Statistical Theory of Extremes

Homepage: http://www.gathacognition.com/book/gcb14
http://dx.doi.org/10.21523/gcb1

Part 1
Probabilistic Patterns of Univariate Statistical Extremes

Chapter 3

Attraction Structure and Speed of Convergence

J. Tiago de Oliveira
Academia das Ciências de Lisboa (Lisbon Academy of Sciences), Lisbon, Portugal.

Abstract

In this chapter results are presented that make it possible to analyse if a sequence of random variables is attracted for maxima to one of the three limiting distributions, for i.i.d. case, how to obtain some normalizing coefficients and how quick the convergence is. This involves necessary and sufficient conditions on the tail of the underlying distribution to be in the respective domain of attraction for maxima; moreover, sufficient conditions and miscellaneous examples are also provided, including normal, geometric, Poisson and lognormal models.

Published Online

23 June 2017

Keywords

Attraction coefficients;
Speed of convergence;
Quantile function;
Slow variation;
Tail equivalent.

Editor(s)

J.C. Tiago de Oliveira

3.1 Introduction

From the utilitarian standpoint this chapter has very little importance: it deals with the i.i.d. case, where the underlying distribution F(x) is known, and we want to know if the sequence of random variables $\{X_k\}$ is attracted for maxima, to one of the three limiting distributions, how to obtain *some* set of attraction coefficients and, finally, how quick is the convergence.

In applications, observations are often not i.i.d., but under some conditions, for large samples, the limiting distributions of maxima of a sequence of non- i.i.d. observations are still the same as in the i.i.d. case. But

Originally published in 'Statistical Analysis of Extremes', 1997, 2016.
http://dx.doi.org/10.21523/gcb1.1704

even if they were independent, F(x) would not be known (in general) in order to analyse if the maxima are attracted to Ψ_α, Λ or Φ_α and thus to obtain a set (not necessarily the best in any sense) of attraction coefficients (λ_k, δ_k). The results here, from a utilitarian standpoint − immediate utility − are of little importance. The analysis of the speed of convergence follows the same lines.

Moreover, the practical importance is not nil; on the contrary. The last section (extensions) of the previous chapter says that the limiting distributions $\tilde{L}(x)$ for maxima (or $\underset{\sim}{L}(x)$ for minima) are valid for outside the i.i.d. case and in many cases (say for instance the discharges of a river) the total sample can be divided into "chunks" (say the hydrological year) where the maximum (the flood) has a distribution function, at least approximately stable, close to one of the $\tilde{L}((x-\lambda)/\delta)$ which will then be used to study maxima, thus avoiding the "computable" values (λ_k, δ_k) and in some cases α, substituted by its statistical estimation. As in oceanographical and hydrological situations data a few days apart (sometimes less than two days) behave in an independent − like way, the knowledge of the speed of convergence gives an idea how the approximation $\tilde{L}((x-\lambda)/\delta)$ is valid.

In consequence, so far *through analogy*, this chapter, for the i.i.d. case plus the previous extensions section, illustrates the use of asymptotic distributions in dealing with a dependent, finite set of data. And this is its practical, though not utilitarian, justification.

3.2 The attraction conditions and attraction coefficients

Let there be a sequence of i.i.d. random variables $\{X_i\}$ with distribution function F(x).

We will prove the following statements for $F \in \mathcal{D}(\tilde{L})$; the dual statements of $F \in \mathcal{D}(\underset{\sim}{L})$ can be obtained by a simple conversion of the previous results. The quantile function $Q(v) = \inf\{x \mid F(x) \geq v\}$; Q (v) is very important in this study and, by the results given here, also in the problems of the right tail estimation using the largest values of a sample (Annex 5 to Part 2).

The statements for $F \in \mathcal{D}(\tilde{L})$, taking $t > 0$, are:

- *F is attracted, for maxima, to the Weibull distribution $\Psi_\alpha(x)$ iff $\bar{w} < +\infty$ and $\frac{1-F(\bar{w}-tx)}{1-F(\bar{w}-t)} \to x^\alpha$ as $t \to 0^+$; a system (λ_k, δ_k) of attraction coefficients is $(\bar{w}, \bar{w} - Q(1-1/k))$.*
- *F is attracted, for maxima, to the Gumbel distribution $\Lambda(x)$ iff*

$\frac{Q(1-tx)-Q(1-t)}{Q(1-te)-Q(1-t)} \to \log x$ *as* $t \to 0^+$, *or*

$\frac{Q(1-tx)-Q(1-t)}{Q(1-ty)-Q(1-t)} \to \frac{\log x}{\log y}$ *as* $t \to 0^+$, *if* $y \neq 1$, *or*

$\frac{Q(1-(1-v)x)-Q(v)}{Q(1-(1-v)e)-Q(v)} \to \log x$ *as* $v \to 1^-$, *or, also*

$\frac{Q(v^x)-Q(v)}{Q(v^e)-Q(v)} \to \log x$ *as* $v \to 1^-$;

a system of attraction coefficients is $\lambda_k = Q(1-1/k), \delta_k = Q(1-1/ek) - Q(1-1/k)$.

- *F is attracted, for maxima, to the Fréchet distribution* $\Phi_\alpha(x)$ *iff* $\overline{w} = +\infty$ *and* $\frac{1-F(t)}{1-F(tx)} \to x^\alpha$ *as* $t \to +\infty$; *a system of attraction coefficients is* $(\lambda_k, \delta_k) = (0, Q(1-1/k))$ *(ultimately positive as* $\overline{w} = +\infty$ *).*

  Analogously to the attraction condition, for maxima, to Gumbel distribution given in terms of the quantile function we have also

- *F is attracted, for maxima, to Fréchet distribution* $\Phi_\alpha(x)$ *iff*

  $$\frac{Q(1-tx)-Q(1-t)}{Q(1-ty)-Q(1-t)} \to \frac{x^{-1/\alpha}-1}{y^{-1/\alpha}-1} \text{ as } t \to 0^+ \text{ if } y \neq 1 .$$

  and to Weibull distribution $\Psi_\alpha(x)$ *iff*

  $$\frac{Q(1-tx)-Q(1-t)}{Q(1-ty)-Q(1-t)} \to \frac{x^{-1/\alpha}-1}{y^{-1/\alpha}-1} \text{ as } t \to 0^+ \text{ if } y \neq 1.$$

