Generalized Densities of Order Statistics

Charles M. Goldie and Ross A. Maller

University of Sussex
and University of Western Australia

June 30, 1997

Let $X_1, \dots, X_n$ be independent identically distributed random variables with distribution F. We derive expressions for generalized joint 'densities' of order statistics of $X_1, \dots, X_n$, for arbitrary distributions F, in terms of Radon-Nikodym derivatives with respect to product measures based on F. We then give formulae for conditional distributions of order statistics and use them to derive results concerning Markov properties of order statistics, formulae for distributions of trimmed sums, and other useful representations. Our approach leads to simple and natural expressions which appear not to have been given before.

Keywords & Phrases: Markov property, Radon-Nikodym derivative, regular conditional probability distribution, trimmed sum
AMS 1991 mathematics subject classification: 62E15

1. Introduction

Let X, $X_1, \dots, X_n$ be independent r.v.s (random variables) in $\mathbb{R}$ with common distribution F, and let $M_n^{(1)} \geq \cdots \geq M_n^{(n)}$ be the sample $X_1, \dots, X_n$ in decreasing order. We will derive here expressions for generalized joint 'densities' of order statistics for arbitrary distributions F, in particular, without making continuity assumptions on F. These are defined in terms of Radon-Nikodym (R-N) derivatives of the joint distributions of the order statistics with respect to product measures based on F, which in our view is the most natural formulation for a general F that may have continuous and discrete components.

Armed with these expressions, we give formulae for conditional distributions of order statistics for general F, which unify and extend currently available theory. We demonstrate that these formulae are more than just notational conveniences by deriving easily a number of far-reaching results concerning Markov properties of order statistics, formulae for distributions of trimmed sums, and other representations that have been used in the literature.

An important starting point for our development is that an order statistic *always equals a member of the original sample.* In §2 we give a general formula for the distribution of a statistic with that property, which in §3 we specialize to the case of a single order statistic $M_n^{(r)}$, giving its F-density (the Radon-Nikodym derivative of its distribution with respect to F) in several equivalent forms.

The foregoing idea extends readily to higher-dimensional settings. In §4 we give the joint density of the n-tuple of order statistics with respect to n-fold product F-distribution. The result is a very simple expression which appears, however, not to have been given before. (See Theorem 4.1 below.) Our methods also allow ready

The authors are grateful to Simon Fitzpatrick for some useful comments, and to a referee for accurate error-detection. The first author thanks the Department of Mathematics, University of Western Australia, for hospitality and the provision of facilities.

calculation of conditional distributions, as we demonstrate in §5 for several sub-families of the order statistics. For ready applicability we exhibit each of these conditional distributions as a regular conditional probability distribution (rcpd) on the *value-* or *image-space* of the random objects concerned, instead of on the base probability space as is usual in the literature. Again, our view is that this is much the most natural approach for an applied setting.

As applications we give in §6 an expeditious proof of a characterization in [Arnold *et al.* 1984] of when the order statistics $M_n^{(r)}$ form a Markov sequence in r, and, in §7, bounds for the distributions of trimmed sums which have been used in various applications. Other authors often approach order statistics for general distributions through quantile constructions based on uniformly distributed samples that live on an enriched probability space. In §8 we compare this technique with our R-N approach.

Finally in this section, let us explain some notation, which leads to symmetric, easily remembered formulae: we usually regard F as a probability measure rather than as a cumulative distribution function, so that $F[y,\infty)$ stands for $P(X \in [y,\infty)) = P(X \geq y)$, and $F^j[y,\infty)$ is its j^{th} power, $(F[y,\infty))^j$. Likewise $F(y,\infty) := P(X \in (y,\infty)) = P(X > y)$, and similarly for $F(-\infty,y] = P(X \leq y)$ and $F(-\infty,y) = P(X < y)$.

We denote the σ-algebra of Borel sets in a topological space S by $\mathcal{B}(S)$.

2. Radon-Nikodym Derivative.

With $X_1, \ldots, X_n$ as in §1 we consider a r.v. Y for which the property

$$Y \text{ equals one of } X_1, \ldots, X_n \tag{A}$$

holds with probability 1. We seek the distribution G of Y, which again we write as a probability measure on $\mathcal{B}(\mathbb{R})$:

$$G(B) = P(Y \in B) \qquad (B \in \mathcal{B}(\mathbb{R})).$$

For each $j = 1, 2, \ldots$ let $\mathcal{I}_j$ denote the collection of all dyadic intervals $((k-1)2^{-j}, k2^{-j}]$ with $k \in \mathbb{Z}$. Choose $x \in \mathbb{R}$, then for each j there is a unique $I_j(x) \in \mathcal{I}_j$ such that $x \in I_j(x)$. We set

$$d_j(x) := \frac{G(I_j(x))}{F(I_j(x))} = \frac{P(Y \in I_j(x))}{P(X \in I_j(x))}$$

if the denominator is non-zero. If

$$D_F G(x) := \lim_{j\to\infty} d_j(x)$$

exists it is called the *derivative* of the measure G, with respect to the measure F and the 'sequence of nets' $(\mathcal{I}_j)$. At points x where the above does not yield a value for $D_F G(x)$ we define $D_F G(x) := 0$. Then we have the following:

THEOREM 2.1. *G is given by*

$$G(B) = \int_B D_F G(x)\, dF(x) \qquad (B \in \mathcal{B}(\mathbb{R})), \tag{2.1}$$

that is, $D_F G$ is a version of the Radon-Nikodym derivative dG/dF.

Remark 2.2. The point is that we get a straightforward *limiting procedure* to calculate the R-N derivative. Use of dyadic rationals is not essential, but is convenient. One can reasonably call this procedure 'differentiating G with respect to F'. We shall henceforth write dG/dF in place of $D_F G$ and refer to dG/dF as the F-*density* of G or of Y. By our construction, this exists at each $x \in \mathbb{R}$.

Remark 2.3. Then also, for any measurable function h,

$$Eh(Y) = \int_{\mathbb{R}} h \frac{dG}{dF} \, dF \tag{2.2}$$

in the sense that if either side is well-defined, so is the other and the two are equal.

Remark 2.4. Let A_F denote the set of atoms of F: the set of all x such that $F\{x\} = P(X = x) > 0$. In the limit procedure that defines $\frac{dG}{dF}(x)$ the denominators $P(X \in I_j(x))$ converge to $P(X = x)$ and the numerators to $P(Y = x)$, so we have the formula

$$\frac{dG}{dF}(x) = \frac{P(Y = x)}{P(X = x)} = \frac{G\{x\}}{F\{x\}} \qquad (x \in A_F) \tag{2.3}$$

for those x that are atoms of F.

Remark 2.5. By property (A),

$$P(Y \in B) \leq \sum_{i=1}^{n} P(Y = X_i \in B) \leq nF(B); \tag{2.4}$$

hence

$$0 \leq \frac{dG}{dF}(x) \leq n \qquad (x \in \mathbb{R}).$$

PROOF OF THEOREM 2.1: (2.4) implies that $G(B) = 0$ whenever $F(B) = 0$, that is, G is absolutely continuous with respect to F. In that circumstance the R-N theorem gives the existence of a non-negative $h(\cdot)$ such that

$$G(B) = \int_B h \, dF \qquad (B \in \mathcal{B}(\mathbb{R})).$$

A sufficiently general form of the theorem on differentiation of an indefinite integral, e.g. [Zaanen 1958, §32, Theorem 3], now gives $h(x) = \lim_{j\to\infty} d_j(x)$ for F-a.a. x, and our conclusion follows. (We abbreviate 'almost all' by 'a.a.') Zaanen's theorem applies here once we know that the sequence $(\mathcal{I}_k)$ is 'regular', in the sense he defines, with respect to F. To check that we need only, in view of [Zaanen 1958, §32, Lemma 4], establish the following for each member A of a generating semi-ring of sets for F:

> given $\varepsilon > 0$ there exists a sequence (I_j) of dyadic intervals from $\bigcup_1^\infty \mathcal{I}_k$, such that $F\big(A \setminus \bigcup_1^\infty I_j\big) = 0$ and $F\big(\bigcup_1^\infty I_j \setminus A\big) \leq \varepsilon$.

However, since $\{\emptyset\} \cup \bigcup_1^\infty \mathcal{I}_k$ is itself a generating semi-ring for F this property holds trivially. $\square$

Remark 2.6. Property (A) is clearly more restrictive than absolute continuity of G with respect to F, $G \ll F$, which we have noted that (A) implies. Indeed (A) implies that G and F are *mutually absolutely continuous:* $G(B) = 0$ if and only if $F(B) = 0$. To see that, note that $F(B) > 0$ implies $P(X_i \in B$ for all $i = 1, \dots, n) = F^n(B) > 0$, so that, by (A), $G(B) = P(Y \in B) > 0$.

Theorem 2.1 needs only $G \ll F$ rather than (A), as the proof makes clear. However we put forward (A) as a useful sufficient condition, easy to check, for the context of functions of a finite sample.

The mutual absolute continuity of F and G implies that G is concentrated on a Borel set B if and only if F is also concentrated on B. (A measure ν is said to be *concentrated on* B if $\nu(B^c) = 0$.) One way to define the *support* of a probability measure is as the smallest closed set of measure 1. G and F then have the same supports by this definition.

A third consequence of the mutual absolute continuity is that not only does G have an F-density $dG/dF \geq 0$, but the F-density is positive F-a.e.

