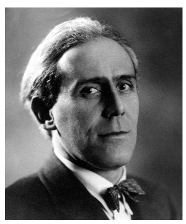
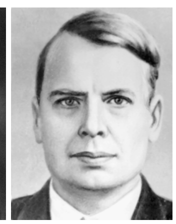
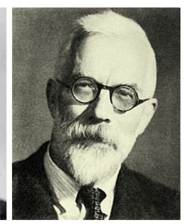
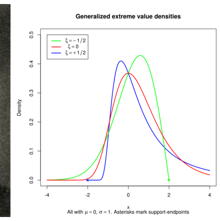
Randal Douc

MAT 4513: A short introduction to Extreme Value distributions.









March 18, 2022

Telecom SudParis

Contents

Part I
Foundations

Introduction

These lecture notes are strongly inspired from the courses "Mesures de risques en finance" by Jean-François Delmas, given at Ponts et Chaussées, and "Extremes", given at Telecom Paristech, (the latter has been teached by Anne Sabourin during the M2,"Mathématiques de l'aléatoire",Orsay). Interested readers may also skim over the following references (but it is not mandatory):

- Resnick, 2007.
- Leadbetter et al., 2012.
- Hahn and Rosenthal, 1948, mainly chapter 1.

Do not hesitate to point out the errors or typos that still remain and to propose any improvements on the content of this course. My email: `randal.douc@it-sudparis.eu`

We start with some notation. In what follows,

- i.i.d means independent and identically distributed.
- r.v. means random variables.
- for $r,s \in \mathbb{N}$ such that $r \leq s$, we write $[r:s] = \{r, r+1, \ldots, s\}$,
- $\mathscr{S}_n$ is the set of permutations on $[1:n]$,
- $X \perp\!\!\!\perp Y$ means X and Y are independent random variables,
- $X \stackrel{\mathscr{L}}{=} Y$ means X and Y have the same law.
- Let $(\mathsf{Z}, \mathscr{Z})$ and $(\mathsf{X}, \mathscr{X})$ two measurable spaces. Assume that for all $w \in \mathsf{Z}$, the two random vectors X and $Y(w)$ take values in X and assume the existence of a third random variable W taking values in Z, then the notation: $X|_{W=w} \stackrel{\mathscr{L}}{=} Y(w)$ means that for all $A \in \mathscr{X}$,

  $$\mathbb{P}(X \in A|W)|_{W=w} = \mathbb{P}(Y(w) \in A)$$

  In words, the distribution of X conditionally on W taken on $W = w$ is the same as the unconditonal distribution of $Y(w)$.
- $\liminf_n a_n = \lim_{n\to\infty} (\inf_{k\geq n} a_k)$ and similarly, $\limsup_n a_n = \lim_{n\to\infty} \left(\sup_{k\geq n} a_k\right)$. Moreover, $\lim_n a_n$ exists if and only if $\liminf_n a_n = \limsup_n a_n$.
- for any $a \in \mathbb{R}$, $a^+ = \max(a,0)$ and $a^- = \max(-a,0) = -\min(a,0)$ and we have $|a| = a^+ + a^-$ and $a = a^+ - a^-$.

Moreover, the following notions of convergence for random variables is used throughout these lecture notes.

▶ $\boxed{X_n \stackrel{w}{\Rightarrow} X}$ means *convergence in distribution* (or "convergence en loi" in French). It is equivalent to any of the following statements.

  (a) for all bounded continuous functions h, we have $\lim_n \mathbb{E}[h(X_n)] = \mathbb{E}[h(X)]$.
  (b) for all $A \in \mathscr{B}(\mathbb{R})$ such that $\mathbb{P}(X \in \partial A) = 0$, we have $\lim_n \mathbb{P}(X_n \in A) = \mathbb{P}(X \in A)$.
  (c) for all $x \in \mathbb{R}$ such that $\mathbb{P}(X = x) = 0$, we have $\lim_n \mathbb{P}(X_n \leq x) = \mathbb{P}(X \leq x)$.
  (d) for all $u \in \mathbb{R}$, we have $\lim_n \mathbb{E}\left[\mathrm{e}^{\mathrm{i}uX_n}\right] = \mathbb{E}\left[\mathrm{e}^{\mathrm{i}uX}\right]$

  In (b), the notation ∂A means the frontier of A, that is, the set of points x such that in any neighborhood of x, there are an infinite number of distinct points of A and an infinite number of points in A^c. By abuse of terminology, we may also say that $\boxed{X_n \textbf{ weakly converges to } X}$ instead of saying the distribution of X_n converges weakly to the distribution of X.

▶ $\boxed{X_n \stackrel{\mathbb{P}\text{-prob}}{\longrightarrow} X}$ means *convergence in probability*: for all $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}(|X_n - X| > \varepsilon) = 0.$$

▶ $\boxed{X_n \stackrel{\mathbb{P}\text{-a.s.}}{\longrightarrow} X}$ means *almost sure convergence*:

$$\mathbb{P}(\lim_{n\to\infty} X_n = X) = 1 \; .$$

$\boxed{\text{Almost sure convergence implies convergence in probability}}$. We recall the following properties.

(i) If $X_n \overset{w}{\Rightarrow} X$ then for all continuous functions f, $f(X_n) \overset{w}{\Rightarrow} f(X)$. Note that this property holds, when f is continuous (and not necessarily bounded), for example $f(u) = u^2$ so that $X_n^2 \overset{w}{\Rightarrow} X^2$.

(ii) **The Slutsky Lemma** If $X_n \overset{\mathbb{P}\text{-prob}}{\longrightarrow} c$ where c is a constant and if $Y_n \overset{w}{\Rightarrow} Y$, then $(X_n, Y_n) \overset{w}{\Rightarrow} (c, Y)$ that is for all continuous functions f, $f(X_n, Y_n) \overset{w}{\Rightarrow} f(c, Y)$.

(iii) $X \sim \mathrm{N}(0,1)$ iif for all $u \geq 0$, $\mathbb{E}\left[\mathrm{e}^{\mathrm{i}uX}\right] = \mathrm{e}^{-u^2/2}$. Moreover, $X \sim \mathrm{N}(\mu, \sigma^2)$ iif for all $u \geq 0$, $\mathbb{E}\left[\mathrm{e}^{\mathrm{i}uX}\right] = \mathrm{e}^{-u^2 \mathrm{Var}(X)/2 + \mathrm{i}u\mathbb{E}(X)}$ and in that case, $\sigma^2 = \mathrm{Var}(X)$ and $\mu = \mathbb{E}(X)$.

(iv) Let c be a constant. Then, $X_n \overset{\mathbb{P}\text{-prob}}{\longrightarrow} c$ if and only if $X_n \overset{w}{\Rightarrow} c$ (in words, convergence in probability to a **constant** is equivalent to convergence in distribution to this **constant**)

Some usual distributions

Name	Acronym	Parameter	density function: $f_X(x)$	cdf: $F_X(x) = \int_{-\infty}^{x} f_X(u)\mathrm{d}u$	Other properties
Gaussian	$\mathrm{N}(\mu, \sigma^2)$	(μ, σ^2)	$\frac{1}{\sqrt{2\pi\sigma^2}}\mathrm{e}^{-(x-\mu)^2/(2\sigma^2)}$	No explicit expression	$\mathbb{E}[X] = \mu,\ \mathrm{Var}(X) = \sigma^2$
Exponential	$\exp(\lambda)$	$\lambda > 0$	$\lambda \mathrm{e}^{-\lambda x} \mathbb{1}_{\mathbb{R}+}(x)$	$\left(1 - \mathrm{e}^{-\lambda x}\right) \mathbb{1}_{\mathbb{R}+}(x)$	$\mathbb{E}[X] = 1/\lambda,\ \mathrm{Var}(X) = 1/\lambda^2$
Gamma	$\Gamma(k, \theta)$	$(k, \theta) \in \left(\mathbb{R}_+^*\right)^2$	$\frac{x^{k-1}\mathrm{e}^{-x/\theta}}{\Gamma(k)\theta^k}$	$\frac{\Gamma_{x/\theta}(k)}{\Gamma(k)}$	

In the above description,

(i) if $X_i \sim \Gamma(k_i, \theta)$ and (X_i) are independent, then $\sum_{i=1}^{n} X_i \sim \Gamma(\sum_{i=1}^{n} k_i, \theta)$.

(ii)

$$\Gamma(k) = \begin{cases} \int_0^\infty t^{k-1}\mathrm{e}^{-t}\mathrm{d}t & \text{if } k \in \mathbb{R}_+^* \\ k! & \text{if } k \in \mathbb{N}. \end{cases} \quad (\blacktriangleright\text{GAMMA FUNCTION})$$

$$\Gamma_x(k) = \int_0^x t^{k-1}\mathrm{e}^{-t}\mathrm{d}t \ (\blacktriangleright\text{INCOMPLETE GAMMA FUNCTION})$$

Chapter 1

Order statistics

Contents

Keywords 1.1 *Order statistics, maximum, minimum, consistency, asymptotic normality.*

Let $X_1, X_2, \ldots$ be independent and identically distributed random variables. The Law of Large Numbers (LLN) and the Central Limit Theorem (CLT) deal with the limiting behavior of the empirical mean value

$$\bar{X}_n = \frac{X_1 + \ldots + X_n}{n}$$

as n tends to infinity. Instead, the Extreme Value theory (EVT) is concerned with the limiting behavior, as n tends to infinity, of extreme values as

$$\hat{X}_n = \max(X_1, \ldots, X_n)\ , \quad \check{X}_n = \min(X_1, \ldots, X_n).$$

1.1 Order statistics

Let $X_1, X_2, \ldots$ be a sequence of iid random variables and let F_X be the associated cumulative distribution function (cdf) defined by $F_X(x) = \mathbb{P}(X \le x)$ for all $x \in \mathbb{R}$. The law of the Maximum and the Minimum of these random variables can be easily obtained as shown in the next lemma.

Lemma 1.1 *For all $x \in \mathbb{R}$,*

$$\mathbb{P}(\hat{X}_n \le x) = F_X(x)^n\ , \quad \mathbb{P}(\check{X}_n \le x) = 1 - (1 - F_X(x))^n\ .$$

PROOF. Since (X_i) are iid, we have for all $x \in \mathbb{R}$,

$$\mathbb{P}(\hat{X}_n \le x) = \mathbb{P}(X_1 \le x, \ldots, X_n \le x) = \prod_{i=1}^{n} \mathbb{P}(X_i \le x) = F_X(x)^n\ .$$

This proves the first part of the Lemma. To prove the second part, note that $\check{X}_n = -\max(-X_1, \ldots, -X_n)$. This implies

$$\mathbb{P}(\check{X}_n > x) = \mathbb{P}(\max(-X_1, \ldots, -X_n) < -x) = \prod_{i=1}^{n} \mathbb{P}(-X_i < -x) = (1 - F_X(x))^n\ .$$

Another way to prove the latter equation is to write $\mathbb{P}(\check{X}_n > x) = \mathbb{P}(X_1 > x, \ldots, X_n > x)$ and to use a similar reasoning as in the first part. ■

If X has a density f_X wrt the Lebesgue measure, we can easily deduce the associated densities of $\hat{X}_n$ and $\check{X}_n$ (at all points x where f_X is continuous) by differentiating their cdfs:

- $\hat{X}_n$ has the density $x \mapsto n f_X(x) F_X(x)^{n-1}$ with respect to the Lebesgue measure,
- $\check{X}_n$ has the density $x \mapsto n f_X(x)(1 - F_X(x))^{n-1}$with respect to the Lebesgue measure.

We are now interested in the distribution of the k^{th} term of the order statistics. We start with some definitions.

