Aio cursus Einmahl / de Haan
17 december 1999

1 Extreme value theory: basics

Suppose $X_1, X_2, \ldots,$ i.i.d. with distribution function F.

Suppose that there exist sequences of numbers $a_n > 0$ and b_n such that the distribution of

$$\frac{\max(X_1, X_2, \ldots, X_n) - b_n}{a_n}$$

converges to a non-degenerate distribution, i.e.

$$\lim_{n\to\infty} F^n(a_n x + b_n) = G(x), \tag{1}$$

with G a non-degenerate distribution function, for all except countably many x. We discuss two issues:

1. What is the class of distribution functions G that can occur in the limit relation (1)?
2. For each one of those G, what are the conditions on F such that (1) holds? (i.e. domain of attraction theory).

(i) Preliminary calculation.
Take logs in (1):

$$\lim_{x\to\infty} -n \log F(a_n x + b_n) = -\log G(x)$$

and for $0 < G(x) < 1$, since then we have $F(a_n x + b_n) \to 1$,

$$\lim_{n\to\infty} n\{1 - F(a_n x + b_n n)\} = -\log G(x) \tag{2}$$

or,

$$\lim_{n\to\infty} -\log\{1 - F(a_n x + b_n)\} - \log n = -\log\{-\log G(x)\}. \tag{3}$$

Now we go to inverse functions. Write

$$V(x) := \left(\frac{1}{1-F}\right)^{\leftarrow}(e^x).$$

In (3) we have a sequence of non-decreasing functions of x converging to some limit function. This is equivalent to the sequence of inverse functions converging to the inverse of the limit functions.
Hence (3) is equivalent to

$$\lim_{n\to\infty} \frac{V(x+\log n)-b_n}{a_n} = (-\log\{-\log G\})^{\leftarrow}(x)$$

for all but countably many x. Clearly this implies

$$\lim_{n\to\infty} \frac{V(x+\log n)-V(\log n)}{a_n}$$
$$= (-\log\{-\log G\})^{\leftarrow}(x) - (-\log\{-\log G\})^{\leftarrow}(0) =: \Psi(x)$$

Some simple but tedious inequalities show that in fact we can take the running variable n going through the reals rather than the integers. Hence for some positive function $a(t)$

$$\lim_{t\to\infty} \frac{V(t+x)-V(t)}{a(t)} = \Psi(x) \tag{4}$$

for all but countably many x. We are now able to attack the two problems stated in the beginning.

(ij) The limit distribution

In order to find the possible limit distributions in (1), we consider the limit function Ψ in (4).
By writing

$$\frac{V(t+x+y)-V(t)}{a(t)} = \frac{V(t+x+y)-V(t+y)}{a(t+y)} \cdot \frac{a(t+y)}{a(t)} + \frac{V(t+y)-V(t)}{a(t)} \tag{5}$$

and choosing x, y and $x+y$ such that for those values the limit in (4) is valid, we find that

$$\lim_{t\to\infty} \frac{a(t+y)}{a(t)} =: A(y)$$

must exist, positive, for all but countably many y. Then from (5) we get

$$\Psi(x+y) = A(y)\Psi(x) + \Psi(y) \tag{6}$$

and, with x and y interchanged,

$$\Psi(x+y) = A(x)\Psi(y) + \Psi(x).$$

Hence

$$\Psi(x)(1-A(y)) = \Psi(y)(1-A(x)). \tag{7}$$

There are two cases:

1. $A(t) \equiv 1$. Then by (6) and the monotonicity of Ψ we obtain for some $c > 0$ and d

   $$\Psi(x) = cx + d.$$

   This leads to

   $$G(x) = \exp\left(-\exp\{(x-d)/c\}\right) \quad \text{for all real } x.$$

2. There exists x_0 with $A(x_0) \neq 1$. I claim that then $A(x) \neq 0$ for all $x \neq 0$. Suppose by way of contradiction that $A(x_1) = 1$ for some $x_1 \neq 0$, then by (7) $\Psi(x_1) = 0$, hence by (6) $\Psi(x_1 + y) = \Psi(y)$ for all y which is impossible since Ψ cannot be constant.
   Hence $A(x) \neq 1$ for all $x \neq 0$ and then by (7) for some constant $c_1 \neq 0$

   $$\Psi(x) = c_1\left(1 - A(x)\right).$$

   Substitution in (6) gives

   $$1 - c_1^{-1}\Psi(x+y) = \left(1 - c_1^{-1}\Psi(x)\right)\left(1 - c_1^{-1}\Psi(y)\right)$$

   which implies as before

   $$1 - c_1^{-1}\Psi(x) = e^{\gamma x}$$

   for some real constant $\gamma \neq 0$. We conclude that

   $$\begin{aligned} \Psi(x) &= c_1(1 - e^{\gamma x}) \\ &= c\gamma^{-1}(e^{\gamma x} - 1) \end{aligned}$$

   with $c > 0$ in view of the monotonicity of Ψ. Hence

   $$G(x) = \exp - \left(1 + \gamma \frac{x-d}{c}\right)^{1/\gamma}$$

   with $c > 0$ and d real constants. This formula holds for $0 < G(x) < 1$ i.e. for $1 + \gamma\frac{x-d}{c} > 0$. In summary: apart from shift and scale constants we have a parametric family of possible limit functions:

   $$G_\gamma(x) := \exp -(1+\gamma x)^{-1/\gamma} \quad \text{for } \gamma \in \mathbb{R}$$

   (for $\gamma = 0$ interpret $(1+\gamma x)^{1/\gamma}$ as its limit for $\gamma \to 0 : e^{-x}$). These distributions are called extreme–value distributions.