If we do use the integrated von Mises-Jenkinson form we can say that $F(.)$ *is attracted, for maxima, to* $G(.|\theta)$ *iff*

$$\frac{Q(1-tx)-Q(1-t)}{Q(1-ty)-Q(1-t)} \to \frac{x^{-\theta}-1}{y^{-\theta}-1}\left(to \ \frac{\log x}{\log y} \text{ if } \theta = 0\right) \text{as } t \to 0^+ \text{ if } y \neq 1.$$

Notes:

1. There are other conditions for the attraction to the Gumbel distribution, but these seem to be the most operational;
2. A function SV(x) is said to be of *slow variation* (or *slowly varying*) at $+\infty$ if it is defined at least in a right half-line (i.e., for all $x > x_0$) and $SV(tx)/SV(x) \to 1$ as $x \to +\infty$; the latter nomenclature can be used to say that F(x) is attracted to the Weibull distribution if $\overline{w} < +\infty$ and

$SV(x) = x^{\alpha}(1 - F(\overline{w} - 1/x))$ is slowly varying at $+\infty$ and that F(x) is attracted to the Fréchet distribution if $\overline{w} = +\infty$ and $SV(x) = x^{\alpha}(1 - F(x))$ is slowly varying at $+\infty$;

3. To obtain the Fréchet distribution as a limit we must have $\overline{w} = +\infty$ and to obtain the Weibull distribution as a limit we must have $\overline{w} < +\infty$; the Gumbel distribution can be attained either with $\overline{w} < +\infty$ or $\overline{w} = +\infty$;
4. The conditions for Weibull and Fréchet distributions to be limiting distributions for maxima are easy to work out; this is not the case for the Gumbel distribution;
5. At the end of the section we will give conditions using F'(x), when it exists, which are in general very handy to use.
6. Dually we obtain the statements for $F \in \mathcal{D}(\underset{\sim}{L})$, supposing, also, $t > 0$. They are:
   - *F is attracted, for minima, to the Weibull distribution* $W_{\alpha}(x) = 1 - \Psi_{\alpha}(-x)$ *iff* $\underline{w} = -\infty$ *and* $\frac{F(\underline{w}+tx)}{1-F(\underline{w}+x)} \to t^{\alpha}$ *as* $x \to 0^{+}$ *; a system of attraction coefficients is* $(\underline{w}, Q(1/k) - \underline{w})$.
   - *F is attracted, for minima, to the Gumbel distribution* $1 - \Lambda(-x)$ *iff*

     $\frac{Q(t\,x) - Q(t)}{Q(e\,t) - Q(t)} \to \log x$ *as* $x \to 0^{+}$, *or*

     $\frac{Q(1-v^{x}) - Q(1-v)}{Q(1-v^{e}) - Q(1-v)} \to \log x$ *as* $v \to 1^{-}$;

     a system of attraction coefficients is $(Q(1/k), Q(1/k) - Q(1/ek))$.
   - *F is attracted, for minima, to the Fréchet distribution* $1 - \Phi_{\alpha}(-x)$ *iff* $\underline{w} = -\infty$ *and* $\frac{F(t)}{F(tx)} \to x^{\alpha}$ *as* $t \to -\infty$ *; a system of attraction coefficients is* $(0, -Q(1/k))$ $(-Q(1/k)$ *is ultimately positive as* $\underline{w} = -\infty$*).*
7. Before going further, let us recall that each of the limiting distributions is attracted to itself; this is shown immediately using the above criteria. Notice also that the conditions on $\overline{w}$ and $\underline{w}$ can help to eliminate one case for each F(x) ; we used this in the previous examples.

Proofs:

We will follow a different order in the proofs for convenience. Let us consider the necessary and sufficient condition for attraction of maxima to $\Phi_{\alpha}(x)$.

Suppose $\overline{w} = +\infty$ and $\frac{1-F(t)}{1-F(t\,x)} \to x^{\alpha}$ as $t \to +\infty$. To prove the result we have to show that for some $\{\lambda_k, \delta_k\}$, $F^{k}(\lambda_k + \delta_k\,x) \to \Phi_{\alpha}(x)$ or, equivalently as said, that $k(1-F(\lambda_k+\delta_k\,x))\to x^{-\alpha}$ if $x>0$, and $k\big(1-F(\lambda_k+\delta_k\,x)\big)\to+\infty$ if $x<0$.

Let us take $\lambda_k=0$ and $\delta_k=Q(1-l/k)$. As $\overline{w}=+\infty$, we have $\delta_k\to+\infty$ and for $x<0, \delta_k\,x\to-\infty$, implying $k\big(1-F(\delta_k\,x)\big)\to+\infty$ for $x<0$.

Consider now $k(1-F(\delta_k\,x))$ for $x>0$. As $\lim k(1-F(\delta_k\,x))=\lim\{k(1-F(\delta_k))\frac{1-F(\delta_k\,x)}{1-F(\delta_k)}\}$ and by hypothesis when $k\to+\infty$, $\delta_k\to+\infty$ the second factor $\frac{1-F(\delta_k\,x)}{1-F(\delta_k)}\to x^{-\alpha}$; it remains to show that $k(1-F(\delta_k))\to 1$. By the definition of δ_k one gets $F(\delta_k^-)\le 1-1/k=F(\delta_k)$ and so $k\big(1-F(\delta_k)\big)\le 1$; but $F(\delta_k\,x)\le F(\delta_k^-)\le 1-1/k$ for $0<x<1$ and so $\frac{1-F(\delta_k)}{1-F(\delta_k\,x)}\le k(1-F(\delta_k))$ and $\frac{1-F(\delta_k)}{1-F(\delta_k\,x)}\to x^{\alpha}$ and thus $k(1-F(\delta_k))\to 1$ with $k\to\infty$, as desired. We have shown up to now that if the attraction conditions for $\Phi_\alpha(x)$ are valid we can take as attraction coefficients $\lambda_k=0, \delta_k=Q(1-1/k)$ obtaining as limit $\Phi_\alpha(x)$.

Let us now prove the converse. As $F^k(\lambda_k+\delta_k\,x)\to\Phi_\alpha(x)$ or, equivalently, $k(1-F\,(\lambda_k+\delta_k\,x))\to x^{-\alpha}$ we have also

$[k\beta]\,(1-F(\lambda_k+\delta_k\,x))\to\beta\,x^{-\alpha}$,

and thus, with $x=\beta^{1/\alpha}\,z$,

$[k\beta]\,(1-F(\lambda_k+\delta_k\,\beta^{1/\alpha}\,z))\to z^{-\alpha}$,

from which, with

$[k\beta]\,(1-F(\lambda_{[k\beta]}+\delta_{[k\beta]}z))\to z^{-\alpha}$,

by Khintchine's convergence of types theorem we get $(\lambda_{[k\beta]}-\lambda_k)/\delta_k\to 0$ and $\delta_{[k\beta]}/\delta_k\to\beta^{1/\alpha}$.

