ON ORDER STATISTICS CLOSE TO THE MAXIMUM

JEF L. TEUGELS

Katholieke Universiteit Leuven

We investigate the asymptotic properties of order statistics in the immediate vicinity of the maximum of a sample. The usual domain of attraction condition for the maximum needs to be replaced by a continuity condition. We illustrate the potential of the approach by a number of examples.

AMS subject classifications: 62G30, 60F99.
Keywords and phrases: Extremal law, domain of extremal attraction, extreme value statistics, geometric range of a sample, extremal quotient.

1 Introduction

Let $X_1, X_2, \ldots, X_n$ be a sample from X with distribution F. For convenience, we assume that $F(x) = P(X \leq x)$ is ultimately continuous for large values of x. We denote the order statistics of the sample by

$$X_1^* \leq X_2^* \leq \ldots \leq X_n^*.$$

Under the *(extremal domain of) attraction condition*, we mean the condition on F inducing the convergence in distribution of the normalized maximum $(X_n^* - b_n)/a_n$ to a non-degenerate limit law. Here a_n are positive constants while the constants b_n, $(n = 1, 2, \ldots)$ are real. The attraction condition can be given in terms of the *tail quantile function* $U(y) := inf\{x : F(x) \geq 1 - \frac{1}{y}\}$ as shown by de Haan [10]. F belongs to an extremal domain of attraction if and only if there exists an ultimately positive *auxiliary function* g and a real *extremal index* γ such that for all $y > 0$ the condition

$$\text{(1)} \qquad \lim_{x \to \infty} \frac{U(xy) - U(x)}{g(x)} = \int_1^y w^{\gamma - 1}\, dw =: h_\gamma(y)$$

holds. In particular, $h_0(y) = \log y$. The constants can then be taken as $b_n = U(n)$ and $a_n = g(n)$ while the one-parameter family of possible limit laws is given by the *(class of) extreme value distributions*

$$G_\gamma(x) := \exp -(1 + \gamma x)_+^{-\frac{1}{\gamma}}.$$

If the tail quantile function U of F satisfies (1), we will write $F \in \mathcal{C}_\gamma(g)$. Note that the function g can be taken to be continuous. Furthermore g satisfies the regular variation property for $u > 0$

$$\frac{g(xu)}{g(x)} \to u^\gamma$$

when $x \uparrow \infty$. If $\gamma \geq 0$, then $g(x)/U(x) \to \gamma$; if $\gamma < 0$, $U(x) \uparrow x_+$, the right-end point of F, and $g(x)/(x_+ - U(x)) \to -\gamma$.

In what follows we are interested in order statistics $X^*_{n-\ell+1}$ that are very close to the maximum. To be more precise, we assume that n and ℓ tend to ∞ but that $\ell/n \to 0$. We search for a centering sequence $\{b_n\}$ of real numbers and a normalizing sequence $\{a_n\}$ of positive reals for which $a_n^{-1}(X^*_{n-\ell+1} - b_n)$ converges in distribution. The solution to this kind of problem depends on the type of condition one would like to impose. One could for example quantify the dependence of ℓ on n explicitly, imposing some strong regularity conditions on ℓ. Non-normal laws are then possible as shown by Chibisov [5] and more recently by Cheng, de Haan and Huang [4]. Alternatively, conditions can be imposed on the underlying distribution F or on its tail-quantile function U. This approach has been followed by Mason and van Zwet [14] and more particularly by Falk in [6,7]. For a comprehensive treatment, see the books by Reiss [15] and Leadbetter, Lindgren and Rootzén [12].

In this paper we offer a sufficient but unifying condition to arrive at asymptotic normality of the intermediate order statistics close to the maximum. We also illustrate the potential of the condition with a variety of different examples. After developing a rationale for its introduction in the next section, the condition is described and illustrated in section 3. The remaining sections contain applications to one and two order statistics.

2 Rationale for the condition

We outline two approaches. The first is based on the Helly-Bray theorem [11] while the second proceeds along a transformation.

