

## Chapter III

### ATTRACTION STRUCTURE AND SPEED OF CONVERGENCE

#### 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 even if they were independent,  $F(x)$  would not be known (in general) in order to analyse if the maxima are attracted to  $\Psi_\alpha$ ,  $\Lambda$  or  $\Phi_\alpha$  and thus to obtain a set (not necessarily the best in any sense) of attraction coefficients  $(\lambda_k, \delta_k)$ . 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  $\underline{L}(x)$  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  $(\lambda_k, \delta_k)$  and in some cases  $\alpha$ , 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.

### 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  $F \in \mathcal{D}(\bar{L})$ ; the dual statements of  $F \in \mathcal{D}(\underline{L})$  can be obtained by a simple conversion of the previous results. The quantile function  $Q(v) = \inf\{x | F(x) \geq v\}$ ;  $Q(v)$  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 V to Part II).

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

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

-  $F$  is attracted, for maxima, to the Gumbel distribution  $\Lambda(x)$  iff

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

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

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

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

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  $\bar{w} = +\infty$  and  $\frac{1-F(t)}{1-F(tx)} \rightarrow x^\alpha$  as  $t \rightarrow +\infty$ ; a system of attraction coefficients is  $(\lambda_k, \delta_k) = (0, Q(1-1/k))$  (ultimately positive as  $\bar{w} = +\infty$ ).

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

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

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

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

If we do use the integrated von Mises-Jenkinson form we can say that  $F(\cdot)$  is attracted, for maxima, to  $G(\cdot|\theta)$  iff  $\frac{Q(1-tx) - Q(1-t)}{Q(1-ty) - Q(1-t)} \rightarrow \frac{x^{-\theta} - 1}{y^{-\theta} - 1}$  (to  $\frac{\log x}{\log y}$  if  $\theta = 0$ ) as  $t \rightarrow 0^+$  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  $SV(tx)/SV(x) \rightarrow 1$  as  $x \rightarrow +\infty$ ; the latter nomenclature can be used to say that  $F(x)$  is attracted to the Weibull distribution if  $\bar{w} < +\infty$  and  $SV(x) = x^\alpha(1 - F(\bar{w} - 1/x))$  is slowly varying at  $+\infty$  and that  $F(x)$  is attracted to the Fréchet distribution if  $\bar{w} = +\infty$  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  $\bar{w} = +\infty$  and to obtain the Weibull distribution as a limit we must have  $\bar{w} < +\infty$ ; the Gumbel distribution can be attained either with  $\bar{w} < +\infty$  or  $\bar{w} = +\infty$ ;
- 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  $F \in \mathcal{D}(L)$ , 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  $\frac{F(\underline{w} + tx)}{1 - F(\underline{w} + x)} \rightarrow t^\alpha$  as  $x \rightarrow 0^+$ ; 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

$\frac{Q(tx) - Q(t)}{Q(e^t) - Q(t)} \rightarrow \log x$  as  $x \rightarrow 0^+$ , or  $\frac{Q(1-v^x) - Q(1-v)}{Q(1-v^e) - Q(1-v)} \rightarrow \log x$  as  $v \rightarrow 1^-$ ;  
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  $\bar{w} = -\infty$  and  $\frac{F(t)}{F(tx)} \rightarrow x^\alpha$  as  $t \rightarrow -\infty$ ; a system of attraction coefficients is  $(0, -Q(1/k))$  ( $-Q(1/k)$  is ultimately positive as  $\bar{w} = -\infty$ ).

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

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

Consider now  $k(1 - F(\delta_k x))$  for  $x > 0$ . As  $\lim k(1 - F(\delta_k x)) = \lim \left\{ k(1 - F(\delta_k)) \frac{1 - F(\delta_k x)}{1 - F(\delta_k)} \right\}$  and by hypothesis when  $k \rightarrow +\infty$ ,  $\delta_k \rightarrow +\infty$  the second factor

$\frac{1 - F(\delta_k x)}{1 - F(\delta_k)} \rightarrow x^{-\alpha}$ ; it remains to show that  $k(1 - F(\delta_k)) \rightarrow 1$ . By the definition of  $\delta_k$  one gets  $F(\delta_k^-) \leq 1 - 1/k = F(\delta_k)$  and so  $k(1 - F(\delta_k)) \leq 1$ ; but  $F(\delta_k x) \leq F(\delta_k^-) \leq 1 - 1/k$  for  $0 < x < 1$

and so  $\frac{1 - F(\delta_k)}{1 - F(\delta_k x)} \leq k(1 - F(\delta_k))$  and  $\frac{1 - F(\delta_k)}{1 - F(\delta_k x)} \rightarrow x^\alpha$  and thus  $k(1 - F(\delta_k)) \rightarrow 1$

with  $k \rightarrow \infty$ , 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  $F^k(\lambda_k + \delta_k x) \rightarrow \Phi_\alpha(x)$  or, equivalently,  $k(1 - F(\lambda_k + \delta_k x)) \rightarrow x^{-\alpha}$  we have also

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

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

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

from which, with

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

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

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  $k(1) = [k\beta]$ ,  $k(s+1) = [k(s) \cdot \beta]$ , and so we get  $\lambda_{k(1)} = \lambda_k$ ,  $\delta_{k(s)} = \delta_k \beta^{(s-1)/\alpha}$  ( $\rightarrow +\infty$  as  $\beta > 1$ ). Then  $\lambda_{k(1)}/\delta_{k(s)} \rightarrow 0$  and we can write, as  $\lambda_{k(s)}$  can be taken to be zero,

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

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

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

and so

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

and as

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

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

$$\frac{1}{\beta} t^\alpha \leq \lim_{y \rightarrow \infty} \frac{1 - F(y)}{1 - F(y t)} \leq \beta t^\alpha$$

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

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

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

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

Consequently for  $F$  to be attracted, for maxima, to  $\Psi_\alpha(x)$  we must have  $\bar{w} < +\infty$ , satisfy the attraction condition  $\frac{1 - F(\bar{w} - t/x)}{1 - F(\bar{w} - t)} \rightarrow x^\alpha$  as  $t \rightarrow 0^+$ , and take as attraction coefficients

$$\lambda_k = \bar{w}, \delta_k = \bar{\delta}_k^{-1}, \text{ or, more simply, } \delta_k = \bar{w} - Q(1-1/k).$$

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

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

$$\frac{\lambda_{[k t]} - \delta_{[k t]} \log t - \lambda_k}{\delta_k} \rightarrow 0$$

and  $\delta_{[k t]} / \delta_k \rightarrow 1,$

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

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

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

the last evidently defining  $\delta_k$ , asymptotically.

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

$$\frac{\lambda_{[k t]} - \lambda_k}{\delta_k} \sim \frac{\lambda_{[k t]} - \lambda_k}{\lambda_{[k e]} - \lambda_k} \rightarrow \log t \quad \text{with } \delta_k \sim \lambda_{[k e]} - \lambda_k,$$

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

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

$$\lambda_k^* + \delta_k^* (\log t - \epsilon) < \lambda_k^* = Q\left(1 - \frac{1}{[k t]}\right) < \lambda_k^* + \delta_k^* (\log t + \epsilon),$$

and so  $F(\lambda_k^* + \delta_k^* (\log t - \epsilon)) \leq F(Q(1 - 1/[k t])) \leq F(\lambda_k^* + \delta_k^* (\log t + \epsilon))$ .

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

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

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

$$e^{-1/t} \leq \liminf F^k(\lambda_k^* + \delta_k^* (\log t + \epsilon))$$

and so  $\exp\{-e^{-(z-\epsilon)}\} \leq \liminf F^k(\lambda_k^* + \delta_k^* z)$ .

