Fiche
Risk measure

Antoine Falck

December 12, 2017

Contents

Part I

1 Monetary risk measures

1.1 Introduction

We want to measure the risk associated to a given portfolio. We assume that $X : \Omega \to \mathbb{R}$ is a function that describes the value of the portfolio at a certain time horizon and given a management rule, on the sceanario $\omega \in \Omega$.

We want to define a function $\rho(X)$ s.t.

(i) the position X is acceptable[1] if $\rho(X) \le 0$;

(ii) if $\rho(X) > 0$, then $\rho(X)$ is the minimal capital requirement, *i.e.* the minimal amount of cash to add to the portfolio so that he becomes acceptable.

1.2 Risk measures and set of acceptable positions

Let Ω describe the space of scenarios and $X : \Omega \to \mathbb{R}$ the updated portfolio value. Assume that $\exists M > 0$, $\forall \omega \in \Omega$, $|X(\omega)| \le M$. We note $\mathcal{X} := \{X : \Omega \to \mathbb{R}, X \text{ bounded}\}$ the vector space of bounded functions.

Definition 1.1 (Risk measure)**.** The function $\rho : \mathcal{X} \to \mathbb{R}$ is a risk measure if;

(i) $X, Y \in \mathcal{X}$ s.t. $X \le Y$, $\rho(X) \ge \rho(Y)$ (monotony);

(ii) $\forall X \in \mathcal{X}$, $\forall m \in \mathbb{R}$, $\rho(X + m) = \rho(X) - m$ (cash invariance) ;

(iii) $\rho(0) = 0$ (normalization).

Remark. The cash invariance property enables us to see $\rho(X)$ as the minimal capital requirement, indeed

$$\rho(X + \rho(x)) \;=\; \rho(X) - \rho(X) \;=\; 0.$$

And $\forall m \ge \rho(X)$, $\rho(X + m) = \rho(X) - m \le 0$.

For a given risk measure ρ we define

$$\mathcal{A}_\rho \;:=\; \{X \in \mathcal{X}, \rho(X) \le 0\}$$

the set of acceptable positions.

Proposition 1.1. *We have the following properties:*

(i) if $X \in \mathcal{A}_\rho$ and $Y \in \mathcal{X}$ s.t. $X \le Y$, then $Y \in \mathcal{A}_\rho$;

(ii) $\inf\{m \in \mathbb{R} : m \in \mathcal{A}_\rho\} = 0$

Proposition 1.2 (Lipschitz property)**.** *If ρ is a risk measure, then $\forall X, Y \in \mathcal{X}$, $|\rho(X) - \rho(Y)| \le \|X - Y\|_\infty$.*

[1] From a regulatory point of view.

Proof.

$$\begin{array}{rrcl} & X & \le & Y + \|X - Y\|_\infty \\ \Leftrightarrow & \rho(X) & \ge & \rho(Y + \|X - Y\|_\infty) \\ \Leftrightarrow & \rho(X) & \ge & \rho(Y) - \|X - Y\|_\infty \\ \Leftrightarrow & \rho(X) - \rho(Y) & \ge & -\|X - Y\|_\infty \end{array}$$

We show the same way that $\rho(X) - \rho(Y) \;\le\; \|X - Y\|_\infty$. □

Definition 1.2 (Convexity)**.** A risk measure ρ is said convex if $\forall X, Y \in \mathcal{X}$, $\forall \lambda \in [0, 1]$, $\rho(\lambda X + (1 - \lambda)Y) \le \lambda\rho(X) + (1 - \lambda)\rho(Y)$.

Remark. The interest of a convex risk measure is that it takes into account protfolio diversification. Indeed suppose $X, Y \in \mathcal{X}$ with $\rho(X) = \rho(Y)$, then for all $\lambda \in [0, 1]$, $\rho(\lambda X + (1 - \lambda)Y) \le \rho(X)$.

Remark. If ρ is convex, $\psi(n) = \rho(nX)$ is also convex, and in particular $n \mapsto \rho((n + 1)X) - \rho(nX)$ is growing with n. Which is desirable because for liquidity reasons it is riskier to buy the 1000-th equity than the first one.

Definition 1.3 (Positive homogeneity)**.** A risk measure ρ is said positively homogeneous if $\forall \lambda > 0$, $\forall X \in \mathcal{X}$, $\rho(\lambda X) = \lambda\rho(X)$.

We say that a risk measure is consistent if it is a convex and positively homogeneous risk measure.

Example 1.1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.

$$\rho(X) \;:=\; \mathbb{E}[-X]\,, \quad X \in \mathcal{X}\,,$$

is a consistent risk measure.

Example 1.2. $\rho(X) = \mathbb{E}[-X] + \alpha\sqrt{\mathrm{Var}[X]}$, for $\alpha > 0$, is not monotonic. Indeed if $\frac{X}{\mu} \sim \mathcal{B}(p)$, $\rho(X) = -\mu p + \alpha\mu\sqrt{p(1 - p)}$.

1.3 Value-at-Risk (VaR)

We fixe a threshold $\lambda \in]0, 1[$. We say that a position $X : \Omega \to \mathbb{R}$ is acceptable if

$$\mathbb{P}\{X < 0\} \;\le\; \lambda.$$

Definition 1.4 (Value-at-Risk)**.** We can then define the VaR,

$$\mathrm{VaR}_\lambda(X) \;:=\; \inf\{m \in \mathbb{R} : \mathbb{P}\{X + m < 0\} \le \lambda\}\,.$$

Proposition 1.3. *The VaR is consistent a risk measure.*

Proof. (i). If $X \le Y$, $\{m \in \mathbb{R}, \mathbb{P}\{X + m < 0\} \le \lambda\} \subset \{m \in \mathbb{R}, \mathbb{P}\{Y + m < 0\} \le \lambda\}$, follows by taking the infimum $\mathrm{Var}_\lambda(X) \ge \mathrm{Var}_\lambda(Y)$.

(ii). Let $m' \in \mathbb{R}$,

$$\begin{aligned}\mathrm{VaR}_\lambda(X+m') &= \inf\{m \in \mathbb{R} : \mathbb{P}\{X+m'+m<0\} \le \lambda\} \\ &= \inf\{m'' \in \mathbb{R} : \mathbb{P}\{X+m''<0\} \le \lambda\} \\ &\qquad - m' , \qquad \text{with } m'' := m+m' \\ &= \mathrm{VaR}_\lambda(X) - m'.\end{aligned}$$

(Positive homogeneity). Let $\alpha > 0$, $X \in \mathcal{X}$,

$$\begin{aligned}\mathrm{VaR}_\lambda(\alpha X) &= \inf\{m \in \mathbb{R} : \mathbb{P}\{\alpha X+m<0\} \le \lambda\} \\ &= \alpha \inf\{m' \in \mathbb{R} : \mathbb{P}\{X+m'<0\} \le \lambda\} \\ &\qquad \text{with } m' := \frac{m}{\alpha} \\ &= \alpha \mathrm{VaR}_\lambda(X).\end{aligned}$$

□

Remark. The VaR is not convex.

Example 1.3. $(\Omega, \mathcal{F})$ is a probability space. Let $\mathcal{Q}$ be the set of all probability measures on $(\Omega, \mathcal{F})$. Define $\gamma : \mathcal{Q} \to \mathbb{R}$ s.t. $\sup_{Q \in \mathcal{Q}} \gamma(Q) = 0$. We define

$$\rho(X) = \sup_{Q \in \mathcal{Q}} (\gamma(Q) - \mathbb{E}_Q[X]).$$

If $\mathcal{Q} = \{\mathbb{P}\}$ we find the previous exemple $\mathbb{E}[-X]$. And ρ is a convex risk measure.

1.4 Characterization of risk measures with set of acceptable positions

Definition 1.5. $\mathcal{A} \subset \mathcal{X}$ is a set of acceptable positions if:

(i) $\mathcal{A} \neq \varnothing$ and $\inf\{m \in \mathbb{R}, m \in \mathcal{A}\} = 0$;

(ii) $X \in \mathcal{A}$, $Y \in \mathcal{X}$ s.t. $X \le Y$ then, $Y \in \mathcal{A}$.

If ρ is a risk measure, $\mathcal{A}_\rho$ is a set of acceptable positions. Reciprocally if $\mathcal{A}$ is a set of acceptable positions, then

$$\rho_{\mathcal{A}} := \inf\{m \in \mathbb{R}, X+m \in \mathcal{A}\} , \quad X \in \mathcal{X} ,$$

is a risk measure.

Proof. (i). If $X \le Y$, $X+m \in \mathcal{A} \Rightarrow Y+m \in \mathcal{A}$, so $\rho_{\mathcal{A}}(X) \ge \rho_{\mathcal{A}}(Y)$.

(ii). If $X \in \mathcal{X}$, $m' \in \mathbb{R}$,

$$\begin{aligned}\rho_{\mathcal{A}}(X+m') &= \inf\{m \in \mathbb{R}, X+m'+m \in \mathcal{A}\} \\ &= \inf\{m'' \in \mathbb{R}, X+m' \in \mathcal{A}\} - m' \\ &\qquad \text{with } m'' := m'+m \\ &= \rho_{\mathcal{A}}(X) - m'.\end{aligned}$$

□

Proposition 1.4. *If ρ is a risk measure, $\rho = \rho_{\mathcal{A}_\rho}$. In particular $\rho_1 = \rho_2 \Leftrightarrow \mathcal{A}_{\rho_1} = \mathcal{A}_{\rho_2}$.*

Proof. Let $X \in \mathcal{X}$,

$$\begin{aligned}\rho_{\mathcal{A}_\rho}(X) &= \inf\{m \in \mathbb{R}, X+m \in \mathcal{A}_\rho\} \\ &= \inf\{m \in \mathbb{R}, \rho(X+m) \ge 0\} \\ &= \inf\{m \in \mathbb{R}, \rho(X)-m \ge 0\} \\ &= \rho(X).\end{aligned}$$

□

Proposition 1.5. *We have the following properties:*

(i) ρ is convex $\Leftrightarrow$ $\mathcal{A}_\rho$ is convex ;

(ii) ρ is positively homogeneous $\Leftrightarrow$ $\mathcal{A}_\rho$ is a cone.

1.5 Expected Shortfall (ES)

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a space probability.

Definition 1.6 (Quantile)**.** If $X : \Omega \to \mathbb{R}$ a r.v., we say that $q \in \mathbb{R}$ is the λ-order quantile with $\lambda \in [0,1]$ s.t. $\mathbb{P}\{X < q\} \le \lambda$ and $\mathbb{P}\{X \le q\} \ge \lambda$.

