1(x_{k-1}) & \cdots & F_n(x_k) - F_n(x_{k-1}) \\ 1 - F_1(x_k) & \cdots & 1 - F_n(x_k) \end{pmatrix} \begin{matrix} \} j_1 \\ \} j_2 \\ \\ \} j_k \\ \} j_{k+1} \end{matrix}$$

and the sum is over $j_1, j_2, \ldots, j_{k+1}$ with $j_1 \ge r_1, j_1 + j_2 \ge r_2, \ldots, j_1 + j_2 + \cdots + j_k \ge r_k$ and $j_1 + j_2 + \cdots + j_{k+1} = n$.

Remark 4.1. If the condition $x_1 < x_2 < \cdots < x_k$ is not imposed in (17), then some of the inequalities among $X_{r_1:n} \le x_1, X_{r_2:n} \le x_2, \ldots, X_{r_k:n} \le x_k$ will be redundant, and the necessary probability can then be determined after making appropriate reductions.

4.2. Log-concavity

In this section, we shall establish the log-concavity of distribution functions of order statistics by making use of Alexandroff's inequality in Theorem 3.6. This interesting result, as first proved in [40], is presented in the following theorem.

Theorem 4.2. *Let* $X_{1:n} \le X_{2:n} \le \cdots \le X_{n:n}$ *denote the order statistics obtained from* n *INID variables with cumulative distribution functions* $F_1(x), F_2(x), \ldots, F_n(x)$. *Then, for fixed* x, *the sequences* $\{F_{r:n}(x)\}_{r=1}^n$ *and* $\{1 - F_{r:n}(x)\}_{r=1}^n$ *are both log-concave. If, further, the underlying variables are all continuous with respective densities* $f_1(x), f_2(x), \ldots, f_n(x)$, *then the sequence* $\{f_{r:n}(x)\}_{r=1}^n$ *is also log-concave.*

Proof. Let us denote, for $i = 1, 2, \ldots, n$,

$$\alpha_i = \mathrm{Per}\begin{pmatrix} F_1(x) & F_2(x) & \cdots & F_n(x) \\ 1 - F_1(x) & 1 - F_2(x) & \cdots & 1 - F_n(x) \end{pmatrix} \begin{matrix} \} i \\ \} n - i \end{matrix}.$$

Since the above square matrix is non-negative, a simple application of Alexandroff's inequality in Theorem 3.6 implies that

$$\alpha_i^2 \ge \alpha_{i-1}\, \alpha_{i+1}, \qquad i = 2, 3, \ldots, n-1;$$

that is, the sequence $\{\alpha_i\}_{i=1}^n$ is log-concave. After directly verifying that the coefficients $\{\frac{1}{i!(n-i)!}\}_{i=1}^n$ form a log-concave sequence, we have the sequence $\{\frac{\alpha_i}{i!(n-i)!}\}_{i=1}^n$ to be log-concave due to (iii) in Lemma 3.5. Now, from the permanent expression of the cumulative distribution function of $X_{r:n}$ in (15) and statement (v) in Lemma 3.5, we immediately have the log-concavity of the sequence $\{F_{r:n}(x)\}_{r=1}^n$. Realizing that the partial sums of $\{\frac{\alpha_i}{i!(n-i)!}\}_{i=1}^n$ from the left also form a log-concave sequence due to (v) in Lemma 3.5, we have the log-concavity of the sequence $\{1 - F_{r:n}(x)\}_{r=1}^n$. A similar application of Alexandroff's inequality in the permanent expression of the density function of $X_{r:n}$ in (9) will reveal that the sequence $\{f_{r:n}(x)\}_{r=1}^n$ is also log-concave. $\qquad\square$

Remark 4.3. The log-concavity of $\{F_{r:n}(x)\}_{r=1}^n$ established above has an important consequence. Suppose $F_{r:n}(x) > 0$ for $r = 1, 2, \ldots, n$. First of all, observe that

$$\frac{F_{r:n}(x)}{F_{r-1:n}(x)} = \frac{\Pr(X_{r:n} \le x)}{\Pr(X_{r-1:n} \le x)} = \Pr(X_{r:n} \le x \mid X_{r-1:n} \le x).$$

Then, due to (i) in Lemma 3.5, we can conclude that the sequence of conditional probabilities $\{\Pr(X_{r:n} \le x \mid X_{r-1:n} \le x)\}_{r=1}^n$ is non-increasing in r.

Remark 4.4. The log-concavity of $\{F_{r:n}(x)\}_{r=1}^n$ established in Theorem 4.2 has been proved in [84] by direct probability arguments. A stronger log-concavity result has been established in [37] wherein the case when the underlying variables X_i's are possibly dependent has also been considered.

4.3. Case of INID symmetric variables

Suppose the random variables $X_1, X_2, \ldots, X_n$ are independent, non-identically distributed and all symmetric about 0 (without loss of generality). In this section, we establish some properties of order statistics from such a INID symmetric case.

From (15), let us consider

$$
\begin{aligned}
F_{r:n}(-x) &= \Pr(X_{r:n} \le -x) \\
&= \sum_{i=r}^n \frac{1}{i!(n-i)!} \operatorname{Per}\left(\begin{matrix} F_1(-x) & \cdots & F_n(-x) \\ 1 - F_1(-x) & \cdots & 1 - F_n(-x) \end{matrix} \right) \begin{matrix} \} \, i \\ \} \, n-i \end{matrix} \\
&= \sum_{i=r}^n \frac{1}{i!(n-i)!} \operatorname{Per}\left(\begin{matrix} 1 - F_1(x) & \cdots & 1 - F_n(x) \\ F_1(x) & \cdots & F_n(x) \end{matrix} \right) \begin{matrix} \} \, i \\ \} \, n-i \end{matrix} \\
&= \sum_{i=0}^{n-r} \frac{1}{i!(n-i)!} \operatorname{Per}\left(\begin{matrix} F_1(x) & \cdots & F_n(x) \\ 1 - F_1(x) & \cdots & 1 - F_n(x) \end{matrix} \right) \begin{matrix} \} \, i \\ \} \, n-i \end{matrix}
\end{aligned}
$$

$$= \sum_{i=0}^{n} \frac{1}{i!(n-i)!} \operatorname{Per} \begin{pmatrix} F_1(x) & \cdots & F_n(x) \\ 1-F_1(x) & \cdots & 1-F_n(x) \end{pmatrix} \begin{matrix} \} \, i \\ \} \, n-i \end{matrix}$$

$$- \sum_{i=n-r+1}^{n} \frac{1}{i!(n-i)!} \operatorname{Per} \begin{pmatrix} F_1(x) & \cdots & F_n(x) \\ 1-F_1(x) & \cdots & 1-F_n(x) \end{pmatrix} \begin{matrix} \} \, i \\ \} \, n-i \end{matrix}$$

$$= 1 - \sum_{i=n-r+1}^{n} \frac{1}{i!(n-i)!} \operatorname{Per} \begin{pmatrix} F_1(x) & \cdots & F_n(x) \\ 1-F_1(x) & \cdots & 1-F_n(x) \end{pmatrix} \begin{matrix} \} \, i \\ \} \, n-i \end{matrix}. \quad (18)$$

The last equality in (18) follows since

$$\sum_{i=0}^{n} \frac{1}{i!(n-i)!} \operatorname{Per} \begin{pmatrix} F_1(x) & \cdots & F_n(x) \\ 1-F_1(x) & \cdots & 1-F_n(x) \end{pmatrix} \begin{matrix} \} \, i \\ \} \, n-i \end{matrix}$$

$$= \sum_{i=0}^{n} \Pr(\text{exactly } i \text{ of } X\text{'s are} \le x) = 1.$$

Equation (18) simply implies that $-X_{r:n} \overset{d}{=} X_{n-r+1:n}$ for $1 \le r \le n$. This generalizes the corresponding result well-known in the IID case; see [6, p. 26; 53, p. 24].

Lemma 4.5. *Suppose t, t_k, and u are all in $(0,1)$ for $k = s+1, s+2, \ldots, n$. Then, for some r $(1 \le r \le n)$*

$$\sum_{i=0}^{r-1} \frac{1}{i!(n-i)!} \operatorname{Per} \begin{pmatrix} t & \cdots & t & t_{s+1} & \cdots & t_n \\ 1-t & \cdots & 1-t & 1-t_{s+1} & \cdots & 1-t_n \end{pmatrix} \begin{matrix} \} \, n-i \\ \} \, i \end{matrix}$$

$$= \sum_{i=0}^{r-1} \frac{1}{i!(n-i)!} \operatorname{Per} \begin{pmatrix} u & \cdots & u & t_{s+1} & \cdots & t_n \\ 1-u & \cdots & 1-u & 1-t_{s+1} & \cdots & 1-t_n \end{pmatrix} \begin{matrix} \} \, n-i \\ \} \, i \end{matrix}$$

if and only if $t = u$.

Proof. For t, $t_k \in (0,1)$, $k = s+1, s+2, \ldots, n$, let

$$h(t) = \sum_{i=0}^{r-1} \frac{1}{i!(n-i)!} \operatorname{Per} \begin{pmatrix} t & \cdots & t & t_{s+1} & \cdots & t_n \\ 1-t & \cdots & 1-t & 1-t_{s+1} & \cdots & 1-t_n \end{pmatrix} \begin{matrix} \} \, n-i \\ \} \, i \end{matrix}. \quad (19)$$

Differentiating $h(t)$ in (19) with respect to t, we get

$$
h'(t)
$$

$$
= \sum_{i=0}^{r-1} \frac{(n-i)}{i!(n-i)!} \operatorname{Per} \begin{pmatrix} 1 & \cdots & 1 & 0 & \cdots & 0 \\ t & \cdots & t & t_{s+1} & \cdots & t_n \\ 1-t & \cdots & 1-t & 1-t_{s+1} & \cdots & 1-t_n \end{pmatrix} \begin{matrix} \} \, 1 \\ \} \, n-i \\ \} \, i \end{matrix}
$$

$$
- \sum_{i=0}^{r-1} \frac{i}{i!(n-i)!} \operatorname{Per} \begin{pmatrix} t & \cdots & t & t_{s+1} & \cdots & t_n \\ 1 & \cdots & 1 & 0 & \cdots & 0 \\ 1-t & \cdots & 1-t & 1-t_{s+1} & \cdots & 1-t_n \end{pmatrix} \begin{matrix} \} \, n-i \\ \} \, 1 \\ \} \, i-1 \end{matrix}
$$

$$
= \sum_{i=0}^{r-1} \frac{1}{i!(n-i-1)!} \operatorname{Per} \begin{pmatrix} 1 & \cdots & 1 & 0 & \cdots & 0 \\ t & \cdots & t & t_{s+1} & \cdots & t_n \\ 1-t & \cdots & 1-t & 1-t_{s+1} & \cdots & 1-t_n \end{pmatrix} \begin{matrix} \} \, 1 \\ \} \, n-i-1 \\ \} \, i \end{matrix}
$$

$$
- \sum_{i=1}^{r-1} \frac{1}{(i-1)!(n-i)!} \operatorname{Per} \begin{pmatrix} 1 & \cdots & 1 & 0 & \cdots & 0 \\ t & \cdots & t & t_{s+1} & \cdots & t_n \\ 1-t & \cdots & 1-t & 1-t_{s+1} & \cdots & 1-t_n \end{pmatrix} \begin{matrix} \} \, 1 \\ \} \, n-i \\ \} \, i-1 \end{matrix}
$$

$$
= \sum_{i=0}^{r-1} \frac{1}{i!(n-i-1)!} \operatorname{Per} \begin{pmatrix} 1 & \cdots & 1 & 0 & \cdots & 0 \\ t & \cdots & t & t_{s+1} & \cdots & t_n \\ 1-t & \cdots & 1-t & 1-t_{s+1} & \cdots & 1-t_n \end{pmatrix} \begin{matrix} \} \, 1 \\ \} \, n-i-1 \\ \} \, i \end{matrix}
$$

$$
- \sum_{i=0}^{r-2} \frac{1}{i!(n-i-1)!} \operatorname{Per} \begin{pmatrix} 1 & \cdots & 1 & 0 & \cdots & 0 \\ t & \cdots & t & t_{s+1} & \cdots & t_n \\ 1-t & \cdots & 1-t & 1-t_{s+1} & \cdots & 1-t_n \end{pmatrix} \begin{matrix} \} \, 1 \\ \} \, n-i-1 \\ \} \, i \end{matrix}
$$

$$
= \frac{1}{(r-1)!(n-r)!} \operatorname{Per} \begin{pmatrix} 1 & \cdots & 1 & 0 & \cdots & 0 \\ t & \cdots & t & t_{s+1} & \cdots & t_n \\ 1-t & \cdots & 1-t & 1-t_{s+1} & \cdots & 1-t_n \end{pmatrix} \begin{matrix} \} \, 1 \\ \} \, n-r \\ \} \, r-1 \end{matrix}
$$

$$
> 0.
$$

Thus, $h(t)$ in (19) is strictly increasing. Hence, $h(t) > h(u)$ for $t > u$, $h(t) < h(u)$ for $t < u$, and $h(t) = h(u)$ if and only if $t = u$. Hence, the lemma. □

The above lemma can be used to prove the following theorem concerning distributions of order statistics.

Theorem 4.6. *Let $X_1, \ldots, X_s,\ Z_{s+1}, \ldots, Z_n$ be independent random variables with each X_i $(1 \le i \le s)$ having an arbitrary distribution function $F(x)$ and Z_i having arbitrary distribution functions $F_i(x)$, $i = s+1, \ldots, n$. Similarly, let $Y_1, \ldots, Y_s$, $Z_{s+1}, \ldots, Z_n$ be independent random variables with each Y_i $(1 \le i \le s)$ having an arbitrary distribution function $G(x)$. Then, for some fixed r $(1 \le r \le n)$, the r-th order statistic $X_{r:n}$ from the first set of n variables has the same distribution as the r-th order statistic $Y_{r:n}$ from the second set of n variables if $F(\cdot) \equiv G(\cdot)$. Conversely, if*

$X_{r:n}$ and $Y_{r:n}$ *are identically distributed for all* x *such that* $0 < F(x), G(x), F_i(x) < 1$, *then* $F(x) \equiv G(x)$.

Proof. From (15), we have the distribution functions of $X_{r:n}$ and $Y_{r:n}$ to be

$$
\Pr(X_{r:n} \le x) = \sum_{i=r}^{n} \frac{1}{i!(n-i)!}
$$
$$
\times \operatorname{Per} \begin{pmatrix} F(x) & \cdots & F(x) & F_{s+1}(x) & \cdots & F_n(x) \\ 1-F(x) & \cdots & 1-F(x) & 1-F_{s+1}(x) & \cdots & 1-F_n(x) \end{pmatrix} \begin{matrix} \} \, i \\ \} \, n-i \end{matrix} \quad (20)
$$

and

$$
\Pr(Y_{r:n} \le x) = \sum_{i=r}^{n} \frac{1}{i!(n-i)!}
$$
$$
\times \operatorname{Per} \begin{pmatrix} G(x) & \cdots & G(x) & F_{s+1}(x) & \cdots & F_n(x) \\ 1-G(x) & \cdots & 1-G(x) & 1-F_{s+1}(x) & \cdots & 1-F_n(x) \end{pmatrix} \begin{matrix} \} \, i \\ \} \, n-i \end{matrix} . \quad (21)
$$

If $F(\cdot) \equiv G(\cdot)$, then it is clear from (20) and (21) that $X_{r:n} \stackrel{d}{=} Y_{r:n}$.

In order to prove the converse, suppose $X_{r:n} \stackrel{d}{=} Y_{r:n}$ for all x such that $0 < F(x), G(x), F_i(x) < 1$. Then, upon equating the right-hand sides of (20) and (21) and invoking Lemma 4.5, we simply get $F(x) \equiv G(x)$. $\qquad \square$

Theorem 4.7. *Let* $X_1, \ldots, X_n$ *be independent random variables with each* X_i *($1 \le i \le s$) having an arbitrary distribution function* $F(x)$ *and* X_i *having arbitrary distribution functions* $F_i(x)$ *for* $i = s+1, s+2, \ldots, n$. *Suppose* X_i *($i = s+1, \ldots, n$) are all symmetric about zero. Then, for fixed* r *($1 \le r \le n$),* $-X_{r:n} \stackrel{d}{=} X_{n-r+1:n}$ *if* X_i *($i = 1, 2, \ldots, s$) are also symmetric about zero. Conversely, if* $-X_{r:n}$ *and* $X_{n-r+1:n}$ *are identically distributed for all* x *such that* $0 < F(x), F_i(x) < 1$, *then* $X_1, \ldots, X_s$ *are also symmetric about zero.*

Proof. The result follows from Theorem 4.6 simply by taking $-X_1, -X_2, \ldots, -X_n$ in place of $Y_1, \ldots, Y_s, Z_{s+1}, \ldots, Z_n$. $\qquad \square$

Remark 4.8. Theorem 4.7, for the case of absolutely continuous distributions and $s = 1$, was proved in [40]. It should be noted that Theorem 4.7 gives a stronger result for distributions of order statistics in the INID symmetric case than the one presented earlier.

Simpler proofs of these results and also some extensions are given in [64]. For example, when the X_i's are symmetric variables (about 0), then by simply noting that $(X_1, X_2, \ldots, X_n)$ and $(-X_1, -X_2, \ldots, -X_n)$ have the same distribution and hence the r-th order statistic of X_i's has the same distribution as the r-th order statistic of $-X_i$'s, the result that $-X_{r:n} \stackrel{d}{=} X_{n-r+1:n}$ (proved in the beginning of this section) follows very easily.

Remark 4.9. We may also note that $-X_{r:n} \overset{d}{=} X_{n-r+1:n}$ for all $r = 1, 2, \ldots, n$ when X_i's are arbitrary random variables (not necessarily independent) such that $(X_1, X_2, \ldots, X_n)$ and $(-X_{P(1)}, -X_{P(2)}, \ldots, -X_{P(n)})$ for some permutation $(P(1), P(2), \ldots, P(n))$ of $(1, 2, \ldots, n)$.

Without assuming absolute continuity for the distribution functions, using simple probability arguments, the following result (due to [40] as indicated above) has been proved in [64].

Theorem 4.10. *Let $X_1, X_2, \ldots, X_n$ be independent random variables. Suppose X_i, $i = 2, \ldots, n$, are all symmetric about 0. If $-X_{r:n} \overset{d}{=} X_{n-r+1:n}$, then X_1 is also symmetric about 0.*

Proceeding similarly, a proof for the more general one-way implication in Theorem 4.7 has also been given in [64].

Definition 4.11. Two random variables X *and* Y *are stochastically ordered if*

$$\Pr(X > t) \geq \Pr(Y > t) \qquad \text{for every } t. \tag{22}$$

If strict inequality holds in (22) for all t, then we say that X *and* Y *are strictly stochastically ordered.*

Then, [64] established an equivalence stated in Theorem 4.13 the proof of which needs the following lemma.

Lemma 4.12. *Let B be the sum of n independent Bernoulli random variables with parameters p_i, $i = 1, 2, \ldots, n$; similarly, let B^* be the sum of n independent Bernoulli random variables with parameters p_i^*, $i = 1, 2, \ldots, n$. If B and B^* have the same distribution, then*

$$(p_1, p_2, \ldots, p_n) = (p_{P(1)}^*, p_{P(2)}^*, \ldots, p_{P(n)}^*)$$

for some permutation $(P_{(1)}, P_{(2)}, \ldots, P_{(n)})$ of $(1, 2, \ldots, n)$.

Theorem 4.13. *Let X_i's be strictly stochastically ordered random variables. Then, the following two statements are equivalent:*

  (i) $(X_1, \ldots, X_n)$ *and* $-(X_{P(1)}, \ldots, X_{P(n)})$ *have the same distribution for some permutation $(P(1), \ldots, P(n))$ of $(1, 2, \ldots, n)$.*

  (ii) $-X_{r:n} \overset{d}{=} X_{n-r+1:n}$ *for all $r = 1, 2, \ldots, n$.*

Proof. Let B_i denote the indicator variable for the event $\{X_i \leq t\}$, and B_i^* denote the indicator variable for the event $\{-X_i \leq t\}$; further, let $B = \sum_{i=1}^{n} B_i$ and $B^* = \sum_{i=1}^{n} B_i^*$. Remark 4.9 showed that (i) $\Rightarrow$ (ii).

Now, suppose $-X_{r:n} \overset{d}{=} X_{n-r+1:n}$ for every r; then B and B^* have the same distribution. Then, (ii) $\Rightarrow$ (i) follows readily from Lemma 4.12 and the fact that X_i's are strictly stochastically ordered. $\qquad\qquad\square$

Remark 4.14. Through this argument, [64] also presented the following simple proof for the log-concavity property of $\{F_{r:n}(t)\}_{r=1}^{n}$ and $\{1 - F_{r:n}(t)\}_{r=1}^{n}$ established earlier in Theorem 4.2. By the simple fact that $B = \sum_{i=1}^{n} B_i$ is the sum of n independent Bernoulli random variables, it is strongly unimodal [59, p. 109]. Consequently, the sequences $\{\Pr(B = r)\}$ and $\{\Pr(B \leq r)\}$ are log-concave. From this, the log-concavity property of $\{F_{r:n}(t)\}_{r=1}^{n}$ and $\{1 - F_{r:n}(t)\}_{r=1}^{n}$ follows at once.

4.4. Characterizations of IID case

In this section, we shall describe some characterizations of the IID case established in [42]. For this purpose, let us denote the pdf, cdf, and the hazard rate (or failure rate) of X_i by $f_i(\cdot)$, $F_i(\cdot)$, and $h_i(\cdot)$, respectively, for $i = 1, 2, \ldots, n$. Let us also define the variables

$$I_{r,n} = i \qquad \text{if} \quad X_{r:n} = X_i \text{ for } 1 \leq r \leq n. \tag{23}$$

Since the random variables X_i's are assumed to be of continuous type, the variables $I_{r,n}$'s in (23) are uniquely defined with probability 1.

Definition 4.15. The variables X_i's are said to have proportional hazard rates if there exist constants $\gamma_i > 0$, $i = 1, 2, \ldots, n$, such that

$$h_i(x) = \gamma_i \, h_1(x) \quad \text{for all } x \text{ and } i = 2, 3, \ldots, n, \tag{24}$$

or equivalently, if the survival functions satisfy

$$1 - F_i(x) = \{1 - F_i(x)\}^{\gamma_i} \quad \text{for all } x \text{ and } i = 2, 3, \ldots, n. \tag{25}$$

The family of distributions satisfying (24) or (25) is then called the *proportional hazard family.*

The following characterization result, which has been used extensively in the theory of competing risks, is due to [2, 4, 86].

Theorem 4.16. *The random variables $X_1, X_2, \ldots, X_n$ belong to the proportional hazard family defined in (24) or (25) if and only if $X_{1:n}$ and $I_{1,n}$ are statistically independent.*

The above theorem simply states that if there are n independent risks acting simultaneously on a system in order to make it fail, then the time to failure of the system is independent of the cause of the failure if and only if the n lifetimes belong to the proportional hazard family.

By assuming that the X_i's belong to the proportional hazard family, a necessary and sufficient condition for the X_i's to be identically distributed has been established in [42]. To present this theorem, we first need the following lemma.

Lemma 4.17. *Let $c_1, c_2, \ldots, c_n$ be real numbers and $0 < d_1 < d_2 < \cdots < d_n$. If*

$$\sum_{i=1}^{n} c_i \, u^{d_i} = 0, \qquad 0 \leq u \leq 1,$$

then $c_i = 0$ for all $i = 1, 2, \ldots, n$.

Proof. The result follows by taking $0 < u_1 < u_2 < \cdots < u_n < 1$, writing the corresponding system of equations as

$$\begin{bmatrix} u_1^{d_1} & \cdots & u_1^{d_n} \\ \cdots & \cdots & \cdots \\ u_n^{d_1} & \cdots & u_n^{d_n} \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix},$$

and using the nonsingularity of the matrix on the L.H.S.; see [82, p. 46]. $\qquad\square$

Theorem 4.18. *Let $X_1, X_2, \ldots, X_n$ be independent random variables with proportional hazard rates. Then, X_i's are IID if and only if $X_{r:n}$ and $I_{r,n}$ are statistically independent for some $r \in \{2, 3, \ldots, n\}$.*

Definition 4.19. The dual family of distributions such that

$$F_i(x) = F_1^{\alpha_i}(x) \quad \text{for all } x \text{ and } i = 2, 3, \ldots, n, \tag{26}$$

or equivalently, if the survival rates satisfy

$$s_i(x) = \alpha_i \, s_1(x) \quad \text{for all } x \text{ and } i = 2, 3, \ldots, n, \tag{27}$$

where the survival rate $s_i(x) = f_i(x)/F_i(x)$, will be called the *proportional survival rate family*.

Analogous to Theorems 4.16 and 4.18, we then have the following two results.

Theorem 4.20. *The random variables $X_1, X_2, \ldots, X_n$ belong to the proportional survival rate family defined in (26) or (27) if and only if $X_{n:n}$ and $I_{n:n}$ are statistically independent.*

Theorem 4.21. *Let $X_1, X_2, \ldots, X_n$ be independent random variables with proportional survival rates. Then, X_i's are IID if and only if $X_{r:n}$ and $I_{r,n}$ are independent for some $r \in \{1, 2, \ldots, n-1\}$.*

Remark 4.22. If X_i's are IID, it is obvious that $X_{r:n}$ and $I_{r,n}$ are independent for any $r \in \{1, 2, \ldots, n\}$. On the other hand, the independence of $X_{r:n}$ and $I_{r,n}$, $r \in \{1, n\}$, and that of $X_{s:n}$ and $I_{s,n}$, $s \in \{1, 2, \ldots, n\} \setminus \{r\}$, will be sufficient to claim the independence of all other pairs $X_{i:n}$ and $I_{i,n}$ from Theorems 4.18 and 4.21.

Another interesting characterization of IID has been presented in [42] based on subsamples of size $n-1$. For describing this result, let us consider the case when $X_1, X_2, \ldots, X_n$ are independent random variables of continuous type with common support, use $X_{r:n-1}^{[i]}$ to denote the r-th order statistic from $n-1$ variables and $F_{r:n-1}^{[i]}(x)$ for the cdf of $X_{r:n-1}^{[i]}$.

Also, let $N_i = \{1, 2, \ldots, n\} \setminus \{i\}$, $N_{ij} = \{1, 2, \ldots, n\} \setminus \{i, j\}$, and $\pi_{k, N_{ij}}(x)$ be the probability that exactly k of the $n-2$ X_ℓ's, $\ell \in N_{ij}$, are less than x (with $\pi_{0, N_{ij}}(x) = 1$).

Lemma 4.23. *For $i, j \in \{1, 2, \ldots, n\}$ and $i \neq j$,*

$$F_{r:n-1}^{[i]}(x) - F_{r:n-1}^{[j]}(x) = \pi_{r-1, N_{ij}}(x) \{F_j(x) - F_i(x)\}, \quad 1 \leq r \leq n-1.$$

Proof. We have

$$\begin{aligned}
F_{r:n-1}^{[i]}(x) &= \Pr(\text{at least } r \text{ of } X_\ell, \ell \in N_i, \text{ are } \leq x) \\
&= \sum_{k=r}^{n-1} \Pr(\text{exactly } k \text{ of } X_\ell, \ell \in N_i, \text{ are } \leq x) \\
&= \sum_{k=r}^{n-1} \Pr(\text{exactly } k-1 \text{ of } X_\ell, \ell \in N_{ij}, \text{ are } \leq x) \, F_j(x) \\
&\quad + \sum_{k=r}^{n-2} \Pr(\text{exactly } k \text{ of } X_\ell, \ell \in N_{ij}, \text{ are } \leq x)\{1 - F_j(x)\} \\
&= \sum_{k=r}^{n-2} \pi_{k, N_{ij}}(x) + \pi_{r-1, N_{ij}}(x) \, F_j(x).
\end{aligned} \tag{28}$$

Similarly, we have

$$F_{r:n-1}^{[j]}(x) = \sum_{k=r}^{n-2} \pi_{k, N_{ij}}(x) + \pi_{r-1, N_{ij}}(x) \, F_i(x). \tag{29}$$

Upon subtracting (29) from (28), the result follows. $\qquad\square$

Theorem 4.24. *The random variables $X_1, X_2, \ldots, X_n$ are IID if and only if the random variables $X_{r:n-1}^{[1]}, \ldots, X_{r:n-1}^{[n]}$ have the same distribution for some fixed $r \in \{1, 2, \ldots, n-1\}$.*

Proof. It is obvious that if $X_1, X_2, \ldots, X_n$ are IID, then $X_{r:n-1}^{[1]}, \ldots, X_{r:n-1}^{[n]}$ will all have the same distribution.

Let $i, j \in \{1, 2, \ldots, n\}$ and $i \neq j$. If x is in the common support of X_ℓ's, then $\pi_{r-1, N_{ij}}(x) > 0$; since $F_{r:n-1}^{[i]}(x) = F_{r:n-1}^{[j]}(x)$, we get $F_i(x) = F_j(x)$ from Lemma 4.23. We may similarly prove that $F_1(x) = \cdots = F_n(x)$ for all x in the common support. $\quad\square$

Remark 4.25. Let $i, j \in \{1, 2, \ldots, n\}$ and $i \neq j$. From Lemma 4.23, we see easily that

$$F_{r:n-1}^{[i]}(x) - F_{r:n-1}^{[j]}(x) \lesseqgtr 0 \quad \text{according as} \quad F_j(x) - F_i(x) \lesseqgtr 0.$$

In particular, $X_i \overset{\text{st}}{\geq} X_j$ if and only if $X_{r:n-1}^{[i]} \overset{\text{st}}{\leq} X_{r:n-1}^{[j]}$.

Remark 4.26. It also follows easily from Lemma 4.23 that $E(X_{r:n-1}^{[i]}) - E(X_{r:n-1}^{[j]}) \geq E(X_j) - E(X_i)$ for $i, j \in \{1, 2, \ldots, n\}$ and $i \neq j$. A similar inequality holds for other moments as well.

5. Relations for order statistics from INID variables

5.1. Introduction

Several recurrence relations and identities for order statistics in the IID case are available in the literature. The book in [53], the survey paper in [76], and the monograph in [5] all provide elaborative and exhaustive treatment to this topic. Since many of these results were extended in [8–10] to the case when the order statistics arise from a sample containing a single outlier, a number of papers have appeared establishing and extending most of the results to (i) the INID case and (ii) the arbitrary case. All the results for (i) are proved through permanents, and they will be discussed here in detail. The results for (ii), on the other hand, are established using a variety of techniques like probabilistic methods, set theoretic arguments, operator methods, and indicator methods.

In this section, we assume that $X_1, X_2, \ldots, X_n$ are INID random variables with X_i having cumulative distribution function $F_i(x)$ and probability density function $f_i(x)$, for $i = 1, 2, \ldots, n$. Let $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ denote the order statistics obtained by arranging the X_i's in increasing order of magnitude. Let S be a subset of $N = \{1, 2, \ldots, n\}$, S^c be the complement of S in N, and $|S|$ denote the cardinality of the set S. Let $X_{r:S}$ denote the r-th order statistic obtained from the variables $\{X_i | i \in S\}$, and $F_{r:S}(x)$ and $f_{r:S}(x)$ denote the cumulative distribution function and density function of $X_{r:S}$, respectively. Occasionally (when there is no confusion, of course), we may even replace S by $|S|$ in the above notations (like, for example, $X_{r:n}$ instead of $X_{r:N}$).

For fixed $x \in \mathbb{R}$, let us denote the row vector $\begin{pmatrix} F_1(x) & F_2(x) & \cdots & F_n(x) \end{pmatrix}_{1 \times n}$ by $\boldsymbol{F}$, $\begin{pmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \end{pmatrix}_{1 \times n}$ by $\boldsymbol{f}$, and $\begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix}_{1 \times n}$ by $\boldsymbol{1}$. Let us use $\boldsymbol{A}_1[S]$ to denote the matrix obtained along the lines of $\boldsymbol{A}_1$ in (10) starting with components corresponding to $i \in S$. Further, let us define for $i = 1, 2, \ldots, r$

$$\overline{F}_{i:r}(x) = \frac{1}{\binom{n}{r}} \sum_{|S|=r} F_{i:S}(x) \tag{30}$$

and

$$\overline{f}_{i:r}(x) = \frac{1}{\binom{n}{r}} \sum_{|S|=r} f_{i:S}(x), \tag{31}$$

where $\sum_{|S|=r}$ denotes the sum over all subsets S of N with cardinality equal to r. We shall also follow notations similar to those in (30) and (31) for the joint distribution functions and density functions of order statistics. Let us also use $F_{r:n-1}^{[i]}(x)$ and $f_{r:n-1}^{[i]}(x)$ to denote the distribution function and density function of the r-th order statistic from $n-1$ variables obtained by deleting X_i from the original n X's. Similar notations also hold for joint distributions as well as for distributions of order statistics from $n-m$ variables obtained by deleting m X's.

