

UNIVERSITÉ MOULOUD MAMMERI DE TIZI OUZOU
FACULTÉ DES SCIENCES
DÉPARTEMENT DE MATHÉMATIQUES
ÉCOLE DOCTORALE DE STATISTIQUES



Mémoire de Magister

Spécialité : Mathématiques

Option : Statistiques

Modélisation des extrêmes spatiaux et applications

Présenté par :

M. AIT YALA Nabil

Devant le jury d'examen composé de :

M. Hocine FELLAG	Professeur	UMMTO	Président
M. Aliou DIOP	Professeur	UGB de Saint -Louis	Rapporteur
M. Djamel HAMADOUCHE	Professeur	UMMTO	Examineur
M. Mohand Arezki BOUDIBA	MC classe A	UMMTO	Examineur

Année universitaire : 2012-2013

Dédicaces

À la mémoire de mon grand-père

À mes grand-mères

À mes parents

REMERCIEMENTS

Louange à Allah le Seigneur des mondes. Avant tout, je remercie le bon Dieu de m'avoir aidé et donné de la force pour terminer ce mémoire, et sans qui ce travail n'aurait, sans doute, pas pu voir le jour.

Les remerciements ont toujours un parfum de convenu qui peut paraître parfois éculé et pourtant ils sont oh! combien importants. En effet, un travail, quel qu'il soit, n'est jamais individuel, de nombreuses personnes m'ont aidé et ont contribué, chacune à sa manière, à la réalisation de ce travail :

Je m'incline respectueusement devant les deux personnes qui sont chères pour moi, mon père et ma mère. Je leur exprime mes profonds signes de reconnaissance pour leurs efforts durant ces longues années d'études.

Je remercie beaucoup ma grand-mère **LAMARA MOHAMED Sma** qui n'a jamais cessé de prier pour ses petits-enfants et qui, seulement la regarder, inspire la confiance et la volonté de continuer malgré tous les obstacles qui s'opposent.

Dans le cadre de ce travail, je remercie **M. Aliou DIOP**, professeur à l'Université Gaston BERGER de Saint-Louis, pour avoir proposé le sujet. Je remercie également **M. Hocine FELLAG**, professeur à l'UMMTO, **M. Djamel HAMADOUCHE**, professeur à l'UMMTO, et **M. Mohand Arezki BOUDIBA**, maître de conférences classe A à l'UMMTO, pour avoir pris le temps de lire et juger ce travail.

Il est important pour moi de remercier tous mes frères et sœurs : **Esma, Mohamed, Oum Essaade, Billel, Khadidja, Sofiane, Idir** et **Amina** qui n'ont ménagé aucun effort pour que je puisse terminer ce travail.

Et enfin, je remercie tous ceux qui m'ont aidé, de près ou de loin, dans ce travail. En particulier, mes amis, sans citer leurs noms, qui m'ont soutenu et avec qui j'ai partagé les bons et les mauvais moments tout le long de la réalisation de ce travail.

TABLE DES MATIÈRES

Remerciements	ii
Table des matières	vi
Introduction générale	vii
I Théorie des valeurs extrêmes	1
1 Limite d'une somme de variables aléatoires	2
Introduction	2
1.1 Lois des grands nombres	3
1.1.1 Loi faible des grands nombres	4
1.1.2 Loi forte des grands nombres	10
1.2 Théorème de la limite centrale	15
1.2.1 Convergence en loi (Convergence en distribution)	16
1.2.2 Théorème de la limite centrale classique	17
1.2.3 Théorème de la limite centrale généralisé	21
Référence	31
2 Théorie des valeurs extrêmes univariée	32
Introduction	32
2.1 Loi max-stable	34
2.2 Distributions des valeurs extrêmes	36
2.2.1 Théorème de FISHER-TIPPETT (1928) & GNEDENKO (1943)	36
2.2.2 Exemples	38
2.2.3 Résultats avancés	39
2.3 Domaine d'attraction du maximum	39
2.3.1 Max-domaine d'attraction de FRÉCHET ($\text{MDA}(\Phi_\alpha)$)	42
2.3.2 Max-domaine d'attraction de WEIBULL ($\text{MDA}(\Psi_\alpha)$)	46
2.3.3 Max-domaine d'attraction de GUMBEL ($\text{MDA}(\Lambda)$)	48

2.4	Distribution des valeurs extrêmes généralisée (GEVD)	53
2.4.1	Représentation de VON MISES -JENKINSON	54
2.4.2	Propriétés de la GEVD	55
2.4.3	Max-domaine d'attraction de la GEVD	56
2.5	Distribution de PARÉTO généralisée (GPD)	58
2.5.1	Distribution des excès	58
2.5.2	Distribution de PARÉTO généralisée	59
2.5.3	Fonction moyenne des excès (MEF)	62
	Référence	64
II Extrêmes spatiaux et applications		65
3 Champs aléatoires		66
	Introduction	66
3.1	Champ aléatoire	68
3.1.1	Loi spatiale d'un champ aléatoire	68
3.1.2	Moments d'un champ aléatoire	70
3.2	Champ aléatoire au second ordre	72
3.2.1	Continuité en moyenne quadratique	73
3.2.2	Champ aléatoire gaussien	73
3.3	Stationnarité d'un champ aléatoire	75
3.3.1	Stationnarité stricte	76
3.3.2	Stationnarité au second ordre	77
3.3.3	Stationnarité intrinsèque	82
	Conclusion	89
4 Champs aléatoires max-stables		90
	Repères historiques et bibliographiques	90
4.1	Max-stabilité	91
4.1.1	Définition	91
4.1.2	Caractérisation	92
4.1.3	Standardisation	94
4.2	Représentations spectrales	96
4.2.1	Représentation de DE HAAN & PICKANDS	96
4.2.2	Représentation de SCHLATHER	97
4.3	Modèles des champs aléatoires max-stables	98
4.3.1	Modèle de SMITH	98
4.3.2	Modèle de SCHLATHER	100
4.3.3	Modèle de BROWN -RESNICK	102
4.4	Dépendance spatiale	104
4.4.1	Coefficient extrêmeal	105
4.4.2	Madogramme	108
4.5	Méthodes d'inférence	112
4.5.1	Estimation du coefficient extrêmeal	112
4.5.2	Vraisemblance composite	114

4.6	Exemple d'application	118
4.6.1	Description des données	118
4.6.2	Transformation des données	120
4.6.3	Modèle des données	126
4.6.4	Coefficient extrémal	127
4.6.5	Estimation des paramètres du modèle de SCHLATHER	128
	Conclusion	129
III Annexes		130
A Théorème de convergence des types		131
	Introduction	131
A.1	Type de distributions	132
A.2	Théorème de convergence des types	134
B Théorie de variation régulière		135
	Introduction	135
B.1	Fonction à variation lente	135
B.1.1	Représentation de KARAMATA	135
B.1.2	Exemples	136
B.1.3	Propriétés des fonctions aux variations lentes	136
B.1.4	Théorème de KARAMATA	137
B.2	Fonction à variation régulière	138
B.2.1	Théorème de caractérisation	138
B.2.2	Représentation de KARAMATA	139
B.2.3	Exemples	139
B.2.4	Propriétés des fonctions aux variations régulières	140
B.2.5	Théorème de KARAMATA	141
B.2.6	Théorème de densité monotone	142
B.2.7	Variation régulière et théorie des probabilités	143
B.3	Fonction à variation rapide	144
B.3.1	Théorème de représentation	144
B.3.2	Exemples	145
B.3.3	Propriétés des fonctions aux variations rapides	146
	Référence	146
C Processus Ponctuels de POISSON		147
	Introduction	147
C.1	Généralités	148
C.1.1	Mesure ponctuelle	148
C.1.2	Processus ponctuel	149
C.1.3	Loi de probabilité d'un processus ponctuel	152
C.1.4	Mesure d'intensité d'un processus ponctuel	153
C.2	Processus ponctuel de POISSON	153
C.3	PPP à espace d'état \mathbb{R}^d	155

D Introduction au logiciel R	159
D.1 Présentation du logiciel R	159
D.2 L'interface graphique de R	160
D.3 R et les statistiques	160
D.4 R et les graphiques	160
D.5 R et les packages	161
D.5.1 Installation des packages	161
D.5.2 Chargement des packages installés	162
Bibliographie	163
Résumé	171
Abstract	172

INTRODUCTION GÉNÉRALE

Avant de commencer ce mémoire, il convient de s'interroger un instant sur la notion de catastrophe. Est-ce un événement caractérisé par un coût financier colossal? Peut-être, mais rappelons que sur les marchés financiers, une variation de quelques pourcentages d'un indice boursier peut correspondre à une perte de plusieurs milliards d'euros, et personne ne parle de catastrophe. Faut-il alors la caractériser par le nombre de victimes? Là aussi, rappelons que de dizaines de milliers de personnes perdent chaque année la vie sur les routes, et que l'on parle d'accidents de la route, pas de catastrophe. En revanche, si un avion de ligne s'écrase, avec à son bord une centaine de passagers, on parle alors de catastrophe. De là, on peut définir une catastrophe comme étant un événement qui se produise pas fréquemment et qui a des conséquences funestes (événement rare et extrême).

Par définition, les événements rares sont des événements ayant une faible probabilité d'apparition. Ils sont dits extrêmes quand il s'agit de valeurs beaucoup plus grandes ou plus petites que celles observées habituellement. Lorsque le comportement de ces événements est dû au hasard on peut étudier leurs loi.

Étudier l'occurrence des événements extrêmes tels les catastrophes naturelles est de première importance pour les assureurs, ou les crises boursières pour les financiers ou, encore, les épidémies pour un état. Ce sont des événements rares aux conséquences désastreuses. C'est pourquoi est-il important de prévoir l'occurrence de tels événements. Et cela est l'objectif de la théorie des valeurs extrêmes.

Historiquement, l'étude de la loi de probabilité du maximum d'un échantillon de n variables aléatoires a été la première approche pour décrire les événements extrêmes. En 1928, FISHER & TIPPETT ont, les premiers, déduit d'une manière heuristique les lois limites possibles pour le maximum d'une suite de variables aléatoires indépendantes et de même loi, avant que GNEDENKO, en 1943, n'obtienne rigoureusement la convergence (elle a été simplifiée en 1976 par DE HAAN). Les applications ont commencé suite aux travaux de GUMBEL en 1958, en particulier en hydrologie.

Depuis quelques années, la théorie des valeurs extrêmes a reçu beaucoup d'attention aussi bien sur le plan théorique que sur le plan pratique. Les domaines d'application sont, en effet, très variés : hydrologie, météorologie, biologie, ingénierie, gestion de l'environnement, finance, assurance, sciences sociales, etc. En effet, la gestion des risques est devenue aujourd'hui fondamentale dans tous ces domaines.

En hydrologie, domaine dans lequel la prévision des crues, par exemple, est particulièrement importante [DAVISON & SMITH (1990)] et [KATZ, PARLANGE & NAVEAU (2002)]. En assurance dont l'une des préoccupations est la prise en compte des grands sinistres [ROOTZÉN & TAJVIDI (1997)] et [MC NEIL & SALADIN (1997)]. Leur introduction en finance [DANIELSSON & DE VRIES (1997)], [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)], [MC NEIL (1998)], [LONGIN (1998)], [LONGIN (2000)], [EMBRECHTS, RESNICK & SAMORODNITSKY (1999)] et [GENCAY & SELCUK (2004)] est une réponse immédiate à la remise en cause de l'hypothèse de normalité surtout avec les observations en hautes fréquences. En météorologie [COLES & WALSHAW (1994)], [SMITH (2001)] et [KLAJNMIC (2003)] où l'étude de la vitesse du vent, par exemple, permet d'évaluer le degré de résistance des matériaux face à la pression exercée par le vent (au cours d'une tempête par exemple) sur les bâtiments ou les structures de génie civil.

Dans les domaines des sciences humaines et sociales, la théorie des valeurs extrêmes pourra également contribuer à la compréhension de nombreux problèmes sociaux. À titre d'exemple, dans le domaine de la démographie, tout un débat qui a été initié par GUMBEL en 1937, auquel FRÉCHET a prit une part active, sur la notion de « durée extrême de la vie humaine » et sur sa mesure [THATCHER (1999)]. Des études récentes ont mis en lumière l'allure remarquable de la mortalité aux grands âges, à savoir une décroissance du taux de croissance de la mortalité à partir d'un certain âge. En 1998, HORIUCHI & WILMOTH ont ainsi montré que sur l'échelle logarithmique, la courbe des taux de mortalité présentait une allure concave aux grands âges. Ceci a conduit les démographes à rechercher des modèles compatibles avec cette réalité. Les démographes peuvent également se demander s'il existe un âge limite au-delà duquel nul ne survit ? Si oui, quel est cet âge ultime, auquel chacun peut rêver de parvenir ? La théorie des valeurs extrêmes permet de prévoir et d'extrapoler la distribution de probabilité de l'âge maximum que l'être humain pourrait atteindre [HAN (2005)]. La théorie des valeurs extrêmes va contribuer vivement au débat sur l'existence d'un âge limite de longévité de l'être humain. Une telle étude est souvent utile pour les organismes de sécurité sociale (dépenses destinées aux personnes âgées, coût de traitement, ...).

Exemple. [LEKINA (2010)].

Le premier février 1953, lors d'une forte tempête, la mer passe par-dessus plusieurs digues aux Pays-Bas, les détruit et inonde la région. Il s'agit d'un accident majeur. Un comité est mis en place pour étudier le phénomène et proposer des recommandations sur les hauteurs de digues. Il doit tenir compte des facteurs économiques (coût de construction, coût des inondations, etc.), des facteurs physiques (rôle du vent sur la marée, etc.), et aussi des données enregistrées sur les hauteurs de marées. En fait, il est plus judicieux de considérer les surcotes, c'est-à-dire la différence entre la hauteur réelle et la hauteur prévue de la marée, que les hauteurs des marées. En effet, on peut supposer, dans une première approximation, que les surcotes des marées lors des tempêtes sont des réalisations de variables aléatoires de même loi. Si on regarde les surcotes pour des marées de tempêtes séparées par quelques jours d'accalmie, on peut même supposer que les variables aléatoires sont indépendantes. L'étude statistique sur des surcotes a pour but de répondre aux questions suivantes :

1. Trouver la hauteur de digues dont la probabilité d'être dépassée par la surcote est $\alpha \in]0, 1[$.
2. Étant donné une hauteur des digues, trouver la probabilité qu'elle soit dépassée par la plus haute surcote annuelle.

La réponse aux questions soulevées dans cet exemple peut fournir des éléments indispensables pour construire aux endroits critiques des digues d'une hauteur appropriée, déterminer les zones inconstructibles, définir la périodicité des opérations de nettoyage des grands cours d'eaux et des estuaires afin de protéger efficacement la population et pourquoi pas les biens.

D'un point de vue mathématique, il s'agit de modéliser le phénomène aléatoire, en s'intéressant principalement non pas au « corps » de la distribution, mais à la « queue » de la distribution. C'est-à-dire qu'on modélise le phénomène uniquement d'un point de vue asymptotique, dans le régime des valeurs extrêmes. Cette idée a donné lieu à de fructueux développements mathématiques, notamment :

- La notion de variation régulière qui joue un rôle important. En supposant que les distributions ont une certaine régularité dans le domaine des valeurs extrêmes, elle permet d'extrapoler le comportement des extrêmes d'un échantillon de taille modérée. Ceci est particulièrement important dans la pratique car les extrêmes sont par essence même peu fréquemment observés : Comment prédire par exemple l'ampleur d'un phénomène arrivant tous les 100 ans à partir d'observation sur les 10 dernières années ? La notion de variation régulière permet une telle extrapolation.
- Les théorèmes limites pour les maximums renormalisés, la caractérisation des lois GEVD (Generalized Extreme Value Distribution) et de leurs max-domaines d'attraction et la notion de max-stabilité.
- La modélisation des excès au delà d'un seuil élevé, la méthode POT (Peaks Over Threshold) et les lois GPD (Generalized PARÉTO Distribution).

Ces notions ont d'abord été développées pour la théorie univariée des extrêmes et les échantillons indépendants, puis étendues au cadre multivarié et aux échantillons dépendants.

Aujourd'hui, une direction de recherche particulièrement active est celle de la théorie spatiale des extrêmes, où l'on quitte le contexte fini-dimensionnel pour se placer dans un cadre fonctionnel et où les processus et les champs aléatoires occupent une place prépondérante. Cette composante spatiale est présente naturellement dans les applications aux sciences de l'environnement, sciences de la terre et sciences humaines citées plus haut, et un enjeu important est celui de la modélisation de la dépendance spatiale des extrêmes et des problèmes d'inférence relatifs. Dans ce mémoire on essayera de donner une introduction générale à tout ces concepts, en particulier la théorie spatiale des valeurs extrêmes et la théorie des champs aléatoires max-stables.

Organisation du mémoire

Ce mémoire s'organise en trois parties. Les deux premières parties sont principales, et elles ont été rédigées d'une façon à être lues l'une après l'autre ainsi que les chapitres qui les composent. La troisième partie comporte des outils nécessaires qui aident à mieux comprendre les résultats des deux premières parties.

La première partie s'intitule « **Théorie des valeurs extrêmes** » dont l'objectif est d'exposer la théorie probabiliste des valeurs extrêmes dans le cas univarié sans présenter la théorie statistique et ses applications. Elle se compose de deux chapitres :

Le premier chapitre : Dans ce chapitre, on étudie le comportement asymptotique d'une somme de variables aléatoires comme une introduction à la théorie des valeurs extrêmes. Car la théorie des valeurs extrêmes et la théorie de la limite centrale sont développées en parallèle et, en fait, les deux théories portent quelques ressemblances : Le résultat fondamental de la théorie des valeurs extrêmes, qui est le théorème de FISHER-TIPPETT (1928) & GNEDENKO (1943), est l'analogie du théorème de la limite centrale pour le maximum d'une suite de variables aléatoires indépendantes et de même loi. Du coup, on va essayer de revoir les différents modes de convergence d'une suite de variables aléatoires, surtout la convergence en loi (Convergence faible), et leurs propriétés.

Le deuxième chapitre : Dans ce chapitre, on expose la théorie probabiliste des valeurs extrêmes dans le cas univarié. Au début, on donne quelques résultats et quelques définitions de certains outils nécessaires dans cette théorie. Ensuite, dans la première section, on définira la loi max-stable sans donner ses propriétés. Dans la deuxième section, on présentera le résultat principal de la théorie des valeurs extrêmes qui montre que la loi limite du maximum d'une suite de variables aléatoires indépendantes et de même loi ne peut appartenir qu'à l'un des trois types de distributions suivants : GUMBEL, WEIBULL ou FRÉCHET. Et dans ce cas, on dit que la loi appartient au max-domaine d'attraction de type GUMBEL, WEIBULL ou FRÉCHET, et la classe de ces lois limites coïncide avec la classe des lois max-stables. La troisième section présente les conditions nécessaires et/ou suffisantes pour qu'une loi de probabilité appartienne à l'un des max-domaines d'attraction cités ci-dessus. Dans la quatrième section, on verra qu'on peut unifier les trois types de distributions précédents dans un seul type de distributions qui contient les lois de probabilité de VON MISES-JENKINSON qui sont plus connues sous le nom « Distributions des valeurs extrêmes généralisées (GEVD) », et qui jouent un rôle important dans la méthode des « Maximums par blocs » qui sert à estimer l'**indice de queue**. Ensuite, on donnera les conditions nécessaires et/ou suffisantes pour qu'une loi de probabilité appartienne au max-domaine d'attraction de la GEVD. Enfin, dans la cinquième section, on définit la distribution de PARÉTO généralisée « Generalized PARÉTO Distribution (GPD) » qui est un outil fondamental pour la méthode des « Excès au-delà d'un seuil », une méthode fondée sur la théorie des valeurs extrêmes également connue sous le nom « Peaks Over Threshold (POT) ». Cette méthode permet la modélisation des queues de distributions d'une série de données à partir de laquelle il devient possible d'estimer la probabilité d'occurrence des événements rares au-delà des plus grandes valeurs observées. Après, on donnera le théorème de BALKEMA-DE HAAN & PICKANDS qui permet de noter le lien qui existe entre le max-domaine d'attraction de la GEVD et le comportement limite de la GPD.

La deuxième partie, qui est la partie essentielle de ce mémoire, s'intitule « **Extrêmes saptiaux et applications** ». Elle se compose de deux chapitres :

Le troisième chapitre : Dans ce chapitre, on présentera une introduction à la théorie des champs aléatoire, qui une généralisation de la théorie des processus stochastiques et de celle des séries temporelles, en donnant les outils essentiels de cette théorie de qui nous aurons besoin au quatrième chapitre qui est le cœur du sujet de ce mémoire. Dans la première section, on donnera la définition d'un champ aléatoire ainsi que ses moments, lorsque ceux-ci existent, qui sont des grandeurs nécessaires pour l'analyse statistique des données. Dans ce cadre on va introduire un nouveau outil qui est « **le variogramme** » qu'on utilisera à la place de la fonction de covariance lorsque celle-ci n'existe pas. Dans la deuxième section on présentera la classe des « **champs aléatoires au second ordre** » qui est une classe très importante, car ses éléments sont des champs aléatoires ayant tout

les moments définis dans la première section. Et comme exemple de cette classe, on donnera « **le champ aléatoire gaussien** » qui est très important dans la théorie des champs aléatoires et ses applications. Dans la troisième section, on présentera quelques modes de stationnarité d'un champ aléatoire surtout « **la stationnarité intrinsèque** » qui est une stationnarité non propre au champ aléatoire lui-même mais à ses accroissements. Cette stationnarité qui engendre une classe plus vaste que celle des champs aléatoires stationnaires au second ordre nous assure l'existence du variogramme de ses éléments et pas de leurs fonctions de covariance. À la fin, on donnera quelques modèles des variogrammes souvent rencontrés dans la littérature.

Le quatrième chapitre : Dans ce chapitre, qui est le chapitre essentiel du sujet de ce mémoire, on présente les outils de base pour l'étude et l'application des champs aléatoires max-stables. Dans la première section, on donne la définition des champs aléatoires max-stables et on va voir que ce n'est que la généralisation de la max-stabilité, définie dans le deuxième chapitre pour les variables aléatoires réelles, au cas des champs aléatoires. Ensuite, on donnera un résultat important dans la théorie des valeurs extrêmes spatiale, qui est en quelques sortes une extension du théorème de FISHER-TIPPETT (1928) & GNEDENKO (1943) au cas spatial, et qui nous montre que les champs aléatoires max-stables sont la généralisation des distributions des valeurs extrêmes au contexte spatial et donc des candidats idéaux pour la modélisation des extrêmes spatiaux. Après, on va voir qu'on peut transformer, sans perte de généralisation, les lois de probabilités marginales d'un champ aléatoire max-stable en lois de probabilité de FRÉCHET standards unitaires, ce qui est important pour les applications. Dans la deuxième section, on va donner les représentations spectrales des champs aléatoires max-stables, en particulier la représentation de DE HAAN & PICKANDS et la représentation de SCHLATHER. Ces représentations spectrales sont à la base de la construction de certains modèles des champs aléatoires max-stables, qu'on présentera dans la troisième section, qui nous permette de mettre en applications les champs aléatoires max-stables. Dans la quatrième section, on définira « **le coefficient extrémal** » qui est un outil qui caractérise la structure, non pas entière, de dépendance entre les extrêmes spatiaux, et qui mesure le degré de dépendance entre les réalisations d'un champ aléatoire max-stable. Ensuite, on définira deux autres moments d'un champ aléatoire qui caractérise la dépendance entre ses réalisations, « **le madogramme** » et « **le F -madogramme** », qui seront utilisés pour l'estimation du coefficient extrémal d'un champ aléatoire max-stable qui est l'objet de la cinquième section où on présente, aussi, la méthode de « **vraisemblance composite** » qui sert à estimer les paramètres d'un champ aléatoire max-stable. Enfin, dans la sixième section, on terminera ce chapitre par un exemple d'application sur des données réelles où l'objectif est de déterminer un modèle des champs aléatoires max-stables approprié à ces données.

La troisième partie s'intitule « **Annexes** ». Et comme son nom l'indique, elle comporte les annexes de ce mémoire qui sont des outils probabilistes et mathématiques qui complètent les résultats des deux premières parties. Elle se compose de quatre chapitres dont les titres sont les suivants : « Théorème de convergence des types », « Théorie de variation régulière », « Processus ponctuels de POISSON » et « Introduction au logiciel **R** ».

Première partie

Théorie des valeurs extrêmes

CHAPITRE 1

LIMITE D'UNE SOMME DE VARIABLES ALÉATOIRES

Introduction

En guise d'exemple introductif, considérons cet exemple qui montre l'importance d'étudier le comportement asymptotique d'une somme de variables aléatoires dans la théorie des probabilités :

On se trouve en présence d'une urne dont on sait seulement qu'elle contient des boules vertes et des boules rouges. On tire une boule de cette urne et on cherche la probabilité d'avoir une boule verte.

Si on note A l'événement « avoir une boule verte » et p la probabilité que l'événement A se réalise, alors, d'après **la définition classique de la probabilité**, on a :

$$p = P(A) = \frac{N_A}{N}$$

où : N_A est le nombre de boules vertes dans l'urne et N le nombre total des boules.

Malheureusement, on connaît pas les nombres N et N_A , alors pour connaître (approximativement) la proportion p des boules vertes dans l'urne on procède de la manière suivante : on effectue n tirages d'une boule avec remise et, à chaque tirage, on note sa couleur. On modélise cette expérience en prenant $\Omega = \{\text{verte}, \text{rouge}\}$ et X_i la variable aléatoire définie, sur l'espace probabilisé $[\Omega, \mathcal{P}(\Omega), P]$ dans l'ensemble $\{0, 1\}$, comme suit :

$$\forall \omega \in \Omega, \quad X_i(\omega) = \begin{cases} 1 & \text{si la boule tirée au } i^{\text{ème}} \text{ tirage est verte,} \\ 0 & \text{si la boule tirée au } i^{\text{ème}} \text{ tirage est rouge.} \end{cases}$$

Alors les X_i sont des variables aléatoires indépendantes qui suivent une loi de probabilité de BERNOULLI de paramètre p , car :

$$P(X_i = 1) = P(\text{la boule tirée au } i^{\text{ème}} \text{ tirage est verte}) = P(A) = p$$

Soit $S_n = \sum_{i=1}^n X_i$ la variable aléatoire somme des variables aléatoires X_i , qui prend ses valeurs dans l'ensemble $\{0, 1, 2, \dots, n\}$, et qui compte le nombre d'apparition de la boule verte dans les n tirages effectués. Alors d'après le théorème de BERNOULLI, on a :

$$\forall \varepsilon > 0, \quad \mathbb{P}\left(\left|\frac{S_n}{n} - p\right| \geq \varepsilon\right) \leq \frac{1}{4n\varepsilon^2}$$

Ce théorème montre que la probabilité que S_n/n , somme des variables aléatoires X_i divisée par le nombre de tirages effectués n , s'écarte de la probabilité p d'avoir une boule verte tend vers 0 lorsque n tend vers l'infini.

Donc pour connaître, approximativement, la probabilité d'un événement issu d'une expérience aléatoire, il suffit de répéter cette expérience un très grand nombre de fois et diviser le nombre de réalisations de cet événement dans les expériences effectuées sur le nombre total des expériences. Et on appelle cette définition, basée sur le théorème de BERNOULLI ci-dessus, **la définition fréquentiste de la probabilité**.

En applications, l'étude des sommes de variables aléatoires est très importante, car une population est souvent décrite statistiquement ou résumée par la donnée d'une statistique moyenne (âge moyen, poids moyen, taille moyenne, prix moyen, etc.), où les données récoltées sont supposées être des réalisations de certaines variables aléatoires indépendantes et identiquement distribuées et la statistique moyenne, dite empirique, est la somme de ces variables aléatoires divisée par le nombre de données récoltées.

Ce chapitre présente deux résultats fondamentaux dans la théorie des probabilités et des statistiques, **la loi des grands nombres** et **le théorème de la limite centrale**. Pour voir en quoi ils consistent ces résultats, soit $\{X_n, n \in \mathbb{N}^*\}$ une suite de variables aléatoires définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans \mathbb{R} et soit $S_n = \sum_{i=1}^n X_i$ la somme des n premiers termes de cette suite.

1. La loi des grands nombres : Sous certaines conditions sur les variables aléatoires X_i , la loi des grands nombres, qui nous donne ces conditions, nous affirme que la suite de variables aléatoires $\left\{n^{-1} \sum_{i=1}^n X_i, n \in \mathbb{N}^*\right\}$ converge en probabilité, même presque sûrement, vers une constante qu'on précisera plus loin.

2. Le théorème de la limite centrale : Sous certaines conditions sur les variables aléatoires X_i et s'il existe des constantes de normalisations $a_n \in \mathbb{R}_+^*$ et $b_n \in \mathbb{R}$ telles que la suite de variables aléatoires normalisées $\left\{a_n^{-1}(S_n - b_n), n \in \mathbb{N}^*\right\}$ converge en loi vers une variable aléatoire, non dégénérée Y , alors le théorème de la limite centrale nous donne ces conditions qui assurent cette convergence et nous précise la forme de la loi de probabilité de Y .

1.1 Lois des grands nombres

Avant d'exposer les résultats envisagés pour ce chapitre, on a vu qu'il est essentiel de définir, au fur et à mesure, certains modes de convergence d'une suite de variables aléatoires qui sont essentiels

pour les applications.

Une variable aléatoire X définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ dans $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ étant une fonction de Ω dans \mathbb{R} . Il serait naturel de considérer qu'une suite de variables aléatoires $(X_n)_{n \in \mathbb{N}^*}$ définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ dans $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ converge vers la variable aléatoire X si et seulement si pour tout élément ω de Ω , la suite réelle $(X_n(\omega))_{n \in \mathbb{N}^*}$ converge vers le réel $X(\omega)$, c'est à dire, la suite des fonctions $(X_n)_{n \in \mathbb{N}^*}$ converge simplement vers la fonction X .

Malheureusement, cette définition est très restrictive, car elle ne fait pas intervenir la probabilité sous-jacente.

1.1.1 Loi faible des grands nombres

I Convergence en probabilité (convergence stochastique)

Lorsque la suite réelle $(X_n(\omega))_{n \in \mathbb{N}^*}$ converge vers le réel $X(\omega)$, on doit s'attendre à ce que pour tout réel strictement positif ε , il existe un entier n_0 tel que, pour tout entier n supérieur à n_0 , $|X_n(\omega) - X(\omega)|$ soit inférieur à ε . Dans la droite ligne de ce raisonnement, on peut ainsi avoir l'idée d'exprimer que la probabilité de tout événement de la forme $(|X_n - X| > \varepsilon)$, où ε est un réel strictement positif, tend vers 0 quand n tend vers l'infini. D'où la définition suivante :

Définition 1.1.1 [CAMINADE & NICOLAS (2008)].

Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires réelles définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, et soit X une variable aléatoire réelle définie, elle aussi, sur le même espace.

On dit que la suite $(X_n)_{n \in \mathbb{N}^*}$ converge **en probabilité (stochastiquement)** vers X , et on note $X_n \xrightarrow{p} X$, si et seulement si, pour tout réel strictement positif ε , la probabilité de l'événement $(|X_n - X| > \varepsilon)$ tend vers 0 quand n tend vers l'infini, i.e :

$$X_n \xrightarrow{p} X \iff \forall \varepsilon \in \mathbb{R}_+^*, \quad \lim_{n \rightarrow +\infty} \mathbb{P}(|X_n - X| > \varepsilon) = 0$$

I.1 Problème de l'unicité de la limite

Dans le cas de limites de suites réelles, de suites de vecteurs ou de suites de fonctions, si une suite converge vers une limite, sa limite est unique. Dans le cas de la convergence en probabilité des suites de variables aléatoires, la situation est un peu différente. On dispose du théorème suivant :

Proposition 1.1.1 [CAMINADE & NICOLAS (2008)].

Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires réelles définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, et soient X et Y deux variables aléatoires réelles définies, elles aussi, sur le même espace telles que $X_n \xrightarrow{p} X$ et $X_n \xrightarrow{p} Y$. Alors $\mathbb{P}(X \neq Y) = 0$.

Cette proposition nous dit que la limite, au sens de la convergence en probabilité, d'une suite de variables aléatoires est presque sûrement unique pour la probabilité \mathbb{P} .

I.2 Composition par une fonction continue

Proposition 1.1.2 [OUVRARD (2000)].

Soit f une fonction continue sur \mathbb{R}^d ($d \geq 1$) dans \mathbb{R} . Si $(X_{1,n}), (X_{2,n}), \dots, (X_{d,n})$ ($n \in \mathbb{N}^*$) sont d suites de variables aléatoires réelles qui convergent en probabilité respectivement vers les variables aléatoires réelles X_1, X_2, \dots, X_d . Alors la suite de variables aléatoires réelles $\left(f(X_{1,n}, X_{2,n}, \dots, X_{d,n})\right)_{n \in \mathbb{N}^*}$ converge en probabilité vers la variable aléatoire réelle $f(X_1, X_2, \dots, X_d)$.

I.3 Inégalités de MARKOV, de TCHEBYCHEV et de BIENAYMÉ-TCHEBYCHEV

Pour démontrer une convergence en probabilité, on a souvent recours à des inégalités que les règles élémentaires de probabilités nous fournissent. Parmi les inégalités probabilistes qui jouent un grand rôle théorique, on distingue les inégalités de MARKOV, et de BIENAYMÉ-TCHEBYCHEV qui servent essentiellement à démontrer des convergences en probabilité.

Théorème 1.1.1 (1^{ère} inégalité de MARKOV) [ATHREYA & LAHIRI (2006)].

Soit X une variable aléatoire réelle **positive** admettant une espérance $E(X)$, alors :

$$\forall \varepsilon > 0, \quad P(X \geq \varepsilon) \leq \frac{E(X)}{\varepsilon}$$

Corollaire 1.1.1 (2^{ème} inégalité de MARKOV) [ATHREYA & LAHIRI (2006)].

Soient X une variable aléatoire réelle et p un nombre réel strictement positif tels que $E(|X|^p) < \infty$, alors :

$$\forall \varepsilon > 0, \quad P(|X| \geq \varepsilon) \leq \frac{E(|X|^p)}{\varepsilon^p}$$

Corollaire 1.1.2 (Inégalité de MARKOV générale) [ATHREYA & LAHIRI (2006)].

Soient X une variable aléatoire réelle et $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ une fonction mesurable et croissante telles que $E[g(|X|)] < \infty$, alors :

$$\forall \varepsilon > 0, \quad P(|X| \geq \varepsilon) \leq \frac{E[g(|X|)]}{g(\varepsilon)}$$

Corollaire 1.1.3 (Inégalité de TCHEBYCHEV) [ATHREYA & LAHIRI (2006)].

Soit X une variable aléatoire réelle admettant une espérance $E(X)$ et une variance $V(X)$, alors :

$$\forall \varepsilon > 0, \quad P(|X - E(X)| \geq \varepsilon) \leq \frac{V(X)}{\varepsilon^2}$$

On remarque que l'inégalité de TCHEBYCHEV est une conséquence directe de la 2^{ème} inégalité de MARKOV en remplaçant X par $|X - E(X)|$ et en prenant $p = 2$.

Maintenant, l'inégalité qui sert à démontrer des convergences en probabilité et, en particulier, la loi faible des grands nombres est l'inégalité de BIENAYMÉ-TCHEBYCHEV présentée dans la proposition qui suit :

Proposition 1.1.3 (Inégalité de BIENAYMÉ-TCHEBYCHEV) [BARBE & LEDOUX (2007)].

Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires réelles, deux à deux **non corrélées** et admettent, toutes, un moment de second ordre, et soit $S_n = \sum_{i=1}^n X_i$, alors :

$$\forall \varepsilon > 0, \quad \mathbb{P}\left(|S_n - \mathbb{E}(S_n)| \geq \varepsilon\right) \leq \frac{\mathbb{V}(S_n)}{\varepsilon^2}$$

Remarque : Rappelons que deux variables aléatoires réelles X et Y sont non corrélées si leur covariance est nulle. Donc, si X et Y sont indépendantes, alors elles sont non corrélées mais l'inverse n'est pas toujours vrai sauf dans le cas gaussien.

I.4 Critère de la convergences en probabilité

Grâce aux inégalités de MARKOV et de TCHEBYCHEV, on peut déduire des conditions suffisantes pour qu'une suite de variables aléatoires converge en probabilité. On résume ces conditions dans la proposition suivante :

Proposition 1.1.4 [FOATA & FUCHS (1998) et CAMINADE & NICOLAS (2008)].

Soient $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires réelles définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et X une variable aléatoire réelle définie, elle aussi, sur le même espace. On a :

1.

$$\left(\exists p > 0 \quad \text{tel que} \quad \lim_{n \rightarrow +\infty} \mathbb{E}(|X_n - X|^p) = 0\right) \implies X_n \xrightarrow{p} X$$

2.

$$\left(\lim_{n \rightarrow +\infty} \mathbb{E}(X_n - X) = 0 \quad \text{et} \quad \lim_{n \rightarrow +\infty} \mathbb{V}(X_n - X) = 0\right) \implies X_n \xrightarrow{p} X$$

II Loi faible des grands nombres

Soient $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans \mathbb{R} et $S_n = \sum_{i=1}^n X_i$ la somme des n premiers termes de la suite $(X_n)_{n \in \mathbb{N}^*}$.

Définition 1.1.2 [FOATA & FUCHS (1998)].

On dit que la suite $(X_n)_{n \in \mathbb{N}^*}$ satisfait **la loi faible des grands nombres** si la suite de variables aléatoires réelles $\left\{n^{-1} \sum_{i=1}^n X_i, n \in \mathbb{N}^*\right\}$ converge en probabilité vers une constante.

Remarque : Il existe plusieurs versions de la loi faible des grands nombres, où, chacune donne la constante vers laquelle la suite $\left\{n^{-1} \sum_{i=1}^n X_i, n \in \mathbb{N}^*\right\}$ converge en probabilité et donne les conditions suffisantes, sur les variables aléatoires X_i , pour que cette convergence aura lieu.

Sachant que toute les versions de la loi faible des grands nombres exigent, au moins, la non corrélation, on s'intéresse dans ce mémoire aux versions les plus célèbres et les plus utilisées. En voici quelques unes :

Théorème 1.1.2 (Loi faible des grands nombres) [FOATA & FUCHS (1998)].

Soit $\{X_n, n \in \mathbb{N}^*\}$ une suite de variables aléatoires, **deux à deux non corrélées**, définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans \mathbb{R} et telle que :

$$\forall i \in \mathbb{N}^*, \quad \mathbb{E}(X_i^2) < \infty$$

Si :

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathbb{E}(X_i) = \mu < \infty \quad \text{et} \quad \lim_{n \rightarrow \infty} \frac{1}{n^2} \sum_{i=1}^n \mathbb{V}(X_i) = 0$$

alors :

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{p} \mu$$

Corollaire 1.1.4 [FOATA & FUCHS (1998)].

Soit $\{X_n, n \in \mathbb{N}^*\}$ une suite de variables aléatoires, **deux à deux non corrélées et identiquement distribuées**, définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans \mathbb{R} et telle que $\mathbb{E}(X_i^2) < \infty$, **alors** :

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{p} \mathbb{E}(X_i) = \mu$$

En fait, on peut avoir le résultat du corollaire ci-dessus avec des conditions restrictives, mais en remplaçant la condition « deux à deux non corrélées » par « indépendantes ».

Proposition 1.1.5 (Théorème de KHINTCHINE) [FOATA & FUCHS (2000)].

Soit $\{X_n, n \in \mathbb{N}^*\}$ une suite de variables aléatoires, **indépendantes et identiquement distribuées**, définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans \mathbb{R} et telle que $\mathbb{E}(|X_i|) < \infty$, **alors** :

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{p} \mathbb{E}(X_i) = \mu$$

III Polynôme de BERNSTEIN et loi faible des grands nombres

On rencontre assez souvent des résultats d'analyse qui se démontrent à l'aide d'arguments probabilistes. En voici un exemple, qui est l'approximation polynomiale de BERNSTEIN, qui se démontre à l'aide de la loi faible des grands nombres :

Théorème 1.1.3 [ATHREYA & LAHIRI (2006)].

Soit f une fonction continue de $[0, 1]$ dans \mathbb{R} . On définit le polynôme :

$$B_n(x) = \sum_{k=0}^n f\left(\frac{k}{n}\right) C_n^k x^k (1-x)^{n-k}, \quad x \in [0, 1]$$

dit polynôme de BERNSTEIN associé à la fonction f . Alors la suite de fonctions $(B_n)_{n \in \mathbb{N}^*}$ converge **uniformément** sur $[0, 1]$ vers la fonction f .

Démonstration.

• Puisque la fonction f est continue sur $[0, 1]$, alors f est bornée sur $[0, 1]$ et atteint ses bornes. Soit

$$M = \sup_{x \in [0, 1]} |f(x)|$$

• Puisque la fonction f est continue sur $[0, 1]$, alors f est uniformément continue sur $[0, 1]$. Donc :

$$\forall \varepsilon > 0; \quad \exists \delta(\varepsilon) > 0 \quad \text{tel que} \quad \forall (x, y) \in [0, 1] \times [0, 1], \quad |y - x| \leq \delta(\varepsilon) \implies |f(x) - f(y)| \leq \varepsilon$$

Soit $\varepsilon > 0$ et $x \in [0, 1]$. Définissons l'ensemble A_ε par :

$$A_\varepsilon = \{z \in [0, 1] \quad \text{tel que} \quad |z - x| \leq \delta(\varepsilon)\}$$

On veut montrer que la suite de fonctions $(B_n)_{n \in \mathbb{N}^*}$ converge uniformément sur $[0, 1]$ vers la fonction f , c'est-à-dire que :

$$\lim_{n \rightarrow \infty} \sup_{x \in [0, 1]} |B_n(x) - f(x)| = 0$$

Pour se faire, introduisons, pour chaque $x \in [0, 1]$, une suite $(X_n)_{n \in \mathbb{N}^*}$ de variables aléatoires indépendantes et identiquement distribuées selon la loi de BERNOULLI de paramètre x , c'est-à-dire :

$$\forall n \in \mathbb{N}^*, \quad \mathbb{P}(X_n = 1) = x \quad \text{et} \quad \mathbb{P}(X_n = 0) = 1 - x$$

Et soit $S_n = \sum_{i=1}^n X_i$ la somme des n premiers termes de la suite $(X_n)_{n \in \mathbb{N}^*}$. Alors S_n est une variable aléatoire qui suit une loi de probabilité binomiale de paramètre (n, x) , c'est-à-dire :

$$\forall k \in \{0, 1, 2, \dots, n\}, \quad \mathbb{P}(S_n = k) = C_n^k x^k (1 - x)^{n-k}$$

On a alors :

$$\mathbb{E} \left[f \left(\frac{S_n}{n} \right) \right] = \sum_{k=0}^n f \left(\frac{k}{n} \right) C_n^k x^k (1 - x)^{n-k} = B_n(x)$$

Donc :

$$|B_n(x) - f(x)| = \left| \mathbb{E} \left[f \left(\frac{S_n}{n} \right) \right] - f(x) \right|$$

Or, puisque $f(x)$ n'est pas aléatoire et l'espérance est linéaire, on a :

$$\left| \mathbb{E} \left[f \left(\frac{S_n}{n} \right) \right] - f(x) \right| = \left| \mathbb{E} \left[f \left(\frac{S_n}{n} \right) \right] - \mathbb{E}[f(x)] \right| = \left| \mathbb{E} \left[f \left(\frac{S_n}{n} \right) - f(x) \right] \right|$$

D'où :

$$|B_n(x) - f(x)| = \left| \mathbb{E} \left[f \left(\frac{S_n}{n} \right) - f(x) \right] \right|$$

Il s'ensuit que :

$$\begin{aligned}
|B_n(x) - f(x)| &\leq \mathbb{E} \left\{ \left| f\left(\frac{S_n}{n}\right) - f(x) \right| \right\} \\
&\leq \mathbb{E} \left\{ \left| \left[f\left(\frac{S_n}{n}\right) - f(x) \right] \cdot \mathbb{1}_{A_\varepsilon}\left(\frac{S_n}{n}\right) + \left[f\left(\frac{S_n}{n}\right) - f(x) \right] \cdot \mathbb{1}_{\bar{A}_\varepsilon}\left(\frac{S_n}{n}\right) \right| \right\} \\
&\leq \mathbb{E} \left\{ \left| f\left(\frac{S_n}{n}\right) - f(x) \right| \cdot \mathbb{1}_{A_\varepsilon}\left(\frac{S_n}{n}\right) + \left| f\left(\frac{S_n}{n}\right) - f(x) \right| \cdot \mathbb{1}_{\bar{A}_\varepsilon}\left(\frac{S_n}{n}\right) \right\} \\
&\leq \mathbb{E} \left\{ \varepsilon \cdot \mathbb{1}_{A_\varepsilon}\left(\frac{S_n}{n}\right) + \left[\left| f\left(\frac{S_n}{n}\right) \right| + |f(x)| \right] \cdot \mathbb{1}_{\bar{A}_\varepsilon}\left(\frac{S_n}{n}\right) \right\} \\
&\leq \mathbb{E} \left\{ \varepsilon \cdot \mathbb{1}_{A_\varepsilon}\left(\frac{S_n}{n}\right) + 2M \cdot \mathbb{1}_{\bar{A}_\varepsilon}\left(\frac{S_n}{n}\right) \right\} \\
&\leq \varepsilon \cdot \mathbb{E} \left\{ \mathbb{1}_{A_\varepsilon}\left(\frac{S_n}{n}\right) \right\} + 2M \cdot \mathbb{E} \left\{ \mathbb{1}_{\bar{A}_\varepsilon}\left(\frac{S_n}{n}\right) \right\} \\
&\leq \varepsilon \cdot \mathbb{P} \left(\left| \frac{S_n}{n} - x \right| \leq \delta(\varepsilon) \right) + 2M \cdot \mathbb{P} \left(\left| \frac{S_n}{n} - x \right| > \delta(\varepsilon) \right) \\
&\leq \varepsilon + 2M \cdot \mathbb{P} \left(\left| \frac{S_n}{n} - x \right| > \delta(\varepsilon) \right) \quad (\text{Car : } \mathbb{P}(\cdot) \leq 1) \\
&\leq \varepsilon + 2M \cdot \frac{V(n^{-1}S_n)}{[\delta(\varepsilon)]^2} \quad (\text{Inégalité de TCHEBYCHEV}) \\
&\leq \varepsilon + 2M \cdot \frac{x(1-x)}{n[\delta(\varepsilon)]^2}
\end{aligned}$$

Donc :

$$\begin{aligned}
\sup_{x \in [0,1]} |B_n(x) - f(x)| &\leq \sup_{x \in [0,1]} \left\{ \varepsilon + 2M \cdot \frac{x(1-x)}{n[\delta(\varepsilon)]^2} \right\} \\
&\leq \varepsilon + 2M \cdot \sup_{x \in [0,1]} \frac{x(1-x)}{n[\delta(\varepsilon)]^2} \\
&\leq \varepsilon + \frac{2M}{4n[\delta(\varepsilon)]^2} \quad \left(\text{Car : } \sup_{x \in [0,1]} x(1-x) = \frac{1}{4} \right)
\end{aligned}$$

D'où :

$$\begin{aligned}
\lim_{n \rightarrow \infty} \sup_{x \in [0,1]} |B_n(x) - f(x)| &\leq \lim_{n \rightarrow \infty} \left\{ \varepsilon + \frac{2M}{4n[\delta(\varepsilon)]^2} \right\} \\
&\leq \varepsilon + \lim_{n \rightarrow \infty} \frac{2M}{4n[\delta(\varepsilon)]^2} \\
&\leq \varepsilon
\end{aligned}$$

Pour conclure : Il suffit de choisir ε suffisamment petit, et la convergence uniforme est établie. ■

1.1.2 Loi forte des grands nombres

I Convergence presque sûre (convergence forte)

Un autre mode de convergence pour les suites de variables aléatoires consiste à *négliger* l'ensemble des ω pour lesquels la suite réelle $(X_n(\omega))_{n \in \mathbb{N}^*}$ ne converge pas vers le réel $X(\omega)$ dès lors que sa probabilité est nulle. Alors, dans ce cas, la probabilité de l'ensemble des ω pour lesquels la suite réelle $(X_n(\omega))_{n \in \mathbb{N}^*}$ converge vers le réel $X(\omega)$ est égale à 1. Comme un événement de probabilité nulle est dit *presque impossible*, et qu'un événement de probabilité 1 est dit *presque certain*, on dira que la convergence obtenue est **presque sûre**.

Définition 1.1.3 [CAMINADE & NICOLAS (2008)].

Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires réelles définies sur un même espace probabilisé (Ω, \mathcal{A}, P) , et X une variable aléatoire réelle définie, elle aussi, sur le même espace.

On dit que la suite $(X_n)_{n \in \mathbb{N}^*}$ converge **presque sûrement** (**fortement**) vers X , et on note $X_n \xrightarrow{p.s.} X$, si et seulement si :

$$P\left(\left\{\omega \in \Omega \text{ tel que } \lim_{n \rightarrow +\infty} X_n(\omega) = X(\omega)\right\}\right) = 1$$

I.1 Problème de l'unicité de la limite

Comme dans la convergence en probabilité, la limite au sens de la convergence presque sûre est presque sûrement unique pour la probabilité P .

Proposition 1.1.6 [OUVRARD (2000)].

Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires réelles définies sur un même espace probabilisé (Ω, \mathcal{A}, P) , et soient X et Y deux variables aléatoires réelles définies, elles aussi, sur le même espace et telles que $X_n \xrightarrow{p.s.} X$ et $X_n \xrightarrow{p.s.} Y$. Alors $P(X \neq Y) = 0$.

I.2 Composition par une fonction continue

Proposition 1.1.7 [OUVRARD (2000)].

Soit f une fonction continue sur \mathbb{R}^d ($d \geq 1$) dans \mathbb{R} . Si $(X_{1,n}), (X_{2,n}), \dots, (X_{d,n})$ ($n \in \mathbb{N}^*$) sont d suites de variables aléatoires réelles qui convergent, presque sûrement, respectivement vers les variables aléatoires réelles X_1, X_2, \dots, X_d , **alors** la suite de variables aléatoires réelles $(f(X_{1,n}, X_{2,n}, \dots, X_{d,n}))_{n \in \mathbb{N}^*}$ converge presque sûrement vers la variable aléatoire réelle $f(X_1, X_2, \dots, X_d)$.

I.3 Lemme de BOREL -CANTELLI

Le lemme de BOREL-CANTELLI est l'un des résultats les plus importants de la théorie des probabilités, il est à la base d'une très utile condition suffisante de la convergence presque sûre.

Avant de présenter le lemme, on présente la définition suivante, qui sera utilisée dans son énoncé :

Définition 1.1.4 [BHATTACHARYA & WAYMIRE (2007)].

Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et $(A_n)_{n \in \mathbb{N}^*}$ une suite d'événements de \mathcal{A} .

1. On appelle **limite supérieure** de la suite $(A_n)_{n \in \mathbb{N}^*}$, l'ensemble A des événements élémentaires défini par :

$$A = \limsup_n A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$$

En d'autres termes :

$$\omega \in A \iff \forall n \in \mathbb{N}^*, \quad \exists k \geq n \text{ tel que } \omega \in A_k.$$

2. On appelle **limite inférieure** de la suite $(A_n)_{n \in \mathbb{N}^*}$, l'ensemble B des événements élémentaires défini par :

$$B = \liminf_n A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k$$

En d'autres termes :

$$\omega \in B \iff \exists n \in \mathbb{N}^* \text{ tel que } \forall k \geq n, \quad \omega \in A_k.$$

Remarque : Dans le cadre de la définition ci-dessus, on peut constater que A est l'ensemble des événements élémentaires ω qui appartiennent à une infinité d'événements A_k , et B est l'ensemble des événements élémentaires ω qui appartiennent à tous les événements A_k à partir d'un certain rang n .

Maintenant, on présente une proposition qui donne une condition nécessaire et suffisante pour qu'une suite de variables aléatoires converge presque sûrement :

Proposition 1.1.8 [BARBE & LEDOUX (2007)].

Soient $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires réelles définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et X une variable aléatoire réelle définie, elle aussi, sur le même espace. On a :

$$X_n \xrightarrow{p.s.} X \iff \forall \varepsilon \in \mathbb{R}_+^*, \quad \mathbb{P} \left\{ \limsup_n \left(|X_n - X| > \varepsilon \right) \right\} = 0$$

Théorème 1.1.4 (Lemme de BOREL-CANTELLI) [BHATTACHARYA & WAYMIRE (2007)].

Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et $(A_n)_{n \in \mathbb{N}^*}$ une suite d'événements de \mathcal{A} . On a :

1.

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty \implies \mathbb{P} \left(\limsup_n A_n \right) = 0$$

2. **Si** les événements A_n sont indépendants, **alors** :

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty \implies \mathbb{P} \left(\limsup_n A_n \right) = 1$$

Remarque : Dans le cadre du théorème ci-dessus et en passant au complémentaire, on trouve :

$$\left(\limsup_n A_n\right)^c = \left[\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k\right]^c = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k^c = \liminf_n A_n^c$$

Or :

$$P\left(\limsup_n A_n\right) = 0 \iff P\left[\left(\limsup_n A_n\right)^c\right] = 1$$

Donc le résultat du lemme de BOREL -CANTELLI peut s'interpréter comme suit : « presque sûrement, aucun des événements A_n ne se réalise au-delà d'un certain rang ».

I.4 Inégalité de KOLMOGOROV

L'une des inégalités pertinentes de la théorie des probabilités est celle de KOLMOGOROV, qui est utilisée souvent dans la démonstration de la loi forte des grands nombres.

Théorème 1.1.5 (Inégalité de KOLMOGOROV) [BHATTACHARYA & WAYMIRE (2007)].

Soient $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires réelles **indépendantes** et admettant, toutes, un moment de second ordre et $S_n = \sum_{i=1}^n X_i$. Alors :

$$\forall \varepsilon > 0, \quad P\left(\max_{1 \leq k \leq n} |S_k - E(S_k)| \geq \varepsilon\right) \leq \frac{V(S_n)}{\varepsilon^2}$$

Remarque : On remarque que l'inégalité de KOLMOGOROV ressemble formellement à celle de BIENAYMÉ-TCHEBYCHEV. Elle est cependant beaucoup plus puissante, puisqu'elle permet de contrôler, en probabilité, les déviations de toute la suite finie $(S_k - E(S_k))_{1 \leq k \leq n}$ au lieu de seulement son dernier terme $(S_n - E(S_n))$ pour l'inégalité BIENAYMÉ-TCHEBYCHEV. L'hypothèse est aussi plus restrictive car elle exige l'indépendance au lieu de la non corrélation deux à deux.

I.5 Critère de la convergence presque sûre

Comme on l'a mentionné avant, le lemme de BOREL -CANTELLI nous donne une condition suffisante pour qu'une suite de variables aléatoires converge presque sûrement. La voici dans la proposition suivante :

Proposition 1.1.9 [FOATA & FUCHS (1998)].

Soient $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires réelles définies sur un même espace probabilisé (Ω, \mathcal{A}, P) , et X une variable aléatoire réelle définie, elle aussi, sur le même espace. On a :

1. **Si :**

$$\forall \varepsilon > 0, \quad \sum_{n=1}^{\infty} P\{|X_n - X| > \varepsilon\} < \infty$$

alors :

$$X_n \xrightarrow{p.s.} X$$

2. *Si* :

$$\exists p > 0 \quad \text{tel que} \quad \sum_{n=1}^{\infty} \mathbb{E}\{|X_n - X|^p\} < \infty$$

alors :

$$X_n \xrightarrow{p.s.} X$$

Remarque : Dans le cadre de la proposition ci-dessus, le premier résultat est obtenu grâce au lemme de BOREL-CANTELLI, et il nous donne une condition suffisante pour que la suite $(X_n)_{n \in \mathbb{N}^*}$ converge presque sûrement vers X . Elle est également nécessaire si, de plus, les variables aléatoires X_i sont indépendantes. Tandis que le deuxième résultat est obtenu grâce au premier et à la 2^{ème} inégalité de MARKOV appliquée à la variable aléatoire $(X_n - X)$.

II Loi forte des grands nombres

Soient $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans \mathbb{R} , et $S_n = \sum_{i=1}^n X_i$ la somme des n premiers termes de la suite $(X_n)_{n \in \mathbb{N}^*}$.

Définition 1.1.5 [FOATA & FUCHS (1998)].

On dit que la suite $(X_n)_{n \in \mathbb{N}^*}$ satisfait **la loi forte des grands nombres** si la suite de variables aléatoires réelles $\left\{n^{-1} \sum_{i=1}^n X_i, n \in \mathbb{N}^*\right\}$ converge presque sûrement vers une constante.

Remarque : Comme pour la loi faible des grands nombres, il existe plusieurs versions de la loi forte des grands nombres, où chacune donne la constante vers laquelle la suite $\left\{n^{-1} \sum_{i=1}^n X_i, n \in \mathbb{N}^*\right\}$ converge presque sûrement et donne les conditions suffisantes (même nécessaires et suffisantes) sur les variables aléatoires X_i pour que cette convergence aura lieu.

Sachant que toutes les versions exigent, au moins, la non corrélation, on s'intéresse dans ce mémoire aux versions où Andreï Nikolaïevitch KOLMOGOROV à une grande part de travail.

Théorème 1.1.6 (Loi forte des grands nombres de KOLMOGOROV) [OUVRARD (2000)].

Soit $\{X_n, n \in \mathbb{N}^*\}$ une suite de variables aléatoires **indépendantes** définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans \mathbb{R} et telle que :

$$\forall i \in \mathbb{N}^*, \quad \mathbb{E}(X_i^2) < \infty$$

Si :

$$\lim_{n \rightarrow \infty} \mathbb{E}(X_n) = \mu < \infty \quad \text{et} \quad \lim_{n \rightarrow \infty} \sum_{i=1}^n \frac{\mathbb{V}(X_i)}{i^2} < \infty$$

alors :

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{p.s.} \mu$$

Corollaire 1.1.5 [ASH & DOLÉANS-DADE (2000)].

Soit $\{X_n, n \in \mathbb{N}^*\}$ une suite de variables aléatoires **indépendantes** et **identiquement distribuées** définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans \mathbb{R} et telle que $E(X_i^2) < \infty$, **alors** :

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{p.s.} E(X_i)$$

En fait, on peut avoir la convergence presque sûre avec une condition nécessaire et suffisante et moins forte que celle du corollaire ci-dessus :

Théorème 1.1.7 (KOLMOGOROV-KHINTCHINE) [OUVRARD (2000)].

Soit $\{X_n, n \in \mathbb{N}^*\}$ une suite de variables aléatoires **indépendantes** et **identiquement distribuées** définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans \mathbb{R} , alors on a l'équivalence suivante :

$$\left(\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{p.s.} \mu < \infty \right) \iff \left(E(|X_i|) < \infty \text{ et } E(X_i) = \mu \right)$$

III Application de la loi forte des grands nombres

La loi forte des grands nombres a une application très importantes et fondamentale en statistiques : elle permet de justifier la convergence de la suite des fonctions de répartition empiriques.

Définition 1.1.6 [OUVRARD (2000)].

Soit (X_1, X_2, \dots, X_n) un échantillon de taille n d'une variable aléatoire réelle X définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ de fonction de répartition F . On appelle **fonction de répartition empirique** associée à la variable aléatoire X et basée sur l'échantillon (X_1, X_2, \dots, X_n) la fonction F_n définie sur \mathbb{R} dans $[0, 1]$ par :

$$\forall x \in \mathbb{R}, \quad F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{X_i \leq x\}}$$

Remarque : Dans le cadre de la définition ci-dessus, si on fixe x dans \mathbb{R} , alors $F_n(x)$ est une variable aléatoire. Et si x varie, alors $F_n(x)$ est une fonction de répartition.

En utilisant les notations de la définition ci-dessus et grâce à la loi forte des grands nombres, on a le résultat suivant :

Proposition 1.1.10

Pour tout $x \in \mathbb{R}$, la suite de variables aléatoires $\{F_n(x), n \in \mathbb{N}^*\}$ converge presque sûrement vers la variable aléatoire $F(x)$, c'est-à-dire :

$$\forall x \in \mathbb{R}, \quad F_n(x) \xrightarrow{p.s.} F(x)$$

Preuve.

Soit $x \in \mathbb{R}$ et posons $Y_i = \mathbb{1}_{\{X_i \leq x\}}$ ($i = 1, 2, \dots, n$), alors :

$$Y_i = \mathbb{1}_{\{X_i \leq x\}} = \begin{cases} 1 & \text{si } X_i \leq x \\ 0 & \text{si } X_i > x \end{cases}$$

est une variable aléatoire réelle qui suit une loi de probabilité de BERNOULLI de paramètre p tel que :

$$p = P(X_i \leq x) = P(X \leq x) = F(x), \quad \forall i = 1, 2, \dots, n$$

Et, puisque les variables aléatoires X_i ($i = 1, 2, \dots, n$) sont indépendantes, alors les variables aléatoires Y_i ($i = 1, 2, \dots, n$) sont indépendantes et identiquement distribuées selon une loi de probabilité de BERNOULLI de paramètre $F(x)$.

Alors, d'après la loi forte des grands nombres :

$$\frac{1}{n} \sum_{i=1}^n Y_i = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{X_i \leq x\}} \xrightarrow{p.s.} F(x)$$

D'où le résultat. ■

Ainsi une loi de probabilité inconnue peut être reconstituée, approximativement, à partir de l'observation d'un échantillon de grande taille. En fait on peut obtenir, mieux que la convergence simple presque sûre, la convergence uniforme presque sûre de la suite $\{F_n(x), n \in \mathbb{N}^*\}$ vers $F(x)$:

Théorème 1.1.8 (Théorème de GLIVENKO-CANTELLI) [RESNICK (1998)].