We also set

$$\begin{aligned}q_X^-(\lambda) &= \sup\{x \in \mathbb{R}, \mathbb{P}\{X<x\} < \lambda\} \\ &= \inf\{x \in \mathbb{R}, \mathbb{P}\{X \le x\} \ge \lambda\}\end{aligned}$$

$$\begin{aligned}q_X^+(\lambda) &= \inf\{x \in \mathbb{R}, \mathbb{P}\{X \le x\} > \lambda\} \\ &= \sup\{x \in \mathbb{R}, \mathbb{P}\{X<x\} \le \lambda\}.\end{aligned}$$

So $\mathrm{VaR}_\lambda(X) = -q_X^+(X) = q_{-X}^-(1-\lambda)$.

We saw that the VaR is not convex. We will see that the Expected Shortfall is more restrictive than the VaR and is convex.

Definition 1.7 (Expected Shortfall)**.** Let $\lambda \in [0,1]$, $X \in \mathcal{X}$, we define the ES associated to the threshold λ,

$$\begin{aligned}\mathrm{ES}_\lambda(X) &:= \frac{1}{\lambda}\int_0^\lambda \mathrm{VaR}_\alpha(X)\,\mathrm{d}\alpha \\ &= -\frac{1}{\lambda}\int_0^\lambda q_X^+(\alpha)\,\mathrm{d}\alpha\end{aligned}$$

Remark. On the other hand $q_X^+ \nearrow$ and $q_X^+(\alpha) \le q_X^+(\lambda)$ for all $\alpha \in]0,\lambda]$. So ES is always defined and $\mathrm{ES}_\lambda(X) \ge \frac{1}{\lambda}\int_0^\lambda \mathrm{VaR}_\lambda(X)\mathrm{d}\alpha = \mathrm{VaR}_\lambda(X)$.

Remark. If X is integrable, ES $< \infty$. Let $U \sim \mathcal{U}([0,1])$, $q_X^+(U) \sim X$. So $\mathbb{E}[|X|] = \int_0^1 |q_X^+(u)|\mathrm{d}u < \infty$.

We also have $\mathbb{E}[-q_X^+(U) \mid U \le \lambda] = \frac{1}{\lambda}\int_0^\lambda -q_X^+(u)\mathrm{d}u = \mathrm{ES}_\lambda(X)$. If X has a density, q_X^+ is bijective and

$$\mathbb{E}[-q_X^+(U) \mid U \le \lambda] = \mathbb{E}[-X \mid -X \ge \mathrm{VaR}_\lambda(X)].$$

Proposition 1.6. *Assume that X is integrable, $\lambda \in [0,1]$, and q is the λ-order quantile of X, then*

$$\mathrm{ES}_\lambda(X) = \frac{1}{\lambda}\mathbb{E}[(q-X)^+] - q.$$

Proof.

$$\begin{aligned}
\mathrm{ES}_\lambda(x) &\triangleq -\frac{1}{\lambda}\int_0^\lambda q_X(u)\,\mathrm{d}u \\
&= \frac{1}{\lambda}\int_0^\lambda \left(q - q_X^+(u)\right)\,\mathrm{d}u \;-\; q \\
&= \frac{1}{\lambda}\int_0^1 \left(q - q_X^+(u)\right)\mathbb{1}_{\{u<\lambda\}}\,\mathrm{d}u \;-\; q \\
&= \frac{1}{\lambda}\mathbb{E}\left[(q-X)^+\right] \;-\; q
\end{aligned}$$

as $q_X^+(U) \sim X$ when $U \sim \mathcal{U}([0,1])$ and $q_X^+(\lambda) = q$. □

2 Introduction to the Fenchel–Legendre transform

2.1 Recall on topology

Definition 2.1 (Topology)**.** Let $\mathcal{X}$ be a space. $\mathcal{T}$ is a topology on $\mathcal{X}$ if $\mathcal{T} \subset \mathcal{P}(\mathcal{X})$ verifies:

(i) $\varnothing, \mathcal{X} \in \mathcal{T}$;

(ii) all subset of elements of $\mathcal{T}$ is in $\mathcal{T}$, *i.e.* $\forall \theta \in \Theta$ s.t. $A_\theta \in \mathcal{T}$, $\bigcup_{\theta\in\Theta} A_\theta \in \mathcal{T}$;

(iii) if $A_1, \dots, A_n \in \mathcal{T}$ then $\bigcap_{i=1}^n A_i \in \mathcal{T}$.

The elements of $\mathcal{T}$ are called open set of $\mathcal{X}$; $A \subset \mathcal{X}$ is a closed set if $\mathcal{X}\backslash A \in \mathcal{T}$.

Definition 2.2 (Topology space)**.** Consider $\mathcal{X}$ with a topology $\mathcal{T}$, then $(\mathcal{X}, \mathcal{T})$ is a topology space.

Definition 2.3 (Hausdorff space)**.** A topology space $(\mathcal{X}, \mathcal{T})$ is a Hardoff space if

(i) $\forall x \in \mathcal{X}$, $\{x\}$ is a closed set for $\mathcal{T}$;

(ii) $\forall x, y \in \mathcal{X}$, $x \neq y$, $\exists O_x, O_y \in \mathcal{T}$ s.t. $x \in O_x$, $y \in O_y$, $O_x \cap O_y = \varnothing$.

Definition 2.4 (Base)**.** A set $\mathcal{B} \subset \mathcal{T}$ is a base for the topology $\mathcal{T}$ if $\forall A \in \mathcal{T}$, $\exists (B_\theta)_{\theta\in\Theta} \in (\mathcal{B})^\Theta$ s.t. $A = \bigcup_{\theta\in\Theta} B_\theta$.

Definition 2.5 (Compact space)**.** A set $A \subset \mathcal{X}$ is a compact space if for all covering of A, one can extract a finite sub-covering, *i.e.* $A \subset \bigcup_{\theta\in\Theta} B_\theta$, $B_\theta \in \mathcal{T}$ for all $\theta \in \Theta$, then $\exists n \in \mathbb{N}^*$, $\theta_1, \dots, \theta_n \in \Theta$ s.t. $A \subset \bigcup_{i=1}^n B_{\theta_i}$.

Theorem 2.1 (Bolzano–Weierstrass)**.** *Let A be a compact space and $(x_n)_{n\in\mathbb{N}} \in A^{\mathbb{N}}$, then $\exists \varphi$ s.t. $(x_{\varphi(n)})$ has a limit.*

If $x \in \mathcal{X}$, we call neighbourhood of x all opened space including x. We say that $(x_n)_n \in \mathcal{X}^{\mathbb{N}}$, $x_n \xrightarrow[n\to\infty]{} x$ if $\forall V$ neighbourhood of x, $\exists N$, $\forall n \geq N$, $x_n \in V$.

Let $f : (\mathcal{X}, \mathcal{T}) \to (\tilde{\mathcal{X}}, \tilde{\mathcal{T}})$ be continuous if $\forall \tilde{\theta} \in \tilde{\mathcal{T}}$, $f^{-1}(\tilde{\theta}) \in \mathcal{T}$.

A function $f : (\mathcal{X}, \mathcal{T}) \to \overline{\mathbb{R}}$ is lower semi-continuous (l.s.c.) if $\forall c \in \mathbb{R}$, $\{x \in \mathcal{X}, f(x) > c\} \in \mathcal{T}$.

Proposition 2.2. *If f is l.s.c. and $x_n \to x$ then,*

$$\liminf_{n\to\infty} f(x_n) \quad\geq\quad f(x).$$

Proof. Let $\varepsilon > 0$, $\{y \in \mathcal{X}, f(y) > f(x) - \varepsilon\}$ is an open set including x. So $\exists N$, $\forall n \geq N$, $f(x_n) \geq f(x) - \varepsilon$, and so $\liminf_n f(x_n) \geq f(x) - \varepsilon$. □

Consequently if f is l.s.c., F a closed set and $F \subset K$ where K is a compact space. Then $\exists x \in F$ s.t. $f(x) = \inf_{y\in F} f(y)$.

If $(f_\theta)_{\theta\in\Theta}$ is a set of l.s.c. functions, $f_\theta : \mathcal{X} \to \overline{\mathbb{R}}$, then $\sup_\theta f_\theta$ is l.s.c.

Definition 2.6 (Topological vector space)**.** $(\mathcal{X}, \mathcal{T})$ is a topological vector space if $\mathcal{X}$ is a $\mathbb{R}$-v.s. and

(i) $(\mathcal{X}, \mathcal{T})$ is a Hausdorff space ;

(ii) $(x, y) \in \mathcal{X} \times \mathcal{X} \mapsto x + y \in \mathcal{X}$ is $\mathcal{C}^0$;

(iii) $(\lambda, x) \in \mathbb{R} \times \mathcal{X} \mapsto \lambda x \in \mathcal{X}$ is $\mathcal{C}^0$.

Theorem 2.3 (Hahn–Banach)**.** *Let $(\mathcal{X}, \mathcal{T})$ be a topological v.s. locally convex. Let $K, C \in \mathcal{X}$ two convex sets s.t.*

(a) K is a compact space ;

(b) C is a closed set et $K \cap C = \varnothing$.

Then there exists $l : (\mathcal{X}, \mathcal{T}) \to \mathbb{R}$ linear and $\mathcal{C}^0$ s.t.

$$\sup_{x\in C} l(x) \quad<\quad \inf_{x\in K} l(x).$$

2.2 Fenchel–Legendre transform

Proposition 2.4. *Let $(\mathcal{X}, \mathcal{T})$ and $(\mathcal{X}', \mathcal{T}')$ be two topological v.s. locally convex and a bilinear form*

$$\begin{array}{rcl}
\mathcal{X} \times \mathcal{X} & \to & \mathbb{R} \\
(x, x') & \mapsto & \langle x, x' \rangle\ ,
\end{array}$$

s.t. $\forall x' \in \mathcal{X}'$, $x \mapsto \langle x, x' \rangle$ is linear and $\mathcal{C}^0$ and $\forall x \in \mathcal{X}$, $x' \mapsto \langle x, x' \rangle$ is linear and $\mathcal{C}^0$. And assume that $l : \mathcal{X} \to \mathbb{R}$ is linear and $\mathcal{C}^0$. Then

$$\exists x' \in \mathcal{X}' \;/\; \forall x \in \mathcal{X},\; l(x) \;=\; \langle x, x' \rangle.$$

Definition 2.7 (Convex function)**.** A function $f : \mathcal{X} \to \overline{\mathbb{R}}$ is convex if its epigraph is a convex set. With

$$\mathrm{epi} f \quad=\quad \{(x, \alpha) \in \mathcal{X} \times \mathbb{R},\; f(x) \leq \alpha\}.$$

Example 2.1. $x \mapsto x^2$ is convex.

But also f s.t.

$$f(x) \quad=\quad \begin{cases} -\infty & \text{if } x \in]a, b] \\ +\infty & \text{else}\ , \end{cases}$$

with $a < b$.