2.1 Helly-Bray approach

Take m to be any real-valued bounded and continuous function on $\Re$. We investigate the limiting behaviour of $E_n := E\left\{m\left(a_n^{-1}(X^*_{n-\ell+1} - b_n)\right)\right\}$. By a classical combinatorial argument one writes

$$E_n = \frac{n!}{(\ell-1)!(n-\ell)!}\int_{-\infty}^{\infty} m\left(\frac{x-b_n}{a_n}\right) F^{n-\ell}(x)\{1-F(x)\}^{\ell-1} dF(x).$$

Note that the two exponents in the integrand are tending to ∞. So we need to rewrite the integrand in such a way that both factors can be handled simultaneously. To achieve that, we follow a procedure suggested by Smirnov [16]. Substitute $1 - F(x) = q + pv$ where the sequences $q = q(\ell, n)$ and $p = p(\ell, n)$ will be determined soon. Here and in the sequel, we write $\bar{q} = 1 - q$. Then an easy calculation yields

$$E_n := \frac{n!\, q^\ell\, \bar{q}^{n-\ell}\, p}{(\ell-1)!(n-\ell)!} \int_{-\frac{q}{p}}^{\frac{\bar{q}}{p}} m(\tau_n(v))(1+\frac{p}{q}v)^{\ell-1}(1-\frac{p}{\bar{q}}v)^{n-\ell}\, dv$$

where we used the abbreviation

$$\tau_n(v) := \frac{U(\frac{1}{q+pv}) - b_n}{a_n}$$

for convenience. The form of the integrand suggests to take $q = q(\ell, n) = \frac{\ell}{n}$ and $p^2 = p^2(\ell, n) = q\,\bar{q}\,n^{-1}$. We now follow the same approach as in [16,18]. Subdivide the integration in E_n over the three intervals $(-\frac{q}{p}, -T)$, $[-T, T]$ and $(T, \frac{\bar{q}}{p})$ where T is a fixed quantity. It is then easy to show that with the notations above, the central part can be controlled by the condition

(A): $\quad \tau_n(v) := \frac{U(\frac{1}{q+pv})-b_n}{a_n} \to \tau(v)$ uniformly on bounded v-intervals.

By taking T large enough, the two remaining pieces ultimately vanish since $\min(\frac{q}{p}, \frac{\bar{q}}{p}) \to \infty$. This then leads to the following result.

Lemma 2.1 *Under condition (A)*

$$(2) \qquad E\left\{m\left(\frac{X^*_{n-\ell+1} - b_n}{a_n}\right)\right\} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} m(\tau(v))\, e^{-\frac{1}{2}v^2}\, dv.$$

In the subsequent applications of the lemma we have the freedom of choosing the constants a_n and b_n in such a way that condition (A) is satisfied. The choice of b_n is usually automatic. If we put $v = 0$ in (A), then it is almost obvious that we should take $b_n = U(q^{-1}) = U(n/\ell)$. Then, the choice of a_n has to be made by requiring the convergence of

$$(3) \qquad \tau_n(v) = \frac{1}{a_n}\left\{U\left(\frac{1}{q+pv}\right) - U\left(\frac{n}{\ell}\right)\right\} \to \tau(v)$$

for $|v| \le T$ and T any positive constant. The choice of a_n will of course be determined by the limiting behaviour of the ratio ℓ/n.

Trying to understand the kind of condition needed on U, let us look at the quantity $q + pv$ which, by the choice of q and p, satisfies the relation

$$\frac{1}{q+pv} = \frac{n}{\ell}\left(1 - \sqrt{\frac{1-\frac{\ell}{n}}{\ell}}v + o\left(\sqrt{\frac{1-\frac{\ell}{n}}{\ell}}\right)\right)$$

when $\ell, n \to \infty$. Introduce this approximation in (3) to obtain

$$\tau_n(v) = \frac{1}{a_n}\left\{U\left(\frac{n}{\ell}\left(1-\sqrt{\frac{1-\frac{\ell}{n}}{\ell}}v+o\left(\sqrt{\frac{1-\frac{\ell}{n}}{\ell}}\right)\right)\right)-U\left(\frac{n}{\ell}\right)\right\} \tag{4}$$
$$\to \tau(v).$$

Let us compare this condition with the extremal condition (1). We then clearly can take $x = n/\ell$ which tends to ∞ by our assumptions. However, the fixed quantity y in (1) has to be replaced in (4) by a quantity that tends to 1 together with x. The resulting condition is discussed in the next section.

Apart from the case where $\ell \to \infty, \ell/n \to 0$, there are at least two other situations.

(*i*) First, ℓ could be taken fixed. Then results for fixed ℓ can be obtained if and only the same result can be found for $\ell = 1$. The extremal domain of attraction will play a predominant role. See for example [3,9,12,15].