Take now, for fixed $\beta>1, \lambda_{[k\beta]}=\lambda_k, \delta_{[k\beta]}=\delta_k\,\beta^{1/\alpha}$. Define, now, a integer sequence $[k(s)]$by $k(1)=[k\beta]\,, k(s+1)=[k(s).\beta]$, and so we get $\lambda_{k(1)}=\lambda_k\,, \delta_{k(s)}=\delta_k\,\beta^{(s-1)/\alpha}(\to+\infty$ as $\beta>1)$. Then $\lambda_{k(1)}/\delta_{k(s)}\to 0$ and we can write, as $\lambda_{k(s)}$ can be taken to be zero,

$F^{k(s)}(\delta_{k(s)}x)\to\Phi_\alpha(x)$.

Let us fix, now, x, choose y — to increase indefinitely — and obtain s such that $\delta_{k(s)}\, x \le y \le \delta_{k(s+1)}\, x$. Then we have

$$1 - F(\delta_{k(s+1)}\, x) \le 1 - F(y) \le 1 - F(\delta_{k(s)} x)$$

and so

$$\frac{1 - F(\delta_{k(s+1)}\, x)}{1 - F(\delta_{k(s)}\, t\, x)} \le \frac{1 - F\,(y)}{1 - F\,(t\ \ y)} \le \frac{1 - F(\delta_{k(s)}\, x)}{1 - F(\delta_{k(s)}\, t\ \ x)}$$

and as

$$\frac{k(s+1)}{k(s)} = \frac{\beta\, k(s) - r}{k(s)} \to \beta$$

($0 \le r < 1$ is the fractional part) and as $k(s)(1 - F(\delta_{k(s)}\, x)) \to x^{-\alpha}$ as $k \to +\infty$ we get, finally letting $y \to +\infty$ and so $k(s) \to +\infty$,

$$\frac{1}{\beta} t^{\alpha} \le \lim_{y\to\infty} \frac{1 - F(y)}{1 - F(y\ \ t)} \le \beta\, t^{\alpha}$$

and thus, as $\beta > 1$, as close to 1 as wished,

$$\frac{1 - F\,(y)}{1 - F\,(y\ \ t)} \to t^{\alpha} \text{ as } y \to \infty.$$

We have shown, also, that as attraction coefficients we can take $(0, Q(1 - 1/k))$ in choosing δ_k such that $k(1 - F(\delta_k)) \to 1$.

The conversion of this result concerning $\Phi_\alpha(x)$ to $\Psi_\alpha(x)$ is very easy. As seen, $\bar{F}^k(\lambda_k + \delta_k\, x) \to \Phi_\alpha(x)$ is equivalent to $\bar{F}^k(\delta_k\, x) \to \Phi_\alpha(x)$; in addition, if we define $\bar{F}(x) = F(\bar{w} - 1/x)$ we see that with $\bar{\delta}_k = \bar{Q}\,(1 - 1/k)$, $\bar{F}^k(\bar{\delta}_k\, x) \to \Phi_\alpha(x)$ iff $F^k(\bar{w} - 1/\bar{\delta}_k\, x) \to \Phi_\alpha(x)$ or $F^k(\bar{w} - 1/\bar{\delta}_k\, x) \to \Phi_\alpha(-1/x) = \Psi_\alpha(x)$.

Consequently for F to be attracted, for maxima, to $\Psi_\alpha(x)$ we must have $\bar{w} < +\infty$, satisfy the attraction condition $\frac{1-F(\bar{w}-t\,x)}{1-F(\bar{w}-t\,)} \to x^{\alpha}$ as $t \to 0^{+}$, and take as attraction coefficients $\lambda_k = \bar{w}$, $\delta_k = \bar{\delta}_k^{-1}$, or, more simply, $\delta_k = \bar{w} - Q(1 - 1/k)$.

Let us now go to the proof of the attraction condition for maxima, to the Gumbel distribution.

We have from $F^k(\lambda_k + \delta_k\, x) \to \Lambda(x)$ also $F^{[kt]}(\lambda_{[kt]} + \delta_{[kt]}\, x\,) \to \Lambda(x)$ and thus $(F^k(\lambda_{[kt]} + \delta_{[kt]}\, x))^{[kt]/k} \to \Lambda(x)$ or $F^k(\lambda_{[kt]} + \delta_{[kt]} x) \to \Lambda^{1/t}(x) = \Lambda\,(x + \log t)$ and so $F^k(\lambda_{[kt]} + \delta_{[kt]}(x - \log t)) \to \Lambda(x)$. By the Khintchine's convergence of types theorem we get

$$\frac{\lambda_{[kt]} - \delta_{[kt]} \log t - \lambda_k}{\delta_k} \to 0$$

and

$$\delta_{[kt]}/\delta_k \to 1,$$

or equivalently $(\lambda_{[kt]} - \lambda_k)/\delta_k \sim \log t$ and $\delta_{[kt]} \sim \delta_k$.

Taking $t = e$ we get $\delta_k \sim \lambda_{[ke]} - \lambda_k$ and so the conditions are

$$\frac{\lambda_{[k\,t]} - \lambda_k}{\lambda_{[k\,e]} - \lambda_k} \to \log t \text{ and } \delta_k \sim \lambda_{[k\,e]} - \lambda_k,$$

the last evidently defining δ_k , asymptotically.

Thus if $F \in \mathcal{D}\,(\Lambda)$, with attraction coefficients $(\lambda_k\,, \delta_k)$, we know that

$$\frac{\lambda_{[kt]} - \lambda_k}{\delta_k} \sim \frac{\lambda_{[kt]} - \lambda_k}{\lambda_{[ke]} - \lambda_k} \to \log t \text{ with } \delta_k \sim \lambda_{[ke]} - \lambda_k.$$

The relation $k(1 - F(\delta_k + \lambda_k\, x)) \to e^{-x}$ suggests the use of $\lambda_k^* = Q(1 - 1/k), \delta_k^* = \lambda_{[ke]}^* - \lambda_k^*$, with Q the quantile function. Let us show that if $(\lambda_{[kt]}^* - \lambda_k^*)/\delta_k^* \to \log t$ then $F \in \mathcal{D}\,(\Lambda)$ and by the Khintchine's convergence of types theorem any $(\lambda_k\,, \delta_k)$ such that $F^k(\lambda_k + \delta_k\, x) \to \Lambda(x)$ and $(\lambda_k^*\,,\ \delta_k^*)$ are equivalent for the limit. For fixed t, from

$$\frac{\lambda_{[kt]}^* - \lambda_k^*}{\delta_k^*} \to \log t \quad \text{we get, for large k,}$$

$$\lambda_k^* + \delta_k^*\,(\log t - \in) < \lambda_k^* = Q(1 - \frac{1}{[kt]}) < \lambda_k^* + \delta_k^*\,(\log t + \in),$$

and so

$$F(\lambda_k^* + \delta_k^*\,(\log t - \in)) \leq F(Q(1 - 1/[kt])) \leq F(\lambda_k^* + \delta_k^*\,(\log t + \in)).$$

