Théorie des valeurs extrêmes

Laurent Gardes

Université de Strasbourg

Chapitre 1

Introduction

La théorie des valeurs extrêmes voit ses origines dans la résolution du problème suivant. Au Pays-Bas, 40 % du territoire est situé en dessous du niveau de la mer. Les phénomènes météorologiques extrêmes peuvent donc avoir des conséquences désastreuses comme en témoigne la tempête de 1953. Pour protéger les habitants et les infrastructures du pays, le gouvernement néerlandais a lancé un vaste plan de construction de digues : le plan Delta. En tenant compte des coûts de construction et de l'efficacité du réseau de digues, il a été décidé de construire des digues de telle sorte que, sur une année donnée, la probabilité de submersion soit de $p = 10^{-4}$. La question à présent est de savoir à quelle hauteur de digue cela correspond. La question a été posée à une équipe de statisticiens qui disposaient pour ce faire des valeurs des hauteurs d'eau maximales relevées durant 111 ans lors de 1877 tempêtes. Ces valeurs sont notées dans la suite $x_1, \ldots, x_n$ avec $n = 1877$.
La première étape a consisté à proposer un modèle pour ces observations. Ils ont supposé qu'elles étaient issues de n variables aléatoires $X_1, \ldots, X_n$ indépendantes et de même loi. L'hypothèse d'indépendance est facilement admissible, la force d'une tempête ne dépendant pas de celle de la précédente. Concernant l'hypothèse d'equi-distribution, on peut l'admettre si on ne tient pas compte du réchauffement climatique.
En supposant que le nombre de tempêtes sur une année est constant [1] et égal à N, on cherche donc la valeur h telle que

$$\mathbb{P}\left(\max_{i=1,\ldots,N} X_i > h\right) = p.$$

En notant $F(\cdot)$ la fonction de répartition commune aux X_i, on cherche donc h telle que

$$\mathbb{P}\left(\max_{i=1,\ldots,N} X_i \leq h\right) = 1 - p \;\Leftrightarrow\; F^N(h) = 1 - p \;\Leftrightarrow\; F(h) = (1-p)^{1/N}.$$

1. Ce qui est évidemment faux en toute rigueur mais c'est malgré tout une bonne approximation de la réalité.

La valeur de N étant inconnue, on la remplace par son estimation $\hat{N} = 1877/111$. En conclusion, en supposant que la fonction de répartition $F(\cdot)$ soit inversible, la hauteur h recherchée est

$$h = F^{-1}\left((1-p)^{1/\hat{N}}\right).$$

La valeur h est donc le quantile d'ordre $1-\alpha$ avec $\alpha = 1-(1-p)^{1/\hat{N}} \approx 5.914 \times 10^{-6}$. Cette valeur de α est très proche de zéro [2] rendant l'estimation du quantile compliquée (voir Figure 1.1). La valeur h est un quantile dit extrême et son estimation requiert d'extrapoler au-delà de la plus grande observation.

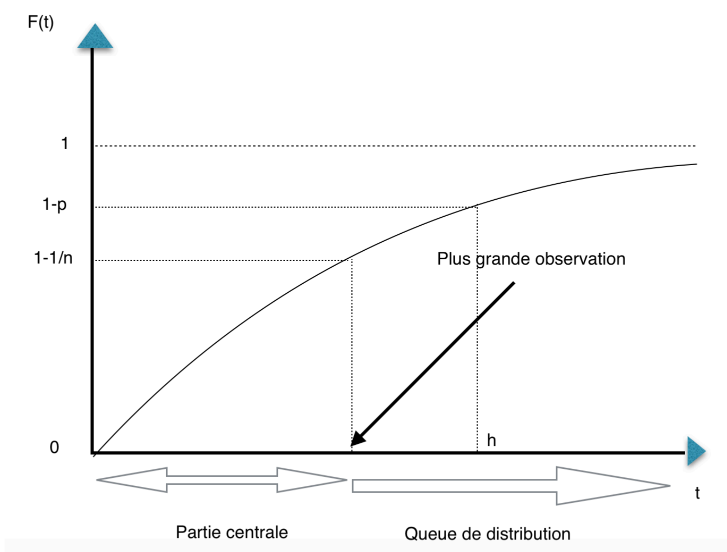

FIGURE 1.1 – Représentation de la fonction de répartition. Les valeurs observées de l'échantillon se situent dans la partie centrale alors que la valeur h du quantile se situe en queue de distribution.

Alors que la statistique "classique" s'intéresse principalement à la partie dite centrale de la loi modélisant au mieux le phénomène considéré (calcul de l'espérance, la médiane, la variance, utilisation du théorème central limite, etc.), nous souhaitons ici étudier les "grandes" valeurs c'est-à-dire la queue de distribution de la loi. La théorie des valeurs extrêmes propose un cadre théorique solide pour l'étude de ces valeurs dites extrêmes. La difficulté principale réside dans l'application concrète de la théorie (estimation, inférence, etc.) puisque par nature,

2. Elle est inférieure à $1/n = 5.328 \times 10^{-4}$, ce qui en fait un ordre de quantile extrême comme nous le verrons plus loin.

un évènement extrême est très peu observé. On parle d'évènement rare.

