Ann Inst Stat Math (2015) 67:515–540
DOI 10.1007/s10463-014-0466-9

Spacings around an order statistic

H. N. Nagaraja · Karthik Bharath · Fangyuan Zhang

Received: 30 July 2013 / Revised: 31 January 2014 / Published online: 26 April 2014

Abstract We determine the joint limiting distribution of adjacent spacings around a central, intermediate, or an extreme order statistic $X_{k:n}$ of a random sample of size n from a continuous distribution F. For central and intermediate cases, normalized spacings in the left and right neighborhoods are asymptotically i.i.d. exponential random variables. The associated independent Poisson arrival processes are independent of $X_{k:n}$. For an extreme $X_{k:n}$, the asymptotic independence property of spacings fails for F in the domain of attraction of Fréchet and Weibull ($\alpha \neq 1$) distributions. This work also provides additional insight into the limiting distribution for the number of observations around $X_{k:n}$ for all three cases.

Keywords Spacings · Uniform distribution · Central order statistics · Intermediate order statistics · Extremes · Poisson process

1 Introduction

Let $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ be order statistics of a random sample $X_1, \ldots, X_n$ from a continuous cdf F. For $1 \leq k \leq n$, we examine the clustering of data around the order statistic $X_{k:n}$. This is done by an investigation into the limiting properties of the right

H. N. Nagaraja (✉)
Division of Biostatistics, College of Public Health, The Ohio State University, 1841 Neil Avenue, Columbus, OH 43210, USA
e-mail: nagaraja.1@osu.edu

K. Bharath · F. Zhang
Department of Statistics, The Ohio State University, 1958 Neil Avenue, Columbus, OH 43210, USA
e-mail: karthikbharath@gmail.com

F. Zhang
e-mail: zhang.1243@buckeyemail.osu.edu

and left neighborhoods formed by the adjacent spacings $(X_{k+1:n} - X_{k:n}, \ldots, X_{k+r:n} - X_{k+r-1:n})$ and $(X_{k:n} - X_{k-1:n}, \ldots, X_{k-s+1:n} - X_{k-s:n})$ for fixed r and s. We let $n \to \infty$ and consider three scenarios: (i) Central case where $k/n \to p, 0 < p < 1$; (ii) Intermediate case where $k, n-k \to \infty$ and $k/n \to 0$ or 1; (iii) Extreme case where k or $n-k$ is held fixed. In the first two cases we show that, under some mild assumptions, these $(r+s)$ spacings appropriately scaled with a common scale parameter converge weakly to a set of i.i.d. standard exponential random variables (rvs). In the extreme case, this conclusion holds only when F is in the domain of attraction of the Gumbel cdf G_3, or the Weibull type cdf $G_{2;\alpha}$ with $\alpha = 1$. A direct and useful consequence of such a result is that order statistics around a selected one arrive as in a homogeneous Poisson process.

Neighborhoods around a selected order statistic have been investigated by several authors in recent years. Almost all these results, starting with $X_{n:n}$, have concentrated on the distribution of counts around it. We refer to a few, relevant to our results, from an exhaustive list: (Balakrishnan and Stepanov 2005; Dembińska et al. 2007; Pakes and Steutel 1997; Pakes 2009; Dembińska and Balakrishnan 2010). These authors typically consider neighborhoods of the form $(X_{k:n} - d, X_{k:n})$ or $(X_{k:n}, X_{k:n} + d)$ where the lengths of the intervals may or may not depend on n; in some papers, the d's are induced by the quantile function F^{-1} or are chosen to be random. While these approaches are beneficial from a technical perspective, it is more natural and practical to consider neighborhoods that are in the scale of the data collected. This is our motivation for considering the joint distribution of adjacent spacings. Our approach allows us to characterize the process governing the distribution of counts and provides additional insight into the asymptotic properties of the counts of cluster sizes around a specified order statistic.

Section 2 contains preliminaries that explore the properties of uniform and exponential order statistics; it introduces the von Mises conditions and the associated extreme value distributions. Section 3 is concerned with the joint distribution of a central order statistic and spacings adjacent to it on its right and left neighborhoods. The Poisson arrival process of adjacent order statistics is established there. Assuming von Mises conditions, Sect. 4 reaches a similar conclusion for the neighborhood of an intermediate order statistic. Section 5 displays the distributional structure of the extreme spacings assuming that F is in the domain of attraction of an extreme value distribution. Section 6 applies our results and describes the limiting distribution of the counts of observations around an order statistic. Section 7 discusses further applications of our results and contains concluding remarks.

Let $f(x)$ denote the pdf and $F^{-1}(p), 0 \le p \le 1$, be the quantile function associated with $F(x)$, where $F^{-1}(p) = \inf\{x : F(x) \ge p\}$ for $0 < p \le 1$, $F^{-1}(0) = \sup\{x : F(x) = 0\}$. We interchangeably use x_p and $F^{-1}(p)$ as the pth quantile. It is well known that if F is differentiable at x_p with finite and positive pdf $f(x_p)$, F^{-1} is differentiable at p with derivative $1/f(x_p)$. Standard uniform and exponential rvs are, respectively, denoted by U and Z. An exponential rv with rate parameter λ will be denoted by Exp(λ), and Poi(λ) represents a Poisson rv with mean parameter λ. The sum of r i.i.d. standard exponentials is a Gamma rv, to be denoted as Gam(r). A Weibull rv with shape parameter δ will be denoted by Wei(δ). Further, a standard

normal rv will be denoted by $N(0, 1)$ and its pdf by $\phi(\cdot)$. The Z_i's and Z_i^*'s are i.i.d. Exp(1) rvs. The symbol $\sim$ indicates asymptotic equivalence.

The rv $U_{i:n}$, $1 \le i \le n$, is the ith order statistic from a random sample of size n from a standard uniform population. The distributional equivalence, $X_{i:n} \stackrel{d}{=} F^{-1}(U_{i:n})$, for *any* collection of order statistics from an arbitrary cdf F is helpful in our investigations.

2 Preliminaries

2.1 Spacings near a uniform order statistic

The key to our approach is the following well-known exchangeable property of the uniform order statistics. Let $U_{0:n} = 0$ and $U_{n+1:n} = 1$, and define the *uniform spacing*

$$\Delta_{i,n} = U_{i+1:n} - U_{i:n}, 0 \le i \le n. \tag{1}$$

Then, it is well known that the $\Delta_{i,n}$'s are exchangeable and for any fixed r, and for constants $v_i \ge 0$, $i = 1, \ldots r$ with $r \le n$, and $\sum_{i=1}^{r} v_i \le 1$, the joint survival function of $\Delta_{1,n}, \ldots, \Delta_{r,n}$ (and hence any collection of $r\Delta_{i,n}$'s) is given by (see, e.g., David and Nagaraja 2003, p. 135)

$$P(\Delta_{1,n} > v_1, \ldots, \Delta_{r,n} > v_r) = \left(1 - \sum_{i=1}^{r} v_i\right)^{n-1}.$$

This means

$$P(n\Delta_{1,n} > v_1, \ldots, n\Delta_{r,n} > v_r) \to \prod_{i=1}^{r} \{e^{-v_i}\}, v_1, \ldots, v_r > 0. \tag{2}$$

That is, $n\Delta_{i,n}$ forms an i.i.d. Exp(1) sequence Z_i as $n \to \infty$. The convergence is fast; Problem P.5.19 of Reiss (1989, p. 201) notes that there exists a constant C such that for every positive integer n and $r \le n$,

$$\sup_{\mathcal{B} \in \mathbb{B}} |P\{(n\Delta_{1,n}, \ldots, n\Delta_{r,n}) \in \mathcal{B}\} - P\{(Z_1, \ldots, Z_r) \in \mathcal{B}\}| \le C \cdot (r/n),$$

where $\mathbb{B}$ denotes the family of all Borel sets. We record the implications of (2) and the exchangeability of the $\Delta_{i,n}$'s as a lemma given below; it uses the fact that the inter-arrival times being i.i.d. Exp(λ) rvs is a defining property of a homogeneous Poisson process with rate λ.

Lemma 1 *Let $U_{i:n}$ denote the ith order statistic from a random sample of size n from a standard uniform distribution, and assume $n \to \infty$. Then, for any k such that $n - k \to \infty$,*

$$(n(U_{k+1:n} - U_{k:n}), \ldots n(U_{k+r:n} - U_{k+r-1:n})) \stackrel{d}{\to} (Z_1, \ldots, Z_r),$$

for any fixed r, and for any $k \to \infty$

$$(n(U_{k:n} - U_{k-1:n}), \ldots n(U_{k-s+1:n} - U_{k-s:n})) \xrightarrow{d} (Z_1^*, \ldots, Z_s^*),$$

for any fixed s, where the Z_i*'s and* Z_i^**'s are all mutually independent Exp*(1) *rvs. That is, inter-arrival times of successive order statistics in the right and left neighborhoods of kth uniform order statistic, upon scaling by n, produce asymptotically independent homogeneous Poisson processes if n, k, and* $n - k$ *approach infinity. If k [resp.* $n - k$*] is bounded, the right [resp. left] neighborhood produces a Poisson process in the limit.*

2.2 Spacings near an exponential order statistic

When F is standard exponential, it is well known that

$$X_{i:n} \stackrel{d}{=} \frac{Z_1}{n} + \cdots + \frac{Z_i}{n-i+1}, i = 1, \ldots, n, \tag{3}$$

where the Z_i's are i.i.d. Exp(1) rvs. From this representation, it follows that $(n - i + 1)(X_{i:n} - X_{i-1:n})$ are i.i.d. Exp(1) rvs. So, if $k/n \to p, 0 \le p < 1, n(1-p)(X_{k+j:n} - X_{k+j-1:n}), j = 1, \ldots, r$ will be asymptotically i.i.d Exp(1) rvs.