From the RHS inequality, as $v \leq F(Q(v))$, we get

$$1 - 1/[kt] \leq F\big(\lambda_k^* + \delta_k^*(\log t + \in)\big)\;;$$

raising to power k and letting $k \to \infty$ we get

$e^{-1/t} \leq \underline{\lim}\, F^k(\lambda_k^* + \delta_k^* (\log t + \in))$

and so $\exp\{-e^{-(z-\in)}\} \leq \underline{\lim}\, F^k(\lambda_k^* + \delta_k^* z)$.

For the LHS inequality, as $\lambda_k^* + \delta_k^*(\log t + \in) < \lambda_k^*$ we get, from $F(Q(v)^-) \leq v$,

$F(\lambda_k^* + \delta_k^* (\log t - \in)) \leq 1 - 1/[kt]$;

raising to power k and letting $k \to +\infty$ we get

$\overline{\lim}\, F^k(\lambda_k^* + \delta_k^* (\log t - \in)) \leq e^{-1/t}$

and in the same way

$\overline{\lim}\, F^k(\lambda_k^* + \delta_k^* z) \leq \exp\{-e^{-(z+\in)}\}$.

Consequently

$\Lambda(z - \in) \leq \underline{\lim}\, F^k(\lambda_k^* + \delta_k^* z) \leq \overline{\lim}\, F^k(\lambda_k^* + \delta_k^* z) \leq \Lambda(z + \in)$

and thus

$F^k(\lambda_k^* + \delta_k^* z) \to \Lambda(z)$,

and any system of attraction coefficients is equivalent to $(\lambda_k^*, \delta_k^*)$.

As a consequence we see that $F \in \mathcal{D}(\Lambda)$ iff

$$\frac{Q(1 - 1/[kx]) - Q(1 - 1/k)}{Q(1 - 1/[ke]) - Q(1 - 1/k)} \to \log x \text{ as } k \to \infty$$

and we can use as attraction coefficients $\lambda_k^* = Q(1 - 1/k)$ and $\delta_k^* = \lambda_{[ke]}^* - \lambda_k^*$

Let us now give a continuous form to this discrete result. We will show that the condition above is equivalent to

$$\frac{Q(1 - t\,x) - Q(1 - t)}{Q(1 - t\,e) - Q(1 - t)} \to \log x \text{ as } t \to 0^+.$$

Let $k = [1/t] + 1$; we will show that

$$\frac{Q(1-t\,x)-Q(1-t)}{\delta_k^*} \to -\log x \ \text{ as } t \to 0^+ ;$$

by division we obtain

$$\frac{Q(1-t\,x)-Q(1-t)}{Q(1-t\,e)-Q(1-t)} \to \log x \ \text{ as } t \to 0^+.$$

From $1/t < k \le 1/t+1$ we get $1/k < t \le 1/(k-1), x/k < t\,x \le x/(k-1)$ and so $[(k/x)+1]^{-1} < t\,x < [(k-1)/x]^{-1}$ and as $Q(1-\xi)$ is non-increasing in ξ we have

$$\frac{\lambda^*_{[(k-1)/x]}-\lambda^*_{k+1}}{\delta_k^*} \le \frac{Q(1-t\,x)-Q(1-t)}{\delta_k^*} \le \frac{\lambda^*_{[(k/x)+1]}-\lambda^*_{k-1}}{\delta_k^*}.$$

But

$$\frac{\lambda^*_{[(k/x)+1]}-\lambda^*_{k-1}}{\delta_k^*} = \frac{\lambda^*_{[k/x]+1}-\lambda^*_{[k/x]}}{\delta_k^*} + \frac{\lambda^*_{[k/x]}-\lambda^*_{k}}{\delta_k^*} + \frac{\lambda^*_{k}-\lambda^*_{k-1}}{\delta_k^*}.$$

The last summand converges to zero and the second to $\log(1/x) = -\log x$. The first summand can be written as $\frac{\lambda^*_{[k/x]+1}-\lambda^*_{[k/x]}}{\delta^*_{[k/x]}} \cdot \frac{\delta^*_{[k/x]}}{\delta_k^*}$. The first factor converges to zero and the second to 1 as seen in the beginning of the proof because $\delta_{[kx]}/\delta_k \to 1$ and so $\delta^*_{[k/x]}/\delta^*_k \to 1$. Thus the RHS of the inequality converges to $-\log x$ and the same can be proved for the LHS. Thus $\frac{Q\,(1-t\,x)-Q(1-t)}{\delta_k^*} \to -\log x$ and so the continuous condition for attraction is

$$\frac{Q(1-t\,x)-Q(1-t)}{Q(1-t\,e)-Q(1-t)} \to \log x \ \text{ as } t \to 0^+.$$

This condition was given by Mejzler (1949) and was also stated by Marcus and Pinsky (1969) independently in another form; other important texts are de Haan (1970), (1971) and Balkema and de Haan (1972).

The proof of the results concerning attraction to $\Phi_\alpha(x)$, $\Psi_\alpha(x)$ and $G(z\,\theta)$ in terms of the quantile function runs in the same lines.

3.3 Sufficiency conditions and examples

Before going on to some examples, some old and revisited and some new, let us state only sufficiency conditions — which are almost necessary — for attraction, but supposing the existence of the derivative F'(x), at least in a right half-line (i.e., for $x > x_0$). The proofs are simple and can be found in Gnedenko (1943) and de Haan (1976).