Enfin, la théorie des valeurs extrêmes est évidemment utile dans d'autres champs d'applications comme en finance, pour l'étude des "log-return" qui sont définis par $r_t := \log(p_t/p_{t-1})$ où p_t est la cotation en bourse à la date t d'une action donnée. Elle est également utilisée pour estimer (si elle existe) la durée maximale pour la vie d'un être humain.

Avant de présenter en détail cette théorie des valeurs extrêmes, le paragraphe suivant est consacré à la notion de quantile.

1.1 Introduction à la notion de quantile

Nous commencons par donner la définition de l'inverse généralisée d'une fonction croissante et continue à droite.

Définition 1 *Soit Φ une fonction croissante et continue à droite sur $\mathbb{R}$. L'inverse généralisée de Φ est définie par :*

$$\Phi^{\leftarrow}(y) := \inf\{x|\Phi(x) \geq y\},$$

avec la convention $\inf\{\emptyset\} = +\infty$.

Il est clair que dans le cas où Φ est une fonction continue, l'inverse généralisée coïncide avec l'inverse classique. On peut aussi définir l'inverse généralisée d'une fonction décroissante et continue à droite Ψ par

$$\Psi^{\leftarrow}(y) := \inf\{x|\Psi(x) \leq y\}$$

L'inverse généralisée vérifie les propriétés utiles suivantes :

Proposition 1 *Soit Φ une fonction croissante et continue à droite. L'inverse généralisée $\Phi^{\leftarrow}$ est une fonction croissante et continue à gauche. On a de plus les propriétés ci-dessous :*

i) $\Phi(\Phi^{\leftarrow}(y)) \geq y$,

ii) $\Phi^{\leftarrow}(y) \leq x \Leftrightarrow y \leq \Phi(x)$,

iii) $x < \Phi^{\leftarrow}(y) \Leftrightarrow y > \Phi(x)$.

Cette Proposition nous permet de démontrer facilement un résultat très utile tant d'un point de vu théorique que pratique pour simuler des réalisations de variables aléatoires.

Lemme 1 (Lemme de transformation des quantiles) *Soit $U : (\Omega, \mathcal{F}, \mathbb{P}) \mapsto ([0,1], \mathcal{B}[0,1])$ une variable aléatoire de loi uniforme sur $[0,1]$. Soit $F : \mathbb{R} \mapsto [0,1]$ une fonction de répartition. Alors la variable aléatoire $X := F^{\leftarrow} \circ U : (\Omega, \mathcal{F}, \mathbb{P}) \mapsto (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ est une variable aléatoire de fonction de répartition F.*

Nous pouvons à présent donner la définition d'un quantile.

Définition 2 *Soit F une fonction de répartition. Le quantile d'ordre α de la fonction de répartition F est défini pour $\alpha \in [0,1]$ par*

$$F^{\leftarrow}(\alpha) = \inf\{x | F(x) \geq \alpha\}.$$

La fonction quantile correspond donc à l'inverse généralisée de la fonction de répartition F au point α. Noter que le quantile d'ordre 1 est égal au point terminal de F défini par $x_F := \inf\{x|F(x) \geq 1\}$. Le point terminal est la plus grande valeur possible pour la réalisation d'un échantillon issu d'une distribution F. En théorie des valeurs extrêmes, on s'intéressera donc à des quantiles dont l'ordre α est proche de 1. Comme l'on préfère manipuler des quantités qui sont proches de 0 (plutôt que de 1), on appelera par abus de language quantile d'ordre α la quantité :

$$q(\alpha) = F^{\leftarrow}(1-\alpha) = \inf\{x|\bar{F}(x) \leq \alpha\},$$

où $\bar{F} = 1 - F$ est appelée fonction de survie. Il est à noter que d'après le point $i)$ de la Proposition 1,

$$\mathbb{P}(X > q(\alpha)) = 1 - F(F^{\leftarrow}(1-\alpha)) \leq \alpha.$$

On a l'égalité si $q(\alpha)$ est un point de continuité de F. Pour estimer $q(\alpha)$, l'idée la plus naturelle (sans faire d'hypothèse supplémentaire) est d'estimer F par la fonction de répartition empirique définie par :

$$\hat{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}\{X_i \leq x\}.$$

On montre alors facilement qu'un estimateur de $q(\alpha)$ est donné par :

$$\hat{F}_n^{\leftarrow}(1-\alpha) = \sum_{i=1}^{n} X_{n-i+1,n}\mathbb{I}\left\{\alpha \in \left[\frac{i-1}{n}, \frac{i}{n}\right[\right\} = X_{n-\lfloor n\alpha\rfloor,n},$$

où $\lfloor x \rfloor$ est le plus grand entier inférieur ou égal à x et $X_{1,n} \leq \ldots \leq X_{n,n}$ est l'échantillon ordonné (la variable aléatoire $X_{i,n}$, $i \in \{1,\ldots,n\}$ est une statistique d'ordre).