When $(n - k)$ is bounded, $(n - k - j + 1)(X_{k+j:n} - X_{k+j-1:n}), j = 1, \ldots, r$ for $r \le n - k$ turn out to be i.i.d. Exp(1) rvs. Only in this scenario, we need finite and distinct scaling constants for the spacings to transform them into i.i.d. exponential rvs for any n, and hence asymptotically as well.

2.3 Extremes and von Mises conditions

Suppose there exist sequences of constants a_n and $b_n > 0$ such that $P\{(X_{n:n} - a_n)/b_n \le x\}$ converges to a nondegenerate cdf $G(x)$ corresponding to a rv W. Then, we say F is in the domain of maximal attraction to G and we write $F \in \mathcal{D}(G)$. Then, it is known that G is necessarily of one of the three types given below.

$$\begin{aligned}
\text{(Fréchet) } G_{1;\alpha}(x) &= \begin{cases} 0 & x \le 0, \alpha > 0, \\ \exp(-x^{-\alpha}) & x > 0; \end{cases} \\
\text{(Weibull) } G_{2;\alpha}(x) &= \begin{cases} \exp[-(-x)^{\alpha}] & x \le 0, \alpha > 0, \\ 1 & x > 0; \end{cases} \\
\text{(Gumbel) } G_3(x) &= \exp(-\mathrm{e}^{-x}), \quad -\infty < x < \infty.
\end{aligned} \tag{4}$$

The following are necessary and sufficient conditions on the right tail of F in order that $F \in \mathcal{D}(G)$. The first two are due to Gnedenko (1943) and the last one is due to de Haan (1970).

(a) $F \in \mathcal{D}(G_{1;\alpha})$ iff for all $t > 0$,

$$\lim_{x\to\infty} \frac{1 - F(tx)}{1 - F(x)} = t^{-\alpha}.$$

(b) $F \in \mathcal{D}(G_{2;\alpha})$ iff $x_1 (= F^{-1}(1))$ is finite, and the following condition holds for every $t > 0$:

$$\lim_{x\to 0+} \frac{1 - F(x_1 - tx)}{1 - F(x_1 - x)} = t^{\alpha}.$$

(c) $F \in \mathcal{D}(G_3)$ iff the following hold: $E(X|X > c)$ is finite for some c, and for all real t,

$$\lim_{x\to x_1} \frac{1 - F(x + tm(x))}{1 - F(x)} = \mathrm{e}^{-t},$$

where $m(x) = E(X - x|X > x)$.

Our approach for the intermediate case assumes the following sufficient conditions that are applicable to absolutely continuous cdf's. The first two are due to von Mises (1936), and the last one is due to Falk (1989) and is weaker than the corresponding von Mises condition that assumes differentiability of the pdf f (see, e. g., David and Nagaraja 2003, p. 300).

(a) $F \in \mathcal{D}(G_{1;\alpha})$ if $f(x) > 0$ for all large x and for some $\alpha > 0$,

$$\lim_{x\to\infty} \frac{xf(x)}{1 - F(x)} = \alpha. \tag{5}$$

(b) $F \in \mathcal{D}(G_{2;\alpha})$ if $x_1 < \infty$ and for some $\alpha > 0$,

$$\lim_{x\to x_1-} \frac{(x_1 - x)f(x)}{1 - F(x)} = \alpha. \tag{6}$$

(c) $F \in \mathcal{D}(G_3)$ if $f(x) > 0$ for all x in (c, x_1) and $E(X|X > c)$ is finite for some c, and

$$\lim_{x\to x_1-} \frac{f(x)m(x)}{1 - F(x)} = 1, \tag{7}$$

where $m(x) = E(X - x|X > x)$.

The family of limiting distribution for normalized $X_{1:n}$ corresponds to that of $-W$ where W has one of the above three types of cdfs; parallel necessary and sufficient, and sufficient conditions exist that impose conditions on the left tail of F.

3 Spacings around a central order statistic

3.1 Joint distribution of spacings

For $0 < p < 1$, $X_{k:n}$ is a central order statistic if $\frac{k}{n} \to p$. For such an $X_{k:n}$, Smirnov (1952; Theorem 3, p. 12) has shown (as pointed out by a reviewer) that

$$X_{k:n} \stackrel{a.s.}{\to} x_p \tag{8}$$

if the condition $F(x) = p$ has a unique solution x_p. Since for any fixed j,

$$n(X_{k+j+1:n} - X_{k+j:n}) \stackrel{d}{=} \frac{F^{-1}(U_{k+j+1:n}) - F^{-1}(U_{k+j:n})}{(U_{k+j+1:n} - U_{k+j:n})} n(U_{k+j+1:n} - U_{k+j:n}), \tag{9}$$

the limiting joint distribution of the spacings from an arbitrary cdf F can be linked to that of a collection of i.i.d. standard uniform rvs provided the first factor on the right in (9) above converges in probability to a nonzero constant.

From (8), it follows that $\Delta_{k+j,n} = U_{k+j+1:n} - U_{k+j:n}$ (defined in (1)) almost surely converges to 0. The first factor on the right in (9),

$$\frac{F^{-1}(U_{k+j:n} + \Delta_{k+j,n}) - F^{-1}(U_{k+j:n})}{\Delta_{k+j,n}} \stackrel{a.s.}{\to} \frac{1}{f(x_p)}, \tag{10}$$

if the following condition holds:

$$f \text{ is positive, finite and continuous at } x_p. \tag{11}$$

This conclusion follows from the definition of the derivative of F^{-1} and its assumed continuity at p.

Upon using (10), (9), Slutsky's Theorem, and Lemma 1, we conclude that jointly

$$nf(x_p)(X_{k+j+1:n} - X_{k+j:n}) \stackrel{d}{\to} Z_{j+1},\ j = -s, \ldots 0, \ldots, r-1,$$

where the Z_j's are i.i.d. Exp(1) rvs if (11) holds.

We can weaken the continuity assumption for f in (11) with the following condition:

$$\lim_{(p_1,h)\to(p,0^+)} \frac{F^{-1}(p_1 + h) - F^{-1}(p_1)}{F^{-1}(p + h) - F^{-1}(p)} = 1, \tag{12}$$

where $0 < p < 1$. This assumption is similar to (17) in Dembińska et al. (2007) (given as (38) in Sect. 6 later). The condition (11) implies that (12) holds since the latter is satisfied upon dividing the numerator and denominator by h and taking the double limit; the converse is not true.

On the other hand, we can weaken the requirement for a finite nonzero $f(x_p)$ by modifying a condition on F used by Chanda (1975) [see also Ghosh and Sukhatme 1981]. We assume that

$$\lim_{h\to 0} \frac{|F^{-1}(p+h) - F^{-1}(p)|}{|h|^{\theta}} = M(p,\theta) \in (0,\infty), \tag{13}$$

for some $\theta > 0$. If f is indeed finite and nonzero at x_p, then the above condition is satisfied with $\theta = 1$. Whenever $f(x_p)$ is finite and positive or (13) holds, there is a unique solution to $F(x) = p$ and (8) holds.

Based on the above discussion, we can now formally state the result for the central case.

Theorem 1 *Let $k/n \to p \in (0,1)$, and r and s be fixed positive integers.*

(a) *If condition (11) holds, or if (12) holds and $f(x_p)$ is finite and positive,*

$$\{nf(x_p)(X_{k+j:n} - X_{k+j-1:n}),\ j = -(s-1), \ldots, 0, \ldots, r\}$$
$$\xrightarrow{d} \{Z_s^*, \ldots, Z_1^*, Z_1, \ldots, Z_r\}$$

where the Z's are i.i.d. Exp(1) rvs. Thus, the two counting processes defined by setting the jth event to occur, respectively, at times

$$nf(x_p) \sum_{i=k}^{k+j-1} (X_{i+1:n} - X_{i:n}) \quad \text{and} \quad nf(x_p) \sum_{i=k-1}^{k-j} (X_{i+1:n} - X_{i:n})$$

converge weakly to two independent homogeneous Poisson processes with unit intensity.

(b) *Assume (12) and (13) hold. Then,*

$$\{n^{\theta}(X_{k+j:n} - X_{k+j-1:n})/M(p,\theta),\ j = -(s-1), \ldots, 0, \ldots, r\}$$
$$\xrightarrow{d} \{(Z_s^*)^{\theta}, \ldots, (Z_1^*)^{\theta}, (Z_1)^{\theta}, \ldots, (Z_r)^{\theta}\}. \tag{14}$$

That is, the counting processes defined by setting the jth event of the process to occur at times

$$n^{\theta} \sum_{i=k}^{k+j-1} (X_{i+1:n} - X_{i:n})/M(p,\theta) \text{ and } n^{\theta} \sum_{i=k-1}^{k-j} (X_{i+1:n} - X_{i:n})/M(p,\theta)$$

converge to i.i.d. renewal processes with Wei$(1/\theta)$ renewal distribution. They reduce to homogeneous Poisson processes with unit intensity only when $\theta = 1$ and $f(x_p)$ is finite and positive.

Proof To prove part (a), we need to show that (10) holds whenever (11) holds, or if (12) holds and $f(x_p)$ is finite and positive. Then, we would use (10), (9), Slutsky's

Theorem, and Lemma 1. We have shown earlier that (10) holds whenever (11) is satisfied.