Definition 2.8 (Effective domain)**.** We call effective domain of a convex function f the set

$$\text{dom}f \quad = \quad \{x \in \mathcal{X},\ f(x) < \infty\}.$$

It's a convex set.

Proposition 2.5. *If $(f_\theta)_{\theta\in\Theta}$ is a set of convex functions, then $\sup_\theta f_\theta$ is a convex function.*

Proof. Indeed,

$$\begin{aligned} \text{epi}\sup_\theta f_\theta &= \{(x,\alpha) \in \mathcal{X}\times\mathbb{R},\ \sup_\theta f_\theta(x) \le \alpha\} \\ &= \bigcap_{\theta\in\Theta} \{(x,\alpha) \in \mathcal{X}\times\mathbb{R},\ f_\theta(x) \le \alpha\}\ , \end{aligned}$$

is convex as intersection of convex sets. □

Lemma 2.6. *If $f : \mathcal{X} \to \overline{\mathbb{R}}$ is l.s.c., epi$f$ is a closed set.*

Proof.

$$\begin{aligned} (\text{epi}f)^C &= \{(x,\alpha) \in \mathcal{X}\times\mathbb{R},\ f(x) - \alpha > 0\} \\ &= \bigcup_{c_1+c_2=0} \underbrace{\{x \in \mathcal{X},\ f(x) > c_1\}}_{\text{open}} \\ &\qquad \times \underbrace{\{\alpha \in \mathbb{R},\ -\alpha > c_2\}}_{\text{open}}. \end{aligned}$$

So $(\text{epi}f)^C$ is an open set. □

Proposition 2.7. *If $f : \mathcal{X} \to \overline{\mathbb{R}}$ is convex and l.s.c. s.t. $\exists x,\ f(x) = -\infty$, then $\forall x \in \mathcal{X},\ f(x) \in \{-\infty, +\infty\}$.*

Proof. Let $y \in \mathcal{X}$, $f(y) = -\infty$ and $y \in \text{dom}f$. If $y \in \text{dom}f$,

$$f(\lambda x + (1-\lambda)y) \quad \le \quad \lambda f(x) + (1-\lambda)f(y) \ = \ -\infty\ ,$$

for $\lambda \in [0,1[$. Now let $\lambda_n \downarrow 0$ and f is l.s.c. so

$$f(y) \quad \le \quad \liminf f(\lambda_n x + (1-\lambda_n)y) \ = \ -\infty.$$

□

Proposition 2.8. *Let $f : \mathcal{X} \to \mathbb{R}\cup\{\infty\}$ a convex function l.s.c. then*

$$f(x) \quad = \quad \sup_{\substack{x'\in\mathcal{X} \\ \alpha\in\mathbb{R} \\ \forall\tilde{x},\ \langle\tilde{x},x'\rangle - \alpha \le f(\tilde{x})}} \langle x, x'\rangle \ - \ \alpha.$$

Proof. □

Definition 2.9 (Legendre transform)**.** Let $f : \mathcal{X} \to \overline{\mathbb{R}}$. One defines its Legendre transform by

$$\begin{aligned} f^* : \mathcal{X}' &\to \overline{\mathbb{R}} \\ x' &\mapsto \sup_{x\in\mathcal{X}} \langle x, x'\rangle - f(x). \end{aligned}$$

f^* is convex l.s.c. as supremum of convex l.s.c. functions.

So,

$$\begin{aligned} f^{**}(x) &= \sup_{x'\in\mathcal{X}'} \langle x, x'\rangle - f^*(x') \\ &\le f(x). \end{aligned}$$

Theorem 2.9. *If $f : \mathcal{X} \to \mathbb{R}\cup\{\infty\}$ is convex l.s.c., then $f^{**}(x) = f(x)$.*

Example 2.2. With $\mathcal{X} = \mathbb{R}$, we have $x \mapsto \frac{x^2}{2}$ and $x \mapsto ax$.

Remark. f^{**} is the convex hull of f.

2.3 Exemples of dual spaces

2.3.1 Spaces L^p and L^q

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. And $p \in [1,\infty[$, $q \in]1,\infty]$ s.t. $\frac{1}{p} + \frac{1}{q} = 1$ or $p = 1$ and $q = \infty$.

For $p \in [1,\infty[$ we define

$$L^p \quad = \quad \{X : \Omega \to \mathbb{R} \text{ measurable } / \ \mathbb{E}[|X|^p] < \infty\}\ ,$$

and

$$L^\infty \quad = \quad \{X : \Omega \to \mathbb{R} \text{ meas. } / \ \exists M > 0,\ \mathbb{P}\{|X| \le M\} = 1\}.$$

With these definitions we set

$$\begin{aligned} \|X\|_p &= \mathbb{E}\left[|X|^p\right]^{\frac{1}{p}}\ ; \\ \|X\|_\infty &= \inf\{M > 0,\ \mathbb{P}\{|X| \le M\} = 1\}. \end{aligned}$$

$(L^p, \|\cdot\|_p)$ and $(L^\infty, \|\cdot\|_\infty)$ are Banach spaces.

Theorem 2.10. *Let $p \in [1,\infty[$ and $l : L^p \to \mathbb{R}$ is a linear $\mathcal{C}^0$ function iff $\exists Y \in L^q \ / \ \forall X \in L^p,\ l(X) = \mathbb{E}[XY]$ and Y is unique $\mathbb{P}$-a.s.*

Consequently with $p \in [1,\infty[$, $q \in]1,\infty]$ s.t. $\frac{1}{p} + \frac{1}{q} = 1$ or $p = 1$ and $q = \infty$. And with $\mathcal{X} = L^p$, $\mathcal{X}' = L^q$, $\langle X, X'\rangle = \mathbb{E}[XX']$ a bilinear $\mathcal{C}^0$ form. If $f : L^p \to \mathbb{R}$ is convex l.s.c. according to the topology of the norm $\|\cdot\|_p$ then for all $X \in L^p$,

$$\begin{aligned} f(X) &= \sup_{X'\in L^q} \mathbb{E}[XX'] - f^*(X') \\ &= f^{**}(X)\ , \end{aligned}$$

where $f^*(X) = \sup_{X\in L^p} \mathbb{E}[XX'] - f(X)$, $X' \in L^q$.

2.3.2 Duality L^∞/L^1

The Theorem 2.10 assures that L^∞ is the topological dual of L^1, *i.e.* $l : L^1 \to \mathbb{R}$ linear $\mathcal{C}^0$, $\exists Y \in L^\infty$, $\forall X \in L^p$, $l(X) = \mathbb{E}[XY]$. But L^1 is not the dual of L^∞.

But if $Y \in L^1$ and $X \in L^\infty$, $|\mathbb{E}[XY]| \le \|X\|_\infty \|Y\|_1$, then $\forall Y \in L^1$, $X \mapsto \mathbb{E}[XY]$ is $\mathcal{C}^0$ for the norm $\|\cdot\|_\infty$.

We will equip L^∞ of an other topology, the weak $*$ topology, that we note $\sigma(L^\infty, L^1)$, which is engendered by the base

$$\{Y \in L^\infty \ / \ \forall i \in [\![1,n]\!],\ |\mathbb{E}[X_i X] - \mathbb{E}[X_i Y]| < r\}.$$

With this definition $(X_p)_{p\in\mathbb{N}} \in (L^\infty)^{\mathbb{N}}$ converges weakly $*$ towards $X \in L^\infty$ if $\forall Z \in L^1$, $\mathbb{E}[X_p Z] \xrightarrow[p\to\infty]{} \mathbb{E}[XZ]$.

We admit that $(L^\infty, \sigma(L^\infty, L^1))$ is locally convex and L^1 is the dual of this space. Then $l : L^\infty \to \mathbb{R}$ is linear and $\mathcal{C}^0$ for $\sigma(L^\infty, L^1)$ iff $\exists Y \in L^1$, $\forall X \in L^\infty$, $l(X) = \mathbb{E}[XY]$.

2.3.3 Duality of measurable functions – Finite additive measures

Let $(\Omega, \mathcal{F})$ be a measurable space. And define

$$\mathcal{X} \;=\; \{F : (\Omega, \mathcal{F}) \to \mathbb{R},\ \text{measurable bounded } \forall \omega \in \Omega\}.$$

For $F \in \mathcal{X}$ we set $\|F\| := \sup_{\omega\in\Omega} |F(\omega)|$. $(\mathcal{X}, \|\cdot\|)$ is a Banach space.

Definition 2.10 (Finite additivity)**.** The application $\mu : \mathcal{F} \to \mathbb{R}$ if finite additive if

(i) $\mu(\varnothing) = 0$;

(ii) $\forall n \in \mathbb{N}^*$, $A_1, \dots, A_n \in \mathcal{F}$ disjointed, $\mu\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n \mu(A_i)$.

Definition 2.11 (Total variation)**.** The total variation of a measure (or finite additive function) is defined by:

$$\|\mu\|_{TV} \;=\; \sup\left\{\sum_{i=1}^n |\mu(A_i)|,\ n \in \mathbb{N}^*,\ A_{1:n} \in \mathcal{F} \text{ disjointed}\right\}$$

Definition 2.12 (Bounded additive)**.** We note

$$\mathrm{ba}(\Omega, \mathcal{F}) \;=\; \{\mu : \mathcal{F} \to \mathbb{R} \text{ finite additive } / \ \|\mu\|_{TV} < \infty\}\ ,$$

and

$$\mathcal{M}_{1,f}(\Omega, \mathcal{F}) \;=\; \{\mu \in \mathrm{ba}(\Omega, \mathcal{F})\ /\ \mu \geq 0,\ \mu(\Omega) = 1\}.$$

Remark. If μ is a probability measure on $(\Omega, \mathcal{F})$ then $\mu \in \mathcal{M}_{1,f}(\Omega, \mathcal{F})$.

Let $F \in \mathcal{X}$ be a simple function *i.e.*

$$F(\omega) \;=\; \sum_{i=1}^n \alpha_i \mathbb{1}_{A_i}(\omega)\ ,$$

with $A_1, \dots, A_n \in \mathcal{F}$ disjointed. We then define

$$\int F \,\mathrm{d}\mu \;=\; \sum_{i=1}^n \alpha_i \mu(A_i).$$

Theorem 2.11. *$l : \mathcal{X} \to \mathbb{R}$ linear and continuous iff $\exists \mu \in \mathrm{ba}(\Omega, \mathcal{F})$ s.t. $\forall F \in \mathcal{X}$, $l(F) = \int F \,\mathrm{d}\mu$.*

Consequently with $\mathcal{X} = \{X : (\Omega, \mathcal{F}) \to \mathbb{R} \text{ bounded measurable}\}$ and $\mathcal{X}' = \mathrm{ba}(\Omega, \mathcal{F})$. Let $X \in \mathcal{X}$, $\mu \in \mathcal{X}'$, $\langle X, \mu\rangle = \int X \,\mathrm{d}\mu$. If $f : \mathcal{X} \to \mathbb{R}$ convex l.s.c. for $\|\cdot\|$ then

$$\begin{aligned} f(X) &= f^{**}(X) \\ &= \sup_{\mu\in\mathrm{ba}(\Omega,\mathcal{F})} \int X \,\mathrm{d}\mu \ -\ f^*(\mu)\ , \end{aligned}$$

where $f^*(\mu) = \sup_{X\in\mathcal{X}} \int X \,\mathrm{d}\mu - f(X)$.