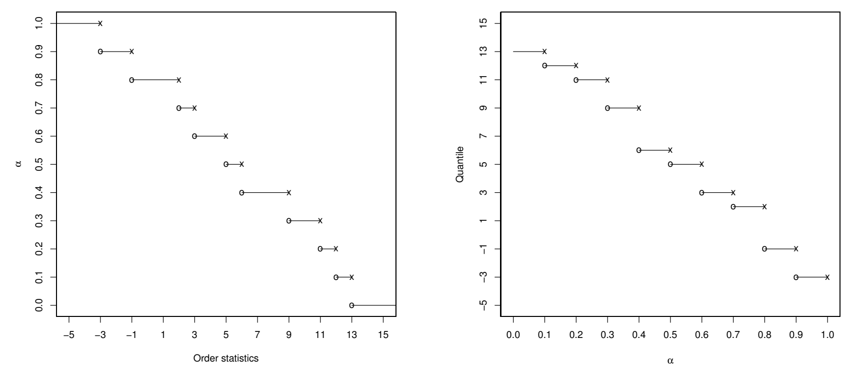

FIGURE 1.2 – Représentation de la fonction de survie empirique pour l'échantillon $x = (-3, -1, 2, 3, 5, 6, 9, 11, 12, 13)$ (figure de gauche) et du quantile empirique associé (figure de droite).

On remarque que l'utilisation de la fonction de répartition empirique n'est pas conseillée lorsque l'on s'intéresse à l'estimation de quantile d'ordre α proche de 0. En effet, dès que $\alpha < 1/n$, $\hat{F}_n^{\leftarrow}(1-\alpha) = X_{n,n}$ et donc l'estimateur du quantile est dans ce cas borné au maximum de l'échantillon. La zone $[X_{n,n}, \infty)$ est appelée la queue de distribution de F et un quantile appartenant à la queue de distribution est appelé un quantile extrême. Nous classons ci-dessous plus précisément les différents types de quantiles en fonction de leur ordre α_n (qui peut donc dépendre de la taille de l'échantillon).

Définition 3 *On classe les quantiles en trois catégories selon leur ordre* α_n *:*
i) On dira d'un quantile qu'il est **classique** *si* $n\alpha_n \to \infty$.
i) On dira d'un quantile qu'il est **intermédiaire** *si* $n\alpha_n \to c \in [1, \infty[$.
iii) On dira d'un quantile qu'il est **extrême** *si* $n\alpha_n \to c \in [0, 1[$.

Noter que les quantiles "classiques" et "intermédiaires" peuvent être estimés à l'aide de la fonction de survie empirique.

1.2 Quantile et assurance

La notion de quantile est très largement utilisée en actuariat pour quantifier le risque d'un portefeuille d'assurances. La mesure de risque la plus connue est la "Value at Risk (VaR)" définie ci-dessous.

Définition 4 *Soit X une variable aléatoire réelle de fonction de répartition F et soit $\alpha \in [0, 1]$. La Value at Risk de X au niveau α est donnée par :*

$$VaR_\alpha(X) := F^{\leftarrow}(1-\alpha).$$

La Value at Risk de niveau α n'est donc rien d'autre que le quantile d'ordre α que l'on avait noté $q(\alpha)$ dans le paragraphe précédent. Généralement, X est une variable aléatoire modélisant un risque pour la compagnie d'assurance (par exemple, le niveau d'eau d'une rivière, etc.) et la Value at Risk représente donc un seuil du risque ayant une probabilité α d'être dépassé. Cependant, bien que très utilisée, la Value at Risk ne donne aucune information sur ce qui se passe lorsque ce seuil est dépassé. On peut par exemple avoir deux variables aléatoires X et Y ayant une même valeur de la Value at Risk au niveau α mais présentant des queues de distribution très différentes (voir Figure 1.2).

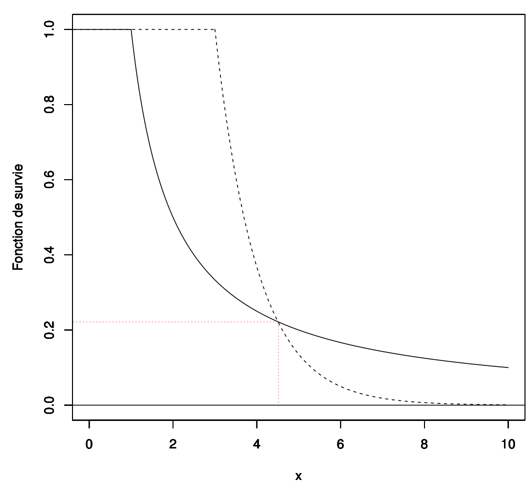

FIGURE 1.3 – Représentation de deux fonctions de survie $\bar{F}(x) = 1/x$ (trait continu) et $\bar{G}(x) = \exp(-(x-3))$ (pointillés) ayant une même Value at Risk au niveau $\alpha \approx 0.22$.