3. Density of the r^{th} Order Statistic.

We shall apply §2 to order statistics. First some remarks on their definition. While $M_n^{(1)}, \dots, M_n^{(n)}$ as defined in §1 are well-defined r.v.s, there is not by that definition an identification of each $M_n^{(r)}$ with a unique X_i from $X_1, \dots, X_n$. For that we need first a permutation $i_n : \{1, \dots, n\} \to \{1, \dots, n\}$ with the property that $X_{i_n(1)} \geq X_{i_n(2)} \geq \cdots \geq X_{i_n(n)}$, and we can then set $M_n^{(r)} := X_{i_n(r)}$ for $r = 1, \dots, n$. This fuller 'tie-breaking' definition of the order statistics is useful in some circumstances, although not essential for our development. One possible definition of $i_n(\cdot)$ is as follows. Let $i_n(1) := \min\{i : X_i \geq X_j,\ 1 \leq j \leq n\}$, and for $r = 1, \dots, n-1$ let

$$i_n(r+1) := \min\big\{i \notin \{i_n(1), \dots, i_n(r)\} : X_i \geq X_j \text{ for all } j \notin \{i_n(1), \dots, i_n(r)\}\big\}.$$

We now fix $r \in \{1, \dots, n\}$ and take Y to be $M_n^{(r)}$, the r^{th} order-statistic. This r.v. certainly has property (A), so by Theorem 2.1 its distribution $G = F_n^{(r)}$ is given by (2.1), and we need now to give an explicit formula for the integrand

$$\frac{dF_n^{(r)}}{dF}(x),$$

the F-density of $M_n^{(r)}$. It is natural to denote this function $f_n^{(r)}$.

Recall that A_F denotes the set of points of discontinuity of F.

THEOREM 3.1. *The F-density of $M_n^{(r)}$ is given for F-a.a. $x \in \mathbb{R}$ by*

$$f_n^{(r)}(x) = \frac{dF_n^{(r)}}{dF}(x) = r\binom{n}{r} F^{n-r}(-\infty, x) F^{r-1}(x, \infty) \quad (x \notin A_F) \tag{3.1}$$

$$= \frac{P(M_n^{(r)} = x)}{F\{x\}} \quad (x \in A_F). \tag{3.2}$$

The proof depends on the following result. This is of value in its own right, as it gives three alternative single-formula evaluations of $f_n^{(r)}$, and a fuller evaluation of $f_n^{(r)}(x)$ for $x \in A_F$.

PROPOSITION 3.2. *The F-density of $M_n^{(r)}$ is given for F-a.a. $x \in \mathbb{R}$ by*

$$f_n^{(r)}(x) = \frac{dF_n^{(r)}}{dF}(x) = \sum_{j=0}^{r-1} \binom{n}{j} \left(F^j(x,\infty) \sum_{i=0}^{n-j-1} F^i(-\infty,x] F^{n-j-i-1}(-\infty,x) \right. \\ \left. - F^{n-j}(-\infty,x) \sum_{i=0}^{j-1} F^i[x,\infty) F^{j-i-1}(x,\infty) \right) \tag{3.3}$$

(interpret $\sum_{i=0}^{-1}$ as 0), or alternatively

$$f_n^{(r)}(x) = \frac{dF_n^{(r)}}{dF}(x) = \sum_{j=0}^{n-r} \binom{n}{j} \left(F^j(-\infty,x) \sum_{i=0}^{n-j-1} F^i[x,\infty) F^{n-j-i-1}(x,\infty) \right. \\ \left. - F^{n-j}(x,\infty) \sum_{i=0}^{j-1} F^i(-\infty,x] F^{j-i-1}(-\infty,x) \right), \tag{3.4}$$

or

$$f_n^{(r)}(x) = \frac{dF_n^{(r)}}{dF}(x) = \sum_{j=0}^{r-1} \sum_{i=0}^{n-r} \frac{n!}{i!j!(n-i-j)!} F^i(-\infty,x) F^{n-i-j-1}\{x\} F^j(x,\infty). \tag{3.5}$$

When $x \in A_F$,

$$f_n^{(r)}(x) = \frac{dF_n^{(r)}}{dF}(x) = \sum_{j=0}^{r-1} \binom{n}{j} \frac{F^j(x,\infty)F^{n-j}(-\infty,x] - F^j[x,\infty)F^{n-j}(-\infty,x)}{F\{x\}}. \tag{3.6}$$

Remark 3.3. While (3.3) and (3.4) look more complicated than (3.5), the $r = n$ case of (3.4) yields

$$P(M_n^{(n)} \in B) = \int_B \sum_{i=0}^{n-1} F^i[x,\infty) F^{n-i-1}(x,\infty)\, dF(x) \qquad (B \in \mathcal{B}(\mathbb{R})) \tag{3.7}$$

which seems to be the simplest evaluation of the distribution of $\min(X_1, \ldots, X_n)$. For general F, such evaluations are not readily found in the literature, and indeed [Kesten & Maller 1992] for example felt obliged to give a proof of this formula. Equally, the $r = 1$ case of (3.3) gives the distribution of the maximum in the form:

$$P(M_n^{(1)} \in B) = \int_B \sum_{i=0}^{n-1} F^i(-\infty,x] F^{n-i-1}(-\infty,x)\, dF(x) \qquad (B \in \mathcal{B}(\mathbb{R})). \tag{3.8}$$

Remark 3.4. The formulae (3.1)–(3.5) can be stated in 'intuitive' forms; for instance, after rewriting (3.1) in the form

$$P(M_n^{(r)} \in dx) = \frac{n!}{(r-1)!1!(n-r)!} F^{n-r}(-\infty, x) F(dx) F^{r-1}(x, \infty)$$

(which is completely equivalent to (3.1)), we can attempt to put this identity in words as 'for $M_n^{(r)}$ to be in dx you need $n-r$ of the X_i below x, one of them in dx, and the remaining $r-1$ above x'. However, this verbal explanation works only when F is continuous at x; we have to supplement it with (3.6) for other x. In general, one cannot be sure that the intuitive reasoning leads to the right answer. What we are presenting is a workable calculus that reliably leads to correct formulae.

Proof of Proposition 3.2: The event $M_n^{(r)} \le b$ occurs if and only if at most $r-1$ of $X_1, \dots, X_n$ exceed b. Hence

$$G(-\infty, b] = P(M_n^{(r)} \le b) = \sum_{j=0}^{r-1} \binom{n}{j} F^{n-j}(-\infty, b] F^j(b, \infty). \tag{3.9}$$

Pick $a < b$ and subtract $G(-\infty, a]$, expanded similarly, to obtain

$$G(a, b] = \sum_{j=0}^{r-1} \binom{n}{j} \Big(F^j(b, \infty) \big(F^{n-j}(-\infty, b] - F^{n-j}(-\infty, a] \big) \\ - F^{n-j}(-\infty, a] \big(F^j(a, \infty) - F^j(b, \infty) \big) \Big).$$

Thus when $F(a, b] = F(-\infty, b] - F(-\infty, a] > 0$ we find

$$\frac{G(a, b]}{F(a, b]} = \sum_{j=0}^{r-1} \binom{n}{j} \left(F^j(b, \infty) \sum_{i=0}^{n-j-1} F^i(-\infty, b] F^{n-j-i-1}(-\infty, a] \right. \\ \left. - F^{n-j}(-\infty, a] \sum_{i=0}^{j-1} F^i(a, \infty) F^{j-i-1}(b, \infty) \right).$$

Fix $x \in \mathbb{R}$. The derivative $\frac{dG}{dF}(x)$ is the limit of this quotient as a and b run through certain dyadic rationals $a_l \to x$ and $b_l \to x$, where $a_l < x \le b_l$. So (3.3) follows.

The proofs of (3.4) and (3.6) are similar and are omitted.

For (3.5), pick $a < b$ and observe that $a < M_n^{(r)} \le b$ if and only if i of the r.v.s $X_1, \dots, X_n$ are at most a, where $i \le n-r$, and j of them exceed b, where $j \le r-1$, the remaining variables being in $(a, b]$. Thus

$$G(a, b] = \sum_{i=0}^{n-r} \sum_{j=0}^{r-1} \frac{n!}{i!j!(n-i-j)!} F^i(-\infty, a] F^{n-i-j}(a, b] F^j(b, \infty).$$

On dividing by $F(a, b]$ and letting a and b converge to a chosen x (in the support of F) from below and above through dyadic rationals, we deduce (3.5). □

Proof of Theorem 3.1: (3.1) follows from (3.5), while (3.2) is immediate from (3.6) and (3.9), or directly from the definition of R-N derivative. □

COROLLARY 3.5. *For $a < b$,*

$$P(a < M_n^{(r)} \le b) = r\binom{n}{r}\int_{F(-\infty,a]}^{F(-\infty,b]} (1-u)^{r-1}u^{n-r}\,du \tag{3.10}$$

$$= r\binom{n}{r}\int_{F(b,\infty)}^{F(a,\infty)} u^{r-1}(1-u)^{n-r}\,du. \tag{3.11}$$

Hence

$$P(M_n^{(r)} = x) = F_n^{(r)}\{x\} = r\binom{n}{r}\int_{F(-\infty,x)}^{F(-\infty,x]} (1-u)^{r-1}u^{n-r}\,du. \tag{3.12}$$

PROOF: By (3.9),

$$F_n^{(r)}(a,b] = \sum_{j=0}^{r-1}\binom{n}{j}\big(F^j(b,\infty)F^{n-j}(-\infty,b] - F^j(a,\infty)F^{n-j}(-\infty,a]\big).$$

Use the identity

$$r\binom{n}{r}\int_a^b u^{r-1}(1-u)^{n-r}\,du = \sum_{j=0}^{r-1}\binom{n}{j}\big(a^j(1-a)^{n-j} - b^j(1-b)^{n-j}\big)$$

on this to obtain (3.11). For (3.10), difference the formula

$$F_n^{(r)}(a,\infty) = \sum_{j=0}^{n-r}\binom{n}{j}F^{n-j}(a,\infty)F^j(-\infty,a],$$

and continue similarly. □

This last Corollary has an explanation and alternative proof in terms of uniform order statistics, which we will come to in §8. (3.12) was obtained in [Khatri 1962].

As another application of (3.12), (3.1), and (3.2), we obtain

COROLLARY 3.6. *The F-density $f_n^{(r)}$ of $M_n^{(r)}$ satisfies, for all $x \in \mathbb{R}$,*

$$r\binom{n}{r}F^{r-1}(x,\infty)F^{n-r}(-\infty,x) \le f_n^{(r)}(x) \le r\binom{n}{r}F^{r-1}[x,\infty)F^{n-r}(-\infty,x], \tag{3.13}$$

with equality if and only if $F\{x\} = 0$.