(iij) Domains of attraction

Take a fixed $\gamma \in \mathbb{R}$. We want to give necessary and sufficient conditions on F such that

$$\lim_{n\to\infty} F^n(a_n x + b_n) = G_\gamma(x)$$

for all $x \in \mathbb{R}$ and appropriate sequences $a_n > 0$ and b_n (i.e. such that $F \in D(G_\gamma)$, the domain of attraction of G_γ). We start from (4): with $V(x) := \left(\frac{1}{1-F}\right)^{\leftarrow}(e^x)$ we have

$$\lim_{t\to\infty} \frac{V(t+x) - V(t)}{a(t)} = \Psi(x) = \frac{e^{\gamma x} - 1}{\gamma} \quad \text{for all } x. \tag{8}$$

Define for a distribution function F the endpoint $x^* = \sup\{x | F(x) < 1\}$.

Lemma 1.1. *1. For $\gamma > 0$ (8) is equivalent to*

$$\lim_{t\to\infty} \frac{V(t+x)}{V(t)} = e^{\gamma x} \quad \textit{for all } x.$$

and this in turn is equivalent to

$$\lim_{t\to\infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-1/\gamma} \quad \textit{for } x > 0.$$

(i.e. the function $1 - F$ is regularly varying of index $-1/\gamma$).

2. For $\gamma < 0$ (8) is equivalent to:

$$\lim_{t\to\infty} V(t) \quad \textit{is finite; call this limit } V(\infty);$$

and

$$\lim_{t\to\infty} \frac{V(\infty) - V(t+x)}{V(\infty) - V(t)} = e^{\gamma x} \quad \textit{for all } x$$

and this in turn is equivalent to:

$$\lim_{t\downarrow 0} \frac{1 - F(x^* - tx)}{1 - F(x^* - x)} = x^{-1/\gamma}$$

(i.e. the function $1 - F(x^* - 1/x)$ is regularly varying of index $1/\gamma$).

(iv) Sufficient conditions

Theorem 1.2. *1. Suppose F is differentiable.*

If for $\gamma > 0$

$$\lim_{t \to infty} \frac{xF'(x)}{1-F(x)} = 1/\gamma,$$

then $F \in D(G_\gamma)$.

If $x^*(F) < \infty$ and for $\gamma < 0$

$$\lim_{x \to x^*} \frac{(x^* - x)F'(x)}{1-F(x)} = -1/\gamma,$$

then $F \in D(G_\gamma)$.

2. Suppose F is twice differentiable. Now $x^*(F)$ may be finite or infinite. If

$$\lim_{x \uparrow x^*} \frac{F''(x)(1-F(x))}{(F'(x))^2} = -1,$$

then $F \in D(G_o)$.

Remark 1.3. *These conditions are rather close to the necessary conditions.*

Remark 1.4. *If $F \in D(G_o)$, then $E(X_+)^\alpha < \infty$ for all $\alpha > 0$. If $F \in D(G_\gamma)$ with $\gamma > 0$, then $x^*(F) = \infty$ and $E(X_+)^\alpha$ is finite for $\alpha < 1/\gamma$ and infinite for $\alpha > 1/\gamma$. If $F \in D(G_\gamma)$ with $\gamma < 0$, then $x_F^* < \infty$.*

Remark 1.5. *If $F \in D(G_\gamma)$, $\gamma \leq 0$ and $x^*(F) < \infty$ then $F(x^* - \frac{1}{x}) \in D(G_{-\gamma})$. If $F \in D(G_\gamma), \gamma \geq 0$ and $x^*(F) = \infty$, then for any $b : F(\frac{1}{b-x}) \in D(G_{-\gamma})$.*

Examples

Check via the theorem Cauchy's distribution ($\gamma = 1$), student $-n$ ($\gamma = n$); exponential, gamma and normal ($\gamma = 0$), uniform ($\gamma = -1$), beta (m, n) ($\gamma = -1/m$), arcsin ($\gamma = -2$). Many discrete distributions satisfy the conditions, but not the geometric one and not Poisson's distribution. An example of a continuous distribution function not satisfying the conditions, is $F(x) = 1 - \exp\{-x + \sin x\}$ for $x > 0$.

It should be clear that probability distributions with quite different characters satisfy the conditions.

Now I turn to the use of these conditions in statistics. In order to do so, we shall reformulate the results in a suitable way.