Soit $\{X_n, n \in \mathbb{N}^*\}$ une suite de variables aléatoires définies sur le même espace probabilisé (Ω, \mathcal{A}, P) à valeurs dans \mathbb{R} , **indépendantes et identiquement distribuées** selon une loi de probabilité de fonction de répartition F , et soit $\{F_n(x), n \in \mathbb{N}^*\}$ ($x \in \mathbb{R}$) la suite des fonctions de répartition empiriques associées. On a :

$$\sup_{x \in \mathbb{R}} |F_n(x) - F(x)| \xrightarrow{p.s.} 0$$

Remarque : Le théorème ci-dessus est très utile en statistiques, c'est pour ça qu'on le trouve des fois sous le nom « **Théorème fondamental de la statistique** ».

1.2 Théorème de la limite centrale

Comme son nom l'indique, le théorème de la limite centrale occupe la place centrale dans la théorie des probabilités et les statistiques, certains probabilistes le décrivent comme « **le trône** » de ces disciplines. Le nom du théorème fait référence au document scientifique, écrit par George POLYÀ en 1920, « Sur le théorème central du calcul des probabilités et le problème des moments ».

Il existe deux versions générales du théorème de la limite centrale, « Théorème de la limite centrale **classique** » et « Théorème de la limite centrale **généralisé** » qui comportent quelques différences qu'on va voir dans la suite.

1.2.1 Convergence en loi (Convergence en distribution)

Considérons une suite $(X_n)_{n \in \mathbb{N}^*}$ de variables aléatoires à valeurs dans \mathbb{R} , et X une variable aléatoire, elle aussi, à valeurs dans \mathbb{R} . Une autre façon d'aborder le problème de la convergence de la suite de variables aléatoires $(X_n)_{n \in \mathbb{N}^*}$ vers la variable aléatoire X est de considérer que les variables aléatoires de la suite $(X_n)_{n \in \mathbb{N}^*}$ sont définies par leurs lois de probabilité, et d'envisager la convergence de la suite de ces lois de probabilité vers la loi de probabilité de la variable aléatoire X . On dira, alors, que la convergence obtenue est une **convergence en loi**.

Définition 1.2.1 [BILLINGSLEY (1999)].

Soient $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires à valeurs dans \mathbb{R} et X une variable aléatoire, elle aussi, à valeurs dans \mathbb{R} .

On dit que la suite $(X_n)_{n \in \mathbb{N}^*}$ converge en loi (*en distribution*) vers X , et on note $X_n \xrightarrow{d} X$, si et seulement si :

$$\forall f : \mathbb{R} \longrightarrow \mathbb{R} \text{ continue et bornée, } \quad \lim_{n \rightarrow \infty} \mathbb{E}[f(X_n)] = \mathbb{E}[f(X)].$$

Remarque : Les variables aléatoires de la suite $(X_n)_{n \in \mathbb{N}^*}$ ne sont pas nécessairement définies sur le même espace probabilisé.

I Convergence en loi et fonctions de répartition

Dans la proposition suivante, qui peut être considérée comme une autre définition de la convergence en loi, on va voir la relation entre la convergence en loi, les fonctions de répartition des variables aléatoires de la suite $(X_n)_{n \in \mathbb{N}^*}$ et la fonction de répartition de la variable aléatoire X .

Proposition 1.2.1 [OUVRARD (2000)].

Soient $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires à valeurs dans \mathbb{R} et X une variable aléatoire, elle aussi, à valeurs dans \mathbb{R} . On note F_n la fonction de répartition de la variable aléatoire X_n et F la fonction de répartition de X . On a :

$$X_n \xrightarrow{d} X \iff \forall x \in \mathcal{C}(F) = \{x \in \mathbb{R}, F \text{ est continue en } x\}, \quad \lim_{n \rightarrow \infty} F_n(x) = F(x)$$

Ou, d'une autre manière :

$$X_n \xrightarrow{d} X \iff \forall x \in \mathbb{R} \text{ tel que } \mathbb{P}(X = x) = 0, \quad \lim_{n \rightarrow \infty} \mathbb{P}(X_n \leq x) = \mathbb{P}(X \leq x)$$

II Problème de l'unicité de la limite

La convergence en loi de la suite de variables aléatoires $(X_n)_{n \in \mathbb{N}^*}$ vers la variable aléatoire X n'est pas une vraie convergence de suite de variables aléatoires en tant qu'applications, c'est seulement la convergence de la suite de lois de probabilité des X_n vers la loi de probabilité de X . On peut donc dire que la suite de variables aléatoires $(X_n)_{n \in \mathbb{N}^*}$ converge aussi en loi vers n'importe quelle variable aléatoire Y de même loi de probabilité que la variable aléatoire X . Par conséquent, **il n'y a pas unicité de la limite** pour la convergence en loi.

III Composition par une fonction continue

Proposition 1.2.2 [BILLINGSLEY (1995)].

Soit f une fonction continue sur \mathbb{R} dans \mathbb{R} . Si $(X_n)_{n \in \mathbb{N}^*}$ est une suite de variables aléatoires réelles qui converge en loi vers la variable aléatoire réelle X , alors la suite de variables aléatoires réelles $\{f(X_n)\}_{n \in \mathbb{N}^*}$ converge en loi vers la variable aléatoire réelle $f(X)$.

VI Convergence en loi et fonctions caractéristiques

Dans le théorème suivant, où sa première partie peut être considérée, aussi, comme une autre définition de la convergence en loi, on va voir la relation entre la convergence en loi, les fonctions caractéristiques des variables aléatoires de la suite $(X_n)_{n \in \mathbb{N}^*}$ et la fonction caractéristique de la variable aléatoire X .

Théorème 1.2.1 (Théorème de Paul LÉVY) [BILLINGSLEY (1995)].

Soient $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires à valeurs dans \mathbb{R} et X une variable aléatoire, elle aussi, à valeurs dans \mathbb{R} . On note φ_n la fonction caractéristique de la variable aléatoire X_n et φ la fonction caractéristique de X . On a :

1.

$$X_n \xrightarrow{d} X \iff \forall t \in \mathbb{R}, \quad \lim_{n \rightarrow \infty} \varphi_n(t) = \varphi(t)$$

2. Si, pour tout $t \in \mathbb{R}$, $\lim_{n \rightarrow \infty} \varphi_n(t) = \phi(t)$ (ϕ est une fonction continue en $t = 0$), alors ϕ est une fonction caractéristique d'une loi de probabilité, et cette loi de probabilité est la limite de la suite $(X_n)_{n \in \mathbb{N}^*}$ au sens de la convergence en loi.

Remarque : La deuxième partie du théorème ci-dessus fournit un outil puissant pour établir la convergence en loi d'une suite de variables aléatoires. Elle est principalement utilisée pour démontrer des versions du théorème de la limite centrale.

Référence : Pour voir les résultats des différents modes de convergence d'une suite de variables aléatoires avec leurs démonstrations bien détaillées et avec des exemples illustratifs, et voir les relations qui lient ces différents modes de convergence, le lecteur pourra se référer aux ouvrages suivants : [ATHREYA & LAHIRI (2006)], [BILLINGSLEY (1995)], [BILLINGSLEY (1999)], [OUVRARD (2000)], ou, encore, [RESNICK (1998)].

1.2.2 Théorème de la limite centrale classique

Le théorème de la limite centrale classique montre que sous certaines conditions plus ou moins générales, la suite de sommes finies, normalisées, de variables aléatoires indépendantes et non dégénérées converge en loi vers une loi de probabilité normale. La première démonstration de ce théorème, faite dans le cas où les variables aléatoires à sommer suivent, toutes, une loi de probabilité de BERNOULLI de paramètre p , est due à Pierre Simon DE LAPLACE, en 1809. Mais le cas particulier où $p = 0.5$ était connu depuis les travaux d'Abraham DE MOIVRE, en 1733. C'est pourquoi on le rencontre des fois sous le nom de « **Théorème de DE MOIVRE-LAPLACE** ».

Le théorème de la limite centrale classique admet plusieurs généralisations qui donnent la convergence en loi de la suite de sommes finies, normalisées, de variables aléatoires vers une loi de probabilité normale sous des hypothèses beaucoup plus faibles. Ces généralisations n'exigent pas que les variables aléatoires à sommer suivent des lois de probabilité identiques, mais font appel à des conditions telles la condition de **LINDEBERG** et la condition de **LYAPOUNOV**.

Ainsi, ce théorème et ses généralisations offrent une explication de l'omniprésence de la loi de probabilité normale dans la nature, car de nombreux phénomènes sont dus à l'addition d'un grand nombre de petites perturbations aléatoires. C'est ce qui rend son champ d'applications extrêmement vaste et établit le rôle universel de la loi de probabilité normale, la fameuse « courbe en cloche ».

I Énoncé du théorème

Dans le théorème qui suit, on va donner la version élémentaire et la plus utilisée du théorème de la limite centrale classique :

Théorème 1.2.2 (Théorème de la limite centrale classique).

Soient $\{X_n, n \in \mathbb{N}^*\}$ une suite de variables aléatoires **non dégénérées, indépendantes et identiquement distribuées** définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans \mathbb{R} et $S_n = \sum_{i=1}^n X_i$. **Supposons que :**

$$\forall i \in \mathbb{N}^*, \quad \mathbb{E}(X_i^2) < \infty$$

Si on pose :

$$\mathbb{E}(X_i) = \mu \quad \text{et} \quad \mathbb{V}(X_i) = \sigma^2, \quad \text{avec} \quad (\mu, \sigma) \in \mathbb{R} \times \mathbb{R}_+^*$$

alors :

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{d} \mathcal{N}(0, 1)$$

Notation : On note $\mathcal{N}(m, v^2)$ la variable aléatoire qui suit une loi de probabilité normale d'espérance m et de variance v^2 .

Comme on l'a déjà dit, le théorème de la limite centrale classique admet des généralisations avec des hypothèses plus faibles que celles données dans le théorème ci-dessus. En voici deux de ces généralisations données dans les deux théorèmes suivants :

Théorème 1.2.3 (Théorème de la limite centrale classique de LINDEBERG) [SHIRYAEV (1984)].

Soient $\{X_n, n \in \mathbb{N}^*\}$ une suite de variables aléatoires **non dégénérées et indépendantes** à valeurs dans \mathbb{R} et $S_n = \sum_{i=1}^n X_i$. **Supposons que :**

$$\forall i \in \mathbb{N}^*, \quad \mathbb{E}(X_i^2) < \infty$$

Si :

$$\forall \varepsilon > 0, \quad \lim_{n \rightarrow \infty} \frac{1}{\mathbb{V}(S_n)} \sum_{i=1}^n \mathbb{E} \left\{ [X_i - \mathbb{E}(X_i)]^2 \cdot \mathbb{1}_{[|X_i - \mathbb{E}(X_i)| \geq \varepsilon \sqrt{\mathbb{V}(S_n)}]} \right\} = 0 \quad (1.1)$$

alors :

$$\frac{S_n - \mathbb{E}(S_n)}{\sqrt{V(S_n)}} \xrightarrow{d} \mathcal{N}(0, 1)$$

Remarque : On appelle (1.1) « la condition de LINDEBERG ».

Théorème 1.2.4 (Théorème de la limite centrale classique de LYAPOUNOV) [SHIRYAEV (1984)].

Soient $\{X_n, n \in \mathbb{N}^*\}$ une suite de variables aléatoires *non dégénérées* et *indépendantes* à valeurs dans \mathbb{R} et $S_n = \sum_{i=1}^n X_i$. **Supposons que :**

$$\exists \delta > 0 \quad \text{tel que} \quad \forall i \in \mathbb{N}^*, \quad \mathbb{E}(|X_i|^{2+\delta}) < \infty$$

Si :

$$\lim_{n \rightarrow \infty} \frac{1}{[V(S_n)]^{\frac{2+\delta}{2}}} \sum_{i=1}^n \mathbb{E}\{|X_i - \mathbb{E}(X_i)|^{2+\delta}\} = 0 \quad (1.2)$$

alors :

$$\frac{S_n - \mathbb{E}(S_n)}{\sqrt{V(S_n)}} \xrightarrow{d} \mathcal{N}(0, 1)$$

Remarque : On appelle (1.2) « la condition de LYAPOUNOV ».

II Méthode de Monte-Carlo et théorème de la limite centrale classique

On désigne par « méthode de Monte-Carlo » toute méthode numérique utilisant le tirage de nombres aléatoires. Elles sont très utilisées dans de nombreux domaines, en particulier en physique nucléaire, en physique statistique et en statistique. La naissance de la méthode de Monte-Carlo est liée à l'apparition des premiers ordinateurs et à leur utilisation dans le cadre des projets secrets du département de la défense des États Unis (Projet « MANHATTAN ») durant la deuxième guerre mondiale en vue de la conception des premières bombes atomiques.

Pour donner une idée générale et illustrative de la méthode de Monte-Carlo et de l'intérêt du théorème de la limite centrale classique dans cette méthode, soit à calculer l'intégrale suivante :

$$I = \int_0^1 f(x) dx$$

où f est une fonction réelle définie et intégrable sur $[0, 1]$ mais dont l'intégrale I ne se calcule pas, ou difficile à calculer.

On sait qu'il existe de nombreuses méthodes d'approximation numérique de l'intégrale I par des formules du type $\sum_{i=1}^n y_i f(x_i)$, avec les y_i sont des nombres positifs de somme 1 et les x_i sont des points de l'intervalle $[0, 1]$, telles la méthode des trapèzes, de GAUSS, de SIMPSON, etc. La méthode de Monte-Carlo est du même type : On choisit $y_i = 1/n$ et l'on tire les x_i « au hasard », mais pas n'importe comment, il faut tirer les points x_i selon la loi de probabilité uniforme $\mathcal{U}([0, 1])$.

En effet, si on pose X la variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et qui suit

la loi de probabilité uniforme $\mathcal{U}([0, 1])$, alors $f(X)$ est une variable aléatoire d'espérance finie, et l'intégrale I n'est rien d'autre que son espérance, c'est-à-dire :

$$I = \int_0^1 f(x) dx = \mathbb{E}[f(X)]$$

Or, si (X_1, X_2, \dots, X_n) est un échantillon de taille $n \in \mathbb{N}^*$ de la variable aléatoire X , alors $\{f(X_1), f(X_2), \dots, f(X_n)\}$ est un échantillon de taille n de la variable aléatoire $f(X)$. Et, d'après la loi forte des grands nombres :

$$\mathbb{P} \left\{ \omega \in \Omega \quad \text{tel que} \quad \lim_{n \rightarrow \infty} \frac{f[X_1(\omega)] + f[X_2(\omega)] + \dots + f[X_n(\omega)]}{n} = \mathbb{E}[f(X)] \right\} = 1$$

Donc, pour donner une valeur approximative de l'intégrale I , on simule n réalisations d'une variable aléatoire X (x_1, x_2, \dots, x_n) selon la loi de probabilité uniforme $\mathcal{U}([0, 1])$ et on aura :

$$I = \int_0^1 f(x) dx \simeq \frac{1}{n} \sum_{i=1}^n f(x_i)$$

Pour avoir une idée de l'intérêt de la méthode Monte-Carlo dans le calcul d'intégrales, il faut être en mesure d'évaluer l'erreur ϵ_n donnée par :

$$|\epsilon_n| = \left| \frac{1}{n} \sum_{i=1}^n f(x_i) - I \right| = \left| \frac{1}{n} \sum_{i=1}^n f(x_i) - \mathbb{E}[f(X)] \right|$$

Pour cela, soit ϵ_n la variable aléatoire définie par :

$$\epsilon_n = \frac{1}{n} \sum_{i=1}^n f(X_i) - \mathbb{E}[f(X)] = \frac{1}{n} \sum_{i=1}^n f(X_i) - I$$

Si on suppose que $f(X)$ est une variable aléatoire ayant une variance finie σ^2 ($\sigma > 0$), alors, d'après le théorème de la limite centrale classique :

$$\frac{\sqrt{n}}{\sigma} \epsilon_n = \frac{\sum_{i=1}^n f(X_i) - n\mathbb{E}[f(X)]}{\sigma\sqrt{n}} \xrightarrow{d} \mathcal{N}(0, 1)$$

Ce qui veut dire que :

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{\sqrt{n}}{\sigma} |\epsilon_n| \leq 1.96 \right) = \lim_{n \rightarrow \infty} \mathbb{P} \left(|\epsilon_n| \leq 1.96 \frac{\sigma}{\sqrt{n}} \right) = \lim_{n \rightarrow \infty} \mathbb{P} \left(\left| \frac{1}{n} \sum_{i=1}^n f(X_i) - I \right| \leq 1.96 \frac{\sigma}{\sqrt{n}} \right) = .95$$

Il en résulte que, lorsque n est grand, avec probabilité 0.95, l'intervalle :

$$\left[\frac{1}{n} \sum_{i=1}^n f(x_i) - 1.96 \frac{\sigma}{\sqrt{n}}, \frac{1}{n} \sum_{i=1}^n f(x_i) + 1.96 \frac{\sigma}{\sqrt{n}} \right]$$

contient la vraie valeur de l'intégrale I . C'est ce qu'on appelle un intervalle de confiance au niveau 95% pour l'intégrale I .

1.2.3 Théorème de la limite centrale généralisé

En statistique, une des lois de probabilité les plus utilisées est la loi de probabilité normale. Ses différentes propriétés, comme la stabilité ou le fait que deux paramètres, seulement, (moyenne et variance) suffisent à la caractériser ou, encore, le théorème de la limite centrale classique font que cette loi s'adapte bien dans de nombreuses situations.

Mais, lorsque nous faisons du traitement de données, il nous arrive fréquemment d'utiliser la variance empirique. Comme le nombre d'observations est fini, cette variance est évidemment finie. Il semble alors logique d'employer des méthodes, de proposer des modèles, de faire des estimations utilisant une variance finie. Dans certains cas cependant, la variance empirique est très grande et la question se pose alors de savoir si la variance théorique est finie ou infinie.

Dans le cadre des lois de probabilité à variance infinie sont apparues les lois de probabilité **stables**, dont la variance peut être infinie et qui donne une généralisation au théorème de la limite centrale. C'est ce qui fait que ces lois de probabilité sont utilisées dans de nombreux domaines tels que la télécommunication, la finance, etc.

I Loi de probabilité indéfiniment divisible

Avant de définir les lois de probabilité stables, nous allons introduire une famille de lois de probabilité plus générale, les lois de probabilité **indéfiniment divisibles**. C'est à partir de ces lois que sera précisée la forme de la fonction caractéristique des lois de probabilité stables. L'idée principale de telles lois vient du problème suivant : comment peut-on déterminer toutes les lois de probabilité qui s'expriment comme limites (en loi) d'une suite de sommes de n ($n \in \mathbb{N}^*$) variables aléatoires indépendantes et identiquement distribuées à valeurs dans \mathbb{R} ?

Définition 1.2.2 [SHIRYAEV (1984)].

Une variable aléatoire réelle X ou la loi de probabilité de X est dite **indéfiniment divisible** si :

$$\forall n \in \mathbb{N}^*, \quad \exists X_1, X_2, \dots, X_n \text{ telles que } X \stackrel{d}{=} X_1 + X_2 + \dots + X_n$$

avec : X_1, X_2, \dots, X_n des variables aléatoires **indépendantes et identiquement distribuées** à valeurs dans \mathbb{R} .

Une autre définition des lois de probabilité indéfiniment divisibles, très utile en application et qui se déduit, facilement, de la définition ci-dessus, se fait à l'aide des fonctions caractéristiques des lois de probabilité, est la suivante :

Définition 1.2.3 [SHIRYAEV (1984)].

Soient X une variable aléatoire à valeurs dans \mathbb{R} et $\varphi : \mathbb{R} \rightarrow \mathbb{C}$ la fonction caractéristique de sa loi de probabilité. On a :

$$X \text{ est indéfiniment divisible} \iff \forall n \in \mathbb{N}^*, \quad \exists \phi_n : \mathbb{R} \rightarrow \mathbb{C} \text{ tel que } \forall t \in \mathbb{R}, \quad \varphi(t) = [\phi_n(t)]^n$$

avec : ϕ_n est une fonction caractéristique d'une certaine loi de probabilité.

Remarque : Dans le cadre de la définition ci-dessus, on dit que φ est une fonction caractéristique indéfiniment divisible. Donc, une fonction caractéristique est indéfiniment divisible si elle peut s'exprimer comme puissance $n^{\text{ème}}$ d'une autre fonction caractéristique.

La classe des variables aléatoires indéfiniment divisible permet de résoudre le problème posé au début de cette section. En effet, on a le théorème suivant :

Théorème 1.2.5 [SHIRYAEV (1984)].

La variable aléatoire réelle X est indéfiniment divisible si et seulement si elle est la limite (en loi) d'une suite de variables aléatoires réelles $(S_n)_{n \in \mathbb{N}^}$, où S_n est la somme de n variables aléatoires réelles indépendantes et identiquement distribuées.*

I.1 Propriétés des lois de probabilité indéfiniment divisibles

On récapitule quelques propriétés relatives aux lois de probabilité indéfiniment divisibles dans le théorème qui suit :

Théorème 1.2.6 [GNEDENKO & KOLMOGOROV (1954)].

1. **Si** $\varphi : \mathbb{R} \rightarrow \mathbb{C}$ est la fonction caractéristique d'une loi de probabilité indéfiniment divisible, **alors** :

$$\forall t \in \mathbb{R}, \quad \varphi(t) \neq 0$$

2. **Si** $(X_n)_{n \in \mathbb{N}^*}$ est une suite de variables aléatoires réelles **indéfiniment divisibles** telle que :

$$X_n \xrightarrow{d} X$$

alors : X est une variable aléatoire réelle indéfiniment divisible.

3. **Si** $X_1, X_2, X_3, \dots, X_n$ ($n \in \mathbb{N}^*$) sont n variables aléatoires réelles **indépendantes** et **indéfiniment divisibles**, **alors** $\sum_{i=1}^n X_i$ est une variable aléatoire réelle indéfiniment divisible.

I.2 Exemples

Beaucoup de lois de probabilité connues sont indéfiniment divisibles. Soient X une variable aléatoire à valeurs dans \mathbb{R} et $n \in \mathbb{N}^*$.

Exemple 1.2.1

Si X est une variable aléatoire qui suit une loi de probabilité normale $\mathcal{N}(\mu, \sigma^2)$ ($\mu \in \mathbb{R}$ et $\sigma > 0$), alors sa fonction caractéristique φ s'écrit comme suit :

$$\forall t \in \mathbb{R}, \quad \varphi(t) = \exp \left\{ i\mu t - \frac{t^2}{2} \sigma^2 \right\} = \left[\exp \left\{ i \frac{\mu}{n} t - \frac{t^2}{2} \frac{\sigma^2}{n} \right\} \right]^n, \quad n \in \mathbb{N}^*$$

Or : la fonction φ_n ($n \in \mathbb{N}^*$) définie par :

$$\forall t \in \mathbb{R}, \quad \varphi_n(t) = \exp \left\{ i \frac{\mu}{n} t - \frac{t^2}{2} \frac{\sigma^2}{n} \right\}$$

est la fonction caractéristique de la loi de probabilité normale $\mathcal{N}\left(\frac{\mu}{n}, \frac{\sigma^2}{n}\right)$.

Donc : la loi de probabilité normale est une loi indéfiniment divisible.

Exemple 1.2.2

Si X suit une loi de probabilité de POISSON $\mathcal{P}(\lambda)$ ($\lambda > 0$), alors sa fonction caractéristique φ s'écrit comme suit :

$$\forall t \in \mathbb{R}, \quad \varphi(t) = \exp \{ \lambda(e^{it} - 1) \} = \left[\exp \left\{ \frac{\lambda}{n}(e^{it} - 1) \right\} \right]^n, \quad n \in \mathbb{N}^*$$

Or : la fonction φ_n ($n \in \mathbb{N}^*$) définie par :

$$\forall t \in \mathbb{R}, \quad \varphi_n(t) = \exp \left\{ \frac{\lambda}{n}(e^{it} - 1) \right\}$$

est la fonction caractéristique de la loi de probabilité de POISSON $\mathcal{P}\left(\frac{\lambda}{n}\right)$.

Donc : la loi de probabilité de POISSON est une loi indéfiniment divisible.

Exemple 1.2.3

Si X est une variable aléatoire qui suit une loi de probabilité de CAUCHY $\mathcal{C}(a, h)$ ($a \in \mathbb{R}$ et $h > 0$), alors sa fonction caractéristique φ s'écrit comme suit :

$$\forall t \in \mathbb{R}, \quad \varphi(t) = \exp(iat - h|t|) = \left[\exp \left(i\frac{a}{n}t - \frac{h}{n}|t| \right) \right]^n, \quad n \in \mathbb{N}^*$$

Or : la fonction φ_n ($n \in \mathbb{N}^*$) définie par :

$$\forall t \in \mathbb{R}, \quad \varphi_n(t) = \exp \left(i\frac{a}{n}t - \frac{h}{n}|t| \right)$$

est la fonction caractéristique de la loi de probabilité de CAUCHY $\mathcal{C}\left(\frac{a}{n}, \frac{h}{n}\right)$.

Donc : la loi de probabilité de CAUCHY est une loi indéfiniment divisible.

I.3 Représentation spectrale d'une loi de probabilité indéfiniment divisible

Il existe plusieurs représentations de la fonction caractéristique d'une loi de probabilité indéfiniment divisible. Parmi ces représentations, il y a celle de KOLMOGOROV, celle de LÉVY et celle de LÉVY-KHINTCHINE. Le théorème suivant donne la représentation de LÉVY-KHINTCHINE :

Théorème 1.2.7 (LÉVY-KHINTCHINE) [BREIMAN (1992)].

Une variable aléatoire X à valeurs dans \mathbb{R} est indéfiniment divisible **si et seulement si** la fonction caractéristique de sa loi de probabilité, qu'on note φ , est donnée comme suit :

$$\forall t \in \mathbb{R}, \quad \varphi(t) = \exp \left\{ iat - \frac{t^2}{2}b^2 + \int_{-\infty}^{+\infty} \left(e^{itx} - 1 - \frac{itx}{1+x^2} \right) \frac{1+x^2}{x^2} d\nu(x) \right\} \quad (1.3)$$

où : $a \in \mathbb{R}$, $b^2 \in \mathbb{R}_+$ et ν est une mesure finie définie sur l'espace mesurable $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ telle que $\nu(\{0\}) = 0$.

Remarque : Dans le cadre du théorème ci-dessus, on appelle ν la **mesure spectrale** de LÉVY-KHINTCHINE et (1.3) la **représentation spectrale** de LÉVY-KHINTCHINE d'une loi de probabilité indéfiniment divisible.

II Loi de probabilité stable

Pour se rapprocher du théorème de la limite centrale classique, on va définir une famille de lois de probabilité dont les propriétés seront intéressantes, « **les lois de probabilité stables** ». Ces lois, qu'on appelle aussi alpha-stables ou PARÉTO-stables ou, encore, LÉVY-stables, ont été introduites par Paul LÉVY, dans les années 20, dans son livre « Calcul des probabilités », durant ses recherches sur le comportement asymptotique de la suite des sommes, normalisées, de variables aléatoires indépendantes et identiquement distribuées.

Définition 1.2.4 [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

Une variable aléatoire réelle X , ou la loi de probabilité de X , est dite **stable** si :

$$\forall a \in \mathbb{R}_+ \quad \text{et} \quad \forall b \in \mathbb{R}_+, \quad \exists c \in \mathbb{R}_+^* \quad \text{et} \quad \exists d \in \mathbb{R} \quad \text{telles que} \quad aX_1 + bX_2 \stackrel{d}{=} cX + d$$

où : X_1 et X_2 sont deux variables aléatoires réelles **indépendantes** et de même loi de probabilité que celle de X .

Une autre définition, équivalente à la précédente, de la loi de probabilité stable qu'on rencontre beaucoup plus dans la littérature est la suivante :

Définition 1.2.5 [SHIRYAEV (1984)].

Une variable aléatoire réelle X , ou la loi de probabilité de X , est dite **stable** si :

$$\forall n \in \mathbb{N}^*, \quad \exists a_n \in \mathbb{R}_+^* \quad \text{et} \quad \exists b_n \in \mathbb{R} \quad \text{telles que} \quad \frac{S_n - b_n}{a_n} \stackrel{d}{=} X$$

où : S_n est une somme de n variables aléatoires réelles **indépendantes** et de même loi de probabilité que celle de X , et a_n , b_n sont des constantes.

Remarque : Dans le cadre des deux définitions ci-dessus, si $d = 0$ ou $b_n = 0$ (pour tout $n \in \mathbb{N}^*$), on dit que la variable aléatoire X , ou sa loi de probabilité, est **strictement stable**.

Dans la proposition qui suit, on va donner le paramètre le plus important d'une loi de probabilité stable et sa relation avec les constantes utilisées dans les deux définitions ci-dessus :

Proposition 1.2.3 [FELLER (1971)].

Soit X une variable aléatoire à valeurs dans \mathbb{R} qui suit une loi de probabilité stable, alors il existe un nombre réel $\alpha \in]0, 2]$ tel que :

- Dans Définition 1.2.4 : $c^\alpha = a^\alpha + b^\alpha$.
- Dans Définition 1.2.5 : $a_n = n^{1/\alpha}$.

Définition 1.2.6 [FELLER (1971)].

On appelle le nombre réel $\alpha \in]0, 2]$, qui apparaît dans la proposition ci-dessus, **exposant caractéristique** de la loi de probabilité stable de X .

Dans la proposition qui suit, on va voir la relation entre la loi de probabilité stable et la loi de probabilité indéfiniment divisible :

Proposition 1.2.4

Une loi de probabilité stable est une loi indéfiniment divisible.

Preuve.

En utilisant les notations de **Définition 1.2.5**, soient X une variable aléatoire réelle stable et $S_n = X_1 + X_2 + \dots + X_n$.

On pose :

$$Y_i = \frac{X_i - (b_n/n)}{a_n}, \quad i = 1, 2, \dots, n$$

Puisque les X_i sont indépendantes et identiquement distribuées, alors il en est de même pour les Y_i .

Donc, on a :

$$Y_1 + Y_2 + \dots + Y_n = \frac{X_1 + X_2 + \dots + X_n - b_n}{a_n} = \frac{S_n - b_n}{a_n}$$

Or :

$$\frac{S_n - b_n}{a_n} \stackrel{d}{=} X \quad (\text{Car } X \text{ est stable})$$

D'où :

$$Y_1 + Y_2 + \dots + Y_n \stackrel{d}{=} X$$

On déduit que, pour tout $n \in \mathbb{N}^*$, il existe des variables aléatoires réelles Y_1, Y_2, \dots, Y_n indépendantes et identiquement distribuées telles que la loi de probabilité de leur somme est la même que celle de X .

Donc : X est une variable aléatoire réelle indéfiniment divisible. ■

Remarque : L'inverse de la proposition ci-dessus n'est pas toujours vrai. On peut facilement vérifier, en utilisant **Définition 1.2.4**, que la loi de probabilité de POISSON $\mathcal{P}(\lambda)$ ($\lambda > 0$), qui est une loi indéfiniment divisible, n'est pas stable.

II.1 Représentation spectrale d'une loi de probabilité stable

La plupart des lois de probabilité stables souffrent du fait que leurs fonctions de répartition et leurs fonctions de densité n'ont pas une forme explicite, sauf dans trois cas qu'on va voir comme exemples. Néanmoins, elles peuvent être décrites par leurs fonctions caractéristiques, qui sont un outil pertinent pour les étudier et pour extraire leurs propriétés. Il existe plusieurs représentations de la fonction caractéristique d'une loi de probabilité stable, la plus utilisée est celle de LÉVY-KHINTCHINE qu'on donne dans le théorème qui suit :

Théorème 1.2.8 (LÉVY-KHINTCHINE).

Une variable aléatoire X à valeurs dans \mathbb{R} a une loi de probabilité stable si et seulement si sa fonction caractéristique, qu'on note φ , est donnée comme suit :

$$\forall t \in \mathbb{R}, \quad \varphi(t) = \exp \left\{ i\mu t - \sigma^\alpha |t|^\alpha \left(1 - i\beta \cdot \text{sign}(t) \varpi(t, \alpha) \right) \right\} \quad (1.4)$$

où : $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}_+^*$, $\alpha \in]0, 2]$, $\beta \in [-1, 1]$, et :

$$\forall t \in \mathbb{R}, \quad \forall \alpha \in]0, 2], \quad \varpi(t, \alpha) = \begin{cases} \tan\left(\frac{\pi}{2}\alpha\right) & \text{si } \alpha \neq 1, \\ -\frac{2}{\pi} \ln|t| & \text{si } \alpha = 1. \end{cases} \quad \text{et } \text{sign}(t) = \begin{cases} -1 & \text{si } t < 0, \\ 0 & \text{si } t = 0, \\ 1 & \text{si } t > 0. \end{cases}$$

Démonstration.

Pour une démonstration de ce théorème, le lecteur pourra se référer aux ouvrages [GNEDEKO & KOLMOGOROV (1954)] et [BREIMAN (1992)]. ■

Notation : D'après le théorème ci-dessus, si la loi de probabilité de X est stable, alors sa fonction caractéristique dépend de quatre paramètres : $\alpha \in]0, 2]$, $\beta \in [-1, 1]$, $\mu \in \mathbb{R}$ et $\sigma \in \mathbb{R}_+^*$. On note alors : $X \sim \mathcal{S}_\alpha(\beta, \mu, \sigma)$.

II.2 Exemples

Maintenant, on va voir les trois lois de probabilité stables où leurs fonctions de densité ont des formules explicites. Pour cela, soit X une variable aléatoire à valeurs dans \mathbb{R} :

Exemple 1.2.4

Si X a une loi de probabilité normale $\mathcal{N}(m, v^2)$ ($m \in \mathbb{R}$ et $v > 0$), alors :

- Sa fonction de densité f_X est donnée par :

$$\forall x \in \mathbb{R}, \quad f_X(x) = \frac{1}{v\sqrt{2\pi}} \exp\left\{-\frac{(x-m)^2}{2v^2}\right\}$$

- Sa fonction caractéristique φ_X est donnée par :

$$\forall t \in \mathbb{R}, \quad \varphi_X(x) = \exp\left\{imt - \frac{t^2}{2}v^2\right\}$$

Donc : une loi de probabilité normale $\mathcal{N}(m, v^2)$ ($m \in \mathbb{R}$ et $v > 0$) est une loi stable $\mathcal{S}_2(\beta, m, v/\sqrt{2})$.

Et inversement, une loi de probabilité stable $\mathcal{S}_2(\beta, \mu, \sigma)$ est une loi normale $\mathcal{N}(\mu, 2\sigma^2)$. On constate, alors, qu'une des lois de probabilité les plus importantes dans la théorie des probabilités et des statistiques, qui est la loi de probabilité normale, est une loi stable.

Exemple 1.2.5

Si X a une loi de probabilité de CAUCHY $\mathcal{C}(a, h)$ ($a \in \mathbb{R}$ et $h > 0$), alors :

- Sa fonction de densité f_X est donnée par :

$$\forall x \in \mathbb{R}, \quad f_X(x) = \frac{h}{\pi[h^2 + (x-a)^2]}$$

- Sa fonction caractéristique φ_X est donnée par :

$$\forall t \in \mathbb{R}, \quad \varphi_X(x) = \exp(iat - h|t|)$$

Donc : une loi de probabilité de CAUCHY $\mathcal{C}(a, h)$, ($a \in \mathbb{R}$ et $h > 0$) est une loi stable $\mathcal{S}_1(0, a, h)$. Et inversement, une loi de probabilité stable $\mathcal{S}_1(0, \mu, \sigma)$ est une loi de CAUCHY $\mathcal{C}(\mu, \sigma)$.

Exemple 1.2.6 [APPLEBAUM (2004)].

Si X a une loi de probabilité stable $\mathcal{S}_{1/2}(1, \mu, \sigma)$, on dit alors que X suit une loi de probabilité de LÉVY. Dans ce cas, sa fonction de densité f_X est définie par :

$$\forall x \in \mathbb{R}, \quad f_X(x) = \left(\frac{\sigma}{2\pi}\right)^{1/2} \frac{1}{(x - \mu)^{3/2}} \exp\left\{-\frac{\sigma}{2(x - \mu)}\right\} \cdot \mathbb{1}_{] \mu, +\infty[}(x)$$

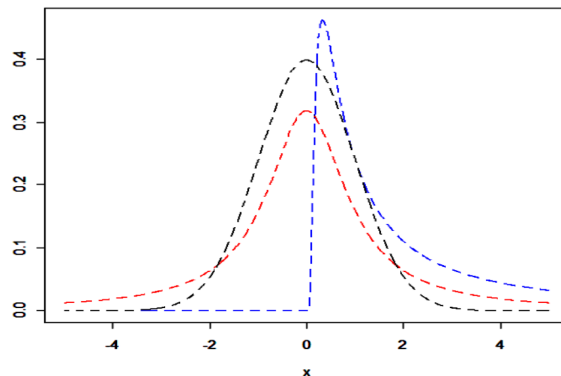


Figure 1.2.1 Graphes des fonctions de densité : Le noir de $\mathcal{N}(0, 1)$, le rouge de $\mathcal{C}(0, 1)$ et le bleu de $\mathcal{S}_{1/2}(1, 0, 1)$ (LÉVY).

II.3 Interprétation des paramètres d'une loi de probabilité stable

Comme on vient de voir, une loi de probabilité stable $\mathcal{S}_\alpha(\beta, \mu, \sigma)$ est caractérisée par quatre paramètres :

1) $\alpha \in]0, 2]$: Le paramètre α , appelé **exposant caractéristique** ou **indice de stabilité** de la loi de probabilité $\mathcal{S}_\alpha(\beta, \mu, \sigma)$, est le plus important des paramètres. Il détermine les propriétés de base de cette loi de probabilité (moments, etc.) et il décrit la forme de la distribution ou le degré d'épaisseur des queues de la distribution : plus α est petit, plus les queues de la distribution sont épaisses. Comme on la vu dans les exemples ci-dessus, la loi de probabilité normale a la valeur maximum de α , soit $\alpha = 2$.

2) $\beta \in [-1, 1]$: Le paramètre β est le paramètre d'**asymétrie** de la loi de probabilité $\mathcal{S}_\alpha(\beta, \mu, \sigma)$. Si $\beta < 0$, la distribution est asymétrique à gauche et sa queue est plus épaisse à gauche et si $\beta > 0$, la distribution est asymétrique à droite et sa queue est plus épaisse à droite. Si $\beta = 0$, la distribution est symétrique par rapport à μ , et, dans ce cas, sa fonction caractéristique donnée dans (1.4) devient :

$$\forall t \in \mathbb{R} \quad \text{et} \quad \forall (\alpha, \mu, \sigma) \in]0, 2] \times \mathbb{R} \times \mathbb{R}_+^*, \quad \varphi(t) = \exp\{i\mu t - \sigma^\alpha |t|^\alpha\}$$

Remarque : Quand le paramètre α se rapproche de 2, le paramètre β perd son effet sur la distribution et la loi de probabilité $\mathcal{S}_\alpha(\beta, \mu, \sigma)$ se rapproche de la loi de probabilité normale qui ne dépend pas de β .

3) $\mu \in \mathbb{R}$: Le paramètre μ est le paramètre de **position** ou de **localisation** de la loi de probabilité $\mathcal{S}_\alpha(\beta, \mu, \sigma)$. Lorsque $\beta = \mu = 0$, la loi de probabilité $\mathcal{S}_\alpha(0, 0, \sigma)$ est dite **symétrique alpha-stable** et on la note $\mathcal{S}_\alpha\mathcal{S}(\sigma)$, et, dans ce cas, sa fonction caractéristique donnée dans (1.4) devient :

$$\forall t \in \mathbb{R} \quad \text{et} \quad \forall (\alpha, \sigma) \in]0, 2] \times \mathbb{R}_+^*, \quad \varphi(t) = \exp \left\{ -\sigma^\alpha |t|^\alpha \right\}$$

4) $\sigma \in \mathbb{R}_+^*$: Le paramètre σ est le paramètre de **dispersion** ou d'**échelle** de la loi de probabilité $\mathcal{S}_\alpha(\beta, \mu, \sigma)$. Si $\mu = 0$ et $\sigma = 1$, on appelle la loi de probabilité $\mathcal{S}_\alpha(\beta, 0, 1)$ une loi stable **standard**.

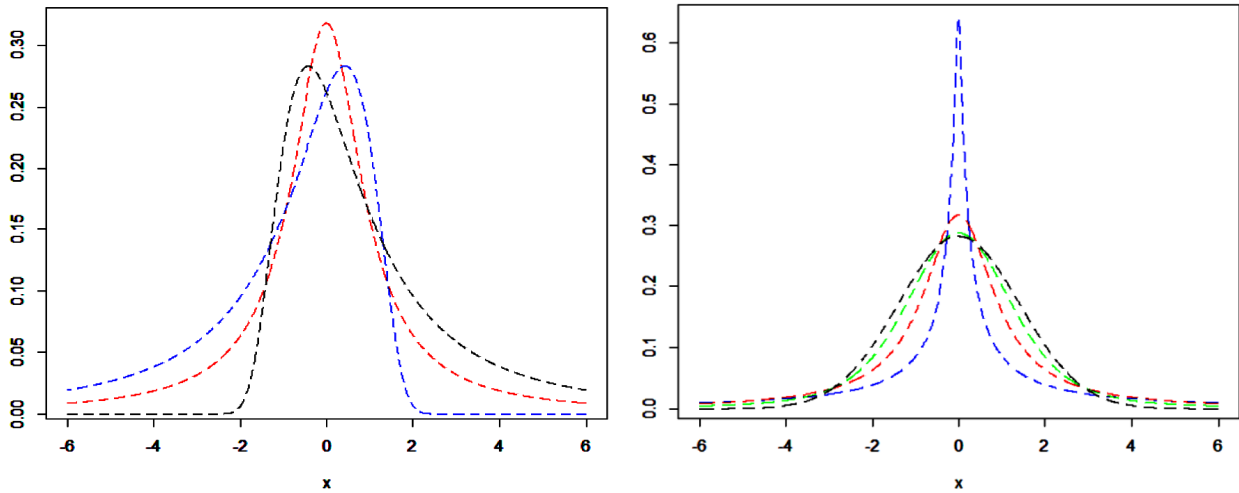


Figure 1.2.2 Graphes des densités : Le panneau gauche pour $\mathcal{S}_1(\beta, 0, 1)$ et le panneau droit pour $\mathcal{S}_\alpha\mathcal{S}(1)$.

II.4 Propriétés d'une loi de probabilité stable

II.4.1 Propriétés arithmétiques

Proposition 1.2.5 [MERAGHNI (2008)].

Soient X et Y deux variables aléatoires réelles *indépendantes* telles que $X \sim \mathcal{S}_\alpha(\beta_1, \mu_1, \sigma_1)$ et $Y \sim \mathcal{S}_\alpha(\beta_2, \mu_2, \sigma_2)$, et soient $a \in \mathbb{R}^*$ et $b \in \mathbb{R}$. On a :

- $X + Y \sim \mathcal{S}_\alpha(\beta, \mu, \sigma)$, avec :

$$\beta = \frac{\beta_1 \sigma_1^\alpha + \beta_2 \sigma_2^\alpha}{\sigma_1^\alpha + \sigma_2^\alpha}, \quad \mu = \mu_1 + \mu_2 \quad \text{et} \quad \sigma = (\sigma_1^\alpha + \sigma_2^\alpha)^{1/\alpha}$$

- $aX \sim \mathcal{S}_\alpha(\beta, \mu, \sigma)$, avec :

$$\begin{cases} \beta = \text{sign}(a)\beta_1, & \mu = a\mu_1 \quad \text{et} \quad \sigma = |a|\sigma_1 & \text{si } \alpha \neq 1, \\ \beta = \text{sign}(a)\beta_1, & \mu = a\mu_1 - \frac{2}{\pi}a\beta_1\sigma_1 \ln |a| \quad \text{et} \quad \sigma = |a|\sigma_1 & \text{si } \alpha = 1. \end{cases}$$

- $X + b \sim \mathcal{S}_\alpha(\beta, \mu, \sigma)$, avec :

$$\beta = \beta_1, \quad \mu = \mu_1 + b \quad \text{et} \quad \sigma = \sigma_1$$

De la proposition ci-dessus, on déduit un résultat très important pour la simulation des lois de probabilité stables (voir [CHAMBERS, MALLOWS & STUCK (1976)]), qui est le suivant :

Corollaire 1.2.1 [MERAGHNI (2008)].

Soit X une variable aléatoire à valeurs dans \mathbb{R} , et soient $\alpha \in]0, 2]$, $\beta \in [-1, 1]$, $\mu \in \mathbb{R}$ et $\sigma \in \mathbb{R}_+^*$. On a :

- Si $\alpha \neq 1$:

$$X \sim \mathcal{S}_\alpha(\beta, \mu, \sigma) \iff Y = \frac{X - \mu}{\sigma} \sim \mathcal{S}_\alpha(\beta, 0, 1)$$

- Si $\alpha = 1$:

$$X \sim \mathcal{S}_1(\beta, \mu, \sigma) \iff Y = \frac{X - [\mu + 2(\pi)^{-1}\beta\sigma \ln(\sigma)]}{\sigma} \sim \mathcal{S}_1(\beta, 0, 1)$$

Donc, pour simuler une loi de probabilité stable $\mathcal{S}_\alpha(\beta, \mu, \sigma)$, il suffit de simuler une loi de probabilité $\mathcal{S}_\alpha(\beta, 0, 1)$ et, en suite, faire le changement de variables.

II.4.2 Moments

Proposition 1.2.6 [D'ESTAMPES (2003)].

Soit X une variable aléatoire à valeurs dans \mathbb{R} qui a une loi de probabilité stable $\mathcal{S}_\alpha(\beta, \mu, \sigma)$, et soit $p \in \mathbb{R}_+^*$. On a :

1. Si $\alpha = 2$, alors :

$$\forall p \in \mathbb{R}_+^*, \quad \mathbb{E}(|X|^p) < \infty$$

2. Si $p < \alpha$, alors :

$$\mathbb{E}(|X|^p) < \infty$$

3. Si $\alpha \in]0, 2[$ et $p > \alpha$, alors :

$$\mathbb{E}(|X|^p) = \infty$$

$0 < \alpha \leq 1$		$1 < \alpha < 2$		$\alpha = 2$	
$\mathbb{E}(X) = \infty$	$\mathbb{V}(X) = \infty$	$\mathbb{E}(X) = \mu$	$\mathbb{V}(X) = \infty$	$\mathbb{E}(X) = \mu$	$\mathbb{V}(X) = 2\sigma^2$

Table 1.2.1 Espérance et variance de la variable aléatoire $X \sim \mathcal{S}_\alpha(\beta, \mu, \sigma)$ selon les valeurs de α .

II.4.3 Queues lourdes

Il n'existe pas une caractérisation type des lois de probabilité aux queues lourdes. La notion caractérise généralement la qualité de la déviation d'une loi de probabilité par rapport à la loi normale dont la fonction de queue décroît exponentiellement.

Définition 1.2.7 [BARRO (2010) et BOUZIANE (2011)].

Soit X une variable aléatoire à valeurs dans \mathbb{R} de fonction de répartition F . On dit que X ou la loi de probabilité de X ou, encore, la fonction de répartition F de X est à queue lourde à droite s'il existe $b \in \mathbb{R}$ tel que :

$$\forall x \in]b, +\infty[, \quad \exists \eta > 0 \text{ et } \kappa > 0 \text{ tels que } \bar{F}(x) = 1 - F(x) > \eta e^{-\kappa x}$$

Remarque : Dans la théorie des probabilités, parmi les lois de probabilité aux queues lourdes à droite, on trouve celles qui ont une fonction de queue qui varie régulièrement à l'infini d'indice $\rho \in \mathbb{R}$ (voir ANNEXE B).

Proposition 1.2.7 [MERAGHNI (2008)].

Soit X une variable aléatoire réelle qui suit une loi de probabilité stable $\mathcal{S}_\alpha(\beta, \mu, \sigma)$ ($\alpha < 2$), alors on a :

$$\lim_{x \rightarrow \infty} x^\alpha \mathbb{P}(X > x) = C_\alpha \frac{1 + \beta}{2} \sigma^\alpha \quad \text{et} \quad \lim_{x \rightarrow \infty} x^\alpha \mathbb{P}(X < -x) = C_\alpha \frac{1 - \beta}{2} \sigma^\alpha$$

$$C_\alpha = \left(\int_0^{+\infty} t^{-\alpha} \sin t \, dt \right)^{-1} = \frac{2}{\pi} \Gamma(\alpha) \sin\left(\frac{\pi}{2}\alpha\right)$$

Remarque : Dans le cadre de la proposition ci-dessus, Γ est la fonction Gamma définie par :

$$\forall x > 0, \quad \Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} \, dt$$

II.4.4 Fonctions de répartition et fonctions de densité

Comme on l'a déjà mentionné en donnant des exemples de lois de probabilité stables, la plupart de ces lois n'ont pas de fonctions de densité ayant des expressions analytiques explicites. Ce manque de formes explicites des fonctions de densité a été considéré comme un inconvénient majeur pour les applications des lois de probabilité stables. Grâce à la méthode « **transformée de FOURIER rapide (FFT¹)** » qui exploite la relation entre la fonction de densité f d'une loi de probabilité et sa fonction caractéristique φ :

$$f(x) = \frac{1}{2\pi} \int e^{-itx} \varphi(t) \, dt$$

il existe, maintenant, des programmes informatiques qui permettent de calculer numériquement toute quantité ayant rapport avec une loi de probabilité stable.

III Théorème de la limite centrale généralisé

Le théorème de la limite centrale généralisé est l'un des résultats les plus importants dans la théorie des probabilités et ses applications, il caractérise et donne du poids à la loi de probabilité stable. Le voici :

1. Fast FOURIER Transform

Théorème 1.2.9 (Théorème de la limite centrale généralisé) [SHIRYAEV (1984)].

Soient $\{X_n, n \in \mathbb{N}^*\}$ une suite de variables aléatoires réelles **non dégénérées, indépendantes et identiquement distribuées** et Y une variable aléatoire réelle **non dégénérée** telles qu'il existe des constantes $a_n \in \mathbb{R}_+^*$ et $b_n \in \mathbb{R}$ vérifiant :

$$\frac{S_n - b_n}{a_n} \xrightarrow{d} Y \quad \text{avec} \quad S_n = \sum_{i=1}^n X_i \quad (1.5)$$

alors :

$$(1.5) \iff Y \sim \mathcal{S}_\alpha(\beta, \mu, \sigma)$$

IV Domaine d'attraction

Dans ce qui suit, soient X une variable aléatoire réelle non dégénérée de fonction de répartition F , $\{X_n, n \in \mathbb{N}^*\}$ une suite de variables aléatoires réelles non dégénérées, indépendantes et de même loi de probabilité que celle de X et S_n la somme des n premiers termes de la suite $(X_n)_{n \in \mathbb{N}^*}$.

La question qu'on peut poser est la suivante : Étant donné une loi de probabilité stable $\mathcal{S}_\alpha(\beta, \mu, \sigma)$, quelles sont les conditions nécessaires et suffisantes sur la variables aléatoire X et quelles sont les constantes de normalisation $a_n \in \mathbb{R}_+^*$ et $b_n \in \mathbb{R}$ pour que X soit dans le domaine d'attraction de $\mathcal{S}_\alpha(\beta, \mu, \sigma)$?

Définition 1.2.8 [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

On dit que la variable aléatoire X ou la loi de probabilité de X ou, encore, la fonction de répartition F de X appartient au **domaine d'attraction** d'une loi de probabilité stable $\mathcal{S}_\alpha(\beta, \mu, \sigma)$ s'il existe des constantes $a_n \in \mathbb{R}_+^*$ et $b_n \in \mathbb{R}$ telles que :

$$\frac{S_n - b_n}{a_n} \xrightarrow{d} \mathcal{S}_\alpha(\beta, \mu, \sigma)$$

Référence

La plupart des résultats de ce chapitre peuvent être trouvés dans beaucoup de livres de probabilités dont certains sont cités dans la bibliographie de ce mémoire. Ce qui concerne la question posée ci-dessus, le lecteur peut consulter les ouvrages suivants : [FELLER (1971)], [GNEDENKO & KOLMOGOROV (1954)] et [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

CHAPITRE 2

THÉORIE DES VALEURS EXTRÊMES UNIVARIÉE

Introduction

Ce chapitre expose la théorie classique des valeurs extrêmes, qui s'intéresse non pas à la modélisation totale d'une distribution mais seulement à la queue de cette distribution, et, par conséquent, il est fondamental pour beaucoup de résultats dans ce mémoire. Le résultat principal de cette théorie est le théorème de FISHER-TIPPETT (1928) & GNEDENKO (1943) qui spécifie la forme de la loi limite de la suite des maximums, normalisés, de n variables aléatoires indépendantes et identiquement distribuées, quand n tend vers l'infini.

Dans toute la suite de ce chapitre, soient X une variable aléatoire non dégénérée définie sur l'espace probabilisé (Ω, \mathcal{A}, P) à valeurs dans \mathbb{R} de fonction de distribution (de répartition) F et $\{X_n, n \in \mathbb{N}^*\}$ une suite de variables aléatoires non dégénérées, indépendantes et de même loi que celle de X . On pose $M_n = \max(X_1, X_2, \dots, X_n)$ et $S_n = \sum_{k=1}^n X_k$ ($n \in \mathbb{N}^*$).

La théorie des valeurs extrêmes a été développée en parallèle avec la théorie de la limite centrale, et, en fait, les deux théories portent quelques ressemblances : loi stable correspond à loi max-stable, domaine d'attraction correspond au max-domaine d'attraction et le théorème limite pour le maximum (Théorème de FISHER-TIPPETT (1928) & GNEDENKO (1943)) exige des constantes de normalisation ainsi que le théorème limite pour la somme (Théorème de la limite centrale).

L'étude de la loi de probabilité du maximum d'un échantillon de n variables aléatoires a été la première approche pour décrire les événements extrêmes. Nous allons mentionner un résultat qui permet de voir le lien entre le comportement asymptotique d'une somme de variables aléatoires indépendantes et identiquement distribuées avec le comportement asymptotique du maximum de ces variables. Il s'agit du **Théorème de HEYDE (1967)** sur *les grandes déviations* qui montre que sous certaines hypothèses sur la fonction de distribution F de X , on a :

$$\lim_{n \rightarrow \infty} \frac{\mathbb{P}(S_n > x_n)}{\mathbb{P}(M_n > x_n)} = 1 \quad \text{avec} \quad \lim_{n \rightarrow \infty} x_n = \infty$$

Autrement dit, pour un nombre entier n grand, la probabilité pour qu'une somme de réalisations d'une suite de n variables aléatoires indépendantes et identiquement distribuées dépasse un seuil x (grand) est proche de la probabilité pour que la plus grandes des réalisations de cette suite dépasse x . Ainsi, pour qu'une somme de variables aléatoires indépendantes et identiquement distribuées dépasse un seuil x , il suffit qu'une de ses composantes prenne une valeur extrême très supérieure aux valeurs prises par chacune des autres composantes. Cette réalisation est l'événement rare dont on cherche à identifier et à estimer la distribution.

Donc, l'objectif principal de la théorie des valeurs extrêmes est de trouver les lois limites possibles pour la suite des maximums $\{M_n = \max(X_1, X_2, \dots, X_n), n \in \mathbb{N}^*\}$. Bien sûr, si on connaît la loi exacte de la variable aléatoire X , il n'est pas difficile de déterminer la loi exacte de M_n donnée dans la proposition suivante :

Proposition 1 [ARNOLD, BALAKRISHNAN & NAGARAJA (2008)].

- La fonction de distribution (de répartition) F_{M_n} de M_n est donnée par :

$$\forall x \in \mathbb{R}, \quad F_{M_n}(x) = \mathbb{P}(M_n \leq x) = F^n(x) \quad (2.1)$$

- Si X est absolument continue de densité f , alors la fonction de densité f_{M_n} de M_n est donnée par :

$$\forall x \in \mathbb{R}, \quad f_{M_n}(x) = nF^{n-1}(x)f(x)$$

Remarque : D'après la formule (2.1), les propriétés statistiques de M_n , données par sa fonction de distribution, dépendent principalement des propriétés de F pour les grandes valeurs de x . Pour les autres valeurs de x , l'influence de F décroît rapidement avec la croissance de n . Donc, l'information la plus importante à propos des extrêmes est contenue dans la queue de la loi de probabilité de X .

La loi de probabilité de M_n devrait nous fournir des informations au sujet des événements extrêmes. Or que, M_n tend vers un nombre réel qui peut être infini, qu'on appelle *point terminal* ou *point le plus à droite* de la fonction de distribution F , quand $n \rightarrow \infty$.

Définition 1 [LEKINA (2010) et AVOUYI-DOVI & GUÉGAN (2001)].

On appelle le **point terminal** ou le **point le plus à droite** de la fonction de distribution F , noté x_F , la borne supérieure du support de F défini par :

$$x_F = \sup \{x \in \mathbb{R}, F(x) < 1\} \leq \infty$$

Proposition 2 [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

La suite des maximums $\{M_n = \max(X_1, X_2, \dots, X_n), n \geq 1\}$ converge presque-sûrement vers x_F quand $n \rightarrow \infty$, i.e :

$$M_n \xrightarrow{p.s.} x_F \quad \text{quand} \quad n \rightarrow \infty \quad (2.2)$$

Le résultat suivant découle automatiquement de (2.2) :

Corollaire 1

La suite des maximums $\{M_n, n \in \mathbb{N}^*\}$ converge en loi vers une variable aléatoire dégénérée concentrée en x_F , car :

$$\lim_{n \rightarrow \infty} F_{M_n}(x) = \lim_{n \rightarrow \infty} \mathbb{P}(M_n \leq x) = \lim_{n \rightarrow \infty} F^n(x) = \begin{cases} 0 & \text{si } x < x_F, \\ 1 & \text{si } x \geq x_F. \end{cases} \quad (2.3)$$

Remarque : Les formules exactes de la loi du terme maximum M_n donnée dans (2.1) et de la loi limite donnée dans (2.3) présentent toutefois un intérêt limité, et elles ne fournissent pas beaucoup d'informations sur les extrêmes. En pratique, la loi de la variable aléatoire parente X est rarement connue avec précision, et, par conséquent, il en est de même pour la loi du terme maximum M_n . De plus, même si la loi de X est connue avec exactitude, la loi de M_n n'est pas toujours facilement calculable. Par exemple, la fonction de distribution d'une variable aléatoire normale ne possède pas d'expression analytique, puisque c'est une intégrale incalculable. Donc, sa puissance $n^{\text{ème}}$ est difficile à calculer et conduit à de sérieux problèmes numériques pour de grandes valeurs de n et pour de grandes valeurs de x . Pour ces raisons, il est intéressant de considérer le comportement asymptotique du terme maximum M_n normalisé.

Supposons qu'il existe une suite $\{a_n, n \in \mathbb{N}^*\}$ de nombres réels strictement positifs et une suite $\{b_n, n \in \mathbb{N}^*\}$ de nombres réels telles que la suite des maximums normalisés $\{a_n^{-1}(M_n - b_n), n \in \mathbb{N}^*\}$ ait une loi limite non dégénérée quand $n \rightarrow \infty$, i.e :

$$\forall x \in \mathcal{C}(H), \quad \lim_{n \rightarrow \infty} \mathbb{P}\left(\frac{M_n - b_n}{a_n} \leq x\right) = \lim_{n \rightarrow \infty} F^n(a_n x + b_n) = H(x) \quad (2.4)$$

où H est une fonction de distribution non dégénérée et $\mathcal{C}(H) = \{x \in \mathbb{R}, H \text{ continue en } x\}$. Donc, le but est de trouver toutes les fonctions de distribution possibles H non dégénérées et qui peuvent réaliser la limite dans (2.4) (en terme exact, trouver les types de distributions possibles (voir ANNEXE A)). Ces distributions sont appelées « **distributions des valeurs extrêmes** ». Ensuite, pour chacune de ces lois limites, trouver les conditions nécessaires et suffisantes sur la fonction de distribution initiale F pour que (2.4) se réalise. La classe des fonctions de distributions F satisfaisant (2.4) est appelée « **max-domaine d'attraction** ». Et, après, identifier toutes les distributions des valeurs extrêmes ainsi que leurs max-domaines d'attraction.

Définition 2

Les suites $\{a_n > 0, n \geq 1\}$ et $\{b_n, n \geq 1\}$ sont appelées *suites de normalisation*, les constantes $a_n \in \mathbb{R}_+^*$ et $b_n \in \mathbb{R}$ sont appelées *constantes de normalisation* et la variable aléatoire $[a_n^{-1}(M_n - b_n)]$ est appelée *maximum normalisé*.

2.1 Loi max-stable

Comme dans la théorie de la limite centrale, nous avons défini les lois stables qui sont les seules lois limites possibles pour la suite des sommes, normalisées, de n variables aléatoires indépendantes

et identiquement distribuées quand n tend vers l'infini, il existe une notion similaire dans la théorie des valeurs extrêmes :

Définition 2.1.1 [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

La variable aléatoire, non dégénérée, X ou la loi de probabilité de X ou, encore, la fonction de distribution F de X est dite **max-stable**, s'il existe des constantes $a_n \in \mathbb{R}_+^*$ et $b_n \in \mathbb{R}$ telles que :

$$M_n \stackrel{d}{=} a_n X + b_n \quad \text{pour tout } n \in \mathbb{N}^*$$

Ou, ce qui est équivalent, s'il existe des constantes $a_n \in \mathbb{R}_+^*$ et $b_n \in \mathbb{R}$ telles que :

$$F^n(a_n x + b_n) = F(x) \quad \text{pour tout } n \in \mathbb{N}^* \text{ et } x \in \mathbb{R}$$

Exemple 2.1.1 [COLES (2001)].

