

République Algérienne Démocratique et Populaire  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique  
Université des Sciences et de la Technologie «Houari Boumediene»  
Faculté de Mathématiques



## Mémoire

Présenté pour l'obtention du diplôme de Magister

En : MATHEMATIQUES

Spécialité : Recherche Opérationnelle (Méthodes Stochastiques)

Par

TEMAM Kamel

Sujet

Analyse des séries chronologiques à queues lourdes et  
valeurs extrêmes

Soutenu, publiquement, le 21/07/2010, devant le jury composé de :

M. BENTARZI Mohamed	Professeur	à l'U. S. T. H. B.,	Président
M. AKNOUCHE Abdelhakim	M. C. A	à l'U. S. T. H. B.,	Directeur de Mémoire
M. KHALDI Khaled	M. C. A	à l'U. M. B. B.,	Examineur
M. BOUKHEBOUZE Ahmed	M. A. A	à l'U. S. T. H. B.,	Invité

# Table des matières

<b>Introduction générale</b>	<b>1</b>
<b>1 Outils mathématiques</b>	<b>5</b>
1.1 Convergence stochastique . . . . .	5
1.1.1 Convergence en un issu . . . . .	6
1.1.2 Convergence simple d'une suite de variables aléatoires . . . . .	6
1.1.3 Convergence presque sûre . . . . .	7
1.1.4 Convergence en probabilité . . . . .	7
1.1.5 Convergence en distribution . . . . .	7
1.1.6 Convergence en moyenne quadratique . . . . .	8
1.1.7 Relation entre modes de convergence . . . . .	8
1.2 Loi des grands nombres dans le cas iid . . . . .	9
1.2.1 Loi faible des grands nombres . . . . .	9
1.2.2 Loi forte des grands nombres . . . . .	10
1.3 Théorème central limite classique . . . . .	10
1.3.1 Cas des variables aléatoires indépendantes et identiquement distribuées (iid) . . . . .	10
1.3.2 Cas des variables aléatoires indépendantes ( non nécessairement (iid))	11
1.3.3 Théorème central limite de Lyapunov . . . . .	12

1.3.4	Théorème central limite de Lindeberg . . . . .	12
1.4	Lois infiniment divisibles . . . . .	13
1.4.1	Variable aléatoire infiniment divisible . . . . .	13
1.5	La stabilité et le domaine d'attraction . . . . .	15
1.5.1	Lois stables . . . . .	15
1.5.2	Forme des lois stables . . . . .	16
1.5.3	Calcul de la fonction caractéristique stable . . . . .	16
1.5.4	Domaine d'attraction d'une loi stable . . . . .	17
1.6	Convergence uniforme . . . . .	19
1.6.1	Convergence d'une fonction . . . . .	19
1.7	Variations régulières . . . . .	20
1.7.1	Propriétés de base . . . . .	21
1.7.2	Théorème de convergence uniforme pour les fonctions à variations régulières . . . . .	22
1.7.3	Théorème de Karamata . . . . .	23
1.8	Variables aléatoires à variations régulières . . . . .	24
1.8.1	Distribution d'une variable aléatoire à variations régulières . . . . .	24
1.9	Processus stochastiques . . . . .	28
1.9.1	Introduction . . . . .	28
1.9.2	Processus aléatoire . . . . .	29
1.9.3	Densité spectrale . . . . .	36
1.9.4	Equations aux récurrences stochastiques . . . . .	37
1.9.5	Comportement des queues . . . . .	39
<b>2</b>	<b>Analyse des phénomènes à queues lourdes : cas iid</b>	<b>43</b>
2.1	Introduction . . . . .	43
2.2	Classes importantes des distributions à queues lourdes . . . . .	46

2.2.1	Distributions à variations régulières d'indice $\alpha > 0$	46
2.2.2	Distributions sous-exponentielles	47
2.3	Conclusion	52
2.4	Estimation de l'indice de queue (Estimateur de Hill)	52
2.4.1	Introduction	52
2.4.2	Estimation d'une queue régulièrement variable	52
2.4.3	Propriétés probabilistes	56
2.4.4	Simulations et applications	57
2.5	Théorie des valeurs extrêmes	63
2.5.1	Introduction	63
2.5.2	Domaine d'attraction maximum	65
2.5.3	Distribution de Fréchet $\Phi_\alpha, \alpha > 0$	65
2.5.4	Distribution de Weibull $\Psi_\alpha, \alpha > 0$	66
2.5.5	Distribution de Gumbel $\Lambda$	66
2.5.6	Distribution de valeurs extrêmes généralisée ( <i>GEV</i> )	67
2.5.7	Estimation des paramètres	68
2.5.8	Application	69
2.5.9	Conclusion	75
<b>3</b>	<b>Processus à queues lourdes : cas stationnaire linéaire</b>	<b>76</b>
3.1	Introduction	76
3.2	Processus Bruit Blanc (White Noise)	77
3.3	Classe des modèles ARMA	77
3.4	Processus autorégressif d'ordre $p$	79
3.5	Processus Moyenne Mobile d'ordre $q$ (Moving Average)	82
3.6	Processus Moyenne Mobile d'ordre infinie ( $MA(\infty)$ )	84

3.7	Densité spectrale et le processus $ARMA(p, q)$ . . . . .	85
3.7.1	Densité spectrale de processus $ARMA(p, q)$ . . . . .	86
3.7.2	Densité spectrale de processus $MA(q)$ . . . . .	87
3.7.3	Densité spectrale de processus $AR(p)$ . . . . .	88
3.8	Processus linéaires de variance infinie . . . . .	90
3.8.1	Introduction . . . . .	90
3.8.2	Propriétés d'une variable aléatoire stable $\varepsilon$ . . . . .	91
3.8.3	Processus moyenne mobile d'ordre $q$ à variance infinie (Moving Average) . . . . .	98
3.8.4	Prédiction linéaire de processus $ARMA$ à variance infinie . . . . .	101
3.8.5	Prédiction linéaire de processus $MA(1)$ à variance infinie . . . . .	103
3.9	Valeurs extrêmes pour les processus linéaires . . . . .	104
3.9.1	Valeurs extrêmes pour la classe des processus $ARMA(p, q)$ . . . . .	104
3.9.2	Valeurs extrêmes de processus $MA(\infty)$ . . . . .	105
3.9.3	Estimation de l'indice extrême $\zeta$ (Estimateur de Hill) . . . . .	106
3.9.4	Simulations et applications . . . . .	106
3.10	Estimation de l'indice de queue (Estimateur de Hill) . . . . .	107
3.10.1	Simulations et applications . . . . .	107
3.10.2	Conclusion . . . . .	112
<b>4</b>	<b>Processus à queues lourdes : cas stationnaire non linéaire</b>	<b>113</b>
4.1	Introduction . . . . .	113
4.2	Les propriétés de base des processus $ARCH$ . . . . .	113
4.2.1	Propriétés du processus $ARCH(1)$ carré . . . . .	119
4.2.2	Comportement d'une queue d'un processus $ARCH(1)$ . . . . .	120
4.2.3	Estimation des paramètres de modèle $ARCH(1)$ par la méthode MV . . . . .	121

4.3	Valeurs extrêmes pour les processus non linéaires . . . . .	123
4.3.1	Notion d'indice extrême . . . . .	123
4.3.2	Valeurs extrêmes pour le processus $ARCH(1)$ . . . . .	124
4.3.3	Estimation de l'indice extrême $\zeta$ (Estimateur de Hill) . . . . .	126
4.4	Estimation de l'indice de queue (Estimateur de Hill) . . . . .	127
4.4.1	Simulations et applications . . . . .	128

**Bibliographie****130**

# Introduction générale

Dans la quasi-totalité des domaines scientifiques tels que les finances, les assurances, la météorologie, les sciences d'environnement et le traitement de signal, il est des phénomènes aléatoires qui sont caractérisés par de grandes probabilités de prendre de grandes valeurs. Les lois de probabilités correspondantes sont alors qualifiées de distributions à queues lourdes (ou épaisses). Cette notion de lourdeur de queue peut avoir plusieurs sens mais le point commun est qu'on fait toujours référence à la loi normale qui, elle, est qualifiée d'avoir des queues légères. Ainsi, on peut dire qu'une distribution est à queues lourdes soit parce qu'elle a un coefficient d'aplatissement (Kurtosis) plus grand que celui de la loi normale, soit parce que, contrairement à la loi normale, la décroissance des queues de la distribution vers zéro est moins rapide que la fonction exponentielle, soit encore parce que la décroissance vers zéro est asymptotiquement équivalente à une décroissance polynomiale avec une vitesse appropriée. Pour ce type de distributions, certains moments d'ordres supérieurs peuvent être infinis. Lorsqu'on désire modéliser de tels phénomènes, les procédures d'inférence basées sur la théorie asymptotique, à savoir le théorème central limite et loi des grands nombres, peuvent souffrir de cette inexistence de moments. Pourtant, il existe pour certaines classes des distributions à queues lourdes, dites alpha-stables, des variantes des théorèmes centraux limites et de la loi des grands nombres qui s'appliquent même pour le cas où certains moments sont infinis. Néanmoins, l'étape d'exploitation d'un modèle qui consiste le plus souvent en la prédiction, le lissage, ou le filtrage est généralement basée sur le principe de moyennisation (espérance, espérance conditionnelle) et peut être dans ce cas inappropriée, voire inefficace en raison de la présence de grandes dispersions des valeurs. Aussi, des outils plus adéquats et mieux adaptés pour l'exploitation des modèles de queues lourdes ont été introduits et sont fondés sur les quantiles (par exemple, la médiane, les quantiles extrêmes, la valeur à risque ou encore en anglais Value at Risk, VaR) dont les valeurs sont insensibles à la présence

de valeurs abberantes. Or ces mesures sont estimées et analysées à partir d'hypothèses semi-paramétriques concernant l'épaisseur asymptotique des queues, épaisseur qui le plus souvent est formulée au moyen d'un polynôme avec indice réel mesurant la vitesse de décroissance vers zéro et qu'on doit estimer à travers les données.

Jusqu'aux années 70, des théories et méthodes assez importantes ont été développées à propos de la modélisation des phénomènes de queues lourdes. L'accent a été mis tant sur le point de vue probabiliste des modèles proposés que sur les méthodes d'estimation des indices des queues et l'étude de leurs propriétés asymptotiques. Cependant, la plupart des résultats concernent des modèles de variables indépendantes et identiquement distribuées (i.i.d.), qui bien qu'ils modélisent certaines situations réelles, s'avèrent insuffisants pour décrire divers phénomènes financiers, actuariels, ou environnementales, phénomènes qui sont caractérisés par une structure de dépendance marquante. Ce cas de dépendance, a été abondamment abordé ces deux dernières décennies dans le cadre de l'hypothèse de stationnarité. En particulier, la forme linéaire des modèles a été prédominante dans un premier temps mais actuellement des modèles de forme non linéaires sont de plus en plus étudiés.

Par ailleurs, une problématique étroitement liée à l'étude des phénomènes de queues lourdes est celle de la modélisation des phénomènes extrêmes qui surgissent de manière rare mais dont les conséquences sont très considérables. L'étude de la distribution des maxima (ou minima) de phénomènes aléatoires est primordiale pour se prévenir contre les événements extrêmes tels que catastrophes naturelles, crash boursiers, ruines d'assurances,... Bien que ces distributions extrêmes soient connues dans le cas i.i.d., les distributions asymptotiques sont souvent dégénérées et il existe une théorie dite des valeurs extrêmes qui représente pour le maximum de variables ce que représente le théorème central limite pour la somme de variables. En effet, cette théorie se base sur le théorème de Fisher et Tippett (1928) qui stipule que la distribution asymptotique d'une suite de maxima (associée à une suite de variables aléatoires i.i.d.), centrées et réduites par des coefficients appropriées, converge en distribution vers une distribution limite qui est soit de Fréchet, soit de Weibul soit de Gumbel. La première est à queues lourdes, la seconde est à queues légères et la troisième est à queues rompues. Ces trois distributions dites *distributions extrêmes* (Extreme Value distributions, *EV*) ont été formulées de manière unifiée par Von Mises (1936) et Jenkinson (1955) sous le cadre d'une distribution générale appelée distribution extrême généralisée

(Generalized Extreme Value *GEV* distribution). Grâce à ce théorème, on peut alors estimer des caractéristiques extrêmes comme la Value at Risk mais à condition de connaître par quel type de distribution on a été attiré. A ce stade, l'estimation de l'indice de queue en est la clé. Une des méthodes les plus utilisées pour l'estimation de cet indice est la méthode des maxima par bloc qui décompose l'échantillon en plusieurs blocs dont pour chacun on détermine le maximum. On obtient ainsi une série de maxima avec laquelle on peut estimer l'indice par des méthodes appropriées (estimation de Hill, (1975), estimation de Pickands (1975), estimation du maximum de vraisemblance).

Cependant, dans le cas d'échantillons de petite taille, la théorie des valeurs extrêmes basée sur le théorème de Fisher-Tippett présente une limite sérieuse, en particulier pour l'application de la méthode des maxima par bloc. Il existe pourtant une deuxième approche en théorie des valeurs extrêmes qui est très utilisée en finance et actuariat et qui se base sur des méthodes comptant le nombre de dépassements au delà d'un certain seuil (méthode des "Peaks over threshold", *POT*).

Comme pour la modélisation des phénomènes à queues lourdes, la théorie des valeurs a d'abord été développée initialement pour le cas de variables i.i.d. En présence de dépendance dans les données, le problème se complique mais les premières généralisations ont été faites pour le cas de processus stationnaires. Pour certaines structures de dépendance, il existe des analogues du théorème de Fisher-Tippett. La deuxième approche basée sur les excédances (méthode *POT*), a été également généralisée en cas de dépendance en utilisant la théorie des processus ponctuels et des mesures aléatoires.

L'objectif de notre mémoire est d'étudier théorie, méthodes et applications des modèles de séries chronologiques stationnaires caractérisées par des queues lourdes. Nous en étudions également la théorie des valeurs extrêmes.

Dans un premier temps, nous étudions la problématique de phénomènes à queues lourdes ainsi que la théorie des valeurs extrêmes dans le cas de variables aléatoires i.i.d. En particulier, les propriétés probabilistes des distributions à queues lourdes notamment les distributions sous-exponentielles et les distributions à variations régulières sont présentées. Nous passons en revue plusieurs méthodes d'estimation des indices de queues ainsi que leurs propriétés asymptotiques. De plus, la théorie des valeurs extrêmes du point de vue du théorème de

Fisher-Tippett est également présentée. Une étude de simulation est effectuée pour déterminer les performances des différents estimateurs et une application à des données réelles confirme la pertinence de cette modélisation.

Dans une deuxième phase, nous étudions la même problématique relativement à des données dépendantes, mais plutôt stationnaires, vérifiant une équation linéaire *ARMA*. En particulier les modèles *ARMA* à variance infinie sont présentés. Nous concluons que pour avoir un processus *ARMA* à queues lourdes, il faudrait impérativement que le processus d'innovation (entrée du modèle) associé le soit. De plus la théorie des valeurs extrêmes est étudiée tant du point de vue de Fisher-Tippett que de celui de l'approche des excédances (Peaks Over Threshold, *POT*). Une étude de simulation et une application à des données réelles est alors donnée.

Enfin, dans une troisième étape, nous nous intéressons toujours à la modélisation des phénomènes de queues lourdes en présence de dépendance stationnaire mais formulée de façon non linéaire. Spécifiquement, nous nous focalisons sur le cas de modèles *ARCH*. Nous concluons que contrairement aux modèles linéaires, la distribution du processus solution peut être à queue lourde, régulièrement variée, même si le processus d'innovation (d'entrée) est à queue légère. La théorie des valeurs extrêmes concernant les modèles *ARCH* est également étudiée du point de vue de l'approche des excédances (méthode *POT*).

# Chapitre 1

## Outils mathématiques

L'analyse des queues lourdes est une branche de la théorie des valeurs extrêmes consacrée à l'étude des phénomènes régis par des grands mouvements. Elle englobe à la fois la modélisation probabiliste ainsi que l'inférence statistique. Ses outils mathématiques sont basés sur le concept des variations régulières, de convergence faible des mesures, de théorie des probabilités et des processus ponctuels.

La quasi-totalité des statistiques classiques sont basées sur des moyennes et des moments d'ordre supérieurs, si bien qu'il est difficile d'imaginer un monde dans lequel la statistique n'utilise pas les moments. Mais comment faire lorsque ces moyennes n'existent pas ?

Dans ce chapitre nous allons présenter quelques définitions et résultats de base concernant les distributions à queues lourdes, pour cela on a besoin de connaître quelques résultats sur la théorie de probabilité (la convergence stochastique et la loi des grand nombres, le théorème central limite classique, les lois infiniment divisibles, la stabilité et le domaine d'attraction, les variables sous-exponentielle, ...) et quelques outils d'analyse mathématiques (les variations régulières, la convergence uniforme ...).

### 1.1 Convergence stochastique

Une suite de variables aléatoires  $X = (X_0(\cdot), X_1(\cdot), \dots)$  (notée aussi  $(X_n, n \in \mathbb{N})$ ) définie sur l'espace de probabilité  $(\Omega, F, P)$  est une application probabilisable de  $\Omega$  dans  $\mathbb{R}^{\mathbb{N}}$  qui pour tout issu  $\omega$  associe une suite de nombres réels  $X = (X_0(\omega), X_1(\omega), \dots)$ . Comme

pour le cas de suite réelles, il est donc très important de connaître le comportement des suites  $X(\omega) = (X_n, n \in \mathbb{N})$  lorsque  $n$  croît indéfiniment. Cependant la question n'est pas aussi simple que pour les suites réelles, puisque dans le cas de suite de variables aléatoires il existe plusieurs façons (ou modes) selon lesquelles une suite de variables aléatoires  $X$  peut s'approcher ou tendre vers une limite  $X$ . Un premier critère est celui de convergence d'une suite  $X = (X_0(\omega), X_1(\omega), \dots)$  pour un  $\omega \in \Omega$  donné.

### 1.1.1 Convergence en un issu

**Définition 1.1.1** Une suite de variables aléatoires  $(X_n, n \in \mathbb{N})$  définie sur  $(\Omega, F, P)$  est dite avoir une limite (ou convergente) en un point  $\omega \in \Omega$  si la suite de nombres réels  $X(\omega) = (X_n, n \in \mathbb{N})$  converge vers une limite  $X(\omega)$ .

Mais cette définition n'est pas utile en soi puisqu'on ne s'intéresse pas au comportement d'une suite en un résultat spécifique  $\omega$  de l'expérience. Cependant elle peut servir pour d'autres concepts de limite plus importante.

Il est utile de noter qu'une suite de variables aléatoires étant une suite de fonctions sur  $\Omega$ , on peut en définir le concept de convergence simple.

### 1.1.2 Convergence simple d'une suite de variables aléatoires

**Définition 1.1.2** On dit qu'une suite de variables aléatoires  $(X_n, n \in \mathbb{N})$  définie sur  $(\Omega, F, P)$  converge simplement vers une variable aléatoire  $X$  définie sur  $(\Omega, F, P)$ , si pour tout  $\omega \in \Omega$ ,

$$\lim_{n \rightarrow \infty} X_n(\omega) = X(\omega).$$

autrement dit,

$$\forall \omega \in \Omega, \forall \varepsilon > 0, \exists N(\omega, \varepsilon) > 0 : \forall n \geq N(\omega, \varepsilon), |X_n(\omega) - X(\omega)| < \varepsilon.$$

Une définition de convergence moins forte que la convergence simple mais qui reste quand même assez forte est le concept de convergence presque sûre.

### 1.1.3 Convergence presque sûre

**Définition 1.1.3** On dit qu'une suite de variables aléatoires  $(X_n, n \in \mathbb{N})$  définie sur  $(\Omega, F, P)$  converge presque sûrement vers une variable aléatoire  $X$  définie sur  $(\Omega, F, P)$  et on écrit  $X_n \xrightarrow[n \rightarrow \infty]{p.s.} X$ , si son ensemble de convergence est de probabilité 1. Autrement dit, si elle converge sauf dans un ensemble  $E \in F$  de probabilité nulle. On écrit alors :

$$p\left(\omega \in \Omega : \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\right) = 1.$$

On peut exhiber des exemples dans lesquels la limite  $\lim_{n \rightarrow \infty} X_n(\omega)$  ne converge nulle part (pour tout  $\omega$ ) mais que  $p(\omega : |X_n(\omega) - X(\omega)| \geq \varepsilon)$  tend à être petite pour tout  $\varepsilon > 0$ .

Ceci conduit à une définition plus faible de convergence : la convergence en probabilité.

### 1.1.4 Convergence en probabilité

**Définition 1.1.4** On dit qu'une suite de variables aléatoires  $(X_n, n \in \mathbb{N})$  définie sur  $(\Omega, F, P)$  converge en probabilité vers une variable aléatoire  $X$  définie sur  $(\Omega, F, P)$  et on écrit  $X_n \xrightarrow[n \rightarrow \infty]{p} X$ , si :

$$\forall \varepsilon > 0 : \lim_{n \rightarrow \infty} p(|X_n - X| \geq \varepsilon) = 0.$$

Un autre mode de convergence très répandu et qui n'implique pas directement les valeurs de  $X_n$  mais plutôt leurs distributions est connu sous le nom convergence en distribution.

### 1.1.5 Convergence en distribution

**Définition 1.1.5** On dit qu'une suite de variables aléatoires  $(X_n, n \in \mathbb{N})$  définie sur  $(\Omega, F, P)$  converge en distribution vers une variable aléatoire  $X$  définie sur  $(\Omega, F, P)$  et on écrit  $X_n \xrightarrow[n \rightarrow \infty]{D} X$ , si :

$$\lim_{n \rightarrow \infty} F_n(x) = F(x), \forall x \in C_F,$$

avec  $C_F = \{x \in \mathbb{R} : F \text{ est continue en } x\}$ .

Pour tout  $x$  où  $F$  est continue, où  $F_n(\cdot)$  et  $F(\cdot)$  sont respectivement les fonctions de répartition de  $X_n$  et  $X$ .

Plusieurs concept de limite sont basés sur la notion de distance. Un cas particulier dans le cas de convergence stochastique est le suivant.

### 1.1.6 Convergence en moyenne quadratique

**Définition 1.1.6** On dit qu'une suite de variables aléatoires  $(X_n, n \in \mathbb{N})$  définie sur  $(\Omega, F, P)$  converge en moyenne quadratique vers une variable aléatoire  $X$  définie sur  $(\Omega, F, P)$  et on écrit  $X_n \xrightarrow[n \rightarrow \infty]{m.q} X$ , si :

$$\lim_{n \rightarrow \infty} E(|X_n - X|^2) = 0.$$

En général, lorsque pour  $r \geq 1$ , on a  $\lim_{n \rightarrow \infty} E(|X_n - X|^r) = 0$ , alors on dit que  $(X_n, n \in \mathbb{N})$  converge en moyenne d'ordre  $r$  vers  $X$  et on écrit  $X_n \xrightarrow[n \rightarrow \infty]{r} X$ .

### 1.1.7 Relation entre modes de convergence

Les résultats suivants font apparaître la relation entre les différents modes étudiés.

#### Théorème 1.1.1

i) La convergence presque sûrement implique la convergence en probabilité

$$X_n \xrightarrow[n \rightarrow \infty]{p.s} X \text{ implique } X_n \xrightarrow[n \rightarrow \infty]{p} X.$$

ii) La convergence en moyenne quadratique implique la convergence en probabilité

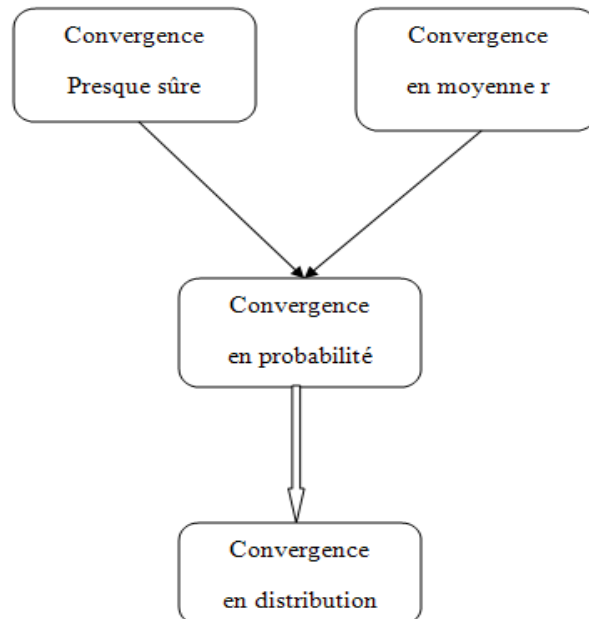
$$X_n \xrightarrow[n \rightarrow \infty]{m.q} X \text{ implique } X_n \xrightarrow[n \rightarrow \infty]{p} X.$$

iii) La convergence en probabilité implique la convergence en loi

$$X_n \xrightarrow[n \rightarrow \infty]{p} X \text{ implique } X_n \xrightarrow[n \rightarrow \infty]{D} X.$$

**Remarque**

Il n'existe pas de relation entre la convergence presque sûre et la convergence en moyenne quadratique, sauf si l'on considère d'autres hypothèses sur la suite en question.



## 1.2 Loi des grands nombres dans le cas iid

Il existe autant de versions de lois des grands nombres qu'il y ait de modes de convergence et de structure de dépendance de suite de variables aléatoires. Dans le cas des variables aléatoires iid, il existe deux lois des grands nombres : forte et faible.

### 1.2.1 Loi faible des grands nombres

#### Théorème 1.2.1

Soit  $(X_n, n \in \mathbb{N})$  une suite de variables aléatoires indépendantes et identiquement distribuées (iid) définies sur  $(\Omega, \mathcal{F}, P)$  à valeurs dans  $\mathbb{R}$ . On suppose que  $\mu = E(X_n) \leq \infty$ . Alors, la suite de sommes partielles  $(S_n, n \in \mathbb{N})$  définie par  $S_n = X_1 + X_2 + \cdots + X_n$  converge en probabilité vers  $\mu$ . Autrement dit,  $\frac{S_n}{n} \xrightarrow[n \rightarrow \infty]{P} \mu$ .

**Remarque 1.2.1**

Pour la démonstration de ce théorème dans le cas de variable aléatoire de variance finie on utilise l'application de l'inégalité de Tchebychev. Dans le cas d'une variable aléatoire de variance infinie, la démonstration est plus délicate et se base sur la loi forte des grands nombres.

**1.2.2 Loi forte des grands nombres****Théorème 1.2.2**

Soit  $(X_n, n \in \mathbb{N})$  une suite de variables aléatoires indépendantes et identiquement distribuées (iid) définies sur  $(\Omega, F, P)$  à valeurs dans  $\mathbb{R}$ . On suppose que  $\mu = E(X_n) \leq \infty$ . Alors, la suite de sommes partielles  $(S_n, n \in \mathbb{N})$  définie par  $S_n = X_1 + X_2 + \dots + X_n$  converge presque sûrement vers  $\mu$ . Autrement dit,  $\frac{S_n}{n} \xrightarrow[n \rightarrow \infty]{p.s.} \mu$ .

**1.3 Théorème central limite classique**

Pour appliquer ce théorème il faut que la somme des variables aléatoires soit distribuée asymptotiquement suivant une loi normale. Il existe en fait plusieurs versions du théorème central limite en fonction de la structure de dépendance de la suite de variables aléatoires. (ici on ne considère les suites indépendantes) mais aussi de la distribution de ses termes. On considère d'abord le cas le plus simple de suite de variables aléatoires iid.

**1.3.1 Cas des variables aléatoires indépendantes et identiquement distribuées (iid)****Théorème 1.3.1**

Soit  $(X_n, n \in \mathbb{N})$  une suite de variables aléatoires indépendantes et identiquement distribuées (iid) définies sur  $(\Omega, F, P)$  à valeurs dans  $\mathbb{R}$ . On suppose que  $\mu = E(X_n) \leq \infty$  et

que  $\sigma^2 = E(X_n^2) < \infty$ . Alors,

$$\frac{(X_1 + X_2 + \cdots + X_n) - n\mu}{\sigma\sqrt{n}} \xrightarrow[n \rightarrow \infty]{D} N(0, 1).$$

**Remarque** La preuve la plus simple de ce théorème se base sur le théorème de Levy.

**Théorème 1.3.2** (Levy)

$X_n \xrightarrow[n \rightarrow \infty]{D} X$  si et seulement si  $\varphi_{X_n}(t) = E(e^{itX_n}) \longrightarrow \varphi_X(t) = E(e^{itX})$  pour tout  $t \in \mathbb{R}$ .

### 1.3.2 Cas des variables aléatoires indépendantes ( non nécessairement (iid))

Si les termes de la suite de variables aléatoires ne sont pas identiquement distribués mais toujours indépendants, les termes de la somme  $S_n$  peuvent ne pas jouer le même rôle et la convergence en loi si elle a lieu peut ne pas être vers la loi normale.

Pour éviter que la somme ne converge vers une loi normale des conditions ont été données (d'abord par Lyapunov, puis par Lindeberg, Levy et Feller) assurant la « petitesse uniforme » des termes centrés (par leurs moyennes respectives) et réduits (par la somme des variances).

Formellement soit  $(X_n, n \in \mathbb{N})$  une suite de variables aléatoires indépendantes définies sur  $(\Omega, F, P)$  à valeurs dans  $\mathbb{R}$ . On suppose que  $\mu = E(X_n) \leq \infty$  et que  $\sigma_n^2 = E(X_n^2) < \infty$

et on pose  $S_n^2 = \sum_{k=1}^n \sigma_k^2$ . Une condition de petitesse uniforme peut être exprimée par le fait

que les variances des termes sont petites par rapport à la somme des variances, i.e.

$$\lim_{n \rightarrow \infty} \max_{1 \leq k \leq n} \frac{\sigma_k}{S_n} = 0.$$

Malheureusement cette condition n'est pas suffisante à la réalisation du théorème central limite. Une première condition assurant la « petitesse uniforme » des termes sommés est la condition de Lyapunov,

$$\lim_{n \rightarrow \infty} \frac{1}{S_n^{2+\delta}} \sum_{k=1}^n E(|X_k - \mu_k|^{2+\delta}) = 0.$$

pour un certain  $\delta > 0$  (dans certain ouvrage on énonce le théorème pour  $\delta = 1$  mais ça semble restrictif). Sous cette condition Lyapunov démontre le résultat suivant :

### 1.3.3 Théorème central limite de Lyapunov

#### Théorème 1.3.3 (Lyapunov)

S'il existe  $\delta > 0$  tel que :

$$\lim_{n \rightarrow \infty} \frac{1}{S_n^{2+\delta}} \sum_{k=1}^n E \left( |X_k - \mu_k|^{2+\delta} \right) = 0,$$

alors :

$$\frac{(X_1 - \mu_1) + \cdots + (X_n - \mu_n)}{S_n} \xrightarrow[n \rightarrow \infty]{D} N(0, 1).$$

### 1.3.4 Théorème central limite de Lindeberg

#### Théorème 1.3.4 (Lindeberg)

Si

$$\forall \varepsilon, \lim_{n \rightarrow \infty} \frac{1}{S_n^2} \sum_{k=1}^n E \left[ (|X_k - \mu_k|) 1_{|X_k - \mu_k| > \varepsilon S_n} \right] = 0,$$

alors :

$$\frac{(X_1 - \mu_1) + \cdots + (X_n - \mu_n)}{S_n} \xrightarrow[n \rightarrow \infty]{D} N(0, 1).$$

On vérifie que la condition de Lindeberg garantie la condition de petitesse uniforme :

$$\lim_{n \rightarrow \infty} \max_{1 \leq k \leq n} \frac{\sigma_k}{S_n} = 0.$$

## 1.4 Lois infiniment divisibles

Pour inclure la convergence de Poisson et  $N(0, 1)$  on doit poser la question suivante :

Soit :

$$S_n = X_1^{(n)} + X_2^{(n)} + \cdots + X_n^{(n)}, \quad (1.4.1)$$

où les  $X_k^{(n)}$ ,  $k = 1, \dots, n$  sont indépendantes et identiquement distribuées.

Si  $\frac{S_n}{n} \xrightarrow[n \rightarrow \infty]{D} X$ , quelles sont les distributions possibles de  $X$  ?

$S_n$  est la somme de plusieurs composantes indépendantes, heuristiquement  $X$  doit avoir cette même propriété.

### 1.4.1 Variable aléatoire infiniment divisible

**Définition 1.4.1** On dira qu'une variable aléatoire  $X$  a une distribution infiniment divisible si pour chaque  $n$ , il y a des variables aléatoires indépendantes et identiquement distribuées  $X_1^{(n)}, X_2^{(n)}, \dots, X_n^{(n)}$  telles que :

$$L(X) = L\left(X_1^{(n)} + X_2^{(n)} + \cdots + X_n^{(n)}\right). \quad (1.4.2)$$

#### Propositions 1.4.1

a) Une variable aléatoire  $X$  est une limite en distribution des sommes du type (1.4.1) si et seulement si elle a une distribution infiniment divisible.

b) Si  $\frac{S_n}{n} \xrightarrow[n \rightarrow \infty]{D} X$ , alors  $X_1^{(n)} \xrightarrow[n \rightarrow \infty]{D} 0$ . Mais le problème est : comment caractériser les distributions infiniment divisibles?

Soit  $f(u)$  la fonction caractéristique de  $X$ .

Puisque  $L(X) = L\left(X_1^{(n)} + X_2^{(n)} + \cdots + X_n^{(n)}\right)$ , il ya une fonction caractéristique tel que :  $f_n(u) = [f_n(u)]^n$ , et par (1.4.1) on a  $f_n(u) \xrightarrow{uc} 1$  (cette notation signifie la convergence uniforme sur les ensembles compacts).

$$\log f(u) = n \log [1 - (1 - f_n(u))] = n(f_n(u) - 1)(1 + \varepsilon_n(u)). \quad (1.4.3)$$

$$\log f(u) = (1 + \varepsilon_n(u)) \int (e^{iux} - 1) ndF_n. \quad (1.4.4)$$

$$\overline{\lim} u_n [-a, +a]^c \leq \alpha a \int_0^{1/\alpha} |Rl \log f(v)| dv. \quad (1.4.5)$$

$$\overline{\lim} \int_{[-1, +1]} x^2 u_n(dx) < \infty. \quad (1.4.6)$$

$$\log f(u) = (1 + \varepsilon_n) \alpha_n \int \left( e^{iux} - 1 - \frac{iux}{1 + x^2} \right) \frac{1 + x^2}{x^2} G_n(dx) + i(1 + \varepsilon_n) \beta_n u. \quad (1.4.7)$$

$$\log f(u) = i\beta u - \frac{\sigma^2 u^2}{2} + \alpha \int_{\{0\}^c} \varphi(x, u) G(dx). \quad (1.4.8)$$

**Théorème 1.4.1** (e.g Breiman, 1968)

Une variable aléatoire  $X$  est de distribution infiniment divisible si et seulement si sa fonction caractéristique  $f(u)$  est donné par :

$$\log f(u) = i\beta u - \frac{\sigma^2 u^2}{2} + \int \left( e^{iux} - 1 - \frac{iux}{1 + x^2} \right) \frac{1 + x^2}{x^2} v(dx), \quad (1.4.9)$$

où  $v$  est une mesure finie.

## 1.5 La stabilité et le domaine d'attraction

### 1.5.1 Lois stables

Soient  $X_1, X_2, \dots$  des variables aléatoires non dégénérées indépendantes identiquement distribuées. Quelle est la classe de toutes les lois limites possibles des sommes normées?

$$S_n = \frac{X_1 + \dots + X_n}{A_n} - B_n, A_n > 0? \quad (1.5.1)$$

Comme on peut écrire  $S_n$  sous la forme suivante :  $S_n = X_1^{(n)} + X_2^{(n)} + \dots + X_n^{(n)}$  où  $X_k^{(n)} = \frac{X_k}{A_n} - \frac{B_n}{n}$ .

La condition  $S_n \xrightarrow[n \rightarrow \infty]{D} X$  implique que  $X$  est infiniment divisible. La condition  $X_1^{(n)} \xrightarrow{D} 0$  implique que  $A_n \rightarrow \infty, \frac{B_n}{n} \rightarrow 0$ .

Cette classe des lois limites est la plus intéressante de l'ensemble des distributions suivant la loi normale et Poisson. Naturellement, si  $E(X_1^2) < \infty$ , donc  $A_n \rightarrow \sqrt{n}$ , dans ce cas  $X$  doit être normale. Ainsi le seul cas intéressant est  $E(X_1^2) = \infty$ . Deux questions importantes se posent :

La Première : quelle est la forme de toutes les distributions limites de  $X$  telles que  $S_n \xrightarrow{D} X$  ?

La Deuxième : Trouvez les conditions nécessaires et suffisantes de la fonction de distribution conjointe de  $X_1, X_2, \dots$  de sorte que  $S_n \xrightarrow{D} X$  ?

Ces deux questions nous amènent à des lois stables et les domaines d'attractions des lois stables.

#### Définition 1.5.1 (Loi stable)

On dit qu'une variable aléatoire  $X$  a une loi stable si pour chaque nombre entier  $K > 0$ , et  $X_1, X_2, \dots, X_k$  indépendants de même distribution que  $X$ , il existe des constantes  $a_k > 0, b_k$  telles que :

$$L(X_1 + X_2 + \dots + X_k) = L(a_k X + b_k). \quad (1.5.2)$$

Cette approche est similaire à la façon dont nous caractérisons intrinsèquement la loi conjointe, en divisant en blocs  $K$ , nous avons conclu que la limite de  $\frac{S_n}{\sqrt{n}}$  doit satisfaire :

$$L(X_1 + X_2 + \dots + X_k) = L(\sqrt{k}X). \quad (1.5.3)$$

**Proposition 1.5.1**

$X$  représente la distribution limite des sommes normées de la représentation (1.5.2) si et seulement si  $X$  a une loi stable.