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

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

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

$$\liminf F^k(\lambda_k^* + \delta_k^* (\log t - \epsilon)) \leq e^{-1/t}$$

and in the same way

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

Consequently

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

and thus

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

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

$$\frac{Q(1 - 1/[k x]) - Q(1 - 1/k)}{Q(1 - 1/[k e]) - Q(1 - 1/k)} \rightarrow \log x \text{ as } k \rightarrow \infty$$

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

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

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

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

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

by division we obtain

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

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



$$\frac{\lambda_{[(k-1)/x]}^* - \lambda_{k+1}^*}{\delta_k^*} \leq \frac{Q(1-tx) - Q(1-t)}{\delta_k^*} \leq \frac{\lambda_{[(k/x)+1]}^* - \lambda_{k-1}^*}{\delta_k^*}.$$

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  $\delta_{[kx]}/\delta_k \rightarrow 1$  and so  $\delta_{[k/x]}^*/\delta_k^* \rightarrow 1$ .

Thus the RHS of the inequality converges to  $-\log x$  and the same can be proved for the LHS. Thus  $\frac{Q(1-tx) - Q(1-t)}{\delta_k^*} \rightarrow -\log x$  and so the continuous condition for attraction is

$$\frac{Q(1-tx) - Q(1-t)}{Q(1-tx) - Q(1-t)} \rightarrow \log x \text{ as } t \rightarrow 0^+.$$

This condition was given by Mejlzer (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.

### 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  $\bar{w} < +\infty$ ,  $F(x)$  has positive density  $F'(x)$  in some interval  $]x_0, \bar{w}[$ , and  $\lim_{x \uparrow \bar{w}} \frac{(\bar{w} - x) F'(x)}{1 - F(x)} = \alpha$ , then  $F(x)$  is attracted to  $\Psi_\alpha(x)$  and we can use the coefficients  $\lambda_k = \bar{w}$ ,  $\delta_k = \bar{w} - Q(1 - 1/k)$ , as  $F(x)$  defines uniquely  $Q(1 - 1/k) (\rightarrow \bar{w})$  for large  $k$ ;

— If  $F''(x) < 0$  for some interval  $]x_0, \bar{w}[$ , ( $\bar{w} < +\infty$  or  $\bar{w} = +\infty$ ) and  $\frac{d}{dx} (\mu(x)^{-1}) \rightarrow 0$ , where  $\mu(x)$  is the force of mortality or hazard rate, or  $\frac{F''(x)(1 - F(x))}{(F'(x))^2} \rightarrow -1$  as  $x \uparrow \bar{w}$  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  $\bar{w} = +\infty$ ,  $F(x)$  has a positive density for  $x > x_0$ , and  $\frac{x F'(x)}{1 - F(x)} \rightarrow \alpha$ , as  $x \rightarrow +\infty$ , 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  $\delta_k / \lambda_k \rightarrow 0$ ;

— If  $\bar{w} = +\infty$ , and  $F(x)$  is attracted for maxima to  $\Lambda(x)$  then  $\frac{1 - F(tx)}{1 - F(x)} \rightarrow 0$  as  $t \rightarrow +\infty$ , with  $t > 0$ ; this result implies the MLLN;

— If ...  $x_1 < x_2 < x_k < \dots \bar{w} (\leq +\infty)$  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  $\bar{w} = +\infty$ ,  $F^k(\lambda_k + \delta x) \rightarrow \Lambda(x)$  ( $\delta_k = \delta$  constant) is equivalent to

$$\frac{1 - F(\log t)}{1 - F(\log t x)} \rightarrow x^\alpha (\alpha = 1/\delta) \text{ as } t \rightarrow \infty \text{ or } \frac{1 - F(y)}{1 - F(y + k)} \rightarrow e^{k/\delta} \text{ as } y \rightarrow +\infty;$$

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  $\bar{w} = +\infty$  it cannot be attracted to  $\Psi_\alpha(x)$ . But  $\frac{1 - N(t)}{1 - N(tx)} \sim x^{-1} e^{t^2(x^2-1)/2}$  which does not converge to  $x^\alpha$  and so  $\Phi_\alpha(x)$  is excluded. The Meizler criterion for the convergence to  $\Lambda(x)$  in this example is not very operational. Let us consider the von Mises criterion  $\frac{N''(x)(1 - N(x))}{(N'(x))^2} \rightarrow -1$  as  $x \rightarrow +\infty$ . But  $1 - N(x) \sim N'(x)/x$ ; as used before, and  $N''(x) = -x N'(x)$  and so  $\frac{N''(x)(1 - N(x))}{(N'(x))^2} \rightarrow -1$ , 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  $\frac{1 - F(k)}{1 - F(k-1)} = \frac{1 - F(k-1)}{1 - F(k-1)} = e^\lambda > 1$  and so  $F(x)$  is not attracted to  $\Lambda(x)$ , as opposed to the exponential. As  $\bar{w} = +\infty$  the limit could be  $\Phi_\alpha(x)$  if  $\frac{1 - F(t)}{1 - F(tx)} \rightarrow x^\alpha$  as  $t \rightarrow +\infty$ . But  $\frac{1 - F(t)}{1 - F(tx)} = e^{\lambda\{[tx] - [t]\}} = e^{\lambda\{[tx-r] - [t-r]\}}$ , with  $0 \leq r, r' < 1$ , which does not converge to  $x^\alpha$ , 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  $j = 0, 1, \dots$ . Then  $\frac{1 - F(j)}{1 - F(j-1)} = \frac{1 - F(j-1)}{1 - F(j-1)} = 1 + \frac{p_j}{1 - F(j)} \geq$

$$1 + \frac{\lambda^j/j!}{\sum_{k=j}^{[\infty]} \lambda^k/k!} = 1 + \frac{1}{\sum_0^{[\infty]} \frac{j!}{(r+j)!} \lambda^r} = 1 + \frac{1}{\sum_0^{[\infty]} \binom{x+j-1}{r} \frac{\lambda^r}{r!}} > 1 + \frac{1}{\sum_0^{[\infty]} \frac{\lambda^r}{r!}} = 1 + e^{-\lambda}$$

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  $\alpha$  and  $\beta > 0$  for which  $e^{\alpha+\beta x(1-F(x))} \rightarrow 1$  as  $x \rightarrow +\infty$  ( $\bar{w} = +\infty$ ), then  $F(x)$  is attracted for maxima to  $\Lambda(x)$  as seen from either that  $\frac{1 - F(y)}{1 - F(y+t)} \rightarrow e^{\beta t}$  and so  $\delta_k = \beta^{-1}$  with  $\lambda_k$  to be determined. By observing that  $e^{\alpha+\beta(\lambda_k + \delta_k x)} (1 - F(\lambda_k + \delta_k x)) \rightarrow 1$  and also that we should have  $k(1 - F(\lambda_k + \delta_k x)) \rightarrow e^{-x}$  which implies  $e^{\alpha+\beta(\lambda_k + \delta_k x)/k} \rightarrow e^x$  or  $e^{\alpha+\beta\lambda_k/k} \rightarrow 1$  and  $e^{\beta\delta_k x} \rightarrow e^x$  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  $\bar{w} = +\infty$  we can try the attraction for maxima to  $\Lambda(x)$  or  $\Phi_\alpha(x)$ . Let us use the von Mises criterion:

$$\frac{LN''(x)(1-LN(x))}{(LN'(x))^2} \rightarrow -1 \text{ as happened before for the normal tail.}$$

### 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  $\bar{w}(F) = \bar{w}(F_1) = \bar{w}$  and  $\frac{1-F(x)}{1-F_1(x)} \rightarrow 1$  as  $x \uparrow \bar{w}$ , and *tail equivalent for minima* if  $\underline{w}(F) = \underline{w}(F_1) = \underline{w}$  and  $\frac{F(x)}{F_1(x)} \rightarrow 1$  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  $F^k(\lambda_k + \delta_k x) \rightarrow \tilde{L}(x)$  or  $k(1 - F(\lambda_k + \delta_k x)) \rightarrow -\log \tilde{L}(x)$ ; as  $\lambda_k + \delta_k x \rightarrow \bar{w}$  we have  $k(1 - F_1(\lambda_k + \delta_k x)) \frac{1 - F(\lambda_k + \delta_k x)}{1 - F_1(\lambda_k + \delta_k x)} \rightarrow -\log \tilde{L}(x)$  and by the condition of equivalence  $k(1 - F_1(\lambda_k + \delta_k x)) \rightarrow -\log \tilde{L}(x)$  or  $F_1^k(\lambda_k + \delta_k x) \rightarrow \tilde{L}(x)$  as said. Let us prove the converse: if  $F^k(\lambda_k + \delta_k x) \rightarrow \tilde{L}(x)$  and  $F_1^k(\lambda_k + \delta_k x) \rightarrow \tilde{L}(x)$  (the same coefficients) we have  $\frac{1 - F(\lambda_k + \delta_k x)}{1 - F_1(\lambda_k + \delta_k x)} \rightarrow 1$  and as  $\lambda_k + \delta_k x \rightarrow \bar{w}$  we get  $\frac{1 - F(y)}{1 - F_1(y)} \rightarrow 1$  as  $x \uparrow \bar{w}$ .