Pour remédier à ce problème, une autre mesure de risque a été proposée : la "Tail Value at Risk (TVaR)".

Définition 5 *Soit X une variable aléatoire réelle de fonction de répartition F et soit $\alpha \in [0,1]$. La Tail Value at Risk de X au niveau α est donnée par :*

$$TVaR_\alpha(X) = \frac{1}{\alpha}\int_0^\alpha VaR_s(X)ds, \tag{1.1}$$

si cette intégrale existe.

La Tail Value at Risk de niveau α est donc la moyenne des Value at Risk de niveaux inférieurs à α. Ceci permet d'avoir de l'information sur la queue de distribution au delà du niveau α. Le résultat ci-dessous donne une autre écriture de la Tail Value at Risk dans le cas d'une fonction de répartition dérivable presque-partout.

Proposition 2 *Soit X une variable aléatoire intégrable, de fonction de répartition F continue et dérivable presque-partout. On a :*

$$TVaR_\alpha(X) = \frac{1}{\alpha}\mathbb{E}\left(X\mathbb{I}\{X \geq VaR_\alpha(X)\}\right) = \mathbb{E}\left(X | X \geq VaR_\alpha(X)\right).$$

Ce résultat a motivé la définition d'une autre mesure de risque : la "Conditional Tail Expectation" définie pour une variable aléatoire X quelconque par

$$CTE_\alpha(X) := \mathbb{E}\left(X | X \geq VaR_\alpha(X)\right).$$

Cette mesure donne donc la valeur moyenne du risque X sachant que l'on a dépassé le seuil $VaR_\alpha(X)$. La Proposition 2 assure que pour une variable aléatoire admettant une densité par rapport à la mesure de Lebesgue, la Conditional Tail Expectation coïncide avec la Tail Value at Risk.

En conclusion, les mesures de risque en assurance sont intimement liées à la notion de quantile. Lorsque le niveau de risque α est proche de zéro, l'estimation des mesures de risque passe donc par la théorie des valeurs extrêmes.

Chapitre 2

Théorie des valeurs extrêmes

2.1 Convergence en loi du maximum d'un échantillon

Soient $X_1, \ldots, X_n$ un échantillon de n variables aléatoires indépendantes et identiquement distribuées de fonction de répartition commune F. En statistique "classique", on s'intéresse essentiellement à la variable aléatoire $S_n := X_1 + \ldots + X_n$. La loi faible des grands nombres assure que si X_1 est intégrable alors $S_n/n \xrightarrow{\mathrm{P}} \mathbb{E}(X_1)$. De plus, si X_1 est de carré intégrable, le Théorème de la Limite Centrale assure que

$$\left(\frac{n}{\mathrm{Var}(X_1)}\right)^{1/2}\left[\frac{S_n}{n} - \mathbb{E}(X_1)\right] \xrightarrow{\mathcal{L}} \mathcal{N}(0,1).$$

La somme S_n permet donc d'obtenir des résultats sur la partie centrale de la distribution. Dans notre cas, on s'intéresse à la queue de distribution et c'est donc tout naturellement que l'on souhaite étudier le comportement du maximum de l'échantillon. Ce maximum est noté dans la suite $X_{n,n} = \max(X_1, \ldots, X_n)$. On rappelle que la fonction de répartition de $X_{n,n}$ est donnée par :

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

Si on note $x_F := q(0) = \inf\{x;\ F(x) = 1\}$, on montre que $X_{n,n} \xrightarrow{\mathrm{P}} x_F$ lorsque $n \to \infty$. Pour obtenir une convergence en loi, il faut donc, à l'instar du TCL, normaliser le maximum.

Définition 6 *Soit H une fonction de répartition non-dégénérée (c'est-à-dire une fonction de répartion qui n'est pas associée à une variable constante presque sûrement). On dit que F* **appartient au domaine d'attraction de** *H (et on note $F \in \mathcal{DA}(H)$) s'il existe deux suites $a_n > 0$ et b_n telles que en tout point de continuité x de H,*

$$\lim_{n\to\infty} F^n(a_n x + b_n) = H(x) \;\Leftrightarrow\; \frac{X_{n,n} - b_n}{a_n} \xrightarrow{\mathcal{L}} Y,$$

où la variable aléatoire Y admet H pour fonction de répartition.

Nous nous intéressons à présent à la question de l'unicité du domaine d'attraction. Commençons par définir la notion de fonctions de répartition de même type.