Definition 1.2 *The order statistics of $(X_1, \ldots, X_n)$ is defined by $(X_{(1,n)}, \ldots, X_{(n,n)})$ where $\{X_{(i,n)} : i \in [1:n]\}$ are the $\{X_1, \ldots, X_n\}$ reordered according to their increasing values. Therefore, there exists a random permutation σ_n on $[1:n]$ such that*

$$(X_{(1,n)}, \ldots, X_{(n,n)}) = (X_{\sigma_n(1)}, \ldots, X_{\sigma_n(n)})$$

where $X_{\sigma_n(1)} \leq \ldots \leq X_{\sigma_n(n)}$.

Without further assumptions, some of the random variables (X_i) may take the same values. This means that the random permutation σ_n is not unique in general. We will add some assumptions in a next step that guarantee the uniqueness of the permutation σ. Before that, we may link the order statistics with the maximum and minimum of a data set by noting that

$$\hat{X}_n = \max(X_1, \ldots, X_n) = X_{(n,n)} , \quad \check{X}_n = \min(X_1, \ldots, X_n) = X_{(1,n)}$$

Definition 1.3 *For all $x \in (0,1)$, the generalized inverse distribution function (or the quantile function) is defined by*

$$F_X^{\leftarrow}(x) = \inf\{y \in \mathbb{R} : F_X(y) \geq x\} ,$$

Of course whenever F_X is invertible, the generalized inverse distribution function is the classical inverse function of F_X that is $F_X^{\leftarrow} = F_X^{-1}$. As a cumulative distribution function, F_X is nondecreasing but this function "jumps" at points $x \in \mathbb{R}$ where $\mathbb{P}(X = x) > 0$ and remains constant on intervals $[\alpha, \beta]$ where $\mathbb{P}(X \in [\alpha, \beta]) = 0$. These two situations typically appear for discrete variables or "continuous" variables with disconnected supports. For these type of non-decreasing functions F_X, inverse functions are not defined in general and that explains the introduction of generalized inverse functions.

Exercise 1.1. (i) Show that the cdf F_X is cadlag (in french, continue à droite, limitée à gauche), that is, for all $x \in \mathbb{R}$, F_X is continuous in a right neighbourhood of x and has a limit in a left neighbourhood of x.
(ii) Show that for all $u \in (0,1)$, we have $F_X(F_X^{\leftarrow}(u)) \geq u$.
(iii) Show that $F_X^{\leftarrow}$ is nondecreasing on $(0,1)$.
(iv) Show that $F_X^{\leftarrow}$ is left-continuous on $(0,1)$.
(v) Show that the discontinuity points of F_X are countable. Deduce that the continuity points of F_X are dense.

Even though it is not possible to write $(F_X^{\leftarrow}(u) = v) \Leftrightarrow (u = F_X(v))$ since $F_X^{\leftarrow}$ is not the inverse of F_X the following lemma shows a useful identity, which in general is far enough to obtain the desired properties on $F_X^{\leftarrow}$.

Lemma 1.4 *For all $u, v \in \mathbb{R}$,*

$$(F_X^{\leftarrow}(u) \leq v) \Leftrightarrow (u \leq F_X(v))$$

PROOF. Define $A_u = \{y \in \mathbb{R} : F_X(y) \geq u\}$.

$\Leftarrow$ if $u \leq F_X(v)$, we have $v \in A_u$ so that $v \geq \inf_{y \in A_u} y = F_X^{\leftarrow}(u)$.

$\Rightarrow$ if $F_X^{\leftarrow}(u) \leq v$, then applying the nondecreasing function F_X on both sides of the inequality, we get $F_X(F_X^{\leftarrow}(u)) \leq F_X(v)$. To completes the proof, it remains to show that $F_X(F_X^{\leftarrow}(u)) \geq u$. But this follows from Exercise 5.1.

■

The following proposition is particularly useful for sampling a random variable of arbitrary cdf from a uniform random variable on $[0,1]$, provided that the inverse generalized cdf is explicit.

Proposition 1.5 *If* $U \sim \mathrm{Unif}[0,1]$*, then*

$$F_X^{\leftarrow}(U) \stackrel{\mathscr{L}}{=} X \ .$$

PROOF. By Lemma 1.4, for all $u, v \in \mathbb{R}$,

$$(F_X^{\leftarrow}(u) \leq v) \Leftrightarrow (u \leq F_X(v))$$

Therefore, if U is a random variable with uniform distribution on $[0,1]$, we have

$$\mathbb{P}(F_X^{\leftarrow}(U) \leq y) = \mathbb{P}(U \leq F_X(y)) = F_X(y) \ ,$$

which completes the proof ■

There is a simple converse to Proposition 1.5. This can be seen in the next exercise.

Exercise 1.2. Show that if F_X is increasing and continuous and if X is a random variable with cdf F_X, then $F_X(X) \sim \mathrm{Unif}[0,1]$.

Exercise 1.3. Let (X_i) be iid random variables with exponential distribution of parameter λ. How can we sample $\hat{X}_n$ using only one random variable uniformly distributed on $(0,1)$?

From Proposition 1.5, it can be easily seen that

Lemma 1.6 *Let* $\{X_i,\ i \in [1:n]\}$ *be iid random variables with cdf* F_X*. Let* $\{U_i,\ i \in [1:n]\}$ *be iid random variables such that* $U_i \sim \mathrm{Unif}[0,1]$*. Then,*

$$(F_X^{\leftarrow}(U_{(1,n)}), \ldots, F_X^{\leftarrow}(U_{(n,n)})) \stackrel{\mathscr{L}}{=} (X_{(1,n)}, \ldots, X_{(n,n)})$$

The proof is easy and left to the reader. We now assume that F_X is increasing and continuous, that is $P(X = x) = 0$ for all $x \in \mathbb{R}$. This additional assumption allows to consider that the data can be uniquely ordered as seen in the next result.

Lemma 1.7 *If* F_X *is continuous and increasing, then, a.s. we have* $X_{(1,n)} < \cdots < X_{(n,n)}(n)$ *and therefore, the random permutation* σ_n *in Definition 1.2 is unique.*

PROOF. It is sufficient to prove that $\mathbb{P}(\exists i \neq j, X_i = X_j) = 0$. Using the representation with $U_i \sim \mathrm{Unif}[0,1]$ and noting that $F_X^{\leftarrow}$ is one-to-one,

$$\begin{aligned}\mathbb{P}(\exists i \neq j, X_i = X_j) &= \mathbb{P}(\exists i \neq j, F_X^{\leftarrow}(U_i) = F_X^{\leftarrow}(U_j)) \\ &= \mathbb{P}(\exists i \neq j, U_i = U_j) = 0\end{aligned}$$

since $\mathbb{P}(U_i = U_j) = \int_{[0,1]^2} \mathbb{1}_{\{u \neq v\}} \mathrm{d}u\mathrm{d}v = 0$. Therefore, a.s., for all $i \neq j$, $X_i \neq X_j$. ■

The main result of this section is the following:

Theorem 1.8. *Let* (X_i) *be a sequence of i.i.d. random variables with cdf* F_X*. If* F_X *is continuous and increasing and if* X *has a density* f_X *with respect to the Lebesgue measure. Then,* $\sigma_n \perp\!\!\!\perp (X_{(1,n)}, \ldots, X_{(n,n)})$ *and for all* $\sigma \in \mathscr{S}_n$ *and* $A \in \mathscr{B}(\mathbb{R})^{\otimes n}$*,*

$$\mathbb{P}(\sigma_n = \sigma) = \frac{1}{n!}$$

$$\mathbb{P}((X_{(1,n)}, \ldots, X_{(n,n)}) \in A) = n! \int_A \mathbb{1}_{\{x_1 < \ldots < x_n\}} f_X(x_1) \ldots f_X(x_n) \mathrm{d}x_1 \ldots \mathrm{d}x_n$$

PROOF. Let $\sigma \in \mathscr{S}_n$. Since (X_i) are iid, $(X_1,\ldots,X_n) \stackrel{\mathscr{L}}{=} (X_{\sigma(1)},\ldots,X_{\sigma(n)})$ for any $\sigma \in \mathscr{S}_n$. This implies that

$$\begin{aligned}
\mathbb{P}(\sigma_n = \sigma,(X_{(1,n)},\ldots,X_{(n,n)}) \in A) \\
&= \mathbb{E}\left[\mathbb{1}\left\{X_{\sigma(1)} < \ldots < X_{\sigma(n)}\right\}\mathbb{1}_A(X_{\sigma(1)},\ldots,X_{\sigma(n)})\right] \\
&= \mathbb{E}\left[\mathbb{1}\left\{X_1 < \ldots < X_n\right\}\mathbb{1}_A(X_1,\ldots,X_n)\right] \\
&= \int_A \mathbb{1}\left\{x_1 < \ldots < x_n\right\} f_X(x_1)\ldots f_X(x_n)\mathrm{d}x_1\ldots\mathrm{d}x_n
\end{aligned} \tag{1.1}$$

Thanks to this equation, we have the joint law of $(\sigma_n,(X_{(1,n)},\ldots,X_{(n,n)}))$. We can get the two marginal distributions by marginalizing with respect to σ_n and $(X_{(1,n)},\ldots,X_{(n,n)})$ and we can prove easily the independence of these random variables. Let us do that. Setting $A = \mathbb{R}^n$ in (1.1), we get

$$\mathbb{P}(\sigma_n = \sigma) = \mathbb{E}\left[\mathbb{1}\left\{X_1 < \ldots < X_n\right\}\right]$$

which does not depend on σ. Since $\mathrm{card}(\mathscr{S}_n) = n!$, we deduce $\mathbb{P}(\sigma_n = \sigma) = 1/n!$. Moreover, summing the identity (1.1) over $\sigma \in \mathscr{S}_n$, we get

$$\begin{aligned}
\mathbb{P}\left((X_{(1,n)},\ldots,X_{(n,n)}) \in A\right) &= \sum_{\sigma\in\mathscr{S}_n} \mathbb{P}(\sigma_n = \sigma,(X_{(1,n)},\ldots,X_{(n,n)}) \in A) \\
&= n!\int_A \mathbb{1}_{x_1<\ldots<x_n} f_X(x_1)\ldots f_X(x_n)\mathrm{d}x_1\ldots\mathrm{d}x_n
\end{aligned}$$

Plugging this equality and $\mathbb{P}(\sigma_n = \sigma) = 1/n!$ into (1.1) yields

$$\mathbb{P}(\sigma_n = \sigma,(X_{(1,n)},\ldots,X_{(n,n)}) \in A) = \mathbb{P}(\sigma_n = \sigma)\mathbb{P}((X_{(1,n)},\ldots,X_{(n,n)}) \in A)$$

This shows that $\sigma_n \perp\!\!\!\perp (X_{(1,n)},\ldots,X_{(n,n)})$ and the proof is completed. ■

The fact that $\sigma_n \perp\!\!\!\perp (X_{(1,n)},\ldots,X_{(n,n)})$ was not obvious at all since σ_n and $(X_{(1,n)},\ldots,X_{(n,n)})$ are both constructed from the data set in an intricate way, but we have to believe the proof : they are independent! From Theorem 1.8, we may deduce the marginal law of k^{th} term $X_{(k,n)}$ of the order statistics where $k \in [1:n]$. The proof is, at least theoretically, streamlined: since we have the joint law, we only need to use marginalisation to get the marginal distribution. This is a nice plan of action. Still, by doing so, multivariate integrals appear on the surface. To obtain explicit calculations of these integrals, some tricks using the law of the maximum and the minimum are needed as can be seen in the proof.