5.2. Relations for single order statistics

In this section, we derive several recurrence relations and identities satisfied by distributions of single order statistics. These generalize many well-known results for order statistics in the IID case discussed in detail in [5, 6, 53, 57]. For a review of all these results, one may refer to [14]. Even though the results are given in terms of distributions or densities, they hold equally well for moments (if they exist).

Result 5.1. *For $n \geq 2$ and $x \in \mathbb{R}$,*

$$\sum_{r=1}^{n} F_{r:n}(x) = \sum_{r=1}^{n} F_r(x) = n\,\overline{F}_{1:1}(x) \ . \tag{32}$$

Proof. The result follows simply by noting that

$$\sum_{r=1}^{n} \Pr(X_{r:n} \leq x) = \sum_{r=1}^{n} \Pr(X_r \leq x). \qquad \square$$

Result 5.2 (Triangle Rule). *For $1 \leq r \leq n-1$ and $x \in \mathbb{R}$,*

$$r\,f_{r+1:n}(x) + (n-r)\,f_{r:n}(x) = n\,\overline{f}_{r:n-1}(x). \tag{33}$$

Proof. By considering the expression of $r\,f_{r+1:n}(x)$ from (9) and expanding the permanent by its first row, we get

$$r\,f_{r+1:n}(x) = \sum_{i=1}^{n} F_i(x)\,f_{r:n-1}^{[i]}(x). \tag{34}$$

Next, by considering the expression of $(n-r)\,f_{r:n}(x)$ from (9) and expanding the permanent by its last row, we get

$$(n-r)\,f_{r:n}(x) = \sum_{i=1}^{n} \{1 - F_i(x)\}\,f_{r:n-1}^{[i]}(x). \tag{35}$$

On adding (34) and (35) and simplifying, we derive the relation in (33). $\qquad \square$

Remark 5.3. It is easy to note from Result 5.2 that one just needs the distribution of a single order statistic arising from n variables in order to determine the distributions of the remaining $n-1$ order statistics, assuming that the distributions of order statistics arising from $n-1$ (and less) variables are known. This result was first proved in [11] and independently in [40].

Result 5.4. *For $x \in \mathbb{R}$,*

$$\frac{1}{2}\left\{F_{n+1:2n}(x) + F_{n:2n}(x)\right\} = \overline{F}_{n:2n-1}(x). \tag{36}$$

Proof. The result follows from Result 5.2 upon taking $2n$ in place of n and n in place of r. $\qquad\square$

In terms of expected values, the relation in (36) simply implies that the expected value of the median from $2n$ variables is exactly the same as the average of the expected values of the medians from $2n-1$ variables (obtained by deleting one variable at a time).

Result 5.5. *For $m = 1, 2, \ldots, n - r$ and $x \in \mathbb{R}$,*

$$f_{r:n}(x) = \sum_{j=0}^{n}(-1)^j\left\{\frac{\binom{r+j-1}{j}\binom{n}{m-j}}{\binom{n-r}{m}}\right\}\overline{f}_{r+j:n-m+j}(x). \tag{37}$$

Proof. By considering the expression of $f_{r:n}(x)$ in (8), writing

$$\{1 - F_{i_{r+1}}(x)\}\cdots\{1 - F_{i_{r+m}}(x)\} = \sum_{j=0}^{m}(-1)^j\sum_{|S|=j}F_{\ell_1}(x)\cdots F_{\ell_j}(x),$$

where $\sum_{|S|=j}$ denotes the sum over all $\binom{m}{j}$ subsets $S = \{\ell_1, \ell_2, \ldots, \ell_j\}$ of $\{i_{r+1}, i_{r+2}, \ldots, i_{r+m}\}$ (with cardinality j), and simplifying the resulting expression, we derive the relation in (37). $\qquad\square$

Result 5.2 may be deduced from (37) by setting $m = 1$.

Proceeding as we did in proving Result 5.5, we can establish the following dual relation.

Result 5.6. *For $m = 1, 2, \ldots, r - 1$ and $x \in \mathbb{R}$,*

$$f_{r:n}(x) = \sum_{j=n-m}^{n}(-1)^{j-n+m}\left\{\frac{\binom{j+m-r}{n-r}\binom{n}{j}}{\binom{r-1}{m}}\right\}\overline{f}_{r-m:j}(x).$$

Upon setting $m = n - r$ and $m = r - 1$ in Results 5.5 and 5.6, respectively, we derive the following relations.

Result 5.7. *For $1 \le r \le n-1$ and $x \in \mathbb{R}$,*

$$f_{r:n}(x) = \sum_{j=r}^{n} (-1)^{j-r} \binom{j-1}{r-1} \binom{n}{j} \overline{f}_{j:j}(x). \tag{38}$$

Result 5.8. *For $2 \le r \le n$ and $x \in \mathbb{R}$,*

$$f_{r:n}(x) = \sum_{j=n-r+1}^{n} (-1)^{j-n+r-1} \binom{j-1}{n-r} \binom{n}{j} \overline{f}_{1:j}(x).$$

Remark 5.9. Results 5.7 and 5.8 are both very useful as they express the distribution of the r-th order statistic arising from n variables in terms of the distributions of the largest and smallest order statistics arising from n variables or less, respectively. We, therefore, note once again from Results 5.7 and 5.8 that we just need the distribution of a single order statistic (either the largest or the smallest) arising from n variables in order to determine the distributions of the remaining $n-1$ order statistics, given the distributions of order statistics arising from at most $n-1$ variables. This agrees with the comment made earlier in Remark 5.3, which is only to be expected as both Results 5.7 and 5.8 could be derived by repeated application of Result 5.2 as shown in $[11, 40]$. This was observed in $[60, 90]$ for the IID case.

Theorem 5.10. *If any one of Results 5.2, 5.7, or 5.8 is used in the computation of single distributions (or single moments) of order statistics arising from n INID variables, then the identity given in Result 5.1 will be automatically satisfied and hence should not be applied to check the computational process.*

Proof. We shall prove the theorem first by starting with Result 5.7, and the proof for Result 5.8 is quite similar. From (38), we have

$$\sum_{r=1}^{n-1} f_{r:n}(x) = \sum_{r=1}^{n-1} \sum_{j=r}^{n} (-1)^{j-r} \binom{j-1}{r-1} \binom{n}{j} \overline{f}_{j:j}(x)$$

$$= \binom{n}{1} \overline{f}_{1:1}(x) + \sum_{r=2}^{n-1} \binom{n}{r} \overline{f}_{r:r}(x) \left\{ \sum_{j=0}^{r-1} (-1)^{r-1-j} \binom{r-1}{j} \right\}$$

$$+ \sum_{j=0}^{n-2} (-1)^{n-1-j} \binom{n-1}{j} f_{n:n}(x)$$

$$= \binom{n}{1} \overline{f}_{1:1}(x) - f_{n:n}(x), \tag{39}$$

where the last equality follows from the fact that $\overline{f}_{n:n}(x) \equiv f_{n:n}(x)$ and upon using the combinatorial identities

$$\sum_{j=0}^{r-1} (-1)^{r-1-j} \binom{r-1}{j} = 0 \quad \text{and} \quad \sum_{j=0}^{n-2} (-1)^{n-1-j} \binom{n-1}{j} = -1.$$

Equation (39), when rewritten, gives the identity presented in Result 5.1. This proof was given in [34] for the IID case.

In order to prove the theorem with Result 5.2, consider the relation in (33) and set $r = 1$, $r = 2$, ..., $r = n - 1$, and add the resulting $n - 1$ equations, to get

$$(n-1)\sum_{r=1}^{n} f_{r:n}(x) = n \sum_{r=1}^{n-1} \overline{f}_{r:n-1}(x)$$

or

$$\frac{1}{n}\sum_{r=1}^{n} f_{r:n}(x) = \frac{1}{n-1}\sum_{r=1}^{n-1} \overline{f}_{r:n-1}(x)$$
$$= \cdots = \overline{f}_{1:1}(x)$$

which is simply the identity in (32). $\qquad\qquad\square$

Through repeated application of Result 5.2, [14] proved the following relation which was established in [89] for the IID case.

Result 5.11. *For* $1 \le r \le m \le n - 1$ *and* $x \in \mathbb{R}$,

$$\sum_{j=0}^{n-m}\binom{r-1+j}{j}\binom{n-r-j}{n-m-j}f_{r+j:n}(x) = \binom{n}{m}\overline{f}_{r:m}(x).$$

Result 5.12. *For* $n \ge 2$ *and* $x \in \mathbb{R}$,

$$\sum_{r=1}^{n}\frac{1}{r}f_{r:n}(x) = \sum_{r=1}^{n}\frac{1}{r}\overline{f}_{1:r}(x) \qquad\qquad (40)$$

and

$$\sum_{r=1}^{n}\frac{1}{n-r+1}f_{r:n}(x) = \sum_{r=1}^{n}\frac{1}{r}\overline{f}_{r:r}(x). \qquad\qquad (41)$$

Proof. We shall prove here the identity in (40), and the proof for (41) is quite similar. By using Result 5.8, we can write

$$\sum_{r=1}^{n}\frac{1}{r}f_{r:n}(x) = \sum_{r=1}^{n}\frac{1}{r}\sum_{j=n-r+1}^{n}(-1)^{j-n+r-1}\binom{j-1}{n-r}\binom{n}{j}\overline{f}_{1:j}(x)$$
$$= \sum_{r=1}^{n}\binom{n}{r}\overline{f}_{1:r}(x)\left\{\sum_{j=0}^{r-1}(-1)^{j}\binom{r-1}{j}\Big/(n-r+1-j)\right\}. \qquad (42)$$

The identity in (40) follows readily from (42) upon using the combinatorial identity that

$$\sum_{j=0}^{r-1} (-1)^j \binom{r-1}{j} \Big/ (n-r+1-j) = \int_0^1 (1-t)^{r-1} \, t^{n-r} \, dt = B(r, n-r+1),$$

where $B(a,b) = \Gamma(a)\,\Gamma(b)/\Gamma(a+b)$ is the complete beta function. $\qquad\square$

The above result, established in [70] for the IID case, was proved in [8] for a single-outlier model and in the INID case in [41]. The result has also been extended to the arbitrary case in [23]. The extensions of Joshi's identities given in [33] for the IID case can also be generalized as follows.

Result 5.13. *For $i, j = 1, 2, \ldots$ and $x \in \mathbb{R}$,*

$$\sum_{r=1}^{n} \frac{1}{(r+i+j-2)^{(j)}} \, f_{r:n}(x)$$

$$= \frac{1}{(n+i+j-2)^{(j-1)}} \sum_{r=1}^{n} \frac{1}{r} \binom{r+j-2}{j-1} \frac{(n-r+i-1)^{(i-1)}}{(n+i-1)^{(i-1)}} \overline{f}_{1:r}(x) \quad (43)$$

and

$$\sum_{r=1}^{n} \frac{1}{(n-r+i+j-1)^{(j)}} \, f_{r:n}(x)$$

$$= \frac{1}{(n+i+j-2)^{(j-1)}} \sum_{r=1}^{n} \frac{1}{r} \binom{r+j-2}{j-1} \frac{(n-r+i-1)^{(i-1)}}{(n+i-1)^{(i-1)}} \overline{f}_{r:r}(x); \quad (44)$$

for $i = 1, 2, \ldots$,

$$\sum_{r=1}^{n} \frac{1}{(r+i-1)^{(i)}(n-r+i)^{(i)}} \, f_{r:n}(x)$$

$$= \frac{1}{(n+2i-1)^{(2i-1)}} \sum_{r=1}^{n} \frac{1}{r} \binom{r+2i-2}{i-1} \{\overline{f}_{1:r}(x) + \overline{f}_{r:r}(x)\};$$

and for $i, j = 1, 2, \ldots$,

$$\sum_{r=1}^{n} \frac{1}{(r+i-1)^{(i)}(n-r+j)^{(j)}} \, f_{r:n}(x)$$

$$= \frac{1}{(n+i+j-1)^{(i+j-1)}}$$

$$\times \sum_{r=1}^{n} \frac{1}{r} \left\{ \binom{r+i+j-2}{i-1} \overline{f}_{1:r}(x) + \binom{r+i+j-2}{j-1} \overline{f}_{r:r}(x) \right\},$$

where

$$m^{(i)} = \begin{cases} m(m-1)\cdots(m-i+1) & for \quad i = 1, 2, \ldots, \\ 1 & for \quad i = 0. \end{cases}$$

The two identities in Result 5.12 may be deduced from (43) and (44) by setting $i = j = 1$. A different type of extension of Result 5.12 derived in [14] is presented below.

Result 5.14. *For $\ell = 0, 1, \ldots, n-2$ and $x \in \mathbb{R}$,*

$$\sum_{r=\ell+1}^{n} \frac{1}{r} f_{r:n}(x) = \sum_{r=\ell+1}^{n} \frac{1}{r} \overline{f}_{\ell+1:r}(x) \tag{45}$$

and

$$\sum_{r=1}^{n-\ell} \frac{1}{n-r+1} f_{r:n}(x) = \sum_{r=1}^{n-\ell} \frac{1}{r+\ell} \overline{f}_{r:r+\ell}(x). \tag{46}$$

Proof. We shall prove here the identity in (45), and the proof for (46) is quite similar. From Result 5.6, upon setting $m = r - 1 - \ell$ we have

$$f_{r:n}(x) = \sum_{j=n-r+1+\ell}^{n} (-1)^{j-n+r-1-\ell} \left\{ \frac{\binom{j-1-\ell}{n-r}\binom{n}{j}}{\binom{r-1}{\ell}} \right\} \overline{f}_{\ell+1:j}(x)$$

$$= \sum_{j=0}^{r-1-\ell} (-1)^{r-1-\ell-j} \left\{ \frac{\binom{n-1-\ell-j}{n-r}\binom{n}{j}}{\binom{r-1}{\ell}} \right\} \overline{f}_{\ell+1:n-j}(x). \tag{47}$$

Upon making use of the expression of $f_{r:n}(x)$ in (47) and rewriting, we get

$$\sum_{r=\ell+1}^{n} \frac{1}{r} f_{r:n}(x) = \sum_{r=\ell+1}^{n} C_r \binom{n}{r} \overline{f}_{\ell+1:r}(x), \tag{48}$$

where the coefficients C_r $(\ell + 1 \leq r \leq n)$ are given by

$$C_{\ell+1+s} = \sum_{j=n-s}^{n} (-1)^{j-n+s} \frac{1}{j} \binom{s}{n-j} \Big/ \binom{j-1}{\ell}$$

$$= \sum_{j=n-s}^{n} (-1)^{j-n+s} \binom{s}{n-j} B(\ell+1, j-\ell)$$

$$= \sum_{j=0}^{s} (-1)^{s-j} \binom{s}{j} B(\ell+1, n-j-\ell)$$

$$= \int_{0}^{1} t^s \, t^\ell \, (1-t)^{n-\ell-s-1} \, dt$$

$$= B(\ell+s+1, n-\ell-s).$$

The identity in (45) follows readily if we substitute the above expression of C_r in (48). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \Box$

Proceeding in an analogous manner, the following identities have been established in [14].

Result 5.15. *For* $\ell_1, \ell_2 \geq 0$, $\ell_1 + \ell_2 \leq n-1$, *and* $x \in \mathbb{R}$,

$$\sum_{r=\ell_1+1}^{n-\ell_2} \frac{1}{r} f_{r:n}(x) = \sum_{r=\ell_1+1}^{n-\ell_2} \frac{1}{r+\ell_2} \overline{f}_{\ell_1+1:r+\ell_2}(x)$$
$$+ \sum_{r=\ell_1+1}^{n-\ell_2} \left(\frac{1}{r} - \frac{1}{r+\ell_2} \right) \overline{f}_{r:r+\ell_2}(x)$$

and

$$\sum_{r=\ell_1+1}^{n-\ell_2} \frac{1}{n-r+1} f_{r:n}(x) = \sum_{r=\ell_1+1}^{n-\ell_2} \frac{1}{r+\ell_2} \overline{f}_{r:r+\ell_2}(x)$$
$$+ \sum_{r=\ell_1+1}^{n-\ell_2} \left(\frac{1}{r+\ell_2-\ell_1} - \frac{1}{r+\ell_2} \right) \overline{f}_{\ell_1+1:r+\ell_2}(x).$$

When X_i's are IID with distribution function $F(x)$, [58] presented the relations

$$F_{r:n}(x) = F_{r+1:n}(x) + \binom{n}{r} \{F(x)\}^r \{1-F(x)\}^{n-r}$$

and

$$F_{r:n}(x) = F_{r:n-1}(x) + \binom{n-1}{r-1} \{F(x)\}^r \{1-F(x)\}^{n-r}$$

wherein analogous results are also presented for the single-outlier model. The generalizations of these results to the INID case, as proved in [40], are presented below.

Result 5.16. *For* $1 \leq r \leq n-1$ *and* $x \in \mathbb{R}$,

$$F_{r:n}(x) = F_{r+1:n}(x) + \binom{n}{r} \frac{1}{n!} \operatorname{Per} \left(\begin{matrix} \boldsymbol{F} \\ \boldsymbol{1-F} \end{matrix} \right) \begin{matrix} \} \, r \\ \} \, n-r \end{matrix} \qquad (49)$$

and

$$F_{r:n}(x) = \overline{F}_{r:n-1}(x) + \binom{n-1}{r-1} \frac{1}{n!} \operatorname{Per} \left(\begin{matrix} \boldsymbol{F} \\ \boldsymbol{1-F} \end{matrix} \right) \begin{matrix} \} \, r \\ \} \, n-r \end{matrix}. \qquad (50)$$

Proof. The relation in (49) follows quite simply from (15) by writing

$$
F_{r:n}(x) = \sum_{i=r}^{n} \frac{1}{i!(n-i)!} \operatorname{Per} \begin{pmatrix} \boldsymbol{F} \\ \boldsymbol{1-F} \end{pmatrix} \begin{matrix} \} \, i \\ \} \, n-i \end{matrix}
$$

$$
= \sum_{i=r+1}^{n} \frac{1}{i!(n-i)!} \operatorname{Per} \begin{pmatrix} \boldsymbol{F} \\ \boldsymbol{1-F} \end{pmatrix} \begin{matrix} \} \, i \\ \} \, n-i \end{matrix}
$$

$$
+ \frac{1}{r!(n-r)!} \operatorname{Per} \begin{pmatrix} \boldsymbol{F} \\ \boldsymbol{1-F} \end{pmatrix} \begin{matrix} \} \, r \\ \} \, n-r \end{matrix}
$$

$$
= F_{r+1:n}(x) + \binom{n}{r} \frac{1}{n!} \operatorname{Per} \begin{pmatrix} \boldsymbol{F} \\ \boldsymbol{1-F} \end{pmatrix} \begin{matrix} \} \, r \\ \} \, n-r \end{matrix}.
$$

Next, in order to prove the relation in (50) let us consider

$$
n! \, F_{r:n}(x) = \sum_{i=r}^{n} \binom{n}{i} \operatorname{Per} \begin{pmatrix} \boldsymbol{F} \\ \boldsymbol{1-F} \end{pmatrix} \begin{matrix} \} \, i \\ \} \, n-i \end{matrix}
$$

$$
= \sum_{i=r}^{n} \left\{ \binom{n-1}{i-1} + \binom{n-1}{i} \right\} \operatorname{Per} \begin{pmatrix} \boldsymbol{F} \\ \boldsymbol{1-F} \end{pmatrix} \begin{matrix} \} \, i \\ \} \, n-i \end{matrix}
$$

$$
= \binom{n-1}{r-1} \operatorname{Per} \begin{pmatrix} \boldsymbol{F} \\ \boldsymbol{1-F} \end{pmatrix} \begin{matrix} \} \, r \\ \} \, n-r \end{matrix}
$$

$$
+ \sum_{i=r}^{n-1} \binom{n-1}{i} \operatorname{Per} \begin{pmatrix} \boldsymbol{F} \\ \boldsymbol{1-F} \end{pmatrix} \begin{matrix} \} \, i+1 \\ \} \, n-i-1 \end{matrix}
$$

$$
+ \sum_{i=r}^{n-1} \binom{n-1}{i} \operatorname{Per} \begin{pmatrix} \boldsymbol{F} \\ \boldsymbol{1-F} \end{pmatrix} \begin{matrix} \} \, i \\ \} \, n-i \end{matrix}
$$

$$
= \binom{n-1}{r-1} \operatorname{Per} \begin{pmatrix} \boldsymbol{F} \\ \boldsymbol{1-F} \end{pmatrix} \begin{matrix} \} \, r \\ \} \, n-r \end{matrix}
$$

$$
+ \sum_{i=r}^{n-1} \binom{n-1}{i} \sum_{|S|=n-1} \operatorname{Per} \begin{pmatrix} \boldsymbol{F} \\ \boldsymbol{1-F} \end{pmatrix} \begin{matrix} \} \, i \\ \} \, n-i-1 \end{matrix} [S] \cdot F[S^c]
$$

$$
+ \sum_{i=r}^{n-1} \binom{n-1}{i} \sum_{|S|=n-1} \operatorname{Per} \begin{pmatrix} \boldsymbol{F} \\ \boldsymbol{1-F} \end{pmatrix} \begin{matrix} \} \, i \\ \} \, n-i-1 \end{matrix} [S] \cdot (1-F)[S^c]
$$

$$
= \binom{n-1}{r-1} \operatorname{Per} \begin{pmatrix} \boldsymbol{F} \\ \boldsymbol{1-F} \end{pmatrix} \begin{matrix} \} \, r \\ \} \, n-r \end{matrix}
$$

$$
+ \sum_{i=r}^{n-1} \binom{n-1}{i} \sum_{|S|=n-1} \operatorname{Per} \begin{pmatrix} \boldsymbol{F} \\ \boldsymbol{1-F} \end{pmatrix} \begin{matrix} \} \, i \\ \} \, n-i-1 \end{matrix} [S]
$$

$$
= \binom{n-1}{r-1} \operatorname{Per} \begin{pmatrix} \boldsymbol{F} \\ \boldsymbol{1-F} \end{pmatrix} \begin{matrix} \} \, r \\ \} \, n-r \end{matrix} + (n-1)! \sum_{|S|=n-1} F_{r:S}(x). \qquad (51)
$$

In the above, $F[S^c]$ denotes the distribution function of the X-variable corresponding to the index $\{1, 2, \ldots, n\} \setminus S$ (viz., S^c). The relation in (50) then follows readily upon simplifying (51). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

5.3. Relations for pairs of order statistics

In this section, we establish several recurrence relations and identities satisfied by joint distributions of pairs of order statistics. These generalize many well-known results for order statistics in the IID case discussed in detail in [5, 6, 53, 57]. For a review of all these results, one may refer to [18, 44]. Even though most of the results in this section are presented in terms of joint densities or distribution functions, they hold equally well for the product moments (if they exist) of order statistics.

For convenience, let us denote

$$\mathbb{R}^2_U = \{(x, y) : -\infty < x \le y < \infty\}, \qquad \mathbb{R}^2_L = \{(x, y) : -\infty < y < x < \infty\},$$

and

$$\mathbb{R}^2 = \mathbb{R}^2_U \cup \mathbb{R}^2_L = \{(x, y) : -\infty < x < \infty, \ -\infty < y < \infty\}.$$

One may then note that the product moment of $X_{r:n}$ and $X_{s:n}$ can be written as

$$E(X_{r:n} X_{s:n}) = \iint\limits_{\mathbb{R}^2_U} xy \, f_{r,s:n}(x, y) \, dx \, dy, \qquad 1 \le r < s \le n,$$

and more generally

$$E\{g_1(X_{r:n}) \, g_2(X_{s:n})\} = \iint\limits_{\mathbb{R}^2_U} g_1(x) \, g_2(y) \, f_{r,s:n}(x, y) \, dx \, dy, \quad 1 \le r \le s \le n,$$

where $f_{r,s:n}(x, y)$ is as given in (13).

Result 5.17. *For $n \ge 2$,*

$$\sum_{r=1}^{n} \sum_{s=1}^{n} E(X_{r:n} X_{s:n}) = \sum_{i=1}^{n} \mathrm{Var}(X_i) + \left\{ \sum_{i=1}^{n} E(X_i) \right\}^2. \tag{52}$$

Proof. By starting with the identity

$$\sum_{r=1}^{n} \sum_{s=1}^{n} X_{r:n} X_{s:n} = \sum_{i=1}^{n} \sum_{j=1}^{n} X_i X_j = \sum_{i=1}^{n} X_i^2 + \sum\sum_{i \ne j} X_i X_j$$

and taking expectation on both sides, we get

$$\sum_{r=1}^{n} \sum_{s=1}^{n} E(X_{r:n} X_{s:n}) = \sum_{i=1}^{n} E(X_i^2) + \sum\sum_{i \ne j} E(X_i) \, E(X_j)$$

$$= \sum_{i=1}^{n} \{\mathrm{Var}(X_i) + [E(X_i)]^2\} + \sum\sum_{i \ne j} E(X_i) \, E(X_j)$$

which gives the identity in (52). □

Result 5.18. *For* $n \geq 2$,

$$\sum_{r=1}^{n-1} \sum_{s=r+1}^{n} E(X_{r:n} X_{s:n}) = \frac{1}{2} \left\{ \left[\sum_{i=1}^{n} E(X_i) \right]^2 - \sum_{i=1}^{n} [E(X_i)]^2 \right\}. \tag{53}$$

Proof. Since

$$\sum_{r=1}^{n} \sum_{s=1}^{n} E(X_{r:n} X_{s:n}) = \sum_{r=1}^{n} E(X_{r:n}^2) + 2 \sum_{r=1}^{n-1} \sum_{s=r+1}^{n} E(X_{r:n} X_{s:n})$$

and

$$\sum_{r=1}^{n} E(X_{r:n}^2) = \sum_{i=1}^{n} E(X_i^2),$$

the identity in (53) follows easily. □

Result 5.19 (Tetrahedron Rule). *For* $2 \leq r < s \leq n$ *and* $(x, y) \in \mathbb{R}_U^2$,

$$(r-1) \, f_{r,s:n}(x,y) + (s-r) \, f_{r-1,s:n}(x,y) + (n-s+1) \, f_{r-1,s-1:n}(x,y)$$
$$= n \, \overline{f}_{r-1,s-1:n-1}(x,y) \, . \tag{54}$$

Proof. By considering the expression of $(r-1) \, f_{r,s:n}(x,y)$ from (13) and expanding the permanent by its first row, we get

$$(r-1) \, f_{r,s:n}(x,y) = \sum_{i=1}^{n} F_i(x) \, f_{r-1,s-1:n-1}^{[i]}(x,y). \tag{55}$$

Next, by considering the expression of $(s-r) \, f_{r-1,s:n}(x,y)$ from (13) and expanding the permanent by its r-th row, we get

$$(s-r) \, f_{r-1,s:n}(x,y) = \sum_{i=1}^{n} \{F_i(y) - F_i(x)\} \, f_{r-1,s-1:n-1}^{[i]}(x,y). \tag{56}$$

Finally, by considering the expression of $(n-s+1) \, f_{r-1,s-1:n}(x,y)$ from (13) and expanding the permanent by its last row, we get

$$(n-s+1) \, f_{r-1,s-1:n}(x,y) = \sum_{i=1}^{n} \{1 - F_i(y)\} \, f_{r-1,s-1:n-1}^{[i]}(x,y). \tag{57}$$

Upon adding (55), (56), and (57), we get the recurrence relation in (54). □

Remark 5.20. It is easy to note that Result 5.19 will enable one to determine all the joint distributions of pairs of order statistics with the knowledge of $n-1$ suitably chosen ones like, for example, the distributions of contiguous order statistics $f_{r,r+1:n}(x,y)$ $(1 \le r \le n-1)$. This bound can be improved as shown in Theorem 5.30. For the IID case, Result 5.19 was proved in [60], and in its general INID form in [11,40].

By repeated application of Result 5.19, the following three recurrence relations can be proved as shown in [23].

Result 5.21. *For $1 \le r < s \le n$ and $(x,y) \in \mathbb{R}_U^2$,*

$$f_{r,s:n}(x,y) = \sum_{i=r}^{s-1} \sum_{j=n-s+i+1}^{n} (-1)^{j+n-r-s+1} \binom{i-1}{r-1} \binom{j-i-1}{n-s} \binom{n}{j} \overline{f}_{i,i+1:j}(x,y),$$

$$f_{r,s:n}(x,y) = \sum_{i=s-r}^{s-1} \sum_{j=n-s+i+1}^{n} (-1)^{n-j-r+1} \binom{i-1}{s-r-1} \binom{j-i-1}{n-s} \binom{n}{j} \overline{f}_{1,i+1:j}(x,y),$$

$$f_{r,s:n}(x,y) = \sum_{i=s-r}^{n-r} \sum_{j=r+i}^{n} (-1)^{s+j} \binom{i-1}{s-r-1} \binom{j-i-1}{r-1} \binom{n}{j} \overline{f}_{j-i,j:j}(x,y).$$

Theorem 5.22. *If either Result 5.19 or 5.21 is used in the computation of product moments of pairs of order statistics arising from n INID variables, then the identities in Results 5.17 and 5.18 will be automatically satisfied and hence should not be applied to check the computational process.*

Proof. We shall prove the theorem by starting with Result 5.19 and the proof for 5.21 is very similar. From (54), upon setting $r = 2$, $s = 3, 4, \ldots, n$, $r = 3$, $s = 4, 5, \ldots, n$, $\ldots$, and $r = n-1$, $s = n$, and adding the resulting $\binom{n-1}{2}$ equations, we get

$$\sum_{r=1}^{n-1} \sum_{s=r+1}^{n} E(X_{r:n} X_{s:n}) = \frac{1}{n-2} \sum_{i_1=1}^{n} \sum_{r=1}^{n-2} \sum_{s=r+1}^{n-1} E(X_{r:n-1}^{[i_1]} X_{s:n-1}^{[i_1]}).$$

By repeating this process, we obtain the identity in Result 5.18 which proves the theorem. □

The above theorem was proved in [34] for the IID case, and in [18] in the INID case.

Result 5.23. *For $1 \le r < s \le m \le n-1$ and $(x,y) \in \mathbb{R}_U^2$,*

$$\sum_{j=0}^{n-m} \sum_{k=j}^{n-m} \binom{r-1+j}{j} \binom{s-r-1+k-j}{k-j} \binom{n-s-k}{n-m-k} f_{r+j,s+k:n}(x,y)$$

$$= \binom{n}{m} \overline{f}_{r,s:m}(x,y).$$

This result can be proved by repeated application of the recurrence relation in (54). Note that Result 5.19 is a special case of this result when $m = n - 1$.

Result 5.24. *For $n \geq 2$ and $(x, y) \in \mathbb{R}_U^2$,*

$$\sum_{r=1}^{n-1} \sum_{s=r+1}^{n} \frac{1}{r} f_{r,s:n}(x,y) = n \sum_{r=1}^{n-1} \sum_{s=r+1}^{n} \frac{1}{(s-1)s} \overline{f}_{1,r+1:s}(x,y),$$

$$\sum_{r=1}^{n-1} \sum_{s=r+1}^{n} \frac{1}{s-r} f_{r,s:n}(x,y) = n \sum_{r=1}^{n-1} \sum_{s=r+1}^{n} \frac{1}{(s-1)s} \overline{f}_{r,r+1:s}(x,y),$$

and

$$\sum_{r=1}^{n-1} \sum_{s=r+1}^{n} \frac{1}{n-s+1} f_{r,s:n}(x,y) = n \sum_{r=1}^{n-1} \sum_{s=r+1}^{n} \frac{1}{(s-1)s} \overline{f}_{r,s:s}(x,y).$$

These identities are bivariate extensions of Joshi's identities and were established in [23]. They can be proved along the lines of Result 5.12.

