$\longrightarrow$ using $Y_1, \dots, Y_k$ we can apply MLE to estimate the parameters $\gamma, \beta = \beta(u)$ of the distr. $P_{\gamma,\beta}$:

The prob. density of $P_{\gamma,\beta}$ is given by

$$f_{\gamma,\beta}(x) = \begin{cases} \frac{1}{\beta}\left(1+\frac{\gamma x}{\beta}\right)^{-\frac{1}{\gamma}-1}, & x>0, \text{ if } \gamma>0 \\ \frac{1}{\beta}\left(1+\frac{\gamma x}{\beta}\right)^{-\frac{1}{\gamma}-1}, & 0 \le x \le -\frac{\beta}{\gamma}, \text{ if } \gamma<0 \\ \beta^{-1}e^{-x/\beta}, & x \in \mathbb{R}, \text{ if } \gamma=0 \end{cases}$$

$\longrightarrow$ log-likelihood function:

$$l(\gamma,\beta) = \sum_{i=1}^{k} \log\left(f_{\gamma,\beta}(Y_i)\right) = -k\log\beta - \left(1+\frac{1}{\gamma}\right)\sum_{i=1}^{k}\log\left(1+\frac{\gamma Y_i}{\beta}\right)$$

for $\gamma \neq 0$

$\longrightarrow$ ML estimator: $(\hat{\gamma},\hat{\beta}) = \operatorname{argmax}_{\gamma,\beta} l(\gamma,\beta)$

We can now use $\hat{\gamma},\hat{\beta}$ to estimate the dike height z_p which is necessary to avoid flooding on some day with prob. $1-p$

Since $F_u(x) = P(X_1-u \le x \mid X_1 > u)$, $x>u$ we get from (4.10) that

$$P(X_1-u>x) = P(X_1>u)\,\underbrace{\bar{P}_{\hat{\gamma},\hat{\beta}}(x)}_{\text{tail distr.}} = P(X_1>u)\left(1+\frac{\hat{\gamma}x}{\hat{\beta}}\right)^{-\frac{1}{\hat{\gamma}}}$$

$$x = z_p - u \ge 0 \Rightarrow \quad \underbrace{P(X_1>z_p)}_{\text{Def. } p} = P(X_1>u)\left(1+\hat{\gamma}\frac{(z_p-u)}{\hat{\beta}}\right)^{-\frac{1}{\hat{\gamma}}}$$

$$\Rightarrow \quad z_p = u + \frac{\hat{\beta}}{\hat{\gamma}}\left(\left(\frac{P(X_1>u)}{p}\right)^{\hat{\gamma}} - 1\right)$$

We can estimate $P(X_1>u) = \bar{F}(u)$ by

$$\bar{F}_n(u) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{(u,\infty)}(x_i) \quad (F_n \text{ empirical (distr.)})$$

and obtain that

$$z_p = u + \frac{\hat{\sigma}}{\hat{\xi}} \left(\left(\frac{1}{n \cdot p} \sum_{i=1}^{n} 1_{(u,\infty)}(x_i) \right)^{\hat{\xi}} - 1 \right) \qquad (4.11)$$

5. Weak convergence of point processes

We now want to study the convergence of extremal events from a more general point of view, that is from the perspective of the weak convergence of point processes.

In order to get an idea about this new framework for the analysis of extremal events, consider the following example:

Ex. 5.1 (Point process of exceedances)
Let $X_n, n \geq 1$ be a sequence of r.v.'s and $u_n, n \geq 1$ a sequence of real numbers.
Further, denote by ε_x the Dirac measure in x, that is

$$\varepsilon_x(A) = \begin{cases} 1, & \text{if } x \in A \\ 0 & \text{else} \end{cases}$$

for $A \in \mathcal{A}$ ($\mathcal{A}$-algebra)
Then the point process of exceedances to the threshold u_n is defined as a "stochastic" measure N_n given by

$$N_n(A) = \sum_{i=1}^{n} \varepsilon_{n^{-1}\cdot i}(A) \cdot \mathbb{1}_{\{X_i > u_n\}}, \quad A \in \mathcal{A} = \mathcal{B}((0,1])$$

$\longrightarrow$ for each outcome of $(X_1, \dots, X_n)$ N_n as a finite linear combination of measures on $\mathcal{B}((0,1])$ is a measure

$A = (0,1]$ $\longrightarrow$

$$N_n((0,1]) = \#\{i : n^{-1}\cdot i \leq 1 \text{ and } X_i > u_n\} \quad (\#: \text{number of})$$
$$= \#\{i \leq n : X_i > u_n\}$$

$\longrightarrow$ N_n counts the number of exceedances of the threshold u_n by the sequence $X_1, \dots, X_n$

In order to see the link between the "convergence" of N_n and the (weak) convergence of $M_n = \max(X_1, \dots, X_n)$ or $X_{(n-k+1)}$ (k-th largest order statistics), let us have a look at the events $\{N_n((0,1]) = 0\}$ and $\{N_n((0,1]) < k\}$, respectively:

$$\begin{aligned}\{N_n((0,1]) = 0\} &= \{\#\{i \le n : X_i > u_n\} = 0\} \\ &= \{\text{none of the } X_i\text{'s}, i \le n \text{ exceeds } u_n\} \\ &= \{\max\{X_1, \dots, X_n\} \le u_n\} \qquad (5.1)\end{aligned}$$

$$\begin{aligned}\{N_n((0,1]) < k\} &= \{\#\{i \le n : X_i > u_n\} < k\} \\ &\quad \{\text{fewer than } k \text{ among } X_i\text{'s}, i \le n \\ &\quad \text{exceed } u_n\} \\ &= \{X_{(n-k+1)} \text{ cannot exceed } u_n\} \\ &= \{X_{(n-k+1)} \le u_n\} \qquad (5.2)\end{aligned}$$