Supposons que X est une variable aléatoire suivant une loi de probabilité de FRÉCHET standard, i.e :

$$\forall x \in \mathbb{R}, \quad F(x) = P(X \leq x) = \begin{cases} \exp(-x^{-1}) & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases}$$

Posons $a_n = n$ et $b_n = 0$, on a alors :

$$\begin{aligned} \forall x \in \mathbb{R}, \quad \forall n \in \mathbb{N}^*, \quad F^n(a_n x + b_n) &= F^n(nx) \\ &= \begin{cases} \left[\exp(-(nx)^{-1}) \right]^n & \text{si } nx > 0 \\ 0 & \text{si } nx \leq 0 \end{cases} \\ &= \begin{cases} \exp(-x^{-1}) & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases} \\ &= F(x) \end{aligned}$$

Donc :

$$\forall x \in \mathbb{R}, \quad \forall n \in \mathbb{N}^*, \quad \exists a_n = n \text{ et } b_n = 0 \text{ tels que } F^n(a_n x + b_n) = F(x)$$

On déduit que la loi de probabilité de FRÉCHET standard est une loi max-stable.

Théorème 2.1.1 [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

La classe des lois max-stables coïncide avec la classe de toutes les lois limites possibles, non dégénérées, de la suite de maximums, normalisés, d'une suite de variables aléatoires indépendantes et identiquement distribuées.

Remarque : Le théorème ci-dessus nous donne une propriété importante des lois max-stables, dite « propriété limite des lois max-stables », qui nous montre que toute loi max-stable est une loi limite pour la suite des maximums, normalisés, d'une suite de variables aléatoires indépendantes et identiquement distribuées et, en plus, les lois max-stables sont les seules et uniques lois limites possibles.

2.2 Distributions des valeurs extrêmes

Maintenant, supposons qu'il existe une suite $(a_n)_{n \in \mathbb{N}^*}$ de nombres réels strictement positifs et une suite $(b_n)_{n \in \mathbb{N}^*}$ de nombres réels telles que la suite des maximums normalisés $\{a_n^{-1}(M_n - b_n), n \in \mathbb{N}^*\}$ converge en distribution vers une variable aléatoire non dégénérée de fonction de distribution H , i.e :

$$\forall x \in \mathcal{C}(H), \quad \lim_{n \rightarrow \infty} \mathbb{P}\left(\frac{M_n - b_n}{a_n} \leq x\right) = \lim_{n \rightarrow \infty} F^n(a_n x + b_n) = H(x) \quad (2.5)$$

Alors, FISHER et TIPPETT trouvent en 1928, au moyen d'un théorème qui porte leurs noms et le nom de GNEDENKO qui donne la première preuve rigoureuse de celui-ci en 1943, toutes les lois limites possibles qui vérifient (2.5). Ce théorème, qu'on appelle aussi « **Théorème des trois types de distribution extrêmes** », est l'un des résultats fondamentaux de la théorie des valeurs extrêmes.

2.2.1 Théorème de FISHER-TIPPETT (1928) & GNEDENKO (1943)

Théorème 2.2.1 (Théorème de FISHER-TIPPETT (1928) & GNEDENKO (1943)).

S'il existe des suites de normalisation $\{a_n, n \in \mathbb{N}^\}$ de nombres réels strictement positifs, $\{b_n, n \in \mathbb{N}^*\}$ de nombres réels et une fonction de distribution non dégénérée H , définie et continue sur \mathbb{R} , telles que :*

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\frac{M_n - b_n}{a_n} \leq x\right) = \lim_{n \rightarrow \infty} F^n(a_n x + b_n) = H(x) \quad (2.6)$$

alors H appartient à l'un des trois types de distributions (à un paramètre d'échelle et de localisation près) suivants :

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

Démonstration.

Pour une démonstration du **Théorème 2.2.1**, le lecteur pourra se référer aux ouvrages suivants : [RESNICK (1998)], [CHARPENTIER & DENUIT (2005)] ou [RESNICK (1987)]. ■

Définition 2.2.1 [RONCALLI (2002)].

*Les types de distributions de FRÉCHET, de WEIBULL et de GUMBEL sont dits « **extrêmes** », les fonctions de distribution Φ_α , Ψ_α et Λ sont appelées « **les distributions des valeurs extrêmes standards** » et les variables aléatoires correspondantes sont « **extrémales** ».*

En fonction de la distribution des valeurs extrêmes Φ_α , Ψ_α ou Λ , on dit que la fonction de distribution F de la variable aléatoire X appartient au **max-domaine d'attraction** de Φ_α , de Ψ_α ou de Λ .

Définition 2.2.2 [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

Le paramètre $\alpha > 0$ qui apparaisse dans les types de distribution de FRÉCHET et de WEIBULL est appelé « *paramètre de forme* ».

Remarque : Le **Théorème 2.2.1** n'est valable que si les suites $\{a_n \in \mathbb{R}_+^*, n \in \mathbb{N}^*\}$ et $\{b_n \in \mathbb{R}, n \in \mathbb{N}^*\}$ de normalisation existent et admettent une limite, et, en plus, ces suites de normalisation ne sont pas uniques (voir ANNEXE A). Dans ce cas, il nous donne un résultat très intéressant, car si le cardinal de l'ensemble des distributions (lois de probabilité) est grand, le cardinal de l'ensemble des distributions des valeurs extrêmes est lui très petit et il est vrai pour la majorité des lois usuelles. Si l'on fait un parallèle avec le théorème de la limite centrale, la constante de normalisation $a_n \in \mathbb{R}_+^*$ joue le rôle de $\sigma\sqrt{n}$, où σ est l'écart type de la variable aléatoire X , et la constante de normalisation $b_n \in \mathbb{R}$ joue le rôle de $n\mu$, où μ est l'espérance de la variable aléatoire X . La constante de normalisation $a_n \in \mathbb{R}_+^*$ (respectivement. $b_n \in \mathbb{R}$) s'interprète comme un paramètre d'échelle (respectivement. un paramètre de **position**).

Proposition 2.2.1 (Fonctions de densité des distributions des valeurs extrêmes standards).

Les fonctions de densité des distributions des valeurs extrêmes standards, des différents types de distributions extrêmes, sont les suivantes :

$$\begin{aligned} \text{FRÉCHET : } \quad \phi_\alpha(x) &= \begin{cases} \alpha x^{-(\alpha+1)} \exp(-x^{-\alpha}) & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases}, \quad \alpha > 0 \\ \text{WEIBULL : } \quad \psi_\alpha(x) &= \begin{cases} \alpha(-x)^{\alpha-1} \exp\{-(-x)^\alpha\} & \text{si } x \leq 0 \\ 0 & \text{si } x > 0 \end{cases}, \quad \alpha > 0 \\ \text{GUMBEL : } \quad \lambda(x) &= \exp\{- (x + e^{-x})\}, \quad x \in \mathbb{R} \end{aligned}$$

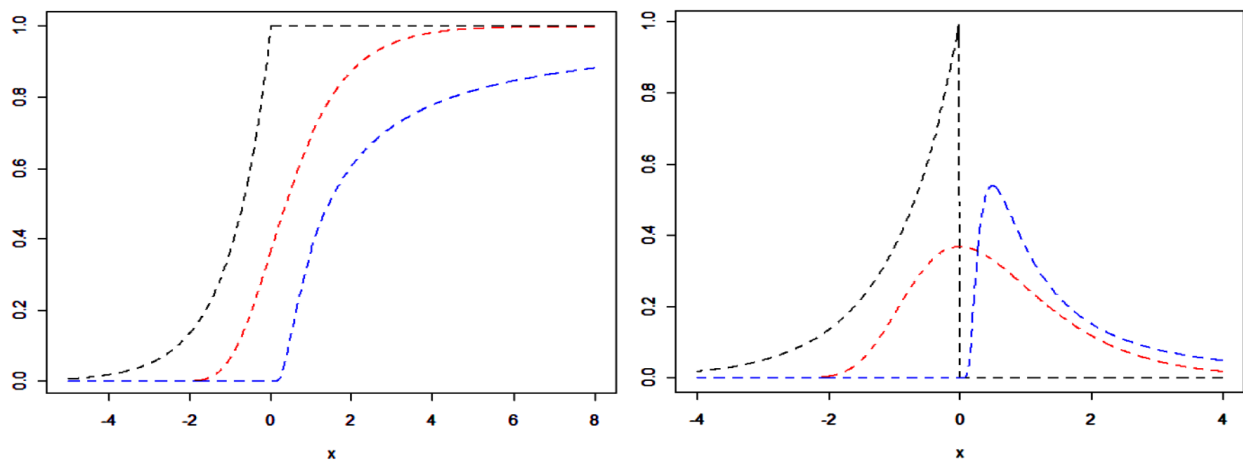


Figure 2.2.1 Panneau gauche représente les graphes des distributions des valeurs extrêmes et le panneau droit représente les graphes de leurs densités ($\alpha = 1$) : Le bleu pour le type de FRÉCHET, le rouge pour le type de GUMBEL et le noir pour le type de WEIBULL.

Proposition 2.2.2 [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

Soient Y une variable aléatoire strictement positive et $\alpha > 0$, alors :

$$(Y \sim \Phi_\alpha) \iff (\ln Y^\alpha \sim \Lambda) \iff (-Y^{-1} \sim \Psi_\alpha)$$

2.2.2 Exemples

Exemple 2.2.1 [COLES (2001)].

Supposons que X suit une loi de probabilité exponentielle standard $\mathcal{E}(1)$, i.e :

$$\forall x \in \mathbb{R}, \quad F(x) = P(X \leq x) = \begin{cases} 1 - e^{-x} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases}$$

Si nous posons $a_n = 1$ et $b_n = \ln n$, nous aurons :

$$\forall x \in \mathbb{R}, \quad \forall n \in \mathbb{N}^*, \quad P\left(\frac{M_n - b_n}{a_n} \leq x\right) = P(M_n \leq a_n x + b_n) = F^n(x + \ln n)$$

Alors :

$$\begin{aligned} P\left(\frac{M_n - b_n}{a_n} \leq x\right) &= \begin{cases} [1 - e^{-(x + \ln n)}]^n & \text{si } x + \ln n > 0 \\ 0 & \text{si } x + \ln n \leq 0 \end{cases} \\ &= \begin{cases} \left[1 + \frac{(-e^{-x})}{n}\right]^n & \text{si } x > -\ln n \\ 0 & \text{si } x \leq -\ln n \end{cases} \end{aligned}$$

Donc :

$$\forall x \in \mathbb{R}, \quad \lim_{n \rightarrow \infty} P\left(\frac{M_n - b_n}{a_n} \leq x\right) = \exp\{-e^{-x}\} = \Lambda(x) \quad \text{Car } \forall x \in \mathbb{R}, \quad \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n = e^x$$

On déduit que la loi exponentielle standard $\mathcal{E}(1)$ appartient au max-domaine d'attraction de Λ .

Exemple 2.2.2 [COLES (2001)].

Supposons que X suit une loi de probabilité uniforme $\mathcal{U}([0, 1])$, i.e :

$$\forall x \in \mathbb{R}, \quad F(x) = P(X \leq x) = \begin{cases} 0 & \text{si } x < 0 \\ x & \text{si } 0 \leq x \leq 1 \\ 1 & \text{si } x > 1 \end{cases}$$

Si nous posons $a_n = n^{-1}$ et $b_n = 1$, nous aurons :

$$\forall x \in \mathbb{R}, \quad \forall n \in \mathbb{N}^*, \quad P\left(\frac{M_n - b_n}{a_n} \leq x\right) = P(M_n \leq a_n x + b_n) = F^n\left(1 + \frac{x}{n}\right)$$

Alors :

$$\begin{aligned} \mathbb{P}\left(\frac{M_n - b_n}{a_n} \leq x\right) &= \begin{cases} 0 & \text{si } 1 + \frac{x}{n} < 0 \\ \left(1 + \frac{x}{n}\right)^n & \text{si } 0 \leq 1 + \frac{x}{n} \leq 1 \\ 1 & \text{si } 1 + \frac{x}{n} > 1 \end{cases} \\ &= \begin{cases} 0 & \text{si } x < -n \\ \left(1 + \frac{x}{n}\right)^n & \text{si } -n \leq x \leq 0 \\ 1 & \text{si } x > 0 \end{cases} \end{aligned}$$

Donc :

$$\forall x \in \mathbb{R}, \quad \lim_{n \rightarrow \infty} \mathbb{P}\left(\frac{M_n - b_n}{a_n} \leq x\right) = \begin{cases} e^x & \text{si } x \leq 0 \\ 1 & \text{si } x > 0 \end{cases} = \Psi_1(x)$$

On déduit que la loi uniforme $\mathcal{U}([0, 1])$ appartient au max-domaine d'attraction de Ψ_1 .

2.2.3 Résultats avancés

La version du **Théorème 2.2.1** présentée ci-dessus repose sur des hypothèses de base. Le même résultat concernant la forme de la loi de probabilité limite est trouvé si l'on relâche ces hypothèses. Citons quelques extensions de ce théorème : WATSON montre, en 1954, que le résultat du théorème est encore valide si les variables aléatoires de la suite $\{X_n, n \in \mathbb{N}^*\}$ sont m -dépendantes. En 1964, BERMAN montre que le résultat tient toujours si les variables aléatoires de la suite $\{X_n, n \in \mathbb{N}^*\}$ sont corrélées, la série des coefficients de corrélation au carré étant convergente. DE HAAN, RESNICK, ROOTZÈN & DE VRIES, en 1989, montrent que si le processus suivi par la suite des variables aléatoires $\{X_n, n \in \mathbb{N}^*\}$ est un processus ARCH, alors la fonction de distribution H est du type de FRÉCHET. D'après LEADBETTER, LINDGREN & ROOTZÈN, en 1983, un mélange de variables aléatoires normales produit une loi de probabilité limite du type de GUMBEL. Ces extensions montrent que les hypothèses d'indépendance et d'identité de la loi de probabilité parente ne sont pas cruciales.

2.3 Domaine d'attraction du maximum

Dans toute la suite de ce chapitre, soit H une distribution des valeurs extrêmes, c'est à dire que H est la fonction de distribution (de répartition) de la loi de probabilité limite de la suite des maxima normalisés $\{a_n^{-1}(M_n - b_n), n \in \mathbb{N}^*\}$ ($a_n \in \mathbb{R}_+^*$ et $b_n \in \mathbb{R}$) et qui est, bien sûr, continue sur \mathbb{R} .

La question qui vient à l'esprit est la suivante : Étant donné H , quelles sont les conditions nécessaires et/ou suffisantes sur la variables aléatoire X pour que la limite donnée dans (2.6) se réalise ?

Définition 2.3.1 [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

On dit que la variable aléatoire X ou la loi de probabilité de X ou, encore, la fonction de distribution F de X appartient au **max-domaine d'attraction** de H , et on note $X \in \text{MDA}(H)$ ($F \in \text{MDA}(H)$), s'il existe des constantes $a_n \in \mathbb{R}_+^*$ et $b_n \in \mathbb{R}$ telles que :

$$\forall x \in \mathbb{R}, \quad \lim_{n \rightarrow \infty} \mathbb{P}\left(\frac{M_n - b_n}{a_n} \leq x\right) = \lim_{n \rightarrow \infty} F^n(a_n x + b_n) = H(x)$$

Le résultat qui suit, qu'on peut considérer comme une autre définition du max-domaine d'attraction d'une distribution des valeurs extrêmes, est très important en applications :

Proposition 2.3.1 [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

$F \in \text{MDA}(H)$ avec des constantes de normalisations $a_n \in \mathbb{R}_+^*$ et $b_n \in \mathbb{R}$ **si et seulement si** :

$$\forall x \in \mathbb{R}, \quad \lim_{n \rightarrow \infty} n[1 - F(a_n x + b_n)] = -\ln H(x) \quad (2.7)$$

Remarque : Quand $H(x) = 0$, la limite dans (2.7) s'interprète comme la limite de $x \mapsto -\ln x$ quand x tend vers 0 qui égale à ∞ .

Exemple 2.3.1 [RESNICK (1987)].

Supposons que X suit une loi de probabilité de PARÉTO, i.e :

$$\forall x \in \mathbb{R}, \quad 1 - F(x) = \begin{cases} x^{-\alpha} & \text{si } x > 1 \\ 1 & \text{si } x \leq 1 \end{cases} \quad \text{avec } \alpha > 0$$

Si nous posons $a_n = n^{1/\alpha}$ et $b_n = 0$, nous aurons :

$$\forall x \in \mathbb{R}, \quad \forall n \in \mathbb{N}^*, \quad n[1 - F(a_n x + b_n)] = n[1 - F(n^{1/\alpha} x)] = \begin{cases} x^{-\alpha} & \text{si } x > n^{-1/\alpha} \\ n & \text{si } x \leq n^{-1/\alpha} \end{cases}$$

Alors :

$$\begin{aligned} \forall x \in \mathbb{R}, \quad \lim_{n \rightarrow \infty} n[1 - F(a_n x + b_n)] &= \begin{cases} x^{-\alpha} & \text{si } x > 0 \\ \infty & \text{si } x \leq 0 \end{cases} \\ &= \begin{cases} -\ln[\exp(-x^{-\alpha})] & \text{si } x > 0 \\ -\ln 0 & \text{si } x \leq 0 \end{cases} \\ &= -\ln \Phi_\alpha(x) \end{aligned}$$

Donc :

$$F \in \text{MDA}(\Phi_\alpha)$$

Avant de donner les conditions nécessaires et/ou suffisantes sur la variable aléatoire X pour qu'elle appartienne au max-domaine d'attraction de H , on va donner des définitions de quelques outils nécessaires pour cela :

Définition 2.3.2 (Fonction de survie ou fonction de queue).

On appelle **fonction de survie** ou **fonction de queue** de la variable aléatoire X ou de la loi de probabilité de X , la fonction $\bar{F} : \mathbb{R} \rightarrow [0, 1]$ définie par :

$$\forall x \in \mathbb{R}, \quad \bar{F}(x) = 1 - F(x)$$

Définition 2.3.3 (Fonction de hasard).

Si la variable aléatoire X est absolument continue de densité f , alors on appelle **fonction du hasard** de X ou de la loi de probabilité de X , la fonction $r : \mathbb{R} \rightarrow \mathbb{R}$ définie par :

$$\forall x \in \mathbb{R}, \quad r(x) = \frac{f(x)}{\bar{F}(x)} = \frac{f(x)}{1 - F(x)}$$

Définition 2.3.4 (Inverse généralisée) [RESNICK (2006)].

Soit $G : \mathbb{R} \rightarrow]a, b[$ ($-\infty \leq a < b \leq +\infty$) une fonction croissante. On appelle **inverse généralisée** de la fonction G , la fonction $G^{\leftarrow} :]a, b[\rightarrow \mathbb{R}$ définie par :

$$\forall y \in]a, b[, \quad G^{\leftarrow}(y) = \inf \{x \in \mathbb{R}, G(x) \geq y\}$$

Avec la convention : $\inf \emptyset = \infty$.

Remarque : Dans le cadre de la définition ci-dessus, l'inverse généralisée G^{\leftarrow} coïncide avec l'inverse classique G^{-1} lorsque G est strictement croissante.

Proposition 2.3.2 (Propriétés de la fonction inverse généralisée) [RESNICK (1987)].

Soient $G : \mathbb{R} \rightarrow]a, b[$ ($-\infty \leq a < b \leq +\infty$) une fonction croissante et G^{\leftarrow} son inverse généralisée, alors G^{\leftarrow} est une fonction continue à gauche. **Si**, en plus, G est continue à droite, **alors** on a :

1.

$$G(x) \geq y \iff G^{\leftarrow}(y) \leq x$$

2.

$$G(x) < y \iff G^{\leftarrow}(y) > x$$

3.

$$\forall y \in]a, b[, \quad G[G^{\leftarrow}(y)] \geq y \quad \text{avec égalité si } G \text{ est continue.}$$

4.

$$\forall x \in \mathbb{R}, \quad G^{\leftarrow}[G(x)] \leq x \quad \text{avec égalité si } G \text{ est strictement croissante.}$$

5. Si Y est une variable aléatoire de fonction de répartition G alors les variables aléatoires Y et $G^{\leftarrow}[G(Y)]$ sont presque sûrement égales, *i.e* :

$$P\{G^{\leftarrow}[G(Y)] \neq Y\} = 0$$

Définition 2.3.5 (Fonction quantile) [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

On appelle **fonction quantile** de la variable aléatoire X ou de la loi de probabilité de X , la fonction Q définie par :

$$\forall p \in]0, 1[, \quad Q(p) = F^{\leftarrow}(p) = \inf \{x \in \mathbb{R}, F(x) \geq p\}$$

Remarque : Dans le cadre de la définition ci-dessus, la quantité $x_p = Q(p)$ définit le p -quantile de la variable aléatoire X .

Proposition 2.3.3 [DE HAAN & FERREIRA (2006)].

Soient $\{Y_n, n \in \mathbb{N}^*\}$ une suite de variables aléatoires et Y une variable aléatoire telles que G_n est la fonction de distribution de Y_n et G la fonction de distribution de Y . Si la suite $\{Y_n, n \in \mathbb{N}^*\}$ converge en distribution (en loi) vers Y , alors :

$$\forall y \in \mathcal{C}(G^{\leftarrow}) = \{y \in]0, 1[, G^{\leftarrow} \text{ continue en } y\}, \quad \lim_{n \rightarrow \infty} G_n^{\leftarrow}(y) = G^{\leftarrow}(y)$$

Définition 2.3.6 (Fonction quantile de queue).

On appelle **fonction quantile de queue** de la variable aléatoire X ou de la loi de probabilité de X , la fonction $U :]1, +\infty[\rightarrow \mathbb{R}$ définie par :

$$\forall y \in]1, +\infty[, \quad U(y) = Q\left(1 - \frac{1}{y}\right) = \left(\frac{1}{1 - F}\right)^{\leftarrow}(y)$$

où : Q est la fonction quantile de X .

Remarque : Avant de lire les résultats exposés dans cette section, on conseille le lecteur de voir ANNEXE B où on a exposé, en détail, la théorie de variation régulière qui est un outil important pour ces résultats.

2.3.1 Max-domaine d'attraction de FRÉCHET (MDA(Φ_α))

Dans cette section, on cherche à caractériser le max-domaine d'attraction de FRÉCHET, c'est à dire trouver les conditions nécessaires et/ou suffisantes sur la fonction de répartition F de la variable aléatoire X pour qu'elle appartienne à MDA(Φ_α), où Φ_α est la distribution des valeurs extrêmes standard de FRÉCHET définie comme suit :

$$\forall x \in \mathbb{R}, \quad \Phi_\alpha(x) = \begin{cases} \exp(-x^{-\alpha}) & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases}, \quad \alpha > 0$$

Remarque : On peut écrire Φ_α sous une forme plus générale, qu'on note $\Phi_{\alpha, \mu, \sigma}$, en faisant introduire un paramètre de localisation $\mu \in \mathbb{R}$ et un paramètre d'échelle $\sigma > 0$ comme suit :

$$\forall x \in \mathbb{R}, \quad \Phi_{\alpha, \mu, \sigma}(x) = \Phi_\alpha\left(\frac{x - \mu}{\sigma}\right) = \begin{cases} \exp\left\{-\left(\frac{x - \mu}{\sigma}\right)^{-\alpha}\right\} & \text{si } x > \mu \\ 0 & \text{si } x \leq \mu \end{cases}, \quad \alpha > 0$$

Et on appelle $\Phi_{\alpha, \mu, \sigma}$ « distribution des valeurs extrêmes de FRÉCHET ».

A. Propriétés de Φ_α ($\alpha > 0$)

A.1 Densité

Proposition 2.3.4

Φ_α est une fonction de répartition absolument continue de fonction de densité ϕ_α définie par :

$$\forall x \in \mathbb{R}, \quad \phi_\alpha(x) = \begin{cases} \alpha x^{-(\alpha+1)} \exp(-x^{-\alpha}) & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases}, \quad \alpha > 0$$

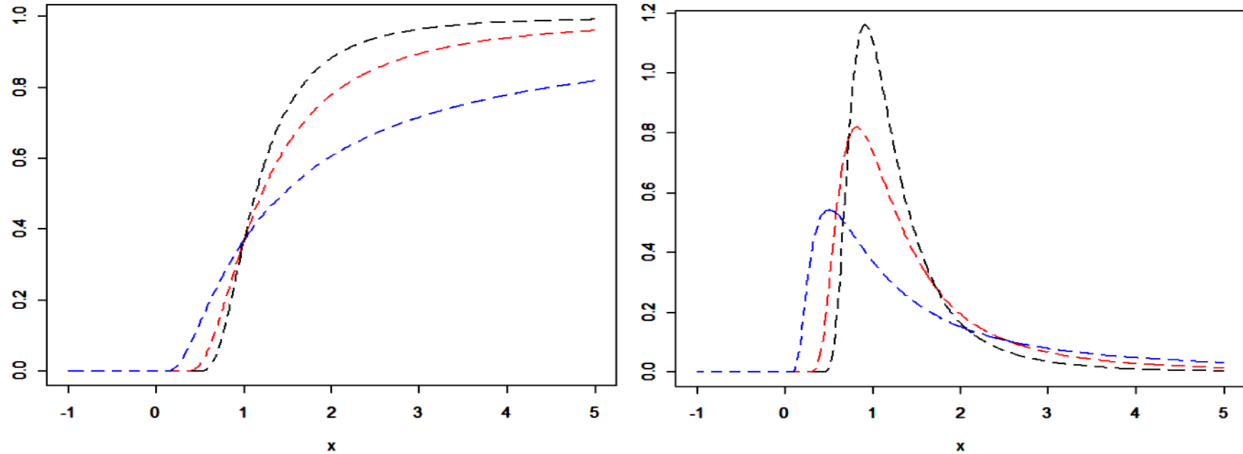


Figure 2.3.1 Panneau gauche représente les graphes de la distribution des valeurs extrêmes standard Φ_α selon les valeurs de α et le panneau droit représente les graphes de leurs densités :
Le bleu pour $\alpha = 1$, Le rouge pour $\alpha = 2$ et Le noir pour $\alpha = 3$.

A.2 Moments

Proposition 2.3.5 [BALAKRISHNAN & NEVZOROV (2003)].

Soit Y la variable aléatoire extrême associée à Φ_α , i.e Φ_α est la fonction de répartition de Y , alors :

$$\forall p \in \mathbb{R}_+^*, \quad E(Y^p) = \begin{cases} \Gamma\left(1 - \frac{p}{\alpha}\right) & \text{si } p < \alpha \\ \infty & \text{si } p > \alpha \end{cases}$$

$0 < \alpha < 1$	$1 < \alpha < 2$	$\alpha > 2$
$E(Y) = \infty$	$E(Y) = \Gamma\left(1 - \frac{1}{\alpha}\right)$	$E(Y) = \Gamma\left(1 - \frac{1}{\alpha}\right)$
$V(Y) = \infty$	$V(Y) = \infty$	$V(Y) = \Gamma\left(1 - \frac{2}{\alpha}\right) - \Gamma^2\left(1 - \frac{1}{\alpha}\right)$

Table 2.3.1 Espérance et variance de la variable aléatoire extrême Y associée à Φ_α selon les valeurs de α .

Remarque : Dans le cadre de la proposition ci-dessus, Γ est la fonction Gamma définie par :

$$\forall x > 0, \quad \Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt$$

A.3 Queue lourde

L'une des caractéristiques de la distribution des valeurs extrêmes de FRÉCHET est que sa queue est lourde à droite, ce qui signifie que les événements extrêmes d'une loi de probabilité qui appartient au max-domaine d'attraction de FRÉCHET ont une probabilité « importante » d'apparition (selon les valeurs de α). En effet, on a la proposition suivante :

Proposition 2.3.6 [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

$$\forall \alpha > 0, \quad \bar{\Phi}_\alpha(x) = 1 - \Phi_\alpha(x) = 1 - \exp(-x^{-\alpha}) \sim x^{-\alpha} \quad \text{quand } x \rightarrow \infty$$

Remarque : De la proposition ci-dessus, on déduit que la queue de Φ_α est une fonction à variation régulière à l'infini d'indice $-\alpha$. Et, en plus, **Figure 2.3.1** nous montre que plus α est petit, plus la queue de Φ_α s'alourdit à droite.

B. Caractérisation du max-domaine d'attraction de FRÉCHET ($\text{MDA}(\Phi_\alpha)$)

Maintenant, revenons à la question principale de cette section qui consiste à trouver les conditions nécessaires et/ou suffisantes sur la fonction de répartition F de la variable aléatoire X ainsi que les constantes de normalisation $a_n > 0$ et $b_n \in \mathbb{R}$ pour que F appartienne au max-domaine d'attraction de Φ_α ($\alpha > 0$).

Théorème 2.3.1 [RESNICK (1987)].

La fonction de distribution F de la variable aléatoire X appartient au max-domaine d'attraction de Φ_α ($\alpha > 0$) si et seulement si son point terminal x_F est infini et sa fonction de queue est à variation régulière à l'infini d'indice $(-\alpha)$, i.e :

$$F \in \text{MDA}(\Phi_\alpha) (\alpha > 0) \iff x_F = \infty \quad \text{et} \quad \bar{F} = 1 - F \in \text{RV}_{-\alpha}$$

Proposition 2.3.7 [DE HAAN & FERREIRA (2006)].

Si $F \in \text{MDA}(\Phi_\alpha)$ ($\alpha > 0$), alors les constantes de normalisations $a_n > 0$ et $b_n \in \mathbb{R}$ telles que :

$$\forall x \in \mathbb{R}, \quad \lim_{n \rightarrow \infty} F^n(a_n x + b_n) = \Phi_\alpha(x)$$

peuvent être choisies de la manière suivante :

$$a_n = U(n) \quad \text{et} \quad b_n = 0$$

où : U est la fonction quantile de queue de la variable aléatoire X .

Proposition 2.3.8 (Condition de VON MISES) [RESNICK (1987)].

Supposons que la variable aléatoire X est absolument continue de fonction de densité f et $\alpha > 0$, alors on a :

1.

$$\lim_{x \rightarrow \infty} \frac{xf(x)}{1 - F(x)} = \alpha \implies F \in \text{MDA}(\Phi_\alpha)$$

2. *Si, de plus, f est décroissante sur $]b, +\infty[$ ($b \in \mathbb{R}$), alors :*

$$F \in \text{MDA}(\Phi_\alpha) \implies \lim_{x \rightarrow \infty} \frac{xf(x)}{1 - F(x)} = \alpha$$

C. Exemples

Dans le tableau qui suit, on présente quelques exemples de lois de probabilité qui appartiennent au max-domaine d'attraction de FRÉCHET $\text{MDA}(\Phi_\alpha)$ avec $\alpha > 0$:

Distribution de X	Fonction de queue \bar{F}	Constantes de normalisation
$X \sim \text{CAUCHY}(0, 1)$ $\alpha = 1$	$\bar{F}(x) = \frac{1}{2} - \frac{1}{\pi} \arctan x$ $x \in \mathbb{R}$	$a_n = \pi^{-1}n$ $b_n = 0$
$X \sim \text{PARÉTO}(\alpha)$	$\bar{F}(x) = x^{-\alpha}$ $x > 1$	$a_n = n^{1/\alpha}$ $b_n = 0$
$X \sim \text{BURR}(\alpha, \beta, \tau)$ $\beta > 0, \tau > 0$	$\bar{F}(x) = \left(\frac{\beta}{\beta + x^\tau}\right)^{\alpha/\tau}$ $x > 0$	$a_n = n^{1/\alpha}$ $b_n = 0$
$X \sim \mathcal{S}_\alpha(\beta, \mu, \sigma)$ $\alpha < 2, \beta \in [-1, 1]$ $\mu \in \mathbb{R}, \sigma > 0$	$\bar{F}(x) \sim Cx^{-\alpha}$ quand $x \rightarrow \infty$ $C > 0$	$a_n = (Cn)^{1/\alpha}$ $b_n = 0$
$X \sim \text{Loggamma}(\alpha, \beta)$ $\beta > 0$	$\bar{F}(x) = \frac{\alpha^\beta}{\Gamma(\beta)} \int_x^{+\infty} (\ln t)^{\beta-1} t^{-\alpha-1} dt$ $x > 1$	$a_n = \left[(\Gamma(\beta))^{-1} (\ln n)^{\beta-1} n \right]^{1/\alpha}$ $b_n = 0$

Table 2.3.2 Exemples de lois de probabilité qui appartiennent au max-domaine d'attraction de FRÉCHET ($\text{MDA}(\Phi_\alpha)$).

2.3.2 Max-domaine d'attraction de WEIBULL (MDA(Ψ_α))

Dans cette section, on cherche à caractériser le max-domaine d'attraction de WEIBULL, c'est à dire trouver les conditions nécessaires et/ou suffisantes sur la fonction de répartition F de la variable aléatoire X pour qu'elle appartienne à MDA(Ψ_α), où Ψ_α est la distribution des valeurs extrêmes standard de WEIBULL définie comme suit :

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

Remarque : On peut donner une forme plus générale à Ψ_α , qu'on note $\Psi_{\alpha,\mu,\sigma}$, en ajoutant un paramètre de localisation $\mu \in \mathbb{R}$ et un paramètre d'échelle $\sigma > 0$ comme suit :

$$\forall x \in \mathbb{R}, \quad \Psi_{\alpha,\mu,\sigma}(x) = \Psi_\alpha\left(\frac{x-\mu}{\sigma}\right) = \begin{cases} \exp\left\{-\left(\frac{\mu-x}{\sigma}\right)^\alpha\right\} & \text{si } x \leq \mu \\ 1 & \text{si } x > \mu \end{cases}, \quad \alpha > 0$$

On appelle $\Psi_{\alpha,\mu,\sigma}$ « distribution des valeurs extrêmes de WEIBULL ».

A. Propriétés de Ψ_α ($\alpha > 0$)

A.1 Densité

Proposition 2.3.9

Ψ_α est une fonction de répartition absolument continue de fonction de densité ψ_α définie par :

$$\forall x \in \mathbb{R}, \quad \psi_\alpha(x) = \begin{cases} \alpha(-x)^{\alpha-1} \exp\{-(-x)^\alpha\} & \text{si } x \leq 0 \\ 0 & \text{si } x > 0 \end{cases}, \quad \alpha > 0$$

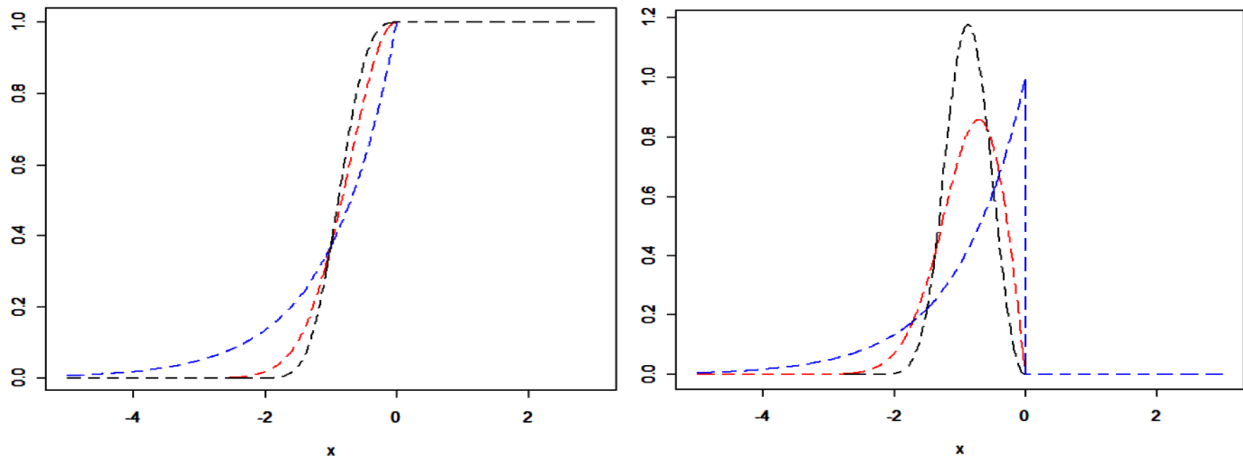


Figure 2.3.2 Panneau gauche représente les graphes de la distribution des valeurs extrêmes standard Ψ_α selon les valeurs de α et le panneau droit représente les graphes de leurs densités :
Le bleu pour $\alpha = 1$, Le rouge pour $\alpha = 2$ et Le noir pour $\alpha = 3$.

A.2 Moments

Proposition 2.3.10 [BALAKRISHNAN & NEVZOROV (2003)].

Soit Y la variable aléatoire extrême associée à Ψ_α , i.e Ψ_α est la fonction de répartition de Y , alors :

$$\forall p \in \mathbb{R}_+^*, \quad \mathbb{E}(Y^p) = (-1)^p \Gamma\left(1 + \frac{p}{\alpha}\right)$$

et, en particulier, l'espérance et la variance de Y sont données comme suit :

$$\mathbb{E}(Y) = -\Gamma\left(1 + \frac{1}{\alpha}\right) \quad \text{et} \quad \mathbb{V}(Y) = \Gamma\left(1 + \frac{2}{\alpha}\right) - \Gamma^2\left(1 + \frac{1}{\alpha}\right)$$

B. Caractérisation du max-domaine d'attraction de WEIBULL ($\text{MDA}(\Psi_\alpha)$)

Revenons à la question qui consiste à trouver les conditions nécessaires et/ou suffisantes que la fonction de répartition F de la variable aléatoire X doit satisfaire pour qu'elle appartienne à $\text{MDA}(\Psi_\alpha)$ ($\alpha > 0$).

Théorème 2.3.2 [RESNICK (1987)].

La fonction de distribution F de la variable aléatoire X appartient au max-domaine d'attraction de Ψ_α ($\alpha > 0$) **si et seulement si** son point terminal x_F est fini et la fonction $x \mapsto \bar{F}(x_F - x^{-1})$ est à variation régulière à l'infini d'indice $(-\alpha)$, i.e :

$$F \in \text{MDA}(\Psi_\alpha) (\alpha > 0) \iff x_F < \infty \quad \text{et} \quad x \mapsto \bar{F}(x_F - x^{-1}) = 1 - F(x_F - x^{-1}) \in \text{RV}_{-\alpha}$$

Proposition 2.3.11 [DE HAAN & FERREIRA (2006)].

Si $F \in \text{MDA}(\Psi_\alpha)$ ($\alpha > 0$), **alors** les constantes de normalisations $a_n > 0$ et $b_n \in \mathbb{R}$ telles que :

$$\forall x \in \mathbb{R}, \quad \lim_{n \rightarrow \infty} F^n(a_n x + b_n) = \Psi_\alpha(x)$$

peuvent être choisies de la manière suivante :

$$a_n = x_F - U(n) \quad \text{et} \quad b_n = x_F$$

où : U est la fonction quantile de queue de la variable aléatoire X .

Proposition 2.3.12 (Condition de VON MISES) [RESNICK (1987)].

Supposons que la variable aléatoire X est absolument continue de fonction de densité f telle que f est positive sur $]a, x_F[$ ($a < x_F$) et soit $\alpha > 0$, **alors** on a :

1.

$$\lim_{x \uparrow x_F} \frac{(x_F - x)f(x)}{1 - F(x)} = \alpha \implies F \in \text{MDA}(\Psi_\alpha)$$

2. **Si**, de plus, f est décroissante sur $]b, x_F[$ ($a \leq b < x_F$), **alors** :

$$F \in \text{MDA}(\Psi_\alpha) \implies \lim_{x \uparrow x_F} \frac{(x_F - x)f(x)}{1 - F(x)} = \alpha$$

C. Exemples

Notons que les lois de probabilité qui appartiennent au max-domaine d'attraction de WEIBULL sont des lois bornées à droite et, donc, elles ont peu d'intérêt dans l'étude des événements extrêmes. Dans le tableau qui suit, on présente quelques exemples de ces lois avec $\alpha > 0$:

Distribution de X	Fonction de queue \bar{F}	Constantes de normalisation
$X \sim \text{Béta}(\beta, \alpha)$ $\beta > 0$	$\bar{F}(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_x^1 t^{\beta-1}(1-t)^{\alpha-1} dt$ $0 < x < 1$	$a_n = \left[\frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha + 1)\Gamma(\beta)} n \right]^{-1/\alpha}$ $b_n = 1$
$X \sim \mathcal{U}([0, 1])$ $\alpha = 1$	$\bar{F}(x) = 1 - x$ $0 \leq x \leq 1$	$a_n = n^{-1}$ $b_n = 1$

Table 2.3.3 Exemples de lois de probabilité qui appartiennent au max-domaine d'attraction de WEIBULL (MDA(Ψ_α)).

2.3.3 Max-domaine d'attraction de GUMBEL (MDA(Λ))

Dans cette section, on cherche à caractériser le max-domaine d'attraction de GUMBEL, c'est à dire trouver les conditions nécessaires et/ou suffisantes sur la fonction de répartition F de la variable aléatoire X ainsi que les constantes de normalisations $a_n > 0$ et $b_n \in \mathbb{R}$ pour que F appartienne au max-domaine d'attraction de Λ , où Λ est la distribution des valeurs extrêmes standard de GUMBEL définie comme suit :

$$\forall x \in \mathbb{R}, \quad \Lambda(x) = \exp(-e^{-x})$$

Remarque : Une forme plus générale à Λ , qu'on note $\Lambda_{\mu, \sigma}$, est donnée en introduisant un paramètre de localisation $\mu \in \mathbb{R}$ et un paramètre d'échelle $\sigma > 0$ comme suit :

$$\forall x \in \mathbb{R}, \quad \Lambda_{\mu, \sigma}(x) = \Lambda\left(\frac{x - \mu}{\sigma}\right) = \exp\left\{-\exp\left(\frac{\mu - x}{\sigma}\right)\right\}$$

Et on appelle $\Psi_{\alpha, \mu, \sigma}$ « **distribution des valeurs extrêmes de GUMBEL** ».

A. Propriétés de Λ

A.1 Densité

Proposition 2.3.13

Λ est une fonction de répartition absolument continue de fonction de densité λ définie par :

$$\forall x \in \mathbb{R}, \quad \lambda(x) = \exp\{- (x + e^{-x})\}$$

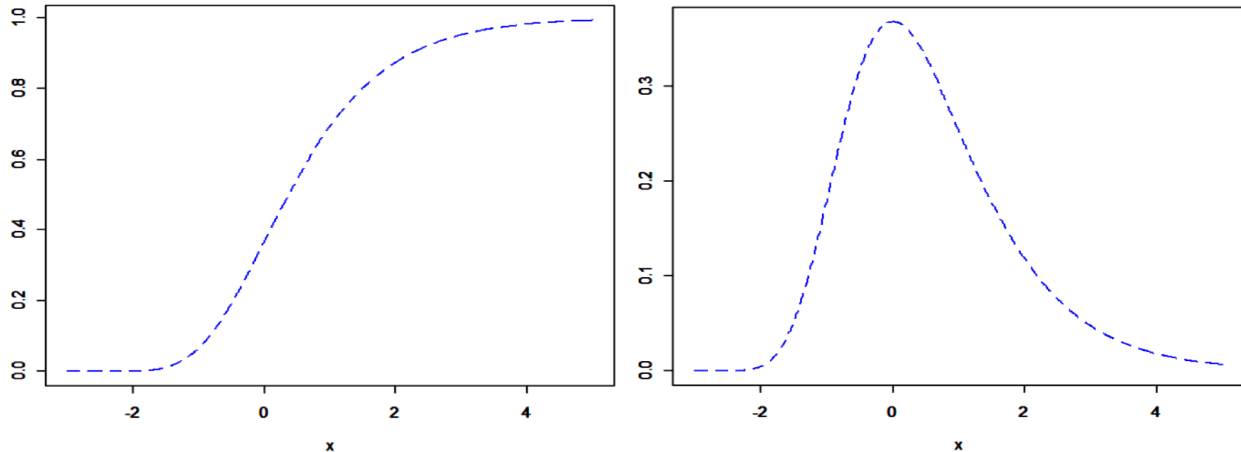


Figure 2.3.3 Panneau gauche représente le graphe de la distribution des valeurs extrêmes standard Λ et le panneau droit représente le graphe de sa densité λ .

A.2 Moments

Proposition 2.3.14 [BALAKRISHNAN & NEVZOROV (2003)].

Soit Y la variable aléatoire extrême associée à Λ , alors :

$$\forall k \in \mathbb{N}^*, \quad \mathbb{E}(Y^k) = (-1)^k \Gamma^{(k)}(1)$$

où : $\Gamma^{(k)}$ dénote la $k^{\text{ème}}$ dérivée de la fonction Γ . En particulier, l'espérance et la variance de Y sont données par :

$$\mathbb{E}(Y) = \gamma \quad \text{et} \quad \mathbb{V}(Y) = \frac{\pi^2}{6}$$

avec : γ est la constante d'EULER donnée par : $\gamma = -\int_0^{+\infty} e^{-t} \ln t \, dt \cong 0.57722\dots$

A.3 Queue légère

Il n'existe pas une caractérisation type des lois de probabilité aux queues légères. La notion caractérise généralement la qualité de la déviation d'une loi de probabilité par rapport à la loi normale dont la fonction de queue décroît exponentiellement.

Définition 2.3.7 [BARRO (2010) et BOUZIANE (2011)].

Soit Z une variable aléatoire à valeurs dans \mathbb{R} de fonction de distribution Υ . On dit que Z ou la loi de probabilité de Z ou, encore, la fonction de répartition Υ de Z est à queue légère à droite s'il existe $b \in \mathbb{R}$ tel que :

$$\forall x \in]b, +\infty[, \quad \exists \eta > 0 \quad \text{et} \quad \kappa > 0 \quad \text{tels que} \quad \bar{\Upsilon}(x) = 1 - \Upsilon(x) \leq \eta e^{-\kappa x}$$

Remarque : Parmi les lois de probabilité aux queues légères, on trouve celles qui ont une fonction de queue qui varie rapidement à l'infini d'indice $-\infty$ (voir ANNEXE B).

Les événements extrêmes d'une loi de probabilité qui appartient au max-domaine d'attraction de GUMBEL ont une probabilité « faible » d'apparition. Cela se traduit par le fait que la distribution des valeurs extrêmes de GUMBEL est une loi de probabilité à queue légère à droite. En effet, on a la proposition suivante :

Proposition 2.3.15 [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

$$\bar{\Lambda}(x) = 1 - \Lambda(x) = 1 - \exp(-e^{-x}) \sim e^{-x} \quad \text{quand } x \rightarrow \infty$$

Remarque : De la proposition ci-dessus, on déduit que la queue de Λ est une fonction à variation rapide à l'infini d'indice $-\infty$.

B. Caractérisation du max-domaine d'attraction de GUMBEL (MDA(Λ))

Le max-domaine d'attraction de GUMBEL est le plus difficile à caractériser car il n'existe pas une condition nécessaire et suffisante, relativement simple, pour que la fonction de répartition F de la variable aléatoire X soit dedans. Généralement, les auteurs présentent une condition suffisante basée sur la représentation d'une fonction de VON MISES qui est difficile à manipuler, tout comme la condition nécessaire et suffisante formulée par GNEDENKO en 1943.

Théorème 2.3.3 [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

La fonction de distribution F de la variable aléatoire X avec le point terminal $x_F \leq \infty$ appartient au max-domaine d'attraction de Λ **si et seulement** s'il existe $a < x_F$ tel que :

$$\forall x \in]a, x_F[, \quad \bar{F}(x) = 1 - F(x) = c(x) \exp \left\{ - \int_a^x \frac{\theta(u)}{\delta(u)} du \right\}$$

où : c et θ sont des fonctions définies mesurables sur $]a, x_F[$ vérifiant

$$\lim_{x \uparrow x_F} c(x) = c > 0 \quad \text{et} \quad \lim_{x \uparrow x_F} \theta(x) = 1$$

et : δ est une fonction absolument continue sur $]a, x_F[$ de densité δ' telle que

$$\forall x \in]a, x_F[, \quad \delta(x) > 0 \quad \text{et} \quad \lim_{x \uparrow x_F} \delta'(x) = 0$$

Proposition 2.3.16 [DE HAAN & FERREIRA (2006)].

Si $F \in \text{MDA}(\Lambda)$, **alors** les constantes de normalisations $a_n > 0$ et $b_n \in \mathbb{R}$ telles que :

$$\forall x \in \mathbb{R}, \quad \lim_{n \rightarrow \infty} F^n(a_n x + b_n) = \Lambda(x)$$

peuvent être choisies de la manière suivante :

$$a_n = \delta(b_n) \quad \text{et} \quad b_n = U(n)$$

où : U est la fonction quantile de queue de la variable aléatoire X et δ la fonction définie dans le théorème précédent.

Définition 2.3.8 (Fonction de VON MISES) [RESNICK (1987)].

La fonction de répartition F de la variable aléatoire X qui a un point terminal x_F est appelée **fonction de VON MISES** s'il existe $a < x_F$ tel que

$$\forall x \in]a, x_F[, \quad \bar{F}(x) = 1 - F(x) = c \exp \left\{ - \int_a^x \frac{1}{\delta(u)} du \right\} \quad \text{avec } c > 0$$

où : δ est une fonction absolument continue sur $]a, x_F[$ de densité δ' telle que

$$\forall x \in]a, x_F[, \quad \delta(x) > 0 \quad \text{et} \quad \lim_{x \uparrow x_F} \delta'(x) = 0$$

Remarque : Dans le cadre de la définition ci-dessus, le point terminal x_F de F peut être fini ou infini, et on appelle δ « la fonction auxiliaire » de F .

Proposition 2.3.17 [RESNICK (1987)].

Si la variable aléatoire X est absolument continue de densité f dérivable sur $]a, x_F[$ ($a < x_F$) de dérivée f' telle que

$$\forall x \in]a, x_F[, \quad f'(x) < 0$$

alors : F est une fonction de VON MISES **si et seulement si**

$$\lim_{x \uparrow x_F} \frac{f'(x)[1 - F(x)]}{[f(x)]^2} = -1$$

et, dans ce cas, la fonction auxiliaire δ de F est donnée par :

$$\forall x \in]a, x_F[, \quad \delta(x) = \frac{1 - F(x)}{f(x)}$$

Proposition 2.3.18 [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

Si la fonction de répartition F de la variable aléatoire X est une fonction de VON MISES, **alors** elle est absolument continue sur $]a, x_F[$ ($a < x_F$) de densité positive f et les assertions suivantes sont vérifiées :

1. **Si** $x_F = \infty$, **alors :** $\bar{F} \in \text{RV}_{-\infty}$ et

$$\lim_{x \rightarrow \infty} \frac{xf(x)}{1 - F(x)} = \infty$$

2. **Si** $x_F < \infty$, **alors :** la fonction $x \mapsto \bar{F}(x_F - x^{-1}) \in \text{RV}_{-\infty}$ et

$$\lim_{x \uparrow x_F} \frac{(x_F - x)f(x)}{1 - F(x)} = \infty$$

Proposition 2.3.19 (Condition de VON MISES) [RESNICK (1987)].

Si la fonction de répartition F de la variable aléatoire X est une fonction de VON MISES, alors F appartient au max-domaine d'attraction de Λ .

Exemple 2.3.2 [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

Supposons que X suit une loi de probabilité exponentielle $\mathcal{E}(\lambda)$, i.e :

$$\forall x \in \mathbb{R}, \quad F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases}, \quad \lambda > 0$$

alors : F est une fonction de VON MISES.

Vérification :

En effet, on a : $x_F = \infty$ et

$$\forall x \in]0, +\infty[, \quad \bar{F}(x) = 1 - F(x) = e^{-\lambda x} = \exp \left\{ - \int_0^x \frac{1}{\lambda^{-1}} du \right\}$$

Donc : F est une fonction de VON MISES de fonction auxiliaire δ donnée par :

$$\forall x \in]0, +\infty[, \quad \delta(x) = \lambda^{-1}$$

On déduit que : $F \in \text{MDA}(\Lambda)$.

Exemple 2.3.3 [RESNICK (1987)].

Supposons que X suit une loi de probabilité normale standard $\mathcal{N}(0,1)$ de fonction de répartition Φ et de fonction de densité φ , i.e :

$$\forall x \in \mathbb{R}, \quad \Phi(x) = \int_{-\infty}^x \varphi(u) du = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-u^2/2} du$$

alors : Φ est une fonction de VON MISES.

Vérification :

On a : $x_F = +\infty$ et

$$\forall x \in]0, +\infty[, \quad \varphi'(x) = -\frac{1}{\sqrt{2\pi}} x e^{-x^2/2} = -x\varphi(x) < 0$$

Et :

$$1 - \Phi(x) \sim x^{-1}\varphi(x) \quad \text{quand } x \rightarrow \infty$$

Donc :

$$\lim_{x \rightarrow \infty} \frac{\varphi'(x)[1 - \Phi(x)]}{[\varphi(x)]^2} = \lim_{x \rightarrow \infty} \frac{-x\varphi(x)x^{-1}\varphi(x)}{[\varphi(x)]^2} = -1$$

On déduit que : Φ est une fonction de VON MISES, et par conséquent $\Phi \in \text{MDA}(\Lambda)$.

C. Exemples

Dans le tableau qui suit, on a résumé quelques exemples de lois de probabilité qui appartiennent au max-domaine d'attraction de GUMBEL $\text{MDA}(\Lambda)$:

Distribution de X	Fonction de queue \bar{F}	Constantes de normalisation
$X \sim \mathcal{N}(0, 1)$	$\bar{F}(x) = \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} e^{-t^2/2} dt$ $x \in \mathbb{R}$	$a_n = (2 \ln n)^{-1/2}$ $b_n = \sqrt{2 \ln n} - \frac{\ln(4\pi) + \ln \ln n}{2\sqrt{2 \ln n}}$
$X \sim \mathcal{E}(\lambda)$ $\lambda > 0$	$\bar{F}(x) = e^{-\lambda x}$ $x > 0$	$a_n = \lambda^{-1}$ $b_n = \lambda^{-1} \ln n$
$X \sim \text{Gamma}(\alpha, \beta)$ $\alpha > 0, \beta > 0$	$\bar{F}(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_x^{+\infty} t^{\alpha-1} e^{-\beta t} dt$ $x > 0$	$a_n = \beta^{-1}$ $b_n = \frac{\ln n + (\alpha - 1) \ln \ln n - \ln \Gamma(\alpha)}{\beta}$
$X \sim \text{Lognormale}(0, 1)$	$\bar{F}(x) = \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} \frac{e^{-(\ln t)^2/2}}{t} dt$ $x > 0$	$a_n = b_n (2 \ln n)^{-1/2}$ $b_n = \exp \left\{ \sqrt{2 \ln n} - \frac{\ln(4\pi) + \ln \ln n}{2\sqrt{2 \ln n}} \right\}$

Table 2.3.4 Exemples de lois de probabilité qui appartiennent au max-domaine d'attraction de GUMBEL ($\text{MDA}(\Lambda)$).

2.4 Distribution des valeurs extrêmes généralisée (GEVD)

Comme on vient de le voir, les trois types de distributions extrêmes FRÉCHET, WEIBULL et GUMBEL ont des comportements différents qui correspondent aux différents comportements de la fonction de queue \bar{F} de la variable aléatoire X . Cela a entraîné, dans les premières applications de la théorie des valeurs extrêmes, d'adopter l'un de ces trois types pour l'analyse des données. Mais cette méthode comporte des inconvénients car, en premier lieu, on doit avoir une technique pour choisir lequel des trois types de distributions extrêmes est plus approprié aux données qu'on a, et, deuxièmement, une fois une telle décision est prise, les déductions subséquentes confirment que notre choix est correct, et ne tient pas compte de l'incertitude qu'une telle sélection implique, bien que cette incertitude puisse être substantielle.

Une meilleure analyse est offerte grâce aux travaux de VON MISES (1936) et de JENKINSON (1955) qui ont montré que les trois types de distributions extrêmes FRÉCHET, WEIBULL et GUMBEL peuvent être combinées dans un seul type de distributions qu'on appelle « *type de distributions des valeurs extrêmes généralisées* » ou « *type de distributions des valeurs extrêmes de VON MISES - JENKINSON* ».

2.4.1 Représentation de VON MISES-JENKINSON

Définition 2.4.1 [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

Soit $\xi \in \mathbb{R}$. On appelle **distribution des valeurs extrêmes généralisée standard** toute fonction de répartition H_ξ ou toute loi de probabilité qui a H_ξ comme fonction de répartition telle que :

$$\forall x \in \mathbb{R} \quad \text{tel que} \quad 1 + \xi x > 0, \quad H_\xi(x) = \begin{cases} \exp \left\{ - (1 + \xi x)^{-1/\xi} \right\} & \text{si } \xi \neq 0 \\ \exp \left\{ - e^{-x} \right\} & \text{si } \xi = 0 \end{cases}$$

Pour $\xi \neq 0$ tel que $1 + \xi x \leq 0$, on peut deviner les valeurs de H_ξ puisque c'est une fonction de répartition.

Remarque : Pour balayer tout le type de distributions des valeurs extrêmes généralisées, on peut donner une forme plus générale à la fonction de répartition H_ξ donnée dans la définition ci-dessus, qu'on note $H_{\xi, \mu, \sigma}$, en lui introduisant un paramètre de position $\mu \in \mathbb{R}$ et un paramètre d'échelle $\sigma > 0$ comme suit :

$$\forall x \in \mathbb{R} \quad \text{tel que} \quad 1 + \xi \left(\frac{x - \mu}{\sigma} \right) > 0, \quad H_{\xi, \mu, \sigma}(x) = \begin{cases} \exp \left\{ - \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{-1/\xi} \right\} & \text{si } \xi \neq 0 \\ \exp \left\{ - \exp \left(\frac{\mu - x}{\sigma} \right) \right\} & \text{si } \xi = 0 \end{cases}$$

Et on appelle $H_{\xi, \mu, \sigma}$ « **distribution des valeurs extrêmes généralisée (GEVD¹)** ».

Notation : Dans toute la suite de ce chapitre, on note $H_{\xi, \mu, \sigma}$ la distribution des valeurs extrêmes généralisée et H_ξ sa version standard.

Dans la proposition qui suit, on va donner un résultat important, facile à vérifier, qui nous permettra de voir la relation qui existe entre la distribution des valeurs extrêmes généralisée et les distributions des valeurs extrêmes :

Proposition 2.4.1 [DE HAAN & FERREIRA (2006)].

Soient H_ξ ($\xi \in \mathbb{R}$) la distribution des valeurs extrêmes généralisée standard et Φ_α , Ψ_α , Λ les distributions des valeurs extrêmes standards avec $\alpha > 0$. On a :

$$\forall x \in \mathbb{R} \quad \text{tel que} \quad 1 + \xi x > 0, \quad H_\xi(x) = \begin{cases} \Phi_{1/\xi}(1 + \xi x) & \text{si } \xi > 0 \\ \Psi_{-1/\xi}[-(1 + \xi x)] & \text{si } \xi < 0 \\ \Lambda(x) & \text{si } \xi = 0 \end{cases}$$

En terme de type de distribution (voir ANNEXE A), la proposition ci-dessus nous donne un résultat très important, dans les applications de la théorie des valeurs extrêmes (**Méthode des maximums par blocs**), qui permette d'unifier les trois types de distributions extrêmes FRÉCHET, WEIBULL et GUMBEL dans un seul type qui est le type de distributions des valeurs extrêmes généralisées. En effet, on a la proposition suivante :

1. Generalized Extreme Value Distribution

Proposition 2.4.2 [MERAGHNI (2008)].

Si on note par H_ξ ($\xi \in \mathbb{R}$) le type de distributions des valeurs extrêmes généralisées et Φ_α , Ψ_α , Λ , avec $\alpha > 0$, les types de distributions des valeurs extrêmes qui sont, respectivement, de FRÉCHET, de WEIBULL et de GUMBEL, alors on a :

$$H_\xi = \begin{cases} \Phi_{1/\xi} & \text{si } \xi > 0 \\ \Psi_{-1/\xi} & \text{si } \xi < 0 \\ \Lambda & \text{si } \xi = 0 \end{cases}$$

2.4.2 Propriétés de la GEVD

A. Densité

Proposition 2.4.3

H_ξ ($\xi \in \mathbb{R}$) est une fonction de répartition absolument continue de densité h_ξ définie par :

$$\forall x \in \mathbb{R}, \quad h_\xi(x) = \begin{cases} (1 + \xi x)^{-\frac{1}{\xi}-1} \exp \left\{ - (1 + \xi x)^{-1/\xi} \right\} \cdot \mathbb{1}_{[1+\xi x > 0]}(x) & \text{si } \xi \neq 0 \\ \exp \left\{ - (x + e^{-x}) \right\} & \text{si } \xi = 0 \end{cases}$$

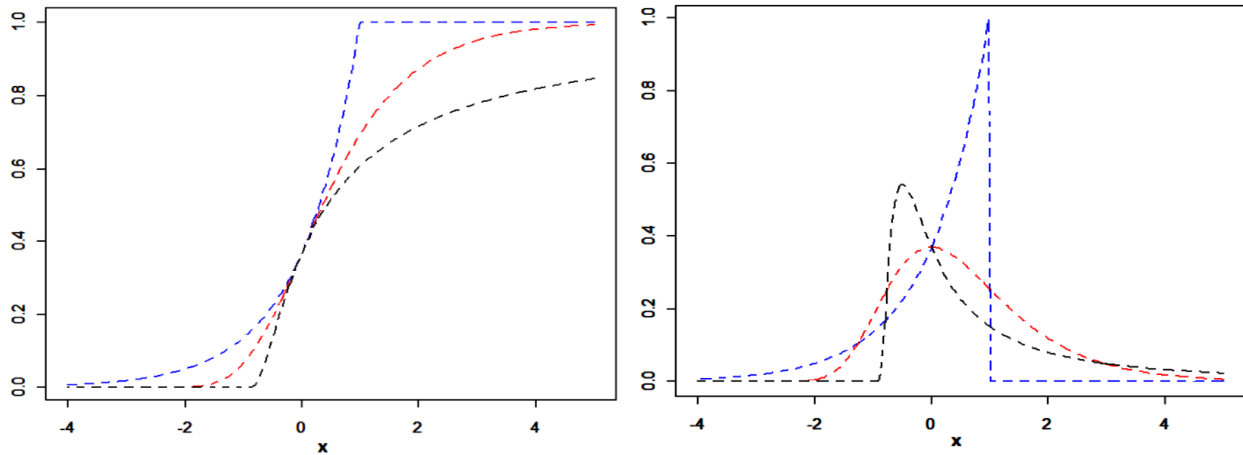


Figure 2.4.1 Panneau gauche représente les graphes de la distribution des valeurs extrêmes généralisée standard H_ξ selon les valeurs de ξ et le panneau droit représente les graphes de leurs densités : Le bleu pour $\xi = -1$, Le rouge pour $\xi = 0$ et Le noir pour $\xi = 1$.