We refer to [Arnold & Balakrishnan 1989] for other bounds on order statistics, mainly involving moments.

4. Joint Distributions.

Joint distributions of all or some of the order statistics can equally well be obtained by the method of §2, since it is completely general. Write $\mathbf{M}_n^{(r)} := (M_n^{(1)}, \ldots, M_n^{(r)})$ for $r \le n$.

Now $\mathbf{M}_n^{(n)}$ takes values in the set $\mathbb{R}^{n\downarrow}$ of all $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ with $x_{i-1} \geq x_i$ for $i = 2, \dots, n$. We decompose this set into the union of the disjoint sets $E_{\mathbf{s}}^{(n)}$ where $\mathbf{s} = (s_1, \dots, s_k)$ is a multi-index with $k \leq n$ components such that the s_i are positive integers with $\sum_1^k s_i = n$. Let $\mathbf{S}_n$ denote the set of such multi-indices. The generic element of $E_{\mathbf{s}}^{(n)}$ is $\mathbf{x} = \mathbf{x_y}$ where $\mathbf{y} = (y_1, \dots, y_k)$ with the y_i real and $y_1 > y_2 > \cdots > y_k$; then $\mathbf{x_y}$ is the element of $\mathbb{R}^{n\downarrow}$ for which the first s_1 components all have value y_1, the next s_2 components all have value y_2, and so on until the final s_k components which all have value y_k. That is, $\mathbf{x_y} = (x_1, \dots, x_n)$ with $x_j := y_i$ for those j such that $\sum_1^{i-1} s_l < j \leq \sum_1^i s_l$. When F is continuous $\mathbf{M}_n^{(n)}$ lies with probability 1 in $E_{\mathbf{1}_n}^{(n)}$ where $\mathbf{1}_n$ is an n-tuple of 1s; that is, in the set of $\mathbf{x} = (x_1, \dots, x_n)$ with $x_{i-1} > x_i$ for $i = 2, \dots, n$. But this is not so for general F.

Let $F^{n\otimes}$ denote product measure on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ built from components F. Denote by G_n the distribution of $\mathbf{M}_n^{(n)}$ on $\mathbb{R}^n$. We first need to evaluate the $F^{n\otimes}$-density of G_n, i.e. the R-N derivative $dG_n/dF^{n\otimes}$, with defining property

$$P(\mathbf{M}_n^{(n)} \in B) = G_n(B) = \int_B \frac{dG_n}{dF^{n\otimes}}\, dF^{n\otimes} \qquad (B \in \mathcal{B}(\mathbb{R}^n)). \tag{4.1}$$

For $\mathbf{s} = (s_1, \dots, s_k) \in \mathbf{S}_n$ we write $\mathbf{s}! := s_1! \cdots s_k!$, a 'multi-factorial', and $\binom{n}{\mathbf{s}} := n!/\mathbf{s}!$, a multinomial coefficient. We also set $1\{T\}$ to be 1 if statement T is true and 0 if T is false.

THEOREM 4.1. *For $F^{n\otimes}$-a.a. $\mathbf{x} \in \mathbb{R}^n$,*

$$\frac{dG_n}{dF^{n\otimes}}(\mathbf{x}) = \sum_{\mathbf{s}\in\mathbf{S}_n} \binom{n}{\mathbf{s}} 1\{\mathbf{x} \in E_{\mathbf{s}}^{(n)}\}. \tag{4.2}$$

PROOF: We have (4.1) and will evaluate the integrand by the method of §2. Thus for each $j = 1, 2, \dots$ let $\mathcal{I}_j^n$ denote the collection of all n-dimensional dyadic rectangles $\prod_{i=1}^n ((k_i - 1)2^{-j}, k_i 2^{-j}]$ with the $k_i \in \mathbb{Z}$. Choose $\mathbf{x} \in \mathbb{R}^{n\downarrow}$, then for each j there is a unique $I_j(\mathbf{x}) \in \mathcal{I}_j^n$ such that $\mathbf{x} \in I_j(\mathbf{x})$. Set

$$d_j(\mathbf{x}) := \frac{G_n(I_j(\mathbf{x}))}{F^{n\otimes}(I_j(\mathbf{x}))} = \frac{P(\mathbf{M}_n^{(n)} \in I_j(\mathbf{x}))}{P((X_1, \dots, X_n) \in I_j(\mathbf{x}))},$$

if the denominator is non-zero. Then exactly as in §2 we find that by letting

$$\frac{dG_n}{dF^{n\otimes}}(\mathbf{x}) := \lim_{j\to\infty} d_j(\mathbf{x})$$

if the limit exists, and 0 otherwise, we get a correct evaluation of the R-N derivative.

We must have $\mathbf{x} \in E_{\mathbf{s}}^{(n)}$ for some $\mathbf{s} \in \mathbf{S}_n$. By construction, identical components of $\mathbf{x}$ will give rise to identical component intervals of $I_j(\mathbf{x})$. So $I_j(\mathbf{x})$ will have as its first s_1 component intervals an interval $((l_1 - 1)2^{-j}, l_1 2^{-j}] \ni y_1$, as its next s_2 component intervals an interval $((l_2 - 1)2^{-j}, l_2 2^{-j}] \ni y_2$, and so on. As soon as j is so large that

$2^{-j+1} < \min_{i=2,\ldots,k}(y_i - y_{i-1})$ these k real dyadic intervals will be disjoint. Write $t_0 := 0$ and $t_i := \sum_1^i s_l$, then

$$\begin{aligned} G_n(I_j(\mathbf{x})) &= P\left(\bigcap_{i=1}^{k}\bigcap_{s=1}^{s_i}\left\{\frac{l_i-1}{2^j} < M_n^{(t_{i-1}+s)} \le \frac{l_i}{2^j}\right\}\right) \\ &= \binom{n}{\mathbf{s}}\prod_{i=1}^{k} F^{s_i}\left(\frac{l_i-1}{2^j}, \frac{l_i}{2^j}\right] \\ &= \binom{n}{\mathbf{s}} F^{n\otimes}(I_j(\mathbf{x})), \end{aligned}$$

and the result follows. $\square$

Next we calculate the joint distribution of the r largest order statistics, where $r < n$.

THEOREM 4.2. *Fix $r \in \{1, \ldots, n-1\}$. The distribution G_{nr} of $\mathbf{M}_n^{(r)} = (M_n^{(1)}, \ldots, M_n^{(r)})$ has $F^{r\otimes}$-density, for $F^{r\otimes}$-a.a. $\mathbf{x} \in \mathbb{R}^{r\downarrow}$,*

$$\begin{aligned} &\frac{dG_{nr}}{dF^{r\otimes}}(\mathbf{x}) = \\ &= \sum_{\mathbf{s}\in\mathbf{S}_r} 1\{\mathbf{x} \in E_{\mathbf{s}}^{(r)}\} \sum_{u=0}^{n-r} \frac{n!}{s_1!\cdots s_{k-1}!(s_k+u)!(n-r-u)!} F^u\{x_r\} F^{n-r-u}(-\infty, x_r) \end{aligned}$$

(recalling that $\mathbf{x} \in E_{\mathbf{s}}^{(r)}$ consists of components $y_1 > \cdots > y_k$ repeated $s_1, \ldots, s_k$ times respectively).

PROOF: We could integrate (4.2) but we choose to use the method of §2 again. Pick $\mathbf{x} \in \mathbb{R}^{r\downarrow}$, then $\mathbf{x} = \mathbf{x_y} \in E_{\mathbf{s}}^{(r)}$ for some $\mathbf{s} \in \mathbf{S}_r$, where now $\mathbf{s} = (s_1, \ldots, s_k)$ with $s_1 + \cdots + s_k = r$. Again, $I_j(\mathbf{x}) \ni \mathbf{x}$ will be the product of s_1 intervals $((l_1-1)2^{-j}, l_1 2^{-j}] \ni y_1$, and s_2 intervals $((l_2-1)2^{-j}, l_2 2^{-j}] \ni y_2$, etc. We decompose the event $\mathbf{M}_n^{(n)} \in I_j(\mathbf{x})$ according to how many further order statistics $M_n^{(r+1)}, \ldots, M_n^{(r+u)}$ also belong to the final interval $((l_k-1)2^{-j}, l_k 2^{-j}] \ni y_k$. For j so large that $2^{-j+1} < \min_i(y_i - y_{i-1})$,

$$\begin{aligned} &G_{nr}(I_j(\mathbf{x})) = \\ &\quad = P\left(\bigcap_{i=1}^{k}\bigcap_{s=1}^{s_i}\left\{\frac{l_i-1}{2^j} < M_n^{(t_{i-1}+s)} \le \frac{l_i}{2^j}\right\}\right) \\ &\quad = \sum_{u=0}^{n-r} \frac{n!}{s_1!\cdots s_{k-1}!(s_k+u)!(n-r-u)!}\left(\prod_{i=1}^{k-1} F^{s_i}\left(\frac{l_i-1}{2^j}, \frac{l_i}{2^j}\right]\right) \times \\ &\qquad\qquad \times F^{s_k+u}\left(\frac{l_k-1}{2^j}, \frac{l_k}{2^j}\right] F^{n-r-u}\left(-\infty, \frac{l_k-1}{2^j}\right] \end{aligned}$$

(interpret $\prod_{i=1}^{0}$ as 1)

$$= \sum_{u=0}^{n-r} \frac{n!}{s_1! \cdots s_{k-1}!(s_k+u)!(n-r-u)!} F^{r\otimes}(I_j(\mathbf{x})) F^u \left(\frac{l_k-1}{2^j}, \frac{l_k}{2^j}\right] \times$$
$$\times F^{n-r-u}\left(-\infty, \frac{l_k-1}{2^j}\right].$$

Divide this by $F^{r\otimes}(I_j(\mathbf{x}))$ and let $j \to \infty$; the claimed evaluation is the limit, as wanted. $\square$

5. Conditional Distributions.

Let $\mathcal{M}_{n-r}$ be the set of *finite* measures on $\mathcal{B}(\mathbb{R}^{n-r})$; thus every $\nu \in \mathcal{M}_{n-r}$ has $\nu(\mathbb{R}^{n-r}) < \infty$. By a *regular conditional probability distribution* (rcpd) for $\tilde{\mathbf{M}}_n^{(n-r)} = (M_n^{(r+1)}, \dots, M_n^{(n)})$ given $M_n^{(r)}$ we mean a Borel-measurable map $\mathbb{R} \to \mathcal{M}_{n-r} : x \mapsto p_x$ such that

$$P(\tilde{\mathbf{M}}_n^{(n-r)} \in B,\ M_n^{(r)} \in C) = \int_C p_x(B)\, P(M_n^{(r)} \in dx) \quad (B \in \mathcal{B}(\mathbb{R}^{n-r}),\ C \in \mathcal{B}(\mathbb{R})). \tag{5.1}$$

It follows from (5.1) that

(a) for any $h : \mathbb{R}^{n-r} \to \mathbb{R}$ for which $E|h(\tilde{\mathbf{M}}_n^{(n-r)})| < \infty$, $\int h\, dp_x$ as a function of x is a version of the *conditional expectation* $x \mapsto E\big(h(\tilde{\mathbf{M}}_n^{(n-r)}) \mid M_n^{(r)} = x\big)$.