(*ii*) If $\ell \to \infty$ but $\frac{\ell}{n} \to \lambda \in (0,1)$. Condition (4) can then be replaced by

$$\tau_n(v) = \frac{1}{a_n}\left\{U\left(\frac{1}{\lambda}(1-\sqrt{\frac{1-\lambda}{\lambda}}\frac{v}{\sqrt{n}}+o(\frac{1}{\sqrt{n}}))\right)-U\left(\frac{1}{\lambda}\right)\right\} \to \tau(v). \tag{5}$$

The latter condition is a differentiability condition of U in a neighbourhood of $\frac{1}{\lambda}$ and is classical in the theory of order statistics [7,15,17].

The condition that we need should be intermediate between conditions (1) and (5).

2.2 Transformation approach

Assume that Z has a standard exponential distribution and let $Z_1^* \leq Z_2^* \leq \ldots \leq Z_n^*$ be the order statistics of a sample of size n from Z. It is well known that for this specific distribution

$$\sqrt{\ell}\{Z_{n-\ell+1}^* - \log(\frac{n}{\ell})\} \xrightarrow{\mathcal{D}} Y \sim \mathcal{N}(0,1)$$

when $\ell \to \infty$ and $n - \ell \to \infty$. See for example [15, p.108] where the result is given for the equivalent case of uniform random variables. In order to transfer this asymptotic normality to a more general situation, we can identify X with $U(e^Z) := \phi(Z)$. This transfer function ϕ should then be approximately linear on intervals of size of order $\ell^{-1/2}$ around $\log(n/\ell)$. More explicitly, we need a condition of the form

$$\frac{\phi(x+t\delta(x))-\phi(x)}{\phi'(x)\delta(x)} \to t \tag{6}$$

when $\delta(x) \to 0$ for $x \to \infty$. We will transform (6) into a condition on U.

3 The class C^*

Motivated by the arguments above, we now introduce our working condition.

Definition 3.1 *Assume that the tail quantile function U is uniformly differentiable at ∞ with ultimately positive derivative u. The distribution F belongs to C^* iff for all $y \in \Re$*

$$\frac{U(x+yx\epsilon(x))-U(x)}{\epsilon(x)xu(x)} \to y \tag{7}$$

whenever $\epsilon(x) \to 0$ when $x \uparrow \infty$.

We first show how this condition emerges from the two approaches above.

(*i*) From the Helly-Bray approach. Take the definition of $\mathcal{C}_\gamma(g)$ and replace y by $1+y\epsilon(x)$ in (1). Then approximately

$$\frac{U(x+yx\epsilon(x))-U(x)}{\epsilon(x)g(x)} \sim \frac{(1+y\epsilon(x))^\gamma - 1}{\gamma\epsilon(x)}.$$

Expanding and taking limits on the right hand side yields y. Note that the quantity γ disappeared from the expression.

(*ii*) From the transformation approach. Replace $\phi(x)$ by $U(e^x)$ in (6), then (7) emerges naturally.

Before embarking on the applicability of condition (7), we make a number of remarks.

Remark 3.1 a. Note first that (7) is satisfied if $\log(xu(x))$ is uniformly continuous on a neighbourhood of ∞. Alternatively, $xu'(x)/u(x)$ is bounded on such a neighbourhood. The latter condition is satisfied if the distribution F has a density that satisfies a Von Mises condition $xu'(x)/u(x) \to c \in \Re$. As such, (7) slightly generalizes the conditions given by Falk [6]. Alternatively, look in [15, p.164].

Remark 3.2 b. The condition for C^* is equivalently transformed into a condition in terms of F itself. For such comparisons in general, see [3,9].

Proposition 3.1 *Assume that F has an ultimately positive density f. Then $F \in C^*$ iff for all $y \in \Re$*

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

whenever $\delta(x) \to 0$ for $x \uparrow x_+$.

Proof Choose $g(v) = vu(v)$ and write $-y_x$ for the expression on the left. Then

$$1 - F(x + y\delta(x)\frac{1 - F(x)}{f(x)}) = (1 - F(x))(1 - y_x\delta(x)).$$

Put $x = U(v)$ and $y_x = y'_v$. Then

$$U\left(\frac{v}{1 - y'_v\delta(U(v))}\right) - U(v) = y\,\frac{\delta(U(v))}{vf(U(v))}.$$

Now, define $\epsilon(v)$ by the equation $1 + y\epsilon(v) := \{1 - y'_v\delta(U(v))\}^{-1}$ and note that u and f are linked by the equality $vu(v) = (vf(U(v)))^{-1}$. Then

$$\frac{U(v + yv\epsilon(v)) - U(v)}{\epsilon(v)vu(v)} = y\frac{\delta(U(v))}{\epsilon(v)} \to y.$$

As all steps can be reversed, also the converse holds. ■

Remark 3.3 c. A useful implication of the condition $F \in C^*$ is given in the next result.