Lemma 1.9 *For all $k \in [1:n]$, and all $x \in \mathbb{R}$,*

$$\mathbb{P}(X_{(k,n)} \le x) = \frac{n!}{(k-1)!(n-k)!}\int_0^{F_X(x)} t^{k-1}(1-t)^{n-k}\mathrm{d}t\ .$$

PROOF. Combining again Lemma 1.1 with Theorem 1.8 yields

$$F_X(x)^n = \mathbb{P}(\hat{X}_n \le x) = \mathbb{P}(X_{(n,n)} \le x) = n!\int_{\mathbb{R}^n} \mathbb{1}\left\{x_1 < \ldots < x_n \le x\right\} f_X(x_1)\ldots f_X(x_n)\mathrm{d}x_1\ldots\mathrm{d}x_n \tag{1.2}$$

Similarly, combining again Lemma 1.1 with Theorem 1.8,

$$[1-F_X(x)]^n = \mathbb{P}(\check{X}_n > x) = \mathbb{P}(X_{(1,n)} > x) = n!\int_{\mathbb{R}^n} \mathbb{1}\left\{x < x_1 < \ldots < x_n\right\} f_X(x_1)\ldots f_X(x_n)\mathrm{d}x_1\ldots\mathrm{d}x_n \tag{1.3}$$

Moreover, by Theorem 1.8,

$$\begin{aligned}
\mathbb{P}(X_{(k,n)} \le x) &= n!\int_{\mathbb{R}^n} \mathbb{1}\left\{x_1 < \ldots < x_k \le x\right\}\mathbb{1}\left\{x_k < x_{k+1} < \ldots < x_n\right\} f_X(x_1)\ldots f_X(x_n)\mathrm{d}x_1\ldots\mathrm{d}x_n \\
&= n!\int_{-\infty}^x \left(\int_{\mathbb{R}^{k-1}} \mathbb{1}\left\{x_1 < \ldots < x_k\right\} f_X(x_1)\ldots f_X(x_{k-1})\mathrm{d}x_1\ldots\mathrm{d}x_{k-1}\right) \qquad (1.4) \\
&\quad \times \left(\int_{\mathbb{R}^{n-k}} \mathbb{1}\left\{x_k < \ldots < x_n\right\} f_X(x_{k+1})\ldots f_X(x_n)\mathrm{d}x_{k+1}\ldots\mathrm{d}x_n\right) f_X(x_k)\mathrm{d}x_k \qquad (1.5)
\end{aligned}$$

Applying (1.2) with n replaced by $k-1$ and (1.3) with n replaced by $n-k$ and plugging these equalities into (1.5), we obtain

$$\mathbb{P}(X_{(k,n)} \le x) = \frac{n!}{(k-1)!(n-k)!}\int_{-\infty}^x F_X(y)^{k-1}(1-F_X(y))^{n-k} f_X(y)\mathrm{d}y\ .$$

Setting $t = F_X(y)$, we get the desired result. ■

Exercise 1.4. Using Lemma 1.9, check that we can obtain Lemma 1.1.

1.2 Asymptotic properties of the empirical quantiles

1.2.1 Consistency

By definition of $X_{(k,n)}$, there are k indexes j in the data set such that $X_j \leq X_{(k,n)}$. The proportion of X_i which are less or equal than $X_{(k,n)}$ is therefore k/n.

$$\frac{\#\left\{i \in [1:n] \,:\, X_i \leq X_{(k,n)}\right\}}{n} = \frac{k}{n}$$

This leads to the following idea: if we let k depend on n and if we let k_n/n tends to p, then we may expect that $X_{(k,n)}$ tends to the point x_p such that when X follows a distribution of cdf F_X, the event $\{X \leq x_p\}$ holds with probability p. That is $F_X(x_p) = p$. This will be the main result of this section, which is now stated and proved.

Proposition 1.10 *Let $p \in (0,1)$. Assume that F_X is continuous and that there exists a unique solution x_p to the equation $F_X(x) = p$. Let $\{k_n, n \in \mathbb{N}\}$ be a sequence of integers such that $k_n \in [1:n]$, and $\lim_{n\to\infty} k_n/n = p$. Then, $\{X_{(k_n,n)},\, n \in \mathbb{N}\}$ converges a.s. to x_p*

PROOF. Fix $x \in \mathbb{R}$. Denote $S_n(x) = \sum_{i=1}^n \mathbb{1}\{X_i \leq x\}$. Note that for all $k \in \mathbb{N}$, we have the equivalence $\{S_n(x) \geq k\}$ if and only if there is at least k indexes $i \in [1:n]$ such that $X_i \leq x$. We deduce

$$\{S_n(x) \geq k\} = \{X_{(k,n)} \leq x\} \tag{1.6}$$

By taking the complement, we also have

$$\{S_n(x) < k\} = \{X_{(k,n)} > x\}$$

This implies the equality between these two events

$$\{\exists n_0 \in \mathbb{N}, \forall n \geq n_0 \quad X_{(k_n,n)} \leq x\} = \{\exists n_0 \in \mathbb{N}, \forall n \geq n_0 \quad S_n(x) \geq k_n\}$$

Using $\lim_n k_n/n = p$ and the strong law of large numbers,

$$\lim_{n\to\infty} \frac{S_n(x)}{k_n} = \lim_{n\to\infty} \frac{S_n(x)}{n}\frac{n}{k_n} = \frac{\mathbb{E}[\mathbb{1}\{X_1 \leq x\}]}{p} = \frac{F_X(x)}{F_X(x_p)} \quad \mathbb{P}-\text{a.s.}$$

Therefore if $x > x_p$, i.e. if $F_X(x) > F_X(x_p)$,

$$\mathbb{P}(\exists n_0 \in \mathbb{N}, \forall n \geq n_0, X_{(k_n,n)} \leq x) = \mathbb{P}(\exists n_0 \in \mathbb{N}, \forall n \geq n_0, S_n(x) \geq k_n) = 1$$

Conversely, with a similar reasoning, if $x < x_p$ i.e. if $F_X(x) < F_X(x_p)$,

$$\mathbb{P}(\exists n_0 \in \mathbb{N}, \forall n \geq n_0, X_{(k_n,n)} > x) = \mathbb{P}(\exists n_0 \in \mathbb{N}, \forall n \geq n_0, S_n(x) < k_n) = 1$$

This shows that $\lim_n X_{(k_n,n)} = x_p$, $\mathbb{P}-$a.s. ■

An inspection of the previous proof shows that if we want to have properties on $X_{(k,n)}$, it is more convenient to express these properties in terms of $S_n(x) = \sum_{i=1}^n \mathbb{1}\{X_i \leq x\}$ which are much more easy to deal with, since it is just a sum of iid random variables... We now consider the extremal cases: $p \in \{0,1\}$.

Lemma 1.11 *We have the two following results:*

(i) If $\lim_n k_n/n = 1$ then,

$$X_{(k_n,n)} \overset{\mathbb{P}\text{-a.s.}}{\longrightarrow} \begin{cases} \infty & \textit{if } F_X(x) < 1 \textit{ for all } x \in \mathbb{R} \\ \inf\{x \in \mathbb{R} \,:\, F_X(x) = 1\} & \textit{otherwise} \end{cases}$$

(ii) *If* $\lim_n k_n/n = 0$ *then,*

$$X_{(k_n,n)} \xrightarrow{\mathbb{P}\text{-a.s.}} \begin{cases} -\infty & \text{if } F_X(x) > 0 \text{ for all } x \in \mathbb{R} \\ \sup\{x \in \mathbb{R} \,:\, F_X(x) = 0\} & \text{otherwise} \end{cases}$$

Exercise 1.5. Prove Lemma 1.11 along the lines of Proposition 1.10.

1.3 Asymptotic normality

The main result of this section is the following.

Proposition 1.12 *Let* $p \in (0,1)$. *Assume that* X *has a density* f_X *which is continuous at* x_p *such that* $F_X(x_p) = p$ *and* $f_X(x_p) > 0$. *Assume moreover that* $k_n = np + o(n^{1/2})$. *Then,*

(i)

$$n^{1/2}\left(X_{(k_n,n)} - x_p\right) \overset{w}{\Rightarrow}_{n\to\infty} \mathrm{N}\left(0, \frac{p(1-p)}{f_X^2(x_p)}\right)$$

(ii) *Let* $\alpha > 0$ *and denote by* a_α *the quantile associated to* $1-\alpha/2$ *for the standard gaussian distribution. Then, the random interval*

$$\left[X_{(k_n,n)} - a_\alpha \left(\frac{p(1-p)}{f_X^2(X_{(k_n,n)})}\right)^{1/2}, X_{(k_n,n)} + a_\alpha \left(\frac{p(1-p)}{f_X^2(X_{(k_n,n)})}\right)^{1/2}\right]$$

is a confidence interval for x_p *associated to the asymptotic level* $1-\alpha$.

PROOF.

(i) We use the same notation as in the proof of Proposition 1.10. Denote for $x \in \mathbb{R}$, $S_n(x) = \sum_{i=1}^n \mathbb{1}\{X_i \le x\}$ and recall that (1.6) yields

$$\{S_n(x) \ge k_n\} = \{X_{(k_n,n)} \le x\}$$

Therefore,

$$\mathbb{P}(n^{1/2}\left(X_{(k_n,n)} - x_p\right) \le x) = \mathbb{P}\left(S_n\left(\underbrace{x_p + n^{-1/2}x}_{y_n}\right) \ge k_n\right) = \mathbb{P}\left(\underbrace{n^{1/2}\left(\frac{S_n(y_n)}{n} - p_n\right)}_{V_n} \ge \underbrace{n^{1/2}\left(\frac{k_n}{n} - p_n\right)}_{z_n}\right) = \mathbb{P}(V_n \ge z_n)$$

where we have set $p_n = F_X(y_n)$. We will show that $\lim_n z_n = -xf_X(x_p)$ and $V_n \overset{w}{\Rightarrow} \mathrm{N}(0, p(1-p))$, and we will use the Slutsky Lemma. We start with $\lim_n z_n = -xf_X(x_p)$. Since $k_n = np + o(n^{1/2})$, we have by definition of z_n , p_n and y_n,

$$z_n = n^{1/2}\left(\frac{k_n}{n} - p_n\right) = n^{1/2}(p - p_n) + o(1) = n^{1/2}\left(p - F_X(\underbrace{y_n}_{x_p + n^{-1/2}x})\right) + o(1) = n^{1/2}\left(p - \underbrace{F_X(x_p)}_{p} - n^{-1/2}x\underbrace{F_X'(x_p)}_{f_X(x_p)}\right) + o(1)\,.$$

Thus, $\lim_n z_n = -xf_X(x_p)$. We now turn to $V_n \overset{w}{\Rightarrow} \mathrm{N}(0, p(1-p))$. For all $u \in \mathbb{R}$,

$$\mathbb{E}\left[\mathrm{e}^{\mathrm{i}uV_n}\right] = \mathbb{E}\left(\mathrm{e}^{\mathrm{i}u\sum_{j=1}^n \left(\mathbb{1}_{\{X_j \le y_n\}} - p_n\right)/n^{1/2}}\right) = \left(\mathbb{E}\left(\mathrm{e}^{\mathrm{i}u\left(\mathbb{1}_{\{X_1 \le y_n\}} - p_n\right)/n^{1/2}}\right)\right)^n = \left[p_n \mathrm{e}^{\mathrm{i}u(1-p_n)/n^{1/2}} + (1-p_n)\mathrm{e}^{-\mathrm{i}up_n/n^{1/2}}\right]^n \tag{1.7}$$

By a Taylor expansion of the term between brackets,

$$
\begin{aligned}
& p_n \mathrm{e}^{\mathrm{i}u(1-p_n)/n^{1/2}} + (1-p_n)\mathrm{e}^{-\mathrm{i}up_n/n^{1/2}} \\
&\qquad = p_n\left(1+\frac{\mathrm{i}u}{n^{1/2}}(1-p_n)-\frac{u^2}{2n}(1-p_n)^2+o(n^{-1})\right)+(1-p_n)\left(1-\frac{\mathrm{i}u}{n^{1/2}}p_n-\frac{u^2}{2n}p_n^2+o(n^{-1})\right) \\
&\qquad = 1-\frac{u^2}{2n}p_n(1-p_n)+o(n^{-1}) = 1-\frac{u^2}{2n}p(1-p)+O(n^{-3/2})
\end{aligned}
$$

Therefore, plugging into (1.7) yields

$$
\mathbb{E}\left[\mathrm{e}^{\mathrm{i}uV_n}\right] = \left(1-\frac{u^2}{2n}p(1-p)+O(n^{-3/2})\right)^n \to_{n\to\infty} \mathrm{e}^{-u^2p(1-p)/2}
$$

Finally, $V_n \overset{w}{\Rightarrow} V$ where $V \sim \mathrm{N}(0, p(1-p))$ and by Slutsky's lemma,

$$
\mathbb{P}(n^{1/2}\left(X_{(k_n,n)}-x_p\right) \le x) = \mathbb{P}(V_n \ge z_n) \to \mathbb{P}(V \ge -xf_X(x_p)) = \mathbb{P}\left(\frac{-V}{f_X(x_p)} \le x\right)
$$

Therefore,

$$
n^{1/2}\left(X_{(k_n,n)}-x_p\right) \overset{w}{\Rightarrow}_{n\to\infty} \frac{-V}{f_X(x_p)} \sim \mathrm{N}\left(0, \frac{p(1-p)}{f_X^2(x_p)}\right)
$$

(ii) The density function f_X being continuous at x_p , we obtain by Proposition 1.10, $\lim_n f_X(X_{(k,n)}) = f_X(x_p)$ $\mathbb{P}-$a.s.This implies by Slutsky's lemma:

$$
n^{1/2}\frac{f_X(X_{(k_n,n)})}{(p(1-p))^{1/2}}\left(X_{(k_n,n)}-x_p\right) \overset{w}{\Rightarrow} \mathrm{N}(0,1)
$$

which completes the proof.