If (12) holds and $f(x_p)$ is finite and positive, the left side expression in (10) can be written as

$$\frac{F^{-1}(U_{k+j:n}+\Delta_{k+j,n})-F^{-1}(U_{k+j:n})}{F^{-1}(p+\Delta_{k+j,n})-F^{-1}(p)}\cdot\frac{F^{-1}(p+\Delta_{k+j,n})-F^{-1}(p)}{\Delta_{k+j,n}},$$

where the first factor converges to 1 and the second factor converges to $1/f(x_p)$, both almost surely. Thus, (10) is established.

For (b), the idea is similar. We note that $n^{\theta}(F^{-1}(U_{k+j:n}+\Delta_{k+j,n})-F^{-1}(U_{k+j:n}))$ can be written as

$$\frac{F^{-1}(U_{k+j:n}+\Delta_{k+j,n})-F^{-1}(U_{k+j:n})}{\Delta_{k+j,n}^{\theta}}(n\Delta_{k+j,n})^{\theta}.$$

Assumption (13) coupled with (12) ensures that the first factor above converges almost surely to $M(p,\theta)$. Since $P(n\Delta_{k+j,n} > w) \to \exp\{-w\}$ for all $w > 0$, $P((n\Delta_{k+j,n})^{\theta} > w) \to \exp\{-w^{1/\theta}\}$; that is, $(n\Delta_{k+j,n})^{\theta} \xrightarrow{d} Z^{\theta}$. The claim in (14) now readily follows. If $\theta = 1$ and (13) holds, then $M(p;1) = (F^{-1}(x_p))' = 1/f(x_p)$ has to be positive and finite, and the limiting arrival process would be Poisson. □

Remark The condition (13) does not imply (12); nor does it ensure that $f(x_p)$ is finite and positive. Consider the pdf

$$f(x) = (\eta+1)|x|^{\eta}/2, \quad |x| \le 1, \eta > -1.$$

This is a corrected version of the pdf given in Chanda (1975), and discussed in Ghosh and Sukhatme (1981) (we thank a reviewer for noticing the error). The associated quantile function is given by

$$F^{-1}(u) = \begin{cases} -(1-2u)^{1/(\eta+1)}, & 0 < u \le \frac{1}{2} \\ (2u-1)^{1/(\eta+1)}, & \frac{1}{2} \le u < 1. \end{cases}$$

This quantile function fails to satisfy the condition in (12) when $p = 0.5$ and η is a positive even integer, but satisfies (13) with $\theta = 1/(\eta+1)$. Here, $x_p = 0$, $M(p,\theta) = 2^{\theta}$; $f(x_p)$ is 0 or ∞ depending on whether $\eta > 0$ or $\eta \in (-1,0)$.

3.2 Asymptotic independence of a central order statistic and spacings in its neighborhood

We now assume $k/n = p + o(n^{-1/2})$ and establish the independence of $X_{k:n}$ and spacings around it.

3.2.1 The uniform parent

Using the (well-known) joint pdf of the consecutive standard uniform order statistics $U_{k-s:n}, \ldots, U_{k:n}, \ldots, U_{k+r:n}$, we first obtain the joint pdf of appropriately normalized $U_{k:n}$ and the vector

$$\{n(U_{k+j:n} - U_{k+j-1:n}),\ j = -(s-1), \ldots, 0, \ldots, r\},$$

and thus determine the limiting form of the joint pdf.

The joint pdf of $U_{k-s:n}, \ldots, U_{k+r:n}$ is given by

$$\frac{n!}{(k-s-1)!(n-k-r)!} u_{k-s}^{k-s-1} (1-u_{k+r})^{n-k-r},$$

for $0 < u_{k-s} < \cdots < u_k < \cdots < u_{k+r} < 1$. With $t_n = \sqrt{[p(1-p)]/n}$, consider the transformation

$$\begin{aligned} v_0 &= (u_{k:n} - p)/t_n, \\ v_1^* &= n(u_{k:n} - u_{k-1:n}), \ldots, v_s^* = n(u_{k-s+1:n} - u_{k-s:n}); \\ v_1 &= n(u_{k+1:n} - u_{k:n}), \ldots, v_r = n(u_{k+r:n} - u_{k+r-1:n}). \end{aligned} \tag{15}$$

Hence,

$$\begin{aligned} u_{k:n} &= p + t_n v_0 \\ u_{k+1:n} &= p + t_n v_0 + (1/n) v_1 \\ &\cdots \\ u_{k+r:n} &= p + t_n v_0 + (1/n)(v_1 + v_2 + \cdots + v_r) \\ &\cdots \\ u_{k-1:n} &= p + t_n v_0 - (1/n) v_1^* \\ &\cdots \\ u_{k-s:n} &= p + t_n v_0 - (1/n)(v_1^* + v_2^* + \cdots + v_s^*), \end{aligned}$$

and the Jacobian is $\left|\frac{\partial \mathbf{u}}{\partial \mathbf{v}}\right| = t_n / n^{r+s}$. The joint pdf of $V_0, V_1, \ldots, V_r, V_1^*, \ldots, V_s^*$ is

$$\begin{aligned} &\frac{n!}{(k-s-1)!(n-k-r)!} \frac{t_n}{n^{r+s}} \{p + t_n v_0 - (1/n)(v_1^* + v_2^* + \cdots + v_s^*)\}^{k-s-1} \\ &\quad \times \{1 - p - t_n v_0 - (1/n)(v_1 + v_2 + \cdots + v_r)\}^{n-k-r}, \\ &\sim \frac{n!}{(k-1)!(n-k)!} t_n p^s (1-p)^r (p + t_n v_0)^{k-s-1} (1 - p - t_n v_0)^{n-k-r} \\ &\quad \times \left\{1 - \frac{(v_1^* + v_2^* + \cdots + v_s^*)}{n(p + t_n v_0)}\right\}^{k-s-1} \\ &\quad \times \left\{1 - \frac{(v_1 + v_2 + \cdots + v_r)}{n(1 - p - t_n v_0)}\right\}^{n-k-r}, \\ &= \eta_1 \times \eta_2 \times \eta_3, \end{aligned}$$

say. Since $k = np + o(\sqrt{n})$ and $t_n \to 0$,

$$\eta_2 \to \exp\{-(v_1^* + v_2^* + \cdots + v_s^*)\} \text{ and } \eta_3 \to \exp\{-((v_1 + v_2 + \cdots + v_r)\}$$

for fixed r and s. Further,

$$\begin{aligned}\eta_1 &= \frac{n!}{(k-1)!(n-k)!} t_n p^s (1-p)^r (p + t_n v_0)^{k-s-1} (1 - p - t_n v_0)^{n-k-r} \\ &\sim \frac{n!}{(k-1)!(n-k)!} t_n (p + t_n v_0)^{k-1} (1 - p - t_n v_0)^{n-k},\end{aligned}$$

the pdf of $V_0 = (U_{k:n} - p)/t_n$. With $k = np + o(\sqrt{n})$, using Stirling's approximations for the factorials and the expansion $\log(1 + x) = x - x^2/2 + o(x^2)$ for x close to 0, it can be shown that this pdf converges to $\phi(v_0)$.

The conclusion of the above discussion is summarized below.

Lemma 2 *With $k = np + o(\sqrt{n})$, asymptotically (i) $V_0, V_1, \ldots, V_r, V_1^*, \ldots, V_s^*$ are mutually independent, (ii) V_0 is $N(0, 1)$, and (iii) the remaining rvs are i.i.d. standard exponential, where the rvs involved are defined in (15).*

3.2.2 Arbitrary parent

By establishing density convergence under the assumption that $k/n = p + o(n^{-1/2})$, we have shown above that

$$\frac{\sqrt{n}}{\sqrt{p(1-p)}} (U_{k:n} - p) \xrightarrow{d} N(0, 1). \tag{16}$$

The conclusion in (16) also follows from Ghosh (1971) who has shown that if $f(x_p)$ is positive and finite,

$$\frac{\sqrt{n} f(x_p)}{\sqrt{p(1-p)}} (X_{k:n} - x_p) \xrightarrow{d} N(0, 1).$$

We have shown in Sect. 3.1 that when $k/n = p + o(1)$, if condition (11) holds or if (12) holds and $f(x_p)$ is finite and positive,

$$\frac{X_{k+j+1:n} - X_{k+j:n}}{U_{k+j+1:n} - U_{k+j:n}} = \frac{F^{-1}(U_{k+j+1:n}) - F^{-1}(U_{k+j:n})}{U_{k+j+1:n} - U_{k+j:n}} \xrightarrow{a.s.} \frac{1}{f(x_p)};$$

and if (12) and (13) hold,

$$\frac{X_{k+j+1:n} - X_{k+j:n}}{\Delta^{\theta}_{k+j:n}} = \frac{F^{-1}(U_{k+j:n} + \Delta_{k+j:n}) - F^{-1}(U_{k+j:n})}{\Delta^{\theta}_{k+j:n}} \xrightarrow{a.s.} M(p, \theta),$$

for $j = -s, \ldots, 0, \ldots, r-1$. Further,

$$\frac{X_{k:n} - x_p}{U_{k:n} - p} = \frac{F^{-1}(U_{k:n}) - F^{-1}(p)}{U_{k:n} - p} \overset{a.s.}{\to} \frac{1}{f(x_p)}$$

whenever $f(x_p)$ is finite and positive;

$$\frac{|X_{k:n} - x_p|}{|U_{k:n} - p|^\theta} = \frac{|F^{-1}(U_{k:n}) - F^{-1}(p)|}{|U_{k:n} - p|^\theta} \overset{a.s.}{\to} M(p, \theta),$$

if (13) holds.

In view of Lemma 2, assuming $k = np + o(\sqrt{n})$, we have established the asymptotic independence of the normalized spacings $(X_{k+j:n} - X_{k+j-1:n})$ introduced in Theorem 1 and appropriately normalized $X_{k:n}$ under the conditions stated there. This discussion leads to the following result.