Lemma 3.1 *Let $x_n \to \infty$, $y_n \to y$ and $r_n \neq 0$, a sequence tending to 0. If $F \in C^*$ then for $n \to \infty$,*

$$\frac{U(x_n(1 + y_n r_n)) - U(x_n)}{r_n\, x_n\, u(x_n)} \to y. \tag{8}$$

Proof First assume that the sequence $\{y_n\}$ is constant and equal to y. Suppose now on the contrary that (8) does not hold. Then there exists a subsequence $\{x_n\}$ and a positive δ for which $x_{n+1} > x_n + 1$ and

$$\left|\frac{U(x_n(1 + y_n r_n)) - U(x_n)}{r_n\, x_n\, u(x_n)} - y\right| > \delta$$

for all n. Define $\epsilon(x) = r_n$ when $x_n \leq x < x_{n+1}$ and $\epsilon(x) = 0$ when $x < x_1$. As $r_n \to 0$ and $x_{n+1} > x_n + 1$, $\epsilon(.)$ is well defined and $\epsilon(x) \to 0$ as $x \to \infty$. Nevertheless

$$\left|\frac{U(x_n(1 + y\epsilon(x_n))) - U(x_n)}{\epsilon(x_n)\, x_n\, u(x_n)} - y\right| > \delta$$

for all n, leading to a contradiction with the definition of C^*. The sequence of increasing functions

$$f_n(y) := \frac{U(x_n(1 + yr_n)) - U(x_n)}{r_n\, x_n\, u(x_n)}$$

converges pointwise to the function $f(y) := y$. But then the convergence is uniform as follows from Polyá's extension of Dini's theorem [13]. Hence the result follows. ■

Remark 3.4 d. All of the results in the next sections can also be derived by following a transformation approach. To avoid duplication we only deal with the Helly-Bray procedure.

Remark 3.5 e. A link between the conditions (1) and (7) is given by the relation $g(x) = xu(x)$ which has already been used in **b.** above. In what follows both functions g and u will be used repeatedly.

4 One large order statistic

We illustrate the above concept first in the easiest possible situation, i.e. that of one order statistic close to the maximum. Recall a well-known weak law under the condition $F \in \mathcal{C}_\gamma(g)$. For then

$$\tau_n(v) = \frac{U\left(\frac{n}{\ell}(1 - \frac{v}{\sqrt{\ell}}(1 + o(1)))\right) - U(\frac{n}{\ell})}{g(\frac{n}{\ell})} \to 0 = \tau(v).$$

But then by lemma 1, $g^{-1}(n/\ell)(X^*_{n-\ell+1} - U(n/\ell)) \xrightarrow{P} 0$. Note that in the case $\gamma > 0$ we can go one step further. Since $\frac{U(x)}{g(x)} \to \gamma^{-1}$, we also have $g^{-1}(n/\ell)\ X^*_{n-\ell+1} \xrightarrow{P} \gamma^{-1}$. When $\gamma < 0$, we find similarly that $g^{-1}(n/\ell)(x_+ - X^*_{n-\ell+1}) \xrightarrow{P} -\gamma^{-1}$.

A key point for introducing the class $\mathcal{C}^*$ is illustrated in the following result, which specifies the speed of convergence in the above weak law. Because of its basic importance, we formulate the result in the form of a theorem.

Theorem 4.1 *Let $F \in \mathcal{C}^*$. If $\ell, n \to \infty$ such that $\frac{\ell}{n} \to 0$, then*

$$\ell^{3/2}\frac{X^*_{n-\ell+1} - U(\frac{n}{\ell})}{nu(n/\ell)} \xrightarrow{\mathcal{D}} Y \sim \mathcal{N}(0,1).$$

Proof Look anew at $\tau_n(v)$ above. Then with $x = n/\ell$ and $\epsilon(x) = \ell^{-\frac{1}{2}}$ the condition $F \in \mathcal{C}^*$ shows that $\tau_n(v) \to -v$. By the symmetry of a normal random variable the result follows. ∎

The above result is precisely of the form deduced by Falk [6] under the traditional Von Mises conditions. See also Reiss [15]. We can expect that the speed of convergence in the above result might be very slow. For fixed ℓ the limit law for $n \to \infty$ is linked to the classical extreme value distribution while for $\ell \to \infty$ and $\ell/n \to 0$ we get a very different distribution in the normal law.

5 Two large order statistics

We turn to the case of $X^*_{n-s-t+1}$ and X^*_{n-s+1} where s and t are two integers both converging to infinity. From lemma 1 we know that we have to use a specific normalisation inspired by a reduction of the kernel of the integrand to a bivariate normal density.