3 Risk measure representation

The objective of this section is to show that any convex risk measure $\rho : \mathcal{X} \to \mathbb{R}$ with $\mathcal{X} := \{F : (\Omega, \mathcal{F}) \to \mathbb{R} \text{ measurable bounded}\}$ can be written as:

$$\rho(X) \;=\; \sup_{Q\in\mathcal{M}_{1,f}} \mathbb{E}_Q[-X] - \alpha(Q)\ ,$$

where $\mathcal{M}_{1,f} := \{\mu \text{ finite additive } / \ \mu \geq 0,\ \mu(\Omega) = 1\}$. And we note for $\mu \in \mathcal{M}_{1,f}$, $x \in \mathcal{X}$, $\int X \mathrm{d}\mu = \mathbb{E}_\mu[X]$ and $\alpha : \mathcal{M}_{1,f} \to \mathbb{R}_+$ is a penalty function s.t. $\inf_{Q\in\mathcal{M}_{1,f}} \alpha(Q) = 0$.

We will say that a risk measure is represented by a penalty function α if:

$$\rho(X) \;=\; \sup_{Q\in\mathcal{M}_{1,f}} \mathbb{E}_Q[-X] - \alpha(Q).$$

Then we will refine this representation result by making more assumptions on ρ.

Theorem 3.1. *Let $(\Omega, \mathcal{F})$ be a measurable space, $\mathcal{X} := \{F : (\Omega, \mathcal{F}) \to \mathbb{R} \text{ measurable bounded}\}$. Then any convex risk measure $\rho : \mathcal{X} \to \mathbb{R}$ can be written as*

$$\rho(X) \;=\; \max_{Q\in\mathcal{M}_{1,f}} \mathbb{E}_Q[-X] - \alpha_{min}(Q)\ ,$$

for $X \in \mathcal{X}$, where $\alpha_{min}(Q) = \sup_{X\in\mathcal{A}_\rho} \mathbb{E}_Q[-X]$, $Q \in \mathcal{M}_{1,f}$.

Besides, α_{min} is the lowest penalty function that represents ρ, *i.e.* if $\rho(X) = \sup_{Q\in\mathcal{M}_{1,f}} \mathbb{E}_Q[-X] - \alpha(Q)$ then $\forall Q \in \mathcal{M}_{1,f}$, $\alpha(Q) \geq \alpha_{min}(Q)$.

Proof. □

When the risk measure is consistent, α_{min} has an elementary form.

Corollary. *If $\rho : \mathcal{X} \to \mathbb{R}$ is a consistent risk measure, $\forall Q \in \mathcal{M}_{1,f}$, $\alpha_{min}(Q) \in \{0, \infty\}$ and we note $\mathcal{Q}_{max} := \{Q \in \mathcal{M}_{1,f},\ \alpha_{min}(Q) = 0\}$. Then $\rho(X) = \max_{Q\in\mathcal{Q}_{max}} \mathbb{E}_Q[-X]$ and $\mathcal{Q}_{max}$ is the biggest set $\mathcal{Q}$ s.t. $\rho(X) = \sup_{Q\in\mathcal{Q}} \mathbb{E}_Q[-X]$.*

Proof. ρ is consistent so $\forall \lambda > 0$, $X \in \mathcal{X}$, $\rho(\lambda X) = \lambda\rho(X)$. And $\mathcal{X}$ is a vector space so $X \in \mathcal{X}$ iff $\lambda X \in \mathcal{X}$.

$$\begin{aligned} \alpha_{min}(Q) &= \sup_{X\in\mathcal{X}} \mathbb{E}[-\lambda X] - \rho(\lambda X) \\ &= \lambda \sup_{X\in\mathcal{X}} \mathbb{E}[-X] - \rho(X) \\ &= \lambda\alpha_{min}(Q). \end{aligned}$$

And this for all $\lambda > 0$, so $\alpha_{min}(Q) \in \{0, \infty\}$. □

3.1 Convex risk measures in $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. We are looking at the risk measures $\rho : \mathcal{X} \to \mathbb{R}$ s.t. $\rho(X) = \rho(Y)$ if $X = Y$ a.s.

Hence we can see ρ as a function of L^∞ to $\mathbb{R}$. We note

$$\begin{aligned} \mathcal{M}_{1,f}(\mathbb{P}) &:= \{Q \in \mathcal{M}_{1,f},\ Q \ll \mathbb{P}\} \ ; \\ \mathcal{M}_1(\mathbb{P}) &:= \{\mathbb{Q} \in \mathcal{M}_1,\ \mathbb{Q} \ll \mathbb{P}\}. \end{aligned}$$

Where $Q \ll \mathbb{P}$ if $\forall A \in \mathcal{F}$, $\mathbb{P}(A) = 0 \Rightarrow Q(A) = 0$.

Lemma 3.2. *If ρ is a convex risk measure represented by $\alpha : \mathcal{M}_{1,f} \to [0, \infty]$ and s.t. $\rho(X) = \rho(Y)$ if $X = Y$ a.s. Then $\alpha(Q) = \infty$ for $Q \in \mathcal{M}_{1,f} \setminus \mathcal{M}_{1,f}(\mathbb{P})$*

Proof. □

If we want to have a risk measure from a penalty function α s.t. $\alpha(Q) = \infty$ if $Q \in \mathcal{M}_{1,f} \setminus \mathcal{M}_1$ we have have more hypothesis on ρ ; typically properties on the convergence.

We will now work with $\mathcal{X} = L^\infty(\Omega, \mathcal{F}, \mathbb{P})$.

Theorem 3.3. *Let $\rho : L^\infty \to \mathbb{R}$ be a convex risk measure. Define $\alpha_{min}(Q) = \sup_{X \in \mathcal{A}_\rho} \mathbb{E}_Q[-X]$ and $Q \in \mathcal{M}_1(\mathbb{P})$. Then the following conditions are equivalent*

(i) ρ *is l.s.c. according to the weak∗ topology* $\sigma(L^\infty, L^1)$;

(ii) $\rho(X) = \sup_{Q \in \mathcal{M}_1(\mathbb{P})} \mathbb{E}_Q[-X] - \alpha_{min}(Q)$;

(iii) ρ *is* $\mathcal{C}^0$ *from above, i.e. if* $X_n \xrightarrow{a.s.} X$ *and* $\mathbb{P}\{X_n \leq X_{n+1}\} = 1$ *then* $\rho(X_n) \xrightarrow[n\to\infty]{} \rho(X)$;

(iv) ρ *satisfies the Fatou property, i.e. if* $\forall (X_n)_n \in (L^\infty)^{\mathbb{N}}$, $X_n \xrightarrow{a.s.} X$, $\exists M$, $\|X\|_\infty \leq M$, *then* $\liminf_n \rho(x_n) \geq \rho(X)$.

In this case α_{min} is the lowest penalty function $\alpha : \mathcal{M}_1(\mathbb{P}) \to [0, \infty]$ s.t. $\rho(X) = \sup_{Q \in \mathcal{M}_1(\mathbb{P})} \mathbb{E}_Q[-X] - \alpha(Q)$.

Proof. □

3.2 Invariant distribution risk measures

We now focus on risk measures s.t. $\rho(X) = \rho(Y)$ if X and Y have the same distribution under $\mathbb{P}$. This hypothesis is reasonable as we give the same risk to two portfolios that have the same distribution.

It is clear that VaR_λ and ES_λ are invariant according to the distribution. We will show that the representation of convex risk measures that are invariant according to the distribution is made of an elementary piece that is ES_λ.

In this paragraph we will need the technical hypothesis: $\exists U \in L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ s.t. $U \sim \mathcal{U}([0, 1])$ under $\mathbb{P}$.

Lemma 3.4. *Let $X \in L^\infty$, $Y \in L^1$. We define $q_X(\lambda) = q_X^+(\lambda)$ and $q_Y(\lambda) = q_Y^+(\lambda)$. Then*

$$\begin{aligned} \sup_{\tilde{X} \sim X} \mathbb{E}[\tilde{X} Y] &= \sup_{\tilde{Y} \sim Y} \mathbb{E}[X \tilde{Y}] \\ &= \int_0^1 q_X(\lambda) q_Y(\lambda)\ \mathrm{d}\lambda. \end{aligned}$$

Remark. If we take $\tilde{X} = q_X(U)$, $\tilde{Y} = q_Y(U)$ and $U \sim \mathcal{U}([0, 1])$, then

$$\mathbb{E}[\tilde{X}\tilde{Y}] = \int_0^1 q_X(\lambda) q_Y(\lambda)\ \mathrm{d}\lambda.$$

Theorem 3.5. *Any distribution invariant risk measure satisfies the Fatou property.*

Theorem 3.6. *Let $\rho : L^\infty \to \mathbb{R}$ be a convex risk measure and that is invariant according to the distribution. Then,*

$$\rho(X) = \sup_{Q \in \mathcal{M}_1(\mathbb{P})} \int_0^1 q_X(\lambda) q_{\frac{\mathrm{d}Q}{\mathrm{d}\mathbb{P}}}(\lambda)\ \mathrm{d}\lambda \ - \ \alpha_{min}(Q) \ ,$$

with

$$\begin{aligned} \alpha_{min}(Q) &= \sup_{X \in \mathcal{A}_\rho} \int_0^1 q_{-X}(\lambda) q_{\frac{\mathrm{d}Q}{\mathrm{d}\mathbb{P}}}(\lambda)\ \mathrm{d}\lambda \\ &= \sup_{X \in \mathcal{X}} \int_0^1 q_{-X}(\lambda) q_{\frac{\mathrm{d}Q}{\mathrm{d}\mathbb{P}}}(\lambda)\ \mathrm{d}\lambda \ - \ \rho(X). \end{aligned}$$

Proof. □

Corollary. *If ρ is a convex risk measure and that is invariant according to the distribution,*

$$\rho(X) = \sup_{\mu \in \mathcal{M}_1(]0,1])} \int_{]0,1]} \mathrm{ES}_\lambda(X) \mu(\mathrm{d}\lambda) - \beta_{min}(\mu) \ ,$$

where $\beta_{min}(\mu) = \sup_{X \in \mathcal{A}_\rho} \int_{]0,1]} \mathrm{ES}_\lambda(X) \mu(\mathrm{d}\lambda)$.