**1.5.2 Forme des lois stables**

**Théorème 1.5.1** (e.g Breiman, 1968)

Soit  $X$  une variable aléatoire de loi stable et de distribution normale où il existe un nombre  $0 < \alpha < 2$ , appelé l'exposant de la loi et les constantes  $m_1 \geq 0, m_2 \geq 0, \beta$  tels que :

$$\log f_x(u) = iu\beta - m_1 \int_0^\infty \left( e^{iu\beta} - 1 - \frac{iu\beta}{1+x^2} \right) \frac{dx}{x^{1+\alpha}} + m_2 \int_{-\infty}^0 \left( e^{iu\beta} - 1 - \frac{iu\beta}{1+x^2} \right) \frac{dx}{|x|^{1+\alpha}}. \quad (1.5.4)$$

**Proposition 1.5.2**

$a_k = k^\lambda, \lambda > 0$ , et :

$$M^+(x) = x^{-1/\lambda} M^+(1) \text{ et } M^-(x) = |x|^{-1/\lambda} M^-(1). \quad (1.5.5)$$

**Remarque** La preuve de cette proposition se trouve dans Breiman, (1968).

**1.5.3 Calcul de la fonction caractéristique stable**

Lorsque l'exposant  $\alpha$  est inférieur à 2, la forme de la fonction caractéristique stable est donnée dans le théorème précédent. En faisant quelques calculs, on peut évaluer ces intégrales sous forme explicite.

**Théorème 1.5.2** (e.g Breiman, 1968)

On dit que  $f(u) = e^{\varphi(u)}$  est une fonction caractéristique d'une loi stable d'exposant  $\alpha$ , avec  $0 < \alpha < 1$  et  $1 < \alpha < 2$  si et seulement si  $\varphi(u)$  de la forme :

$$\varphi(u) = iuc - d|u| \left( 1 + i\theta \frac{u}{|u|} \frac{2}{\pi} \tan \frac{\pi}{2} \alpha \right), \quad (1.5.6)$$

où  $C$  est réel,  $d$  réel et positif, et  $\theta$  réel, tel que  $|\theta| \leq 1$ . Pour  $\alpha = 1$ , la forme de la fonction caractéristique est donnée par :

$$\varphi(u) = iuc - d|u| \left( 1 + i\theta \frac{u}{|u|} \frac{2}{\pi} \log \frac{\pi}{2} \alpha \right), \quad (1.5.7)$$

avec  $C, d, \theta$  comme ci-dessus.

#### 1.5.4 Domaine d'attraction d'une loi stable

Soient  $X_1, X_2, \dots$  des variables aléatoires indépendantes et identiquement distribuées. Quelles sont les conditions nécessaires et suffisantes sur leur fonction de distribution  $F(x)$  telle que  $F$  convenablement normalisée converge vers la distribution de  $X$ , où  $X$  est non dégénérée ? bien entendu, la limite d'une variable aléatoire  $X$  doit être stable.

**Définition 1.5.2** La distribution  $F(x)$  est dite dans le domaine d'attraction d'une loi stable avec l'exposant  $\alpha < 2$  s'il y a des constantes  $A_n, B_n$  tels que :

$$\frac{S_n}{A_n} - B_n \xrightarrow{D} X, \quad (1.5.9)$$

et  $X$  a l'exposant  $\alpha$ . On note ceci par  $F \in MAD(\alpha)$ .

Les conditions complètes sur  $F(x)$  sont donnés par :

**Théorème 1.5.3** (e.g Breiman, 1968)

On dit que  $F(x)$  est dans le domaine d'attraction d'une loi stable avec un exposant  $\alpha < 2$  si et seulement si il existe des constantes  $M^+, M^- \geq 0, M^+ + M^- > 0$  telles que quand  $y \rightarrow \infty$  :

a)  $\lim \frac{F(-y)}{1-F(y)} = \frac{M^+}{M^-}$ .

b) pour chaque  $\zeta > 0$  :  $M^+ > 0 \Rightarrow \lim \frac{1-F(\zeta y)}{1-F(y)} = \frac{1}{\zeta^\alpha}$  et  $M^- > 0 \Rightarrow \lim \frac{F(-\zeta y)}{F(-y)} = \frac{1}{\zeta^\alpha}$ .

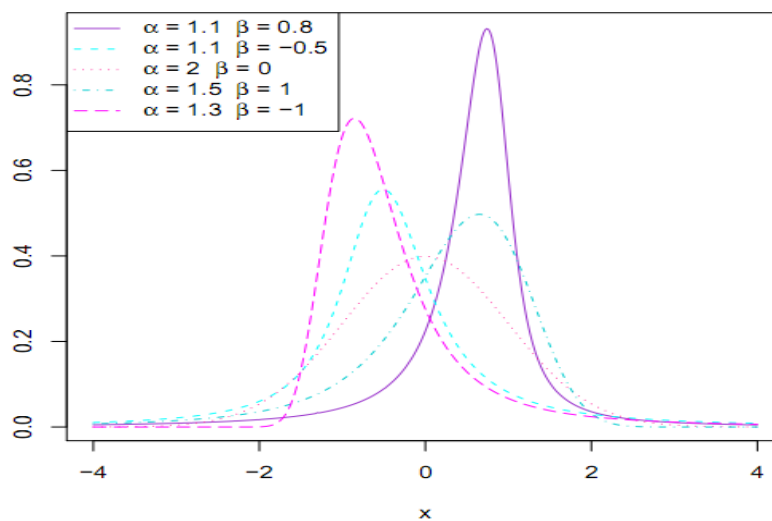


Figure lois  $\alpha$ -stables

L'analyse des phénomènes à queues lourdes est aussi basée sur des outils d'analyse mathématique, parmi ces outils on a : la variation régulière et la convergence uniforme . . . etc.

## 1.6 Convergence uniforme

### 1.6.1 Convergence d'une fonction

**Définition 1.6.1** Soit  $\{f_n, n \geq 0\}$  des fonctions à valeurs réelles dans  $\mathbb{R}$  (ou, dans n'importe quel espace métrique), alors  $f_n$  converge uniformément sur  $A \subset \mathbb{R}$  vers  $f_0$

si seulement si :

$$\sup_{x \in A} |f_0(x) - f_n(x)| \rightarrow 0, \text{ si } n \rightarrow \infty. \quad (1.6.1)$$

Cette définition serait toujours dans ce sens si les fonctions  $\{f_n, n \geq 0\}$ , sont définies dans un espace métrique, mais  $|f_0(x) - f_n(x)|$  devrait être remplacée par  $d(f_0(x), f_n(x))$ , où  $d(., .)$  est le système métrique. Pour les fonctions définies sur  $\mathbb{R}$ , la convergence uniforme signifie que (1.6.1) est vraie pour tout les intervalles compacts  $A$  avec  $(A \subset \mathbb{R})$ . Pour les fonctions monotones qui convergent simplement vers une limite continue donc sont convergentes localement uniformément.

#### Proposition 1.6.1

Supposons que  $\{U_n, n \geq 0\}$  sont des fonctions à valeurs réelles dans  $\mathbb{R}$  non décroissantes et  $U_0$  continue.

Si, pour tout  $x$  on a :

$$U_n(x) \rightarrow U_0(x), \text{ si } n \rightarrow \infty, \quad (1.6.2)$$

alors  $U_n \rightarrow U_0$  localement uniformément ; c'est-à-dire, pour tout  $a < b$  on a :

$$\sup_{x \in [a, b]} |U_n(x) - U_0(x)| \rightarrow 0.$$

## 1.7 Variations régulières

Dans les domaines d'application mathématiques nous observons le comportement de la loi de puissance. Les lois de puissance se produisent généralement sous la forme perturbée. Pour décrire la forme écart des lois de puissance pure, la notion de variation régulière a été introduite.

**Définition 1.7.1** Une fonction mesurable positive  $f$  est dite variée régulièrement avec l'indice  $\alpha \in \mathbb{R}$  si :

- Elle est définie sur l'intervalle  $[x_0, \infty[$  ou voisinage de l'infini.

et :

$$\lim_{x \rightarrow \infty} \frac{f(tx)}{f(x)} = t^\alpha \text{ pour tout } t > 0. \quad (1.7.1)$$

### Remarque 1.7.1

a) la définition de la variation régulière peut être étendue de plusieurs manières par exemple, il suffit que la limite dans l'expression (1.7.1) existe, positive et finie pour tout  $t > 0$ . Alors la fonction limite  $\chi$  satisfait la relation  $\chi(ts) = \chi(t) \cdot \chi(s)$  qui implique que  $\chi$  est une fonction de puissance.

b) il est facile de voir que chaque fonction régulièrement variable  $f$  de l'indice  $\alpha$  représentées sous la forme suivante :  $f(x) = x^\alpha L(x)$ , dans ce cas  $L$  est une fonction qui varie lentement.

c) La variation régulière d'une fonction  $f$  peut être définie sur un point quelconque  $x_0 \in \mathbb{R}$  mais en exigeant que  $f(x_0 - x^{-1})$  varie régulièrement à l'infini. Dans ce qui suit nous traitons les cas où la variation régulière est à l'infini.

### Exemple

Pour tout réel  $\alpha$  les fonctions :  $x^\alpha, x^\alpha \ln(1+x), (x \ln(1+x))^\alpha, x^\alpha \ln(\ln(e+x))$  sont des fonctions qui varient régulièrement à  $\infty$  avec l'indice  $\alpha$ . Mais les fonctions :  $2 + \sin x, e^{\ln(1+x)}$  ne varient pas régulièrement.

Dans les applications de probabilité, nous sommes concernés par les distributions dont les queues varient régulièrement par exemples :

$$1 - F(x) = x^{-\alpha}, x \geq 1, \alpha > 0.$$

Et la distribution de la valeur extrême est :  $\Phi_\alpha(x) = \exp\{-x^{-\alpha}\}, x \geq 0$ .

Avec  $\Phi_\alpha(x)$  vérifie la propriété suivante :  $1 - \Phi_\alpha(x) \sim x^{-\alpha}$  si  $x \rightarrow \infty$ .

La densité de Cauchy  $f(x) = \frac{1}{\pi(1+x^2)}$  et une fonction de répartition  $F$  avec la propriété suivante :  $1 - F(x) \sim (\pi x)^{-1}$ .

Si  $N(x)$  est la fonction de distribution normale standard, puis  $1 - N(x)$  n'est pas régulièrement variant est la queue de la valeur extrême Gumbel est de distribution  $1 - \exp(-e^{-x})$ .

### 1.7.1 Propriétés de base

Pour donner une définition claire sur les fonctions à variations régulières, on doit présenter dans cette section quelques propriétés très importantes sur ces fonctions.

**Théorème 1.7.1** (e.g Mikosch, T, 1999)

On dit qu'une fonction positive mesurable  $L$  varie lentement sur  $[x_0, \infty[$  si  $L$  est de la forme suivante :

$$L(x) = c(x) \exp \left\{ \int_{x_0}^x \frac{\varepsilon(y)}{y} dy \right\}, \quad (1.7.2)$$

avec  $c(\cdot)$  est une fonction mesurable non négative telle que  $\lim_{x \rightarrow \infty} c(x) = c_0 \in [0, \infty[$  et  $c(x) \rightarrow \infty$  si  $x \rightarrow \infty$ .

**Remarque 1.7.2**

a) D'après le théorème précédent on dit que  $f$  est une fonction régulièrement variable avec l'indice  $\alpha$  si  $f$  est de représentation :

$$f(x) = x^\alpha c(x) \exp \left\{ \int_{x_0}^x \frac{\varepsilon(y)}{y} dy \right\},$$

avec  $c(\cdot)$  et  $\varepsilon(\cdot)$  sont comme ci-dessus.

b) D'après la représentation du théorème précédent nous pouvons conclure que la fonction  $f$  varie régulièrement avec l'indice  $\alpha \neq 0$  pour tout  $x \rightarrow \infty$  comme suit :

$$f(x) = \begin{cases} \infty & \text{si } \alpha > 0 \\ 0 & \text{si } \alpha < 0 \end{cases} ,$$

de plus, si  $L$  varie lentement alors pour tout  $\delta > 0$  on a :  $x^{-\delta}L(x) \rightarrow 0$  et  $x^\delta L(x) \rightarrow \infty$  pour  $x \rightarrow \infty$ .

La dernière propriété donne une certaine signification intuitive à la notion " de variation lente".

Un résultat important est le fait que la convergence dans (1.5.1) est uniforme sur chaque sous-ensemble compact  $]0, \infty[$ .

## 1.7.2 Théorème de convergence uniforme pour les fonctions à variations régulières

**Théorème 1.7.2** (e.g Mikosch,T,1999)

Si  $f$  est une fonction à variation régulière avec l'indice  $\alpha$  (dans le cas  $\alpha > 0$ , en supposant que  $f$  est bornée pour chaque intervalle  $]0, x]$ ,  $x > 0$ ), puis pour  $0 < a \leq b < \infty$  :

$\lim_{x \rightarrow \infty} \frac{f(tx)}{f(x)} = t^\alpha$ , uniformément en  $t$  :

- a) Sur chaque  $[a, b]$  si  $\alpha = 0$ .
- b) Sur chaque  $]0, b]$  si  $\alpha > 0$ .
- c) Sur chaque  $]0, \infty[$  si  $\alpha < 0$ .

Dans ce qui suit, pour toutes fonctions positives  $f$  et  $g$  :  $f(x) \sim g(x)$  pour tout  $x \rightarrow x_1$  signifie que  $\lim_{x \rightarrow x_1} \frac{f(x)}{g(x)} = 1$ .

Dans les applications la question suivante est importante. Supposons que  $f$  est une fonction à variation régulière avec l'indice  $\alpha$ . Comment trouver une autre fonction à variation régulière  $f_1$  avec le même indice de sorte que  $f(x) \sim f_1(x)$  pour  $x \rightarrow \infty$ ? Dans la représentation (1.7.2) nous avons une certaine flexibilité pour construire les deux fonctions  $c$  et

$\varepsilon$ . En prenant la fonction  $c$  par exemple, nous avons déjà une réponse positive à la question ci-dessus.

### Proposition 1.7.1

Supposons que  $L$  varie lentement, alors il existe une autre fonction  $L_1 \in C^\infty$  ( $C^\infty$  est l'espace des fonctions infiniment différentiables) qui, elle aussi varie lentement de sorte que  $L(x) \sim L_1(x)$  pour  $x \rightarrow \infty$ . Finalement si  $L$  est monotone,  $L_1$  est aussi monotone.

Le résultat suivant du Karamata est applicable. Il indique essentiellement que les intégrales des fonctions régulièrement variables varient encore régulièrement, ou plus avec précision, on peut prendre la fonction lentement variable hors de l'intégrale.

### 1.7.3 Théorème de Karamata

#### Théorème 1.7.3 (Karamata)

Soit  $L$  est une fonction qui varie lentement et bornée sur  $[x_0, \infty)$  pour  $x_0 \geq 0$ . Alors

a) Pour  $\alpha > -1$  on a :

$$\int_{x_0}^x t^\alpha L(t) dt \sim (\alpha + 1)^{-1} x^{\alpha+1} L(x), x \rightarrow \infty.$$

b) Pour  $\alpha < -1$  on a :

$$\int_{x_0}^x t^\alpha L(t) dt \sim -(\alpha + 1)^{-1} x^{\alpha+1} L(x), x \rightarrow \infty.$$

#### Remarque 1.7.3

1) le résultat reste toujours vrai pour  $\alpha = -1$  dans le sens que :  $\frac{1}{L(x)} \int_{x_0}^x \frac{L(t)}{t} dt \rightarrow \infty, x \rightarrow \infty$  et  $\int_{x_0}^x (L(t)/t) dt$  varie lentement. Si  $\int_{x_0}^x (L(t)/t) dt < \infty$  alors  $\frac{1}{L(x)} \int_{x_0}^x \frac{L(t)}{t} dt \rightarrow \infty, x \rightarrow \infty$  et  $\int_{x_0}^x (L(t)/t) dt$  varie lentement.

2) les conclusions du théorème de Karamata peuvent également être formulées comme suit : supposons que  $f$  varie régulièrement avec l'indice  $\alpha$  et  $f$  est localement bornée sur  $[x_0, \infty)$  pour  $x_0 \geq 0$  alors :

a) pour  $\alpha > -1$  on a :

$$\lim_{x \rightarrow \infty} \frac{\int_{x_0}^x f(t) dt}{x f(x)} = \frac{1}{\alpha + 1}.$$

b) pour  $\alpha < -1$  on a :

$$\lim_{x \rightarrow \infty} \frac{\int_{x_0}^x f(t) dt}{x f(x)} = -\frac{1}{\alpha + 1}.$$

## 1.8 Variables aléatoires à variations régulières

La notion de variation régulière apparaît d'une manière naturelle dans plusieurs domaines d'applications dans (la probabilité appliquée, la théorie des files d'attente, la théorie des valeurs extrêmes, la théorie de renouvellement, la théorie de processus ponctuel...). Dans ces domaines, les variables aléatoires (et plus généralement, vecteurs aléatoires) avec les queues à variation régulières ont beaucoup d'applications.

Dans ce qui suit, on note par :  $F_X(x) = p(X \leq x)$ ,  $x \in \mathbb{R}$ , pour la fonction de répartition de  $X$  et par :  $\bar{F}_X(x) = p(X > x)$ ,  $x \in \mathbb{R}$ , pour la fonction de distribution de la queue de  $X$ .

### 1.8.1 Distribution d'une variable aléatoire à variations régulières

**Définition 1.8.1** On dit qu'une variable aléatoire non négative  $X$  est de fonction de distribution à variations régulières pour chaque indice  $\alpha \geq 0$ , si sa fonction de distribution de queue à droite  $\bar{F}_X$  varie régulièrement pour chaque indice  $-\alpha$ .

Pour clarifier les choses de plus en plus, on doit faire un rappel sur les propriétés de base de distributions à variations régulières.

#### **Proposition 1.8.1 (Distribution à variation régulière)**

Supposons que  $X$  est une variable aléatoire de fonction de répartition  $F_X$  avec  $F(x) < 1$  pour tout  $x \geq 0$ .

a) Si les suites  $(a_n)$ ,  $(x_n)$  telles que  $a_n/a_{n+1} \rightarrow 1$ ,  $x_n \rightarrow \infty$ , et si pour une fonction réelle  $g$  et tout  $\lambda$  d'un sous-ensemble de  $]0, \infty[$ ,  $\lim_{n \rightarrow \infty} a_n \bar{F}(\lambda x_n) = g(\lambda) \in ]0, \infty[$ .

alors :

$$g(\lambda) = \lambda^{-\alpha},$$

pour  $\alpha \geq 0$  et  $\bar{F}$  est à variations régulières.

b) supposons que  $F$  est absolument continue avec la densité  $f$  telle que pour

$$\alpha > 0, \lim_{n \rightarrow \infty} x f(x) / \bar{F}(x) = \alpha,$$

alors  $f$  varie régulièrement avec l'indice  $-(\alpha + 1)$  et par conséquent  $\bar{F}$  varie régulièrement avec l'indice  $-\alpha$ .

c) Supposons que la densité  $f$  de  $F$  varie régulièrement avec l'indice  $-(\alpha + 1)$  pour  $\alpha > 0$ .

Alors :

$$\lim_{n \rightarrow \infty} x f(x) / \bar{F}(x) = \alpha.$$

Ce résultat reste toujours vrai si  $\bar{F}$  varie régulièrement avec l'indice  $-\alpha$  pour  $\alpha > 0$  et la densité  $f$  est finalement monotone.

d) Si  $X$  est une variable aléatoire non négative à variations régulières avec l'indice  $\alpha > 0$ .

alors :

$$\begin{cases} EX^\beta < \infty & \text{si } \beta < \alpha \\ EX^\beta = \infty & \text{si } \beta > \alpha \end{cases} .$$

e) Si  $\bar{F}$  varie régulièrement avec l'indice  $(-\alpha)$  pour  $\alpha > 0, \beta \geq \alpha$ . Alors :

$$\lim_{n \rightarrow \infty} \frac{x^\beta \bar{F}(x)}{\int_0^x y^\beta dF(Y)} = \frac{\beta - \alpha}{\alpha}.$$

La réciproque est vraie dans le cas où  $\beta > \alpha$ . Si  $\beta = \alpha$  on peut seulement conclure que  $\bar{F}(x) = o(x^{-\alpha} L(x))$  pour  $L$  varie lentement.

f) Les mentions suivantes sont équivalentes :

1)  $\int_0^x y^\beta dF(Y)$  varie lentement.

2)  $o(x^{-2} \int_0^x y^\beta dF(Y)), x \rightarrow \infty$ .

**Propriétés 1.8.1**

L'une des raisons de la popularité des fonctions régulièrement variables dans la théorie des probabilités est la propriété élémentaire suivante : les variables aléatoires à variation régulière.

**Lemme 1.8.1**

Soit  $X$  et  $Y$  deux variables aléatoires indépendantes non négatives à variations régulières avec l'indice  $\alpha \geq 0$ . Alors  $X + Y$  est aussi à variations régulières avec l'indice et

$$p(X + Y \geq x) \sim p(X > x) + p(Y > x),$$

pour  $x \rightarrow \infty$ .

**Remarque 1.8.1**

a) Si  $X$  et  $Y$  deux variables aléatoires non négatives et pas nécessairement indépendantes telles que  $p(Y > x) = o(p(X > x))$  et  $X$  varie régulièrement avec l'indice  $\alpha$ , alors

$$p(X + Y \geq x) \sim p(X > x) \text{ pour } x \rightarrow \infty.$$

b) En particulier, si  $X$  et  $Y$  varient régulièrement avec l'indice  $\alpha_X, \alpha_Y$  et, respectivement, et si  $\alpha_X < \alpha_Y$ , alors  $X + Y$  varie régulièrement avec l'indice  $\alpha_X$ .

**Corollaire 1.8.1**

Soient  $X, X_1, \dots, X_n$  des variables aléatoires indépendantes identiquement distribuées (iid) non négatives et à variations régulières et  $S_n = X_1 + X_2 + \dots + X_n$ ,  $n \geq 1$  alors :

$$p(S_n > x) \sim np(X > x), \quad (1.8.1)$$

pour  $x \rightarrow \infty$ .

**Remarque 1.8.2**

La relation (1.8.1) a une interprétation intuitive. Soit  $M_n = \max_{i=1, \dots, n} X_i$ ,  $n \geq 1$ . Alors il est facile de voir que pour tout  $n \geq 1$ ,

$$p(S_n > x) \sim np(X > x) \sim p(M_n > x).$$

Ceci signifie que, pour un  $x$  grand on a l'événement  $\{S_n > x\}$  doit être l'événement  $\{M_n > x\}$ . La relation  $np(X > x) \sim p(M_n > x)$  pour  $x \rightarrow \infty$  implique aussi que  $M_n$  varie

régulièrement avec le même indice que  $X$ . Ceci est une autre propriété de classe des variables aléatoires à variations régulières.

Une autre conséquence immédiate du lemme précédent est la suivante :

**Corollaire 1.8.2**

Soient  $X, X_1, \dots, X_n$  des variables aléatoires indépendantes identiquement distribuées (iid) non négatives et à variations régulières avec l'indice  $\alpha$  et  $\Psi_1, \Psi_2, \dots, \Psi_n$  sont des constantes positives. Alors :

$$p(\Psi_1 X_1 + \dots + \Psi_n X_n) \sim p(X > x) (\Psi_1^\alpha + \Psi_2^\alpha + \dots + \Psi_n^\alpha).$$

Dans des conditions supplémentaires, ce résultat peut être prolongé pour les séries infinies du type suivant :  $Y = \sum_{j=0}^{\infty} \Psi_j X_j$  avec  $\{X_n, n \in \mathbb{N}\}$  est une suite de variables aléatoires indépendantes identiquement distribuées (iid) à variations régulières et  $(\Psi_j)$  sont des constantes positives. Pour assurer la convergence presque sûre de  $Y$ , il faut que les  $\{\Psi_j\}$  vérifient la condition de sommabilité dans leur ordre.

Dans les applications on doit traiter les produits des variables aléatoires où ces variables varient régulièrement.

**Proposition 1.8.2**

soient  $X$  et  $Y$  deux variables aléatoires indépendantes non négatives.

a) Supposons que  $X$  et  $Y$  varient régulièrement avec l'indice  $\alpha > 0$ . Donc  $XY$  aussi varient régulièrement avec le même indice.

b) Supposons que  $X$  varie régulièrement avec l'indice  $\alpha$  et pour tout  $EY^{\alpha+\varepsilon} < \infty$ . Donc  $XY$  varie aussi régulièrement avec l'indice  $\alpha > 0$ . de plus :

$$p(XY > x) \sim EY^\alpha p(X > x),$$

pour  $x \rightarrow \infty$ .

La variation régulière des queues des distributions apparaît comme un état naturel dans plusieurs résultats théoriques de la théorie des probabilités. Parmi ces résultats on a : les lois stables, le domaine d'attraction . . . .

## 1.9 Processus stochastiques

### 1.9.1 Introduction

La théorie des probabilités a pour but ultime la représentation mathématique des phénomènes aléatoires. En fait, on peut classer les phénomènes aléatoires en deux catégories : les phénomènes statiques et les phénomènes évolutifs (ou dynamiques). Dans le premier cas, il s'agit d'un fait qu'on admet fixe qui ne se rattache ou ne dépend d'aucun autre fait, dusse le même fait à un autre instant ou à une autre région. Le modèle mathématique de phénomènes aléatoires statiques est la variable (ou vecteur) aléatoire. Lorsqu'on veut inférer des informations concernant la distribution de probabilité des valeurs possibles d'un phénomène aléatoire statique, on peut répéter des observations identiques et indépendantes. Cependant, si l'on désire appréhender certains phénomènes comme le prix du pétrole à un instant donné, la taille d'une population à une génération ou la consommation d'électricité à une région donnée, on peut encore obtenir un ensemble d'observations prises chaque jour, chaque génération, ou chaque région, mais à ce moment là les hypothèses d'indépendance et d'équidistributivité ne sont plus admissibles. Il est plus judicieux, ici, de considérer que les observations prises à des instants ou à des générations ou encore à des régions différentes font partie d'une même famille infinie représentant tout le phénomène à tout instant, à chaque génération ou à chaque région. Le modèle de variable aléatoire s'avère ainsi insuffisant pour décrire les phénomènes qui évoluent par rapport à un ensemble renfermant une infinité de membres, parfois non dénombrable. Ainsi, on est amené à considérer plutôt une famille de variables aléatoires, évoluant sur un domaine, et dont chacune est destinée à décrire un moment correspondant de l'évolution du phénomène. Un tel modèle est appelé processus aléatoire ou stochastique et est donc caractérisé par une structure de dépendance entre membres de la famille et une distribution de probabilité des valeurs, en raison de l'aléa qui le caractérise. La théorie des processus aléatoires a pour but de donner les règles de base pour le calcul des probabilités d'évènements se rattachant plutôt à des processus aléatoires. La statistique des processus, quant à elle, permet sur la base d'une observation d'une partie du phénomène d'induire des informations concernant la distribution de probabilité d'un ou de plusieurs membres du processus ayant généré cette observation.

Soit  $(\Omega, A, P)$  un espace de probabilité, où  $\Omega$  est l'espace des réalisations possibles,  $A$  est

une tribu adaptée à (c'est l'ensemble qui contient les combinaisons possibles d'évènements) et  $P$  est une mesure de probabilité définie sur  $A$ .

Les processus aléatoires qui seront présentés dans la suite de ce travail sont définis sur l'espace de probabilité  $(\Omega, A, P)$  et à valeurs dans un espace d'états  $(\mathbb{R}, B(\mathbb{R}))$ .

## 1.9.2 Processus aléatoire

**Définition 1.9.1** Un processus aléatoire de domaine d'évolution  $T$ , défini sur un espace de probabilité  $(\Omega, A, P)$  et à valeurs dans un espace d'état  $E$  qui est muni d'une tribu  $\varepsilon$ , est une famille de variables aléatoires  $(Y_t, t \in T)$  chacune définie sur  $(\Omega, A, P)$  à valeur dans  $(E, \varepsilon)$ .

On peut classifier les processus aléatoires selon la dénombrabilité ou non de l'espace d'état  $E$  et du domaine d'évolution  $T$ . En effet, on distingue :

- Les processus à temps discret et à espace d'état discret.
- Les processus à temps continu et à espace d'état discret.
- Les processus à temps discret et à espace d'état continu.
- Les processus à temps continu et à espace d'état continu.

Dans la suite on s'intéresse au cas des processus à temps discret ( $T = \mathbb{Z}$  ou  $T = \mathbb{N}$ ) et à espace d'état fini ou infini et  $E = \mathbb{Z}$  soit  $\mathbb{R}^d$  ( $d \in \mathbb{N}^*$ ) selon le cas.

### Distribution de probabilité d'un processus aléatoire

**Distributions fini-dimensionnelles** Il est bien connu que la structure d'un processus aléatoire  $(Y_t, t \in \mathbb{Z})$  défini sur un espace de probabilité  $(\Omega, A, P)$ , à valeurs dans un espace d'états  $(\mathbb{R}, B(\mathbb{R}))$  et de domaine d'évolution  $\mathbb{Z}$  est caractérisée par la distribution infini-dimensionnelle  $P_Y(\cdot)$  défini comme suit :

$$\begin{aligned}
P_Y(\cdot) &: (B(\mathbb{R}))^{\otimes \mathbb{Z}} \rightarrow [0, 1] \\
\mathbb{B} &= \prod_{t \in T} B_t \rightarrow P_Y(\mathbb{B}) = P(\omega : Y(\omega) \in \mathbb{B}) \\
&= P\left(\omega : \bigcap_{t \in T} Y_t(\omega) \in B_t\right),
\end{aligned}$$

avec  $B_t \in B(\mathbb{R})$ . Comme la distribution infini-dimensionnelle est généralement difficile à manipuler, on fait recours à un outil plus simple qui est la distribution fini-dimensionnelle. L'introduction de cette distribution est justifiée par un théorème qui stipule que dans le cas à temps discret la distribution infini-dimensionnelle est uniquement déterminée par les probabilités fini-dimensionnelles et donc la structure probabiliste d'un processus à temps discret est entièrement déterminée par les distributions fini-dimensionnelles.

**Définition 1.9.2 (Fonction de répartition fini-dimensionnelle)**

Pour  $n \in \mathbb{N}^*$  et pour tout  $t_1, t_2, \dots, t_n \in \mathbb{Z}$ , la fonction de répartition fini-dimensionnelle de  $(Y_t, t \in \mathbb{Z})$  est la fonction de répartition de toute sous-suite finie  $(Y_{i1}, Y_{i2}, \dots, Y_{in})$ . Autrement dit, c'est la fonction  $F_{Y_{t_1}, \dots, Y_{t_n}}(\cdot)$  définie sur  $\mathbb{R}^n$  à valeurs dans  $[0, 1]$  par :

$$\begin{aligned}
F_{Y_{t_1}, \dots, Y_{t_n}}(\cdot) &: \mathbb{R}^n \rightarrow [0, 1] \\
(y_1, y_2, \dots, y_n) &\rightarrow F_{Y_{t_1}, \dots, Y_{t_n}}(y_1, y_2, \dots, y_n) = P\left(\omega : \bigcap_{j=1}^n Y_{t_j}(\omega) \leq y_j\right).
\end{aligned}$$

A partir de la fonction de répartition on peut définir :

**i)** La fonction de masse fini-dimensionnelle  $P_{Y_{t_1}, \dots, Y_{t_n}}(\cdot)$  lorsque l'espace d'état  $\mathbb{R}$  est dénombrable.

**ii)** La fonction densité de masse fini-dimensionnelle  $f_{Y_{t_1}, \dots, Y_{t_n}}(\cdot)$  lorsque les variables aléatoires  $(Y_t, t \in \mathbb{Z})$  sont absolument continues.

## Caractéristiques de la distribution d'un processus aléatoire

Comme pour les variables aléatoires, la distribution infini-dimensionnelle d'un processus aléatoire  $(Y_t, t \in \mathbb{Z})$  est aussi caractérisée par certaines familles particulières définies sur  $\mathbb{Z}$  à valeurs dans  $\mathbb{R}$ , à savoir : la fonction moyenne, la fonction variance, la fonction d'autocovariance.

**a. Fonction moyenne :** La fonction moyenne  $\mu = E(Y_t)$  est une fonction de  $\mathbb{Z}$  dans  $\mathbb{R}$  qui pour tout  $t \in \mathbb{Z}$  associe l'espérance mathématique du membre  $Y_t$ . Ayant comme domaine de définition l'ensemble  $D_\mu = \{t \in \mathbb{Z} : E(Y_t) < \infty\}$ .

**b. Fonction variance :** La fonction variance  $\sigma^2$  est une fonction de  $\mathbb{Z}$  dans  $\mathbb{R}$  qui pour tout  $t \in \mathbb{Z}$  associe la variance mathématique du membre  $Y_t$ . Ayant comme domaine de définition l'ensemble  $D_{\sigma^2} = \{t \in \mathbb{Z} : E(Y_t^2) < \infty\}$ .

**c. Fonction d'autocovariance :** La fonction d'autocovariance  $\gamma(.,.)$  est une fonction de  $\mathbb{Z} \times \mathbb{Z}$  dans  $\mathbb{R}$  qui pour tout couple  $(t, s) \in \mathbb{Z} \times \mathbb{Z}$  associe la covariance entre les membres  $Y_t$  et  $Y_s$ . Ayant comme domaine de définition l'ensemble  $D_\gamma = \{(t, s) \in \mathbb{Z} \times \mathbb{Z} : E(Y_t Y_s) < \infty\}$ .

**d. Fonction d'autocorrélation :** On définit également la fonction d'autocorrélation  $\rho(.,.)$  comme une fonction de  $\mathbb{Z} \times \mathbb{Z}$  dans  $[-1, 1]$  qui pour tout couple  $(t, s) \in \mathbb{Z} \times \mathbb{Z}$  associe la corrélation entre les membres  $Y_t$  et  $Y_s$ . Ayant comme domaine de définition l'ensemble  $D_\rho = D_\gamma$ .

**e. Fonction d'autocorrélation partielle :** L'autocorrélation partielle  $\delta_h$  de retard  $h$  mesure la corrélation entre deux observations éloignées de  $h$  périodes quand l'influence linéaire des observations à l'intérieur a été neutralisée. Donc c'est la corrélation entre  $Y_t$  et  $Y_{t-h}$  qui n'est pas expliquée par  $Y_{t+1}, \dots, Y_{t-h-1}$ .

**f. Fonction génératrice :** On appelle fonction génératrice des moments de  $(Y_t, t \in \mathbb{Z})$  la fonction  $M(.,.)$  définie de  $\mathbb{R}$  dans  $\mathbb{R}$ . Le domaine de définition de  $M(.,.)$  est alors donné par  $D_M = \{(t, z) \in T \times \mathbb{R} : E(e^{zY_t}) < \infty\}$ .

## Structure de dépendance associée aux processus aléatoires

### a. Stationnarité

La propriété de stationnarité (stochastique), qui caractérise plutôt une certaine régularité stochastique dans l'évolution, joue un rôle crucial dans la théorie des processus aléatoires. Dans plusieurs problèmes du monde réel, on rencontre des phénomènes aléatoires qui évoluent dans un régime "d'équilibre stochastique" dans le sens où les caractéristiques fréquentistes du phénomène ne changent pas dans le domaine d'évolution. De tels phénomènes peuvent être représentés par lesdits processus stationnaires.

Dans la suite nous allons évoquer la définition de la stationnarité qui a été introduite par Kintchine (1931) pour le cas d'un domaine d'évolution  $\mathbb{Z}$  quelconque et qui est dite stationnarité stricte.

#### Définition 1.9.3 (Stationnarité stricte )

Un processus aléatoire  $(Y_t, t \in \mathbb{Z})$  défini sur un espace de probabilité  $(\Omega, A, P)$  à valeurs dans un espace d'état  $(\mathbb{R}, B(\mathbb{R}))$  et de domaine d'évolution  $\mathbb{Z}$  est dit strictement stationnaire, si  $\forall n, h \in \mathbb{Z}, \forall y_i \in \mathbb{R}$ , on a

$$P(Y_{t_1} \leq y_1, Y_{t_2} \leq y_2, \dots, Y_{t_n} \leq y_n) = P(Y_{t_1+h} \leq y_1, Y_{t_2+h} \leq y_2, \dots, Y_{t_n+h} \leq y_n),$$

tel que  $t_i + h \in \mathbb{Z}, i = 1, \dots, n$ . Autrement dit, si les distributions fini-dimensionnelles sont invariantes par translation admissible dans le domaine d'évolution.

Le processus strictement stationnaire le plus simple est la suite i.i.d.

Il est important de noter que tous les moments d'un processus strictement stationnaire, lorsqu'ils existent, sont invariants dans le temps. Le processus est ainsi dit stationnaire en tous les moments ou à l'ordre 1. Ainsi, la définition de stationnarité stricte semble contraignante puisque lorsque tous les moments existent, elle exige l'invariance de tous ces moments par rapport au temps. De plus, elle repose sur la connaissance des lois fini-dimensionnelles du processus qui ne peuvent être connues en pratique, sauf dans des cas très spéciaux. Cependant, plusieurs propriétés probabilistes essentielles des processus aléatoires peuvent être obtenues juste à partir des deux premiers moments (lorsqu'ils existent) et pour les moments restants, la distinction est souvent négligeable. La stationnarité de ces deux premiers mo-

ments peut être suffisante pour expliquer du moins avec bonne précision, la stationnarité dans la distribution du processus. En particulier, pour les processus dits gaussiens qui sont très répandus en pratique, la stationnarité de ces deux premiers moments est équivalente à la stationnarité stricte. C'est pourquoi, on a souvent besoin d'un concept de stationnarité moins fort mais qui peut être rencontré en pratique.

#### Définition 1.9.4 (Stationnarité au second ordre)

Un processus aléatoire  $(Y_t, t \in \mathbb{Z})$  défini sur un espace de probabilité  $(\Omega, A, P)$  à valeurs dans un espace d'états  $(\mathbb{R}, B(\mathbb{R}))$  et de domaine d'évolution  $\mathbb{Z}$  est dit stationnaire au second ordre (*faiblement stationnaire, covariance stationnaire*), si

- i) La fonction variance  $\sigma^2(t) = \text{var}(Y_t) < \infty$  existe pour tout  $t$  dans le domaine d'évolution  $\mathbb{Z}$  (et donc nécessairement la fonction moyenne  $\mu(t) = E(Y_t)$  existe :  $E(Y_t^2) < \infty$ )
- ii) La fonction moyenne  $\mu(t)$  est constante sur le domaine d'évolution et la fonction d'autocovariance  $\gamma(Y_t, Y_{t+h}) = \text{cov}(Y_t, Y_{t+h})$  dépend seulement de  $h$ , ce qui entraîne que la fonction variance  $\sigma^2(t)$  est constante dans le domaine d'évolution  $\mathbb{Z}$ .

Le processus faiblement stationnaire le plus simple est le processus bruit blanc.

#### b. Ergodicité

L'ergodicité exprime le fait que le processus peut prendre n'importe quelle valeur dans l'espace des états indépendamment de sa valeur initiale, et qu'il ne peut s'isoler dans un sous ensemble strictement inclus dans l'espace des états, c'est donc une propriété d'irréductibilité mais associée à des processus stationnaires. Cette propriété est très importante pour l'établissement du théorème ergodique.