■

In the previous proposition, the confidence interval is expressed according to f_X but this quantity is most presumably not known if we are interested in the estimation of an empirical quantile. In the next result, we obtain a confidence interval for x_p without any further information on f_X.

Proposition 1.13 *Let $p \in (0,1)$ and denote by a_α the quantile associated to $1-\alpha/2$ for the standard gaussian distribution. Define*

$$
i_n = \lfloor np - a_\alpha [np(1-p)]^{1/2} \rfloor, \quad \textit{and} \quad j_n = \lfloor np + a_\alpha [np(1-p)]^{1/2} \rfloor
$$

Then, for sufficiently large n, $(i_n, j_n) \in [1:n]\times[1:n]$. Moreover, the random interval $[X_{(i_n,n)}, X_{(j_n,n)}]$ is a confidence interval for x_p with asymptotic level $1-\alpha$.

PROOF. Set $Z_n = n^{1/2}\frac{S_n/n-p}{[p(1-p)]^{1/2}}$ where $S_n = \sum_{i=1}^n \mathbb{1}_{\{X_i \le x_p\}}$. From the central limit theorem, the sequence of random variables $\{Z_n,\ n \in \mathbb{N}\}$ converges in distribution to a standard gaussian variable, that is $Z_n \overset{w}{\Rightarrow} Z$ where $Z \sim N(0,1)$. For sufficiently large n, we have $1 \le i_n \le j_n \le n$ and

$$
\begin{aligned}
\mathbb{P}(X_{(i_n,n)} \le x_p < X_{(j_n,n)}) &= \mathbb{P}(i_n \le S_n < j_n) = \mathbb{P}\left(\underbrace{n^{1/2}\frac{i_n/n-p}{[p(1-p)]^{1/2}}}_{a_n} \le Z_n < \underbrace{n^{1/2}\frac{j_n/n-p}{[p(1-p)]^{1/2}}}_{b_n}\right) \\
&= \mathbb{P}(Z_n < b_n) - \mathbb{P}(Z_n < a_n) = \mathbb{P}(Z_n - b_n < 0) - \mathbb{P}(Z_n - a_n < 0)
\end{aligned}
$$

We will use twice Slutsky's lemma to show that

$$
\begin{pmatrix} Z_n \\ b_n \end{pmatrix} \overset{w}{\Rightarrow} \begin{pmatrix} Z \\ a_\alpha \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} Z_n \\ a_n \end{pmatrix} \overset{w}{\Rightarrow} \begin{pmatrix} Z \\ -a_\alpha \end{pmatrix} \tag{1.8}
$$

Since

$$
np - a_\alpha[np(1-p)]^{1/2} - 1 < i_n \le np - a_\alpha[np(1-p)]^{1/2}
$$

we get $-a_\alpha - 1/\sqrt{np(1-p)} < a_n \le -a_\alpha$, which implies that $\lim_n a_n = -a_\alpha$. Similarly, $\lim_n b_n = a_\alpha$. Finally, Slutsky's lemma yields (1.8). And this, in turn, implies that $Z_n - b_n \overset{w}{\Rightarrow} Z - a_\alpha$ and $Z_n - a_n \overset{w}{\Rightarrow} Z + a_\alpha$. Finally

$$\lim_{n\to\infty} \mathbb{P}(X_{(i_n,n)} \leq x_p < X_{(j_n,n)}) = \mathbb{P}(Z - a_\alpha < 0) - \mathbb{P}(Z + a_\alpha < 0) = \mathbb{P}(-a_\alpha \leq Z < a_\alpha) = 1 - \alpha$$

which concludes the proof. ■

Note that in Proposition 1.13, $p \in (0,1)$, that is, we exclude the extremal values for p. A natural question would be the following: can we expect the same type of CLT when $p \in \{0,1\}$? It turns out that the situation is more complicated. First we have to find sequences of constants (a_n) and (b_n) such that $a_n > 0$ and $a_n^{-1}(\hat{X}_n - b_n)$ converges in distribution to some limiting distribution and second, we will show that this limiting distribution is not gaussian!!!

1.4 Highlights

1.4.1 North Sea flood of 1953

The 1953 North Sea flood was a major flood caused by a heavy storm that occurred on the night of Saturday, 31 January 1953 and morning of Sunday, 1 February 1953. The floods struck the Netherlands, Belgium, England and Scotland.

Fig. 1.1 North Sea flood of 1953.

A combination of a high spring tide and a severe European windstorm over the North Sea caused a storm tide; the combination of wind, high tide, and low pressure led to a water level of more than 5.6 metres (18.4 ft) above mean sea level in some locations. The flood and waves overwhelmed sea defences and caused extensive flooding. The Netherlands, a country with 20% of its territory below mean sea level and 50% less than 1 metre (3.3 ft) above sea level and which relies heavily on sea defences, was worst affected, recording 1,836 deaths and widespread property damage. Most of the casualties occurred in the southern province of Zeeland. In England, 307 people were killed in the counties of Lincolnshire, Norfolk, Suffolk and Essex. Nineteen were killed in Scotland. Twenty-eight people were killed in West Flanders, Belgium.

In addition, more than 230 deaths occurred on water craft along Northern European coasts as well as on ships in deeper waters of the North Sea. The ferry MV Princess Victoria was lost at sea in the North Channel east of Belfast with 133 fatalities, and many fishing trawlers sank.

Realising that such infrequent events could recur, the Netherlands particularly, and the United Kingdom carried out major studies on strengthening of coastal defences. The Netherlands developed the Delta Works, an extensive system of dams and storm surge barriers. The UK constructed storm surge barriers on the River Thames below London and on the River Hull where it meets the Humber Estuary.

Chapter 2

Examples of convergence for the renormalized maxima

Contents

Keywords 2.1 *Uniform, exponential, Cauchy distributions. Extreme distributions: Weibull, Gumbel, Frechet.*

Let (X_i) be a sequence of iid random variables . If $\mathbb{E}\left[X^2\right] < \infty$, the CLT may be written as

$$\sqrt{n}(\bar{X}_n - \mathbb{E}[X]) \stackrel{w}{\Rightarrow}_{n\to\infty} \mathrm{N}(0, \mathrm{Var}(X))$$

We are adopting the same approach in this chapter but we replace $\bar{X}_n$. To be specific, we find examples where, setting $\hat{X}_n = \max\{X_i : i \in [1:n]\}$, $a_n^{-1}(\hat{X}_n - b_n)$ weakly converges and we give the form of the limiting distribution.

2.1 The uniform distribution

Let $X_1 \sim \mathrm{Unif}[0,\theta]$ where $\theta > 0$. The cdf of this distribution is $F_X : x \mapsto x/\theta$. Since $\inf\{x \in \mathbb{R} : F_X(x) = 1\} = \theta$, $\lim_n \hat{X}_n = \theta$ $\mathbb{P}-$a.s. by Lemma 1.11.

Lemma 2.1 *The sequence* $\{n\left(\frac{\hat{X}_n}{\theta} - 1\right),\ n \in \mathbb{N}^*\}$ *converges in distribution to* W *with cdf*

$$x \mapsto \mathbb{P}(W \le x) = \mathrm{e}^x\,, \qquad x \le 0\,.$$

This limiting distribution is a special case of the **Weibull** *distribution.*

PROOF. Let F_n be the cdf of the random variable $n\left(\frac{\hat{X}_n}{\theta}-1\right)$. As $\hat{X}_n < \theta$, we have $F_n(x) = 1$ for all $x \geq 0$. Now, consider the case $x < 0$,

$$F_n(x) = \mathbb{P}\left(\hat{X}_n \leq \theta + \theta\frac{x}{n}\right) = \left[\mathbb{P}\left(X_1 \leq \theta + \theta\frac{x}{n}\right)\right]^n = \left(1+\frac{x}{n}\right)^n$$

Therefore $\lim_n F_n(x) = \mathrm{e}^x$ for $x < 0$. We deduce that $n\left(\frac{\hat{X}_n}{\theta}-1\right)$ converges in distribution to a random variable W with cdf $x \mapsto \min(\mathrm{e}^x, 1)$. ■

2.2 The exponential distribution

Let $X_1 \sim \exp(\lambda)$ where $\lambda > 0$. The cdf of this distribution is $F_X : x \mapsto (1-\exp(-\lambda x))\,\mathbb{1}_{\mathbb{R}^+}(x)$. Since $\inf\{x \in \mathbb{R} :\ F_X(x) = 1\} = \infty$, $\lim_n \hat{X}_n = \infty\ \ \mathbb{P}-$a.s. by Lemma 1.11.

Lemma 2.2 *The sequence* $\{\left[\lambda\hat{X}_n - \log n\right],\ n \in \mathbb{N}^*\}$ *converges in distribution to G with cdf*

$$x \mapsto \mathbb{P}(G \leq x) = \mathrm{e}^{-\mathrm{e}^{-x}}\,, \qquad x \in \mathbb{R}\,.$$

This limiting distribution is a special case of the **Gumbel** *distribution.*

PROOF. Let F_n be the cdf of $\lambda\hat{X}_n - \log n$. We have

$$F_n(x) = \mathbb{P}(\lambda\hat{X}_n - \log n \leq x) = \mathbb{P}\left(\hat{X}_n \leq (x+\log n)/\lambda\right) = \left[\mathbb{P}\left(X_1 \leq (x+\log n)/\lambda\right)\right]^n = \left(1-\frac{\mathrm{e}^{-x}}{n}\right)^n\,.$$

Letting n goes to infinity, we get $\lim_n F_n(x) = \exp(-\mathrm{e}^{-x})$ which completes the proof. ■

2.3 The Cauchy distribution

Let $X_1 \sim \mathrm{C}(a)$ where a is the parameter of the Cauchy distribution. For simplicity, we suppose $a = 1$. The density of this distribution is $x \mapsto \frac{1}{\pi(1+x^2)}$. As the support of this density is $\mathbb{R}$, it is clear that $\inf\{x \in \mathbb{R} :\ F_X(x) = 1\} = \infty$. Thus, $\lim_n \hat{X}_n = \infty\ \ \mathbb{P}-$a.s. by Lemma 1.11.