In particular if ρ is consistent, $\beta_{min} = \{0, \infty\}$,

$$\rho(X) = \sup_{\substack{\mu \in \mathcal{M}_1(]0,1]) \\ \mu \ / \ \beta_{min}(\mu) < \infty}} \int_{]0,1]} \mathrm{ES}_\lambda(X) \mu(\mathrm{d}\lambda)$$

Example 3.1. We have the following:

- Let
$$\beta(\mu) = \begin{cases} 0 & \text{if} \quad \mu = \delta_\lambda \\ \infty & \text{else} \end{cases}$$
Then $\rho(X) = \mathrm{ES}_\lambda(X)$.

- Let
$$\beta(\mu) = \begin{cases} 0 & \text{if} \quad \mu = \sum_{i=1}^n p_i \delta_{\lambda_i} \\ \infty & \text{else} \end{cases}$$
where $\sum_i p_i = 1$, $p_i > 0$, $\lambda_i \in]0, 1]$. Then $\rho(X) = \sum_{i=1}^n p_i \mathrm{ES}_{\lambda_i}(X)$.

- Let
$$\beta(\mu) = \begin{cases} \beta_i & \text{if} \quad \mu = \mu_i,\ i \in [\![1, n]\!] \\ \infty & \text{else} \end{cases}$$
where $\mu_i = \sum_{j=1}^n q_{ij} \delta_{\lambda_{ij}}$, $\sum q_{ij} = 1$, $q_{ij} \geq 0$, $\lambda_{ij} \in]0, 1]$ and $\min_i \beta_i = 0$. Then
$$\rho(X) = \max_{i \in [\![1,n]\!]} \sum_{j=1}^n q_{ij} \mathrm{ES}_{\lambda_{ij}}(x) \ - \ \beta_i.$$

Part II

Notations We consider $X_1, \dots, X_n$ i.i.d. r.v. of density μ. In this course X represents a loss. The question is how can one estimate a quantile q_α with $\alpha \in]0,1[$?

Let F be the cdf associated to μ:

$$\begin{aligned} q_\alpha &= F^{-1}(\alpha) \\ &= \inf\{x : F(x) \geq \alpha\}. \end{aligned}$$

In the rest of the course we note $\bar{F} = 1 - F$.

Non-parametric approach Let us replace F with $F_n := \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\{X_i \leq x\}}$. The problem that we could meet is when we don't have a lot of observations, because $\alpha \approx 1$, so it is possible that $M_n := \max_i X_i < q_\alpha$.

So this method theoretically robust suffers from the sample size.

Parametric approach We suppose that we know μ but it depends on unknown parameters. Indeed, $\forall x \in \mathbb{R}$,

$$\mathbb{P}\{M_n \leq x\} = F(x)^n.$$

Let us define

$$x_\mu := \sup\{x : F(x) < 1\} \leq \infty$$

so for all $x \in]-\infty, x_\mu[$, $\mathbb{P}\{M_n \leq x\} = F(x)^n \xrightarrow[n\to\infty]{} 0$.

If $x_\mu < \infty$, for all $x \in [x_\mu, \infty[$, $\mathbb{P}\{M_n \leq x\} = 1$. So for all $\varepsilon > 0$,

$$\begin{aligned} \mathbb{P}\{M_n \leq x_\mu - \varepsilon\} &\xrightarrow[n\to\infty]{} 0 \\ \mathbb{P}\{M_n > x_\mu + \varepsilon\} = 1 - F(x_\mu + \varepsilon)^n &= 0 \end{aligned}$$

and so $M_n \xrightarrow{\mathbb{P}} x_\mu$. And as $M_n \uparrow M_\infty$, we have a.s. $M_\infty = x_\mu$ a.s. and so $M_n \xrightarrow{a.s.} x_\mu$.

It is then obvious to see the speed of convergence. Putting aside the fact that $x_\mu = \infty$ this is equivalent to look for to sequences $(c_n)_n$, $(d_n)_n$ with $c_n > 0$ s.t.

$$c_n(M_n - d_n) \xrightarrow{\mathcal{L}} H\,, \tag{3.1}$$

where H is characterised by it cdf. We have (3.1) which is equivalent to

$$\begin{aligned} \mathbb{P}\{c_n(M_n - d_n) \leq x\} &= F^n(c_n^{-1}x + d_n) \\ &\xrightarrow[n\to\infty]{} G(x). \end{aligned}$$

This limit is a punctual limit and is expressed thanks to the convergence on the continuous points of G. We will then say that $F \in D(G)$ which is the attraction domain of G.

4 Extreme value theory

4.1 Extreme values categories

The goal of this section is to prove the following result

Theorem 4.1 (Extreme value)**.** *If there exists two sequences $c_n > 0$ and d_n verifying (3.1) with G non degenerated (i.e. $\neq \mathbb{1}_{\{X \geq x_0\}}$) then G is in one of the following extreme value categories: $G \in \{\phi_\alpha, \phi_\alpha, \Lambda\}$. With*

- *Frechet: $\phi_\alpha(x) = \exp(-x^{-\alpha})\mathbb{1}_{x>0}$ $(\alpha > 0)$.*
- *Weibull: $\psi_\alpha(x) = \exp(-(-x)^{\alpha})\mathbb{1}_{x\leq 0} + \mathbb{1}_{x>0}$ $(\alpha > 0)$.*
- *Gumbel: $\Lambda(x) = \exp(-e^{-x})$.*

Inversely every category can be obtained as limit of (3.1).

The next three figures are showing the density function of the categories.

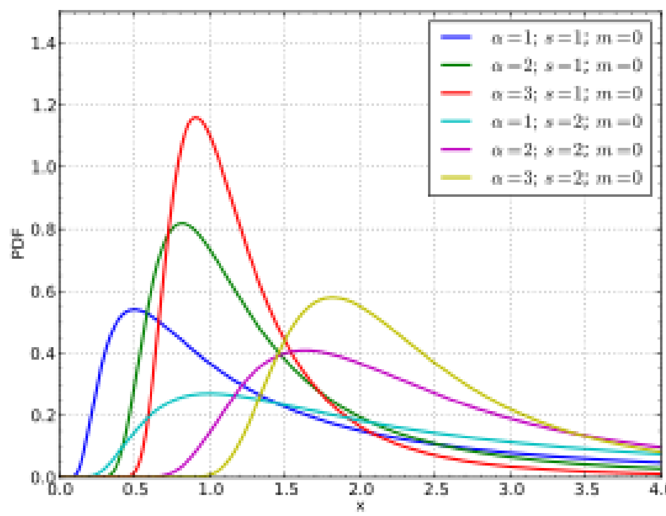

Figure 1: Frechet

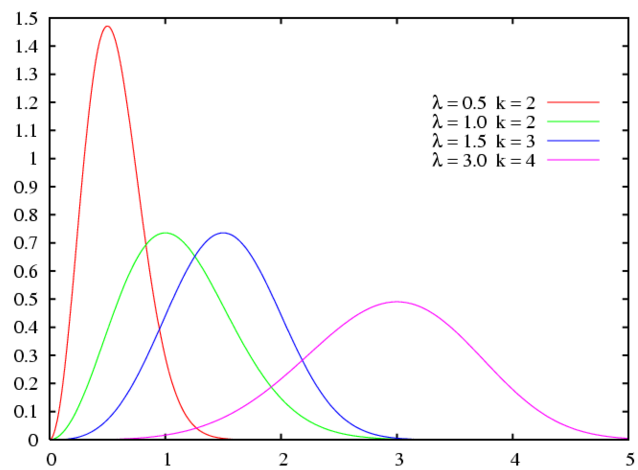

Figure 2: Weibull

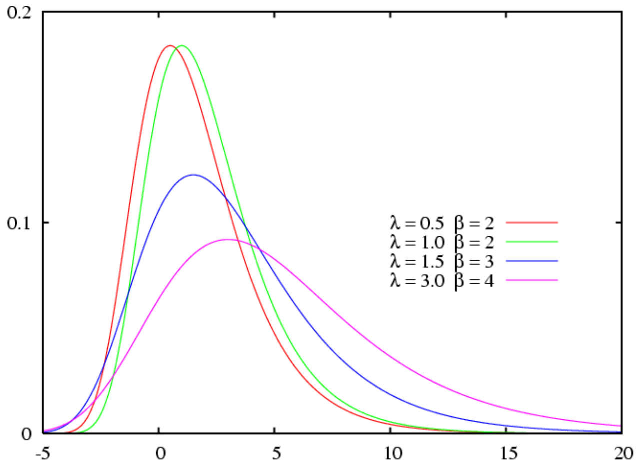

Figure 3: Gumbel

To prove this result we first need some results on the generalized inverse of growing right continuous functions.

Proposition 4.2. *Let f be a growing right continuous function and $f^{-1}(y) = \inf\{x : f(x) \geq y\}$.*

(i) Let $a > 0$, b and c constants and $g(x) := f(ax+b) - c$, then $g^{-1}(y) = a^{-1}f^{-1}(y+c) - b$.

(ii) If f^{-1} is continuous, then $f^{-1}(f(x)) = x$.

(iii) If f is a non degenerated cdf and $a, \alpha > 0$, b and β s.t. $f(ax+b) = f(\alpha x + \beta)$ for all x then $a = \alpha$ and $b = \beta$.

Proof. □

From (3.1) it is natural to restrain the specification of G to the set of cdf max stable define as follow.

Definition 4.1 (Max stable cdf)**.** A cdf is max stable if it is non degenerated and for all $k \in \mathbb{N}^*$, $\exists b_k, a_k > 0$ s.t.

$$f^k(a_k x + b_k) = f(x).$$

This definition is justified by the point (iii) of the following theorem.

Theorem 4.3. *(i) A cdf f non degenerated is max stable iff $\exists (f_n)_n$ a sequence of cdf and two sequences b_n and $a_n > 0$ s.t.*

$$f(a_{nk}^{-1}x + b_{nk}) \xrightarrow[n\to\infty]{} f^{\frac{1}{k}}(x).$$

(ii) If f is max stable, there exists real functions $b(s)$ and $a(s) > 0$ defined on $]0,\infty[$ s.t.

$$f(x) = f^s(a(s)x + b(s)),$$

for all $x \in \mathbb{R}$, $s > 0$.

(iii) If G is a non degenerated cdf then $D(G) \neq \varnothing$ iff G is max stable. In this case $G \in D(G)$.

Lemma 4.4. *Let f be a non degenerated cdf and f_n a sequence of cdf. Define the sequences b_n and $a_n > 0$ s.t. $f_n(a_n x + b_n) \to f(x)$. Then $\exists \tilde{f}$ a non degenerated cdf and two sequences $\tilde{a}_n > 0$, $\tilde{b}_n$ s.t. $f_n(\tilde{a}_n x + \tilde{b}_n) \to \tilde{f}(x)$ iff*

$$\frac{\tilde{a}_n}{a_n} \xrightarrow[n\to\infty]{} a \quad \text{and}$$
$$\frac{\tilde{b}_n - b_n}{a_n} \xrightarrow[n\to\infty]{} b$$

for a given couple (a,b) with $a > 0$ s.t. $\tilde{f}(x) = f(ax+b)$.