Theorem 2 *Let $k/n = p + o(n^{-1/2})$, $0 < p < 1$, and r and s be fixed positive integers.*

(a) *If condition (11) holds, or if (12) holds and $f(x_p)$ is finite and positive,*

$$\left(\frac{\sqrt{n} f(x_p)(X_{k:n} - x_p)}{\sqrt{p(1-p)}}, nf(x_p)(X_{k+j:n} - X_{k+j-1:n}), -(s-1) \le j \le r \right)$$
$$\overset{d}{\to} (N(0,1), Z_s^*, \ldots, Z_1^*, Z_1, \ldots, Z_r).$$

(b) Assume (12) and (13) hold. Then,

$$\left(\frac{\sqrt{n}}{\sqrt{p(1-p)}} \left| \frac{(X_{k:n} - x_p)}{M(p,\theta)} \right|^{1/\theta}, n \left(\frac{X_{k+j+1:n} - X_{k+j:n}}{M(p,\theta)} \right)^{1/\theta}, -s \le j \le r-1 \right)$$
$$\overset{d}{\to} (|N(0,1)|, Z_s^*, \ldots, Z_1^*, Z_1, \ldots, Z_r).$$

In both cases, the Z_i's and Z_i^*'s are Exp(1) rvs, and the $r + s + 1$ components in the limit vector are mutually independent.

3.3 Remarks—the central case

Siddiqui (1960) considered higher order spacings around a central order statistic and showed that when F is continuously twice differentiable and $f(x_p)$ is finite and positive, the rvs

$$\frac{\sqrt{n} f(x_p)(X_{k:n} - x_p)}{\sqrt{p(1-p)}}, \quad nf(x_p)(X_{k:n} - X_{k-s:n}) \quad \text{and } nf(x_p)(X_{k+r:n} - X_{k:n})$$

are asymptotically independent when $k/n \to p \in (0,1)$ with r/n and s/n tending to zero; further, asymptotically the higher order spacings are Gam(r) and Gam(s),

respectively. We have proved a more refined result here with less assumptions on the properties of F, but have taken r and s to be fixed.

Pyke's (1965) classic paper on spacings shows (Theorem 5.1) that $nf(x_{p_1})(X_{i:n} - X_{i-1:n})$ and $nf(x_{p_2})(X_{j:n} - X_{j-1:n})$ with $i/n \to p_1$ and $j/n \to p_2$ where $0 < p_1 \neq p_2 < 1$ are asymptotically i.i.d. Exp(1) rvs. The key difference is that the spacings considered there are far apart, while our focus is on adjacent spacings around $X_{i:n}$.

The asymptotic half-normal distribution of the normalized central order statistic under the conditions of part (b) of Theorem 2 is comparable to Chanda's (1975) conclusion; our condition (13) is on F^{-1}, whereas his comparable condition (given as (6) there) is on F.

4 The intermediate case

Here, we lean on the work of Falk (1989) and directly examine the convergence of the joint pdf of an intermediate order statistic $X_{k:n}$ and spacings around it. We assume $k \to \infty$ but $k/n \to 1$ as $n \to \infty$ such that $n - k \to \infty$ and assume one of the von Mises sufficient conditions stated in (5)–(7) holds. Theorem 2.1 of Falk (1989) shows that

$$(X_{k:n} - a_n)/b_n \xrightarrow{d} N(0,1) \text{ with } a_n = F^{-1}(k/n);\ b_n = \sqrt{n-k}/(nf(a_n)). \quad (17)$$

This is established by showing that the pdf of $(X_{k:n} - a_n)/b_n$ at x,

$$\frac{n!}{(k-1)!(n-k)!}[F(a_n + b_n x)]^{k-1}[1 - F(a_n + b_n x)]^{n-k} b_n f(a_n + b_n x) \to \phi(x), \quad (18)$$

for all real x.

Consider the joint pdf of $X_{k-s:n}, \ldots, X_{k:n}, \ldots, X_{k+r:n}$:

$$\frac{n!}{(k-s-1)!(n-k-r)!}\left[\prod_{j=1}^{s} f(x_j^*)\right][F(x_s^*)]^{k-s-1}$$
$$f(x_0)\left[\prod_{j=1}^{r} f(x_j)\right][1 - F(x_r)]^{n-k-r},$$

for $x_s^* < \cdots < x_1^* < x_0 < x_1 < \cdots < x_r$. Define

$$c_n = \frac{1}{nf(a_n)} = \frac{b_n}{\sqrt{n-k}} = \frac{1}{nf(F^{-1}(k/n))} \quad (19)$$

and consider the transformation

$$\begin{aligned} y_0 &= (x_{k:n} - a_n)/b_n, \\ y_1 &= (x_{k+1:n} - x_{k:n})/c_n, \ldots, y_r = (x_{k+r:n} - x_{k+r-1:n})/c_n, \\ y_1^* &= (x_{k:n} - x_{k-1:n})/c_n, \ldots, y_s^* = (x_{k-s+1:n} - x_{k-s:n})/c_n, \end{aligned} \tag{20}$$

so that

$$\begin{aligned} x_{k:n} &= a_n + b_n y_0; \\ x_{k+j:n} &= a_n + b_n y_0 + c_n(y_1 + \cdots + y_j), \ j = 1, \ldots, r; \\ x_{k-j:n} &= a_n + b_n y_0 - c_n(y_1^* + \cdots + y_j^*), \ j = 1, \ldots, s. \end{aligned}$$

The Jacobian is $\left|\frac{\partial \mathbf{x}}{\partial \mathbf{y}}\right| = b_n c_n^{r+s}$ and the joint pdf of $Y_0, Y_1, \ldots, Y_r, Y_1^*, \ldots, Y_s^*$, defined in (20), is

$$\frac{n!}{(k-1)!(n-k)!}[F(a_n + b_n y_0)]^{k-1}[1 - F(a_n + b_n y_0)]^{n-k} b_n f(a_n + b_n y_0) \tag{21}$$

$$\times \left(\frac{F(a_n + b_n y_0 - c_n(y_1^* + \cdots + y_s^*))}{F(a_n + b_n y_0)}\right)^{k-s-1} \tag{22}$$

$$\times \prod_{j=1}^{s} \left\{ c_n(k-j) \frac{f(a_n + b_n y_0 - c_n(y_1^* + \cdots + y_j^*))}{F(a_n + b_n y_0)} \right\} \tag{23}$$

$$\times \prod_{j=1}^{r} \left\{ c_n(n-k-j+1) \frac{f(a_n + b_n y_0 + c_n(y_1 + \cdots + y_j))}{1 - F(a_n + b_n y_0)} \right\} \tag{24}$$

$$\times \left(\frac{1 - F(a_n + b_n y_0 + c_n(y_1 + \cdots + y_r))}{1 - F(a_n + b_n y_0)}\right)^{n-k-r} \tag{25}$$

$$= \tau_1 \times \tau_2 \times \tau_3 \times \tau_4 \times \tau_5, \tag{26}$$

where τ_1–τ_5 are, respectively, given by (21)–(25), y_0 is real, and $y_i, y_i^* > 0$.

From (18), it follows that $\tau_1 \to \phi(y_0)$. We will establish the limits for remaining τ_i using the following lemma.

Lemma 3 *Suppose one of the von Mises conditions stated in (5)–(7) holds, and $n \to \infty$, $k/n \to 1$ such that $n - k \to \infty$. Then, with a_n and b_n given by (17), the following statements hold.*

(a) *For any real y_0, $\frac{1-F(a_n+b_n y_0)}{1-F(a_n)} = \frac{n(1-F(a_n+b_n y_0))}{n-k} \to 1$.*

(b) *If $c_n = b_n/\sqrt{n-k}$, for any y_0, y_1 real, $\frac{f(a_n+b_n y_0+c_n y_1)}{f(a_n)} \to 1$.*

Proof (a) From Theorem 2.1 of Falk (1989), it follows that (17) holds under the conditions we have assumed, and the limit distribution is $N(0, 1)$. From Theorem 1

of Smirnov (1967) [Remark (ii) of Falk (1989)], it then follows that $[n-k+1+n(F(a_n+b_n y_0)-1)]/\sqrt{n-k+1} \to x$ for all real y_0. Thus,

$$\sqrt{n-k+1}\cdot\left\{1-\frac{1-F(a_n+b_n y_0)}{1-F(a_n)}\right\}\to y_0$$

since $1-F(a_n)=(n-k)/n$. This implies that $(1-F(a_n+b_n y_0))/(1-F(a_n))\to 1$ for any real y_0.

(b) In the proof of his Theorem 2.1, Falk establishes that whenever one of the sufficient conditions stated in (5)–(7) holds, for any real y for which $F(a_n+b_n y)\to 1$ (or equivalently $a_n+b_n y\to x_1$) as $n\to\infty$,

$$f(a_n+b_n\theta y)/f(a_n)\to 1 \tag{27}$$

uniformly for $\theta\in(0,1)$ where a_n and b_n are given in (17). Part (a) that we just proved implies that for any real y_0, $F(a_n+b_n y_0)\to 1$ as $n\to\infty$. Thus, from (27), it follows that $f(a_n+b_n\theta y_0)/f(a_n)\to 1$ for all real y_0.