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  $\tilde{L}(x)$ , then the other is also attracted to  $\tilde{L}(x)$  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).

### The convergence of quantiles

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

$$L(x) - \epsilon \leq F_k(x) \leq L(x) + \epsilon$$

and so

$$F_k^{-1}(L(x) - \epsilon) \leq x \leq F_k^{-1}(L(x) + \epsilon) \quad \text{and with} \quad L(x) = p$$

we get

$$F_k^{-1}(p - \epsilon) \leq L^{-1}(p) \leq F_k^{-1}(p + \epsilon).$$

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

$$L^{-1}(p - \epsilon) \leq F_k^{-1}(p) \leq L^{-1}(p + \epsilon);$$

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

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

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  $|F^k(\log k + \chi_p) - p| = |(1 + \frac{\log p}{k})^k - p| \rightarrow 0$  as  $k \rightarrow \infty$ , and also the linear error  $|F^{-1}(\log k + \chi_p) - (p^{1/k})| = |\log \frac{k(1 - p^{1/k})}{-\log p}| \rightarrow 0$  as  $k \rightarrow +\infty$ .

### 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  $\epsilon = 10^{-2}$  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  $\epsilon = 10^{-2}$  is attained only for  $k \geq 10^6$  the result has no practical use.

Suppose that  $F^k(\lambda_k + \delta_k x) \rightarrow \tilde{L}(x)$ . 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  $\chi_p$  is  $p$ -quantile of  $\tilde{L}(x)$  we see that  $p - \eta < F^k(\lambda_k + \delta_k \chi_p) < p + \eta$ , and if  $\eta$  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  $(\lambda_k, \delta_k)$ , 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  $F^k(\lambda_k + \delta_k x) \rightarrow \Lambda(x)$  we have, approximately  $F^k(y) \approx \Lambda((y - \lambda_k)/\delta_k)$  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  $d_k = \sup_p |\log k - \log(-\log p) + \log(1 - p^{1/k})| = \sup_p \left| \log \frac{k(1 - p^{1/k})}{-\log p} \right| = +\infty$ , the value attained when  $p \rightarrow 0$ . A careful study of the linear error (dependent also on  $(\lambda_k, \delta_k)$ ) was not made but it leads to the computations being made in a shorter interval  $\epsilon \leq p \leq 1 - \epsilon$  chiefly because, for maxima, we are essentially interested in values of  $p$  close to 1 and not to 0.

$$\text{The relative probability error } \sup_x \left| \frac{F^k(\lambda_k + \delta_k x) - \tilde{L}(x)}{1 - \tilde{L}(x)} \right| = \sup_x \left| \frac{1 - F^k(\lambda_k + \delta_k x)}{1 - \tilde{L}(x)} - 1 \right|$$

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 Hann and Resnick (1984), Galambos (1984) and Beirlant and Willekens (1999).

If  $F^k(\lambda_k + \delta_k x) \rightarrow \tilde{L}(x)$  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

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

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

with  $q < 1$ ,  $s < 1$  such that  $z_k^2(x) \leq 3 k q/2$  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  $1 - (1 - F(\lambda_k + \delta_k x))^k \rightarrow \tilde{L}(x)$  denote by  $z_k(x) = k F(\lambda_k + \delta_k x)$  and for  $x < \bar{w}$ ,  $\rho_k(x) = z_k(x) + \log(1 - L(x))$ : then for  $x < \bar{w}$ ,  $z_k(x) \leq k/2$  we have

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

$r_{1,k}$  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  $|a^k - b^k| \leq k |a - b| \max(a^{k-1}, b^{k-1})$  for  $0 \leq a, b \leq 1$ . If  $F^k(\lambda_k + \delta_k x) \rightarrow \tilde{L}(x)$ , as we see that  $|F^k(\lambda_k + \delta_k x) - \tilde{L}(x)| = |F^k(\lambda_k + \delta_k x) - (\tilde{L}^{1/k}(x))^k| \leq k |F(\lambda_k + \delta_k x) - \tilde{L}^{1/k}(x)| \max(F^{k-1}(\lambda_k + \delta_k x), \tilde{L}^{1-1/k}(x)) \leq k |F(\lambda_k + \delta_k x) - \tilde{L}^{1/k}(x)|$ . 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 ( $\tilde{L}(x) = 1$ ).

This the basic result is  $|F^k(\lambda_k + \delta_k x) - \tilde{L}(x)| \leq k |F(\lambda_k + \delta_k x) - \tilde{L}^{1/k}(x)|$ , 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,  $k \tau_k(x) \rightarrow 0$  we can give a formulation analogous to the one of Davis (1982)  $|F^k(\lambda_k + \delta_k x) - \tilde{L}(x)| \leq k |\tau_k(x)| \tilde{L}(x) \max((1 + \tau_k(x))^{k-1}, 1)$  and we see, once more, that the order of convergence of  $F^k(\lambda_k + \delta_k x)$  to  $\tilde{L}(x)$  to is the one of  $k \tau_k(x) \rightarrow 0$

For the relative error of the tail evaluation we have

$$\begin{aligned} \left| \frac{1 - F^k(\lambda_k + \delta_k x)}{1 - \tilde{L}(x)} - 1 \right| &\leq k \left| \frac{1 - F(\lambda_k + \delta_k x)}{1 - \tilde{L}^{1/k}(x)} - 1 \right| \frac{1 - \tilde{L}^{1/k}(x)}{1 - \tilde{L}(x)} \\ &\leq \left| \frac{1 - F(\lambda_k + \delta_k x)}{1 - \tilde{L}^{1/k}(x)} - 1 \right| \left(1 + \frac{1 - 1/k}{2} (1 - \tilde{L}(x))\right) \end{aligned}$$



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  $|F^k(\lambda_k + \delta_k x) - \tilde{L}(x)| \leq k |F(\lambda_k + \delta_k x) - \tilde{L}^{1/k}(x)|$ , as  $\tilde{L}^{1/k}(x) = \exp\left(\frac{1}{k} \log \tilde{L}(x)\right) = 1 + \frac{1}{k} \log \tilde{L}(x) + O\left(\frac{1}{k^2}\right)$ , under the form  $|F^k(\lambda_k + \delta_k x) - \tilde{L}(x)| \leq k |F(\lambda_k + \delta_k x) + \log \tilde{L}(x) + O(1/k)|$  which shows that the order of convergence is, at most,  $O(1/k)$  and is of that order only if  $k(1 - F(\lambda_k + \delta_k x) + \log \tilde{L}(x)) = O(1/k)$ . 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  $\theta_m \rightarrow 0$  conveniently chosen, can give a better approximation to  $F^k(\lambda'_k + \delta'_k x)$  ( $(\lambda'_k, \delta'_k)$  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 IV and VIII).

In Part I 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 III, we will use  $n (> 0)$  integer, which will be the sample size.

## ANNEX I

### On The "Duality" between Extremes and Sums

For simplicity we will deal with some "duality" between sums (or averages) and maxima, the translation to minima being obvious from the relation  $\min_1^n \{ X_i \} = - \max_1^n \{ - X_i \}$ .

The "duality" is expressed by the two columns in correspondence, where there are various gaps.  $F(\cdot)$ ,  $F(\cdot, \cdot)$ , ... and  $\varphi(\cdot)$ ,  $\varphi(\cdot, \cdot)$ , ... will denote the distribution functions and the characteristic functions.