Le théorème ergodique représente une généralisation du théorème des grands nombres au cas des processus non i.i.d. mais stationnaires. Ce théorème stipule que sous certaines conditions, la moyenne empirique d'une fonction du processus converge vers une certaine variable aléatoire. Lorsque le processus est ergodique cette variable aléatoire, si elle existe n'est autre que la variable dégénérée représentant la moyenne commune à tous les membres.

Avant de définir formellement la notion d'ergodicité, donnons d'abord le concept d'invariance d'un espace par rapport à une transformation.

**Définition 1.9.5**

Soit  $\mathbb{R}^{\mathbb{Z}}$  l'espace de toutes les séquences  $(Y_t, t \in \mathbb{Z})$  de nombres réels.

i) L'application  $\mathcal{T} : \mathbb{R}^{\mathbb{Z}} \rightarrow \mathbb{R}^{\mathbb{Z}}$  est dite opérateur de translation si

$$\mathcal{T} (\dots, Y_0, Y_1, Y_2, \dots) = (\dots, Y_1, Y_2, Y_3, \dots).$$

ii) Un sous ensemble  $A \subset \mathbb{R}^{\mathbb{Z}}$  est dit  $\mathcal{T}$ -invariant si  $\mathcal{T}(A) = A$ . Autrement dit, le seul espace  $\mathcal{T}$ -invariant est  $\mathbb{R}$  lui-même.

Il existe deux types d'ergodicité; ergodicité faible et ergodicité forte.

**Définition 1.9.6 (Ergodicité)**

Un processus stochastique stationnaire  $(Y_t, t \in \mathbb{Z})$  est dit ergodique si

$$P((Y_1, Y_2, Y_3, \dots) \in A) = 0 \text{ ou } 1,$$

pour chaque ensemble  $A$   $\mathcal{T}$ -invariant.

Autrement dit, le seul sous espace invariant par la translation  $\mathcal{T}$  à quelques ensembles de mesures nulles près est l'ensemble  $\mathbb{R}^d$  lui-même, ce qui exprime une certaine irréductibilité.

Lorsque le processus aléatoire  $(Y_t, t \in \mathbb{Z})$  est strictement stationnaire, la suite définie par  $\frac{1}{n} \sum_{t=1}^n Y_t$  converge vers une variable aléatoire, mais lorsque le processus  $(Y_t, t \in \mathbb{Z})$  est ergodique, une telle variable n'est rien d'autres que la moyenne (une constante). D'où le théorème ergodique suivant :

**Théorème ergodique pour processus stationnaire et ergodique**

**Théorème 1.9.1**

Soit  $(Y_t, t \in \mathbb{Z})$  un processus ergodique strictement stationnaire avec une moyenne finie  $m$ . Alors

$$P \left( \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n Y_k = m \right) = 1.$$

Le théorème suivant nous donne différentes caractérisations de l'ergodicité.

**Théorème de Karlin et Taylor, (1975)****Théorème 1.9.2** (Karlin et Taylor, 1975)

Soit  $(Y_t, t \in \mathbb{Z})$  un processus stationnaire. Les conditions suivantes sont équivalentes :

(a)  $(Y_t, t \in \mathbb{Z})$  est ergodique.

(b) Pour chaque ensemble  $A$   $\mathcal{T}$ -invariant par translation

$$P((Y_0, Y_1, \dots) \in A) = 0 \text{ ou } 1.$$

(c) Pour chaque ensemble  $A$  de séquences réelles  $(Y_0, Y_1, \dots)$ ,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n I((Y_j, Y_{j+1}) \in A) = P((Y_0, Y_1, \dots) \in A),$$

tel que  $I(A)$  est la variable aléatoire indicatrice de l'événement  $A$ .

$$I(A) = \begin{cases} 1 & \text{si } A \text{ se produit} \\ 0 & \text{si non} \end{cases}.$$

(d) Pour chaque  $k = 1, 2, \dots$  et chaque ensemble  $A$  de vecteurs réels  $(Y_0, \dots, Y_k)$ ,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n I((Y_j, \dots, Y_{j+k}) \in A) = P((Y_0, Y_1, \dots, Y_k) \in A).$$

(e) Pour chaque  $k$  et chaque fonction  $g$  de  $k + 1$  variables,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n g(Y_j, \dots, Y_{j+k}) = E(g(Y_0, Y_1, \dots, Y_k)),$$

à condition que l'espérance existe.

(f) Pour chaque fonction  $\phi$  de  $\mathbb{R}^{\mathbb{Z}}$  dans  $\mathbb{R}$ , on a

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \phi(Y_j, Y_{j+1}, \dots) = E(\phi(Y_0, Y_1, \dots)),$$

sous condition que l'espérance existe.

L'ergodicité déjà définie implique le théorème ergodique dans le sens de la convergence presque sûre. On peut avoir une version du théorème plus faible dans le sens de convergence en probabilité et en moyenne quadratique.

### 1.9.3 Densité spectrale

Jusque là, nous avons considéré une série chronologique du point de vue temporel, i.e. on regarde une série comme une suite ordonnée d'observations dont la représentation graphique peut suffisamment décrire l'évolution. En revanche, on peut regarder une série chronologique d'un autre point de vue dit fréquentiel, selon lequel la série étant fonction du temps peut être décomposée, selon la théorie de Fourier, comme combinaison (infinie) de fonctions sinusoïdales de périodes (et donc de fréquences) variées. On mesure alors pour chaque fréquence (et donc pour chaque période) l'importance de sa contribution dans la composition de la série. La fonction qui associe à chaque fréquence sa contribution dans la série s'appelle densité spectrale de la série. Le graphique correspondant est appelé spectre.

Mathématiquement, la densité spectrale d'un processus stationnaire est la transformée de Fourier de la fonction d'autocorrélation. Elle est donnée par

$$\begin{aligned} f(\omega) &= \frac{1}{2\pi} \sum_{h=-\infty}^{+\infty} \rho_h e^{-i\omega h} \\ &= \frac{1}{2\pi} \left( \rho_0 + 2 \sum_{h=-\infty}^{+\infty} \rho_h \cos(h\omega) \right), \quad -\pi \leq \omega \leq \pi, \end{aligned}$$

où par convention  $e^{-i\omega h} = \cos(h\omega) + i \sin(\omega h)$ .

C'est une fonction paire et périodique de période  $2\pi$ .

### 1.9.4 Equations aux récurrences stochastiques

La quasi-totalité des modèles de séries chronologiques sont souvent exprimés par des équations aux récurrences stochastiques de la forme suivante :

$$Y_{n+1} = A_{n+1}Y_n + B_{n+1}, \quad n \in \mathbb{Z}. \quad (1.9.1)$$

Notre premier objectif dans cette section est de déterminer les conditions d'existence d'une unique solution stationnaire de l'équation (1.9.1), dans le cas où la séquence  $(A_n, B_n)$  est stationnaire et ergodique.

Les propriétés de stationnarité de la quasi totalité des modèles de séries chronologiques sont liées directement aux propriétés de stationnarité de la solution de l'équation (1.9.1). Il est suffisant donc d'étudier cette équation. Brandt (1986) a établi un théorème dans lequel il a donné les conditions d'existence d'une solution strictement stationnaire pour (1.9.1). Il l'a utilisé pour le cas unidimensionnel ( $d = 1$ ), mais la démonstration est valable pour tout  $d$ .

#### **Théorème 1.9.3 (Brandt, 1986)**

Soit  $\{(A_n, B_n), n \in \mathbb{Z}\}$  un processus strictement stationnaire et ergodique tel que  $E(\log^+(A_0))$  et  $E(\log^+(B_0))$  sont toutes les deux finies. Supposons que l'exposant de Lyapounov  $\gamma$  défini par

$$\gamma = \inf \left\{ E \left( \frac{1}{n+1} \log \|A_0 A_{-1}, \dots, A_{-n}\|, n \in \mathbb{N} \right) \right\},$$

est strictement négatif. Alors, pour tout  $n \in \mathbb{N}$ , la série

$$Y_n = \sum_{k=0}^{+\infty} A_n A_{n-1}, \dots, A_{n-k+1} B_{n-k},$$

converge presque sûrement, et le processus  $(Y_n, n \in \mathbb{Z})$  est l'unique solution strictement stationnaire de (1.9.1).

On se place dans le cas où  $(A_n, B_n)$  est i.i.d. et on donne quelques définitions relatives à l'équation (1.9.1).

**Définition 1.9.7** Un modèle autorégressif généralisé avec des coefficients i.i.d. est un modèle

$$Y_{n+1} = A_{n+1}Y_n + B_{n+1}, \quad n \in \mathbb{Z}, \quad (1.9.2)$$

avec  $\{(A_n, B_n), n \in \mathbb{Z}\}$  est une séquence de variables aléatoires i.i.d. définies dans l'espace de probabilité  $(\Omega, A, P)$  qui prennent leurs valeurs dans  $M(d) \times \mathbb{R}^d$ . La solution de cette équation est une séquence  $(Y_n, n \in \mathbb{Z})$  de  $\mathbb{R}^d$  variables aléatoires pour laquelle (1.9.2) est vérifiée.

Dans la suite nous allons définir une solution stationnaire qui ne dépend pas du futur.

**Définition 1.9.8 (Solution non anticipative)**

Une solution strictement stationnaire non anticipative de (1.9.2) est un processus strictement stationnaire  $(Y_n, n \in \mathbb{Z})$  qui est une solution de (1.9.2), tel que pour chaque  $p \in \mathbb{Z}$ ,  $Y_p$  est indépendant des variables aléatoires  $\{(A_n, B_n), n \in \mathbb{Z}\}$ .

Passons à la définition suivante qui introduit l'irréductibilité de (1.9.2).

**Définition 1.9.9** Soit  $H$  un sous espace affine de  $\mathbb{R}^d$ . Il serait dit invariant sous le modèle (1.9.2) si  $\{A_0y + B_0, y \in H\}$  est contenue dans  $H$  presque sûrement. Le modèle (2.3.2) est appelé irréductible si  $\mathbb{R}^d$  est le seul sous espace affine invariant.

**Théorème 1.9.4 (Bougerol et Picard, 1992)**

Considérons le modèle (1.9.2) avec des coefficients i.i.d. Supposons que ce modèle est irréductible et qu'il a une solution strictement stationnaire et non anticipative  $(Y_n, n \in \mathbb{Z})$ .

Donc il s'ensuit que :

- i)  $A_0, A_{-1}, \dots, A_{-k}$  converge vers 0 presque sûrement quand  $k \rightarrow +\infty$ .
- ii) Pour tout entier  $n$ ,

$$Y_n = \sum_{k=0}^{+\infty} A_n A_{n-1} \dots A_{n-k+1} B_{n-k},$$

où la série converge presque sûrement.

- iii) Cette solution est l'unique solution strictement stationnaire de (1.9.2).

Le théorème que nous allons donner par la suite stipule que, sous la condition d'irréductibilité, s'il existe une solution strictement stationnaire non anticipative de (1.9.2) alors l'exposant de Lyapounov  $\gamma$  est strictement négatif.

**Théorème 1.9.5 (Bougerol et Picard, 1992)**

Supposons que le modèle (1.9.2) avec les coefficients i.i.d est irréductible et que  $E(\log^+ \|A_0\|)$  et  $E(\log^+ \|B_0\|)$  sont toutes les deux finies. Alors (1.9.2) a une solution strictement stationnaire non anticipative si et seulement si l'exposant de Lyapounov est strictement négatif.

**Proposition 1.9.1 (Bougerol et Picard, 1992)**

Supposons que le modèle (1.9.2) a une solution strictement stationnaire non anticipative  $(Y_n, n \in \mathbb{Z})$ . Soit  $H$  le sous espace affine minimal de  $\mathbb{R}^d$  tel que  $P(Y_0 \in H) = 1$ . Alors  $H$  est invariant sous le modèle et n'importe quel sous espace invariant de  $H$  porte une solution non anticipative strictement stationnaire.

Le corollaire suivant est une conséquence immédiate du théorème 1.9.3, théorème 1.9.5 et de la proposition 1.9.1.

**Corollaire 1.9.1 (Bougerol et Picard, 1992)**

Considérons le modèle (2.3.2) avec  $E(\log^+ \|A_0\|)$  et  $E(\log^+ \|B_0\|)$  sont finies. Supposons qu'il existe une solution non anticipative strictement stationnaire qui n'est pas portée par l'hyperplan affine. Alors les conditions suivantes sont équivalentes :

- i) L'exposant de Lyapounov est strictement négatif.
- ii) Le modèle est irréductible.
- iii) Il existe une unique solution stationnaire.

## 1.9.5 Comportement des queues

L'étude de la queue de la solution stationnaire de l'équation aux récurrences stochastiques passe par celle des produits  $A_1, \dots, A_n$ . Dans la suite nous donnons deux résultats, le premier concerne le maximum de ces produits, et le deuxième porte sur la queue de la solution stationnaire de cette équation.

Soient  $(\Omega, A, P)$  un espace probabilisé, et  $(A_n, B_n)$  une suite de variables aléatoires indépendantes et identiquement distribuées sur  $\Omega$  où les  $(A_n)$  sont des matrices carrées de taille  $d$ , et les  $(B_n)$  des vecteurs colonne de  $\mathbb{R}^d$ . Soit  $\mathcal{S}^{d-1}$  la sphère de  $\mathbb{R}^d$ , i.e. l'ensemble des vecteurs lignes de norme 1, et  $\mathcal{S}_+$  ses éléments positifs :

$$\mathcal{S}^{d-1} = \{x \in \mathbb{R}^d / \|x\| = 1\}, \mathcal{S}_+ = \{x \in \mathcal{S}^{d-1} / x \geq 0\}$$

avec  $x$  est l'élément de la sphère.

Si  $\mu$  est une mesure, on notera  $\mathcal{S}_\mu$  son support. Et enfin on note  $R = \sum_{k=0}^{+\infty} A_1, \dots, A_{k-1} B_k$

qui suit la loi stationnaire de l'équation aux récurrences stochastiques.

**Théorème 1.9.6 (Kesten, 1973)**

Soit  $(A_n)$  une suite de matrices réelles i.i.d. de taille  $d \times d$  et de loi  $\mu$  telles que :

$$P(A_0 \geq 0) = 1, P(A_0 \text{ a une ligne de } 0) = 0.$$

$$E [\log^+ \|A_0\|] < +\infty$$

Supposons aussi que le sous-groupe de  $\mathbb{R}$  engendré par

$$\{\log \rho(a) / a = a_1, \dots, a_n, n \in \mathbb{N}, a_i \in \mathcal{S}_\mu, a > 0\},$$

est dense. Alors, il existe une constante (exposant de Lyapounov)  $\gamma < +\infty$  telle que pour tout  $x \in \mathcal{S}_+$  on ait :

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \|A_1 \dots A_n\| = \lim_{n \rightarrow \infty} \log \|x A_1 \dots A_n\| = \gamma.$$

presque sûrement. Si de plus  $\gamma$  est strictement négatif, et s'il existe  $\sigma > 0$  tel que

$$E \left[ \min_{1 \leq i \leq d} \left( \sum_{j=1}^d A_0(i, j) \right) \right]^\sigma \geq d^{\sigma/2}, \text{ et } E [\|A_0\|^\sigma \log^+ \|A_0\|] < \infty,$$

alors il existe un exposant  $k$  dans l'intervalle  $]0, \sigma]$  tel que la limite

$$\lim_{k \rightarrow \infty} t^k P \left( \max_n \|x A_1 \dots A_n\| > t \right),$$

existe et est strictement positive pour tout  $x$  dans  $\mathcal{S}_+$ .

On donne maintenant le théorème sur la queue de la loi stationnaire de l'équation aux récurrences stochastiques (voir Saporta, 2004).

**Théorème 1.9.7 (Kesten, 1973)**

Soit  $(A_n, B_n)$  une suite de variables aléatoires i.i.d. où les  $A_n$  sont des matrices carrées et les  $B_n$  des vecteurs colonne. Supposons que la suite  $(A_n)$  vérifie toutes les hypothèses du théorème précédent, et que, avec les notions de ce théorème :

$$P(B_0 = 1) < 1, P(B_0 \geq 0) = 1, E \|B_0\|^k < \infty,$$

alors, pour tout  $x \in \mathcal{S}^{d-1}$ , la limite

$$\lim_{t \rightarrow \infty} t^k P(xR > t),$$

existe et est finie. De plus pour tout  $x$  dans  $\mathcal{S}_+$ , cette limite est strictement positive.

**Théorème 1.9.8 (Goldie, 1991)**

Soit  $A$  et  $B$  deux variables aléatoires définies sur un espace de probabilité commune.

Supposons que :

- i) Il existe  $\alpha > 0$  tels que  $E |A|^\alpha = 1, E |A|^\alpha \ln^+ |A|^\alpha < \infty$  et  $E |A|^\beta < \infty$ .
- ii) La loi conditionnelle de  $\ln |A|$  donnée  $\{A \neq 0\}$  est non arithmétique.

Alors il existe une variable aléatoire  $Y$  telle que :  $Y$  et  $(A, B)$  sont indépendants et  $Y \stackrel{d}{=} AY + b$  ( $\stackrel{d}{=}$  désigne l'égalité dans la distribution).

De plus, il existe des constantes non-négatives  $c^+$  et  $c^-$  telles que  $p(Y > t) \sim c^+ t^{-\alpha}$  et  $p(Y < -t) \sim c^- t^{-\alpha}$  si  $t \rightarrow \infty$ , où  $c^+ + c^- > 0$  si seulement si pour tout  $c$  réel,  $p(B = (1 - A)/c) < 1$ . Supposons en outre que  $(Y_t, t \in \mathbb{Z})$  est une solution de l'équation

(1.9.1), où la paire iid  $(A_t, B_t)$  a la même distribution que  $(A, B)$ . Alors  $Y_t \xrightarrow{d} Y$  indépendamment de la valeur de départ  $Y_0$ . En particulier, si  $Y_t \stackrel{d}{=} Y$  alors  $(Y_t, t \in \mathbb{Z})$  est une séquence stationnaire.

# Chapitre 2

## Analyse des phénomènes à queues lourdes et valeurs extrêmes dans le cas iid

### 2.1 Introduction

Les distributions à queues lourdes sont liées à la théorie des valeurs extrêmes et permettent de modéliser plusieurs phénomènes rencontrés dans différentes disciplines : finances, télécommunication, . . . Plusieurs définitions ont été associées à ces distributions en fonction du critère de classification. La caractérisation la plus simple et celle basée sur la comparaison avec la loi normale : On dit, alors, qu'une distribution a une queue lourde si :

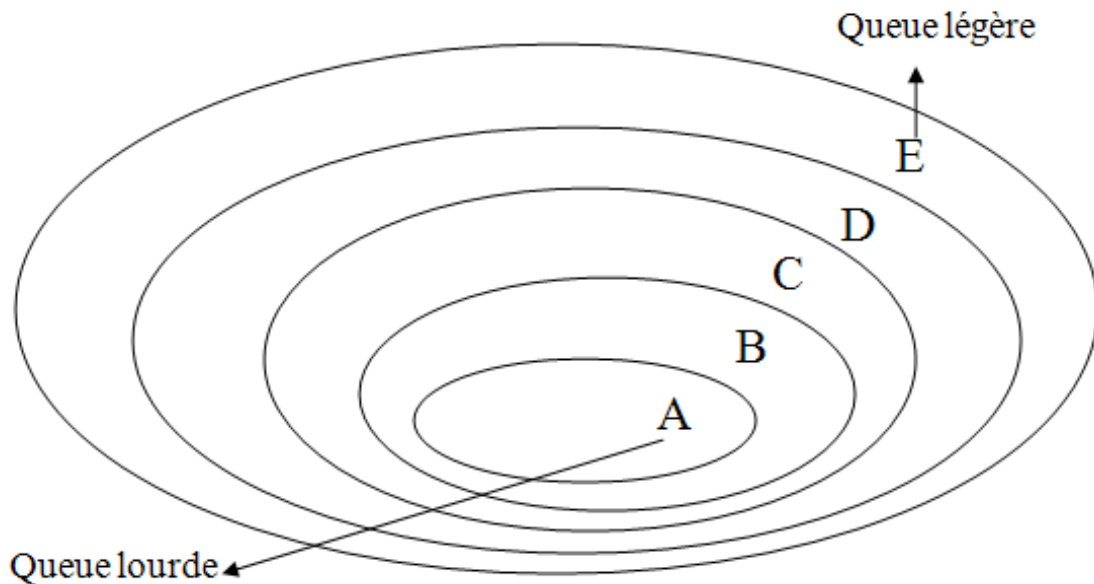
$$C_k = E \left( \frac{(Y - \mu)^4}{\sigma} \right) > 4.$$

Ce qui est équivalent à dire qu'une distribution a une queue lourde si et seulement si son coefficient d'aplatissement,  $C_k$ , est supérieur à celui de la loi normale (*pour laquelle*  $C_k = 3$ ).

La caractérisation, donnée par l'équation précédente, est très générale et ne peut être appliquée que si le moment d'ordre 4 existe. Par conséquent aucune discrimination, pour les distributions ayant un moment d'ordre 4 infini, ne peut être faite si on ne considère que ce critère. Malheureusement, il n'y a pas de critère pour classer toutes les distributions selon la queue droite. Cependant, on obtient un tel classement pour certaines classes de distributions :

- Les distributions avec des moments exponentielles inexistants ( $E$ ),
- Les distributions sous-exponentielles ( $D$ ),
- Les distributions à variations régulières ( $C$ ),
- Les distributions avec un comportement de Pareto ( $B$ ),
- Les distributions  $\alpha$ -Stables avec  $\alpha < 2$  ( $A$ ).

Toutes ces classes sont emboîtées :  $A \subset B \subset C \subset D \subset E$  et peuvent être représentées selon la figure suivante (Werner et Upper, 2002).



Diférentes classes de distributions de queues très légères ( $E$ ) à très lourdes ( $A$ ).

Ils existent plusieurs définitions concernant les distributions à queues lourdes, mais on peut considérer la définition simple suivante :

**Définition 1.1** Soit  $X$  une variable aléatoire de fonction de répartition  $F(x) = p(X \leq x)$ , on dit que  $X$  est à queue lourde (de type régulièrement varié) s'il existe un paramètre positif  $\alpha \geq 0$  tels que :

$$\bar{F}(x) := p[X > x] \sim x^{-\alpha}, x \rightarrow \infty,$$

ou " $\sim$ " désigne le symbole d'équivalence asymptotique défini par :  $f(x) \sim g(x)$  quand

$x \rightarrow \infty$  si seulement si :  $\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = 1$ .

**Exemples de queues de quelques distributions :**

Soit le tableaux suivant :

Distributions	$\bar{F} = 1 - F(x)$ ou la queue de la densité $f$	Paramètres
Exponentiel	$\bar{F} = \exp\{-\lambda x\}$	$\lambda > 0$
Weibull	$\bar{F} = \exp\{-cx^r\}$	$c > 0, r \geq 1$

Le tableau suivant représente quelques distributions à queues lourdes :

Distributions	$\bar{F} = 1 - F(x)$ ou la queue de la densité $f$	Paramètres
Pareto	$\bar{F} = \left(\frac{k}{k+x}\right)^\alpha$	$\alpha, k > 0$
Burr	$\bar{F} = \left(\frac{k}{k+x^r}\right)^\alpha$	$\alpha, k, r > 0$

## 2.2 Classes importantes des distributions à queues lourdes

Parmi les classes les plus importantes des distributions à queues lourdes, on en distingue :

### 2.2.1 Distributions à variations régulières d'indice $\alpha > 0$

La classe  $C$ , des distributions à variations régulières, est une sous classe des distributions sous-exponentielles. Ces distributions vérifient la condition suivante (cf, Feller, 1968 ; Embrechts et al,1997) :

$$\lim_{t \rightarrow \infty} \frac{\bar{F}(ty)}{\bar{F}(y)} = y^{-\alpha},$$

avec  $\bar{F}$  la fonction de probabilité au dépassement (connue aussi sous le nom de la fonction de survie en fiabilité) :

$$\bar{F}(y) = P(Y \geq y) = 1 - F(y) .$$

Ce qui est équivalent à dire que, en ce qui concerne les valeurs extrêmes ( $t$  tend vers l'infini) la distribution a le même comportement que celui de la loi de Pareto. Par conséquent, la probabilité au dépassement des extrêmes décroît suivant une fonction puissance (appelée aussi décroissance géométrique). Le paramètre  $\alpha$  est appelé «l'indice des valeurs extrêmes» et peut être utilisé comme critère pour classer les distributions par rapport au comportement de la fonction de probabilité au dépassement  $\bar{F}$  au niveau des valeurs extrêmes.

La classe des distributions à variations régulières est liée à la théorie des valeurs extrêmes. En effet, la théorie des valeurs extrêmes est basée sur le théorème de Fisher-Tippett (Fisher et Tippett, 1928) qui permet de déterminer la distribution du maximum d'un échantillon de  $n$  variables i.i.d :

Si  $Y_1, \dots, Y_n$  une séquence de variables aléatoires i.i.d, et  $M_n$  le maximum définit par :

$$M_n = \max(Y_1, Y_2, \dots, Y_n),$$

alors, le théorème de Fisher-Tippett montre que si la distribution du maximum de l'échantillon converge vers une distribution non dégénérée alors elle suit une des trois fonctions de

densité de probabilité suivantes :

$$\text{Gumbel} : \Lambda(y) = \exp(-e^{-y}), \quad y \in ]-\infty, +\infty[$$

$$\text{Fréchet} : \Phi_\alpha(y) = \begin{cases} 0 & \text{si } y \leq 0 \\ \exp(-y^{-\alpha}) & \text{si } y \geq 0, \alpha > 0 \end{cases}$$

$$\text{Weibull} : \Psi_\alpha(y) = \begin{cases} \exp(-(-y^{-\alpha})) & \text{si } y \leq 0, \alpha > 0 \\ 0 & \text{si } y > 0 \end{cases}.$$

Lorsque la distribution du maximum converge vers l'une de ces trois distributions, on dit qu'elle appartient au domaine d'attraction maximum de Fréchet, Weibull ou Gumbel. Les distributions appartenant au domaine d'attraction maximum de Gumbel ont une queue légère.

Alors que, les distributions appartenant au domaine d'attraction maximum de Weibull ont un support borné supérieurement. D'autres caractérisations des distributions appartenant à chacun de ces domaines d'attraction maximum sont données dans la littérature. On peut donner le domaine d'attraction maximum de la majorité des distributions usuelles. Comme par exemple pour le domaine d'attraction maximum de Gumbel on trouve les lois : Normale, Exponentielle, Log-normale, Gamma. Pour le domaine d'attraction maximum de Fréchet : les lois Cauchy, Pareto, Chi-Deux et Student et pour le domaine d'attraction maximum de Weibull : les lois Uniforme et Beta.

### 2.2.2 Distributions sous-exponentielles

Les distributions sous-exponentielles  $D$  sont des classes très importantes de distributions à queues lourdes, on peut comparer entre ces distributions et les distributions exponentielles telles que dans les distributions sous-exponentielles la queue décroît plus lentement qu'une queue à distribution exponentielle.

Cette classe de distributions à été introduite indépendamment par Chistyakov (1964, p.15) et Chver, Ney et Wainger (1972, p.16) et ce principalement dans le contexte du processus de branchement. Une introduction indépendante de  $D$  à travers des questions en théorie des files d'attente se trouve dans Borovkov (1984, p.7-8), voir aussi Pakes (1975, p.92).

Comme on peut considérer  $D$  comme une classe utile des fonctions de distribution à queues lourdes dans le contexte de probabilités appliquées en général, et dans les mathématiques d'assurances en particulier, ceci a été réalisé par Teugels (1975, p.107).

Pour bien définir les distributions sous-exponentielles on va présenter quelques définitions et quelques résultats sur ces distributions.

**Définitions 2.2.1** Soient  $X, X_1, \dots, X_n$  des variables aléatoires indépendantes identiquement distribuées (iid) non négatives.

On a vu dans le corollaire (1.8.1) et la remarque (1.8.2) que pour dire qu'une variable aléatoire  $X$  soit à variation régulière si est seulement si :

$$p(S_n > x) \sim np(X > x) \sim p(M_n > x), \quad (2.2.1)$$

pour  $x \rightarrow \infty$  et pour  $n = 2, 3, \dots$ , où  $S_n = X_1 + X_2 + \dots + X_n$  et  $M_n = \max_{i=1, \dots, n} X_i, n \geq 1$ .

L'interprétation intuitive de l'expression (2.2.1) est que le maximum  $M_n$  de  $X_1, \dots, X_n$  apporte une contribution importante à sa somme : les dépassements des seuils élevés par la somme  $S_n$  sont dues au dépassement de ce seuil par la plus grande valeur dans l'échantillon. Cette interprétation propose une façon de définir une distribution à queue lourde : la queue de la somme  $S_n$  est essentiellement déterminée par la queue du maximum. Cette approche intuitive conduit à la définition d'une classe suffisamment large des distributions à queues lourdes.

**Définitions 2.2.2 (Distribution sous-exponentielles)**

On dit qu'une variable aléatoire non négative  $X$  est de distribution sous-exponentielles si les variables aléatoires  $X_1, X_2, \dots, X_n$  indépendantes identiquement distribuées sont de même distribution que  $X$  satisfait la relation (2.2.1).

**Remarque 2.2.1**

a) Chistyakov (1964, p.15) a prouvé que (2.2.1) est valable pour tout  $n \geq 2$  si et seulement si elle est valable pour  $n = 2$ .

b) Embrechts et Goldie (1980, p.32) a montré que (2.2.1) est valable pour tout  $n \geq 2$ , si elle est valable pour  $n \geq 2$ .

Par ailleurs, il suffit d'exiger que la relation :  $\limsup_{x \rightarrow \infty} \frac{p(S_n > x)}{np(X > x)} \leq 1$  est valable pour  $n \geq 2$ , voir les deux lemmes (1.3.4) et (A3.14) dans Embrechts et al [34].

**Propriétés de base** Dans ce qui suit, nous donnons quelques propriétés élémentaires des distributions sous-exponentielles. Ce qui suit est le lemme (1.3.5) Embrechts et al (1997, p.34).

**Lemme 2.2.1 (Propriétés de base des distributions sous-exponentielles)**

a) Si  $F \in D$ , alors elle définit uniformément sur les ensembles compacts de  $]0, \infty[$ , tel que :

$$\lim_{x \rightarrow \infty} \frac{\overline{F}(x-y)}{\overline{F}(x)} = 1. \tag{2.2.2}$$

b) Si (2.2.2) est valable, donc pour toutes  $\varepsilon > 0$  on a :  $e^{\varepsilon x} \overline{F}(x) \rightarrow \infty, x \rightarrow \infty$ .

c) Si  $F \in D$  et pour toutes  $\varepsilon > 0$ , il existe une infinité de constantes  $K$  de sorte que pour tout  $n \geq 2$  :

$$\frac{\overline{F^{2*}}(x)}{\overline{F}(x)} \leq K(1 + \varepsilon)^n, x \geq 0. \tag{2.2.3}$$

**Remarque 2.2.2**

1. La propriété (b) du lemme précédent justifie le nom sous-exponentielles pour  $F \in D$ , en effet  $\overline{F}(x)$  décroît vers 0 plus lentement que n'importe quelle exponentielle de type  $e^{-\varepsilon x}$

pour  $\varepsilon > 0$ . En outre, puisque pour tout  $\varepsilon > 0$  on a :

$$\int_y^\infty e^{\varepsilon x} dF(x) \geq e^{\varepsilon y} \bar{F}(y), y \geq 0.$$

Il suit la forme de la propriété (b) du lemme précédent qui dit pour tout  $F \in D$ , au voisinage de 0 la fonction génératrice des moments de  $F$  n'existe pas. Pour cette raison Laplace-Stieltjes ont fait une transformation de fonction de distribution sous-exponentielles a une singularité essentielle vers 0. Ce résultat a été prouvé pour la première fois par Chistyuyakov (1984, p.15), théorème 2.

2) La condition (1.7.2) peut être prise à une autre définition pour les distributions à queues lourdes. Notons que pour une variable aléatoire  $X$  avec la distribution  $F$  la relation (1.7.2) peut être écrite sous la forme suivante :

$$\lim_{x \rightarrow \infty} p(X > x + y / X > x) = \lim_{x \rightarrow \infty} \frac{\bar{F}(x + y)}{\bar{F}(x)} = 1, y > 0.$$

Intuitivement ceci signifie que si  $X$  ne dépasse pas une grande valeur mais il est possible de dépasser une autre grande valeur. Notons que pour une distribution exponentielle avec  $\bar{F}(x) = \exp\{-\lambda x\}$ ,  $x > 0$  pour  $\lambda > 0$  on a :

$$\lim_{x \rightarrow \infty} \frac{\bar{F}(x + y)}{\bar{F}(x)} = e^{-\lambda y}, y > 0.$$

### Exemple

Nous avons déjà établi que la classe des distributions à variations régulières la queue droite est un sous-ensemble de  $D$ , voir le corollaire (1.6.1), et que la distribution exponentielle n'appartient pas à  $D$ , voir la remarque précédente.

**Proposition 2.2.1 (Théorème de caractérisation pour  $D$ )**

Supposons que  $F$  est absolument continue de densité  $f$  et le taux de risque  $q(x)$  éventuellement décroît vers 0. Alors :

a)  $F \in D$  si et seulement si :

$$\lim_{x \rightarrow \infty} \int_0^x e^{yq(x)} f(y) dy = 1. \quad (2.2.4)$$

b) Si la fonction  $x \rightarrow \exp \{xq(x)\} f(x)$  est intégrable sur  $[0, \infty[$  alors  $F \in D$ .

**Exemple (Exemples sur les distributions sous-exponentielles)**

a) Soit  $F$  une distribution de Weibull avec les paramètres  $0 < \tau < 1$  et  $c > 0$  c'est-à-dire :  $\bar{F}(x) = e^{-\tau x^\tau}, x \geq 0$  alors :

$$f(x) = c\tau x^{\tau-1} e^{-\tau x^\tau}, Q(x) = cx^\tau \text{ et } q(x) = c\tau x^{\tau-1},$$

qui décroît vers 0 si  $\tau < 1$ . Nous pouvons appliquer immédiatement la proposition précédente (b) alors :

$$x \rightarrow \exp \{xq(x)\} f(x) = \exp \{c(\tau - 1)x^\tau\} c\tau x^{\tau-1},$$

est intégrable sur  $[0, \infty[$  pour  $0 < \tau < 1$ . donc  $F \in D$ .

b) En utilisant la proposition précédente, on peut prouver aussi pour :  $\bar{F}(x) \sim e^{-x(\ln x)^{-\beta}}, x \rightarrow \infty, \beta > 0$  que  $F \in D$ . Cet exemple montre que l'on peut venir au comportement de la queue exponentielle et on reste toujours dans  $D$ .

c) A ce point, on peut espérer que, pour  $\bar{F}(x) \sim e^{-x^\tau L(x)}, x \rightarrow \infty, 0 < \tau < 1$ , avec  $L$  varie lentement, et  $F \in D$ . De plus, dans cette généralité, la réponse à cette question est non. On peut construire des exemples avec la fonction  $L$  qui varie lentement de sorte que  $F$  ne correspond pas à la condition (1.7.2). Un exemple pour  $\tau = 0$  à été communiqué par Charles Goldie et Cline [18].

## 2.3 Conclusion

D’après l’étude précédente, nous avons été en mesure de connaître la différence entre les queues des distributions comme suit :

Le tableau suivant représente les distributions qui sont à queues lourdes :

sous-exponentielles	à variations régulières
Pareto, Lognormale, Weibull	Pareto, Cauchy, Burr, Fréchet, loi de Zipf-Mandelbrot

Le tableau suivant représente les distributions qui sont à queues fines :

Exponentielle, Gamma, Weibull, Normale, Les distributions à supports compacts
---

## 2.4 Estimation de l’indice de queue (Estimateur de Hill)

### 2.4.1 Introduction

Il existe toute une machinerie pour l’estimation de l’indice de queue  $\alpha$  et l’estimation des quantiles extrêmes. En particulier, il existe plusieurs estimateurs du paramètre  $\alpha$  dont parmi figure l’estimateur de Hill, (1995), l’estimateur de Pickands (1975), l’estimateurs Dekkers-Einmahl et l’estimateur de de Hann (1998) comme estimateurs les plus populaires. Ils sont construits à partir des statistiques d’ordres dans un échantillon.

Les différentes méthodes d’estimations de paramètre  $\alpha$  sont basés sur l’état de la queue, on commence par exemple l’estimation de parametre  $\alpha$  d’une queue régulièrement variable.

### 2.4.2 Estimation d’une queue régulièrement variable

Parmi les estimateurs statistiques pour le parametre  $\alpha$  d’une queue régulièrement variable  $\bar{F}(x) = x^{-\alpha}L(x)$ , l’estimateur de Hill est devenu particulièrement populaire. Dans ce qui

suit, nous expliquons le raisonnement suivant :

Supposons que  $X_1, \dots, X_n$  sont des variables aléatoires iid de distribution  $F$  satisfaisant  $\bar{F}(x) = x^{-\alpha}L(x), x > 0$ , pour une fonction  $L$  qui varie lentement et pour  $\alpha > 0$ . La connaissance de l'indice  $\alpha$  dans les applications est très importante. Si par exemple pour  $\alpha < 2$  alors  $EX_1^2 = \infty$ . Ce cas est souvent observé dans la modélisation des données d'assurance, voir par exemple Hogg et Klugman (1984, p.57). Les études empiriques sur les queues dans les finances ont indiqué qu'on rencontre fréquemment les valeurs  $\alpha$  entre 3 et 4, voir par exemple Guillaume et al (1994, p.50), Longin (1996, p.72) ou Loretan et Phillips (1994, p.73).

L'estimateur de Hill de  $\alpha$  prend essentiellement la forme suivante :

$$\hat{\alpha}^{(H)} = \hat{\alpha}_{k,n}^{(H)} = \left( \frac{1}{k} \sum_{j=1}^k \ln X_{j,n} - \ln X_{k,n} \right)^{-1}, \quad (2.3.1)$$

où  $k = k(n) \rightarrow \infty$  d'une manière appropriée et  $X_{n,n} \leq \dots \leq X_{1,n} = M_n$  désignent les statistiques d'ordre de l'échantillon  $X_1, \dots, X_n$ . Ceci signifie qu'une suite croissante de statistiques d'ordre supérieure est utilisée. L'un des faits intéressants concernant (2.3.1) est que les différentes versions asymptotiquement équivalentes de  $\alpha$  peuvent être obtenus à travers des méthodes essentiellement différentes, montrant que l'estimateur de Hill est très simple. Ci-dessous nous discuterons de certaines dérivations.