5.1 General approach

As in section 2.1 we start from the expression for the joint expectation. Then with

$$E_n := E\left\{m\left(\frac{X^*_{n-s-t+1}-b_n}{a_n}, \frac{X^*_{n-s+1}-b'_n}{a'_n}\right)\right\}$$

it is easy to show that

$$(9)\quad E_n = c^{(n)}_{s,s+t}\int_0^1 a^{s+t-1}(1-a)^{n-s-t}\int_0^1 b^{s-1}(1-b)^{t-1}m_n(a,b)\,db\,da$$

where in general

$$c^{(n)}_{k,m} := \frac{n!}{(n-m)!(k-1)!(m-k-1)!}$$

and where

$$m_n(a,b) := m\left(\frac{U(\frac{1}{a})-b_n}{a_n}, \frac{U(\frac{1}{ab})-b'_n}{a'_n}\right)$$

keeps all the references to the original distribution.

We make the change of variables $a = q + pu$ and $b = q' + p'v$ where q' and p' are determined as before. We easily find the identifications

$$q = \frac{s+t}{n},\ q' = \frac{s}{s+t},\ p^2 = \frac{q(1-q)}{n},\ p'^2 = \frac{q'(1-q')}{t+s}.$$

With these choices the remaining steps are easy when deriving possible asymptotic distributions. We can formulate an auxiliary result which is of the same form as lemma 1. Actually, the proof is a bivariate version of that of lemma 1.

Lemma 5.1 *Assume $n, s, t \to \infty$ and let m be continuous and bounded on $\Re_2$. If, with the above choices of q, p, q' and p',*

$$\tau_n(u) := \frac{1}{a_n}\left(U((q+pu)^{-1}) - b_n\right) \to \tau(u)$$

and

$$\sigma_n(u,v) := \frac{1}{a'_n}\left(U((q+pu)^{-1}(q'+p'v)^{-1}) - b'_n\right) \to \sigma(u,v)$$

for appropriate constants a_n, b_n, a'_n and b'_n, and uniformly on bounded u and v intervals, then

$$E\left\{m\left(\frac{X^*_{n-s-t+1}-b_n}{a_n},\frac{X^*_{n-s+1}-b'_n}{a'_n}\right)\right\}$$
$$\to \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{1}{2}(u^2+v^2)}m\left(\tau(u),\sigma(u,v)\right)\,dv\,du.$$

We apply the lemma under the condition that $F\in\mathcal{C}^*$. We assume from the outset that $\frac{t+s}{n}\to 0$. To deal with $\tau_n(u)$, compare with (4). Take $b_n = U(\frac{n}{s+t})$ and $a_n = g(\frac{n}{s+t})/\sqrt{s+t}$; then apply lemma 2 with the choice $x_n = \frac{n}{s+t}, r_n = (s+t)^{-1/2}$ and $y_n = -u + o(1)$. It then easily follows that $\tau_n(u)\to -u$. For $\sigma_n(u,v)$, the situation is more complicated. The argument of U equals

$$\begin{aligned}
&\frac{1}{(q+pu)(q'+p'v)}\\
&= \frac{n}{s+t}\left(1-\frac{u}{\sqrt{s+t}}+o(\frac{1}{\sqrt{s+t}})\right)\frac{s+t}{s}\times\\
&\qquad\times\left(1-\sqrt{\frac{t}{s(s+t)}}v+o\left(\sqrt{\frac{t}{s(s+t)}}\right)\right)\\
&= \frac{n}{s}\left(1-\frac{u}{\sqrt{s+t}}-\sqrt{\frac{t}{s(s+t)}}v+o\left(\frac{1}{\sqrt{s+t}},\sqrt{\frac{t}{s(s+t)}}\right)\right).
\end{aligned}$$

It therefore seems natural to assume that while $s,t\to\infty$, their ratio $\frac{t}{s}\to\theta\in[0,\infty)$. With this choice, we can apply lemma 2. Take $b'_n = U(n/s)$, $x_n = n/s$, $r_n = s^{-1/2}$, $a'_n = g(n/s)/\sqrt{s}$ and $y_n = -(u+\sqrt{\theta}v)(1+\theta)^{-1/2}+o(1)$. We then easily find that $\sigma(u,v) = -(u+\sqrt{\theta}v)(1+\theta)^{-1/2}$. Introducing all of this and $g(v) = vu(v)$ in the expression E_n we arrive at the following result.