For large $n-k$ and real y_1,

$$y_0+\frac{y_1}{\sqrt{n-k}}\in(0,2y_0) \text{ if } y_0>0;\ y_0+\frac{y_1}{\sqrt{n-k}}\in(2y_0,0) \text{ if } y_0<0.$$

Using (27) with $y=2y_0$, we conclude that $f(a_n+2y_0\theta b_n)\{f(a_n)\}^{-1}\to 1$ uniformly for all $0\le\theta<1$. Since $a_n+b_n y_0+c_n y_1=a_n+b_n(y_0+y_1/\sqrt{n-k})$ is in $(a_n, a_n+2y_0 b_n)$ if $y_0>0$ and in $(a_n+2y_0 b_n, a_n)$ if $y_0<0$. Hence, $f(a_n+b_n y_0+c_n y_1)/f(a_n)\to 1$ for all real $y_0\neq 0$ and real y_1. When $y_0=0$, for any real y_1, $f(a_n+b_n(1/\sqrt{n-k})y_1)/f(a_n)\to 1$ since $1/\sqrt{n-k}\in(0,1)$ and (27) holds. This completes the proof of the claim in (b). □

With $y=y_1^*+\cdots+y_s^*>0$, consider the following component of τ_2 in (22):

$$\begin{aligned}\frac{F(a_n+b_n y_0-c_n y)}{F(a_n+b_n y_0)}&=1-\frac{F(a_n+b_n y_0)-F(a_n+b_n y_0-c_n y)}{F(a_n+b_n y_0)}\\&=1-\frac{f(a_n+b_n y_0-\theta^* c_n y)}{F(a_n+b_n y_0)}\cdot c_n y\\&=1-\frac{y d_{k,n}}{k}\end{aligned}$$

where the second form above follows from the mean value theorem, $\theta^*\in(0,1)$, and

$$d_{k,n}=\frac{1}{F(a_n+b_n y_0)}\frac{f(a_n+b_n y_0+\theta^* c_n(-y))}{f(a_n)}\frac{k}{n},$$

where we have used the fact that $c_n=1/nf(a_n)$. From part (a) of Lemma 3, the first factor of $d_{k,n}$ above converges to 1; from part (b), the second factor approaches 1, and from our assumptions about k and n made in the intermediate case, the third factor

also approaches 1. Hence, $d_{k,n} \to 1$ as $n, k \to \infty$. Thus, upon recalling (22), we obtain

$$\tau_2 = \left(1 - \frac{y d_{k,n}}{k}\right)^{k-s-1} \to \exp\{-y\} = \exp\left\{-\left(\sum_{j=1}^{s} y_j^*\right)\right\}, \ y_j^* > 0.$$

With $y = y_1^* + \cdots + y_j^*$, the jth term in the product representing τ_3 in (23) is given by

$$c_n(k-j)\frac{f(a_n + b_n y_0 + c_n(-y))}{F(a_n + b_n y_0)} = \frac{k-j}{n}\frac{f(a_n + b_n y_0 + c_n(-y))}{f(a_n)}\frac{1}{F(a_n + b_n y_0)}.$$

Using Lemma 3 as we did in proving $d_{k,n} \to 1$ as n and $n - k \to \infty$, we conclude that the jth factor of $\tau_3 \to 1$ for all j and so does τ_3.

With $y = y_1 + y_2 + \cdots + y_j$, the jth term in the product representing τ_4 in (24) is given by

$$\begin{aligned}&\left\{c_n(n-k-j+1)\frac{f(a_n + b_n y_0 + c_n y)}{1 - F(a_n + b_n y_0)}\right\} \\ &= \frac{n-k-j+1}{n(1-F(a_n))} \cdot \frac{f(a_n + b_n y_0 + c_n y)}{f(a_n)} \cdot \frac{1 - F(a_n)}{1 - F(a_n + b_n y_0)}.\end{aligned}$$

Since $F(a_n) = k/n$, the first factor above converges to 1, and Lemma 3 shows that the other two factors also approach 1 as n and $n - k \to \infty$. Thus, $\tau_4 \to 1$.

Finally, with $y = y_1 + \cdots + y_r$, consider the following component of τ_5 in (25):

$$\begin{aligned}\frac{1 - F(a_n + b_n y_0 + c_n y)}{1 - F(a_n + b_n y_0)} &= 1 - \frac{F(a_n + b_n y_0 + c_n y) - F(a_n + b_n y_0)}{1 - F(a_n + b_n y_0)} \\ &= 1 - \frac{f(a_n + b_n y_0 + \theta^* c_n y)}{1 - F(a_n + b_n y_0)} \cdot c_n y \quad \text{from the Mean Value Theorem} \\ &= 1 - y\frac{f(a_n + b_n y_0 + \theta^* c_n y)}{f(a_n)} \cdot \frac{1 - F(a_n)}{1 - F(a_n + b_n y_0)} \cdot \frac{1}{n(1 - F(a_n))}\end{aligned}$$

where $\theta^* \in (0, 1)$ and we have used the fact that $c_n = 1/nf(a_n)$. Lemma 3 implies that the first two factors of y converge to 1 as $n, n - k \to \infty$, and the denominator of the last factor, $n(1 - F(a_n))$, is $n - k$. Hence,

$$\tau_5 \sim \left(1 - \frac{y_1 + \cdots + y_r}{n-k}\right)^{n-k-r} \to \mathrm{e}^{-(y_1 + \cdots + y_r)}, \ y_1, \ldots, y_r > 0,$$

as $n, n - k \to \infty$.

Thus, we have formally proved the following theorem.

Theorem 3 *Whenever one of the von Mises condition (5)–(7) holds, and $n, k, n-k \to \infty$ such that $k/n \to 1$, with $a_n = F^{-1}(k/n)$ and $c_n = 1/nf(a_n)$,*

$$((X_{k-s:n} - X_{k-s+1:n})/c_n, \ldots, (X_{k:n} - X_{k-1:n})/c_n, (X_{k+1:n} - X_{k:n})/c_n, \ldots,$$
$$(X_{k+r:n} - X_{k+r-1:n})/c_n) \xrightarrow{d} (Z_s^*, \ldots, Z_1^*, Z_1, \ldots, Z_r);$$
$$\frac{1}{c_n\sqrt{n-k}}(X_{k:n} - a_n) \xrightarrow{d} N(0,1),$$

where the Z_i^*'s and Z_i's are Exp(1) rvs, and the $r + s + 1$ limiting rvs are mutually independent.

4.1 Remarks—the intermediate case

When $F \in \mathcal{D}(G_{1;\alpha})$,

$$\frac{a_n f(a_n)}{1 - F(a_n)} = \frac{F^{-1}(k/n)c_n}{n-k} \to \alpha$$

and hence $\alpha(n-k)/F^{-1}(k/n)$ can be chosen as c_n. When $F \in \mathcal{D}(G_{2;\alpha})$, the von Mises condition implies that $\alpha(n-k)/(x_1 - F^{-1}(k/n))$ can be used as c_n. When $F \in \mathcal{D}(G_3)$, $[nf(a_n)m(a_n)/(n-k)] \to 1$ and we can use $m(a_n)/(n-k)$ as our c_n.

From Theorem 3, it follows that as in the central case, asymptotically, any two spacings, possibly of higher order, formed by nonoverlaping collections of order statistics around an intermediate order statistic $X_{k:n}$ are independent, and the collection is independent of $X_{k:n}$.

Teugels (2001) has introduced a family $\mathcal{C}^*$ of cdfs F with the following property: F has an ultimately positive pdf f and for all real y,

$$\frac{1}{h(x)}\left\{\frac{F\left(x + yh(x)\frac{1-F(x)}{f(x)}\right) - F(x)}{1 - F(x)}\right\} \to y,$$

whenever $h(x) \to 0$ as $x \to x_1$. He states that the condition $F \in \mathcal{C}^*$ 'slightly generalizes' Falk's (1989) version of von Mises conditions (i.e., (5)–(7)). Assuming $F \in \mathcal{C}^*$, Teugels shows that upon normalization described above (i) $X_{k:n}$ is asymptotically normal and (ii) $(X_{k:n} - X_{k-s:n})$ is asymptotically Gam(s). Their joint distribution and the asymptotic independence are not discussed there.

5 The upper extreme case

We now assume that $k \to \infty$ such that $n - k$ is fixed. It is well known that when $F \in D(G)$ for G given in (4),

$$((X_{n:n} - a_n)/b_n, \ldots, (X_{k:n} - a_n)/b_n) \xrightarrow{d} (W_1, \ldots, W_{n-k+1}) \tag{28}$$

where for any finite fixed j the vector $(W_1, \ldots, W_j)$ has the same joint distribution as

$$(Z_1^{-1/\alpha}, (Z_1 + Z_2)^{-1/\alpha}, \ldots, (Z_1 + \cdots + Z_j)^{-1/\alpha}), \text{ if } G = G_{1;\alpha}; \tag{29}$$
$$(-Z_1^{1/\alpha}, -(Z_1 + Z_2)^{1/\alpha}, \ldots, -(Z_1 + \cdots + Z_j)^{1/\alpha}), \text{ if } G = G_{2;\alpha}; \tag{30}$$

and if $G = G_3$,

$$(W_1, \ldots, W_j) \stackrel{d}{=} (-\ln Z_1, -\ln(Z_1 + Z_2), \ldots, -\ln(Z_1 + \cdots + Z_j)) \quad (31)$$

$$(W_1, \ldots, W_j) \stackrel{d}{=} \left(\sum_{i=k}^{\infty} \frac{Z_i - 1}{i} + \gamma - \sum_{i=1}^{k-1} \frac{1}{i}, k = 1, \ldots, j \right) \quad (32)$$

where the Z_i's are i.i.d. Exp(1) rvs, and γ is the Euler's constant. The first three representations above are from Nagaraja (1982) who also shows that the joint limiting distribution $(W_1, \ldots, W_j)$ is identical to the joint distribution of the first j lower record values from the cdf G. The representation in (32) is due to Hall (1978), and is more convenient when $G = G_3$.