B. Moments

Proposition 2.4.4

Soit Z la variable aléatoire associée à H_ξ , i.e H_ξ est la fonction de répartition de Z , alors l'espérance et la variance de Z sont données comme suit :

$$E(Z) = \begin{cases} \gamma & \text{si } \xi = 0 \\ \frac{\Gamma(1-\xi)}{\xi} & \text{si } \xi < 1 \\ \infty & \text{si } \xi > 1 \end{cases} \quad \text{et} \quad V(Z) = \begin{cases} \pi^2/6 & \text{si } \xi = 0 \\ \frac{\Gamma(1-2\xi) - \Gamma^2(1-\xi)}{\xi^2} & \text{si } \xi < 1/2 \\ \infty & \text{si } \xi > 1/2 \end{cases}$$

Remarque : Dans le cadre de la proposition ci-dessus, Γ est la fonction Gamma et γ est la constante d'EULER.

C. Quantiles

Proposition 2.4.5 [CASTILLO, HADI, BALAKRISHNAN & SARABIA (2005)].

Soient $p \in]0, 1[$ et Z la variable aléatoire associée à la distribution des valeurs extrêmes généralisée standard H_ξ ($\xi \in \mathbb{R}$), alors le p -quantile x_p de Z est donnée par :

$$x_p = \begin{cases} \frac{(-\ln p)^{-\xi} - 1}{\xi} & \text{si } \xi \neq 0 \\ -\ln(-\ln p) & \text{si } \xi = 0 \end{cases}$$

Remarque : Dans le cadre de la proposition ci-dessus, on voit bien que le quantile x_p est fortement influencé par le paramètre $\xi \in \mathbb{R}$, plus ξ est grand plus le quantile x_p est élevé. En applications, cela veut dire que plus ξ est grand plus la probabilité qu'un événement extrême se produise est importante.

2.4.3 Max-domaine d'attraction de la GEVD

A. Théorème de FISHER-TIPPETT (1928) & GNEDENKO (1943)

Les propriétés de la distribution des valeurs extrêmes généralisée nous permette d'avoir une version plus raffinée du **Théorème 2.2.1** qui est la suivante :

Théorème 2.4.1 (Théorème de FISHER-TIPPETT (1928) & GNEDENKO (1943)).

S'il existe des suites de normalisation $\{a_n, n \in \mathbb{N}^\}$ de nombres réels strictement positifs, $\{b_n, n \in \mathbb{N}^*\}$ de nombres réels et une fonction de distribution non dégénérée H , définie et continue sur \mathbb{R} , telles que :*

$$\lim_{n \rightarrow \infty} P\left(\frac{M_n - b_n}{a_n} \leq x\right) = \lim_{n \rightarrow \infty} F^n(a_n x + b_n) = H(x)$$

alors H appartient au type de distributions des valeurs extrêmes généralisées donné, à un paramètre d'échelle et de localisation près, par la distribution des valeurs extrêmes généralisée standard H_ξ ($\xi \in \mathbb{R}$) suivante :

$$\forall x \in \mathbb{R} \quad \text{tel que } 1 + \xi x > 0, \quad H_\xi(x) = \begin{cases} \exp\left\{- (1 + \xi x)^{-1/\xi}\right\} & \text{si } \xi \neq 0 \\ \exp\{-e^{-x}\} & \text{si } \xi = 0 \end{cases}$$

Démonstration.

Pour une démonstration du **Théorème 2.4.1**, le lecteur pourra se référer à l'ouvrage [DE HAAN & FERREIRA (2006)]. ■

Définition 2.4.2 [LEKINA (2010)].

Le paramètre $\xi \in \mathbb{R}$ qu'on voit dans la distribution des valeurs extrêmes généralisée est un paramètre de forme qu'on appelle « **indice des valeurs extrêmes** » ou « **indice de queue** ».

Remarque : On appelle le paramètre $\xi \in \mathbb{R}$ comme on l'a cité dans la définition ci-dessus car il mesure l'épaisseur de la queue de la distribution des valeurs extrêmes généralisée. Plus ξ est grand en valeur absolue, plus la queue de H_ξ est épaisse, et, par conséquent, le poids des événements extrêmes dans la distribution initiale F est important.

B. Caractérisation du max-domaine d'attraction de la GEVD ($\text{MDA}(H_\xi)$)

Si les conditions du **Théorème 2.4.1** sont vérifiées, on dira que la variable aléatoire X ou la fonction de distribution F de X appartient au max-domaine d'attraction de H , et on note $F \in \text{MDA}(H)$.

Théorème 2.4.2 [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

La fonction de distribution F de la variable aléatoire X appartient au max-domaine d'attraction de la distribution des valeurs extrêmes généralisée standard H_ξ ($\xi \in \mathbb{R}$) **si et seulement si** il existe une fonction strictement positive δ telle que :

$$\forall x > 0 \quad \text{tel que} \quad 1 + \xi x > 0, \quad \lim_{t \uparrow x_F} \frac{1 - F[t + x\delta(t)]}{1 - F(t)} = \begin{cases} (1 + \xi x)^{-1/\xi} & \text{si } \xi \neq 0 \\ e^{-x} & \text{si } \xi = 0 \end{cases} \quad (2.8)$$

Remarque : Dans le cadre du théorème ci-dessus, la relation (2.8) peut s'écrire de la manière suivante :

$$\forall x > 0 \quad \text{tel que} \quad 1 + \xi x > 0, \quad \lim_{t \uparrow x_F} \text{P}\left(\frac{X - t}{\delta(t)} > x \mid X > t\right) = \begin{cases} (1 + \xi x)^{-1/\xi} & \text{si } \xi \neq 0 \\ e^{-x} & \text{si } \xi = 0 \end{cases}$$

Ce qui nous permette de donner une approximation de la loi de probabilité de l'excès de X au-delà d'un seuil t lorsque ce seuil est élevé (proche du point terminal x_F). Cette interprétation est très utile dans les applications de la théorie des valeurs extrêmes, en particulier, dans une méthode qui s'appelle « **excès au-delà d'un seuil** ».

Une autre caractérisation du max-domaine d'attraction $\text{MDA}(H_\xi)$ est donnée dans la théorème suivant :

Théorème 2.4.3 [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

La fonction de distribution F de la variable aléatoire X appartient au max-domaine d'attraction de la distribution des valeurs extrêmes généralisée standard H_ξ ($\xi \in \mathbb{R}$) **si et seulement si** :

$$\forall x > 0, \quad \forall y > 0 \quad \text{tel que} \quad y \neq 1, \quad \lim_{t \rightarrow \infty} \frac{U(tx) - U(t)}{U(ty) - U(t)} = \begin{cases} \frac{x^\xi - 1}{y^\xi - 1} & \text{si } \xi \neq 0 \\ \frac{\ln x}{\ln y} & \text{si } \xi = 0 \end{cases}$$

où : U est la fonction quantile de queue de la variable aléatoire X .

2.5 Distribution de PARÉTO généralisée (GPD)

Comme on l'a mentionné dans la section précédente, la distribution des valeurs extrêmes généralisée est très utile en applications de la théorie des valeurs extrêmes, car c'est la seule et unique loi de probabilité qui modélise le comportement du maximum d'un échantillon. Pour estimer ses paramètres, les statisticiens ont souvent recours à une méthode qui s'appelle « **maximums par blocs** » qui consiste à construire un échantillon de maximums à partir d'un échantillon de données en formant des blocs de même dimension. Cette méthode a un inconvénient majeur qui entraîne une perte de certaines informations, en particulier, certains blocs peuvent contenir plusieurs valeurs extrêmes, alors que d'autres blocs peuvent ne pas en contenir.

Pour remédier à l'inconvénient de la méthode des maximums par blocs, les statisticiens utilisent une autre méthode qui permet de prendre en compte beaucoup plus de données pour assurer beaucoup plus de précision dans l'estimation des paramètres de la distribution des valeurs extrêmes généralisée, en particulier, l'indice des valeurs extrêmes $\xi \in \mathbb{R}$. Cette méthode qu'on appelle « **excès au-delà d'un seuil (POT²)** », basée sur la distribution de PARÉTO généralisée (GPD³), consiste à étudier le comportement non pas du maximum des données qu'on a en main, mais de toutes les données qui dépassent un seuil élevé u , et plus précisément, les différences entre ces données et le seuil u , appelées « **excès** ».

2.5.1 Distribution des excès

Définition 2.5.1 [LEKINA (2010)].

On appelle *excès* de la variable aléatoire X **au-delà d'un seuil** $u < x_F$ la variable aléatoire Y , qui prend ses valeurs sur $]0, x_F - u[$, définie par :

$$Y = X - u \mid X > u, \quad u < x_F$$

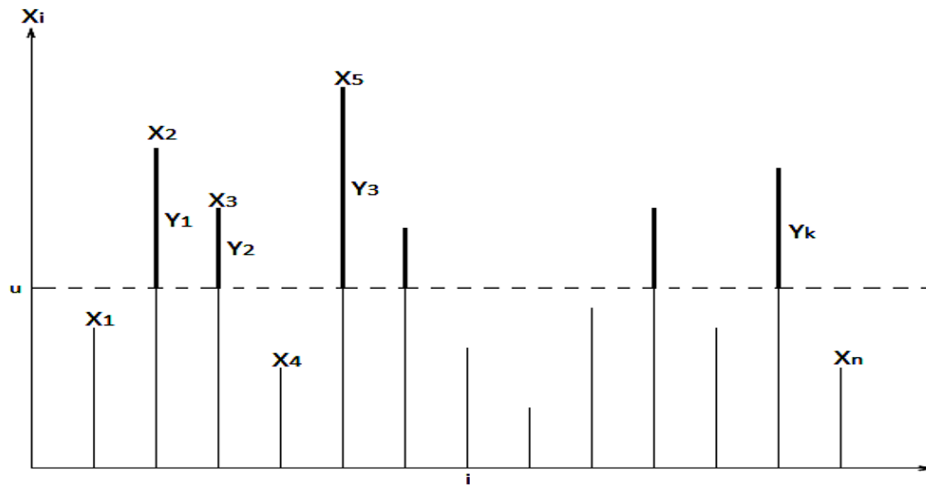


Figure 2.5.1 Les données X_1, X_2, \dots, X_n et leurs k excès au-delà du seuil u correspondants Y_1, Y_2, \dots, Y_k ($k \leq n$).

2. Peaks Over Threshold
3. Generalized PARÉTO Distribution

Définition 2.5.2 (Distribution des excès) [LEKINA (2010)].

On appelle **distribution des excès** de la variable aléatoire X par rapport à un seuil $u < x_F$ la loi de probabilité de la variable aléatoire Y excès de X au-delà du seuil $u < x_F$, donnée par sa fonction de répartition F_u , qu'on appelle **fonction de distribution des excès**, suivante :

$$\forall y \in \mathbb{R}, \quad F_u(y) = P(X - u \leq y \mid X > u) = \begin{cases} 0 & \text{si } y \leq 0 \\ 1 - \frac{1 - F(u + y)}{1 - F(u)} & \text{si } 0 < y < x_F - u \\ 1 & \text{si } y \geq x_F - u \end{cases}$$

Notation : Dans toute la suite de ce chapitre, on note F_u la fonction de distribution des excès donnée dans la définition ci-dessus.

2.5.2 Distribution de PARÉTO généralisée

Comme on l'a mentionné dans la section précédente, la relation (2.8) joue un rôle important dans les applications de la théorie des valeurs extrêmes. En effet, elle nous confirme que si la fonction de distribution F de la variable aléatoire X est dans le domaine d'attraction de la distribution des valeurs extrêmes généralisée H_ξ ($\xi \in \mathbb{R}$), alors quand le seuil u est très élevé (proche du point terminal x_F) on peut rapprocher la fonction de distribution des excès F_u par une fonction de répartition d'une loi de probabilité qu'on appelle distribution de PARÉTO généralisée dont la définition est la suivante :

Définition 2.5.3 [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

Soit $\zeta \in \mathbb{R}$. On appelle **distribution de PARÉTO généralisée standard** toute fonction de répartition G_ζ ou toute loi de probabilité qui a G_ζ comme fonction de répartition telle que :

$$\forall x > 0 \quad \text{tel que } 1 + \zeta x > 0, \quad G_\zeta(x) = \begin{cases} 1 - (1 + \zeta x)^{-1/\zeta} & \text{si } \zeta \neq 0 \\ 1 - e^{-x} & \text{si } \zeta = 0 \end{cases}$$

Pour les autres valeurs de x , on peut deviner les valeurs de G_ζ puisque c'est une fonction de répartition.

Remarque : On peut donner une forme plus générale à la fonction de répartition G_ζ donnée dans la définition ci-dessus, qu'on note $G_{\zeta, \mu, \sigma}$, en faisant apparaître un paramètre de position $\mu \in \mathbb{R}$ et un paramètre d'échelle $\sigma > 0$ comme suit :

$$\forall x > \mu \quad \text{tel que } 1 + \zeta \left(\frac{x - \mu}{\sigma} \right) > 0, \quad G_{\zeta, \mu, \sigma}(x) = \begin{cases} 1 - \left[1 + \zeta \left(\frac{x - \mu}{\sigma} \right) \right]^{-1/\zeta} & \text{si } \zeta \neq 0 \\ 1 - \exp \left(-\frac{x - \mu}{\sigma} \right) & \text{si } \zeta = 0 \end{cases}$$

Et on appelle $G_{\zeta, \mu, \sigma}$ « **distribution de PARÉTO généralisée (GPD)** ».

Notation : Dans toute la suite de ce chapitre, on note $G_{\zeta, \mu, \sigma}$ la distribution de PARÉTO généralisée, G_ζ sa version standard et $G_{\zeta, \sigma}$ la distribution $G_{\zeta, 0, \sigma}$.

Définition 2.5.4 [LEKINA (2010)].

Le paramètre $\zeta \in \mathbb{R}$ qu'on voit dans la distribution de PARÉTO généralisée est un paramètre de forme qu'on appelle « *indice de queue* ».

Remarque : Le paramètre $\zeta \in \mathbb{R}$ de la distribution de PARÉTO généralisée mesure l'épaisseur de sa queue. Le cas $\zeta > 0$ correspond aux distributions aux queues épaisses, le cas $\zeta = 0$ correspond à la distribution exponentielle à queue légère et le cas $\zeta < 0$ correspond à la distribution de PARÉTO type II à support borné.

A. Propriétés de la GPD

A.1 Densité

Proposition 2.5.1

G_ζ ($\zeta \in \mathbb{R}$) est une fonction de répartition absolument continue de densité g_ζ définie par :

$$\forall x > 0, \quad g_\zeta(x) = \begin{cases} (1 + \zeta x)^{-\frac{1}{\zeta}-1} \cdot \mathbb{1}_{[1+\zeta x > 0]}(x) & \text{si } \zeta \neq 0 \\ e^{-x} & \text{si } \zeta = 0 \end{cases}$$

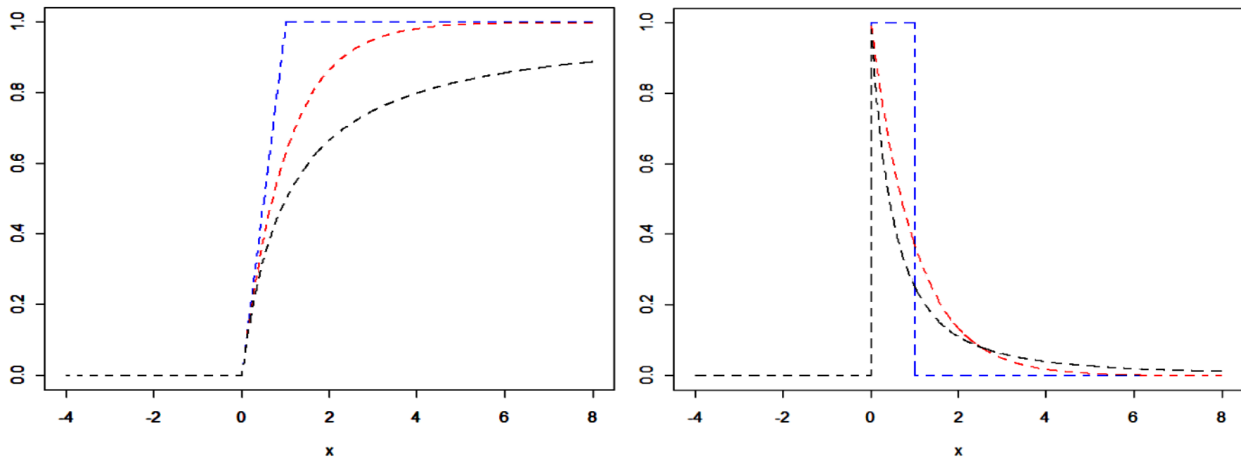


Figure 2.5.2 Panneau gauche représente les graphes de la distribution de PARÉTO généralisée standard G_ζ selon les valeurs de ζ et le panneau droit représente les graphes de leurs densités : Le bleu pour $\zeta = -1$, Le rouge pour $\zeta = 0$ et Le noir pour $\zeta = 1$.

A.2 Moments

Proposition 2.5.2 [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

Soit W la variable aléatoire associée à $G_{\zeta, \sigma}$ ($\zeta \in \mathbb{R}$ et $\sigma > 0$), i.e $G_{\zeta, \sigma}$ est la fonction de répartition de W , alors :

$$\forall k \in \mathbb{N}^* \quad \text{tel que } \zeta < \frac{1}{k}, \quad \mathbb{E}(W^k) = \frac{\sigma^k \Gamma(\zeta^{-1} - k)}{\zeta^{k+1} \Gamma(\zeta^{-1} + 1)} k!$$

et, en particulier, l'espérance et la variance de W sont données comme suit :

$$\zeta < 1 \implies \mathbb{E}(W) = \frac{\sigma}{1 - \zeta} \quad \text{et} \quad \zeta < \frac{1}{2} \implies \mathbb{V}(W) = \frac{\sigma^2}{(1 - \zeta)^2 (1 - 2\zeta)}$$

Remarque : Dans le cadre de la proposition ci-dessus, Γ est la fonction Gamma définie par :

$$\forall x > 0, \quad \Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt$$

A.3 Quantiles

Proposition 2.5.3 [CASTILLO, HADI, BALAKRISHNAN & SARABIA (2005)].

Soient $p \in]0, 1[$ et W la variable aléatoire associée à la distribution de PARÉTO généralisée $G_{\zeta, \sigma}$ ($\zeta \in \mathbb{R}$ et $\sigma > 0$), alors le p -quantile x_p de W est donnée par :

$$x_p = \begin{cases} \frac{\sigma \left[(1-p)^{-\zeta} - 1 \right]}{\zeta} & \text{si } \zeta \neq 0 \\ -\sigma \ln(1-p) & \text{si } \zeta = 0 \end{cases}$$

B. La GPD et la méthode POT

En termes mathématiques, la relation (2.8) donnée dans **Théorème 2.4.2** nous assure que si la fonction de distribution F de la variable aléatoire X est dans le max-domaine d'attraction de la distribution des valeurs extrêmes généralisée H_ξ ($\xi \in \mathbb{R}$), alors la fonction de distribution des excès F_u converge simplement vers la distribution de PARÉTO généralisée $G_{\xi, \delta(u)}$ quand u tend vers le point terminal x_F de la fonction de répartition F , où δ est une fonction strictement positive.

En fait, il y a mieux que la convergence simple de la fonction de distribution des excès F_u vers la distribution de PARÉTO généralisée $G_{\xi, \delta(u)}$. Dans les années 1974 et 1975, BALKEMA, DE HAAN et PICKANDS ont proposé un théorème, qui porte leurs noms, montrant que cette convergence est uniforme. Le voici :

Théorème 2.5.1 [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

La fonction de distribution F de la variable aléatoire X appartient au max-domaine d'attraction de la distribution des valeurs extrêmes généralisée standard H_ξ ($\xi \in \mathbb{R}$) **si et seulement si** il existe une fonction strictement positive δ telle que la fonction de distribution des excès F_u de X par rapport au seuil $u < x_F$ converge uniformément vers une distribution de PARÉTO généralisée $G_{\xi, \delta(u)}$ lorsque u tend vers x_F . Formellement :

$$F \in \text{MDA}(H_\xi) \iff \lim_{u \uparrow x_F} \sup_{0 < y < x_F - u} |F_u(y) - G_{\xi, \delta(u)}| = 0$$

Récapitulation : Pour résumer ce qu'on vient de voir jusqu'ici, si pour une variable aléatoire X de fonction de répartition F inconnue, l'échantillon des maximums normalisés converge en distribution vers une loi de probabilité non dégénérée, alors il est équivalent de dire que F est dans le max-domaine d'attraction d'une distribution GEVD d'indice de queue $\xi \in \mathbb{R}$ (Théorème de FISHER-TIPPET & GNEDENKO). Dans ce cas, il s'en déduit que la distribution des excès de X au-delà d'un certain seuil u converge uniformément vers une distribution GPD, de même indice de queue que celui de la distribution GEVD ($\xi \in \mathbb{R}$), lorsque ce seuil tend vers le point terminal de la fonction de répartition F (Théorème de BALKEMA-DE HAAN & PICKANDS).

À partir de ces deux résultats de probabilité, nous obtenons deux méthodes différentes de modélisation des événements extrêmes : Le premier théorème est à la base de la méthode des maximums par blocs, tandis que le deuxième est à la base de la méthode des excès au-delà d'un seuil (POT). C'est pour cette raison que le théorème de BALKEMA-DE HAAN & PICKANDS est considéré comme le deuxième théorème fondamental de la théorie des valeurs extrêmes, ce qui a donné de l'importance à la distribution de PARÉTO généralisée dans cette théorie.

2.5.3 Fonction moyenne des excès (MEF)

Comme on l'a déjà dit, le théorème de BALKEMA-DE HAAN & PICKANDS nous propose une technique de modélisation des événements extrêmes, qu'on note POT, dont on l'a résumé comme suit : Étant avoir des données x_1, x_2, \dots, x_n ($n \in \mathbb{N}^*$), on suppose qu'elles sont des réalisations de la suite de variables aléatoires X_1, X_2, \dots, X_n indépendantes et de même loi de probabilité que celle de la variable aléatoire X , alors les événements extrêmes seront considérés comme les données qui ont dépassé un certain seuil élevé u qu'on note $x_{(1)}, x_{(2)}, \dots, x_{(k)}$ ($k \leq n$). Donc, d'après le théorème de BALKEMA-DE HAAN & PICKANDS, les excès $y_1 = x_{(1)} - u, \dots, y_k = x_{(k)} - u$ peuvent être considérés comme des réalisations d'une suite de variables aléatoires Y_1, Y_2, \dots, Y_k indépendantes et identiquement distribuées selon une distribution de PARÉTO généralisée.

L'estimation des paramètres d'une distribution GPD pose le problème de la détermination du seuil u , car il doit être suffisamment grand pour que l'on puisse appliquer le résultat précédent, mais il ne doit pas être trop grand afin d'avoir suffisamment de données pour obtenir des estimateurs de bonne qualité. Une méthode, baptisée « graphique », est donnée pour déterminer le bon seuil u en se servant de « la fonction moyenne des excès (MEF⁴) ».

Définition 2.5.5 [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

On appelle **fonction moyenne des excès** de la variable aléatoire X par rapport au seuil $u < x_F$, et on l'a note $e(u)$, la fonction espérance de la variable aléatoire Y excès de X au-delà du seuil $u < x_F$, définie par :

$$\forall u < x_F, \quad e(u) = E(X - u \mid X > u) = \frac{1}{\bar{F}(u)} \int_u^{x_F} \bar{F}(t) dt$$

Dans la proposition qui suit, on va donner un résultat qui justifie la manière dont on sert de la fonction moyenne des excès pour déterminer le bon seuil u :

Proposition 2.5.4 [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

Si W est une variable aléatoire qui a comme fonction de répartition une distribution de PARÉTO généralisée $G_{\zeta, \sigma}$ ($\zeta < 1$, $\sigma > 0$), **alors** sa fonction moyenne des excès $e(u)$ au-delà d'un seuil $u < w_0$ (w_0 est le point terminal de $G_{\zeta, \sigma}$) est donnée par :

$$\forall u < w_0, \quad e(u) = E(W - u \mid W > u) = \frac{\sigma + \zeta u}{1 - \zeta} \quad \text{avec} \quad \sigma + \zeta u > 0$$

4. Mean Excess Function

Remarque : Dans le cadre de la proposition ci-dessus, on remarque que la fonction moyenne des excès $e(u)$ est linéaire en u .

L'idée principale de la manière dont on détermine le bon seuil $u < x_F$ par la méthode graphique est la suivante : Si à partir d'un certain seuil $u_0 < x_F$, l'excès de la variable aléatoire X au-delà du seuil u_0 suit une distribution de PARÉTO généralisée $G_{\zeta, \delta(u_0)}$ ($\zeta < 1$, $\delta(u_0) > 0$) (Théorème de BALKEMA-DE HAAN & PICKANDS), alors la fonction moyenne des excès $e(u_0)$ de X au-delà du seuil u_0 n'est rien d'autre que l'espérance de la distribution de PARÉTO généralisée $G_{\zeta, \delta(u_0)}$, c-à-d :

$$\forall u_0 < x_F, \quad e(u_0) = \mathbb{E}(X - u_0 \mid X > u_0) = \frac{\delta(u_0)}{1 - \zeta}$$

Or, si cette approximation est vraie pour le seuil u_0 , elle sera vraie pour un autre seuil u tel que $u_0 < u < x_F$. Dans ce cas, la fonction moyenne des excès $e(u)$ de X au-delà du seuil u n'est rien d'autre que la fonction moyenne des excès de la distribution de PARÉTO généralisée $G_{\zeta, \delta(u_0)}$ au-delà du seuil u . Formellement :

$$\forall u_0 < u < x_F, \quad e(u) = \mathbb{E}(X - u \mid X > u) = \frac{\delta(u_0) + \zeta u}{1 - \zeta} \quad \text{avec} \quad \delta(u_0) + \zeta u > 0$$

Et pour déterminer le seuil u , on exploite la linéarité en u de la fonction moyenne des excès $e(u)$.

Pour mettre les résultats précédents en application, on doit définir une fonction, qu'on appelle « **fonction moyenne des excès empirique** », qui permette d'estimer la fonction moyenne des excès d'une variable aléatoire au-delà d'un seuil donné :

Définition 2.5.6 [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

Soient (X_1, X_2, \dots, X_n) l'échantillon de taille $n \in \mathbb{N}^*$ de la variable aléatoire X et F_n sa fonction de répartition empirique basée sur cet échantillon. On appelle **fonction moyenne des excès empirique** de la variable aléatoire X par rapport au seuil $u < x_F$ basée sur l'échantillon (X_1, X_2, \dots, X_n) la fonction $e_n(u)$ définie par :

$$\forall u < x_F, \quad e_n(u) = \frac{1}{\overline{F}_n(u)} \int_u^{x_F} \overline{F}_n(t) dt = \frac{1}{\text{card}[\Delta_n(u)]} \sum_{i \in \Delta_n(u)} (X_i - u)$$

avec :

$$\Delta_n(u) = \{i = 1, 2, \dots, n \text{ tel que } X_i > u\} \quad \text{et} \quad \frac{0}{0} = 0 \quad (\text{conventionnellement})$$

Résumé : Si on suppose que nos données x_1, x_2, \dots, x_n ($n \in \mathbb{N}^*$) sont une réalisation de l'échantillon (X_1, X_2, \dots, X_n) de la variable aléatoire X , alors on procède de la manière suivante pour déterminer le seuil $u < x_F$: On calcule l'estimateur $\hat{e}_n(u)$ de la fonction moyenne des excès $e(u)$ de X au-delà du seuil u , en utilisant sa version empirique $e_n(u)$, par :

$$\hat{e}_n(u) = \frac{1}{\text{card}\{i : x_i > u\}} \sum_{i : x_i > u} (x_i - u) \quad \text{avec} \quad \min_{1 \leq i \leq n} x_i \leq u < \max_{1 \leq i \leq n} x_i$$

Ensuite, on trace le graphe :

$$\mathcal{G}_u = \left\{ [u, \hat{e}_n(u)], \min_{1 \leq i \leq n} x_i \leq u < \max_{1 \leq i \leq n} x_i \right\}$$

Une fois le graphe \mathcal{G}_u est tracé, on exploite la linéarité en u de la fonction moyenne des excès d'une distribution de PARÉTO généralisée $G_{\zeta, \sigma}$ ($\zeta < 1$, $\sigma > 0$) au-delà du seuil u en choisissant $u \leq x$, où x est le point à partir duquel le graphe \mathcal{G}_u est approximativement une droite.

Référence

Pour une bonne compréhension de la théorie des valeurs extrêmes univariée et ses applications, le lecteur peut trouver dans la bibliographie de ce mémoire une multitude d'ouvrages dans ce sujet. Parmi eux : [RESNICK (1987)], [DE HAAN & FERREIRA (2006)], [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)], [REISS & THOMAS (2007)] et [BEIRLANT, GOEGEBEUR & TEUGELS (2004)].

Deuxième partie

Extrêmes spatiaux et applications

CHAPITRE 3

CHAMPS ALÉATOIRES

Introduction

Les recherches actuelles dans des nombreux domaines tels l'exploitation minière, les sciences de l'environnement et de la terre, la météorologie, l'écologie, la biologie, la géographie, l'économie, l'épidémiologie, l'agronomie, etc. sont souvent confrontées à l'analyse d'importantes quantités d'informations qui présentent une composante **spatiale** (position géographique). Ces informations sont, de nos jours, faciles à recueillir et à gérer, mais elles ne peuvent pas être analysées avec les outils standards de la statistique classique dont l'une des hypothèses de base est l'indépendance entre les observations. Donc, il faut avoir recours à des nouvelles méthodes d'analyse qui prennent en compte **la dépendance spatiale** entre les observations.

La statistique spatiale, née dans ce cadre, étudie des phénomènes dont les observations sont des données collectées en des lieux dont on a relevé la position géographique, dans le but d'utiliser cette information spatiale dans la modélisation statistique. En particulier, deux données proches géographiquement tendent à être similaires en valeur. Par exemple, la température mesurée à Tizi Ouzou nous permet d'avoir une bonne idée de la température à Boumerdes, car, par expérience, nous savons qu'entre ces deux villes proches les températures varient peu. Mais cela nous renseigne nettement moins sur la température à Blida, et pas du tout sur la température à Tlemcen ou Biskra. Cela traduit le fait que les températures mesurées en des lieux géographiquement proches sont moins variables que les températures mesurées en des lieux éloignés. Plus généralement, dans beaucoup de phénomènes naturels, les corrélations diminuent et la variabilité augmente avec la distance. Cette modélisation nous permettra de réaliser des prédictions spatiales ou de tester des hypothèses en intégrant explicitement cette dépendance spatiale dans les calculs.

Comme les séries temporelles, la statistique spatiale se différencie de la statistique classique par le fait que les observations sont dépendantes. Cela traduit le fait que les méthodes d'analyse spatiales sont fortement demandées de la part de nombreux domaines d'application, ce qui a impliqué un développement important de ces méthodes et a fait la richesse du sujet.

La statistique spatiale s'intéresse à trois grands domaines, qui correspondent à des types de

données distincts : Les données géostatistiques, les données sur un réseau fixé et les données ponctuelles.

1. **Les données géostatistiques** : Ici, les données sont recueillies sur un domaine spatial continu référencé à la Terre. Elles sont caractérisées par des observations associées à une variation continue dans l'espace, généralement en fonction de la distance entre les sites d'observation qui peuvent être disposés régulièrement ou non. Par exemple, les teneurs en matière organique relevées sur des échantillons de sol recueillis à travers tout un champ agricole.
2. **Les données sur un réseau fixé** : Ici, les données sont recueillies sur un réseau régulier ou irrégulier qui comprend un nombre fini de sites et l'étendue de ce réseau est limitée. Les valeurs des données sont conférées à des points et les emplacements de ces points sont connus. Par exemple, dans l'analyse d'images, le réseau est un ensemble régulier de pixels et les données sont les nombres de bits intervenant dans le codage de chaque pixel, le nombre de personnes à l'intérieur de chaque comté d'un État, etc.
3. **Les données ponctuelles** : Ici, c'est l'ensemble des sites où ont lieu les observations qui est aléatoire tout comme leur nombre (voir ANNEXE C). Ces données se concentrent sur l'emplacement de chacun de ces sites (points). Avec ce type de données spatiales, il n'est pas possible de prédire avec exactitude l'emplacement de l'occurrence de l'événement. Par exemple, l'emplacement des unités d'habitation, les positions des arbres dans une forêt, les points d'impacts de la foudre dans une région, etc.

Bien entendu, les frontières ou la distinction entre ces trois domaines n'est pas toujours aussi nette. Par exemple, lorsque un champ aléatoire ponctuel porte une variable attachée au point (le diamètre d'un tronc d'arbre, par exemple), le formalisme adopté pour étudier ce type de données utilise à la fois le formalisme de données ponctuelles et le formalisme de données géostatistiques.

La théorie des champs aléatoires est l'outil probabiliste permettant de modéliser ce fait que nous expérimentons tous les jours. La statistique spatiale est l'application de cette théorie pour répondre à des questions de prédiction spatiale. L'objectif de ce chapitre est de fournir une introduction à la théorie des champs aléatoires et à la géostatistique en tant que domaine de la statistique spatiale.

La géostatistique est née dans le contexte très particulier de l'estimation des réserves des gisements miniers. Georges MATHERON (1930-2000) en a posé les fondations dans l'ouvrage « Les variables régionalisées et leur estimation » paru en 1965. Il s'agissait d'une contribution pionnière en statistiques appliquées visant à poser un cadre théorique et méthodologique répondant à deux difficultés majeures :

1. Comment définir un cadre statistique cohérent pour étudier un phénomène sans répétition ?
2. Comment prendre en compte les corrélations spatiales existantes entre des données récoltées sur un événement unique ?

Les années 1965 à 1985 ont essentiellement été des années de consolidation et de diffusion auprès des utilisateurs de ces concepts (principalement, en géologie et en sciences de l'environnement), tout en restant en marge de la communauté statistique. De ce contexte minier est resté un vocabulaire assez spécifique : **krigeage**, **effet de pépité**, **effet de support**, etc. En 1989 est paru

« Estimating and Choosing », un livre de réflexion épistémologique sur la place de la modélisation probabiliste en géoscience. À partir des années 1980, la géostatistique a été redécouverte par les statisticiens, principalement américains. Le livre « Statistics for Spatial Data » de Noël CRESSIE a marqué l’ancrage définitif de la géostatistique dans le domaine plus vaste des statistiques spatiales et a entraîné sa pleine reconnaissance par la communauté statistique. Les recherches les plus récentes dans ce domaine concernent la pratique des tests, la modélisation spatio-temporelle, la géostatistique multivariée et fonctionnelle, la géostatistique sur les réseaux et la modélisation géostatistique dans les modèles hiérarchiques.

Remarque : Le préfixe « **géo** » qu’on voit dans le mot « **géostatistique** » sert à indiquer la Terre, ce qui signifie que le mot « géostatistique » veut dire les statistiques concernant la Terre.

3.1 Champ aléatoire

Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, (E, \mathcal{E}) un espace mesurable et S un sous espace d’un espace euclidien¹ de dimension finie.

Définition 3.1.1 [GAETAN & GUYON (2008)].

Un champ aléatoire, ou un processus stochastique, à espace d’état E est une famille $X = \{X(s), s \in S\}$ de variables aléatoires définies sur $(\Omega, \mathcal{A}, \mathbb{P})$ dans (E, \mathcal{E}) .

Remarque : Dans le cadre de la définition ci-dessus, l’espace d’état E du champ aléatoire X est l’ensemble des valeurs de X . Dans ce chapitre, on ne considère que le cas où $E = \mathbb{R}$ et dans ce cas l’espace mesurable (E, \mathcal{E}) s’identifie à l’espace mesurable $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ où $\mathcal{B}_{\mathbb{R}}$ est la tribu borélienne² de \mathbb{R} . Bien sûr, on peut envisager d’autres espaces tels l’espace des nombres complexes \mathbb{C} ou l’espace \mathbb{R}^p ($p \geq 1$) ou, encore, un espace fonctionnel, mais cela nous entrainera dans des calculs qui sortent de l’objectif de ce travail.

Définition 3.1.2 [GAETAN & GUYON (2008)].

On appelle l’ensemble d’indices S du champ aléatoire X , qui apparait dans la définition précédente, l’ensemble spatial des sites sur lequel est défini X .

Remarque : Dans ce chapitre, on ne considère que le cas où l’ensemble spatial des sites, sur lequel est défini le champ aléatoire X , $S \subset \mathbb{R}^d$ ($d \geq 1$), ou, tout simplement, $S = \mathbb{R}^d$.

3.1.1 Loi spatiale d’un champ aléatoire

Soit $X = \{X(s), s \in \mathbb{R}^d\}$ un champ aléatoire défini sur l’espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à espace d’état $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$.

Définition 3.1.3 [GAETAN & GUYON (2008)].

On appelle loi spatiale du champ aléatoire X la loi de probabilité de X .

1. Un espace euclidien est un espace vectoriel muni de la norme euclidienne $\|\cdot\|_2$.

2. La tribu borélienne d’un espace topologique est la tribu engendrée par la famille des ouverts de cet espace.

D'après la définition du champ aléatoire $X = \{X(s), s \in \mathbb{R}^d\}$, on sait que pour $s \in \mathbb{R}^d$ fixé, $X(s)$ est une variable aléatoire définie sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$. Donc, on peut déterminer la loi spatiale de X si on connaît la famille de toutes les lois de probabilité des vecteurs aléatoires de la forme :

$$\left(X(s_1), X(s_2), \dots, X(s_n) \right), \quad s_i \in \mathbb{R}^d \quad \text{avec} \quad i \in \{1, 2, \dots, n\} \quad \text{et} \quad n \in \mathbb{N}^*$$

Définition 3.1.4 [GAETAN & GUYON (2008)].

On appelle **distribution finie-dimensionnelle** du champ aléatoire $X = \{X(s), s \in \mathbb{R}^d\}$, la famille de toutes les lois de probabilité des vecteurs aléatoires de la forme :

$$\left(X(s_1), X(s_2), \dots, X(s_n) \right), \quad s_i \in \mathbb{R}^d \quad \text{avec} \quad i \in \{1, 2, \dots, n\} \quad \text{et} \quad n \in \mathbb{N}^*$$

données par :

$$F_{s_1, s_2, \dots, s_n}(A_1 \times A_2 \times \dots \times A_n) = \mathbb{P}\left(X(s_1) \in A_1, X(s_2) \in A_2, \dots, X(s_n) \in A_n\right), \quad A_i \in \mathcal{B}_{\mathbb{R}} \quad (3.1)$$

Remarque : Puisque $\mathcal{B}_{\mathbb{R}}$ la tribu borélienne de \mathbb{R} est engendrée par la famille des intervalles de la forme $]-\infty, x]$ ($x \in \mathbb{R}$), alors on peut remplacer (3.1) par :

$$F_{s_1, s_2, \dots, s_n}(x_1, x_2, \dots, x_n) = \mathbb{P}\left(X(s_1) \leq x_1, X(s_2) \leq x_2, \dots, X(s_n) \leq x_n\right), \quad x_i \in \mathbb{R}$$

Donc, on peut poser le résultat suivant :

Proposition 3.1.1 [RESNICK (1992)].

La loi spatiale du champ aléatoire $X = \{X(s), s \in \mathbb{R}^d\}$ est complètement déterminée par sa distribution finie-dimensionnelle.

Les lois de probabilité de la distribution finie-dimensionnelle du champ aléatoire X donnée par (3.1) ne peuvent pas être quelconques. Elles doivent vérifier des conditions minimales de consistance, appelées **conditions de consistance de KOLMOGOROV** données dans la définition suivante :

Définition 3.1.5 [BHATTACHARYA & WAYMIRE (2007)].

On dit que la loi de probabilité, donnée dans (3.1), de la distribution finie-dimensionnelle du champ aléatoire $X = \{X(s), s \in \mathbb{R}^d\}$ vérifie **les conditions de consistance de KOLMOGOROV**, si elle vérifie les deux conditions suivantes :

1. Condition de symétrie :

$$F_{s_1, s_2, \dots, s_n}(A_1 \times A_2 \times \dots \times A_n) = F_{s_{\pi(1)}, s_{\pi(2)}, \dots, s_{\pi(n)}}(A_{\pi(1)} \times A_{\pi(2)} \times \dots \times A_{\pi(n)})$$

π est une permutation quelconque de $(1, 2, \dots, n)$

2. Condition de consistance par les marginales :

$$\forall s_{n+1} \in \mathbb{R}^d, \quad F_{s_1, s_2, \dots, s_n}(A_1 \times A_2 \times \dots \times A_n) = F_{s_1, s_2, \dots, s_n, s_{n+1}}(A_1 \times A_2 \times \dots \times A_n \times \mathbb{R})$$

Le théorème suivant, qu'on appelle **Théorème d'existence de KOLMOGOROV** ou **Théorème de consistance de KOLMOGOROV**, nous affirme que si on a une distribution finie-dimensionnelle dont les éléments vérifient les conditions de consistance de KOLMOGOROV, alors c'est une distribution finie-dimensionnelle d'un champ aléatoire. En fait, dans le théorème qui suit, on va voir que même l'inverse est vraie :

Théorème 3.1.1 [BILLINGSLEY (1995)].

Soit $\mathcal{P} = \left\{ G_{t_1, t_2, \dots, t_n} : t_i \in \mathbb{R}^d (1 \leq i \leq n), n \in \mathbb{N}^* \right\}$ une famille de lois de probabilité définies sur $[\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)]$. \mathcal{P} est une distribution finie-dimensionnelle d'un champ aléatoire défini sur l'ensemble spoutial des sites \mathbb{R}^d à espace d'état $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ **si et seulement si** les lois de probabilité de \mathcal{P} vérifient les conditions de consistance de KOLMOGOROV.

Résultat : Du théorème ci-dessus, on peut constater que si la famille \mathcal{P} de lois de probabilité G_{t_1, t_2, \dots, t_n} ne vérifient pas les conditions de consistance de KOLMOGOROV, alors il ne peut pas exister un champ aléatoire dont \mathcal{P} est la distribution finie-dimensionnelle. Donc, toute distribution finie-dimensionnelle d'un champ aléatoire doit vérifier les conditions de consistance de KOLMOGOROV.

3.1.2 Moments d'un champ aléatoire

Soit $X = \{X(s), s \in \mathbb{R}^d\}$ un champ aléatoire défini sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à espace d'état $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$. Pour $s \in \mathbb{R}^d$, on note F_s la loi de probabilité de la variable aléatoire $X(s)$.

Définition 3.1.6 [LE & ZIDEK (2006)].

On appelle **moment d'ordre** $k \in \mathbb{N}^*$ (**s'il existe**) du champ aléatoire $X = \{X(s), s \in \mathbb{R}^d\}$ la fonction, qu'on note m_k , définie sur \mathbb{R}^d dans \mathbb{R} par :

$$\forall s \in \mathbb{R}^d, \quad m_k(s) = \mathbb{E}[X^k(s)] = \int_{\mathbb{R}} x^k dF_s(x), \quad k \in \mathbb{N}^*$$

Remarque : Comme on l'a signalé dans la définition ci-dessus, il se peut que le moment d'ordre $k \in \mathbb{N}^*$ d'un champ aléatoire X n'existe pas. Mais s'il existe, alors le champ aléatoire X admet des moments d'ordres inférieurs à k .

A. Espérance

Définition 3.1.7 [LE & ZIDEK (2006)].

On appelle **espérance (si elle existe)** du champ aléatoire $X = \{X(s), s \in \mathbb{R}^d\}$ le moment d'ordre un de X qui est une fonction, qu'on note m , définie sur \mathbb{R}^d dans \mathbb{R} par :

$$\forall s \in \mathbb{R}^d, \quad m(s) = \mathbb{E}[X(s)]$$

Remarque : On appelle, souvent, l'espérance d'un champ aléatoire, puisque c'est une fonction, par **fonction de moyenne** de ce champ aléatoire.

Définition 3.1.8 [GAETAN & GUYON (2008)].

On dit que le champ aléatoire $X = \{X(s), s \in \mathbb{R}^d\}$ est **centré** si sa fonction de moyenne m est nulle, c-à-d :

$$\forall s \in \mathbb{R}^d, \quad m(s) = 0$$

B. Variance

Définition 3.1.9 [LE & ZIDEK (2006)].

On appelle **variance** ou **fonction de variance** (si elle existe) du champ aléatoire $X = \{X(s), s \in \mathbb{R}^d\}$ le moment centré d'ordre **deux** de X qui est une fonction, qu'on note v , définie sur \mathbb{R}^d dans \mathbb{R}_+ par :

$$\forall s \in \mathbb{R}^d, \quad v(s) = V[X(s)] = E\left([X(s) - m(s)]^2\right)$$

C. Covariance

La covariance de deux variables aléatoires mesure la force de la relation qui existe entre elles. Dans un champ aléatoire, la covariance mesure la force de la relation qui existe entre les variables aléatoires qui le représentent dans les différents sites d'observation. C'est pour cette raison qu'elle joue un rôle important dans le problème de prédiction.

Définition 3.1.10 [LE & ZIDEK (2006)].

On appelle **covariance** ou **fonction de covariance** (si elle existe) du champ aléatoire $X = \{X(s), s \in \mathbb{R}^d\}$ la fonction, qu'on note c , définie sur $\mathbb{R}^d \times \mathbb{R}^d$ dans \mathbb{R} par :

$$\forall (s, t) \in \mathbb{R}^d \times \mathbb{R}^d, \quad c(s, t) = \text{Cov}[X(s), X(t)] = E\left([X(s) - m(s)][X(t) - m(t)]\right)$$

La propriété caractéristique d'une fonction de covariance d'un champ aléatoire, si elle existe, est d'être **semi-définie positive** :

Proposition 3.1.2 [GAETAN & GUYON (2008)].

Supposons que la fonction de covariance c du champ aléatoire $X = \{X(s), s \in \mathbb{R}^d\}$ existe, **alors** c est une fonction **semi-définie positive**, c-à-d :

$$\forall n \in \mathbb{N}^*, \quad \forall a = (a_1, a_2, \dots, a_n) \in \mathbb{R}^n, \quad \forall (s_1, s_2, \dots, s_n) \in (\mathbb{R}^d)^n :$$

$$\sum_{i=1}^n \sum_{j=1}^n a_i a_j c(s_i, s_j) \geq 0 \tag{3.2}$$

Remarque : De la proposition ci-dessus, on constate que la condition (3.2) est une condition nécessaire pour que la fonction c soit une fonction de covariance du champ aléatoire X . C'est, également, une condition suffisante dans le cas gaussien qu'on va voir plus loin. On dira que c est **définie positive** si l'inégalité dans (3.2) est stricte.

D Fonction de corrélation

Comme on va le voir dans la définition qui suit, l'existence de la fonction de corrélation d'un champ aléatoire dépend directement de l'existence de la fonction de covariance de ce champ. C'est pour cette raison qu'elles ont, presque, le même rôle dans la statistique spatiale.

Définition 3.1.11 [GAETAN & GUYON (2008)].

On appelle **fonction de corrélation** (si elle existe) du champ aléatoire $X = \{X(s), s \in \mathbb{R}^d\}$ la fonction, qu'on note ρ , définie sur $\mathbb{R}^d \times \mathbb{R}^d$ dans $[-1, 1]$ par :

$$\forall (s, t) \in \mathbb{R}^d \times \mathbb{R}^d, \quad \rho(s, t) = \frac{\text{Cov}[X(s), X(t)]}{\sqrt{V[X(s)]}\sqrt{V[X(t)]}} = \frac{c(s, t)}{\sqrt{v(s)}\sqrt{v(t)}}$$

E Variogramme

En 1971, Georges MATHERON a proposé l'utilisation du **semi-variogramme** ou **variogramme** comme une alternative à la fonction de covariance d'un champ aléatoire, dont la définition est la suivante :

Définition 3.1.12 [LE & ZIDEK (2006)].

On appelle **semi-variogramme** (s'il existe) du champ aléatoire $X = \{X(s), s \in \mathbb{R}^d\}$ la fonction, qu'on note γ , définie sur $\mathbb{R}^d \times \mathbb{R}^d$ dans \mathbb{R}_+ par :

$$\forall (s, t) \in \mathbb{R}^d \times \mathbb{R}^d, \quad \gamma(s, t) = \frac{1}{2} V[X(s) - X(t)] = \frac{1}{2} E\left\{([X(s) - X(t)] - [m(s) - m(t)])^2\right\}$$

Définition 3.1.13 [LE & ZIDEK (2006)].

On appelle **variogramme** (s'il existe) du champ aléatoire $X = \{X(s), s \in \mathbb{R}^d\}$ la fonction, qu'on note ϑ , définie sur $\mathbb{R}^d \times \mathbb{R}^d$ dans \mathbb{R}_+ par :

$$\forall (s, t) \in \mathbb{R}^d \times \mathbb{R}^d, \quad \vartheta(s, t) = V[X(s) - X(t)]$$

Remarque : Le variogramme d'un champ aléatoire, s'il existe, est souvent utilisé dans le cas où sa fonction de covariance n'existe pas. En fait, en 1973, Georges MATHERON a introduit la classe des champs aléatoires « **intrinsèquement stationnaires** », qu'on va voir plus loin, qui ont des accroissements stationnaires au second ordre et dont le variogramme existe et pas nécessairement la fonction de covariance.

3.2 Champ aléatoire au second ordre

Soit $X = \{X(s), s \in \mathbb{R}^d\}$ un champ aléatoire défini sur l'espace probabilisé (Ω, \mathcal{A}, P) à espace d'état $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$. L'une des classes des champs aléatoires, très importante en applications, est la classe des champs aléatoires au **second ordre**. Son importance réside dans le fait que de tels champs aléatoires admettent les moments qu'on a définis dans la section précédente.

Définition 3.2.1 [GAETAN & GUYON (2008)].

On dit que le champ aléatoire $X = \{X(s), s \in \mathbb{R}^d\}$ est **au second ordre** si :

$$\forall s \in \mathbb{R}^d, \quad \mathbb{E}[X^2(s)] < \infty$$

Remarque : De la définition ci-dessus, on constate que le champ aléatoire $X = \{X(s), s \in \mathbb{R}^d\}$ est au second ordre si et seulement si, pour tout $s \in \mathbb{R}^d$, la variable aléatoire $X(s)$ est de carré intégrable, c'est-à-dire $X(s) \in \mathbb{L}^2(\Omega, \mathcal{A}, \mathbb{P})$. Dans ce cas, on écrit $X \in \mathbb{L}^2$ pour dire que le champ aléatoire $X = \{X(s), s \in \mathbb{R}^d\}$ est au second ordre.

3.2.1 Continuité en moyenne quadratique

Dans cette section, on va voir la relation qui existe entre la propriété géométrique « **continuité en moyenne quadratique** » d'un champ aléatoire et les caractéristiques de sa fonction de covariance. Pour cela, on considère que le champ aléatoire $X = \{X(s), s \in \mathbb{R}^d\}$ est au second ordre, c'est-à-dire que $X \in \mathbb{L}^2$.

Définition 3.2.2 [GAETAN & GUYON (2008)].

On dit que le champ aléatoire $X \in \mathbb{L}^2$ est **continu en moyenne quadratique** au point $s \in \mathbb{R}^d$, si, pour toute suite $\{s_n, n \in \mathbb{N}\}$ de \mathbb{R}^d convergente vers s , on a :

$$\lim_{n \rightarrow \infty} \mathbb{E}\left[\left(X(s_n) - X(s)\right)^2\right] = 0$$

La proposition suivante caractérise la continuité en moyenne quadratique d'un champ aléatoire au second ordre :

Proposition 3.2.1 [GAETAN & GUYON (2008)].

Soit $X = \{X(s), s \in \mathbb{R}^d\}$ un champ aléatoire au second ordre de fonction de covariance c . **Si** c est continue sur la diagonale de $\mathbb{R}^d \times \mathbb{R}^d$, c'est-à-dire c est continue en tout point $(s, s) \in \mathbb{R}^d \times \mathbb{R}^d$, **alors** X est un champ aléatoire continu en moyenne quadratique partout dans \mathbb{R}^d .

Remarque : La condition de continuité quadratique du champ aléatoire X donnée dans la proposition ci-dessus est une condition suffisante. Elle est, également, nécessaire si la fonction moyenne de X est constante.

3.2.2 Champ aléatoire gaussien

Comme exemple des champs aléatoires au second ordre, les champs aléatoires **gaussiens** constituent une classe très importante. Ces champs aléatoires sont largement utilisés comme modèles des données géostatistiques, et cela est dû, essentiellement, à deux raisons. La première est une raison pragmatique, car, sous l'hypothèse du champ aléatoire gaussien, la prédiction, l'estimation et la distribution sont moins difficiles. La deuxième raison vient du fait que beaucoup de phénomènes sont gaussiens, car ils sont dus à l'addition d'un grand nombre de petites perturbations

aléatoires qui sont pas, nécessairement, gaussiennes (Théorème de la limite centrale).

Soit $X = \{X(s), s \in \mathbb{R}^d\}$ un champ aléatoire défini sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à espace d'état $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$:

Définition 3.2.3 [DIGGLE & RIBEIRO JR (2007)].

On dit que la champ aléatoire $X = \{X(s), s \in \mathbb{R}^d\}$ est **gaussien** si, pour toute suite d'éléments $\{s_1, s_2, \dots, s_n\}$ de \mathbb{R}^d ($n \in \mathbb{N}^*$), le vecteur aléatoire $(X(s_1), X(s_2), \dots, X(s_n))$ est un vecteur aléatoire gaussien, c'est-à-dire :

$$\forall \lambda = (\lambda_1, \lambda_2, \dots, \lambda_n) \in \mathbb{R}^n, \quad \sum_{i=1}^n \lambda_i X(s_i) \sim \mathcal{N}(\mu, \sigma^2), \quad \mu \in \mathbb{R} \text{ et } \sigma \in \mathbb{R}_+^*$$

Remarque : De la définition ci-dessus, on constate que le champ aléatoire X est gaussien si et seulement si les lois de probabilité de sa distribution finie-dimensionnelle sont des lois de probabilité normales multidimensionnelles.

Proposition 3.2.2 [GAETAN & GUYON (2008)].

Si le champ aléatoire $X = \{X(s), s \in \mathbb{R}^d\}$ est gaussien, alors X est au second ordre.

Remarque : Le résultat de la proposition ci-dessus découle des propriétés des lois de probabilité gaussiennes. Et en fait, un champ aléatoire gaussien n'admet pas seulement des moments d'ordre inférieur ou égal à deux, mais il admet des moments de tout ordre.

Donc, un champ aléatoire gaussien est un champ aléatoire dont les lois de probabilité de sa distribution finie-dimensionnelle sont des lois de probabilité normales multidimensionnelles. Par conséquent, un tel champ est complètement déterminé par sa fonction moyenne et sa fonction de covariance. Et, dans ce cas particulier, les conditions de consistance du « Théorème d'existence de KOLMOGOROV » se réduisent à l'exigence que la fonction de covariance soit semi-définie positive :

Théorème 3.2.1 [BILLINGSLEY (1995)].

Soit c une fonction définie sur $\mathbb{R}^d \times \mathbb{R}^d$ dans \mathbb{R} . La fonction c est semi-définie positive **si et seulement** s'il existe un champ aléatoire gaussien défini sur l'ensemble spatial des sites \mathbb{R}^d à espace d'état $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ dont la fonction de covariance est c .

Exemple

Comme exemple des champs aléatoires gaussiens, les **mouvements browniens** définis sur \mathbb{R}_+ constituent une classe très importante.

Le mouvement brownien est, avec le processus ponctuel de POISSON, le champ aléatoire (processus stochastique) le plus important dans la théorie des probabilités et ses applications. Sa découverte est due au botaniste britannique ROBERT BROWN, en 1827, lorsqu'il observa le mouvement

irrégulier des particules du pollen totalement immergées dans l'eau. Ainsi, en raison de leurs mouvement naturel, les molécules d'un liquide heurtent, en effet, de façon aléatoire les particules en suspension, les mettant ainsi en mouvement. Depuis, le mouvement brownien est largement utilisé dans de nombreux domaines. En 1905, Albert EINSTEIN établit une description mathématique du phénomène et l'intégra dans la théorie cinétique. En chimie, l'un des premiers calculs du nombre d'AVOGADRO, constante fondamentale en chimie, fut réalisé, en 1910, par le physicien français Jean Baptiste PERRIN, à partir d'une étude quantitative du mouvement brownien. Cependant, c'est Norbert WIENER qui, en 1921, donna sa définition concise suivante :

Définition 3.2.4 [KANNAN (1979)].

Soit $B = \{B(t), t \in \mathbb{R}_+\}$ un champ aléatoire défini sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à espace d'état $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$. On dit que le champ aléatoire B est un **mouvement brownien** s'il vérifie les conditions suivantes :

1. $B(0) = 0$.
2. B est à **accroissements stationnaires**, c'est-à-dire :

$$\forall t \in \mathbb{R}_+, \quad \forall h > 0, \quad B(t+h) - B(t) \stackrel{d}{=} B(h)$$

3. B est à **accroissements indépendants**, c'est-à-dire, pour $n \in \mathbb{N}^*$, on a :

$$\forall 0 \leq t_1 < t_2 < \dots < t_{n-1} < t_n, \quad B(t_2 - t_1) \perp \dots \perp B(t_n - t_{n-1})$$

4. $B(t)$ suit une loi de probabilité normale de paramètre $(0, \sigma^2 t)$ ($\sigma > 0$), c'est-à-dire :

$$\forall x \in \mathbb{R}, \quad \mathbb{P}[B(t) \leq x] = \frac{1}{\sigma\sqrt{2t\pi}} \int_{-\infty}^x \exp\left(-\frac{u^2}{2\sigma^2 t}\right) du, \quad t \in \mathbb{R}_+^*$$

Remarque : Le signe \perp qu'on voit dans la troisième condition de la définition ci-dessus veut dire l'indépendance. C'est-à-dire, si on écrit $Y \perp Z$ (Y et Z sont des variables aléatoires), alors cela voudra dire que Y et Z sont indépendantes.

Proposition 3.2.3 [KANNAN (1979)].

Soit $B = \{B(t), t \in \mathbb{R}_+\}$ un champ aléatoire défini sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à espace d'état $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$. Si B est un mouvement brownien, **alors** B est un champ aléatoire gaussien.

3.3 Stationnarité d'un champ aléatoire

En statistique classique, l'inférence des paramètres est rendue possible par la répétition indépendante des données. En statistique spatiale, on observe très souvent une réalisation unique des données. Par exemple, **un** épisode de pollution à l'ozone, **une** région agricole particulière, **une** épidémie végétale, etc. Pour pouvoir réaliser l'inférence statistique pour un événement unique, il faut donc en quelques sortes remplacer l'hypothèse sur les répétitions indépendantes par une hypothèse sur le champ aléatoire qui considère que certaines de ses caractéristiques sont identiques d'un point à l'autre de l'espace. On pose donc des hypothèses de stationnarité.

3.3.1 Stationnarité stricte

Soit $X = \{X(s), s \in \mathbb{R}^d\}$ un champ aléatoire défini sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à espace d'état $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$.

Définition 3.3.1 [GAETAN & GUYON (2008)].

On dit que le champ aléatoire $X = \{X(s), s \in \mathbb{R}^d\}$ est **strictement stationnaire** si, pour tout $h \in \mathbb{R}^d$ et pour toute suite d'éléments $\{s_1, s_2, \dots, s_n\}$ de \mathbb{R}^d ($n \in \mathbb{N}^*$), les vecteurs aléatoires $(X(s_1), X(s_2), \dots, X(s_n))$ et $(X(s_1 + h), X(s_2 + h), \dots, X(s_n + h))$ sont identiquement distribués.

Remarque : De la définition ci-dessus, on constate que le champ aléatoire $X = \{X(s), s \in \mathbb{R}^d\}$ est strictement stationnaire si et seulement si les lois de probabilité de sa distribution finie-dimensionnelle sont invariantes pour toute translation $h \in \mathbb{R}^d$. Formellement, si F_{s_1, s_2, \dots, s_n} ($n \in \mathbb{N}^*$) représente les lois de probabilité de la distribution finie-dimensionnelle du champ aléatoire $X = \{X(s), s \in \mathbb{R}^d\}$, alors la stationnarité stricte de X veut dire :

$$\forall h \in \mathbb{R}^d, \quad F_{s_1, s_2, \dots, s_n}(A_1 \times A_2 \times \dots \times A_n) = F_{s_1+h, s_2+h, \dots, s_n+h}(A_1 \times A_2 \times \dots \times A_n), \quad A_i \in \mathcal{B}_{\mathbb{R}}$$

Exemple

Comme exemple de champs aléatoires strictement stationnaires, les **bruits blancs forts** constituent une classe très importante surtout dans la théorie des séries temporelles :

Définition 3.3.2 [GAETAN & GUYON (2008)].

Supposons que le champ aléatoire $X = \{X(s), s \in \mathbb{R}^d\} \in \mathbb{L}^2$. On dit que X est un **bruit blanc fort** s'il vérifie les deux conditions suivantes :

1. X est un champ aléatoire **centré** :

$$\forall s \in \mathbb{R}^d, \quad m(s) = \mathbb{E}[X(s)] = 0$$

2. Les variables aléatoires $\{X(s), s \in \mathbb{R}^d\}$ sont **indépendantes et identiquement distribuées**.

Remarque : Dans le cadre de la définition ci-dessus et si, de plus, les variables aléatoires $\{X(s), s \in \mathbb{R}^d\}$ sont gaussiennes, on dit que X est un **bruit blanc gaussien**.

Proposition 3.3.1 [GAETAN & GUYON (2008)].

Un bruit blanc fort est un champ aléatoire strictement stationnaire.

Remarque : Le résultat de la proposition ci-dessus découle, automatiquement, de la deuxième condition de la définition précédente.