Sums	Maxima
$S_k = \sum_1^k X_i$	$M_k = \max_1^k \{ X_i \}$
$\varphi_X(t) = M_X(e^{itX})$ : ch. f. of $X$	$F_X(x) = \text{Prob}\{X \leq x\}$ : d. f. of $X$
$X_i$ indep. : $\varphi_{S_k}(t) = \prod_1^k \varphi_i(t)$	$X_i$ indep. : $F_{M_k}(x) = \prod_1^k F_i(x)$
$X_i$ i. i. d. : $\varphi_{S_k}(t) = \varphi^k(t)$	$X_i$ i. i. d. : $F_{M_k}(x) = F^k(x)$
$(X, Y)$ indep: $\varphi_{aX+bY}(t) = \varphi_X(at) \varphi_Y(bt)$	$(X, Y)$ indep: $F_{\max(X+a, Y+b)} = F_X(x-a) F_Y(x-b)$
$(+, \cdot)$	$(\max, +)$
$M(aX + b) = a M(X) + b$	...
$V(aX + b) = a^2 V(X)$	...

If  $\{X_i\}$  i.i.d. have  $\mu, \sigma^2$  then

$$\varphi(t) \rightarrow (e^{-t^2/2})$$

$$(s_k - k\mu)/\sqrt{k} \sigma$$

(Central Limit Theorem); in the general case "sometimes" the ch. f. of the normal law may be substituted by that of an indefinitely divisible law;  $\mu = \varphi'_X(0)/i, \sigma^2 = \varphi''_X(0)^2 - \varphi'''_X(0)$ , in the usual case.

If  $(X, Y)$  has a binormal distribution with standard normal margins and correlation coefficient then  $Z = aX + bY$  has a normal distribution  $N(x/\sigma(a, b))$  with  $\sigma(a, b) = a^2 + b^2 + 2\rho a b$ .

If  $\rho = 0$  (independence),  $\sigma(a, b) = 1$  iff  $a^2 + b^2 = 1$ ; in that case  $C(X, Z) = a$ .

If  $\{X_i\}$  are i.i.d. there "sometimes" exist  $(\lambda_k, \delta_k > 0)$  such that  $\text{Prob}\{(M_k - \lambda_n)/\delta_n \leq x\} = F_k(\lambda_k + \delta_k x) \rightarrow \tilde{L}(x), \tilde{L}(x)$  then being  $\Psi_\alpha(x), \Lambda(x)$  or  $\Phi_\alpha(x)$ ;

for  $\Lambda(x)$  we have  $n(1 - F(\lambda_k)) \rightarrow 1, k(1 - F(\lambda_k + \delta_k)) \rightarrow e^{-1}$  or  $\delta_n \sim 1/n F'(\lambda_n)$ ; there are corresponding results for  $\Phi_\alpha$  and  $\Psi_\alpha$ .

If  $(X, Y)$  has a bivariate distribution with reduced Gumbel margins then  $Z = \max(X - a, Y - b)$  has a Gumbel distribution  $\Lambda(z - \lambda(a, b))$  with  $\lambda(a, b) = \log\{(e^{-a} + e^{-b})k(b - a)\}$

If  $k(w) = 1$  (independence),  $\lambda(a, b) = 0$  if  $e^{-a} + e^{-b} = 1$ ; in that case  $\text{Prob}\{Z \leq X - a\} = \text{Prob}\{X - b \leq X - a\} = e^{-a}$

## ANNEX II

### ON THE KHINTCHINE'S CONVERGENCE OF TYPES THEOREM

As this theorem has been referred to and will be important to the proof of asymptotic univariate results — and to subsequent generalizations — we will give a proof of it. Let  $\{V_k(x)\}$  be any sequence of distribution functions and suppose that for two sets  $\{(\lambda_k, \delta_k), \delta_k > 0\}$  and  $\{(\lambda_k^*, \delta_k^*), \delta_k^* > 0\}$  we have  $V_k(\lambda_k + \delta_k x) \xrightarrow{w} L(x)$  and  $V_k(\lambda_k^* + \delta_k^* x) \xrightarrow{w} L^*(x)$ . Then:

$L^*(x) = L(A + Bx)$  with  $(\lambda_k^* - \lambda_k) / \delta_k \rightarrow A$ ,  $\delta_k^* / \delta_k \rightarrow B$  ( $-\infty < A < +\infty$ ,  $0 < B < +\infty$ ) and conversely.

Essentially the theorem says that no sequence of distribution functions is attracted or converges after convenient positive linear transformations to distribution functions not belonging to the same type, if the limits are proper and non-degenerate.

We will closely follow Feller (1966) for the proof. Assume that  $(\lambda_k^* - \lambda_k) / \delta_k \rightarrow A$ ,  $\delta_k^* / \delta_k \rightarrow B$  and we will show that  $L^*(x) = L(A + Bx)$ . In fact, given  $\varepsilon > 0$ , for  $k$  sufficiently large, we have

$$V_k(\lambda_k + \delta_k(A + Bx - \varepsilon)) \leq V_k(\lambda_k^* + \delta_k^* x) \leq V_k(\lambda_k + \delta_k(A + Bx + \varepsilon)).$$

Letting  $k \rightarrow +\infty$  and at the points  $x$  that are continuity points of  $L^*(x)$  and  $A + Bx \pm \varepsilon$  that are continuity points of  $L(x)$  — which excludes at most a denumerable set — we get

$$L(A + Bx - \varepsilon) \leq L^*(x) \leq L(A + Bx + \varepsilon).$$

Letting now  $\varepsilon \rightarrow 0$  we get  $L^*(x) = L(A + Bx)$  except possibly at a denumerable set and thus everywhere owing to the right-continuity of distribution functions.

Let us now assume that  $L^*(x) = L(A + Bx)$ . As  $L$  and  $L^*$  are proper and non-degenerate, choose two continuity points  $x' < x''$  of  $L^*(x)$  ( $0 < L^*(x') < L^*(x'') < 1$ ); then  $A + Bx'$  and  $A + Bx''$  are continuity points of  $L(x)$ . For sufficiently large  $k$  we have from  $V_k(\lambda_k + \delta_k x) \rightarrow L(x)$  and  $V_k(\lambda_k^* + \delta_k^* x) \rightarrow L^*(x)$  that there exist also continuity points  $y' < y''$  of  $L(x)$  such that

$$V_k(\lambda_k + \delta_k y') < V_k(\lambda_k^* + \delta_k^* x') \leq V_k(\lambda_k^* + \delta_k^* x'') < V_k(\lambda_k + \delta_k y'')$$

and so 
$$\lambda_k + \delta_k y' < \lambda_k^* + \delta_k^* x' < \lambda_k^* + \delta_k^* x'' < \lambda_k + \delta_k y''$$

Thus  $(0 <) \frac{\delta_k^*}{\delta_k} < \frac{y'' - y'}{x'' - x'}$  and consequently for some subsequence of  $k$  we have  $\frac{\delta_k^*}{\delta_k} \rightarrow B' (> 0$  because of the non-degeneracy). Also  $y' - \frac{\delta_k^*}{\delta_k} x' < (\lambda_k^* - \lambda_k)/\delta_k < y'' - \frac{\delta_k^*}{\delta_k} x''$  and so, as  $(\lambda_k^* - \lambda_k)/\delta_k$  remains bounded and in a subsequence of the previous one, it has a limit  $A'$ .

Evidently  $A'$  and  $B'$  are uniquely determined and we have  $A' = A, B' = B$ .

For the situation described where  $L^*(x) = L(A + Bx)$  we say that  $L(x)$  and  $L^*(x)$  belong to the same type and the coefficients  $(\lambda_k, \delta_k)$  and  $(\lambda_k^*, \delta_k^*)$  are *equivalent*; if  $L^*(x) = L(x)$  or  $A = 0, B = 1$  will say that they are *totally equivalent*.