Proof. (Of the Theorem 4.3). □

We will now formalise the notion of category as well as some properties

Definition 4.2. Two cdf G_1 and G_2 are in the same category if for some constants b and $a > 0$, $G_2(x) = G_1(ax+b)$.

This definition implies the next few properties.

Proposition 4.5. *(i) A cdf G is max stable if for all $n \in \mathbb{N}^*$, G^n and G are in the same category.*

(ii) If the set of cdf F_n verifies $F_n(a_n x + b_n) \to G_1(x)$ and $F_n(\alpha_n x + b_n) \to G_2(x)$ with $a_n, \alpha_n > 0$ and if G_1, G_2 are non degenerated, then they are in the same category.

Proof. (Of the Theorem 4.1). □

Example 4.1. Let $X_1, \ldots, X_n \sim \mathcal{E}(\theta)$ i.i.d. r.v. with $\theta > 0$. So for all $x \geq 0$,

$$\begin{aligned}
\mathbb{P}\left\{M_n - \frac{\ln n}{\theta} \leq x\right\} &= \mathbb{P}\left\{X_1 - \frac{\ln n}{\theta} \leq x\right\}^n \\
&= \left(1 - e^{-\theta\left(x + \frac{\ln n}{\theta}\right)}\right)^n \\
&= \left(1 - \frac{e^{-\theta x}}{n}\right) \\
&\xrightarrow[n\to\infty]{} \Lambda(\theta x).
\end{aligned}$$

4.2 Attraction domain

We know that in the additive case, most[2] of the densities have: $\exists c_n > 0$, $d_n \in \mathbb{R}$ s.t. $c_n(X_1 + \cdots + X_n - d_n) \xrightarrow[n\to\infty]{\mathcal{L}} Z$ (CLT). Hence when $X \in L^2$, we have $c_n = (\sigma_X\sqrt{n})^{-1}$ and $d_n = nm_X$ with $Z \sim \mathcal{N}(0,1)$.

For the maximum, numerous classical densities don't verify the attraction principle (3.1). Indeed we can take $X \sim \mathcal{P}(\theta)$ ([3]) for instance.

Theorem 4.6. *The exists a sequence (u_n) and $\tau \in]0,\infty[$ s.t. $n\bar{F}(u_n) \xrightarrow[n\to\infty]{} \tau$ iff*

$$\frac{\bar{F}(x)}{\bar{F}(x^-)} \xrightarrow[x\to x_\mu]{} 1.$$

With $\bar{F} = 1 - F$ the survival function.

[2] Those where the second order moment exists.
[3] This is the Poisson density.

In the case of integer r.v., $\mu(\mathbb{N}) = 1$ and $x_\mu = \infty$, hence

$$\begin{aligned} & \frac{\bar{F}(n)}{\bar{F}(n-1)} \xrightarrow[x\to\infty]{} 1 \\ \Leftrightarrow \quad & \frac{\bar{F}(n) - \bar{F}(n-1)}{\bar{F}(n-1)} \xrightarrow[x\to\infty]{} 0. \end{aligned}$$

Lemma 4.7. *Let $\tau \in \bar{\mathbb{R}}_+$ and (u_n) a sequence of real numbers, then*

$$\begin{aligned} & n\bar{F}(u_n) \xrightarrow[n\to\infty]{} \tau \\ \Leftrightarrow \quad & \mathbb{P}\{M_n \le u_n\} \xrightarrow[n\to\infty]{} e^{-\tau}. \end{aligned}$$

Proof. □

We define $u_n := \frac{x}{a_n} + b_n$, $a_n > 0$, then

$$\mathbb{P}\{M_n \le u_n\} = F^n\left(\frac{x}{a_n} + b_n\right).$$

So if $\lim_{x\to x_\mu} \frac{\bar{F}(x)}{\bar{F}(x^-)} \neq 1$ then $\lim_{n\to\infty} \mathbb{P}\{M_n \le u_n\} \neq e^{-\tau} \in]0,1[$.

In the case of the Poisson density, $X \sim \mathcal{P}(\theta)$, $\mathbb{P}\{X = n\} = e^{-\theta}\frac{\theta^n}{n!}$,

$$\begin{aligned} \frac{\bar{F}(n) - \bar{F}(n-1)}{\bar{F}(n-1)} &= -\frac{\theta^n}{n!}e^{-\theta}\frac{1}{e^{-\theta}\sum_{k\ge n}\frac{\theta^k}{k!}} \\ &= \frac{-1}{1 + S_n^\theta}, \end{aligned}$$

with

$$\begin{aligned} S_n^\theta &:= \sum_{k\ge n+1} \frac{\theta^{k-n}}{k\dots(n+1)} \\ &= \sum_{k\ge 1} \frac{\theta^k}{(n+k)\dots(n+1)} \\ &\le \sum_{k\ge 1}\left(\frac{\theta}{n}\right)^k = \frac{\frac{\theta}{n}}{1-\frac{\theta}{n}} \to 0. \end{aligned}$$

Hence $\frac{\bar{F}(n)-\bar{F}(n-1)}{\bar{F}(n-1)} \xrightarrow[n\to\infty]{} -1 \neq 0$. Then the Poisson density doesn't verify (3.1).

Corollary. *Let $F \in D(G)$ with the coefficients $c_n > 0$ and d_n for (3.1) iff*

$$n\bar{F}\left(\frac{x}{c_n} + d_n\right) \xrightarrow[n\to\infty]{} -\ln G(x),$$

for all $x \in \mathbb{R}$.

4.2.1 Attraction domain of the Frechet density

We recall that

$$\begin{aligned} 1 - \phi_\alpha(x) &= 1 - \exp\left(-x^{-\alpha}\right)\mathbb{1}_{\{x>0\}} \\ &\sim x^{-\alpha} \qquad \text{when } x \to \infty. \end{aligned}$$

So densities in its attraction domain can only be densities which end queue are "near" a power density, in a sense to determine.

Definition 4.3 (Slow variation)**.** A function L is said with slow variations if for x big enough, $L(x) > 0$ and $\forall t > 0$,

$$\frac{L(tx)}{L(x)} \xrightarrow[x\to\infty]{} 1.$$

Example 4.2. $L(x) = C$, $L(x) = \ln x$.

Theorem 4.8. *We have $F \in D(\phi_\alpha)$, $\alpha > 0$ iff $L(x) = x^\alpha \bar{F}(x)$ is with slow variations.*

Moreover, if $F \in D(\phi_\alpha)$ then $c_n = \left(F^{-1}(1-\frac{1}{n})\right)^{-1}$ and $d_n = 0$ are coefficients for (3.1).

Remark. With this theorem, $F \in D(\phi_\alpha) \Rightarrow x_\mu = \infty$.

Example 4.3. The Pareto density, $\bar{F}(x) = \left(\frac{x}{\theta}\right)^{-\alpha}$ for $x \ge \theta$, then $\bar{F}(x) \sim Kx^{-\alpha}$ when $x \to \infty$, then we can take $c_n = (Kn)^{-\frac{1}{\alpha}}$.

4.2.2 Attraction domain of the Weibull density

We recall

$$\psi_\alpha(x) = e^{-(-x)^\alpha}\mathbb{1}_{\{x<0\}} + \mathbb{1}_{\{x\ge 0\}}.$$

Theorem 4.9. *$F \in D(\psi_\alpha)$ iff $x_\mu < \infty$ and $\bar{F}\left(x_\mu - \frac{1}{x}\right) = x^{-\alpha}L(x)$, where L is with slow variations.*

Moreover if $F \in D(\psi_\alpha)$ then $c_n = \left(x_\mu - F^{-1}(1-\frac{1}{n})\right)^{-1}$ and $d_n = x_\mu$ are coefficients for (3.1).

Example 4.4. $\mathcal{U}([0,1]) \Rightarrow x_\mu = 1$. $\bar{F}(1-x^{-1}) = x^{-1} \Rightarrow F \in D(\psi_1)$ and $c_n = n$.

4.2.3 Attraction domain of the Gumbel density

We recall

$$\Lambda(x) = \exp(-e^{-x}).$$

Theorem 4.10. *$F \in D(\Lambda)$ iff $\exists x_0 \in]-\infty, x_\mu[$ s.t. $\forall x \in]x_0, x_\mu[$,*

$$\bar{F}(x) = \rho(x)\exp\left(\int_{x_0}^x \frac{g(t)}{a(t)}\,\mathrm{d}t\right),$$

where ρ, g are s.t. $\rho(x) \xrightarrow[x\to x_\mu]{} \rho > 0$, $g(x) \xrightarrow[x\to x_\mu]{} 1$, $a : \mathbb{R} \to \mathbb{R}_+$ is absolutely continuous with $a'(x) \xrightarrow[x\to x_\mu]{} 0$.

In this case we can choose $d_n = F^{-1}(1-\frac{1}{n})$ and $c_n = \frac{1}{a(d_n)}$.

The reprentation of the theorem 4.10 is not unique, in the sens that we have to make a choice for ρ, g.

Example 4.5 (Normal density)**.**

4.2.4 Summary

Definition 4.4 (Generalized formulation)**.** We call generalized formulation the following expression of max stable densities

$$H_\xi \quad = \quad \begin{cases} \exp\left(-(1+x\xi)^{-\frac{1}{\xi}}\right) & \text{if } \xi \neq 0 \text{ with } 1+\xi x > 0 \\ \exp(-e^{-x}) & \text{else.} \end{cases}$$

We can then summarise the previous theorems with this one:

Theorem 4.11. *Let $U(t) = F^{-1}(1-\frac{1}{t})$, $t>0$ et let $\xi \in \mathbb{R}$ fixed. The three following assertions are equivalent:*

(i) $F \in D(H_\xi)$

(ii) $\exists a$ *s.t.*

$$\frac{\bar{F}(u+xa(u))}{\bar{F}(u)} \quad \xrightarrow[x\to x_\mu]{} \quad \begin{cases} (1+\xi x)^{-\frac{1}{\xi}} & \text{if } \xi \neq 0 \\ e^{-x} & \text{if } \xi = 0. \end{cases}$$

(iii) $\forall x, y > 0,\ y \neq 1$,

$$\frac{U(sx)-U(s)}{U(sy)-U(s)} \quad \xrightarrow[s\to\infty]{} \quad \begin{cases} \frac{x^\xi - 1}{y^\xi - 1} & \text{if } \xi \neq 0 \\ \frac{\ln x}{\ln y} & \text{if } \xi = 0. \end{cases}$$

5 Application to quantile calculation

The empiric quantile is very linked to the order statistic of the observation $(X_1,\dots,X_n)$. To simplify we assume that F is continuous. Hence μ has no atom and $\mathbb{P}\{X_i = X_j\} = 0$ for $i \neq j$. Then $\left(X_k^{(n)}\right)_{1\le k\le n}$ is well defined and verifies

(i) $\min_k X_k = X_1^{(n)} < \cdots < X_n^{(n)} = \max_k X_k$;

(ii) $\{X_k^{(n)},\ k \in [\![1,n]\!]\} = \{X_k,\ k \in [\![1,n]\!]\}$.