La stationnarité stricte est une hypothèse rarement utilisée. Car, d'une part et comme elle le montre sa définition, cette hypothèse est totalement invérifiable puisque elle est faite pour tout

$n \in \mathbb{N}^*$. D'autre part, même si cette hypothèse est vérifiable, elle ne dit rien sur l'existence ou non des moments du champ aléatoire $X = \{X(s), s \in \mathbb{R}^d\}$, alors que la moyenne et la variance sont des grandeurs nécessaires pour l'analyse statistique des données. D'où on a recours à d'autres types de stationnarité.

3.3.2 Stationnarité au second ordre

Soit $X = \{X(s), s \in \mathbb{R}^d\} \in \mathbb{L}^2$ un champ aléatoire au second ordre défini sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à espace d'état $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$. Dans cette section, on va définir un autre type de stationnarité, largement utilisé dans l'analyse statistique des données, qu'on appelle **stationnarité au second ordre** ou **stationnarité faible** ou, encore, **stationnarité au sens large**.

Définition 3.3.3 [CRESSIE (1991)].

On dit que le champ aléatoire $X = \{X(s), s \in \mathbb{R}^d\} \in \mathbb{L}^2$ est **stationnaire au second ordre** s'il vérifie les conditions suivantes :

1. La fonction de moyenne de X est constante :

$$\forall s \in \mathbb{R}^d, \quad m(s) = \mathbb{E}[X(s)] = \mu \in \mathbb{R}$$

2. La fonction de covariance de X en deux sites s et t ne dépend que de vecteur $(t - s)$:

$$\forall (s, t) \in \mathbb{R}^d \times \mathbb{R}^d, \quad c(s, t) = \text{Cov}[X(s), X(t)] = C(t - s)$$

Remarque : De la deuxième condition de la définition ci-dessus, on peut constater que même la fonction de variance d'un champ aléatoire $X = \{X(s), s \in \mathbb{R}^d\} \in \mathbb{L}^2$ stationnaire au second ordre est constante, car :

$$\forall s \in \mathbb{R}^d, \quad v(s) = \mathbb{V}[X(s)] = c(s, s) = \text{Cov}[X(s), X(s)] = C(0_{\mathbb{R}^d})$$

Dans la proposition qui suit, on va établir la relation qui existe entre la stationnarité stricte et la stationnarité au second ordre :

Proposition 3.3.2 [GAETAN & GUYON (2008)].

Si le champ aléatoire $X = \{X(s), s \in \mathbb{R}^d\} \in \mathbb{L}^2$ est strictement stationnaire, **alors** X est stationnaire au second ordre.

Remarque : De la proposition ci-dessus, on déduit que la stationnarité au second ordre est une condition nécessaire pour qu'un champ aléatoire $X = \{X(s), s \in \mathbb{R}^d\} \in \mathbb{L}^2$ soit strictement stationnaire. Elle est, également, suffisante si X est un champ aléatoire gaussien.

Proposition 3.3.3

Soit $X = \{X(s), s \in \mathbb{R}^d\} \in \mathbb{L}^2$ un champ aléatoire gaussien. X est un champ aléatoire strictement stationnaire **si et seulement si** X est stationnaire au second ordre.

A Exemple

Comme exemple de champs aléatoires stationnaires au second ordre, la classe des **bruits blancs faibles** est un bon et simple exemple :

Définition 3.3.4 [GAETAN & GUYON (2008)].

Soit $X = \{X(s), s \in \mathbb{R}^d\} \in \mathbb{L}^2$ un champ aléatoire au second ordre. On dit que X est un **bruit blanc faible** s'il vérifie les conditions suivantes :

1. X est un champ aléatoire centré :

$$\forall s \in \mathbb{R}^d, \quad m(s) = \mathbb{E}[X(s)] = 0$$

2. La fonction de variance de X est constante :

$$\forall s \in \mathbb{R}^d, \quad v(s) = \mathbb{V}[X(s)] = \sigma^2 \in \mathbb{R}_+$$

3. Les variables aléatoires $\{X(s), s \in \mathbb{R}^d\}$ sont deux à deux non corrélées :

$$\forall (s, t) \in \mathbb{R}^d \times \mathbb{R}^d, \quad c(s, t) = \text{Cov}[X(s), X(t)] = 0$$

Remarque : De la définition même d'un bruit blanc faible, on peut en déduire qu'il est stationnaire au second ordre.

B Covariogramme

Comme on l'a dit auparavant, la fonction de covariance d'un champ aléatoire mesure la force de la relation qui existe entre les variables aléatoires qui le représentent dans les différents sites d'observation. Donc, elle joue un rôle très important dans le problème de prédiction.

De la définition de la stationnarité au second ordre, on constate que lorsque un champ aléatoire $X = \{X(s), s \in \mathbb{R}^d\} \in \mathbb{L}^2$ est stationnaire au second ordre, alors sa fonction de covariance ne dépend que d'un seul argument au lieu de deux. Et, dans ce cas, on appelle cette fonction de covariance « **le covariogramme** » de X .

Définition 3.3.5 [CRESSIE (1991)].

Soit $X = \{X(s), s \in \mathbb{R}^d\} \in \mathbb{L}^2$ un champ aléatoire stationnaire au second ordre. On appelle **covariogramme** du champ aléatoire X la fonction C définie sur \mathbb{R}^d dans \mathbb{R} par :

$$\forall h \in \mathbb{R}^d, \quad C(h) = \text{Cov}[X(s), X(s+h)], \quad s \in \mathbb{R}^d$$

Remarque : De la définition ci-dessus, on comprend qu'un covariogramme d'un champ aléatoire $X = \{X(s), s \in \mathbb{R}^d\} \in \mathbb{L}^2$ stationnaire au second ordre est sa fonction de covariance qu'on appelle, des fois, **fonction de covariance stationnaire** de X .

B.1 Propriétés d'un covariogramme

Dans la proposition qui suit, on va résumer quelques propriétés du covariogramme d'un champ aléatoire stationnaire au second ordre :

Proposition 3.3.4 [GAETAN & GUYON (2008)].

Soient $X = \{X(s), s \in \mathbb{R}^d\} \in \mathbb{L}^2$ un champ aléatoire stationnaire au second ordre et C son covariogramme, alors :

1. C est une fonction bornée :

$$\forall h \in \mathbb{R}^d, \quad |C(h)| \leq C(0_{\mathbb{R}^d}) = V[X(s)], \quad s \in \mathbb{R}^d$$

2. C est une fonction paire :

$$\forall h \in \mathbb{R}^d, \quad C(h) = C(-h)$$

3. C est une fonction semi-définie positive :

$$\forall n \in \mathbb{N}^*, \quad \forall a = (a_1, a_2, \dots, a_n) \in \mathbb{R}^n, \quad \forall (s_1, s_2, \dots, s_n) \in (\mathbb{R}^d)^n :$$

$$\sum_{i=1}^n \sum_{j=1}^n a_i a_j C(s_j - s_i) \geq 0$$

4. Si ℓ est une application linéaire définie sur \mathbb{R}^d dans \mathbb{R}^d , alors le champ aléatoire $X^\ell = \{X[\ell(s)], s \in \mathbb{R}^d\}$ est stationnaire au second ordre de covariogramme C^ℓ défini par :

$$\forall h \in \mathbb{R}^d, \quad C^\ell(h) = C[\ell(h)] = C \circ \ell(h)$$

5. Si C est continue à l'origine $0_{\mathbb{R}^d}$, alors C est uniformément continue sur \mathbb{R}^d .

Le résultat de la proposition suivante donne, en quelques sortes, la structure de l'espace des covariogrammes :

Proposition 3.3.5 [GAETAN & GUYON (2008)].

Si $\{C_n, n \in \mathbb{N}^*\}$ est une suite de covariogrammes définis sur \mathbb{R}^d dans \mathbb{R} , alors les fonctions C suivantes, définis sur \mathbb{R}^d dans \mathbb{R} , sont aussi des covariogrammes :

1.

$$\forall h \in \mathbb{R}^d, \quad C(h) = \lambda_1 C_1(h) + \lambda_2 C_2(h), \quad (\lambda_1, \lambda_2) \in \mathbb{R}_+ \times \mathbb{R}_+$$

2.

$$\forall h \in \mathbb{R}^d, \quad C(h) = C_1(h)C_2(h)$$

3.

$$\forall h \in \mathbb{R}^d, \quad C(h) = \lim_{n \rightarrow \infty} C_n(h)$$

B.2 Covariogramme isotropique

Dans toute la suite de ce chapitre, notons par $\|\cdot\|$ la norme euclidienne définie sur \mathbb{R}^d dans \mathbb{R}_+ par :

$$\forall s = (s_1, s_2, \dots, s_d) \in \mathbb{R}^d, \quad \|s\| = \|s\|_2 = \left(\sum_{i=1}^d s_i^2 \right)^{1/2}$$

Définition 3.3.6 [GAETAN & GUYON (2008)].

Soient $X = \{X(s), s \in \mathbb{R}^d\} \in \mathbb{L}^2$ un champ aléatoire stationnaire au second ordre et C son covariogramme. On dit que C est **isotropique** si, pour tout $h \in \mathbb{R}^d$, $C(h)$ ne dépend que de la norme $\|h\|$ du vecteur h et non de sa direction. C'est-à-dire :

$$\forall h \in \mathbb{R}^d, \quad C(h) = \text{Cov}[X(s), X(s+h)] = C(\|h\|), \quad s \in \mathbb{R}^d$$

Remarque : Lorsque le covariogramme d'un champ aléatoire stationnaire au second ordre est isotropique, on dit que ce champ aléatoire est **isotropiquement stationnaire**.

Le théorème suivant, dû à SCHOENBERG (1938), nous donne un résultat qui caractérise les covariogrammes isotropiques pour toutes les dimensions $d \in \mathbb{N}^*$ de l'ensemble spatial des sites \mathbb{R}^d :

Théorème 3.3.1

Soit C une fonction définie sur \mathbb{R}^d dans \mathbb{R} . C est un covariogramme isotropique d'un champ aléatoire isotropiquement stationnaire défini sur l'ensemble spatial des sites \mathbb{R}^d à espace d'état $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ **si et seulement** s'il existe une mesure finie ν strictement positive sur $[\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+)]$ telle que :

$$\forall h \in \mathbb{R}^d, \quad C(h) = C(\|h\|) = \int_0^{+\infty} \exp\left(-\frac{\|h\|^2}{x^2}\right) \nu(dx)$$

Une extension au théorème précédent, est le théorème suivant qui caractérise les covariogrammes isotropiques strictement monotones pour toutes les dimensions $d \in \mathbb{N}^*$ de l'ensemble spatial des sites \mathbb{R}^d :

Théorème 3.3.2

Soit C une fonction définie sur \mathbb{R}^d dans \mathbb{R} . C est un covariogramme isotropique **strictement monotone** d'un champ aléatoire isotropiquement stationnaire défini sur l'ensemble spatial des sites \mathbb{R}^d à espace d'état $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ **si et seulement** s'il existe une mesure finie ν strictement positive sur $[\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+)]$ telle que :

$$\forall h \in \mathbb{R}^d, \quad C(h) = C(\|h\|) = \int_0^{+\infty} \exp\left(-\frac{\|h\|}{x}\right) \nu(dx)$$

C Corrélogramme

De la définition de la stationnarité au second ordre, on a constaté que lorsque un champ aléatoire $X = \{X(s), s \in \mathbb{R}^d\} \in \mathbb{L}^2$ est stationnaire au second ordre, alors sa fonction de covariance ne dépend que d'un seul argument au lieu de deux. Il en est de même pour sa fonction de corrélation. Et, dans ce cas, on appelle cette fonction de corrélation « **le corrélogramme** » de X .

Définition 3.3.7 [CRESSIE (1991)].

Soient $X = \{X(s), s \in \mathbb{R}^d\} \in \mathbb{L}^2$ un champ aléatoire stationnaire au second ordre tel que sa fonction de variance est strictement positive et C son covariogramme. On appelle **corrélogramme** du champ aléatoire X la fonction ϱ définie sur \mathbb{R}^d dans $[-1, 1]$ par :

$$\forall h \in \mathbb{R}^d, \quad \varrho(h) = \frac{\text{Cov}[X(s), X(s+h)]}{\sqrt{V[X(s)]}\sqrt{V[X(s+h)]}} = \frac{C(h)}{C(0_{\mathbb{R}^d})}, \quad s \in \mathbb{R}^d$$

Remarque : Dans le cadre de la définition ci-dessus et si, de plus, le champ aléatoire X est isotropiquement stationnaire, alors la fonction ϱ sera appelé **corrélogramme isotropique** du champ aléatoire X . Et dans ce cas, pour tout $h \in \mathbb{R}^d$, $\varrho(h)$ ne dépendra que de la norme du vecteur h et non de sa direction. C'est-à-dire :

$$\forall h \in \mathbb{R}^d, \quad \varrho(h) = \frac{C(\|h\|)}{C(\|0_{\mathbb{R}^d}\|)} = \frac{C(\|h\|)}{C(0)} = \varrho(\|h\|)$$

D Variogramme

Comme dans le cas de la fonction de covariance et de la fonction de corrélation, le variogramme d'un champ aléatoire stationnaire au second ordre ne dépend que d'un seul argument au lieu de deux. Et dans ce cas, il sera défini comme suit :

Définition 3.3.8 [CRESSIE (1991)].

Soient $X = \{X(s), s \in \mathbb{R}^d\} \in \mathbb{L}^2$ un champ aléatoire stationnaire au second ordre et C son covariogramme. On appelle **semi-variogramme** du champ aléatoire X la fonction γ définie sur \mathbb{R}^d dans \mathbb{R}_+ par :

$$\forall h \in \mathbb{R}^d, \quad \gamma(h) = \frac{1}{2} V[X(s+h) - X(s)] = C(0_{\mathbb{R}^d}) - C(h), \quad s \in \mathbb{R}^d \quad (3.3)$$

Définition 3.3.9 [CRESSIE (1991)].

Soient $X = \{X(s), s \in \mathbb{R}^d\} \in \mathbb{L}^2$ un champ aléatoire stationnaire au second ordre et C son covariogramme. On appelle **variogramme** du champ aléatoire X la fonction ϑ définie sur \mathbb{R}^d dans \mathbb{R}_+ par :

$$\forall h \in \mathbb{R}^d, \quad \vartheta(h) = V[X(s+h) - X(s)] = 2[C(0_{\mathbb{R}^d}) - C(h)], \quad s \in \mathbb{R}^d \quad (3.4)$$

Les définitions respectives (3.3) et (3.4) de semi-variogramme et de variogramme du champ aléatoire $X = \{X(s), s \in \mathbb{R}^d\} \in \mathbb{L}^2$ se justifient par le fait que :

$$\forall (s, h) \in \mathbb{R}^d \times \mathbb{R}^d, \quad V[X(s+h) - X(s)] = V[X(s+h)] + V[X(s)] - 2\text{Cov}[X(s+h), X(s)]$$

Et comme $X = \{X(s), s \in \mathbb{R}^d\} \in \mathbb{L}^2$ est un champ aléatoire stationnaire au second ordre de covariogramme C , alors :

$$\forall (s, h) \in \mathbb{R}^d \times \mathbb{R}^d, \quad \text{Cov}[X(s+h), X(s)] = C(h) \quad \text{et} \quad V[X(s+h)] = V[X(s)] = C(0_{\mathbb{R}^d})$$

Remarque : Dans le cadre des deux définitions ci-dessus et si, de plus, le champ aléatoire X est isotropiquement stationnaire, alors la fonction γ sera appelé **semi-variogramme isotropique** du champ aléatoire X et ϑ **variogramme isotropique** de X . Et dans ce cas, pour tout $h \in \mathbb{R}^d$, $\gamma(h)$ et $\vartheta(h)$ ne dépendront que de la norme du vecteur h et non de sa direction. C'est-à-dire :

$$\forall h \in \mathbb{R}^d, \quad \gamma(h) = C(\|0_{\mathbb{R}^d}\|) - C(\|h\|) = C(0) - C(\|h\|) = \gamma(\|h\|)$$

$$\forall h \in \mathbb{R}^d, \quad \vartheta(h) = 2[C(\|0_{\mathbb{R}^d}\|) - C(\|h\|)] = 2[C(0) - C(\|h\|)] = \vartheta(\|h\|)$$

E Continuité en moyenne quadratique

Dans la proposition qui suit, qui est une conséquence directe de **Proposition 4.2.1**, on va donner un condition nécessaire et suffisante pour qu'un champ aléatoire stationnaire au second ordre soit continu en moyenne quadratique :

Proposition 3.3.6 [GAETAN & GUYON (2008)].

Soient $X = \{X(s), s \in \mathbb{R}^d\} \in \mathbb{L}^2$ un champ aléatoire stationnaire au second ordre et C son covariogramme. X est continu en moyenne quadratique partout sur \mathbb{R}^d si et seulement si C est continu à l'origine $0_{\mathbb{R}^d}$.

3.3.3 Stationnarité intrinsèque

La propriété de stationnarité au second ordre ne peut pas être satisfaite à tout les coups (voir [GAETAN & GUYON (2008)]). En plus, comme le montre **Définition 3.1.13**, le variogramme d'un champ aléatoire en deux sites d'observation ne se définit pas directement à partir des valeurs de ce champ sur ces deux sites, mais plutôt à partir de leur différence. Donc, l'hypothèse de stationnarité au second ordre, qui exige l'existence des deux premiers moments d'un champ aléatoire, est restrictive si l'on souhaite n'utiliser que le variogramme et non la fonction de covariance. Pour ces raisons, les géostatisticiens, en particulier Georges MATHERON (1973), qui favorisent l'utilisation du variogramme plutôt que la fonction de covariance, ont affaibli l'hypothèse de stationnarité au second ordre en introduisant un nouveau type de stationnarité qu'on appelle « **stationnarité intrinsèque** ».

Soit $X = \{X(s), s \in \mathbb{R}^d\}$ un champ aléatoire défini sur l'espace probabilisé (Ω, \mathcal{A}, P) à espace

d'état $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$. Introduisons, pour tout $h \in \mathbb{R}^d$, le champ aléatoire ΔX^h défini sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à espace d'état $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ tel que :

$$\Delta X^h = \left\{ \Delta X^h(s) = X(s+h) - X(s), s \in \mathbb{R}^d \right\}, \quad h \in \mathbb{R}^d$$

Définition 3.3.10 [GAETAN & GUYON (2008)].

On appelle ΔX^h **champ aléatoire des accroissement** du champ aléatoire X .

Définition 3.3.11 [GAETAN & GUYON (2008)].

On dit que le champ aléatoire $X = \{X(s), s \in \mathbb{R}^d\}$ est **intrinsèque** ou **intrinsèquement stationnaire** si, pour tout $h \in \mathbb{R}^d$, son champ aléatoire des accroissement ΔX^h est un champ aléatoire stationnaire au second ordre.

Le résultat de la proposition ci-dessous, qui est évident, nous donne la relation qui existe entre la stationnarité au second ordre et la stationnarité intrinsèque et montre que la première stationnarité est plus forte que la deuxième :

Proposition 3.3.7 [GAETAN & GUYON (2008)].

Si le champ aléatoire $X = \{X(s), s \in \mathbb{R}^d\}$ est stationnaire au second ordre, alors X est un champ aléatoire intrinsèque.

Remarque : De la proposition ci-dessus, on déduit que l'hypothèse de stationnarité au second ordre est plus forte que l'hypothèse de stationnarité intrinsèque et l'entraîne. L'inverse, en revanche, n'est pas vrai, car la stationnarité intrinsèque d'un champ aléatoire n'exige pas l'existence des moments de ce champ, mais seulement l'existence des moments des accroissements de ce champ. Ainsi, la classe des champs aléatoires intrinsèquement stationnaires est plus vaste que la classe des champs aléatoires stationnaires au second ordre.

A Exemples

On en déduit que tous les champs aléatoires strictement stationnaires et tous les champs aléatoires stationnaires au second ordre sont des champs aléatoires intrinsèquement stationnaires.

B Variogramme

Comme on l'a déjà dit, le variogramme d'un champ aléatoire est une alternative de sa fonction de covariance lorsque celle-ci n'existe pas. Dans le cas de stationnarité intrinsèque, nous sommes sûr que le variogramme existe, mais pas la fonction de covariance. C'est pour ça, en applications, il est préférable de travailler avec le variogramme plutôt que la fonction de covariance, car la classe des champs aléatoires intrinsèquement stationnaires est plus large que les autres classes des champs aléatoires de stationnarité différente.

Définition 3.3.12 [GAETAN & GUYON (2008)].

Soit $X = \{X(s), s \in \mathbb{R}^d\}$ un champ aléatoire intrinsèque. On appelle **semi-variogramme** du champ aléatoire X la fonction γ définie sur \mathbb{R}^d dans \mathbb{R}_+ par :

$$\forall h \in \mathbb{R}^d, \quad \gamma(h) = \frac{1}{2} V[X(s+h) - X(s)], \quad s \in \mathbb{R}^d$$

Définition 3.3.13 [GAETAN & GUYON (2008)].

Soit $X = \{X(s), s \in \mathbb{R}^d\}$ un champ aléatoire intrinsèque. On appelle **variogramme** du champ aléatoire X la fonction ϑ définie sur \mathbb{R}^d dans \mathbb{R}_+ par :

$$\forall h \in \mathbb{R}^d, \quad \vartheta(h) = V[X(s+h) - X(s)], \quad s \in \mathbb{R}^d$$

B.1 Propriétés d'un variogramme

Dans la proposition qui suit, on va résumer quelques propriétés du variogramme d'un champ aléatoire intrinsèquement stationnaire :

Proposition 3.3.8 [GAETAN & GUYON (2008)].

Soient $X = \{X(s), s \in \mathbb{R}^d\}$ un champ aléatoire intrinsèquement stationnaire et ϑ son variogramme, alors :

1. ϑ est nul à l'origine :

$$\vartheta(0_{\mathbb{R}^d}) = 0$$

2. ϑ est une fonction positive :

$$\forall h \in \mathbb{R}^d, \quad \vartheta(h) \geq \vartheta(0_{\mathbb{R}^d}) = 0$$

3. ϑ est une fonction paire :

$$\forall h \in \mathbb{R}^d, \quad \vartheta(h) = \vartheta(-h)$$

4. ϑ est une fonction **conditionnellement définie négative** :

$$\forall n \in \mathbb{N}^*, \quad \forall a = (a_1, a_2, \dots, a_n) \in \mathbb{R}^n \quad \text{tel que} \quad \sum_{i=1}^n a_i = 0 \quad \text{et} \quad \forall (s_1, s_2, \dots, s_n) \in (\mathbb{R}^d)^n :$$

$$\sum_{i=1}^n \sum_{j=1}^n a_i a_j \vartheta(s_j - s_i) \leq 0$$

5. **Si** ℓ est une application linéaire définie sur \mathbb{R}^d dans \mathbb{R}^d , **alors** le champ aléatoire $X^\ell = \{X[\ell(s)], s \in \mathbb{R}^d\}$ est intrinsèquement stationnaire de variogramme ϑ^ℓ défini par :

$$\forall h \in \mathbb{R}^d, \quad \vartheta^\ell(h) = \vartheta[\ell(h)] = \vartheta \circ \ell(h)$$

6. **Si** ϑ est borné au voisinage de l'origine $0_{\mathbb{R}^d}$, **alors** :

$$\exists a \in \mathbb{R}_+, \quad \exists b \in \mathbb{R}_+ \quad \text{tels que :} \quad \forall h \in \mathbb{R}^d, \quad \vartheta(h) \leq a\|h\|^2 + b$$

7. **Si** ϑ est continu à l'origine $0_{\mathbb{R}^d}$, **alors** ϑ est continu en tout $h \in \mathbb{R}^d$ où ϑ est localement borné.

Remarque : Dans le cadre de la proposition ci-dessus, on remarque que la quatrième assertion est une condition nécessaire pour que ϑ soit un variogramme d'un champ aléatoire intrinsèquement stationnaire. En fait, elle est également suffisante.

Théorème 3.3.3

Soit ϑ une fonction définie sur \mathbb{R}^d dans \mathbb{R}_+ . La fonction ϑ est un variogramme d'un champ aléatoire intrinsèque défini sur l'ensemble spatial des sites \mathbb{R}^d à espace d'état $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ **si et seulement si** ϑ est conditionnellement définie négative, c'est-à-dire :

$$\forall n \in \mathbb{N}^*, \quad \forall a = (a_1, a_2, \dots, a_n) \in \mathbb{R}^n \quad \text{tel que} \quad \sum_{i=1}^n a_i = 0 \quad \text{et} \quad \forall (s_1, s_2, \dots, s_n) \in (\mathbb{R}^d)^n :$$

$$\sum_{i=1}^n \sum_{j=1}^n a_i a_j \vartheta(s_j - s_i) \leq 0$$

Le résultat de la proposition suivante donne, en quelques sortes, la structure de l'espace des variogrammes :

Proposition 3.3.9 [GAETAN & GUYON (2008)].

Si $\{\vartheta_n, n \in \mathbb{N}^*\}$ est une suite de variogrammes définis sur \mathbb{R}^d dans \mathbb{R}_+ , **alors** les fonctions ϑ suivantes, définies sur \mathbb{R}^d dans \mathbb{R}_+ , sont aussi des variogrammes :

1.

$$\forall h \in \mathbb{R}^d, \quad \vartheta(h) = \lambda_1 \vartheta_1(h) + \lambda_2 \vartheta_2(h), \quad (\lambda_1, \lambda_2) \in \mathbb{R}_+ \times \mathbb{R}_+$$

2.

$$\forall h \in \mathbb{R}^d, \quad \vartheta(h) = \lim_{n \rightarrow \infty} \vartheta_n(h)$$

B.2 Variogramme et stationnarité au second ordre

Comme on l'a déjà dit, un champ aléatoire stationnaire au second ordre est un champ aléatoire intrinsèque. Donc, son variogramme existe. Dans la théorème qui suit, on va donner une condition nécessaire et suffisante pour qu'un champ aléatoire intrinsèque soit stationnaire au second ordre :

Théorème 3.3.4

Soit $X = \{X(s), s \in \mathbb{R}^d\}$ un champ aléatoire intrinsèque. X est un champ aléatoire stationnaire au second ordre **si et seulement si** son variogramme est borné.

B.3 Modèles d'un variogramme

Comme on l'a déjà dit, la fonction de covariance (si elle existe) d'un champ aléatoire mesure la force de la relation qui lie les variables aléatoires qui représentent ce champ aléatoire. Dans le cas de stationnarité au second ordre, on est sûr que cette fonction existe et on l'appelle covariogramme. Donc, on peut caractériser un modèle de champs aléatoires stationnaires au second ordre par son

covariogramme ou par son variogramme puisque si C est son covariogramme et ϑ son variogramme, on a :

$$\forall h \in \mathbb{R}^d, \quad \vartheta(h) = 2[C(0_{\mathbb{R}^d}) - C(h)] \iff C(h) = C(0_{\mathbb{R}^d}) - \frac{1}{2}\vartheta(h)$$

Dans le cas où le champ aléatoire est, simplement, intrinsèquement stationnaire, on le caractérisera par son variogramme.

Dans cette section, on va essayer de donner quelques modèles d'un semi-variogramme d'un champ aléatoire intrinsèquement stationnaire qui sont aussi des modèles de son variogramme, car si γ est le semi-variogramme de ce champ et ϑ son variogramme, alors on a :

$$\forall h \in \mathbb{R}^d, \quad \gamma(h) = \frac{1}{2}\vartheta(h) \iff \vartheta(h) = 2\gamma(h)$$

B.3.1 Modèle effet de pépité (Nugget Effect Model)

$$\forall h \in \mathbb{R}^d, \quad \gamma(h) = \begin{cases} C_0 & \text{si } h \neq 0_{\mathbb{R}^d} \\ 0 & \text{si } h = 0_{\mathbb{R}^d} \end{cases}, \quad C_0 \geq 0$$

B.3.2 Modèle stable (Modèle exponentiel généralisé)

$$\forall h \in \mathbb{R}^d, \quad \gamma(h) = \begin{cases} C_0 + b \left[1 - \exp \left(- \frac{\|h\|^\alpha}{a^\alpha} \right) \right] & \text{si } h \neq 0_{\mathbb{R}^d} \\ 0 & \text{si } h = 0_{\mathbb{R}^d} \end{cases}$$

$$C_0 \geq 0, \quad a > 0, \quad b \geq 0 \quad \text{et} \quad 0 < \alpha \leq 2$$

B.3.3 Modèle gaussien

Si on pose $\alpha = 2$ dans le modèle stable défini ci-dessus, alors on obtient un nouveau modèle qu'on appelle « modèle gaussien » dont la définition est la suivante :

$$\forall h \in \mathbb{R}^d, \quad \gamma(h) = \begin{cases} C_0 + b \left[1 - \exp \left(- \frac{\|h\|^2}{a^2} \right) \right] & \text{si } h \neq 0_{\mathbb{R}^d} \\ 0 & \text{si } h = 0_{\mathbb{R}^d} \end{cases}$$

$$C_0 \geq 0, \quad a > 0 \quad \text{et} \quad b \geq 0$$

B.3.4 Modèle exponentiel

Si on pose $\alpha = 1$ dans le modèle stable, alors on obtient un nouveau modèle qu'on appelle « modèle exponentiel » dont la définition est la suivante :

$$\forall h \in \mathbb{R}^d, \quad \gamma(h) = \begin{cases} C_0 + b \left[1 - \exp \left(- \frac{\|h\|}{a} \right) \right] & \text{si } h \neq 0_{\mathbb{R}^d} \\ 0 & \text{si } h = 0_{\mathbb{R}^d} \end{cases}$$

$$C_0 \geq 0, \quad a > 0 \quad \text{et} \quad b \geq 0$$

B.3.5 Modèle sphérique

Le modèle sphérique est un modèle qui n'est valide que sur \mathbb{R}^d où $d \in \{1, 2, 3\}$. Il est défini comme suit :

$$\forall h \in \mathbb{R}^d (d = 1, 2, 3), \quad \gamma(h) = \begin{cases} C_0 + b \left(\frac{3\|h\|}{2a} + \frac{\|h\|^3}{2a^3} \right) & \text{si } 0 < \|h\| \leq a \\ C_0 + b & \text{si } \|h\| > a \\ 0 & \text{si } h = 0_{\mathbb{R}^d} \end{cases}$$

$C_0 \geq 0, a > 0$ et $b \geq 0$

B.3.6 Modèle de WHITTLE-MATÉRN

La classe des modèles de WHITTLE-MATÉRN est une classe très importante en applications, car elle engendre beaucoup d'autres modèles (il suffit de jouer sur ses paramètres). Elle est définie comme suit :

$$\forall h \in \mathbb{R}^d, \quad \gamma(h) = \begin{cases} C_0 + b \left[1 - \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\|h\|}{a} \right)^\nu \mathcal{K}_\nu \left(\frac{\|h\|}{a} \right) \right] & \text{si } h \neq 0_{\mathbb{R}^d} \\ 0 & \text{si } h = 0_{\mathbb{R}^d} \end{cases}$$

$C_0 \geq 0, a > 0, b \geq 0$ et $\nu > 0$

Remarque : \mathcal{K}_ν qui apparait dans le modèle de WHITTLE-MATÉRN est une fonction de BESSEL modifiée de deuxième espèce de paramètre ν . Et Γ est la fonction Gamma définie par :

$$\forall x > 0, \quad \Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$$

B.3.7 Modèle puissance

$$\forall h \in \mathbb{R}^d, \quad \gamma(h) = \begin{cases} C_0 + b\|h\|^\lambda & \text{si } h \neq 0_{\mathbb{R}^d} \\ 0 & \text{si } h = 0_{\mathbb{R}^d} \end{cases}, \quad C_0 \geq 0, b \geq 0 \text{ et } 0 \leq \lambda < 2$$

B.3.8 Modèle linéaire

$$\forall h \in \mathbb{R}^d, \quad \gamma(h) = \begin{cases} C_0 + b\|h\| & \text{si } h \neq 0_{\mathbb{R}^d} \\ 0 & \text{si } h = 0_{\mathbb{R}^d} \end{cases}, \quad C_0 \geq 0 \text{ et } b \geq 0$$

B.3.9 Modèle triangulaire

$$\forall h \in \mathbb{R}^d, \quad \gamma(h) = \begin{cases} C_0 + b \left(\frac{\|h\|}{a} \right) & \text{si } 0 \leq \|h\| \leq a \\ C_0 + b & \text{si } \|h\| > a \end{cases}, \quad C_0 \geq 0, \quad a > 0 \quad \text{et} \quad b \geq 0$$

B.3.10 Modèle de CAUCHY

$$\forall h \in \mathbb{R}^d, \quad \gamma(h) = \begin{cases} C_0 + b \left[1 - \left(1 + \frac{\|h\|^2}{a^2} \right)^{-\lambda} \right] & \text{si } h \neq 0_{\mathbb{R}^d} \\ 0 & \text{si } h = 0_{\mathbb{R}^d} \end{cases}$$

$$C_0 \geq 0, \quad a > 0, \quad b \geq 0 \quad \text{et} \quad \lambda \geq 0$$

B.3.11 Modèle de DE WIJSIAN

$$\forall h \in \mathbb{R}^d, \quad \gamma(h) = \begin{cases} \frac{3}{2} \ln(\|h\|^2 + a) & \text{si } h \neq 0_{\mathbb{R}^d} \\ 0 & \text{si } h = 0_{\mathbb{R}^d} \end{cases}, \quad a \geq 0 \quad \text{et} \quad b \geq 0$$

B.3.12 Modèle rationnel quadratique (Rational Quadratic Model)

$$\forall h \in \mathbb{R}^d, \quad \gamma(h) = \begin{cases} C_0 + b \left(\frac{\|h\|^2}{1 + \|h\|^2} \right) & \text{si } h \neq 0_{\mathbb{R}^d} \\ 0 & \text{si } h = 0_{\mathbb{R}^d} \end{cases}, \quad C_0 \geq 0 \quad \text{et} \quad b \geq 0$$

B.3.13 Modèle effet de trou (Hole Effect Model)

$$\forall h \in \mathbb{R}^d, \quad \gamma(h) = \begin{cases} C_0 + b \left[1 - a \|h\|^{-1} \sin \left(\frac{\|h\|}{a} \right) \right] & \text{si } h \neq 0_{\mathbb{R}^d} \\ 0 & \text{si } h = 0_{\mathbb{R}^d} \end{cases}$$

$$C_0 \geq 0, \quad a > 0 \quad \text{et} \quad b \geq 0$$

Référence : Pour une bonne compréhension et interprétation des paramètres des modèles définis ci-dessus, voir l'utilité de ces modèles dans les applications et qui sont issus de la stationnarité au second ordre et ceux qui sont issus de la stationnarité intrinsèque seulement, le lecteur pourra consulter les références suivantes : [JOURNAL & HUIJBREGTS (1978)], [DIGGLE & RIBEIRO JR (2007)], [GAETAN & GUYON (2008)], [LE & ZIDEK (2006)] et [CRESSIE (1991)].

C Continuité en moyenne quadratique

Le résultat de la proposition ci dessous nous donne une condition suffisante pour qu'un champ aléatoire intrinsèque soit continu en moyenne quadratique :

Proposition 3.3.10 [CRESSIE (1991)].

Soient $X = \{X(s), s \in \mathbb{R}^d\}$ un champ aléatoire intrinsèque et ϑ son variogramme. **Si** ϑ est continue à l'origine $0_{\mathbb{R}^d}$, **alors** X est un champ aléatoire continu en moyenne quadratique partout sur \mathbb{R}^d .

Conclusion

La théorie des champs aléatoires est une théorie très vaste. Elle est, en quelques sortes, la généralisation de la théorie des processus stochastiques et des séries temporelles. Dans ce chapitre, on s'est limité à donner les outils de base des champs aléatoires, à savoir la définition, les moments, la stationnarité, les propriétés de la fonction de covariance et les propriétés du variogramme, car ces outils sont fondamentaux pour mettre en application la théorie des champs aléatoires. Tout de même, on tient à dire qu'il y a beaucoup de choses qui ne sont pas dites, comme l'ergodicité, la représentation spectrale, les propriétés géométriques (continuité et différentiabilité), etc.

CHAPITRE 4

CHAMPS ALÉATOIRES MAX-STABLES

Repères historiques et bibliographiques

La théorie des valeurs extrêmes a, d'abord, été développée dans un contexte univarié, c'est-à-dire pour des observations réelles unidimensionnelles, et pour les échantillons indépendants. La théorie probabiliste concernant la caractérisation des lois max-stables et de leurs max-domaines d'attraction débute, en 1927, par les travaux [FRÉCHET (1927)], [FISHER & TIPPETT (1928)], [VON MISES (1936)], [GNEDENKO (1943)], [GUMBEL (1958)], etc. Les aspects statistiques de la théorie univariée ont été étudiés plus tardivement, à partir de 1975 dans les travaux [PICKANDS (1975)] et [HILL (1975)].

Le cas des extrêmes d'échantillons dépendants (suites stationnaires de variables aléatoires réelles) a été considéré dans les années 50 et a une histoire riche. Différentes conditions de dépendance ont été étudiées. Les livres [GALAMBOS (1987)] et [LEADBETTER, LINDGREN & ROOTZÉN (1983)] offrent une synthèse de ces développements.

La théorie multivariée des extrêmes débute dans les années 60 avec les travaux [GUMBEL (1958)], [SIBUYA (1960)] et [TIAGO DE OLIVEIRA (1962)] dans un cadre bivarié. Mais elle ne prend son essor qu'à la fin des années 70 et est aujourd'hui encore un thème de recherche dynamique. La caractérisation des lois max-indéfiniment divisibles en dimension finie est due à [BROWN & RESNICK (1977)], le cas max-stable à [DE HAAN (1978)], [Deheuvels (1978)] et [Deheuvels (1984)]. La caractérisation des extrêmes multivariés est également abordée dans [HEFFERNAN & TAWN (2004)] et [HEFFERNAN & RESNICK (2007)]. Les applications statistiques dans un cadre multivarié sont considérées entre autre dans [COLES & TAWN (1991)], [SCHLATHER & TAWN (2002)] et [SCHLATHER & TAWN (2003)] avec l'étude des propriétés des coefficients extrémaux, ou encore [ABDOUS, FOUGÈRES, GHOUDI & SOULIER (2008)] et [FOUGÈRES & SOULIER (2010)] avec des modèles de dépendance en présence de conditionnement extrêmes.

Le contexte fonctionnel et la théorie des champs aléatoires max-stables connaissent un important développement du milieu des années 80 jusqu'à aujourd'hui. La structure des champs aléatoires max-stables a été donnée dans [DE HAAN (1984)] et le cas max-indéfiniment divisible dans [GINÉ,

HAHN & VATAN (1990)]. Citons également les travaux [DE HAAN & PICKANDS (1986)], [RESNICK & ROY (1991)]. Différents modèles de champs aléatoires max-stables ont été introduits dans [SMITH (1990)], [SCHLATHER (2002)] et [KABLUCHKO, SCHLATHER & DE HAAN (2009)] avec les fameux processus de BROWN-RESNICK. Les contributions récentes sur les propriétés des champs aléatoires max-stables incluent les travaux [KABLUCHKO & SCHLATHER (2010)] sur les questions d'ergodicité et de mélange, [KABLUCHKO (2011)] en lien avec les extrêmes d'un champ aléatoire gaussien ou encore [STOEV (2010)] et [KABLUCHKO (2009)] sur les représentations des champs aléatoires max-stables, etc. Cependant, une importante différence entre le cas univarié et le cas spatial est qu'il n'y a pas de modèles paramétriques qui peuvent représenter entièrement les champs aléatoires max-stables. D'où l'inférence statistique s'est concentrée principalement sur les cas bivariés spéciaux et la modélisation de la structure de dépendance spatiale des extrêmes qui va avec qui sont à la base de beaucoup de travaux ces dernières années. Concernant les applications et les aspects statistiques dans un cadre spatial, [COLES & TAWN (1996)], [DE HAAN & PEREIRA (2006)], [BUISHAND, DE HAAN & ZHOU (2008)] ou encore [COOLEY, NYCHKA & NAVEAU (2007)], etc. [DAVISON, PADOAN & RIBATET (2012)] présente une revue récente des méthodes statistiques pour les extrêmes spatiaux et contient de nombreuses références bibliographiques sur cet aspect de la théorie.

Note : Faisons remarque que l'introduction ci-dessus est inspirée de la thèse [DOMBRY (2012)] présentée le 08 Novembre 2012 par Clément DOMBRY sous la direction de Laurens DE HAAN, Gennady SAMORODNITSKY et Philippe SOULIER pour l'obtention de l'habilitation à diriger des recherches. Donc toutes les références citées dans cette introduction sont incluses dans la bibliographie de cette thèse.

Remarque : Beaucoup de résultats de la théorie des champs aléatoires max-stables sont liés directement aux résultats de la la théorie des valeurs extrêmes multivariée. On conseille le lecteur de voir cette théorie qu'il peut trouver dans les ouvrages suivants : [RESNICK (1987)], [DE HAAN & FERREIRA (2006)], [COLES (2001)], [FALK, HÜSLER & REISS (2011)], [FINKENSTÄDT & ROOTZÉN (2004)] et [KOTZ & NADARAJAH (2000)].

4.1 Max-stabilité

4.1.1 Définition

Soit $Y = \{Y(x), x \in \mathbb{R}^d\}$ ($d \geq 1$) un champ aléatoire ou un processus stochastique défini sur l'espace probabilisé (Ω, \mathcal{A}, P) à espace d'état $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$.

Définition 4.1.1 [STOEV (2010)].

*On dit que le champ aléatoire $Y = \{Y(x), x \in \mathbb{R}^d\}$ est **max-stable**, ou Y est un **processus max-stable**, si toute les lois de probabilité de sa distribution finie-dimensionnelle sont max-stables.*

Remarque : On sait que les lois de probabilité de la distribution finie-dimensionnelle d'un champ aléatoire sont des lois de probabilité des vecteurs aléatoires de dimensions finies. Pour voir la définition et les propriétés de la max-stabilité d'une loi de probabilité multidimensionnelle, qui est une extension de la max-stabilité d'une loi de probabilité unidimensionnelle au cas multidimensionnel, le lecteur peut se référer au CHAPITRE 5 de l'ouvrage [RESNICK (1987)].

La définition d'un champ aléatoire max-stable, souvent rencontrée dans la littérature et qui est plus proche de la définition d'une loi de probabilité max-stable unidimensionnelle donnée dans **Définition 2.1.1** de ce mémoire, est la suivante :

Définition 4.1.2 [SMITH (1990)].

On dit que le champ aléatoire $Y = \{Y(x), x \in \mathbb{R}^d\}$ est **max-stable** si, pour toute n **copies indépendantes** Y_1, Y_2, \dots, Y_n ($n \in \mathbb{N}^*$) de Y , il existe une fonction continue a_n strictement positive et une fonction continue b_n réelle définies sur \mathbb{R}^d telles que :

$$\bigvee_{i=1}^n \frac{Y_i - b_n}{a_n} = \frac{\bigvee_{i=1}^n Y_i - b_n}{a_n} \stackrel{d}{=} Y \quad \text{avec} \quad \bigvee_{i=1}^n Y_i = \max_{1 \leq i \leq n} Y_i$$

Remarque : Bien sûr, la définition d'un maximum de n champs aléatoires ($n \in \mathbb{N}^*$) n'est pas vraiment claire. Dans le cadre de la définition ci-dessus, on parle de **maximum par composante** (Componentwise Maxima). C'est-à-dire, le champ aléatoire $\bigvee_{i=1}^n Y_i$ maximum des champs aléatoires Y_i ($i = 1, 2, \dots, n$), où $Y_i = \{Y_i(x), x \in \mathbb{R}^d\}$, est défini comme suit :

$$\bigvee_{i=1}^n Y_i = \left\{ \bigvee_{i=1}^n Y_i(x_j), \quad x_j \in \mathbb{R}^d \right\}, \quad j \in J \quad \text{et} \quad J \subset \mathbb{N}^*$$

Donc, le champ aléatoire $Y = \{Y(x), x \in \mathbb{R}^d\}$ est max-stable si, pour tout n champs aléatoires Y_1, Y_2, \dots, Y_n ($n \in \mathbb{N}^*$) indépendants et de même loi spatiale que celle de Y , il existe une fonction continue a_n strictement positive définie sur \mathbb{R}^d et une fonction continue b_n réelle définie, elle aussi, sur \mathbb{R}^d telles que la loi spatiale du champ aléatoire Y est la même que celle du champ aléatoire $a_n^{-1}(\bigvee_{i=1}^n Y_i - b_n)$ maximum par composante normalisé des Y_i ($i = 1, 2, \dots, n$), avec :

$$\frac{\bigvee_{i=1}^n Y_i - b_n}{a_n} = \left\{ \frac{\bigvee_{i=1}^n Y_i(x_j) - b_n(x_j)}{a_n(x_j)}, \quad x_j \in \mathbb{R}^d \right\}, \quad j \in J \quad \text{et} \quad J \subset \mathbb{N}^*$$

On peut clairement voir que dans le cas où le cardinal $|J|$ de l'ensemble d'indices J est fini, c'est-à-dire que l'ensemble spatial des sites contient un nombre fini de sites, alors la définition précédente est précisément la définition d'un vecteur aléatoire max-stable à valeurs dans $\mathbb{R}^{|J|}$ (voir [DE HAAN & RESNICK (1977)]). Et si $|J| = 1$, c'est-à-dire que l'ensemble spatial des sites contient un seul site, alors elle est exactement **Définition 2.1.1** du deuxième chapitre de ce mémoire qui est la définition d'une variable aléatoire max-stable à valeurs dans \mathbb{R} .

De ce qu'on vient de voir, on peut dire que la définition précédente (**Définition 4.1.2**) est une extension de la max-stabilité au cas spatial.

4.1.2 Caractérisation

Dans le théorème suivant, on va donner une caractérisation importante d'un champ aléatoire max-stable. Cette caractérisation, qu'on rencontre souvent sous une forme de définition de la max-stabilité, est un résultat très important dans la théorie des valeurs extrêmes spatiale qu'on peut considérer comme une généralisation de « Théorème de FISHER-TIPPETT (1928) & GNEDENKO (1943) ».

Théorème 4.1.1 [RIBATET (2009) et PADOAN, RIBATET & SISSON (2009)].

Soient Y_1, Y_2, \dots, Y_n ($n \in \mathbb{N}^*$) n champs aléatoires, définis sur l'ensemble spatial des sites \mathbb{R}^d à espace d'état \mathbb{R} , **indépendants** et **de même loi spatiale** et Z un champ aléatoire non dégénéré défini, lui aussi, sur l'ensemble spatial des sites \mathbb{R}^d à espace d'état \mathbb{R} . **S'il existe une fonction continue a_n strictement positive définie sur \mathbb{R}^d et une fonction continue b_n réelle définie, elle aussi, sur \mathbb{R}^d telles que :**

$$\bigvee_{i=1}^n \frac{Y_i - b_n}{a_n} \xrightarrow{d} Z \quad \text{quand } n \rightarrow \infty \quad (4.1)$$

alors :

$$(4.1) \iff Z \text{ est un champ aléatoire } \mathbf{max-stable}$$

Interprétation : Dans ce qui suit et en utilisant les notations du théorème ci-dessus, on va essayer d'expliquer le résultat donné dans ce théorème :

Posons :

$$Z = \left\{ Z(x), \quad x \in \mathbb{R}^d \right\} \quad \text{et} \quad Y_i = \left\{ Y_i(x), \quad x \in \mathbb{R}^d \right\} \quad \text{avec } i = 1, 2, \dots, n \quad (n \in \mathbb{N}^*)$$

Sachant que le champ aléatoire maximum des champs aléatoires Y_i ($i = 1, 2, \dots, n$) est donné par :

$$\bigvee_{i=1}^n \frac{Y_i - b_n}{a_n} = \left\{ \frac{\bigvee_{i=1}^n Y_i(x) - b_n(x)}{a_n(x)}, \quad x \in \mathbb{R}^d \right\}$$

Alors la limite donnée dans (4.1) s'interprète comme suit :

$$(4.1) \iff \forall z \in \mathbb{R}, \quad \left\{ \lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{\bigvee_{i=1}^n Y_i(x) - b_n(x)}{a_n(x)} \leq z \right), \quad x \in \mathbb{R}^d \right\} = \left\{ \mathbb{P}[Z(x) \leq z], \quad x \in \mathbb{R}^d \right\}$$

C'est-à-dire :

$$\forall z \in \mathbb{R}, \quad \lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{\bigvee_{i=1}^n Y_i(x) - b_n(x)}{a_n(x)} \leq z \right) = \mathbb{P}[Z(x) \leq z], \quad x \in \mathbb{R}^d$$

Or, les hypothèses du **Théorème 4.1.1** stipulent que, pour tout $x \in \mathbb{R}^d$, les variables aléatoires réelles $Y_i(x)$ ($i = 1, 2, \dots, n$) sont indépendantes et identiquement distribuées et que la limite donnée dans (4.1) existe. Alors d'après Théorème de FISHER-TIPPETT (1928) & GNEDENKO (1943), les variables aléatoires réelles $Z(x)$ sont extrémales, c'est-à-dire qu'elles sont des variables aléatoires ayant des distributions des valeurs extrêmes généralisées (GEVD) comme lois de probabilité.

Conclusion : De ce qu'on vient de voir, on peut déduire que les variables aléatoires réelles qui représentent un champ aléatoire max-stable ou les réalisations d'un champ aléatoire max-stable ou, encore, les lois de probabilité marginales d'un champ aléatoire max-stable sont des **distributions des valeurs extrêmes généralisées (GEVD)**.

Formellement : Si $Z = \{Z(x), x \in \mathbb{R}^d\}$ est un champ aléatoire max-stable, alors les variables aléatoires réelles $Z(x)$ ($x \in \mathbb{R}^d$) ont des GEVD comme lois de probabilité. C'est-à-dire qu'il existe des fonctions continues ξ réelle, μ réelle et σ strictement positive définies sur \mathbb{R}^d telles que :

$$\forall z \in \mathbb{R}, \quad \mathbb{P}[Z(x) \leq z] = \exp \left\{ - \left[1 + \xi(x) \left(\frac{z - \mu(x)}{\sigma(x)} \right) \right]^{-1/\xi(x)} \right\}, \quad x \in \mathbb{R}^d$$

avec :

$$1 + \xi(x) \left(\frac{z - \mu(x)}{\sigma(x)} \right) > 0$$

Donc, pour un certain $x \in \mathbb{R}^d$:

1. Si $\xi(x) > 0$, la variable aléatoire réelle $Z(x)$ aura une distribution des valeurs extrêmes de type FRÉCHET. C'est-à-dire qu'il existe des constantes $\alpha_1 > 0$, $\mu_1 \in \mathbb{R}$ et $\sigma_1 > 0$ telles que :

$$\forall z \in \mathbb{R}, \quad \mathbb{P}[Z(x) \leq z] = \begin{cases} \exp \left\{ - \left(\frac{z - \mu_1}{\sigma_1} \right)^{-\alpha_1} \right\} & \text{si } z > \mu_1 \\ 0 & \text{si } z \leq \mu_1 \end{cases}$$

2. Si $\xi(x) < 0$, la variable aléatoire réelle $Z(x)$ aura une distribution des valeurs extrêmes de type WEIBULL. C'est-à-dire qu'il existe des constantes $\alpha_2 > 0$, $\mu_2 \in \mathbb{R}$ et $\sigma_2 > 0$ telles que :

$$\forall z \in \mathbb{R}, \quad \mathbb{P}[Z(x) \leq z] = \begin{cases} \exp \left\{ - \left(\frac{\mu_2 - z}{\sigma_2} \right)^{\alpha_2} \right\} & \text{si } z \leq \mu_2 \\ 1 & \text{si } z > \mu_2 \end{cases}$$

3. Si $\xi(x) = 0$, la variable aléatoire réelle $Z(x)$ aura une distribution des valeurs extrêmes de type GUMBEL. C'est-à-dire qu'il existe des constantes $\mu_3 \in \mathbb{R}$ et $\sigma_3 > 0$ telles que :

$$\forall z \in \mathbb{R}, \quad \mathbb{P}[Z(x) \leq z] = \exp \left\{ - \exp \left(\frac{\mu_3 - z}{\sigma_3} \right) \right\}$$

Conclusion générale : Du théorème précédent on peut tirer une conclusion importante dans la théorie des valeurs extrêmes : Les champs aléatoires max-stables sont la généralisation des distributions des valeurs extrêmes au contexte spatial. Ce sont donc des candidats idéaux pour la modélisation des extrêmes spatiaux.

4.1.3 Standardisation

Comme on vient de le voir, les lois de probabilité marginales d'un champ aléatoire max-stable sont un mélange de distributions des valeurs extrêmes. Pour les applications, il est toujours préférable de standardiser ces marginales en leur donnant une forme unique.

Dans la proposition qui suit, on va donner un résultat qui nous permettra, sans perte de généralisation et sans perte de la propriété de max-stabilité, de standardiser les lois de probabilité marginales d'un champ aléatoire max-stable en les transformant en distributions de **FRÉCHET unitaires** ou **standards**.

Soit $X = (X_1, X_2, \dots, X_m)$ ($m \in \mathbb{N}^*$) un vecteur aléatoire défini sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans \mathbb{R}^m de fonction de répartition F , c'est-à-dire :

$$\forall (x_1, x_2, \dots, x_m) \in \mathbb{R}^m, \quad \mathbb{P}(X_1 \leq x_1, X_2 \leq x_2, \dots, X_m \leq x_m) = F(x_1, x_2, \dots, x_m)$$

Et soit F_i ($i = 1, 2, \dots, m$) les fonctions de répartition des lois de probabilité marginales du vecteur aléatoire X définies comme suit :

$$\forall x \in \mathbb{R}, \quad F_i(x) = \mathbb{P}(X_i \leq x), \quad i = 1, 2, \dots, m$$

Proposition 4.1.1 [RESNICK (1987)].

Supposons que les lois de probabilité marginales du vecteur aléatoire X sont continues, et définissons pour tout $i \in \{1, 2, \dots, m\}$ les fonctions ψ_i et G par :

$$\forall y > 0, \quad \psi_i(y) = \left(\frac{1}{-\ln G_i} \right)^{\leftarrow} (y)$$

$$\forall (y_1, y_2, \dots, y_m) \in \mathbb{R}_+^m, \quad G(y_1, y_2, \dots, y_m) = F(\psi_1(y_1), \psi_2(y_2), \dots, \psi_m(y_m))$$

alors :

1. G est une fonction de répartition d'une loi de probabilité de dimension m avec des marginales de **FRÉCHET unitaires**, c'est-à-dire :

$$\forall y > 0, \quad G_i(y) = e^{-1/y}, \quad i = 1, 2, \dots, m$$

2. La loi de probabilité du vecteur aléatoire X est max-stable **si et seulement si** la loi de probabilité ayant G comme fonction de répartition est max-stable.

Remarque : Dans le cadre de la proposition ci-dessus, on désigne par $(\cdot)^{\leftarrow}$ l'inverse généralisée définie dans le CHAPITRE 2 de ce mémoire. Et pour plus de détails sur le résultat, le lecteur peut consulter le livre [RESNICK (1987)].

Résultat : De la proposition ci-dessus, on peut déduire qu'on peut transformer les lois de probabilité multidimensionnelles de la distribution finie-dimensionnelle d'un champ aléatoire max-stable, qui sont max-stables, en lois de probabilité multidimensionnelles max-stables ayant des marginales de FRÉCHET unitaires. D'où, sans perte de généralisation et de la propriété de max-stabilité, on peut standardiser les lois de probabilité marginales d'un champ aléatoire max-stable en les transformant en lois de probabilité de FRÉCHET unitaires.

Définition 4.1.3 [DE HAAN & FERREIRA (2006)].

Soit $Y = \{Y(x), x \in \mathbb{R}^d\}$ un champ aléatoire max-stable défini sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à espace d'état \mathbb{R} . On dit que Y est **max-stable simple** si ses lois de probabilité marginales sont de FRÉCHET unitaires, c'est-à-dire :

$$\forall y \in \mathbb{R}, \quad \mathbb{P}[Y(x) \leq y] = \begin{cases} e^{-1/y} & \text{si } y > 0 \\ 0 & \text{si } y \leq 0 \end{cases}, \quad \forall x \in \mathbb{R}^d$$

Dans toute la suite de ce chapitre, sauf mention du contraire, on ne travaillera qu'avec les champs aléatoires max-stables simples. Dans ce cas, les constantes ou les fonctions de normalisation a_n et b_n utilisées dans **Définition 4.1.2** et **Théorème 4.1.1** seront données par :

$$\forall x \in \mathbb{R}^d, \quad a_n(x) = n \quad \text{et} \quad b_n(x) = 0$$

Définition 4.1.4 [DE HAAN & FERREIRA (2006)].

Soit $Y = \{Y(x), x \in \mathbb{R}^d\}$ un champ aléatoire défini sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à espace d'état \mathbb{R} . Y est un champ aléatoire **max-stable simple** si et seulement si, pour toute n **copies indépendantes** Y_1, Y_2, \dots, Y_n ($n \in \mathbb{N}^*$) de Y , on a :

$$\frac{1}{n} \bigvee_{i=1}^n Y_i \stackrel{d}{=} Y$$

Théorème 4.1.2 [RIBATET (2009) et PADOAN, RIBATET & SISSON (2009)].

Soient Y_1, Y_2, \dots, Y_n ($n \in \mathbb{N}^*$) n champs aléatoires, définis sur l'ensemble spatial des sites \mathbb{R}^d à espace d'état \mathbb{R} , **indépendants** et **de même loi spatiale** et Z un champ aléatoire non dégénéré défini, lui aussi, sur l'ensemble spatial des sites \mathbb{R}^d à espace d'état \mathbb{R} . **Supposons que :**

$$\frac{1}{n} \bigvee_{i=1}^n Y_i \xrightarrow{d} Z \quad \text{quand} \quad n \rightarrow \infty \quad (4.2)$$

alors :

$$(4.2) \iff Z \text{ est un champ aléatoire } \mathbf{max-stable simple}$$

4.2 Représentations spectrales

Dans cette section, on va voir qu'on peut caractériser les champs aléatoires max-stable simples stationnaires au second ordre par leur représentations spectrales. Ces représentations spectrales nous donnent des méthodes pour construire de tels champs aléatoires. Mais avant de voir ces représentations, le lecteur est conseillé de lire **ANNEXE C** de ce mémoire qui parle des processus ponctuels de POISSON.

Remarque : Dans toute la suite de ce chapitre, lorsque on veut parler de stationnarité au second ordre, on utilise seulement « stationnarité » au lieu de « stationnarité au second ordre ».

4.2.1 Représentation de DE HAAN & PICKANDS

Soit ν une mesure strictement positive définie sur l'espace mesurable $[\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)]$ et soit la suite $\{(U_i, V_i), i \in \mathbb{N}^*\}$ des réalisations d'un processus ponctuel de POISSON N défini sur l'espace d'état $\mathbb{R}_+^* \times \mathbb{R}^d$ ($d \geq 1$) de mesure d'intensité μ définie par :

$$\forall (u, v) \in \mathbb{R}_+^* \times \mathbb{R}^d, \quad d\mu(u, v) = u^{-2} du \times \nu(dv)$$

Théorème 4.2.1 (Représentation de DE HAAN & PICKANDS) [RIBATET (2009)].

Soient N le processus ponctuel de POISSON défini ci-dessus, $Y = \{Y(x), x \in \mathbb{R}^d\}$ un champ aléatoire définie sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à espace d'état \mathbb{R} et f une fonction positive définie sur $\mathbb{R}^d \times \mathbb{R}^d$ telle que :

$$\int_{\mathbb{R}^n} f(v, x) \nu(dv) = 1, \quad \forall x \in \mathbb{R}^d$$

Supposons que :

$$\forall x \in \mathbb{R}^d, \quad Y(x) = \bigvee_{i=1}^{\infty} U_i f(V_i, x) \quad (4.3)$$

alors :

$$(4.3) \iff Y \text{ est un champ aléatoire } \mathbf{max-stable simple stationnaire}$$

Démonstration.

Pour une démonstration du **Théorème 4.2.1**, le lecteur pourra consulter l'article [SMITH (1990)] ou [RIBATET (2009)]. ■

Interprétation : Dans l'article [SMITH (1990)], Richard L. SMITH a interprété le champ aléatoire $Y = \{Y(x), x \in \mathbb{R}^d\}$ ayant la représentation (4.3) en terme d'épisodes environnementaux tels les tempêtes. Et depuis, on rencontre souvent le champ aléatoire ayant la représentation (4.3) sous le nom de « **Rainfall Storm Process** ». SMITH a considéré les V_i comme étant les centres de tempêtes et ν la distribution de ces centres de tempêtes dans l'espace \mathbb{R}^d . Chaque U_i représente la magnitude de la $i^{\text{ème}}$ tempête et $U_i f(V_i, x)$ la quantité de pluie produite par cette tempête au site x . À savoir que $f(V_i, \cdot)$ décrit la manière de déplacement de la $i^{\text{ème}}$ tempête dans l'espace \mathbb{R}^d , c'est-à-dire f représente la forme des tempêtes.

4.2.2 Représentation de SCHLATHER

Soit $\{U_i, i \in \mathbb{N}^*\}$ la suite des réalisations d'un processus ponctuel de POISSON N défini sur l'espace d'état \mathbb{R}_+^* de mesure d'intensité μ définie par :

$$\forall u \in \mathbb{R}_+^*, \quad d\mu(u) = u^{-2} du$$

Théorème 4.2.2 (Représentation de SCHLATHER) [RIBATET (2009)].

Soient N le processus ponctuel de POISSON défini ci-dessus et $\{Y_i, i \in \mathbb{N}^*\}$ une suite de **copies indépendantes** du champ aléatoire stationnaire $Y = \{Y(x), x \in \mathbb{R}^d\}$ défini sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à espace d'état \mathbb{R} tels que les Y_i ($i \in \mathbb{N}^*$) et N sont indépendants et :

$$\forall x \in \mathbb{R}^d, \quad \mathbb{E} \left(\max [0, Y(x)] \right) = 1$$

Soit $Z = \{Z(x), x \in \mathbb{R}^d\}$ un champ aléatoire défini sur l'espace probabilisé $(\Omega', \mathcal{A}', \mathbb{P})$ à espace d'état \mathbb{R} tel que :

$$\forall x \in \mathbb{R}^d, \quad Z(x) = \bigvee_{i=1}^{\infty} U_i \max [0, Y_i(x)] \quad (4.4)$$

alors :

$$(4.4) \iff Z \text{ est un champ aléatoire } \mathbf{max-stable simple stationnaire}$$

Démonstration.