### a. L'approche MLE (Hill [56])

Supposons pour le moment que  $X$  est une variable aléatoire de distribution  $F$  de sorte que pour  $\alpha > 0$  :

$$p(X > x) = \bar{F}(x) \sim x^{-\alpha}, x \geq 1.$$

Alors il est clair que  $Y = \ln X$  est de distribution suivante :  $p(Y > y) = e^{-\alpha y}, y \geq 0$ . C'est-à-dire  $Y$  est  $\exp(\alpha)$  et donc le MLE de  $\alpha$  est donnée par :

$$\hat{\alpha}_n = \bar{Y}_n^{-1} = \left( \frac{1}{n} \sum_{j=1}^n \ln X_j \right)^{-1} = \left( \frac{1}{n} \sum_{j=1}^n \ln X_{j,n} \right)^{-1}.$$

Une généralisation triviale concerne :

$$\bar{F}(x) = Cx^{-\alpha}, x \geq u > 0, \quad (2.3.2)$$

avec  $u$  connus. Si nous interprétons (2.3.2) comme complètement spécifié, c'est-à-dire  $C = u^{-\alpha}$ , alors nous obtenons immédiatement comme MLE de  $\alpha$  :

$$\hat{\alpha}_n = \left( \frac{1}{n} \sum_{j=1}^n \ln \left( \frac{X_{j,n}}{u} \right) \right)^{-1} = \left( \frac{1}{n} \sum_{j=1}^n \ln X_{j,n} - \ln u \right)^{-1}. \quad (2.3.3)$$

Maintenant nous ne possédons pas souvent les informations paramétriques précises de ces exemples, mais dans l'esprit de  $MDA(\Phi_\alpha)$ , nous supposons que  $\bar{F}$  se comporte comme une fonction de distribution de Pareto au-dessus d'un seuil  $u$  connue.

Soit :

$$k = \text{card} \{i : X_{i,n} > u, i = 1, \dots, n\}. \quad (2.3.4)$$

Conditionnellement à l'événement  $\{K = k\}$ , l'estimation par le maximum de vraisemblance de  $\alpha$  et  $C$  en (2.3.2) réduit à la maximisation de la densité conjointe de  $(X_{k,n}, \dots, X_{1,n})$ . On peut montrer que :

$$f_{X_{k,n}, \dots, X_{1,n}}(x_k, \dots, x_1) = \frac{n!}{(n-k)!} (1 - Cx_k^{-\alpha})^{n-k} C^k \alpha^k \prod_{i=1}^k x_i^{-(\alpha+1)}, u < x_k < \dots < x_1.$$

Un simple calcul donne le MLE conditionnel comme suit :

$$\hat{\alpha}_{k,n}^{(H)} = \left( \frac{1}{k} \sum_{j=1}^k \ln \left( \frac{X_{j,n}}{X_{k,n}} \right) \right)^{-1} = \left( \frac{1}{k} \sum_{j=1}^k \ln X_{j,n} - \ln X_{k,n} \right)^{-1}, \hat{C}_{k,n} = \frac{k}{n} X_{k,n}^{\hat{\alpha}_{k,n}^{(H)}}.$$

Donc, l'estimateur de Hill a la même forme que le MLE dans le modèle exact sous-jacente (3.3), mais ayant maintenant le  $u$  déterministe remplacé par le seuil aléatoire  $X_{k,n}$ , où  $k$  est définie par (2.3.3).

Nous avons également obtenu immédiatement une estimation de la queue  $\overline{F}(x)$  comme suit :

$$\widehat{\overline{F}}(x) = \frac{k}{n} \left( \frac{x}{X_{k,n}} \right)^{-\widehat{\alpha}_{k,n}^{(H)}}, \quad (2.3.4)$$

et pour le quantile d'ordre  $p$  on a :

$$\widehat{x}_p = \left( \frac{k}{n} (1-p) \right)^{-1/\widehat{\alpha}_{k,n}^{(H)}} X_{k,n}. \quad (2.3.5)$$

D'après la formule (2.3.4), on obtient un estimateur pour la fonction de répartition excessive  $F_u(x-u)$ ,  $x \geq u$ , en utilisant  $F_u(x-u) = 1 - \overline{F}(x)/\overline{F}(u)$ .

### b. L'approche de Haan ([53])

Cette approche est basée sur une reformulation adaptée de la condition de variation régulière sur la queue droite. En effet,  $\overline{F}$  est variant régulièrement avec l'indice  $-\alpha$  si et seulement si :

$$\lim_{t \rightarrow \infty} \frac{\overline{F}(tx)}{\overline{F}(t)} = x^{-\alpha}, x > 0.$$

en utilisant une intégration partielle, on obtient :

$$\int_t^\infty (\ln x - \ln t) dF(x) = \int_t^\infty \frac{\overline{F}(x)}{x} dF(x),$$

alors que d'après le théorème de Karamata (1.7.3) on a :

$$\frac{1}{\overline{F}(x)} \int_t^\infty (\ln x - \ln t) dF(x) \rightarrow \frac{1}{\alpha}, t \rightarrow \infty. \quad (2.3.6)$$

Comment peut-on trouver un estimateur à partir de ce résultat? Deux choix doivent être faits :

a) Remplacer  $F$  par un estimateur, le candidat évident ici est la fonction de distribution empirique :

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n I_{(X_i \leq x)} = \frac{1}{n} \sum_{i=1}^n I_{(X_{i,n} \leq x)}.$$

b) Prendre  $t = X_{k,n}$  pour  $k = k(n)$  et  $t \rightarrow \infty$ .

Le choix de  $t$  est motivé par le fait que  $X_{k,n} \xrightarrow{a.s.} \infty$  a condition que  $k = k(n)$  et  $k/n \rightarrow 0$ .  
à partir (2.3.6) on a le resultat suivant :

$$\frac{1}{\overline{F}(X_{k;n})} \int_{X_{k,n}}^{\infty} (\ln x - \ln X_{k;n}) dF_n(x) = \frac{1}{k-1} \sum_{j=1}^{k-1} \ln X_{j,n} - \ln X_{k,n},$$

dans (2.3.1) en remplace le facteur  $k$  par  $k-1$  pour obtenir  $(\widehat{\alpha}^{(H)})^{-1}$ , mais pour le remplacement de  $k$  par  $k-1$  est asymptotiquement négligeable.

### Conclusion

Supposons que  $X_1, \dots, X_n$  sont des variables aléatoires iid de distribution  $F$  telle que  $\overline{F}$  varie régulièrement avec l'indice  $-\alpha$ , pour tout  $\alpha > 0$ . Alors un estimateur simple pour  $\alpha$  est donné par l'estimateur de Hill suivant :

$$\widehat{\alpha}_{k,n}^{(H)} = \left( \frac{1}{k} \sum_{j=1}^k \ln X_{j,n} - \ln X_{k,n} \right)^{-1}, \quad (2.3.7)$$

où  $k = k(n)$  satisfait  $k \rightarrow \infty$  et  $k = o(n)$ .

Ci-dessous, nous résumons les principales propriétés de l'estimateur de Hill.

### 2.4.3 Propriétés probabilistes

#### Théorème (Propriétés de l'estimateur de Hill)

Supposons que  $(X_n)$  est strictement stationnaire avec une distribution marginale  $F$ , pour  $\alpha > 0$  et  $L$  varie lentement, i.e.

$$\overline{F}(x) = p(X > x) = x^{-\alpha} L(x), x > 0.$$

a) (Uniformité faible) si on a l'une des conditions suivantes :

\*  $(X_n)$  est iid.

\*  $(X_n)$  est faiblement dépendante.

\*  $(X_n)$  est un processus linéaire.

si  $k \rightarrow \infty, k/n \rightarrow 0$  pour  $n \rightarrow \infty$ , alors :

$$\hat{\alpha}^{(H)} \xrightarrow{p} \alpha.$$

b) (Uniformité forte) si  $k/n \rightarrow 0, k/\ln \ln n \rightarrow \infty$  pour  $n \rightarrow \infty$  et  $(X_n)$  est iid, alors :

$$\hat{\alpha}^{(H)} \xrightarrow{a.s} \alpha.$$

c) (Normalité asymptotique) si d'autres conditions sur  $k$  et  $F$  sont satisfaites et  $(X_n)$  est iid, alors :

$$\sqrt{k} \left( \hat{\alpha}^{(H)} - \alpha \right) \xrightarrow{D} N(0, \alpha^2).$$

#### 2.4.4 Simulations et applications

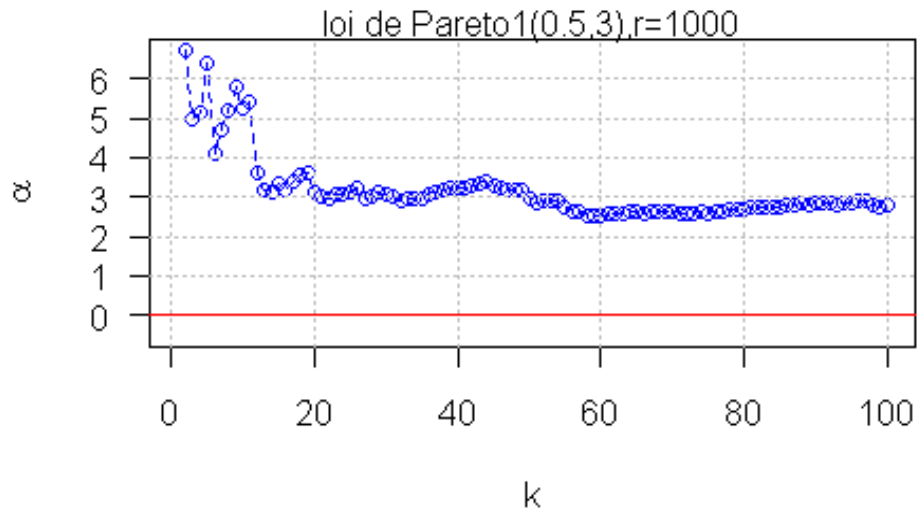
Pour bien clarifier la différence entre l'indice des distributions à queues lourdes et l'indice des distributions à queues fines, on a besoin d'une simple application sur l'estimateur de l'indice de queue  $\alpha$  (l'estimateur de Hill) et on commence par :

##### a. Indice des distributions à queues lourdes

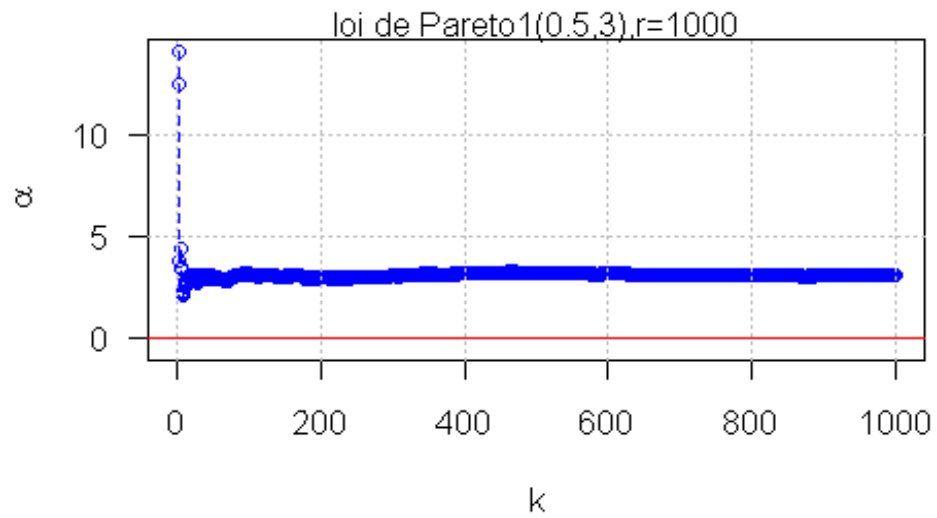
Nous générons des échantillons de variables aléatoires iid de différentes tailles

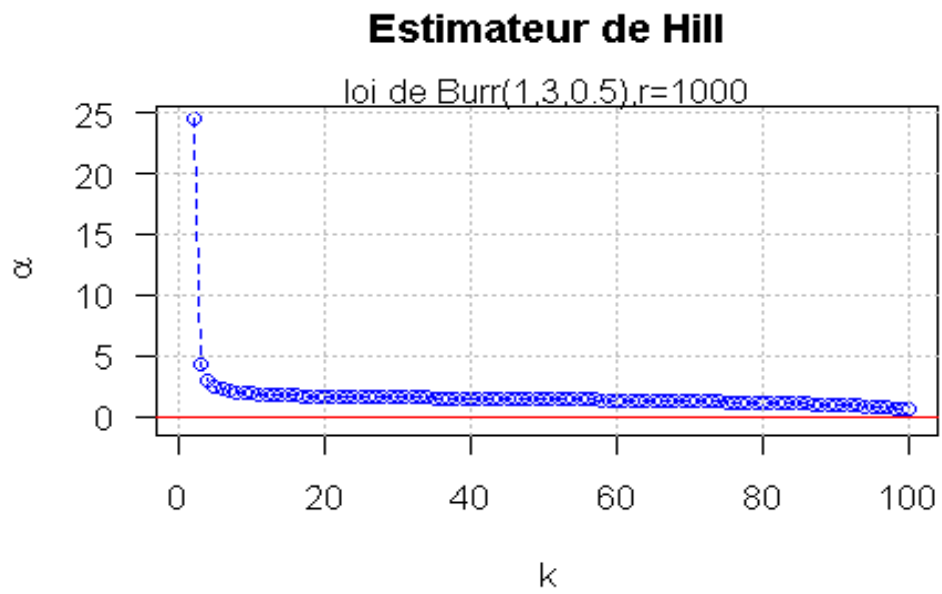
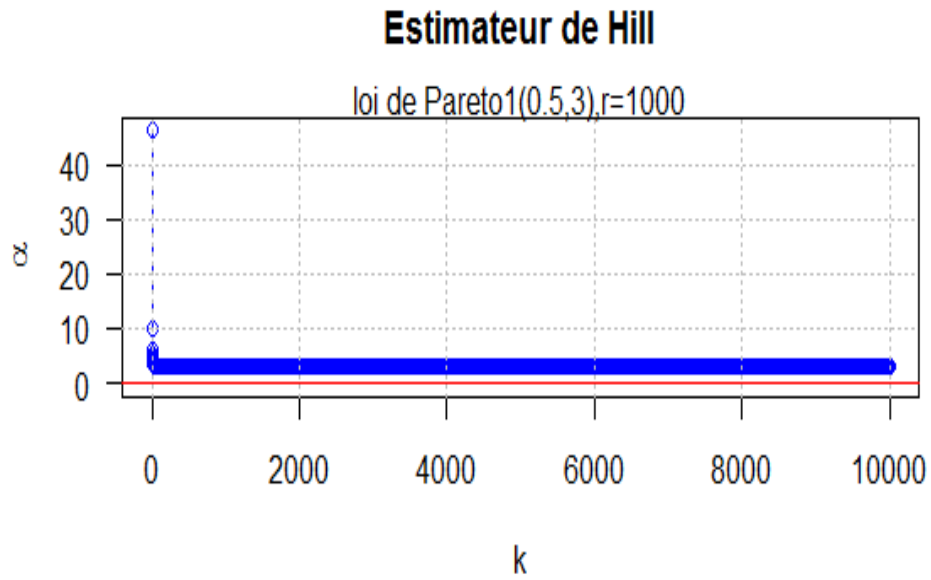
( $n = 100, n = 1000, n = 10000$ ) avec  $r$  ( $r = 1000$ ) répétitions pour chaque échantillon, pour la distribution de Pareto avec  $\alpha = 3$  et la distribution de Burr avec  $\alpha = 0, 5$ .

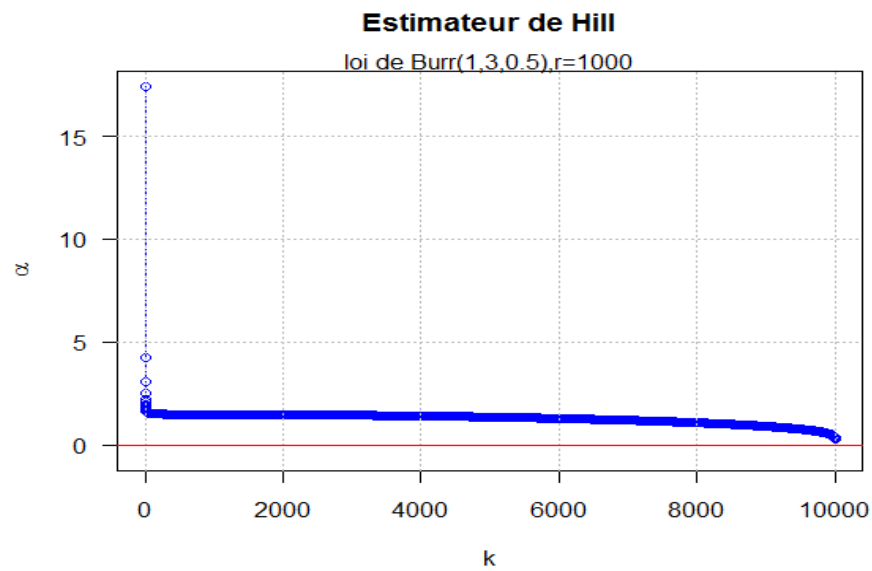
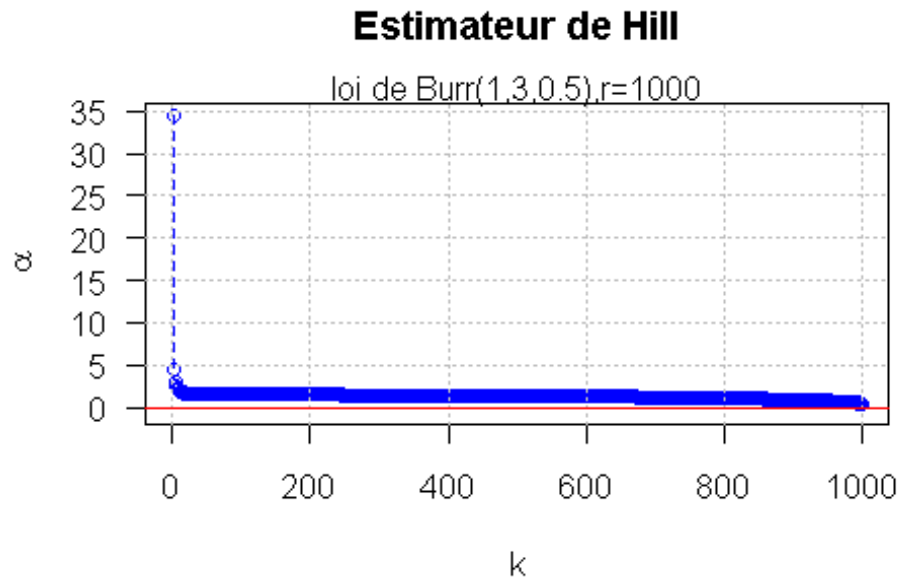
### Estimateur de Hill



### Estimateur de Hill







D'après les graphes on a :

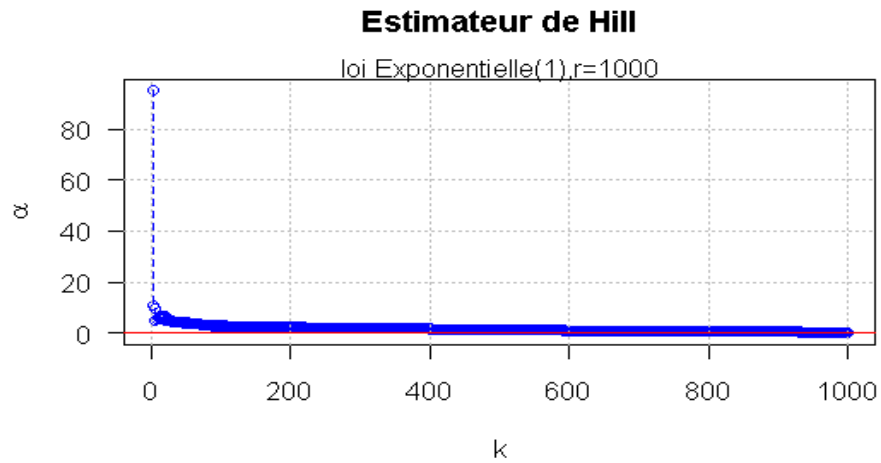
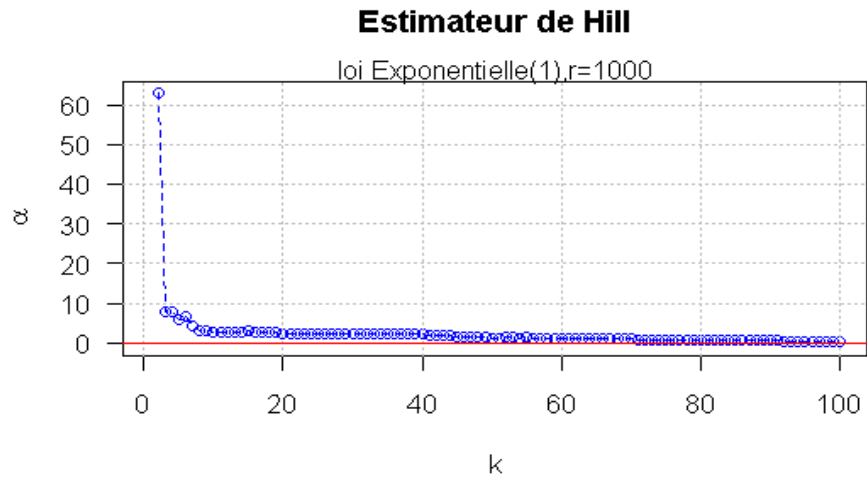
\* L'estimateur de l'indice de queue de distribution de Pareto est stable ( $\hat{\alpha} = 3$ ) à partire de  $n = 20$ .

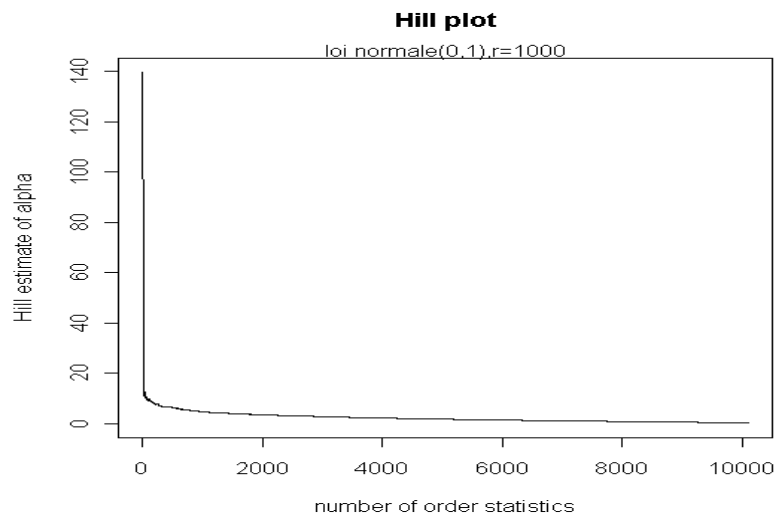
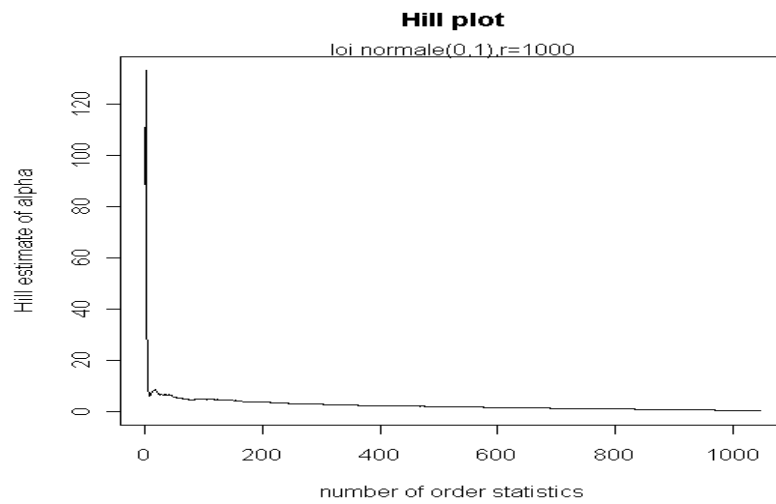
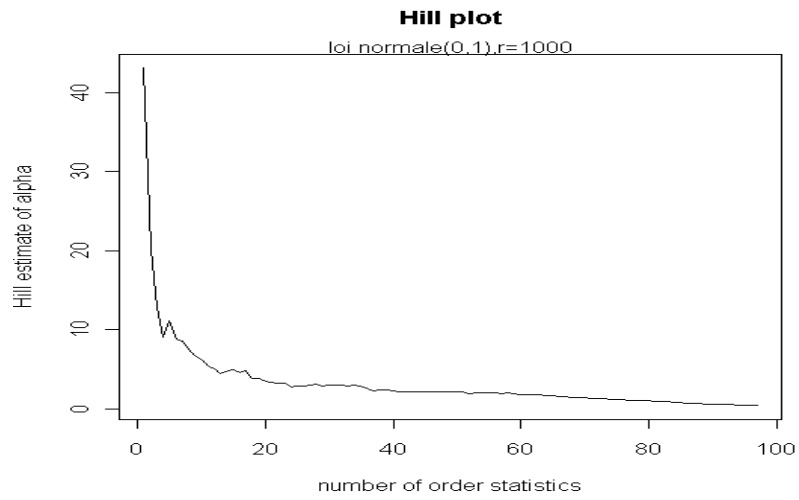
\* Pour l'indice de queue de distribution de Burr est aussi stable ( $\hat{\alpha} = 0,5$ ) à partire de  $n = 20$ .

**b. Indice des distributions à queues fines :**

Nous générons des échantillons de variables aléatoires iid de différentes tailles

( $n = 100, n = 1000, n = 10000$ ) avec  $r$  ( $r = 1000$ ) répétitions pour chaque échantillon, pour la distribution exponentielle ( $\lambda = 1$ ) et la distribution normale  $(0, 1)$ .





D'après les graphes on a :

\* L'estimateur de l'indice de queue de distribution exponentielle décroît vers zéro ( $\hat{\alpha} = 0$ ) à partir de  $n = 45$ .

\* Pour l'indice de queue de distribution normale décroît aussi vers zéro ( $\hat{\alpha} = 0$ ) à partir de  $n = 45$ .

## 2.5 Théorie des valeurs extrêmes

### 2.5.1 Introduction

La théorie des valeurs extrêmes classique cherche à comprendre le comportement des maxima ou minima d'une variable aléatoire sur une période de temps fixe. Par exemple, nous pourrions nous intéresser à la valeur maximale prise par un portefeuille pendant une semaine. Une bonne référence pour la théorie et les applications des valeurs extrêmes est le livre de Embrechts et al, (1997, p.22).

Nous nous concentrerons sur le cas où le maximum de variables puisque le cas du minimum est symétrique. Une variable aléatoire est appelée max-stable si elle est stable par rapport au maximum :

**Définition 2.5.1** Soient  $X_1, \dots, X_n$  des copies d'une variable aléatoire  $X$  de fonction de répartition  $F$ . Posons :

$$\begin{aligned} M_1 &= X_1 \\ M_n &= \max(X_1, \dots, X_n). \end{aligned}$$

Une variable aléatoire non-dégénérée  $X$  est appelée max-stable si elle satisfait l'égalité en distribution suivante,  $\forall n > 2$  :

$$M_n \stackrel{d}{=} c_n X + d_n,$$

où  $c_n > 0$  et  $d_n \in \mathbb{R}$ .

Le théorème fondamental de la théorie des valeurs extrêmes est le théorème de Fisher-Tippett qui stipule les lois limites possibles pour les maxima de variables aléatoires. S'il est possible de centrer et de réduire le maximum  $M_n$  d'une variable aléatoire de sorte qu'il converge en loi, alors cette loi est une des trois lois de valeurs extrêmes (Fréchet, Weibull et Gumbel), à une transformation affine près. Le théorème de Fisher-Tippett est analogue au théorème de la limite centrale et à sa généralisation.

**Théorème 2.5.1 (Fisher-Tippett, 1928)**

Si  $\exists c_n > 0$  et  $d_n \in \mathbb{R}$  et une fonction de répartition  $H$  telle que

$$c_n^{-1} (M_n - d_n) \xrightarrow{d} H,$$

alors  $H$  appartient à un des types de distributions suivantes :

$$\begin{aligned} \Phi_\alpha(x) &= \begin{cases} 0 & \text{si } x \leq 0 \\ \exp\{-x^{-\alpha}\} & \text{si } x > 0, \alpha > 0 \end{cases} \\ \Psi_\alpha(x) &= \begin{cases} \exp\{-(-x)^\alpha\} & \text{si } x \leq 0, \alpha > 0 \\ 1 & \text{si } x > 0 \end{cases} \\ \Lambda(x) &= \exp\{-e^{-x}\}, \quad \text{si } x \in \mathbb{R}. \end{aligned}$$

Les trois distributions limites du théorème (2.5.1) sont appelées les distributions de valeurs extrêmes standard. La première est la distribution de Fréchet, dont la fonction de répartition est donnée par  $\Phi_\alpha(x)$ , la deuxième est la distribution de Weibull, dont la fonction de répartition est  $\Psi_\alpha(x)$  et la troisième est la distribution de Gumbel, notée par  $\Lambda(x)$ . La famille des distributions centrées en  $\mu$  et de dispersion  $\sigma$  de type valeurs extrêmes est donnée en remplaçant  $x$  par

$$\frac{x-\mu}{\sigma} \text{ dans les expressions pour } \Phi_\alpha(x), \Psi_\alpha(x) \text{ et } \Lambda(x).$$

Deux fonctions de répartitions  $F_1$  et  $F_2$  sont dites du même type si  $\exists a, b > 0 \in \mathbb{R}$  tels que  $F_1(x) \stackrel{d}{=} F_2(\frac{x-a}{b})$ . D'autre part, les distributions limites du théorème (2.5.1) sont reliées de la façon suivante :

$$X \sim \Phi_\alpha \iff \ln X^\alpha \sim \Lambda \iff -X^{-1} \sim \Psi_\alpha.$$

Les distributions limites pour les minima sont obtenues en remplaçant  $x$  par  $-x$  dans les distributions du théorème (2.5.1).

**Preuve**

En effet, posons  $Y = -X^{-1}$ . Cherchons une expression pour  $F_Y(y)$ . Nous avons  $F_Y(y) = P(Y \leq y) = P(-X^{-1} \leq y) = P(X \leq -y^{-1})$  à condition que  $y \leq 0$ . Nous obtenons donc que  $F_Y(y) = F_X(-y^{-1}) = \exp\{-(-y^{-1})^{-\alpha}\} = \exp\{-(-y)^\alpha\} = \Phi_\alpha(y)$ . De même, posons maintenant  $Z = \ln X^\alpha$ . Nous avons  $F_Z(z) = P(Z \leq z) = P(\ln X^\alpha \leq z) = P(X \leq e^{z/\alpha}) = FX(e^{z/\alpha}) = \exp\{-(e^{z/\alpha})^{-\alpha}\} = \exp\{-e^{-z}\} = \Lambda(z)$ , ce qui est vrai puisque  $e^{z/\alpha} > 0$ .

**2.5.2 Domaine d’attraction maximum**

La variable aléatoire  $X$  est dans le domaine d’attraction maximum de la fonction de répartition  $H$ , ce qui s’écrit  $X \in MDA(H)$ , si

$$\exists c_n > 0 \text{ et } d_n \in \mathbb{R} \text{ tel que } c_n^{-1} (M_n - d_n) \xrightarrow{d} H.$$

Il est possible de caractériser le domaine d’attraction maximum de chacune des fonctions de répartition limites du théorème (2.5.1).

**2.5.3 Distribution de Fréchet  $\Phi_\alpha, \alpha > 0$**

Les variables aléatoires qui sont dans le domaine d’attraction maximum de la distribution de Fréchet sont caractérisées par des queues décroissant comme une puissance de  $x$  (power-law tail) :

$$1 - F(x) = x^{-\alpha} L(x),$$

où  $L(x)$  est une fonction à variation lente.

Ces distributions ont des queues épaisses, les moments d’ordre supérieurs à  $\alpha$  n’existent pas ( $E[(X^+)^\delta] = \infty, \forall \alpha > \alpha, \text{ où } X^+ = \max(0, X)$ ). Ce qui signifie que pour  $\alpha \geq 2$ , la variance est infinie et pour  $\alpha \geq 1$ , l’espérance aussi est infinie. De plus, toute distribution dans le domaine d’attraction maximum de Fréchet a un point extrême droit infini ( $x_F = \infty, \text{ où } x_F = \sup\{x \in \mathbb{R} : F(x) < 1\}$  (est le point extrême droit du support de la distribution).

Les distributions suivantes appartiennent au domaine d'attraction maximum de Fréchet :

- la distribution de Pareto :

$$1 - F(x) \approx Kx^{-\alpha}, x \rightarrow \infty,$$

pour  $K, \alpha > 0$ .

- la distribution de Cauchy :

$$f(x) = \frac{1}{\pi\beta(1 + (x - \mu)^2/\beta^2)},$$

où  $\mu, \beta$  sont les paramètres d'emplacement et de dispersion.

- les distributions  $\alpha$ -stables avec  $\alpha > 2$ .

#### 2.5.4 Distribution de Weibull $\Psi_\alpha, \alpha > 0$

Les distributions appartenant au domaine d'attraction maximum de Weibull ont le point extrême droit fini ( $x_F < \infty$ ) et la queue supérieure de la distribution se comporte comme une puissance de  $x$  :

$$1 - F(x_F - x^{-1}) = x^{-\alpha}L(x),$$

où  $L(x)$  est une fonction à variation lente.

Les distributions suivantes appartiennent au domaine d'attraction maximum de Weibull :

- la distribution uniforme sur  $(0, 1)$ .

- les distributions décroissant comme une puissance de  $x$  au point extrême droit fini.

- la distribution Bêta.

#### 2.5.5 Distribution de Gumbel $\Lambda$

Le domaine d'attraction maximale de Gumbel est plus difficile à caractériser. Des distributions à queues modérément épaisses comme la distribution

lognormale et des distributions à queues légères comme la famille des distributions exponentielles (qui comprend la distribution normale) s'y retrouvent. Le point extrême droit peut être aussi bien fini qu'infini .

### 2.5.6 Distribution de valeurs extrêmes généralisée (*GEV*)

La distribution de valeurs extrêmes généralisée est simplement une reparamétrisation des distributions apparaissant dans le théorème (2.5.1) :

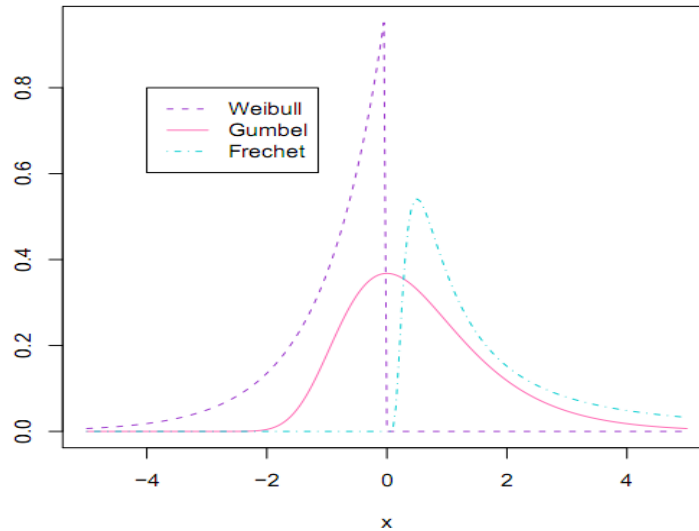
$$H_{\xi,\mu,\beta}(x) = \begin{cases} \exp \left\{ - \left( 1 + \xi \left( \frac{x-\mu}{\beta} \right) \right)^{-1/\xi} \right\} & \text{si } \xi \neq 0 \\ \exp \left\{ - \exp \left\{ - \frac{x-\mu}{\beta} \right\} \right\} & \text{si } \xi = 0 \end{cases},$$

où  $1 + \xi \left( \frac{x-\mu}{\beta} \right) > 0$ ,  $\mu$  et  $\beta$  sont les paramètres d'emplacement et de dispersion respectivement.

Dépendant des valeurs prises par le paramètre  $\xi$ , nous retrouvons chacun des types de distributions du théorème (2.5.1) :

$$\begin{aligned} \xi > 0 &\implies \text{Distribution de Fréchet } \Phi_\alpha \\ \xi = 0 &\implies \text{Distribution de Gumbel } \Lambda \\ \xi < 0 &\implies \text{Distribution de Weibull } \Psi_\alpha. \end{aligned}$$

Ces densités sont illustrées à la figure suivante :



Loi des valeurs extrêmes généralisée

### 2.5.7 Estimation des paramètres

Selon le contexte, les distributions de valeurs extrêmes peuvent être utilisées pour modéliser les maxima ou minima d'une variable aléatoire sur une période donnée. Pour les applications en finance et en assurance, les distributions de la famille de Gumbel et de Fréchet sont les modèles les plus importants pour tenir compte des événements extrêmes. Les distributions dans le domaine d'attraction maximum de Weibull ont un point extrême droit fini ( $x_F < \infty$ ) ce qui contraint les observations à être bornées supérieurement.

Soient  $X_1, \dots, X_n$  des variables aléatoires indépendantes et identiquement distribuées (i.i.d) ayant pour fonction de répartition  $F \in MDA(H_\xi)$  (où  $H_\xi = H_{\xi, \mu=0, \beta=1}$ ).

Un estimateur possible pour le paramètre de queue  $\xi$  est l'estimateur de Hill :

$$\hat{\zeta}_{n,k}^{(H)} = \frac{1}{k} \sum_{j=1}^k \ln X_{j,n} - \ln X_{k+1,n},$$

où  $X_{1,n} \geq X_{2,n}, \dots, \geq X_{n,n}$  sont les statistiques d'ordre et  $k = k(n)$  doit être choisi de

façon appropriée. Un autre estimateur possible est l'estimateur de Pickands :

$$\widehat{\zeta}_{n,k}^{(p)} = \frac{1}{\ln 2} \ln \left( \frac{X_{k,n} - X_{2k,n}}{X_{2k,n} X_{4k,n}} \right),$$

où  $k = k(n)$  comme pour l'estimateur de Hill. Embrechts et al.(1997, p. 22) propose de choisir  $k$  de sorte que l'estimateur soit relativement stable.