Result 5.25. *For $1 \leq r < s \leq n$,*

$$E(X_{r:n}X_{s:n}) + \sum_{j=0}^{r-1} \sum_{k=0}^{n-s} (-1)^{n-j-k} \binom{j+k}{j} \sum_{|S|=n-j-k} E(X_{n-s-k+1:S}X_{n-r-k+1:S})$$

$$= \sum_{j=1}^{s-r} (-1)^{s-r-j} \binom{s-1-j}{r-1} \sum_{|S|=s-j} E(X_{s-j:S})\, E(X_{j:S^c}). \quad (58)$$

Proof. For $1 \leq r < s \leq n$, let us consider

$$I = \frac{1}{(r-1)!(s-r-1)!(n-s)!} \iint_{\mathbb{R}^2} xy \operatorname{Per} \begin{pmatrix} \boldsymbol{F}(x) \\ \boldsymbol{f}(x) \\ \boldsymbol{F}(y) - \boldsymbol{F}(x) \\ \boldsymbol{f}(y) \\ \boldsymbol{1} - \boldsymbol{F}(y) \end{pmatrix} \begin{matrix} \} \, r-1 \\ \} \, 1 \\ \} \, s-r-1 \\ \} \, 1 \\ \} \, n-s \end{matrix} \, dy\, dx \quad (59)$$

$$= \frac{1}{(r-1)!(s-r-1)!(n-s)!} \sum_{j=0}^{s-r-1} (-1)^{s-r-1-j} \binom{s-r-1}{j}$$

$$\times \iint_{\mathbb{R}^2} xy \operatorname{Per} \begin{pmatrix} \boldsymbol{F}(x) \\ \boldsymbol{f}(x) \\ \boldsymbol{F}(y) \\ \boldsymbol{f}(y) \\ \boldsymbol{1} - \boldsymbol{F}(y) \end{pmatrix} \begin{matrix} \} \, s-j-2 \\ \} \, 1 \\ \} \, j \\ \} \, 1 \\ \} \, n-s \end{matrix} \quad dy\, dx$$

$$= \frac{1}{(r-1)!(s-r-1)!(n-s)!} \sum_{j=0}^{s-r-1} (-1)^{s-r-1-j} \binom{s-r-1}{j}$$

$$\times \sum_{|S|=s-j-1} \int_{-\infty}^{\infty} x \operatorname{Per} \begin{pmatrix} \boldsymbol{F}(x) \\ \boldsymbol{f}(x) \end{pmatrix} \begin{array}{l} \} \, s-j-2 \\ \} \, 1 \end{array} [S] \, dx$$

$$\times \int_{-\infty}^{\infty} y \operatorname{Per} \begin{pmatrix} \boldsymbol{F}(y) \\ \boldsymbol{f}(y) \\ \boldsymbol{1}-\boldsymbol{F}(y) \end{pmatrix} \begin{array}{l} \} \, j \\ \} \, 1 \\ \} \, n-s \end{array} [S^c] \, dy$$

which, when simplified gives the R.H.S. of (58).

Alternatively, by noting that $\mathbb{R}^2 = \mathbb{R}_U^2 \cup \mathbb{R}_L^2$, we can write from (59) that

$$I = E(X_{r:n} X_{s:n}) + J, \tag{60}$$

where

$$J = \frac{1}{(r-1)!(s-r-1)!(n-s)!} \iint_{\mathbb{R}_L^2} xy \operatorname{Per} \begin{pmatrix} \boldsymbol{F}(x) \\ \boldsymbol{f}(x) \\ \boldsymbol{F}(y)-\boldsymbol{F}(x) \\ \boldsymbol{f}(y) \\ \boldsymbol{1}-\boldsymbol{F}(y) \end{pmatrix} \begin{array}{l} \} \, r-1 \\ \} \, 1 \\ \} \, s-r-1 \\ \} \, 1 \\ \} \, n-s \end{array} dx \, dy.$$

Upon writing the first $r-1$ rows in terms of $\boldsymbol{1}-\boldsymbol{F}(x)$ and the last $n-s$ rows in terms of $\boldsymbol{F}(y)$, we get

$$J = \frac{1}{(r-1)!(s-r-1)!(n-s)!} \sum_{j=0}^{r-1} \sum_{k=0}^{n-s} (-1)^{n-j-k} \binom{r-1}{j} \binom{n-s}{k}$$

$$\times \iint_{\mathbb{R}_L^2} xy \operatorname{Per} \begin{pmatrix} \boldsymbol{1} \\ \boldsymbol{F}(x) \\ \boldsymbol{f}(y) \\ \boldsymbol{F}(x)-\boldsymbol{F}(y) \\ \boldsymbol{f}(x) \\ \boldsymbol{1}-\boldsymbol{F}(x) \end{pmatrix} \begin{array}{l} \} \, j+k \\ \} \, n-s-k \\ \} \, 1 \\ \} \, s-r-1 \\ \} \, 1 \\ \} \, r-1-j \end{array} dx \, dy$$

$$= \frac{1}{(r-1)!(s-r-1)!(n-s)!} \sum_{j=0}^{r-1} \sum_{k=0}^{n-s} (-1)^{n-j-k} \binom{r-1}{j} \binom{n-s}{k}$$

$$\times \sum_{|S|=n-j-k} (j+k)! \iint_{\mathbb{R}_L^2} xy \operatorname{Per} \begin{pmatrix} \boldsymbol{F}(y) \\ \boldsymbol{f}(y) \\ \boldsymbol{F}(x)-\boldsymbol{F}(y) \\ \boldsymbol{f}(x) \\ \boldsymbol{1}-\boldsymbol{F}(x) \end{pmatrix} \begin{array}{l} \} \, n-s-k \\ \} \, 1 \\ \} \, s-r-1 \\ \} \, 1 \\ \} \, r-1-j \end{array} dx \, dy$$

$$= \frac{1}{(r-1)!(s-r-1)!(n-s)!} \sum_{j=0}^{r-1} \sum_{k=0}^{n-s} (-1)^{n-j-k} \binom{r-1}{j} \binom{n-s}{k}$$

$$\times \sum_{|S|=n-j-k} (j+k)!(n-s-k)!(s-r-1)!(r-1-j)!$$

$$\times E(X_{n-s-k+1:S} X_{n-r-k+1:S})$$

which, when simplified and substituted in (60), yields the L.H.S. of (58). Hence, the result.   □

Upon setting $s = r + 1$ in (58), we obtain the following result.

Result 5.26. *For $r = 1, 2, \ldots, n - 1$,*

$$E(X_{r:n} X_{r+1:n}) + (-1)^n E(X_{n-r:n} X_{n-r+1:n})$$

$$= \sum_{j=0}^{r-1} \sum_{k=1}^{n-r-1} (-1)^{n+1-j-k} \binom{j+k}{j} \sum_{|S|=n-j-k} E(X_{n-r-k:S} X_{n-r-k+1:S})$$

$$+ \sum_{j=1}^{r-1} (-1)^{n+1-j} \sum_{|S|=n-j} E(X_{n-r:S} X_{n-r+1:S}) + \sum_{|S|=r} E(X_{r:S}) E(X_{1:S^c}).$$

Similarly, upon setting $s = n - r + 1$ in (58), we obtain the following result.

Result 5.27. *For $r = 1, 2, \ldots, [n/2]$,*

$$\{1 + (-1)^n\} E(X_{r:n} X_{n-r+1:n})$$

$$= \sum_{j=0}^{r-1} \sum_{k=1}^{r-1} (-1)^{n+1-j-k} \binom{j+k}{j} \sum_{|S|=n-j-k} E(X_{r-k:S} X_{n-r-k+1:S})$$

$$+ \sum_{j=1}^{r-1} (-1)^{n+1-j} \sum_{|S|=n-j} E(X_{r:S} X_{n-r+1:S})$$

$$+ \sum_{j=1}^{n-2r+1} (-1)^{n+1-j} \binom{n-r-j}{r-1} \sum_{|S|=n-r+1-j} E(X_{n-r+1-j:S}) E(X_{j:S^c}). \quad (61)$$

In particular, upon setting $n = 2m$ and $r = 1$ in (61), we obtain the following relation.

Result 5.28. *For $m = 1, 2, \ldots$,*

$$2 E(X_{1:2m} X_{2m:2m}) = \sum_{j=1}^{2m-1} (-1)^{j-1} \sum_{|S|=2m-j} E(X_{2m-j:S}) E(X_{j:S^c}).$$

Similarly, upon setting $n = 2m$ and $r = m$ in (61), we obtain the following relation.

Result 5.29. *For $m = 1, 2, \ldots$,*

$$2\ E(X_{m:2m}X_{m+1:2m})$$

$$= \sum_{j=0}^{m-1} \sum_{k=1}^{m-1} (-1)^{j+k-1} \binom{j+k}{j} \sum_{|S|=2m-j-k} E(X_{m-k:S}X_{m-k+1:S})$$

$$+ \sum_{j=1}^{m-1} (-1)^{j-1} \sum_{|S|=2m-j} E(X_{m:S}X_{m+1:S}) + \sum_{|S|=m} E(X_{m:S})\ E(X_{1:S^c}).$$

Theorem 5.30. *In order to find the first two single moments and the product moments of all order statistics arising from n INID variables, given these moments of order statistics arising from $n-1$ and less INID variables (for all subsets of the n variables), one needs to find at most two single moments and $(n-2)/2$ product moments when n is even, and two single moments and $(n-1)/2$ product moments when n is odd.*

Proof. In view of Remark 5.3 or 5.9, it is sufficient to find two single moments in order to compute the first two single moments of all order statistics, viz., $E(X_{r:n})$ and $E(X_{r:n}^2)$ for $r = 1, 2, \ldots, n$. Also, as pointed out in Remark 5.20, the knowledge of $n-1$ immediate upper-diagonal product moments $E(X_{r:n}X_{r+1:n})$, $1 \le r \le n-1$, is sufficient for the calculation of all the product moments. For even values of n, say $n = 2m$, Results 5.26 and 5.29 imply that the knowledge of $(n-2)/2 = m-1$ of the immediate upper-diagonal product moments, viz., $E(X_{r:2m}X_{r+1:2m})$ for $r = 1, 2, \ldots, m-1$, is sufficient for the determination of all the product moments. Finally, for odd values of n, say, $n = 2m+1$, Result 5.26 implies that the knowledge of $(n-1)/2 = m$ of the immediate upper-diagonal product moments, viz., $E(X_{r:2m+1}X_{r+1:2m+1})$ for $r = 1, 2, \ldots, m$, is sufficient for the computation of all the product moments. $\qquad\square$

Remark 5.31. It is of interest to mention here that the bounds established for the number of single and product moments while determining the means, variances and covariances of order statistics arising from n INID variables are exactly the same as the bounds established in [73] for the IID case; see also [5].

5.4. Relations for covariances of order statistics

In this section, we establish several recurrence relations and identities satisfied by the covariances of order statistics. These generalize several well-known results on covariances of order statistics in the IID case discussed in detail in [5].

Result 5.32. *For $n \ge 2$,*

$$\sum_{r=1}^{n} \sum_{s=1}^{n} \mathrm{Cov}(X_{r:n}, X_{s:n}) = \sum_{i=1}^{n} \mathrm{Var}(X_i). \tag{62}$$

Proof. By writing

$$\sum_{r=1}^{n}\sum_{s=1}^{n}\operatorname{Cov}(X_{r:n},X_{s:n}) = \sum_{r=1}^{n}\sum_{s=1}^{n}E(X_{r:n}\,X_{s:n}) - \sum_{r=1}^{n}E(X_{r:n})\sum_{r=1}^{n}E(X_{s:n})$$

and then using Result 5.17 on the R.H.S., we get the identity in (62). $\qquad\square$

Result 5.33. *For $2 \le r < s \le n$,*

$$
\begin{aligned}
(r-1)\,&\operatorname{Cov}(X_{r:n},X_{s:n})\\
&+ (s-r)\operatorname{Cov}(X_{r-1:n},X_{s:n}) + (n-s+1)\operatorname{Cov}(X_{r-1:n},X_{s-1:n})\\
&= \sum_{i=1}^{n}\operatorname{Cov}(X_{r-1:n-1}^{[i]},X_{s-1:n-1}^{[i]})\\
&\quad + \sum_{i=1}^{n}\{E(X_{r-1:n-1}^{[i]}) - E(X_{r-1:n})\}\{E(X_{s-1:n-1}^{[i]}) - E(X_{s:n})\}. \qquad (63)
\end{aligned}
$$

Proof. Using Result 5.19, we have for $2 \le r < s \le n$

$$
\begin{aligned}
(r-1)\,&\operatorname{Cov}(X_{r:n},X_{s:n}) + (s-r)\operatorname{Cov}(X_{r-1:n},X_{s:n}) + (n-s+1)\operatorname{Cov}(X_{r-1:n},X_{s-1:n})\\
&= \sum_{i=1}^{n}\operatorname{Cov}(X_{r-1:n-1}^{[i]},X_{s-1:n-1}^{[i]}) + \sum_{i=1}^{n}E(X_{r-1:n-1}^{[i]})E(X_{s-1:n-1}^{[i]})\\
&\quad - (r-1)E(X_{r:n})\,E(X_{s:n}) - (s-r)E(X_{r-1:n})\,E(X_{s:n})\\
&\quad - (n-s+1)E(X_{r-1:n})\,E(X_{s-1:n})\\
&= \sum_{i=1}^{n}\operatorname{Cov}(X_{r-1:n-1}^{[i]},X_{s-1:n-1}^{[i]}) + \sum_{i=1}^{n}E(X_{r-1:n-1}^{[i]})E(X_{s-1:n-1}^{[i]})\\
&\quad - E(X_{s:n})\{(r-1)E(X_{r:n}) + (n-r+1)E(X_{r-1:n})\}\\
&\quad - (n-s+1)\,E(X_{r-1:n})\{E(X_{s-1:n}) - E(X_{s:n})\}\\
&= \sum_{i=1}^{n}\operatorname{Cov}(X_{r-1:n-1}^{[i]},X_{s-1:n-1}^{[i]}) + \sum_{i=1}^{n}E(X_{r-1:n-1}^{[i]})E(X_{s-1:n-1}^{[i]})\\
&\quad - E(X_{s:n})\sum_{i=1}^{n}E(X_{r-1:n-1}^{[i]}) - E(X_{r-1:n})\sum_{i=1}^{n}\{E(X_{s-1:n-1}^{[i]}) - E(X_{s:n})\}
\end{aligned}
$$

upon using Result 5.2. The relation in (63) is derived by simplifying the above equation. $\qquad\square$

The above result was established in its general form in [12].

Result 5.34. *For* $1 \le r \le n-1$ *and* $1 \le \ell \le n-r$,

$$\sum_{k=r+1}^{n-\ell+1} \binom{n-k}{\ell-1} E(X_{r:n}X_{k:n}) + \sum_{j=1}^{r} \sum_{k=r+1}^{r+\ell} \binom{k-j-1}{k-r-1}\binom{n-k}{n-\ell-r} E(X_{j:n}X_{k:n})$$

$$= \sum_{|S|=n-\ell} E(X_{r:S})E(X_{1:S^c}). \quad (64)$$

Proof. For $1 \le r \le n-1$ and $1 \le \ell \le n-r$, let us consider

$$I = \frac{1}{(r-1)!(\ell-1)!(n-r-\ell)!}$$

$$\times \iint_{\mathbb{R}^2} xy \operatorname{Per} \begin{pmatrix} \boldsymbol{F}(x) \\ \boldsymbol{f}(x) \\ \boldsymbol{1}-\boldsymbol{F}(x) \\ \boldsymbol{f}(y) \\ \boldsymbol{1}-\boldsymbol{F}(y) \end{pmatrix} \begin{matrix} \} \, r-1 \\ \} \, 1 \\ \} \, n-\ell-r \\ \} \, 1 \\ \} \, \ell-1 \end{matrix} \quad dx\,dy \qquad (65)$$

$$= \frac{1}{(r-1)!(\ell-1)!(n-r-\ell)!}$$

$$\times \sum_{|S|=n-\ell} \int_{-\infty}^{\infty} x \operatorname{Per} \begin{pmatrix} \boldsymbol{F}(x) \\ \boldsymbol{f}(x) \\ \boldsymbol{1}-\boldsymbol{F}(x) \end{pmatrix} \begin{matrix} \} \, r-1 \\ \} \, 1 \\ \} \, n-\ell-r \end{matrix} [S]\,dx$$

$$\times \int_{-\infty}^{\infty} y \operatorname{Per} \begin{pmatrix} \boldsymbol{f}(y) \\ \boldsymbol{1}-\boldsymbol{F}(y) \end{pmatrix} \begin{matrix} \} \, 1 \\ \} \, \ell-1 \end{matrix} [S^c]\,dy,$$

which, when simplified, yields the R.H.S. of (64).

Alternatively, by noting that $\mathbb{R}^2 = \mathbb{R}_U^2 \cup \mathbb{R}_L^2$ we may write I in (65) as

$$I = J_1 + J_2,$$

where J_1 and J_2 have the same expressions as I in (65) with the integration being over the regions $\mathbb{R}_U^2$ and $\mathbb{R}_L^2$ (instead of $\mathbb{R}^2$), respectively. By considering the expression for J_1 and writing $\boldsymbol{1}-\boldsymbol{F}(x)$ as $(\boldsymbol{F}(y)-\boldsymbol{F}(x)) + (\boldsymbol{1}-\boldsymbol{F}(y))$, we get

$$J_1 = \frac{1}{(r-1)!(\ell-1)!(n-r-\ell)!}$$

$$\times \sum_{j=0}^{n-r-\ell} \binom{n-r-\ell}{j} \iint_{\mathbb{R}_U^2} xy \operatorname{Per} \begin{pmatrix} \boldsymbol{F}(x) \\ \boldsymbol{f}(x) \\ \boldsymbol{F}(y)-\boldsymbol{F}(x) \\ \boldsymbol{f}(y) \\ \boldsymbol{1}-\boldsymbol{F}(y) \end{pmatrix} \begin{matrix} \} \, r-1 \\ \} \, 1 \\ \} \, j \\ \} \, 1 \\ \} \, n-r-1-j \end{matrix} \quad dx\,dy,$$

which, when simplified, gives the first term on the L.H.S. of (64). Similarly, by considering the expression for J_2 and writing $\boldsymbol{F}(x)$ as $\boldsymbol{F}(y) + (\boldsymbol{F}(x) - \boldsymbol{F}(y))$ and $\boldsymbol{1} - \boldsymbol{F}(y)$ as $(\boldsymbol{F}(x) - \boldsymbol{F}(y)) + (\boldsymbol{1} - \boldsymbol{F}(x))$, we get

$$
J_2 = \frac{1}{(r-1)!(\ell-1)!(n-r-\ell)!} \sum_{j=0}^{r-1} \sum_{k=0}^{\ell-1} \binom{r-1}{j}\binom{\ell-1}{k}
$$

$$
\times \iint_{\mathbb{R}_L^2} xy \operatorname{Per} \begin{pmatrix} \boldsymbol{F}(y) \\ \boldsymbol{f}(y) \\ \boldsymbol{F}(x) - \boldsymbol{F}(y) \\ \boldsymbol{f}(x) \\ \boldsymbol{1} - \boldsymbol{F}(x) \end{pmatrix} \begin{matrix} \} \, j \\ \} \, 1 \\ \} \, r-1-j+k \\ \} \, 1 \\ \} \, n-r-k-1 \end{matrix} \quad dx\,dy,
$$

which, when simplified, gives the second term on the L.H.S. of (64). Hence, the result. $\qquad\square$

Upon setting $r = 1$ in (64), we obtain the following result.

Result 5.35. *For* $1 \le \ell \le n-1$,

$$
\sum_{k=2}^{n-\ell+1} \binom{n-k}{\ell-1} E(X_{1:n} X_{k:n}) + \sum_{k=2}^{\ell+1} \binom{n-k}{n-\ell-1} E(X_{1:n} X_{k:n})
$$

$$
= \sum_{|S|=n-\ell} E(X_{1:S})\, E(X_{1:S^c}). \quad (66)
$$

Remark 5.36. If we replace ℓ by $n - \ell$ in Result 5.35, the relation in (66) remains unchanged and, therefore, there are exactly $[n/2]$ equations in $n-1$ product moments, viz., $E(X_{1:n} X_{s:n})$ for $s = 2, 3, \dots, n$. For even values of n, there are $n/2$ equations in $n-1$ unknowns and hence a knowledge of $(n-2)/2$ of these product moments is sufficient. For odd values of n, there are $(n-1)/2$ equations in $n-1$ unknowns and hence a knowledge of $(n-1)/2$ of these product moments is sufficient. These are exactly the same bounds as presented in Theorem 5.30 which is not surprising since the product moments $E(X_{1:n} X_{s:n})$, for $s = 2, 3, \dots, n$, are also sufficient for the determination of all the product moments through Result 5.19.

Upon setting $\ell = 1$ in (64), we obtain the following result.

Result 5.37. *For* $1 \le r \le n-1$,

$$
\sum_{k=r+1}^{n} E(X_{r:n} X_{k:n}) + \sum_{j=1}^{r} E(X_{j:n} X_{r+1:n}) = \sum_{|S|=n-1} E(X_{r:S})\, E(X_{1:S^c}).
$$

Similarly, upon setting $\ell = n - r$ in (64), we obtain the following result.

Result 5.38. *For* $1 \le r \le n-1$,

$$
E(X_{r:n}X_{r+1:n}) + \sum_{j=1}^{r} \sum_{k=r+1}^{n} \binom{k-j-1}{k-r-1} E(X_{j:n}X_{k:n})
$$
$$
= \sum_{|S|=r} E(X_{r:S})E(X_{1:S^c}).
$$

Result 5.39. *For* $1 \le r \le n-1$,

$$
\sum_{k=r+1}^{n} \mathrm{Cov}(X_{r:n}, X_{k:n}) + \sum_{j=1}^{r} \mathrm{Cov}(X_{j:n}, X_{r+1:n})
$$
$$
= \sum_{|S|=n-1} \{ E(X_{r:S}) - E(X_{r:n}) \} E(X_{1:S^c})
$$
$$
- \left(\sum_{j=1}^{r} E(X_{j:n}) \right) \{ E(X_{r+1:n}) - E(X_{r:n}) \}. \tag{67}
$$

Proof. From Result 5.37, we have for $1 \le r \le n-1$

$$
\sum_{k=r+1}^{n} \mathrm{Cov}(X_{r:n}, X_{k:n}) + \sum_{j=1}^{r} \mathrm{Cov}(X_{j:n}, X_{r+1:n})
$$
$$
= \sum_{|S|=n-1} E(X_{r:S})E(X_{1:S^c}) - E(X_{r:n}) \sum_{k=r+1}^{n} E(X_{k:n})
$$
$$
- E(X_{r+1:n}) \sum_{j=1}^{r} E(X_{j:n})
$$
$$
= \sum_{|S|=n-1} E(X_{r:S})E(X_{1:S^c}) - E(X_{r:n}) \sum_{i=1}^{n} E(X_i)
$$
$$
- \{ E(X_{r+1:n}) - E(X_{r:n}) \} \sum_{j=1}^{r} E(X_{j:n})
$$

from which the relation in (67) follows directly. $\square$

Upon setting $r = 1$ in (67), we obtain the following result.

Result 5.40. *For $n \geq 3$,*

$$2\,\mathrm{Cov}(X_{1:n}, X_{2:n}) + \sum_{k=3}^{n} \mathrm{Cov}(X_{1:n}, X_{k:n})$$

$$= \sum_{|S|=n-1} \{E(X_{1:S}) - E(X_{1:n})\}E(X_{1:S^c})$$

$$- E(X_{1:n})\{E(X_{2:n}) - E(X_{1:n})\}.$$

Upon setting $r = n - 1$ in (67), we obtain the following result.

Result 5.41. *For $n \geq 3$,*

$$\sum_{j=1}^{n-2} \mathrm{Cov}(X_{j:n}, X_{n:n}) + 2\,\mathrm{Cov}(X_{n-1:n}, X_{n:n})$$

$$= \sum_{|S|=n-1} \{E(X_{n-1:S}) - E(X_{n-1:n})\}E(X_{1:S^c})$$

$$- \{E(X_{n:n}) - E(X_{n-1:n})\}\left\{\sum_{i=1}^{n} E(X_i) - E(X_{n:n})\right\}.$$

5.5. Results for the symmetric case

In this section, we consider the special case when the variables X_i $(i = 1, 2, \ldots, n)$ are all symmetric about zero and present some recurrence relations and identities satisfied by the single and the product moments of order statistics. These results enable us to determine improved bounds (than the ones presented in Theorem 5.30) for the number of the single and the product moments to be determined for the calculation of means, variances and covariances of order statistics arising from n INID variables, assuming these quantities to be known for order statistics arising from $n-1$ (and less) variables (for all subsets of the n X_i's). These results generalize several well-known results for the IID case developed in [60, 68, 73], and presented in detail in [5, 53].

As already shown in section 4.3, when the X_i's are all symmetric about zero, then $-X_{r:n} \stackrel{d}{=} X_{n-r+1:n}$ and $(-X_{s:n}, -X_{r:n}) \stackrel{d}{=} (X_{n-r+1:n}, X_{n-s+1:n})$. From these distributional relations, it is clear that

$$E(X_{r:n}^k) = (-1)^k E(X_{n-r+1:n}^k), \tag{68}$$

$$E(X_{r:n}X_{s:n}) = E(X_{n-s+1:n}X_{n-r+1:n}), \tag{69}$$

$$\mathrm{Var}(X_{r:n}) = \mathrm{Var}(X_{n-r+1:n}),$$

and

$$\mathrm{Cov}(X_{r:n}, X_{s:n}) = \mathrm{Cov}(X_{n-s+1:n}, X_{n-r+1:n}).$$

Result 5.42. *For $m \geq 1$ and $k = 1, 2, \ldots,$*

$$E(X_{m:2m-1}^{[i]^k}) = 0 \qquad \text{for odd values of } k \tag{70}$$

and

$$E(X_{m:2m}^k) = \frac{1}{2m} \sum_{i=1}^{2m} E(X_{m:2m-1}^{[i]^k}) \qquad \text{for even values of } k. \tag{71}$$

Proof. Equation (70) follows simply from (68). Equation (71) follows from Result 5.2 upon using the fact that

$$E(X_{m+1:2m}^k) = E(X_{m:2m}^k) \qquad \text{for even values of } k,$$

observed easily from (68). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Result 5.43. *For $1 \leq r < s \leq n$,*

$$
\begin{aligned}
\{1 + (-1)^n\} & E(X_{r:n} X_{s:n}) \\
&= \sum_{j=0}^{r-1} \sum_{k=1}^{n-s} (-1)^{n-j-k-1} \binom{j+k}{j} \sum_{|S|=n-j-k} E(X_{n-s-k+1:S} X_{n-r-k+1:S}) \\
&\quad + \sum_{j=1}^{r-1} (-1)^{n-j-1} \sum_{|S|=n-j} E(X_{n-s+1:S} X_{n-r+1:S}) \\
&\quad + \sum_{j=1}^{s-r} (-1)^{s-r-j-1} \binom{s-1-j}{r-1} \sum_{|S|=s-j} E(X_{1:S}) \, E(X_{j:S^c}).
\end{aligned}
\tag{72}
$$

Proof. Equation (72) follows directly from Result 5.25 upon using the symmetry relations in (68) and (69). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Result 5.44. *For even values of n and $1 \leq r < s \leq n$,*

$$
\begin{aligned}
2 \, E(X_{r:n} X_{s:n}) &= 2 \, E(X_{n-s+1:n} X_{n-r+1:n}) \\
&= \sum_{j=0}^{r-1} \sum_{k=1}^{n-s} (-1)^{j+k-1} \binom{j+k}{j} \sum_{|S|=n-j-k} E(X_{n-s-k+1:S} X_{n-r-k+1:S}) \\
&\quad + \sum_{j=1}^{r-1} (-1)^{j-1} \sum_{|S|=n-j} E(X_{n-s+1:S} X_{n-r+1:S}) \\
&\quad + \sum_{j=1}^{s-r} (-1)^{s-r-j-1} \binom{s-1-j}{r-1} \sum_{|S|=s-j} E(X_{1:S}) \, E(X_{j:S^c}).
\end{aligned}
$$

The above relation follows simply from Result 5.43 when n is even.

Result 5.45. *For even values of* n,

$$2\,E(X_{1:n}X_{2:n}) = \sum_{k=1}^{n-2}(-1)^{k-1}\sum_{|S|=n-k}E(X_{1:S}X_{2:S}). \tag{73}$$

Proof. Equation (73) follows from Result 5.44 when we set $r = 1$ and $s = 2$ and use the facts that

$$\sum_{|S|=n-k}E(X_{n-k-1:S}X_{n-k:S}) = \sum_{|S|=n-k}E(X_{1:S}X_{2:S})$$

and

$$\sum_{|S|=1}E(X_{1:S}X_{j:S^c}) = 0$$

(since $E(X_i) = 0$, $i = 1, \dots, n$). $\hfill\square$

Theorem 5.46. *In order to find the first two single moments and the product moments of all order statistics arising from* n *INID symmetric variables, given these moments of order statistics arising from* $n-1$ *and less variables (for all subsets of the* n *variables), one needs to find at most one single moment when* n *is even, and one single moment and* $(n-1)/2$ *product moments when* n *is odd.*

Proof. In view of Results 5.2 and 5.42, it is sufficient to find one single moment ($E(X_{n:n}^2)$ for odd values of n and $E(X_{n:n})$ for even values of n) in order to compute the first two single moments of all order statistics. The theorem is then proved by simply noting that there is no need to find any product moment when n is even due to Result 5.44. $\hfill\square$

Open Problem 5.47. *For the case when* n *is even, the assumption of symmetry for the distributions of* X_i*'s reduced the upper bound for the number of product moments from* $(n-2)/2$ *to* 0*. However, when* n *is odd, the assumption of symmetry had no effect on the upper bound* $(n-1)/2$ *for the number of product moments. Is this upper bound the best in this case or can it be improved?*

5.6. Some comments

First of all, it should be mentioned that all the results presented in this section (even though stated for continuous random variables) hold equally well for INID discrete random variables. This may be proved by either starting with the permanent expressions of distributions of discrete order statistics, or establishing all the results in terms of distribution functions of order statistics (which cover both continuous and discrete cases) instead of density functions. These will generalize the corresponding results in the IID discrete case established in [7]; see also [81].

It is worth concluding this section by stating that all the results presented in terms of moments of order statistics may very well be established in terms of expectations of functions of order statistics (assuming that they exist), with minor changes in the ensuing results.

6. Additional results for order statistics from INID variables

6.1. Introduction

In the last section, we established several recurrence relations and identities for distributions of single order statistics and joint distributions of pairs of order statistics. We also presented bounds for the number of single and product moments to be determined for the calculation of means, variances and covariances of all order statistics, and also improvements to those bounds in the case when the underlying variables are all symmetric.

In this section, we discuss some additional properties satisfied by order statistics from INID variables. These include some formulae in the case of distributions closed under extrema, a duality principle in order statistics for reflective families, relationships between two related sets of INID variables, some inequalities among the distributions of order statistics, and simple expressions for the variances of trimmed and Winsorized means.

6.2. Results for distributions closed under extrema

Suppose a random variable X has an arbitrary distribution function $F(x)$. Let us define the following two families of distribution functions with a parameter λ:

$$\text{Family I}: \qquad F^{(\lambda)}(x) = (F(x))^\lambda, \qquad\qquad \lambda > 0, \qquad (74)$$

$$\text{Family II}: \qquad F_{(\lambda)}(x) = 1 - (1 - F(x))^\lambda, \qquad \lambda > 0. \qquad (75)$$

It is clear from (74) and (75) that Family I is the family of distributions closed under maxima, while Family II is the family of distributions closed under minima.

Now, suppose $X_1, X_2, \ldots, X_n$ are INID random variables from Family I with X_i having parameter λ_i. Then, it may be noted that

$$F_{|S|:S}(x) = \prod_{i \in S} F^{(\lambda_i)}(x) = \prod_{i \in S} (F(x))^{\lambda_i} = (F(x))^{\lambda_S}, \qquad (76)$$

where $\lambda_S = \sum_{i \in S} \lambda_i$. Then, from Result 5.7 we have

$$F_{r:n}(x) = \sum_{j=r}^{n} (-1)^{j-r} \binom{j-1}{r-1} \sum_{|S|=j} F_{|S|:S}(x). \qquad (77)$$

Upon using (76) in (77), we get

$$F_{r:n}(x) = \sum_{j=r}^{n}(-1)^{j-r}\binom{j-1}{r-1}\sum_{|S|=j}(F(x))^{\lambda_S}. \tag{78}$$

For example, if X_i's are distributed as power function with parameter λ_i and with cumulative distribution function

$$F_i(x) = x^{\lambda_i}, \qquad 0 < x < 1, \quad \lambda_i > 0,$$

Equation (78) gives the cumulative distribution function of $X_{r:n}$ as

$$F_{r:n}(x) = \sum_{j=r}^{n}(-1)^{j-r}\binom{j-1}{r-1}\sum_{|S|=j}x^{\lambda_S}, \quad 0 < x < 1,$$

the density function of $X_{r:n}$ as

$$f_{r:n}(x) = \sum_{j=r}^{n}(-1)^{j-r}\binom{j-1}{r-1}\sum_{|S|=j}\lambda_S\,x^{\lambda_S-1}, \quad 0 < x < 1,$$

and the single moments of $X_{r:n}$ as

$$E(X_{r:n}^k) = \sum_{j=r}^{n}(-1)^{j-r}\binom{j-1}{r-1}\sum_{|S|=j}\lambda_S/(\lambda_S+k), \quad k = 1, 2, \ldots .$$

Suppose $X_1, X_2, \ldots, X_n$ are INID random variables from Family II with X_i having parameter λ_i. Then, it may be noted that

$$F_{1:S}(x) = 1 - \prod_{i \in S}(1 - F_{(\lambda_i)}(x)) = 1 - \prod_{i \in S}(1 - F(x))^{\lambda_i}$$

$$= 1 - (1 - F(x))^{\lambda_S}, \tag{79}$$

where $\lambda_S = \sum_{i \in S}\lambda_i$. Then, from Result 5.8 we have

$$F_{r:n}(x) = \sum_{j=n-r+1}^{n}(-1)^{j-n+r-1}\binom{j-1}{n-r}\sum_{|S|=j}F_{1:S}(x). \tag{80}$$

Upon using (79) in (80), we get

$$F_{r:n}(x) = \sum_{j=n-r+1}^{n}(-1)^{j-n+r-1}\binom{j-1}{n-r}\sum_{|S|=j}\{1 - (1 - F(x))^{\lambda_S}\}$$

$$= 1 - \sum_{j=n-r+1}^{n}(-1)^{j-n+r-1}\binom{j-1}{n-r}\sum_{|S|=j}(1 - F(x))^{\lambda_S}, \tag{81}$$

due to the combinatorial identity

$$\sum_{j=n-r+1}^{n}(-1)^{j-n+r-1}\binom{j-1}{n-r}\binom{n}{j}=1.$$

For example, if X_i's are distributed as exponential with parameter λ_i and with cumulative distribution function

$$F_i(x)=1-e^{-\lambda_i x},\qquad x\geq 0,\quad \lambda_i>0,$$

Equation (81) gives the cumulative distribution function of $X_{r:n}$ as

$$F_{r:n}(x)=1-\sum_{j=n-r+1}^{n}(-1)^{j-n+r-1}\binom{j-1}{n-r}\sum_{|S|=j}e^{-\lambda_S x},\quad 0\leq x<\infty,$$

the density function of $X_{r:n}$ as

$$f_{r:n}(x)=\sum_{j=n-r+1}^{n}(-1)^{j-n+r-1}\binom{j-1}{n-r}\sum_{|S|=j}\lambda_S e^{-\lambda_S x},\quad 0\leq x<\infty,$$

and the single moments of $X_{r:n}$ as

$$E(X_{r:n}^k)=\sum_{j=n-r+1}^{n}(-1)^{j-n+r-1}\binom{j-1}{n-r}\sum_{|S|=j}k!/\lambda_S^k,\quad k=1,2,\dots$$

These results and many more examples are given in [39].