(v) Alternative formulation; estimation of a small set

We know that $F \in D(G_\gamma)$ if and only if

$$\lim_{t\to\infty} \frac{V(t+x)-V(t)}{a(t)} = \frac{e^{\gamma x}-1}{\gamma}$$

for all x where

$$V(x) = \left(\frac{1}{1-F}\right)^{\leftarrow}(e^x) = F^{\leftarrow}(1-e^{-x}),$$

or, equivalently,

$$\lim_{t\to\infty} \frac{U(tx)-U(t)}{a(t)} = \frac{x^\gamma-1}{\gamma} \tag{9}$$

for all x with a **different** function a and

$$U(x) := \left(\frac{1}{1-F}\right)^{\leftarrow}(x) = F^{\leftarrow}(1-\frac{1}{x}).$$

Now upon replacing t by $\frac{1}{1-F(s)}$ in (9) and using an asymptotic continuity property for U, we get

$$\lim_{s\uparrow x^*} \frac{U\left(x\cdot\frac{1}{1-F(s)}\right)-s}{a\left(\frac{1}{1-F(s)}\right)} = \frac{x^\gamma-1}{\gamma}.$$

Inverting this relation as before we get

$$\lim_{s\uparrow x^*} \frac{1}{1-F\left(s+xa(\frac{1}{1-F(s)})\right)} \Big/ \frac{1}{1-F(s)} = (1+\gamma x)^{1/\gamma}$$

i.e.

$$\lim_{s\uparrow x^*} \frac{1-F(s+xa(s))}{1-F(s)} = (1+\gamma x)^{-1/\gamma}. \tag{10}$$

A liberal interpretation of (10) is the following: suppose $v >> s$, then

$$1-F(v) \approx \{1-F(s)\}\cdot\left(1+\gamma\frac{v-s}{a(s)}\right)^{-1/\gamma}. \tag{11}$$

Now look at the following problem: here is a sample of 1759 i.i.d. observations

```
* Vlissingen            winters 1882 t/m 1991
* Daan S4-selectie op hoogwateropzet ( min. 30 cm )
* Stormseizoen 1 oktober t/m 15 maart
* Resp. datum ( JJJJMMDD ), tijd ( UUMM , in MET ),
* stand hoogwater in cm + NAP, stand herleid naar
* 1992 in cm + NAP, en hoogwateropzet in cm.
* Trend rijzing gemiddeld hoogwater is 0,3129 cm/jaar.
* Ontwerppeil: 540 cm. Aantal stormen: 1759.
*                    vvv
18811015  605  256  291  138
18811028 1710  217  252   31
18811113 1840  174  209   34
18811117 2305  230  265   83
18811121  110  228  263   38
18811126  430  241  276   57
18811128 1910  214  249   54
18811202 1100  194  229   38
```

from some unknown distribution and we want to know the probability of having a value above 5.40m. Now the largest value of all the observations is 4.67m. So we want to estimate the distribution function beyond the range of the observations. Then formula (10) can provide an estimator (since (10) tells us how to extrapolate) provided we can estimate γ and the scale function a.

Next I discuss the dual problem starting from another reformulation.

(vi) Alternative formulation: estimation of extreme quantile

Relation (9) reads

$$\lim_{t\to\infty} \frac{F^{\leftarrow}(1-1/(tx)) - F^{\leftarrow}(1-1/t)}{a(t)} = \frac{x^\gamma - 1}{\gamma},$$

i.e. for $0 < p << q$

$$F^{\leftarrow}(1-p) \approx F^{\leftarrow}(1-q) + a(\frac{1}{q}) \cdot \frac{\left(\frac{q}{p}\right)^\gamma - 1}{\gamma}. \tag{12}$$

With the same sample of 2000 i.i.d. observations, suppose that we want to know the $1-10^{-6}$ quantile of the distribution (i.e. $p = 10^{-6}$). Then

(12) offers an opportunity: take $q = \frac{k}{n}$ with $k < n$, then $F^{\leftarrow}(1-q)$ can be estimated using the inverse empirical distribution function (i.e. by the $(n-k)$-th order statistic) and we can estimate $F^{\leftarrow}(1-p)$ provided we have estimators for γ and the scale function a.

So let us see how we can estimate γ and the scale function.

(vij) Estimator, γ positive

For simplicity we start with the case $\gamma > 0$.
There are two ways:

a. Start from

$$\lim_{t\to\infty} \frac{U(tx) - U(t)}{a(t)} = \frac{x^\gamma - 1}{\gamma}$$

which for $\gamma > 0$ simplifies to

$$\lim_{t\to\infty} \frac{U(tx)}{U(t)} = x^\gamma.$$

It is useful to take logs:

$$\lim_{t\to\infty} \log U(t/x) - \log U(t) = -\gamma \log x \tag{13}$$

with $U(t) := F^{\leftarrow}(1 - 1/t)$. We want to give a sample version of this relation. Now, upon replacing t by $\frac{n}{k}$ as before and replacing F by the empirical distribution function F_n, we have

$$U_n(\frac{1}{x}) := F_n^{\leftarrow}(1-x) = X_{n(1-x)n} = X_{n-xn,n}$$

(writing $X_{s,n}$ for $X_{[s],n}$)
and hence

$$U_n(\frac{n}{k}) = X_{n-k,n}$$

$$U_n(\frac{n}{kx}) = X_{n-kx,n}.$$

So the left hand side of (13) becomes for $0 < x < 1$

$$\log X_{n-kx,n} - \log X_{n-k,n} \tag{14}$$

Now in (13) we are allowed to integrate: (13) implies

$$\gamma = -\gamma \int_0^1 \log x dx = \lim_{t\to\infty} \int_0^1 \{\log U(t/x) - \log U(t)\} dx.$$