Thus, whenever $F \in D(G)$, the limiting form of the joint distribution of the normalized spacings and the concerned extreme order statistics can be described as follows:

$$\begin{aligned} &((X_{n:n} - X_{n-1:n})/b_n, \ldots, (X_{k+1:n} - X_{k:n})/b_n, (X_{k:n} - a_n)/b_n, \\ &\quad (X_{k:n} - X_{k-1:n})/b_n, \ldots, (X_{k-s+1:n} - X_{k-s:n})/b_n)) \\ &\stackrel{d}{\to} (W_1 - W_2, \ldots, W_{n-k} - W_{n-k+1}, W_{n-k+1}, \\ &\quad W_{n-k+1} - W_{n-k+2}, \ldots, W_{n-k+s} - W_{n-k+s+1}) \end{aligned} \quad (33)$$

where the W_j's have one of the forms given in (29)–(32). We now specialize our results for each of the three domains.

5.1 The Fréchet domain

In this case, a_n can be chosen to be 0 and b_n to be $F^{-1}(1 - 1/n)$ $(= x_{1-n^{-1}})$. The representation in (33) for the limiting joint distribution along with (29) suggests that an extreme spacing is not asymptotically exponential, and the adjacent spacings are neither independent, nor identically distributed in the limit. The asymptotic independence of the spacings and the extreme order statistic also fail. Hence, when $F \in D(G_{1;\alpha})$, the asymptotic distributional structure for the extreme spacings differs from that for the central and intermediate cases.

From (28) and (29), we conclude that when $F \in \mathcal{D}(G_{1;\alpha})$, the normalized higher order spacing,

$$(X_{n:n} - X_{n-j:n})/b_n \stackrel{d}{\to} W_1 - W_{j+1} \stackrel{d}{=} Z_1^{-1/\alpha} - (Z_1 + S_j)^{-1/\alpha},$$

where the sum $S_j = Z_2 + \cdots + Z_{j+1}$ is a Gam(j) rv that is independent of Z_1. This distributional representation complements the work of Pakes and Steutel (1997) who have given an expression for the cdf of the limiting rv as (p. 192)

$$P(W_1 \le w) + P(W_1 - (S_j + W_1^{-\alpha})^{-1/\alpha} \le w, W_1 > w), \ w > 0.$$

They comment that this expression for the cdf 'does not seem susceptible to simplification for any choice of the parameter α'; for the other two domains, they provide explicit distributional representation that is equivalent to ours (see below).

5.2 The Weibull domain

Here, $x_1 (= F^{-1}(1))$ is finite and can be chosen to be our a_n and the scaling constant b_n can be chosen to be $x_1 - x_{1-1/n}$. From (33) and (30), we can conclude that the normalized adjacent spacings are asymptotically i.i.d. exponential iff $\alpha = 1$ when $F \in \mathcal{D}(G_{2;\alpha})$. Otherwise, they are all dependent. When $\alpha = 1$ and $k < n$, the joint asymptotic distributional structure of $((X_{n:n} - X_{n-1:n})/b_n, \ldots, (X_{k+1:n} - X_{k:n})/b_n, (X_{k:n} - a_n)/b_n, (X_{k:n} - X_{k-1:n})/b_n, \ldots, (X_{k-s+1:n} - X_{k-s:n})/b_n)$ is that of

$$(Z_2, \ldots, Z_{n-k+1}, -(Z_1 + \cdots + Z_{n-k+1}), Z_{n-k+2}, \ldots, Z_{n-k+s+1}). \tag{34}$$

Thus, when $\alpha = 1$, $X_{k:n}$ is asymptotically independent of the spacings in its left neighborhood, but is symmetrically dependent on the ones on its right. This conclusion is formalized in the following result.

Theorem 4 *When $F \in \mathcal{D}(G_{2;\alpha=1})$, for each fixed $n - k$ and s, the asymptotic joint distribution of $((X_{n:n} - X_{n-1:n})/b_n, \ldots, (X_{k+1:n} - X_{k:n})/b_n, (X_{k:n} - a_n)/b_n, (X_{k:n} - X_{k-1:n})/b_n, \ldots, (X_{k-s+1:n} - X_{k-s:n})/b_n)$ has the representation given by (34), where the Z_j's are i.i.d. Exp(1) rvs. The b_n can be chosen as $x_1 - F^{-1}(1 - 1/n)$. When $F \in \mathcal{D}(G_{2;\alpha})$, the adjacent spacings are i.i.d. exponential iff $\alpha = 1$.*

The standard uniform distribution is in $\mathcal{D}(G_{2;\alpha})$ with $\alpha = 1$ and hence has asymptotically i.i.d. extreme spacings. We had reached this conclusion earlier in Lemma 1 (recall $x_1 - F^{-1}(1 - 1/n) = 1/n$). But we have a more general result now that describes the symmetric dependence of the right neighborhood spacings on $X_{k:n}$ and is applicable to all $F \in \mathcal{D}(G_{2;1})$. In fact with $n - k$ fixed, given $(X_{k:n} - x_1)/b_n = u \ (< 0), (X_{n:n} - X_{n-1:n})/b_n, \ldots, (X_{k+1:n} - X_{k:n})/b_n$ behave asymptotically like the spacings from a random sample of size $n - k$ from a uniform distribution over $(u, 0)$. For the form of the joint distribution, see, e.g., David and Nagaraja (2003; Sec. 6.4).

From (28) and (30), we conclude that when $F \in \mathcal{D}(G_{2;\alpha})$, the normalized higher order spacing

$$(X_{n:n} - X_{n-j:n})/b_n \xrightarrow{d} W_1 - W_{j+1} \stackrel{d}{=} (Z_1 + S_j)^{1/\alpha} - Z_1^{1/\alpha},$$

where the sum $S_j = Z_2 + \cdots + Z_{j+1}$ is a Gam(j) rv that is independent of Z_1. This is the conclusion of Theorem 7.2 in Pakes and Steutel (1997).

5.3 The Gumbel domain

Using (33) along with the representation for W_j in (32), we conclude the following.

Theorem 5 *When $F \in \mathcal{D}(G_3)$, for each fixed $n-k$ and s, the asymptotic joint distribution of $((X_{n:n} - X_{n-1:n})/b_n, \ldots, (X_{k+1:n} - X_{k:n})/b_n, (X_{k:n} - a_n)/b_n, (X_{k:n} - X_{k-1:n})/b_n, \ldots, (X_{k-s+1:n} - X_{k-s:n})/b_n)$ has the following distributional representation:*

$$\left(Z_1, \frac{Z_2}{2}, \ldots, \frac{Z_{n-k}}{n-k}, \sum_{i=n-k+1}^{\infty} \frac{Z_i - 1}{i} + \gamma - \sum_{i=1}^{n-k} \frac{1}{i}, \frac{Z_{n-k+1}}{n-k+1}, \ldots, \frac{Z_{n-k+s}}{n-k+s}\right) \tag{35}$$

where the Z_i's are i.i.d. Exp(1) rvs. The b_n can be chosen as $m(F^{-1}(1-1/n))$, and if (7) holds, as $\{nf(F^{-1}(1-1/n))\}^{-1}$.

The representation in (35) shows that while $X_{k:n}$ is independent of spacings in its right neighborhood, it is correlated with the spacings in its left neighborhood, and this correlation decreases at the rate of $1/(n-k+j)$, $j \geq 1$ as one moves away from it. This is in contrast with the situation when $F \in \mathcal{D}(G_{2,1})$.

Weissman (1978) has considered the limit distribution of

$$((X_{n:n} - X_{n-1:n})/b_n, \ldots, (X_{k+1:n} - X_{k:n})/b_n, (X_{k:n} - a_n)/b_n),$$

and has given the representation given by the first $(n-k)$ components of the vector in (35). He has also noted the independence of these spacings and W_{n-k+1} (in his Theorem 2).

The above theorem shows that using varying scaling sequences for the spacings, we can obtain i.i.d. standard exponential distributions in the limit. In particular, we have the following:

$$((n-j+1)(X_{n-j+1:n} - X_{n-j:n})/b_n, j = 1, \ldots, n-k+s) \xrightarrow{d} (Z_1, \ldots, Z_{n-k+s}).$$

From (28) and (32) or from the representation (35), we can conclude that when $F \in \mathcal{D}(G_3)$, the normalized higher order spacing,

$$(X_{n:n} - X_{n-j:n})/b_n \xrightarrow{d} W_1 - W_{j+1} \stackrel{d}{=} Z_1 + \frac{Z_2}{2} + \cdots + \frac{Z_j}{j} \stackrel{d}{=} Z_{j:j}.$$

The last equality above follows from the representation for exponential order statistics given in (3). This is the conclusion reached in Theorem 7.1 of Pakes and Steutel (1997).

5.4 Remarks—the extreme case

5.4.1 Cox processes

Whenever $F \in \mathcal{D}(G)$, we can say that the arrival process of order statistics in the left neighborhood of an upper extreme order statistic is asymptotically a Poisson process and is independent of its value only when $F \in G_2$ with $\alpha = 1$. The arrival processes

on both sides of the extreme order statistics are pure birth processes when $F \in G_3$; only the arrival process on the right side is independent of the order statistic.