Lemma 2.3 *The sequence* $\{\frac{\pi\hat{X}_n}{n},\ n \in \mathbb{N}^*\}$ *converges in distribution to* W *with cdf*

$$x \mapsto \mathbb{P}(W \leq x) = \mathrm{e}^{-1/x}\,, \qquad x > 0\,.$$

This limiting distribution is a special case of the **Frechet** *distribution.*

PROOF. Let F_n be the cdf of $\frac{\pi\hat{X}_n}{n}$. We have

$$F_n(x) = \mathbb{P}(\hat{X}_n \leq nx/\pi) = \left[\mathbb{P}\left(X_1 \leq nx/\pi\right)\right]^n = \left(1-\int_{nx/\pi}^{\infty}\frac{1}{\pi(1+y^2)}\mathrm{d}y\right)^n\,.$$

For $x > 0$, we have

$$\int_{nx/\pi}^{\infty}\frac{1}{\pi(1+y^2)}\mathrm{d}y = \int_{nx/\pi}^{\infty}\frac{1}{\pi y^2}\mathrm{d}y + \int_{nx/\pi}^{\infty}\left[\frac{1}{\pi(1+y^2)}-\frac{1}{\pi y^2}\right]\mathrm{d}y = \frac{1}{nx}+o(1/n)\,.$$

This implies that for $x > 0$, $F_n(x) = \left(1 - \frac{1}{nx} + o(1/n)\right)^n$. Therefore, $\lim_n F_n(x) = \mathrm{e}^{-1/x}$ for $x > 0$. And the sequence $\{\frac{\pi \hat{X}_n}{n},\ n \in \mathbb{N}^*\}$ converges in distribution to W with cdf

$$\mathbb{P}(W \leq x) = \mathrm{e}^{-1/x}\,, \qquad x \geq 0\,.$$

This completes the proof of the Lemma. ■

2.4 The Bernoulli distribution

Let $\{X_n,\ n \in \mathbb{N}^*\}$ be a sequence of iid random variables such that $X_1 \sim \mathrm{Be}(p)$ where $p \in (0,1)$. We have $\hat{X}_n = 1$ if $n \geq T = \inf\{k \geq 1\ :\ X_k = 1\}$. Thus T follows a geometric distribution with success probability p. Therefore, T is $\mathbb{P}-$a.s. finite and this implies that $\hat{X}_n$ is $\mathbb{P}-$a.s. constant and equal to 1 for sufficiently large n. It means that there does not exist a sequence of constants $\{(a_n, b_n),\ n \in \mathbb{N}^*\}$ such that $a_n^{-1}(\hat{X}_n - b_n)$ converges in distribution to a non trivial limit, that is, a limit which is different from an $\mathbb{P}-$a.s. constant random variable.

2.5 Highlights

2.5.1 Gumbel (1891-1966). Source: wikipedia.

Emil Julius Gumbel (18 July 1891, Munich – 10 September 1966, New York City) was a German mathematician and political writer.

Born in Munich, he graduated from the University of Munich shortly before the outbreak of the First World War. He was Professor of Mathematical Statistics at the University of Heidelberg.

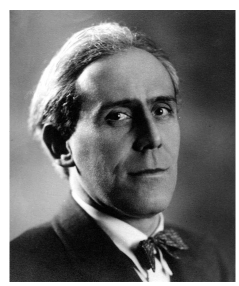


Fig. 2.1 Emil Gumbel.

Following the murder of a friend, he attended the trial where he saw that the judge completely ignored evidence against the Nazi Brownshirts. Horrified, he ardently investigated many similar political murders that had occurred and published his findings in Four Years of Political Murder in 1922. In 1928, he published Causes of Political Murder and also tried to create a political group to counter Nazism. Gumbel was also one of the 33 signers of the 1932 Dringender Appell.

Among the Nazis' most-hated public intellectuals, he was forced out of his position in Heidelberg in 1932. Gumbel then moved to France, where he taught in Paris and Lyon, and then to the United States in 1940. He taught at the École Libre Des Hautes Études in Paris and at the New School for Social Research and Columbia University in New York City until his death in 1966. As a mathematician, Gumbel was instrumental in the development of extreme value theory, along with Leonard Tippett and Ronald Fisher. In 1958, Gumbel published a key book on the topic: Statistics of Extremes. He derived and analyzed the probability distribution that is now known as the Gumbel distribution in his honor.

When he died, Gumbel's papers were made a part of The Emil J. Gumbel Collection, Political Papers of an Anti-Nazi Scholar in Weimar and Exile. These papers include reels of microfilm that document his activities against the Nazis.

2.5.2 Weibull (1887-1979)

Ernst Hjalmar Waloddi Weibull (18 June 1887 – 12 October 1979) was a Swedish engineer, scientist, and mathematician.

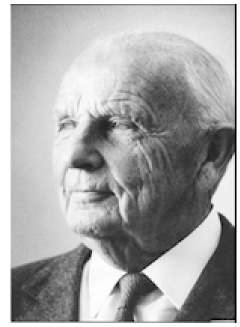


Fig. 2.2 Weibull.

Weibull came from a family that had strong ties to Scania. He was a cousin of the historian brothers Lauritz, Carl Gustaf and Curt Weibull, of whom especially the first is noteworthy for introducing a stricter criticism in the interpretation of medieval Scandinavian sources.

He joined the Swedish Coast Guard in 1904 as a midshipman. Weibull moved up the ranks with promotion to sublieutenant in 1907, Captain in 1916 and Major in 1940. While in the coast guard he took courses at the Royal Institute of Technology. In 1924 he graduated and became a full professor. Weibull obtained his doctorate from the University of Uppsala in 1932. He was employed in Swedish and German industry as a consulting engineer.

In 1914, while on expeditions to the Mediterranean, the Caribbean and the Pacific Ocean on the research ship Albatross, Weibull wrote his first paper on the propagation of explosive waves. He developed the technique of using explosive charges to determine the type of ocean bottom sediments and their thickness. The same technique is still used today in offshore oil exploration.

In 1939 he published his paper on the Weibull distribution in probability theory and statistics. In 1941 he received a personal research professorship in Technical Physics at the Royal Institute of Technology in Stockholm from the arms producer Bofors.

Weibull published many papers on strength of materials, fatigue, rupture in solids, bearings, and of course, the Weibull distribution, as well as one book on fatigue analysis in 1961. Twenty seven of these papers were reports to the US Air Force at Wright Field on Weibull analysis.

In 1951 he presented his most famous paper to the American Society of Mechanical Engineers (ASME) on the Weibull distribution, using seven case studies.

The American Society of Mechanical Engineers awarded Weibull their gold medal in 1972. The Great Gold Medal from the Royal Swedish Academy of Engineering Sciences was personally presented to him by King Carl XVI Gustaf of Sweden in 1978.

Weibull died on October 12, 1979 in Annecy, France.

2.5.3 Frechet (1878-1973). Source: wikipedia.

Maurice Fréchet (2 September 1878 – 4 June 1973) was a French mathematician. He made major contributions to the topology of point sets and introduced the entire concept of metric spaces. He also made several important contributions to the field of statistics and probability, as well as calculus. His dissertation opened the entire field of functionals on metric spaces and introduced the notion of compactness. Independently of Riesz, he discovered the representation theorem in the space of Lebesgue square integrable functions.

Fig. 2.3 Maurice Frechet

He was born to a Protestant family in Maligny, Yonne to Jacques and Zoé Fréchet. At the time of his birth, his father was a director of a Protestant orphanage in Maligny and was later in his youth appointed a head of a Protestant school. However, the newly established Third Republic was not sympathetic to religious education and so the laws were enacted requiring all education to be secular. As a result, his father lost his job. To generate some income his mother set up a boarding house for foreigners in Paris. His father was able later to obtain another teaching position within the secular system – it was not a job of a headship, however, and the family could not expect as high standards as they might have otherwise.

Maurice attended the secondary school Lycée Buffon in Paris where he was taught mathematics by Jacques Hadamard. Hadamard recognised the potential of young Maurice and decided to tutor him on an individual basis. After Hadamard moved to the University of Bordeaux in 1894, Hadamard continuously wrote to Fréchet, setting him mathematical problems and harshly criticising his errors. Much later Fréchet admitted that the problems caused him to live in a continual fear of not being able to solve some of them, even though he was very grateful for the special relationship with Hadamard he was privileged to enjoy.

After completing high-school Fréchet was required to enroll in the military service. This is the time when he was deciding whether to study mathematics or physics – he chose mathematics out of dislike of chemistry classes he would have had to take otherwise. Thus in 1900 he enrolled to École Normale Supérieure to study mathematics.

He started publishing quite early, having published four papers in 1903. He also published some of his early papers in the American Mathematical Society due to his contact with American mathematicians in Paris—particularly Edwin Wilson.

Fréchet served at many different institutions during his academic career. From 1907–1908 he served as a professor of mathematics at the Lycée in Besançon, then moved in 1908 to the Lycée in Nantes to stay there for a year. After that he served at the University of Poitiers between 1910–1919.

He married in 1908 to Suzanne Carrive (1881-1945) and had four children: Hélène, Henri, Denise and Alain.

Fréchet was planning to spend a year in the United States at the University of Illinois but his plan was disrupted when the First World War broke out in 1914. He was mobilised on 4 August the same year. Because of his diverse language skills, gained when his mother ran the establishment for foreigners, he served as an interpreter for the British Army. However, this was not a safe job; he spent two and a half years very near to or at the front. French egalitarian ideals caused many academics to be mobilised. They served in the trenches and many of them were lost during the war. It is remarkable that during his service in the war, he still managed to produce cutting edge mathematical papers frequently, despite having little time to devote to mathematics.

After the end of the war, Fréchet was chosen to go to Strasbourg to help with the reestablishment of the university. He served as a professor of higher analysis and Director of the Mathematics Institute. Despite being burdened with administrative work, he was again able to produce a large amount of high quality research.

In 1928 Fréchet decided to move back to Paris, thanks to encouragement from Borel, who was then Chair in the Calculus of Probabilities and Mathematical Physics at the Sorbonne. Fréchet briefly held a position of lecturer at the Sorbonne's Rockefeller Foundation and from 1928 was a Professor (without a Chair). Fréchet was promoted to tenured Chair of General Mathematics in 1933 and to Chair of Differential and Integral Calculus in 1935. In 1941 Fréchet succeeded Borel as Chair in the Calculus of Probabilities and Mathematical Physics, a position Fréchet held until he retired in 1949. From 1928 to 1935 Fréchet was also put in charge of lectures at the École Normale Supérieure; in this latter capacity Fréchet was able to direct a significant number of young mathematicians toward research in probability, including Doeblin, Fortet, Loeve, and Ville.

Despite his major achievements, Fréchet was not overly appreciated in France. As an illustration, while being nominated numerous times, he was not elected a member of the Academy of Sciences until the age of 78.[citation needed]. In 1929 he became foreign member of the Polish Academy of Science and Arts and in 1950 foreign member of the Royal Netherlands Academy of Arts and Sciences.

Fréchet was an Esperantist, publishing some papers and articles in that constructed language. He also served as president of the Internacia Scienca Asocio Esperantista ("International Scientific Esperantist Association") from 1950–53.

Chapter 3

Limiting distributions of the maxima

Contents

Keywords 3.1 *Convergence in distribution, extreme value distributions, quantile functions, max-stable distributions.*

In this chapter, setting $\hat{X}_n = \max\{X_i \,:\, i \in [1:n]\}$, we will obtain *a general form* for the limiting distributions of $a_n^{-1}(\hat{X}_n - b_n)$, whenever the weak limit of this quantity exists as n goes to infinity. This generalizes the results of the previous chapter.

3.1 The Fisher-Tippet-Gnedenko theorem

We start with a useful and intuitive technical lemma.

Lemma 3.1 *Suppose that $\{f_n,\, n \in \mathbb{N}\}$ is a sequence of nondecreasing functions and g is a nondecreasing function. Let $a < b$ and suppose that for each $x \in (a,b)$ that is a continuity point of g, $\lim_n f_n(x) = g(x)$. Let $f_n^{\leftarrow}$ and $g^{\leftarrow}$ be the generalized inverse of f_n and g. Then, for each $x \in (g(a), g(b))$ that is a continuity point of $g^{\leftarrow}$, we have*

$$\lim_n f_n^{\leftarrow}(x) = g^{\leftarrow}(x)\ .$$

PROOF. Let x be a continuity point of $g^{\leftarrow}$ and fix $\varepsilon > 0$. We have to prove that there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$,

$$f_n^{\leftarrow}(x) - \varepsilon \leq g^{\leftarrow}(x) \leq f_n^{\leftarrow}(x) + \varepsilon\ .$$

We only prove the right inequality. The proof of the left-hand inequality is similar.