In particular,

(b) for each fixed $B \in \mathcal{B}(\mathbb{R}^{n-r})$, $p_x(B)$ as a function of x is a version of the *conditional probability* $x \mapsto P(\tilde{\mathbf{M}}_n^{(n-r)} \in B \mid M_n^{(r)} = x)$.

The point of an rcpd is that more is true. Taking $B := \mathbb{R}^{n-r}$, (5.1) implies that the set $D := \{x : p_x(\mathbb{R}^{n-r}) \neq 1\}$ has the property $P(M_n^{(r)} \in D) = 0$. So D is F-null. Thus

(c) *for all x not in the F-null set D, p_x is a proper probability distribution on $\mathbb{R}^{n-r}$.*

We shall discuss the concept of rcpd further in Remarks 5.2–5.6 below. For the present we concentrate on proving the existence of the rcpd specified above.

Let $x_- := \inf\{x : F(-\infty, x] > 0\}$ and $x_+ := \sup\{x : F[x, \infty) > 0\}$, so that $-\infty \leq x_- \leq x_+ \leq \infty$. Let I be the open interval (x_-, x_+), together with the point x_- if $x_- > -\infty$ and $F\{x_-\} > 0$, and the point x_+ if $x_+ < \infty$ and $F\{x_+\} > 0$. Then F is concentrated on I (see Remark 2.6), though the support of F may be a proper subset of I. In fact I is the smallest *connected* closed set in $\mathbb{R}$ on which F is concentrated.

For fixed $x \in I \setminus \{x_-\}$, let $X_1(x)$, $X_2(x)$, $\dots$ be independent, each with the distribution of X given $X \leq x$. Similarly, let $X_1(x-)$, $X_2(x-)$, $\dots$ be independent, each with the distribution of X given $X < x$. If $F\{x_-\} > 0$, let all $X_i(x_-)$ and $X_i(x_- -)$ be degenerate at x_-. For every $x \in I$ and positive integer i, write the order statistics of $X_1(x)$, $\dots$, $X_i(x)$ as $M_i^{(1)}(x) \geq \cdots \geq M_i^{(i)}(x)$ and the order statistics of $X_1(x-)$, $\dots$, $X_i(x-)$ as $M_i^{(1)}(x-) \geq \cdots \geq M_i^{(i)}(x-)$.

THEOREM 5.1. *Fix $r \in \{1, \dots, n-1\}$. For $x \in \mathbb{R}$ and $B \in \mathcal{B}(\mathbb{R}^{n-r})$ let*

$$p_x(B) := \begin{cases} P\big((M_{n-r}^{(1)}(x), \dots, M_{n-r}^{(n-r)}(x)) \in B\big) & \text{if } x \in I \setminus A_F, \\ \frac{1}{P(M_n^{(r)}=x)} \sum_{i=0}^{n-r} \sum_{j=0}^{r-1} \frac{n!}{i!j!(n-i-j)!} F^i(-\infty, x) F^{n-i-j}\{x\} \times & \\ \qquad \times F^j(x,\infty) P\big((x, \dots, x, M_i^{(1)}(x-), \dots, M_i^{(i)}(x-)) \in B\big) & \\ & \text{if } x \in A_F, \\ 0 & \text{if } x \notin I. \end{cases} \tag{5.2}$$

Then p_x is an rcpd for $\tilde{\mathbf{M}}_n^{(n-r)} = (M_n^{(r+1)}, \dots, M_n^{(n)})$ given $M_n^{(r)} = x$.

PROOF OF THEOREM 5.1: Take $x \in I$. The $F^{(n-r+1)\otimes}$-density of $(M_n^{(r)}, \dots, M_n^{(n)})$ can be calculated similarly to that for $(M_n^{(1)}, \dots, M_n^{(r)})$ in Theorem 4.2, yielding

$$\begin{aligned} &P(M_n^{(r)} \in dx_1, \dots, M_n^{(n)} \in dx_{n-r+1}) = \\ &= \sum_{j=0}^{r-1} \frac{n!}{j!(r-1-j+s_1)! s_2! \cdots s_k!} \times \\ &\qquad \times F^j(x_1, \infty) F^{r-1-j}\{x_1\} F^{(n-r+1)\otimes}(dx_1, \dots, dx_{n-r+1}) \end{aligned} \tag{5.3}$$

for $(x_1, \dots, x_{n-r+1}) \in E_{\mathbf{s}}^{(n-r+1)}$ and $\mathbf{s} = (s_1, \dots, s_k) \in \mathbf{S}_{n-r+1}$. On the subset of $\mathbb{R}^{(n-r+1)\downarrow}$ in which $x_1 \notin A_F$ this reduces to

$$\frac{n!}{(r-1)!1!s_2! \cdots s_k!} F^{r-1}(x_1, \infty) F^{(n-r+1)\otimes}(dx_1, \dots, dx_{n-r+1}),$$

and only those $E_{\mathbf{s}}^{(n-r+1)}$ of the form $\mathbf{s} = (1, s_2, \dots, s_k)$ need be considered. By re-indexing we thus obtain, for $x \notin A_F$ and $x > x_1 \geq x_2 \geq \cdots \geq x_{n-r}$,

$$\begin{aligned} &P(M_n^{(r)} \in dx, M_n^{(r+1)} \in dx_1, \dots, M_n^{(n)} \in dx_{n-r}) = \\ &= r\binom{n}{r} F^{r-1}(x, \infty) F(dx) \binom{n-r}{\mathbf{s}} F^{(n-r)\otimes}(dx_1, \dots, dx_{n-r}), \end{aligned}$$

where also $(x_1, \dots, x_{n-r}) \in E_{\mathbf{s}}^{(n-r)}$ and $\mathbf{s} \in \mathbf{S}_{n-r}$. Since by (3.1) the F-density of $M_n^{(r)}$ is $f_n^{(r)}(x) = r\binom{n}{r} F^{r-1}(x,\infty) F^{n-r}(-\infty, x]$ at points $x \notin A_F$, the right-hand side is

$$\begin{aligned} &f_n^{(r)}(x) F(dx) \binom{n-r}{\mathbf{s}} \frac{F^{(n-r)\otimes}(dx_1, \dots, dx_{n-r})}{F^{n-r}(-\infty, x]} \\ &= f_n^{(r)}(x) F(dx) P\big(M_{n-r}^{(1)}(x) \in dx_1, \dots, M_{n-r}^{(n-r)}(x) \in dx_{n-r}\big). \end{aligned}$$

Thus for all $C \in \mathcal{B}(\mathbb{R})$ such that $C \cap A_F = \emptyset$, and all $B \in \mathcal{B}(\mathbb{R}^{n-r})$,

$$\begin{aligned} &P(M_n^{(r)} \in C,\ (M_n^{(r+1)}, \dots, M_n^{(n)}) \in B) = \\ &\quad = \int_C P\big((M_{n-r}^{(1)}(x), \dots, M_{n-r}^{(n-r)}(x)) \in B\big) f_n^{(r)}(x)\, F(dx). \end{aligned}$$

By (5.1), we thus have (5.2) in the case $x \in I \setminus A_F$.

To complete the proof of (5.2) we take any $x \in A_F$ and must prove that

$$\begin{aligned} &P\big(M_n^{(r)} = x, (M_n^{(r+1)}, \dots, M_n^{(n)}) \in B\big) = \\ &\quad = \sum_{i=0}^{n-r} \sum_{j=0}^{r-1} \frac{n!}{i!j!(n-i-j)!} F^i(-\infty, x) F^{n-i-j}\{x\} F^j(x, \infty) \times \\ &\qquad \times P\big((x, \dots, x, M_i^{(1)}(x-), \dots, M_i^{(i)}(x-)) \in B\big) \end{aligned}$$

for all $B \in \mathcal{B}(\mathbb{R}^{n-r})$. The left-hand side decomposes as

$$\sum_{i=0}^{n-r} P\big(x = M_n^{(r)} = \cdots = M_n^{(n-i)} > M_n^{(n-i+1)}, (x, \dots, x, M_n^{(n-i+1)}, \dots, M_n^{(n)}) \in B\big),$$

where a convention $M_n^{(n+1)} := -\infty$ is assumed for the $i = 0$ term in the sum. We prove that the i^{th} terms on the two sides are equal to each other. In the case $i = 0$ the assertion is that

$$\begin{aligned} &P(x = M_n^{(r)} = \cdots = M_n^{(n)}) 1\{x\mathbf{1}_{n-r} \in B\} = \\ &\quad = \sum_{j=0}^{r-1} \binom{n}{j} F^{n-j}\{x\} F^j(x, \infty) 1\{x\mathbf{1}_{n-r} \in B\}, \end{aligned}$$

which is correct. For the other cases we use the notation

$$B_{x,i} := \{(x_1, \dots, x_i) \in \mathbb{R}^i : (x, \dots, x, x_1, \dots, x_i) \in B\}$$

for sections of B, and our task then reduces to proving that

$$\begin{aligned} &P\big(x = M_n^{(r)} = \cdots = M_n^{(n-i)} > M_n^{(n-i+1)}, (M_n^{(n-i+1)}, \dots, M_n^{(n)}) \in B_{x,i}\big) = \\ &\quad = \sum_{j=0}^{r-1} \frac{n!}{i!j!(n-i-j)!} F^i(-\infty, x) F^{n-i-j}\{x\} F^j(x, \infty) \times \\ &\qquad \times P\big((M_i^{(1)}(x-), \dots, M_i^{(i)}(x-)) \in B_{x,i}\big). \end{aligned} \tag{5.4}$$