## ANNEX III

### On the $\delta$ -Method

Essentially the  $\delta$ -method is based on the fact that if  $\{X_k\}$  and  $\{Y_k\}$  are sequences of random variables such that either  $X_k - Y_k \xrightarrow{P} 0$  or  $X_k / Y_k \xrightarrow{P} 1$  and  $\{X_k\}$  has a limiting proper and non-degenerate distribution, the same happens to  $\{Y_k\}$ ; see Cramér (1946), p. 254-255. Also if  $\{(X_k, Z_k)\}$  and  $\{(Y_k, Z_k)\}$  are sequences of random pairs and  $X_k - Y_k \xrightarrow{P} 0$  or  $X_k / Y_k \xrightarrow{P} 1$  and  $\{(X_k, Z_k)\}$  has a limiting (bivariate) proper and non-degenerate distribution, the same happens to  $\{(Y_k, Z_k)\}$ . The multivariate generalization is obvious. In many applications the  $\delta$ -method leads to a linearization (use of the terms up to the first order of Taylor development). For details see Tiago de Oliveira (1982).

The  $\delta$ -method was underlying some reasoning in this Part and will be used more extensively in the following Parts.

Let us give two examples, used in the text, that help to clarify the use of the  $\delta$ -method.

#### Example 1:

We wish to study the asymptotic behaviour of  $(\bar{x}, s)$  and to show that  $C(\bar{x}, s) \sim \beta_1 \sigma^2/2k$ , as obtained by Gumbel and Carlson (1956).

By the Central Limit Theorem, if  $\{X_i\}$  have moments up to the fourth order, we know that

$$A_k = \sqrt{k} \frac{\frac{1}{k} \sum_{i=1}^k x_i - \mu}{\sigma} \quad \text{and} \quad B_k = \sqrt{k} \frac{\frac{1}{k} \sum_{i=1}^k (x_i - \mu)^2 - \sigma^2}{\sqrt{\beta_2 - 1} \sigma^2}$$

are asymptotically standard normal and the random pair is asymptotically binormal with standard margins and correlation coefficient  $\rho = \frac{\beta_1}{\sqrt{\beta_2 - 1}}$ . But  $B_k$ , as  $x_i - \mu = (x_i - \bar{x}) + (\bar{x} - \mu)$ , takes the form

$$B_k = \sqrt{k} \frac{s^2 + (\bar{x} - \mu)^2 - \sigma^2}{\sqrt{\beta_2 - 1} \sigma^2}$$

Then as  $\sqrt{k} (\bar{x} - \mu)^2 = \sigma^2 A_k^2 / \sqrt{k} \xrightarrow{P} 0$  and so

$$\left( \sqrt{k} \frac{\bar{x} - \mu}{\sigma}, \sqrt{k} \frac{s^2 - \sigma^2}{\sqrt{\beta_2 - 1} \sigma^2} \right)$$

is also an asymptotically binormal pair with standard margins and correlation coefficient  $\rho = \beta_1 / \sqrt{\beta_2 - 1}$ . Also  $s^2 \xrightarrow{P} \sigma^2$  and, thus,  $\frac{s^2 - \sigma^2}{2 \sigma(s - \sigma)} \xrightarrow{P} 1$ .

Consequently  $\left( \sqrt{k} \frac{\bar{x} - \mu}{\sigma}, 2 \sqrt{k} \frac{s - \sigma}{\sqrt{\beta_2 - 1} \sigma} \right)$  is also an asymptotically binormal pair with standard margins, correlation coefficient  $\rho$ , and

$$C(\bar{x}, s) \sim \rho \frac{\sigma}{\sqrt{k}} \cdot \frac{\sqrt{\beta_2 - 1} \sigma}{2 \sqrt{k}} = \frac{\beta_1 \sigma^2}{2k}$$

### Example 2:

Let  $f(x, y)$  be a twice-differentiable function and  $\{(X_k, Y_k)\}$  a sequence of random pairs such that  $\left( \sqrt{k} \frac{X_k - \mu_x}{\sigma_x}, \sqrt{k} \frac{Y_k - \mu_y}{\sigma_y} \right)$  is asymptotically binormal with standard margins and correlation coefficient  $\rho$ . Then:

$\sqrt{k} \{f(X_k, Y_k) - f(\mu_x, \mu_y)\}$  is asymptotically normal with mean value zero and variance  $V$  below.

In fact, as

$$\sqrt{k} \{f(X_k, Y_k) - f(\mu_x, \mu_y)\} = \sqrt{k} \{(X_k - \mu_x) \frac{\partial f}{\partial x} \Big|_{\mu_x, \mu_y} + (Y_k - \mu_y) \frac{\partial f}{\partial y} \Big|_{\mu_x, \mu_y}$$

+ 2nd order terms } ,

as the second order terms are  $o_p(n^{-1/2})$  we see that  $\sqrt{k} \{ f(X_k, Y_k) - f(\mu_x, \mu_y) \}$  has the same asymptotic distribution of the linearized form

$$\sqrt{k} \left\{ (X_k - \mu_x) \frac{\partial f}{\partial x} \Big|_{\mu_x, \mu_y} + (Y_k - \mu_y) \frac{\partial f}{\partial y} \Big|_{\mu_x, \mu_y} \right\}$$

and is thus normal with mean value zero and variance

$$V = \sigma_x^2 \left( \frac{\partial f}{\partial x} \Big|_{\mu_x, \mu_y} \right)^2 + 2 \rho \sigma_x \sigma_y \frac{\partial f}{\partial x} \Big|_{\mu_x, \mu_y} \frac{\partial f}{\partial y} \Big|_{\mu_x, \mu_y} + \sigma_y^2 \left( \frac{\partial f}{\partial y} \Big|_{\mu_x, \mu_y} \right)^2 .$$

The multivariate generalization is obvious.

This result presupposes that not all the first derivatives are null; if this is not the case (i.e., all first derivatives are null) then we have different asymptotic distributions, such as the  $\chi^2$  — see Tiago de Oliveira (1982).



### Exercises — Part I

1.1 — Show that  $\{\min_{1 \leq i \leq k} X_i\} \xrightarrow{P} \underline{w}$ .

1.2 — Study the Cauchy distribution for the LLN.

1.3 — Prove the conditions for the validity of MLLN.

1.4 — Consider a sample of  $n$  observations  $\{X_i\}$  and a possible future sample of  $m$  observations  $\{Y_j\}$ , all of them i.i.d. Show that

$$\text{Prob} \left\{ \min_{1 \leq i \leq n} X_i \leq \min_{1 \leq j \leq m} Y_j \leq \max_{1 \leq j \leq m} Y_j \leq \max_{1 \leq i \leq n} X_i \right\} = \frac{n(n-1)}{(n+m)(n+m-1)}$$

1.5 — Show that  $\max_{1 \leq i \leq k} \{X_i\} \leq a + \max_{1 \leq i \leq k} \{(X_i - a)_+\}$ . If the  $\{X_i\}$  have the margins  $F_i(x)$  then

$$M(\max_{1 \leq i \leq k} X_i) \leq a + \sum_{i=1}^k \int_a^{+\infty} (1 - F_i(x)) dx; \text{ no assumption of independence is made.}$$

1.6 — Obtain the constants  $(\alpha_m, \beta_m > 0)$  that satisfy the stability equation for maxima and for minima.

1.7 — Obtain the transformations interchanging  $\tilde{L}(x)$  between themselves,  $\underline{L}(x)$  between themselves and between  $\underline{L}(x)$  and  $\tilde{L}(x)$ ; compare the different graphs of the densities to see the analogies.

1.8 — Translate maxima results to minima results and vice-versa; use, in particular, the relation between the distribution function and the survival function and with the reduced Gumbel distribution.

1.9 —  $F(x) = 0$  if  $x \leq 0$ ,  $F(x) = 1 - e^{-\sqrt{x}}$  is attracted for maxima to the Gumbel distribution; obtain one system of attraction coefficients.

- 1.10 — Show that if  $\bar{w} = +\infty$ , and for some  $u$  and  $v > 0$ , as  $x \rightarrow +\infty$ ,  $x^u \exp(x^v) (1 - F(x)) \rightarrow c$ ,  $0 < c < +\infty$ , then  $\{(\max_{1 \leq k} X_k - \lambda_k)/\delta_k\}$  is attracted to  $\Lambda(x)$ ; a system of coefficients can be  $\lambda_k = (\log(kc))^{1/v} - \frac{u \log \log(kc)}{v^2 (\log(kc))^{(v-1)/v}}$  and  $\delta_k = \frac{(\log(kc))^{(1-v)/v}}{v}$ .