So a possible estimator for γ is obtained by integrating (14):

$$\begin{aligned} H_{k,n} : &= \int_0^1 (\log X_{n-kx,n} - \log X_{n-k,n}) dx \qquad (15) \\ &= \frac{1}{k} \sum_{i=0}^{k-1} \log X_{n-i,n} - \log X_{n-k,n}, \qquad (16) \end{aligned}$$

the latter equality is obtained since the integrand is constant on intervals of length $\frac{1}{k}$. This is the well-known estimator of B. Hill (1975, Annals of Statistics).
It is useful to look at an alternative derivation.

b. Start from

$$\lim_{t\to\infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-1/\gamma} \quad \text{for } x > 0.$$

Again we are allowed to integrate: for $t \to \infty$.

$$\begin{aligned} \gamma &= \int_q^\infty x^{1/\gamma} \frac{dx}{x} \longleftarrow \int_1^\infty \frac{1 - F(tx)}{1 - F(t)} \frac{dx}{x} \\ &= \frac{1}{1 - F(t)} \int_t^\infty (1 - F(s)) \frac{ds}{s} = \frac{1}{1 - F(t)} \int_t^\infty (\log s - \log t) dF(s). \end{aligned}$$

Replace in the latter expression F by the empirical distribution F_n and t by the order statistic $X_{n-k,n}$, then by the discrete character of F_n the integral becomes a sum and we get the same estimator

$$H_{k,n} := \frac{1}{k} \sum_{i=0}^{k-1} \log X_{n-i,n} - \log X_{n-k,n}.$$

Theorem 1.6. *(D. Mason, 1982, Annals of Statistics) If $F \in (G_\gamma)$ with $\gamma > 0$, we have*

$$H_{k,n} \xrightarrow{P} \gamma$$

provided $k = k(n) \to \infty$, $k(n)/n \to 0 (n \to \infty)$.

Remark 1.7. *This is a somewhat usual and easy-to-explain estimator. There are others. A similar remark applies to what follows.*

Now we proceed to construct an estimator for γ without the restriction $\gamma > 0$. In fact it can be shown that Hill's estimator tends to zero if $\gamma < 0$ (provided of course $x^*(F) > 0$) so that we only have to estimate $\gamma_- := \min(0, \gamma)$.

(viij) Estimator, general γ.

It is not difficult, but somewhat tedious, to see that

$$\lim_{t\to\infty} \frac{U(tx) - U(t)}{a(t)} = \frac{x^\gamma - 1}{\gamma}$$

implies

$$\lim_{t\to\infty} \frac{\log U(tx) - \log U(t)}{a(t)/U(t)} = \frac{x^{\gamma_-} - 1}{\gamma_-}.$$

This again implies

$$\lim_{t\to\infty} \int_0^1 \frac{\log U(t/x) - \log U(t)}{a(t)/U(t)} = \int_0^1 \frac{x^{-\gamma_-} - 1}{\gamma_-} dx = \frac{1}{1-\gamma_-}.$$

"Hence"

$$\frac{\frac{1}{k}\sum_{i=0}^{k-1} \log X_{n-i,n} - \log X_{n-k,n}}{a(\frac{n}{k})/U(\frac{n}{k})} \xrightarrow{P} \frac{1}{1-\gamma_-} \tag{17}$$

Similarly

$$\lim_{t\to\infty} \frac{(\log U(t/x) - \log U(t))^2}{\{a(t)/U(t)\}^2} = \left(\frac{x^{-\gamma_-} - 1}{\gamma_-}\right)^2,$$

"hence"

$$\frac{\frac{1}{k}\sum_{i=n}^{k-1} (\log X_{n-i,n} - \log X_{n-k,n})^2}{\{a(\frac{n}{k})/U(\frac{n}{k})\}^2} \xrightarrow{p} \frac{2}{(1-\gamma_-)(1-2\gamma_-)} \tag{18}$$

Combining (16) and (17) we get

$$\frac{M_{k,n}}{(H_{k,n})^2} \xrightarrow{P} \frac{2}{1-2\gamma_-}$$

with $H_{k,n}$ as before and

$$M_{k,n} := \frac{1}{k}\sum_{i=0}^{k-1} (\log X_{n-i,n} - \log X_{n-k,n})^2.$$

Theorem 1.8. *If $F \in D(G_\gamma), \gamma \in \mathbb{R}$ and $x^*(F) > 0$,*

$$\hat{\gamma}_{k,n} \xrightarrow{p} \gamma$$

with

$$\hat{\gamma}_{k,n} := H_{k,n} + 1 - \frac{1}{2}\left(1 - \frac{H_{k,n}^2}{M_{k,n}}\right)^{-1}$$

provided $k = k(n) \to \infty,\ k(n)/n \to o\ (n \to \infty)$.