Définition 7 *Soient H et G deux fonctions de répartition non-dégénérées. On dit que H et G* **sont de même type** *s'il existe $a > 0$ et b tels que pour tout $x \in \mathbb{R}$, $H(x) = G(ax + b)$.*

Proposition 3 *Soient H et G deux fonctions de répartition non-dégénérées. Si $F \in \mathcal{DA}(H)$ (c'est-à-dire s'il existe deux suites $a_n > 0$ et b_n telles que $F^n(a_n x + b_n) \to H(x)$) et si $F \in \mathcal{DA}(G)$ (c'est-à-dire s'il existe deux suites $u_n > 0$ et v_n telles que $F^n(u_n x + v_n) \to G(x)$), alors, les fonctions de répartition H et G sont de même type ($G(x) = H(ax + b)$). De plus, $u_n/a_n \to a > 0$ et $(v_n - b_n)/a_n \to b$.*

Nous donnons à présent le résultat fondamental de la théorie des valeurs extrêmes. Ce résultat est à rapprocher du théorème de la limite centrale.

Théorème 1 *Si $F \in \mathcal{DA}(G)$ où G est non dégénérée, alors G est de même type que la fonction de répartition définie pour tout x tel que $1 + \gamma x > 0$ par*

$$H_\gamma(x) = \begin{cases} \exp[-(1+\gamma x)^{-1/\gamma}] & \text{si } \gamma \neq 0, \\ \exp(-\mathrm{e}^{-x}) \text{ si } \gamma = 0. \end{cases}$$

La loi limite du maximum dépend donc du seul paramètre γ appelé l'indice des valeurs extrêmes. Selon le signe de γ, on définit trois types de domaines d'attraction : domaine d'attraction de Fréchet lorsque $\gamma > 0$; de Weibull lorsque $\gamma < 0$ et de Gumbel lorsque $\gamma = 0$. Le tableau suivant regroupe quelques lois usuelles classées en fonction de leur domaine d'attraction.

Gumbel	**Fréchet**	**Weibull**
Normale	Pareto	Uniforme
Exponentielle	Log-gamma	Beta
Log-normale	Student	
Gamma		
Weibull		

On rappelle que la loi de Weibull de paramètre $\beta > 0$ admet pour fonction de répartition $F(x) = 1 - \exp(-x^\beta)$ pour $x > 0$. La fonction de répartition de la loi de Pareto de paramètres $\gamma > 0$ et $a > 0$ est $F(x) = 1-(x/a)^{-1/\gamma}$ pour $x > a$.

Le paramètre γ est une paramètre de forme. Il contrôle la forme de la queue de distribution. Nous verrons un peu plus loin que

1) Si $\gamma > 0$, la fonction de survie $1 - F(x)$ converge vers 0 lorsque $x \to x_F^* = +\infty$ à une vitesse polynomiale (proportionnelle à une puissance de x). Ce domaine d'attraction regroupe les distributions *heavy-tailed* (à queue "lourde')'.
2) Si $\gamma = 0$, la fonction de survie $1 - F(x)$ converge vers 0 lorsque $x \to x_F^*$ à une vitesse exponentielle. On parle ici de light-tailed distributions (à queue légère)
3) Si $\gamma < 0$, la fonction de survie $1 - F(x)$ converge vers 0 lorsque $x \to x_F^* < +\infty$ à une vitesse polynomiale (proportionnelle à une puissance de x). Ce domaine d'attraction regroupe la plupart des distributions ayant un point terminal fini.

Il est donc intuitivement clair que si deux fonctions de survie sont asymptotiquement proportionnelles, elles appartiennent au même domaine d'attraction avec le même indice des valeurs extrêmes.

Définition 8 *Soient F et G deux fonctions de répartition. On dit que F et G ont des queues de distribution proportionnelles si elles ont le même point terminal x^* et si*

$$\lim_{x \to x^*} \frac{\bar{F}(x)}{\bar{G}(x)} = c \in]0, \infty[.$$

Proposition 4 *Si F et G ont des queues de distribution proportionnelles alors si $F \in \mathcal{DA}(H_\gamma)$, il en est de même pour G.*

2.2 Caractérisation des domaines d'attraction

2.2.1 Quelques résultats sur les fonctions à variations régulières

Définition 9 *Une fonction mesurable $U : \mathbb{R} \mapsto [0, \infty[$ est dite à variations régulières d'indice $\alpha \in \mathbb{R}$ à l'infini si pour tout $t > 0$,*

$$\lim_{x \to \infty} \frac{U(tx)}{U(x)} = t^\alpha.$$

On notera dans la suite $\mathcal{RV}_\alpha$ l'ensemble des fonctions à variations régulières d'indice α.

Définition 10 *Si une fonction $L : \mathbb{R} \mapsto [0, \infty[$ est à variations régulières d'indice 0 ($L \in \mathcal{RV}_0$), on dit que L est à variations lentes.*

Le résultat ci-dessous fournit une représentation des fonctions à variations régulières.