Take the particular case of (5.3) where $s_1 = n - r + 1 - i$. This corresponds to $x = x_1 = \cdots = x_{n-r+1-i} > x_{n-r+2-i}$, and (5.3) becomes

$$\begin{aligned} &P(x = M_n^{(r)} = \cdots = M_n^{(n-i)}, M_n^{(n-i+1)} \in dx_{n-r+2-i}, \dots, M_n^{(n)} \in dx_{n-r+1}) = \\ &\quad = \sum_{j=0}^{r-1} \frac{n!}{j!(n-i-j)!s_2! \cdots s_k!} F^j(x, \infty) F^{n-i-j}\{x\} F^{i\otimes}(dx_{n-r+2-i}, \dots, dx_{n-r+1}) \end{aligned}$$

for $(x, \ldots, x, x_{n-r+2-i}, \ldots, x_{n-r+1}) \in E_{\mathbf{s}}^{(n-r+1)}$ and $\mathbf{s} = (n-r+1-i, s_2, \ldots, s_k) \in \mathbf{S}_{n-r+1}$. In this, rename $x_{n-r+2-i}, \ldots, x_{n-r+1}$ as $x_1, \ldots, x_i$ respectively; also put $l := k-1$ and rename $s_2, \ldots, s_k$ as $s_1, \ldots, s_l$ respectively. Thus

$$\begin{aligned} &P(x = M_n^{(r)} = \cdots = M_n^{(n-i)}, M_n^{(n-i+1)} \in dx_1, \ldots, M_n^{(n)} \in dx_i) = \\ &\quad = \sum_{j=0}^{r-1} \frac{n!}{j!(n-i-j)!s_1! \cdots s_l!} F^j(x,\infty) F^{n-i-j}\{x\} F^{i\otimes}(dx_1, \ldots, dx_i) \end{aligned}$$

for $(x_1, \ldots, x_i) \in E_{\mathbf{s}}^{(i)}$ and $\mathbf{s} = (s_1, \ldots, s_l) \in \mathbf{S}_i$. The right-hand side can be rewritten as

$$\sum_{j=0}^{r-1} \frac{n!}{i!j!(n-i-j)!} F^i(-\infty,x) F^{n-i-j}\{x\} F^j(x,\infty) \binom{i}{\mathbf{s}} \frac{F^{i\otimes}(dx_1, \ldots, dx_i)}{F^i(-\infty,x)},$$

so we obtain

$$\begin{aligned} &P\big(x = M_n^{(r)} = \cdots = M_n^{(n-i)} > M_n^{(n-i+1)}, (M_n^{(n-i+1)}, \ldots, M_n^{(n)}) \in B_{x,i}\big) = \\ &\quad = \sum_{j=0}^{r-1} \frac{n!}{i!j!(n-i-j)!} F^i(-\infty,x) F^{n-i-j}\{x\} F^j(x,\infty) \times \\ &\qquad \times \sum_{\mathbf{s}\in\mathbf{S}_i} \binom{i}{\mathbf{s}} \frac{F^{i\otimes}\big(B_{x,i} \cap E_{\mathbf{s}}^{(i)} \cap (-\infty,x)^{i\otimes}\big)}{F^i(-\infty,x)}, \end{aligned}$$

which we recognize as (5.4) on employing Theorem 4.1 to evaluate the distribution of $\big(M_i^{(1)}(x-), \ldots, M_i^{(i)}(x-)\big)$. This establishes (5.2).

The measurability specified in our definition of an rcpd reduces to the requirement that

(d) *for each $B \in \mathcal{B}(\mathbb{R}^{n-r})$ the map $x \mapsto p_x(B)$ is Borel-measurable.*

Now, from (5.2) and Theorem 4.1,

$$p_x(B) = \sum_{\mathbf{s}\in\mathbf{S}_{n-r}} \binom{n-r}{\mathbf{s}} \frac{F^{(n-r)\otimes}\big(B \cap E_{\mathbf{s}}^{(n-r)} \cap (-\infty,x]^{(n-r)\otimes}\big)}{F^{n-r}(-\infty,x]}$$

when $x \in I \setminus A_F$. For fixed B this is a measurable function of $x \in I \setminus A_F$, and since any assignment of values of $p_x(B)$ for x in the countable set A_F preserves this measurability, we conclude that (d) is satisfied. $\square$

Remark 5.2. The technical reason why the measurability requirement on an rcpd reduces to (d) is that the topology on the set $\mathcal{M}_{n-r}$, the usual 'weak' topology, is the smallest that makes all the evaluation maps $B \mapsto \nu(B)$, for $\nu \in \mathcal{M}_{n-r}$, measurable. We already made use of (d) in stating property (c).

Remark 5.3. An rcpd is usually set up on the underlying probability space Ω (a 'regular conditional law') rather than on the value space as in our set-up. There are theorems ensuring existence of regular conditional laws on well-behaved spaces; see, e.g., [Dudley 1989, §10.2] or [Hoffmann-Jørgensen 1994, Chapter 10]. We do not need these results, as our approach is simply to exhibit a candidate rcpd, visibly having the requisite measurability property, and prove it has the remaining required properties.

Remark 5.4. 'Conditional probability distributions seem to have a bad reputation when it comes to rigorous treatment of conditioning' [Chang & Pollard 1997, Abstract]. One way to avoid using the conditioning formalism is to work always with integral expressions like (5.1), proving enough good properties of the integrand to make manipulation easy. This is the basis of the *disintegration* approach to conditioning, a good account of which is given in the Chang-Pollard paper from which the above quotation is taken. Our (5.1) can be regarded as a disintegration, which we perform on the value-space of the random objects concerned, in accordance with our general approach laid down in §1.

Remark 5.5. (5.2) looks technical but is not hard to use. Let's calculate for example a 'regular' conditional d.f. of $M_n^{(r+1)}$, given $M_n^{(r)}$, for $r \geq 1$. Fix $x \in A_F$ and let $y \geq x$. Take $B := (-\infty, y] \times \mathbb{R}^{n-r-1}$ in (5.2) and note that the value of the probability in the sum in (5.2) is then 1, for each $i \leq n-r$. But by (3.5)

$$\sum_{j=0}^{r-1} \sum_{i=0}^{n-r} \frac{n!}{i!j!(n-i-j)!} F^i(-\infty, x) F^{n-i-j}\{x\} F^j(x, \infty) = f_n^{(r)}(x) F\{x\},$$

while, by (3.2),

$$f_n^{(r)}(x) F\{x\} = P(M_n^{(r)} = x). \tag{5.5}$$

Thus we get from (5.2), when $y \geq x$,

$$P(M_n^{(r+1)} \leq y | M_n^{(r)} = x) = 1, \tag{5.6}$$

as indeed we should. When $y < x$, take B as before in (5.2). This time only the term with $i = n-r$ survives, and we get for the right hand side of (5.2)

$$\sum_{j=0}^{r-1} \frac{n!}{(n-r)!j!(r-j)!} \frac{F^{n-r}(-\infty, x) F^{r-j}\{x\} F^j(x, \infty) P(M_{n-r}^{(1)}(x-) \leq y)}{P(M_n^{(r)} = x)}$$

$$= \binom{n}{r} \frac{F^{n-r}(-\infty, x) P(M_{n-r}^{(1)}(x-) \leq y)(F^r[x, \infty) - F^r(x, \infty))}{P(M_n^{(r)} = x)}.$$

Substitute for the denominator using (5.5) and notice that, by (3.6),

$$f_r^{(r)}(x) = \frac{(1 - F^r(x, \infty)) - (1 - F^r[x, \infty))}{F\{x\}} = \frac{F^r[x, \infty) - F^r(x, \infty)}{F\{x\}},$$

whilst

$$P(M_{n-r}^{(1)}(x-) \leq y) = \frac{F^{n-r}(-\infty, y]}{F^{n-r}(-\infty, x)},$$

to obtain finally

$$P(M_n^{(r+1)} \le y | M_n^{(r)} = x) = \binom{n}{r} \frac{F^{n-r}(-\infty, y] f_r^{(r)}(x)}{f_n^{(r)}(x)} \tag{5.7}$$

when $y < x$. Notice that if we extend this formula to $x \in I \setminus A_F$, by (3.1) it reduces to

$$P(M_n^{(r+1)} \le y | M_n^{(r)} = x) = \frac{F^{n-r}(-\infty, y]}{F^{n-r}(-\infty, x)},$$

which agrees with the right hand side of (5.2) in this case.

Thus we can use (5.6) and (5.7) for *any* $x \in I$. They give

$$P(M_n^{(r+1)} \le y | M_n^{(r)} = x) = 1\{y \ge x\} + 1\{y < x\} \binom{n}{r} \frac{F^{n-r}(-\infty, y] f_r^{(r)}(x)}{f_n^{(r)}(x)}. \tag{5.8}$$

This is a *regular* conditional distribution function in that

- for each $y \in \mathbb{R}$ it is a version of the conditional probability $P(M_n^{(r+1)} \le y \mid M_n^{(r)} = x)$;
- for each $x \in I$ it is a (cumulative) distribution function in y;
- the map from x to $P(M_n^{(r+1)} \le y \mid M_n^{(r)} = x)$ is (Borel-)measurable for each $y \in \mathbb{R}$.