*J. Villaseñor (1976), Ph.D. Thesis*

- 1.11 — Show that the conditions given for  $F \in \mathcal{D}(\Lambda)$  are equivalent.
- 1.12 — Obtain the continuous condition on  $Q(v)$  for  $F$  to be attracted to the Gumbel distribution of minima.
- 1.13 — The maximum likelihood estimators of  $(\lambda, \delta)$  for the Gumbel distribution  $\Lambda((x - \lambda)/\delta)$  are (Chapter V, Part II)

$$\hat{\delta} = \bar{x} - \frac{\sum_{i=1}^n x_i e^{-x_i/\hat{\delta}}}{\sum_{i=1}^n e^{-x_i/\hat{\delta}}}$$

$$\hat{\lambda} = -\hat{\delta} \log\left(\frac{1}{n} \sum_{i=1}^n e^{-x_i/\hat{\delta}}\right).$$

Show that if  $\Phi(\theta) = \int_{-\infty}^{+\infty} e^{-\theta x} d\Lambda((x - \lambda)/\delta)$  then

$$\delta = -\Phi'(0) + \frac{\Phi''(1/\delta)}{\Phi(1/\delta)}$$

$$\lambda = -\delta \log \Phi(1/\delta),$$

corresponding to the limiting structure of the maximum likelihood estimators.

- 1.14 — If  $F \in \mathcal{D}(\Lambda)$  then  $F^k(x) \approx \Lambda\left(\frac{x - \lambda_k}{\delta_k}\right)$ . Define then  $\Phi_k(\theta)$ . Can this be used to obtain attraction coefficients? Can this be extended to the Weibull and Fréchet distributions?
- 1.15 — Show that Fuller approximations to the return period of overpassing the level  $a$  for the Gumbel and Fréchet distributions for maxima are

$$\tilde{T}(a) = [-\log \Lambda\left(\frac{a - \lambda}{\delta}\right)]^{-1} = e^{(a - \lambda)/\delta} \quad \text{and} \quad \tilde{T}(a) = [-\log \Phi_\alpha\left(\frac{a - \lambda}{\delta}\right)]^{-1} = \left(\frac{a - \lambda}{\delta}\right)^\alpha$$

( $a > \lambda$ ) for large  $a$  and that the return period of underpassing the level  $a$  for the Weibull distribution is  $\tilde{T}(a) = [-\log(1 - W_\alpha(\frac{a-\lambda}{\delta}))]^{-1} = (\frac{a-\lambda}{\delta})^{-\alpha}$  ( $a > \lambda$ ) for small  $a - \lambda$ .

I.16— Show, for the Gumbel distribution, that  $\tilde{T}((1 + \eta)a)$  and  $\tilde{T}(a)$  verify the relation  $\tilde{T}((1 + \eta)a) = e^{\eta \lambda \delta} \tilde{T}^{1 + \eta}(a)$ . This approximation, with a very small relative error, shows the nonsense of the rule “to tame your river and build the dam strong enough to withstand anything, double the largest flood so far observed”. Then  $\eta = 1$ ,  $\tilde{T}(2a) = e^{\lambda \delta} \tilde{T}^2(a) > \tilde{T}^2(a)$ , as  $\lambda > 0$ , and so if the number of years of observation is 50 years, we have the return period  $\tilde{T}(2a) > 2500$  years which is not a realistic planning time.

Note : See in Part II, for each limiting distribution, the status of statistical analysis of extreme data for risk evaluation and control and design planning.

I.17— Using the fact that, for the standard normal distribution,  $\frac{1 - N(x)}{N'(x)/x} \rightarrow 1$  as  $x \rightarrow \infty$ , and defining  $c_n$  by  $c_n = n N'(c_n)$ , show that  $N^n(c_n + x/c_n) \rightarrow \Lambda(x)$ .

I.18— Using Mill's inequalities for the standard normal distribution

$$\left(\frac{1}{x} - \frac{1}{x^3}\right) N'(x) < 1 - N(x) < \frac{1}{x} N'(x) \text{ for } x > 0,$$

show that  $N^k(\lambda_k + \delta_k x) \rightarrow \Lambda(x)$  as  $k \rightarrow \infty$  if  $\lambda_k \delta_k \rightarrow 1$  and  $\frac{k}{\sqrt{2\pi} \lambda_k} e^{-\lambda_k^2/2} \rightarrow 1$  as  $k \rightarrow \infty$ , thus directly obtaining the results in the text.

I.19— Let  $\{X_k\}$  be a sequence of i.i.d. reduced Gumbel random variables. Fix an integer  $m$  and define  $Y_k = \max(Y_i, X_{k+1}, \dots, X_{k+m-1})$ . Show that  $\text{Prob}\{Y_k - \log m \leq x\} = \Lambda(x)$  and, also, that

$$\text{Prob}\left\{\max_{1 \leq i \leq k} (Y_i - \log(k+m-1)) \leq x\right\} = \Lambda(x).$$

Note that the sequence  $\{Y_k\}$  is  $(m-1)$ -dependent.

I.20— Let  $\{X_i\}$  be a sequence of independent random variables such that  $a_i X_i + b_i$  ( $a_i > 0$ ) have the distribution function  $F \in \mathcal{D}(\Lambda)$ ; let a system of attraction coefficients for  $F$  be  $(\lambda_k, \delta_k)$ . Show then, as

$$F(y) = 1 - \frac{\exp(-(y - \lambda_k)/\delta_k) + o(1)}{k}, \text{ that}$$

$$\text{Prob}\{(\max(X_1, \dots, X_k) - \lambda'_k)/\delta'_k \leq x\} \rightarrow \Lambda(x)$$

$$\text{if } \frac{1}{k} \sum_1^k e^{-(\eta_i + \eta_i x)} \rightarrow e^{-x} = 1$$

where  $\xi_i (= \xi_i(k)) = a_i \delta'_k / \delta_k$  and  $\eta_i (= \eta_i(k)) = (a_i \lambda'_k + b_i - \lambda_k) / \delta_k$ .

We can have the same attraction coefficients (i.e.,  $(\lambda'_k, \delta'_k) = (\lambda_k, \delta_k)$ ) if  $\eta_i = 0$  and  $\xi_i = 1$  or  $a_i = 1$  and  $b_i = 0$ , which is the already classical case: but if  $a_i = 1$  and  $b_i = 0$  is false there may be a (non-equivalent) change in the attraction coefficients.

I.21— Compute the mean value and variance of  $\Psi_\alpha$  and  $\Phi_\alpha$  (when they exist) up to  $O(\alpha^{-1})$  and  $O(\alpha^{-2})$  and obtain thus the linear transformation such that the transformed variables converge to the corresponding ones of the reduced Gumbel distribution (whose mean value and variance are  $\gamma$  and  $\frac{\pi^2}{6}$ ); see, for more details, in Part II, the Gumbel distribution.

I.22— Show that if  $Z_\alpha$  has the reduced Weibull distribution for maxima, with shape parameter  $\alpha > 0$ , then  $\alpha(1 + Z_\alpha)$  has the Gumbel distribution for maxima as a limit when  $\alpha \rightarrow \infty$ . Show also that if  $Z_\alpha$  has the reduced Fréchet distribution, with shape parameter  $\alpha > 0$ , then  $\alpha(-1 + Z_\alpha)$  has the Gumbel distribution for maxima as a limit when  $\alpha \rightarrow +\infty$ .

I.23— The graphs suggest that  $\Phi_\alpha(\cdot)$  and  $W_\alpha(z)$  are very close. Show, although, that

$$\sup_{z \geq 0} |W_\alpha(z) - \Phi_\alpha(z)| = |W_\alpha(1) - \Phi_\alpha(1)| = 1 - 2e^{-1} = .2642411,$$

with  $W'_\alpha(1) = \Phi'_\alpha(1)$ . Note that  $\sup_{z \geq 0} |W'_\alpha(z) - \Phi'_\alpha(z)| = +\infty$ , for  $\alpha < 1$  owing to the fact

$W'_\alpha(0) = +\infty$ , but for  $\alpha \geq 1$  we have  $\sup_{z \geq 0} |W'_\alpha(z) - \Phi'_\alpha(z)| < +\infty$ . Compute this

bound.