Pour une démonstration du **Théorème 4.2.2**, le lecteur pourra se référer à [RIBATET (2009)] ou à l'ouvrage [DE HAAN & FERREIRA (2006)]. ■

Remarque : La représentation spectrale de SCHLATHER n'a pas d'interprétation physique nette comme celle de la représentation spectrale de DE HAAN & PICKANDS. Cependant, il y a une relation forte entre les deux représentations. Si on considère le cas particulier où les champs aléatoires Y_i ($i \in \mathbb{N}^*$) du **Théorème 4.2.2** sont donnés par :

$$\{Y_i(x), \quad x \in \mathbb{R}^d\} \stackrel{d}{=} \{g(x - X_i), \quad x \in \mathbb{R}^d\}$$

Où $\{X_i, i \in \mathbb{N}^*\}$ est la suite de réalisations d'un processus de POISSON homogène à espace d'état \mathbb{R}^d et g une fonction de densité d'une loi de probabilité multidimensionnelle de dimension d . Dans ce cas, la représentation spectrale de DE HAAN & PICKANDS donnée dans (4.3) et la représentation spectrale de SCHLATHER donnée dans (4.4) sont identiques.

4.3 Modèles des champs aléatoires max-stables

Les représentations spectrales des champs aléatoires max-stables simples stationnaires sont à la base de construction de beaucoup de modèles des champs aléatoires max-stables stationnaires. Dans cette section, on présentera les modèles de ces champs aléatoires les plus utilisés.

4.3.1 Modèle de SMITH

A. Premier modèle de SMITH

En utilisant le résultat de **Section 4.2.1** de ce chapitre ainsi que ses notations, on peut dire que la représentation spectrale de DE HAAN & PICKANDS donnée par (4.3) est une définition plutôt générale d'un champ aléatoire max-stable simple stationnaire. En 1990, en utilisant la représentation spectrale de DE HAAN & PICKANDS, Richard L. SMITH a construit un modèle des champs aléatoires max-stables simples stationnaires, qui porte son nom, en considérant le cadre particulier où :

1. La mesure ν est la mesure de LEBESGUE sur \mathbb{R}^d :

$$\forall v \in \mathbb{R}^d, \quad \nu(dv) = dv$$

2. La fonction f est donnée par :

$$\forall (v, x) \in \mathbb{R}^d \times \mathbb{R}^d, \quad f(v, x) = f_0(v - x)$$

Où f_0 est la fonction de densité de la loi de probabilité normale multidimensionnelle de dimension d d'espérance $0_{\mathbb{R}^d}$ et de matrice variance-covariance Σ :

$$\forall (v, x) \in \mathbb{R}^d \times \mathbb{R}^d, \quad f(v, x) = \frac{1}{\sqrt{(2\pi)^d |\Sigma|}} \exp \left\{ -\frac{1}{2} {}^T(v - x) \Sigma^{-1} (v - x) \right\}, \quad |\Sigma| = \det(\Sigma)$$

Avec ces conditions additionnelles, Richard L. SMITH a montré que la loi de probabilité marginale

bidimensionnelle du champ aléatoire max-stable simple stationnaire $Y = \{Y(x), x \in \mathbb{R}^d\}$ en deux sites $x_1 \in \mathbb{R}^d$ et $x_2 \in \mathbb{R}^d$ est donnée, pour tout $y_1 \in \mathbb{R}_+^*$ et $y_2 \in \mathbb{R}_+^*$, par :

$$P[Y(x_1) \leq y_1, Y(x_2) \leq y_2] = \exp \left\{ -\frac{1}{y_1} \Phi \left(\frac{a}{2} + \frac{1}{a} \ln \frac{y_2}{y_1} \right) - \frac{1}{y_2} \Phi \left(\frac{a}{2} + \frac{1}{a} \ln \frac{y_1}{y_2} \right) \right\} \quad (4.5)$$

Avec :

$$a^2 = {}^T(x_2 - x_1)\Sigma^{-1}(x_2 - x_1)$$

et Φ la fonction de répartition de la loi de probabilité normale standard $\mathcal{N}(0, 1)$:

$$\forall s \in \mathbb{R}, \quad \Phi(s) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^s e^{-t^2/2} dt$$

Remarque : Le modèle de SMITH présenté ci-dessus est appelé dans l'article [SMITH (1990)] par « Processus des valeurs extrême gaussien ». Pour voir la construction bien détaillée de ce modèle ainsi que la démonstration de résultat (4.5), le lecteur peut consulter l'article [SMITH (1990)] ou [RIBATET (2009)].

B. Deuxième modèle de SMITH

En se basant, encore, sur la représentation spectrale de DE HAAN & PICKANDS et en utilisant les notations de **Section 4.2.1** de ce chapitre, Richard L. SMITH a construit un autre modèle de champs aléatoire max-stables simples stationnaires où il a considéré le cas particulier où :

1. La mesure ν est la mesure de LEBESGUE sur \mathbb{R}^d :

$$\forall v \in \mathbb{R}^d, \quad \nu(dv) = dv$$

2. La fonction f est donnée par :

$$\forall (v, x) \in \mathbb{R}^d \times \mathbb{R}^d, \quad f(v, x) = f_0(v - x)$$

Où f_0 est la fonction de densité de la loi de probabilité de STUDENT multidimensionnelle de dimension d à m degré de liberté ($m \in \mathbb{N}^*$) :

$$\forall (v, x) \in \mathbb{R}^d \times \mathbb{R}^d, \quad f(v, x) = \frac{1}{\sqrt{(\pi m)^d |\Sigma|}} \frac{\Gamma(m/2)}{\Gamma[(m-d)/2]} \left(1 + \frac{{}^T(v-x)\Sigma^{-1}(v-x)}{m} \right)^{-m/2}$$

$$|\Sigma| = \det(\Sigma) \quad \text{et} \quad m > d$$

Remarque : Le modèle de SMITH présenté ci-dessus est appelé dans l'article [SMITH (1990)] par « Processus des valeurs extrême de STUDENT ». Le champ d'applications de ce modèle est très limité, c'est pour ça qu'on le trouve pas souvent dans la littérature. Et la loi de probabilité marginale bidimensionnelle du champ aléatoire max-stable simple stationnaire ayant la représentation spectrale de DE HAAN & PICKANDS avec les conditions additionnelles ci-dessus s'avère plus difficile à calculer que celle du premier modèle.

4.3.2 Modèle de SCHLATHER

En utilisant les notations de **Section 4.2.2** de ce chapitre, on peut dire que la représentation spectrale de SCHLATHER donnée dans (4.4) est un modèle très général des champs aléatoires max-stables simples stationnaires. Pour les applications, on a besoin d'ajouter quelques hypothèses pour avoir des modèles plus adéquats. En 2002, Martin SCHLATHER a proposé de prendre le champ aléatoire $Y = \{Y(x), x \in \mathbb{R}^d\}$ du **Théorème 4.2.2** comme étant un champ aléatoire gaussien standard stationnaire de corrélogramme ϱ tel que :

$$\forall x \in \mathbb{R}^d, \quad \mathbb{E}\left(\max [0, Y(x)]\right) = 1 \quad \text{et} \quad \mathbb{V}[Y(x)] = 1$$

Avec ces hypothèses supplémentaires, Martin SCHLATHER a montré que la loi de probabilité marginale bivariée du champ aléatoire max-stable simple stationnaire $Z = \{Z(x), x \in \mathbb{R}^d\}$ en deux sites $x_1 \in \mathbb{R}^d$ et $x_2 \in \mathbb{R}^d$ est donnée, pour tout $z_1 \in \mathbb{R}_+^*$ et $z_2 \in \mathbb{R}_+^*$, par :

$$\mathbb{P}[Z(x_1) \leq z_1, Z(x_2) \leq z_2] = \exp \left\{ -\frac{1}{2} \left(\frac{1}{z_1} + \frac{1}{z_2} \right) \left(1 + \sqrt{1 - 2[\varrho(h) + 1] \frac{z_1 z_2}{(z_1 + z_2)^2}} \right) \right\} \quad (4.6)$$

$$h = \|x_2 - x_1\|_2 \in \mathbb{R}_+$$

Généralement, le corrélogramme ϱ du champ aléatoire stationnaire $Y = \{Y(x), x \in \mathbb{R}^d\}$ est choisi parmi l'une des familles paramétriques suivantes :

A. Famille de WHITTLE-MATHÉRN :

$$\forall h \in \mathbb{R}_+, \quad \varrho(h) = m_1 \cdot \frac{2^{1-\ell}}{\Gamma(\ell)} \left(\frac{h}{m_2} \right)^\ell \mathcal{K}_\ell \left(\frac{h}{m_2} \right), \quad 0 \leq m_1 \leq 1, \quad m_2 > 0 \quad \text{et} \quad \ell > 0$$

Où \mathcal{K}_ℓ est une fonction de BESSEL modifiée de deuxième espèce de paramètre ℓ et Γ est la fonction Gamma définie par :

$$\forall x > 0, \quad \Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$$

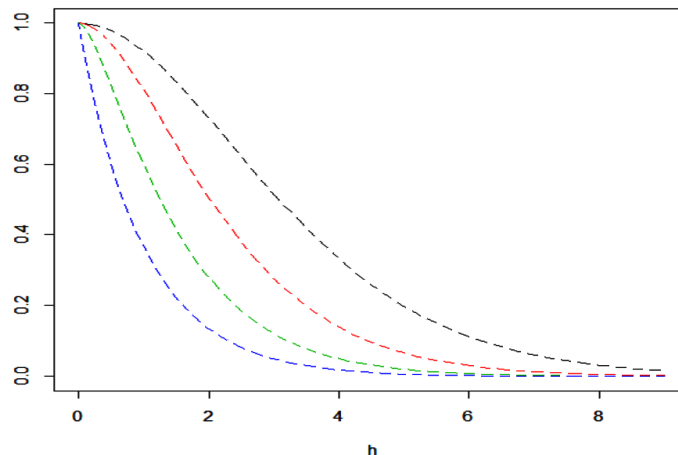


Figure 4.3.1 Graphes des corrélogrammes de la famille WHITTLE-MATHÉRN selon les valeurs de $h \in \mathbb{R}_+$ pour différentes valeurs de ℓ , avec $m_1 = m_2 = 1$: Le noir pour $\ell = 4$, le rouge pour $\ell = 2$, le vert pour $\ell = 1$ et le bleu pour $\ell = 0.5$.

B. Famille de CAUCHY :

$$\forall h \in \mathbb{R}_+, \quad \varrho(h) = m_1 \cdot \left[1 + \left(\frac{h}{m_2} \right)^2 \right]^{-\ell}, \quad 0 \leq m_1 \leq 1, \quad m_2 > 0 \quad \text{et} \quad \ell > 0$$

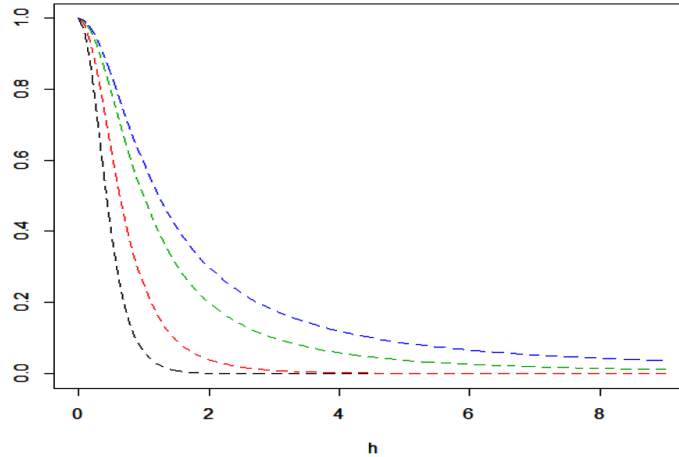


Figure 4.3.2 Graphes des corrélogrammes de la famille de CAUCHY selon les valeurs de $h \in \mathbb{R}_+$ pour différentes valeurs de ℓ , avec $m_1 = m_2 = 1$: Le noir pour $\ell = 4$, le rouge pour $\ell = 2$, le vert pour $\ell = 1$ et le bleu pour $\ell = 0.75$.

C. Famille stable :

$$\forall h \in \mathbb{R}_+, \quad \varrho(h) = m_1 \cdot \exp \left\{ - \left(\frac{h}{m_2} \right)^\ell \right\}, \quad 0 \leq m_1 \leq 1, \quad m_2 > 0 \quad \text{et} \quad 0 < \ell \leq 2$$

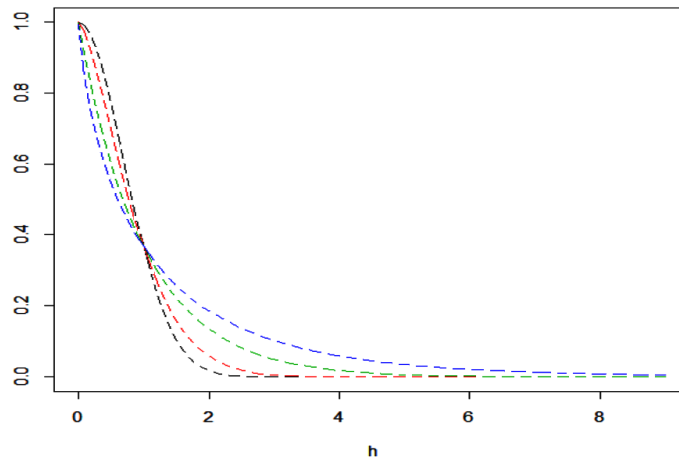


Figure 4.3.3 Graphes des corrélogrammes de la famille stable selon les valeurs de $h \in \mathbb{R}_+$ pour différentes valeurs de ℓ , avec $m_1 = m_2 = 1$: Le noir pour $\ell = 2$, le rouge pour $\ell = 1.5$, le vert pour $\ell = 1$ et le bleu pour $\ell = 0.75$.

Remarque : Pour voir la construction bien détaillée du modèle de SCHLATHER ainsi que la démonstration de résultat (4.6), le lecteur peut consulter l'article [SCHLATHER (2002)].

4.3.3 Modèle de BROWN-RESNICK

Dans cette section, on va définir un autre modèle des champs aléatoires max-stable qu'on appelle « Modèle de BROWN-RESNICK » construit dans l'article [KABLUCHKO, SCHLATHER & DE HAAN (2009)]. Donc les résultats de cette section sont entièrement tirés de cet article.

A. Construction du modèle

Soit $\{U_i, i \in \mathbb{N}^*\}$ la suite des réalisations d'un processus ponctuel de POISSON N à espace d'état \mathbb{R} de mesure d'intensité μ définie par :

$$\forall u \in \mathbb{R}, \quad \mu(du) = e^{-u} du$$

Théorème 4.3.1 [KABLUCHKO, SCHLATHER & DE HAAN (2009)].

Soient N le processus ponctuel de POISSON défini ci-dessus et $\{Y_i, i \in \mathbb{N}^*\}$ une suite de **copies indépendantes** du champ aléatoire **gaussien et intrinsèque** $Y = \{Y(x), x \in \mathbb{R}^d\}$ défini sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à espace d'état \mathbb{R} de fonction de variance v et de variogramme ϑ et tels que les Y_i ($i \in \mathbb{N}^*$) et N sont indépendants. On définit sur l'ensemble spatial des sites \mathbb{R}^d le champ aléatoire $Z = \{Z(x), x \in \mathbb{R}^d\}$ à espace d'état \mathbb{R} tel que :

$$\forall x \in \mathbb{R}^d, \quad Z(x) = \bigvee_{i=1}^{\infty} \left[U_i + Y_i(x) - \frac{v(x)}{2} \right] \quad (4.7)$$

alors Z est un champ aléatoire **max-stable stationnaire** de lois de probabilité marginales de GUMBEL standards, c'est-à-dire :

$$\forall z \in \mathbb{R}, \quad \mathbb{P}[Z(x) \leq z] = \exp \{ -e^{-z} \}, \quad \forall x \in \mathbb{R}^d$$

Remarque : On appelle le modèle des champs aléatoires max-stables de BROWN-RESNICK construit dans le théorème ci-dessus « **Processus de BROWN-RESNICK associé au variogramme ϑ** ». Donc, un champ aléatoire $Z = \{Z(x), x \in \mathbb{R}^d\}$ défini sur l'ensemble spatial des sites \mathbb{R}^d à espace d'état \mathbb{R} est un processus de BROWN-RESNICK si c'est un champ aléatoire max-stable stationnaire ayant la représentation (4.7) avec des lois de probabilité marginales de GUMBEL standards.

En fait, le nom « Modèle de BROWN-RESNICK » du modèle construit ci-dessus fait référence aux noms des deux statisticiens Bruce M. BROWN et Sidney I. RESNICK qui ont établi, en 1977, un résultat important dans l'article [BROWN & RESNICK (1977)]. Et **Théorème 4.3.1** n'est qu'une généralisation de ce résultat que nous présentons dans le théorème suivant :

Théorème 4.3.2 [BROWN & RESNICK (1977)].

Soient N le processus ponctuel de POISSON défini ci-dessus et $\{B_i, i \in \mathbb{N}^*\}$ une suite de **copies indépendantes** d'un mouvement brownien standard $B = \{B(t), t \in \mathbb{R}\}$ tels que les B_i ($i \in \mathbb{N}^*$) et N sont indépendants. On définit le processus stochastique $\eta = \{\eta(t), t \in \mathbb{R}\}$ tel que :

$$\forall t \in \mathbb{R}, \quad \eta(t) = \bigvee_{i=1}^{\infty} \left(U_i + B_i(t) - \frac{|t|}{2} \right)$$

alors η est un processus stochastique **max-stable** de lois de probabilité marginales de GUMBEL standards, c'est-à-dire :

$$\forall \xi \in \mathbb{R}, \quad \mathbb{P}[\eta(t) \leq \xi] = \exp \{ -e^{-\xi} \}, \quad \forall t \in \mathbb{R}$$

B. Stationnarité au sens de BROWN-RESNICK

On veut pas terminer cette section sans parler d'un autre mode de stationnarité des champs aléatoires qui est la **stationnarité au sens de BROWN-RESNICK** définie dans l'article [KABLUCHKO, SCHLATHER & DE HAAN (2009)]. Du coup, on présentera quelques résultats concernant cette stationnarité et qui sont en relation directe avec la construction du modèle de BROWN-RESNICK.

Soit $\{U_i, i \in \mathbb{N}^*\}$ la suite des réalisations d'un processus ponctuel de POISSON N à espace d'état \mathbb{R} de mesure d'intensité μ définie par :

$$\forall u \in \mathbb{R}, \quad \mu(du) = e^{-u} du$$

Définition 4.3.1 [KABLUCHKO, SCHLATHER & DE HAAN (2009)].

Soient N le processus ponctuel de POISSON défini ci-dessus et $\{Y_i, i \in \mathbb{N}^*\}$ une suite de **copies indépendantes** du champ aléatoire $Y = \{Y(x), x \in \mathbb{R}^d\}$ défini sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à espace d'état \mathbb{R} tels que les Y_i ($i \in \mathbb{N}^*$) et N sont indépendants et :

$$\forall x \in \mathbb{R}^d, \quad \mathbb{E}[e^{Y(x)}] < \infty \tag{4.8}$$

Soit $Z = \{Z(x), x \in \mathbb{R}^d\}$ un champ aléatoire défini sur l'ensemble spatial des sites \mathbb{R}^d à espace d'état \mathbb{R} tel que :

$$\forall x \in \mathbb{R}^d, \quad Z(x) = \bigvee_{i=1}^{\infty} [U_i + Y_i(x)] \tag{4.9}$$

On dit que le champ aléatoire $Y = \{Y(x), x \in \mathbb{R}^d\}$ qui satisfait (4.8) est **stationnaire au sens de BROWN-RESNICK** si le champ aléatoire $Z = \{Z(x), x \in \mathbb{R}^d\}$ donné par (4.9) est stationnaire.

Remarque : Dans le cadre de la définition ci-dessus, on a un résultat très important concernant la max-stabilité et qui est le suivant :

$$(4.9) \iff Z \text{ est un champ aléatoire } \mathbf{max-stable} \tag{4.10}$$

Théorème 4.3.3 [KABLUCHKO, SCHLATHER & DE HAAN (2009)].

Soient $Y = \{Y(x), x \in \mathbb{R}^d\}$ un champ aléatoire **gaussien intrinsèque** à espace d'état \mathbb{R} de fonction de variance v et $Z = \{Z(x), x \in \mathbb{R}^d\}$ un champ aléatoire défini sur l'ensemble spatial des sites \mathbb{R}^d à espace d'état \mathbb{R} tel que :

$$\forall x \in \mathbb{R}^d, \quad Z(x) = Y(x) - \frac{v(x)}{2}$$

alors le champ aléatoire $Z = \{Z(x), x \in \mathbb{R}^d\}$ est un champ aléatoire **stationnaire au sens de BROWN-RESNICK**.

Remarque : On peut constater que le résultat du théorème ci-dessus et le résultat donné dans (4.10) nous assurent la stationnarité et la max-stabilité du modèle de BROWN-RESNICK construit dans **Théorème 4.3.1**.

C. Loi de probabilité marginale bivariée

Soit $Z = \{Z(x), x \in \mathbb{R}^d\}$ le champ aléatoire max-stable stationnaire donné dans **Théorème 4.3.1**, c'est-à-dire que Z est le modèle de BROWN-RESNICK construit dans **Théorème 4.3.1**. Avec les hypothèses de ce théorème, Zakhar KABLUCHKO, Martin SCHLATHER et Laurens DE HAAN ont montré que la loi de probabilité marginale bivariée du champ aléatoire $Z = \{Z(x), x \in \mathbb{R}^d\}$ en deux sites $x_1 \in \mathbb{R}^d$ et $x_2 \in \mathbb{R}^d$ est donnée, pour tout $z_1 \in \mathbb{R}$ et $z_2 \in \mathbb{R}$, par :

$$\begin{aligned} & \mathbb{P}[Z(x_1) \leq z_1, Z(x_2) \leq z_2] \\ & \quad \parallel \\ & \exp \left\{ -e^{-z_1} \Phi \left(\frac{\sqrt{\vartheta(x_1 - x_2)}}{2} + \frac{z_2 - z_1}{\sqrt{\vartheta(x_1 - x_2)}} \right) - e^{-z_2} \Phi \left(\frac{\sqrt{\vartheta(x_1 - x_2)}}{2} + \frac{z_1 - z_2}{\sqrt{\vartheta(x_1 - x_2)}} \right) \right\} \end{aligned}$$

Avec Φ la fonction de répartition de la loi de probabilité normale standard $\mathcal{N}(0, 1)$:

$$\forall s \in \mathbb{R}, \quad \Phi(s) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^s e^{-t^2/2} dt$$

Remarque : Pour voir la construction bien détaillée du modèle de BROWN-RESNICK ainsi que les démonstrations des résultat vus dans cette section, le lecteur peut se référer à l'article [KABLUCHKO, SCHLATHER & DE HAAN (2009)].

4.4 Dépendance spatiale

Lorsque l'on s'intéresse au comportement des extrêmes dans un cadre multivarié tel les champs aléatoires, la caractérisation de la dépendance des réalisations extrêmes est une notion fondamentale pour toute inférence statistique sur la distribution jointe des extrêmes. Quand les données ne sont pas extrêmes, les outils permettant de mesurer cette dépendance sont la fonction de covariance, la fonction de corrélation et le variogramme. Mais dès que notre intérêt se porte sur les valeurs

extrêmes, ces outils ne seront plus d'une grande utilité et ils peuvent, même, ne pas exister. Ainsi, des différentes caractérisations ont été proposées dans la littérature qui ont été à la base de différents travaux pour la caractérisation de la dépendance spatiale des extrêmes.

Dans cette section, on va présenter un outil, souvent rencontré dans la littérature concernant les champs aléatoires max-stables, qui mesure la dépendance spatiale des extrêmes. Cet outil s'appelle « **coefficient extrémal** ».

4.4.1 Coefficient extrémal

Soit $Y = \{Y(x), x \in \mathbb{R}^d\}$ un champ aléatoire max-stable. D'après la caractérisation des champs aléatoires max-stables donnée dans **Théorème 4.1.1**, on peut constater que non seulement les lois de probabilité marginales univariées du champ aléatoire Y qui sont des distributions des valeurs extrêmes, mais même les lois de probabilité de sa distribution finie-dimensionnelle sont des distributions des valeurs extrêmes multivariées (voir [RESNICK (1987)]).

Sans entrer dans les détails pour ne pas sortir de l'objectif de cette section, lorsque les lois de probabilité marginales univariées du champ aléatoire max-stable $Y = \{Y(x), x \in \mathbb{R}^d\}$ sont standardisées, on peut donner une expression générale des lois de probabilité de la distributions finie-dimensionnelle de Y en n sites d'observations x_1, x_2, \dots, x_n dans \mathbb{R}^d :

1. Lorsque les lois de probabilité marginale univariées du champ aléatoire max-stable $Y = \{Y(x), x \in \mathbb{R}^d\}$ sont de FRÉCHET unitaires, comme celles des modèles de SMITH et de SCHLATHER, les lois de probabilité de sa distribution finie-dimensionnelle sont données, pour y_1, y_2, \dots, y_n dans \mathbb{R}_+^* , par :

$$P[Y(x_1) \leq y_1, Y(x_2) \leq y_2, \dots, Y(x_n) \leq y_n] = \exp \{ -\psi_1(y_1, y_2, \dots, y_n) \} \quad (4.11)$$

Où ψ_1 est une fonction **homogène** d'ordre -1 , c'est-à-dire :

$$\forall (y_1, y_2, \dots, y_n) \in (\mathbb{R}_+^*)^n, \quad \psi_1(\lambda y_1, \lambda y_2, \dots, \lambda y_n) = \frac{1}{\lambda} \psi_1(y_1, y_2, \dots, y_n), \quad \lambda > 0$$

2. Lorsque les lois de probabilité marginale univariées du champ aléatoire max-stable $Y = \{Y(x), x \in \mathbb{R}^d\}$ sont de GUMBEL, comme celles du modèle de BROWN-RESNICK, les lois de probabilité de sa distribution finie-dimensionnelle sont données, pour y_1, y_2, \dots, y_n dans \mathbb{R} , par :

$$P[Y(x_1) \leq y_1, Y(x_2) \leq y_2, \dots, Y(x_n) \leq y_n] = \exp \{ -\psi_2(y_1, y_2, \dots, y_n) \} \quad (4.12)$$

Où ψ_2 est aussi une fonction **homogène** d'ordre -1 , c'est-à-dire :

$$\forall (y_1, y_2, \dots, y_n) \in \mathbb{R}^n, \quad \psi_2(\lambda y_1, \lambda y_2, \dots, \lambda y_n) = \frac{1}{\lambda} \psi_2(y_1, y_2, \dots, y_n), \quad \lambda > 0$$

Donc les fonctions homogènes ψ_1 et ψ_2 données, respectivement, dans (4.11) et (4.12) déterminent complètement la relation et la structure de dépendance entre les différentes variables aléatoires $Y(x_1), Y(x_2), \dots, Y(x_n)$. D'autre part, pour $y_1 = y_2 = \dots = y_n = y$, les résultats donnés dans (4.11) et (4.12) seront équivalents à :

$$P[Y(x_1) \leq y, Y(x_2) \leq y, \dots, Y(x_n) \leq y] = \left(P[Y(x_1) \leq y] \right)^{\theta_n}, \quad 1 \leq \theta_n \leq n \quad (4.13)$$

Remarque : Pour plus de détails sur les résultats donnés dans (4.11), (4.12) et (4.13) qui sont des résultats importants dans la théorie des valeurs extrêmes multivariée, le lecteur peut consulter les références citées dans la note de l'introduction de ce chapitre.

Définition 4.4.1 [COOLEY, NAVEAU & PONCET (2006)].

On appelle la paramètre θ_n qui apparaisse dans (4.13), qui mesure le degré de dépendance entre les différentes réalisations du champ aléatoire max-stable standardisé $Y = \{Y(x), x \in \mathbb{R}^d\}$, « le coefficient extrémal ».

Remarque : Comme les fonctions homogènes ψ_1 et ψ_2 données dans (4.11) et (4.12), le coefficient extrémal θ_n donné dans (4.13) détermine, mais non complètement, la structure de la dépendance spatiale entre les différentes réalisations $Y(x_1), Y(x_2), \dots, Y(x_n)$ du champ aléatoire max-stable $Y = \{Y(x), x \in \mathbb{R}^d\}$. Lorsque $\theta_n = 1$, cela signifie qu'il existe une dépendance complète entre les réalisations. Et quand $\theta_n = n$, ça veut dire que les réalisations sont indépendantes. Donc, plus le coefficient extrémal θ_n se rapproche de 1, plus les variables aléatoires $Y(x_1), Y(x_2), \dots, Y(x_n)$ présentent une forte dépendance. Et plus θ_n se rapproche de n , plus les variables aléatoires $Y(x_1), Y(x_2), \dots, Y(x_n)$ sont indépendantes.

Après avoir vu d'une manière récapitulative et générale la définition du coefficient extrémal et son utilité dans l'étude de la dépendance spatiale des extrêmes, on va donner, dans ce qui suit, les différentes expressions analytiques des coefficients extrémaux bivariés qui correspondent aux différents modèles des champs aléatoires max-stables donnés dans **Section 4.3** de ce chapitre. Cela va se faire aisément puisque les expressions analytiques des lois de probabilité marginales bivariées de ces modèles sont connues explicitement.

Remarque : On entend par « coefficient extrémal bivarié » le coefficient extrémal qui caractérise la structure de dépendance entre deux réalisations d'un champ aléatoire max-stable standardisé en deux sites d'observations. Pour un champ aléatoire max-stable standardisé $Y = \{Y(x), x \in \mathbb{R}^d\}$, le coefficient extrémal bivarié θ en deux sites d'observations $x_1 \in \mathbb{R}^d$ et $x_2 \in \mathbb{R}^d$ sera donné pour $y \in \mathbb{R}$ par :

$$P[Y(x_1) \leq y, Y(x_2) \leq y] = \left(P[Y(x_1) \leq y] \right)^{\theta(x_2 - x_1)}$$

Donc, ce coefficient extrémal ne dépend que de la distance qui sépare les deux sites d'observations, et par conséquent même le degré de dépendance entre les deux réalisations en ces deux sites ne dépendra que de cette distance.

A. Coefficient extrémal du modèle de SMITH

En utilisant les notations de **Section 4.3.1**, soit $Y = \{Y(x), x \in \mathbb{R}^d\}$ le modèle de SMITH construit dans cette dernière. D'après (4.5), la loi de probabilité marginale bivariée de ce modèle en deux sites $x_1 \in \mathbb{R}^d$ et $x_2 \in \mathbb{R}^d$ est donnée, pour $y \in \mathbb{R}_+^*$, par :

$$P[Y(x_1) \leq y, Y(x_2) \leq y] = \exp \left\{ -\frac{1}{y} 2\Phi \left(\frac{h}{2} \right) \right\}, \quad h = [{}^T(x_2 - x_1)\Sigma^{-1}(x_2 - x_1)]^{1/2}$$

Par conséquent, le coefficient extrémal θ du modèle de SMITH en deux sites d'observations $x_1 \in \mathbb{R}^d$ et $x_2 \in \mathbb{R}^d$ est donné par :

$$\theta(h) = \theta(x_2 - x_1) = 2\Phi\left(\frac{h}{2}\right), \quad h = \left[{}^T(x_2 - x_1)\Sigma^{-1}(x_2 - x_1)\right]^{1/2}$$

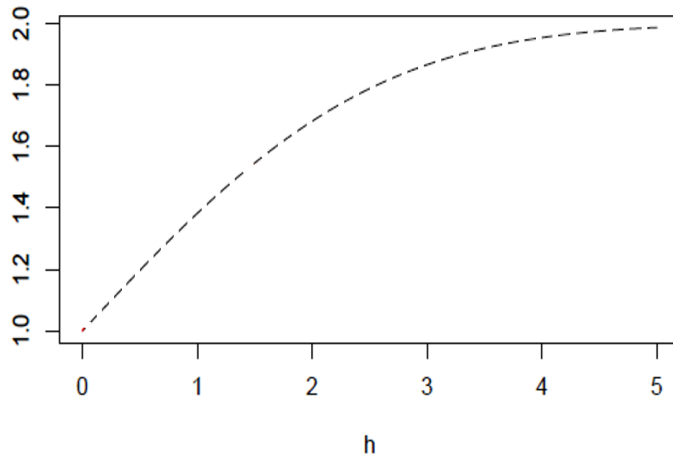


Figure 4.4.1 Graphe du coefficient extrémal du modèle de SMITH avec Σ la matrice identité d'ordre 2.

B. Coefficient extrémal du modèle de SCHLATHER

En utilisant les notations de **Section 4.3.2**, soit $Z = \{Z(x), x \in \mathbb{R}^d\}$ le modèle de SCHLATHER construit dans cette dernière. D'après (4.6), la loi de probabilité marginale bivariée de ce modèle en deux sites $x_1 \in \mathbb{R}^d$ et $x_2 \in \mathbb{R}^d$ est donnée, pour $z \in \mathbb{R}_+^*$, par :

$$P[Z(x_1) \leq z, Z(x_2) \leq z] = \exp\left\{-\frac{1}{z}\left(1 + \sqrt{\frac{1 - \varrho(h)}{2}}\right)\right\}, \quad h = \|x_2 - x_1\|_2$$

Par conséquent, le coefficient extrémal θ du modèle de SCHLATHER en deux sites d'observations $x_1 \in \mathbb{R}^d$ et $x_2 \in \mathbb{R}^d$ est donné par :

$$\theta(h) = \theta(x_2 - x_1) = 1 + \sqrt{\frac{1 - \varrho(h)}{2}}, \quad h = \|x_2 - x_1\|_2 \quad (4.14)$$

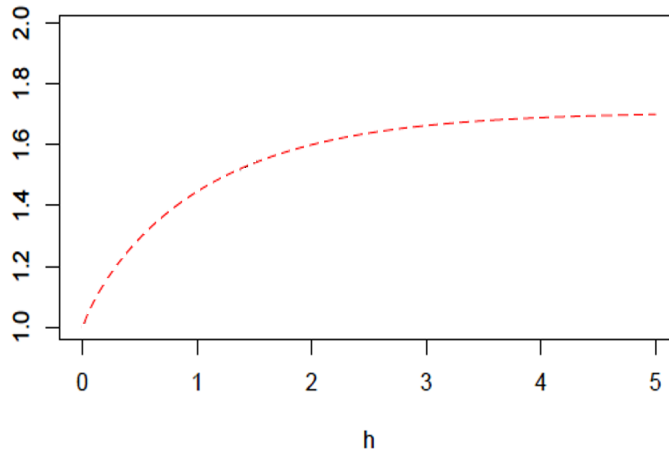


Figure 4.4.2 Graphe du coefficient extrémal du modèle de SCHLATHER où ϱ le corrélogramme de WHITTLE-MATHÉRN avec $m_1 = m_2 = \ell = 1$.

C. Coefficient extrémal du modèle de BROWN-RESNICK

En utilisant les notations de **Section 4.3.3**, soit $Z = \{Z(x), x \in \mathbb{R}^d\}$ le modèle de BROWN-RESNICK construit dans **Théorème 4.3.1**. Donc la loi de probabilité marginale bivariable de ce modèle en deux sites $x_1 \in \mathbb{R}^d$ et $x_2 \in \mathbb{R}^d$ est donnée, pour $z \in \mathbb{R}$, par :

$$P[Z(x_1) \leq z, Z(x_2) \leq z] = \exp \left\{ -2 e^{-z} \Phi \left(\frac{\sqrt{\vartheta(x_1 - x_2)}}{2} \right) \right\}$$

Par conséquent, le coefficient extrémal θ du modèle de BROWN-RESNICK est donné en deux sites d'observations $x_1 \in \mathbb{R}^d$ et $x_2 \in \mathbb{R}^d$ par :

$$\theta(h) = \theta(x_2 - x_1) = 2 \Phi \left(\frac{h}{2} \right), \quad h = \sqrt{\vartheta(x_1 - x_2)}$$

4.4.2 Madogramme

Comme on l'a déjà dit auparavant, en géostatistique le variogramme est un bon outil pour mesurer la dépendance spatiale entre deux réalisations d'un champ aléatoire. Cependant, quand il s'agit des observations extrêmes, son usage est très limité car il se peut qu'il n'existera pas pour de telles observations. Pour voir cela, considérons un champ aléatoire max-stable avec des lois de probabilité marginales univariées de FRÉCHET unitaires. Pour un tel champ aléatoire, l'espérance et la variance ne sont pas finies et par conséquent le variogramme n'existe pas théoriquement. D'où la nécessité d'introduire de nouveaux outils qui remplacent le variogramme.

En 1987, dans le cadre de la géostatistique, Georges MATHERON a introduit un nouveau outil similaire au variogramme qu'on appelle « **madogramme** », ou quelques fois rencontré sous le nom de « variogramme au premier ordre ». Dans l'article [COOLEY, NAVEAU & PONCET (2006)], Dan COOLEY, Philippe NAVEAU et Paul PONCET ont montré qu'il y a une forte relation entre le coefficient extrémal bivarié d'un champ aléatoire max-stable et son madogramme, lorsque ce dernier existe. Et

par conséquent, ils ont utilisé le madogramme comme un moyen pour estimer le coefficient extrémal bivarié d'un champ aléatoire max-stable.

Définition 4.4.2 [COOLEY, NAVEAU & PONCET (2006)].

Soit $Y = \{Y(x), x \in \mathbb{R}^d\}$ un champ aléatoire **stationnaire** défini sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à espace d'état \mathbb{R} tel que son espérance est finie. On appelle « **madogramme** » du champ aléatoire Y la fonction, qu'on note M , définie sur \mathbb{R}^d dans \mathbb{R}_+ par :

$$\forall h \in \mathbb{R}^d, \quad M(h) = \frac{\mathbb{E}(|Y(x+h) - Y(x)|)}{2}, \quad x \in \mathbb{R}^d \quad (4.15)$$

Dans la proposition suivante, on va essayer de récapituler les propriétés de base, qui sont faciles à vérifier, du madogramme d'un champ aléatoire stationnaire :

Proposition 4.4.1 [COOLEY, NAVEAU & PONCET (2006)].

Soit M le madogramme d'un champ aléatoire stationnaire $Y = \{Y(x), x \in \mathbb{R}^d\}$ défini sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à espace d'état \mathbb{R} . On a :

1.

$$M(0_{\mathbb{R}^d}) = 0$$

2. M est une fonction positive :

$$\forall h \in \mathbb{R}^d, \quad M(h) \geq M(0_{\mathbb{R}^d}) = 0$$

3. M est une fonction paire :

$$\forall h \in \mathbb{R}^d, \quad M(h) = M(-h)$$

4.

$$\forall (h, k) \in \mathbb{R}^d \times \mathbb{R}^d, \quad M(h+k) \leq M(h) + M(k)$$

A. Madogramme et coefficient extrémal

Dans cette section, on va voir la relation qui existe entre le madogramme d'un champ aléatoire max-stable stationnaire et son coefficient extrémal bivarié. En fait, il y a le même problème avec le madogramme que celui avec le variogramme lorsque les lois de probabilité marginales du champ aléatoire max-stable sont de FRÉCHET unitaires. Dans ce cas, l'espérance d'un tel champ aléatoire est infinie, et par conséquent son madogramme n'existe pas. Cela a mené, en 2006, Dan COOLEY, Philippe NAVEAU et Paul PONCET à standardiser, sans perte de généralisation et de la max-stabilité, les lois de probabilité marginales univariées d'un champ aléatoire max-stable en les transformant en distributions des valeurs extrêmes généralisées identiques avec un paramètre de forme $\xi < 1$. Dans ce cas, l'espérance d'un tel champ aléatoire max-stable standardisé est finie, ce qui entraîne l'existence de son madogramme. Bien sûr, il est possible de suivre la même stratégie pour assurer l'existence du variogramme d'un champ aléatoire max-stable, mais on va voir qu'en travaillant avec le madogramme est plus convenable car il a des liens forts avec le coefficient extrémal.

Soit $Z = \{Z(x), x \in \mathbb{R}^d\}$ un champ aléatoire max-stable stationnaire ayant des distributions des valeurs extrêmes généralisées identiques comme lois de probabilité marginales univariées de paramètre de forme $\xi \in \mathbb{R}$, de paramètre de localisation $\mu \in \mathbb{R}$ et de paramètre d'échelle $\sigma > 0$, c'est-à-dire :

$$\forall z \in \mathbb{R}, \quad \mathbb{P}[Z(x) \leq z] = \exp \left\{ - \left[1 + \xi \left(\frac{z - \mu}{\sigma} \right) \right]_+^{-1/\xi} \right\}, \quad x \in \mathbb{R}^d$$

avec :

$$\left[1 + \xi \left(\frac{z - \mu}{\sigma} \right) \right]_+ = \begin{cases} 1 + \xi \left(\frac{z - \mu}{\sigma} \right) & \text{si } 1 + \xi \left(\frac{z - \mu}{\sigma} \right) > 0 \\ 0 & \text{si } 1 + \xi \left(\frac{z - \mu}{\sigma} \right) \leq 0 \end{cases}$$

Dans la proposition suivante, on va présenter un résultat intéressant qui nous donne une méthode directe pour estimer le coefficient extrémal bivarié d'un champ aléatoire max-stable stationnaire et qu'on verra dans **Section 4.5** :

Proposition 4.4.2 [COOLEY, NAVEAU & PONCET (2006)].

Soient $Z = \{Z(x), x \in \mathbb{R}^d\}$ le champ aléatoire max-stable stationnaire défini ci-dessus et θ son coefficient extrémal bivarié. **Si** le paramètre de forme ξ est strictement inférieur à 1 ($\xi < 1$), **alors** le madogramme M du champ aléatoire Z et son coefficient extrémal θ en deux sites d'observations $x_1 \in \mathbb{R}^d$ et $x_2 \in \mathbb{R}^d$ vérifient la relation suivante :

$$\theta(x_2 - x_1) = \begin{cases} u_\beta(\cdot) \left[\mu + \frac{M(x_2 - x_1)}{\Gamma(1 - \xi)} \right] & \text{si } \xi \neq 0 \\ \exp \left\{ \frac{M(x_2 - x_1)}{\sigma} \right\} & \text{si } \xi = 0 \end{cases} \quad (4.16)$$

avec :

$$\forall z \in \mathbb{R}, \quad u_\beta(z) = \left[1 + \xi \left(\frac{z - \mu}{\sigma} \right) \right]_+^{1/\xi} \quad \text{et} \quad \forall y > 0, \quad \Gamma(y) = \int_0^{+\infty} t^{y-1} e^{-t} dt$$

Remarque : Dans le cadre de la proposition ci-dessus et afin de voir l'application directe de son résultat, on a :

1. Lorsque les lois de probabilité marginales univariées du champ aléatoire max-stable stationnaire $Z = \{Z(x), x \in \mathbb{R}^d\}$ sont de WEIBULL standards ($\xi = -1$, $\mu = 0$ et $\sigma = 1$), on a :

$$(4.16) \iff M(x_2 - x_1) = 1 - \frac{1}{\theta(x_2 - x_1)}$$

2. Lorsque les lois de probabilité marginales univariées du champ aléatoire max-stable stationnaire $Z = \{Z(x), x \in \mathbb{R}^d\}$ sont de GUMBEL standards ($\xi = 0$, $\mu = 0$ et $\sigma = 1$), comme c'est le cas du modèle de BROWN-RESNICK, on a :

$$(4.16) \iff M(x_2 - x_1) = \ln \theta(x_2 - x_1)$$

B. F-Madogramme

De ce qu'on a vu jusque-là, on constate qu'il y a un lien fort entre le coefficient extrémal bivarié d'un champ aléatoire max-stable stationnaire standardisé et son madogramme. Ce qui nous a donné une méthode directe pour estimer ce coefficient extrémal. Mais lorsque les lois de probabilité marginales univariées d'un champ aléatoire max-stable stationnaire sont de FRÉCHET unitaires, le madogramme tel qu'il est défini ci-dessus ne sera pas d'une grande utilité, car ce madogramme n'existe pas ($\xi = 1$). Dans le même article [COOLEY, NAVEAU & PONCET (2006)], Dan COOLEY, Philippe NAVEAU et Paul PONCET ont proposé, au lieu du madogramme, le madogramme modifié qu'ils ont appelé « le *F*-madogramme ».

Soit $Z = \{Z(x), x \in \mathbb{R}^d\}$ un champ aléatoire max-stable stationnaire ayant des distributions des valeurs extrêmes généralisées identiques comme lois de probabilité marginales univariées de paramètre de forme $\xi \in \mathbb{R}$, de paramètre de localisation $\mu \in \mathbb{R}$ et de paramètre d'échelle $\sigma > 0$.

Définition 4.4.3 [COOLEY, NAVEAU & PONCET (2006)].

Soit $Z = \{Z(x), x \in \mathbb{R}^d\}$ le champ aléatoire max-stable stationnaire défini ci-dessus. On appelle *F*-madogramme du champ aléatoire max-stable Z la fonction, qu'on note M_F , définie sur \mathbb{R}^d dans \mathbb{R}_+ par :

$$\forall h \in \mathbb{R}^d, \quad M_F(h) = \frac{\mathbb{E}\left(\left|F[Z(x+h)] - F[Z(x)]\right|\right)}{2}, \quad x \in \mathbb{R}^d \quad (4.17)$$

avec :

$$\forall z \in \mathbb{R}, \quad F(z) = \mathbb{P}[Z(x) \leq z] = \exp\left\{-\left[1 + \xi\left(\frac{z - \mu}{\sigma}\right)\right]_+^{-1/\xi}\right\}, \quad x \in \mathbb{R}^d$$

Dans la proposition suivante, on va présenter un résultat qui nous montre le lien entre le coefficient extrémal bivarié d'un champ aléatoire max-stable stationnaire et son *F*-madogramme :

Proposition 4.4.3 [COOLEY, NAVEAU & PONCET (2006) et RIBATET (2009)].

Soient $Z = \{Z(x), x \in \mathbb{R}^d\}$ le champ aléatoire max-stable stationnaire défini ci-dessus, θ son coefficient extrémal bivarié et M_F son *F*-madogramme. On a :

$$2M_F(x_2 - x_1) = \frac{\theta(x_2 - x_1) - 1}{\theta(x_2 - x_1) + 1}, \quad x_1 \in \mathbb{R}^d \quad \text{et} \quad x_2 \in \mathbb{R}^d$$

Où, ce qui est équivalent :

$$\theta(x_2 - x_1) = \frac{1 + 2M_F(x_2 - x_1)}{1 - 2M_F(x_2 - x_1)}, \quad x_1 \in \mathbb{R}^d \quad \text{et} \quad x_2 \in \mathbb{R}^d \quad (4.18)$$

Remarque : De la proposition ci-dessus, on constate que pour les champs aléatoires max-stables stationnaires ayant des lois de probabilité marginales univariées de FRÉCHET unitaires, comme celles des modèles de SMITH et de SCHLATHER, il serait plus convenable d'utiliser le *F*-madogramme pour l'estimation du coefficient extrémal bivarié au lieu du madogramme.

4.5 Méthodes d'inférence

Dans cette section, on va essayer de donner quelques méthodes d'estimation du coefficient extrémal bivarié d'un champ aléatoire max-stable ainsi que les paramètres de sa loi de probabilité marginale bivariée :

4.5.1 Estimation du coefficient extrémal

A. Approche SMITH

L'estimation du coefficient extrémal par l'approche SMITH est une méthode proposée par Richard L. SMITH dans son article [SMITH (1990)]. Elle sert à estimer le coefficient extrémal bivarié d'un champ aléatoire max-stable simple. Cette méthode peut être résumée comme suit :

Soit $Y = \{Y(x), x \in \mathbb{R}^d\}$ un champ aléatoire max-stable simple, c'est-à-dire que ses lois de probabilité marginales univariées sont des lois de probabilité de FRÉCHET standards. Et soient Y_1, Y_2, \dots, Y_n n copies indépendantes du champ aléatoire max-stable simple Y .

Alors, pour tout $x \in \mathbb{R}^d$, la variable aléatoire $1/Y(x)$ suit une loi de probabilité exponentielle standard $\mathcal{E}(1)$. Et, pour tout $x_1 \in \mathbb{R}^d$ et $x_2 \in \mathbb{R}^d$, la variable aléatoire $\min[1/Y(x_1), 1/Y(x_2)]$ suit une loi de probabilité exponentielle de paramètre $\theta(x_2 - x_1)$, où $\theta(x_2 - x_1)$ est le coefficient extrémal bivarié du champ aléatoire max-stable $Y = \{Y(x), x \in \mathbb{R}^d\}$ entre les deux sites d'observations x_1 et x_2 .

Richard L. SMITH a proposé un estimateur $\hat{\theta}(x_2 - x_1)$ du coefficient extrémal bivarié $\theta(x_2 - x_1)$ donné par :

$$\hat{\theta}(x_2 - x_1) = \frac{n}{\sum_{i=1}^n \min[1/Y_i(x_1), 1/Y_i(x_2)]}$$

B. Approche F-madogramme

L'estimation du coefficient extrémal par l'approche F -madogramme est une méthode proposée par Dan COOLEY, Philippe NAVEAU et Paul PONCET dans leur article [COOLEY, NAVEAU & PONCET (2006)]. Cette méthode est basée sur les deux résultats (4.17) et (4.18) et elle sert à estimer le coefficient extrémal bivarié d'un champ aléatoire max-stable ayant des lois de probabilité marginales univariées identiques. Donc elle peut être utilisée pour l'estimation du coefficient extrémal bivarié du modèle de SMITH, du modèle de SCHLATHER, ainsi que celui du modèle de BROWN-RESNICK.

Pour résumer, soit $Y = \{Y(x), x \in \mathbb{R}^d\}$ un champ aléatoire max-stable ayant des lois de probabilité marginales univariées identiques. Et soient Y_1, Y_2, \dots, Y_n n copies indépendantes du champ aléatoire max-stable Y . Notons, pour x_1 et x_2 dans \mathbb{R}^d , $\theta(x_2 - x_1)$ le coefficient extrémal bivarié du champ aléatoire max-stable Y entre les deux sites d'observations x_1 et x_2 .

Alors, d'après (4.18) on a :

$$\hat{\theta}(x_2 - x_1) = \frac{1 + 2\hat{M}_F(x_2 - x_1)}{1 - 2\hat{M}_F(x_2 - x_1)} \quad (4.19)$$

Avec, d'après (4.17) :

$$\hat{M}_F(x_2 - x_1) = \frac{1}{2n} \sum_{i=1}^n \left| \hat{F}[Y_i(x_2)] - \hat{F}[Y_i(x_1)] \right|$$

Où $\hat{F}[Y(x_j)]$ est la fonction de répartition empirique associée à la variable aléatoire $Y(x_j)$ basée sur l'échantillon $\{Y_1(x_j), Y_2(x_j), \dots, Y_n(x_j)\}$.

C. Approche madogramme

L'estimation du coefficient extrémal par l'approche madogramme est une méthode proposée, elle aussi, par Dan COOLEY, Philippe NAVEAU et Paul PONCET dans leur article [COOLEY, NAVEAU & PONCET (2006)]. Cette méthode est basée sur les deux résultats (4.15) et (4.16) et elle sert à estimer le coefficient extrémal bivarié d'un champ aléatoire max-stable ayant des GEVD identiques comme lois de probabilité marginales univariées dont le paramètre de forme $\xi < 1$. Donc elle peut être utilisée pour l'estimation du coefficient extrémal bivarié du modèle de BROWN-RESNICK.

Pour cela, soit $Y = \{Y(x), x \in \mathbb{R}^d\}$ un champ aléatoire max-stable ayant des GEVD identiques comme lois de probabilité marginales univariées dont le paramètre de forme $\xi < 1$. Et soient Y_1, Y_2, \dots, Y_n n copies indépendantes du champ aléatoire max-stable Y . Notons, pour x_1 et x_2 dans \mathbb{R}^d , $\theta(x_2 - x_1)$ le coefficient extrémal bivarié du champ aléatoire max-stable Y entre les deux sites d'observations x_1 et x_2 .

Alors, d'après (4.16) on a :

$$\hat{\theta}(x_2 - x_1) = \begin{cases} u_\beta(\cdot) \left[\mu + \frac{\hat{M}(x_2 - x_1)}{\Gamma(1 - \xi)} \right] & \text{si } \xi \neq 0 \\ \exp \left\{ \frac{\hat{M}(x_2 - x_1)}{\sigma} \right\} & \text{si } \xi = 0 \end{cases}$$

Et lorsque le champ aléatoire max-stable $Y = \{Y(x), x \in \mathbb{R}^d\}$ est un modèle de BROWN-RESNICK, on aura :

$$\hat{\theta}(x_2 - x_1) = \exp [\hat{M}(x_2 - x_1)]$$

Avec, d'après (4.15) :

$$\hat{M}(x_2 - x_1) = \frac{1}{2n} \sum_{i=1}^n |Y_i(x_2) - Y_i(x_1)|$$

Remarque : Pour mettre en application les résultats précédents, il est souvent préférable de suivre la procédure suivante : Supposons qu'on a m sites d'observation où, sur une période de temps donnée, on a récolté sur chaque site un ensemble d'observations. Ensuite, on répartit l'ensemble de données de chaque site en n sous ensemble, où n est souvent pris comme étant le nombre d'années de la période de temps sur laquelle on a relevé les observations. Après, sur chaque site d'observation, on construit un autre ensemble de données composé de maximum de données relevées durant chacune des n années d'observation. Ainsi, on pourra dire qu'on a construit, approximativement, n champs aléatoires max-stables Y_1, Y_2, \dots, Y_n indépendants et identiquement distribués, avec :

$$\forall i \in \{1, 2, \dots, n\}, \quad Y_i = \{Y_i(x_1), Y_i(x_2), \dots, Y_i(x_m)\}$$

Où, pour $j \in \{1, 2, \dots, m\}$, $Y_i(x_j)$ représente le maximum des données récoltées sur le $j^{\text{ème}}$ site durant la $i^{\text{ème}}$ année.

4.5.2 Vraisemblance composite

Dans cette section, on va donner une méthode d'estimation des paramètres d'un champ aléatoire max-stable. En particulier les paramètres des modèles de SMITH, de SCHLATHER et de BROWN-RESNICK. Cette méthode qu'on appelle « **méthode de vraisemblance composite** » est proche de « **la méthode du maximum de vraisemblance** » introduite, en 1922, par Ronald A. FISHER dans son article « On the Mathematical Foundations of Theoretical Statistics » qui est très connue et beaucoup utilisée pour l'estimation des paramètres d'un modèle statistique dont l'expression de la fonction de vraisemblance totale est facile à calculer. Cependant, il y a des modèles statistiques où l'expression de la fonction de vraisemblance totale n'est pas connue ou difficile à calculer, et dans ce cas la méthode du maximum de vraisemblance n'a plus d'intérêt. En 1975, J. E. BESAG a introduit dans son article « Statistical Analysis of Non Lattice Data » le concept de « pseudo-vraisemblance » ou « vraisemblance composite ». En 1988, Bruce G. LINDSAY a utilisé ce concept dans son article « Composite Likelihood Methods » afin de donner une méthode pour l'estimation des paramètres d'un modèle statistique dont l'expression de la fonction de vraisemblance totale n'est pas connue. Et comme on a vu, les expressions des fonctions de vraisemblance des modèles des champs aléatoires max-stables présentés dans ce chapitre ne sont connues que dans le cas bivarié, c'est-à-dire que les expressions de leurs fonctions de vraisemblance totales sont inconnues ou presque impossible de les calculer lorsque le nombre des sites d'observations est très grand (supérieur ou égal à 3). Cela a mené les statisticiens à employer la méthode de vraisemblance composite, au lieu de celle du maximum de vraisemblance, pour estimer les paramètres d'un champ aléatoire max-stable.

Dans ce qui suit, sauf mention du contraire, soit X un vecteur aléatoire à valeurs dans \mathbb{R}^m ($m \in \mathbb{N}^*$) de fonction de densité $f_X(\cdot, \psi)$ avec $\psi \in \Psi \subseteq \mathbb{R}^p$ ($p \in \mathbb{N}^*$) est le paramètre inconnu qu'on veut estimer, et soient X_1, X_2, \dots, X_n n copies indépendantes du vecteur aléatoire X . On note par $f_{i_1, \dots, i_r}(\cdot, \psi)$ ($r = 1, 2, \dots, m$) la fonction de densité marginale r -dimensionnelle du vecteur aléatoire X .

Définition 4.5.1 [GHOLAM REZAEI (2010)].

On appelle **vraisemblance composite d'ordre** $r \in \{1, 2, \dots, m\}$ de notre modèle statistique $\{f_X(x, \Psi), x \in \mathbb{R}^m \text{ et } \psi \in \Psi \subseteq \mathbb{R}^p\}$ la fonction VC_r définie sur \mathbb{R}^p par :

$$\forall \psi \in \Psi, \quad VC_r(\psi) = \prod_{k=1}^n \prod_{i_1 < \dots < i_r} f_{i_1, \dots, i_r}(x_{i_1, k}, \dots, x_{i_r, k}, \psi), \quad (x_{i_1, k}, \dots, x_{i_r, k}) \in \mathbb{R}^r$$

avec :

$$i_j \in \{1, 2, \dots, m\} \quad \text{où} \quad j \in \{1, 2, \dots, r\}$$

Remarque : En restant toujours dans le cadre de la définition ci-dessus et en utilisant ses notations, on peut définir la **log-vraisemblance composite** d'ordre r , qu'on note ℓ_r , de notre modèle statistique comme suit :

$$\forall \psi \in \Psi, \quad \ell_r(\psi) = \log [VC_r(\psi)] = \sum_{k=1}^n \sum_{i_1 < \dots < i_r} \log [f_{i_1, \dots, i_r}(x_{i_1, k}, \dots, x_{i_r, k}, \psi)]$$

Remarque : D'après la définition ci-dessus et en utilisant ses notations, on peut remarquer que la vraisemblance totale du modèle statistique est un cas particulier de sa vraisemblance composite, et cela dans le cas où $r = m$.

Remarque : Toujours en restant dans le cadre de la définition précédente, on peut constater que la vraisemblance composite d'ordre r de notre modèle statistique est une combinaison de ses vraisemblances marginales r -dimensionnelles. Dans le cas des modèles des champs aléatoires max-stables vus dans ce chapitre, on a vu que seules les expressions de leurs lois de probabilité marginales bivariées sont connues analytiquement. Ce qui nous conduit à utiliser les vraisemblances composites d'ordre 2 de tels modèles pour estimer leurs paramètres. Et dans ce cas, on appelle cette vraisemblance composite « **vraisemblance par paires** ». Pour le modèle statistique donné dans la définition précédente, sa vraisemblance par paires est la fonction VP définie sur \mathbb{R}^p comme suit :

$$\forall \psi \in \Psi, \quad \text{VP}(\psi) = \text{VC}_2(\psi) = \prod_{k=1}^n \prod_{i=1}^{m-1} \prod_{j=i+1}^m f_{i,j}(x_{i,k}, x_{j,k}, \psi), \quad (x_{i,k}, x_{j,k}) \in \mathbb{R}^2$$

Et sa « **log-vraisemblance par paires** » est la fonction ℓ_2 donnée par :

$$\forall \psi \in \Psi, \quad \ell_2(\psi) = \log [\text{VP}(\psi)] = \sum_{k=1}^n \sum_{i=1}^{m-1} \sum_{j=i+1}^m \log [f_{i,j}(x_{i,k}, x_{j,k}, \psi)], \quad (x_{i,k}, x_{j,k}) \in \mathbb{R}^2$$

Définition 4.5.2 [PADOAN, RIBATET & SISSON (2010)].

On appelle **fonction de score composite d'ordre r** , qu'on note \mathbb{D}_r , du modèle statistique $\{f_X(x, \Psi), x \in \mathbb{R}^m \text{ et } \psi \in \Psi \subseteq \mathbb{R}^p\}$ le gradient de sa fonction log-vraisemblance composite d'ordre r , c'est-à-dire pour $\psi = (\psi_1, \psi_2, \dots, \psi_p) \in \Psi \subseteq \mathbb{R}^p$ on a :

$$\mathbb{D}_r(\psi) = \nabla \ell_r(\psi) = \left(\frac{\partial}{\partial \psi_1} \ell_r(\psi), \frac{\partial}{\partial \psi_2} \ell_r(\psi), \dots, \frac{\partial}{\partial \psi_p} \ell_r(\psi) \right)^T$$

avec :

$$\frac{\partial}{\partial \psi_i} \ell_r(\psi) = \sum_{k=1}^n \sum_{i_1 < \dots < i_r} \frac{\partial}{\partial \psi_i} \log [f_{i_1, \dots, i_r}(x_{i_1, k}, \dots, x_{i_r, k}, \psi)]$$

Remarque : Dans le cadre de la définition ci-dessus et si $r = 2$, on appelle \mathbb{D}_2 **la fonction de score par paires** du modèle statistique $\{f_X(x, \Psi), x \in \mathbb{R}^m \text{ et } \psi \in \Psi \subseteq \mathbb{R}^p\}$:

$$\mathbb{D}_2(\psi) = \nabla \ell_2(\psi) = \left(\frac{\partial}{\partial \psi_1} \ell_2(\psi), \frac{\partial}{\partial \psi_2} \ell_2(\psi), \dots, \frac{\partial}{\partial \psi_p} \ell_2(\psi) \right)^T$$

avec :

$$\frac{\partial}{\partial \psi_i} \ell_2(\psi) = \sum_{k=1}^n \sum_{i=1}^{m-1} \sum_{j=i+1}^m \frac{\partial}{\partial \psi_i} \log [f_{i,j}(x_{i,k}, x_{j,k}, \psi)]$$

A. Estimateur du maximum de vraisemblance composite (EMVC)

D'une manière analogue à celle de construction d'un estimateur de maximum de vraisemblance classique, on peut construire un estimateur du maximum de vraisemblance composite en utilisant la vraisemblance composite d'un modèle statistique :

Définition 4.5.3 [PADOAN, RIBATET & SISSON (2010)].