Remark 5.6. A variant of Theorem 5.1 in which the conditioning is on $\mathbf{M}_n^{(r)} = (M_n^{(1)}, \dots, M_n^{(r)})$ rather than just on $M_n^{(r)}$ can be given. Write $\mathbf{x}_r := (x_1, \dots, x_r)$ and set $\phi(\mathbf{x}_r) := \sum_{j=1}^r 1\{x_j = x_r\}$. Then the variant form of (5.2) is

$$p_{\mathbf{x}_r}(B) := \begin{cases} P\big((M_{n-r}^{(1)}(x_r), \dots, M_{n-r}^{(n-r)}(x_r)) \in B\big) & \text{if } x_r \in I \setminus A_F, \\ \frac{1}{P(M_n^{(r)} = x_r,\ \phi(\mathbf{M}_n^{(r)}) = j)} \sum_{i=0}^{n-r} \frac{n!}{i!(n-r+j-i)!(r-j)!} F^i(-\infty, x_r) F^{n-r+j-i}\{x_r\} \times & \\ \qquad \times F^{r-j}(x_r, \infty) P\big((x_r, \dots, x_r, M_i^{(1)}(x_r-), \dots, M_i^{(i)}(x_r-)) \in B\big) & \\ & \text{if } x_r \in A_F \text{ and } \phi(\mathbf{x}_r) = j, \\ 0 & \text{if } x_r \notin I. \end{cases} \tag{5.9}$$

We omit the proof that p is an rcpd for $\tilde{\mathbf{M}}_n^{(n-r)}$ given $\mathbf{M}_n^{(r)}$, which can be done by an extension of that of Theorem 5.1. It is easy to deduce Theorem 5.1, as well as Theorem 6.1 below, from (5.9).

Let $\Phi_n^{(r)} := \phi(\mathbf{M}_n^{(r)})$. Since the right-hand side of (5.9) depends on $\mathbf{x}_r$ only through $(x_r, \phi(\mathbf{x}_r))$, we obtain that $P(\tilde{\mathbf{M}}_n^{(n-r)} \in B \mid \mathbf{M}_n^{(r)}) = P(\tilde{\mathbf{M}}_n^{(n-r)} \in B \mid M_n^{(r)}, \Phi_n^{(r)})$ a.s. Since $\Phi_n^{(r+1)}$ is a function of $M_n^{(r+1)}$, $M_n^{(r)}$ and $\Phi_n^{(r)}$, indeed $\Phi_n^{(r+1)} = (\Phi_n^{(r)} + 1)1\{M_n^{(r+1)} = M_n^{(r)}\} + 1\{M_n^{(r+1)} \ne M_n^{(r)}\}$, it follows that $P(M_n^{(r+1)} \in B,\ \Phi_n^{(r+1)} = j \mid \mathbf{M}_n^{(r)}) = P(M_n^{(r+1)} \in B,\ \Phi_n^{(r+1)} = j \mid M_n^{(r)}, \Phi_n^{(r)})$ a.s., for all $B \in \mathcal{B}(\mathbb{R})$ and $j = 1, \dots, r+1$. We thus retrieve the conclusion of [Rüschendorf 1985] that the sequence $\big((M_n^{(r)}, \Phi_n^{(r)})\big)_{r=1,\dots,n}$ is Markov, no matter what F is. In the next section we investigate further Markov properties of the sequence of partial maxima.

6. Applications: Markov Properties.

The methods developed above give a new proof of the following known result, which is from [Arnold *et al.* 1984], where applications are given.

Recall that we defined x_- and x_+ in §4. In addition to the notation $\mathbf{x} = (x_1, \dots, x_n)$ and $\mathbf{M}_n^{(r)} = (M_n^{(1)}, \dots, M_n^{(r)})$ employed earlier, we now also write

$$\begin{aligned} \mathbf{x}_r &:= (x_1, \dots, x_r), \\ \tilde{\mathbf{x}}_{n-r} &:= (x_{r+1}, \dots, x_n), \\ \tilde{\mathbf{M}}_n^{(n-r)} &:= (M_n^{(r+1)}, \dots, M_n^{(n)}). \end{aligned}$$

THEOREM 6.1. *Fix $n \in \{3, 4, \dots\}$. The finite sequence $(M_n^{(r)})_{r=1,\dots,n}$ is Markov if and only if F is continuous on the open interval (x_-, x_+).*

PROOF: Suppose first F is continuous on (x_-, x_+). For non-triviality, assume $x_- < x_+$. Pick $r \in \{2, \dots, n\}$. We shall exhibit non-negative measurable functions $a(\tilde{\mathbf{x}}_{n-r}|x_r)$ and $b(\mathbf{x}_{r-1}|x_r)$ that factor the $F^{n\otimes}$-density of $\mathbf{M}_n^{(n)}$: setting

$$p(\mathbf{x}) := a(\tilde{\mathbf{x}}_{n-r}|x_r) b(\mathbf{x}_{r-1}|x_r) f_n^{(r)}(x_r) \qquad (x \in \mathbb{R}^n), \tag{6.1}$$

we shall prove that

$$\frac{dG_n}{dF^{n\otimes}}(\mathbf{x}) = p(\mathbf{x}) \tag{6.2}$$

for $F^{n\otimes}$-a.a. $\mathbf{x} \in \mathbb{R}^n$. This proves 'conditional independence of past and future given the present', which is well-known to imply the Markov property, in either time direction—see Remark 6.2, below.

For $x_- < x_r < x_+$ we have $F(x_r, \infty) > 0$ by definition of x_+ and can define

$$b(\mathbf{x}_{r-1}|x_r) := \left(\prod_{i=1}^{r-1} \frac{1\{x_i > x_r\}}{F(x_r, \infty)} \right) \sum_{\mathbf{s} \in \mathbf{S}_{r-1}} \binom{r-1}{\mathbf{s}} 1\{\mathbf{x}_{r-1} \in E_{\mathbf{s}}^{(r-1)}\} \qquad (\mathbf{x}_{r-1} \in \mathbb{R}^{r-1}). \tag{6.3}$$

If $F\{x_+\} > 0$ set

$$b(\mathbf{x}_{r-1}|x_+) := \prod_{i=1}^{r-1} \frac{1\{x_i = x_+\}}{F\{x_+\}} \qquad (\mathbf{x}_{r-1} \in \mathbb{R}^{r-1}).$$

If $F\{x_-\} > 0$ consider the probability distribution H_{r-1} of $\mathbf{M}_n^{(r-1)}$ given $M_n^{(r)} = x_-$. The conditioning event has positive probability so H_{r-1} is elementary to define. Set

$$b(\mathbf{x}_{r-1}|x_-) := \frac{dH_{r-1}}{dF^{(r-1)\otimes}}(\mathbf{x}_{r-1}) \qquad (\mathbf{x}_{r-1} \in \mathbb{R}^{r-1}).$$

Define $b(\mathbf{x}_{r-1}|x_r)$ to be identically 0 for all $x_r \in \mathbb{R}$ not dealt with so far.

The definition of $a(\tilde{\mathbf{x}}_{n-r}|x_r)$ is similar. For $x_- < x_r < x_+$ set

$$a(\tilde{\mathbf{x}}_{n-r}|x_r) := \left(\prod_{i=r+1}^{n} \frac{1\{x_i < x_r\}}{F(-\infty, x_r)}\right) \sum_{\mathbf{s}\in\mathbf{S}_{n-r}} \binom{n-r}{\mathbf{s}} 1\{\tilde{\mathbf{x}}_{n-r} \in E_{\mathbf{s}}^{(n-r)}\} \qquad (\tilde{\mathbf{x}}_{n-r} \in \mathbb{R}^{n-r}).$$

If $F\{x_-\} > 0$ set

$$a(\tilde{\mathbf{x}}_{n-r}|x_-) := \prod_{i=r+1}^{n} \frac{1\{x_i = x_-\}}{F\{x_-\}} \qquad (\tilde{\mathbf{x}}_{n-r} \in \mathbb{R}^{n-r}),$$

and if $F\{x_+\} > 0$ set

$$a(\tilde{\mathbf{x}}_{n-r}|x_+) := \frac{dJ_{n-r}}{dF^{(n-r)\otimes}}(\tilde{\mathbf{x}}_{n-r}) \qquad (\tilde{\mathbf{x}}_{n-r} \in \mathbb{R}^{n-r}),$$

where J_{n-r} is the probability distribution of $\tilde{\mathbf{M}}_n^{(n-r)}$ given $M_n^{(r)} = x_+$. Let $a(\tilde{\mathbf{x}}_{n-r}|x_r)$ be identically 0 for other x_r.

We shall prove (6.2) separately for the subset C of $\mathbb{R}^n$ on which $x_- < x_r < x_+$, the subset C_- on which $x_r = x_-$ and the subset C_+ on which $x_r = x_+$, the latter two needing to be considered only if $F\{x_-\} > 0$ or $F\{x_+\} > 0$ respectively.

Take C first. It needs $\mathbf{s} \in \mathbf{S}_n$ to be such that $\mathbf{x}$ has $x_+ > x_{i+1} > \cdots > x_{n-j} > x_-$ for some $i \in \{0, \ldots, r-1\}$ and $j \in \{0, \ldots, n-r\}$, the highest i components of $\mathbf{x}$ being x_+ and the lowest j being x_-. On this $E_{\mathbf{s}}^{(n)}$ the first defining expressions for $a(\cdot|\cdot)$ and $b(\cdot|\cdot)$ apply, and we obtain

$$p(\mathbf{x}) = \frac{1}{F^{n-r}(-\infty, x_r)} \frac{(n-r)!}{j!} \frac{1}{F^{r-1}(x_r, \infty)} \frac{(r-1)!}{i!} f_n^{(r)}(x_r).$$

Use of (3.1) brings this to $n!/(i!j!)$ which is the right-hand side of (4.2), so (6.2) is proved on all such sets $E_{\mathbf{s}}^{(n)}$. Thus (6.2) holds on C.