(ix) Estimator of the scale function.

Relation (16) also shows how to estimate the scale function $a(\frac{n}{k})$; we only need to estimate $U(\frac{n}{k})$ in order to use (16) for that purpose. So we need to prove that

$$X_{n-k,n}/U(\frac{n}{k}) \xrightarrow{P} 1.$$

This can be done by writing

$$X_{n-k,n} = U(1/Z_{k,n})$$

with $Z_{k,n}$ the $(n-k)$-th order statistics from the uniform distribution (integral transformation). Now as you know from this course

$$\frac{n}{k} Z_{k,n} \xrightarrow{P} 1,$$

$n \to \infty,\ k(n) \to \infty,\ k(n)/n \to 0$. Hence, since

$$\lim_{t\to\infty} \frac{U(tx)}{U(t)} = x^{\gamma \vee 0}$$

locally uniformly for $x > 0$ (this follows from the theory of regularly varying functions),

$$\lim_{n\to\infty} \frac{X_{n-k,n}}{U(\frac{n}{k})} = \lim_{n\to\infty} \frac{U\left(\frac{n}{k}\left(\frac{n}{k} Z_{k,n}\right)^{-1}\right)}{U(\frac{n}{k})} = 1$$

in probability. Hence

$$\left\{ \frac{(1-\hat{\gamma}_-)X_{n-k,n}}{k} \cdot \sum_{i=0}^{k-1} \log X_{n-i,n} - \log X_{n-k,n} \right\} / a(\frac{n}{k}) \xrightarrow{P} 1.$$

2 Extreme value theory: statistics

I now turn to the asymptotic distribution of estimators, needed for confidence statements about the parameter gamma. I shall give three proofs of the asymptotic normality of the estimator

$$H_{k,n} := \frac{1}{k}\sum_{i=0}^{k-1} \log X_{n-i,n} - \log X_{n-k,n} \tag{19}$$

which is suitable for estimating gamma when this parameter is positive. The proof of the asymptotic normality of the estimator for general gamma is more involved, but follows the same lines. Two of the three proofs use empirical process theory.

(i) Proof via tail quantile process

We use the following Theorem from the course.

Theorem 2.1. *With a special construction, (and $k = k(n) \to \infty,\ k(n)/n \to 0$)*

$$\lim_{n\to\infty} t^{\frac{3}{2}+\varepsilon}\left[\sqrt{k}\left(\frac{k}{n}Y_{n-kt,n} - t^{-1}\right) - t^{-2}W(t)\right] = 0$$

in probability uniformly for $0 \le t \le 1$, with W Brownian motion and $\{Y_{i,n}\}_{i=1}^{n}$ the order statistics from a sample with distribution function $1 - 1/x,\ x > 1$.

We need a result like this one for a general class of probability distributions in $D(G_\gamma)$. First we state the extra conditions for this. The conditions are necessary for the proof of the asymptotic normality. We know

$$\lim_{t\to\infty} \frac{1-F(tx)}{1-F(t)} = x^{-1/\gamma} \tag{20}$$

for $x > 0$. We need to specify the rate of convergence in (20). So suppose that there is a function A_1, such that

$$\lim_{t\to\infty} \frac{\frac{1-F(tx)}{1-F(t)} - x^{-1/\gamma}}{A_1(t)}$$

exists for $x > 0$. One can prove that then for a possibly different function A^* which does not change sign eventually and some parameter $\rho^* \le 0$.

$$\lim_{t\to\infty} \frac{\frac{1-F(tx)}{1-F(t)} - x^{-1/\gamma}}{A^*(t)} = x^{-1/\gamma}\frac{x^{\rho^*}-1}{\rho^*} \tag{21}$$

for all $x > 0$. In the sequel we require that the parameter ρ^* which determines the convergence rate (since A^* is regularly varying with index ρ^*), is in fact negative that is, we exclude very slow convergence.

Next we translate this condition into a condition for the quantile function. We have seen that (20) is equivalent to

$$\lim_{t\to\infty} \frac{U(tx)}{U(t)} = x^\gamma \tag{22}$$

for $x > 0$. One can show that (21) is equivalent to a similar relation for U: there exist a function A not changing sign near infinity and a parameter ρ such that

$$\lim_{t\to\infty} \frac{\frac{U(tx)}{U(t)} - x^\gamma}{A(t)} = x^\gamma \cdot \frac{x^\rho - 1}{\rho} \tag{23}$$

for all $x > 0$. We shall also need its consequence

$$\lim_{t\to\infty} \frac{\log U(tx) - \log U(t) + \gamma \log x}{A(t)} = \frac{x^\rho - 1}{\rho} \text{ for all } x > 0. \tag{24}$$

Lemma 2.2. *If (23) holds and the sequence $k = k(n)$ is such that $k(n) \to \infty$ and*

$$\lim_{n\to\infty} \sqrt{k} A\left(\frac{n}{k}\right) = 0, \tag{25}$$

then

$$\lim_{n\to\infty} t^{\gamma+\frac{1}{2}+\varepsilon} \left[\sqrt{k}\left(\frac{X_{n-kt,n}}{U(\frac{n}{k})} - t^{-\gamma}\right) - t^{-\gamma-1}W(t)\right] = 0$$

in probability, uniformly for $0 \le t \le 1$ with W Brownian motion.