- *if* $\overline{w} < +\infty$, $F(x)$ *has positive density* $F'(x)$ *in some interval* $]x_0, \overline{w}[$, *and* $\lim_{x \uparrow \overline{w}} \frac{(\overline{w}- x)F'(x)}{1-F(x)} = \alpha$, *then* $F(x)$ *is attracted to* $\Psi_\alpha(x)$ *and we can use the coefficients* $\lambda_k = \overline{w}, \delta_k = \overline{w} - Q(1-1/k)$, *as* $F(x)$ *defines uniquely* $Q(1-1/k)(\to \overline{w})$ *for large* k;
- *if* $F''(x) < 0$ *for some interval* $]x_0, \overline{w}$ [, $(\overline{w} < +\infty$ *or* $\overline{w} = +\infty)$ *and* $\frac{d}{dx}(\mu(x)^{-1}) \to 0$, *where* $\mu(x)$ *is the force of mortality or hazard rate, or* $\frac{F''(x)\,(1-F(x))}{(F'(x))^2} \to -1$ *as* $x \uparrow \overline{w}$ *then* $F(x)$ *is attracted to* $\Lambda(x)$ *and we can use the coefficients* $\lambda_k = Q(1-1/k$, *and* $\delta_k = 1/k\, F'(\lambda_k)$ *(von Mises criterion);*
- *If* $\overline{w} = +\infty$, $F(x)$ *has a positive density for* $x > x_0$, *and* $\frac{x\,F'(x)}{1-F(x)} \to \alpha$, *as* $x \to +\infty$, *then* $F(x)$ *is attracted to* $\Phi_\alpha(x)$ *and we can use the attraction coefficients* $\lambda_k = 0$, $\delta_k = Q(1-1/k)$.

These maxima results, in the same way as those following, can be converted to minima results.

But also:

- *If* $F(x)$ *is attracted for maxima to* $\Lambda(x)$ *then* $\delta_k/\lambda_k \to 0$;
- *If* $\overline{w} = +\infty$, *and* $F(x)$ *is attracted for maxima to* $\Lambda(x)$ *then* $\frac{1-F(t\,x)}{1-F(t)} \to 0$ *as* $t \to +\infty$, *with* $t > 0$; *this result implies the MLLN;*
- *If* $\ldots x_1 < x_2 < x_k < \ldots \overline{w}\ (\leq +\infty)$ *are jump points of* $F(x)$ *and* $\frac{1-F(x_k^-)}{1-F(x_k)} \geq 1+\beta$, $\beta > 0$, *then* $F(x)$ *is not attracted for maxima to* $\Lambda(x)$;
- *If* $\overline{w} = +\infty$, $F^k(\lambda_k + \delta_k) \to \Lambda(x)$ $(\delta_k = \delta$ *constant) is equivalent to* $\frac{1-F(\log\, t)}{1-F(\log\, t\, x)} \to x^\alpha(\alpha = 1/\delta)$ *as* $t \to \infty$ *or* $\frac{1-F(y)}{1-F(y+k)} \to e^{k/\delta}$ *as* $y \to +\infty$;

 the proof of the last statement is easily reduced to the Fréchet limiting situation by the use of the transform $Y = \log X$ and some adjustments.

Let us now go to the miscellaneous examples. We will not obtain the attraction coefficients, but will only decide on attraction.

1. Let us consider the classical normal distribution. As $\overline{w} = +\infty$ it cannot be attracted to $\Psi_\alpha(x)$. But $\frac{1-N(t)}{1-N(t\,x)} \sim x^{-1} e^{t^2(x^2-1)/2}$ which does not converge to x^α and so $\Phi_\alpha(x)$ is excluded. The Mejzler criterion for the convergence to $\Lambda(x)$ in this example is not very operational. Let us consider the von Mises criterion $\frac{N''(x)(1-N(x))}{(N'(x))^2} \to -1$ as $x \to +\infty$. But $1 - N(x) \sim N'(x)/x$; as used before, and $N''(x) = -x\,N'(x)$ and so $\frac{N''(x)(1-N(x))}{(N'(x))^2} \to -1$, and $N(x)$is attracted, for maxima, to $\Lambda(x)$.
2. Consider the geometric distribution $F(x) = 1 - e^{-\lambda[x]}$ if $x \geq 0$, $F(x) = 0$ if $x < 0$. The jump points are the non-negative integers and we have $\frac{1-F(k^-)}{1-F(k)} = \frac{1-F(k-1)}{1-F(k)} = e^\lambda > 1$ and so $F(x)$ is not attracted to $\Lambda(x)$, as opposed to the exponential. As $\overline{w} = +\infty$ the limit could be $\Phi_\alpha(x)$ if $\frac{1-F(t)}{1-F(t\,x)} \to x^\alpha$ as $t \to +\infty$. But $\frac{1-F(t)}{1-F(t\,x)} = e^{\lambda\{[tx]-[1]\}} = e^{\lambda\{[tx-r]-[t-r']\}}$, with $0 \leq r, r' < 1$, which does not converge to x^α, and so the distribution is not attracted for maxima to $\Phi_\alpha(x)$.
3. Let us now analyse the Poisson distribution. We have $F(x) = \sum_{j=0}^{[x]} e^{-\lambda} \frac{\lambda^j}{j!} = \sum_0^{[x]} p_j$ with jumps at the integers $j = 0,1,\ldots,$. Then $\frac{1-F(j^-)}{1-F(j)} = \frac{1-F(j-1)}{1-F(j)} = 1 + \frac{p_j}{1-F(j)} \geq 1 + \frac{\lambda^j/j!}{\sum_{k=j}^{[\infty]} \lambda^k/k!} = 1 + \frac{1}{\sum_0^{[\infty]} \frac{j!}{(r+j)!}\lambda^r} = 1 + \frac{1}{\sum_0^{[\infty]} \binom{r+j}{r}^{-1} \frac{\lambda^r}{r!}} > 1 + \frac{1}{\sum_0^{[\infty]} \frac{\lambda^r}{r!}} = 1 + e^{-\lambda}$ and Poisson distribution is not attracted for maxima to Gumbel distribution; the non-attraction to Fréchet distribution is left as an exercise.
4. If we have $F(x)$ such that there exist α and $\beta > 0$ for which $e^{\alpha+\beta x}(1 - F(x)) \to 1$ as $x \to +\infty$ $(\overline{w} = +\infty)$, then $F(x)$ is attracted for maxima to $\Lambda(x)$ as seen from either that $\frac{1-F(y)}{1-F(y+t)} \to e^{\beta t}$ and so $\delta_k = \beta^{-1}$ with λ_k to be determined. By observing that $e^{\alpha+\beta}(\lambda_k + \delta_k\,x)\,(1 - F(\lambda_k + \delta_k\,x)) \to 1$ and also that we should have $k(1 - F(\lambda_k + \delta_k\,x)) \to e^{-x}$ which implies $e^{\alpha+\beta(\lambda_k+\delta_k\,x)}/k \to e^x$ or $e^{\alpha+\beta\lambda_k}/k \to 1$ and $e^{\beta\delta_k\,x} \to e^x$ giving thus $\lambda_k = \frac{\log k-\alpha}{\beta}$, $\delta_k = \frac{1}{\beta}$ as a system of attractions coefficients.
5. Consider, finally, the lognormal distribution $LN(x) = N(\log x)$ for $x \geq 0$, $LN(x) = 0$ for $x \leq 0$. As $\overline{w} = +\infty$ we can try the attraction for maxima to $\Lambda(x)$, or $\Phi_\alpha(x)$. Let us use the von Mises criterion: $\frac{LN''(x)(1-LN(x))}{(LN'(x))^2} \to -1$ as happened before for the normal tail.