On appelle *estimateur du maximum de vraisemblance composite* de notre modèle statistique $\{f_X(x, \Psi), x \in \mathbb{R}^m \text{ et } \psi \in \Psi \subseteq \mathbb{R}^p\}$ le vecteur $\hat{\psi}_{\text{MVC}} \in \Psi \subseteq \mathbb{R}^p$ défini par :

$$\hat{\psi}_{\text{MVC}} = \arg \max_{\psi \in \Psi} \text{VC}_r(\psi) = \arg \max_{\psi \in \Psi} \ell_r(\psi)$$

Remarque : En général, lorsque l'EMVC $\hat{\psi}_{\text{MVC}}$ est unique, il est la racine de la fonction score composite de notre modèle statistique, c'est-à-dire :

$$\mathbb{D}_r(\hat{\psi}_{\text{MVC}}) = 0_{\mathbb{R}^p}$$

Définition 4.5.4 [GHOLAM REZAEI (2010)].

On appelle *estimateur du maximum de vraisemblance par paires* de notre modèle statistique $\{f_X(x, \Psi), x \in \mathbb{R}^m \text{ et } \psi \in \Psi \subseteq \mathbb{R}^p\}$ le vecteur $\hat{\psi}_{\text{MVP}} \in \Psi \subseteq \mathbb{R}^p$ défini par :

$$\hat{\psi}_{\text{MVP}} = \arg \max_{\psi \in \Psi} \text{VC}_2(\psi) = \arg \max_{\psi \in \Psi} \ell_2(\psi)$$

Remarque : Comme l'EMVP est un cas particulier de l'EMVC, alors lorsque l'EMVP $\hat{\psi}_{\text{MVP}}$ est unique, il est la racine de la fonction score par paires de notre modèle statistique, c'est-à-dire :

$$\mathbb{D}_2(\hat{\psi}_{\text{MVP}}) = 0_{\mathbb{R}^p}$$

B. Propriétés de l'EMVC

L'intérêt de la méthode de vraisemblance composite réside dans le fait qu'elle peut être employée pour estimer les paramètres marginaux d'un modèle statistique très complexe, tels les modèles des champs aléatoires max-stables, où il est difficile à calculer sa fonction de vraisemblance totale. Et même si cette fonction de vraisemblance totale est calculable, elle nous entraîne dans des calculs trop compliqués lorsque on veut estimer les paramètres par la méthode classique du maximum de vraisemblance à cause de grand nombre des paramètres à estimer. Et sous certaines conditions de régularité (voir [LINDSAY (1988)] et [COX & REID (2004)]), l'estimateur du maximum de vraisemblance composite $\hat{\psi}_{\text{MVC}}$ du modèle statistique $\{f_X(x, \Psi), x \in \mathbb{R}^m \text{ et } \psi \in \Psi \subseteq \mathbb{R}^p\}$ est sans biais, consistant et suit asymptotiquement une loi de probabilité normale p -dimensionnelle d'espérance ψ et de matrice variance-covariance $\text{H}(\psi)\text{J}(\psi)^{-1}\text{H}(\psi)$:

$$\hat{\psi}_{\text{MVC}} \sim \mathcal{N}\left(\psi, \text{H}(\psi)\text{J}(\psi)^{-1}\text{H}(\psi)\right)$$

Avec :

$$H(\psi) = E \left\{ -\nabla [\nabla \ell_r(\psi)]^T \right\} \quad \text{et} \quad J(\psi) = E \left\{ [\nabla \ell_r(\psi)] [\nabla \ell_r(\psi)]^T \right\}$$

Où :

$$\nabla [\nabla \ell_r(\psi)]^T = \left[\frac{\partial^2}{\partial \psi_i \partial \psi_j} \ell_r(\psi) \right]$$

et :

$$[\nabla \ell_r(\psi)] [\nabla \ell_r(\psi)]^T = \left[\left(\frac{\partial}{\partial \psi_i} \ell_r(\psi) \right) \left(\frac{\partial}{\partial \psi_j} \ell_r(\psi) \right) \right]$$

C'est-à-dire que $H(\psi)$ et $J(\psi)$ sont des matrices carrées d'ordre p dont les composantes respectives de leurs $i^{\text{ème}}$ lignes et $j^{\text{ème}}$ colonnes $A_{ij}(\psi)$ et $B_{ij}(\psi)$ ($i = 1, \dots, p$ et $j = 1, \dots, p$) sont données par :

$$A_{ij}(\psi) = -E \left\{ \frac{\partial^2}{\partial \psi_i \partial \psi_j} \ell_r(\psi) \right\}$$

et :

$$B_{ij}(\psi) = E \left\{ \left(\frac{\partial}{\partial \psi_i} \ell_r(\psi) \right) \left(\frac{\partial}{\partial \psi_j} \ell_r(\psi) \right) \right\}$$

En application, les matrices $H(\psi)$ et $J(\psi)$ doivent être estimées, et leurs estimateurs respectifs $\hat{H}(\hat{\psi}_{\text{MVC}})$ et $\hat{J}(\hat{\psi}_{\text{MVC}})$ sont donnés par :

$$\hat{H}(\hat{\psi}_{\text{MVC}}) = \left[\hat{A}_{ij}(\hat{\psi}_{\text{MVC}}) \right] \quad \text{et} \quad \hat{J}(\hat{\psi}_{\text{MVC}}) = \left[\hat{B}_{ij}(\hat{\psi}_{\text{MVC}}) \right]$$

Et comme X_1, X_2, \dots, X_n sont n réalisations indépendantes de notre vecteur aléatoire X , on obtient les $\hat{A}_{ij}(\hat{\psi}_{\text{MVC}})$ et les $\hat{B}_{ij}(\hat{\psi}_{\text{MVC}})$ par :

$$\hat{A}_{ij}(\hat{\psi}_{\text{MVC}}) = -\frac{1}{n} \sum_{k=1}^n \frac{\partial^2}{\partial \psi_i \partial \psi_j} \ell_r(\hat{\psi}_{\text{MVC}})$$

et :

$$\hat{B}_{ij}(\hat{\psi}_{\text{MVC}}) = \frac{1}{n} \sum_{k=1}^n \left(\frac{\partial}{\partial \psi_i} \ell_r(\hat{\psi}_{\text{MVC}}) \right) \left(\frac{\partial}{\partial \psi_j} \ell_r(\hat{\psi}_{\text{MVC}}) \right)$$

Avec :

$$\ell_r(\hat{\psi}_{\text{MVC}}) = \sum_{k=1}^n \sum_{i_1 < \dots < i_r} \log [f_{i_1, \dots, i_r}(X_{i_1, k}, \dots, X_{i_r, k}, \hat{\psi}_{\text{MVC}})]$$

Remarque : Pour les modèles des champs aléatoires max-stables vus dans ce chapitre, on sait que seules les expressions de leurs lois de probabilité marginales bivariées sont connues. Donc la méthode qui convient plus pour estimer les paramètres de tels modèles est la méthode de vraisemblance par paires qui est un cas particulier de la méthode de vraisemblance par composite. Pour cela, il suffit de remplacer dans les expressions précédentes la log-vraisemblance composite ℓ_r par la log-vraisemblance par paires ℓ_2 et l'estimateur de maximum de vraisemblance composite $\hat{\psi}_{\text{MVC}}$ par l'estimateur de maximum de vraisemblance par paires $\hat{\psi}_{\text{MVP}}$.

C. Application de la méthode de VC

Comme on l'a déjà dit, la méthode qui convient plus pour estimer les paramètres des modèles des champs aléatoires max-stables vus dans ce chapitre est la méthode de vraisemblance par paires, car seules les expressions de leurs lois de probabilité marginales bivariées sont connues. Mais avant de mettre en application cette méthode, on doit connaître avec exactitude les expressions des fonctions de densité marginales bivariées de ces modèles.

Pour voir et connaître les expressions des fonctions de densité marginales bivariées des modèles des champs aléatoires max-stables, on renvoie le lecteur à des références où il peut trouver ces expressions. Concernant le modèle de SMITH, le lecteur peut voir l'article [PADOAN, RIBATET & SISSON (2010)]. Et pour le modèle de SCHLATHER, le lecteur est renvoyé à la thèse de Doctorat [GHOLAM REZAEI (2010)] présentée par M.M. GHOLAM REZAEI le 17 décembre 2010 sous la direction de A.C. DAVISON, J. TAWN, D. WALSHAW et S. MORGENTHALER.

4.6 Exemple d'application

Les événements naturels tels les canicules, les précipitations, les marées et les tempêtes constituent des vrais dangers lorsqu'ils sont extrêmes. Bien que ce soit difficile d'attribuer un événement particulier, tel l'Ouragan Katrina ou l'inondation de 2010 au Pakistan, aux effets du changement de climat, les données d'observation et les modèles climatiques existants suggèrent que les dimensions de telles catastrophes augmenteront dans le futur. Les conséquences potentielles incluent des augmentations des tempêtes sévères, des inondations, des feux dans la nature, des échecs de la récolte, des déplacements de la population et augmentation de la mortalité. À part leurs impacts directs, de tels événements auront aussi des effets indirects tels que l'augmentation des coûts pour fortifier les infrastructures et l'augmentation des primes d'assurance. Il y a donc un besoin pressant pour une meilleure compréhension des extrêmes spatiaux et une estimation plus détaillée de leurs conséquences. Et sur les dernières années, le sujet est devenu une interface active entre les scientifiques du climat, du social et des statistiques, en interaction avec les actionnaires tels les compagnies d'assurances et les fonctionnaires de la santé.

La théorie des valeurs extrêmes spatiale est un outil pertinent pour décrire de tels événements et prévenir leurs conséquences. Comme exemple d'application, on va utiliser les données « `rainfall` » incluses dans le package « `SpatialExtremes` » du logiciel **R**.

4.6.1 Description des données

Les données qui seront traitées dans cette section sont des données qui correspondent aux quantités maximales des précipitations journalières (en millimètres) récoltées sur 79 sites en Suisse durant les étés des années 1962 à 2008. En fait, les données `rainfall` du package `SpatialExtremes` comportent deux tableaux de données appelés « `coord` » et « `rain` ». Le tableau `coord` correspond aux positions géographiques des 79 sites sur lesquels on a relevé les données et il comporte 79 lignes et 3 colonnes, où la première colonne correspond aux longitudes, la deuxième aux latitudes et la troisième aux altitudes (en mètres) des 79 sites et chaque ligne correspond à un site. Et le tableau `rain` correspond aux données qu'on va traiter dans cette section et il comporte 47 lignes et 79 colonnes, où chaque colonne correspond à un site et chaque ligne à une année. **Figure 4.6.1** montre les positions de ces sites sur la carte géographique.

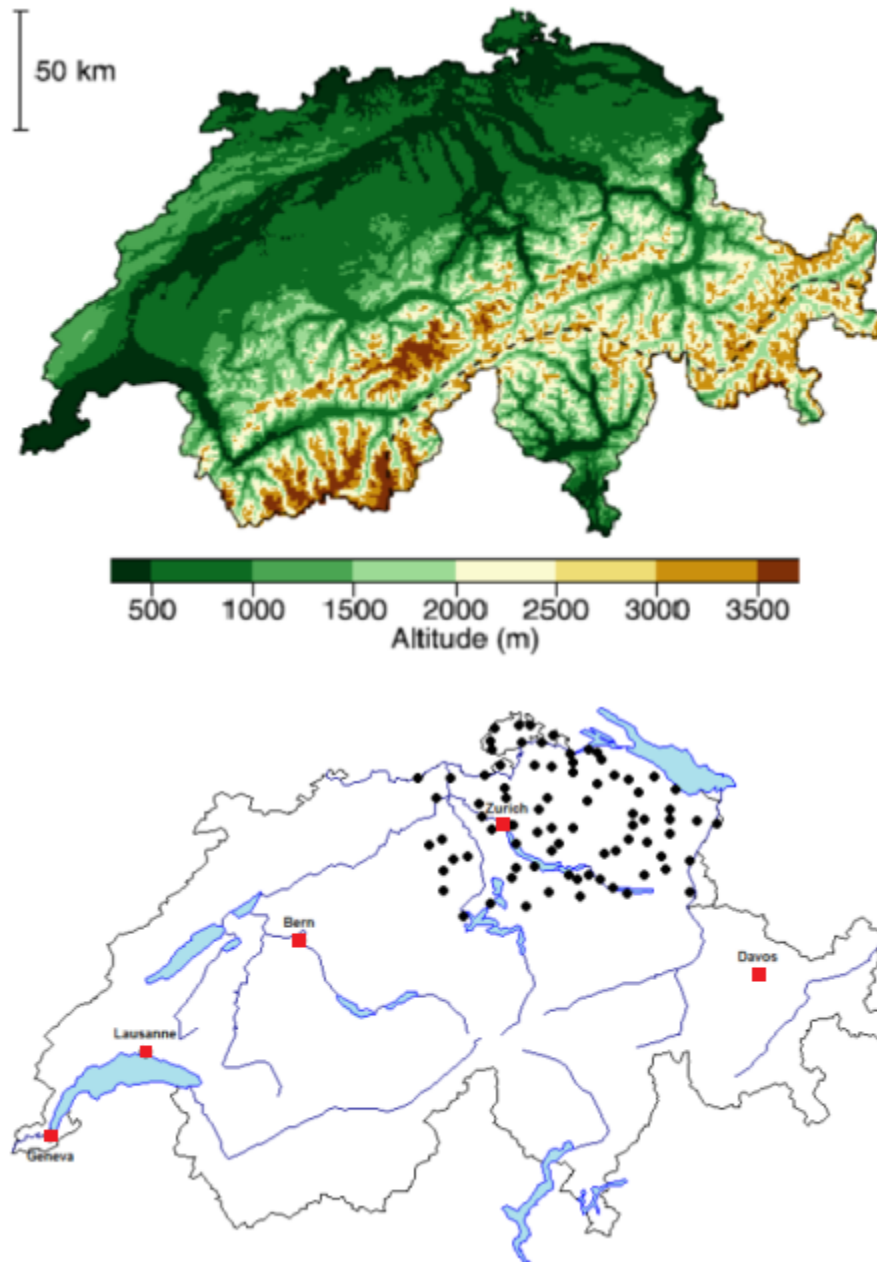


Figure 4.6.1 Carte géographique de la Suisse où on désigne par des points noirs les 79 sites sur lesquels sont relevées les données `rainfall` du package `SpatialExtremes`.

Code R

Pour charger les données `rainfall` du package `SpatialExtremes` dans l'environnement de travail **R**, on utilise le code suivant :

```
> data(rainfall)
> Y <- rain
```

```
> Y
> S1 <- coord
> S1
```

4.6.2 Transformation des données

Notre but dans cette analyse est de poser un modèle des champs aléatoires max-stables approprié aux données qu'on a. Afin d'atteindre cet objectif, on suppose que les quantités maximales des précipitations journalières de chaque site durant les 47 années sont des réalisations d'une certaine variable aléatoire suivant une GEVD. Après avoir posé la GEVD appropriée à chaque site, et cela en estimant ses paramètres, on doit transformer, sans perte de généralisation, les données de chaque site en des données issues d'une variable aléatoire suivant une distribution des valeurs extrêmes de FRÉCHET standard. Bien sûr, on peut envisager d'autres distributions des valeurs extrêmes comme celle de GUMBEL ou celle de WEIBULL.

L'expression de la fonction de répartition d'une GEVD et l'expression de sa fonction de densité sont connues (voir CHAPITRE 2, Section 2.4). Sous l'hypothèse d'indépendance des données de chaque site d'observation, on peut construire la fonction de vraisemblance des quantités maximales des précipitations journalières de chacun des 79 sites. Ensuite on estimera les paramètres de chaque GEVD qui correspond à chaque site en maximisant le fonction de vraisemblance qui correspond à ce site. Ainsi, on aura $3 \times 79 = 237$ paramètres à estimer. Pour cela, on utilisera l'instruction « `gev.fit` » du package « `ismev` » de logiciel **R**.

Après avoir posé la GEVD appropriée à chaque site d'observation, on transformera les données de chacun de ces sites en des données issues d'une variable aléatoire suivant une distribution des valeurs extrêmes de FRÉCHET standard. Et cela en utilisant la transformation :

$$y \mapsto z = \left[1 + \xi \left(\frac{y - \mu}{\sigma} \right) \right]^{1/\xi}$$

Où les paramètres $\xi \in \mathbb{R}$, $\mu \in \mathbb{R}$ et $\sigma \in \mathbb{R}_+^*$ sont respectivement le paramètre de forme, de localisation et d'échelle de la GEVD. Pour cela, on utilisera l'instruction « `gev2frech` » du package `SpatialExtremes`.

Ainsi, on construira un nouveau tableau de données composé de 47 lignes et 79 colonnes, où chacune de ses composantes correspond à une réalisation d'une variable aléatoire suivant une distribution des valeurs extrêmes de FRÉCHET standard.

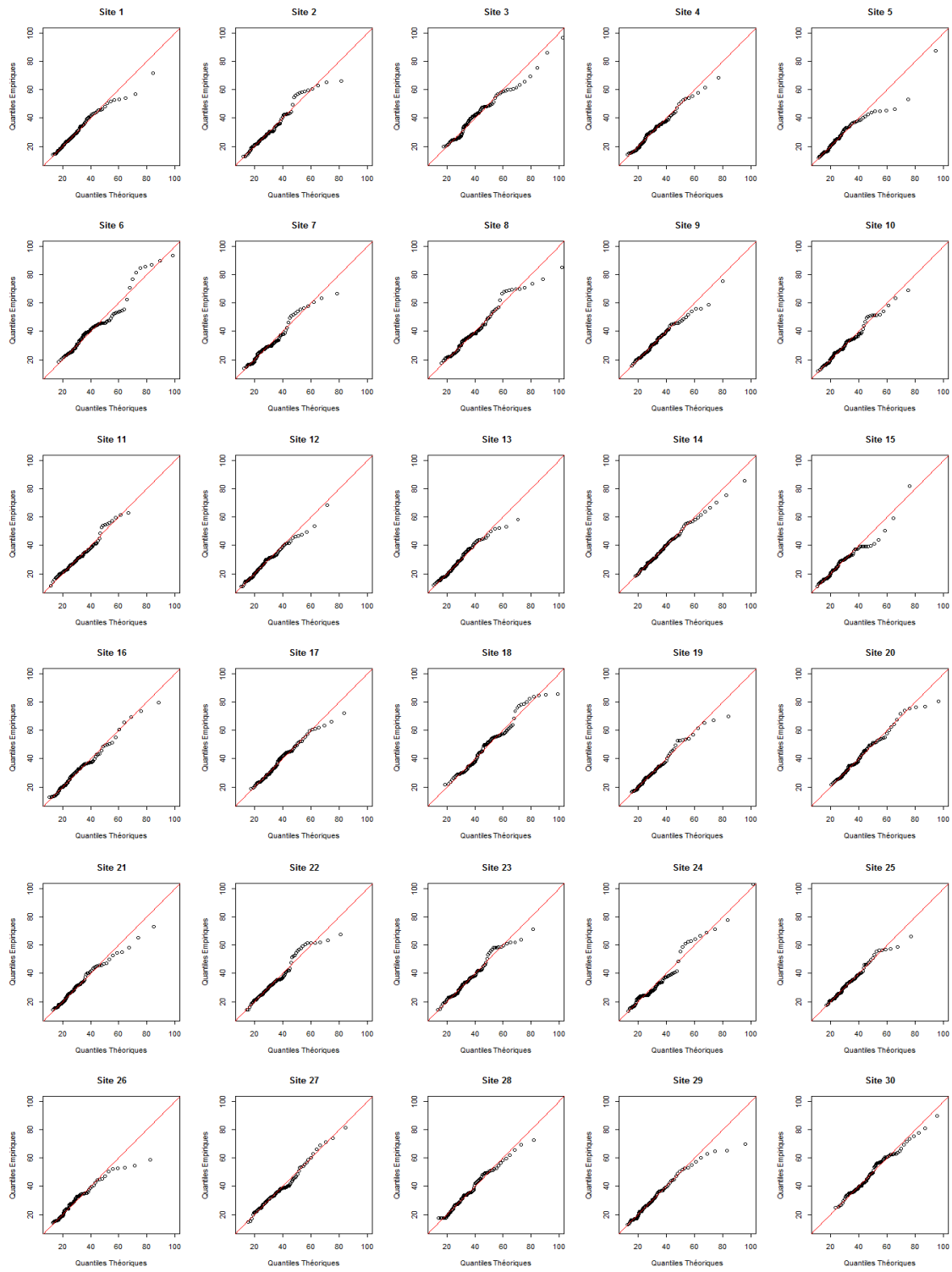
Code R

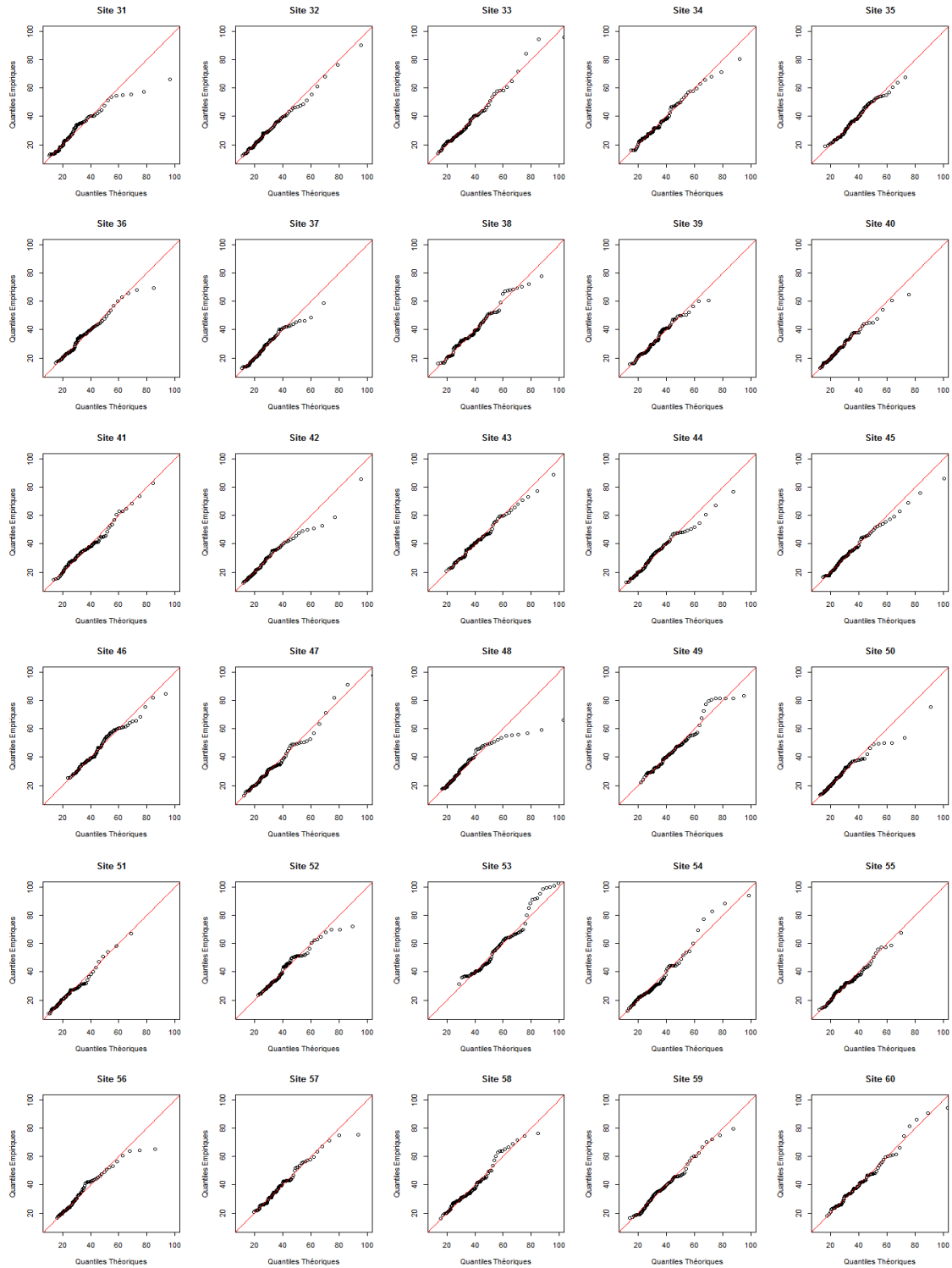
Pour transformer les données du tableau `rain`, qu'on a appelé volontairement « `Y` » pour faciliter les manipulations, on suit les étapes suivantes :

1) On estime le paramètre de localisation, le paramètre d'échelle et le paramètre de forme de la GEVD de chaque site en utilisant le code **R** suivant :

```
> gev.fit(Y[,1])
> gev.fit(Y[,2])
> gev.fit(Y[,3])
```

```
⋮  
> gev.fit(Y[,78])  
> gev.fit(Y[,79])
```





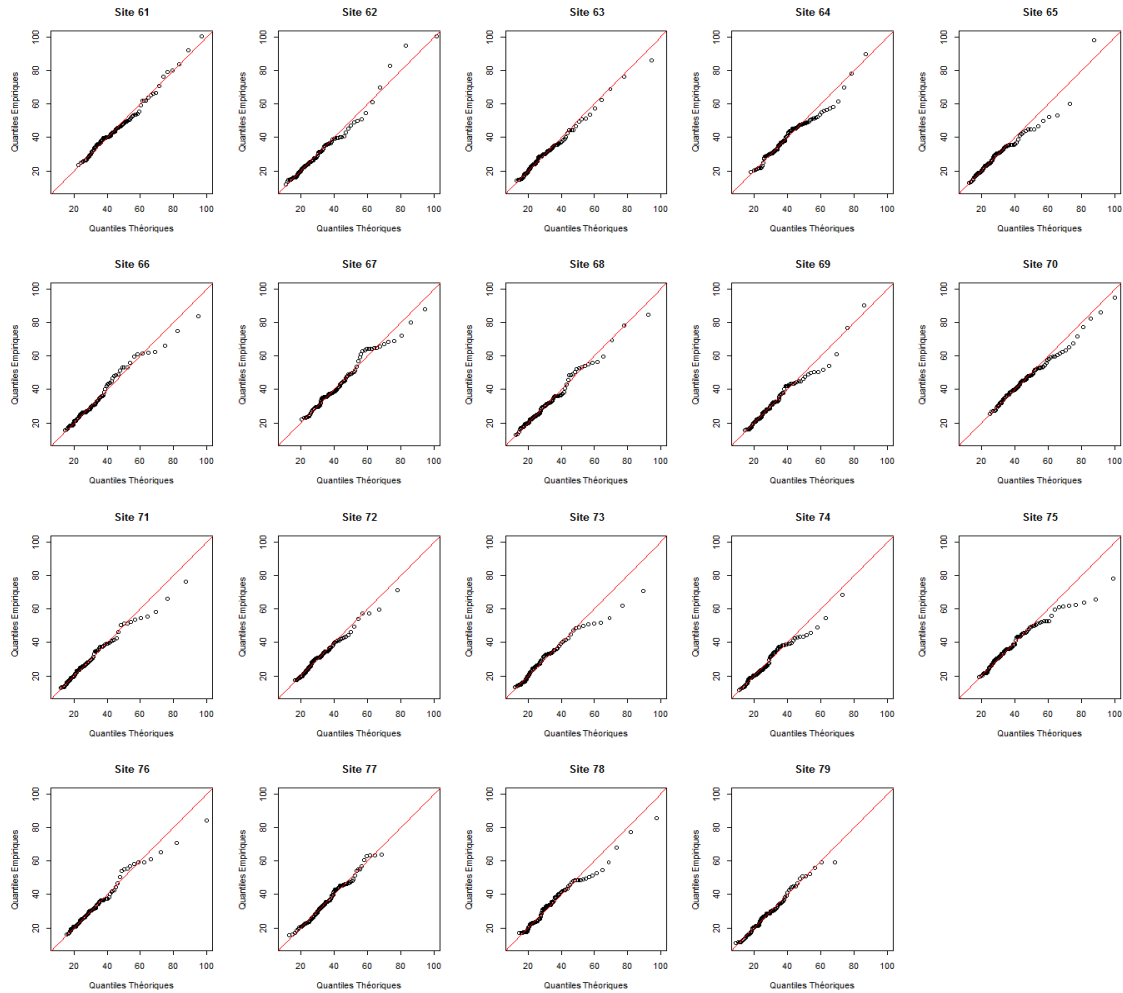


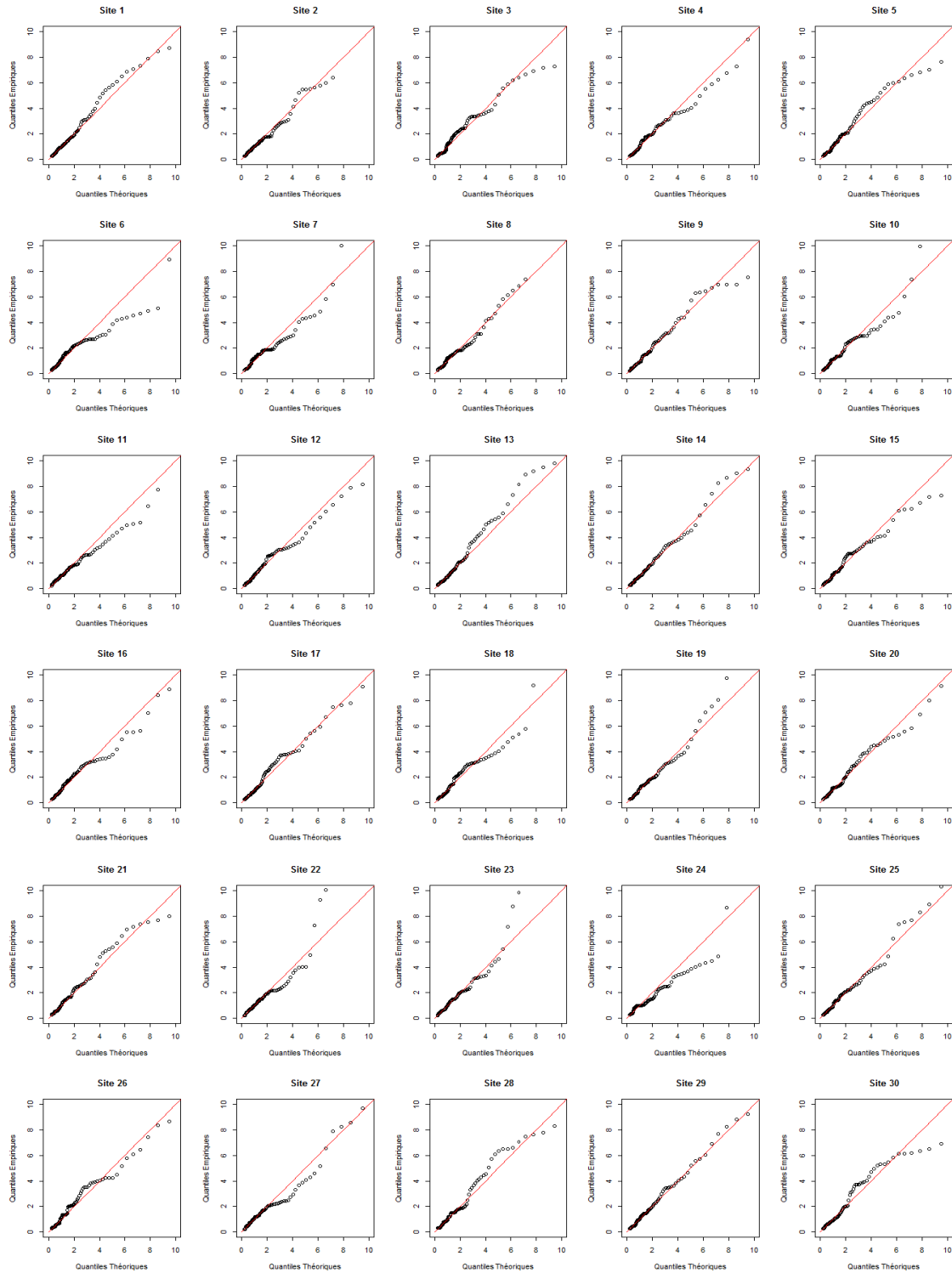
Figure 4.6.2 GEVD Q-Q Plots pour les 79 sites d'observation, où on trace les quantiles des données de chaque site d'observation en fonction des quantiles de sa GEVD estimée.

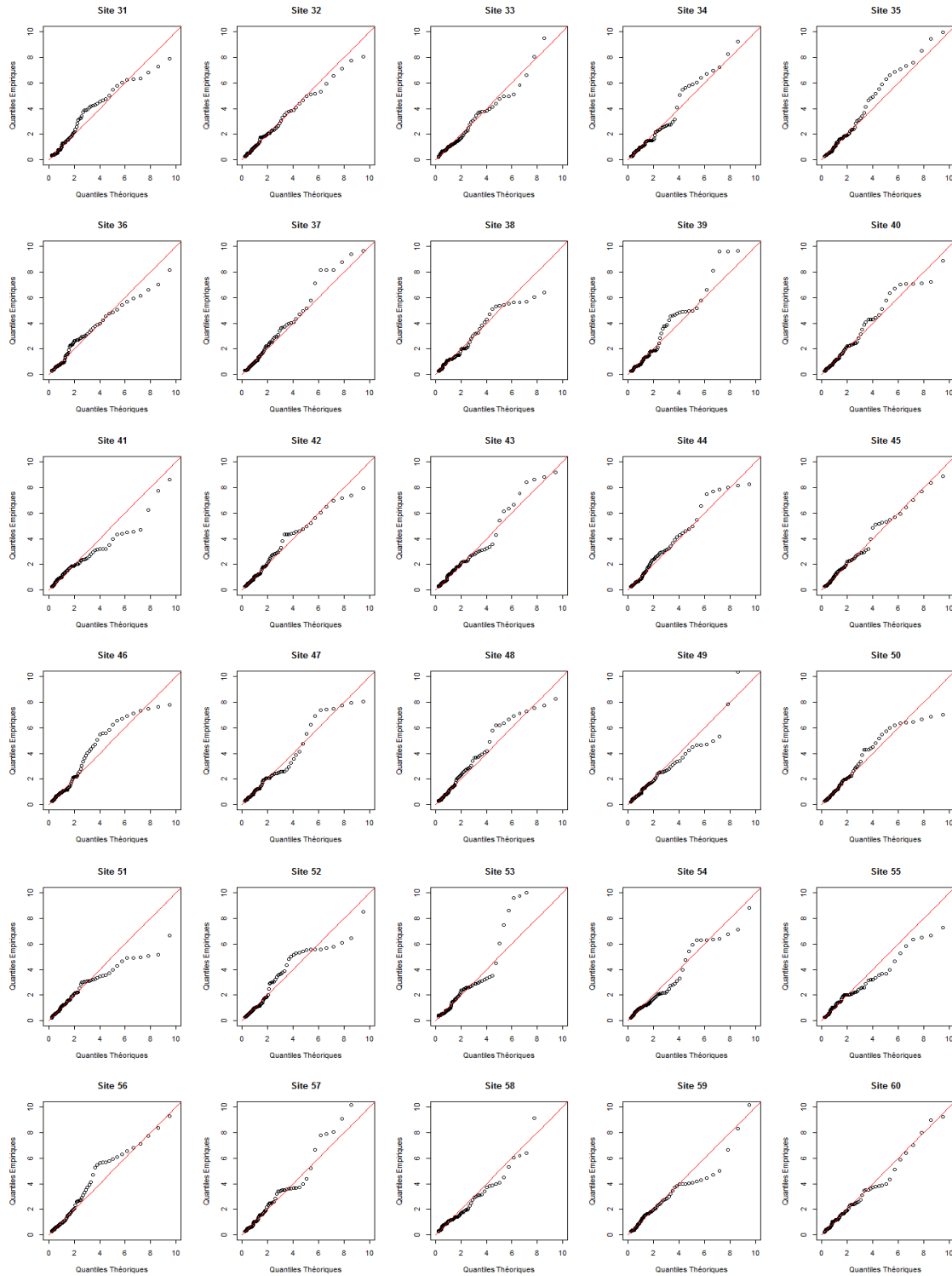
2) On construira une matrice « P » qui contient tous les paramètres estimés. C'est-à-dire que la matrice P comportera 3 lignes et 79 colonnes, où chaque colonne contiendra les paramètres estimés de la GEVD d'un site. La première ligne comportera les paramètres de localisation, la deuxième comportera les paramètres d'échelle et la troisième contiendra les paramètres de forme.

3) Une fois que la matrice P est construite, on construira une autre matrice « Z » composée de 47 lignes et 79 colonnes et qui contiendra les données transformées. Le Code **R** utilisé est le suivant :

```
> Z <- matrix(rep(0,3713), nrow = 47, ncol = 79)
> Z
> for (i in 1:79) Z[,i] <- gev2frech(Y[,i], P[1,i], P[2,i], P[3,i])
> Z
```

Ainsi, les composantes de la matrice Z représenteront des réalisations d'une variable aléatoire suivant une loi de probabilité de FRÉCHET standard.





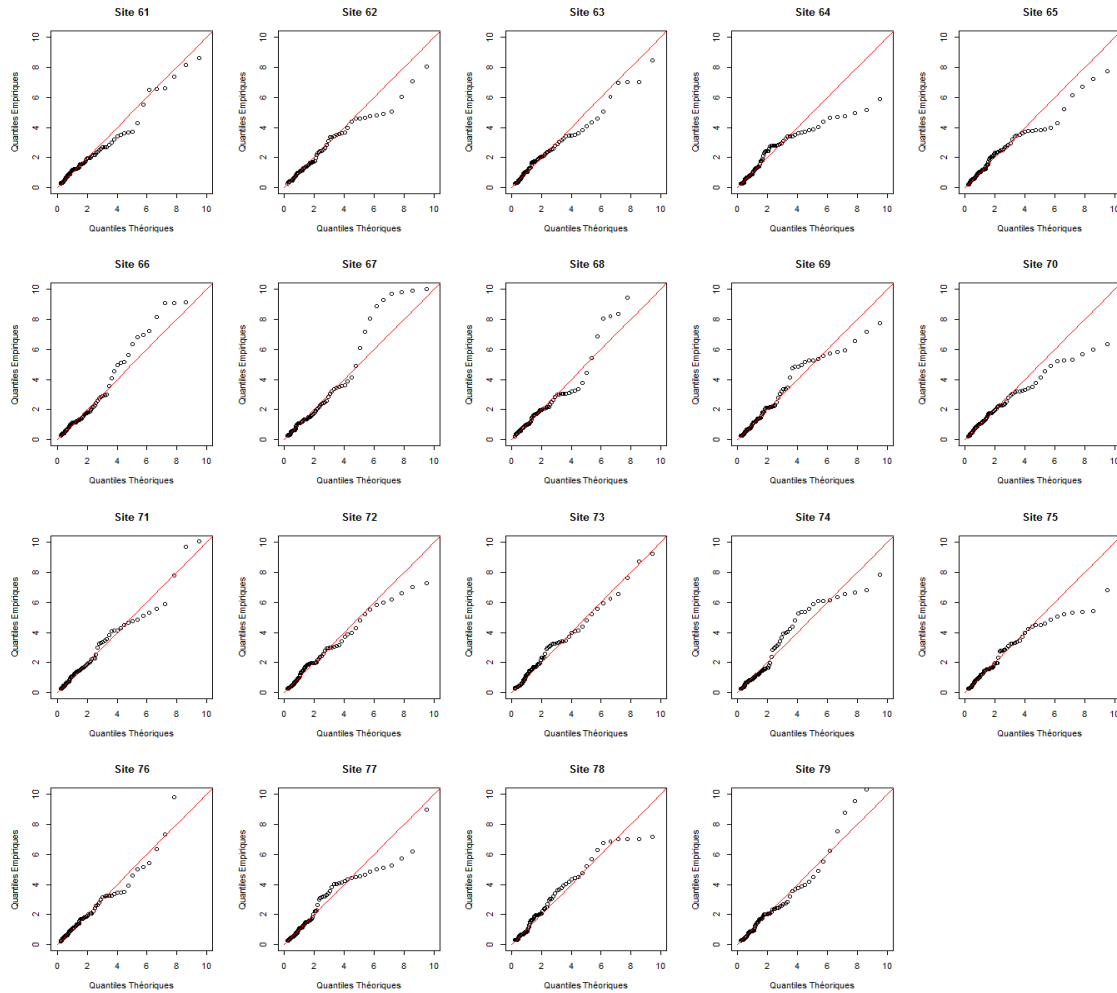


Figure 4.6.3 FRÉCHET unitaire Q-Q Plots pour les 79 sites d'observation après avoir transformé les données de chaque site en des données issues d'une variable aléatoire suivant une loi de probabilité de FRÉCHET unitaire en utilisant les paramètres de la GEVD estimés de chaque site.

4.6.3 Modèle des données

Parmi les modèles des champs aléatoires max-stables on a choisi le modèle de SCHLATHER (voir Section 4.3.2) pour modéliser les quantités maximales des précipitations journalières. On a préféré le modèle de SCHLATHER car ce modèle est flexible et a une interprétation simple (voir [RIBATET (2009)]). Bien sûr, si cela paraît plus raisonnable, le modèle de SCHLATHER pourrait être remplacé par un autre modèle tel le modèle de SMITH ou celui de BROWN-RESNICK, et les mêmes techniques statistiques seront valables pour ce choix alternatif.

La loi de probabilité marginale bivariée du modèle de SCHLATHER est donnée par sa fonction de répartition dans (4.6). Et la fonction de densité marginale bivariée de ce modèle peut être trouvée dans ANNEXE B de la thèse [GHOLAM REZAAE (2010)].

4.6.4 Coefficient extrémal

Comme on l'a déjà dit (voir Section 4.4), il est toujours important de mesurer le degré de dépendance entre les réalisations d'un champ aléatoire. Lorsque les réalisations de ce champ aléatoire ne sont pas extrêmes, les outils qui permettent de mesurer cette dépendance sont la fonction de covariance, la fonction de corrélation et le variogramme. Mais lorsque il s'agit des champs aléatoires max-stables, c'est-à-dire que les réalisations sont extrêmes, ces outils ne seront plus d'une grande utilité et en général ils n'existent pas. Alors les statisticiens ont proposé un autre outil pour mesurer la dépendance des extrêmes qui est le coefficient extrémal. Différentes approches pour estimer le coefficient extrémal sont données dans Section 4.5.1.

Sous l'hypothèse que les données transformées sont des réalisations d'un modèle de SCHLATHER, on estimera le coefficient extrémal par l'approche F -madogramme donnée dans (4.19). Pour cela, on utilisera l'instruction « `fmadogram` » du package `SpatialExtremes`. **Figure 4.6.4** nous montre les valeurs estimées du coefficient extrémal associé aux données en fonction de la distance euclidienne.

Code R

Pour estimer les valeurs du coefficient extrémal associé aux données de la matrice Z , on suit les étapes suivantes :

1) On construit une matrice « $S2$ » qui contiendra seulement les longitudes et les latitudes des 79 sites. C'est-à-dire que $S2$ comportera 79 lignes et 2 colonnes, où la première colonne contiendra les longitudes et la deuxième colonne contiendra les latitudes (chaque ligne correspondra à un site). Le code **R** qu'on utilisera est le suivant :

```
> S2 <- matrix(c(S1[,1],S1[,2]), nrow = 79, ncol = 2)
> S2
```

2) Une fois la matrice $S2$ est construite, on calculera les valeurs du F -madogramme et celles du coefficient extrémal associé aux données en utilisant le code **R** suivant :

```
> fmadogram(Z, S2)
```

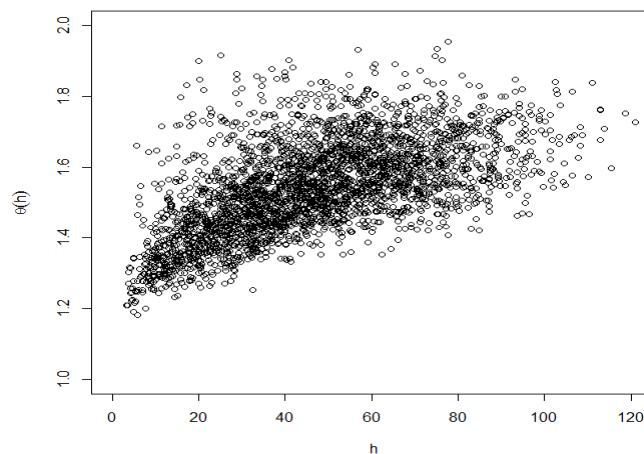


Figure 4.6.4 Valeurs estimées du coefficient extrémal par l'approche F -madogramme en fonction de la distance euclidienne (h représente la distance euclidienne entre deux sites d'observation).

4.6.5 Estimation des paramètres du modèle de SCHLATHER

Après la transformation des données en des données issues d'une variable aléatoire ayant une distribution des valeurs extrêmes de FRÉCHET Standard et avoir choisi le modèle de SCHLATHER pour modéliser les quantités maximales des précipitations journalières, on doit estimer les paramètres de ce modèle pour atteindre l'objectif de cette analyse qui est de trouver le modèle approprié aux données qu'on a. Pour cela, on va utiliser « la méthode de vraisemblance par paires » expliquée dans Section 4.5.2 de ce mémoire.

Comme elle le montre (4.6), la loi de probabilité marginale bivariée d'un modèle de SCHLATHER ne dépend que de corrélogramme ϱ . En ce qui concerne notre exemple, on supposera que le corrélogramme ϱ est issu de la famille de WHITTLE-MATHÉRN définie par :

$$\varrho(h) = m_1 \cdot \frac{2^{1-\ell}}{\Gamma(\ell)} \left(\frac{h}{m_2}\right)^\ell \mathcal{K}_\ell\left(\frac{h}{m_2}\right), \quad 0 \leq m_1 \leq 1, \quad m_2 > 0 \quad \text{et} \quad \ell > 0 \quad (4.20)$$

Où \mathcal{K}_ℓ est une fonction de BESSEL modifiée de deuxième espèce de paramètre ℓ et Γ est la fonction Gamma définie par :

$$\forall x > 0, \quad \Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$$

Remarque : On a choisi de travailler avec la famille des corrélogrammes de WHITTLE-MATHÉRN car elle engendre beaucoup d'autres familles de corrélogrammes. Bien sûr, on peut travailler avec d'autres familles telle la famille stable ou la famille de CAUCHY et les techniques statistiques resteront les mêmes.

Donc, notre objectif est d'estimer les paramètres $0 \leq m_1 \leq 1$, $m_2 > 0$ et $\ell > 0$ qui apparaissent dans (4.20) en employant la méthode de vraisemblance par paires. Pour cela, on utilise l'instruction « fitmaxstab » du package SpatialExtremes. **Table 4.6.1** montre les résultats.

Code R

Pour estimer les paramètres du modèle de SCHLATHER approprié à nos données et comparer son coefficient extrémal avec le coefficient extrémal estimé, on suit les étapes suivantes :

1) On estime les paramètres du corrélogramme du modèle de SCHLATHER par la méthode de vraisemblance par paires. Le code **R** utilisé est le suivant :

```
> Schlather <- fitmaxstab(Z, S2, "whitmat")
> Schlather
```

	m_1	m_2	ℓ
Valeur estimée	0.1967	23.7592	0.9973
Erreur standard	0.06033	17.02832	0.86225

Table 4.6.1 Paramètres estimés du modèle de SCHLATHER approprié aux quantités maximales des précipitations journalières.

2) On compare le coefficient extrême du modèle estimé avec le coefficient extrême estimé précédemment par l'approche F -madogramme (le résultat de la comparaison est donné dans **Figure 4.6.5**). Le code **R** utilisé est le suivant :

```
> fmadogram(fitted = Schlather, which = "ext")
```

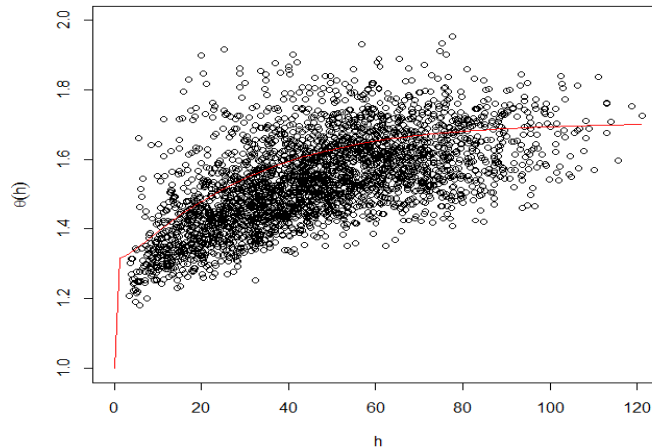


Figure 4.6.5 Les points représentent les valeurs estimées du coefficient extrême par l'approche F -madogramme et la ligne rouge représente le graphe du coefficient extrême du modèle de SCHLATHER approprié aux données (h représente la distance euclidienne entre deux sites d'observation).

Conclusion

Les champs aléatoires max-stables sont une généralisation des distributions des valeurs extrêmes et, par conséquent, sont des candidats idéaux pour la modélisation des extrêmes spatiaux. Dans ce chapitre, qui est le cœur du sujet de ce mémoire, on a essayé de donner une introduction à la théorie des valeurs extrêmes spatiale, où les champs aléatoires max-stables occupent le trône de cette théorie. Le but principal de cette étude est de présenter quelques-uns des modèles, les plus utilisés en applications, des champs aléatoires max-stables et leurs propriétés, ainsi que les méthodes statistiques qui permettent d'estimer leurs paramètres pour terminer avec un exemple d'application où on a illustré la manière dont on peut mettre en application tous les résultats vus dans cette étude. Comme on peut le remarquer, l'approche statistique qui nous a permis de mettre en application la théorie des valeurs extrêmes spatiale (dans ce mémoire) est « l'approche maximums par blocs », alors, avant de terminer ce travail, on signale qu'on peut envisager une autre approche qui est « l'approche excès au-delà d'un seuil ». Cependant, on veut aussi signaler qu'il y a beaucoup de questions non traitées dans ce travail, surtout ce qui concerne l'aspect statistique de la théorie des valeurs extrêmes spatiale, comme l'approche bayésienne, l'approche par copules, les outils de diagnostics, la prédiction, etc. qui sont des sujets d'actualité en plein essor.

Troisième partie

Annexes

ANNEXE A

THÉORÈME DE CONVERGENCE DES TYPES

Introduction

En probabilités et statistiques, beaucoup de résultats de la convergence en loi (convergence en distribution) sont de la forme suivante :

1. Donner une suite de variables aléatoires $\{X_n, n \in \mathbb{N}^*\}$, définies sur l'espace probabilisé (Ω, \mathcal{A}, P) dans $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, et des suites de normalisation $\{a_n, n \in \mathbb{N}^*\}$ de nombres réels strictement positifs et $\{b_n, n \in \mathbb{N}^*\}$ de nombres réels.
2. Ensuite, montrer que la suite de variables aléatoires normalisées $\{a_n^{-1}(X_n - b_n), n \in \mathbb{N}^*\}$ converge en distribution vers une variable aléatoire, non dégénérée, Y .

Si F_n est la fonction de distribution (de répartition) de X_n et G la fonction de distribution de Y , alors la convergence en distribution de la suite de variables aléatoires $\{a_n^{-1}(X_n - b_n), n \in \mathbb{N}^*\}$ vers la variable aléatoire, non dégénérée, Y veut dire que pour tout point x où G est continue :

$$\lim_{n \rightarrow \infty} P\left(\frac{X_n - b_n}{a_n} \leq x\right) = P(Y \leq x) \iff \lim_{n \rightarrow \infty} F_n(a_n x + b_n) = G(x) \quad (\text{A.1})$$

En posant $y = a_n x + b_n$, on aura :

$$(\text{A.1}) \iff \lim_{n \rightarrow \infty} F_n(y) = G\left(\frac{1}{a_n}y - \frac{b_n}{a_n}\right)$$

Exemple. [RESNICK (1998)].

Soit $\{X_n, n \in \mathbb{N}^*\}$ une suite de variables aléatoires indépendantes et identiquement distribuées, définies sur l'espace probabilisé (Ω, \mathcal{A}, P) dans $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ telles que :

$$\forall n \in \mathbb{N}^*, \quad E(X_n) = \mu \in \mathbb{R} \quad \text{et} \quad V(X_n) = \sigma^2 \quad (\sigma > 0)$$

Et soit $\{S_n, n \in \mathbb{N}^*\}$ la suite de variables aléatoires définie par :

$$\forall n \in \mathbb{N}^*, \quad S_n = \sum_{k=1}^n X_k$$

D'après le théorème de la limite centrale, on a :

$$\forall x \in \mathbb{R}, \quad \lim_{n \rightarrow \infty} \mathbb{P}\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$$

Cela veut dire que la suite des sommes partielles normalisées $\left\{\frac{S_n - n\mu}{\sigma\sqrt{n}}, n \in \mathbb{N}^*\right\}$ converge en distribution vers la variable aléatoire qui suit une loi de probabilité normale centrée réduite $\mathcal{N}(0, 1)$.

Maintenant, après avoir vu la forme de la plupart des résultats de la convergence en distribution dans les probabilités et les statistiques, les questions qui viennent à l'esprit sont les suivantes :

- Les constantes de normalisation $a_n \in \mathbb{R}_+^*$ et $b_n \in \mathbb{R}$ sont-elles uniques ?
- Si nous changeons de constantes de normalisation, aurons nous une loi limite différente ? En d'autres termes, est-ce que la loi limite dépend des constantes de normalisation ?
- Si nous avons une loi limite différente en changeant les constantes de normalisation, quel peut être le lien entre ces lois limites ?

La réponse à ces questions sont contenues dans un beau théorème, développé par GNEDENKO et KHINTCHINE, qui s'appelle « **théorème de convergence des types** ».

A.1 Type de distributions

Il est souvent utile de considérer, avec une variable aléatoire X , définie sur un espace probabilisé (Ω, \mathcal{A}, P) à valeurs dans \mathbb{R} , une autre variable aléatoire Y qui est en relation avec X par :

$$Y \stackrel{d}{=} aX + b \quad \text{avec } a \in \mathbb{R}_+^* \text{ et } b \in \mathbb{R} \quad (\text{A.2})$$

Cela veut dire que :

$$\forall x \in \mathbb{R}, \quad (\text{A.2}) \iff \mathbb{P}(Y \leq x) = \mathbb{P}(aX + b \leq x) \quad (\text{A.3})$$

Géométriquement, cette transformation veut dire un changement d'échelle et d'origine.

Si F est la fonction de distribution de X et G la fonction de distribution de Y , alors il est facile de voir que ces deux fonctions sont en relation par :

$$\forall x \in \mathbb{R}, \quad (\text{A.3}) \iff F(x) = G(ax + b) \iff G(x) = F\left(\frac{1}{a}x - \frac{b}{a}\right)$$

Exemple A.1.1 [GNEDENKO & KOLMOGOROV (1954)].

Comme une illustration, supposons que la solution d'un certain problème est la fonction de distribution normale N d'espérance $\mu \in \mathbb{R}$ et de variance σ^2 ($\sigma > 0$) définie par :

$$\forall x \in \mathbb{R}, \quad N(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x \exp\left\{-\frac{(t-\mu)^2}{2\sigma^2}\right\} dt$$

Pour les calculs numériques nous ferons pas une table spéciale pour cette fonction, mais nous utiliserons la table disponible de la fonction de distribution normale centrée réduite Φ définie par :

$$\forall x \in \mathbb{R}, \quad \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$$

qui est en relation avec N par :

$$\forall x \in \mathbb{R}, \quad N(x) = \Phi\left(\frac{1}{\sigma}x - \frac{\mu}{\sigma}\right)$$

De ce que nous avons vu ci-dessus, il est naturel de faire usage du concept suivant :

Définition A.1.1 [RESNICK (1998)].

Soient X et Y deux variables aléatoires réelles qui ont, respectivement, les fonctions de distribution F et G . On dit que F et G sont de **même type** si :

$$\forall x \in \mathbb{R}, \quad \exists a \in \mathbb{R}_+^* \quad \text{et} \quad \exists b \in \mathbb{R} \quad \text{tels que} \quad G(x) = F(ax + b)$$

ou, ce qui est équivalent, X et Y sont de **même type** si :

$$\exists a \in \mathbb{R}_+^* \quad \text{et} \quad \exists b \in \mathbb{R} \quad \text{tels que} \quad X \stackrel{d}{=} aY + b$$

Autrement dit : Les variables aléatoires de même type sont celles qui ont la même distribution, à un facteur de localisation et d'échelle près.

Remarque : Il est facile de vérifier que la propriété « être de même type » est une relation d'équivalence dans l'espace des fonctions de distribution, où chaque classe d'équivalence forme un « type de distributions ».

Nous allons voir dans le résultat qui suit, qu'on va prendre comme exemple, que toutes les fonctions de distribution normales forment un seul type de distributions, qu'on appelle « type normal ».

Exemple A.1.2 [GNEDENKO & KOLMOGOROV (1954)].

Soit X et Y deux variables aléatoires réelles telles que :

$$X \sim \mathcal{N}(\mu_1, \sigma_1^2) \quad \text{et} \quad Y \sim \mathcal{N}(\mu_2, \sigma_2^2) \quad \text{avec} \quad (\mu_1, \mu_2) \in \mathbb{R}^2 \quad \text{et} \quad (\sigma_1, \sigma_2) \in \mathbb{R}_+^* \times \mathbb{R}_+^*$$

alors :

$$Y \stackrel{d}{=} \frac{\sigma_2}{\sigma_1} X + \frac{\mu_2\sigma_1 - \mu_1\sigma_2}{\sigma_1}$$

A.2 Théorème de convergence des types

Le résultat suivant est un outil particulièrement important pour la convergence en distribution. Il nous informe que la loi limite d'une suite de variables aléatoires normalisées est uniquement déterminée à un facteur de position et d'échelle près.

Théorème A.2.1 [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

Soient $\{X_n, n \in \mathbb{N}^*\}$ une suite de variables aléatoires réelles, avec F_n la fonction de distribution de X_n , et Y, Z deux variables aléatoires réelles non dégénérées ayant, respectivement, les fonctions de distribution G et H , et soient les suites de nombres réels $\{a_n > 0, n \in \mathbb{N}^*\}$, $\{\alpha_n > 0, n \in \mathbb{N}^*\}$, $\{b_n \in \mathbb{R}, n \in \mathbb{N}^*\}$ et $\{\beta_n \in \mathbb{R}, n \in \mathbb{N}^*\}$.

Supposons que :

$$\forall x \in \mathcal{C}(G) = \{x \in \mathbb{R}, G \text{ continue en } x\}, \quad \lim_{n \rightarrow \infty} F_n(a_n x + b_n) = G(x) \quad (\text{A.4})$$

Alors :

1. *On a :*

$$\forall x \in \mathcal{C}(H) = \{x \in \mathbb{R}, H \text{ continue en } x\}, \quad \lim_{n \rightarrow \infty} F_n(\alpha_n x + \beta_n) = H(x) \quad (\text{A.5})$$

si et seulement si :

$$\lim_{n \rightarrow \infty} \frac{a_n}{\alpha_n} = a \in \mathbb{R}_+^* \quad \text{et} \quad \lim_{n \rightarrow \infty} \frac{b_n - \beta_n}{\alpha_n} = b \in \mathbb{R}$$

2. *Si (A.5) est vérifiée, alors :*

$$Z \stackrel{d}{=} aY + b \quad (\text{A.6})$$

et a, b sont les seules et uniques constantes qui vérifient (A.6).

3. *Supposons que (A.5) est vérifiée, alors : Y est non dégénérée si et seulement si $a > 0$, et, dans ce cas, Y et Z sont de même type.*

Démonstration.

Pour une démonstration du **Théorème A.2.1**, le lecteur pourra se référer à l'ouvrage [RESNICK (1998)] ou [GNEDENKO & KOLMOGOROV (1954)] ou, encore, à [RESNICK (1987)]. ■

ANNEXE B

THÉORIE DE VARIATION RÉGULIÈRE

Introduction

Bien que des traitements préliminaires ou partiels de la théorie de variation régulière se trouvent dans des travaux qui précèdent celui de Jovan KARAMATA : LANDAU en 1911, VALIRON en 1913, POLYÀ en 1917 et autres, c'est à KARAMATA qu'on doit le développement de cette théorie depuis l'apparition de son papier célèbre « Sur un mode de croissance régulière des fonctions », en 1930, jusqu'aux années soixante. Avec William FELLER, dans son ouvrage « An Introduction to Probability Theory and its Applications », la théorie de variation régulière est reconnue comme un outil indispensable dans la théorie des probabilités et ses applications. Un nouvel élan au sujet a été fourni, en 1970, par Laurens DE HAAN dans sa thèse « On Regular Variation and its Applications to the Weak Convergence of Sample Extremes » où il a introduit une généralisation substantielle de la théorie de variation régulière, en visant, encore, la théorie des probabilités et ses applications. Aujourd'hui, cette théorie a un statut de **connaissance standard** pour tout probabiliste et statisticien.

B.1 Fonction à variation lente

Définition B.1.1 [BINGHAM, GOLDIE & TEUGELS (1987)].

Une fonction mesurable $L :]a, +\infty[\rightarrow \mathbb{R}_+$ ($a \geq 0$) est dite à *variation lente à l'infini*, si :

$$\forall x > 0, \quad \lim_{t \rightarrow +\infty} \frac{L(tx)}{L(t)} = 1 \quad (\text{B.1})$$

Remarque : Dans ce qui suit, lorsque on parle de voisinage de l'infini, on entend par là le voisinage de l'infini qui est de la forme $]a, +\infty[$ où $a \geq 0$. Donc, une fonction L est à variation lente à l'infini si, et seulement si, elle est définie mesurable, positive sur un voisinage de l'infini et vérifie (B.1).

B.1.1 Représentation de KARAMATA

Beaucoup de résultats concernant les fonctions aux variations lentes à l'infini sont dus à Jovan KARAMATA, et l'un des résultats les plus importants est le théorème de représentation suivant :

Théorème B.1.1 (Théorème de représentation) [BINGHAM, GOLDIE & TEUGELS (1987)].