Proposition 5 (Représentation de Karamata) *Toute fonction $U \in \mathcal{RV}_\alpha$ avec $\alpha \in \mathbb{R}$ s'écrit sous la forme*

$$U(x) = c(x) \exp\left\{\int_{x_0}^{x} \frac{\Delta(t)}{t} dt\right\},$$

où $x_0 \in \mathbb{R}$, c et Δ sont des fonctions mesurables avec $c(x) \to c > 0$ et $\Delta(x) \to \alpha$ lorsque x converge vers l'infini.

Définition 11 *Si la fonction c est constante, on dit que U est une fonction à variations régulières (ou rapides) normalisée.*

Nous donnons à présent quelques propriétés utiles sur les fonctions à variations régulières. Le premier résultat assure que la convergence simple de la définition 9 est en fait une convergence localement uniforme (c'est-à-dire uniforme sur les intervalles fermés de $\mathbb{R}$).

Proposition 6 (Convergence uniforme locale) *Si $U \in \mathcal{RV}_\alpha$, $\alpha \in \mathbb{R}$ alors pour tout intervalle $I = [a, b] \subset \mathbb{R}$, $0 < a < b < \infty$, on a :*

$$\sup_{t \in I} \left| \frac{U(tx)}{U(x)} - t^\alpha \right| \to 0 \text{ lorsque } x \to \infty.$$

Si $\alpha < 0$, le résultat de convergence uniforme est vrai pour des intervalles I de la forme $[a, \infty[$, $a > 0$. Si $\alpha > 0$ et si U est bornée sur les intervalles $]0, b]$, $b > 0$, alors la convergence uniforme est vraie sur les intervalles $]0, b]$, $b > 0$.

Une propriété intéressante des fonctions à variations régulières est la conservation des équivalents à l'infini.

Proposition 7 *Soit $U \in \mathcal{RV}_\alpha$, $\alpha \in \mathbb{R}$ et soient (u_n) et (v_n) deux suites positives telles que $u_n \to \infty$ et $v_n \to \infty$ lorsque n tend vers l'infini. Alors si $u_n \sim v_n$, on a $U(u_n) \sim U(v_n)$.*

Le prochain résultat s'intéresse au comportement à l'infini d'une fonction à variations régulières.

Proposition 8 *Si $U \in \mathcal{RV}_\alpha$, $\alpha \in [-\infty, +\infty]$, alors,*

$$\lim_{x \to \infty} \frac{\log U(x)}{\log(x)} = \alpha.$$

Ainsi,

$$\lim_{x \to \infty} U(x) = \begin{cases} 0 & \text{si } \alpha < 0, \\ +\infty & \text{si } \alpha > 0. \end{cases}$$

On s'intéresse à présent à l'inverse généralisée d'une fonction à variations régulières croissante.

Proposition 9 *Si $U \in \mathcal{RV}_\alpha$ avec $\alpha \geq 0$ est une fonction croissante telle que $U(x) \to \infty$ lorsque $x \to \infty$ alors l'inverse généralisée de U est à variations régulières d'indice $1/\alpha$ ($U^{\leftarrow} \in \mathcal{RV}_{1/\alpha}$).*

Remarque 1 *Si $U \in \mathcal{RV}_\alpha$ avec $\alpha < 0$ est une fonction décroissante, alors la fonction $U^{\leftarrow}(1/\cdot)$ est à variations régulières d'indice $-1/\alpha$. En effet, si $U \in \mathcal{RV}_\alpha$ alors $1/U \in \mathcal{RV}_{-\alpha}$. Donc, d'après la Proposition 9, $(1/U)^{\leftarrow} \in \mathcal{RV}_{-1/\alpha}$. Il reste à remarquer que $(1/U(\cdot))^{\leftarrow} = U^{\leftarrow}(1/\cdot)$.*

Pour finir ce paragraphe, nous nous intéressons à la dérivée et l'intégrale d'une fonction à variations régulières.

Dérivation des fonctions à variations régulières

Proposition 10 *Si U est une fonction à variations régulières normalisée d'indice $\alpha \in \mathbb{R}$ qui est dérivable de dérivée u alors*

$$\lim_{x\to\infty} \frac{xu(x)}{U(x)} = \alpha. \tag{2.1}$$

Si de plus $\alpha \neq 0$, $|u| \in \mathcal{RV}_{\alpha-1}$.

2.2.2 Domaine d'attraction de Fréchet

Le résultat ci-dessous assure que toute fonction appartenant au domaine d'attraction de Fréchet est une fonction à variations régulières.