En général, si  $F \in MDA(H_{\xi,\mu,\beta})$ , les paramètres  $\xi$ ,  $\mu$  et  $\beta$  peuvent être estimés par maximum de vraisemblance, ou par la méthode des moments

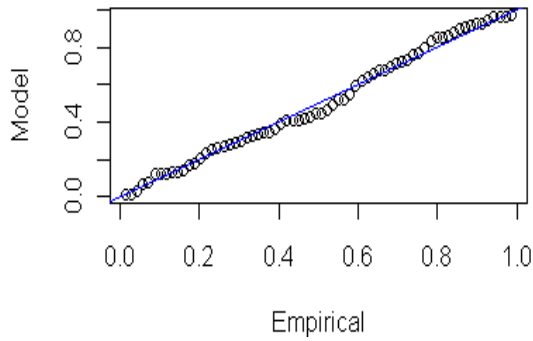
pondérés de façon probabiliste. Lorsque  $\mu$  et  $\beta$  sont estimés, les observations peuvent être standardisées et le paramètre  $\xi$  peut être estimé par une des méthode mentionnée ci-haut (Hill ou Pickands et *MLE*).

Par exemple, la variable aléatoire d'intérêt pourrait être la valeur quotidienne d'un portefeuille notée  $\widehat{X}_j$ . Pour étudier le risque mensuel du portefeuille, la variable suivante est utilisée pour l'estimation des paramètres :  $X_j = \min\{X_{j_1}, \dots, X_{j_n}\}$ . Donc, la variable  $X_j$  représente le maximum de la valeur du portefeuille sur un mois. Nous faisons l'hypothèse que ces maxima sont indépendants et identiquement distribués, ce qui n'est pas souvent justifié. Certaines techniques existes pour éliminer ce problème.

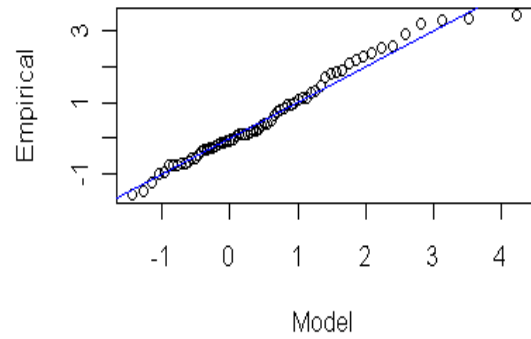
### 2.5.8 Application

On s'est basé sur des données concrètes fournies par l'organisme public, Office National de la Météo (*ONM*). Ces données sont constituées d'observations (Sur la pluie) au niveau de six stations : Alger, Annaba, Bouzaréah, Ghardaïa, Constantine, et Oran.

**Residual Probability Plot**

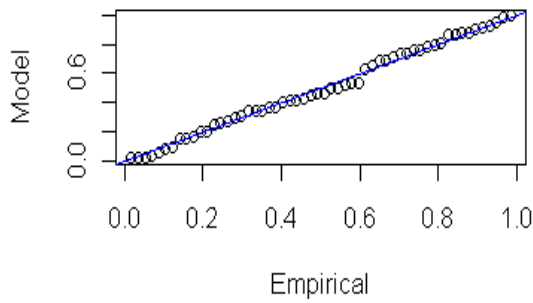


**Residual Quantile Plot (Gumbel Scale)**

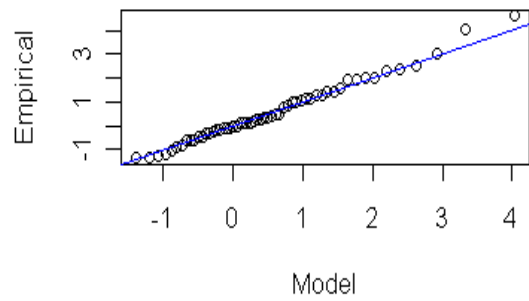


*station Alger*

**Residual Probability Plot**

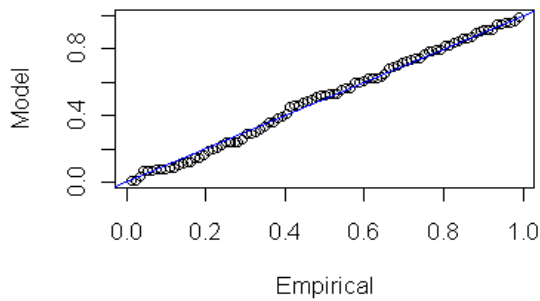


**Residual Quantile Plot (Gumbel Scale)**

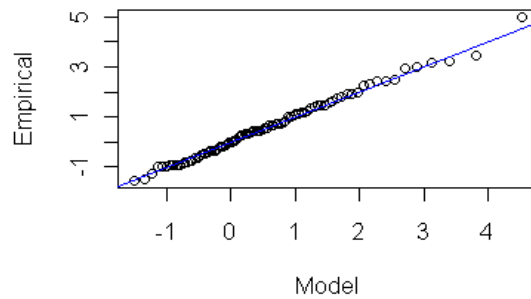


*station Annaba*

**Residual Probability Plot**

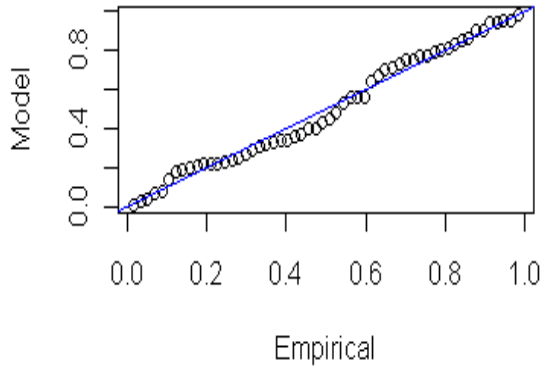


**Residual Quantile Plot (Gumbel Scale)**

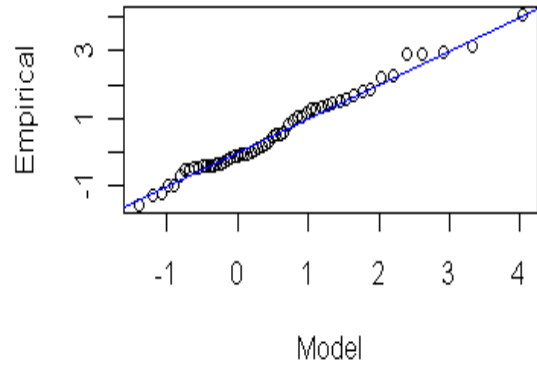


*station Bouzaréah*

**Residual Probability Plot**

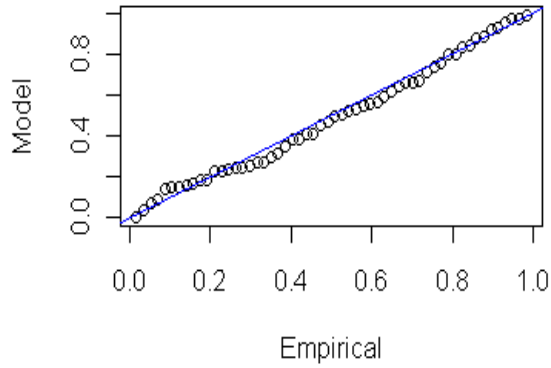


**Residual Quantile Plot (Gumbel Scale)**

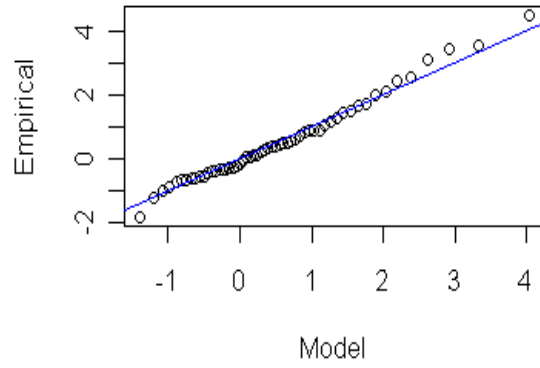


*station Constantine*

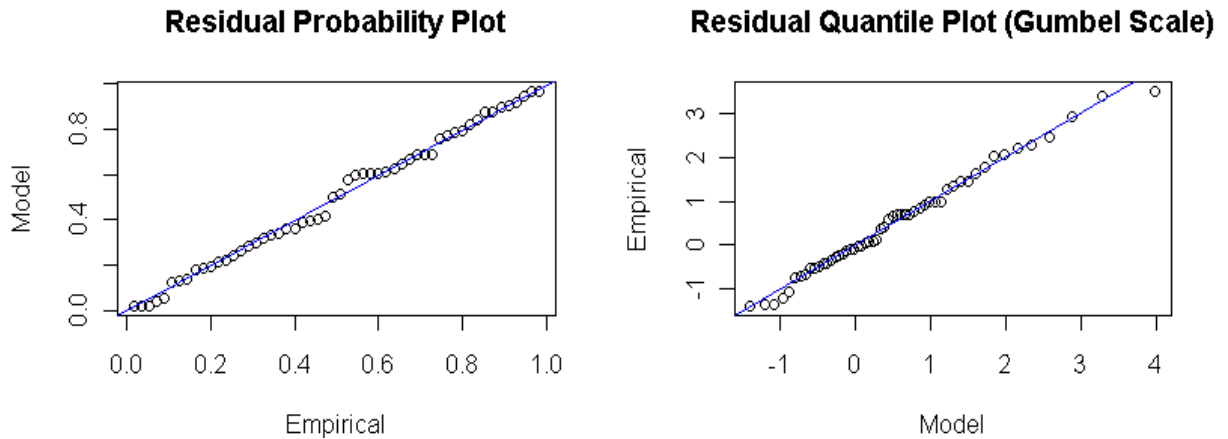
**Residual Probability Plot**



**Residual Quantile Plot (Gumbel Scale)**



*station Oran*



station Ghardaïa

**a. Estimation de paramètre  $\zeta$  de la  $GEV$  par la méthode  $MLE$  ,pour les six tations étudiées**

Les résultats de l'estimation de  $\zeta$  obtenus, pour les données des stations étudiées, sont mentionnés dans le tableau suivant :

	Alger	Annaba	Bouzareah	Constantine	Oran	Ghardaïa
MLE	0.36634	0.01609	0.18565	0.18911	0.19317	0.25455

Estimation de paramètre  $\zeta$  de la  $GEV$  par la méthode  $MLE$  ,pour les six tations étudiées.

**Conclusion**

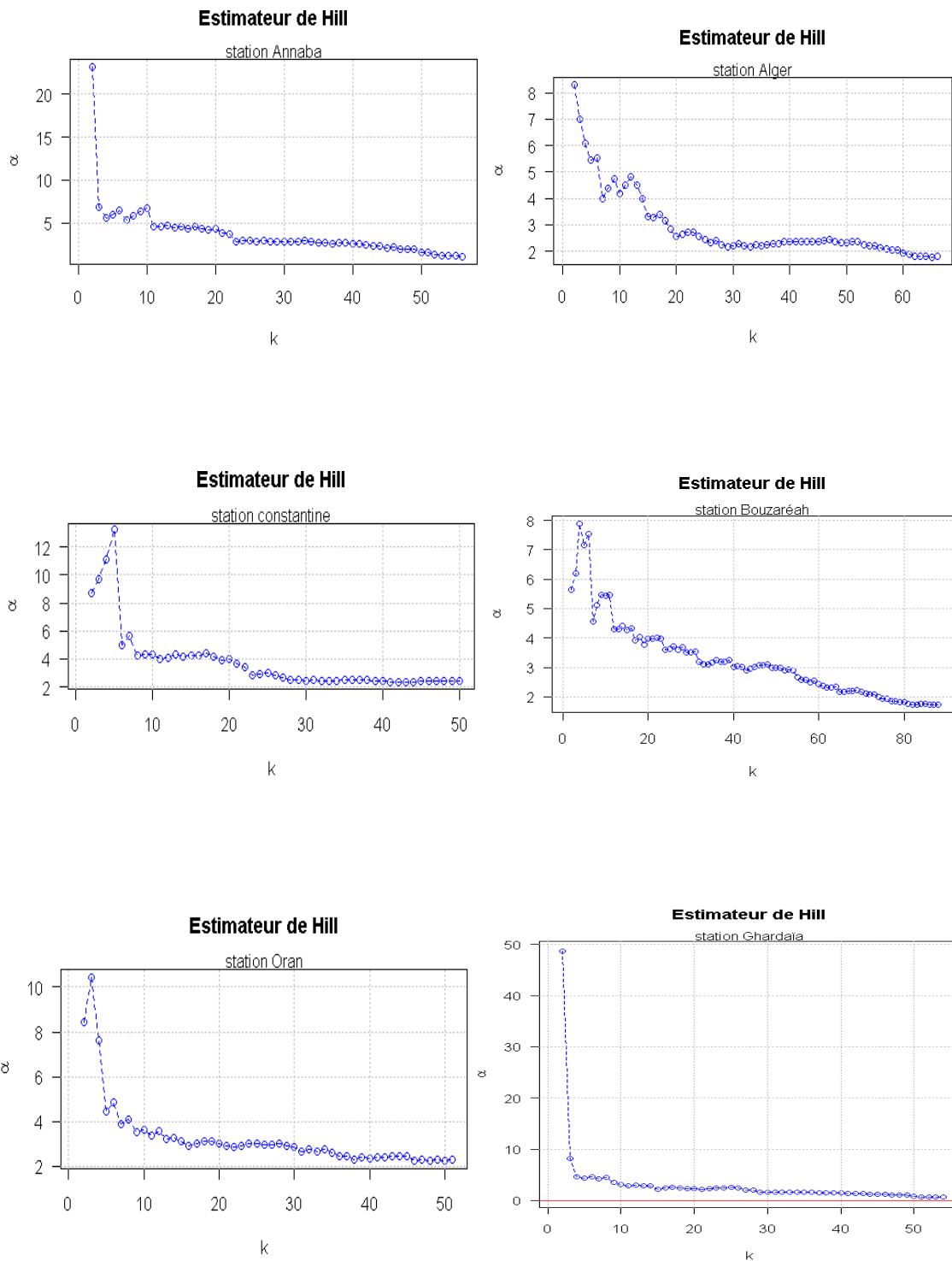
D'après ce tableau l'estimateur de  $\zeta$  est positif ( $\hat{\zeta} > 0$ ), pour toute les stations, on est donc dans le cas de distributions de Fréchet  $\Phi_\alpha$ .

**b. Estimateur de Hill et Pickands de paramètre  $\zeta$ , pour les six tations étudiées**

Les résultats de l'estimation de  $\zeta$  obtenus, pour les données des stations étudiées, sont mentionnés dans le tableau suauivant :

	Alger	Annaba	Bouzareah	Constantine	Oran	Ghardaïa
Hill	0.3944756	0.3600830	0.3651205	0.3470279	0.3598294	0.5942026
Pickands	0.2958	0.2635	0.31245	0.142354	0.21453	0.23564

Estimateur de Hill et Pickands de paramètre  $\zeta$ , pour les six tations étudiées



Les six graphes précédents représentent l'estimateur de Hill ( $\alpha = 1/\zeta$ ) de l'indice de queue pour toute les stations.

### Conclusion

D'après les graphiques de l'estimateur de Hill précédents, nous permet d'avoir des estimations du paramètre ( $\alpha = 1/\zeta$ ) en fonction de l'ordre statistique le plus élevé, nous choisissons ainsi l'indice le plus stable.

La stabilité est atteinte pour  $\alpha$  dans le graphique de chaque station comme suit :

- $\hat{\alpha} = 1.70$  pour Alger.
- $\hat{\alpha} = 1.10$  pour Annaba.
- $\hat{\alpha} = 2.30$  pour Constantine.
- $\hat{\alpha} = 1.70$  pour Bouzaréah.
- $\hat{\alpha} = 0.50$  pour Ghardaïa.
- $\hat{\alpha} = 2.20$  pour Oran.

On peut aussi représenter, le Hill plot en utilisant les seuils (en valeur) au lieu du nombre de valeurs extrêmes.

La représentation ci-dessus nous permet d'estimer le quantile extrême (seuil) en se servant de l'indice de queue de la distribution retrouvé par la même méthode.

L'estimateur de Hill est donc un outil à double utilité :

- L'estimation de l'indice de la queue de la distribution.
- L'estimation du seuil.

de Haan (1994) a montré que l'estimateur de Hill de l'indice de queue converge en probabilité vers sa vraie valeur et cela pour tout  $\alpha$  réel.

L'estimateur de Hill est plus fiable dans le domaine d'attraction de Fréchet pour lequel il fournit un estimateur de l'indice de queue plus efficace que d'autres méthodes et pour confirmer on a le tableau suivant qui présente l'erreur quadratique moyenne de chaque méthode :

	Alger	Annaba	Bouzareah	Constantine	Oran	Ghardaïa
MLE	4.93072	22.43047	44.39004	7.14494	3.62086	3.43292
Hill	1.3156	3.0708	1.3179	2.3165	1.5604	1.5166
Pickands	2.1245	4.1236	3.3562	3.4536	2.3636	2.6213

### 2.5.9 Conclusion

La distribution de valeurs extrêmes généralisée (*GEV*) intervient comme distribution limite de maxima (ou minima) normalisés. Cette distribution peut donc être utilisée pour modéliser le maximum d'une variable aléatoire sur une période donnée.

Une autre alternative, qui ne relève pas de la théorie des valeurs extrêmes traditionnelle, est d'utiliser les distributions limites (Weibull, Gumbel et Fréchet) pour modéliser directement des distributions de variables aléatoires ayant des queues sous-exponentielles.

# Chapitre 3

## Processus stochastiques linéaires à queues lourdes et valeurs extrêmes

### 3.1 Introduction

Nous présentons dans ce chapitre quelques résultats concernant l'analyse des séries chronologiques, caractérisées par des grandes fluctuations notamment les séries financières. De tels séries ont la propriété que la distribution marginale est à queues lourdes, dans le sens où la décroissance vers 0 n'est pas exponentielle. Mais juste avant on a besoin de rappeler quelques définitions et résultats de base concernant l'analyse des séries chronologiques classiques, cela peut être utilisé dans l'analyse des séries chronologiques à queues lourdes.

L'analyse des séries chronologiques est principalement concernée par l'analyse statistique des processus stationnaires et, en particulier, des processus linéaires de la forme suivante :

$$Y_t = \sum_{j=-\infty}^{\infty} \psi_j \varepsilon_{t-j}, t \in \mathbb{Z}, \quad (3.1.1)$$

avec  $\{\varepsilon_t, t \in \mathbb{Z}\}$  est un processus bruit blanc (d'innovation) de moyenne nulle et de variance finie c'est-à-dire  $(Z_t \rightarrow WN(0, \sigma_\varepsilon^2))$ . Pour des raisons de normalisation nous exigeons également que  $\psi_0 = 1$ . En pratique, on dit que les représentations de la forme (3.1.1) sont causals si et seulement si  $\psi_j = 0$  pour  $j < 0$ .

Pour  $t$  fixé les séries dans la représentation (3.1.1) convergent asymptotiquement à condi-

tion que les coefficients à valeurs réelles  $\psi_j$  satisfont la condition suivante :

$$\text{var}(Y_t) = \sigma_\varepsilon^2 \sum_{j=-\infty}^{\infty} \psi_j^2 < \infty. \quad (3.1.2)$$

## 3.2 Processus Bruit Blanc (White Noise)

Le plus simple processus stationnaire en analyse des séries temporelles est appelé : processus bruit blanc  $\{\varepsilon_t, t \in \mathbb{Z}\}$  qui est une suite de variables aléatoires non corrélées de moyenne nulle et de variance constante  $\sigma_\varepsilon^2$ .

Le fait que les variables aléatoires  $(\varepsilon_t)_{t \in \mathbb{Z}}$  soient mutuellement non corrélées (hypothèse d'orthogonalité), nous permet de donner la fonction d'autocovariance de ce processus par :

$$\gamma_\varepsilon(h) = \text{cov}(\varepsilon_t, \varepsilon_{t+h}) = E(\varepsilon_t \varepsilon_{t+h}) = \begin{cases} \sigma_\varepsilon^2 & \text{si } h = 0 \\ 0 & \text{si } h \neq 0 \end{cases}.$$

Par conséquent, la fonction d'autocorrélation est donnée par :

$$\rho_\varepsilon(h) = \begin{cases} 1 & \text{si } h = 0 \\ 0 & \text{si } h \neq 0 \end{cases}.$$

## 3.3 Classe des modèles ARMA

**Définition 3.3.1** Les processus linéaires les plus populaires sont des processus  $ARMA(p, q)$  qui sont données par :

$$Y_t - \phi_1 Y_{t-1} - \dots - \phi_p Y_{t-p} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}, t \in \mathbb{Z}. \quad (3.3.1)$$

Comme on peut exprimer le modèle  $ARMA(p, q)$  par l'équation aux différences :

$$\Phi(B)Y_t = \Theta(B)\varepsilon_t,$$

avec  $\varepsilon_t$  est un bruit blanc de moyenne nulle et de variance  $\sigma_\varepsilon^2$  et :

$$\Theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q \text{ et } \Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$$

$$\theta_j \in \mathbb{R}, \forall j = 1, \dots, q \text{ et } \phi_i \in \mathbb{R}, \forall i = 1, \dots, p.$$

**Théorème 3.3.1 (Condition de stationnarité et d'inversibilité) :**

1. Une condition nécessaire et suffisante pour que le processus autorégressif moyenne mobile d'ordre  $(p, q)$  soit stationnaire est que les racines de l'équation caractéristique suivante :  $\Phi(Z) = 0$  soient à l'extérieur du cercle unitaire.

2. Une condition nécessaire et suffisante pour que le processus autorégressif moyenne mobile d'ordre  $(p, q)$  soit inversible est que les racine de l'équation caractéristique suivante :  $\Theta(Z) = 0$  soient à l'extérieur du cercle unitaire.

**Caractéristique d'un Processus ARMA  $(p, q)$  :**

Les corrélogrammes simple et partiel sont un mélange des fonctions exponentielles et sinusoïdales amorties. Cependant l'identification des paramètres  $p$  et  $q$  à partir de l'étude des fonctions d'autocorrélations empiriques s'avère plus délicate.

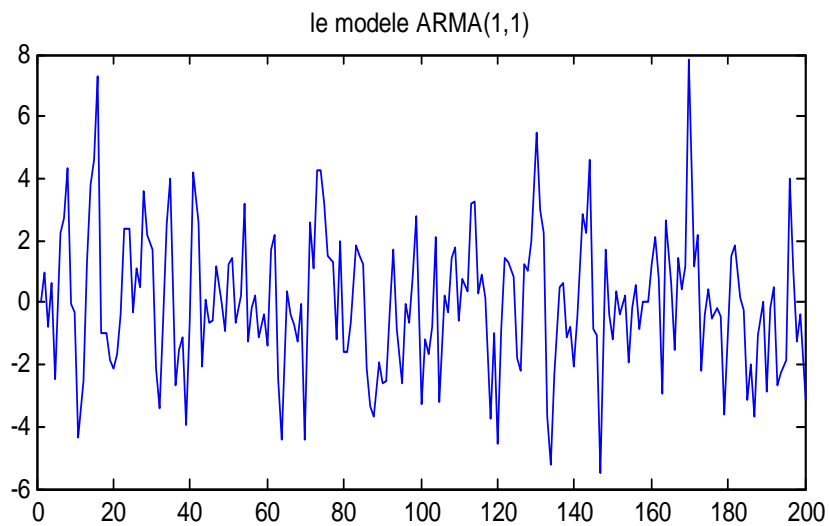
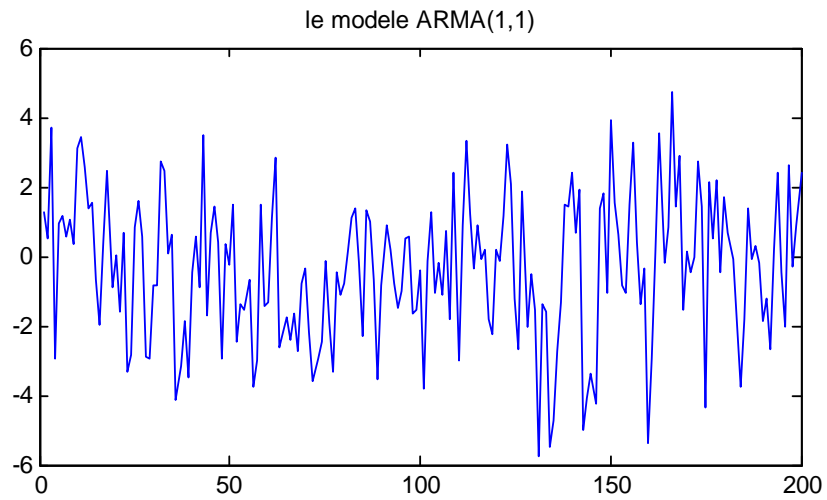
**Remarques :**

1. Les processus *AR*, *MA* et *ARMA* ne sont représentatifs que de séries stationnaires en tendance et corrigées des variations saisonnières. De plus ces modèles ne tiennent pas compte d'une éventuelle variable exogène.

2. dans un contexte ultérieur, nous allons exposer une extension de la classe des modèles *ARMA*.

3. L'ordre  $(p, q)$  sont généralement déterminé par l'intermédiaire d'un critère de sélection d'ordre. Les paramètres  $\phi_i$  et  $\theta_i$  satisfont certaines conditions afin de garantir que l'équation (3.3.1) a une solution ce qui peut être exprimé sous la forme (3.1.1).

Parmi les cas spéciaux de processus *ARMA*  $(p, q)$  on a le processus moyenne mobile d'ordre  $q$  (noté par *MA*  $(q)$ ) et le processus autorégressif d'ordre  $p$  (noté par *AR*  $(p)$ ).



**Figure1.** Estimation du modèle  $ARMA(1, 1)$ ,  $(Y_t - 0, 4Y_{t-1} = 0, 3\varepsilon_{t-1} + \varepsilon_t)$ ,  
avec 200 réalisations de  $\{\varepsilon_t\}$  iid de loi  $N(0, 2)$

### 3.4 Processus autorégressif d'ordre $p$

**Définition 3.4.1** Le processus  $\{Y_t, t \in \mathbb{Z}\}$  satisfait à une représentation  $AR$  d'ordre  $p$ , noté  $AR(p)$ , s'il est solution de l'équation aux différences stochastique suivante :

$$\varepsilon_t = Y_t - \sum_{j=1}^p \phi_j Y_{t-j},$$

ou encore

$$\varepsilon_t = \Phi(B) Y_t \text{ avec } \Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \text{ et } \phi_p \in \mathbb{R}^*,$$

où  $\Phi(B)$  représente le polynôme de retard et  $\varepsilon_t$  est un bruit blanc de moyenne nulle et de variance  $\sigma_\varepsilon^2$ .

**Remarque**

Le degré  $p$  est appelé le degré d'autocorrélation, il indique la profondeur de la mémoire.

**Théorème 3.4.1 (Condition de stationnarité)**

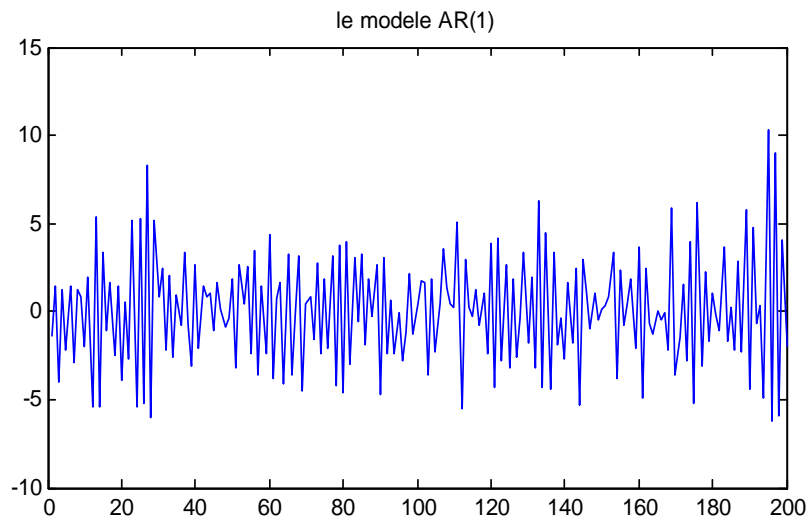
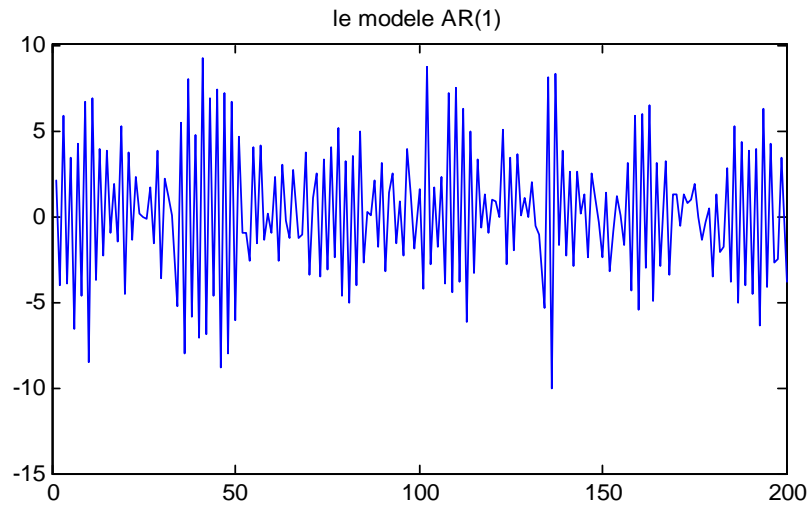
Une condition nécessaire et suffisante pour que le processus autorégressif soit stationnaire du second ordre est que les racines de l'équation caractéristique suivante  $\Phi(Z) = 0$  soient à l'extérieur du cercle unitaire.

**Cas particulier :**

Soit le processus stationnaire  $AR(1)$  : Ce processus est dit processus de Markov car l'observation dépend seulement de l'observation précédente  $Y_{t-1}$  et qui s'écrit comme suit :

$$\begin{aligned} Y_t &= \varepsilon_t + \phi_1 Y_{t-1} & (3.4.1) \\ &= \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_1^2 Y_{t-2} \\ &= \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_1^2 \varepsilon_{t-2} + \phi_1^3 \varepsilon_{t-3} + \dots \\ &= \sum_{j=0}^{\infty} \phi_1^j \varepsilon_{t-j} \end{aligned}$$

la condition  $|\phi_1| < 1$  est évidemment nécessaire afin de justifier la dernière égalité, dans ce cas on dit (3.1.2) converge, et (3.4.1) converge asymptotiquement.



**Figure1** . Estimation du modèle  $AR(1)$ ,  $(Y_t = -1.8Y_{t-1} + \varepsilon_t)$ ,  
avec 200 réalisations de  $\{\varepsilon_t\}$  iid de loi  $N(0, 2)$

### 3.5 Processus Moyenne Mobile d'ordre $q$ (Moving Average)

**Définition 3.5.1** Le processus  $\{Y_t, t \in \mathbb{Z}\}$  satisfait à une représentation moyenne mobile d'ordre  $q$ , noté :  $MA(q)$ , s'il est solution de l'équation aux différences stochastique suivante :

$$Y_t = \varepsilon_t + \sum_{j=1}^q \theta_j \varepsilon_{t-j}.$$

En introduisant le polynôme de retard nous obtenons :

$$Y_t = \Theta(B) \varepsilon_t \text{ où } \Theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q \text{ et } \theta_j \in \mathbb{R}^*,$$

et  $\varepsilon_t$  est un bruit blanc de moyenne nulle et de variance  $\sigma_\varepsilon^2$ .

#### Remarque

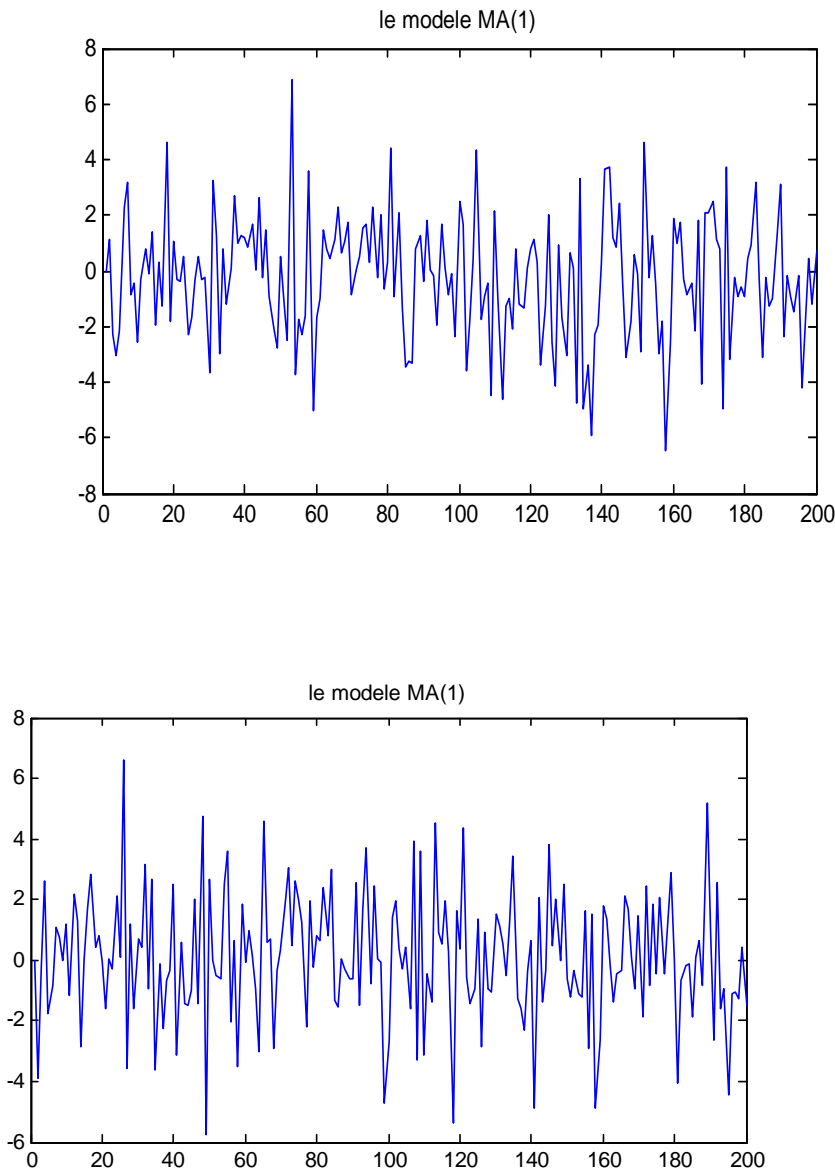
Le modèle moyenne mobile d'ordre  $q$  ( $MA(q)$ ), explique la valeur de la série à l'instant  $t$  par une moyenne pondérée d'aléas  $\varepsilon_t$ , jusqu'à la  $q^{\text{ième}}$  période qui sont supposées être générées par un processus de type bruit blanc.

#### Théorème 3.5.1 (Condition d'inversibilité) :

Une condition nécessaire et suffisante pour que le processus moyenne mobile  $MA(q)$  soit inversible est que les racines de l'équation caractéristique suivante :  $\Theta(Z) = 0$  soient à l'extérieur du cercle unitaire.

#### Exemple :

Soit le processus  $MA(1)$  :  $Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1}$  dans ce cas  $\Theta(Z) = 1 - \theta Z$  alors :  $\Theta(Z) = 0 \Rightarrow 1 - \theta Z = 0 \Rightarrow Z = \frac{1}{\theta} \Rightarrow |Z| > 1 \Rightarrow |\theta| < 1$ .



**Figure 1** . Estimation du modèle moyenne mobile  $MA(1)$ ,  $(Y_t = 0.4\varepsilon_{t-1} + \varepsilon_t)$ ,  
avec 200 réalisations de  $\{\varepsilon_t\}$  iid de loi  $N(0, 2)$

### Caractéristiques d'un Processus $MA(q)$

Un processus moyenne mobile d'ordre  $q$  est toujours stationnaire, car il est une combinaison linéaire finie d'un processus stationnaire  $\{\varepsilon_t, t \in \mathbb{Z}\}$ .

Pour le corrélogramme simple seuls ses  $q$  premiers termes sont différents de zéro. La fonction d'autocorrélation est dite tronquée au-delà du  $q^{\text{ième}}$  retard puisque la fonction d'au-

la corrélation est définie par :

$$\rho_k = \begin{cases} \frac{\sum_{i=0}^{q-k} \theta_i \theta_{i+k}}{\sum_{i=0}^q \theta_i^2} & \text{pour } k = 0, \dots, q \\ 0 & \text{pour } k > q \end{cases} .$$

Le corrélogramme partiel est caractérisé par une décroissance exponentielle de ses termes.

### Remarques

1. Un processus autorégressif est toujours inversible.
2. Il y a une équivalence entre le processus  $MA(1)$  et le processus  $AR(p)$  avec  $p$  infini.