Theorem 5.1 *Assume $F\in\mathcal{C}^*$. Assume $n,s,t\to\infty$, $s/n\to 0$ and $t/s\to\theta\in[0,\infty)$. Then*

$$\left((s+t)^{3/2}\frac{X^*_{n-s-t+1}-U(\frac{n}{s+t})}{nu(\frac{n}{s+t})},\, s^{3/2}\frac{X^*_{n-s+1}-U(\frac{n}{s})}{nu(\frac{n}{s})}\right)\xrightarrow{\mathcal{D}}(V,W)$$

where (V,W) has a bivariate normal distribution with zero means, unit variances and correlation coefficient $\rho_{V,W} = \sqrt{\frac{1}{1+\theta}}$.

5.2 Spacings

We put $m(a,b) = k(b-a)$ for some continuous and bounded k in the general formula (10). Further, we define the *general t − spacing* by

$$T^{(t)}_{n,s} := X^*_{n-s+1} - X^*_{n-s-t+1}.$$

We get the following expression for the normalized expectation

$$E\left\{k\left(\frac{T_{n,s}^{(t)}}{a_n}\right)\right\} = c_{s,s+t}^{(n)} \int_0^1 a^{t+s-1}(1-a)^{n-s-t}$$
$$\times \int_0^1 b^{s-1}(1-b)^{t-1} k\left(\frac{U(\frac{1}{ab}) - U(\frac{1}{a})}{a_n}\right) db\, da.$$

We assume as before that $s \to \infty$ such that $s/n \to 0$ but that t remains fixed. The appropriate substitutions are now $a = q + pu$ with $q = s/n, p^2 = sn^{-2}$ and $b = 1-(v/s)$. With these choices we need the convergence for appropriate a_n of

$$\tau_n(u,v) := a_n^{-1}\left\{U\left(\frac{1}{(q+pu)(1-\frac{v}{s})}\right) - U\left(\frac{1}{q+pu}\right)\right\} \to \tau(u,v)$$

so as to have that

$$E\left\{k\left(\frac{T_{n,s}^{(t)}}{a_n}\right)\right\} \to \frac{1}{\sqrt{2\pi}}\frac{1}{(t-1)!}\int_{-\infty}^{\infty} e^{-\frac{u^2}{2}} \int_0^{\infty} e^{-v}v^{t-1}k(\tau(u,v))\, dv\, du.$$

The natural choice for the convergence of $\tau_n(u,v)$ is of course that $F \in \mathcal{C}^*(g)$. For then the norming constant $a_n = \frac{1}{s}g(\frac{n}{s})$ while $r_n = s^{-1}$. With this choice, $\tau(u,v) = v$ and we obtain the following result

Theorem 5.2 *Assume $F \in \mathcal{C}^*$. Assume $n, s \uparrow \infty$, $s/n \to 0$ and t fixed. Then*

$$\frac{s^2\left(X_{n-s+1}^* - X_{n-s-t+1}^*\right)}{nu(\frac{n}{s})} \xrightarrow{\mathcal{D}} Z$$

where Z has the gamma density $f_Z(v) = \frac{1}{(t-1)!}e^{-v}v^{t-1}$.

6 Symmetric distributions

To show the flexibility of the condition $F \in \mathcal{C}^*$, we give an unusual application to possible centering and dispersion measures for symmetric distributions.

6.1 Preparation

Symmetric distributions have always been of great importance in statistics. It is not surprising that a number of theoretical results can be derived for this specific situation. In this section we deal with some of them as adaptations of results derived by Gumbel [8]. We assume that we are dealing with a *symmetric distribution* for which $F(-x) = 1 - F(x)$ for all x or equivalently for which the tail quantile function satisfies $U(v) = -U(\frac{v}{v-1})$ for $v > 1$.

From the general formula (10), one easily sees that for continuous and bounded m

$$
\begin{aligned}
E\left\{m(X_r^*, X_{n-r+1}^*)\right\} = c_{r,n+1-r}^{(n)} \int_0^1 a^{n-r}(1-a)^{r-1} \\
\times \int_0^1 b^{r-1}(1-b)^{n-2r} m(U(\frac{1}{a}), U(\frac{1}{ab}))\, db\, da.
\end{aligned} \tag{10}
$$

In the formula n and r are the remaining indexes.