Then $F_n^{-1}(\alpha,\omega) = X_{k_n(\alpha)}^{(n)}$ where $\frac{k_n(\alpha)-1}{n} < \alpha \le \frac{k_n(\alpha}{n}$. The problem is when n is small and $\alpha \approx 1$, there is a possibility that $X_n^{(n)} < q_\alpha$.

A solution is to use a parameter: $F \in D(H_\xi)$ and with the corollary,

$$n\bar{F}\left(\frac{x}{c_n}+d_n\right) \quad \xrightarrow[n\to\infty]{} \quad -\ln H_\xi(x).$$

So for $u := \frac{x}{c_n} + d_n$ and n big enough,

$$\bar{F}(u) \quad \approx \quad \frac{1}{n}(1+\xi c_n(n-d_n))^{-\frac{1}{\xi}},$$

so with $u = q_\alpha$ and by approximating $\hat{\xi}$ (as well as $\hat{c}_n$ and $\hat{d}_n$) we have

$$\hat{q}_\alpha \quad = \quad \hat{d}_n \;+\; \frac{1}{\hat{\xi}\hat{c}_n}\left((n(1-\alpha))^{-\hat{\xi}} - 1\right).$$

So if we can estimate $\hat{\xi}$ we have a chance to estimate the quantile.

5.1 Pickands estimator

5.1.1 Founding principle

Theorem 5.1. *Let $F \in D(H_\xi)$, $\xi \in \mathbb{R}$ and $(k_n)_n$ a sequence of integer numbers s.t. $\frac{k_n}{n} \xrightarrow[n\to\infty]{} 0$ and $k_n \xrightarrow[n\to\infty]{} \infty$. Then*

$$\hat{\xi}_{k_n} \quad = \quad \frac{1}{\ln 2}\left(\frac{X_{n-k_n+1}^{(n)} - X_{n-2k_n+1}^{(n)}}{X_{n-2k_n+1}^{(n)} - X_{n-4k_n+1}^{(n)}}\right) \quad \xrightarrow[n\to\infty]{\mathbb{P}} \quad \xi.$$

Lemma 5.2. *If $U_1,\dots,U_n$ iid $\sim \mathcal{U}([0,1])$ then*

$$\left(U_k^{(n)}\right) \quad \sim \quad \frac{\Gamma_k}{\Gamma_{n+1}},$$

where $\Gamma_k = \sum_{i=1}^k E_i$, E_i iid $\sim \mathcal{E}(1)$.

Proof. □

Lemma 5.3. *Let $k_n \in [\![1,n]\!]$, $k_n \xrightarrow[n\to\infty]{\infty}$. Let $(V_n)_{n\ge1}$ sequence of r.v. iid $\sim$ Pareto(1), then*

$$\frac{k_n}{n}V_{n+1-k_n}^{(n)} \quad \xrightarrow{\mathbb{P}} \quad 1.$$

Proof. □

Proof. (of the Theorem). □

5.1.2 Heuristic method

We have by definition $U(x) = F^{-1}(1-\frac{1}{n})$, it is clear that

$$q_\alpha \quad = \quad U\left(\frac{1}{1-\alpha}\right).$$

With theorem 4.11 we know that

$$U(sx) \quad \approx \quad U(s) \;+\; \frac{x^\xi - 1}{y^\xi - 1}(U(sy) - U(s) \quad \text{when } s \to \infty$$

with the convention, $\xi = 0 \Rightarrow \frac{x^\xi-1}{y^\xi-1} = \frac{\ln x}{\ln y}$.

For a given k, the idea is then to set $s := \frac{n}{k-1}$,

$$\begin{aligned} x \quad &= \quad \frac{1}{s(1-\alpha} \\ &= \quad \frac{k-1}{n(1-\alpha)} \\ &\approx \quad \frac{k}{n(1-\alpha)} \end{aligned}$$

and $y = \frac{1}{2}$. Then we have $U(sx) = U(\frac{1}{1-\alpha}) = q_\alpha$, and then

$$\begin{aligned} q_\alpha \quad \approx \quad & U\left(\frac{n}{k-1}\right) \\ & + \frac{\left(\frac{k}{n(1-\alpha)}\right)^\xi - 1}{2^{-\xi}-1}\left(U\left(\frac{n}{2(k-1)}\right) - U\left(\frac{n}{k-1}\right)\right). \end{aligned}$$

We are still in the heuristic method, we can then replace U with its estimator

$$\begin{aligned} U_n(y) &= F_n^{-1}\left(1-\frac{1}{y}\right) \\ F_n(x) &= \frac{1}{n}\sum_{i=1}^{n}\mathbb{1}_{\{X_k\le x\}}. \end{aligned}$$

And as F_n is supposed continuous, we have $F_n\left(X_{k-1}^{(n)}\right) = \frac{k-1}{n}$. So $X_{k-1}^{(n)} = F_n^{-1}\left(\frac{k-1}{n}\right)$ and

$$\begin{aligned} U_n(s) &= U_n\left(\frac{n}{k-1}\right) \\ &= F_n^{-1}\left(\frac{n-k+1}{n}\right) \\ &= X_{n+1-k}^{(n)}. \end{aligned}$$

This conducts to the estimator

$$\hat{q}_\alpha = \frac{\left(\frac{k}{n(1-\alpha)}\right)^{\hat{\xi}}-1}{2^{-\hat{\xi}}-1}\left(X_{n+2-2k}^{(n)}-X_{n+1-k}^{(n)}\right)+X_{n+1-k}^{(n)}$$

5.2 Peaks over threshold

Here the approach is different from the Hill and Pickands estimators.

5.2.1 Theoretical principle

Let u be a fixed threshold, and define

$$N_u := \mathrm{Card}\{i\in[\![1,n]\!],\ X_i>u\}$$

that represents the number over the threshold. And the cdf of $Y_1,\dots,Y_{N_u}$ is given by, $\forall y\ge 0$,

$$\begin{aligned} F_u(y) &= \mathbb{P}\{Y\le y|X>u\} \\ &= \mathbb{P}\{X-u\le y|X>u\}. \end{aligned}$$

We can adopt a constructive approach, let

$$\begin{aligned} \tau_k &:= \inf\{t\in[\![\tau_{k-1}+1,n]\!],\ X_t>u\}\ , \quad \tau_0=0 \\ Y_k &= X_{\tau_k}\ -\ u. \end{aligned}$$

And we can see that $F_u(y)\bar{F}(u) = \mathbb{P}\{X\le u+y \text{ and } X>u\}$, hence

$$\begin{aligned} F_u(y)\bar{F}(y) &= \mathbb{P}\{u<X\le u+y\} \\ &= F(y+u)\ -\ F(u) \\ &= \bar{F}(u)\ -\ \bar{F}(y+u)\,, \end{aligned}$$

so finally the relation, $\forall y\ge 0$, $\forall u\in\mathbb{R}$,

$$\bar{F}(u+y) = \bar{F}(u)\bar{F}_u(y). \tag{5.1}$$

Definition 5.1 (Generalized Pareto)**.** We define the generalized Pareto density $G_{\xi,\beta}$, $\xi\in\mathbb{R}$, $\beta>0$, the density charaterized by the survival function

$$\bar{G}_{\xi,\beta} = \left(1+\xi\frac{x}{\beta}\right)^{-\frac{1}{\xi}}\mathbb{1}_{\{\xi\neq 0\}}\ +\ e^{-\frac{x}{\beta}}\mathbb{1}_{\{\xi=0\}}\,,$$

with

$$x\in D_{\xi,\beta} = \begin{cases} \mathbb{R}_+ & \text{if } \xi\ge 0\ ; \\ \left[0,-\frac{\beta}{\xi}\right] & \text{if } \xi<0. \end{cases}$$

Theorem 5.4. *There exists a function $\beta:\mathbb{R}\to\mathbb{R}_+^*$ s.t.*

$$\lim_{\substack{\mu\to x_\mu \\ <}}\ \sup_{y\in[0,x_\mu-\mu]}\left|\bar{F}_u(y)-\bar{G}_{\xi,\beta(u)}(y)\right| = 0$$

iff $F\in D(H_\xi)$, $\xi\in\mathbb{R}$.

Remark. If $x_\mu=\infty$ the convergence is uniform.

To use the result of this last theorem we need the following proposition.

Proposition 5.5. *Let the observations $Y_1,\dots,Y_n$ i.i.d. $\sim G_{\xi,\beta}$. Then for all $u\in D_{\xi,\beta}$,*

$$\begin{aligned} e(u) &:= \mathbb{E}[Y-u|Y>u] \\ &= \frac{\beta-\xi u}{1-\xi} \qquad \text{for } \xi<1\,, \end{aligned}$$

and the log likelihood

$$\begin{aligned} l\left((\xi,\beta),(Y_1,\dots,Y_n)\right) &= -n\ln\beta \\ &\quad -\left(\frac{1}{\xi}+1\right)\sum_{i=1}^{n}\ln\left(\xi\frac{Y_i}{\beta}+1\right). \end{aligned}$$

Proof. □

5.2.2 Heuristic

When $x_\mu=\infty$ and u big enough, we can approximate $\bar{F}_u\approx\bar{G}_{\hat{\xi},\hat{\beta}(u)}$ where $\hat{\xi}$ and $\hat{\beta}(u)$ are estimators of ξ and β. We can also use the estimator for $\bar{F}_n(u) \approx \frac{1}{n}\sum_{k=1}^{n}\mathbb{1}_{\{X_k>u\}} = \frac{N_u}{n}$. Hence (5.1) implies

$$\bar{F}(u+y) \approx \frac{N_u}{n}\left(1+\hat{\xi}\frac{y}{\hat{\beta}}\right)^{-\frac{1}{\hat{\xi}}}.$$

The by setting $y=\bar{q}_\alpha-u$, we have the estimator

$$\hat{q}_\alpha = u\ +\ \frac{\hat{\beta}}{\hat{\xi}}\left(\left(\frac{n}{N_u}(1-\alpha)\right)^{-\hat{\xi}}-1\right).$$

How to choose u ? We introduce the empirical analog of $e(u)$:

$$e_n(u) = \frac{1}{N_u}\sum_{k=1}^{n}(X_k-u)\mathbb{1}_{\{X_k>u\}}.$$

Then we choose u so that $e_n(u)$ is more or less affine when $x\ge u$.