## 3.6 Processus Moyenne Mobile d'ordre infini ( $MA(\infty)$ )

**Définition 3.6.1** On dit que le processus  $\{Y_t, t \in \mathbb{Z}\}$  est un processus moyenne mobile d'ordre infini ( $MA(\infty)$ ), s'il existe une suite  $\{\Psi_j\}$  avec  $\sum_{j=0}^{\infty} |\Psi_j| < \infty$  tel que :

$$Y_t = \sum_{j=0}^{\infty} \Psi_j \varepsilon_{t-j} \quad \text{avec } \varepsilon_t \rightarrow WN(0, \sigma^2), t = 0, \pm 1, \pm 2, \dots \quad (3.6.1)$$

**Théorème 3.6.1** (e.g Brockwell . Davis, 1991)

Le processus  $MA(\infty)$  qui est définie dans (3.5.1) est de moyenne nulle et de fonction d'autocovariance :

$$\gamma(h) = \sigma^2 \sum_{j=0}^{\infty} \Psi_j \Psi_{j+|h|},$$

et de fonction d'autocorrélation :

$$\rho(h) = \frac{\sum_{j=-\infty}^{\infty} \Psi_j \Psi_{j+|h|}}{\sum_{j=-\infty}^{\infty} \Psi_j^2} .$$

### 3.7 Densité spectrale et le processus $ARMA(p, q)$

**Théorème 3.7.1** (e.g Brockwell . Davis,1991)

1. Soit  $\{\varepsilon_t, t \in \mathbb{Z}\}$  est un processus stationnaire à valeur complexe et de moyenne nulle avec une fonction de distribution spectrale  $F_\varepsilon(\cdot)$ , et soit  $\{Y_t, t \in \mathbb{Z}\}$  est un processus de la forme suivante :

$$Y_t = \sum_{j=-\infty}^{\infty} \Psi_j \varepsilon_{t-j} \quad \text{où} \quad \sum_{j=-\infty}^{\infty} |\Psi_j| < \infty, \quad (3.7.1)$$

alors le processus  $\{Y_t, t \in \mathbb{Z}\}$  est stationnaire et de fonction de distribution spectrale de la forme suivante :

$$F_Y(\lambda) = \int_{-\pi}^{\lambda} \left| \sum_{j=-\infty}^{\infty} \Psi_j e^{-ijv} \right|^2 dF_\varepsilon(v), \quad -\pi \leq \lambda \leq \pi. \quad (3.7.2)$$

2. Si  $\{\varepsilon_t, t \in \mathbb{Z}\}$  est de densité spectrale  $f_\varepsilon(\cdot)$  et si  $\{Y_t, t \in \mathbb{Z}\}$  est défini par la représentation (4.7.1), alors  $\{Y_t, t \in \mathbb{Z}\}$  est aussi de densité spectrale  $f_Y(\cdot)$  qui donnée par la formule suivante :

$$f_Y(\lambda) = |\Psi(e^{-i\lambda})|^2 f_\varepsilon(\lambda), \quad (3.7.3)$$

avec  $\Psi(e^{-i\lambda}) = \sum_{j=-\infty}^{\infty} \Psi_j e^{-ij\lambda}$ . L'opérateur  $\psi(B) = \sum_{j=-\infty}^{\infty} \psi_j B^j$  appliqué à  $\{\varepsilon_t, t \in \mathbb{Z}\}$

dans la représentation (4.7.1) s'appelle un filtre linéaire invariant dans le temps avec des poids  $\{\psi_j\}$ . La fonction  $\psi(e^{-i\lambda})$  est appelé la fonction de transfert du filtre et le module carré  $|\psi(e^{-i\lambda})|^2$  est mentionnée comme la fonction de transfert du filtre.

Comme une application de ce théorème, nous pouvons maintenant tirer la densité spectrale d'un processus  $ARMA(p, q)$ .

### 3.7.1 Densité spectrale de processus $ARMA(p, q)$

**Théorème 3.7.2** (e.g Brockwell . Davis,1991)

Soit  $\{Y_t, t \in \mathbb{Z}\}$  un processus  $ARMA(p, q)$  satisfait à l'équation aux différences stochastique suivante :

$$\Phi(B) Y_t = \Theta(B) \varepsilon_t \quad \text{avec } \varepsilon_t \sim WN(0, \sigma^2), \quad (3.7.4)$$

où  $\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$  et  $\Theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q$  n'ont pas de zéros communs et  $\Phi(B)$  n'a pas de zéros sur le cercle unité. Alors  $\{Y_t, t \in \mathbb{Z}\}$  est de densité spectrale suivante :

$$f_Y(\lambda) = \frac{\sigma^2 |\Theta(e^{-i\lambda})|^2}{2\pi |\Phi(e^{-i\lambda})|^2}, \quad -\pi \leq \lambda \leq \pi. \quad (3.7.5)$$

#### Exemple

Soit  $\{Y_t, t \in \mathbb{Z}\}$  un processus  $ARMA(1, 1)$  qui défini par :

$$Y_t - 2Y_{t-1} = \varepsilon_t + 4\varepsilon_{t-1} \quad \varepsilon_t \sim WN(0, 6.25).$$

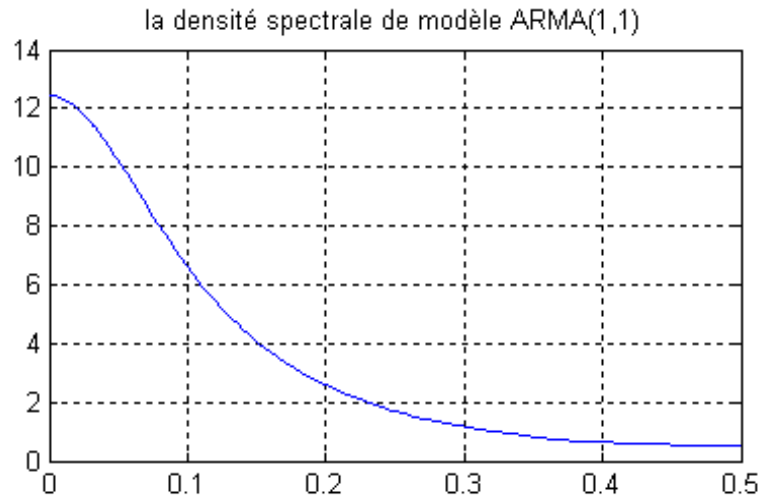
Nous remarquons que  $\{Y_t, t \in \mathbb{Z}\}$  n'est pas causale et n'est pas inversible. Dans ce cas nous remplaçons les polynômes  $\Phi(z) = 1 - 2z$  et  $\Theta(z) = 1 + 4z$  par les polynômes  $\widehat{\Phi}(z) = 1 - 0,5z$  et  $\widehat{\Theta}(z) = 1 + 0,25z$  dans la représentation de  $\{Y_t, t \in \mathbb{Z}\}$  précédente et on obtient :

$$Y_t - 0,5Y_{t-1} = \varepsilon_t + 0,25\varepsilon_{t-1} \quad \varepsilon_t \sim WN(0, 12.5),$$

alors la densité spectrale du processus  $\{Y_t, t \in \mathbb{Z}\}$  est donnée par la formule suivante :

$$f_Y(\lambda) = \frac{\sigma^2}{2\pi} \left( \frac{1 + 0,5 \cos(\lambda) + 0,0625}{1 - \cos(\lambda) + 0.25} \right), \quad -\pi \leq \lambda \leq \pi \text{ et } \lambda = 2\pi c \text{ avec } 0 \leq c \leq 1/2,$$

et sont graphe comme suit :



### 3.7.2 Densité spectrale de processus $MA(q)$

Si  $\{Y_t, t \in \mathbb{Z}\}$  est un processus  $MA(q)$  c'est-à-dire :

$$Y_t = \Theta(B) \varepsilon_t, \quad \text{avec } \varepsilon_t \sim WN(0, \sigma^2) \text{ et } \Theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q,$$

alors  $\{Y_t, t \in \mathbb{Z}\}$  est de densité spectrale suivante :

$$f_Y(\lambda) = |\Theta(e^{-i\lambda})|^2 f_\varepsilon(\lambda), \quad \text{avec } \Theta(e^{-i\lambda}) \neq 0 \text{ et } -\pi \leq \lambda \leq \pi$$

$f_\varepsilon(\lambda)$  et la densité spectrale de processus  $\{\varepsilon_t, t \in \mathbb{Z}\}$ .

#### Exemple

Soit  $\{Y_t, t \in \mathbb{Z}\}$  un processus  $MA(1)$  qui défini par :

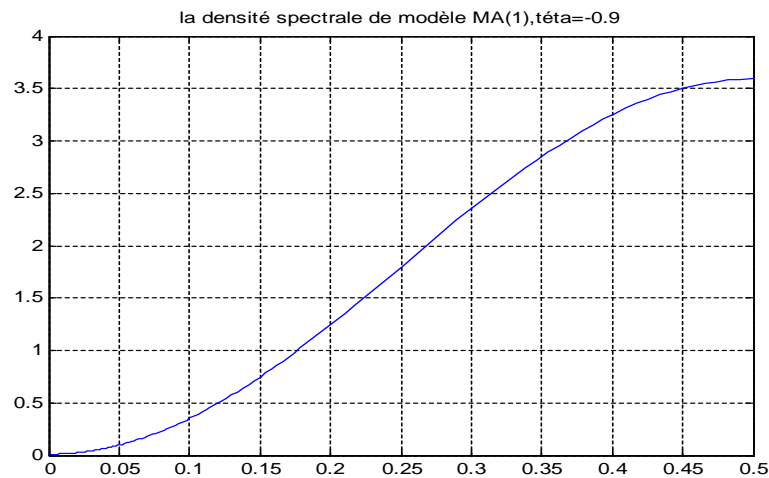
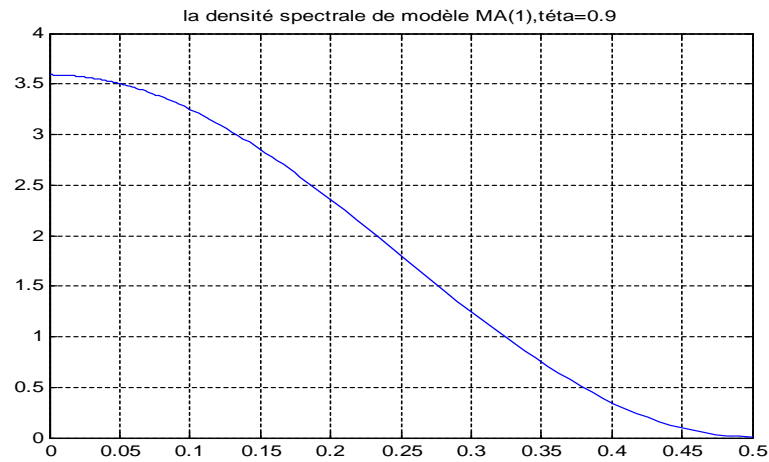
$$Y_t = \varepsilon_t + \theta \varepsilon_{t-1} \quad \varepsilon_t \sim WN(0, 6.25)$$

$$\text{et } \theta = \mp 0.9,$$

alors la densité spectrale de  $\{Y_t, t \in \mathbb{Z}\}$  est donnée par :

$$f_Y(\lambda) = \frac{\sigma^2}{2\pi} |1 + \theta e^{-i\lambda}|^2 = \frac{\sigma^2}{2\pi} (1 + 2\theta \cos(\lambda) + \theta^2), \quad -\pi \leq \lambda \leq \pi, \quad \theta = \mp 0.9,$$

et sont graphé comme suit :



### 3.7.3 Densité spectrale de processus $AR(p)$

Si  $\{Y_t, t \in \mathbb{Z}\}$  est un processus  $AR(p)$  c'est-à-dire :

$$\varepsilon_t = \Phi(B) Y_t \text{ avec } \Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \text{ et } \phi_p \in \mathbb{R}^*,$$

alors  $\{Y_t, t \in \mathbb{Z}\}$  est de densité spectrale suivante :

$$f_Y(\lambda) = |\Phi(e^{-i\lambda})|^2 f_\varepsilon(\lambda), \text{ avec } \Phi(e^{-i\lambda}) \neq 0 \text{ et } -\pi \leq \lambda \leq \pi$$

$f_\varepsilon(\lambda)$  et la densité spectrale de processus  $\{\varepsilon_t, t \in \mathbb{Z}\}$ .

**Exemple**

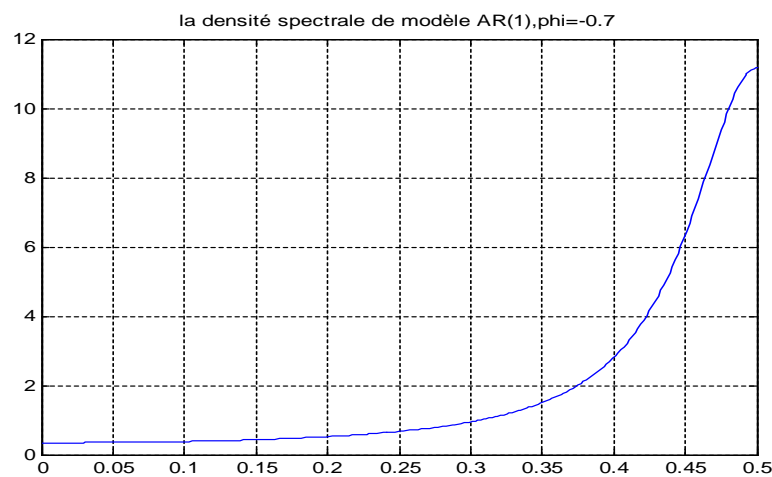
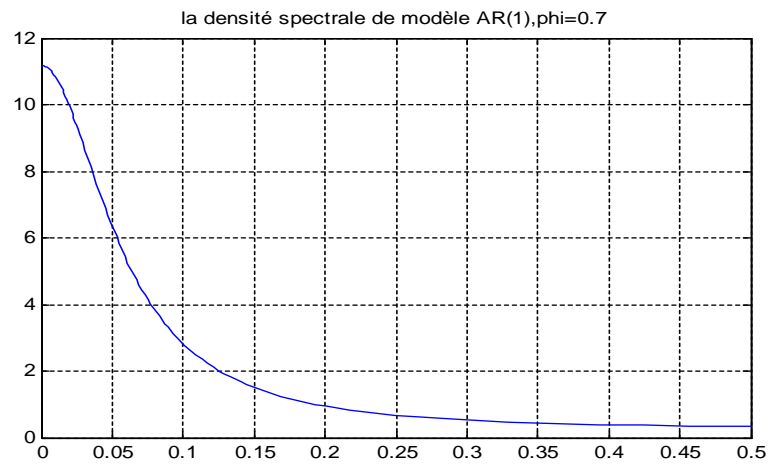
Soit  $\{Y_t, t \in \mathbb{Z}\}$  un processus  $AR(1)$  qui défini par :

$$Y_t = \phi Y_{t-1} + \varepsilon_t, \text{ avec } \varepsilon_t \sim WN(0, 6.25),$$

alors la densité spectrale de  $\{Y_t, t \in \mathbb{Z}\}$  est donnée par :

$$f_Y(\lambda) = \frac{\sigma^2}{2\pi} |1 - \phi e^{-i\lambda}|^2 = \frac{\sigma^2}{2\pi} (1 - 2\phi \cos(\lambda) + \phi^2), \quad -\pi \leq \lambda \leq \pi, \quad \phi = \mp 0.7,$$

et sont graphe comme suit :



## 3.8 Processus linéaires de variance infinie

### 3.8.1 Introduction

Il y a eu récemment beaucoup d'intérêt concernant la modélisation des séries chronologiques en utilisant des processus *ARMA* à variance infinie (Resnick, 2007; Embrechts et al, 1997; Brockwell et Davis, 1991). Des exemples où de tels modèles semblent être appropriés peuvent être trouvés dans Ttuck et Kleiner (1974), qui ont considéré des signaux de téléphone, et Fama (1965), qui a modelisé des prix de marché boursier. N'importe quelle série temporelle qui présente des pointes aiguës (*des piques*) d'occasionnelles observations périphériques suggère l'utilisation possible d'un modèle à variance infini.

Dans cette section, nous nous concentrons sur les processus générés par l'application d'un filtre linéaire à une suite iid,  $\{\varepsilon_t, t \in \mathbb{Z}\}$ , de distribution  $F$ , avec queue de type Pareto, c'est-à-dire :

$$\begin{cases} y^\alpha (1 - F(x)) = y^\alpha p(\varepsilon_t > y) \rightarrow pC, \text{ si } y \rightarrow \infty \\ y^\alpha (F(-x)) = y^\alpha p(\varepsilon_t \leq -y) \rightarrow qC, \text{ si } y \rightarrow \infty \\ 0 < \alpha < 2, 0 \leq p = 1 - q \leq 1 \end{cases}, \quad (3.8.1)$$

où  $C$  est une constante finie positive que nous appellerons la dispersion (noté  $\text{disp}(\varepsilon_t)$ ) de la variable aléatoire  $\varepsilon_t$  et à partir de la représentation (3.8.1) nous pouvons écrire :

$$y^\alpha (1 - F(y) + F(-y)) = y^\alpha p(|\varepsilon_t| > x) \rightarrow C, \text{ si } y \rightarrow \infty, \quad (3.8.2)$$

avec un simple calcul on a montré que :

$$\begin{cases} E|\varepsilon_t|^\delta = \infty & \text{si } \delta \geq \alpha \\ E|\varepsilon_t|^\delta < \infty & \text{si } \delta < \alpha \end{cases}, \quad (3.8.3)$$

ainsi que  $\text{Var}(\varepsilon_t) = 0$  pour  $0 < \alpha < 2$  et  $E|\varepsilon_t| < \infty$  si seulement si  $1 < \alpha < 2$ . Une classe importante des distributions satisfaisant la représentation (3.8.1) comprend les distributions stables non-normales.

**Remarque** La définition des distributions stables se trouve dans le premier chapitre.

### 3.8.2 Propriétés d'une variable aléatoire stable $\varepsilon$

Parmi les propriétés importantes de variable aléatoire stable  $\varepsilon$  on a :

1. La fonction caractéristique,  $\Phi(u) = \exp(iu\varepsilon)$  est donnée par :

$$\Phi(u) = \begin{cases} \exp\{iu\beta - d|u|^\alpha(1 - i\theta \operatorname{sgn}(u) \tan(\pi\alpha/2))\} & \text{si } \alpha \neq 1 \\ \exp\{iu\beta - d|u|(1 + i\theta(\pi/2) \operatorname{sgn}(u) \ln|u|)\} & \text{si } \alpha = 1 \end{cases}, \quad (3.8.4)$$

où

$$\begin{cases} \operatorname{sgn}(u) = u/|u| & \text{si } u \neq 0 \\ 0 & \text{si } u = 0 \end{cases},$$

et les paramètres  $\alpha \in ]0, 2]$ ,  $\beta \in \mathbb{R}$ ,  $d^{1/\alpha} \in [0, \infty[$  et  $\theta \in [-1, +1]$  sont connus.

2. Si  $\alpha = 2$  alors  $\varepsilon \sim N(\beta, 2d)$ .

3. Si  $\theta = 0$  donc la distribution de  $\varepsilon$  est symétrique par rapport à  $\beta$ . La distribution stable symétrique (c'est-à-dire symétrique par rapport à 0) ont des fonctions caractéristiques de la forme suivantes :

$$\Phi(u) = \exp\{-d|u|^\alpha\}. \quad (3.8.5)$$

4. Si  $\alpha = 1$  et  $\theta = 0$ , alors  $\varepsilon$  est de distribution de Cauchy de densité de probabilité suivante :

$$f(\varepsilon) = \left(\frac{d}{\pi}\right) [d^2 + (\varepsilon - \beta)^2]^{-1}, \varepsilon \in \mathbb{R}.$$

5. La distribution symétrique stable satisfait la propriété (3.8.1) avec  $a_n = n^{1/\alpha}$  et  $b_n = 0$ , car si  $\varepsilon, \varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ , sont de fonction caractéristique de la représentation (3.8.5) et  $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$  sont indépendantes, alors :

$$E(\exp[iu(\varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_n)]) = e^{-nd|u|^\alpha} = E(\exp[iu\varepsilon n^{1/\alpha}]).$$

6. Si  $F$  est une fonction de distribution de  $\varepsilon$  et  $\alpha \in ]0, 2[$ , l'expression (1.1) avec  $p = (1 + \theta)/2$  satisfait la forme suivante :

$$C = \begin{cases} d/(\Gamma(1-\alpha)\cos(\pi\alpha/2)) & \text{si } \alpha \neq 1 \\ 2d/\pi & \text{si } \alpha = 1 \end{cases} . \quad (3.8.6)$$

Dans la proposition suivante, nous donnons des conditions dans lesquelles la somme  $\varepsilon_t = \sum_{j=-\infty}^{+\infty} \Psi_j \varepsilon_{t-j}$  existe lorsque  $\{\varepsilon_t\}$  est une suite de variables aléatoires iid satisfaisant la représentation (3.8.1).

**Proposition 3.8.1** Soient  $\{\varepsilon_t\}$  est une suite de variables aléatoires iid satisfaisant la représentation (3.8.1).

Si  $\{\Psi_j\}$  est une suite de constantes telle que :

$$\sum_{j=-\infty}^{\infty} |\Psi_j|^\delta < \infty \quad \text{pour } \delta \in ]0, \alpha[ \cap [0, 1], \quad (3.8.7)$$

alors la série infinie  $\sum_{j=-\infty}^{\infty} \Psi_j \varepsilon_{t-j}$  converge absolument avec une probabilité 1,

$$\left( \text{c'est-à-dire } p \left( \left( \sum_{j=-\infty}^{\infty} \Psi_j \varepsilon_{t-j} \right) < \infty \right) = 1 \right).$$

### Preuve

1. Considérer d'abord le cas où  $1 < \alpha < 2$ . Alors d'après (3.8.3) on a  $E|\varepsilon_t| < \infty$  et par conséquent :

$$\begin{aligned} E \left( \sum_{j=-\infty}^{\infty} |\Psi_j \varepsilon_{t-j}| \right) &= \sum_{j=-\infty}^{\infty} |\Psi_j| E|\varepsilon_{t-j}| \\ &= \sum_{j=-\infty}^{\infty} |\Psi_j| E|\varepsilon_1| \\ &< \infty, \end{aligned}$$

ainsi  $\sum_{j=-\infty}^{\infty} |\Psi_j \varepsilon_{t-j}|$  est finie avec une probabilité 1,

$$\left( \text{c'est-à-dire } p \left( \left( \sum_{j=-\infty}^{\infty} |\Psi_j \varepsilon_{t-j}| \right) < \infty \right) = 1 \right).$$

2. Nous supposons maintenant  $0 < \alpha < 1$ . Ainsi  $0 < \delta < 1$ , nous pouvons appliquer l'inégalité  $|x + y|^\delta \leq |x|^\delta + |y|^\delta$  de triangle à la somme infinie  $\sum_{j=-\infty}^{\infty} \Psi_j \varepsilon_{t-j}$ . Donc d'après (3.8.3) on a :

$$\begin{aligned} E \left( \sum_{j=-\infty}^{\infty} |\Psi_j \varepsilon_{t-j}| \right)^\delta &= \sum_{j=-\infty}^{\infty} |\Psi_j|^\delta E |\varepsilon_{t-j}|^\delta \\ &= \sum_{j=-\infty}^{\infty} |\Psi_j|^\delta E |\varepsilon_1|^\delta \\ &< \infty, \end{aligned}$$

ainsi  $\sum_{j=-\infty}^{\infty} |\Psi_j \varepsilon_{t-j}|$  est finie avec une probabilité 1,

$$\left( \text{c'est-à-dire } p \left( \sum_{j=-\infty}^{\infty} |\Psi_j \varepsilon_{t-j}| < \infty \right) = 1 \right).$$

### Remarque 3.8.1

La distribution de la somme infinie  $\sum_{j=-\infty}^{\infty} \Psi_j Z_{t-j}$  satisfait la représentation (3.1.2) est donnée par :

$$y^\alpha p \left( \left| \sum_{j=-\infty}^{\infty} \Psi_j Z_{t-j} \right| > x \right) \rightarrow \left( \sum_{j=-\infty}^{\infty} |\Psi_j|^\delta \right) C.$$

**Remarque 3.8.2**

Si  $\varepsilon_t$  est de distribution symétrique stable avec la fonction caractéristique  $e^{-d|t|^\alpha}$ .

(et de dispersion  $C$  donnée par la représentation (3.8.6)), alors  $\sum_{j=-\infty}^{\infty} \Psi_j \varepsilon_{t-j}$  est aussi de distribution symétrique stable avec une dispersion :

$$\bar{C} = C \sum_{j=-\infty}^{\infty} |\Psi_j|^\delta.$$

**Remarque 3.8.3** Soit le processus  $\{Y_t, t \in \mathbb{Z}\}$  qui définie par :

$$Y_t = \sum_{j=-\infty}^{\infty} \Psi_j \varepsilon_{t-j}, \tag{3.8.8}$$

où  $\{\Psi_j\}$  et  $\{\varepsilon_t\}$  satisfait les hypothèses de proposition précédente, existe avec une probabilité 1 et strictement stationnaires, c'est -à-dire la distribution conjointe  $(Y_1, \dots, Y_k)'$  est la même que celle de  $(Y_{1+h}, \dots, Y_{k+h})'$  pour tout  $h \in \mathbb{Z}$  et pour tout  $k \in \mathbb{Z}^+$ .

En particulier si les coefficients  $\Psi_j$  sont choisis de sorte que  $\Psi_j = 0$  si  $j < 0$  et :

$$\sum_{j=0}^{\infty} \Psi_j \varepsilon^j = \Theta(z) / \Phi(z), |z| \leq 1, \tag{3.8.9}$$

où  $\Theta(z) = 1 + \theta_1 z + \theta_2 z^2 + \dots + \theta_q z^q$  et  $\Phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \dots - \phi_p z^p \neq 0$  pour  $|z| \leq 1$ , il est alors facile de montrer que  $\{Y_t\}$  définie par la représentation (3.8.8) satisfait aux équations aux différences  $\Phi(B) Y_t = \Theta(B) \varepsilon_t$  de modèle *ARMA*  $(p, q)$ . Nous enregistrons ce résultat comme une proposition.

**Proposition 3.8.1**

Soit  $\{\varepsilon_t\}$  une suite de variables aléatoires iid avec une fonction de distribution satisfaisant la représentation (3.1.1) . Alors si  $\Phi(\cdot)$  et  $\Theta(\cdot)$  deux polynomes tels que  $\Phi(z) \neq 0$  pour  $|z| \leq 1$ , l'équation aux différences :

$$\Phi(B) Y_t = \Theta(B) \varepsilon_t, \tag{3.8.10}$$

donne une solution stationnaire unique de la forme :

$$Y_t = \sum_{j=0}^{\infty} \Psi_j \varepsilon_{t-j}, \quad (3.8.11)$$

où les coefficients  $\{\Psi_j\}$  sont déterminés par la représentation (3.8.9). Si, en plus  $\Phi(z)$  et  $\Theta(z)$  n'ont pas de zéros communs, alors le processus est inversible si et seulement si  $\Phi(z) \neq 0$  pour  $|z| \leq 1$ .

### Preuve

D'après la proposition précédente les séries de la représentation (3.8.11) convergent absolument avec une probabilité 1,

$$\left( \text{c'est-à-dire } p \left( \sum_{j=0}^{\infty} \Psi_j \varepsilon_{t-j} < \infty \right) = 1 \right).$$

Bien que le processus  $\{Y_t, t \in \mathbb{Z}\}$  défini par la relation (3.8.8) soit strictement stationnaire, il n'est pas stationnaire au second ordre puisque  $E|Y_t|^2 = \infty$ . Malgré cela, nous pouvons encore définir, pour un tel processus, un analogue de la fonction d'autocorrélation comme suit :

$$\rho(h) = \frac{\sum_j \Psi_j \Psi_{j+h}}{\sum_j \Psi_j^2}, \quad h = 1, 2, \dots \quad (3.8.12)$$

Nous utilisons les mêmes notations que pour la fonction d'autocorrélation d'un processus stationnaire au second ordre car si  $\{\varepsilon_t\}$  est remplacée dans la relation (3.8.8) par une suite de bruits blancs de variance finie, alors la représentation (3.8.12) coïncide avec la fonction d'autocorrélation du processus  $\{Y_t, t \in \mathbb{Z}\}$ .

Ainsi,  $\rho(h)$  est simplement est une fonction des coefficients  $\{\Psi_j\}$  dans la représentation (3.8.12), ou en fonction des coefficients de  $\{\phi_j\}$  et  $\{\theta_j\}$  si  $\{Y_t\}$  est un processus *ARMA* de représentation (3.8.10).

Nous pouvons donc estimer  $\rho(h)$  en utilisant la fonction d'autocorrélation simple comme

suit :

$$\tilde{\rho}(h) = \frac{\sum_{t=1}^{n-h} Y_t Y_{t+h}}{\sum_{t=1}^n Y_t^2}, \quad h = 1, 2, \dots$$

Mais il n'est en aucun cas clair que  $\tilde{\rho}(h)$  soit même un estimateur consistant de  $\rho(h)$ .

D'après, le théorème de Davis et Resnick (1986), nous constatons que  $\tilde{\rho}(h)$  est non seulement cohérente, mais on peut la considérer comme un bon estimateur de  $\rho(h)$ .

**Théorème 3.8.1** (e.g Brockwell . Davis,1991)

Soit  $\{\varepsilon_t\}$  une suite de variables aléatoires iid avec une fonction de distribution satisfaisant la représentation (3.8.1) et soit  $\{Y_t\}$  un processus strictement stationnaire tel que :

$$Y_t = \sum_{j=-\infty}^{\infty} \Psi_j \varepsilon_{t-j},$$

où

$$\sum_{j=-\infty}^{\infty} |j| |\Psi_j|^\delta < \infty, \quad \text{pour } \delta \in ]0, \alpha[ \cap [0, 1],$$

donc pour tout  $h > 0$ , (avec  $h \in \mathbb{Z}$ ) on a :

$$(n/\ln(n))^{1/\alpha} (\tilde{\rho}(1) - \rho(1), \dots, \tilde{\rho}(h) - \rho(h))' \implies (X_1, \dots, X_h)', \quad (3.8.13)$$

où

$$X_k = \sum_{j=1}^{\infty} (\rho(k+j) + \rho(k-j) - 2\rho(j)\rho(k)) S_j/S_0, \quad k = 1, 2, \dots, h,$$

et  $S_1, \dots, S_n$  sont des variables aléatoires stables indépendantes et  $S_0$  est une variable aléatoire stable positive de fonction caractéristique de la forme suivante :

$$E(\exp(iu S_0)) = \exp \left\{ -C\Gamma(1 - \alpha/2) \cos(\pi\alpha/4) |u|^{\alpha/2} (1 - i \operatorname{sgn}(u) \tan(\pi\alpha/4)) \right\}, \quad (3.8.14)$$

et  $S_1, S_2, \dots$  sont des variables aléatoires iid et de fonction caractéristique :

$$E(\exp(iu S_1)) = \begin{cases} \exp\{-C^2 \Gamma(1-\alpha) \cos(\pi\alpha/2) |u|^\alpha\} & \text{si } \alpha \neq 1 \\ \exp\{-C^2 \pi |u|/2\} & \end{cases} . \quad (3.8.15)$$

Si  $\alpha > 1$  alors la représentation (3.8.13) est également vraie quand  $\tilde{\rho}(h)$  est remplacé par sa version corrigée moyenne,

$$\tilde{\rho}(h) = \frac{\sum_{t=1}^{n-h} (Y_t - \bar{Y})(Y_{t+h} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2} \quad \text{Où } \bar{Y} = \frac{1}{n} \sum_{t=1}^n Y_t.$$

D'après ce théorème on a :  $\tilde{\rho}(h) \xrightarrow{p} \rho(h)$  et plus spécifiquement on a aussi :

$$\tilde{\rho}(h) - \rho(h) = O_p\left([n/\ln(n)]^{-1/\alpha}\right) = O_p\left(n^{-1/\beta}\right) \quad \text{pour tout } \beta > \alpha.$$

Ce taux de convergence vers zéro se compare favorablement avec le rythme plus lent,  $O_p(n^{-1/\beta})$ , pour différence  $\tilde{\rho}(h) - \rho(h)$  dans le cas de variance finie.

La forme de la distribution asymptotique de  $\tilde{\rho}(h)$  peut être quelque peu simplifiée. Pour ce faire, notez que  $Y_k$  a la même distribution que :

$$\left( \sum_{j=1}^{\infty} |\rho(h+j) + \rho(h-j) - 2\rho(j)\rho(h)|^\alpha \right)^{1/\alpha} U/V, \quad (3.8.16)$$

où  $V (\geq 0)$  et  $U$  sont deux variables aléatoires de fonction caractéristique donnée par (3.8.14) et (3.8.15) respectivement avec  $C = 1$ . La distribution de la variable aléatoire de  $U/V$  peut être trouvée avec deux méthodes : soit par une simulation des copies indépendantes de  $U/V$  ou par une intégration numérique de la densité conjointe de  $(U, V)$  sur une région appropriée. Sauf lorsque  $\alpha = 1$ , la densité conjointe de  $U$  et  $V$  ne peut pas être écrite sous forme fermée. Dans le cas où  $\alpha = 1$ , la variable aléatoire  $U$  suit la loi de Cauchy de densité

suivante :

$$f_U(u) = \frac{1}{2} [\pi^2/4 + u^2]^{-1} \quad (\text{voir propriété 4 de la variable aléatoire stable}),$$

et la variable aléatoire  $V$  de densité suivante :

$$f_V(v) = \frac{1}{2} v^{-3/2} e^{-\pi/(4v)}, v \geq 0 \quad (\text{voir Feller (1971)}),$$

alors la densité de distribution de la variable aléatoire  $U/V$  est donnée par :

$$\begin{aligned} p(U/V \leq x) &= \int_0^\infty p(U \leq xy) f_V(y) dy & (3.8.17) \\ &= \int_0^\infty 2^{-1/2} (\pi w)^{-3/2} [\arctan(xw) + (\pi/2)] \exp(-1/(2w)) dw. \end{aligned}$$

Remaquons aussi que  $U/V$  a la même distribution que le produit d'une variable aléatoire standard de Cauchy (avec une densité de probabilité de la forme  $\pi^{-1} (1 + x^2)^{-1}$ )

### 3.8.3 Processus moyenne mobile d'ordre $q$ à variance infinie (Moving Average)

Soit  $\{Y_t\}$  un processus moyenne mobile d'ordre  $q$  c'est-à-dire :

$$Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q},$$

où  $\{\varepsilon_t\}$  est une suite de variables aléatoires satisfaisant les suppositions du théorème précédent. Lorsque  $\rho(h) = 0$  pour  $|h| > q$ , le théorème implique dans ce cas :

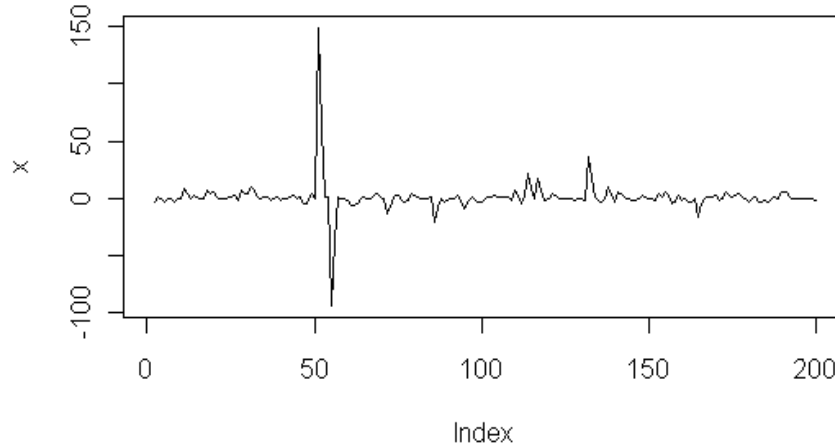
$$(n/\ln(n))^{1/\alpha} (\hat{\rho}(h) - \rho(h)) \implies \left(1 + 2 \sum_{j=1}^q |\rho(j)|^\alpha\right)^{1/\alpha} U/V, \quad h > q,$$

où le côté droit réduit à  $U/V$  si  $q = 0$ .

On va simuler deux cents valeurs ( $n = 200$ ) du processus  $MA(1)$  :

$$Y_t = \varepsilon_t + 0.4\varepsilon_{t-1}, \quad \text{c'est-à-dire } \theta = 0.4, \quad (3.8.18)$$

avec  $\{\varepsilon_t, t \in \mathbb{Z}\}$  est une suite iid de distributions de Cauchy standard (i.e  $Ee^{iu z_1} = e^{-|u|}$ ).



**Figure 1** . Simulation du modèle  $MA(1)$ , ( $Y_t = 0.4\varepsilon_{t-1} + \varepsilon_t$ ),  
avec 200 réalisations de  $\{\varepsilon_t\}$  iid de loi cauchy standard.

On sait que dans la distribution de Cauchy le moment d'ordre 1 ( $E(Y_t)$ ) et le moment d'ordre 2 ( $E(Y_t^2)$ ) n'existent pas, alors dans ce cas on ne peut pas estimer le coefficient d'autocorrélation  $\rho(1)$  ( $\hat{\rho}(1)$ ).

Malgré  $E|Y_t|^2$  n'existe pas mais on peut remplacer l'estimateur  $\hat{\rho}(1)$  avec  $\tilde{\rho}(1)$  qui est défini par la forme suivante :  $\tilde{\rho}(1) = \frac{\sum_{t=1}^{n-1} Y_t Y_{t+h}}{\sum_{t=1}^n Y_t^2}$  car à partir du théorème de Davis et Resnick (1986), nous constatons que  $\tilde{\rho}(1)$  est non seulement cohérente, mais on peut le

considérer comme un bon estimateur de  $\rho(1)$ .

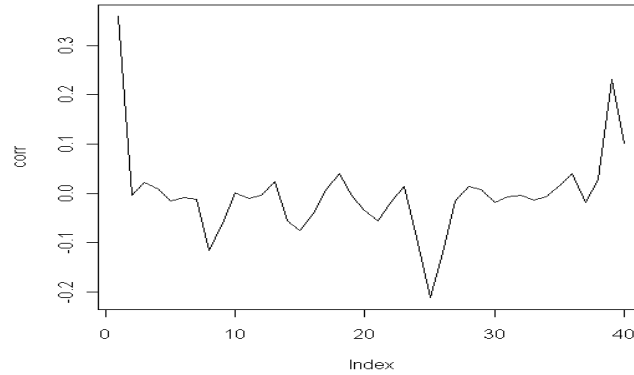


Figure 2. Estimateur du coefficient d'autocorrélation

$\tilde{\rho}(h)$  pour  $h = 1, 2, \dots, 40$ .

Les observations qu'on a générées en haut sont celles du processus de la représentation (3.8.13) ?

Pour montrer que les observations sont des réalisations de processus de la représentation (3.8.13) il faut montrer que  $\tilde{\theta} \xrightarrow{p} \theta$  ?

D'après la figure 2 on a  $\tilde{\rho}(1) = 0.359$  et on a aussi  $\tilde{\theta}$  est une racine dans l'intervalle  $[-1, +1]$  de l'équation  $\tilde{\rho}(1) = \frac{\tilde{\theta}}{(1+\tilde{\theta}^2)}$ . S'il n'ya pas de racines, nous avons définie  $\tilde{\theta} = \text{sgn}(\tilde{\rho}(1))$  (voir la section 8.5 dans Brockwell). D'après les calculs on a trouvé  $\tilde{\theta} = 0.401$ .

De plus l'estimateur  $\tilde{\theta} = 0.401$  est très proche de la vraie valeur  $\theta = 0.40$ , ceci implique que les observations qu'on a générées en haut sont des réalisations du processus de la représentation (3.8.13).

### Intervalle de confiance pour $\rho(1)$

D'après le calcul de  $\tilde{\rho}(1)$  on peut construire un intervalle de confiance pour  $\rho(1)$ .