6.3. Duality principle in order statistics for reflective families

Let $\boldsymbol{V}=(X_1,X_2,\dots,X_n)$ be a random vector, $S\subset\{1,2,\dots,n\}$ and $_{\boldsymbol{V}}F_{\boldsymbol{r}:S}(\boldsymbol{x})$, with $\boldsymbol{r}=(r_1,r_2,\dots,r_k)$ and $\boldsymbol{x}=(x_1,x_2,\dots,x_k)$, be the joint cumulative distribution function of the k order statistics $X_{r_1:S},X_{r_2:S},\dots,X_{r_k:S}$ corresponding to the X_i, $i\in S$, with $1\leq r_1<r_2<\cdots<r_k\leq|S|$. Similarly, let $_{\boldsymbol{V}}\overline{F}_{\boldsymbol{r}:S}(\boldsymbol{x})$ be the joint survival function of the k order statistics $X_{r_1:S},X_{r_2:S},\dots,X_{r_k:S}$ corresponding to the X_i, $i\in S$, with $1\leq r_1<r_2<\cdots<r_k\leq|S|$.

Let $\mathcal{C}$ be a family of random vectors of dimension n such that, if $\boldsymbol{V}=(X_1,X_2,\dots,X_n)$ is in $\mathcal{C}$, then $\overline{\boldsymbol{V}}=(-X_1,-X_2,\dots,-X_n)$ is also in $\mathcal{C}$. Such a family $\mathcal{C}$ is referred to as a "reflective family." For example, the family consisting of all n-dimensional random vectors each of whose components are (a) discrete, (b) continuous, (c) absolutely continuous, (d) IID, (e) symmetric, (f) exchangeable, and (g) INID are all clearly reflective families. Similarly, any meaningful intersection of these collections is also a reflective family.

Then, the following theorem which proves a duality principle in order statistics for reflective families has been established in [38].

Theorem 6.1. *Suppose that a relation of the form*

$$\sum \boldsymbol{c}_{\boldsymbol{r}:S} \; _{\boldsymbol{V}} F_{\boldsymbol{r}:S}(\boldsymbol{x}) \equiv 0 \tag{82}$$

for all $\boldsymbol{V}$ in a reflective family $\mathcal{C}$, for every real $\boldsymbol{x}$, and where the summation is over all subsets S of $\{1, 2, \ldots, n\}$ and over $\boldsymbol{r} = (r_1, r_2, \ldots, r_k)$ with $1 \le r_1 < r_2 < \cdots < r_k \le |S|$, is satisfied. Then, the following dual relation is also satisfied by every $\boldsymbol{V} \in \mathcal{C}$:

$$\sum \boldsymbol{c}_{\boldsymbol{r}:S} \; _{\boldsymbol{V}} F_{\boldsymbol{R}:S}(\boldsymbol{x}) \equiv 0, \tag{83}$$

where $\boldsymbol{R} = (R_1, R_2, \ldots, R_k) = (|S| - r_k + 1, \; |S| - r_{k-1} + 1, \; \ldots, |S| - r_1 + 1)$.

Proof. By changing $\boldsymbol{V}$ to $\overline{\boldsymbol{V}}$ in (82), we simply obtain

$$\sum \boldsymbol{c}_{\boldsymbol{r}:S} \; _{\overline{\boldsymbol{V}}} F_{\boldsymbol{r}:S}(\boldsymbol{x}) = \sum \boldsymbol{c}_{\boldsymbol{r}:S} \; _{\boldsymbol{V}} \overline{F}_{\boldsymbol{R}:S}(-\boldsymbol{x}) \equiv 0. \tag{84}$$

Since the equality in (84) holds for every real $\boldsymbol{x}$, we immediately have

$$\sum \boldsymbol{c}_{\boldsymbol{r}:S} \; _{\boldsymbol{V}} \overline{F}_{\boldsymbol{R}:S}(\boldsymbol{x}) \equiv 0. \tag{85}$$

Now by writing

$$F_{X_1, X_2, \ldots, X_k}(\boldsymbol{x}) = 1 + \sum_{\ell=1}^{k} (-1)^\ell \sum_{1 \le i_1 < \cdots < i_\ell \le k} \overline{F}_{\boldsymbol{X}_{(i)}}(\boldsymbol{x}_{(i)}), \tag{86}$$

where $\boldsymbol{X}_{(i)} = (X_{i_1}, \ldots, X_{i_\ell})$, $\boldsymbol{x}_{(i)} = (x_{i_1}, \ldots, x_{i_\ell})$, and $\boldsymbol{R}_{(i)} = (R_{i_1}, R_{i_2}, \ldots, R_{i_\ell}) = (|S| - r_{i_\ell} + 1, \ldots, |S| - r_{i_1} + 1)$, and observing that (85) implies

$$\sum \boldsymbol{c}_{\boldsymbol{r}:S} = 0 \tag{87}$$

and

$$\sum \boldsymbol{c}_{\boldsymbol{r}:S} \; _{\boldsymbol{V}} \overline{F}_{\boldsymbol{R}_{(i)}:S}(\boldsymbol{x}_{(i)}) \equiv 0 \tag{88}$$

(by setting all or other x_i's as 0), the dual relation in (83) simply follows from (86) on using (87) and (88). $\qquad\square$

For illustration of this duality, let us consider Result 5.7 which gives for $1 \le r \le n - 1$ and $x \in \mathbb{R}$

$$F_{r:n}(x) = \sum_{j=r}^{n} (-1)^{j-r} \binom{j-1}{r-1} \sum_{|S|=j} F_{j:S}(x).$$

Upon using the duality principle, we simply obtain

$$F_{n-r+1:n}(x) = \sum_{j=r}^{n} (-1)^{j-r} \binom{j-1}{r-1} \sum_{|S|=j} F_{1:S}(x),$$

which readily yields the relation

$$F_{r:n}(x) = \sum_{j=n-r+1}^{n} (-1)^{j-n+r-1} \binom{j-1}{n-r} \sum_{|S|=j} F_{1:S}(x).$$

Note that this is exactly Result 5.8.

Similarly, let us consider the first identity in Result 5.12 which gives for $n \geq 2$ and $x \in \mathbb{R}$

$$\sum_{r=1}^{n} \frac{1}{r} F_{r:n}(x) = \sum_{r=1}^{n} \frac{1}{r\binom{n}{r}} \sum_{|S|=r} F_{1:S}(x).$$

Upon using the duality principle in the above identity, we simply obtain

$$\sum_{r=1}^{n} \frac{1}{r} F_{n-r+1:n}(x) = \sum_{r=1}^{n} \frac{1}{r\binom{n}{r}} \sum_{|S|=r} F_{r:S}(x),$$

which readily yields the identity

$$\sum_{r=1}^{n} \frac{1}{n-r+1} F_{r:n}(x) = \sum_{r=1}^{n} \frac{1}{r\binom{n}{r}} \sum_{|S|=r} F_{r:S}(x).$$

Note that this is exactly the second identity in Result 5.12.

Next, let us consider the second relation in Result 5.21 which gives for $1 \leq r < s \leq n$ and $(x,y) \in \mathbb{R}_U^2$

$$F_{r,s:n}(x,y) = \sum_{i=s-r}^{s-1} \sum_{j=n-s+1+i}^{n} (-1)^{n-r+1-j} \binom{i-1}{s-r-1} \binom{j-i-1}{n-s}$$
$$\times \sum_{|S|=j} F_{1,i+1:S}(x,y).$$

Upon using the duality principle in the above relation, we simply obtain

$$F_{n-s+1,n-r+1:n}(x,y) = \sum_{i=s-r}^{s-1} \sum_{j=n-s+1+i}^{n} (-1)^{n-r+1-j} \binom{i-1}{s-r-1} \binom{j-i-1}{n-s}$$
$$\times \sum_{|S|=j} F_{j-i,j:S}(x,y).$$

This readily gives

$$F_{r,s:n}(x,y) = \sum_{i=s-r}^{n-r} \sum_{j=r+i}^{n} (-1)^{s+j} \binom{i-1}{s-r-1} \binom{j-i-1}{r-1} \sum_{|S|=j} F_{j-i,j:S}(x,y),$$

which is exactly the last relation in Result 5.21.

There are many more such dual pairs among the results presented in section 5.

6.4. Results for two related sets of INID variables

Let $X_1, X_2, \ldots, X_n$ be INID random variables with X_i having probability density function $f_i(x)$ symmetric about 0 (without loss of any generality), and cumulative distribution function $F_i(x)$. Then, for $x \geq 0$ let

$$G_i(x) = 2\,F_i(x) - 1 \qquad \text{and} \qquad g_i(x) = 2\,f_i(x). \tag{89}$$

That is, the density functions $g_i(x)$, $i = 1, 2, \ldots, n$, are obtained by folding the density functions $f_i(x)$ at zero (the point of symmetry).

Let $Y_{1:n} \leq Y_{2:n} \leq \cdots \leq Y_{n:n}$ denote the order statistics obtained from n INID random variables $Y_1, Y_2, \ldots, Y_n$, with Y_i having probability density function $g_i(x)$ and cumulative distribution function $G_i(x)$ as given in (89).

In the IID case, some relationships among the moments of these two sets of order statistics were derived in [61]. These relations were then employed successfully in [62] to compute the moments of order statistics from the Laplace distribution by making use of the known results on the moments of order statistics from the exponential distribution. These results were extended in [10] to the case when the order statistics arise from a sample containing a single outlier. In [20], these results were used to examine the robustness properties of various linear estimators of the location and scale parameters of the Laplace distribution in the presence of a single outlier.

All these results were generalized in [13] to the case of INID variables, and these results are presented below.

Result 6.2. *For* $1 \leq r \leq n$ *and* $k = 1, 2, \ldots,$

$$E(X_{r:n}^k) = 2^{-n}\left\{\sum_{\ell=0}^{r-1} \sum_{|S|=n-\ell} E(Y_{r-\ell:S}^k) + (-1)^k \sum_{\ell=r}^{n} \sum_{|S|=\ell} E(Y_{\ell-r+1:S}^k)\right\}. \tag{90}$$

Proof. From (9), we have

$$E(X_{r:n}^k) = \frac{1}{(r-1)!(n-r)!} \int_{-\infty}^{\infty} x^k \operatorname{Per} \begin{pmatrix} \boldsymbol{F}(x) \\ \boldsymbol{f}(x) \\ \boldsymbol{1} - \boldsymbol{F}(x) \end{pmatrix} \begin{matrix} \} \, r-1 \\ \} \, 1 \\ \} \, n-r \end{matrix} \, dx$$

$$= \frac{1}{(r-1)!(n-r)!} \left\{ \int_{0}^{\infty} x^k \operatorname{Per} \begin{pmatrix} \boldsymbol{F}(x) \\ \boldsymbol{f}(x) \\ \boldsymbol{1} - \boldsymbol{F}(x) \end{pmatrix} \begin{matrix} \} \, r-1 \\ \} \, 1 \\ \} \, n-r \end{matrix} \, dx \right.$$

$$\left. + (-1)^k \int_{0}^{\infty} x^k \operatorname{Per} \begin{pmatrix} \boldsymbol{F}(x) \\ \boldsymbol{f}(x) \\ \boldsymbol{1} - \boldsymbol{F}(x) \end{pmatrix} \begin{matrix} \} \, n-r \\ \} \, 1 \\ \} \, r-1 \end{matrix} \, dx \right\} \tag{91}$$

upon using the symmetry properties $f_i(-x) = f_i(x)$ and $F_i(-x) = 1 - F_i(x)$, for

$i = 1, 2, \ldots, n$. Now, upon using (89) in (91), we get

$$
\begin{aligned}
E(X_{r:n}^k) &= \frac{2^{-n}}{(r-1)!(n-r)!} \left\{ \int_0^\infty x^k \operatorname{Per} \begin{pmatrix} \mathbf{1} + \mathbf{G}(x) \\ \mathbf{g}(x) \\ \mathbf{1} - \mathbf{G}(x) \end{pmatrix} \begin{matrix} \} \, r-1 \\ \} \, 1 \\ \} \, n-r \end{matrix} \, dx \right. \\
&\quad \left. + (-1)^k \int_0^\infty x^k \operatorname{Per} \begin{pmatrix} \mathbf{1} + \mathbf{G}(x) \\ \mathbf{g}(x) \\ \mathbf{1} - \mathbf{G}(x) \end{pmatrix} \begin{matrix} \} \, n-r \\ \} \, 1 \\ \} \, r-1 \end{matrix} \, dx \right\} \\
&= \frac{2^{-n}}{(r-1)!(n-r)!} \left\{ \int_0^\infty x^k \, I_{r-1,n-r}(x) \, dx \right. \\
&\quad \left. + (-1)^k \int_0^\infty x^k I_{n-r,r-1}(x) \, dx \right\},
\end{aligned}
\tag{92}
$$

where

$$
I_{r-1,n-r}(x) = \operatorname{Per} \begin{pmatrix} \mathbf{1} + \mathbf{G}(x) \\ \mathbf{g}(x) \\ \mathbf{1} - \mathbf{G}(x) \end{pmatrix} \begin{matrix} \} \, r-1 \\ \} \, 1 \\ \} \, n-r \end{matrix}
$$

and

$$
I_{n-r,r-1}(x) = \operatorname{Per} \begin{pmatrix} \mathbf{1} + \mathbf{G}(x) \\ \mathbf{g}(x) \\ \mathbf{1} - \mathbf{G}(x) \end{pmatrix} \begin{matrix} \} \, n-r \\ \} \, 1 \\ \} \, r-1 \end{matrix} .
$$

By expanding $I_{r-1,n-r}(x)$ by the first row, we obtain

$$
I_{r-1,n-r}(x) = \sum_{i_1=1}^n J_{0,r-2,n-r}^{[i_1]}(x) + J_{1,r-2,n-r}(x),
$$

where $J_{0,r-2,n-r}^{[i_1]}(x)$ is the permanent obtained from $I_{r-1,n-r}(x)$ by dropping the first row and the i_1-th column, and $J_{1,r-2,n-r}(x)$ is the permanent obtained from $I_{r-1,n-r}(x)$ by replacing the first row by $G(x)$. Proceeding in a similar way, we obtain

$$
I_{r-1,n-r}(x) = \sum_{\ell=0}^{r-1} (r-1-\ell)! \binom{r-1}{\ell} \sum_{|S|=n-\ell} \operatorname{Per} \begin{pmatrix} \mathbf{G}(x) \\ \mathbf{g}(x) \\ \mathbf{1} - \mathbf{G}(x) \end{pmatrix} \begin{matrix} \} \, r-1-\ell \\ \} \, 1 \\ \} \, n-r \end{matrix} \, [S]
$$

so that

$$
\frac{1}{(r-1)!(n-r)!} \int_0^\infty x^k \, I_{r-1,n-r}(x) \, dx = \sum_{\ell=0}^{r-1} \sum_{|S|=n-\ell} E(Y_{r-\ell:S}^k).
\tag{93}
$$

Proceeding exactly on the same lines, we also obtain

$$
\frac{1}{(r-1)!(n-r)!} \int_0^\infty x^k \, I_{n-r,r-1}(x) \, dx = \sum_{\ell=r}^n \sum_{|S|=\ell} E(Y_{\ell-r+1:S}^k).
\tag{94}
$$

Making use of the expressions in (93) and (94) on the R.H.S. of (92), we derive the relation in (90).   □

Result 6.3. *For* $1 \le r < s \le n$,

$$
E(X_{r:n}X_{s:n}) = 2^{-n} \Bigg\{ \sum_{\ell=0}^{r-1} \sum_{|S|=n-\ell} E(Y_{r-\ell:S}Y_{s-\ell:S})
$$

$$
- \sum_{\ell=r}^{s-1} \sum_{|S|=n-\ell} E(Y_{s-\ell:S})\, E(Y_{\ell-r+1:S^c})
$$

$$
+ \sum_{\ell=s}^{n} \sum_{|S|=\ell} E(Y_{\ell-s+1:S}Y_{\ell-r+1:S}) \Bigg\}. \tag{95}
$$

Proof. From (13), we have

$$
(r-1)!(s-r-1)!(n-s)!\, E(X_{r:n}X_{s:n})
$$

$$
= \iint\limits_{-\infty < x < y < \infty} xy \operatorname{Per} \begin{pmatrix} \boldsymbol{F}(x) \\ \boldsymbol{f}(x) \\ \boldsymbol{F}(y) - \boldsymbol{F}(x) \\ \boldsymbol{f}(y) \\ \boldsymbol{1} - \boldsymbol{F}(y) \end{pmatrix} \begin{matrix} \} \, r-1 \\ \} \, 1 \\ \} \, s-r-1 \\ \} \, 1 \\ \} \, n-s \end{matrix} \; dx\, dy
$$

$$
= \iint\limits_{0 < x < y < \infty} xy \operatorname{Per} \begin{pmatrix} \boldsymbol{F}(x) \\ \boldsymbol{f}(x) \\ \boldsymbol{F}(y) - \boldsymbol{F}(x) \\ \boldsymbol{f}(y) \\ \boldsymbol{1} - \boldsymbol{F}(y) \end{pmatrix} \begin{matrix} \} \, r-1 \\ \} \, 1 \\ \} \, s-r-1 \\ \} \, 1 \\ \} \, n-s \end{matrix} \; dx\, dy
$$

$$
+ \iint\limits_{0 < y < x < \infty} xy \operatorname{Per} \begin{pmatrix} \boldsymbol{1} - \boldsymbol{F}(x) \\ \boldsymbol{f}(x) \\ \boldsymbol{F}(x) - \boldsymbol{F}(y) \\ \boldsymbol{f}(y) \\ \boldsymbol{F}(y) \end{pmatrix} \begin{matrix} \} \, r-1 \\ \} \, 1 \\ \} \, s-r-1 \\ \} \, 1 \\ \} \, n-s \end{matrix} \; dx\, dy
$$

$$
- \int_0^\infty \int_0^\infty xy \operatorname{Per} \begin{pmatrix} \boldsymbol{1} - \boldsymbol{F}(x) \\ \boldsymbol{f}(x) \\ \boldsymbol{F}(y) - \boldsymbol{1} + \boldsymbol{F}(x) \\ \boldsymbol{f}(y) \\ \boldsymbol{1} - \boldsymbol{F}(y) \end{pmatrix} \begin{matrix} \} \, r-1 \\ \} \, 1 \\ \} \, s-r-1 \\ \} \, 1 \\ \} \, n-s \end{matrix} \; dx\, dy \tag{96}
$$

upon using the symmetry properties of $\boldsymbol{F}$. Now, upon using (89) in (96), we get

$$(r-1)!(s-r-1)!(n-s)!\,E(X_{r:n}X_{s:n})$$

$$= 2^{-n}\left\{ \iint\limits_{0<x<y<\infty} xy\,\mathrm{Per}\begin{pmatrix} \mathbf{1}+\boldsymbol{G}(x) \\ \boldsymbol{g}(x) \\ \boldsymbol{G}(y)-\boldsymbol{G}(x) \\ \boldsymbol{g}(y) \\ \mathbf{1}-\boldsymbol{G}(y) \end{pmatrix} \begin{matrix} \}\,r-1 \\ \}\,1 \\ \}\,s-r-1 \\ \}\,1 \\ \}\,n-s \end{matrix} \quad dx\,dy \right.$$

$$+ \iint\limits_{0<y<x<\infty} xy\,\mathrm{Per}\begin{pmatrix} \mathbf{1}+\boldsymbol{G}(y) \\ \boldsymbol{g}(y) \\ \boldsymbol{G}(x)-\boldsymbol{G}(y) \\ \boldsymbol{g}(x) \\ \mathbf{1}-\boldsymbol{G}(x) \end{pmatrix} \begin{matrix} \}\,n-s \\ \}\,1 \\ \}\,s-r-1 \\ \}\,1 \\ \}\,r-1 \end{matrix} \quad dx\,dy$$

$$\left. - \int_0^\infty\int_0^\infty xy\,\mathrm{Per}\begin{pmatrix} \mathbf{1}-\boldsymbol{G}(x) \\ \boldsymbol{g}(x) \\ \boldsymbol{G}(x)+\boldsymbol{G}(y) \\ \boldsymbol{g}(y) \\ \mathbf{1}-\boldsymbol{G}(y) \end{pmatrix} \begin{matrix} \}\,r-1 \\ \}\,1 \\ \}\,s-r-1 \\ \}\,1 \\ \}\,n-s \end{matrix} \quad dx\,dy \right\}. \quad (97)$$

The recurrence relation in (95) may be proved by expanding the three permanents on the R.H.S. of (97) as we did in proving Result 6.2 and then simplifying the resulting expressions. $\qquad\square$

Remark 6.4. If we set $F_1 = F_2 = \cdots = F_n = F$ and $f_1 = f_2 = \cdots = f_n = f$, Result 6.2 reduces to

$$E(X_{r:n}^k) = 2^{-n}\left\{ \sum_{\ell=0}^{r-1}\binom{n}{\ell}E(Y_{r-\ell:n-\ell}^k) + (-1)^k\sum_{\ell=r}^{n}\binom{n}{\ell}E(Y_{\ell-r+1:\ell}^k) \right\}$$

and Result 6.3 reduces to

$$E(X_{r:n}X_{s:n}) = 2^{-n}\left\{ \sum_{\ell=0}^{r-1}\binom{n}{\ell}E(Y_{r-\ell:n-\ell}Y_{s-\ell:n-\ell}) \right.$$

$$- \sum_{\ell=r}^{s-1}\binom{n}{\ell}E(Y_{s-\ell:n-\ell})\,E(Y_{\ell-r+1:\ell})$$

$$\left. + \sum_{\ell=s}^{n}\binom{n}{\ell}E(Y_{\ell-s+1:\ell}Y_{\ell-r+1:\ell}) \right\},$$

which were the relations derived in [61] for the IID case.

Remark 6.5. If we set $F_1 = F_2 = \cdots = F_{n-1} = F$ and $f_1 = f_2 = \cdots = f_{n-1} = f$, Results 6.2 and 6.3 reduce to the relations derived in [10] for the single-outlier model.

Remark 6.6. It should be mentioned here that Results 6.2 and 6.3 have been extended in [29] after relaxing the assumption of independence for the random variables X_i's.

6.5. Inequalities for distributions of order statistics

For the case when the X_i's are INID continuous random variables with pdf $f_i(x)$ and cdf $F_i(x)$, $i = 1, 2, \ldots, n$, some very interesting inequalities between the distribution of $X_{r:n}$ from $\boldsymbol{F} = (F_1 \, F_2 \cdots F_n)$ and the distribution of $X_{r:n}$ arising from an IID sample from a population with average distribution function $G(x) = \frac{1}{n} \sum_{i=1}^{n} F_i(x)$ have been established in [85]. We shall discuss these results in this section. Some of these results have also been presented in [53, pp. 22–24].

Let us assume that the p-th quantile of the distribution $G(\cdot)$ is uniquely given by $\overline{x}_p$; that is, $G(\overline{x}_p) = p$.

Theorem 6.7. *For $r = 2, 3, \ldots, n - 1$ and all $x < \overline{x}_{\frac{r-1}{n}} < \overline{x}_{\frac{r}{n}} \leq y$,*

$$\Pr\{x < X_{r:n} \leq y \mid \boldsymbol{F}\} \geq \Pr\{x < X_{r:n} \leq y \mid G\} \, , \tag{98}$$

where equality holds only if $F_1 = F_2 = \cdots = F_n = F$ at both x and y.

Proof. The proof of this theorem requires the following result in [67].

Let B_i $(i = 1, 2, \ldots, n)$ be n independent Bernoulli trials with B_i having p_i as the probability of success. Let $B = \sum_{i=1}^{n} B_i$ denote the number of successes in the n trials, and $\boldsymbol{p} = (p_1 \, p_2 \cdots p_n)$. Then, it was established in [67] that if $E(B) = np$ and c is an integer,

$$0 \leq \Pr(B \leq c \mid \boldsymbol{p}) \leq \Pr(B \leq c \mid p) \qquad \text{for } 0 \leq c \leq np - 1 \tag{99}$$

and

$$\Pr(B \leq c \mid p) \leq \Pr(B \leq c \mid \boldsymbol{p}) \leq 1 \qquad \text{for } np \leq c \leq n. \tag{100}$$

Now, by taking B_i's to be the indicator variables for the events $\{X_i \leq x\}$, taking $\boldsymbol{p} = \boldsymbol{F}(x)$ and $p = G(x)$, and observing that

$$\{B \geq r\} \qquad \text{and} \qquad \{X_{r:n} \leq x\}$$

are equivalent events, we obtain from (100) and (99) for the case when $c = r - 1$

$$\Pr\{X_{r:n} \leq x \mid \boldsymbol{F}\} \leq \Pr\{X_{r:n} \leq x \mid G\} \quad \text{for } nG(x) \leq r - 1 \text{ or } x \leq \overline{x}_{\frac{r-1}{n}} \tag{101}$$

and

$$\Pr\{X_{r:n} \leq y \mid \boldsymbol{F}\} \geq \Pr\{X_{r:n} \leq y \mid G\} \quad \text{for } r - 1 \leq nG(y) - 1 \text{ or } \overline{x}_{\frac{r}{n}} \leq y. \tag{102}$$

The inequality in (98) follows by subtracting (101) from (102). $\qquad \square$

Remark 6.8. For the case when $r = 1$ and $r = n$, (98) gives the inequalities

$$\Pr\{X_{1:n} \leq y \mid \boldsymbol{F}\} \geq \Pr\{X_{1:n} \leq y \mid G\} \qquad \text{for } y \geq \overline{x}_{\frac{1}{n}}$$

and

$$\Pr\{X_{n:n} \leq x \mid \boldsymbol{F}\} \leq \Pr\{X_{n:n} \leq x \mid G\} \qquad \text{for } x \leq \overline{x}_{\frac{n-1}{n}},$$

respectively. But, these two inequalities hold for all values as shown in the following theorem.

Theorem 6.9. *For $n \geq 2$ and all x,*

$$\Pr\{X_{1:n} \leq x \mid \boldsymbol{F}\} \geq \Pr\{X_{1:n} \leq x \mid G\} \tag{103}$$

and

$$\Pr\{X_{n:n} \leq x \mid \boldsymbol{F}\} \leq \Pr\{X_{n:n} \leq x \mid G\} \tag{104}$$

with equalities holding if and only if $F_1 = F_2 = \cdots = F_n = F$ at x.

Proof. For proving this theorem, we shall use the A.M.-G.M. (arithmetic mean-geometric mean) inequality given by

$$\left\{ \prod_{i=1}^{n} z_i \right\}^{1/n} \leq \overline{z} \tag{105}$$

with equality holding if and only if all z_i's are equal.

Since

$$\Pr\{X_{n:n} \leq x \mid \boldsymbol{F}\} = \prod_{i=1}^{n} F_i(x) \qquad \text{and} \qquad \Pr\{X_{n:n} \leq x \mid G\} = \{G(x)\}^n \, ,$$

the inequality in (104) follows readily from (105) by taking $z_i = F_i(x)$.

Also, since

$$\Pr\{X_{1:n} \leq x \mid \boldsymbol{F}\} = 1 - \Pr\{X_{1:n} > x \mid \boldsymbol{F}\} = 1 - \prod_{i=1}^{n} \{1 - F_i(x)\}$$

and

$$\Pr\{X_{1:n} \leq x \mid G\} = 1 - \Pr\{X_{1:n} > x \mid G\} = 1 - \{1 - G(x)\}^n,$$

the inequality in (103) follows easily from (105) by taking $z_i = 1 - F_i(x)$. $\qquad \square$

6.6. Variance of a trimmed mean

Simple expressions for the variance of a trimmed mean have been derived in [46, 54], with the former focusing on a symmetrically trimmed mean while the latter focusing on a general trimmed mean. These expressions, derived for the IID case, were generalized to the INID case in [28], and we present these formulas in this section and illustrate them with some examples.

From Result 5.37, we have for $1 \le i \le n-1$,

$$\sum_{j=i+1}^{n} \mu_{i,j:n} + \sum_{h=1}^{i} \mu_{h,i+1:n} = \sum_{h=1}^{n} \mu_{i:n-1}^{[h]} E(X_h), \tag{106}$$

where, as before, $\mu_{i:n-1}^{[h]}$ denotes the mean of the i-th order statistic among $n-1$ variables obtained by deleting X_h from the original n variables.

SCALE-OUTLIER MODEL

Suppose the n distributions have the same mean 0 (without loss of any generality) but have different variances σ_i^2 $(i = 1, 2, \ldots, n)$. In this case, (106) becomes

$$\sum_{j=i+1}^{n} \mu_{i,j:n} + \sum_{h=1}^{i} \mu_{h,i+1:n} = 0. \tag{107}$$

From (107), we readily obtain

$$\sum_{\substack{i=1 \\ i \ne j}}^{k} \sum_{j=1}^{n} \mu_{i,j:n} = -\sum_{i=1}^{k} \mu_{i,k+1:n} \tag{108}$$

and

$$\sum_{i=m}^{n} \sum_{j=1}^{n} \mu_{i,j:n} = -\sum_{i=m}^{n} \mu_{m-1,i:n}. \tag{109}$$

By using (108) and (109), it can be shown that

$$E\left(\sum_{i=k+1}^{m-1} X_{i:n} \right)^2 = \sum_{i=1}^{n} \sigma_i^2 - \sum_{i=1}^{k} \mu_{i,i:n} - \sum_{j=m}^{n} \mu_{j,j:n} + 2 \sum_{i=1}^{k} \sum_{j=i+1}^{k+1} \mu_{i,j:n}$$
$$+ 2 \sum_{i=m-1}^{n-1} \sum_{j=i+1}^{n} \mu_{i,j:n} + 2 \sum_{i=1}^{k} \sum_{j=m}^{n} \mu_{i,j:n}. \tag{110}$$

Now, let $T_n(k,\ell)$ denote the trimmed mean after deleting the smallest k and the largest ℓ order statistics, i.e.,

$$T_n(k,\ell) = \frac{1}{n-\ell-k} \sum_{i=k+1}^{n-\ell} X_{i:n}. \tag{111}$$

Then, from (110) we immediately have the variance of the trimmed mean in (111) as

$$\mathrm{Var}(T_n(k,\ell))$$
$$= \frac{1}{(n-\ell-k)^2} \Bigg[\sum_{i=1}^{n} \sigma_i^2 - \sum_{i=1}^{k} \mu_{i,i:n} - \sum_{j=n-\ell+1}^{n} \mu_{j,j:n} + 2\sum_{i=1}^{k} \sum_{j=i+1}^{k+1} \mu_{i,j:n}$$
$$+ 2\sum_{i=n-\ell}^{n-1} \sum_{j=i+1}^{n} \mu_{i,j:n} + 2\sum_{i=1}^{k} \sum_{j=n-\ell+1}^{n} \mu_{i,j:n}$$
$$- \bigg\{ \sum_{i=1}^{k} \mu_{i:n} + \sum_{j=n-\ell+1}^{n} \mu_{j:n} \bigg\}^2 \Bigg]. \tag{112}$$

Remark 6.10. For the p-outlier scale-model (case when $\sigma_1^2 = \cdots = \sigma_{n-p}^2 = \sigma^2$ and $\sigma_{n-p+1}^2 = \cdots = \sigma_n^2 = \tau^2$), the expression in (112) can be used just with $\sum_{i=1}^{n} \sigma_i^2$ replaced by $(n-p)\sigma^2 + p\tau^2$.