Under the extremal domain of attraction condition (1), the case where r is fixed can be treated as in [8] where $r = 1$ had been taken. If we now allow r to go to ∞ with n but such that $\frac{r}{n} \to 0$, then we get a new situation. From the above formula we derive that for any continuous and bounded k,

$$
\begin{aligned}
& Ek\left\{\frac{1}{a_n}\left(m(X_r^*, X_{n-r+1}^*) - b_n\right)\right\} \\
= \; & c_{r,n+1-r}^{(n)} \int_0^1 a^{n-r}(1-a)^{r-1} \\
& \int_0^1 b^{r-1}(1-b)^{n-2r} k\left(\frac{m(U(\frac{1}{a}), U(\frac{1}{ab})) - b_n}{a_n}\right) db\, da.
\end{aligned}
$$

Both integrands have to be treated. So, put $a = q + pu$ and $b = q' + p'v$. Then it easily follows that $\overline{q} = q' = r/n$ and $p^2 = p'^2 = rn^{-2}$. With the usual calculations and since $U(\frac{1}{a}) = -U(\frac{1}{1-a})$ we arrive at a function $\tau_n(u,v)$ as before.

Lemma 6.1 *Let $F \in C^*$ be symmetric. Assume $n, r \to \infty$ and let k be continuous and bounded on $\Re_2$. If, with the choices $\overline{q} = q' = r/n$ and $p^2 = p'^2 = rn^{-2}$, $\tau_n(u,v)\tau_n(u,v) :=$*

$$
\frac{1}{a_n}\left\{m\left(-U(\frac{1}{\overline{q} - pu}), U(\frac{1}{(q+pu)(\overline{q}' + p'v)})\right) - b_n\right\} \to \tau(u,v) \tag{11}
$$

for bounded u and v, with b_n and a_n centering and normalising constants, then

$$
a_n^{-1}\left\{m(X_r^*, X_{n-r+1}^*) - b_n\right\} \xrightarrow{\mathcal{D}} Y
$$

where for bounded and continuous k

$$
E(k(Y)) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} k(\tau(u,v))\, e^{-\frac{1}{2}(u^2+v^2)}\, dv\, du.
$$

We give two examples of the above result. We could prove the asymptotic independence of the two composing order statistics first and deduce the results by a continuous mapping argument. We use a direct approach that again shows the use of condition (7).

6.2 The pen-ultimate extremal quotient

The *pen-ultimate extremal quotient* is defined for symmetric distributions by the expression $Q_n := X^*_{n-r+1}/(-X^*_r)$. Actually, one should take the absolute value, but since the largest order statistics will tend to x_+ a.s., we omit the absolute value sign. The concept of extremal quotient has been studied for the case $r = 1$ in [8]. For general r the formula for the expectation is easily obtained by the choice $m(a,b) = -b/a$ in (10) and (11). We prove the following.

Theorem 6.1 *Assume that F is symmetric, continuous and belongs to C^*. If $r/n \to 0$, then*

$$\sqrt{\frac{r^3}{2}\frac{U(\frac{n}{r})}{nu(\frac{n}{r})}}\left(-\frac{X^*_{n-r+1}}{X^*_r}-1\right) \xrightarrow{\mathcal{D}} Y \sim \mathcal{N}(0,1).$$

Proof We need to determine the constants and the resulting limit. Clearly, $b_n = 1$ in lemma 4. We then turn to the quantity a_n. Here

$$\tau_n(u,v) = \frac{1}{a_n}\frac{U\left(\frac{1}{(q+pu)(q'+p'v)}\right) - U\left(\frac{1}{\bar{q}-pu}\right)}{U\left(\frac{1}{\bar{q}-pu}\right)} = \frac{g(x_n)}{a_n U(x_n)}\frac{U(x_n z_n) - U(x_n)}{g(x_n)}$$

where

$$x_n = \frac{1}{\bar{q}-pu} = \frac{n}{r}\left(1 + \frac{u}{\sqrt{r}} + o(\frac{1}{\sqrt{r}})\right)$$

and

$$z_n = \frac{\bar{q}-pu}{(q+pu)(q'+p'v)} = 1 - \frac{u+v}{\sqrt{r}} + o(\frac{1}{\sqrt{r}}).$$