3.4 Tail equivalent distributions

The subject of equivalence of distributions is relevant, like the use of attraction conditions, to evaluate what can be the asymptotic behaviour of maxima (or minima) of some distribution by substituting it with another one which is easier to manipulate.

We say that $F(x)$ and $F_1(x)$ are *tail equivalent for maxima* if $\overline{w}(F) = \overline{w}(F_1) = \overline{w}$ and $\frac{1-F(x)}{1-F_1(x)} \to 1$ as $x \uparrow \overline{w}$, and *tail equivalent for minima* if $\underline{w}(F) = \underline{w}(F_1) = \underline{w}$ and $\frac{F(x)}{F_1(x)} \to 1$ as $x \downarrow \underline{w}$; we could, instead of the limit 1, put a constant $c(0 < c < +\infty)$, thus substituting the equality of limits for a power relation using c. This is left as an exercise.

Let us deal with maxima.

If $F(x)$ and $F_1(x)$ are tail equivalent for maxima, then if one of the distributions is attracted to $\tilde{L}(x)$ the other distribution is also attracted to the same limit $\tilde{L}(x)$ and with the same coefficients.

Suppose that $F^k(\lambda_k + \delta_k\, x) \to \tilde{L}(x)$ or $k\big(1 - F(\lambda_k + \delta_k\, x)\big) \to -\log \tilde{L}(x)$; as $\lambda_k + \delta_k\, x \to \overline{w}$ we have $k\big(1 - F_1(\lambda_k + \delta_k\, x)\big)\frac{1-F(\lambda_k+\delta_k\, x)}{1-F_1(\lambda_k+\delta_k\, x)} \to -\log \tilde{L}(x)$ and by the condition of equivalence $k(1 - F_1(\lambda_k + \delta_k\, x)) \to -\log \tilde{L}(x)$ or $F_1^k(\lambda_k + \delta_k\, x) \to \tilde{L}(x)$ as said. Let us prove the converse: if $F^k(\lambda_k + \delta_k\, x) \to \tilde{L}(x)$ and $F_1^k(\lambda_k + \delta_k\, x) \to \tilde{L}(x)$(the same coefficients) we have $\frac{1-F(\lambda_k+\delta_k\, x)}{1-F_1(\lambda_k+\delta_k\, x)} \to 1$ as $\lambda_k + \delta_k\, x \to \overline{w}$ we get $\frac{1- F(y)}{1 - F_1(y)} \to 1$ as $x \uparrow \overline{w}$.

The statement for minima is obvious:

If $F(x)$ and $F_1(x)$ are tail equivalent for minima and one of them is attracted to $\underset{\sim}{L}(x)$, then the other is also attracted to $\underset{\sim}{L}(x)$ with the same coefficients and conversely.

It is immediate that the logistic, the exponential, and Gumbel distributions are tail equivalent for maxima, which shows that in many cases attraction can simply be obtained by seeking equivalence to one of the $\tilde{L}(x)$; the Cauchy distribution is tail equivalent to $\Phi_1(\pi\, x)$. But the normal distribution, although attracted to the Gumbel distribution, is not tail equivalent to it (not the logistic or the exponential); for more details see Resnick (1971) and Tiago de Oliveira and Epstein (1972).

3.5 The convergence of quantiles

Consider a sequence of distribution functions $\{F_k(x)\}$ such that $F_k(x) \overset{w}{\rightarrow} L(x), L(x)$ being continuous and, thus, the convergence being uniform. For large k, given $\in$, we know that

$$L(x) - \in \le F_k(x) \le L(x) + \in$$

and so

$$F_k^{-1}(L(x) - \in) \le x \le F_k^{-1}(L(x) + \in) \text{ and with } L(x) = p$$

we get

$$F_k^{-1}(p - \in) \le L^{-1}(p) \le F_k^{-1}(p + \in).$$

Taking in the RHS inequality $p - \in$ instead of p and in the LHS inequality $p + \in$ instead of p, we get

$$L^{-1}(p - \in) \le F_k^{-1}(p) \le L^{-1}(p + \in);$$

and so, by the continuity of $L(x)$, we get

$$F_k^{-1}(p) \rightarrow L^{-1}(p).$$

Applying the previous results to maxima, where the $\tilde{L}(x)$ are continuous, we see that, as $F_k(x) = F^k(\lambda_k + \delta_k x)$, we get for

$$\chi_p = \tilde{L}^{-1}(p), \frac{F^{-1}(p^{1/k}) - \lambda_k}{\delta_k} \rightarrow \chi_p$$

which shows the practical result $F^{-1}(p^{1/k}) \approx \lambda_k + \delta_k \chi_p$.

For instance if $F(x) = 1/(1 + e^{-x})$ is the standard logistic, attracted for maxima to Gumbel distribution, as we can take $\lambda_k = \log k$ and $\delta_k = 1$ we get $F^{-1}(p^{1/k}) \approx \log k + \chi_p, \chi_p = \Lambda^{-1}(p) = -\log(-\log p)$; the probability error is $|F^k(\log k + \chi_p) - p| = |(1 - \frac{\log p}{k})^{-k} - p| \rightarrow 0$ as $k \rightarrow +\infty$ and the linear error is $|(\log + \chi_p) - F^{-1}(p^{1/k})| = |\log \frac{k(p^{-1/k} - 1)}{-\log p}| \rightarrow 0$ as $k \rightarrow +\infty$.

For the exponential we have, in the same conditions $\lambda_k = \log k, \delta_k = 1$, the probability error is $|F^k(\log k + \chi_p) - p| = |(1 + \frac{\log p}{k})^k - p| \to 0$ as $k \to \infty$, and also the linear error $|F^{-1}(\log k + \chi_p) - (p^{1/k})| = |\log \frac{k(1-p^{1/k})}{-\log p}| \to 0$ as $k \to +\infty$.