Remark 6.11. If all F_i's are symmetric about zero and $k = \ell$, then the variance of the resulting symmetrically trimmed mean can be simplified as

$$\mathrm{Var}(T_n(k,k))$$
$$= \frac{1}{(n-2k)^2} \Bigg[\sum_{i=1}^{n} \sigma_i^2 - 2\sum_{i=1}^{i} \mu_{i,i:n} + 4\sum_{i=1}^{k} \sum_{j=i+1}^{k+1} \mu_{i,j:n} + 2\sum_{i=1}^{k} \sum_{j=n-k+1}^{n} \mu_{i,j:n} \Bigg]. \tag{113}$$

This simplification is achieved by using the symmetry relations $\mu_{i:n} = -\mu_{n-i+1:n}$ and $\mu_{i,j:n} = \mu_{n-j+1,n-i+1:n}$.

LOCATION-OUTLIER MODEL

Suppose the n distributions have different means μ_i $(i = 1, 2, \ldots, n)$ but have same variance σ^2. In this case, (106) readily yields

$$\sum_{\substack{i=1 \\ i \neq j}}^{k} \sum_{j=1}^{n} \mu_{i,j:n} = \sum_{i=1}^{k} \sum_{h=1}^{n} \mu_{i:n-1}^{[h]} E(X_h) - \sum_{i=1}^{k} \mu_{i,k+1:n} \tag{114}$$

and

$$\sum_{\substack{i=m \\ i \neq j}}^{n} \sum_{j=1}^{n} \mu_{i,j:n} = \sum_{i=m-1}^{n-1} \sum_{h=1}^{n} \mu_{i:n-1}^{[h]} E(X_h) - \sum_{i=m}^{n} \mu_{m-1,i:n}. \tag{115}$$

By using (114) and (115), it can be shown that

$$
\begin{aligned}
E\left(\sum_{i=k+1}^{m-1} X_{i:n}\right)^2 = {} & n\sigma^2 + \left(\sum_{i=1}^{n} \mu_i\right)^2 - \sum_{i=1}^{k} \mu_{i,i:n} - \sum_{j=m}^{n} \mu_{j,j:n} \\
& + 2\sum_{i=1}^{k} \sum_{j=i+1}^{k+1} \mu_{i,j:n} + 2\sum_{i=m-1}^{n-1} \sum_{j=i+1}^{n} \mu_{i,j:n} \\
& + 2\sum_{i=1}^{k} \sum_{j=m}^{n} \mu_{i,j:n} - 2\sum_{i=1}^{k} \sum_{h=1}^{n} \mu_{i:n-1}^{[h]} E(X_h) \\
& - 2\sum_{j=m-1}^{n-1} \sum_{h=1}^{n} \mu_{j:n-1}^{[h]} E(X_h).
\end{aligned}
\tag{116}$$

From (116), we immediately have the variance of the trimmed mean in (111) as

$$
\begin{aligned}
\mathrm{Var}(T_n(k,\ell)) = {} & \frac{1}{(n-\ell-k)^2}\Bigg[n\sigma^2 + \left(\sum_{i=1}^{n} \mu_i\right)^2 - \sum_{i=1}^{k} \mu_{i,i:n} - \sum_{j=n-\ell+1}^{n} \mu_{j,j:n} \\
& + 2\sum_{i=1}^{k} \sum_{j=i+1}^{k+1} \mu_{i,j:n} + 2\sum_{i=n-\ell}^{n-1} \sum_{j=i+1}^{n} \mu_{i,j:n} \\
& + 2\sum_{i=1}^{k} \sum_{j=n-\ell+1}^{n} \mu_{i,j:n} - 2\sum_{i=1}^{k} \sum_{h=1}^{n} \mu_{i:n-1}^{[h]} E(X_h) \\
& - 2\sum_{j=n-\ell}^{n-1} \sum_{h=1}^{n} \mu_{j:n-1}^{[h]} E(X_n) \\
& - \left\{\sum_{i=1}^{n} \mu_i - \sum_{i=1}^{k} \mu_{i:n} - \sum_{j=k-\ell+1}^{n} \mu_{j:n}\right\}^2\Bigg].
\end{aligned}
\tag{117}$$

Remark 6.12. For the *p*-outlier location-model (case when $\mu_1 = \cdots = \mu_{n-p} = 0$, $\mu_{n-p+1} = \cdots = \mu_n = \lambda$, and $\sigma_1^2 = \cdots = \sigma_n^2 = \sigma^2$), the expression in (117) simplifies

to

$$
\mathrm{Var}(T_n(k,\ell)) = \frac{1}{(n-\ell-k)^2}\left[n\sigma^2 + p^2\lambda^2 - \sum_{i=1}^{k}\mu_{i,i:n} - \sum_{j=n-\ell+1}^{n}\mu_{j,j:n} \right.
$$

$$
+ 2\sum_{i=1}^{k}\sum_{j=i+1}^{k+1}\mu_{i,j:n} + 2\sum_{i=n-\ell}^{n-1}\sum_{j=i+1}^{n}\mu_{i,j:n}
$$

$$
+ 2\sum_{i=1}^{k}\sum_{j=n-\ell+1}^{n}\mu_{i,j:n} - 2p\lambda\sum_{i=1}^{k}\mu_{i:n-1}[p-1]
$$

$$
\left. - 2p\lambda\sum_{j=n-\ell}^{n-1}\mu_{j:n-1}[p-1] - \left\{ p\lambda - \sum_{i=1}^{k}\mu_{i:n} - \sum_{j=k-\ell+1}^{n}\mu_{j:n} \right\}^2 \right], \quad (118)
$$

where $\mu_{i:n-1}[p-1]$ denotes the mean of the i-th order statistic in a sample of size $n-1$ containing $p-1$ location-outliers.

Remark 6.13. The results presented here reduce to those in [54] for the special case when there are no outliers in the sample; in this situation, the result for the symmetric case presented in Remark 6.11 becomes the same as the formula derived in [46] using an entirely different method.

ILLUSTRATIVE EXAMPLES

For the normal case, the means, variances and covariances of order statistics from a single-outlier model have been tabulated in [56] (determined through extensive numerical integration). Tables for the two cases when (i) the single outlier is a location-outlier, and (ii) the single outlier is a scale-outlier, have been presented and these tables cover all sample sizes up to 20 and different choices of λ and τ.

These tables have been used in [58] for robustness studies and, in particular, for determining the exact bias and variance of various L-estimators including the trimmed mean in (111); see also [5,53]. Their computation of the variance of the trimmed mean made use of the variance-covariance matrix of the order statistics $X_{k+1:n}, \ldots, X_{n-\ell:n}$ under the outlier-model.

However, it should be noted that the trimming involved in the trimmed mean based on robustness considerations is often light (i.e., k and ℓ are small); see, for example, [3]. In such a situation, the expressions in (112) and (118) will be a lot more convenient to use in order to compute the variance of the lightly trimmed mean when the sample contains multiple scale-outliers and location-outliers, respectively.

Example 6.14. Consider the case when $n = 10$, and X_i ($i = 1, 2, \ldots, 9$) are standard normal and X_{10} is distributed as normal with mean 0 and variance 16; that is, we have a single scale-outlier model with $\tau = 4$.

In this case, by using the entire variance-covariance matrix of the order statistics $X_{2:10}, \ldots, X_{9:10}$ taken from the tables in [56], it was determined in [58] that

$$\mathrm{Var}(T_{10}(1,1)) = 0.1342.$$

On the other hand, the expression in (113) gives

$$\begin{aligned}
\mathrm{Var}(T_{10}(1,1)) &= \frac{1}{64}\{9 + \tau^2 - 2\mu_{1,1:10} + 4\mu_{1,2:10} + 2\mu_{1,10:10}\} \\
&= \frac{1}{64}\{25 - 2(9.6007396) + 4(3.0600175) + 2(-4.7257396)\} \\
&= 0.13417.
\end{aligned}$$

Example 6.15. Consider the case when $n = 10$, and X_i $(i = 1, 2, \ldots, 9)$ are standard normal and X_{10} is distributed as normal with mean 4 and variance 1; that is, we have a single location-outlier with $\lambda = 4$.

In this case, once again by using the entire variance-covariance matrix of the order statistics $X_{2:10}, \ldots, X_{9:10}$ taken from the tables in [56], it was determined in [58] that

$$\mathrm{Var}(T_{10}(1,1)) = 0.1145.$$

The expression in (118), on the other hand, gives

$$\begin{aligned}
\mathrm{Var}(T_{10}(1,1)) &= \frac{1}{16}\{26 - \mu_{1,1:10} - \mu_{10,10:10} + 2\mu_{1,2:10} \\
&\quad + 2\mu_{9,10:10} + 2\mu_{1,10:10} - (4 - \mu_{1:10} - \mu_{10:10})^2\} \\
&= \frac{1}{64}\{26 - 2.562625 - 17.031655 + 2(1.5625655) \\
&\quad + 2(5.9421416) + 2(-5.950589) - 2.1833018\} \\
&= 0.11454.
\end{aligned}$$

6.7. Variance of a Winsorized mean

Let us consider the general Winsorized mean

$$\begin{aligned}
W_n(r,s) &= \frac{1}{n}\Bigg\{\sum_{i=r+1}^{n-s} X_{i:n} + rX_{r+1:n} + sX_{n-s:n}\Bigg\} \\
&= \frac{1}{n}\Bigg\{\sum_{i=1}^{n} X_{i:n} + rX_{r+1:n} + sX_{n-s:n} - \sum_{i=1}^{r} X_{i:n} - \sum_{i=n-s+1}^{n} X_{i:n}\Bigg\}. \quad (119)
\end{aligned}$$

Then, by proceeding as in the last section, simple expressions for the mean and variance of the general Winsorized mean in (119) for the INID case were derived in [30], and these are described in this section and illustrated with some examples.

From (119), we readily find the mean of $W_n(r,s)$ as

$$E[W_n(r,s)] = \frac{1}{n}\left\{\sum_{i=1}^{n}\mu_i + \sum_{i=1}^{r}(\mu_{r+1:n}-\mu_{i:n}) - \sum_{i=n-s+1}^{n}(\mu_{i:n}-\mu_{n-s:n})\right\}, \qquad (120)$$

where X_i's are INID random variables with $E(X_i) = \mu_i$ and $\mathrm{Var}(X_i) = \sigma_i^2$, $i = 1,\ldots,n$. Further, upon using (106), (114), and (115), it can be shown that

$$\begin{aligned}
E[W_n(r,s)]^2 = \frac{1}{n^2}\Bigg[&\sum_{i=1}^{n}\sigma_i^2 + \left(\sum_{i=1}^{n}\mu_i\right)^2 + r(r+2)\mu_{r+1:n}^{(2)} + s(s+2)\mu_{n-s:n}^{(2)} \\
&- \sum_{i=1}^{r}\mu_{i:n}^{(2)} - \sum_{i=n-s+1}^{n}\mu_{i:n}^{(2)} + 2\sum_{i=1}^{r}\sum_{j=i+1}^{r+1}\mu_{i,j:n} + 2\sum_{i=n-s}^{n-1}\sum_{j=i+1}^{n}\mu_{i,j:n} \\
&- 2r\sum_{i=1}^{r+1}\mu_{i,r+2:n} - 2s\sum_{i=n-s}^{n}\mu_{n-s-1,i:n} + 2rs\mu_{r+1,n-s:n} \\
&- 2r\sum_{i=n-s+1}^{n}\mu_{r+1,i:n} - 2s\sum_{i=1}^{r}\mu_{i,n-s:n} + 2\sum_{i=1}^{r}\sum_{j=n-s+1}^{n}\mu_{i,j:n} \\
&+ 2\sum_{i=1}^{r}\sum_{h=1}^{n}\mu_h\left(\mu_{r+1:n-1}^{[h]} - \mu_{i:n-1}^{[h]}\right) \\
&- 2\sum_{i=n-s}^{n-1}\sum_{h=1}^{n}\mu_h\left(\mu_{i:n-1}^{[h]} - \mu_{n-s-1:n-1}^{[h]}\right)\Bigg]. \qquad (121)
\end{aligned}$$

From the expressions in (120) and (121), the variance of $W_n(r,s)$ can be readily computed as

$$\mathrm{Var}(W_n(r,s)) = E[W_n(r,s)]^2 - \{E[W_n(r,s)]\}^2.$$

For the remainder of this section, we shall focus on symmetrically Winsorized mean, i.e., the case when $r = s$.

SCALE-OUTLIER MODEL

Suppose the n distributions are all symmetric about 0 (without loss of any generality) but have different variances, say, $\sigma_1^2 = \cdots = \sigma_{n-p}^2 = \sigma^2$ and $\sigma_{n-p+1}^2 = \cdots = \sigma_n^2 = \tau^2$. In this case, because of the symmetry relationships

$$\mu_{i:n}^{(k)} = (-1)^k \mu_{n-i+1:n}^{(k)} \qquad \text{and} \qquad \mu_{i,j:n} = \mu_{n-j+1,n-i+1:n},$$

the expressions in (120) and (121) reduce to

$$E[W_n(r,r)] = \frac{1}{n}\left\{r\mu_{r+1:n} - \sum_{i=1}^{r}\mu_{i:n} + \sum_{i=1}^{r}\mu_{i:n} - r\mu_{r+1:n}\right\} = 0,$$

which simply implies that the symmetrically Winsorized mean $W_n(r,r)$ is an unbiased estimator, and

$$\mathrm{Var}(W_n(r,r)) = E[W_n(r,r)]^2$$

$$= \frac{1}{n^2}\Bigg[(n-p)\sigma^2 + p\tau^2 + 2r(r+2)\mu_{r+1:n}^{(2)} - 2\sum_{i=1}^{r}\mu_{i:n}^{(2)}$$

$$+ 4\sum_{i=1}^{r}\sum_{j=i+1}^{r+1}\mu_{i,j:n} - 4r\sum_{i=1}^{r+1}\mu_{i,r+2:n} + 2r^2\mu_{r+1,n-r:n}$$

$$- 4r\sum_{i=1}^{r}\mu_{i,n-r:n} + 2\sum_{i=1}^{r}\sum_{j=n-r+1}^{n}\mu_{i,j:n}\Bigg], \tag{122}$$

respectively.

LOCATION-OUTLIER MODEL

Suppose the n distributions have different means, say, $\mu_1 = \cdots = \mu_{n-p} = 0$ and $\mu_{n-p+1} = \cdots = \mu_n = \lambda$, and same variance $\sigma_1^2 = \cdots = \sigma_n^2 = \sigma^2$. In this case, the expressions in (120) and (121) reduce to

$$E[W_n(r,r)] = \frac{1}{n}\Bigg\{p\lambda + \sum_{i=1}^{r}(\mu_{r+1:n} - \mu_{i:n}) - \sum_{i=n-r+1}^{n}(\mu_{i:n} - \mu_{n-r:n})\Bigg\} \tag{123}$$

and

$$E[W_n(r,r)]^2 = \frac{\sigma^2}{n} + \frac{1}{n^2}\Bigg[p^2\lambda^2 + r(r+2)\{\mu_{r+1:n}^{(2)} + \mu_{n-r:n}^{(2)}\}$$

$$- \sum_{i=1}^{r}\mu_{i:n}^{(2)} - \sum_{i=n-r+1}^{n}\mu_{i:n}^{(2)} + 2\sum_{i=1}^{r}\sum_{j=i+1}^{r+1}\mu_{i,j:n} + 2\sum_{i=n-r}^{n-1}\sum_{j=i+1}^{n}\mu_{i,j:n}$$

$$- 2r\sum_{i=1}^{r+1}\mu_{i,r+2:n} - 2r\sum_{i=n-r}^{n}\mu_{n-r-1,i:n} + 2r^2\mu_{r+1,n-r:n}$$

$$- 2r\sum_{i=n-r+1}^{n}\mu_{r+1,i:n} - 2r\sum_{i=1}^{r}\mu_{i,n-r:n} + 2\sum_{i=1}^{r}\sum_{j=n-r+1}^{r}\mu_{i,j:n}$$

$$+ 2p\lambda\sum_{i=1}^{r}(\mu_{r+1:n-1}[p-1] - \mu_{i:n-1}[p-1])$$

$$- 2p\lambda\sum_{i=n-r}^{n-1}(\mu_{i:n-1}[p-1] - \mu_{n-r-1:n-1}[p-1])\Bigg], \tag{124}$$

respectively.

ILLUSTRATIVE EXAMPLES

Example 6.16. Suppose we have a sample of size 10, out of which nine are distributed as $N(\mu, 1)$ and one outlier is distributed as $N(\mu + \lambda, 1)$. Then, by using the tables of means, variances and covariances of order statistics from a single location-outlier normal model presented in [56], the bias and mean square error of $W_{10}(1, 1)$ and $W_{10}(2, 2)$ were computed in [30] from (123) and (124) as follows:

	$\lambda = 0.5$	$\lambda = 1.0$	$\lambda = 1.5$	$\lambda = 2.0$	$\lambda = 3.0$	$\lambda = 4.0$
Bias($W_{10}(1, 1)$)	0.04937	0.09505	0.13368	0.16298	0.19406	0.20239
Bias($W_{10}(2, 2)$)	0.04889	0.09155	0.12392	0.14500	0.16216	0.16503
MSE($W_{10}(1, 1)$)	0.10693	0.11403	0.12404	0.13469	0.15038	0.15627
MSE($W_{10}(2, 2)$)	0.11402	0.12106	0.12997	0.13805	0.14715	0.14926

Example 6.17. Suppose we have a sample of size 10, out of which nine are distributed as $N(\mu, 1)$ and one outlier is distributed as $N(\mu, \tau^2)$. Once again, by using the tables of means, variances and covariances of order statistics from a single scale-outlier normal model presented in [56], the variance of the unbiased estimators $W_{10}(1, 1)$ and $W_{10}(2, 2)$ were computed in [30] from (122) as follows:

	$\tau = 0.5$	$\tau = 2.0$	$\tau = 3.0$	$\tau = 4.0$
Var($W_{10}(1, 1)$)	0.09570	0.12214	0.13222	0.13802
Var($W_{10}(2, 2)$)	0.09972	0.12668	0.13365	0.13743

Remark 6.18. These values agree (up to 5 decimal places) in almost all cases with those in [5, pp. 128–130]. The differences that exist are in the fifth decimal place and may be due to the computational error accumulated in the computations in [5], since all the elements of the variance-covariance matrix of order statistics were used in the calculations there. For example, in the case when $n = 10$ and $r = 1$, while the expression in (122) would use only 6 elements of the variance-covariance matrix, the direct computation carried out in [5] would have used $(8 \times 9)/2 = 36$ elements.

7. Robust estimation for exponential distribution

7.1. Introduction

In the last three sections, numerous results have been presented on distributions and moments of order statistics from INID variables, with a special emphasis in many cases on results for multiple-outlier models. In their book *Outliers in Statistical Data*, Barnett and Lewis [43, p. 68] have stated

> "A study of the multiple-outlier model has been recently carried out by Balakrishnan, who gives a substantial body of results on the moments of order statistics. . . . He indicated that these results can in principle be applied to robustness studies in the multiple-outlier situation, but at the time of writing, we are not aware of any published application. There is much work waiting to be done in this important area."

Subsequently, the permanent approach was used successfully in [16] along with a differential equation technique to develop a simple and efficient recursive algorithm for the computation of single and product moments of order statistics from INID exponential random variables. This algorithm was then utilized to address the robust estimation of the exponential mean when multiple outliers are possibly present in the sample. Here, we describe these developments and their applications to robustness issues.

Consider $X_1, \ldots, X_n$ to be INID exponential random variables with X_i (for $i = 1, \ldots, n$) having probability density function

$$f_i(x) = \frac{1}{\theta_i} e^{-x/\theta_i}, \qquad x \geq 0, \quad \theta_i > 0, \tag{125}$$

and cumulative distribution function

$$F_i(x) = 1 - e^{-x/\theta_i}, \qquad x \geq 0, \quad \theta_i > 0. \tag{126}$$

It is clear from (125) and (126) that the distributions satisfy the differential equations

$$f_i(x) = \frac{1}{\theta_i}\{1 - F_i(x)\}, \qquad x \geq 0, \quad \theta_i > 0, \quad i = 1, \ldots, n. \tag{127}$$

7.2. Relations for single moments

The following theorem has been established in [16] for the single moments of order statistics by using the differential equation in (127).

Theorem 7.1. *For $n = 1, 2, \ldots$ and $k = 0, 1, 2, \ldots$,*

$$\mu_{1:n}^{(k+1)} = \frac{k+1}{\left(\sum_{i=1}^{n} 1/\theta_i\right)} \mu_{1:n}^{(k)}; \tag{128}$$

for $2 \leq r \leq n$ and $k = 0, 1, 2, \ldots$,

$$\mu_{r:n}^{(k+1)} = \frac{1}{\left(\sum_{i=1}^{n} 1/\theta_i\right)} \left\{ (k+1)\mu_{r:n}^{(k)} + \sum_{i=1}^{n} \frac{1}{\theta_i} \mu_{r-1:n-1}^{[i](k+1)} \right\}. \tag{129}$$

Proof. We shall present the proof for the relation in (129) while (128) can be proved on similar lines. For $2 \leq r \leq n$ and $k = 0, 1, \ldots$, we can write from (8)

$$(r-1)!(n-r)\mu_{r:n}^{(k)}$$

$$= \sum_P \int_0^\infty x^k F_{i_1}(x) \cdots F_{i_{r-1}}(x) f_{i_r}(x)\{1 - F_{i_{r+1}}(x)\} \cdots \{1 - F_{i_n}(x)\}\, dx$$

$$= \sum_P \frac{1}{\theta_{i_r}} \int_0^\infty x^k F_{i_1}(x) \cdots F_{i_{r-1}}(x)\{1 - F_{i_r}(x)\} \cdots \{1 - F_{i_n}(x)\}\, dx$$

upon using (127). Integrating now by parts treating x^k for integration and the rest of the integrand for differentiation, we obtain

$$(r-1)!(n-r)\mu_{r:n}^{(k)} = \frac{1}{k+1}\sum_P \frac{1}{\theta_{i_r}}\left[-\int_0^\infty x^{k+1} f_{i_1}(x)F_{i_2}(x)\cdots F_{i_{r-1}}(x)\right.$$
$$\times \{1-F_{i_r}(x)\}\cdots\{1-F_{i_n}(x)\}\,dx$$
$$-\cdots$$
$$-\int_0^\infty x^{k+1} F_{i_1}(x)\cdots F_{i_{r-2}}(x)f_{i_{r-1}}(x)$$
$$\times \{1-F_{i_r}(x)\}\cdots\{1-F_{i_n}(x)\}\,dx$$
$$+\int_0^\infty x^{k+1} F_{i_1}(x)\cdots F_{i_{r-1}}(x)f_{i_r}(x)$$
$$\times \{1-F_{i_{r+1}}(x)\}\cdots\{1-F_{i_n}(x)\}\,dx$$
$$+\cdots$$
$$+\int_0^\infty x^{k+1} F_{i_1}(x)\cdots F_{i_{r-1}}(x)$$
$$\left.\times \{1-F_{i_r}(x)\}\cdots\{1-F_{i_{n-1}}(x)\}f_{i_n}(x)\,dx\right]. \qquad (130)$$

Upon splitting the first set of integrals (ones with negative sign) on the RHS of (130) into two each through the term $1-F_{i_r}(x)$, we obtain

$$(r-1)!(n-r)\mu_{r:n}^{(k)} = \frac{1}{k+1}\sum_P \frac{1}{\theta_{i_r}}\left[\left\{\int_0^\infty x^{k+1} f_{i_1}(x)F_{i_2}(x)\cdots F_{i_r}(x)\right.\right.$$
$$\times \{1-F_{i_{r+1}}(x)\}\cdots\{1-F_{i_n}(x)\}\,dx$$
$$+\cdots$$
$$+\int_0^\infty x^{k+1} F_{i_1}(x)\cdots F_{i_{r-2}}(x)f_{i_{r-1}}(x)F_{i_r}(x)$$
$$\times \{1-F_{i_{r+1}}(x)\}\cdots\{1-F_{i_n}(x)\}\,dx$$
$$+\int_0^\infty x^{k+1} F_{i_1}(x)\cdots F_{i_{r-1}}(x)f_{i_r}(x)$$
$$\times \{1-F_{i_{r+1}}(x)\}\cdots\{1-F_{i_n}(x)\}\,dx$$
$$+\cdots$$
$$+\int_0^\infty x^{k+1} F_{i_1}(x)\cdots F_{i_{r-1}}(x)$$
$$\left.\left.\times \{1-F_{i_r}(x)\}\cdots\{1-F_{i_{n-1}}(x)\}f_{i_n}(x)dx\right\}\right.$$

$$- \left\{ \int_0^\infty x^{k+1} f_{i_1}(x) F_{i_2}(x) \cdots F_{i_r}(x) \right.$$

$$\times \left\{ 1 - F_{i_{r+1}}(x) \right\} \cdots \left\{ 1 - F_{i_n}(x) \right\} dx$$

$$+ \cdots$$

$$+ \int_0^\infty x^{k+1} F_{i_1}(x) \cdots F_{i_{r-2}}(x) f_{i_{r-1}}(x)$$

$$\left. \times \left\{ 1 - F_{i_{r+1}}(x) \right\} \cdots \left\{ 1 - F_{i_n}(x) \right\} dx \right\} \right]$$

$$= \frac{1}{k+1} \left[\left(\sum_{i=1}^n \frac{1}{\theta_i} \right) (r-1)!(n-r)! \mu_{r:n}^{(k+1)} \right.$$

$$\left. - (r-2)!(n-r)!(r-1) \sum_{i=1}^n \frac{1}{\theta_i} \mu_{r-1:n-1}^{[i](k+1)} \right]. \tag{131}$$

The relation in (129) is obtained by simply rewriting (131). □

Remark 7.2. The relations in Theorem 7.1 will enable one to compute all the single moments of all order statistics in a simple recursive manner for any specified values of θ_i $(i = 1, \ldots, n)$.

Remark 7.3. For the case when the exponential variables are IID, i.e., $\theta_1 = \cdots = \theta_n = 1$, the relations in Theorem 7.1 readily reduce to those in [71].

7.3. Relations for product moments

The following theorem has been established in [16] for the product moments of order statistics by using the differential equation in (127).

Theorem 7.4. *For $n = 2, 3, \ldots$,*

$$\mu_{1,2:n} = \frac{1}{\left(\sum_{i=1}^n 1/\theta_i \right)} \left\{ \mu_{1:n} + \mu_{2:n} \right\};$$

for $2 \le r \le n-1$,

$$\mu_{r,r+1:n} = \frac{1}{\left(\sum_{i=1}^n 1/\theta_i \right)} \left\{ \left(\mu_{r:n} + \mu_{r+1:n} \right) + \sum_{i=1}^n \frac{1}{\theta_i} \mu_{r-1,r:n-1}^{[i]} \right\}; \tag{132}$$

for $3 \le s \le n$,

$$\mu_{1,s:n} = \frac{1}{\left(\sum_{i=1}^n 1/\theta_i \right)} \left\{ \mu_{1:n} + \mu_{s:n} \right\};$$

for $2 \le r < s \le n$ and $s - r \ge 2$,

$$\mu_{r,s:n} = \frac{1}{\left(\sum_{i=1}^n 1/\theta_i \right)} \left\{ \left(\mu_{r:n} + \mu_{s:n} \right) + \sum_{i=1}^n \frac{1}{\theta_i} \mu_{r-1,s-1:n-1}^{[i]} \right\}.$$

Proof. We shall present the proof for the relation in (132) while the other three relations can be proved on similar lines. For $2 \le r \le n-1$, we can write from (12)

$$
(r-1)!(n-r-1)!\,\mu_{r:n} = (r-1)!(n-r-1)!\,E(X_{r:n}X^0_{r+1:n})
$$
$$
= \sum_P \int_0^\infty \int_x^\infty x F_{i_1}(x) \cdots F_{i_{r-1}}(x) f_{i_r}(x) f_{i_{r+1}}(y)
$$
$$
\times \{1 - F_{i_{r+2}}(y)\} \cdots \{1 - F_{i_n}(y)\}\, dy\, dx
$$
$$
= \sum_P \int_0^\infty x F_{i_1}(x) \cdots F_{i_{r-1}}(x) f_{i_r}(x) I(x)\, dx, \qquad (133)
$$

where

$$
I(x) = \int_x^\infty f_{i_{r+1}}(y)\{1 - F_{i_{r+2}}(y)\} \cdots \{1 - F_{i_n}(y)\}\, dy
$$
$$
= \frac{1}{\theta_{i_{r+1}}} \int_x^\infty \{1 - F_{i_{r+1}}(y)\} \cdots \{1 - F_{i_n}(y)\}\, dy
$$
$$
= \frac{1}{\theta_{i_{r+1}}} \left[\int_x^\infty y f_{i_{r+1}}(y)\{1 - F_{i_{r+2}}(y)\} \cdots \{1 - F_{i_n}(y)\}\, dy \right.
$$
$$
+ \cdots
$$
$$
+ \int_x^\infty y\{1 - F_{i_{r+1}}(y)\} \cdots \{1 - F_{i_{n-1}}(y)\} f_{i_n}(y)\, dy
$$
$$
\left. - x\{1 - F_{i_{r+1}}(x)\} \cdots \{1 - F_{i_n}(x)\} \right].
$$

Upon substituting this in (133), we get

$$
(r-1)!(n-r-1)!\,\mu_{r:n} = \sum_P \frac{1}{\theta_{i_{r+1}}} \left[\int_0^\infty \int_x^\infty xy F_{i_1}(x) \cdots F_{i_{r-1}}(x) f_{i_r}(x) f_{i_{r+1}}(y) \right.
$$
$$
\times \{1 - F_{i_{r+2}}(y)\} \cdots \{1 - F_{i_n}(y)\}\, dy\, dx
$$
$$
+ \cdots
$$
$$
+ \int_0^\infty \int_x^\infty xy F_{i_1}(x) \cdots F_{i_{r-1}}(x) f_{i_r}(x)
$$
$$
\times \{1 - F_{i_{r+1}}(y)\} \cdots \{1 - F_{i_{n-1}}(y)\} f_{i_n}(y)\, dy\, dx
$$
$$
- \int_0^\infty x^2 F_{i_1}(x) \cdots F_{i_{r-1}}(x) f_{i_r}(x)
$$
$$
\left. \times \{1 - F_{i_{r+1}}(x)\} \cdots \{1 - F_{i_n}(x)\}\, dx \right]. \qquad (134)
$$

Next, from (12) let us write for $2 \le r \le n-1$

$$(r-1)!(n-r-1)!\mu_{r+1:n}$$
$$= (r-1)!(n-r-1)!\, E(X_{r:n}^0 X_{r+1:n})$$
$$= \sum_P \int_0^\infty \int_0^y y F_{i_1}(x) \cdots F_{i_{r-1}}(x) f_{i_r}(x) f_{i_{r+1}}(y)$$
$$\times \{1 - F_{i_{r+2}}(y)\} \cdots \{1 - F_{i_n}(y)\}\, dx\, dy$$
$$= \sum_P \int_0^\infty y f_{i_{r+1}}(y)\{1 - F_{i_{r+2}}(y)\} \cdots \{1 - F_{i_n}(y)\} J(y)\, dy, \quad (135)$$

where

$$J(y) = \int_0^y F_{i_1}(x) \cdots F_{i_{r-1}}(x) f_{i_r}(x)\, dx$$
$$= \frac{1}{\theta_{i_r}} \left[\int_0^y F_{i_1}(x) \cdots F_{i_{r-1}}(x)\, dx - \int_0^y F_{i_1}(x) \cdots F_{i_r}(x)\, dx \right]$$
$$= \frac{1}{\theta_{i_r}} \left[y F_{i_1}(y) \cdots F_{i_{r-1}}(y) - \int_0^y x f_{i_1}(x) F_{i_2}(x) \cdots F_{i_{r-1}}(x)\, dx \right.$$
$$- \cdots - \int_0^y x F_{i_1}(x) \cdots F_{i_{r-2}}(x) f_{i_{r-1}}(x)\, dx$$
$$- y F_{i_1}(y) \cdots F_{i_r}(y) + \int_0^y x f_{i_1}(x) F_{i_2}(x) \cdots F_{i_r}(x)\, dx$$
$$\left. + \cdots + \int_0^y x F_{i_1}(x) \cdots F_{i_{r-1}}(x) f_{i_r}(x)\, dx \right].$$

Upon substituting this in (135), we get

$$(r-1)!(n-r-1)!\mu_{r+1:n}$$
$$= \sum_P \frac{1}{\theta_{i_r}} \left[\int_0^\infty y^2 F_{i_1}(y) \cdots F_{i_{r-1}}(y) f_{i_{r+1}}(y)\{1 - F_{i_{r+2}}(y)\} \cdots \{1 - F_{i_n}(y)\}\, dy \right.$$
$$- \int_0^\infty \int_x^\infty xy f_{i_1}(x) F_{i_2}(x) \cdots F_{i_{r-1}}(x) f_{i_{r+1}}(y)$$
$$\times \{1 - F_{i_{r+2}}(y)\} \cdots \{1 - F_{i_n}(y)\}\, dy\, dx$$
$$- \cdots$$
$$- \int_0^\infty \int_x^\infty xy F_{i_1}(x) \cdots F_{i_{r-2}}(x) f_{i_{r-1}}(x) f_{i_{r+1}}(y)$$
$$\times \{1 - F_{i_{r+2}}(y)\} \cdots \{1 - F_{i_n}(y)\}\, dy\, dx$$
$$- \int_0^\infty y^2 F_{i_1}(y) \cdots F_{i_r}(y) f_{i_{r+1}}(y)$$

$$\times \{1 - F_{i_{r+2}}(y)\} \cdots \{1 - F_{i_n}(y)\}\, dy$$

$$+ \int_0^\infty \int_x^\infty xy f_{i_1}(x) F_{i_2}(x) \cdots F_{i_r}(x) f_{i_{r+1}}(y)$$

$$\times \{1 - F_{i_{r+2}}(y)\} \cdots \{1 - F_{i_n}(y)\}\, dy\, dx$$

$$+ \cdots$$

$$+ \int_0^\infty \int_x^\infty xy F_{i_1}(x) \cdots F_{i_{r-1}}(x) f_{i_r}(x) f_{i_{r+1}}(y)$$

$$\times \{1 - F_{i_{r+2}}(y)\} \cdots \{1 - F_{i_n}(y)\}\, dy\, dx \Big]. \tag{136}$$

On adding (134) and (136) and simplifying the resulting expression, we obtain

$$(r-1)!(n-r-1)!(\mu_{r:n} + \mu_{r+1:n})$$

$$= \left(\sum_{i=1}^n \frac{1}{\theta_i} \right)(r-1)!(n-r-1)!\mu_{r,r+1:n}$$

$$-(r-2)!(n-r-1)!(r-1) \sum_{i=1}^n \frac{1}{\theta_i} \mu_{r-1,r:n-1}^{[i]}.$$

The relation in (132) is derived simply by rewriting the above equation. $\qquad\square$

Remark 7.5. The relations in Theorem 7.4 will enable one to compute all the product moments of all order statistics in a simple recursive manner for any specified values of θ_i $(i = 1, \ldots, n)$.