I.24— As known, for the Fréchet distribution (for maxima) it can be shown that  $\bar{w}_F = +\infty$  and that we can take  $\lambda_k = 0$ , i.e.,  $F^k(\delta_k z) \rightarrow \Phi_\alpha(z)$  as  $k \rightarrow \infty$ . Show that if  $\text{Prob}\{(\max(X_1, \dots, X_k) / \delta_k \leq z)\} = F^n(\delta_n z) \rightarrow \Phi_\alpha(z)$ , with  $k \rightarrow \infty$ , then with  $Y_i = \log(\max(X_i, 0)) = \log(X_i)_+$

we have also  $\text{Prob} \left\{ \alpha \left( \max_1^k Y_i - \log \delta_k \right) \leq z \right\} \rightarrow \Lambda(z)$ . Clearly the result stresses the importance of homogeneous estimators.

I.25— As known, for the Weibull distribution (for minima), it can be shown that  $\underline{w}_F < +\infty$  and that we can take  $\lambda_k = \underline{w}_F$ , i.e.,  $1 - (1 - F(\underline{w}_F + \delta_k z))^k \rightarrow W_\alpha(z)$ , as  $k \rightarrow \infty$ ; Show that if  $Y_i = -\log(X_i - \underline{w}_F)$  then  $\text{Prob} \left\{ \alpha \left( \max_1^k Y_i + \log \delta_k \right) \leq z \right\} \rightarrow \Lambda(z)$  also as  $k \rightarrow \infty$ . Consequently, in estimation, the left-end point is very important.

I.26— Using the relation between the Gumbel and Fréchet distributions, show that  $F^n(\lambda_n + a z) \rightarrow \Lambda(z)$  if and only if  $\frac{1 - F(x)}{1 - F(x + t)} \rightarrow e^{t/a}$  and so  $1 - F(\log x)$  is a slowly varying function with index  $1/a$ .

*Gnedenko (1943)*

I.27— Show, in sequence to the previous exercise, that if  $X$  is normal  $(\mu, \sigma)$  then  $Y = X^2$  is also attracted to Gumbel distribution with  $\delta_k = a = 2\sigma^2$ .

I.28— For the examples 1 to 5 and for the attraction coefficients given in the section "The asymptotic distributions of extremes – some examples" (Chapter II) compute numerically the probability and linear errors for probability levels  $p = .95, .99, .999$  and values of  $k = 10, 100, 1000$ . Study the behaviour of example 6.

I.29— Use attraction conditions to show that both  $F(x) = 0$  if  $x < e$ ,  $F(x) = \frac{1}{\log x}$  if  $x \geq e$ , and the Pascal distribution are not attracted to the Fréchet distribution.

I.30— Obtain the constants  $(\alpha_t, \beta_t > 0)$  such that the stability equation for  $G(z|\theta)$ , i.e.,  $G^t(\alpha_t + \beta_t z|\theta) = G(z|\theta)$  is valid.

I.31— Show using the criteria that  $G(z|\theta)$  is attracted to  $G(z|\theta)$  as could be expected; i.e.,  $\Psi_\alpha$  to  $\Psi_\alpha$ ,  $\Lambda$  to  $\Lambda$  and  $\Phi_\alpha$  to  $\Phi_\alpha$ .

I.32— Show that the quantiles of  $G(z|\theta)$  converge to those of the Gumbel distribution  $\Lambda(z)$  when  $\theta \rightarrow 0^+$  or  $k \rightarrow 0^-$ .

I.33— Study the case where the  $\{X_i\}$  are i.i.d. and  $\{X_i^*\}$ , with  $X_i^* = X_i + (-1)^i$ , are both attracted for maxima (or for minima) to the same  $\tilde{L}(x)$  ( $\tilde{L}(x)$ ).

I.34— Show that if  $\{X_i\}$  are i.i.d. such that  $\bar{w} < +\infty$  and  $\text{Prob} \left\{ \left( \max_1^k X_i - \lambda_k \right) / \delta_k \leq x \right\} \rightarrow \tilde{L}(x)$ , as  $k \rightarrow \infty$ , for  $X_i^* = \frac{1}{\bar{w} - X_i}$  there exist  $\lambda_k^*$  and  $\delta_k^*$  such that  $\text{Prob} \left\{ \left( \max_1^k X_i^* - \lambda_k^* \right) / \delta_k^* \leq x \right\}$  also converges to some  $\tilde{L}(x)$ .

I.35— Obtain conditions such that if  $\{X_i\}$  are independent and such that if  $\{(X_i - \lambda_i) / \delta_i\}$  have the same distribution  $\tilde{L}(x)$ , then  $\max_1^k \{X_i\}$  also has the same asymptotic distribution  $\tilde{L}(x)$ ; compute one system of attraction coefficients, asymptotically.

I.36— Using the fact that, if  $F \in \mathcal{D}(\tilde{L})$  then  $F(\lambda_k + \delta_k x) \rightarrow 1$  and if  $F \in \mathcal{D}(\tilde{L})$  then  $F(\lambda_k + \delta_k x) \rightarrow 0$ , show, by the use of Khintchine's convergence of types theorem, that not only  $(\lambda_{k+1}, \delta_{k+1})$  are also attraction coefficients for the same  $\tilde{L}$  or  $\underline{L}$ , i.e.,  $(\lambda_{k+1} - \lambda_k) / \delta_k \rightarrow 0$  and  $\delta_{k+1} / \delta_k \rightarrow 1$ , but, also, that  $F^{n+k_n}(\lambda_n + \delta_n x) \rightarrow L(x)$  iff  $k_n/n$ ; interpret the result.

I.37— Let  $\{X_i\}$  be i.i.d. random variables with distribution function  $F(x)$  attracted to  $\tilde{L}(x)$  for maxima with attraction coefficients  $\lambda_k, \delta_k > 0$  and to  $\underline{L}(x)$  for minima (with attraction coefficients  $\lambda'_k, \delta'_k > 0$ ); recall that  $\underline{L}(x)$  may be different  $1 - \tilde{L}(x)$ , i.e., that the limiting distributions can not be in correspondence. It was shown that as  $k \rightarrow \infty$

$$\text{Prob} \left\{ \min_1^k X_i \leq \lambda'_k + \delta'_k x, \max_1^k X_i \leq \lambda_k + \delta_k y \right\} \rightarrow \underline{L}(x) \tilde{L}(y).$$

Consider the range  $R_k = X'_k - X_1'$  and show that, if  $\delta'_k / \delta_k \rightarrow b$ ,

$$\text{Prob} \left\{ (R_k - (\lambda_k - \lambda'_k)) / \max(\delta_k, \delta'_k) \leq z \right\} \rightarrow \int \int_D d\underline{L}(x) d\tilde{L}(y),$$

with  $D = \left\{ \frac{y}{\max(1,b)} - \frac{b x}{\max(1,b)} \leq z \right\}$ . Conclude the dominance of  $X'_k$  if  $b = 0$ , of  $X_1'$  if  $b = +\infty$ , and a joint influence (a convolution) if  $0 < b < +\infty$ ; compare with the

results of Chapter VIII (Analytical Statistical Choice for Univariate Extremes) of Part II for the Gumbel statistic  $Q$ .

Obtain specific results for  $\Lambda\left(\frac{x-\lambda}{\delta}\right)$ ,  $\Phi_\alpha\left(\frac{x-\lambda}{\delta}\right)$  and  $W_\alpha\left(\frac{x-\lambda}{\delta}\right)$  and other common distributions. Note that the  $\delta$ -method can also be used.

*Gumbel (1958), Galambos (1987)*

I.38— Obtain corresponding results for the mid-point  $(X'_1 + X'_k)/2$ .