Théorème 2 *Une fonction de répartition F appartient au domaine d'attraction $\Psi_\gamma^{(F)}$ (de Fréchet, avec un indice des valeurs extrêmes $\gamma > 0$) si et seulement si la fonction de survie $\bar{F} \in \mathcal{RV}_{-1/\gamma}$. Des suites possibles de normalisation (a_n) et (b_n) sont données par $a_n = \bar{F}^{\leftarrow}(1/n)$ et $b_n = 0$.*

Autrement dit, une fonction de répartition F appartenant au domaine d'attraction de Fréchet s'écrit sous la forme :

$$F(x) = 1 - x^{-1/\gamma} L(x),\ L \in \mathcal{RV}_0. \tag{2.2}$$

Remarque 2 *a) Toutes les fonctions de répartition du domaine d'attraction de Fréchet ont un point terminal infini. En effet, par définition, une fonction à variations régulières ne peut pas être nulle à partir d'un certain rang.*
b) D'après la Remarque 1, l'équation (2.2) est équivalente à :

$$q(\alpha) = \alpha^{-\gamma} \ell(\alpha^{-1}), \tag{2.3}$$

où $\alpha \in [0, 1]$ et ℓ est une fonction à variations lentes.

2.2.3 Domaine d'attraction de Weibull

Le résultat suivant montre que l'on passe du domaine d'attraction de Fréchet à celui de Weibull par un simple changement de variable dans la fonction de répartition.

Théorème 3 *Une fonction de répartition F appartient au domaine d'attraction de $\Psi_\gamma^{(W)}$ (Weibull, avec un indice des valeurs extrêmes $\gamma < 0$) si et seulement si son point terminal x_F est fini et si la fonction de répartition F^* définie par*

$$F^*(x) = \begin{cases} 0 & \text{si } x < 0 \\ F(x_F - 1/x) & \text{si } x \geq 0, \end{cases}$$

appartient au domaine d'attraction de Fréchet avec un indice des valeurs extrêmes $-\gamma > 0$. Des suites possibles de normalisation (a_n) et (b_n) sont données par $a_n = x_F - \bar{F}^{\leftarrow}(1/n)$ et $b_n = x_F$.

Ainsi, une fonction de répartition F du domaine d'attraction de Weibull s'écrit pour $x \leq x_F$:

$$F(x) = 1 - (x_F - x)^{-1/\gamma} L((x_F - x)^{-1}),\ L \in \mathcal{RV}_0. \tag{2.4}$$

En utilisant la Remarque 1, le quantile d'ordre $\alpha \in [0, 1]$ associé s'écrit :

$$q(\alpha) = x_F - \alpha^{-\gamma} \ell(\alpha^{-1}),\ \ell \in \mathcal{RV}_0. \tag{2.5}$$

2.2.4 Domaine d'attraction de Gumbel

La caractérisation des fonctions de répartition du domaine d'attraction de Gumbel est plus complexe.

Théorème 4 *Une fonction de répartition F appartient au domaine d'attraction de $\Psi^{(G)}$ (Gumbel) si et seulement si il existe $z < x_F \leq \infty$ tel que*

$$\bar{F}(x) = c(x) \exp\left\{ -\int_z^x \frac{1}{a(t)} dt \right\},\ z < x < x_F, \tag{2.6}$$

où $c(x) \to c > 0$ lorsque $x \to x_F$ et a est une fonction positive et dérivable de dérivée a' telle que $a'(x) \to 0$ lorsque $x \to x_F$. Des suites possibles de normalisation (a_n) et (b_n) sont données par $b_n = \bar{F}^{\leftarrow}(1/n)$ et $a_n = a(b_n)$.

Le domaine d'attraction de Gumbel regroupe une grande diversité de lois comptant parmi elles la plupart des lois usuelles (loi normale, exponentielle, gamma, log-normale). Cette famille étant difficile à étudier dans toute sa généralité, de nombreux auteurs se sont concentrés sur une sous-famille : les lois à queue de type Weibull.

Définition 12 *On dit qu'une loi est à queue de type Weibull si sa fonction de répartition est de la forme*

$$\bar{F}(x) = \exp\left\{ -V^{\leftarrow}(x) \right\}, \tag{2.7}$$

où $V \in \mathcal{RV}_\theta$, $\theta > 0$ étant appelé l'indice de queue de Weibull.

Cette famille de lois contient par exemple les lois normale, Gamma, exponentielle, etc. Par contre, la loi log-normale qui appartient au domaine d'attraction de Gumbel n'est pas une loi à queue de type Weibull.

2.2.5 Loi de Pareto Généralisée - Théorème de Pickands

Nous donnons tout d'abord la définition de la loi de Pareto Généralisée.

Définition 13 *La loi de Pareto Généralisée de paramètres $\gamma \in \mathbb{R}$ et $\sigma > 0$ est définie par sa fonction de répartition donnée par :*

$$G_{\gamma,\sigma}(x) = 1 + \log H_\gamma\left(\frac{x}{\sigma}\right) = \begin{cases} 1 - (1+\gamma x/\sigma)^{-1/\gamma} & si\ \gamma \neq 0, \\ 1 - \exp(-x/\sigma) & si\ \gamma = 0., \end{cases}$$

pour $x \in \{t \in \mathbb{R};\ 1 + \gamma t/\sigma > 0\} \cap [0, \infty[$.