Harshova and Hüsler (2000) have shown that the arrival processes on the left neighborhood of the sample maximum are special *Cox processes*, when G is of Weibull ($G_{2;\alpha}$) or Gumbel (G_3) type cdf. Cox processes are mixed Poisson processes where the time-dependent intensity $\lambda(t)$ is itself a stochastic process (Daley and Vere-Jones (2003; Sec. 6.2)). Harshova and Hüsler consider the counting process in the left neighborhood of $X_{n:n}$, $N_n(\cdot)$, defined by $N_n([a,b)) = \#\{X_i \in [X_{n:n} - a \cdot d_n, X_{n:n} - b \cdot d_n)\}, 0 < a < b$. From their Theorem 1.2, it follows that $N_n(\cdot) \xrightarrow{d} N(\cdot)$ where $\{N(t), t > 0\}$ is a Cox process with stochastic intensity function $\lambda(t) = \mathrm{e}^{t-W}$ in the Gumbel case where W has cdf G_3; and $\lambda(t) = \alpha(t-W)^{\alpha-1}$ in the Weibull case where W has cdf $G_{2;\alpha}$. The cdfs $G_{2;\alpha}$ and G_3 are given in (4). Representations given in (30), and (31) or (32) provide another characterization of the resulting Cox processes in terms of the distribution of inter-arrival times of order statistics below the maximum.

5.4.2 Higher order extreme spacings

The representation for the special higher order extreme spacing involving the sample maximum ($X_{n:n} - X_{n-j:n}$, discussed above) can be expanded to other extremes. From (28)–(30), and (32), we conclude that as $n \to \infty$, for fixed $1 \le i < j$, $(X_{n-i+1:n} - X_{n-j+1:n})/b_n$ converges in distribution to

$$\begin{aligned}
\left(\sum_{l=1}^{i} Z_l\right)^{-1/\alpha} - \left(\sum_{l=1}^{j} Z_l\right)^{-1/\alpha} &\stackrel{d}{=} W_i - (W_i^{-\alpha} + \mathrm{Gam}((j-i)))^{-1/\alpha}, G \in \mathcal{D}(G_{1;\alpha});\\
\left(\sum_{l=1}^{i} Z_l\right)^{1/\alpha} - \left(\sum_{l=1}^{j} Z_l\right)^{1/\alpha} &\stackrel{d}{=} (W_i^{\alpha} + \mathrm{Gam}((j-i)))^{1/\alpha} - W_i, G \in \mathcal{D}(G_{2;\alpha}),\\
&\stackrel{d}{=} \mathrm{Gam}((j-i)), G \in \mathcal{D}(G_{2;1});\\
\frac{Z_i}{i} + \cdots + \frac{Z_{j-1}}{j-1} &\stackrel{d}{=} Z_{j-i:j-1}, G \in \mathcal{D}(G_3). \qquad (36)
\end{aligned}$$

Here, the last distributional equality follows from (3). The above representations are extremely helpful in providing the asymptotic distribution theory for the number of order statistics around a specified extreme order statistic. This will be illustrated in the next section where all cases (central, extreme, and intermediate) will be considered.

6 Counts of observations around an order statistic

Consider the following count statistics that track the number of observations in the right and left neighborhoods of $X_{k:n}$:

$$\begin{aligned}
K_-(n,k,d) &= \#\{j : X_j \in (X_{k:n} - d, X_{k:n})\},\\
K_+(n,k,d) &= \#\{j : X_j \in (X_{k:n}, X_{k:n} + d)\}.
\end{aligned}$$

Clearly,

$$P(K_-(n,k,d) < i) = P(X_{k:n} - X_{k-i:n} > d),$$
$$P(K_+(n,k,d) < j) = P(X_{k+j:n} - X_{k:n} > d), \tag{37}$$

and thus the asymptotic distribution theory for spacings developed here can be directly applied to determine the limit distributions of the count statistics for appropriately chosen d that is dependent on n. Pakes and Steutel (1997) have used the link in (37) in the reverse direction in the extreme case where they derive the limit distribution of $K_-(n,n,d_n)$ first and use it to determine the limit distribution of the spacing $X_{n:n} - X_{n-k:n}$.

As noted in the introduction, the literature on the investigation into the limit distribution of K_- and K_+ is substantial. Poisson limits are generally obtained when $d = d_n$ is nonrandom but is dependent on the behavior of F around the concerned statistic. We now discuss implications of our results on spacings on the asymptotic distribution of counts and compare our results with only the most relevant results in the literature.

6.1 The central and intermediate cases—the Poisson counts

We have seen in Theorems 1 and 3 that the $(X_{k+i:n} - X_{k+i-1:n})/c_n$ are asymptotically i.i.d. standard exponential for any fixed (positive or negative) integer i, where $c_n = 1/nf(x_{p_n})$ with $p_n \equiv p = \lim(k/n) \in (0,1)$ in the central case, and in the intermediate case, $p_n = k/n \to 0$ or 1 such that, respectively k or $n-k \to \infty$. In other words, $K_-(n,k,\lambda_1 c_n)$ and $K_+(n,k,\lambda_2 c_n)$ are asymptotically independent, and Poi(λ_1) and Poi(λ_2) rvs, respectively. This conclusion matches with that of Pakes (2009) who has established the asymptotic Poisson property under a variety of technical conditions for the central [Theorem 16] and intermediate cases [Theorems 9, 10 (b), 12 (b)].

Dembińska et al. (2007) have established the asymptotic Poisson nature of both K_- and K_+ for d_n that depends on the local property of the quantile function F^{-1} in the intermediate case and an associated technical condition on F in the central case. Their condition for Poisson convergence of $K_+(n,k,d_n)$ for a central order statistic [(17) in their Theorem 5.2] is

$$\lim_{(x,y)\to(x_p,0^+)} \frac{F(x+y) - F(x)}{F(x_p+y) - F(x_p)} = 1. \tag{38}$$

This is comparable to our condition (12), but there are differences in these conditions and their implications. While (12) specifies the behavior of the quantile function F^{-1} around p, (38) puts a similar condition on the property of the cdf F around x_p. Dembińska et al.'s neighborhood is determined by $d_n = F^{-1}(p+\lambda/n) - x_p$, a quantity dependent on the behavior of F^{-1} at $(p+\lambda/n)$. In contrast, our $d_n = \lambda/n(f(x_p))$ depends on the behavior of F^{-1} only at x_p. When f is continuous around x_p and $f(x_p)$ is positive and finite, from L'Hospital's rule it follows that (38) is readily satisfied, and $F^{-1}(p+\lambda/n) - x_p \sim \lambda/n(f(x_p))$.

Under (38) and a similar condition on the left neighborhood, Dembińska and Balakrishnan (2010) have established the asymptotic independence of K_- and K_+. This independence readily follows from our Theorems 2 and 3. The technical conditions used in Dembińska et al. (2007) in the intermediate case (in their Theorem 6.1, for example) appear difficult to verify whereas the familiar von Mises conditions needed here are known to hold for many common distributions. In addition, our results show that counts in disjoint intervals are Poisson and independent, and also that these are independent of the location of $X_{k:n}$. These finer conclusions on the limiting structure of the neighborhood cannot be reached using any of the currently available results in the literature on count statistics for the central and intermediate cases.

6.2 The upper extremes—non-Poisson and Poisson counts

Asymptotic distributions of $K_-(n,k,d)$ and $K_+(n,k,d)$ have been investigated by many authors when k or $n-k$ are held fixed starting from the work of Pakes and Steutel (1997) who looked at $K_-(n,n,d)$. Assuming k is held fixed, Pakes and Li (1998) showed that $K_-(n,n-k,d)$ is asymptotically negative binomial, and Balakrishnan and Stepanov (2005) showed that $K_+(n,n-k,d)$ is asymptotically binomial. The success probability in these distributions is given by

$$\beta(d) = \lim_{x\to x_1} \frac{\overline{F}(x+d)}{\overline{F}(x)} \in (0,1),$$

where x_1 is assumed to be infinite.

Pakes (2009) has considered the limit distribution of $K_+(n,n-k,cb_n)$ with k fixed assuming that F is in the domain of attraction of either Fréchet or Gumbel distribution and the b_n's form the associated scaling sequence. When $F \in \mathcal{D}(G_{1;\alpha})$, he shows that the limit distribution of $K_+(n,n-k,b_n)$ is mixed binomial with parameters k and random success probability that is a function of a Gam$(k+1)$ rv (his Theorem 5, part (a)). When $F \in \mathcal{D}(G_3)$, $K_+(n,n-k,\lambda b_n)$ is shown to be asymptotically a Binomial rv with parameters k and success probability $1-\mathrm{e}^{-\lambda}$ (his Theorem 4).

We now examine the consequences of the representations in (36) and the relations in (37). When k is fixed and $F \in \mathcal{D}(G_3)$, for any j, $1 \le j \le k$,

$$\begin{aligned} P(K_+(n,n-k,\lambda b_n) < j) &= P(X_{n-k+j:n} - X_{n-k:n} > \lambda b_n) \\ &\to P(Z_{j:k} > \lambda) = \sum_{i=0}^{j-1} \binom{k}{i} (1-\mathrm{e}^{-\lambda})^i (\mathrm{e}^{-\lambda})^{k-i}, \end{aligned}$$

as $n \to \infty$. Since the maximum value attainable is k, the limit distribution is binomial, a result noted above. Further,

$$\begin{aligned} P(K_-(n,n-k,\lambda b_n) < j) &= P(X_{n-k:n} - X_{n-k-j:n} > \lambda b_n) \\ &\to P(Z_{j:k+j} > \lambda) = \sum_{i=0}^{j-1} \binom{k+j}{i} (1-\mathrm{e}^{-\lambda})^i (\mathrm{e}^{-\lambda})^{k+j-i}, \end{aligned}$$

resulting in a negative binomial distribution, a result shown by Pakes and Li (1998). They also derive the limit distribution of $K_-(n, n-k, \lambda b_n)$ in other cases; the representations in (36) along with (37) yield us the same results. Of these, a commonly known distribution is obtained only when $F \in \mathcal{D}(G_{2;1})$ in which case $K_+(n, n-k, \lambda b_n)$ has a censored Poi(λ) distribution that is censored on the right at k; this conclusion was reached in Theorem 4.1 of Dembińska et al. (2007) under a set of technical conditions similar to the one given in (38). Further, $K_-(n, n-k, \lambda b_n)$ will have a Poi(λ) distribution and these two statistics are asymptotically independent.