3.6 Speed of convergence

The way a sequence of distributions of maxima converge to its limit is a very important question: either it converges quickly to the limit and this limit can be used as an approximation to the real distribution, or the approach is slow and the limit, from the statistical standpoint, has little relevance: if an error of $\in = 10^{-2}$ is obtained in one case for $k = 50$ the approximation can be used for moderate samples – the approximation for small samples being practically speaking an illusion – but if the error $\in = 10^{-2}$ is a attained only for $k \geq 10^6$ the result has no practical use.

Suppose tha $F^k(\lambda_k + \delta_k x) \to \tilde{L}(x)$. We can think of two different approaches, briefly touched upon in the examples of the preceding section.

We may be interested in $p_k = \sup_x |F^k(\lambda_k + \delta_k x) - \tilde{L}(x)|$: this maximum probability error gives an evaluation of the computation of the probability of overpassing $\lambda_k + \delta_k x$ —i.e. of $1 - F^k(\lambda_k + \delta_k x)$ — by evaluating it by $1 - \tilde{L}(x)$. If $p_k \leq \eta$ and χ_p is p-quantile of $\tilde{L}(x)$ we see that $p - \eta < F^k(\lambda_k + \delta_k \chi_p) < p + \eta$, and if η is very small in comparison with p (in general close to 1) we have good approximations to design, etc. It is evident that $p_k = p_k(\lambda_k + \delta_k)$ and so an open question is to determine the best (λ_k, δ_k), i.e., the values that minimize $p_k(\lambda, \delta)$.

The other error — the linear one — needs some care in its definition. The idea is to study the difference of the quantiles of $F^k(\lambda_k + \delta_k x)$ and of $\tilde{L}(x)$, i.e., to compute $d_k = \sup_x |\lambda_k + \delta_k Q(p^{1/k}) - \chi_p|$. But a simple example shows that this definition can lead to results of no practical use. Suppose $F(x)$ is the exponential distribution which (with $\lambda_k = \log k, \delta_k = 1$) is attracted for maxima to $\Lambda(x)$. As $F^k(\lambda_k + \delta_k x) \to \Lambda(x)$ we have, approximately $F^k(y) \approx \Lambda(\frac{y-\lambda_k}{\delta_k})$ and so the exact p-quantile is $Q(p^{1/k}) = -\log(1 - p^{1/k})$ and the approximate one is $\lambda_k + \delta_k \Lambda^{-1}(p) = \log k - \log(-\log p)$. The maximum resulting linear error is then $d_k = \sup_p |\log k - \log(-\log p) + \log(1 -$

$\mathrm{p}^{1/\mathrm{k}})| = \sup_{\mathrm{p}} |\log \frac{\mathrm{k}(1-\mathrm{p}^{1/\mathrm{k}})}{-\log \mathrm{p}}| = +\infty$, the value attained when $\mathrm{p} \to 0$. A careful study of the linear error (dependent also on (λ_k, δ_k)) was not made but it leads to the computations being made in a shorter interval $\in \leq \mathrm{p} \leq 1 - \in'$ chiefly because, for maxima, we are essentially interested in values of p close to 1 and not to 0.

The relative probability error $\sup_{x} |\frac{F^k(\lambda_k + \delta_k x) - \tilde{L}(x)}{1 - \tilde{L}(x)}| = \sup_{x} |\frac{1 - F^k(\lambda_k + \delta_k x)}{1 - \tilde{L}(x)} - 1|$ for the interval $0 < \tilde{L}(x) < 1$ is also an open problem.

Although the more general results are due to Davis (1982) and Tiago de Oliveira (1991) we will only describe the statements of Galambos (1978); some important results are the ones of Balkema, de Haan and Resnick (1984), Galambos (1984) and Beirlant and Willekens (1990).

If $F^k(\lambda_k + \delta_k\, x) \to \tilde{L}(x)$ *denote by* $z_k(x) = k(1 - F(\lambda_k + \delta_k\, x))$ *and for* $x > \underline{w}$, $\rho_k(x) = z_k(x) + \log \tilde{L}(x)$: *then for* $x > \underline{w}$ *and* $z_k(x) \leq k/2$ *we have*

$$|F^k(\lambda_k + \delta_k\, x) - \tilde{L}(x)| \leq \tilde{L}(x)\, [r_{1,k}(x) + r_{2,k}(x) + r_{1,k}(x)\, r_{2,k}(x)]$$

where

$$r_{1,k}(x) = \frac{2\, z_k^2\,(x)}{k} + \frac{2\, z_k^4\,(x)}{k^2} \frac{1}{1-q}$$

$$r_{2,k}(x) = |\rho k(x)| + \frac{\rho_k^2\,(x)}{2} . \frac{1}{1-s}$$

with $q < 1$, $s < 1$ *such that* $z_k^2\,(x) \leq 3k\, q/2$ *and* $|\rho k(x)| < 3\, s$.

As can be seen, this statement does not apply for all x but only for a part of the domain of F(x) although is valid for all admissible sets $\{(\lambda_k, \delta_k)\}$.

The dual statement for minima is:

If $1 - (1 - F(\lambda_k + \delta_k\, x))^k \to \underset{\sim}{L}(x)$ *denote by* $z_k(x) = k\, F(\lambda_k + \delta_k\, x)$ *and for* $< \overline{w}, \rho_k(x) = z_k^2\,(x) + \log(1 - \underset{\sim}{L}(x))$: *then for* $x < \overline{w}$, $z_k^2\,(x) \leq k/2$ *we have*

$|1-(1-F(\lambda_k+\delta_k\,x))^k \to \underset{\sim}{L}(x)| \le \left(1-\underset{\sim}{L}(x)\right)\left[r_{1,k}(x)+r_{2,k}(x)+r_{1,k}(x), r_{2,k}(x)\right]$ *$r_{1,k}$ and $r_{2,k}$ having the same definition as before.*

Davis (1982) gave a different approach in probability error evaluation using, essentially, the approximation $-\,k\log F^k(\lambda_k+\delta_k\,x) \sim k(1-F(\lambda_k+\delta_k\,x))$ in the interval $0<\tilde{L}(x)<1$. In Tiago de Oliveira (1991) we sketched a similar result, but with a more direct approach, that we will explain.