Le résultat suivant permet de caractériser l'ensemble des trois domaines d'attraction.

Proposition 11 *La fonction de répartition F de point terminal x_F appartient au domaine d'attraction de H_γ si et seulement si il existe une fonction positive a telle que*

$$\lim_{u \uparrow x_F} \frac{\bar{F}(u + xa(u))}{\bar{F}(u)} = 1 - G_{\gamma,1}(x).$$

Nous donnons à présent le résultat important de ce paragraphe assurant que la loi des excès au dessus d'un seuil peut-être approchée par une loi de Pareto Généralisée. Donnons dans un premier temps la définition de la loi des excès.

Définition 14 *Soit X une variable aléatoire de fonction de répartition F et de point terminal x_F. Pour tout $u < x_F$ la fonction de répartition des excès de X au dessus du seuil u est définie pour tout $x \in [0, x_F - u[$ par :*

$$F_u(x) := \mathbb{P}(X - u \leq x | X > u).$$

De même, si X est intégrable, la fonction excès moyen (mean excess function) de X est donnée par :

$$e_u(X) := \mathbb{E}(X - u | X > u).$$

Théorème 5 *La fonction de répartition F de point terminal x_F appartient au domaine d'attraction de H_γ si et seulement si il existe une fonction positive a telle que*

$$\lim_{u \uparrow x_F} \sup_{0 < x < x_F - u} \left| F_u(x) - G_{\gamma,a(u)}(x) \right| = 0.$$

Chapitre 3

Statistiques d'ordres

Lorsqu'on dispose d'un échantillon $X_1, \ldots, X_n$ de variables aléatoires indépendantes et de même fonction de répartition F, les statistiques d'ordres sont les observations ordonnées c'est-à-dire les variables aléatoires $X_{1,n} \leq \ldots \leq X_{n,n}$. Evidemment, les statistiques d'ordre ne sont ni indépendantes ni de même loi. L'objectif de ce chapitre est de donner quelques résultats sur la loi de la statistique $X_{k,n}$ pour $k \in \{1, \ldots, n\}$.

3.1 Distribution des statistiques d'ordre

Proposition 12 *La fonction de répartition de la statistique d'ordre $X_{k,n}$ est donnée pour tout $x \in \mathbb{R}$ par*

$$F_{k,n}(x) := \mathbb{P}(X_{k,n} \leq x) = \sum_{r=k}^{n} \mathrm{C}_n^r F^r(x)(1 - F(x))^{n-r}.$$

Dans le cas où la fonction de répartition F est continue et dérivable presque-partout de dérivée f, les statistiques d'ordres ont une loi qui admet une densité par rapport à la mesure de Lebesgue. La densité de $X_{k,n}$ est donnée par

$$f_{k,n}(x) = \frac{n!}{(k-1)!(n-k)!} f(x) F^{k-1}(x)[1 - F(x)]^{n-k}.$$

Comme autre conséquence directe de ce résultat, on montre facilement que si $U_1, \ldots, U_n$ sont des variables aléatoires indépendantes de loi uniforme alors $U_{k,n}$ suit une loi Beta de paramètres k et $n - k + 1$. En particulier, pour tout $s \in \mathbb{N}$,

$$\mathbb{E}(U_{k,n}^s) = \frac{n!}{(n+s)!} \frac{(k+s-1)!}{(k-1)!}.$$

Terminons ce paragraphe en donnant la densité du vecteur aléatoire des statistiques d'ordre.

Proposition 13 *Si la fonction de répartition F est continue et dérivable presque-partout de dérivée f, la densité du vecteur $(X_{1,n},\dots,X_{n,n})$ est*

$$f_n(x_1,\dots,x_n) = n!\prod_{i=1}^n f(x_i)\mathbb{I}\{x_1 \le \dots \le x_n\}.$$

Le résultat suivant est un résultat essentiel sur les statistiques d'ordre d'un échantillon uniforme ou exponentiel. Ce résultat sera utilisé à plusieurs reprises dans la suite.

Théorème 6 (Représentation de Rényi) *On note $U_1,\dots,U_n$ un échantillon de loi uniforme, $E_1,\dots,E_{n+1}$ et $F_1,\dots,F_n$ deux échantillons de loi exponentielle. On a les deux égalités en lois suivantes :*

$$\{U_{j,n},\ j=1,\dots,n\} \stackrel{\mathcal{L}}{=} \left\{\frac{T_j}{T_{n+1}},\ j=1,\dots,n\right\}$$

où $T_j = E_1+\dots+E_j$ et

$$\{E_{j,n},\ j=1,\dots,n\} \stackrel{\mathcal{L}}{=} \left\{\sum_{r=1}^{j}\frac{F_r}{n-r+1},\ j=1,\dots,n\right\}.$$