Choose $0 < \varepsilon_1 < \varepsilon$ such that $g^{\leftarrow}(x) - \varepsilon_1$ is a continuity point of g. This is possible since the continuity points of g form a dense set. Since $g^{\leftarrow}$ is continuous in x, $g^{\leftarrow}(x)$ is a point of increase for g; hence $g(g^{\leftarrow}(x) - \varepsilon_1) < x$. Choose $\delta < x - g(g^{\leftarrow}(x) - \varepsilon_1)$. Since $g^{\leftarrow}(x) - \varepsilon_1$ is a continuity point of g, there exists n_0 such that $f_n(g^{\leftarrow}(x) - \varepsilon_1) < g(g^{\leftarrow}(x) - \varepsilon_1) + \delta < x$ for $n > n_0$. The definition of the function $f_n^{\leftarrow}$ then implies $g^{\leftarrow}(x) - \varepsilon_1 < f_n^{\leftarrow}(x)$. ■

We have now all the ingredients for proving the main result of the course.

Theorem 3.2. *Let $\{X_i,\, i \in \mathbb{N}^*\}$ be a sequence of iid random variables. Assume that there exists a sequence $\{(a_n, b_n),\, n \in \mathbb{N}^*\}$ and a distribution $\mathscr{L}_0$ such that $a_n > 0$ and*

$$\frac{\max_{i\in[1:n]} X_i - b_n}{a_n} \overset{w}{\Rightarrow} Z \qquad where \qquad Z \sim \mathscr{L}_0 \; .$$

Then, up to a translation and multiplication by a positive constant, the cdf of the distribution $\mathscr{L}_0$ is one of the following cdf:

(i) The **Weibull** *distribution :* $x \mapsto \Psi_\alpha(x) = \begin{cases} \mathrm{e}^{-(-x)^\alpha}, & x \leq 0 \\ 1, & x > 0 \end{cases}$ *and* $\alpha > 0$.

(ii) The **Gumbel** *distribution:* $x \mapsto \Lambda(x) = \mathrm{e}^{-\mathrm{e}^{-x}}$, $x \in \mathbb{R}$.

(iii) The **Frechet** *distribution:* $x \mapsto \Phi_\alpha(x) = \begin{cases} 0, & x \leq 0 \\ \mathrm{e}^{-x^{-\alpha}}, & x > 0 \end{cases}$ *and* $\alpha > 0$.

PROOF. We only sketch the proof and give the main ideas. Denote F_X the cdf of the random variable X and G the cdf of Z. Then, under the assumptions of the theorem, for all continuity point x of G,

$$\lim_{n\to\infty} \mathbb{P}(a_n^{-1}(\hat{X}_n - b_n) \leq x) = \lim_{n\to\infty} F_X^n(a_n x + b_n) = G(x) \; .$$

Taking the log, we get $\lim_{n\to\infty} n \log F_X(a_n x + b_n) = \log G(x)$ and noting that $\log(u) \sim_{u\sim 1} u - 1$, we obtain

$$\lim_{n\to\infty} \frac{1}{n\left[1 - F_X(a_n x + b_n)\right]} = -\frac{1}{\log G(x)}$$

Denote by U the generalized inverse of $1/(1-F_X)$ and by D the generalized inverse of $-\frac{1}{\log G}$. Note that since G is nondecreasing, D is also nondecreasing. Then,

$$U(x) = F_X^{\leftarrow}\left(1 - \frac{1}{x}\right) , \qquad D(x) = G^{\leftarrow}(\mathrm{e}^{-1/x}) \; .$$

By applying Lemma 3.1, we then obtain for all $x > 0$,

$$\lim_n \frac{U(nx) - b_n}{a_n} = D(x)$$

More generally, we will find conditions on D such that there exists two functions a and b such that

$$\lim_{t\to\infty} \frac{U(tx) - b(t)}{a(t)} = D(x)$$

If 1 is a continuity point of D, we can get rid of b by writing

$$\lim_{t\to\infty} \frac{U(tx) - U(t)}{a(t)} = \lim_{t\to\infty} \left(\frac{U(tx) - b(t)}{a(t)} - \frac{U(t) - b(t)}{a(t)} \right) = D(x) - D(1) = E(x) \; ,$$

with the initial condition $E(1) = D(1) - D(1) = 0$. To simplify, we will assume that all the involved functions are differentiable. Write

$$\frac{U(txy) - U(t)}{a(t)} = \frac{U(txy) - U(ty)}{a(ty)} \frac{a(ty)}{a(t)} + \frac{U(ty) - U(t)}{a(t)} \; .$$

Letting t tends to infinity, we deduce that $A(y) = \lim_t a(ty)/a(t)$ exists and

$$E(xy) = E(x)A(y) + E(y) \qquad \text{and} \qquad E(1) = 0 \; .$$

Noting that

$$A(xy) = \lim_{t\to\infty} \frac{a(txy)}{a(t)} = \lim_{t\to\infty} \frac{a(txy)}{a(ty)} \frac{a(ty)}{a(t)} = A(x)A(y)$$

and noting that $A(1) = 1$, we get by differentiating wrt x, $yA'(xy) = A'(x)A(y)$ and taking $x = 1$, we deduce $A'(y)/A(y) = A'(1)/y$ so that $A(y) = y^\xi$ (since we have $A(1) = 1$) for some $\xi \in \mathbb{R}$. Therefore,

$$E(xy) = E(x)y^\xi + E(y) \qquad \text{and} \qquad E(1) = 0 \; .$$

(a) If $\xi \neq 0$, then by differentiating the above equation wrt x, we get $yE'(xy) = E'(x)y^{\xi}$ and taking $x = 1$, it yields $E'(y) = E'(1)y^{\xi-1}$ with $E'(1) > 0$ since D and consequently E is nondecreasing. Then, $E(y) = \frac{c}{\xi}(y^{\xi} - 1)$ for some positive constant $c = E'(1)$. This implies there exists a constant β such that for all $y \geq 0$,

$$G^{\leftarrow}(\mathrm{e}^{-1/y}) = D(y) = D(1) + E(y) = cy^{\xi}/\xi + \beta$$

which implies $\mathrm{e}^{-1/y} = G(\underbrace{cy^{\xi}/\xi + \beta}_{x})$. Thus, $y = \left(\frac{x-\beta}{c/\xi}\right)^{1/\xi}$, which implies $G(x) = \exp\left[-\left(\frac{x-\beta}{c/\xi}\right)^{-1/\xi}\right]$. Then, up to a translation and multiplication by a positive constant,

$$G(x) = H_{\xi}(x) = \mathrm{e}^{-(1+\xi x)^{-1/\xi}}, \qquad \text{where} \qquad 1 + \xi x > 0\,.$$

Moreover, taking $\xi = -1/\alpha$ with $\alpha > 0$,

$$\psi_{\alpha}(x) = H_{-1/\alpha}(\alpha(x+1)), \qquad x < 0 \ \ (\blacktriangleright\text{WEIBULL CDF})$$

And taking $\xi = 1/\alpha$ with $\alpha > 0$,

$$\phi_{\alpha}(x) = H_{1/\alpha}(\alpha(x-1)), \qquad x > 0 \ \ (\blacktriangleright\text{FRECHET CDF})$$

(b) If $\xi = 0$, then $E(x) = c\log(x)$ for some positive constant c (c is positive since D and consequently E is nondecreasing). Then, following the same arguments, we have up to a translation and multiplication by a positive constant,

$$G(x) = \mathrm{e}^{-\mathrm{e}^{-x}}, \qquad x \in \mathbb{R}\,. \ (\blacktriangleright\text{GUMBEL CDF})$$

■

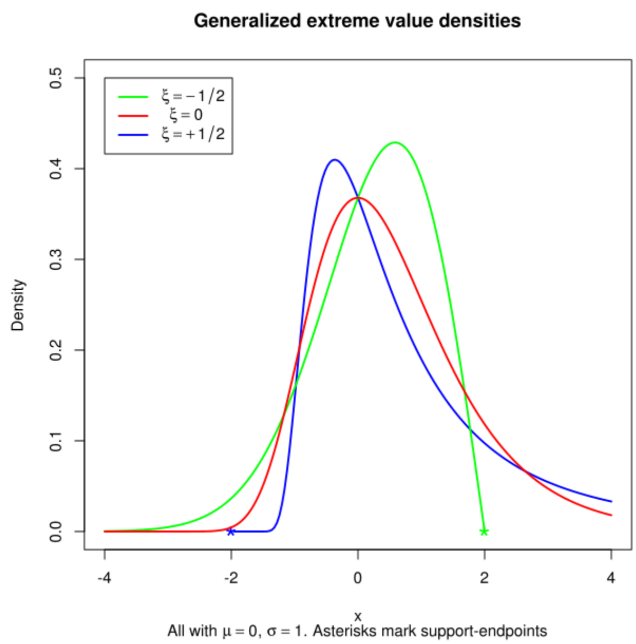


Fig. 3.1 Extreme value densities, i.e. associated to the cdf $H_{\xi}(x) = \mathrm{e}^{-(1+\xi x)^{-1/\xi}}$ where $1 + \xi x > 0$

We end up this chapter with the definition of *max-stable* distributions. Of course, the Weibull, the Gumbel and the Frechet distributions are max-stable.

Definition 3.3 *A distribution $\mathscr{L}_0$ is* max-stable *if for all $n \geq 2$, if $\{W_i,\ i \in [1:n]\}$ are iid with distribution $\mathscr{L}_0$ then there exists constants $a_n > 0$ and $b_n \in \mathbb{R}$ such that $a_n^{-1}\left(\max_{i\in[1:n]} W_i - b_n\right)$ is distributed according to $\mathscr{L}_0$.*

3.2 Highlights

3.2.1 Ronald Fisher (1890-1962). Source: wikipedia.

Sir Ronald Aylmer Fisher FRS (17 February 1890 – 29 July 1962), who published as R. A. Fisher, was an English statistician and biologist who used mathematics to combine Mendelian genetics and natural selection. This contributed to the revival of Darwinism in the early 20th century revision of the theory of evolution known as the modern synthesis. He was a prominent eugenicist in the early part of his life.

Fig. 3.2 Fisher as a child.

He worked at Rothamsted Research for 14 years from 1919, where he developed the analysis of variance (ANOVA) to analyse its immense data from crop experiments since the 1840s, and established his reputation there in the following years as a biostatistician. He is known as one of the three principal founders of population genetics. He outlined Fisher's principle as well as the Fisherian runaway and sexy son hypothesis theories of sexual selection. He also made important contributions to statistics, including the maximum likelihood, fiducial inference, the derivation of various sampling distributions among many others.

Anders Hald called him "a genius who almost single-handedly created the foundations for modern statistical science", while Richard Dawkins named him "the greatest biologist since Darwin":

Not only was he the most original and constructive of the architects of the neo-Darwinian synthesis, Fisher also was the father of modern statistics and experimental design. He therefore could be said to have provided researchers in biology and medicine with their most important research tools, as well as with the modern version of biology's central theorem.

Geoffrey Miller said of him:

"To biologists, he was an architect of the "modern synthesis" that used mathematical models to integrate Mendelian genetics with Darwin's selection theories. To psychologists, Fisher was the inventor of various statistical tests that are still supposed to be used whenever possible in psychology journals. To farmers, Fisher was the founder of experimental agricultural research, saving millions from starvation through rational crop breeding programs."

Fisher was born in East Finchley in London, England, one of twins with the other being still-born and grew up the youngest with three sisters and one brother. From 1896 until 1904 they lived at Inverforth House in London, where English Heritage installed a blue plaque in 2002, before moving to Streatham. His mother, Kate, died from acute peritonitis when he was 14, and his father, George, then lost his business as a successful partner in Robinson & Fisher, auctioneers and fine art dealers, 18 months later.