So, if e.g. $u_n = a_n x + b_n$, $a_n > 0$, $b_n \in \mathbb{R}$, $n \ge 1$, $x \in \mathbb{R}$ in (5.1) and (5.2) then "convergence" of N_n implies (weak) convergence of

$$\frac{M_n - b_n}{a_n}$$

and

$$\frac{X_{(n-k+1)} - b_n}{a_n}$$

On the other hand, if e.g. X_i, $i \ge 1$ is i.i.d. with $F = F_{X_1} \in MDA(\Phi_\alpha)$ then it follows from (5.1) that

$$P(N_n((0,1]) = 0) = P(a_n^{-1}(M_n - b_n) \le x) \xrightarrow[n \to \infty]{} \Phi_\alpha(x)$$

for all $x \in \mathbb{R}$ for some $a_n > 0$, $b_n \in \mathbb{R}$, $n \ge 1$ (see Th. 3.2.5), if we choose $u_n = a_n x + b_n$ in (5.1)

The latter example shows that an appropriate concept of convergence of N_n or more generally of point processes would provide a more general framework for the study of extremal events than in Section 2 and grant a deeper understanding of the behaviour of such events.

→ next aim is to introduce the concepts of a point process and the (weak) convergence of such processes

For this purpose we need some notions and results:

In the sequel (let $E = \mathbb{R}^d$ (or more generally a separable metric space which is locally compact and has a countable base)

In order to define point processes on E, we need the notion of a random measure, that is a r.v. with values in a space of measures

A possible choice of such a space could be the space

$$M_+(E) \tag{5.3}$$

of non-negative Radon measures, that is of measures μ on $\mathcal{A} = \mathcal{B}(E)$ (= smallest σ-algebra containing all open sets of E) such that

$$\mu(K) < \infty$$

for all compact sets $K \subseteq E$.

Ex. For each outcome of $X_1, \dots, X_n$ in Ex. 5.1 the point process of exceedances $N_n \in M_+(E)$ for $E = (0,1]$

(84) Is $N_n : \Omega \longrightarrow M_+(E)$ also a r.v., that is a random measure?

The latter, however presumes that there is a σ-algebra $\mathcal{B}$ on $M_+(E)$ such that

N_n $\tilde{\mathcal{F}}$, $\mathcal{B}$-measurable,

where $\tilde{\mathcal{F}}$ is the σ-algebra on Ω, that is N_n satisfies the condition

$$\{\omega \in \Omega : N_n(\omega) \in \mathcal{B}\} \in \mathcal{F}$$

for all $\mathcal{B} \in \mathcal{B}$

A natural choice of $\mathcal{B}$ could be

$$\mathcal{B} = \mathcal{M}_+(E) := \mathcal{B}(M_+(E)) \qquad (5.4)$$

= smallest σ-algebra which contains all open sets of $M_+(E)$

$\longrightarrow$ $\mathcal{M}_+(E)$ Borel-σ-algebra on $M_+(E)$

But what is an open set $\mathcal{O}$ in $M_+(E)$?

If we have a metric, that is a distance between measures in $M_+(E)$ we can as in the case of $\mathbb{R}^d$ define sets $\mathcal{O} \subseteq M_+(E)$ as open, if

$$\mathcal{O} = \bigcup_{i \in I} U_{r_i}(\mu_i),$$

where

$$U_r(\mu) := \{\tilde{\mu} \in M_+(E) : d(\tilde{\mu}, \mu) < r\}$$

is an open ball with radius $r > 0$ and center μ and where d is a metric on $M_+(E)$, that is

(i) $d(\mu_1, \mu_2) = d(\mu_2, \mu_1)$ for all $\mu_1, \mu_2 \in M_+(E)$

(ii) $d(\mu_1, \mu_2) = 0$, iff $\mu_1 = \mu_2$

(iii) $d(\mu_1, \mu_2) \leq d(\mu_1, \mu_3) + d(\mu_3, \mu_2)$ (triangle inequality)

forall $\mu_1, \mu_2, \mu_3 \in M_+(E)$.

$\longrightarrow$ $d(\mu_1, \mu_2)$ "distance" between μ_1 and μ_2

What could be a natural definition of a metric or distance between measures on $M_+(E)$?

$\longrightarrow$ A natural notion of a metric d on $M_+(E)$ should be <u>compatible with</u> the notion of convergence in distribution on $\mathbb{R}$:

$$F_n \xrightarrow[n\to\infty]{d} F$$

$\Longleftrightarrow$

$$F_n(x) \xrightarrow[n\to\infty]{} F(x)$$

for all continuity points x of F

$\Longleftrightarrow$

$$\int_{\mathbb{R}} f(y)\, m_n(dy) \xrightarrow[n\to\infty]{} \int_{\mathbb{R}} f(y)\, m(dy) \qquad (5.5)$$

for all functions f of the form $\mathbb{1}_{(-\infty, x]}$, x continuity point of F.

Here m_n, m are the probability measures on $\mathcal{B}(\mathbb{R})$ associated with F_n, F, respectively (see Ex. 2.3.4.), that is $m_n((a,b]) = F_n(b) - F_n(a)$ and $m((a,b]) = F(b) - F(a)$ forall $a < b$

In fact, a metric d on $M_+(E)$, which "produces" the type of convergence in (5.5) is the following:

$$d(\mu, \mu') := \sum_{i \geq 1} 2^{-i} \left(1 - \exp\left(-|\mu(h_i) - \mu'(h_i)|\right)\right), \qquad (5.6)$$

where

$$\mu(h) := \int_E h(y)\, \mu(dy)$$