Whenever $F \in \mathcal{D}(G_{1;\alpha}$ or $G_{2;\alpha\neq 1})$, we can obtain the asymptotic distributions of K_- and K_+ using (36) and (37) directly. For example, when $F \in \mathcal{D}(G_{1;\alpha})$, we can use the corresponding representation in (36) to obtain the cdf of $K_+(n, n-k, b_n)$ in terms of Gamma rvs (in contrast with the mixed binomial representation of Pakes (2009) mentioned earlier). While closed form expression for the cdf may not be available, the needed probabilities can be evaluated using tractable univariate integrals that involve gamma type integrands that can be easily evaluated numerically. The link between Gamma and Poisson cdfs comes in handy in this simplification.

7 Discussion

We now provide further illustrations of applications of our results to distribution theory and inference.

7.1 Examples

Our examples thus far were the uniform and exponential populations, but our results are widely applicable since the conditions imposed here are satisfied by several common distributions. In the central case, we need positivity and continuity of the population pdf at x_p to achieve independent Poisson arrival process in both right and left neighborhoods. von Mises conditions are satisfied by the common distributions that are in the domain of attraction of an extreme value cdf G (given in (4)) and thus the intermediate case also leads to independent Poisson arrival process for these distributions. The extreme case does not require the von Mises conditions, and provides interesting examples of situations where we do not get Poisson processes. For example, for $F \in \mathcal{D}(G_{1;\alpha})$, a property satisfied by Pareto and loggamma distributions, the arrival process is no longer Poisson. Tables 3.4.2–3.4.4 of Embrechts et al. (1997) contain a good list of distributions in the domain of attraction of each of the three extreme value distributions along with the necessary norming (scaling) constants needed for the application of our results in the extreme case.

Our intermediate and extreme case discussions focused on the upper end of the sample. Parallel results hold for the lower end of the sample and upper-end and lower-end spacings can exhibit different types of clustering processes. For example, in the exponential parent case, upper extremes are in the Gumbel domain, and the lower extremes are in the Weibull domain with $\alpha = 1$ (i.e., Exp (1)). Thus, for the lower extremes, we have a homogeneous Poisson arrival process in the right neighborhood, whereas for the upper extremes, we have a pure birth process in the left neighborhood of the concerned order statistic.

7.2 Inferential implications

Theorem 2 (a) can be used in the central case to provide (asymptotically) distribution-free estimates of x_p and $f(x_p)$ as noted by Siddiqui (1960) when he studied the joint distribution of $X_{k:n}, X_{k+r:n} - X_{k:n}, X_{k:n} - X_{k-s:n}$ [see Sect. 3.3]. It follows from Theorem 2 that $nf(x_p)(X_{k+r:n} - X_{k-s:n})$ is asymptotically Gam$(r+s)$ and this fact can be used to provide estimates of $f(x_p)$ and confidence intervals for the population pdf at the pth quantile. The asymptotic independence of $nf(x_p)(X_{k+r:n} - X_{k-s:n})$ and $\sqrt{n}f(x_p)(X_{k:n} - x_p)$ and their known familiar distributions can be used to find the distribution of the pivotal quantity

$$\frac{(X_{k:n} - x_p)}{\sqrt{n}(X_{k+r:n} - X_{k-s:n})}.$$

From Theorem 2, it follows that this rv behaves asymptotically as the ratio of a standard normal and an independent gamma rv (or a scaled Chi-square rv) and this distribution is free of $f(x_p)$. It easily leads to an asymptotically distribution-free confidence interval for x_p.

A similar application of Theorem 3 would provide asymptotically distribution-free inference for the intermediate population quantile $F^{-1}(k/n)$ and pdf $f(F^{-1}(k/n))$ when one of the von Mises conditions is assumed to hold.

In the extreme case, we have seen that the limit distributions of the top k order statistics are dependent on the domain of attraction. Weissman (1978) has discussed in detail inference on tail parameters (extreme quantiles and the *tail index* $1/\alpha$) based on these limit results.

7.3 Concluding remarks

It is interesting to note that norming/scaling constants for $X_{k:n}$ and the adjacent spacings are of the same order only for the extreme case (b_n); the limiting distributions are similar as well (functions of Exp(1) rvs). For the central and intermediate cases, the spacings and $X_{k:n}$ are scaled differently, and their limit distributions are different; the spacings are related to exponential, whereas $X_{k:n}$ relates to the normal. For the extreme and intermediate cases, our sufficient conditions that ensure the nondegenerate limit distributions for $X_{k:n}$ and for the adjacent spacings are the same. In the central case, asymptotic normality for $X_{k:n}$ requires $k = np + o(\sqrt{n})$ (actually slightly less restriction on k would work), the asymptotic independence property of spacings holds whenever $k = np + o(n)$.

We have focused here on neighborhoods of a single selected order statistic; this work can easily be extended to multiple neighborhoods. In the case of two or more central order statistics and their neighborhoods, we obtain a multivariate normal limit distribution for the selected order statistics and independent Poisson processes around them. Such a set up is considered in Theorem 3.1 of Dembińska and Balakrishnan (2010) where the independence of Poisson counts in right and left neighborhoods of multiple

central order statistics is derived. In other cases (for example, one upper extreme, and another lower, considered in Theorem 2.1 of Dembińska and Balakrishnan (2010)), the resulting counting processes will turn out to be independent.

Acknowledgments The authors thank two anonymous reviewers for a close reading of the manuscript and many helpful suggestions.

References

Balakrishnan, N., Stepanov, A. (2005). A note on the number of observations near an order statistic. *Journal of Statistical Planning and Inference*, *134*, 1–14.

Chanda, K. C. (1975). Some comments on the asymptotic probability laws of sample quantiles. *Calcutta Statistical Association Bulletin*, *24*, 123–126.

Daley, D. J., Vere-Jones, D. (2003). *An introduction to the theory of point processes. Elementary theory and methods*, Vol. I, (2nd ed.) New York: Springer.

David, H. A., Nagaraja, H. N. (2003). *Order Statistics* (3rd ed.). Hoboken: Wiley.

de Haan, L. (1970). *On regular variation and its application to the weak convergence of sample extremes. Mathematical centre tracts 32*. Amsterdam: Mathematisch Centrum.

Dembińska, A., Balakrishnan, N. (2010). On the asymptotic independence of numbers of observations near order statistics. *Statistics*, *44*, 517–528.

Dembińska, A., Stephanov, A., Wesołowski, J. (2007). How many observations fall in a neighborhood of an order statistic? *Communications in Statistics -Theory and Methods*, *36*, 851–867.

Embrechts, P., Klüppelberg, C., Mikosch, T. (1997). *Modelling extremal events*. Berlin: Springer.

Falk, M. (1989). A note of uniform asymptotic normality of intermediate order statistics. *Annals of Institute of Statistical Mathematics*, *41*, 19–29.

Ghosh, J. K. (1971). A new of proof of Bahadur's representation of quantiles and an application. *Annals of Mathematical Statistics*, *42*, 1957–1961.

Ghosh, M., Sukhatme, S. (1981). On Bahadur's representation of quantiles in nonregular cases. *Communications in Statistics -Theory and Methods*, *10*, 269–282.

Gnedenko, B. V. (1943). Sur la distribution limite du terme maximum d'une série altéatoire. *Annals of Mathematics*, *44*, 423–453.

Hall, P. (1978). Representations and limit theorems for extreme value distributions. *Journal of Applied Probability*, *15*, 639–644.

Harshova, E., Hüsler, J. (2000). On the number of near-maxima. *Rendiconti del Circolo Mathematico di Palermo, Serie II, Supplementi*, *65*, 121–136.

Nagaraja, H. N. (1982). Record values and extreme value distributions. *Journal of Applied Probability*, *19*, 233–239.

Pakes, A. G. (2009). Number of observations near order statistics. *Australian and New Zealand Journal of Statistics*, *51*(4), 375–395.

Pakes, A. G., Li, Y. (1998). Limit laws for the number of near maxima via the Poisson approximation. *Statistics and Probability Letters*, *40*, 395–401.

Pakes, A. G., Steutel, F. (1997). On the number of records near the maximum. *Australian Journal of Statistics*, *39*, 179–192.

Pyke, R. (1965). Spacings. *Journal of Royal Statistical Society, Series B*, *27*, 395–436.

Reiss, R.-D. (1989). *Approximate distributions of order statistics*. New York: Springer.

Siddiqui, M. M. (1960). Distribution of quantiles in samples from a bivariate population. *Journal of Research, National Bureau of Standards*, *64B*, 145–150.

Smirnov, N.V. (1952). *Limit distributions for the terms of a variational series. American Mathematical Translations, Series 1*, no. 67. Original published in 1949.

Smirnov, N. V. (1967). Some remarks on limit laws for order statistics. *Theory of Probability and Its Applications*, *12*, 82–143.

Teugels, J. L. (2001). On order statistics close to the maixmum. In M. de Gunst, C. Klaassen, A. van der Vaart (Eds.), *State of the art in probability and statistics* (pp. 568–581). Beachwood: Institute of Mathematical Statistics.

von Mises, R. (1936). La distribution de la plus grande de n valeurs. *Review of Mathematical Union Interbalkanique*, *1*, 141–160.

Weissman, I. (1978). Estimation of parameters and larger quantiles based on the k largest observations. *Journal of American Statistical Association*, *73*, 812–815.