Lifelong poor eyesight caused his rejection by the British Army for World War I, but also developed his ability to visualize problems in geometrical terms, but not in writing mathematical solutions, or proofs. He entered Harrow School age 14 and won the school's Neeld Medal in mathematics. In 1909, he won a scholarship to Gonville and Caius College, Cambridge.

Fisher worked for six years as a statistician in the City of London and taught physics and maths at a sequence of public schools, and at the Thames Nautical Training College, and Bradfield College where he settled with his new bride, Eileen Guinness, with whom he had two sons and six daughters. In 1919 he began working at Rothamsted Research.

His fame grew and he began to travel and lecture widely. In 1931, he spent six weeks at the Statistical Laboratory at Iowa State College where he gave three lectures per week, and met many American statisticians, including George W. Snedecor before returning again in 1936. In 1937, he visited the Indian Statistical Institute in Calcutta, and its one part-time employee, P. C. Mahalanobis, often returning to encourage its development, being the guest of honour at its 25th anniversary in 1957 when it had 2000 employees.

Memorial plaque over his mortal remains, lectern-side aisle of St Peter's Cathedral, Adelaide His marriage disintegrated during World War II and his oldest son George, an aviator, was killed in combat. His daughter and one of his biographers, Joan, married the noted statistician George E. P. Box. In 1957, a retired Fisher emigrated to Australia where he spent time as a senior research fellow at the Australian Commonwealth Scientific and Industrial Research Organisation (CSIRO) in Adelaide, where he died in 1962, with his remains interred within St Peter's Cathedral.

Chapter 4

The domain of attraction

Contents

Keywords 4.1 *Domain of attraction, slowly-varying functions, representation theorem.*

Definition 4.1 *If $\mathscr{L}$ is a distribution with cdf F_X such that there exist $\{a_n,\, n \in \mathbb{N}\}$ and $\{b_n,\, n \in \mathbb{N}\}$ sequence of constants and $\{X_n,\, n \in \mathbb{N}\}$ a sequence of iid random variables such that $X_i \sim \mathscr{L}$ such that $a_n > 0$ and*

$$\frac{\max_{i\in[1:n]} X_i - b_n}{a_n} \overset{w}{\Rightarrow} \mathscr{L}_0$$

Then we say that $\mathscr{L}$ belongs to the domain of attraction of *$\mathscr{L}_0$ and we write $\mathscr{L} \in \mathscr{D}(\mathscr{L}_0)$. By abuse of notation, we also identify the set of probability measures $\mathscr{D}(\mathscr{L}_0)$ with the set of the associated cdfs and in that case, we will also write $F_X \in \mathscr{D}(\mathscr{L}_0)$.*

4.1 General characterization

Let U be the function defined by

$$U(t) = F_X^{\leftarrow}\left(1 - \frac{1}{t}\right), \qquad t > 1$$

For $\xi \in \mathbb{R}$, define

$$H_\xi(x) = \exp\left\{-(1+\xi x)^{-1/\xi}\right\}$$

if $1+\xi x > 0$. Note that H_ξ is a cdf and write $\mathscr{H}_\xi$ the associated distribution. The following proposition is admitted. The interested reader may get some intuition of the result from inspection of the proof of the Fisher-Tippet Theorem.

Proposition 4.2 *Let $\xi \in \mathbb{R}$. Then, $F_X \in \mathscr{D}\left(\mathscr{H}_\xi\right)$ if and only if for all $x, y > 0$, $y \neq 1$*

$$\lim_{s\to\infty} \frac{U(sx)-U(s)}{U(sy)-U(s)} = \begin{cases} \frac{x^\xi - 1}{y^\xi - 1} & \text{if } \xi \neq 0 \\ \frac{\log(x)}{\log(y)} & \text{otherwise} \end{cases}$$

Define $\overline{F}_X(x) = 1 - F_X(x) = \mathbb{P}(X > x)$.

Proposition 4.3 *Let $\xi \in \mathbb{R}$. Then,*

$$F_X \in \mathscr{D}\left(\mathscr{H}_\xi\right) \qquad \Longleftrightarrow \qquad \lim_{n\to\infty} n\overline{F}_X(xa_n + b_n) = -\log H_\xi(x)$$

for some sequence of constants $\{a_n,\, n \in \mathbb{N}\}$ and $\{b_n,\, n \in \mathbb{N}\}$ such that $a_n > 0$. Moreover, in that case,

$$a_n^{-1}(\hat{X}_n - b_n) \overset{w}{\Rightarrow} Z \qquad \text{where} \qquad Z \sim \mathscr{H}_\xi$$

PROOF. Let $F_X \in \mathscr{D}\left(\mathscr{H}_\xi\right)$. Then, $\lim_n \left(1 - \overline{F}_X(xa_n + b_n)\right)^n = H_\xi(x)$ for some sequence of constants $\{a_n,\, n \in \mathbb{N}\}$ and $\{b_n,\, n \in \mathbb{N}\}$ such that $a_n > 0$. Taking the logarithm, we obtain

$$\lim_{n\to\infty} n \log\left(1 - \overline{F}_X(xa_n + b_n)\right) = \log H_\xi(x) \,.$$

This implies that for $1 + \xi x > 0$, $\overline{F}_X(xa_n + b_n)$ tends to 0 and by a Taylor expansion of the log,

$$\lim_{n\to\infty} n\overline{F}_X(xa_n + b_n) = -\log H_\xi(x) \,.$$

The converse is obvious. ∎

Let $\boxed{x_{F_X} = F_X^{\leftarrow}(1) = \inf\{x \in \mathbb{R} : F_X(x) \geq 1\} = \inf\{x \in \mathbb{R} : F_X(x) = 1\}}$. We have the following more general result.

Proposition 4.4 *Let $\xi \in \mathbb{R}$. Then, $F_X \in \mathscr{D}\left(\mathscr{H}_\xi\right)$ if and only if there exists a function α such that for all x satisfying $1 + \xi x > 0$, we have*

$$\lim_{u\to x_{F_X}^-} \frac{\overline{F}_X(x\alpha(u) + u)}{\overline{F}_X(u)} = \begin{cases} (1+\xi x)^{-1/\xi} & \text{if } \xi \neq 0 \\ \mathrm{e}^{-x} & \text{if } \xi = 0 \end{cases}$$

PROOF. Assume that there exists a function α that satisfies the limit described in the statement of the Proposition. Assume for simplicity that $\overline{F}_X$ is continuous. Then, choosing $b_n = U(n)$, we have $\overline{F}_X(b_n) = 1/n$. Taking $u = b_n$, we thus have $\lim_{n\to\infty} n\overline{F}_X(x\alpha(b_n) + b_n) = -\log H_\xi(x)$. This implies that $F_X \in \mathscr{D}\left(\mathscr{H}_\xi\right)$ by Proposition 4.3

The converse is more involved and is admitted in this course.

∎

4.2 Domain of attraction for the Frechet and Weibull distributions

Definition 4.5 *We say that a function L is slowly-varying if $L(t) > 0$ for sufficiently large t and for all $x > 0$,*

$$\lim_{t\to\infty} \frac{L(tx)}{L(t)} = 1 \ .$$

The following representation theorem gives the general expression of the slowly-varying functions.

Proposition 4.6 *Let L be a slowly-varying function. Then, there exist two measurable functions c and κ such that*

$$\lim_{x\to\infty} c(x) = c_0 \in (0,\infty) \ , \qquad et \qquad \lim_{x\to\infty} \kappa(x) = 0 \ ,$$

and $a \in \mathbb{R}$ such that for all $x \geq a$,

$$L(x) = c(x) \exp \int_a^x \frac{\kappa(u)}{u} \mathrm{d}u$$

Remark 4.7 *If $t \mapsto g(t)$ is positive for sufficiently large t and if for all $x > 0$, $\lim_{t\to\infty} g(tx)/g(t) = x^\beta$, then $g(x) = x^\beta L(x)$ where L is slowly varying. In such a case, we say that $\boxed{g \text{ is varying at order } \beta}$.*

Theorem 4.8. *The cdf function F_X belongs to the domain of attraction of the Frechet distribution with parameter $\alpha > 0$ if and only if $\boxed{\overline{F}_X(x) = x^{-\alpha} L(x)}$ for all $x > 0$ where the function L is slowly-varying. In particular, $x_{F_X} = \infty$. Moreover, if $F_X \in \mathscr{D}(\Phi_\alpha)$, then, letting $a_n = U(n) = F_X^{\leftarrow}\left(1 - \frac{1}{n}\right)$, the sequence $\{a_n^{-1} \hat{X}_n,\ n \in \mathbb{N}^*\}$ converges in distribution to a random variable of cdf Φ_α.*

PROOF. Assume that $\overline{F}_X(x) = x^{-\alpha} L(x)$ where L is slowly-varying. We use the notation of Proposition 4.6. We have $\overline{F}_X(x) \sim_\infty g(x)$ where $g(x) = x^{-\alpha} c_0 \exp \int_a^x \frac{\kappa(u)}{u} \mathrm{d}u$ is a continuous function. Set $a_n = U(n)$. We have $\overline{F}_X(a_n) \leq 1/n \leq \overline{F}_X(a_n^-)$ in general but to simplify, we assume that F_X is continuous at a_n and therefore $\overline{F}_X(a_n) = 1/n$. For $x > 0$, we thus have

$$\lim_{n\to\infty} n\overline{F}_X(a_n x) = \lim_{n\to\infty} \frac{\overline{F}_X(a_n x)}{\overline{F}_X(a_n)} = x^{-\alpha} \ .$$

As in the proof of Proposition 4.3, this implies that $F_X \in \mathscr{D}(\Phi_\alpha)$. The converse is admitted. ■

Theorem 4.9. *The cdf function F_X belongs to the domain of attraction of the Weibull distribution with parameter $\alpha > 0$ if and only if $x_{F_X} < \infty$ and $\boxed{\overline{F}_X\left(x_{F_X} - \frac{1}{x}\right) = x^{-\alpha} L(x)}$ for all $x > 0$ where the function L is slowly-varying. Moreover, if $F_X \in \mathscr{D}(\Psi_\alpha)$, then, letting $a_n = x_{F_X} - U(n) = x_{F_X} - F_X^{\leftarrow}\left(1 - \frac{1}{n}\right)$, the sequence $\{a_n^{-1}(\hat{X}_n - x_{F_X}),\ n \in \mathbb{N}^*\}$ converges in distribution to a random variable of cdf Ψ_α.*

PROOF. The proof is similar to the one of Theorem 4.8 and is omitted for brevity. ■

We admit the following result.

Proposition 4.10 (Von Mises criterium) *(i) Assume that*

$$\lim_{x\to\infty} \frac{xf(x)}{\overline{F}_X(x)} = \alpha > 0 ,$$

then F_X belongs to the domain of attraction of the Frechet distribution of parameter α.

(ii) Suppose that the density f_X is positive on an interval (z, x_{F_X}) where $x_{F_X} < \infty$. Assume that

$$\lim_{x\to x_{F_X}^-} \frac{(x_{F_X} - x) f_X(x)}{\overline{F}_X(x)} = \alpha > 0 ,$$

then F_X belongs to the domain of attraction of the Weibull distribution of parameter α.

4.3 Highlights

4.3.1 Gnedenko (1912-1995). Source: wikipedia.

Boris Vladimirovich Gnedenko (January 1, 1912 – December 27, 1995) was a Soviet mathematician and a student of Andrey Nikolaevich Kolmogorov. He was born in Simbirsk (now Ulyanovsk), Russia, and died in Moscow. He is perhaps best known for his work with Kolmogorov, and his contributions to the study of probability theory, particularly extreme value theory, with such results as the Fisher–Tippett–Gnedenko theorem. Gnedenko was appointed as Head of the Physics, Mathematics and Chemistry Section of the Ukrainian Academy of Sciences in 1949, and also became Director of the Kiev Institute of Mathematics in the same year.