If $F\{x_-\} > 0$ take sets $A \in \mathcal{B}(\mathbb{R}^{n-r})$ and $B \in \mathcal{B}(\mathbb{R}^{r-1})$, then

$$\int_{B\times\{x_-\}\times A} p\,dF^{n\otimes} = \\ = \int_A a(\tilde{\mathbf{x}}_{n-r}|x_-)\,dF^{(n-r)\otimes}(\tilde{\mathbf{x}}_{n-r}) \int_B b(\mathbf{x}_{r-1}|x_-)\,dF^{(r-1)\otimes}(\mathbf{x}_{r-1})\, f_n^{(r)}(x_-)F\{x_-\}.$$

The A integral is just $1\{(x_-, \ldots, x_-) \in A\}$. The B integral is

$$\int_B \frac{dH_{r-1}}{dF^{(r-1)\otimes}}\,dF^{(r-1)\otimes} = \int_B dH_{r-1} = P(\mathbf{M}_n^{(r-1)} \in B \mid M_n^{(r)} = x_-).$$

The product $f_n^{(r)}(x_-)F\{x_-\}$ is, by (3.2), $P(M_n^{(r)}=x_-)>0$. Thus

$$\begin{aligned}\int_{B\times\{x_-\}\times A} p\,dF^{n\otimes} &= 1\{(x_-,\dots,x_-)\in A\}P(\mathbf{M}_n^{(r-1)}\in B,\ M_n^{(r)}=x_-)\\ &= P(\mathbf{M}_n^{(r-1)}\in B,\ M_n^{(r)}=x_-,\ \tilde{\mathbf{M}}_n^{(n-r)}\in A)\\ &= P(\mathbf{M}_n^{(n)}\in B\times\{x_-\}\times A)\\ &= \int_{B\times\{x_-\}\times A}\frac{dG_n}{dF^{n\otimes}}\,dF^{n\otimes}.\end{aligned}$$

Sets of the form $B\times\{x_-\}\times A$ constitute a generating semi-ring for the Borel sets that are subsets of C_-, so this proves (6.2) $F^{n\otimes}$-a.e. on C_-. The proof of (6.2) on C_+, for the case when $F\{x_+\}>0$, is similar and is omitted. Thus (6.2) is proved.

Now suppose $F\{z\}>0$ for some $z\in(x_-,x_+)$. Then easy calculations give

$$P(M_n^{(n)}=z\mid M_n^{(1)}=\cdots=M_n^{(n-1)}=z)=\frac{F\{z\}}{F\{z\}+nF(-\infty,z)},$$

$$P(M_n^{(n)}=z\mid M_n^{(1)}>M_n^{(2)}=\cdots=M_n^{(n-1)}=z)=\frac{F\{z\}}{F\{z\}+(n-1)F(-\infty,z)},$$

and since these are unequal the sequence is not Markov. $\square$

Remark 6.2. Assume F is continuous on (x_-,x_+), so that (6.2) holds. It is straightforward to check that $\int_{\mathbb{R}^{n-r}} a(\tilde{\mathbf{x}}_{n-r}|x_r)\,dF^{(n-r)\otimes}(\tilde{\mathbf{x}}_{n-r})$ and $\int_{\mathbb{R}^{r-1}} b(\mathbf{x}_{r-1}|x_r)\,dF^{(r-1)\otimes}(\mathbf{x}_{r-1})$ are both 1 for F-a.a. $x_r\in\mathbb{R}$. Then from (6.2), by integrating over $B\times\mathbb{R}^{n-r}$ with respect to $F^{n\otimes}$,

$$P(\mathbf{M}_n^{(r)}\in B)=\int_B b(\mathbf{x}_{r-1}|x_r)f_n^{(r)}(x_r)\,dF^{r\otimes}(\mathbf{x}_r)\qquad(B\in\mathcal{B}(\mathbb{R}^r)),$$

so that

$$P(\mathbf{M}_n^{(r)}\in d\mathbf{x}_r)=b(\mathbf{x}_{r-1}|x_r)f_n^{(r)}(x_r)\,dF^{r\otimes}(\mathbf{x}_r)$$

$F^{r\otimes}$-a.e. Again, from (6.2), by integrating over $B\times A$,

$$P(\mathbf{M}_n^{(r)}\in B,\ \tilde{\mathbf{M}}_n^{(n-r)}\in A)=\int_B P(\mathbf{M}_n^{(r)}\in d\mathbf{x}_r)\int_A a(\tilde{\mathbf{x}}_{n-r}|x_r)\,dF^{(n-r)\otimes}(\tilde{\mathbf{x}}_{n-r}),$$

and this identity shows that $\int_A a(\tilde{\mathbf{x}}_{n-r}|x_r)\,dF^{(n-r)\otimes}(\tilde{\mathbf{x}}_{n-r})$ is a version of $P(\tilde{\mathbf{M}}_n^{(n-r)}\in A\mid\mathbf{M}_n^{(r)}=\mathbf{x}_r)$. However as it does not depend on $x_1,\dots,x_{r-1}$ it is also a version of $P(\tilde{\mathbf{M}}_n^{(n-r)}\in A\mid M_n^{(r)}=x_r)$. So the two are the same for $F^{r\otimes}$-a.a. $\mathbf{x}_r$, which is the Markov property (in forward time).

Note that if we set

$$\tilde{q}_x(A):=\int_A a(\tilde{\mathbf{x}}_{n-r}|x)\,dF^{(n-r)\otimes}(\tilde{\mathbf{x}}_{n-r})\qquad(x\in\mathbb{R},\ A\in\mathcal{B}(\mathbb{R}^{n-r}))$$

then the above shows $\tilde{q}$ to be an rcpd (see §5 for the definition) for $\tilde{\mathbf{M}}_n^{(n-r)}$ given $M_n^{(r)}$. Similarly

$$q_x(B):=\int_B b(\mathbf{x}_{r-1}|x)\,dF^{(r-1)\otimes}(\mathbf{x}_{r-1})\qquad(x\in\mathbb{R},\ B\in\mathcal{B}(\mathbb{R}^{r-1}))$$

is an rcpd for $\mathbf{M}_n^{(r-1)}$ given $M_n^{(r)}$.

7. Applications: Trimmed Sums.

One form of trimmed sum of $X_1, \ldots, X_n$ is ${}^{(r)}S_n := \sum_{j=r+1}^n M_n^{(j)}$, that is, the sum $X_1 + \cdots + X_n$ with the r largest values trimmed off. Such sums are of great interest in Statistics and also as versions of random walks from which extreme terms have been removed. The following application of Theorem 5.1 is immediate. With the notations $X_j(x)$ and $X_j(x-)$ as set up for Theorem 5.1, put $S_n(x) := \sum_{j=1}^n X_j(x)$ and $S_n(x-) := \sum_{j=1}^n X_j(x-)$.

THEOREM 7.1. *Fix $r \in \{1, \ldots, n-1\}$. Set*

$$q_x(B) :=
\begin{cases}
P\big(S_{n-r}(x) \in B\big) & \text{if } x \in I \setminus A_F, \\
\frac{1}{P(M_n^{(r)}=x)} \sum_{j=0}^{r-1} \sum_{i=0}^{n-r} \frac{n!}{i!j!(n-i-j)!} F^i(-\infty, x) F^{n-i-j}\{x\} \times & \\
\qquad \times F^j(x,\infty) P\big(S_i(x-) + (n-r-i)x \in B\big) & \\
& \text{if } x \in A_F, \\
0 & \text{if } x \notin I,
\end{cases}$$

for all $x \in \mathbb{R}$ and $B \in \mathcal{B}(\mathbb{R})$. Then q is an rcpd (see §5) for ${}^{(r)}S_n$ given $M_n^{(r)}$.

Remark 7.2. Rather than the particular function ${}^{(r)}S_n$ of $M_n^{(1)}, \ldots, M_n^{(n)}$, one can write down formulae for the distributions of other functions of order statistics just as above. These are certainly of interest in Statistics, but we will not pursue them here.

Remark 7.3. The defining expression for $q_x(B)$ when $x \in A_F$ may be rewritten as

$$\sum_{i=0}^{n-r} P(x = M_n^{(r)} = \cdots = M_n^{(n-i)} > M_n^{(n-i+1)} \mid M_n^{(r)} = x) P\big(S_i(x-) + (n-r-i)x \in B\big) \tag{7.1}$$

This leads to the following bounds on the distribution of the trimmed sum. Alternative bounds for the tail probability will be given in §8.

COROLLARY 7.4. *Fix $r \in \{1, \ldots, n-1\}$. For all $B \in \mathcal{B}(\mathbb{R})$,*

$$\begin{aligned}
\int_I & P\big(S_{n-r}(x-) \in B\big) P(M_n^{(r+1)} < x \mid M_n^{(r)} = x) f_n^{(r)}(x)\, dF(x) \\
&\le P({}^{(r)}S_n \in B) \\
&\le P(M_n^{(r+1)} = M_n^{(r)}) \\
&\qquad + \int_I P\big(S_{n-r}(x-) \in B\big) P(M_n^{(r+1)} < x \mid M_n^{(r)} = x) f_n^{(r)}(x)\, dF(x).
\end{aligned} \tag{7.2}$$

PROOF: In the expression (7.1) for $q_x(B)$ replace $P\big(S_i(x-) + (n-r-i)x \in B\big)$ by 0 and by 1 in each term of the sum except the $i = n-r$ term. Hence, for $x \in A_F$,

$$\begin{aligned}
P(&M_n^{(r+1)} < x \mid M_n^{(r)} = x) P(S_{n-r}(x-) \in B) \\
&\le q_x(B) \\
&\le P(M_n^{(r+1)} = M_n^{(r)} \mid M_n^{(r)} = x) + P(M_n^{(r+1)} < x \mid M_n^{(r)} = x) P(S_{n-r}(x-) \in B).
\end{aligned} \tag{7.3}$$

(7.3) also holds in a trivial sense when $x \notin A_F$, both bounds being then equal to the central term. The result follows on integrating (7.3). □

Remark 7.5. [Kesten & Maller 1994] (see (4.42) of their paper) use a lower bound for $P({}^{(r)}S_n \in B)$ of the form

$$P({}^{(r)}S_n \in B) \geq \binom{n}{r} \int_{(-\infty,\infty)} P(\min_{1\leq i\leq r} X_i \in dx) F^{n-r}(-\infty, x) P(S_{n-r}(x-) \in B).$$

The reader may like to verify that this is the same as the lower bound given by the left hand side of (7.2).