Proof. (sketch) One can prove uniformity in (23) in the following sense: for any $\varepsilon > 0$

$$\lim_{t\to\infty} x^{-\gamma-\rho-\varepsilon} \left[\frac{\frac{U(tx)}{U(t)} - x^\gamma}{A(t)} - x^\gamma \cdot \frac{x^\rho - 1}{\rho}\right] = 0 \tag{26}$$

uniformly for $x \ge 1$.

Now, replacing t in (25) by n/k and x by $\frac{k}{n} Y_{n-kx,n}$ with $\{Y_{i,n}\}$ the order statistics from the distribution function $1-1/x$, $x > 1$, we find (since $\{X_{i,n}\} \stackrel{d}{=} \{U(Y_{i,n})\}$)

$$\sqrt{k}\left[\frac{X_{n-tk,n}}{U(\frac{n}{k})} - (\frac{k}{n} Y_{n-kt,n})^\gamma\right] < \sqrt{k}|A(\frac{n}{k})| \left[H\left(\frac{k}{n}Y_{n-kt,n}\right) + R_\varepsilon\left(\frac{k}{n}Y_{n-kt,n}\right)\right]$$

with $H(x) := x^\gamma(x^\rho - 1)/\rho$ and $R_\varepsilon(x) := \varepsilon x^{\gamma+\rho+\varepsilon}$ and a similar lower inequality. Hence

$$t^{\gamma+\frac{1}{2}+\varepsilon}\sqrt{k}\left\{\left(\frac{X_{n-tk,n}}{U(\frac{n}{k})} - t^{-\gamma}\right) - \left(\left(\frac{k}{n}Y_{n-kt,n}\right)^\gamma - t^{-\gamma}\right)\right\} <$$

$$< \sqrt{k}|A(\frac{n}{k})|\left[t^{\gamma+\frac{1}{2}+\varepsilon}H\left(\frac{k}{n}Y_{n-kt,n}\right) + t^{\gamma+\frac{1}{2}+\varepsilon}R_\varepsilon\left(\frac{k}{n}Y_{n-kt,n}\right)\right].$$

Since by the Theorem the expression between the square brackets to the right is bounded in n and t, we get the result by applying the expansion of the Theorem to the second term to the left. $\square$

Corollary 2.3.

$$\lim_{n\to\infty} t^{\frac{1}{2}+\varepsilon}\left[\sqrt{k}\left(\log X_{n-kt,n} - \log U(\frac{n}{k}) + \gamma\log t\right) - t^{-1}W(t)\right] = 0$$

uniformly for $0 \le t \le 1$.

Now it is easy to get the asymptotic normality now.

Theorem 2.4. *If (23) holds and the sequence $k = k(n)$ is such that $k(n) \to \infty$ and $\sqrt{k}\cdot A(\frac{n}{k}) \to 0$,*

$$\sqrt{k}(H_{k,n} - \gamma)$$

is asymptotically normal with mean zero and variance γ^2.

Proof. We have seen that

$$H_{k,n} = \int_0^1 (\log X_{n-kt,n} - \log X_{n-k,n})\,dt.$$

Hence by the Corollary

$$\sqrt{k}\,(H_{k,n} - \gamma) = \sqrt{k}\;\;\gamma\left(\int_0^1(-\log t)dt - 1\right)$$

$$-\gamma\int_0^1 (t^{-1}W(t) - W(1))dt + o_p(1)\cdot\int_0^1 t^{-\frac{1}{2}-\varepsilon}dt.$$

Note that the first term is zero, that the integral in the second term has variance 1 and that the integral in the last term is finite. The proof is complete. $\square$

(ij) Proof via the tail empirical process

We are going to use Theorem 5.1 page 52, second part. The statement is as follows: for $k = k(n) \to \infty,\ k(n)n \to 0$, with a special construction

$$\lim_{n\to\infty} \frac{w_n(t) - W(t)}{t^\eta} = 0 \tag{27}$$

uniformly for $0 \le t \le 1$, a.s., if $0 \le \eta < 1/2$, where

$$\{w_n(t)\}_t = \left\{ \sqrt{k} \left(\frac{1}{k} \sum_{i=1}^{n} 1_{\{\frac{n}{k}\xi_i \le t\}} - t \right) \right\}_t$$

with $\xi_1, \xi_2, \ldots$ i.i.d. uniform (0,1) and W a suitable Brownian motion.