Fig. 4.1 Boris Gnedenko.

Gnedenko was a leading member of the Russian school of probability theory and statistics. He also worked on applications of statistics to reliability and quality control in manufacturing. He wrote a history of mathematics in Russia (published 1946) and with O. B. Sheynin the section on the history of probability theory in the history of mathematics by Andrei Kolmogorov and Adolph P. Yushkevich (published 1992). In 1958 he was a plenary speaker at the International Congress of Mathematicians in Edinburgh with a talk entitled ”Limit theorems of probability theory”.

Chapter 5

List of exercises

Exercise 5.1. (i) Show that the cdf F_X is cadlag (in french, continue à droite, limitée à gauche), that is, for all $x \in \mathbb{R}$, F_X is continuous in a right neighbourhood of x and has a limit in a left neighbourhood of x.

(ii) Show that for all $u \in (0,1)$, we have $F_X(F_X^{\leftarrow}(u)) \geq u$.

(iii) Show that $F_X^{\leftarrow}$ is nondecreasing on $(0,1)$.

(iv) Show that $F_X^{\leftarrow}$ is left-continuous on $(0,1)$.

(v) Show that the discontinuity points of F_X are countable. Deduce tqqhat the continuity points of F_X are dense.

PROOF.

(i) By the dominated convergence theorem, $\lim_{t \to t_0^+} \mathbb{E}\left[\mathbb{1}_{X \leq t}\right] = \mathbb{E}\left[\lim_{t \to t_0^+} \mathbb{1}_{X \leq t}\right] = \mathbb{E}\left[\mathbb{1}_{X \leq t_0}\right]$. Therefore, $\lim_{t \to t_0^+} \mathbb{P}(X \leq t) = \mathbb{P}(X \leq t_0)$, which shows that F_X is right-continuous. Similarly, using again the dominated convergence theorem, $\lim_{t \to t_0^-} \mathbb{E}\left[\mathbb{1}_{X \leq t}\right] = \mathbb{E}\left[\lim_{t \to t_0^-} \mathbb{1}_{X \leq t}\right] = \mathbb{E}\left[\mathbb{1}_{X < t_0}\right]$. Therefore, $\lim_{t \to t_0^-} \mathbb{P}(X \leq t) = \mathbb{P}(X < t_0)$, which shows that F_X has a left-limit $\mathbb{P}(X < t_0)$, which is in general different from $\mathbb{P}(X \leq t_0)$, except if $\mathbb{P}(X = t_0) = 0$. We can sum up this question by saying that
$$\mathbb{P}(X = t_0) = \mathbb{P}(X \leq t_0) - \mathbb{P}(X < t_0) = \mathbb{P}(X \leq t_0) - \lim_{t \nearrow t_0} \mathbb{P}(X \leq t)$$

(ii) By definition of $F_X^{\leftarrow}$, there exists a sequence (y_n) such that for all $n \in \mathbb{N}$, $F_X(y_n) \geq u$ and $y_n \searrow F_X^{\leftarrow}(u)$. Since F_X is right-continuous, we have that $\lim_n F_X(y_n) = F_X(F_X^{\leftarrow}(u))$. Combining with the fact that $F_X(y_n) \geq u$ holds for all $n \in \mathbb{N}$, we conclude that $F_X(F_X^{\leftarrow}(u)) \geq u$.

(iii) If $u \leq v < 1$. Then,
$$\{y \in \mathbb{R} : F_X(y) \geq v\} \subset \{y \in \mathbb{R} : F_X(y) \geq u\}$$
Taking the infimum in both sides yields:
$$F_X^{\leftarrow}(v) = \inf\{y \in \mathbb{R} : F_X(y) \geq v\} \geq \inf\{y \in \mathbb{R} : F_X(y) \geq u\} = F_X^{\leftarrow}(u)$$
This shows that $F_X^{\leftarrow}$ is nondecreasing.

(iv) Let $u_0 \in (0,1)$. By the previous question, $F_X^{\leftarrow}$ is nondecreasing. This implies that $\lim_{u \nearrow u_0} F_X^{\leftarrow}(u)$ exists and is less than $F_X^{\leftarrow}(u_0)$. We now show by contradiction that this limit is $F_X^{\leftarrow}(u_0)$. Assume that this is not the case. Then, there exists $y \in \mathbb{R}$ such that
$$\lim_{u \nearrow u_0} F_X^{\leftarrow}(u) < y < F_X^{\leftarrow}(u_0) \tag{5.1}$$
Since $F_X^{\leftarrow}$ is nondecreasing, the first inequality above shows that $y > F_X^{\leftarrow}(u)$ for all $u < u_0$. Applying the nondecreasing function F_X, we get:
$$F_X(y) \geq F_X \circ F_X^{\leftarrow}(u) \geq u$$
Letting $u \nearrow u_0$, we obtain $F_X(y) \geq u_0$, this in turn implies that
$$F_X^{\leftarrow}(u_0) = \inf\{x \in \mathbb{R} : F_X(x) \geq u_0\} \leq y$$
which contradicts (5.1)

(v) If x is a discontinuity point of F_X, we have seen in the first question that $\mathbb{P}(X = x) = F_X(x) - \lim_{y \nearrow x} F_X(y) > 0$. We can thus associate to each discontinuity point of F_X a rational number in the interval $(\lim_{y \nearrow x} F_X(y), F_X(x)]$. Since each of these intervals are disjoint, we conclude that discontinuity points are countable. Any countable set has a dense complement and this concludes the proof.

■

Exercise 5.2. Show that if F_X is increasing and continuous and if X is a random variable with cdf F_X, then $F_X(X) \sim \text{Unif}[0,1]$.

Exercise 5.3. Let (X_i) be iid random variables with exponential distribution of parameter λ. How can we sample $\hat{X}_n$ and $\check{X}_n$ using only one random variable uniformly distributed on $(0,1)$?

Exercise 5.4. Let (X_i) be iid random variables with Cauchy distribution $x \to \frac{1}{\pi(1+x^2)}$. How can we sample $\hat{X}_n$ and $\check{X}_n$ using only one random variable uniformly distributed on $(0,1)$?

Exercise 5.5. Prove Lemma 1.6

Exercise 5.6. Using Lemma 1.9, check that we can obtain Lemma 1.1.

Exercise 5.7. Prove Lemma 1.11 along the lines of Proposition 1.10.

Exercise 5.8. Let $(U_i)_{1\le i\le n}$ be iid random variables according to a uniform distribution on $[0,1]$. Write $(U_{(1,n)},\ldots,U_{(n,n)})$ the associated order statistics.

1. Show that $U_{(n,n)}$ has the same distribution as $U_1^{\frac{1}{n}}$.
2. Find the distribution of the vector $(U_{(1,n)},\ldots,U_{(n,n)})$.
3. Show that $(U_{(n,n)}, \frac{U_{(n-1,n)}}{U_{(n,n)}},\ldots,\frac{U_{(1,n)}}{U_{(2,n)}})$ has the same distribution as $(U_n^{\frac{1}{n}}, U_{n-1}^{\frac{1}{n-1}},\ldots,U_1)$.
4. Deduce a way to draw a n-uplet of random variables with the same distribution as $(U_{(i,n)})_{1\le i\le n}$.

Exercise 5.9. We admit that a sequence of random vectors (Z_n) taking values in $\mathbb{R}^p$ converges in distribution to Z if and only if one of the following equivalent conditions is satisfied:

- for all constant vectors $u \in \mathbb{R}^p$; we have $\lim_n \mathbb{E}\left[e^{iu^T Z_n}\right] = \mathbb{E}\left[e^{iu^T Z}\right]$,
- for all real valued bounded functions $f : \mathbb{R}^p \to \mathbb{R}$ such that $\mathbb{P}(Z \in D_f) = 0$ where D_f is the set of discontinuity points of f, we have $\lim_n \mathbb{E}[f(Z_n)] = \mathbb{E}[f(Z)]$.

1. Show the Slutsky lemma: If $X_n \xrightarrow{\mathbb{P}\text{-prob}} c$ where c is a constant and if $Y_n \overset{w}{\Rightarrow} Y$, then $(X_n,Y_n) \overset{w}{\Rightarrow} (c,Y)$ that is for all continuous functions f, $f(X_n,Y_n) \overset{w}{\Rightarrow} f(c,Y)$. **Hint:** We can admit that to show $(X_n,Y_n) \overset{w}{\Rightarrow} (c,Y)$ is equivalent to show that for all $u,v \in \mathbb{R}$, $\lim_{n\to\infty} \mathbb{E}\left[e^{i(uX_n+vY_n)}\right] = \mathbb{E}\left[e^{i(uc+vY)}\right]$.
2. Extend the Slutsky lemma in the case where X_n and Y_n are random vectors.
3. With the notation introduced in Proposition 1.13 and its proof, set

$$A_n = n^{1/2}\frac{i_n/n - p}{[p(1-p)]^{1/2}} \qquad\qquad B_n = n^{1/2}\frac{j_n/n - p}{[p(1-p)]^{1/2}}$$

and show the convergence of these quantities. Apply the extended Slutsky's lemma to obtain an alternative proof of Proposition 1.13.

Exercise 5.10. Let (X_n) be iid random variables such that $X_1 \sim \text{Unif}[0,\theta]$ where $\theta > 0$. Find a confidence interval for θ using $\hat{X}_n$ and Lemma 2.1. Compare with a confidence interval based on $\bar{X}_n$ and the Central Limit Theorem.

Exercise 5.11. Let (X_n) be iid random variables such that $X_1 \sim \text{Unif}[0,\theta]$ where $\theta > 0$. Using Lemma 1.9, show that the sequence $\{n\left(\frac{X_{(n-1,n)}}{\theta} - 1\right),\ n \in \mathbb{N}^*\}$ converges in distribution to some random variable Z and give the expression for the cdf of Z.

Exercise 5.12. Let (X_n) be iid random variables such that $X_1 \sim \mathrm{Unif}[0,\theta]$ where $\theta > 0$. Using several times the Slutsky lemma and Lemma 2.1, proves that $\sqrt{n}\left(\frac{\bar{X}_n}{\hat{X}_n} - 1/2\right)$ converges in distribution to some random variable and give the limiting distribution.

Exercise 5.13. In the proof of Theorem 3.2, show that we get a Gumbel distribution in the case $\xi = 0$.

Exercise 5.14. From inspection of the proof of Theorem 3.2, explain Proposition 4.2

Exercise 5.15. Show the converse implication of Proposition 4.6.

Exercise 5.16. Let $(E_n)_{n\geq 0}$ be iid random variables such that $E_1 \sim \exp(1)$. Set $S_k = E_1 + \ldots + E_k$ for all $k \in \mathbb{N}$. Then, show that

$$\left(\frac{S_1}{S_{n+1}}, \ldots, \frac{S_n}{S_{n+1}}\right) \stackrel{\mathscr{L}}{=} \left(U_{(1,n)}, \ldots, U_{(n,n)}\right)$$

where $\left(U_{(1,n)}, \ldots, U_{(n,n)}\right)$ is the order statistics associated to iid random variables according to the uniform distribution on $[0,1]$.

Exercise 5.17. Use Proposition 4.2, Theorem 4.8, Theorem 4.9 and to obtain the domain of attraction of a uniform distribution on $[0,\theta]$, of an exponential distribution of parameter $\lambda > 0$ and of the Cauchy distribution of density $x \mapsto \frac{1}{\pi(1+x^2)}$.

Show that the geometric distribution does not belong to any domain of attraction.

Exercise 5.18. Show that a discrete distribution with finite support does not belong to any domain of attraction.

$$(\widehat{\lambda}, \widehat{\lambda'}) = \left(\left(\bar{x_k} + \sqrt{\bar{x_k^2} - (\bar{x_k})^2}\right), \frac{1}{\sqrt{\bar{x_k^2} - (\bar{x_k})^2}}\right)$$

Index