Le quantile 0.975 de  $U/V$  pour le processus de la représentation (3.8.13) se trouve numériquement à partir de l'intégrale de la représentation (3.8.17) est donné la valeur 12.46. D'après le théorème (3.8.1), l'intervalle de confiance pour  $\rho(1)$  à 95% est donné par :

$$\tilde{\rho}(1) \pm 12.46 (|1 - 2\tilde{\rho}^2(1)| + |\tilde{\rho}(1)|) (\ln(n)/n) = 0.351 \pm 0.362 = [-0.011, 0.713].$$

### 3.8.4 Prédiction linéaire de processus ARMA à variance infinie

Soit  $\{Y_t, t \in \mathbb{Z}\}$  un processus ARMA strictement stationnaire défini par la représentation (3.8.10) avec  $\Phi(z)\Theta(z) \neq 0$  pour tout  $z \in \mathbb{C}$  tels que  $|z| \leq 1$ . Supposons aussi que la séquence iid  $\{\varepsilon_t, t \in \mathbb{Z}\}$  satisfait la représentation (3.8.1). En évaluant la performance du prédicteur linéaire,

$$\widehat{Y}_{n+1} = a_{n1}Y_n + a_{n2}Y_{n-1} + \cdots + a_{nn}Y_1. \quad (3.8.19)$$

Nous ne pouvons pas considérer le critère  $E\left(Y_{n+1} - \widehat{Y}_{n+1}\right)^2$  comme pour des processus au second ordre puisque cette espérance est infinie. D'autres critères pour choisir un "meilleur" prédicteur qui ont été proposés comprennent la minimisation de l'erreur absolue attendue (lorsque  $\alpha > 1$ ), et l'utilisation d'une technique de pseudo-spectrale (Cambanis et Soltani (1982)). Nous considérons ici un seul critère, à savoir la minimisation de la dispersion de l'erreur (voir aussi Brockwell et Davis, 1991). En utilisant (3.8.11) on peut écrire  $\widehat{Y}_{n+1}$  sous la forme

$$\widehat{Y}_{n+1} = \sum_{j=0}^{\infty} (a_{n1}\Psi_j + a_{n2}\Psi_{j-1} + \cdots + a_{nn}\Psi_{j-n+1}) \varepsilon_{n-j}, \quad (3.8.20)$$

et en utilisant (3.8.11) encore une fois nous obtenons

$$Y_{n+1} - \widehat{Y}_{n+1} = \varepsilon_{n+1} + \sum_{j=0}^{\infty} (\Psi_{j+1} - a_{n1}\Psi_j - a_{n2}\Psi_{j-1} - \cdots + a_{nn}\Psi_{j-n+1}) \varepsilon_{n-j}, \quad (3.8.21)$$

Comme  $\{\varepsilon_t, t \in \mathbb{Z}\}$  est supposée avoir la dispersion  $C$ , alors

$$\text{disp}\left(Y_{n+1} - \widehat{Y}_{n+1}\right) = C \left(1 + \sum_{j=0}^{\infty} |\Psi_{j+1} - a_{n1}\Psi_j - a_{n2}\Psi_{j-1} - \cdots + a_{nn}\Psi_{j-n+1}|^{\alpha}\right). \quad (3.8.22)$$

Dans le cas particulier quand  $\{\varepsilon_t\}$  a la distribution stable symétrique avec l'exposant  $\alpha \in (0, 2)$  et le paramètre d'échelle  $d^{1/\alpha}$  (i.e  $Ee^{i\theta\varepsilon_t} = \exp(-d|\theta^\alpha|)$ ), la dispersion de  $\{\varepsilon_t\}$  est

$$C = \begin{cases} d/[\Gamma(1-\alpha)\cos(\pi\alpha/2)] & \text{si } \alpha \neq 1 \\ 2d/\pi & \text{si } \alpha = 1 \end{cases},$$

(voir la propriétés 6). L'erreur de prédiction est également stable symétrique avec une dispersion donnée par (3.8.22). La minimisation de (3.8.22) est donc équivalente à la minimisation de paramètre d'échelle de la distribution de l'erreur et donc à la minimisation de  $p\left(\left|Y_{n+1} - \widehat{Y}_{n+1}\right| > \varepsilon\right)$  pour tout  $\varepsilon > 0$ .

Le critère minimum de dispersion est utile également dans des problèmes de régression (Blattberg et Sargent (1971) et problèmes de filtrage de Kalman (Stuck (1978)) liés aux séquences stables. Pour des séquences générales  $\{\varepsilon_t\}$  satisfaisant la représentation (3.8.1) le critère minimum de dispersion réduit au minimum des probabilités de queue de la distribution de l'erreur de prévision.

la minimisation de (3.8.22) pour  $\alpha \in (0, 2)$  est un peu plus compliquée que dans le cas  $\alpha = 2$  et le meilleur prédicteur n'est pas en général unique. Pour une discussion générale du problème voir Cline et Brockwell (1985).

Ici nous basons simplement sur les résultats d'un processus  $MA(1)$  et, quand  $\{\varepsilon_t\}$  a une distribution de Cauchy, et en comparant la dispersion minimum du prédicteur

$$\widehat{Y}_{n+1} = \sum_{j=1}^n a_{nj} Y_{n+1-j} \hat{=} Y_{n+1}^* = \sum_{j=1}^n \Phi_{nj} Y_{n+1-j},$$

obtenu en supposant que  $\{\varepsilon_t\}$  est de variance finie.

**Proposition 3.8.2** Si  $Y_t = \varepsilon_t + \theta\varepsilon_{t-1}$  où  $\{\varepsilon_t\}$  est une séquence iid avec une fonction de distribution satisfaisant la représentation (3.8.1), alors la dispersion minimum du prédicteur linéaire  $\widehat{Y}_{n+1}$  de  $Y_{n+1}$  basé sur  $Y_1, Y_2, \dots, Y_n$  est donnée par

$$\widehat{Y}_{n+1} = \begin{cases} -\sum_{j=1}^n (-\theta)^j Y_{n+1-j} & \text{si } \alpha \leq 1 \\ -\sum_{j=1}^n (-\theta)^j \frac{1-\eta^{n+1-j}}{1-\eta^{n+1}} & \text{si } \alpha > 1 \end{cases},$$

où  $\eta = |\theta|^{\alpha/(\alpha-1)}$ . La dispersion de l'erreur de  $\widehat{Y}_{n+1}$  est

$$\begin{cases} C \left[1 + |\theta|^{(n+1)\alpha}\right] & \text{si } \alpha \leq 1 \\ C \left[1 + |\theta|^{(n+1)\alpha} \left(\frac{1-\eta}{1-\eta^n}\right)^{\alpha-1}\right] & \text{si } \alpha > 1 \end{cases}.$$

### 3.8.5 Prédiction linéaire de processus MA(1) à variance infinie

Soit :

$$Y_t = \varepsilon_t + \theta\varepsilon_{t-1}, |\theta| < 1, \quad (3.8.23)$$

où  $\{\varepsilon_t\}$  est une séquence iid de distribution de cauchy standar. Alors le prédicteur de  $\{Y_t\}$  est donné par :

$$\widehat{Y}_{n+1} = - \sum_{j=1}^n (-\theta)^j Y_{n+1-j} \quad (3.8.24)$$

$$p = q = \frac{1}{2} \text{ et } C = \frac{2}{\pi}, \quad (3.1)$$

et

$$\text{disp} \left( Y_{n+1} - \widehat{Y}_{n+1} \right) = \frac{2}{\pi} (1 + |\theta|^{n+1}). \quad (3.8.25)$$

#### Exemple

Soit

$$Y_t = \varepsilon_t + 0.4\varepsilon_{t-1}$$

avec  $\{\varepsilon_t, t \in \mathbb{Z}\}$  est une suites iid de distribution de Cauchy standard (*i.e*  $Ee^{iuz_1} = e^{-|u|}$ ).

D'après la simulation précédente de deux cents valeurs du processus  $\{Y_t, t \in \mathbb{Z}\}$  on a trouvé  $\tilde{\theta} = 0.401$ .

Dans ce cas on peut calculer le prédicteur linéaire de ce processus, en utilisant la représentation (3.8.24) comme suit :

$$\widehat{Y}_{n+1} = - \sum_{j=1}^n (-0.401)^j Y_{n+1-j}, \text{ pour } n = 200, \dots$$

### 3.9 Valeurs extrêmes pour les processus linéaires

Nous commençons par nous intéresser aux processus linéaires vérifiant les deux conditions suivantes :

$D$  : pour tout choix de  $i_1, i_2, \dots, i_p, j_1, j_2, \dots, j_q$  avec  $i_1 < i_2 < \dots < i_p < j_1 < j_2 < \dots < j_q$ , tels que,  $j_1 - i_p \geq l$ , on a :

$$\left| F_{i_1, i_2, \dots, i_p, j_1, j_2, \dots, j_q}(u) - F_{i_1, i_2, \dots, i_p}(u) F_{j_1, j_2, \dots, j_q}(u) \right| \leq \alpha_i(l),$$

où  $\alpha_i(l) \rightarrow 0$  quand  $l \rightarrow \infty$ . (Voir Leadbetter, Lindgren et Rootzèn, 1983).

On introduit maintenant une condition plus faible que  $D$ , appelée la condition  $D(u_n)$ ,

(Voir Leadbetter, Lindgren et Rootzèn, 1983 et Hsing, 1984) posons :

$$A = \{X_{i_1} \leq u, \dots, X_{i_p} \leq u\} \text{ et } B = \{X_{j_1} \leq u, \dots, X_{j_q} \leq u\}.$$

La condition est la suivante :

$D(u_n)$  : étant donnée une suite de réels  $(u_n)$  et une suite d'entiers naturels  $i_1 < i_2 < \dots < i_p < j_1 < j_2 < \dots < j_q \leq n$ , telle que  $j_1 - i_p \geq l$ , la condition  $D(u_n)$  est vérifiée si on a

$$\left| F_{i_1, i_2, \dots, i_p, j_1, j_2, \dots, j_q}(u_n) - F_{i_1, i_2, \dots, i_p}(u_n) F_{j_1, j_2, \dots, j_q}(u_n) \right| \leq \alpha_{n, l_n},$$

où  $\alpha_{n, l_n}$  ne croît pas avec  $l_n$ , et où la suite  $l_n$ , est telle que

$$\lim_{n \rightarrow \infty} \alpha_{n, l_n} = 0, \text{ et } \frac{l_n}{n} \rightarrow 0.$$

Ils englobent, en particulier, les processus stationnaires linéaires Gaussiens.

#### 3.9.1 Valeurs extrêmes pour la classe des processus $ARMA(p, q)$

Pour la classe des processus  $ARMA(p, q)$  Gaussiens, on dispose de deux types de résultats, l'un sur la loi du *max* l'autre sur le comportement asymptotique de la loi du processus, à savoir :

i. Soit  $(Y_n, n \in \mathbb{Z})$  une suite de variables aléatoires issue d'un processus de type  $ARMA(p, q)$  Gaussiens, alors il existe deux suites  $c_n > 0$  et  $d_n$  telles que

$$p(c_n^{-1}(M_n - d_n) \leq y) \rightarrow G_\alpha(y),$$

où  $G_\alpha(\cdot)$  est une distribution de type (*Fréchet* ou *Gumbel* ou *Weibull*), et  $M_n = \max(Y_1, \dots, Y_n)$ .

ii. Soit  $(Y_n, n \in \mathbb{Z})$  une suite de variables aléatoires issues d'un processus de type  $ARMA(p, q)$  Gaussiens, (donc vérifiant la condition  $D(u_n)$ ), alors il existe  $\tau > 0$  tel que :

$$p(M_n \leq u_n) \rightarrow e^{-\tau}, \quad n \rightarrow \infty.$$

**Remarque** Notons que l'existence des suites  $c_n$  et  $d_n$  n'est pas liée aux ordres des processus  $ARMA$ .

### 3.9.2 Valeurs extrêmes de processus $MA(\infty)$

i. Soit le processus moyenne mobile d'ordre infini défini par :

$$Y_t = \sum_{j=1}^{\infty} c_j \varepsilon_{t-j},$$

où  $\varepsilon_t$  représente une suite de variables aléatoires iid. Dans le cas où la queue de la distribution des  $\varepsilon_t$  décroît lentement vers zéro ( $F(y) = \exp(-e^{-y})$ ), on montre que le comportement asymptotique du processus  $(Y_t, t \in \mathbb{Z})$  est proche de celui d'une loi du Gumbel définie pour le minimum d'une série, c'est-à-dire

$$G_\alpha(y) = 1 - \exp(-e^{-y}).$$

Sous les mêmes hypothèses, et dans le cas d'un processus moyenne mobile fini, le comportement de la distribution extrême de la série est proche de celui d'une loi de Fréchet (voir Davis et Resnick, 1991).

ii. Soit le processus ponctuel  $(Y_t, t \in \mathbb{Z})$  à la représentation moyenne mobile défini par l'équation suivante :

$$Y_t = \sum_{j=1}^{\infty} c_j \varepsilon_{t-j},$$

où  $\varepsilon_t$  représente une suite de variables aléatoires iid. Dont la distribution  $F$  appartient au domaine d'attraction de Gumbel.

Sous des conditions classiques de sommabilité pour les coefficients  $c_j$ , on montre que la loi des maxima est proche de celle d'un processus ponctuel, voir Davis et Resnick (1988).

Les résultats sont en général très techniques et peu faciles à utiliser dans les applications.

### 3.9.3 Estimation de l'indice extrême $\zeta$ (Estimateur de Hill)

Parmi les estimateurs possibles pour l'indice extrême  $\xi$  est l'estimateur de Hill suivant :

$$\widehat{\zeta}_{n,k}^{(H)} = \frac{1}{k} \sum_{j=1}^k \ln X_{j,n} - \ln X_{k+1,n},$$

où  $X_{1,n} \geq X_{2,n}, \dots, \geq X_{n,n}$  sont les statistiques d'ordre et  $k = k(n)$  doit être choisi de façon appropriée. Cet estimateur, le plus utilisé en théorie des valeurs extrêmes.

### 3.9.4 Simulations et applications

On va simuler le processus  $ARMA(1, 1)$  suivant :

$$Y_t - 0.4Y_{t-1} = \varepsilon_t + 0.3\varepsilon_{t-1}.$$

On obtient les résultats suivantes :

1. Si l'entrée  $(\varepsilon_t)$  est de distribution marginale de loi de Pareto avec l'indice de queue  $\alpha = 3$  on a trouvé  $\widehat{\zeta} = 0.333$  ( $\zeta = 1/\alpha$ ).

2. Si l'entrée  $(\varepsilon_t)$  est de distribution marginale de loi de Burr avec l'indice de queue  $\alpha = 0.5$  on a trouvé  $\widehat{\zeta} = 2$  ( $\zeta = 1/\alpha$ ).

## 3.10 Estimation de l'indice de queue (Estimateur de Hill)

En particulier, il existe plusieurs estimateurs du paramètre  $\alpha$ . Parmi ces estimateurs on a l'estimateur de Hill, (1995). Après l'estimateur de Hill on trouve l'estimateur de Pickands, (1975), Dekkers-Einmahl et l'estimateur de Hann comme estimateurs les plus populaires.

D'après le chapitre 2 on a expliqué bien le résonnement de l'estimateur de Hill, qui s'écrit essentiellement sous la forme suivante :

$$\hat{\alpha}^{(H)} = \hat{\alpha}_{k,n}^{(H)} = \left( \frac{1}{k} \sum_{j=1}^k \ln X_{j,n} - \ln X_{k,n} \right)^{-1},$$

où  $k = k(n) \rightarrow \infty$  d'une manière appropriée et  $X_{n,n} \leq \dots \leq X_{1,n} = M_n$  désignent les statistiques d'ordre de l'échantillon  $X_1, \dots, X_n$ . Ceci signifie qu'une suite croissante de statistiques d'ordre supérieure est utilisée.

### 3.10.1 Simulations et applications

Pour bien clarifier la différence entre l'indice des distributions marginales à queues lourdes d'un processus linéaires et l'indice des distributions marginales à queues fines, on a besoin d'une simple application sur l'estimateur de l'indice de queue  $\alpha$  (Estimateur de Hill) et on commence par :

#### a. Indice des distributions marginales à queues lourdes

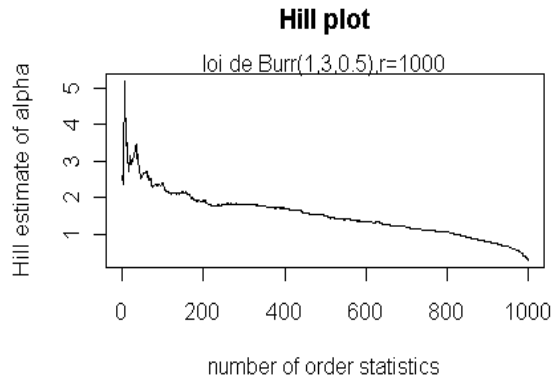
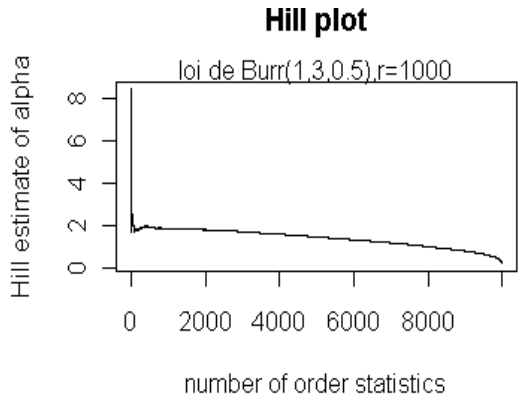
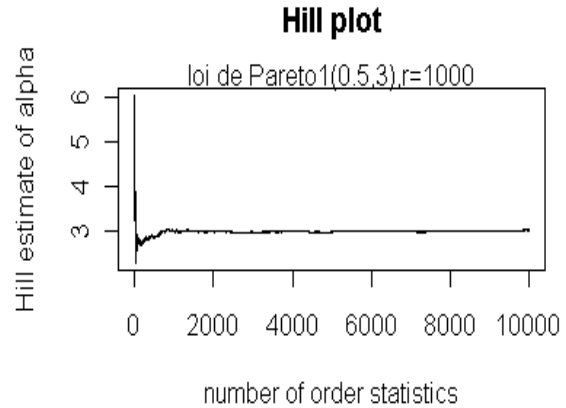
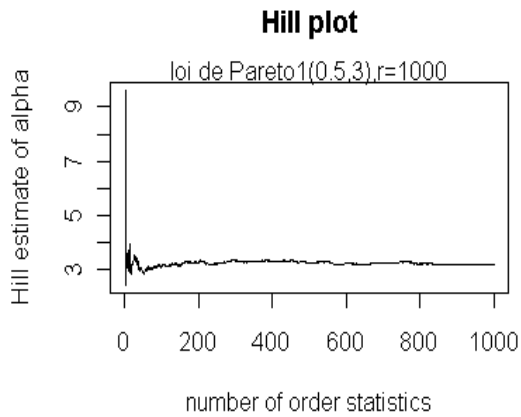
On a fait une simulation des modèles linéaires de classe  $ARMA(p, q)$

$$Y_t - \phi_1 Y_{t-1} - \dots - \phi_p Y_{t-p} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}, t \in \mathbb{Z},$$

où  $\{\varepsilon_t, t \in \mathbb{Z}\}$  est de distribution marginale à queue lourde (*Pareto avec  $\alpha = 3$ , Burr avec  $\alpha = 0, 5$* ) pour différentes tailles ( $n = 1000, n = 10000$ ) avec  $r$  ( $r = 1000$ ) répétitions pour chaque échantillon, et on a obtenu les résultats suivantes :

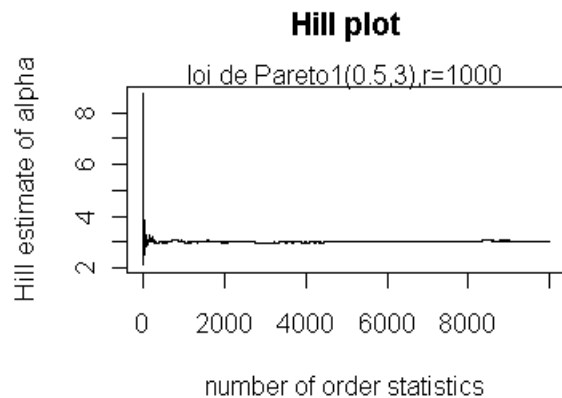
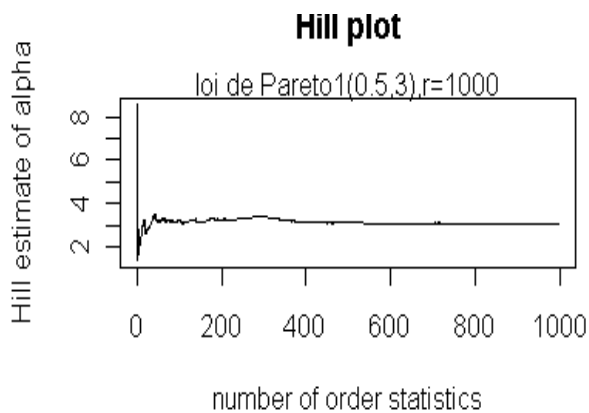
1. Soit le modèle  $ARMA(1, 1)$  suivant :

$$Y_t - 0.4Y_{t-1} = \varepsilon_t + 0.3\varepsilon_{t-1}.$$



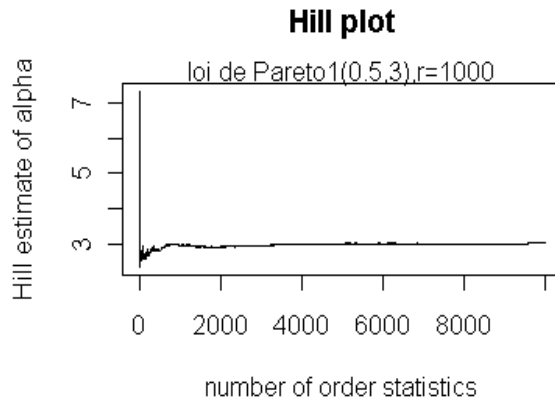
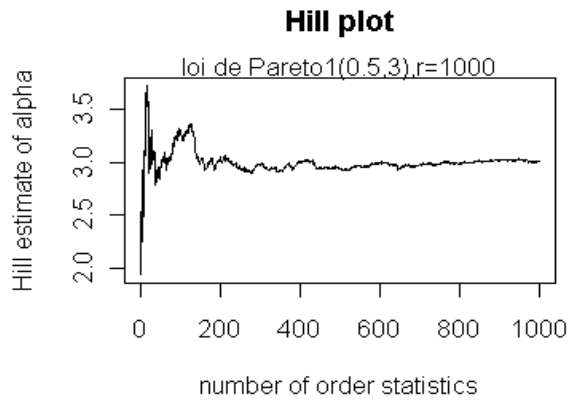
2. Soit le modèle  $AR(1)$  suivant :

$$Y_t = \varepsilon_t + 0.4Y_{t-1}$$



3. Soit le modèle  $MA(1)$  suivant :

$$Y_t = \varepsilon_t + 0.3\varepsilon_{t-1}$$



**b. Indice des distributions marginales à queues fines**

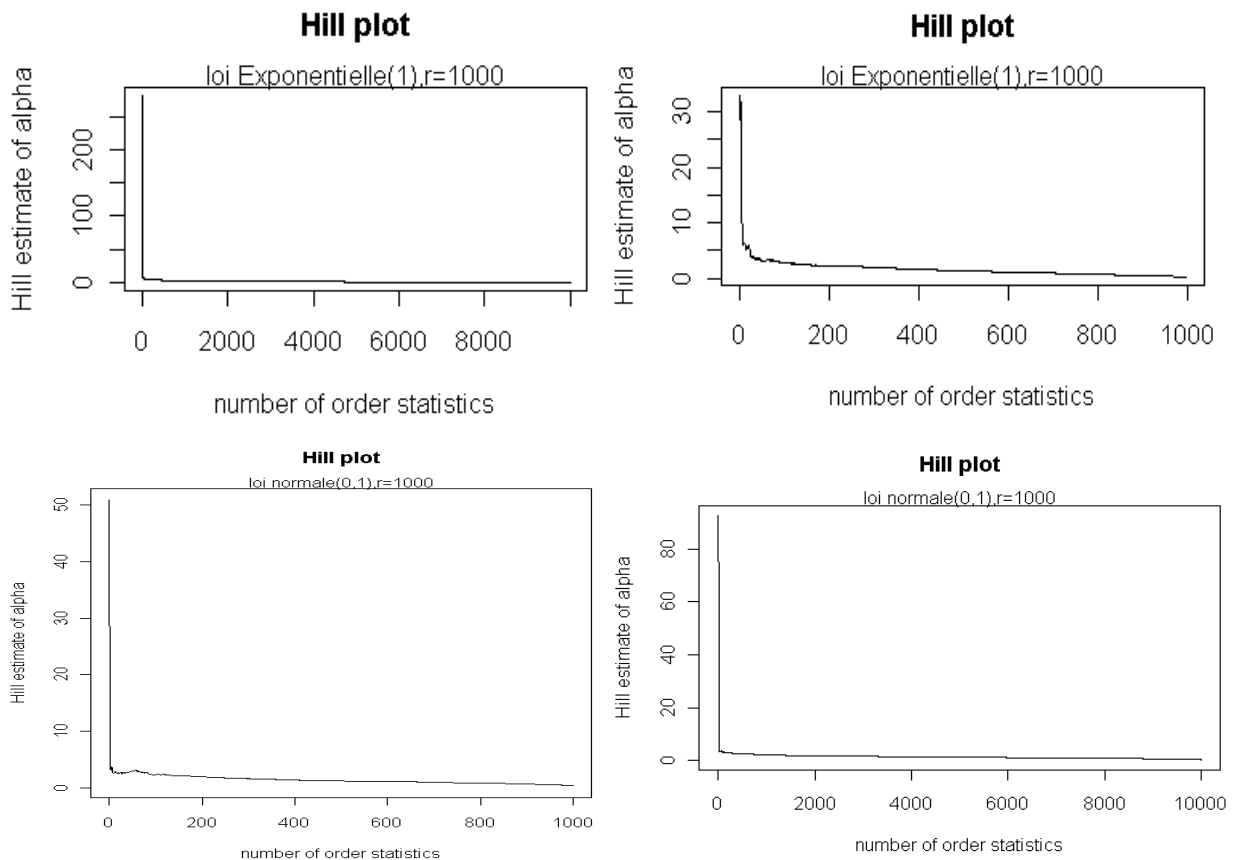
On a fait une simulation des modèles linéaires de classe  $ARMA(p, q)$ ,

$$Y_t - \phi_1 Y_{t-1} - \dots - \phi_p Y_{t-p} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}, t \in \mathbb{Z},$$

où  $\{\varepsilon_t, t \in \mathbb{Z}\}$  est de distribution marginale à queue fine (*Exponentielle avec  $\alpha = 1$ , normale  $(0, 1)$* ) pour différentes tailles ( $n = 1000, n = 10000$ ) avec  $r$  ( $r = 1000$ ) répétitifs pour chaque échantillon, et on a obtenu les résultats suivantes :

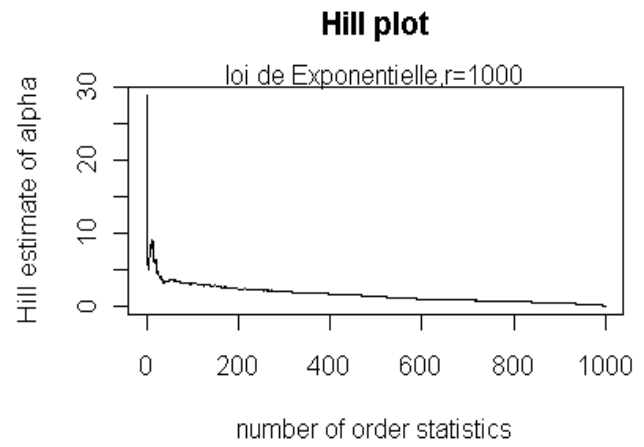
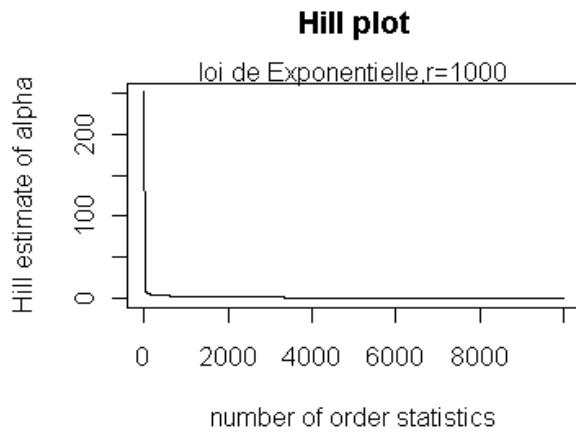
1. Soit le modèle  $ARMA(1, 1)$  suivant :

$$Y_t - 0.4Y_{t-1} = \varepsilon_t + 0.3\varepsilon_{t-1}.$$



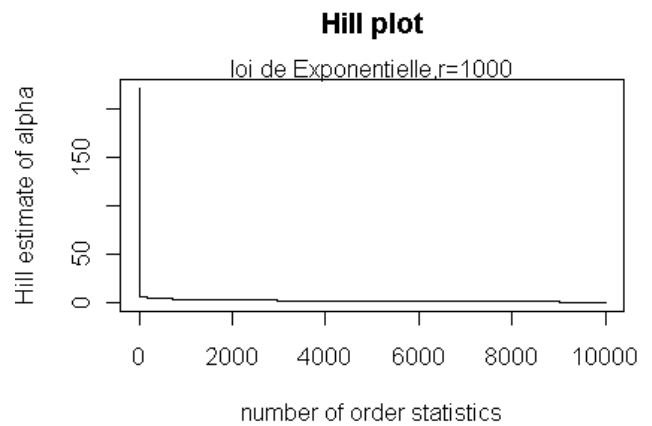
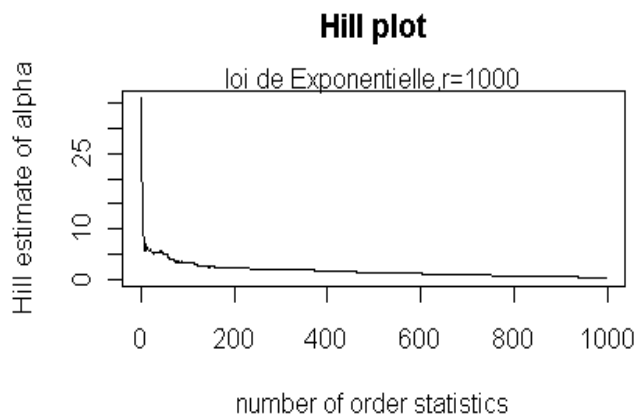
2. Soit le modèle  $AR(1)$  suivant :

$$Y_t = \varepsilon_t + 0.4Y_{t-1}$$



3. Soit le modèle  $MA(1)$  suivant :

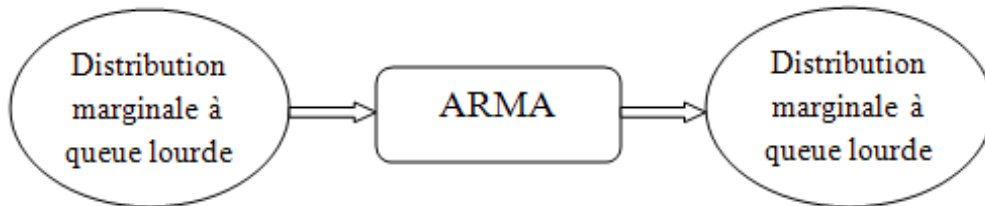
$$Y_t = \varepsilon_t + 0.3\varepsilon_{t-1}.$$



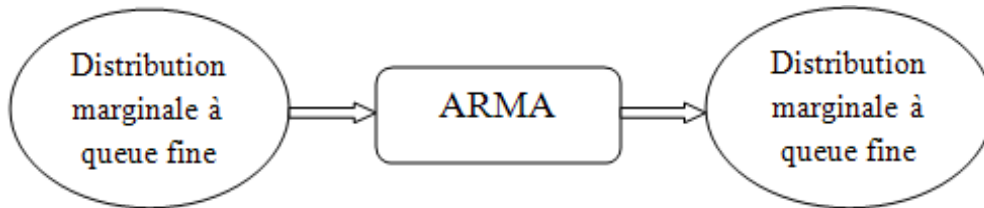
### 3.10.2 Conclusion

D'après les graphes précédents on a :

1. Si l'entrée du modèle est de distribution marginale à queue lourde, alors la sortie aussi est à queue lourde.



2. Si l'entrée du modèle est de distribution marginale à queue fine, alors la sortie aussi est à queue fine.



# Chapitre 4

## Processus stochastiques non linéaires à queues lourdes et valeurs extrêmes

### 4.1 Introduction

Les équations de récurrence stochastique ont été utilisées dans plusieurs domaines de probabilités appliquées ( pour appeler quelques références : Kesten (1994, p. 379) et Vervaat (1979, p.631) ont beaucoup d'articles dans ce domaine, la monographie de Brandt, Franken et L'sek (1986, p.87), la vue d'ensemble par Embrechts et Goldie (1982,p.206). En particulier, les processus *ARCH* et *GARCH* sont données par les équations de récurrence stochastique. Ils servent de taux de change spécial ou d'actifs financiers à volatilité stochastique et sont très populaires en économétrie. Plusieurs documents ont été publiés sur le modèle *ARCH* et les autres modèles, principalement dans le contexte des statistiques.

Notre intérêt pour ces modèles vient du fait que les processus *ARCH* avec une entrée à queue fine (c'est-à-dire les innovations gaussiennes), alors la sorti sont des séries chronologiques à queues lourdes. Ceci a été observé la première fois par Kesten (1973, p.379). C'est le but de ce chapitre pour expliquer d'où les queues lourdes viennent.

### 4.2 Les propriétés de base des processus *ARCH*

Dans le chapitre précédent, nous avons proposé un processus linéaire pour la modélisation des données financières. Ils peuvent être appropriés comme une première approximation, mais

souvent ne tiennent pas compte de la structure plus détaillée des données financières. Ces données présentent souvent les caractéristiques suivantes :

- (a) Presque aucune corrélation dans les données.
- (b) Variations de la volatilité dans le temps.
- (c) Les données sont à queue lourde.
- (d) Une forte corrélation entre les carrés et les valeurs absolues des données.
- (e) Le haut Seuil excédance apparaît dans les groupes.

Certaines de ces caractéristiques ne sont pas capturées par des processus linéaires. Différents modèles ont été mis en place visant à propriétés (a)-(e).

Cette section est consacrée à une classe particulière de tels modèles. En 1982, Engle [219] a introduit le processus Autorégressifs Conditionnellement Hétéroscédastiques d'ordre  $p$  ( $ARCH(p)$ ) pour  $p \in \mathbb{N}$ . Cette classe a été prolongée par Bollerslev (1986, p.77) qui a proposé une structure de dépendance alternative et plus flexible pour décrire les modèles  $ARCH$  ou  $GARCH(p, q)$  ( $p, q \in \mathbb{N}$ ). Le modèle  $GARCH(p, q)$  est définie par l'équations suivante :

$$Y_t = \sigma_t \varepsilon_t, \quad t \in \mathbb{N}, \quad (4.2.1)$$

où  $\{\varepsilon_t\}$  est une séquence de variables aléatoires iid de loi normale standard et  $\sigma_t$  s'écrit sous la forme suivante :

$$\sigma_t^2 = \beta + \sum_{i=1}^p \lambda_i Y_{t-i}^2 + \sum_{j=1}^q \delta_j \sigma_{t-j}^2, \quad (4.2.2)$$

avec  $\beta, \lambda_i$  et  $\delta_i$  sont des constantes positives à estimer et  $\sum_{i=1}^p \lambda_i + \sum_{j=1}^q \delta_j < 1$ .

Pour  $q = 0$  et  $p = 1$ , (4.2.1) et (4.2.2) se réduisent au modèle  $ARCH(1)$ , qui est défini par l'équation suivante :

$$Y_t = \sqrt{\beta + \lambda Y_{t-1}^2} \varepsilon_t, \quad t \in \mathbb{N}. \quad (4.2.3)$$

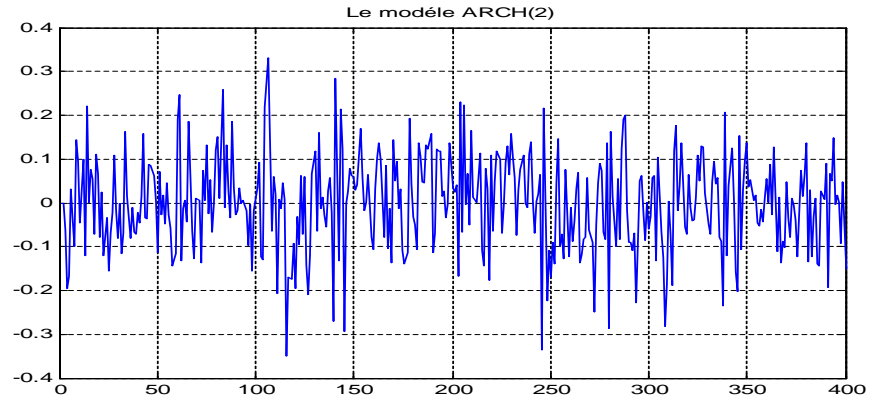


Figure1. Simulation du modèle  $ARCH(2)$

$$Y_t = (0.01 + 0.1Y_{t-1}^2 + 0.01Y_{t-2}^2)^{1/2} \varepsilon_t, n = 400, \varepsilon_t \rightarrow N(0, 1)$$

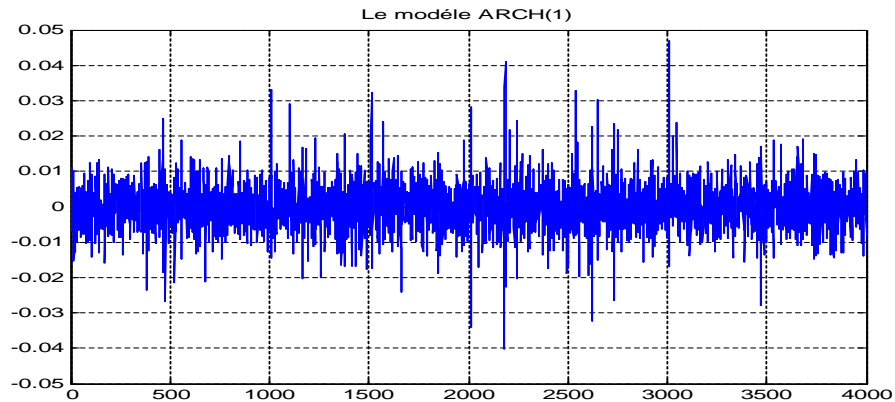


Figure2. Simulation du modèle  $ARCH(1)$

$$Y_t = (1.9 * 10^{-5} + 0.5Y_{t-1}^2)^{1/2} \varepsilon_t, n = 4000, \varepsilon_t \rightarrow N(0, 1)$$

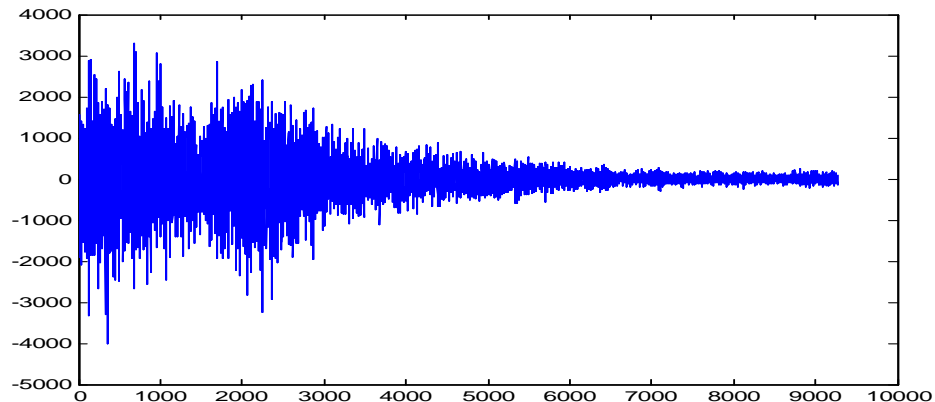


Figure 3. Données SP500 avec le modèle  $ARCH(1)$

$$Y_t = \left(1.9 * 10^{-5} + 0.5Y_{t-1}^2\right)^{1/2} \varepsilon_t, n = 10000, \varepsilon_t \rightarrow N(0, 1)$$

Pour une variable aléatoire initiale  $Y_0$  indépendante de  $\varepsilon_t$ , et les paramètres  $\beta > 0$  et  $\lambda > 0$ . Il s'agit d'un processus de Markov, donnée par la structure explicite autorégressif (4.2.3). Par construction,  $Y_{t-1}$  et  $\varepsilon_t$  sont indépendants pour chaque  $t \in \mathbb{N}$ , et  $Y_t$  sont des variables aléatoires, non corrélées et  $E(Y_0^2) < \infty$ .