Write as before

$$\Gamma_n(t) := \frac{1}{n} \sum_{i=1}^{n} 1_{\{\xi_i \le t\}}.$$

Next we consider $X_1, X_2, \ldots$ i.i.d. F with $F \in D(G_\gamma), \gamma > 0$. Write

$$F_n(x) := \frac{1}{n} \sum_{i=1}^{n} 1_{\{X_i \le x\}}$$

and note that

$$F_n(x) \stackrel{d}{=} \Gamma_n(F(x))$$

so that

$$1 - F_n(x) \stackrel{d}{=} \Gamma_n(1 - F(x)).$$

Next we shall represent $H_{k,n}$ in terms of F_n:

$$\frac{1}{k} \sum_{i=1}^{k-1} \log X_{n-i,n} - \log X_{n-k,n}$$

$$= \frac{n}{k} \int_{X_{n-k,n}}^{\infty} (\log s) dF_n(s) - \frac{n}{k} \int_{X_{n-k,n}}^{\infty} (\log X_{n-k,n}) dF_n(s)$$

$$= \frac{n}{k} \int_{X_{n-k,n}}^{\infty} \int_{X_{n-k,n}}^{s} \frac{dx}{x} dF_n(s) = \frac{n}{k} \int_{X_{n-k,n}}^{\infty} \int_{x}^{\infty} dF_n(s) \frac{dx}{x}$$

$$= \int_{X_{n-k,n}}^{\infty} \frac{n}{k}(1 - F_n(x))\frac{dx}{x}.$$

$$= \frac{n}{k}\int_{X_{n-k,n}/U(\frac{n}{k})}^{\infty} \left\{1 - F_n\left(xU(\frac{n}{k})\right)\right\}\frac{dx}{x}$$

$$= \frac{n}{k}\int_{X_{n-k,n}/U(\frac{n}{k})}^{\infty} \Gamma_n\left(1 - F\left(xU(\frac{n}{k})\right)\right)\frac{dx}{x}$$

$$= \int_{\frac{1}{\frac{n}{k}\{1-F(X_{n-k,n})\}}}^{\infty} \frac{n}{k}\Gamma_n(\frac{k}{ny})d\log U(\frac{n}{k}y)$$

$$= \int_{\frac{k}{n}\{1-F(X_{n-k,n})\}^{-1}} \left\{\frac{1}{y} + \frac{W(\frac{1}{y})}{\sqrt{k}} + o(1)y^{-\eta}\frac{1}{\sqrt{k}}\right\} d\log U(\frac{n}{k}y)$$

$$= \int_1^{\infty} \frac{1}{y}d\log U(\frac{n}{k}y) + \frac{1}{\sqrt{k}}\int_1^{\infty} W(\frac{1}{y})d\log U(\frac{n}{k}y)$$

$$+ \int_{\frac{k}{n}\{1-F(X_{n-k,n})\}^{-1}} y^{-}1d\log U(\frac{n}{k}y)$$

$$\int_{\frac{k}{n}\{1-F(X_{n-k,n})\}^{-1}}^{1} W(\frac{1}{y})d\log U(\frac{n}{k}y) + \frac{o(1)}{\sqrt{k}}\int_{\frac{k}{n}\{1-F(X_{n-k,n})\}^{-1}}^{\infty} y^{-\eta}d\log U(\frac{n}{k}y).$$

Only the first three terms contribute to the result.
First consider

$$\int_1^{\infty} \frac{1}{y}d\log U(\frac{n}{k}y) = \int_1^{\infty} \left(\log U(\frac{n}{k}s) - \log U(\frac{n}{k})\right)\frac{ds}{s^2}.$$

Now one can prove uniformity in relation (24) in the following sense: for any $\varepsilon > 0$.

$$\lim_{t\to\infty} x^{-\gamma-\varepsilon}\left[\frac{\log U(tx) - \log U(t) + \gamma\log x}{A(t)} - \frac{x^{\gamma} - 1}{\gamma}\right] = 0 \tag{28}$$

uniformly for $x \geq 1$. This, combined with (25) $\left(\lim_{n\to\infty} \sqrt{k}\left(\frac{n}{k}\right) = 0\right)$, implies

$$\lim_{n\to\infty} \sqrt{k} \int_1^\infty \left(\log U\left(\frac{n}{k}s\right) - \log U\left(\frac{n}{k}\right) + \gamma \log s\right) ds = 0,$$

i.e.

$$\sqrt{k}\left(\int_1^\infty \frac{1}{y} d\log U(\frac{n}{k}y) - \gamma\right) = 0.$$

Further

$$\int_1^\infty W(\frac{1}{y}) d\log U(\frac{n}{k}y)$$

converges $(n \to \infty)$ to

$$\gamma \int_1^\infty W(\frac{1}{y})\frac{dy}{y} = \gamma \int_0^1 w(y)\frac{dy}{y}$$

(no proof: notice that (28) suggests: $d\log U(tx) \approx \gamma d\log x$). Similary

$$\int_{\frac{k}{n}\{1-F(X_{n-k,n})\}^{-1}}^1 \frac{1}{y} d\log U(\frac{n}{k}y) \sim \gamma \int_{\frac{k}{n}\{1-F(X_{n-k,n})\}^{-1}}^1 \frac{1}{y} \cdot \frac{dy}{y}$$

$$\sim \gamma \left[\frac{k}{n}\left\{1 - F(X_{n-k,n})\right\}^{-1} - 1\right]$$

$(n \to \infty)$. Next note that $\frac{n}{k}\{1 - F(xU(\frac{n}{k}))\}$ is close to $x^{-1/\gamma}$ and that

$$\sqrt{k}\left(\frac{X_{n-k,n}}{U(\frac{n}{k})} - 1\right) \longrightarrow -W(1)$$

in distribution (by the Lemma in (i), note that the W of (i) is minus the W of (ij)!). Putting everything together we obtain

$$\sqrt{k}(H_{k,n} - \gamma) \longrightarrow \gamma \int_0^1 W(y)\frac{dy}{y} - \gamma W(1)$$

which has variance γ^2.