Remark 7.6. For the case when the exponential variables are IID, i.e., $\theta_1 = \cdots = \theta_n = 1$, the relations in Theorem 7.4 readily reduce to relations equivalent to those in [72].

7.4. Results for the multiple-outlier model

Let us consider the multiple-outlier model in which $\theta_1 = \cdots = \theta_{n-p} = \theta$ and $\theta_{n-p+1} = \cdots = \theta_n = \tau$. In this case, the relations in Theorems 7.1 and 7.4 reduce to the following:

(i) For $n \geq 1$ and $k = 0, 1, 2, \ldots$,

$$\mu_{1:n}^{(k+1)}[p] = \frac{k+1}{\left(\frac{n-p}{\theta} + \frac{p}{\tau}\right)}\, \mu_{1:n}^{(k)}[p].$$

(ii) For $2 \leq r \leq n$ and $k = 0, 1, 2, \ldots$,

$$\mu_{r:n}^{(k+1)}[p] = \frac{1}{\left(\frac{n-p}{\theta} + \frac{p}{\tau}\right)} \left\{ (k+1)\mu_{r:n}^{(k)}[p] + \frac{n-p}{\theta}\, \mu_{r-1:n-1}^{(k+1)}[p] \right.$$

$$\left. + \frac{p}{\tau}\, \mu_{r-1:n-1}^{(k+1)}[p-1] \right\}.$$

(iii) For $n \geq 2$,

$$\mu_{1,2:n}[p] = \frac{1}{\left(\frac{n-p}{\theta} + \frac{p}{\tau}\right)}\{\mu_{1:n}[p] + \mu_{2:n}[p]\}.$$

(iv) For $2 \leq r \leq n-1$,

$$\mu_{r,r+1:n}[p] = \frac{1}{\left(\frac{n-p}{\theta} + \frac{p}{\tau}\right)}\left\{\mu_{r:n}[p] + \mu_{r+1:n}[p] + \frac{n-p}{\theta}\,\mu_{r-1,r:n-1}[p]\right.$$
$$\left. + \frac{p}{\tau}\,\mu_{r-1,r:n-1}[p-1]\right\}.$$

(v) For $3 \leq s \leq n$,

$$\mu_{1,s:n}[p] = \frac{1}{\left(\frac{n-p}{\theta} + \frac{p}{\tau}\right)}\{\mu_{1:n}[p] + \mu_{s:n}[p]\}.$$

(vi) For $2 \leq r < s \leq n$ and $s - r \geq 2$,

$$\mu_{r,s:n}[p] = \frac{1}{\left(\frac{n-p}{\theta} + \frac{p}{\tau}\right)}\left\{\mu_{r:n}[p] + \mu_{s:n}[p] + \frac{n-p}{\theta}\,\mu_{r-1,s-1:n-1}[p]\right.$$
$$\left. + \frac{p}{\tau}\,\mu_{r-1,s-1:n-1}[p-1]\right\}.$$

Here, $\mu_{r:n}[p]$ and $\mu_{r:n-1}[p-1]$ denote the mean of the r-th order statistic when there are p and $p-1$ outliers, respectively.

Remark 7.7. Relations (i)–(vi) will enable one to compute all the single and product moments of all order statistics from a p-outlier model in a simple recursive manner. By starting with the IID results (case $p = 0$), these relations will yield the single and product moments of all order statistics from a single-outlier model (case $p = 1$), which in turn can be used to produce the results for $p = 2$, and so on.

7.5. Optimal Winsorized and trimmed means

By allowing a single outlier in an exponential sample, it was shown in [75] that the one-sided Winsorized mean

$$W_{m,n} = \frac{1}{m+1}\left\{\sum_{i=1}^{m-1} X_{i:n} + (n-m-1)X_{m:n}\right\} \tag{137}$$

is optimal in that it has the smallest mean square error among all linear estimators based on the first m order statistics when, in fact, there is no outlier. The determination of an optimal m for given values of n and $h = \theta/\tau$ was subsequently discussed in [69], where τ is the mean of the outlying observation. Making use of the recursive algorithm described above for the moments of order statistics from a p-outlier

n	h	$p = 1$		$p = 2$		$p = 3$		$p = 4$	
		m^*	RE	m^*	RE	m^*	RE	m^*	RE
20	0.05	17	25.071	14	50.296	11	70.707	9	85.163
	0.10	17	6.355	14	12.349	12	17.307	9	20.964
	0.15	18	3.057	15	5.506	12	7.627	10	9.260
	0.20	18	1.964	15	3.189	13	4.320	11	5.210
	0.25	18	1.486	16	2.170	14	2.786	12	3.371
	0.30	19	1.251	17	1.642	15	2.046	13	2.398
	0.35	19	1.132	17	1.352	16	1.601	14	1.833
	0.40	19	1.061	18	1.187	16	1.337	15	1.488
	0.45	19	1.019	18	1.085	17	1.177	16	1.272
	0.50	20	1.000	19	1.036	18	1.082	17	1.137

Table 4 – Optimal Winsorized estimator of θ and relative efficiency when p outliers (with $\theta/\tau = h$) are in the sample

exponential model, the optimal choice m^* of m for various choices of h, n and p were determined in [16]. For $n = 20$, these values are presented in table 4. Similar results are presented in table 5 for the optimal choice m^{**} of m for the one-sided trimmed mean

$$T_{m,n} = \frac{1}{m} \sum_{i=1}^{m} X_{i:n} \qquad (138)$$

that yields the smallest mean square error for given values of n and $h = \theta/\tau$.

The values in tables 4 and 5 reveal that the relative efficiency of the optimal trimmed estimator compared to the optimal Winsorized estimator increases significantly as h decreases and/or p increases. This means that the optimal trimmed estimator provides greater protection than the optimal Winsorized estimator when more (or few pronounced) outliers are present in the sample, but that it comes at a higher premium when at most one or few non-pronounced outliers are present.

7.6. Robustness of various linear estimators

Let us consider the following linear estimators of θ:

(i) Complete sample estimator $W_{n,n}$,

(ii) Winsorized estimator in (137) based on $m = 90\%$ of n,

(iii) Winsorized estimator in (137) based on $m = 80\%$ of n,

(iv) Winsorized estimator in (137) based on $m = 70\%$ of n,

(v) trimmed estimator in (138) based on $m = 90\%$ of n,

n	h	$p = 1$		$p = 2$		$p = 3$		$p = 4$	
		m^{**}	RE	m^{**}	RE	m^{**}	RE	m^{**}	RE
20	0.05	19	1.295	18	1.813	17	2.424	16	3.135
	0.10	19	1.268	18	1.702	17	2.190	16	2.718
	0.15	19	1.230	18	1.574	17	1.930	17	2.541
	0.20	19	1.188	18	1.444	18	1.866	17	2.449
	0.25	19	1.149	18	1.315	18	1.822	17	2.188
	0.30	19	1.109	19	1.316	18	1.694	18	2.036
	0.35	19	1.065	19	1.302	18	1.537	18	1.953
	0.40	19	1.028	19	1.255	19	1.399	18	1.787
	0.45	19	0.998	19	1.206	19	1.371	18	1.586
	0.50	19	0.967	19	1.139	19	1.309	19	1.438

Table 5 – Optimal trimmed estimator of θ and relative efficiency when p outliers (with $\theta/\tau = h$) are in the sample

(vi) trimmed estimator in (138) based on $m = 80\%$ of n,

(vii) trimmed estimator in (138) based on $m = 70\%$ of n, and

(viii) Chikkagoudar-Kunchur [48] estimator defined as

$$\mathrm{CK}_n = \frac{1}{n} \sum_{i=1}^{n} \Big(1 - \frac{2i}{n(n+1)}\Big) X_{i:n}. \tag{139}$$

We have presented in table 6 the values of bias and mean square error for the above eight estimators when $n = 20$, $p = 1(1)4$ and $h = 0.25(0.25)0.75$.

From table 6, we observe that while the complete sample estimator and the Chikkagoudar-Kunchur estimator in (139) are most efficient when there is no outlier or when there are few non-pronounced outliers, they develop serious bias and possess large mean square error when the outliers become pronounced. Once again, from this table we observe that the trimmed estimators provide good protection against the presence of few pronounced outliers, but it is attained with a higher premium.

Remark 7.8. After noting that the Chikkagoudar-Kunchur estimator is non-robust to the presence of outliers, [22] modified the estimator CK_n in (139) by downweighing the larger order statistics. Though this resulted in an improvement, yet pronounced outliers had an adverse effect on this estimator as well.

Remark 7.9. Some recurrence relations for the single and product moments of order statistics from multiple-outlier exponential models were derived directly in [49], and used to carry out a rather extensive evaluation and comparison of several different linear estimators of the exponential mean.

h		$p = 1$		$p = 2$		$p = 3$		$p = 4$	
		Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
0.75	(i)	-0.0317	0.0481	-0.0159	0.0491	0.0000	0.0506	0.0159	0.0527
	(ii)	-0.0381	0.0530	-0.0234	0.0538	-0.0085	0.0550	0.0066	0.0568
	(iii)	-0.0450	0.0591	-0.0309	0.0598	-0.0166	0.0609	-0.0021	0.0625
	(iv)	-0.0534	0.0669	-0.0398	0.0675	-0.0260	0.0686	-0.0120	0.0701
	(v)	-0.2217	0.0841	-0.2101	0.0802	-0.1983	0.0765	-0.1864	0.0731
	(vi)	-0.3697	0.1634	-0.3607	0.1576	-0.3515	0.1519	-0.3422	0.1462
	(vii)	-0.4848	0.2560	-0.4777	0.2497	-0.4703	0.2434	-0.4629	0.2370
	(viii)	-0.0572	0.0479	-0.0418	0.0481	-0.0264	0.0487	-0.0110	0.0498
0.50	(i)	0.0000	0.0522	0.0476	0.0612	0.0952	0.0748	0.1429	0.0930
	(ii)	-0.0157	0.0546	0.0231	0.0600	0.0638	0.0691	0.1061	0.0825
	(iii)	-0.0257	0.0603	0.0093	0.0645	0.0459	0.0717	0.0844	0.0824
	(iv)	-0.0360	0.0679	-0.0037	0.0714	0.0303	0.0775	0.0660	0.0867
	(v)	-0.2052	0.0787	-0.1759	0.0704	-0.1451	0.0637	-0.1129	0.0589
	(vi)	-0.3581	0.1560	-0.3366	0.1429	-0.3139	0.1303	-0.2901	0.1182
	(vii)	-0.4761	0.2484	-0.4596	0.2343	-0.4421	0.2201	-0.4238	0.2058
	(viii)	-0.0266	0.0500	0.0194	0.0560	0.0655	0.0662	0.1116	0.0808
0.25	(i)	0.0952	0.0884	0.2381	0.1701	0.3810	0.2925	0.5238	0.4558
	(ii)	0.0203	0.0595	0.1071	0.0841	0.2131	0.1414	0.3353	0.2402
	(iii)	0.0019	0.0633	0.0708	0.0784	0.1496	0.1089	0.2403	0.1619
	(iv)	-0.0130	0.0701	0.0468	0.0814	0.1139	0.1035	0.1895	0.1405
	(v)	-0.1805	0.0714	-0.1194	0.0596	-0.0471	0.0568	0.0360	0.0686
	(vi)	-0.3426	0.1464	-0.3023	0.1240	-0.2569	0.1030	-0.2054	0.0848
	(vii)	-0.4653	0.2390	-0.4359	0.2150	-0.4032	0.1905	-0.3667	0.1659
	(viii)	0.0645	0.0782	0.2017	0.1459	0.3392	0.2515	0.4770	0.3954

Table 6 – Bias and mean square error of eight estimators of θ when p outliers (with $\theta/\tau = h$) are present in the sample of size $n = 20$

Remark 7.10. In order to illustrate the usefulness of the differential equation technique described in sections 7.2 and 7.3, a "complete set" of recurrence relations for the single and product moments of order statistics from INID right-truncated exponential random variables were derived in [17].

8. Robust estimation for logistic distribution

8.1. Introduction

In the last section, we presented the results of [16] on a recursive algorithm for the computation of single and product moments of order statistics from INID exponential random variables and their use in the robust estimation of the exponential mean in the presence of multiple outliers. Arnold [16, pp. 243–246], in his discussion of this work, presented a direct approach for the computation of these moments and then remarked that

> "*Bala's specialized differential equation techniques may have their finest hour in dealing with X_i's for which minima and maxima are not nice. His proposed work in this direction will be interesting.*"

Motivated by this comment, the differential equation technique was recently used successfully in [51] to derive recurrence relations for the single moments of order statistics from INID logistic variables. These results were then applied to examine the effect of multiple outliers on various linear estimators of the location and scale parameters of the logistic distribution. These results, which extend the discussion in [15] on the robustness issues for a single-outlier model, are described in this section. In this regard, let $X_1, X_2, \ldots, X_n$ be independent logistic random variables having cumulative distribution functions $F_1(x), F_2(x), \ldots, F_n(x)$ and probability density functions $f_1(x), f_2(x), \ldots, f_n(x)$, respectively. Let $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ denote the order statistics obtained by arranging the n X_i's in increasing order of magnitude. Then the density function of $X_{r:n}$ $(1 \leq r \leq n)$ is [see (8)]

$$f_{r:n}(x) = \frac{1}{(r-1)!(n-r)!} \sum_P \prod_{a=1}^{r-1} F_{i_a}(x) f_{i_r}(x) \prod_{b=r+1}^{n} \{1 - F_{i_b}(x)\}, \qquad (140)$$

where $\sum_P$ denotes the summation over all $n!$ permutations $(i_1, i_2, \ldots, i_n)$ of $(1, \ldots, n)$.

Similarly, if another independent random variable $X_{n+1} \stackrel{d}{=} X_i$ (that is, with cumulative distribution function $F_i(x)$ and probability density function $f_i(x)$) is added to the original n variables $X_1, X_2, \ldots, X_n$, then the density function of $X_{r:n+1}$

$(1 \leq r \leq n+1)$ can be written as [see (8)]

$$
f_{r:n+1}^{[i]^+}(x) = \frac{F_i(x)}{(r-2)!(n-r+1)!} \sum_P \prod_{a=1}^{r-2} F_{i_a}(x) f_{i_{r-1}}(x) \prod_{b=r}^{n} \{1 - F_{i_b}(x)\}
$$

$$
+ \frac{f_i(x)}{(r-1)!(n-r+1)!} \sum_P \prod_{a=1}^{r-1} F_{i_a}(x) \prod_{b=r}^{n} \{1 - F_{i_b}(x)\}
$$

$$
+ \frac{1 - F_i(x)}{(r-1)!(n-r)!} \sum_P \prod_{a=1}^{r-1} F_{i_a}(x) f_{i_r}(x) \prod_{b=r+1}^{n} \{1 - F_{i_b}(x)\}, \quad x \in \mathbb{R}, \quad (141)
$$

with the conventions that $\prod_{i=r}^{s} = 1$ if $s - r = -1$ and $\prod_{i=r}^{s} = 0$ if $s - r = -2$, so that the first term is omitted if $r = 1$ and the last term is omitted if $r = n+1$. The superscript $[i]^+$ indicates that the random variable X_i is repeated.

Now, let us consider $X_1, \ldots, X_n$ to be INID logistic random variables, with X_i (for $i = 1, \ldots, n$) having its probability density function as

$$
f_i(x) = \frac{c e^{-c(x-\mu_i)/\sigma_i}}{\sigma_i (1 + e^{-c(x-\mu_i)/\sigma_i})^2}, \qquad x \in \mathbb{R}, \quad \mu_i \in \mathbb{R}, \quad \sigma_i > 0, \qquad (142)
$$

and cumulative distribution function as

$$
F_i(x) = \frac{1}{1 + e^{-c(x-\mu_i)/\sigma_i}}, \qquad x \in \mathbb{R}, \quad \mu_i \in \mathbb{R}, \quad \sigma_i > 0, \qquad (143)
$$

for $i = 1, 2, \ldots, n$, where $c = \pi/\sqrt{3}$.

From (142) and (143), we see that the distributions satisfy the differential equations

$$
f_i(x) = \frac{c}{\sigma_i} F_i(x) \{1 - F_i(x)\}, \qquad x \in \mathbb{R}, \quad \sigma_i > 0, \qquad (144)
$$

for $i = 1, 2, \ldots, n$.

Let us denote the single moments $E(X_{r:n}^k)$ by $\mu_{r:n}^{(k)}$, $1 \leq r \leq n$ and $k = 1, 2, \ldots$. Let us also use $\mu_{r:n-1}^{[i](k)}$ and $\mu_{r:n+1}^{[i]^+(k)}$ to denote the single moments of order statistics arising from $n-1$ variables obtained by deleting X_i from the original n variables $X_1, X_2, \ldots, X_n$ and the single moments of order statistics arising from $n+1$ variables obtained by adding an independent $X_{n+1} \stackrel{d}{=} X_i$ to the original n variables $X_1, X_2, \ldots, X_n$, respectively.

8.2. Relations for single moments

In this section, we present the following recurrence relations for the single moments established in [51] by making use of the differential equations in (144).

Theorem 8.1. *For $n \geq 1$ and $k = 0, 1, 2, \ldots,$*

$$\sum_{i=1}^{n} \frac{1}{\sigma_i} \mu_{1:n+1}^{[i]^+(k+1)} = -\frac{(k+1)}{c} \mu_{1:n}^{(k)} + \left(\sum_{i=1}^{n} \frac{1}{\sigma_i}\right) \mu_{1:n}^{(k+1)}; \tag{145}$$

for $2 \leq r \leq n$ and $k = 0, 1, 2, \ldots,$

$$\sum_{i=1}^{n} \frac{1}{\sigma_i} \mu_{r:n+1}^{[i]^+(k+1)} = \frac{(k+1)}{c} \{\mu_{r-1:n}^{(k)} - \mu_{r:n}^{(k)}\} - \sum_{i=1}^{n} \frac{1}{\sigma_i} \mu_{r-1:n-1}^{[i](k+1)}$$
$$+ \left(\sum_{i=1}^{n} \frac{1}{\sigma_i}\right) \{\mu_{r-1:n}^{(k+1)} + \mu_{r:n}^{(k+1)}\}; \tag{146}$$

for $n \geq 1$ and $k = 0, 1, 2, \ldots,$

$$\sum_{i=1}^{n} \frac{1}{\sigma_i} \mu_{n+1:n+1}^{[i]^+(k+1)} = \frac{(k+1)}{c} \mu_{n:n}^{(k)} + \left(\sum_{i=1}^{n} \frac{1}{\sigma_i}\right) \mu_{n:n}^{(k+1)}. \tag{147}$$

Proof. We shall present here the proof for the recurrence relation in (146), while the relations in (145) and (147) can be proved in a similar manner; see [51].

For $2 \leq r \leq n$, we can first of all write from (141) that

$$f_{r:n+1}^{[i]^+}(x) = \frac{F_i(x)}{(r-2)!(n-r+1)!} \sum_{P\,:\,i_{r-1}=i} \left[\sum_{j=1}^{r-2}\sum_{\substack{a=1 \\ a \neq j}}^{r-1} F_{i_a}(x) f_{i_j}(x) \prod_{b=r}^{n} \{1 - F_{i_b}(x)\}\right.$$

$$+ \sum_{j=r-1}^{n} \prod_{a=1}^{r-2} F_{i_a}(x) f_{i_j}(x) \prod_{\substack{b=r-1 \\ b \neq j}}^{n} \{1 - F_{i_b}(x)\}\right]$$

$$+ \frac{f_i(x)}{(r-1)!(n-r+1)!} \left[\sum_{P:\,i_{r-1}=i} (r-1) \prod_{a=1}^{r-1} F_{i_a}(x) \prod_{b=r}^{n} \{1 - F_{i_b}(x)\}\right.$$

$$+ \sum_{P:\,i_r=i} (n-r+1) \prod_{a=1}^{r-1} F_{i_a}(x) \prod_{b=r}^{n} \{1 - F_{i_b}(x)\}\right]$$

$$+ \frac{1 - F_i(x)}{(r-1)!(n-r)!} \sum_{P:\,i_r=i} \left[\sum_{j=1}^{r-1}\sum_{\substack{a=1 \\ a \neq j}}^{r} F_{i_a}(x) f_{i_j}(x) \prod_{b=r+1}^{n} \{1 - F_{i_b}(x)\}\right.$$

$$+ \sum_{j=r}^{n} \prod_{a=1}^{r-1} F_{i_a}(x) f_{i_j}(x) \prod_{\substack{b=r \\ b \neq j}}^{n} \{1 - F_{i_b}(x)\}\right], \qquad x \in \mathbb{R}. \tag{148}$$

Now, from (140), let us consider for $2 \le r \le n$, and $k = 0, 1, 2, \ldots$,

$$\frac{(k+1)}{c}\{\mu_{r-1:n}^{(k)} - \mu_{r:n}^{(k)}\}$$

$$= \frac{k+1}{c(r-2)!(n-r+1)!}\sum_P \int_{-\infty}^{\infty} x^k \prod_{a=1}^{r-2} F_{i_a}(x) f_{i_{r-1}}(x) \prod_{b=r}^{n}\{1 - F_{i_b}(x)\}\,dx$$

$$- \frac{k+1}{c(r-1)!(n-r)!}\sum_P \int_{-\infty}^{\infty} x^k \prod_{a=1}^{r-1} F_{i_a}(x) f_{i_r}(x) \prod_{b=r+1}^{n}\{1 - F_{i_b}(x)\}\,dx$$

$$= \frac{k+1}{(r-2)!(n-r+1)!}\sum_P \frac{1}{\sigma_{i_{r-1}}} \int_{-\infty}^{\infty} x^k \prod_{a=1}^{r-1} F_{i_a}(x) \prod_{b=r-1}^{n}\{1 - F_{i_b}(x)\}\,dx$$

$$- \frac{k+1}{(r-1)!(n-r)!}\sum_P \frac{1}{\sigma_{i_r}} \int_{-\infty}^{\infty} x^k \prod_{a=1}^{r} F_{i_a}(x) \prod_{b=r}^{n}\{1 - F_{i_b}(x)\}\,dx$$

upon using (144). Integrating now by parts in both integrals above, treating x^k for integration and the rest of the integrand for differentiation, we obtain

$$\frac{(k+1)}{c}\{\mu_{r-1:n}^{(k)} - \mu_{r:n}^{(k)}\}$$

$$= \frac{1}{(r-2)!(n-r+1)!}\sum_P \frac{1}{\sigma_{i_{r-1}}}\left[-\sum_{j=1}^{r-1}\int_{-\infty}^{\infty} x^{k+1} \prod_{\substack{a=1\\a\neq j}}^{r-1} F_{i_a}(x) f_{i_j}(x)\right.$$

$$\left.\times \prod_{b=r-1}^{n}\{1-F_{i_b}(x)\}\,dx + \sum_{j=r-1}^{n}\int_{-\infty}^{\infty} x^{k+1}\prod_{a=1}^{r-1} F_{i_a}(x) f_{i_j}(x) \prod_{\substack{b=r-1\\b\neq j}}^{n}\{1 - F_{i_b}(x)\}\,dx\right]$$

$$- \frac{1}{(r-1)!(n-r)!}\sum_P \frac{1}{\sigma_{i_r}}\left[-\sum_{j=1}^{r}\int_{-\infty}^{\infty} x^{k+1}\prod_{\substack{a=1\\a\neq j}}^{r} F_{i_a}(x) f_{i_j}(x) \prod_{b=r}^{n}\{1 - F_{i_b}(x)\}\,dx\right.$$

$$\left. + \sum_{j=r}^{n}\int_{-\infty}^{\infty} x^{k+1}\prod_{a=1}^{r} F_{i_a}(x) f_{i_j}(x) \prod_{\substack{b=r\\b\neq j}}^{n}\{1 - F_{i_b}(x)\}\,dx\right]. \tag{149}$$

We now split the first term in the first sum above into three by separating out the $j = r - 1$ term from $\sum_{j=1}^{r-1}$ and splitting the remaining sum, $\sum_{j=1}^{r-2}$, into two through $\{1 - F_{i_{r-1}}(x)\}$. We also split the first term in the second sum above by separating out the $j = r$ term from $\sum_{j=1}^{r}$. And we split the second term in the second sum above

into two through $F_{i_r}(x) = 1 - \{1 - F_{i_r}(x)\}$. Equation (149) now becomes

$$\frac{(k+1)}{c}\{\mu_{r-1:n}^{(k)} - \mu_{r:n}^{(k)}\}$$

$$= \frac{1}{(r-2)!(n-r+1)!}\sum_P \frac{1}{\sigma_{i_{r-1}}}\Bigg[\sum_{j=1}^{r-2}\int_{-\infty}^{\infty} x^{k+1}F_{i_{r-1}}(x)\prod_{\substack{a=1\\a\neq j}}^{r-1}F_{i_a}(x)f_{i_j}(x)$$

$$\times \prod_{b=r}^{n}\{1-F_{i_b}(x)\}\,dx - \int_{-\infty}^{\infty} x^{k+1}\prod_{a=1}^{r-2}F_{i_a}(x)f_{i_{r-1}}(x)\prod_{b=r-1}^{n}\{1-F_{i_b}(x)\}\,dx$$

$$- \sum_{j=1}^{r-2}\int_{-\infty}^{\infty} x^{k+1}\prod_{\substack{a=1\\a\neq j}}^{r-1}F_{i_a}(x)f_{i_j}(x)\prod_{b=r}^{n}\{1-F_{i_b}(x)\}\,dx$$

$$+ \sum_{j=r-1}^{n}\int_{-\infty}^{\infty} x^{k+1}F_{i_{r-1}}(x)\prod_{a=1}^{r-2}F_{i_a}(x)f_{i_j}(x)\prod_{\substack{b=r-1\\b\neq j}}^{n}\{1-F_{i_b}(x)\}\,dx\Bigg]$$

$$- \frac{1}{(r-1)!(n-r)!}\sum_P \frac{1}{\sigma_{i_r}}\Bigg[-\sum_{j=1}^{r-1}\int_{-\infty}^{\infty} x^{k+1}\{1-F_{i_r}(x)\}\prod_{\substack{a=1\\a\neq j}}^{r}F_{i_a}(x)f_{i_j}(x)$$

$$\times \prod_{b=r+1}^{n}\{1-F_{i_b}(x)\}dx - \int_{-\infty}^{\infty} x^{k+1}f_{i_r}(x)\prod_{a=1}^{r-1}F_{i_a}(x)\prod_{b=r}^{n}\{1-F_{i_b}(x)\}\,dx$$

$$- \sum_{j=r}^{n}\int_{-\infty}^{\infty} x^{k+1}\{1-F_{i_r}(x)\}\prod_{a=1}^{r-1}F_{i_a}(x)f_{i_j}(x)\prod_{\substack{b=r\\b\neq j}}^{n}\{1-F_{i_b}(x)\}\,dx$$

$$+ \sum_{j=r}^{n}\int_{-\infty}^{\infty} x^{k+1}\prod_{a=1}^{r-1}F_{i_a}(x)f_{i_j}(x)\prod_{\substack{b=r\\b\neq j}}^{n}\{1-F_{i_b}(x)\}\,dx\Bigg].$$

We now split the second term in the first sum above through $\{1 - F_{i_{r-1}}(x)\}$ to get

$$\frac{(k+1)}{c}\{\mu_{r-1:n}^{(k)} - \mu_{r:n}^{(k)}\}$$

$$= \frac{1}{(r-2)!(n-r+1)!}\sum_P \frac{1}{\sigma_{i_{r-1}}}\Bigg[\sum_{j=1}^{r-2}\int_{-\infty}^{\infty} x^{k+1}F_{i_{r-1}}(x)\prod_{\substack{a=1\\a\neq j}}^{r-1}F_{i_a}(x)f_{i_j}(x)$$

$$\times \prod_{b=r}^{n}\{1-F_{i_b}(x)\}\,dx + \int_{-\infty}^{\infty} x^{k+1}\prod_{a=1}^{r-1}F_{i_a}(x)f_{i_{r-1}}(x)\prod_{b=r}^{n}\{1-F_{i_b}(x)\}\,dx$$

$$- \int_{-\infty}^{\infty} x^{k+1}\prod_{a=1}^{r-2}F_{i_a}(x)f_{i_{r-1}}(x)\prod_{b=r}^{n}\{1-F_{i_b}(x)\}\,dx$$

$$-\sum_{j=1}^{r-2}\int_{-\infty}^{\infty}x^{k+1}\prod_{\substack{a=1\\a\neq j}}^{r-1}F_{i_a}(x)f_{i_j}(x)\prod_{b=r}^{n}\{1-F_{i_b}(x)\}\,dx$$

$$+\sum_{j=r-1}^{n}\int_{-\infty}^{\infty}x^{k+1}F_{i_{r-1}}(x)\prod_{a=1}^{r-2}F_{i_a}(x)f_{i_j}(x)\prod_{\substack{b=r-1\\b\neq j}}^{n}\{1-F_{i_b}(x)\}\,dx\Bigg]$$

$$-\frac{1}{(r-1)!(n-r)!}\sum_{P}\frac{1}{\sigma_{i_r}}\Bigg[-\sum_{j=1}^{r-1}\int_{-\infty}^{\infty}x^{k+1}\{1-F_{i_r}(x)\}\prod_{\substack{a=1\\a\neq j}}^{r}F_{i_a}(x)f_{i_j}(x)$$

$$\times\prod_{b=r+1}^{n}\{1-F_{i_b}(x)\}\,dx-\int_{-\infty}^{\infty}x^{k+1}f_{i_r}(x)\prod_{a=1}^{r-1}F_{i_a}(x)\prod_{b=r}^{n}\{1-F_{i_b}(x)\}\,dx$$

$$-\sum_{j=r}^{n}\int_{-\infty}^{\infty}x^{k+1}\{1-F_{i_r}(x)\}\prod_{a=1}^{r-1}F_{i_a}(x)f_{i_j}(x)\prod_{\substack{b=r\\b\neq j}}^{n}\{1-F_{i_b}(x)\}\,dx$$

$$+\sum_{j=r}^{n}\int_{-\infty}^{\infty}x^{k+1}\prod_{a=1}^{r-1}F_{i_a}(x)f_{i_j}(x)\prod_{\substack{b=r\\b\neq j}}^{n}\{1-F_{i_b}(x)\}\,dx\Bigg].$$