How to choose $\hat{\beta}$? $\hat{\xi}$? We differentiate the log likelihood and use a numerical method of resolution, *e.g.* Newton–Raphson.

6 CVA (Credit Valuation Adjustment) and extensions

6.1 Real contract

6.2 Construction of a replication portfolio

6.3 XVA and predefault BSDE

Part III

7 Price an option

7.1 Insurer

Let's say that I'm an insurer, I sell the option at $t = 0$ and buy it at the maturity T, hence my portfolio's value is

$$\Pi_t \quad = \quad Ce^{rT} \; - \; (S_T - K)^+.$$

7.1.1 First approach

My criteria is that I want on average $\Pi_t = 0$, *i.e.* $\mathbb{E}_{\mathbb{P}}[\Pi_t] = 0$, with $\mathbb{P}$ the historical probability. We recall that under this probability, $\mathbb{E}_{\mathbb{P}}[e^{-rT}S_T] \neq S_0$.

Example 7.1. Assume that under $\mathbb{P}$ the asset drives like

$$\frac{\mathrm{d}S_t}{S_t} \quad = \quad \mu \,\mathrm{d}t \; + \; \sigma \,\mathrm{d}W_t \,, \qquad \mu \neq r.$$

We find then

$$\begin{aligned} C \quad &= \quad e^{-rT}\mathbb{E}_{\mathbb{P}}\left[(S_T - K)^+\right] \\ &\neq \quad C^{BS}(S_0, K, T, r, \sigma). \end{aligned}$$

7.1.2 Second approach

Now assume that we want C s.t. $\Pi_t \geq 0$ $\mathbb{P}$-a.s., hence

$$C \quad \geq \quad e^{-rT}(S_T - K)^+ \qquad \mathbb{P}\text{-a.s.}$$

and then $C = \infty$.

7.2 Baby trader

Now we have the right to hedge but just one time at $t = 0$, hence I sell the option and buy Δ of the asset at $t = 0$, and at the maturity I buy the option and sell Δ in the asset, my portfolio's value is

$$\begin{aligned} \Pi_t \quad &= \quad Ce^{rT} \; - \; \Delta S_0 e^{rT} \; - \; (S_T - K)^+ \; + \; \Delta S_T \\ &= \quad Ce^{rT} \; - \; (S_T - K)^+ \; + \; \Delta(S_T - S_0). \end{aligned}$$

Remark. The trader has one parameter more than the insurer, its prices will then be lower.

7.2.1 Variance minimization

We still want $\mathbb{E}_{\mathbb{P}}[\Pi_t] = 0$, and furthermore we want to minimize the variance, the problem writes

$$\min_{\substack{(\Delta, C) \\ \text{s.t. } \mathbb{E}_{\mathbb{P}}[\Pi_t]=0}} \quad \mathbb{E}_{\mathbb{P}}\left[\Pi_t\right] \quad = \quad 0.$$

It's a problem of quadratic optimisation under linear constraint, the solution is

$$\begin{aligned} C_q \quad &= \quad e^{-rT}\mathbb{E}_{\mathbb{P}}\left[(S_T - K)^+ + \Delta^* \left(S_T - S_0 e^{rT}\right)\right] \\ &= \quad e^{-rT}\mathbb{E}_{\mathbb{Q}^q}\left[(S_T - K)^+\right] \,, \end{aligned}$$

where $\mathbb{Q}^q$ s.t. $\mathbb{E}_{\mathbb{Q}^q}\left[e^{-rT}S_T\right] = S_0$.

7.2.2 Super replication

Let's assume that I never want to loose money, hence the problem writes

$$C_{\text{sup}} \quad = \quad \min\{C : \exists \Delta \text{ s.t. } \Pi_t \geq 0 \quad \mathbb{P}\text{-a.s.}\}.$$

But Π_t is piecewise, then the optimum is reached on edges,

$$\Pi_t \geq 0 \quad \mathbb{P}\text{-a.s.}$$

$$\Leftrightarrow \quad \begin{cases} Ce^{rT} \; - \; \Delta S_0 e^{rT} \; \geq \; 0 & (S_T = 0) \\ \Delta \; - \; 1 \; \geq \; 0 & (S_T = \infty) \\ Ce^{rT} \; + \; \Delta\left(K - S_0 e^{rT}\right) \; \geq \; 0 & (S_T = K). \end{cases}$$

It's a linear programming problem (simplex), and we find $\Delta = 1$, $C_{\text{sup}} = S_0$.

Dual problem Now let us see the problem in its dual form. First we can write the problem as

$$\begin{aligned} C_{\text{sup}} \quad = \quad \min_{(C,\Delta)} \max_{q(S)\geq 0} \; & C + \int q(\mathrm{d}S)\left(e^{-rT}(S - K)^+ \right. \\ & \left. + \, \Delta(Se^{-rT} - S_0) - C\right). \end{aligned}$$

Theorem 7.1 (Minimax). *Let $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}^m$ be compact convex sets. If $f : X \times Y \to \mathbb{R}$ is a continuous function that is convex–concave, then we have that*

$$\min_{x\in X} \max_{y \in Y} f(x, y) \quad = \quad \max_{y \in Y} \min_{x \in X} f(x, y).$$

Hence we have

$$\begin{aligned} C_{\text{sup}} \quad = \quad \max_{q(S)\geq 0} \min_{(C,\Delta)} \; & C + \int q(\mathrm{d}S)\left(e^{-rT}(S - K)^+ \right. \\ & \left. + \, \Delta(Se^{-rT} - S_0) - C\right). \end{aligned}$$

We see that the min according to C is like $C\left(1 - e^{rT}\int q(\mathrm{d}s)\right)$, then we want $\int q(\mathrm{d}s)e^{rT} = 1$. For Δ as well we have $\int q(\mathrm{d}s)e^{rT}\left(Se^{-rT} - S_0\right) = 0$. Then there is a probability $\mathbb{Q}$ s.t.

$$\mathbb{E}_{\mathbb{Q}}[1] \quad = \quad 1 \tag{7.1}$$

$$\mathbb{E}_{\mathbb{Q}}\left[S_T e^{-rT}\right] \quad = \quad S_0 \tag{7.2}$$

So finally

$$C_{\text{sup}} \quad = \quad \max_{\mathbb{Q}\in\mathcal{M}_1} \mathbb{E}_{\mathbb{Q}}\left[e^{-rT}(S_T - K)^+\right] \,,$$

where

$$\mathcal{M}_1 \quad := \quad \{\mathbb{Q} \sim \mathbb{P} \quad \text{ s.t. (7.1) and (7.2)}\} \,.$$

Remark. Then the martingale property comes from the dual form of the problem.

Proposition 7.2. *If $\mathcal{M}_1 = \varnothing$, then $C_{\text{sup}} = -\infty$ and there an arbitrage opportunity.*

Remark. $C_{\text{sub}} = \max\{C : \exists \Delta \text{ s.t. } \Pi_t \geq 0 \quad \mathbb{P}\text{-a.s.}\}$.

Theorem 7.3. *The price of the option C is without arbitrage iff*

$$C_{\text{sub}} \;\leq\; C \;\leq\; C_{\text{sup}}.$$

Theorem 7.4. *If there exists a replication strategy, $C_{\text{sub}} = C = C_{\text{sup}}$.*

Theorem 7.5. *There exists a probability $\mathbb{Q}^* \in \mathcal{M}_1$ s.t.*

$$C \;=\; e^{-rT}\mathbb{E}_{\mathbb{Q}^*}\left[(S_T - K)^+\right].$$

7.3 Real trader

We suppose to simplify that $r = 0$, but the following results are true with $r \neq 0$.

Now we can hedge are any time, such that

$$\Pi_t \;=\; C \;-\; (S_t - K)^+ \;+\; \sum_{i=1}^{n} \Delta_i (S_{i+1} - S_i).$$

Hence with the same calculus than previously we find

$$C_{\text{sup}} \;=\; \sup_{\mathbb{Q}\in\mathcal{M}_n} \mathbb{E}_{\mathbb{Q}}\left[(S_T - K)^+\right],$$

where

$$\mathcal{M}_n \;:=\; \{\mathbb{Q} \sim \mathbb{P}, \quad \mathbb{E}[S_i | S_1, \ldots, S_{i-1}] = S_{i-1}\}.$$

And then

$$C \;=\; \mathbb{E}_{\mathbb{Q}^*}\left[(S_T - K)^+\right], \qquad \mathbb{Q}^* \in \mathcal{M}_n.$$

7.4 Trader M2

Now $n \to \infty$,

$$\Pi_t \;=\; C \;-\; (S_T - K)^+ \;+\; \int_0^T \Delta_t \, \mathrm{d}S_t.$$

So we have

$$C_{\text{sup}} \;=\; \sup_{\mathbb{Q}\in\mathcal{M}_\infty} \mathbb{E}_{\mathbb{Q}}\left[(S_T - K)^+\right],$$

with

$$\mathcal{M}_\infty \;:=\; \{\mathbb{Q} \sim \mathbb{P}, \quad \text{where } S \text{ is a } \mathbb{Q}\text{-martingale}\}.$$

Let's assume we know that the dynamic is $\frac{\mathrm{d}S_t}{S_t} = \mu\,\mathrm{d}t + \sigma\,\mathrm{d}W_t$, hence $C = \mathbb{E}_{\mathbb{Q}^*}[(S_T - K)^+]$, $\mathbb{Q}^* \in \mathcal{M}_\infty$. But in fact with price unicity we have $\mathcal{M}_\infty = \{\mathbb{Q}^{BS}\}$ where under this probability $\frac{\mathrm{d}S_t}{S_t} = \sigma\,\mathrm{d}W_t$.

Remark. If we have $\frac{\mathrm{d}S_t}{S_t} = \mu\,\mathrm{d}t + \sigma_t\,\mathrm{d}W_t$ with σ_t a stochastic control, $\mathcal{M}_\infty$ is not just one point.

Now we say that we want C, Δ s.t. $\Pi_t = 0$ $\mathbb{P}$-a.s. So *Ansatz*: $C = u(0, S_0)$, $\Delta_t = \partial_S u(t, S_t)$ where u is solution of

$$\begin{cases} \partial_t u \;+\; \frac{1}{2}\sigma^2 S^2 \partial_S^2 u \;=\; 0 \\ u(T, S_T) \;=\; (S_T - K)^+. \end{cases}$$

We verify, with Itô,

$$\begin{aligned} \Pi_t \;&=\; u(0, S_0) \;+\; u(T, S_T) \;-\; u(0, S_0) \\ &\quad - \int_0^T \underbrace{\left(\partial_t u + \frac{1}{2}\sigma^2 S^2 \partial_S^2 u\right)}_{=\,0} \mathrm{d}t \;-\; (S_T - K)^+ \\ &=\; 0. \end{aligned}$$

7.5 Model independent option