8. A quantile construction and trimmed sums.

To handle joint distributions and trimmed sums a construction from a uniform sample has also commonly been used in the literature. This leads to quite different representations for joint distributions which are less explicit than ours. See [Reiss 1989, Theorem 1.5.6, p. 35], for example. We will compare the two approaches here from the point of view of the representations they give for distributions of trimmed sums.

Consider F for the present as a right-continuous distribution function: $F(x) := F(-\infty, x]$ for all x. The *quantile function* of F is its left-continuous pseudo-inverse $F^{\leftarrow} : [0,1] \to [x_-, x_+]$ given by

$$F^{\leftarrow}(p) := \inf\{x \in [x_-, x_+] : F(x) \geq p\}.$$

The main reason for the opposite continuity conventions is that then

$$F^{\leftarrow}(p) \leq x \quad \text{iff} \quad p \leq F(x) \tag{8.1}$$

for every $p \in [0,1]$ and $x \in [x_-, x_+]$.

Let $U_1, U_2, \ldots$ be independent Unif$(0,1)$ r.v.s, and set $X_i := F^{\leftarrow}(U_i)$, thus constructing independent F-distributed X_i. Let $U_n^{(1)} \geq \cdots \geq U_n^{(n)}$ be the order statistics of $U_1, \ldots, U_n$, so that $U_n^{(1)} > \cdots > U_n^{(n)}$ with probability 1, and by monotonicity $M_n^{(i)} := F^{\leftarrow}(U_n^{(i)})$ are the order statistics of $X_1, \ldots, X_n$.

The first thing to note is that this construction gives an immediate proof and explanation of (3.10). For by (8.1), $a < M_n^{(r)} \leq b$ if and only if $F(a) < U_n^{(r)} \leq F(b)$, and since $U_n^{(r)}$ has probability density

$$g_n^{(r)} := r\binom{n}{r}(1-u)^{r-1}u^{n-r} \qquad (0 < u < 1),$$

the latter event's probability is the right-hand side of (3.10).

Fix $r \in \{1, \ldots, n-1\}$. With a view towards trimmed sums, we now calculate the conditional distribution of $\tilde{\mathbf{M}}_n^{(n-r)}$ given $U_n^{(r)}$. Let $V_1, V_2, \ldots$ be independent Unif$(0,1)$ r.v.s, with the sequences (U_i), (V_i) independent of each other, and as usual let $V_n^{(1)} \geq \cdots \geq V_n^{(n)}$ be the order statistics of $V_1, \ldots, V_n$. It is easy to see that given $U_n^{(r)} = u$, the $n - r$ members of the sample $(U_1, \dots, U_n)$ that are less than u are conditionally distributed as $uV_1, \dots, uV_{n-r}$. Hence the conditional distribution of $U_n^{(r+1)}, \dots, U_n^{(n)}$ given $U_n^{(r)} = u$ is the joint distribution of $uV_{n-r}^{(1)}, \dots, uV_{n-r}^{(n-r)}$. Hence, further, the conditional distribution of $M_n^{(r+1)}, \dots, M_n^{(n)}$ given $U_n^{(r)} = u$ is the joint distribution of $F^{\leftarrow}(uV_{n-r}^{(1)}), \dots, F^{\leftarrow}(uV_{n-r}^{(n-r)})$. That is, if we set

$$p_u(B) := P\big((F^{\leftarrow}(uV_{n-r}^{(1)}), \dots, F^{\leftarrow}(uV_{n-r}^{(n-r)})) \in B\big) \qquad (u \in (0,1),\ B \in \mathcal{B}(\mathbb{R}^{n-r})), \tag{8.2}$$

then p is an rcpd for $\tilde{\mathbf{M}}_n^{(n-r)} = (M_n^{(r+1)}, \dots, M_n^{(n)})$ given $U_n^{(r)}$.

Specializing the above, if we set

$$\begin{aligned} q_u(B) &:= P\left(\sum_{j=1}^{n-r} F^{\leftarrow}(uV_{n-r}^{(j)}) \in B\right) \\ &= P\left(\sum_{j=1}^{n-r} F^{\leftarrow}(uV_j) \in B\right) \qquad (u \in (0,1),\ B \in \mathcal{B}(\mathbb{R})), \end{aligned} \tag{8.3}$$

then q is an rcpd for ${}^{(r)}S_n$ given $U_n^{(r)}$. Hence we obtain the result of [Mori 1984, Lemma 1]:

$$P({}^{(r)}S_n \in B) = \int_0^1 P\left(\sum_{j=1}^{n-r} F^{\leftarrow}(uV_j) \in B\right) g_n^{(r)}(u)\, du \qquad (B \in \mathcal{B}(\mathbb{R})).$$

These considerations lead for instance to stochastic bounds on the trimmed sums, alternatives to those of Corollary 7.4. Conditional probabilities in the following are to be interpreted as values of the rcpd q of (8.3).

Proposition 8.1. *Given $U_n^{(r)} = u \in (0,1)$, the trimmed sum ${}^{(r)}S_n$ is stochastically bounded below and above by $S_{n-r}(x-)$ and $S_{n-r}(x)$ respectively, where $x := F^{\leftarrow}(u)$. Hence*

$$P(S_{n-r}(x-) > s) \le P({}^{(r)}S_n > s \mid U_n^{(r)} = u) \le P(S_{n-r}(x) > s) \qquad (s \in \mathbb{R}), \tag{8.4}$$

and, with $g_n^{(r)}$ the probability density of $U_n^{(r)}$,

$$\begin{aligned} &\int_0^1 P\big(S_{n-r}(F^{\leftarrow}(u)-) > s\big) g_n^{(r)}(u)\, du \\ &\qquad \le P({}^{(r)}S_n > s) \le \int_0^1 P\big(S_{n-r}(F^{\leftarrow}(u)) > s\big) g_n^{(r)}(u)\, du \qquad (s \in \mathbb{R}). \end{aligned} \tag{8.5}$$

Proof: (8.3) says that the distribution $q_u(\cdot)$ is that of $\sum_{j=1}^{n-r} F^{\leftarrow}(uV_j)$. Now $F(x-) < u \le F(x)$ where $x := F^{\leftarrow}(u)$, so because $F^{\leftarrow}$ is non-decreasing,

$$\sum_{j=1}^{n-r} F^{\leftarrow}(F(x-)V_j) \le \sum_{j=1}^{n-r} F^{\leftarrow}(uV_j) \le \sum_{j=1}^{n-r} F^{\leftarrow}(F(x)V_j).$$

Since the r.v.s $F^{\leftarrow}(F(x-)V_j)$ are distributed as the $X_j(x-)$ and the $F^{\leftarrow}(F(x)V_j)$ as the $X_j(x)$, the left of this inequality is distributed as $S_{n-r}(x-)$ and the right as $S_{n-r}(x)$, and we obtain (8.4).

The final line follows by integrating. □

One should note that the left and right bounds in (8.4) coincide when x is a continuity point of F. When F is continuous everywhere the bounds in (7.2) or (8.5) coincide, and one has (cf. [Maller & Resnick 1984, Lemma 1.1])

$$P({}^{(r)}S_n \in B) = \int_{-\infty}^{\infty} P(S_{n-r}(x) \in B) f_n^{(r)}(x)\, dF(x) \qquad (B \in \mathcal{B}(\mathbb{R})).$$

It is tempting for general F to seek bounds that are closer in form to this formula than those of (8.5) are. The bounds

$$\begin{aligned} &\int_I P(S_{n-r}(x-) > s) f_n^{(r)}(x)\, dF(x) \\ &\qquad \le P({}^{(r)}S_n > s) \le \int_I P(S_{n-r}(x) > s) f_n^{(r)}(x)\, dF(x) \qquad (s \in \mathbb{R}) \end{aligned}$$

are valid, as may be proved from (8.4) after integrating it with respect to $P(U_n^{(r)} \in du \mid M_n^{(r)} = x)$, and they may be combined with the bounds (3.13) on $f_n^{(r)}$ to get other inequalities for the distribution of the trimmed sum. It is not clear how they compare with those obtained above because of the extra complication of the F-density in the F-discontinuous case.

References

Arnold, B. C. and N. Balakrishnan, *Relations, Bounds and Approximations for Order Statistics*, Springer-Verlag, New York, 1989. Lecture Notes in Statistics **53**.

Arnold, B. C., A. Becker, U. Gather and H. Zahedi, *On the Markov property of order statistics*, Journal of Statistical Planning and Inference **9** (1984), 147–154.

Chang, J. T. and Pollard, D., *Conditioning as disintegration*, Statistica Neerlandica (1997) (to appear).

Dudley, R. M., *Real Analysis and Probability*, Wadsworth & Brooks/Cole, Pacific Grove, Ca., 1989.

Hoffmann-Jørgensen, J., *Probability with a View toward Statistics, vol. I, II*, Chapman & Hall, New York, 1994.

Kesten, H. and R. A. Maller, *Ratios of trimmed sums and order statistics*, Annals of Probability **20** (1992), 1805–1842.

Kesten, H. and R. A. Maller, *Infinite limits and infinite limit points of random walks and trimmed sums*, Annals of Probability **22** (1994), 1473–1513.

Khatri, C. G., *Distributions of order statistics for discrete case*, Annals of the Institute of Statistical Mathematics **14** (1962), 167–171.

Maller, R. A. and S. Resnick, *Limiting behaviour of sums and the term of maximum modulus.*, Proc. Lond. Math. Soc. **49** (1984), 385–422.

Mori, T., *On the limit distributions of lightly trimmed sums*, Mathematical Proceedings of the Cambridge Philosophical Society **96** (1984), 507–516.

Reiss, R.-D., *Approximate Distributions of Order Statistics*, Springer-Verlag, New York, 1989.

Rüschendorf, L., *Two remarks on order statistics*, Journal of Statistical Planning and Inference **11** (1985), 71–74.

Zaanen, A. C., *An Introduction to the Theory of Integration*, North-Holland, Amsterdam, 1958.

C. M. Goldie, School of Mathematical Sciences, University of Sussex, Brighton BN1 9QH, UK
R. A. Maller, Department of Mathematics, University of Western Australia, Nedlands, W. A. 6907, Australia