*Gumbel (1958), Galambos (1987)*

I.39— Obtain the corresponding results for the extremal quotient  $EQ_k = X'_k / X'_1$  when the  $X_i$  are supposed to be non-negative ( $\text{Prob}\{X_i > 0\} = 1$ ). Note that  $\log EQ_k$  is the range of  $\{\log X_i\}$ .

*Gumbel (1958)*

I.40— Show that if  $F \in \mathcal{D}(\tilde{L})$ , with attraction coefficients  $(\lambda_k, \delta_k)$  and  $G(x)$  is such that  $\frac{1-F(x)}{1-G(x)} \rightarrow c$  ( $0 < c < +\infty$ ) when  $x \rightarrow w_F$  then  $G \in \mathcal{D}(\tilde{L})$  and a system of attraction coefficients is  $(\lambda_{[ck]}, \delta_{[ck]})$ .

I.41— Use the previous result to show that if  $X$  has a distribution function  $F \in \mathcal{D}(\tilde{L})$  then the truncated random variable  $X_a = \max(X, a)$ , with  $a < w_F$ , has a distribution function  $F_a \in \mathcal{D}(\tilde{L})$  also. Compute  $c$ .

- A. A. BALKEMA and L. DE HAAN (1972). On the von Mises condition for the domain of attraction of  $\exp(-e^{-x})$ , *Ann. Math. Statist.*, 43, 1352-1354
- A. A. BALKEMA, L. DE HAAN and S. RESNICK (1984). Uniform rates of convergence to extreme value distributions, in *Statistical Extremes and Applications*, J. Tiago de Oliveira ed., 341-346, D. Reidel, Dordrecht
- J. BEIRLANT and E. WILLEKENS (1990). Rapid variation with remainder and rates of convergence, *Adv Appl. Probab.*, 22, 787-801
- RICHARD A. DAVIS (1982). The rate of convergence in distribution of the maxima, *Statist. Neerl.*, 36, #1, 31-35
- E. L. DODD (1923). The greatest and least variate under general laws of error, *Trans. Am. Math. Soc.*, 25, 525-539
- W. FELLER (1966). *An Introduction to Probability Theory and its Applications*, vol. II, Wiley, New York
- B. DE FINETTI (1932). Sulle legge di probabilità degli estreme, *Metron*, 9, 127-138
- R. A. FISHER and L. M. C. TIPPET (1928). Limiting forms of the frequency distribution of the largest and smallest member of a sample. *Proc. Camb. Phil. Soc.*, 24, 180-190
- M. FRÉCHET (1927). Sur la loi de probabilité de l'écart maximum, *Ann. Soc. Polon. Math.* (Cracovie), 6, 93-116
- J. GALAMBOS (1978). *The Asymptotic Theory of Extreme Order Statistics*, Wiley, New-York; (1987), 2nd ed., Krieger Publ.
- J. GALAMBOS (1984). Introduction, order statistics, exceedances, laws of large numbers, in *Statistical Extremes and Applications*, J. Tiago de Oliveira ed., 13-17, D. Reidel, Dordrecht
- J. GALAMBOS (1984). Rates of convergence in extreme value theory, in *Statistical Extremes and Applications*, J. Tiago de Oliveira ed., 347-352, D. Reidel, Dordrecht
- J. GEFFROY (1958/59). Contribution a l'étude de la théorie des valeurs extrêmes, *Publ. Inst. Statist. Univ. Paris*, 7/8, 37-185, Thèse de Doctorat d'Etat
- B. V. GNEDENKO (1943). Sur la distribution limite du terme maximum d'une série aléatoire, *Ann. Math.*, 44, 423-453
- E. J. GUMBEL (1935). Les valeurs extrêmes des distributions statistiques, *Ann. Inst. H. Poincaré*, 4, 115-168
- E. J. GUMBEL (1955). The calculated risk in flood control, *Appl. Sc. Research*, (Holland), 5, A, 273-280
- E. J. GUMBEL and P.G. CARLSON (1956). On the asymptotic covariance between the sample mean and standard deviation, *Metron*, XVIII, # 1-2, 3-9
- L. DE HAAN (1970). On regular variation and its applications to weak convergence of sample extremes, *Math. Centre Tracts*, # 32, Amsterdam



- L. DE HAAN (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
- L. DE HAAN (1976). Sample extremes; an elementary introduction, *Statist. Neerl.*, 30, 161-172
- M. L. JUNCOSA (1949). On the distribution of the minimum of a sequence of mutually independent random variables, *Duke Math. J.*, 16, 609-618
- M. R. LEADBETTER, GEORG LINDGREN and HOLGER ROOTZÉN (1983). *Extremes and Related Properties in Random Sequences and Processes*, Springer-Verlag, Heidelberg
- DAVID MEJZLER (1949). On a theorem of B. V. Gnedenko, *Sb. Trudov Inst. Math. Akad. Nauk. SSR*, 12, 31-35 (in russian)
- DAVID MEJZLER (1984). Asymptotic behaviour of extreme order statistics in the non-identically distributed case, in *Statistical Extremes and Applications*, J. Tiago de Oliveira ed., 535-547, D. Reidel, Dordrecht
- MICHAEL MARCUS and MARK PINSKY (1969). On the domain of attraction of  $\exp(-e^{-x})$ , *J. Math. Analysis Appl.*, 28, 440-449
- G. F. NEWELL (1964). Asymptotic extremes for m-dependent random variables, *Ann. Math. Statist.*, 35, 1322-1325
- S. I. RESNICK (1971). Tail equivalence and its applications, *J. Appl. Prob.*, 8, 136-156
- A. ROSENGARD (1962). Étude des lois limites jointes et marginales de la moyenne et des valeurs extrêmes d'un échantillon, *Publ. Inst. Statist. Univ. Paris*, 11, 3-53, Thèse de Doctorat de 3e cycle
- A. ROSENGARD (1966). Contribution a l'étude des liaisons limites entre les différents statistiques d'un échantillon, *Université de Paris*, Thèse de Doctorat d'Etat
- H. J. W. RÖSSBERG (1963). Über des asymptotische verhalten der rand-und zentralglieder einer variationsreihe, *Publ. Inst. Math. Acad. Soc. Hungary*, 8, 463-468
- H. J. W. RÖSSBERG (1965). Die asymptotische unababhängigkeit der kleinen und grossten werte einer stichprobe von stichprobenmittel, *Math. Nachr.*, 28, 305-318
- N. V. SMIRNOV (1952). Limit distributions for the terms of a variational series, *Amer. Math. Soc. Transl.*, 67, 1-64
- J. TIAGO DE OLIVEIRA (1961). The asymptotic independence of the sample mean and extremes, *Rev. Fac. Ciências de Lisboa*, A, VIII, 215-222
- J. TIAGO DE OLIVEIRA (1976). Asymptotic behaviour of maxima with periodic disturbances, *Ann. Inst. Statist. Math*, 28, 19-23
- J. TIAGO DE OLIVEIRA (1978). Approximate distribution for sequences of maxima, *Metron*, XXXVI, 3-21
- J. TIAGO DE OLIVEIRA (1982). The  $\delta$ -method for the obtention of asymptotic distributions — applications, *Publ. Inst. Statist. Univ. Paris*, XXVII, 49-70

- J. TIAGO DE OLIVEIRA (ed.) (1984). *Statistical Extremes and Applications*, D. Reidel, Dordrecht
- J. TIAGO DE OLIVEIRA (1991). Perspectivas sobre a estatística de extremos e problemas em aberto, to be publ. in *Rev. Real Acad. Ciencias Exactas, Físicas y Naturales*, Madrid
- J. TIAGO DE OLIVEIRA and B. EPSTEIN (1972). Tail comparison for maxima; applications to mixtures, unpubl. paper, *Technion*
- R. VON MISES (1936). La distribution de la plus grande de  $n$  valeurs, *Rev. Math. Union Interbalkanique*, 1, 141-160 and also in *Selected Papers*, vol.II, 1954, 271-294, Am. Math. Soc., Providence, Rhode Island
- G. S. WATSON (1954). Extreme values in samples for  $m$  independent stationary stochastic processes. *Ann. Math. Statist.*, 25, 798-800



## **PART II**

### **Statistics for univariate extremes**