Pour tirer les propriétés probabilistes du processus  $ARCH(1)$ , nous utiliserons le carré du processus  $ARCH(1)$  c'est-à-dire  $(Y_t^2)$  qui est défini par l'équation de récurrence stochastique suivante :

$$Y_t^2 = (\beta + \lambda Y_{t-1}^2) \varepsilon_t^2 = A_t + B_t Y_{t-1}^2, \quad t \in \mathbb{N}, \quad (4.2.4)$$

où

$$(A_t, B_t) = (\beta \varepsilon_t^2, \lambda \varepsilon_t^2). \quad (4.2.5)$$

**Lemme 4.2.1** Pour une variable aléatoire  $\varepsilon$  de loi normale standard et  $\lambda \in (0, 2e^\gamma)$ , où  $\gamma \approx 0.55772$  est une constante d'Euler, on a

$$h(u) = E(\lambda \varepsilon^2)^u, \quad u \geq 0,$$

alors

$$h(u) = \frac{(2\lambda)^u}{\sqrt{\pi}} \Gamma\left(u + \frac{1}{2}\right), \quad u \geq 0. \quad (4.2.6)$$

La fonction  $h$  est strictement convexe en  $u$ , et il existe une solution unique  $k = k(\lambda) > 0$  de l'équation  $h(u) = 1$ . En plus on a

$$k(\lambda) \begin{cases} > 1, & \lambda \in (0, 1) \\ = 1, & \lambda = 1 \\ < 1, & \lambda \in (1, 2e^\gamma) \end{cases}, \quad (4.2.7)$$

et

$$E \left[ (\lambda \varepsilon^2)^k \ln(\lambda \varepsilon^2) \right] > 0 \quad \text{et} \quad E \left[ \ln(\lambda \varepsilon^2) \right] < 0. \quad (4.2.8)$$

### Preuve

Remarquons d'abord que  $h(0) = 1$  pour tout  $\lambda$ . En outre,  $h$  a des dérivées pour tout les ordres. En particulier,

$$h'(u) = E \left[ (\lambda \varepsilon^2)^u \ln(\lambda \varepsilon^2) \right]. \quad (4.2.9)$$

$$h''(u) = E \left[ (\lambda \varepsilon^2)^u (\ln(\lambda \varepsilon^2))^2 \right] > 0. \quad (4.2.10)$$

La représentation (4.2.9) implique que :

$$\begin{aligned} h'(0) &= E \left[ \ln(\lambda \varepsilon^2) \right] = \ln \lambda + E \left( \ln \varepsilon^2 \right) \\ &= \ln(2\lambda) + \Gamma' \left( \frac{1}{2} \right) / \Gamma \left( \frac{1}{2} \right) = \ln(2\lambda) - \gamma - 2 \ln(2) \\ &= \ln(\lambda) - \ln(2) - \gamma < 0. \end{aligned} \quad (4.2.11)$$

Pour  $0 < \lambda < 2e^\lambda$ , où  $\gamma$  est une constante d'Euler. La représentation (4.2.10) implique que  $h$  est strictement convexe sur  $\mathbb{R}_+$ . Par la symétrie de la densité normale et de l'intégration

partielle, nous obtenons

$$\begin{aligned}
 h(u) &= E(\lambda\varepsilon^2)^u = \frac{\lambda^u}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^{2u} e^{-x^2/2} dx \\
 &= \frac{\lambda^u}{\sqrt{2\pi}} \int_0^{+\infty} (2y)^{u-1/2} e^{-y} dy \\
 &= \frac{(2\lambda)^u}{\sqrt{\pi}} \Gamma\left(u + \frac{1}{2}\right).
 \end{aligned}$$

donner la représentation (4.2.6). En plus,

$$h(u) > E[(\lambda\varepsilon^2)^u I_{\{\lambda\varepsilon^2 > 2\}}] \geq 2^u p(\lambda\varepsilon^2 > 2) \rightarrow \infty, u \rightarrow \infty.$$

Ce dernier fait, avec  $h(0) = 1$  et la convexité de  $h$  implique qu'il existe un unique  $k > 0$  tel que  $h(k) = 1$ . En plus,  $h'(k) > 0$ , ainsi que (4.2.9) et (4.2.10) donnant les inégalités (4.2.8). Depuis  $h(1) = \lambda$ , (4.2.7) suit d'un argument de monotonicity.

La valeur  $k = k(\lambda)$  est cruciale pour le comportement de queue du processus de distribution marginale, l'existence de moments et le comportement extrémales de processus ARCH(1).

L'équation  $h(u) = 1$  ne peut être résolu explicitement, mais les solutions numériques se trouvent dans le tableau suivant :

$\lambda$	0.1	0.3	0.5	0.7	0.9	1.0	1.5	2.0	2.5	3.0	3.5
$k$	13.24	4.18	2.37	1.59	1.15	1.0	0.54	0.31	0.17	0.075	0.007

Table. Les valeurs  $k = k(\lambda)$  pour  $\lambda \in (0, 2e^\gamma)$ .

### 4.2.1 Propriétés du processus ARCH(1) carré

**Théorème 4.2.1** (e.g Embrechts et al ,1997)

Soit  $\{Y_t\}$  un processus ARCH(1) donné par la représentation (4.2.3), pour  $\beta > 0$  et  $\lambda \in (0, 2e^\gamma)$ , où  $\gamma$  est une constante d'Euler, et supposons que  $Y_0$  est indépendante de  $\{\varepsilon_t\}$ .

a. Le processus  $\{Y_t\}$  est strictement stationnaire si

$$Y_0^2 \stackrel{d}{=} \beta \sum_{m=1}^{\infty} \varepsilon_m^2 \prod_{j=1}^{m-1} (\lambda \varepsilon_j^2), \quad (4.2.12)$$

de plus, chaque processus ARCH(1) strictement stationnaire est de distribution marginale suivante :

$$Y_t \stackrel{d}{=} |Y_0| r_0,$$

avec  $Y_0^2$  satisfait la représentation (4.2.12), et  $r_0$  est une variable aléatoire de Bernoulli avec  $p(r_0 = \pm 1) = 0.5$ , et aussi indépendante de  $|Y_0|$ .

b. Supposons que  $\{Y_t\}$  est strictement stationnaire et  $Y = Y_0, \varepsilon = \varepsilon_0$ . Soit  $k$  l'unique solution positive de l'équation  $h(u) = E(\lambda \varepsilon^2)^u = 1$ . Alors,

$$E(Y^2)^u < \infty \quad \text{pour } 0 \leq u < k.$$

Désignons par  $p$  le plus grand entier strictement inférieur à  $k$ . Alors, pour  $m = 1, \dots, p$  on a

$$EY^{2m} = (1 - E(\lambda \varepsilon^2)^m)^{-1} E\varepsilon^{2m} \sum_{k=0}^{m-1} \binom{m}{k} \lambda^k \beta^{m-k} EY^{2k} < \infty. \quad (4.2.13)$$

**Corollaire 4.2.1** Soit  $\{Y_t\}$  un processus stationnaire ARCH(1) avec les paramètres  $\beta > 0$  et  $\lambda \in (0, 1)$ . Alors, les relations suivantes sont valables

a)  $EY^2 = \beta / (1 - \lambda)$ .

b) Si  $\lambda^2 < 1/3$ , alors  $EY^4 < \infty$  et  $corr(Y_t^2, Y_0^2) = \lambda^t$  pour tout  $t \in \mathbb{N}$ .

**Remarque 4.2.1**

1) Pour l’instant, nous avons réalisé que  $\lambda$  est un paramètre essentiel du processus  $ARCH(1)$  (voir la représentation (4.2.3))

- Pour  $\lambda = 0$ ,  $\{Y_t\}$  est le bruit normal.

- Pour  $\lambda \in (0, 1)$ ,  $\{Y_t\}$  est stationnaire avec une variance finie.

- Pour  $1 \leq \lambda \leq 2e^\gamma \approx 3.56856$ ,  $\{Y_t\}$  est stationnaire avec une variance infinie.

2) Nous prévoyons en ce moment que  $EY^4 = \infty$  pour  $\lambda^2 \geq 1/3$ . Ceci suit de  $h(2) = 1$  pour  $\lambda = 1/\sqrt{3} \approx 0.577$  ainsi que la représentation (4.2.14) ci-dessous.

Dans ce cas, la notion d’autocorrélations n’a pas de sens pour le processus carré  $ARCH(1)$ . Néanmoins, l’échantillon des autocorrélations  $\tilde{\rho}_{Y^2}(t)$  sont bien définis pour  $Y_1^2, \dots, Y_n^2$ .

Contrairement aux processus linéaires nous ne sommes pas au courant de la suite de cohérence pour  $\tilde{\rho}_{Y^2}(t)$  ce qui explique son comportement pour de grandes  $n$ .

3) Pour  $\lambda^2 < 1/3$ , il est intéressant d’observer que le processus carré  $ARCH(1)$  a la même structure d’autocorrélation qu’un processus  $AR(1)$  :  $Y_t = \lambda Y_{t-1} + \varepsilon_t$ .

Dans le résultat suivant nous décrivons la queue de la distribution marginale d’un processus  $ARCH(1)$ . Un fait déjà observé par Kesten (1994, p.379) dans le contexte général d’équations aux récurrences stochastiques.

### 4.2.2 Comportement d’une queue d’un processus $ARCH(1)$

**Théorème 4.2.2 (Kesten, 1973 ; Goldie, 1991)**

Soit  $\{Y_t\}$  un processus  $ARCH(1)$  donné par la représentation (4.2.3), pour  $\beta > 0$  et  $\lambda \in (0, 2e^\gamma)$ , où  $\gamma$  est une constante d’Euler. Soit  $k$  l’unique solution positive de l’équation  $h(u) = 1$ . Alors,

$$p(Y > y) \sim \frac{c}{2} y^{-2k}, \quad y \rightarrow \infty, \tag{4.2.14}$$

où

$$c = \frac{E \left[ \left( (\beta + \lambda Y^2)^k - (\lambda Y^2)^k \right) (\varepsilon^2)^k \right]}{k E \left[ (\lambda \varepsilon^2)^k \ln (\lambda \varepsilon^2) \right]}. \quad (4.2.15)$$

pour une variable aléatoire d'une loi normale standard, indépendante de  $Y = Y_0$ .

Avant de démontrer ce résultat, nous devons d'abord assurer que la constante de la représentation (4.2.15) est bien définie. En effet, son numérateur est la différence attendue de deux grandeurs dont chacune ayant l'espérance infinie.

**Lemme 4.2.2** Sous les hypothèses du théorème 4.2.2,

$$0 < E \left[ (\beta + \lambda Y^2)^k - (\lambda Y^2)^k \right] < \infty. \quad (4.2.16)$$

**Lemme 4.2.3** Soit  $X \geq Y$  presque sûrement sont deux variables aléatoires non négatives.

Alors

$$\int_0^{\infty} (p(X > t) - p(Y > t)) t^{\delta-1} dt = \delta^{-1} E [X^\delta - Y^\delta], \quad \delta \neq 0.$$

### 4.2.3 Estimation des paramètres de modèle $ARCH(1)$ par la méthode MV

Soit  $\{Y_t\}$  un processus  $ARCH(1)$  donné :

$$\left\{ \begin{array}{l} Y_t = \sigma_t \varepsilon_t, \quad t \in \mathbb{N} \\ \sigma_t^2 = \beta + \lambda Y_{t-1}^2, \beta > 0, \lambda > 0 \end{array} \right\},$$

où  $(\varepsilon_t)$  est une séquence iid, telle que  $\varepsilon_t \sim L(0, 1)$ . On note  $\theta = (\beta, \lambda)'$ .

Le modèle  $ARCH(1)$  est classiquement estimé par la méthode maximum de vraisemblance. La fonction de vraisemblance prend la forme suivante :

$$\begin{aligned} L(\theta, Y_1, \dots, Y_n) &= f(Y_1, \dots, Y_n / Y_0) \\ &= f(Y_1 / Y_0) \cdot f(Y_2 / Y_1, Y_0) \dots f(Y_n / Y_{n-1}, \dots, Y_1, Y_0). \end{aligned}$$

Calcule de  $f(Y_t/Y_{t-1}, \dots, Y_1, Y_0) = ?$

Si  $\varepsilon_t \sim N(0, 1)$ , alors  $f_{\varepsilon_t}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$  et

$$\begin{aligned} F_{Y_t/Y_{t-1}, \dots, Y_1, Y_0}(x) &= p(Y_t \leq x/Y_{t-1}, \dots, Y_1, Y_0) \\ &= p\left(\frac{Y_t}{\sigma_t} \leq \frac{x}{\sigma_t}/Y_{t-1}, \dots, Y_1, Y_0\right) \\ &= p\left(\varepsilon_t \leq \frac{x}{\sigma_t}\right) \\ &= \Phi_{N(0,1)}\left(\frac{x}{\sigma_t}\right), \end{aligned}$$

alors

$$\begin{aligned} f_{Y_t/Y_{t-1}, \dots, Y_1, Y_0}(y_t) &= \frac{1}{\sigma_t} f\left(\frac{y_t}{\sigma_t}\right) \\ &= \frac{1}{\sigma_t \sqrt{2\pi}} e^{-\frac{1}{2} \frac{y_t^2}{\sigma_t^2}} \end{aligned}$$

et

$$\begin{aligned} f(Y_1/Y_0) &= \frac{1}{\sigma_1 \sqrt{2\pi}} e^{-\frac{1}{2} \frac{y_1^2}{\sigma_1^2}} \\ f(Y_2/Y_1, Y_0) &= \frac{1}{\sigma_2 \sqrt{2\pi}} e^{-\frac{1}{2} \frac{y_2^2}{\sigma_2^2}} \\ &\vdots \\ &\vdots \\ f(Y_n/Y_{n-1}, \dots, Y_1, Y_0) &= \frac{1}{\sigma_n \sqrt{2\pi}} e^{-\frac{1}{2} \frac{y_n^2}{\sigma_n^2}} \end{aligned}$$

$$L(\theta, Y_1, \dots, Y_n) = \frac{1}{(2\pi)^{n/2} \prod_{t=1}^n \sigma_t} \exp\left\{-\frac{1}{2} \sum_{t=1}^n \frac{y_t^2}{\sigma_t^2}\right\}.$$

Le logarithme de la vraisemblance est égale à :

$$\log L = \frac{n}{2} \log(2\pi) - \sum_{t=1}^n \sigma_t - \frac{1}{2} \sum_{t=1}^n \frac{y_t^2}{\sigma_t^2}.$$

## 4.3 Valeurs extrêmes pour les processus non linéaires

Avant de définir les valeurs extrêmes, on a besoin de la notion d'indice extrême.

### 4.3.1 Notion d'indice extrême

Afin d'étudier le comportement asymptotique du maximum d'une série de points issus d'un processus non linéaire, on utilise parfois la notion d'indice extrême. Celui-ci intervient, en général, comme on le verra plus loin, dans le calcul de la distribution limite. L'indice extrême est défini de la manière suivante :

**Définition** Soit  $\{Y_t, t \in \mathbb{Z}\}$  un processus strictement stationnaire de loi  $F(\cdot)$  et  $\theta$  un nombre positif. Supposons que pour tout  $\tau > 0$ , il existe une suite  $(u_n)$  telle que :

$$\lim_{n \rightarrow \infty} n(1 - F(u_n)) = \tau,$$

et

$$\lim_{n \rightarrow \infty} p(M_n \leq u_n) = e^{-\theta\tau},$$

alors  $\theta$  est appelé l'indice extrême de la suite  $\{Y_t\}$ .

Dans le cas de variables aléatoires iid ou de suites faiblement stationnaires,  $\theta$  est égale à 1.

Dans le cas des processus linéaires Gaussiens,  $\theta \in (0, 1]$ . Pour un processus linéaire défini l'équation suivante :

$$Y_t = \sum_{j=-\infty}^{\infty} \Psi_j \varepsilon_{t-j},$$

où les  $\Psi_j$  vérifient la condition  $\sum_{j=-\infty}^{\infty} |\Psi_j|^\alpha < \infty$  et où  $\varepsilon_t$  est un bruit blanc qui suit une

loi  $\alpha$ -stable caractérisée par la condition suivante :

$$p(\varepsilon_1 > y) \sim p(|\varepsilon_1| > y), \quad p(\varepsilon_1 \leq -y) \sim qp(|\varepsilon_1| > y),$$

alors

$$\theta = (\Psi_+^\alpha p + \Psi_-^\alpha q) / \|\Psi\|_\alpha^\alpha,$$

où

$$\Psi_+ = \max_j(\Psi_j, 0) \quad \text{et} \quad \Psi_- = \max_j((-\Psi_j), 0) \quad \text{et} \quad \|\Psi\|_\alpha^\alpha = \sum_{j=-\infty}^{\infty} |\Psi_j|^\alpha (pI_{\Psi_j > 0} + qI_{\Psi_j < 0}).$$

Les paramètres  $p$  et  $q$  dépendent du coefficient de symétrie de la fonction caractéristique de la variable  $\alpha$ -stable. De même si le bruit  $(\varepsilon_t)$  suit une loi de Gumbel, on peut montrer que :

$$\theta = (k^+ p + k^- q)^{-1},$$

où

$$k^+ = \text{card}j : \Psi_j = 1 \quad \text{et} \quad k^- = \text{card}j : \Psi_j = -1.$$

### 4.3.2 Valeurs extrêmes pour le processus $ARCH(1)$

Dans cette section on s'intéresse au comportement extrême du processus  $ARCH(1)$ , avec les paramètres  $\beta > 0$  et  $\lambda \in (0, 2e^\gamma)$ , où  $\gamma$  est la constante d'Euler. D'après la représentation (4.2.14) on a  $p(Y > y) \sim \frac{c}{2}y^{-2k}$ . Dans le cas où  $(Y_t)$  est une séquence iid, avec  $Y_1 \stackrel{d}{=} Y$ , donc  $Y \in MDA(\Phi_{2k})$ . En particulier,

$$\lim_{n \rightarrow \infty} p(n^{-1/(2k)} M_n \leq y) = \exp\left\{-\frac{c}{2}y^{-2k}\right\}, \quad (4.3.1)$$

où  $M_n = \max(Y_1, \dots, Y_n)$ , Il est maintenant naturel de demander : le processus strictement stationnaire  $ARCH(1)$  dont le comportement extrême comme une séquence iid avec la même distribution marginale?

Une réponse complète à cette question a été donnée par de Hann et al [303].

**Théorème 4.3.1 (Embrechts et al, 1997)**

Soit  $(Y_t^2)$  un processus stationnaire défini par la représentation (4.2.4) et  $M_n^2 = \max(Y_1^2, \dots, Y_n^2)$ .

Alors

$$\lim_{n \rightarrow \infty} p(n^{-1/k} M_n^2 \leq y) = \exp\left\{-c\theta^{(2)}y^{-k}\right\}, y > 0, \quad (4.3.2)$$

où  $k$  est une solution positive de l'équation  $E(\lambda\varepsilon^2)^u = 1$ ,  $c$  est définie par la représentation (4.2.15) et

$$\theta^{(2)} = k \int_1^\infty p\left(\max_{n \geq 1} \prod_{t=1}^n (\lambda\varepsilon_t^2) \leq y^{-1}\right) y^{-k-1} dy, \quad (4.3.3)$$

pour  $x > 0$ , soit

$$N_n^{(2)}(\cdot) = \sum_{i=1}^n \varepsilon_{n^{-1}i(\cdot)} I_{\{Y_i^2 > yn^{1/k}\}}.$$

le processus ponctuel des excédances au seuil  $yn^{1/k}$  par  $Y_1^2, \dots, Y_n^2$ . Alors

$$N_n^{(2)} \xrightarrow{d} N^{(2)}, \quad n \rightarrow \infty.$$

en  $M_p((0, 1])$ , où  $N^{(2)}$  est un processus composé de Poisson avec l'intensité  $c\theta^{(2)}y^{-k}$  et les probabilités

$$\pi_k^{(2)} = \frac{\theta_k^{(2)} - \theta_{k+1}^{(2)}}{\theta^{(2)}}, k \in \mathbb{N}, \quad (4.3.4)$$

où

$$\theta_k^{(2)} = k \int_1^\infty p\left(\text{card}\left\{n \in \mathbb{N} : \prod_{t=1}^n (\lambda\varepsilon_t^2) \geq y^{-1}\right\} = k - 1\right) y^{-k-1} dy.$$

en particulier  $\theta_1^{(2)} = \theta^{(2)}$ .

**Théorème 4.3.2 (Embrechts et al, 1997)**

Soit le processus *ARCH*(1) défini par la représentation (4.2.3). Alors

$$\lim_{n \rightarrow \infty} p(n^{-1/(2k)} M_n \leq y) = \exp\left\{-c\theta^{(2)}(1 - \Pi^{(2)}(0.5))y^{-2k}\right\}. \quad (4.3.5)$$

De plus, pour  $y > 0$  soit

$$N_n(\cdot) = \sum_{i=1}^n \varepsilon_{n^{-1}i(\cdot)} I_{\{Y_i > yn^{1/(2k)}\}},$$

le processus ponctuel des excédances au seuil  $yn^{1/(2k)}$  par  $Y_1, \dots, Y_n$ . Alors

$$N_n \xrightarrow{d} N, \quad n \rightarrow \infty.$$

en  $M_p((0, 1])$ , où  $N$  est un processus composé de Poisson avec l'intensité  $c\theta y^{-2k}$  et les probabilités

$$\pi_k = (1 - \Pi^{(2)}(0.5))^{-1} \sum_{m=k}^{\infty} \binom{m}{k} \pi_m^{(2)} 2^{-m}, \quad k \in N_0. \quad (4.3.6)$$

**Remarque 4.3.1** Une comparaison avec la représentation (4.3.1) prouve que  $\theta = 2\theta^{(2)}(1 - \Pi^{(2)}(0.5))$  est l'indice extrême du process  $(Y_t)$ .

### 4.3.3 Estimation de l'indice extrême $\zeta$ (Estimateur de Hill)

Parmi les estimateurs possibles pour l'indice extrême  $\xi$  est l'estimateur de Hill suivant :

$$\widehat{\zeta}_{n,k}^{(H)} = \frac{1}{k} \sum_{j=1}^k \ln X_{j,n} - \ln X_{k+1,n},$$

où  $X_{1,n} \geq X_{2,n}, \dots, \geq X_{n,n}$  sont les statistiques d'ordre et  $k = k(n)$  doit être choisi de façon appropriée. Cet estimateur, le plus utilisé en théorie des valeurs extrêmes.

On va simuler le processus  $ARCH(1)$  suivant :

$$Y_t = (1.9 * 10^{-5} + 0.5Y_{t-1}^2)^{1/2} \varepsilon_t.$$

On obtient les résultats suivants :

1. Si l'entrée  $(\varepsilon_t)$  est de distribution marginale de loi de Pareto avec l'indice de queue  $\alpha = 3$  on a trouvé  $\widehat{\zeta} = 0.333$  ( $\zeta = 1/\alpha$ ).

2. Si l'entrée  $(\varepsilon_t)$  est de distribution marginale de loi de cauchy avec l'indice de queue  $\alpha = 0.5$  on a trouvé  $\widehat{\zeta} = 2$  ( $\zeta = 1/\alpha$ ).

## 4.4 Estimation de l'indice de queue (Estimateur de Hill)

En particulier, il existe plusieurs estimateurs du paramètre  $\alpha$ . Parmi ces estimateurs on a l'estimateur de Hill. Après l'estimateur de Hill on trouve l'estimateur de Pickands , Dekkers-Einmahl et l'estimateur de Hann comme estimateurs les plus populaires.

D'après le chapitre 2 on a expliqué bien le raisonnement de l'estimateur de Hill, qui s'écrit essentiellement sur la forme suivante :

$$\widehat{\alpha}^{(H)} = \widehat{\alpha}_{k,n}^{(H)} = \left( \frac{1}{k} \sum_{j=1}^k \ln X_{j,n} - \ln X_{k,n} \right)^{-1},$$

où  $k = k(n) \rightarrow \infty$  d'une manière appropriée et  $X_{n,n} \leq \dots \leq X_{1,n} = M_n$  désignent les statistiques d'ordre de l'échantillon  $X_1, \dots, X_n$ . Ceci signifie qu'une suite croissante de statistiques d'ordre supérieure est utilisée.

### 4.4.1 Simulations et applications

Pour bien clarifier la différence entre l'indice des distributions marginales à queues lourdes d'un processus non linéaires et l'indice des distributions marginales à queues fines, on a besoin d'une simple application sur l'estimateur de l'indice de queue  $\alpha$  (Estimateur de Hill) et on comence par :

#### a. Indice des distributions marginales à queues lourdes

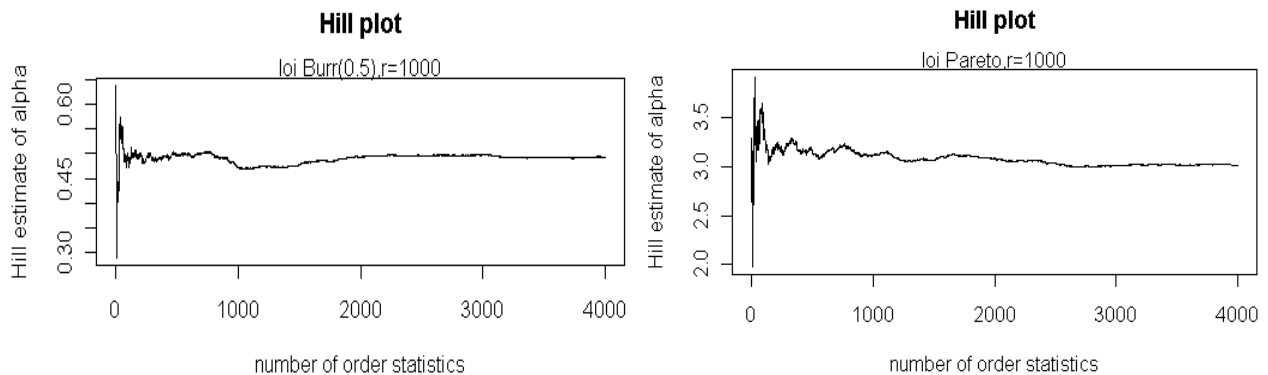
On a fait une simulation des modèles non linéaires de type  $ARCH(1)$

$$Y_t = \left( \sqrt{\beta + \lambda Y_{t-1}^2} \right) \varepsilon_t, t \in \mathbb{Z},$$

où  $\{\varepsilon_t, t \in \mathbb{Z}\}$  est de distribution marginale à queue lourde (*Pareto avec  $\alpha = 3$ , Burr avec  $\alpha = 0,5$* ) pour  $n = 4000$  avec  $r$  ( $r = 1000$ ) répétitions pour chaque échantillon, et on a obtenu les résultats suivantes :

1. Soit le modèle  $ARCH(1)$  suivant :

$$Y_t = \left( 1.9 * 10^{-5} + 0.5Y_{t-1}^2 \right)^{1/2} \varepsilon_t.$$

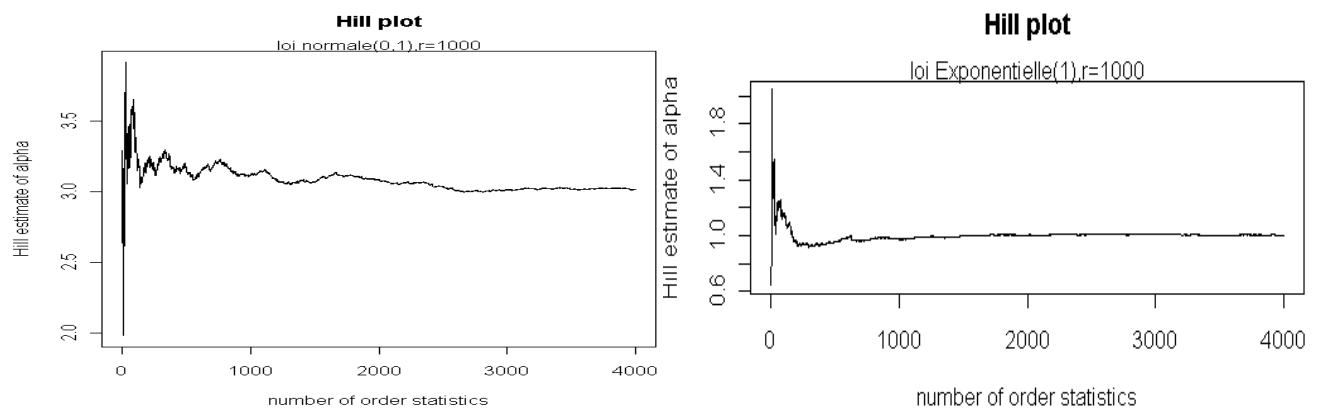


**b. Indice des distributions marginales à queues fines**

On a fait une simulation des modèles non linéaires de type  $ARCH(1)$ ,

$$Y_t = (1.9 * 10^{-5} + 0.5Y_{t-1}^2)^{1/2} \varepsilon_t,$$

où  $\{\varepsilon_t, t \in \mathbb{Z}\}$  est de distribution marginale à queue fine (*Exponentielle avec  $\alpha = 1$ , normale  $(0, 1)$* ) pour  $n = 4000$  avec  $r$  ( $r = 1000$ ) répétitions pour chaque échantillon, et on a obtenu les résultats suivantes :



**Remarque 4.2.2**

i. Dans les processus linéaires on a :

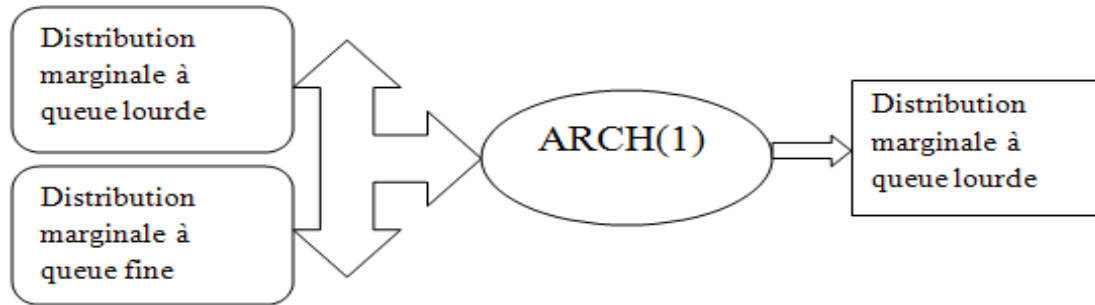
1. Si l'entrée du modèle est de distribution marginale à queue lourde, alors la sortie aussi est à queue lourde.

2. Si l'entrée du modèle est de distribution marginale à queue fine, alors la sortie aussi est à queue fine.

ii. Dans les processus non linéaires on a :

1. Si l'entrée du modèle est de distribution marginale à queue lourde, alors la sortie aussi est à queue lourde.

2. Si l'entrée du modèle est de distribution marginale à queue fine, alors la sortie est à queue lourde.



# Bibliographie

- [1] Aknouche, A. (2009). Processus aléatoires : théorie, méthodes et applications. Faculté de Mathématiques. Université Houari Boumediène, Alger.
- [2] Borovkov, A .A (1984). *Asymptotic Methods in Queueing Theory*. Wiley Chichester.
- [3] Bougerol, P. and Picard, N. (1992). Strict stationarity of generalised autoregressive processes. *Annals of probability*, **20**, 1714-1730.
- [4] Brandt, A. (1986). The stochastic equation  $Y_{n+1} = A_n Y_n + B_n$  with stationary coefficients. *Advanced in applied probability*, **18**, 211-220.
- [5] Breiman, L., (1968). *Probability*. Addison Wesley, California.
- [6] Brockwell, P.J. and Davis, R.A. (2002). *Introduction to time series and forecasting*. Springer-Verlag, New York.
- [7] Brockwell, P.J. and Davis, R.A. (1991). *Time Series : Theory and Methods*, 2nd ed., Springer-Verlag, New York.
- [8] Brockwell, P.J. and Davis, R.A. (1987). *Time series : Theory and Methods*. Springer-Verlag, New York, Berlin, Heidelberg, London, Paris, Tokyo.
- [9] Chistyakov, V. P (1964). A theorem on sums of independent positive random variables and its applications to branching random processes Theory Probab Appl. **9**, 640-648.
- [10] Chver, J. Ney, P and Wainger, S. (1972). Functions of probability measures. *J. Analyse Math.* **26**, 255-302.
- [11] Davis, R.A. and Mikosch, T. (2001). Point process convergence of stochastic volatility processes with application to sample autocorrelation, *J. Appl. Probab.*, **38A**, 93-104.
- [12] Davis, R. A. and Resnick, S. I. (1984) Tail estimates motivated by extreme value theory, *Ann. Statist.*, **12**, 1467-1487.

- [13] Davis, R.A. and Resnick, S. I. (1985). Limit theory for moving averages of random variables with regularly varying tail probabilities, *Ann. Probab.*, 13-1, 179–195.
- [14] Davis, R. A. and Resnick, S. I. (1988). Extremes of moving averages of random variables from the domain of attraction of the double exponential distribution, *Stochastic Process. Appl.*,30-1, 41–68.
- [15] de Haan, L and Ferreira, A. (2006). *Extreme value theory an introduction*. Springer.
- [16] Dekkers,A. L.M.and de Haan, L. (1989). On the estimation of the extreme-value index and large quantile estimation, *Ann. Statist.*, **17**, 1795–1832.
- [17] Dekkers and de Haan, L. (1993). Optimal choice of sample fraction in extreme-value estimation, *J. Multivariate Anal.*, 47-2, 173–195.
- [18] Dekkers, A. L. M, Einmahl, J. H. J and de Haan, L. (1989). A moment estimator for the index of an extreme-value distribution, *Ann. Statist*, **17**, 1833–1855.
- [19] Embrechts, P., Kluppelberg, C. and Mikosch, T. (1997). *Modelling extremal events for insurance and finance*. Springer, Berlin.
- [20] Embrechts, P. and Goldie, C.M. (1980). On closure and factorization theorems for subexponential and related distributions. *J. Austral. Math. Soc. Ser . A.* **29**,243-256.
- [21] Fama, E.F. (1965). The behaviour of stock market prices. *J. Busin. Univ. Chicago*, **38**, 34-105.
- [22] Feller, W (1968). *An Introduction to Probability, Theory and Its Applications*.
- [23] Fisher, R. and Tippett, L. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proceedings of the Cambridge Philosophical Society*, **24**, 180-190.
- [24] Guillaume,D. M, Dacarogna M.M, Dave R.R, Muller U.A,Olsen R.B.and Pictet,O.V. (1994). From the bird’s eye to the microscope A survey of new stylized facts of the intra daily foreign exchange markets Preprint Olsen Associates Zurich
- [25] Hill, M.B. (1975). A simple general approach to inference about the tail of a distribution. *Annals of statistics*. **3**, 1163-1174.
- [26] Hogg, R.V and Klugman, S.A.(1984). *Loss Distributions*.Wiley New York

- [27] Karlin, S. and Taylor, H.M. (1975). *A first course in stochastic processes*. Academic press, USA.
- [28] Longin, F (1996). The Asymptotic Distribution of Extreme Stock Market Returns J Busin "
- [29] Ibratimov, M. and Phillips, C.B. (1994). Testing the covariance stationarity of heavy-tailed time series : An overview of the theory with applications to several financial instruments. *J. Empirical Finance*, **1**,211-248.
- [30] Mikosch, T. (1999). *Regular variation, subexponentiality and their applications in probability theory*. University of Groningen.
- [31] Pakes, A.G. (1975). On the tails of waiting time distributions. *J Appl Probab.***12**, 555-564.
- [32] Resnick, S. (1987). *Extrem values, Regular variation, and point processes*. Springer, New York.
- [33] Resnick, S. (1992). *Adventures in stochastic processes*. Boston, Basel, Berlin.
- [34] Saporta, B. (2004). Etude de la solution stationnaire de l'équation  $Y_{n+1} = a_n Y_n + b_n$  à coefficients aléatoires. Thèse de Doctorat. Institut de recherche mathématique de Rennes. Ecole doctorale MATISSE.
- [35] Teugels, J. (1975). The class of subexponential distributions. *Annals of probability* **3**,1000-1011.