The last proof does not use (inverse) empirical processes.

(iij) Proof via the central limit theorem

Let $Y_1, Y_2, \ldots$ be i.i.d. random variables with distribution function $1 - 1/x, x \geq 1$. Suppose F is a distribution function in $D(G_\gamma)$ and define

$$U := \left(\frac{1}{1-F}\right)^{\leftarrow}.$$

Then the random variables $U(Y_1), U(Y_2), \ldots$ are i.i.d. with distribution function F. So

$$H_{k,n} \stackrel{d}{=} \frac{1}{k}\sum_{i=0}^{k-1} \log U(Y_{n-i,n}) - \log U(Y_{n-k,n}).$$

For the proof of the Theorem we now use (28). Note that $Y_{n-k,n} \to \infty$ almost surely $(n \to \infty)$, hence for sufficiently large n and all $0 \leq i \leq k-1$

$$\log U\left(\frac{Y_{n-i,n}}{Y_{n-k,n}} \cdot Y_{n-k,n}\right) - \log U(Y_{n-k,n}) \gtreqless \gamma \log\left(\frac{Y_{n-i,n}}{Y_{n-k,n}}\right) +$$

$$+A(Y_{n-k,n})\left(\left(\frac{Y_{n-i,n}}{Y_{n-k,n}}\right)^{\rho} - 1\right)/\rho \pm \varepsilon|A(Y_{n-k,n})|\left(\frac{Y_{n-i,n}}{Y_{n-k,n}}\right)^{+\rho+\varepsilon}.$$

It follows that

$$H_{k,n} \gtreqless \gamma\frac{1}{k}\sum_{i=0}^{k-1}\log\left(\frac{Y_{n-i,n}}{Y_{n-k,n}}\right) + A(Y_{n-k,n})\frac{1}{k}\sum_{i=0}^{k-1}\left(\left(\frac{Y_{n-i,n}}{Y_{n-k,n}}\right)^{\rho} - 1\right)/\rho$$

$$\pm\varepsilon|A(Y_{n-k,n})|\frac{1}{k}\sum_{i=0}^{k-1}\left(\frac{Y_{n-i,n}}{Y_{n-k,n}}\right)^{\rho+\varepsilon}.$$

Now note (this is the essential step) that

$$\left\{\frac{Y_{n-i,n}}{Y_{n-k,n}}\right\}_{i=0}^{k-1} \stackrel{d}{=} \{Y_{k-i,k}\}_{i=0}^{k-1}$$

(that is, the distribution of the left hand side does not depend on n).
Hence

$$\frac{1}{k}\sum_{i=0}^{k-1}\log\left(\frac{Y_{n-i,n}}{Y_{n-k,n}}\right) \stackrel{d}{=} \frac{1}{k}\sum_{i=0}^{k-1} Y_{k-i,k}$$

and, since we now take the sum of **all** order statistics, this is the same in distribution as

$$\frac{1}{k}\sum_{i=1}^{k} Y_i$$

with $Y_1, Y_2, \ldots$ i.i.d. as before.
Similarly for the other terms. Hence

$$\sqrt{k}(H_{k,n} - \gamma) \gtrless \gamma\sqrt{k}\left(\frac{1}{k}\sum_{i=1}^{k}\log Y_i - 1\right) +$$

$$\sqrt{k}A(Y_{n-k,n})\frac{1}{k}\sum_{i=1}(Y_i^{\rho} - 1)/\rho \pm \varepsilon|A(Y_{n-k,n})|\frac{1}{k}\sum_{i=1}^{k} Y_i^{\rho+\epsilon}.$$

Now notice that

$$\sqrt{k}\left(\frac{1}{k}\sum_{i=1}^{k}\log Y_i - 1\right)$$

is asymptotically standard normal and that

$$\frac{1}{k}\sum_{i=1}^{k}(Y_i^{\rho} - 1)/\rho \longrightarrow E(Y^{\rho} - 1)/\rho \text{ finite}$$

and

$$\frac{1}{k}\sum_{i=1}^{k} Y_i^{\rho+\varepsilon} \longrightarrow EY^{\rho+\varepsilon} \text{ finite}$$

in probability. Furthermore we know $\frac{k}{n}Y_{n-k,n} \to 1$ in probability so that, since $|A|$ is regularly varying with index ρ,

$$\frac{A(Y_{n-k,n})}{A(\frac{n}{k})} \longrightarrow 1$$

in probability. Combining all of this and noticing that $\sqrt{k}A(\frac{n}{k}) \to 0$ $(n \to \infty)$, we get that

$$\sqrt{k}\left(\frac{H_{k,n}}{\gamma} - 1\right)$$

has asymptotically a standard normal distribution.