As the latter is of the form $z_n = 1 + r_n y_n$ with $r_n = \frac{1}{\sqrt{r}}$ and $y_n = -(u + v) + o(1)$, we can apply lemma 2 to the major portion of the expression for $\tau_n(u,v)$. Choosing $a_n = \sqrt{\frac{2}{r}\frac{g(n/r)}{U(n/r)}}$ we see that

$$\tau_n(u,v) \to \tau(u,v) = -\frac{u+v}{\sqrt{2}}.$$

By lemma 4, we then get that

$$\sqrt{\frac{r}{2}\frac{U(\frac{n}{r})}{g(\frac{n}{r})}}\left(-\frac{X^*_{n-r+1}}{X^*_r}-1\right) \xrightarrow{\mathcal{D}} Y$$

where

$$\begin{aligned} E(k(Y)) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} k(-\frac{u+v}{\sqrt{2}})\, e^{-\frac{1}{2}(u^2+v^2)}\, dv\, du \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} k(x)\, e^{-\frac{x^2}{2}}\, dx \end{aligned}$$

and hence Y follows a $\mathcal{N}(0,1)$ distribution. ∎

A similar result can be derived for the *pen-ultimate midrange* which is defined by the quantity $\frac{1}{2}(X^*_{n-r+1} + X^*_r)$.

6.3 The pen-ultimate geometric mean

This quantity is also introduced by Gumbel [8] for the case $r = 1$ as the geometric mean of the two opposite r−th largest order statistics. Put $m(a,b) = \sqrt{-ab}$ in (10) and (11). As most of the results follow from previous arguments, we omit the details.

Theorem 6.2 *Assume that F is symmetric, continuous and belongs to C^*. Then if $r \to \infty$ and $r/n \to 0$, then*

$$\frac{\sqrt{2r^3}}{nu(\frac{n}{r})}\left(\sqrt{X^*_{n-r+1}(-X^*_r)} - U(\frac{n}{r})\right) \xrightarrow{\mathcal{D}} Y \sim \mathcal{N}(0,1).$$

Acknowledgements. The author thanks Jan Beirlant, Goedele Dierckx and Stefan Vaes for many helpful suggestions. Even more thanks go to the editors and the referees whose comments have been crucial in upgrading an earlier version of the paper.

REFERENCES

[1] Beirlant, J. and Teugels, J.L. (1986). Asymptotics of Hill's estimator. *Teor. Verojatnost. i Primenen.*, **31**, 530–536.

[2] Beirlant, J. and Teugels, J.L. (1996). A simple approach to classical extreme value theory. In *Exploring Stochastic Laws, Festschrift for Academician V.S. Korolyuk*, 457–468, VSP, Zeist.

[3] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1987). *Regular Variation*, Cambridge University Press, Cambridge.

[4] Cheng, S., de Haan, L. and Xin Huang (1997). Rate of convergence for intermediate order statistics. *J. Theoretical Probability*, **10**, 1–23.

[5] Chibisov, D.M. (1964). On limit distributions for order statistics. *Theory Probability Appl.*, **9**, 150–165.

[6] Falk, M. (1989). A note on uniform asymptotic normality of intermediate order statistics. *Ann. Inst. Statist. Math.*, **41**, 19–29.

[7] Falk, M. (1998). Local asymptotic normality of intermediate and central order statistics. *Comm. Stat. Th. Meth.*, **27**, 1729–1755.

[8] Gumbel, E.J. (1958). *Statistics of Extremes*, Columbia University Press, New York.

[9] Haan, L. de (1970). *On Regular Variation and its Applications to the Weak Convergence of Sample Extremes*, Math. Centre Tract, **32**, Amsterdam.

[10] Haan, L. de (1993). A unified criterion for the domain of attraction of extreme-value distributions, Technical Report 9931/A, Erasmus University Rotterdam.

[11] Laha, R.G. and Rohatgi, V.K. (1979). *Probability Theory*, John Wiley & Sons, New York.

[12] Leadbetter, M.R., Lindgren, G. and Rootzén, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*, Springer Series in Statistics, New York, Springer.

[13] Pólya, G. and Szegö, G. (1979). *Problems and Theorems in Analysis*, II, Springer-Verlag, Berlin.

[14] Mason, D. and van Zwet, W. (1987). A refinement of the KMT inequality for the uniform empirical process. *Annals of Statistics*, **18**, 871–884.

[15] Reiss, R.-D. (1989). *Approximate Distributions of Order Statistics*, Springer Series in Statistics, New York, Springer.

[16] Smirnov, N.V. (1962). Limit distributions for the terms of a variational series. *Translations Amer. Math. Soc.*, **11**, 82–143.

[17] Smirnov, N.V. (1967). Some remarks on limit laws for order statistics. *Theory Probability Appl.*, **12**, 337–391.

[18] Teugels, J.L. (1981). Limit theorems on order statistics. *Ann.Probability*, **9**, 868–880.

J. Teugels
University Centre for Statistics
Katholieke Universiteit Leuven
DeCroylaan, 52B, 3001 Heverlee
Belgium
jef.teugels@wis.kuleuven.ac.be