Now, comparison with (148) shows that the first, second, and fifth terms in the first sum above combine with the first, second, and third terms in the second sum above to give

$$\frac{(k+1)}{c}\{\mu_{r-1:n}^{(k)}-\mu_{r:n}^{(k)}\}$$

$$=\sum_{i=1}^{n}\frac{1}{\sigma_i}\mu_{r:n+1}^{[i]^+(k+1)}+\frac{1}{(r-2)!(n-r+1)!}\sum_{P}\frac{1}{\sigma_{i_{r-1}}}$$

$$\times\Bigg[-\int_{-\infty}^{\infty}x^{k+1}\prod_{a=1}^{r-2}F_{i_a}(x)f_{i_{r-1}}(x)\prod_{b=r}^{n}\{1-F_{i_b}(x)\}\,dx$$

$$-\sum_{j=1}^{r-2}\int_{-\infty}^{\infty}x^{k+1}\prod_{\substack{a=1\\a\neq j}}^{r-1}F_{i_a}(x)f_{i_j}(x)\prod_{b=r}^{n}\{1-F_{i_b}(x)\}\,dx\Bigg]$$

$$-\frac{1}{(r-1)!(n-r)!}\sum_{P}\frac{1}{\sigma_{i_r}}\Bigg[\sum_{j=r}^{n}\int_{-\infty}^{\infty}x^{k+1}\prod_{a=1}^{r-1}F_{i_a}(x)f_{i_j}(x)$$

$$\times\prod_{\substack{b=r\\b\neq j}}^{n}\{1-F_{i_b}(x)\}\,dx\Bigg]$$

$$= \sum_{i=1}^{n} \frac{1}{\sigma_i} \mu_{r:n+1}^{[i]^+(k+1)}$$

$$+ \frac{1}{(r-2)!(n-r+1)!} \sum_{P} \frac{1}{\sigma_{i_{r-1}}} \left[- \sum_{j=1}^{r-1} \int_{-\infty}^{\infty} x^{k+1} \prod_{\substack{a=1 \\ a \neq j}}^{r-1} F_{i_a}(x) f_{i_j}(x) \right.$$

$$\left. \times \prod_{b=r}^{n} \{1 - F_{i_b}(x)\}\, dx \right]$$

$$- \frac{1}{(r-1)!(n-r)!} \sum_{P} \frac{1}{\sigma_{i_r}} \left[\sum_{j=r}^{n} \int_{-\infty}^{\infty} x^{k+1} \prod_{a=1}^{r-1} F_{i_a}(x) f_{i_j}(x) \right.$$

$$\left. \times \prod_{\substack{b=r \\ b \neq j}}^{n} \{1 - F_{i_b}(x)\}\, dx \right]. \tag{150}$$

We now use the fact that

$$\frac{1}{(r-1)!(n-r)!} \sum_{P} \frac{1}{\sigma_{i_r}} \left[\sum_{j=1}^{r-1} \int_{-\infty}^{\infty} x^{k+1} \prod_{\substack{a=1 \\ a \neq j}}^{r-1} F_{i_a}(x) f_{i_j}(x) \prod_{b=r}^{n} \{1 - F_{i_b}(x)\}\, dx \right]$$

$$= \frac{1}{(r-2)!(n-r+1)!} \sum_{P} \frac{1}{\sigma_{i_{r-1}}} \left[\sum_{j=r}^{n} \int_{-\infty}^{\infty} x^{k+1} \prod_{a=1}^{r-2} F_{i_a}(x) f_{i_j}(x) \right.$$

$$\left. \times \prod_{\substack{b=r-1 \\ b \neq j}}^{n} \{1 - F_{i_b}(x)\}\, dx \right]$$

to rewrite (150) as follows:

$$\frac{(k+1)}{c} \{ \mu_{r-1:n}^{(k)} - \mu_{r:n}^{(k)} \}$$

$$= \sum_{i=1}^{n} \frac{1}{\sigma_i} \mu_{r:n+1}^{[i]^+(k+1)} + \frac{1}{(r-2)!(n-r+1)!} \sum_{P} \frac{1}{\sigma_{i_{r-1}}}$$

$$\times \left[- \sum_{j=1}^{r-1} \int_{-\infty}^{\infty} x^{k+1} \prod_{\substack{a=1 \\ a \neq j}}^{r-1} F_{i_a}(x) f_{i_r}(x) \prod_{b=r}^{n} \{1 - F_{i_b}(x)\}\, dx \right.$$

$$\left. - \sum_{j=r}^{n} \int_{-\infty}^{\infty} x^{k+1} \prod_{a=1}^{r-2} F_{i_a}(x) f_{i_j}(x) \prod_{\substack{b=r-1 \\ b \neq j}}^{n} \{1 - F_{i_b}(x)\}\, dx \right]$$

$$-\frac{1}{(r-1)!(n-r)!}\sum_{P}\frac{1}{\sigma_{i_r}}\left[\sum_{j=r}^{n}\int_{-\infty}^{\infty}x^{k+1}\prod_{a=1}^{r-1}F_{i_a}(x)f_{i_j}(x)\right.$$

$$\times\prod_{\substack{b=r\\b\neq j}}^{n}\{1-F_{i_b}(x)\}\,dx - \sum_{j=1}^{r-1}\int_{-\infty}^{\infty}x^{k+1}\prod_{\substack{a=1\\a\neq j}}^{r-1}F_{i_a}(x)f_{i_j}(x)$$

$$\times\left.\prod_{b=r}^{n}\{1-F_{i_b}(x)\}\,dx\right].$$

We now recognize the first sum above as $-(\sum_{i=1}^{n}1/\sigma_i)\mu_{r-1:n}^{(k+1)}$, and we split the second term in the second sum above into two through $\{1-F_{i_r}(x)\}$ to obtain

$$\frac{(k+1)}{c}\{\mu_{r-1:n}^{(k)}-\mu_{r:n}^{(k)}\}$$

$$=\sum_{i=1}^{n}\frac{1}{\sigma_i}\mu_{r:n+1}^{[i]^+(k+1)}-\left(\sum_{i=1}^{n}\frac{1}{\sigma_i}\right)\mu_{r-1:n}^{(k+1)}$$

$$-\frac{1}{(r-1)!(n-r)!}\sum_{P}\frac{1}{\sigma_{i_r}}\left[\sum_{j=r}^{n}\int_{-\infty}^{\infty}x^{k+1}\prod_{a=1}^{r-1}F_{i_a}(x)f_{i_j}(x)\prod_{\substack{b=r\\b\neq j}}^{n}\{1-F_{i_b}(x)\}\,dx\right.$$

$$+\sum_{j=1}^{r-1}\int_{-\infty}^{\infty}x^{k+1}\prod_{\substack{a=1\\a\neq j}}^{r}F_{i_a}(x)f_{i_j}(x)\prod_{b=r+1}^{n}\{1-F_{i_b}(x)\}\,dx$$

$$-\left.\sum_{j=1}^{r-1}\int_{-\infty}^{\infty}x^{k+1}\prod_{\substack{a=1\\a\neq j}}^{r-1}F_{i_a}(x)f_{i_j}(x)\prod_{b=r+1}^{n}\{1-F_{i_b}(x)\}\,dx\right]$$

$$=\sum_{i=1}^{n}\frac{1}{\sigma_i}\mu_{r:n+1}^{[i]^+(k+1)}-\left(\sum_{i=1}^{n}\frac{1}{\sigma_i}\right)\mu_{r-1:n}^{(k+1)}-\left(\sum_{i=1}^{n}\frac{1}{\sigma_i}\right)\mu_{r:n}^{(k+1)}+\sum_{i=1}^{n}\frac{1}{\sigma_i}\mu_{r-1:n-1}^{[i](k+1)}.$$

The relation in (146) readily follows when we rewrite the above equation.   □

8.3. Results for the multiple-outlier model

In this section, we consider the special case when $X_1, X_2, \ldots, X_{n-p}$ are independent logistic random variables with location parameter μ and scale parameter σ, while $X_{n-p+1}, \ldots, X_n$ are independent logistic random variables with location parameter μ_1 and scale parameter σ_1 (and independent of $X_1, X_2, \ldots, X_{n-p}$).

Here, we denote the single moments by $\mu_{r:n}^{(k)}[p]$, and the results presented in Theorem 8.1 then readily reduce to the following recurrence relations:

(i) For $n \geq 1$,

$$\mu_{1:n+1}^{(k+1)}[p+1]$$
$$= \frac{\sigma_1}{p}\left\{\left(\frac{n-p}{\sigma}+\frac{p}{\sigma_1}\right)\mu_{1:n}^{(k+1)}[p] - \frac{n-p}{\sigma}\mu_{1:n+1}^{(k+1)}[p] - \frac{(k+1)}{c}\mu_{1:n}^{(k)}[p]\right\}.$$

(ii) For $2 \leq r \leq n$,

$$\mu_{r:n+1}^{(k+1)}[p+1] = \frac{\sigma_1}{p}\left\{\left(\frac{n-p}{\sigma}+\frac{p}{\sigma_1}\right)\left\{\mu_{r:n}^{(k+1)}[p]+\mu_{r-1:n}^{(k+1)}[p]\right\}\right.$$
$$- \frac{n-p}{\sigma}\left\{\mu_{r:n+1}^{(k+1)}[p]+\mu_{r-1:n-1}^{(k+1)}[p]\right\}$$
$$\left.+ \frac{(k+1)}{c}\left\{\mu_{r-1:n}^{(k)}[p]-\mu_{r:n}^{(k)}[p]\right\}\right\} - \mu_{r-1:n-1}^{(k+1)}[p-1].$$

(iii) For $n \geq 1$,

$$\mu_{n+1:n+1}^{(k+1)}[p+1] = \frac{\sigma_1}{p}\left\{\left(\frac{n-p}{\sigma}+\frac{p}{\sigma_1}\right)\mu_{n:n}^{(k+1)}[p] - \frac{n-p}{\sigma}\mu_{n+1:n+1}^{(k+1)}[p]\right.$$
$$\left.+ \frac{(k+1)}{c}\mu_{n:n}^{(k)}[p]\right\}.$$

Note that if we replace p by $n-p$, we get a set of equivalent relations by regarding the first p X_i's as the outliers.

Remark 8.2. If we now multiply each of the above relations by p/σ_1 and then set $p = 0$ and $\sigma = 1$ (or simply set $p = n$ and $\sigma = 1$), we obtain the following recurrence relations for the case when the X_i's are IID standard logistic random variables (which could alternatively be obtained by setting $\sigma_1 = \sigma_2 = \cdots = \sigma_n = 1$, $\mu_1 = \mu_2 = \cdots = \mu_n = 0$ in Theorem 8.1):

$$\mu_{1:n+1}^{(k+1)} = \mu_{1:n}^{(k+1)} - \frac{(k+1)}{cn}\mu_{1:n}^{(k)}, \qquad n \geq 1,$$

$$\mu_{r:n+1}^{(k+1)} = \mu_{r:n}^{(k+1)} + \mu_{r-1:n}^{(k+1)} - \mu_{r-1:n-1}^{(k+1)} + \frac{(k+1)}{cn}\{\mu_{r-1:n}^{(k)}-\mu_{r:n}^{(k)}\}, \quad 2 \leq r \leq n,$$

and

$$\mu_{n+1:n+1}^{(k+1)} = \mu_{n:n}^{(k+1)} + \frac{(k+1)}{cn}\mu_{n:n}^{(k)}, \qquad n \geq 1.$$

These relations are equivalent to those in [88].

Remark 8.3. Assuming that the moments of order statistics for the single-outlier model are known (for example, they can be found in [24]), setting $p = 1$ in relations (i)–(iii), along with the above IID results, will enable one to compute all of the

moments of order statistics from a 2-outlier model. One can then set $p = 2$ in relations (i)–(iii) to obtain all of the moments of order statistics from a 3-outlier model. Continuing in this manner, we see that relations (i)–(iii), the above IID relations, and knowledge of the moments of order statistics for the single-outlier model, will enable one to compute all of the moments of order statistics for the multiple-outlier model in a simple recursive manner.

Remark 8.4. Interestingly, this particular recursive property of moments of order statistics from a logistic multiple-outlier model was made as a conjecture by Balakrishnan [16, pp. 252–253] in his reply to the comments of Arnold [16, pp. 243–246].

8.4. Robustness of estimators of location

Through numerical integration, the means, variances and covariances of order statistics for the single-outlier model were computed in [24], and these values were then used to examine the bias of various linear estimators of the location parameter μ under a single location-outlier logistic model. Here, we discuss the bias of these linear estimators of μ under the multiple location-outlier logistic model.

The omnibus estimators of μ that are considered here are the following:

(i) *Sample mean*:
$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_{i:n}.$$

(ii) *Median*:
$$\text{for } n \text{ odd,} \qquad X_{\frac{n+1}{2}:n},$$
$$\text{for } n \text{ even,} \qquad \frac{1}{2}\big(X_{[n/2]:n} + X_{[n/2]+1:n}\big).$$

(iii) *Trimmed mean*:
$$T_n(r) = \frac{1}{n-2r} \sum_{i=r+1}^{n-r} X_{i:n}.$$

(iv) *Winsorized mean*:
$$W_n(r) = \frac{1}{n}\left[(r+1)(X_{r+1:n} + X_{n-r:n}) + \sum_{i=r+2}^{n-r-1} X_{i:n}\right].$$

(v) *Modified maximum likelihood (MML) estimator*:
$$\mu_c = \frac{1}{m}\left[r\beta(X_{r+1:n} + X_{n-r:n}) + \sum_{i=r+1}^{n-r} X_{i:n}\right],$$

where $m = n - 2r + 2r\beta$, $\beta = (g(h_2) - g(h_1))/(h_2 - h_1)$, $h_1 = F^{-1}(1 - q - \sqrt{q(1-q)/n})$, $h_2 = F^{-1}(1 - q + \sqrt{q(1-q)/n})$, $q = r/n$, $F(h) = \int_{-\infty}^{h} f(z)\,dz$, $f(z) = \frac{1}{\sqrt{2\pi}}e^{-z^2/2}$, and $g(h) = f(h)/(1 - F(h))$.

(vi) *Linearly weighted means*:

for n odd,

$$L_n(r) = \frac{1}{2(\frac{n-1}{2} - r)^2 + (n - 2r)}\left[\sum_{i=1}^{\frac{n-1}{2}-r}(2i-1)(X_{r+i:n} + X_{n-r-i+1:n})\right.$$
$$\left. + (n - 2r)X_{\frac{n+1}{2}:n}\right],$$

for n even,

$$L_n(r) = \frac{1}{2(\frac{n}{2} - r)^2}\left[\sum_{i=1}^{\frac{n}{2}-r}(2i-1)(X_{r+i:n} + X_{n-r-i+1:n})\right].$$

(vii) *Gastwirth mean*:

$$T_n = \frac{3}{10}(X_{[n/3]+1:n} + X_{n-[n/3]:n}) + \frac{2}{5}\tilde{X},$$

where $\tilde{X}$ is the median.

In addition to these omnibus estimators, we also included the following estimators of μ:

(viii) *BLUE*:

$$\hat{\mu}(r) = \frac{\mathbf{1}\boldsymbol{\omega}\boldsymbol{X}'}{\mathbf{1}\boldsymbol{\omega}\mathbf{1}'},$$

where $\mathbf{1} = (1, 1, \dots, 1)$, $\boldsymbol{\omega} = [(\sigma_{i,j:n}[0]); \ r + 1 \leq i, \ j \leq n - r]^{-1}$, and $\boldsymbol{X} = (X_{r+1:n}, X_{r+2;n}, \dots, X_{n-r:n}); \sigma_{i,j:n}[0]$ denotes the covariance between the i-th and j-th order statistics in a sample of size n from the standard logistic distribution (the [0] indicates that there are no outliers).

(ix) *The approximate best linear unbiased estimator in* [45]:

$$\mu' = \frac{6}{n(n+1)(n+2)}\sum_{i=1}^{n}i(n+1-i)X_{i:n}$$

which is also discussed in [63].

(x) *RSE (estimator proposed in [83]):*

$$v_{(i^*)} = \frac{1}{2}(X_{n-i^*+1:n} + X_{i^*:n})$$

where i^* is chosen to minimize the variance of $v_{(i^*)}$ when the X's are IID.

The recursive computational method presented above was utilized in [51] to examine the bias of all these estimators of μ when multiple outliers are possibly present in the sample. In table 7, we have presented the bias of all the estimators for $n = 20$, $p = 1, 2$, different choices of r, and $\mu = 0$, $\mu_1 = 0.5(0.5)3.0, 4.0$, $\sigma = \sigma_1 = 1$. In all cases, we observe that median is the estimator with the smallest bias. For large values of r, the linearly weighted mean and the BLUE are quite comparable to the median in terms of bias, with the linearly weighted mean having a smaller bias than the BLUE. For small values of μ_1, the modified maximum likelihood estimator, the Gastwirth mean and the Winsorized mean are also comparable to the BLUE and linearly weighted mean, but for larger values of μ_1 their bias becomes much larger. For small values of r, however, all of these estimators are quite sensitive to the presence of outliers, as one would expect.

The fact that all of the estimators have similar bias for small values of μ_1 is explained by the fact that these estimators are all unbiased in the IID case. Therefore, all of the biases become the same as μ_1 approaches zero.

Remark 8.5. Note that the outlier model considered in table 7 is a multiple location-outlier model. Since all the linear estimators of μ considered here are symmetric functions of order statistics, they will all be unbiased under the multiple scale-outlier model. Hence, a comparison of these estimators under the multiple-scale outlier model would have to be made on the basis of their variance. Symmetric functions are most appropriate when the direction of the slippage is not known. However, if the direction is known then some asymmetric unbiased estimators will naturally perform better than the symmetric ones. For example, any estimator that gives less weight to the larger order statistics (and more to the smaller ones) will be expected to perform better if μ_1 is positive.

8.5. Robustness of estimators of scale

By considering both the single location-outlier and single scale-outlier model, the bias of the following linear estimators of the scale parameter σ were determined in [15]:

(i) *BLUE*:

$$\hat{\sigma}(r) = \frac{\boldsymbol{\mu}\boldsymbol{\omega}\boldsymbol{X}'}{\boldsymbol{\mu}\boldsymbol{\omega}\boldsymbol{\mu}'},$$

where $\boldsymbol{\mu} = (\mu_{r+1:n}[0], \mu_{r+2:n}[0], \ldots, \mu_{n-r:n}[0])$, $\boldsymbol{\omega} = [(\sigma_{i,j:n}[0]; \ r+1 \leq i, j \leq n-r]^{-1}$, and $\boldsymbol{X} = (X_{r+1:n}, X_{r+2:n}, \ldots, X_{n-r:n})$; $\mu_{s:n}[0]$ denotes the mean of the s-th order statistic in a sample of size n from the standard logistic distribution (the [0] indicates that there are no outliers), and

μ_1	0.5	1.0	1.5	2.0	2.5	3.0	4.0
$n = 20, p = 1$							
BLUE0	0.0245	0.0460	0.0629	0.0751	0.0837	0.0900	0.0990
BLUE2	0.0244	0.0455	0.0612	0.0712	0.0768	0.0769	0.0814
BLUE4	0.0242	0.0438	0.0565	0.0634	0.0666	0.0679	0.0687
BLUE7	0.0236	0.0408	0.0504	0.0550	0.0570	0.0578	0.0583
RSE	0.0241	0.0433	0.0556	0.0620	0.0650	0.0663	0.0670
Mean	0.0250	0.0500	0.0750	0.1000	0.1250	0.1500	0.2000
Trimm2	0.0245	0.0459	0.0622	0.0728	0.0787	0.0817	0.0836
Trimm4	0.0241	0.0434	0.0559	0.0626	0.0658	0.0672	0.0681
Median	0.0236	0.0407	0.0503	0.0548	0.0568	0.0576	0.0581
Winsor2	0.0248	0.0479	0.0673	0.0812	0.0897	0.0943	0.0974
Winsor4	0.0244	0.0451	0.0598	0.0683	0.0726	0.0745	0.0756
Winsor8	0.0237	0.0411	0.0510	0.0558	0.0579	0.0588	0.0593
MML2	0.0247	0.0477	0.0666	0.0801	0.0883	0.0926	0.0956
MML4	0.0243	0.0449	0.0592	0.0675	0.0716	0.0735	0.0746
MML8	0.0237	0.0411	0.0510	0.0558	0.0579	0.0587	0.0592
LinWei2	0.0240	0.0432	0.0556	0.0624	0.0658	0.0673	0.0682
LinWei4	0.0239	0.0420	0.0529	0.0585	0.0610	0.0620	0.0627
LinWei8	0.0236	0.0408	0.0505	0.0551	0.0571	0.0580	0.0584
Gastw	0.0239	0.0423	0.0535	0.0591	0.0617	0.0628	0.0634
Blom	0.0245	0.0464	0.0642	0.0781	0.0889	0.0978	0.1126
$n = 20, p = 2$							
BLUE0	0.0491	0.0933	0.1297	0.1584	0.1811	0.1996	0.2303
BLUE2	0.0490	0.0925	0.1272	0.1526	0.1698	0.1806	0.1904
BLUE4	0.0486	0.0895	0.1182	0.1350	0.1433	0.1470	0.1492
BLUE7	0.0477	0.0838	0.1051	0.1156	0.1203	0.1223	0.1234
RSE	0.0485	0.0887	0.1162	0.1318	0.1394	0.1428	0.1449
Mean	0.0500	0.1000	0.1500	0.2000	0.2500	0.3000	0.4000
Trimm2	0.0491	0.0933	0.1293	0.1562	0.1746	0.1862	0.1968
Trimm4	0.0485	0.0887	0.1167	0.1332	0.1418	0.1458	0.1482
Median	0.0476	0.0836	0.1048	0.1153	0.1200	0.1219	0.1231
Winsor2	0.0496	0.0969	0.1394	0.1751	0.2029	0.2224	0.2420
Winsor4	0.0490	0.0920	0.1249	0.1464	0.1584	0.1643	0.1680
Winsor8	0.0478	0.0843	0.1064	0.1175	0.1226	0.1247	0.1259
MML2	0.0496	0.0964	0.1380	0.1726	0.1991	0.2176	0.2360
MML4	0.0489	0.0916	0.1238	0.1446	0.1561	0.1617	0.1652
MML8	0.0478	0.0843	0.1064	0.1175	0.1225	0.1246	0.1258
LinWei2	0.0484	0.0883	0.1160	0.1328	0.1420	0.1467	0.1500
LinWei4	0.0480	0.0861	0.1105	0.1236	0.1299	0.1327	0.1343
LinWei8	0.0477	0.0838	0.1053	0.1159	0.1207	0.1227	0.1239
Gastw	0.0481	0.0866	0.1116	0.1252	0.1317	0.1345	0.1361
Blom	0.0492	0.0939	0.1318	0.1631	0.1892	0.2121	0.2528

Table 7 – Bias of various estimators of the location of logistic distribution in the presence of multiple location-outliers

(ii) *RSE (estimator proposed in [83]):*

$$C \sum_{i=r+1}^{[n/2]} a_i(X_{n-i+1:n} - X_{i:n}),$$

where each a_i takes the values 0 or 1, and C is a constant. The a_i's and C are chosen so as to make the estimator unbiased and have minimum variance when the X's are IID.

Here, we use the recursive method presented earlier to examine the bias of the above estimators of σ under the multiple location-outlier and multiple scale-outlier models. We also consider the following approximate best linear unbiased estimator presented in [45]:

(iii)

$$\sigma' = \sum_{i=1}^{n} \alpha_i X_{i:n},$$

where $\alpha_i = ci(n+1-i)(c_i-c_{i-1})/(d(n+1)^2)$, $c_i = (ci(n+1-i)/(n+1)^2)\mu_{i:n}[0] - (c(i+1)(n-i)/(n+1)^2)\mu_{i+1:n}[0]$, and $d = \sum_{i=0}^{n} c_i^2$.

(iv) The following modified Jung's estimator in [74]:

$$\hat{\sigma} = c_n \sum_{i=1}^{n} \gamma_i X_{i:n},$$

where $\gamma_i = 9c/(n(n+1)^2(3+\pi^2))\{-(n+1)^2 + 2i(n+1) + 2i(n+1-i)\ln(i/(n+1-i))\}$, and $c_n = 1/(\sum_{i=1}^{n} \gamma_i \mu_{i:n}[0])$ (both of the above estimators are discussed in [63]).

(v) *Winsorized median absolute deviation (WMAD):*

for n odd,

$$\hat{\sigma}(r) = \frac{1}{n-2r}\left[r(X_{n-r:n} - X_{r+1:n}) + \sum_{i=\frac{n-1}{2}+2}^{n-r} X_{i:n} - \sum_{i=r+1}^{\frac{n-1}{2}} X_{i:n}\right],$$

for n even,

$$\hat{\sigma}(r) = \frac{1}{n-2r}\left[r(X_{n-r:n} - X_{r+1:n}) + \sum_{i=\frac{n}{2}+1}^{n-r} X_{i:n} - \sum_{i=r+1}^{\frac{n}{2}} X_{i:n}\right],$$

which, incidentally, is the MLE of σ for a symmetrically Type-II censored sample from a Laplace distribution; see [27].

μ_1	0.5	1.0	1.5	2.0	2.5	3.0	4.0
$n = 20,\ p = 1$							
BLUE0	0.0070	0.0263	0.0542	0.0871	0.1223	0.1584	0.2313
BLUE2	0.0078	0.0275	0.0511	0.0710	0.0842	0.0914	0.0965
BLUE4	0.0082	0.0275	0.0471	0.0606	0.0678	0.0712	0.0733
BLUE8	0.0088	0.0266	0.0413	0.0496	0.0535	0.0552	0.0562
RSE0	0.0070	0.0262	0.0542	0.0871	0.1226	0.1592	0.2337
RSE2	0.0078	0.0275	0.0511	0.0708	0.0838	0.0909	0.0959
RSE4	0.0082	0.0274	0.0468	0.0600	0.0671	0.0704	0.0724
RSE9	0.0088	0.0265	0.0410	0.0490	0.0528	0.0544	0.0554
WMAD0	-0.2573	-0.2432	-0.2233	-0.2006	-0.1765	-0.1519	-0.1022
WMAD2	-0.2213	-0.2061	-0.1885	-0.1739	-0.1645	-0.1594	-0.1558
WMAD4	-0.1860	-0.1707	-0.1556	-0.1454	-0.1401	-0.1376	-0.1361
WMAD6	-0.1626	-0.1474	-0.1341	-0.1261	-0.1222	-0.1205	-0.1195
Blom	0.0066	0.0257	0.0555	0.0940	0.1390	0.1886	0.2955
Jung	0.0069	0.0262	0.0545	0.0882	0.1247	0.1625	0.2393
$n = 20,\ p = 2$							
BLUE0	0.0132	0.0502	0.1043	0.1692	0.2397	0.3131	0.4631
BLUE2	0.0148	0.0540	0.1058	0.1584	0.2032	0.2363	0.2708
BLUE4	0.0158	0.0550	0.0998	0.1357	0.1580	0.1695	0.1770
BLUE8	0.0170	0.0543	0.0884	0.1092	0.1195	0.1240	0.1267
RSE0	0.0132	0.0501	0.1041	0.1687	0.2389	0.3117	0.4604
RSE2	0.0148	0.0541	0.1058	0.1579	0.2021	0.2346	0.2685
RSE4	0.0159	0.0549	0.0992	0.1343	0.1560	0.1672	0.1746
RSE9	0.0171	0.0542	0.0876	0.1078	0.1176	0.1220	0.1245
WMAD0	-0.2526	-0.2253	-0.1865	-0.1415	-0.0936	-0.0445	0.0550
WMAD2	-0.2157	-0.1855	-0.1467	-0.1084	-0.0764	-0.0531	-0.0292
WMAD4	-0.1798	-0.1484	-0.1137	-0.0870	-0.0708	-0.0627	-0.0573
WMAD(r)	-0.1560	-0.1244	-0.0935	-0.0730	-0.0623	-0.0235	0.0812
r	6	6	6	6	6	1	1
Blom	0.0125	0.0485	0.1038	0.1730	0.2511	0.3343	0.5069
Jung	0.0131	0.0500	0.1045	0.1702	0.2423	0.3175	0.4717

Table 8 – Bias of various estimators of the scale of logistic distribution in the presence
of multiple location-outliers

In table 8, we have presented the bias of all these estimators of σ under the
multiple location-outlier model for $n = 20$, $p = 1, 2$, different choices of r, and $\mu = 0$,
$\mu_1 = 0.5(0.5)3.0, 4.0$, $\sigma = \sigma_1 = 1$. For the RSE, BLUE, and WMAD, we have included
$r = 0, 10, 20\%$ of n as well as the value of r that gave the estimator with the smallest
bias. We also observe from table 8 that for small values of μ_1 Blom's estimator is
usually the one with the smallest bias. For these same small values of μ_1, the RSE
and BLUE both increase in bias as r increases while the WMAD decreases in bias as
r increases. On the other hand, as μ_1 increases the RSE and BLUE for larger values
of r begin to decrease in bias while no clear pattern can be seen for the WMAD. In
this same situation, the estimators of Blom and Jung, being approximations to the
full sample BLUE, have a very large bias as well. For larger values of p and μ_1, it is
the WMAD that has the smallest bias.

μ_1	0.5	1.0	1.5	2.0	3.0	4.0
$n = 20,\ p = 1$						
BLUE0	-0.0229	0.0000	0.0260	0.0529	0.1079	0.1633
BLUE2	-0.0273	0.0000	0.0212	0.0360	0.0539	0.0640
BLUE4	-0.0318	0.0000	0.0194	0.0313	0.0445	0.0516
BLUE8	-0.0416	0.0000	0.0175	0.0269	0.0365	0.0414
RSE0	-0.0230	0.0000	0.0260	0.0532	0.1089	0.1654
RSE2	-0.0273	0.0000	0.0212	0.0359	0.0537	0.0638
RSE4	-0.0323	0.0000	0.0193	0.0311	0.0441	0.0510
RSE9	-0.0424	0.0000	0.0174	0.0267	0.0361	0.0409
WMAD0	-0.2804	-0.2626	-0.2439	-0.2251	-0.1871	-0.1490
WMAD2	-0.2493	-0.2273	-0.2112	-0.2002	-0.1869	-0.1795
WMAD4	-0.2197	-0.1927	-0.1773	-0.1681	-0.1580	-0.1526
WMAD6	-0.2015	-0.1698	-0.1548	-0.1465	-0.1377	-0.1332
Blom	-0.0208	0.0000	0.0281	0.0609	0.1343	0.2129
Jung	-0.0225	0.0000	0.0263	0.0540	0.1112	0.1693
$n=20,\ p = 2$						
BLUE2	-0.0548	0.0000	0.0430	0.0743	0.1152	0.1404
BLUE4	-0.0633	0.0000	0.0395	0.0645	0.0935	0.1095
BLUE8	-0.0811	0.0000	0.0357	0.0552	0.0759	0.0866
RSE0	-0.0462	0.0000	0.0519	0.1061	0.2174	0.3301
RSE2	-0.0547	0.0000	0.0430	0.0741	0.1147	0.1397
RSE4	-0.0643	0.0000	0.0393	0.0640	0.0926	0.1084
RSE9	-0.0824	0.0000	0.0355	0.0548	0.0750	0.0855
WMAD0	-0.2984	-0.2626	-0.2253	-0.1876	-0.1116	-0.0355
WMAD2	-0.2713	-0.2273	-0.1946	-0.1713	-0.1412	-0.1228
WMAD4	-0.2462	-0.1927	-0.1614	-0.1420	-0.1198	-0.1077
WMAD6	-0.2321	-0.1698	-0.1393	-0.1218	-0.1029	-0.0929
Blom	-0.0419	0.0000	0.0559	0.1202	0.2624	0.4133
Jung	-0.0452	0.0000	0.0524	0.1077	0.2218	0.3380

Table 9 – Bias of various estimators of the scale of logistic distribution in the presence of multiple scale-outliers

In table 9, we have presented the bias of the above estimators of the scale parameter σ under the multiple scale-outlier model for $n = 10(5)20$, $p = 0(1)3$, and $\mu = \mu_1 = 0$, $\sigma = 1$, $\sigma_1 = 0.5(0.5)2, 3, 4$. For the RSE, BLUE, and WMAD, we have included $r = 0, 10, 20\%$ of n as well as the value of r that gave the estimator with the smallest bias. We also observe from table 9 that each estimator except for the WMAD is quite sensitive to the presence of outliers. As the value of σ_1 increases from 0.5 to 4.0, the bias of each estimator except for the WMAD increases, although much less so for large values of r. On the other hand, the WMAD usually decreases in bias as σ_1 increases. Also, for a given value of σ_1 and n, the bias of each estimator increases considerably as p increases. The bias of the RSE and BLUE are quite comparable, with the forms with large values of r giving the smallest bias (except for the case $\sigma_1 = 0.5$). But the estimators of Blom and Jung, each involving all of the order statistics, both have very large bias compared with the censored forms of RSE

and BLUE. When there are two or more outliers and σ_1 is large, the WMAD usually has the smallest bias, whereas the censored forms of the RSE and BLUE have the smallest bias for small values of p and σ_1.