As $a^k-b^k=(a-b)\sum_{j=0}^{k-1}a^j\,b^{k-1-j}$ we have $|a^k-b^k|\le k|a-b|\max(a^{k-1},\ b^{k-1})$ for $0\le a\,,\,b\le 1$. If $F^k(\lambda_k+\delta_k\,x)\to\tilde{L}(x)$, as we see that $\left|F^k(\lambda_k+\delta_k\,x)-\tilde{L}(x)\right| = |F^k(\lambda_k+\delta_k\,x)-(\tilde{L}^{1/k}(x))^k| \le k|F(\lambda_k+\delta_k\,x)-\tilde{L}^{1/k}(x)|\times\max(F^{k-1}(\lambda_k+\delta_k\,x)\,,\tilde{L}^{1^{-1/k}}(x)) \le k|F(\lambda_k+\delta_k\,x)-\tilde{L}^{1/k}(x)|$. As the third factor (max) in the before last expression converges to $\tilde{L}(x)$ we could substitute it by $\tilde{L}(x)$ but this is practically irrelevant because we are interested in the large quantiles ($\tilde{L}(x)\approx 1$).

This basic result is $|F^k(\lambda_k+\delta_k\,x)-\tilde{L}(x)| \le k|F(\lambda_k+\delta_k\,x)-\tilde{L}^{1/k}(x)|$, *the RHS being the principal part of the error, giving thus the order of convergence of* $|F^k(\lambda_k+\delta_k\,x)-\tilde{L}(x)|$ *to zero.*

In the interval $0<\tilde{L}(x)<1$, introducing $\tau_k(x)=F(\lambda_k+\delta_k\,x)/\tilde{L}^{1/k}(x)-1$, where, as it is immediate, $k\,\tau_k(x)\to 0$ we can give a formulation analogous to the one of Davis (1982) $|F^k(\lambda_k+\delta_k\,x)-\tilde{L}(x)\le k|\tau_k(x)|\tilde{L}(x)\max((1+\tau_k(x))^{k-1},1)$ and we see, once more, that the order of convergence of $F^k(\lambda_k+\delta_k\,x)$ to $\tilde{L}(x)$ is the one of $k\ \tau_k(x)\to 0$.

For the relative error of the tail evaluation we have

$$|\frac{1-F^k(\lambda_k+\delta_k\,x)}{1-\tilde{L}(x)}-1| \le k|\frac{1-F(\lambda_k+\delta_k\,x)}{1-\tilde{L}^{1/k}(x)}-1|\frac{1-\tilde{L}^{1/k}(x)}{1-\tilde{L}(x)}$$
$$\le |\frac{1-F(\lambda_k+\delta_k\,x)}{1-\tilde{L}^{1/k}(x)}-1|(1+\frac{1-1/k}{2}1-\tilde{L}(x))$$

using the development of $1-\tilde{L}^{1/k}(x)$ in the (alternating) Taylor series on $1-\tilde{L}(x)$.

Davis (1982) essential result can be obtained from $\left|F^k(\lambda_k+\delta_k\,x)-\tilde{L}(x)\right| \le k\left|F(\lambda_k+\delta_k\,x)-\tilde{L}^{1/k}(x)\right|$, as $\tilde{L}^{1/k}(x)=\exp(\frac{1}{k}\log\tilde{L}(x))=1+$

$\frac{1}{k}\log\tilde{L}(x)+O(\frac{1}{k^2})$, under the form $|F^k(\lambda_k+\delta_k\,x)-\tilde{L}(x)|\le k|F(\lambda_k+\delta_k\,x)+\log\tilde{L}(x)+O(1/k)|$, which shows that the order of convergence is, at most, $O(1/k)$ and is of that order only if $k(1-\lambda_k+\delta_k\,x)+\log\tilde{L}(x)=O(1/k)$. The convergence is thus slow in general.

Finally it should be noted that in some cases, as for the normal distribution, a sequence of von Mises-Jenkinson forms $G(z|\theta_m)$, with $\theta_m\to 0$ conveniently chosen, can give a better approximation to $F^k\left(\lambda_k^{'}+\delta_k^{'}\,x\right)$ $((\lambda_k^{'},\delta_k^{'})$ also convenient) than $G(z|0)=\Lambda(z)$, the Gumbel distribution. Although theoretically very interesting this point it has a small statistical interest because we can make the statistical choice of a distribution that fits better the data (see Chapter 4 and 8).

In Part 1 we used the traditional notation where $k>0$ (integer) is the index of a sequence; in the next parts, except for the probabilistic chapter of Part 3, we will use $n(>0)$ integer, which will be the sample size.

References

Balkema, A. A., and de Haan, L., 1972. On the von Mises condition for the domain of attraction of $\exp(-e^{-x})$. *Ann. Math. Statist.*, 43, 1352-1354.

Balkema, A. A., de Haan, L., and Resnick, S., 1984. Uniform rates of convergence to extreme value distributions. *Statistical Extremes and Applications*, J. Tiago de Oliveira ed., 341-346, D. Reidel Dordrecht.

Beirlant, J., and Willekens, E., 1990. Rapid variation with remainder and rates of convergence. *Adv. Appl. Probab.*, 22, 787-801.

Davis, R. A., 1982. The rate of convergence in distribution of the maxima. *Statist. Neerl.*, 36(1), 31-35.

Galambos, J., 1978. *The Asymptotic Theory of Extreme Order Statistics*, Wiley, New - York; (1987), 2nd ed., Krieger Publ.

Galambos, J., 1984. Introduction, order statistics, exceedances, laws of large numbers. in *Statistical Extremes and Applications*, J. Tiago de Oliveira ed., 13-17.

Gnedenko, B. V., 1943. Sur la distribution limite du terme maximum d′ une série aléatoire. *Ann. Math.*, 44, 423-453.

de Haan, L., 1970. On regular variation and its applications to weak convergence of sample extremes. *Math. Centre Tracts,* 32, Amsterdam.

de Haan, L., 1971. A form of regular variation and its applications to the domain of attraction of the double exponential distribution, *Z. Wahrsch. verw. Gebiete*, 17, 241-258.

de Haan, L., 1976. Sample extremes: an elementary introduction. *Statist. Neerl.*, 30, 161-172.

Marcus, M., and Pinsky, M., 1969. On the domain of attraction of $\exp(-e^{-x})$. *J. Math. Analysis Appl.*, 28, 440-449.

Mejzler, D., 1949. On a theorem of B. V. Gnedenko. *Sb. Trudov* Inst. *Math. Akad. Nauk. SSR*, 12, 31-35 (in russian).

Resnick, S. I., 1971. Tail equivalence and its applications. *J. Appl. Prob.*, 8, 136-156.

Tiago de Oliveira, J., and Epstein, B., 1972. Tail comparison for maxima: applications to mixtures. Unpubl. paper, Technion.

Tiago de Oliveira, J., 1991. Perspectivas sobre a estatistica de extremos e problemas em aberto, to be publ. in Rev. Real Acad. *Ciencias Exactas*, *Fisicas y Naturales*, Madrid.

*****