8.6. Some comments

We have established in Theorem 8.1 some recurrence relations for the single moments of order statistics from INID logistic random variables which enable the recursive computation of single moments of order statistics from a multiple-outlier model. From the values of the bias computed for various linear estimators of the location and scale parameters of the logistic distribution in the presence of multiple outliers in the sample, it is observed that the sample median is the least-biased estimator of the location while the WMAD is the estimator of the scale with the smallest bias in general.

Open Problem 8.6. *A sufficient condition to verify whether the sample median is the least-biased estimator of the location parameter when a single outlier is present in the sample was presented in [55] wherein it was also shown that this condition is satisfied in the case of the logistic distribution. Table 7 reveals that the sample median remains as the least-biased estimator even when multiple outliers are present in the sample. This then raises a question whether there is a version of the condition in [55] for the multiple-outlier situation and that whether this condition is satisfied in the logistic case?*

It is important to mention that an evaluation of the robustness of estimators by bias alone is not sufficient and that it is necessary to evaluate them by variance or mean square error as well. This would, of course, require the computation of product moments of order statistics.

Open Problem 8.7. *The relations in Theorem 8.1 generalize the recurrence relations for the single moments of logistic order statistics in the IID case established in [88] to the INID case. This leaves a question whether there are similar generalizations of the recurrence relations for the product moments of logistic order statistics in the IID case established in [87] to the INID case?*

9. Robust estimation for Laplace distribution

Results 6.2 and 6.3 presented earlier can be used to evaluate the robustness properties of various linear estimators of the location and scale parameters of the Laplace distribution with regard to the presence of one or more outliers in the sample. To this end, let us assume that $X_1, \ldots, X_{n-p}$ are IID variables from a Laplace distribution with probability density function

$$f(x; \mu, \sigma) = \frac{1}{2\sigma} \exp\left\{ -\frac{|x - \mu|}{\sigma} \right\}, \qquad x \in \mathbb{R}, \quad \mu \in \mathbb{R}, \quad \sigma \in \mathbb{R}^+, \qquad (151)$$

and $X_{n-p+1}, \ldots, X_n$ are IID variables from another Laplace distribution (independently of $X_1, \ldots, X_{n-p}$) with probability density function

$$g(x; \mu, a\sigma) = \frac{1}{2a\sigma} \exp\left\{ -\frac{|x - \mu|}{a\sigma} \right\}, \qquad x \in \mathbb{R}, \quad \mu \in \mathbb{R}, \quad a, \sigma \in \mathbb{R}^+.$$

Thus, the two sets of variables are symmetric about μ; let us denote the corresponding standardized variables by Z_i, i.e., $Z_i = (X_i - \mu)/\sigma$ $(i = 1, \ldots, n)$, and the corresponding order statistics by $Z_{1:n} \leq \cdots \leq Z_{n:n}$. Then, Results 6.2 and 6.3 reduce to

$$\mu_{r:n}^{(k)}[p] = \frac{1}{2^n} \left\{ \sum_{t=0}^{p} \left[\sum_{i=p-t}^{\min(r-1, n-t)} \binom{p}{t} \binom{n-p}{n-i-t} v_{r-i:n-i}^{(k)}[t] \right. \right.$$
$$\left. \left. + (-1)^k \sum_{i=\max(r,t)}^{n-p+t} \binom{p}{t} \binom{n-p}{i-t} v_{i-r+1:i}^{(k)}[t] \right] \right\}, \quad 1 \leq r \leq n, \ k \geq 1, \ (152)$$

and

$$\mu_{r,s:n}[p] = \frac{1}{2^n} \left\{ \sum_{t=0}^{p} \left[\sum_{i=p-t}^{\min(r-1, n-t)} \binom{p}{t} \binom{n-p}{n-i-t} v_{r-i,s-i:n-i}[t] \right. \right.$$
$$- \sum_{i=\max(r,p-t)}^{\min(s-1, n-t)} \binom{p}{t} \binom{n-p}{n-i-t} v_{s-i:n-i}[t] \ v_{i-r+1:i}[p-t]$$
$$\left. \left. + \sum_{i=\max(s,t)}^{n-p+t} \binom{p}{t} \binom{n-p}{i-t} v_{i-s+1,i-r+1:i}[t] \right] \right\}, \quad 1 \leq r < s \leq n, \ (153)$$

where $\mu_{r:n}^{(k)}[p]$ and $\mu_{r,s:n}[p]$ denote the single and product moments of order statistics from the p-outlier Laplace sample $Z_1, \ldots, Z_n$, and $v_{r:m}^{(k)}[t]$ and $v_{r,s:m}[t]$ denote the single and product moments of order statistics from the t-outlier exponential sample (obtained by folding Z_i's around zero).

The relations in (152) and (153) were used in [49] to examine the robustness features of different linear estimators of the parameters μ and σ of the Laplace distribution in (151). Through this study, they observed that the Linearly Weighted Mean

$$L_n(r) = \frac{1}{2(\frac{n}{2} - r)^2} \sum_{i=1}^{\frac{n}{2} - r} (2i - 1)(X_{r+i:n} + X_{n-r-i+1:n}) \qquad \text{for } n \text{ even}$$

and the maximum likelihood estimator (see [27])

$$ML(r) = \frac{1}{n - 2r} \left\{ r(X_{n-r:n} - X_{r+1:n}) + \sum_{i=\frac{n}{2}+1}^{n-r} X_{i:n} - \sum_{i=r+1}^{\frac{n}{2}} X_{i:n} \right\} \qquad \text{for } n \text{ even}$$

p	a	0.50	1.0	2.0	3.0	4.0	6.0	8.0	10.0
1	LinWei 5	0.0591	0.0639	0.0677	0.0693	0.0702	0.0711	0.0716	0.0719
	LinWei 6	0.0589	0.0637	0.0675	0.0690	0.0698	0.0707	0.0711	0.0714
	LinWei 7	0.0591	0.0641	0.0678	0.0693	0.0701	0.0709	0.0713	0.0716
2	LinWei 6	0.0544	0.0637	0.0715	0.0749	0.0768	0.0789	0.0800	0.0806
	LinWei 7	0.0545	0.0641	0.0718	0.0751	0.0769	0.0788	0.0799	0.0805
	LinWei 8	0.0552	0.0651	0.0728	0.0760	0.0778	0.0797	0.0807	0.0813
3	LinWei 6	0.0503	0.0637	0.0759	0.0815	0.0848	0.0885	0.0905	0.0918
	LinWei 7	0.0503	0.0641	0.0761	0.0816	0.0847	0.0882	0.0901	0.0912
	LinWei 8	0.0509	0.0651	0.0771	0.0825	0.0856	0.0890	0.0908	0.0919
4	LinWei 6	0.0466	0.0637	0.0806	0.0891	0.0942	0.1001	0.1034	0.1055
	LinWei 7	0.0465	0.0641	0.0808	0.0890	0.0938	0.0993	0.1024	0.1043
	LinWei 8	0.0470	0.0651	0.0818	0.0899	0.0946	0.0999	0.1028	0.1047

Table 10 – Values of (Variance of $L_n(r)$)$/\sigma^2$ for selected values of r

p	a	0.50	1.0	2.0	3.0	4.0	6.0	8.0	10.0
1	ML0	0.0495	0.0494	0.0566	0.0738	0.1009	0.1851	0.3094	0.4736
	ML1	0.0549	0.0549	0.0592	0.0640	0.0679	0.0733	0.0768	0.0793
	ML2	0.0615	0.0616	0.0654	0.0687	0.0711	0.0742	0.0760	0.0772
2	ML0	0.0508	0.0494	0.0686	0.1176	0.1966	0.4444	0.8122	1.3000
	ML(r)	0.0508	0.0494	0.0686	0.1176	0.0949	0.1104	0.1211	0.1289
	ML(r)	0.0562	0.0549	0.0676	0.0873	0.0969	0.1067	0.1130	0.1172
	ML(r)	0.0629	0.0616	0.0723	0.0845	0.1054	0.1128	0.1173	0.1203
				$r = 0, 1, 2$				$r = 2, 3, 4$	
3	ML0	0.0531	0.0494	0.0855	0.1810	0.3364	0.8271	1.5576	2.5281
	ML(r)	0.0531	0.0494	0.0855	0.1115	0.1302	0.1584	0.1782	0.1928
	ML(r)	0.0589	0.0549	0.0802	0.1186	0.1328	0.1524	0.1650	0.1738
	ML(r)	0.0659	0.0616	0.0829	0.1325	0.1443	0.1598	0.1694	0.1758
				$r = 0, 1, 2$				$r = 3, 4, 5$	

Table 11 – Values of (MSE of $ML(r)$)$/\sigma^2$ for selected values of r

were the most efficient estimators of μ and σ, respectively, in the presence of one or more outliers in the sample.

Table 10 presents the values of $\frac{1}{\sigma^2}\mathrm{Var}(L_n(r))$ for different choices of r, p, and a when the sample size $n = 20$. Similarly, table 11 presents the values of $\frac{1}{\sigma^2}\mathrm{MSE}(ML(r))$ for different choices of r, p, and a when the sample size $n = 20$. The robustness feature of these two estimators is evident from these two tables and it is also clear that larger values of r provide more protection against the presence of pronounced outliers but at the cost of a higher premium.

10. Results for some other distributions

In [50], order statistics arising from INID Pareto random variables with probability density functions

$$f_i(x) = v_i x^{-(v_i+1)}, \qquad x \geq 1, \quad v_i > 0,$$

and cumulative distribution functions

$$F_i(x) = 1 - x^{-v_i}, \qquad x \geq 1, \quad v_i > 0,$$

for $i = 1, 2, \ldots, n$, were considered. In this case, the characterizing differential equations are

$$1 - F_i(x) = \frac{1}{v_i} x f_i(x), \qquad x \geq 1, \quad v_i > 0, \quad i = 1, \ldots, n. \tag{154}$$

By using the permanent approach along with the characterizing differential equations in (154), several recurrence relations were derived in [50] for the single and product moments of order statistics arising from INID Pareto random variables. These results were then used to examine the robustness of the maximum likelihood estimators and best linear unbiased estimators of the scale parameter of a one-parameter Pareto distribution and of the location and scale parameters of a two-parameter Pareto distribution under the presence of multiple outliers. They observed that the estimators based on censored samples possess robustness features in general than those based on complete samples. Of course, these were carried out under the assumption that the shape parameter is known; see also [52] for related work on Lomax distribution.

Open Problem 10.1. *Though robust estimation of the scale or location and scale parameters of the Pareto distribution has been discussed, the robust estimation of the shape parameter remains as an open problem and deserves attention.*

In [21], order statistics arising from INID power function random variables with probability density functions

$$f_i(x) = v_i x^{v_i - 1}, \qquad 0 < x < 1, \quad v_i > 0,$$

and cumulative distribution functions

$$F_i(x) = x^{v_i} \qquad 0 < x < 1, \quad v_i > 0,$$

for $i = 1, 2, \ldots, n$, were considered. The characterizing differential equations in this case are

$$F_i(x) = \frac{1}{v_i} x f_i(x), \qquad 0 < x < 1, \quad v_i > 0, \quad i = 1, \ldots, n. \tag{155}$$

By using the permanent approach along with the characterizing differential equations in (155), several recurrence relations were derived in [21] for the single and product moments of order statistics arising from INID power function random variables. They can be used to determine these moments in the case of multiple-outlier model.

11. Miscellanea

In this section, we briefly describe two recent developments wherein permanent representations of distributions of order statistics have proven to be useful. First one concerns ranked set sampling. The basic procedure of obtaining a ranked set sample is as

follows. First, we draw a random sample of size n from the population and order them (without actual measurement, for example, visually). Then, the smallest observation is measured and denoted as $X_{(1)}$, and the remaining are not measured. Next, another sample of size n is drawn and ordered, and only the second smallest observation is measured and denoted as $X_{(2)}$. This procedure is continued until the largest observation of the n-th sample of size n is measured. The collection $\{X_{(1)}, \ldots, X_{(n)}\}$ is called as a *one-cycle ranked set sample* of size n. If we replicate the above procedure m times, we finally get a ranked set sample of total size $N = mn$. The data thus collected in this case is denoted by $\boldsymbol{X}_{\mathrm{RSS}} = \{X_{1(1)}, X_{2(1)}, \ldots, X_{m(1)}, \ldots, X_{1(n)}, X_{2(n)}, \ldots, X_{m(n)}\}$. The ranked set sampling was first proposed in [77] in order to find a more efficient method to estimate the average yield of pasture. Since then, numerous parametric and nonparametric inferential procedures based on ranked set samples have been developed in the literature. For a comprehensive review of various developments on ranked set sampling, one may refer to [47].

It is evident that if the ranking (done by visual inspection, for example) is perfect, then $X_{(i)}$ is distributed exactly as the i-th order statistic from a random sample of size n from a distribution $F(x)$ and hence has its density function and distribution function as in (2) and (1), respectively. Note that in this case, however, $X_{(1)}, \ldots, X_{(n)}$ are mutually independent. Therefore, if the observations from a one-cycle ranked set sample are ordered, the distributions of these ordered observations can be expressed in the permanent form as in (9) and (13) with $f_i(x)$ and $F_i(x)$ replaced by $f_{i:n}(x)$ and $F_{i:n}(x)$ in (2) and (1), respectively. For example, if $X_{1:n}^{\mathrm{ORSS}} < \cdots < X_{n:n}^{\mathrm{ORSS}}$ denote these order statistics obtained from a one-cycle ranked set sample, the density function of $X_{r:n}^{\mathrm{ORSS}}$ $(1 \leq r \leq n)$ can be expressed as

$$f_{X_{r:n}^{\mathrm{ORSS}}}(x) = \frac{1}{(r-1)!(n-r)!} \operatorname{Per} \begin{pmatrix} F_{1:n}(x) & \cdots & F_{n:n}(x) \\ f_{1:n}(x) & \cdots & f_{n:n}(x) \\ 1 - F_{1:n}(x) & \cdots & 1 - F_{n:n}(x) \end{pmatrix} \begin{matrix} \} \ r-1 \\ \} \ 1 \\ \} \ n-r \end{matrix},$$
$$x \in \mathbb{R}.$$

This set of ordered observations has been referred to as *ordered ranked set sample* in [31,32], and has been used to develop efficient inferential procedures in this context.

Another scenario in which permanent representations arise naturally is in the context of progressive censoring. In the model of progressively Type-II censored order statistics, some of the underlying random variables $X_1, \ldots, X_n$ are censored during the observation. In particular, this means that in a life-testing experiment with n independent units, a pre-fixed number R_1 of surviving units are randomly censored from the sample after the first failure time, $\min\{X_1, \ldots, X_n\}$. Then, at the first failure time of the remaining $n - R_1 - 1$ units, R_2 units are censored, and so on. Finally, at the time of the m-th failure, all the remaining $R_m = n - m - R_1 - \cdots - R_{m-1}$ units are censored. For a detailed description of this progressive censoring scheme and related developments, one may refer to [19].

Note that while carrying out this life-test, it is assumed that the units being

tested have IID life-times $X_1, \ldots, X_n$ with distribution function $F(\cdot)$. Instead, if we assume that $X_1, \ldots, X_n$ are INID random variables, then the joint density of these progressively Type-II censored order statistics $(X_{1:m:n}^{(R_1,\ldots,R_m)}, \ldots, X_{m:m:n}^{(R_1,\ldots,R_m)})$ has been expressed in a permanent form recently in [26] as

$$
f_{X_{1:m:n}^{(R_1,\ldots,R_m)},\ldots,X_{m:m:n}^{(R_1,\ldots,R_m)}}(x_1,\ldots,x_m)
$$

$$
= \frac{1}{(n-1)!} \prod_{j=2}^{m} \gamma_j \operatorname{Per}
\begin{pmatrix}
f_1(x_1) & \cdots & f_n(x_1) \\
1 - F_1(x_1) & \cdots & 1 - F_n(x_1) \\
\cdot & \cdots & \cdot \\
f_1(x_m) & \cdots & f_n(x_m) \\
1 - F_1(x_m) & \cdots & 1 - F_n(x_m)
\end{pmatrix}
\begin{matrix}
\} \ 1 \\
\} \ R_1 \\
\\
\} \ 1 \\
\} \ R_m
\end{matrix},
$$

$$
x_1 < x_2 < \cdots < x_m,
$$

where $\gamma_1 = \sum_{i=1}^{m}(R_i+1) = n$ and $\gamma_j = \sum_{i=j}^{m}(R_i+1)$ is the number of units remaining in the experiment after the $(j-1)$-th failure for $j = 2, 3, \ldots, m$. Such permanent forms have been used in [26] to establish some interesting properties of these progressively censored order statistics arising from INID random variables, and have also been applied to discuss the robustness of the maximum likelihood estimator of the mean of an exponential distribution when one or more outliers are possibly present in the observed progressively Type-II censored sample.

References

[1] A. D. Alexandroff, *Zur Theorie der gemischten Volumina von konvexen Körpern. IV: Die gemischten Diskriminanten und die gemischten Volumina*, Mat. Sbornik (3) **45** (1938), 227–251 (Russian with German summary).

[2] W. R. Allen, *A note on the conditional probability of failure when the hazards are proportional*, Oper. Res. **11** (1963), no. 4, 658–659.

[3] D. F. Andrews, P. J. Bickel, F. R. Hampel, P. J. Huber, W. H. Rogers, and J. W. Tukey, *Robust estimates of location: Survey and advances*, Princeton University Press, Princeton, N.J., 1972.

[4] P. Armitage, *The comparison of survival curves*, J. Roy. Statist. Soc. Ser. A **122** (1959), no. 3, 279–300.

[5] B. C. Arnold and N. Balakrishnan, *Relations, bounds and approximations for order statistics*, Lecture Notes in Statistics, vol. 53, Springer-Verlag, New York, 1989.

[6] B. C. Arnold, N. Balakrishnan, and H. N. Nagaraja, *A first course in order statistics*, John Wiley & Sons Inc., New York, 1992.

[7] N. Balakrishnan, *Order statistics from discrete distributions*, Comm. Statist. Theory Methods **15** (1986), no. 3, 657–675.

[8] _____, *Two identities involving order statistics in the presence of an outlier*, Comm. Statist. Theory Methods **16** (1987), no. 8, 2385–2389.

[9] _____, *Relations and identities for the moments of order statistics from a sample containing a single outlier*, Comm. Statist. Theory Methods **17** (1988), no. 7, 2173–2190.

[10] _____, *Recurrence relations among moments of order statistics from two related outlier models*, Biometrical J. **30** (1988), no. 6, 741–746.

[11] _____, *Recurrence relations for order statistics from n independent and non-identically distributed random variables*, Ann. Inst. Statist. Math. **40** (1988), no. 2, 273–277.

[12] _____, *A relation for the covariances of order statistics from n independent and nonidentically distributed random variables*, Statist. Hefte **30** (1989), no. 2, 141–146.

[13] _____, *Recurrence relations among moments of order statistics from two related sets of independent and non-identically distributed random variables*, Ann. Inst. Statist. Math. **41** (1989), no. 2, 323–329.

[14] _____, *Relationships between single moments of order statistics from non-identically distributed variables*, Order Statistics and Nonparametrics: Theory and Applications (Alexandria, 1991) (P. K. Sen and I. A. Salama, eds.), North-Holland, Amsterdam, 1992, pp. 65–78.

[15] _____ (ed.), *Handbook of the logistic distribution*, Statistics: Textbooks and Monographs, vol. 123, Marcel Dekker Inc., New York, 1992.

[16] _____, *Order statistics from non-identical exponential random variables and some applications*, with a discussion and a reply, Comput. Statist. Data Anal. **18** (1994), no. 2, 203–253.

[17] _____, *On order statistics from non-identical right-truncated exponential random variables and some applications*, Comm. Statist. Theory Methods **23** (1994), no. 12, 3373–3393.

[18] _____, *Relationships between product moments of order statistics from non-identically distributed variables*, 4th International Meeting of Statistics in the Basque Country (San Sebastián, Spain, August 4–7, 1992), Recent advances in statistics and probability (M. L. Puri and J. Vilaplana, eds.), VSP Publishers, Amsterdam, 1994.

[19] N. Balakrishnan and R. Aggarwala, *Progressive censoring: Theory, methods, and applications*, Birkhäuser Boston Inc., Boston, MA, 2000.

[20] N. Balakrishnan and R. S. Ambagaspitiya, *Relationships among moments of order statistics in samples from two related outlier models and some applications*, Comm. Statist. Theory Methods **17** (1988), no. 7, 2327–2341.

[21] N. Balakrishnan and K. Balasubramanian, *Order statistics from non-identical power function random variables*, Comm. Statist. Theory Methods **24** (1995), no. 6, 1443–1454.

[22] N. Balakrishnan and V. Barnett, *Outlier-robust estimation of the mean of an exponential distribution under the presence of multiple outliers*, Statistical methods and practice: Recent advances (N. Balakrishnan, N. Kannan, and M. R. Srinivasan, eds.), Narosa Publishing House, New Delhi, 2003, pp. 191–202.

[23] N. Balakrishnan, S. M. Bendre, and H. J. Malik, *General relations and identities for order statistics from nonindependent nonidentical variables*, Ann. Inst. Statist. Math. **44** (1992), no. 1, 177–183.

[24] N. Balakrishnan, P. S. Chan, K. L. Ho, and K. K. Lo, *Means, variances and covariances of logistic order statistics in the presence of an outlier*, technical report, McMaster University, Hamilton, Canada, 1992.

[25] N. Balakrishnan and A. C. Cohen, *Order statistics and inference*: *Estimation methods*, Statistical Modeling and Decision Science, Academic Press Inc., Boston, MA, 1991.

[26] N. Balakrishnan and E. Cramer, *Progressive censoring from heterogeneous distributions with applications to inference*, Ann. Inst. Statist. Math. (2007), to appear.

[27] N. Balakrishnan and C. D. Cutler, *Maximum likelihood estimation of Laplace parameters based on type-II censored samples*, Statistical theory and applications: Papers in honor of Herbert A. David (H. N. Nagaraja, P. K. Sen, and D. F. Morrison, eds.), Springer, New York, 1996, pp. 145–151.

[28] N. Balakrishnan and H. A. David, *A note on the variance of a lightly trimmed mean when multiple outliers are present in the sample*, Statist. Probab. Lett. **55** (2001), no. 4, 339–343.

[29] N. Balakrishnan, Z. Govindarajulu, and K. Balasubramanian, *Relationships between moments of two related sets of order statistics and some extensions*, Ann. Inst. Statist. Math. **45** (1993), no. 2, 243–247.

[30] N. Balakrishnan and N. Kannan, *Variance of a Winsorized mean when the sample contains multiple outliers*, Comm. Statist. Theory Methods **32** (2003), no. 1, 139–149.

[31] N. Balakrishnan and T. Li, *Confidence intervals for quantiles and tolerance intervals based on ordered ranked set samples*, Ann. Inst. Statist. Math. **58** (2006), no. 4, 757—777.

[32] _____, *Ordered ranked set samples and applications to inference*, J. Statist. Plann. Inference (2007), to appear.

[33] N. Balakrishnan and H. J. Malik, *Some general identities involving order statistics*, Comm. Statist. Theory Methods **14** (1985), no. 2, 333–339.

[34] _____, *A note on moments of order statistics*, Amer. Statist. **40** (1986), no. 2, 147–148.

[35] N. Balakrishnan and C. R. Rao (eds.), *Order statistics*: *Theory & methods*, Handbook of Statistics, vol. 16, North-Holland, Amsterdam, 1998.

[36] _____ (ed.), *Order statistics*: *Applications*, Handbook of Statistics, vol. 17, North-Holland, Amsterdam, 1998.

[37] K. Balasubramanian and N. Balakrishnan, *A log-concavity property of probability of occurrence of exactly r arbitrary events*, Statist. Probab. Lett. **16** (1993), no. 3, 249–251.

[38] _____, *Duality principle in order statistics*, J. Roy. Statist. Soc. Ser. B **55** (1993), no. 3, 687–691.

[39] K. Balasubramanian, M. I. Beg, and R. B. Bapat, *On families of distributions closed under extrema*, Sankhyā Ser. A **53** (1991), no. 3, 375–388.

[40] R. B. Bapat and M. I. Beg, *Order statistics for nonidentically distributed variables and permanents*, Sankhyā Ser. A **51** (1989), no. 1, 79–93.

[41] _____, *Identities and recurrence relations for order statistics corresponding to nonidentically distributed variables*, Comm. Statist. Theory Methods **18** (1989), no. 5, 1993–2004.

[42] R. B. Bapat and S. C. Kochar, *Characterizations of identically distributed independent random variables using order statistics*, Statist. Probab. Lett. **17** (1993), no. 3, 225–230.

[43] V. Barnett and T. Lewis, *Outliers in statistical data*, 3rd ed., John Wiley & Sons Ltd., Chichester, 1994.

[44] M. I. Beg, *Recurrence relations and identities for product moments of order statistics corresponding to non-identically distributed variables*, Sankhyā Ser. A **53** (1991), no. 3, 365–374.

[45] G. Blom, *Statistical estimates and transformed beta-variables*, John Wiley and Sons, Inc., New York, 1958.

[46] P. Capéraà and L.-P. Rivest, *On the variance of the trimmed mean*, Statist. Probab. Lett. **22** (1995), no. 1, 79–85.

[47] Z. Chen, Z. Bai, and B. K. Sinha, *Ranked set sampling: Theory and applications*, Lecture Notes in Statistics, vol. 176, Springer-Verlag, New York, 2004.

[48] M. S. Chikkagoudar and S. H. Kunchur, *Estimation of the mean of an exponential distribution in the presence of an outlier*, Canad. J. Statist. **8** (1980), no. 1, 59–63.

[49] A. Childs and N. Balakrishnan, *Some extensions in the robust estimation of parameters of exponential and double exponential distributions in the presence of multiple outliers*, Robust inference (G. S. Maddala and C. R. Rao, eds.), Handbook of Statistics, vol. 15, North-Holland, Amsterdam, 1997, pp. 201–235.

[50] ———, *Generalized recurrence relations for moments of order statistics from non-identical Pareto and truncated Pareto random variables with applications to robustness*, Order statistics: Theory & methods, Handbook of Statistics, vol. 16, North-Holland, Amsterdam, 1998, pp. 403–438.

[51] ———, *Relations for order statistics from non-identical logistic random variables and assessment of the effect of multiple outliers on the bias of linear estimators*, J. Statist. Plann. Inference **136** (2006), no. 7, 2227–2253.

[52] A. Childs, N. Balakrishnan, and M. Moshref, *Order statistics from non-identical right-truncated Lomax random variables with applications*, Statist. Papers **42** (2001), no. 2, 187–206.

[53] H. A. David, *Order statistics*, 2nd ed., John Wiley & Sons Inc., New York, 1981.

[54] H. A. David and N. Balakrishnan, *Product moments of order statistics and the variance of a lightly trimmed mean*, Statist. Probab. Lett. **29** (1996), no. 1, 85–87.

[55] H. A. David and J. K. Ghosh, *The effect of an outlier on L-estimators of location in symmetric distributions*, Biometrika **72** (1985), no. 1, 216–218.

[56] H. A. David, W. J. Kennedy, and R. D. Knight, *Means, variances, and covariances of normal order statistics in the presence of an outlier*, Selected Tables in Mathematical Statistics, vol. 5, 1977, pp. 75–204.

[57] H. A. David and H. N. Nagaraja, *Order statistics*, 3rd ed., Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, 2003.

[58] H. A. David and V. S. Shu, *Robustness of location estimators in the presence of an outlier*, Contributions to survey sampling and applied statistics: Papers in honour of H. O. Hartley (H. A. David, ed.), Academic Press, New York, 1978, pp. 235–250.

[59] S. Dharmadhikari and K. Joag-Dev, *Unimodality, convexity, and applications*, Academic Press Inc., Boston, MA, 1988.

[60] Z. Govindarajulu, *On moments of order statistics and quasi-ranges from normal populations*, Ann. Math. Statist. **34** (1963), 633–651.

[61] ———, *Relationships among moments of order statistics in samples from two related populations*, Technometrics **5** (1963), 514–518.

[62] ———, *Best linear estimates under symmetric censoring of the parameters of a double exponential population*, J. Amer. Statist. Assoc. **61** (1966), 248–258.

[63] S. S. Gupta and M. Gnanadesikan, *Estimation of the parameters of the logistic distribution*, Biometrika **53** (1966), 565–570.

[64] S. Hande, *A note on order statistics for nonidentically distributed variables*, Sankhyā Ser. A **56** (1994), no. 2, 365–368.

[65] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 1988. Reprint of the 1952 edition.

[66] H. L. Harter and N. Balakrishnan, *CRC handbook of tables for the use of order statistics in estimation*, CRC Press, Boca Raton, FL, 1996.

[67] W. Hoeffding, *On the distribution of the number of successes in independent trials*, Ann. Math. Statist. **27** (1956), 713–721.

[68] P. C. Joshi, *Recurrence relations for the mixed moments of order statistics*, Ann. Math. Statist. **42** (1971), 1096–1098.

[69] ———, *Efficient estimation of the mean of an exponential distribution when an outlier is present*, Technometrics **14** (1972), 137–143.

[70] ———, *Two identities involving order statistics*, Biometrika **60** (1973), 428–429.

[71] ———, *Recurrence relations between moments of order statistics from exponential and truncated exponential distributions*, Sankhyā Ser. B **39** (1978), no. 4, 362–371.

[72] ———, *A note on the mixed moments of order statistics from exponential and truncated exponential distributions*, J. Statist. Plann. Inference **6** (1982), no. 1, 13–16.

[73] P. C. Joshi and N. Balakrishnan, *Recurrence relations and identities for the product moments of order statistics*, Sankhyā Ser. B **44** (1982), no. 1, 39–49.

[74] J. Jung, *On linear estimates defined by a continuous weight function*, Ark. Mat. **3** (1956), 199–209.

[75] B. K. Kale and S. K. Sinha, *Estimation of expected life in the presence of an outlier observation*, Technometrics **13** (1971), 755–759.

[76] H. J. Malik, N. Balakrishnan, and S. E. Ahmed, *Recurrence relations and identities for moments of order statistics*, I: *Arbitrary continuous distribution*, Comm. Statist. Theory Methods **17** (1988), no. 8, 2623–2655.

[77] G. A. McInyre, *A method for unbiased selective sampling, using ranked sets*, Aust. J. Agric. Res. **3** (1952), 385–390.

[78] H. Minč, *Permanents*, with a foreword by M. Marcus, Encyclopedia of Mathematics and its Applications, vol. 6, Addison-Wesley Publishing Co., Reading, Mass., 1978.

[79] ———, *Theory of permanents, 1978–1981*, Linear Multilinear Algebra **12** (1983), no. 4, 227–263.

[80] ———, *Theory of permanents, 1982–1985*, Linear Multilinear Algebra **21** (1987), no. 2, 109–148.

[81] H. N. Nagaraja, *Order statistics from discrete distributions*, with a discussion and a rejoinder by the author, Statistics **23** (1992), no. 3, 189–216.

[82] G. Pólya and G. Szegő, *Problems and theorems in analysis*, II: *Theory of functions, zeros, polynomials, determinants, number theory, geometry*, Revised and enlarged translation by C. E. Billigheimer of the fourth German edition, Die Grundlehren der Mathematischen Wissenschaften, vol. 216, Springer-Verlag, New York, 1976.

[83] K. Raghunandanan and R. Srinivasan, *Simplified estimation of parameters in a logistic distribution*, Biometrika **57** (1970), 677–678.

[84] Y. S. Sathe and S. M. Bendre, *Log-concavity of probability of occurrence of at least r independent events*, Statist. Probab. Lett. **11** (1991), no. 1, 63–64.

[85] P. K. Sen, *A note on order statistics for heterogeneous distributions*, Ann. Math. Statist. **41** (1970), 2137–2139.

[86] J. Sethuraman, *On a characterization of the three limiting types of the extreme*, Sankhyā Ser. A **27** (1965), 357–364.

[87] B. K. Shah, *On the bivariate moments of order statistics from a logistic distribution*, Ann. Math. Statist. **37** (1966), 1002–1010.

[88] _____, *Note on moments of a logistic order statistics*, Ann. Math. Statist. **41** (1970), 2150–2152.

[89] G. P. Sillitto, *Some relations between expectations of order statistics in samples of different sizes*, Biometrika **51** (1964), 259–262.

[90] K. S. Srikantan, *Recurrence relations between the pdf's of order statistics, and some applications*, Ann. Math. Statist. **33** (1962), 169–177.

[91] B. L. van der Waerden, *Aufgabe 45*, Jber d. D.M.V. **35** (1926), 117.

[92] J. H. van Lint, *Notes on Egoritsjev's proof of the van der Waerden conjecture*, Linear Algebra Appl. **39** (1981), 1–8.

[93] R. J. Vaughan and W. N. Venables, *Permanent expressions for order statistic densities*, J. Roy. Statist. Soc. Ser. B **34** (1972), 308–310.