Une fonction L est à variation lente à l'infini **si et seulement si** elle peut être représentée sous la forme suivante :

$$\forall x \geq b > a, \quad L(x) = c(x) \exp \left\{ \int_b^x \frac{\delta(u)}{u} du \right\} \quad (\text{B.2})$$

où : c et δ sont des fonctions définies mesurables sur $]a, +\infty[$ ($a \geq 0$) telles que :

$$\lim_{x \rightarrow +\infty} c(x) = c_0 \in]0, +\infty[\quad \text{et} \quad \lim_{x \rightarrow +\infty} \delta(x) = 0$$

Remarque : On appelle (B.2) **représentation de KARAMATA** de la fonction à variation lente à l'infini.

B.1.2 Exemples

Avec la définition ou le théorème de représentation d'une fonction à variation lente à l'infini, on peut vérifier que toute fonction définie mesurable et positive sur un voisinage de l'infini qui a une limite positive à l'infini est une fonction qui varie lentement à l'infini. La fonction logarithme ($x \mapsto \ln x$), les itérations du logarithme ($x \mapsto \ln \ln \dots \ln x$) et la fonction $x \mapsto \exp \{ (\ln x)^\gamma \}$ ($0 < \gamma < 1$) sont, aussi, des fonctions qui varient lentement à l'infini.

Exemple B.1.1

Soit $L :]1, +\infty[\rightarrow \mathbb{R}_+$ une fonction définie par :

$$\forall x \in]1, +\infty[, \quad L(x) = (1 + x^{-1}) \ln x$$

Alors L est une fonction à variation lente à l'infini.

Vérification :

1) En utilisant (B.1) :

$$\forall x > 0, \quad \lim_{t \rightarrow +\infty} \frac{L(tx)}{L(t)} = \lim_{t \rightarrow +\infty} \frac{[1 + (tx)^{-1}] \ln(tx)}{(1 + t^{-1}) \ln t} = 1$$

2) En utilisant (B.2) :

$$\forall x \geq e > 1, \quad L(x) = (1 + x^{-1}) \exp \left\{ \int_e^x \frac{(\ln u)^{-1}}{u} du \right\}$$

B.1.3 Propriétés des fonctions aux variations lentes

Dans la proposition qui suit, on va donner quelques propriétés élémentaires des fonctions aux variations lentes :

Proposition B.1.1 [DE HAAN & FERREIRA (2006) et BINGHAM, GOLDIE & TEUGELS (1987)].
Soient L, L_1, L_2, \dots, L_n des fonctions définies mesurables et positives sur un voisinage de l'infini.
On a :

1. **Si** L est à variation lente à l'infini, **alors** :

$$\lim_{x \rightarrow +\infty} \frac{\ln L(x)}{\ln x} = 0$$

2. **Si** L est à variation lente à l'infini et $\alpha > 0$, **alors** :

$$\lim_{x \rightarrow +\infty} x^\alpha L(x) = +\infty \quad \text{et} \quad \lim_{x \rightarrow +\infty} x^{-\alpha} L(x) = 0$$

3. **Si** L est à variation lente à l'infini, **alors** $L^\alpha : x \mapsto [L(x)]^\alpha$ est une fonction à variation lente pour tout $\alpha \in \mathbb{R}$.

4. **Si** L_1 et L_2 sont aux variations lentes à l'infini, **alors** $L_1 + L_2 : x \mapsto L_1(x) + L_2(x)$ et $L_1 \cdot L_2 : x \mapsto L_1(x) \cdot L_2(x)$ sont des fonctions aux variations lentes à l'infini, et **si**, de plus, $\lim_{x \rightarrow +\infty} L_2(x) = +\infty$, **alors** $L_1 \circ L_2 : x \mapsto L_1[L_2(x)]$ est aussi une fonction à variation lente à l'infini.

5. **Si** L_1, L_2, \dots, L_n sont aux variations lentes à l'infini et R une fraction rationnelle définie sur \mathbb{R}_+^n à coefficients positifs, **alors** $R(L_1, L_2, \dots, L_n) : x \mapsto R[L_1(x), L_2(x), \dots, L_n(x)]$ est une fonction à variation lente à l'infini.

B.1.4 Théorème de KARAMATA

Théorème B.1.2 (Théorème de KARAMATA) [BINGHAM, GOLDIE & TEUGELS (1987)].

Soit L une fonction à variation lente à l'infini définie sur le voisinage $]a, +\infty[$ ($a \geq 0$) de l'infini.
Alors :

• Pour $\alpha \geq -1$:

$$\exists b > a \quad \text{tel que} \quad \forall x \geq b, \quad \lim_{x \rightarrow +\infty} \frac{x^{\alpha+1} L(x)}{\int_b^x t^\alpha L(t) dt} = \alpha + 1$$

• Pour $\alpha < -1$ (ou : $\alpha = -1$ et $\int_x^{+\infty} t^{-1} L(t) dt < \infty$) :

$$\forall x > a, \quad \int_x^{+\infty} t^\alpha L(t) dt < \infty \quad \text{et} \quad \lim_{x \rightarrow +\infty} \frac{x^{\alpha+1} L(x)}{\int_x^{+\infty} t^\alpha L(t) dt} = -\alpha - 1$$

Remarque : Dans le cadre du théorème ci-dessus, on a un résultat supplémentaire pour le cas de $\alpha = -1$: Les fonctions $\ell_1 : x \mapsto \int_b^x t^{-1} L(t) dt$ et $\ell_2 : x \mapsto \int_x^{+\infty} t^{-1} L(t) dt$ sont des fonctions aux variations lentes à l'infini.

B.2 Fonction à variation régulière

Définition B.2.1 [BINGHAM, GOLDIE & TEUGELS (1987)].

Une fonction mesurable $h :]a, +\infty[\rightarrow \mathbb{R}_+$ ($a \geq 0$) est dite à variation régulière à l'infini d'indice $\rho \in \mathbb{R}$, et on note $h \in \text{RV}_\rho$, si :

$$\forall x > 0, \quad \lim_{t \rightarrow +\infty} \frac{h(tx)}{h(t)} = x^\rho \quad (\text{B.3})$$

Remarques :

1. Une fonction h est à variation régulière à l'infini d'indice $\rho \in \mathbb{R}$ si, et seulement si, elle est définie mesurable, positive sur un voisinage de l'infini et vérifie (B.3).
2. Si $h \in \text{RV}_0$ alors h est une fonction à variation lente à l'infini. Donc une fonction à variation lente à l'infini est une fonction à variation régulière à l'infini d'indice $\rho = 0$.

Comme on a défini la notion de variation régulière à l'infini, on peut, aussi, la définir à l'origine :

Définition B.2.2 [MARIĆ (2000)].

Une fonction mesurable $h :]0, a[\rightarrow \mathbb{R}_+$ ($a > 0$) est dite à variation régulière à l'origine (à droite de l'origine) d'indice $\rho \in \mathbb{R}$, et on note $h \in \text{RV}_\rho^0$, si :

$$\forall x > 0, \quad \lim_{t \rightarrow 0^+} \frac{h(tx)}{h(t)} = x^\rho \quad (\text{B.4})$$

Remarque : Dans le cadre de la définition ci-dessus, dire que $h \in \text{RV}_\rho^0$ est équivalent à dire que la fonction $x \mapsto h(1/x)$ est à variation régulière à l'infini d'indice $-\rho$.

B.2.1 Théorème de caractérisation

Dans le théorème qui suit, on va donner des caractérisations ou des définitions équivalentes de la notion de variation régulière à l'infini :

Théorème B.2.1 (Théorème de caractérisation) [BINGHAM, GOLDIE & TEUGELS (1987)].

Soient $h :]a, +\infty[\rightarrow \mathbb{R}_+$ ($a \geq 0$) une fonction mesurable et $\rho \in \mathbb{R}$. Alors les assertions suivantes sont équivalentes :

- 1) $h \in \text{RV}_\rho$.
- 2) Il existe $L \in \text{RV}_0$ telle que :

$$\forall x > a, \quad h(x) = x^\rho L(x) \quad (\text{B.5})$$

- 3) Il existe une fonction g positive telle que :

$$\forall y > 0, \quad \lim_{t \rightarrow +\infty} \frac{h(ty)}{h(t)} = g(y)$$

et dans ce cas $g(y) = y^\rho$.

Proposition B.2.1 [MARIĆ (2000)].

Soient f et g deux fonctions définies mesurables et positives sur un voisinage de l'infini et soit $\rho \in \mathbb{R}$. Alors :

$$(g \in \text{RV}_\rho \text{ et } f \sim g \text{ à } +\infty) \implies f \in \text{RV}_\rho$$

B.2.2 Représentation de KARAMATA

Comme dans le cas de la notion de variation lente à l'infini, grâce à (B.5) et (B.2), on peut donner une représentation des fonctions qui varient régulièrement à l'infini :

Théorème B.2.2 (Théorème de représentation) [DE HAAN & FERREIRA (2006)].

Une fonction h est à variation régulière à l'infini d'indice $\rho \in \mathbb{R}$ **si et seulement si** elle peut être représentée sous la forme suivante :

$$\forall x \geq b > a, \quad h(x) = c(x) \exp \left\{ \int_b^x \frac{\delta(u)}{u} du \right\} \quad (\text{B.6})$$

où : c et δ sont des fonctions définies mesurables sur $]a, +\infty[$ ($a \geq 0$) telles que :

$$\lim_{x \rightarrow +\infty} c(x) = c_0 \in]0, +\infty[\quad \text{et} \quad \lim_{x \rightarrow +\infty} \delta(x) = \rho$$

Remarque : On appelle (B.6) **représentation de KARAMATA** de la fonction $h \in \text{RV}_\rho$.

B.2.3 Exemples

En utilisant (B.3), (B.5) ou (B.6), on peut vérifier que pour ρ et γ dans \mathbb{R} , les fonctions : $x \mapsto x^\rho$, $x \mapsto x^\rho \ln(1+x)$, $x \mapsto [x \ln(1+x)]^\rho$, $x \mapsto x^\rho (\ln x)^\gamma$ et $x \mapsto x^\rho (\ln x \ln x)^\gamma$ varient régulièrement à l'infini d'indice ρ .

Dans la théorie des probabilités et ses applications, on s'intéresse aux lois de probabilité qui ont des fonctions de queue (fonctions de survie) qui varient régulièrement à l'infini. Dès lors, on dit que ces lois de probabilité ont des queues lourdes ou épaisses (à droite) :

Exemple B.2.1 [RESNICK (2006)].

Soit X une variable aléatoire à valeurs dans \mathbb{R} qui suit une loi de probabilité de PARÉTO de fonction de répartition F , i.e :

$$\forall x \in \mathbb{R}, \quad \text{P}(X \leq x) = F(x) = \begin{cases} 0 & \text{si } x < 1 \\ 1 - x^{-\alpha} & \text{si } x \geq 1 \end{cases} \quad \text{avec } \alpha > 0$$

Alors :

$$\forall x \geq 1, \quad \text{P}(X > x) = 1 - F(x) = x^{-\alpha}$$

Donc : $(1 - F) \in \text{RV}_{-\alpha}$ et la loi de probabilité de PARÉTO est une loi à queue lourde.

Exemple B.2.2 [RESNICK (2006)].

Soit X une variable aléatoire à valeurs dans \mathbb{R} qui suit une loi de probabilité stable $\mathcal{S}_\alpha(\beta, \mu, \sigma)$ ($0 < \alpha < 2$) de fonction de répartition F , alors :

$$\lim_{x \rightarrow +\infty} x^\alpha \mathbb{P}(X > x) = \frac{1 + \beta}{2} \sigma^\alpha C_\alpha \quad \text{avec} \quad C_\alpha = \frac{2}{\pi} \Gamma(\alpha) \sin\left(\frac{\pi}{2} \alpha\right)$$

En posant : $C = \frac{1 + \beta}{2} \sigma^\alpha C_\alpha$, on aura :

$$\mathbb{P}(X > x) = 1 - F(x) \sim Cx^{-\alpha} \quad \text{quand} \quad x \rightarrow +\infty$$

Donc : $(1 - F) \in \text{RV}_{-\alpha}$ et la loi de probabilité stable $\mathcal{S}_\alpha(\beta, \mu, \sigma)$ ($0 < \alpha < 2$) est à queue lourde.

Exemple B.2.3 [RESNICK (2006)].

Soit X une variable aléatoire à valeurs dans \mathbb{R} qui suit une loi de probabilité de valeurs extrêmes de type FRÉCHET de fonction de répartition Φ_α ($\alpha > 0$), i.e :

$$\forall x \in \mathbb{R}, \quad \mathbb{P}(X \leq x) = \Phi_\alpha(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ \exp(-x^{-\alpha}) & \text{si } x > 0 \end{cases}$$

Alors :

$$\forall x > 0, \quad \mathbb{P}(X > x) = 1 - \Phi_\alpha(x) = 1 - \exp(-x^{-\alpha})$$

Or :

$$1 - \exp(-x^{-\alpha}) \sim x^{-\alpha} \quad \text{quand} \quad x \rightarrow +\infty$$

Donc : $(1 - \Phi_\alpha) \in \text{RV}_{-\alpha}$ et la loi de probabilité de valeurs extrêmes de type FRÉCHET est une loi à queue lourde.

B.2.4 Propriétés des fonctions aux variations régulières

On va résumer, dans la proposition qui suit, quelques propriétés élémentaires des fonctions aux variations régulières à l'infini :

Proposition B.2.2 [DE HAAN & FERREIRA (2006) et BINGHAM, GOLDIE & TEUGELS (1987)].

Soient h, h_1, h_2, \dots, h_n des fonctions définies mesurables et positives sur un voisinage de l'infini et soient $\rho, \rho_1, \rho_2, \dots, \rho_n$ dans \mathbb{R} . On a :

1. **Si** $h \in \text{RV}_\rho$, **alors** :

$$\lim_{x \rightarrow +\infty} \frac{\ln h(x)}{\ln x} = \rho$$

2. **Si** $h \in \text{RV}_\rho$ et $\rho \neq 0$, **alors** :

$$\lim_{x \rightarrow +\infty} h(x) = \begin{cases} 0 & \text{si } \rho < 0 \\ +\infty & \text{si } \rho > 0 \end{cases}$$

3. Si $h \in \text{RV}_\rho$ et $\alpha \in \mathbb{R}$, alors :

$$h^\alpha : x \mapsto [h(x)]^\alpha \in \text{RV}_{\alpha\rho}$$

4. Si $h_1 \in \text{RV}_{\rho_1}$ et $h_2 \in \text{RV}_{\rho_2}$, alors :

$$h_1 + h_2 : x \mapsto h_1(x) + h_2(x) \in \text{RV}_{\max(\rho_1, \rho_2)}$$

et si, de plus, $\lim_{x \rightarrow +\infty} h_2(x) = +\infty$, alors :

$$h_1 \circ h_2 : x \mapsto h_1[h_2(x)] \in \text{RV}_{\rho_1\rho_2}$$

5. Si $h_i \in \text{RV}_{\rho_i}$ ($i = 1, 2, \dots, n$) et R une fraction rationnelle définie sur \mathbb{R}_+^n à coefficients positifs, alors :

$$R(h_1, h_2, \dots, h_n) : x \mapsto R[h_1(x), h_2(x), \dots, h_n(x)] \in \text{RV}_\sigma \quad \text{avec} \quad \sigma \in \mathbb{R}$$

B.2.5 Théorème de KARAMATA

Théorème B.2.3 (Théorème de KARAMATA) [BINGHAM, GOLDIE & TEUGELS (1987)].

Soit h une fonction définie mesurable et positive sur le voisinage $]a, +\infty[$ ($a \geq 0$) de l'infini et soit $(\rho, \sigma) \in \mathbb{R}^2$. Alors $h \in \text{RV}_\rho$ si et seulement si :

- Pour $\sigma \geq -(\rho + 1)$:

$$\exists b > a \quad \text{tel que} \quad \forall x \geq b, \quad \lim_{x \rightarrow +\infty} \frac{x^{\sigma+1}h(x)}{\int_b^x t^\sigma h(t) dt} = \sigma + \rho + 1$$

- Pour $\sigma < -(\rho + 1)$ (ou : $\sigma = -(\rho + 1)$ et $\int_x^{+\infty} t^\sigma h(t) dt < \infty$) :

$$\forall x > a, \quad \int_x^{+\infty} t^\sigma h(t) dt < \infty \quad \text{et} \quad \lim_{x \rightarrow +\infty} \frac{x^{\sigma+1}h(x)}{\int_x^{+\infty} t^\sigma h(t) dt} = -(\sigma + \rho + 1)$$

Corollaire B.2.1 [DE HAAN & FERREIRA (2006)].

Soient h une fonction définie sur le voisinage $]a, +\infty[$ ($a \geq 0$) de l'infini et $\rho \in \mathbb{R}$.

- Si $h \in \text{RV}_\rho$ et $\rho \geq -1$, alors :

$$\exists b > a \quad \text{tel que} \quad \forall x \geq b, \quad x \mapsto \int_b^x h(t) dt \in \text{RV}_{\rho+1}$$

- Si $h \in \text{RV}_\rho$ et $\rho < -1$ (ou : $\rho = -1$ et $\int_x^{+\infty} h(t) dt < \infty$), alors :

$$\forall x > a, \quad \int_x^{+\infty} h(t) dt < \infty \quad \text{et} \quad x \mapsto \int_x^{+\infty} h(t) dt \in \text{RV}_{\rho+1}$$

Corollaire B.2.2 (VON MISES) [RESNICK (2006) et MARIĆ (2000)].

Soit H une fonction définie positive et dérivable (de dérivée h) sur le voisinage $]a, +\infty[$ ($a \geq 0$) de l'infini, i.e :

$$\forall x > a, \quad H'(x) = h(x)$$

Si :

$$\lim_{x \rightarrow +\infty} \frac{xh(x)}{H(x)} = \rho, \quad \rho \in \mathbb{R}$$

Alors :

$$H \in \text{RV}_\rho$$

B.2.6 Théorème de densité monotone

Théorème B.2.4 (Théorème de densité monotone) [BINGHAM, GOLDIE & TEUGELS (1987)].

Soit H une fonction définie sur le voisinage $]a, +\infty[$ ($a \geq 0$) de l'infini telle que :

$$\forall x \geq b > a, \quad H(x) = \int_b^x h(u) du \quad \text{ou} \quad H(x) = \int_x^{+\infty} h(u) du$$

où h est une fonction monotone sur $]b, +\infty[$ ($b > a$), et soient $L \in \text{RV}_0$, $C \in \mathbb{R}$ et $\rho \in \mathbb{R}$.

Si :

$$H(x) \sim Cx^\rho L(x) \quad \text{quand} \quad x \rightarrow +\infty \tag{B.7}$$

Alors :

$$h(x) \sim C\rho x^{\rho-1} L(x) \quad \text{quand} \quad x \rightarrow +\infty$$

Remarque : Dans le cadre du théorème ci-dessus, si $C > 0$ alors (B.7) veut dire que $H \in \text{RV}_\rho$. Mais (B.7) n'implique pas que $h \in \text{RV}_{\rho-1}$, sauf dans le cas où $C\rho > 0$.

Proposition B.2.3 [DE HAAN & FERREIRA (2006)].

Soient H et h deux fonctions définies sur le voisinage $]a, +\infty[$ ($a \geq 0$) de l'infini et $\rho \in \mathbb{R}$.

- Si $H \in \text{RV}_\rho$ ($\rho \geq 0$), h monotone et :

$$\forall x \geq b > a, \quad H(x) = H(b) + \int_b^x h(u) du$$

alors :

$$\lim_{x \rightarrow +\infty} \frac{xh(x)}{H(x)} = \rho$$

- Si $H \in \text{RV}_\rho$ ($\rho \leq 0$), h décroissante et :

$$\forall x > a, \quad H(x) = \int_x^{+\infty} h(u) du$$

alors :

$$\lim_{x \rightarrow +\infty} \frac{xh(x)}{H(x)} = -\rho$$

Remarque : Dans le cadre de la proposition ci-dessus, si $\rho \neq 0$, on aura, de plus, $h \in \text{RV}_{\rho-1}$.

B.2.7 Variation régulière et théorie des probabilités

Dans la proposition qui suit, on va résumer quelques-uns des résultats utiles dans la théorie des probabilités et ses applications :

Proposition B.2.4 [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

Soit X une variable aléatoire à valeurs dans \mathbb{R} de fonction de répartition F et de fonction de survie $\bar{F} = 1 - F$.

1. X a une densité f , alors : $f \in \text{RV}_{-\rho-1}$ ($\rho > 0$) *si et seulement si* :

$$\lim_{x \rightarrow +\infty} \frac{xf(x)}{\bar{F}(x)} = \rho$$

2. *Si* X a une densité f telle que :

$$\lim_{x \rightarrow +\infty} \frac{xf(x)}{\bar{F}(x)} = \rho, \quad \rho \in \mathbb{R}$$

alors :

$$\bar{F} \in \text{RV}_{-\rho}$$

3. *Si* X a une densité f décroissante sur $]b, +\infty[$ ($b \in \mathbb{R}$) et $\bar{F} \in \text{RV}_{-\rho}$ ($\rho \geq 0$), *alors* :

$$\lim_{x \rightarrow +\infty} \frac{xf(x)}{\bar{F}(x)} = \rho$$

Si $\rho \neq 0$, on aura, de plus, $f \in \text{RV}_{-\rho-1}$.

4. Soient $\rho > 0$ et $\gamma > \rho$, alors : $\bar{F} \in \text{RV}_{-\rho}$ *si et seulement si* :

$$\lim_{x \rightarrow +\infty} \frac{x^\gamma \bar{F}(x)}{\int_0^x t^\gamma dF(t)} = \frac{\gamma - \rho}{\rho}$$

5. *Si* $\bar{F} \in \text{RV}_{-\rho}$ ($\rho > 0$), *alors* :

$$\lim_{x \rightarrow +\infty} \frac{x^\rho \bar{F}(x)}{\int_0^x t^\rho dF(t)} = 0$$

6. *Si* X est positive et $\bar{F} \in \text{RV}_{-\rho}$ ($\rho > 0$), *alors* :

$$\forall p \in \mathbb{R}, \quad \mathbb{E}(X^p) = \begin{cases} \mu < \infty & \text{si } p < \rho \\ +\infty & \text{si } p > \rho \end{cases}$$

B.3 Fonction à variation rapide

Dans la théorie des probabilités et ses applications, la notion de variation régulière à l'infini d'indice $\rho \in \mathbb{R}$ est plus utilisée, car elle caractérise les lois de probabilité à queue lourde. Il y a une autre notion de variation des fonctions qui caractérise les lois de probabilité à queue légère (fine), baptisée variation rapide à l'infini, définie comme suit :

Définition B.3.1 [BINGHAM, GOLDIE & TEUGELS (1987)].

Une fonction mesurable $g :]a, +\infty[\rightarrow \mathbb{R}_+$ ($a \geq 0$) est dite **à variation rapide à l'infini d'indice $-\infty$** , et on note $g \in \text{RV}_{-\infty}$, si :

$$\forall x > 0, \quad \lim_{t \rightarrow +\infty} \frac{g(tx)}{g(t)} = x^{-\infty} = \begin{cases} +\infty & \text{si } x < 1 \\ 1 & \text{si } x = 1 \\ 0 & \text{si } x > 1 \end{cases} \quad (\text{B.8})$$

Remarque : Une fonction g est à variation rapide à l'infini d'indice $-\infty$, si et seulement si elle est définie mesurable, positive sur un voisinage de l'infini et vérifie (B.8).

Comme on a défini la notion de variation rapide à l'infini d'indice $-\infty$, on peut, aussi, la définir d'indice $+\infty$.

Définition B.3.2 [BINGHAM, GOLDIE & TEUGELS (1987)].

Une fonction mesurable $g :]a, +\infty[\rightarrow \mathbb{R}_+$ ($a \geq 0$) est dite **à variation rapide à l'infini d'indice $+\infty$** , et on note $g \in \text{RV}_{+\infty}$, si :

$$\forall x > 0, \quad \lim_{t \rightarrow +\infty} \frac{g(tx)}{g(t)} = x^{+\infty} = \begin{cases} 0 & \text{si } x < 1 \\ 1 & \text{si } x = 1 \\ +\infty & \text{si } x > 1 \end{cases}$$

B.3.1 Théorème de représentation

Dans le théorème qui suit, on va donner une représentation des fonctions qui varient rapidement à l'infini d'indice $-\infty$:

Théorème B.3.1 [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

Une fonction g est à variation rapide à l'infini d'indice $-\infty$ **si et seulement si** elle peut être représentée sous la forme suivante :

$$\forall x \geq b > a, \quad g(x) = c(x) \exp \left\{ \int_b^x \frac{\delta(u)}{u} du \right\} \quad (\text{B.9})$$

où : c et δ sont des fonctions définies mesurables sur $]a, +\infty[$ ($a \geq 0$) telles que :

$$\lim_{x \rightarrow +\infty} c(x) = c_0 \in]0, +\infty[\quad \text{et} \quad \lim_{x \rightarrow +\infty} \delta(x) = -\infty$$

B.3.2 Exemples

Exemple B.3.1

Soit $g :]0, +\infty[\rightarrow \mathbb{R}_+$ une fonction définie par :

$$\forall x \in]0, +\infty[, \quad g(x) = e^{-x}$$

Alors $g \in \text{RV}_{-\infty}$.

Vérification :

1) En utilisant (B.8) :

$$\forall x > 0, \quad \lim_{t \rightarrow +\infty} \frac{g(tx)}{g(t)} = \lim_{t \rightarrow +\infty} \frac{e^{-tx}}{e^{-t}} = \lim_{t \rightarrow +\infty} e^{-t(x-1)} = \begin{cases} +\infty & \text{si } x < 1 \\ 1 & \text{si } x = 1 \\ 0 & \text{si } x > 1 \end{cases}$$

2) En utilisant (B.9) :

$$\forall x \geq 1 > 0, \quad g(x) = e^{-1} \exp \left\{ \int_1^x \frac{(-u)}{u} du \right\}$$

Exemple B.3.2

Soit X une variable aléatoire à valeurs dans \mathbb{R} qui suit une loi de probabilité normale standard $\mathcal{N}(0, 1)$ de fonction de répartition Φ et de fonction de densité φ i.e :

$$\forall x \in \mathbb{R}, \quad \Phi(x) = \int_{-\infty}^x \varphi(u) du = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-u^2/2} du$$

Alors :

$$\text{P}(X > x) = 1 - \Phi(x) \sim \frac{\varphi(x)}{x} = \frac{1}{\sqrt{2\pi}} x^{-1} e^{-x^2/2} \quad \text{quand } x \rightarrow +\infty$$

Or :

$$\forall y > 0, \quad \lim_{t \rightarrow +\infty} \frac{(ty)^{-1} \varphi(ty)}{t^{-1} \varphi(t)} = \lim_{t \rightarrow +\infty} y^{-1} \exp \left\{ -\frac{t^2}{2} (y^2 - 1) \right\} = \begin{cases} +\infty & \text{si } y < 1 \\ 1 & \text{si } y = 1 \\ 0 & \text{si } y > 1 \end{cases}$$

Donc : $(1 - \Phi) \in \text{RV}_{-\infty}$ et la loi de probabilité normale standard $\mathcal{N}(0, 1)$ est à queue légère.

Exemple B.3.3

Soit X une variable aléatoire à valeurs dans \mathbb{R} qui suit une loi de probabilité exponentielle standard $\mathcal{E}(1)$ de fonction de répartition F , i.e :

$$\forall x \in \mathbb{R}, \quad \text{P}(X \leq x) = F(x) = 1 - e^{-x} \cdot \mathbb{1}_{[0, +\infty[}(x)$$

Alors :

$$\forall x \geq 0, \quad \text{P}(X > x) = 1 - F(x) = e^{-x}$$

Donc : $(1 - F) \in \text{RV}_{-\infty}$ et la loi de probabilité exponentielle standard $\mathcal{E}(1)$ est à queue légère.

B.3.3 Propriétés des fonctions aux variations rapides

Dans la proposition qui suit, on va donner quelques propriétés élémentaires des fonctions qui varient rapidement à l'infini d'indice $-\infty$:

Proposition B.3.1 [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997) et RESNICK (2006)].
Soit g une fonction définie mesurable et positive sur le voisinage $]a, +\infty[$ ($a \geq 0$) de l'infini.

- Si g est décroissante et $g \in \text{RV}_{-\infty}$, alors :

$$\exists b > a \quad \text{tel que} \quad \forall \alpha \in \mathbb{R}, \quad \int_b^{+\infty} t^\alpha g(t) dt < \infty \quad \text{et} \quad \lim_{x \rightarrow +\infty} \frac{x^{\alpha+1} g(x)}{\int_x^{+\infty} t^\alpha g(t) dt} = +\infty, \quad x \geq b$$

- Si :

$$\exists \alpha \in \mathbb{R} \quad \text{tel que} \quad \int_b^{+\infty} t^\alpha g(t) dt < \infty \quad (b > a) \quad \text{et} \quad \lim_{x \rightarrow +\infty} \frac{x^{\alpha+1} g(x)}{\int_x^{+\infty} t^\alpha g(t) dt} = +\infty \quad (x \geq b)$$

alors :

$$g \in \text{RV}_{-\infty}$$

- Si $g \in \text{RV}_{-\infty}$, alors :

$$\lim_{x \rightarrow +\infty} \frac{\ln g(x)}{\ln x} = -\infty \quad \text{et} \quad \lim_{x \rightarrow +\infty} g(x) = 0$$

- Si g est dérivable sur $]a, +\infty[$ de dérivée g' , alors :

$$\lim_{x \rightarrow +\infty} \frac{xg'(x)}{g(x)} = -\infty \implies g \in \text{RV}_{-\infty}$$

Référence

Pour voir les résultats de la théorie de variation régulière avec leurs démonstrations bien détaillées et leurs applications dans la théorie des probabilités et ses applications, le lecteur pourra se référer aux ouvrages suivants : [BINGHAM, GOLDIE & TEUGELS (1987)], [FELLER (1971)], [DE HAAN & FERREIRA (2006)], [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)], ou, encore, [RESNICK (2006)].

ANNEXE C

PROCESSUS PONCTUELS DE POISSON

Introduction

Les techniques de processus ponctuels sont largement utilisées dans la théorie des valeurs extrêmes parce qu'elles fournissent des outils qui peuvent être utilisés pour prouver plusieurs résultats asymptotiques avec élégance et perspicacité.

Pour donner une idée générale, soit à construire des modèles pour la distribution aléatoire des points dans un espace qu'on prend souvent comme une partie de \mathbb{R}_+ , de \mathbb{R} ou de \mathbb{R}^d ($d \geq 1$). Par exemple, les emplacements des arbres dans une forêt, les localisations et les instants où se produit un ouragan ou une réplique à la suite d'un tremblement de terre, les instants d'arrivée des clients dans une file d'attente, les positions des réservoirs dans un champs de bataille, les instants de panne d'une machine ou d'un groupe de machines dans une usine, etc. Alors, les processus ponctuels sont des candidats idéaux pour ce genre de modélisation.

Pour être un peu plus précis, soient E un espace et $x = \{x_i, i \in I \subset \mathbb{N}^*\}$ un ensemble de points de E considérés comme des réalisations de certaines variables aléatoires aux valeurs dans E . Un processus ponctuel N est une distribution aléatoire des points de x dans l'espace E qui à chaque ensemble $A \subseteq E$, $N(A)$ est le nombre de points de x dans A .

Les plus importants des processus ponctuels dans la théorie des valeurs extrêmes sont ceux pour lesquels le nombre de points de x qui appartiennent à l'ensemble A suit une loi de probabilité de POISSON. Ces processus qu'on appelle « **Processus Ponctuels de POISSON** », qu'on va voir plus loin, sont une généralisation des processus de POISSON classiques (homogènes) où $E = \mathbb{R}_+$ est considéré comme l'espace de temps, x est l'ensemble des instants d'occurrence d'un certain événement aléatoire, A est un intervalle de temps et $N(A)$ est le nombre de réalisations de l'événement pendant la période du temps $|A|$ ($|\cdot|$ est la mesure de LEBESGUE sur \mathbb{R}_+).

Remarque : La processus ponctuels ne sont pas importants seulement dans la théorie des valeurs extrêmes mais dans toute la théorie des probabilités et ses applications. L'utilisation de ces processus peuvent être observés dans des domaines aussi variés que la démographie, la biologie, l'épidémiologie,

l'astronomie, la mécanique, la géophysique, la physique des particules, l'ingénierie au sens large, etc. Pour exemple, voir [KERESZTES (2009)].

Afin de définir rigoureusement les processus ponctuels, soit, dans toute la suite de ce chapitre, (E, \mathcal{E}) un espace mesurable où E est un sous espace compact d'un espace **euclidien**¹ de dimension finie et \mathcal{E} sa tribu **borélienne**².

C.1 Généralités

C.1.1 Mesure ponctuelle

Pour certains concepts de cette section, le lecteur est conseillé de revoir certains résultats qui concernent la théorie de la mesure et qu'il peut trouver dans n'importe quel ouvrage standard de probabilité.

Définition C.1.1 (Mesure de DIRAC).

Soit x un point de l'espace E . On appelle **mesure de DIRAC** ou **masse de DIRAC** au point x , la mesure δ_x définie sur l'espace mesurable (E, \mathcal{E}) dans $\{0, 1\}$ par :

$$\forall A \in \mathcal{E}, \quad \delta_x(A) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{si } x \notin A \end{cases}$$

Remarque : La mesure de DIRAC est une mesure très importante dans la théorie des processus ponctuels qu'on rencontre souvent dans l'écriture symbolique de ces processus. C'est pour cette raison qu'on désigne, dans toute la suite de ce chapitre, par δ_x la mesure de DIRAC au point x dont la définition est celle donnée ci-dessus.

Définition C.1.2 (Mesure ponctuelle) [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

Soit $\{x_i, i \in I \subset \mathbb{N}^*\}$ une suite de points (non nécessairement distincts) de l'espace E . On appelle **mesure de dénombrement** ou **mesure de comptage** de la suite $\{x_i, i \in I \subset \mathbb{N}^*\}$, la mesure m définie sur l'espace mesurable (E, \mathcal{E}) dans $\mathbb{N} \cup \{+\infty\}$ par :

$$\forall A \in \mathcal{E}, \quad m(A) = \sum_{i \in I} \delta_{x_i}(A) = \text{card}\{i \in I \text{ tel que } x_i \in A\}$$

et on écrit :

$$m = \sum_{i \in I} \delta_{x_i}$$

On dit que la mesure de comptage m est **une mesure ponctuelle** si elle est finie sur tout **compact** C de \mathcal{E} , i.e :

$$\forall C \in \mathcal{E}, \quad m(C) < \infty$$

1. Un espace euclidien est un espace vectoriel muni de la norme euclidienne $\|\cdot\|_2$.

2. La tribu borélienne de E est la tribu engendrée par la famille des ouverts de E , qu'on note souvent par \mathcal{B}_E sauf mention du contraire, comme dans notre cas.

Remarque : Dans le cadre de la définition ci-dessus, la mesure ponctuelle m est une mesure qui compte le nombre d'éléments de la suite $\{x_i, i \in I \subset \mathbb{N}^*\}$ qui appartiennent au **borélien**³ A .

Dans toute la suite de ce chapitre, soient $M_p(E)$ l'espace de toutes les mesures ponctuelles définies sur l'espace mesurable (E, \mathcal{E}) muni de sa tribu $\mathcal{M}_p(E)$, où $\mathcal{M}_p(E)$ est la plus petite tribu (au sens de l'inclusion) de $M_p(E)$ dont les éléments sont des ensembles de la forme :

$$\{m \in M_p(E) \text{ telle que } m(A) \in B\} \quad \text{avec} \quad A \in \mathcal{E} \quad \text{et} \quad B \in \mathcal{B}_{[0,+\infty]}$$

Remarque : D'une autre manière, la tribu $\mathcal{M}_p(E)$ est la plus petite tribu (au sens de l'inclusion) de $M_p(E)$ qui, pour tout borélien A de \mathcal{E} , rend mesurable l'application :

$$\begin{array}{ccc} [M_p(E), \mathcal{M}_p(E)] & \longrightarrow & ([0, +\infty], \mathcal{B}_{[0,+\infty]}) \\ m & \longmapsto & m(A) \end{array}$$

C.1.2 Processus ponctuel

Comme on les a décrit dans l'introduction de ce chapitre, les processus ponctuels servent à modéliser la distribution aléatoire des points dans un espace. Pour cela, soit $\{x_i, i \in I \subset \mathbb{N}^*\}$ une suite de points aléatoires de E , alors un processus ponctuel N sur E est une fonction qui, à chaque borélien A de \mathcal{E} , compte le nombre de points aléatoires de la suite $\{x_i, i \in I \subset \mathbb{N}^*\}$ qui tombent dans A , c-à-d :

$$\forall A \in \mathcal{E}, \quad N(A) = \sum_{i \in I} \delta_{x_n}(A)$$

Pour donner une définition générale des processus ponctuels qui permette d'utiliser les résultat de la théorie des probabilités, soient, dans toute la suite de ce chapitre, $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et $\{X_i, i \in I \subset \mathbb{N}^*\}$ une suite de variables aléatoires définies sur $(\Omega, \mathcal{A}, \mathbb{P})$ dans (E, \mathcal{E}) .

Définition C.1.3 [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

Un processus ponctuel à **espace d'état** E est une application mesurable N définie sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ dans l'espace mesurable $[M_p(E), \mathcal{M}_p(E)]$ des mesures ponctuelles définies sur (E, \mathcal{E}) :

$$N : (\Omega, \mathcal{A}, \mathbb{P}) \longrightarrow [M_p(E), \mathcal{M}_p(E)]$$

Remarque : D'après la définition ci-dessus, un processus ponctuel N n'est qu'une variable aléatoire qui prend ses valeurs dans l'espace des mesures ponctuelles définies sur (E, \mathcal{E}) . Donc la loi de probabilité de N , qu'on note \mathbb{P}_N , est la mesure image $\mathbb{P} \circ N^{-1}$ définie sur l'espace mesurable $[M_p(E), \mathcal{M}_p(E)]$ dans $[0, 1]$ par :

$$\forall D \in \mathcal{M}_p(E), \quad \mathbb{P}_N(D) = \mathbb{P} \circ N^{-1}(D) = \mathbb{P}[N^{-1}(D)] = \mathbb{P}(\{\omega \in \Omega, N(\omega) \in D\})$$

3. Un borélien est un élément de la tribu borélienne.

En fait, il est souvent difficile de vérifier qu'une application N , définie sur $(\Omega, \mathcal{A}, \mathbb{P})$ dans $[M_p(E), \mathcal{M}_p(E)]$, est un processus ponctuel juste de sa définition à cause de l'abstraction de celle-ci. Mais, en utilisant la définition de $\mathcal{M}_p(E)$, on a la proposition suivante qui rend la tâche un peu plus facile :

Proposition C.1.1 [RESNICK (1987)].

Une application N définie sur $(\Omega, \mathcal{A}, \mathbb{P})$ dans $[M_p(E), \mathcal{M}_p(E)]$ est un processus ponctuel **si et seulement si** l'application N_A définie sur $(\Omega, \mathcal{A}, \mathbb{P})$ dans $([0, +\infty], \mathcal{B}_{[0, +\infty]})$ par :

$$\forall \omega \in \Omega, \quad N_A(\omega) = N(\omega, A)$$

est mesurable pour tout borélien A de \mathcal{E} .

D'après la définition précédente, un processus ponctuel N à espace d'état E est une variable aléatoire qui prend des mesures ponctuelles définies sur (E, \mathcal{E}) comme valeurs. Alors, il est plus pratique de représenter N comme étant une collection de variables aléatoires $[N(A)]_{A \in \mathcal{E}}$. Dans ce cas, en utilisant la définition d'une mesure ponctuelle sur (E, \mathcal{E}) , on peut poser la définition suivante qui est plus claire que celles ci-dessus :

Définition C.1.4 [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

Soit $\{X_i, i \in I \subset \mathbb{N}^*\}$ une suite de variables aléatoires définies sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ aux valeurs dans (E, \mathcal{E}) . Un processus ponctuel à espace d'état E est une application mesurable N définie sur $\Omega \times \mathcal{E}$ dans $\mathbb{N} \cup \{+\infty\}$ par :

$$\forall \omega \in \Omega, \quad \forall A \in \mathcal{E}, \quad N(\omega, A) = \sum_{i \in I} \delta_{X_i(\omega)}(A)$$

telle que :

$$\forall C \in \mathcal{E}, \quad N(\omega, C) < \infty \quad \text{avec } C \text{ compact}$$

et on écrit :

$$N = \sum_{i \in I} \delta_{X_i}$$

Remarque : Dans le cadre de la définition ci-dessus, on appelle les variables aléatoires de la suite $\{X_i, i \in I \subset \mathbb{N}^*\}$ **les réalisations** de N , et on peut constater que :

1. Pour $\omega \in \Omega$ fixé : N est une mesure ponctuelle définie sur (E, \mathcal{E}) .
2. Pour $A \in \mathcal{E}$ fixé : N est une variable aléatoire qui prend ses valeurs dans $\mathbb{N} \cup \{+\infty\}$.

Notation : Dans le cadre de la définition précédente et par abus de notation, on supprime souvent la dépendance du processus ponctuel N de ω , et on écrit :

$$N(A) = \sum_{i \in I} \delta_{X_i}(A), \quad A \in \mathcal{E}$$

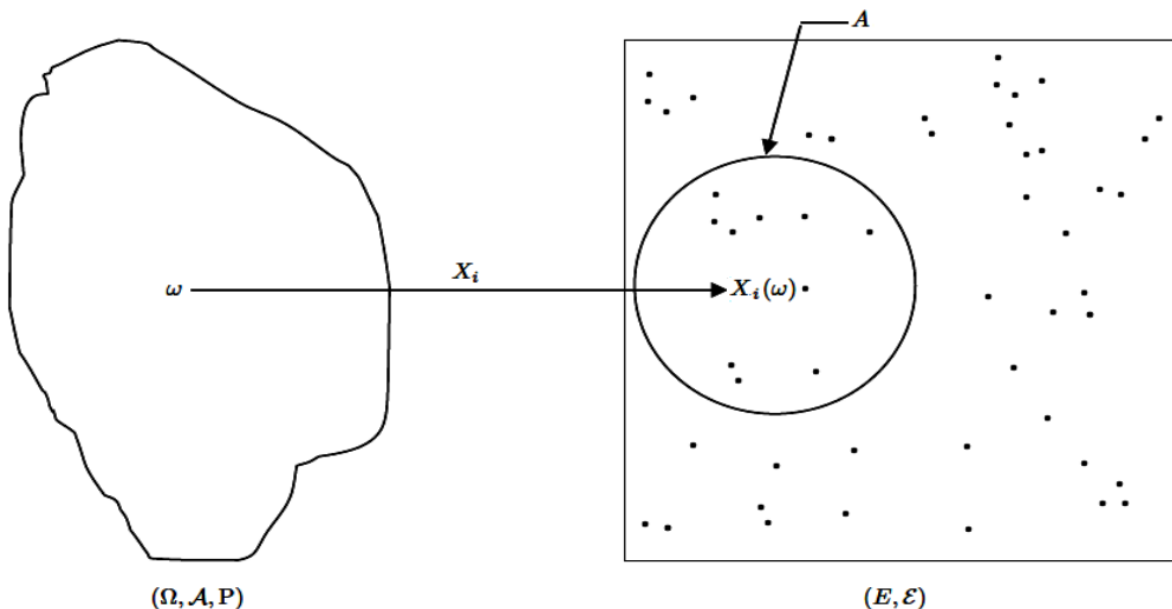


Figure C.1.1 Ensemble de points aléatoires $\{X_i(\omega), i \in I \subset \mathbb{N}^*\}$ dans (E, \mathcal{E}) pour $\omega \in \Omega$. Le nombre de ces points qui appartiennent au borélien A de \mathcal{E} est $N(A)$, dans ce cas $N(A) = 10$.

Exemple C.1.1 [RESNICK (1992)].

On veut modéliser les instants et les localisations où se produisent des séismes. Un espace d'état convenable pour cette modélisation est $E = [0, +\infty[\times \mathbb{R}^2$, où $[0, +\infty[$ représente l'espace temps et \mathbb{R}^2 le globe terrestre, et le processus ponctuel N associé est :

$$N = \sum_i \delta_{X_i}, \quad X_i = [T_i, (L_{i1}, L_{i2})]$$

où T_i représente l'instant du $i^{\text{ème}}$ séisme et (L_{i1}, L_{i2}) la latitude et la longitude de la localisation où s'est produit ce séisme.

Si $t > 0$ et $F \subset \mathbb{R}^2$, alors :

$$N([0, t] \times F) = \sum_i \delta_{X_i}([0, t] \times F)$$

est le nombre de séismes qui se produisent dans la région F durant l'intervalle de temps $[0, t]$.

On peut ajouter des paramètres qui rendent le modèle plus précis. Comme exemple, on veut ajouter l'intensité des séismes au modèle, alors l'espace d'état devient $E = [0, +\infty[\times \mathbb{R}^2 \times [0, +\infty[$ et le processus ponctuel N associé sera :

$$N = \sum_i \delta_{X_i}, \quad X_i = [T_i, (L_{i1}, L_{i2}), I_i]$$

où I_i représente la magnitude du $i^{\text{ème}}$ séisme.

Exemple C.1.2 [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

L'un des processus ponctuels importants dans la théorie des valeurs extrêmes est le **processus ponctuel des excès** qui compte le nombre des excès d'une suite finie de variables aléatoires au-delà d'un certain seuil (voir CHAPITRE 2, Section 2.5).

Soient $\{Y_n, n \in \mathbb{N}^*\}$ une suite de variables aléatoires réelles **iid** et $u \in \mathbb{R}$. Le processus ponctuel des excès de l'échantillon de variables aléatoires $\{Y_i, i = 1, 2, \dots, n\}$ au-delà du seuil u est un processus ponctuel, qu'on note N_n , à espace d'état $E =]0, 1]$ qui compte le nombre de points aléatoires Y_i ($i \in I \subset \{1, 2, \dots, n\}$) qui dépassent le seuil u :

$$N_n(A) = \sum_{i=1}^n \delta_{i/n}(A) \cdot \mathbb{1}_{]u, +\infty[}(Y_i), \quad A \subset]0, 1] \quad \text{et} \quad n \in \mathbb{N}^*$$

avec :

$$\mathbb{1}_{]u, +\infty[}(Y_i) = \begin{cases} 1 & \text{si } Y_i > u \\ 0 & \text{si } Y_i \leq u \end{cases}$$

Si $A = E =]0, 1]$, alors $N_n(]0, 1])$ est le nombre de variables aléatoires de tout l'échantillon $\{Y_n, n \in \mathbb{N}^*\}$ qui dépassent u , car :

$$N_n(]0, 1]) = \text{card}\{i, 0 < i/n \leq 1 \text{ et } Y_i > u\} = \text{card}\{i \leq n, Y_i > u\}$$

C.1.3 Loi de probabilité d'un processus ponctuel

Comme on l'a déjà mentionné, juste après la définition d'un processus ponctuel, la loi de probabilité ou la distribution d'un processus ponctuel N à espace d'état E , qu'on note P_N , est la probabilité image $P \circ N^{-1}$ définie sur l'espace mesurable $[M_p(E), \mathcal{M}_p(E)]$ dans $[0, 1]$ par :

$$\forall D \in \mathcal{M}_p(E), \quad P_N(D) = P \circ N^{-1}(D) = P[N^{-1}(D)] = P(\{\omega \in \Omega, N(\omega) \in D\})$$

Malheureusement, cette loi de probabilité n'est pas facile à imaginer. Cependant, on sait que pour tout borélien A fixé dans \mathcal{E} , $N(A)$ est une variable aléatoire définie sur l'espace probabilisé (Ω, \mathcal{A}, P) à valeurs dans $\mathbb{N} \cup \{+\infty\}$. Donc, on peut déterminer la loi de probabilité de N si on connaît la famille de toute les lois de probabilité des vecteurs aléatoires de la forme :

$$(N(A_1), N(A_2), \dots, N(A_n)), \quad A_i \in \mathcal{E} \quad \text{avec} \quad i \in \{1, 2, \dots, n\} \quad \text{et} \quad n \in \mathbb{N}^*$$

Définition C.1.5 [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

Soit N un processus ponctuel à espace d'état E . On appelle **distribution finie-dimensionnelle** de N la famille de toute les lois de probabilité des vecteurs aléatoires de la forme :

$$(N(A_1), N(A_2), \dots, N(A_n)), \quad A_i \in \mathcal{E} \quad \text{avec} \quad i \in \{1, 2, \dots, n\} \quad \text{et} \quad n \in \mathbb{N}^*$$

données par :

$$P(N(A_1) = k_1, N(A_2) = k_2, \dots, N(A_n) = k_n), \quad k_i \in \mathbb{N} \cup \{+\infty\}$$

Proposition C.1.2 [RESNICK (1987)].

Soit N un processus ponctuel à espace d'état E . La loi de probabilité P_N de N est complètement déterminée par la donnée de sa distribution finie-dimensionnelle.

Corollaire C.1.1 [SERFOZO (2009)].

*Deux processus ponctuels N et N' définis sur l'espace d'état E ont la même loi de probabilité **si et seulement si** leurs distributions finie-dimensionnelles sont identiques.*

Remarque : Dans le cadre du corollaire ci-dessus, les processus ponctuels N et N' peuvent ne pas être définis sur le même espace probabilisé (Ω, \mathcal{A}, P) .

C.1.4 Mesure d'intensité d'un processus ponctuel

Dans la définition qui suit, on va définir un paramètre appelé « **mesure d'intensité** » qui caractérise tout processus ponctuel :

Définition C.1.6 [RESNICK (1987)].

*Soit N un processus ponctuel à espace d'état E . On appelle **mesure d'intensité** du processus N la mesure μ définie sur l'espace mesurable (E, \mathcal{E}) dans $[0, +\infty]$ par :*

$$\forall A \in \mathcal{E}, \quad \mu(A) = \mathbf{E}[N(A)]$$

Remarque : D'après la définition ci-dessus, la mesure d'intensité μ d'un processus ponctuel N à espace d'état E n'est que l'espérance, pour tout borélien A de \mathcal{E} , de la variable aléatoire $N(A)$. Donc, elle nous donne le nombre moyen de points aléatoires qui se trouvent dans A . C'est pour cette raison qu'on l'appelle, parfois, « **mesure moyenne** » du processus N .

Note : Dans le cadre de la définition ci-dessus, le nombre $\mu(A)$ peut être infini, même si A est un borélien compact de \mathcal{E} .

C.2 Processus ponctuel de POISSON

Les processus ponctuels de POISSON est une classe très importante des processus ponctuels. Ils jouent un rôle similaire pour les processus ponctuels à celui que jouent les lois de probabilité normales pour les lois de probabilité, car ils sont les limites naturelles des sommes de beaucoup de processus ponctuels. C'est pour cette raison que de tels processus sont largement utilisés dans la théorie des probabilités et les domaines de son application.

Comme illustration, supposons que l'espace E est une forêt qui contienne $n \in \mathbb{N}$ arbres et $\{x_i, i = 1, 2, \dots, n\}$ les positions aléatoires de ces arbres dans cette forêt, et soit A un site dans E . Soient p_A la probabilité qu'un arbre de position x_i soit dans A et $N(A)$ le nombre total des arbres qui sont dans A , c'est-à-dire :

$$P(x_i \in A) = p_A \quad \text{et} \quad N(A) = \sum_{i=1}^n \delta_{x_i}(A)$$

Alors le nombre $N(A)$ est une variable aléatoire qui suit une loi de probabilité binomiale de paramètre (n, p_A) , c'est-à-dire :

$$\forall k \in \{1, 2, \dots, n\}, \quad \mathbb{P}[N(A) = k] = C_n^k p_A^k (1 - p_A)^{n-k} \quad \text{avec} \quad C_n^k = \frac{n!}{k!(n-k)!}$$

Et, d'après l'approximation de la loi de probabilité binomiale par une loi de probabilité de POISSON, on a :

$$\lim_{n \rightarrow \infty} \mathbb{P}[N(A) = k] = \lim_{n \rightarrow \infty} C_n^k p_A^k (1 - p_A)^{n-k} = \frac{[\mu(A)]^k}{k!} e^{-\mu(A)} \quad \text{avec} \quad \mu(A) = \lim_{n \rightarrow \infty} n p_A$$

Donc, pour n grand (la forêt E est grande), le nombre d'arbres qui se trouvent dans le site A suit, approximativement, une loi de probabilité de POISSON de paramètre $\mu(A)$ où $\mu(A)$ est le nombre moyen de ces arbres qui se trouvent dans A .

Définition C.2.1 [RESNICK (1987)].

Soit μ une mesure définie sur l'espace mesurable (E, \mathcal{E}) dans $[0, +\infty]$. On dit que μ est une mesure de **RADON** si elle est finie sur tout compact de \mathcal{E} , c'est-à-dire :

$$\forall C \in \mathcal{E}, \quad \mu(C) < \infty \quad \text{avec} \quad C \text{ compact}$$

Dans toute la suite de ce chapitre, on désigne par μ une mesure de RADON définie sur l'espace mesurable (E, \mathcal{E}) dans $[0, +\infty]$.

Définition C.2.2 [RESNICK (1987)].

Soit N un processus ponctuel à espace d'état E . On dit que N est un **processus ponctuel de POISSON** de mesure d'intensité μ , et on le note $\text{PPP}(\mu)$, s'il vérifie les assertions suivantes :

1.

$$\forall k \in \mathbb{N} \cup \{+\infty\}, \quad \mathbb{P}[N(A) = k] = \begin{cases} \frac{[\mu(A)]^k}{k!} e^{-\mu(A)} & \text{si } \mu(A) < \infty \\ 0 & \text{si } \mu(A) = \infty \end{cases}, \quad A \in \mathcal{E}$$

avec :

$$\mu(A) = \infty \implies \mathbb{P}[N(A) = \infty] = 1$$

2. Quel que soit $n \in \mathbb{N}$, si A_1, A_2, \dots, A_n sont des boréliens de \mathcal{E} deux à deux disjoints, alors $N(A_1), N(A_2), \dots, N(A_n)$ sont des variables aléatoires indépendantes.

Remarque : On peut constater, de la définition ci-dessus, que pour tout borélien A de \mathcal{E} , le nombre $N(A)$ est une variable aléatoire qui suit une loi de probabilité de POISSON de moyenne $\mu(A)$. Or, une loi de probabilité de POISSON est totalement déterminée par sa moyenne. Donc, d'après les deux conditions de la définition ci-dessus, la distribution finie-dimensionnelle ou la loi de probabilité d'un processus ponctuel de POISSON est complètement déterminée par sa mesure d'intensité.

Proposition C.2.1 [SERFOZO (2009)].

Deux processus ponctuels de POISSON définis sur l'espace d'état E ont la même loi de probabilité si et seulement s'ils ont la même mesure d'intensité.

C.3 PPP à espace d'état \mathbb{R}^d

Dans les applications, les PPP les plus utilisés sont ceux qui ont comme espace d'état un sous espace E de $\overline{\mathbb{R}}^d$ ($d \geq 1$ et $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$). Dans toute la suite de ce chapitre, ça sera le cas.

Définition C.3.1 [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

Soit N un PPP(μ) à espace d'état $E \subset \overline{\mathbb{R}}^d$. On dit que N est un processus de POISSON **non homogène** de **fonction d'intensité** f , et on le note par NHPP(f), si μ est absolument continue par rapport à la mesure de LEBESGUE de densité f , c-à-d :

$$\forall A \in \mathcal{E}, \quad \mu(A) = \int_A f(x) dx, \quad f : E \longrightarrow \mathbb{R}_+$$

Définition C.3.2 [EMBRECHTS, KLÜPPELBERG & MIKOSCH (1997)].

Soit N un PPP(μ) à espace d'état $E \subset \overline{\mathbb{R}}^d$. On dit que N est un processus de POISSON **homogène** d'**intensité** $\lambda \in \mathbb{R}_+$, et on le note par HPP(λ), si μ est absolument continue par rapport à la mesure de LEBESGUE de densité constante λ , c-à-d :

$$\forall A \in \mathcal{E}, \quad \mu(A) = \int_A \lambda dx = \lambda |A|, \quad \lambda \in \mathbb{R}_+ \quad \text{et} \quad |\cdot| \quad \text{est la mesure de LEBESGUE}$$

Remarque : Si N est un NHPP(f)⁴, alors on appelle, parfois, f **fonction de taux** de N . Et si N est un HPP(λ)⁵, on appelle λ **taux** de N .

En particulier, les NHPP et les HPP les plus connus et qui sont largement utilisés dans les applications de la théorie des probabilités sont ceux qui ont l'espace d'état \mathbb{R}_+ . Ces processus modélisent, généralement, la distribution aléatoire des points dans le temps. Par exemple, les instants des émissions radioactives, les instants d'arrivée des clients dans une file d'attente, les instants de panne d'une machine ou d'un groupe de machines dans une usine, etc.

Soit $\{T_n, n \in \mathbb{N}^*\}$ une suite de variables aléatoires définies sur le même espace probabilisé (Ω, \mathcal{A}, P) dans $(\mathbb{R}_+, \mathcal{B}_{\mathbb{R}_+})$ qui représente les instants (points) d'occurrence d'un certain événement, c'est-à-dire T_n est l'instant de la $n^{\text{ème}}$ réalisation de l'événement. Et soit N le processus ponctuel à espace d'état \mathbb{R}_+ qui compte le nombre de réalisations de l'événement durant un intervalle de temps donné.

L'avantage que nous présente l'espace d'état \mathbb{R}_+ est qu'on peut représenter N de deux manières, un peu, différentes :

La première, tirée de la définition même d'un processus ponctuel, considère N comme étant une collection de variables aléatoires $\{N(A), A \in \mathcal{B}_{\mathbb{R}_+}\}$, où $N(A)$ est le nombre de réalisations de l'événements dans l'intervalle de temps A . Formellement :

4. Non Homogeneous POISSON Process with intensity function f .

5. Homogeneous POISSON Process with intensity λ .

$$N(A) = \sum_{n=1}^{\infty} \delta_{T_n}(A), \quad A \in \mathcal{B}_{\mathbb{R}_+}$$

Tandis que la seconde, qui est plus pratique, nous représente N comme étant une autre suite de variables aléatoires $\{N(t), t \in \mathbb{R}_+\}$, où $N(t)$ est le nombre de réalisations de l'événements dans l'intervalle de temps $[0, t]$. Formellement :

$$N(t) = \sum_{n=1}^{\infty} \delta_{T_n}([0, t]), \quad t \in \mathbb{R}_+$$

On peut voir, facilement, que si $A =]s, t] \subset \mathbb{R}_+$ est un intervalle de temps, alors le nombre de réalisations de l'événement durant cet intervalle est donné par :

$$N(A) = N(]s, t]) = N(t) - N(s)$$

Dans la suite de ce chapitre, on va donner la définition **classique**⁶ d'un HPP(λ) ($\lambda > 0$) à espace d'état \mathbb{R}_+ . Ce processus est l'un des modèles de processus de POISSON les plus connus et les plus utilisés dans les applications. Il joue un rôle très important dans la théorie des files d'attente qui étudie les solutions optimales de gestion des files d'attente. Ensuite, on annoncera les résultats qui nous permettent d'extraire les caractéristiques de ce processus et de voir que c'est un cas particulier des PPP.

Définition C.3.3

Soit $N = \{N(t), t \in \mathbb{R}_+\}$ un processus ponctuel à espace d'état \mathbb{R}_+ . On dit que N est un processus de POISSON **homogène** de **taux** $\lambda > 0$, et le note par HPP(λ), s'il vérifie les assertions suivantes :

1. $N(0) = 0$.
2. N est un processus à **accroissements stationnaires**, c'est-à-dire :

$$\forall t \in \mathbb{R}_+, \quad \forall h > 0, \quad N(t+h) - N(t) \stackrel{d}{=} N(h)$$

3. N est un processus à **accroissements indépendants**, c'est-à-dire, pour $n \in \mathbb{N}^*$, on a :

$$\forall 0 \leq t_1 < t_2 < \dots < t_{n-1} < t_n, \quad N(t_2 - t_1) \perp \dots \perp N(t_n - t_{n-1})$$

4. $N(t)$ suit une loi de probabilité de POISSON de paramètre λt , c'est-à-dire :

$$\forall k \in \mathbb{N}, \quad \mathbb{P}[N(t) = k] = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad t \in \mathbb{R}_+$$

Remarque : Le signe \perp qu'on voit dans la troisième assertion de la définition ci-dessus veut dire l'indépendance. C'est-à-dire, si on écrit $Y \perp Z$ (Y et Z sont des variables aléatoires), alors cela voudra dire que Y et Z sont indépendantes.

6. La plus connue et la plus rencontrée dans la littérature

Interprétation : On peut interpréter les assertions de la définition ci-dessus que doit vérifier un HPP(λ) ($\lambda > 0$) comme suit :

1. À l'instant $t = 0$, il est certain qu'on observera aucune réalisation de l'événement.
2. Un HPP(λ) est un processus **homogène** : Le nombre de réalisations de l'événement durant un intervalle de temps $]t, t + h]$ ne dépend que de la longueur h de cet intervalle, et non de t .
3. Un HPP(λ) est un processus **sans mémoire** : Les réalisations de l'événement avant une date t n'influencent en rien sur les réalisations après cette date.
4. En moyenne, le nombre de réalisations de l'événement par unité de temps est λ . C'est pour cette raison qu'on appelle λ **taux** du processus.

A. Mesure d'intensité d'un HPP(λ)

Dans la proposition qui suit, on va donner un résultat, déduit directement de la définition ci-dessus, qui nous permette de voir la forme de la mesure d'intensité d'un HPP(λ) :

Proposition C.3.1 [SERFOZO (2009)].

Soit $N = \{N(t), t \in \mathbb{R}_+\}$ un processus ponctuel à espace d'état \mathbb{R}_+ . N est un HPP(λ) ($\lambda > 0$) si et seulement s'il vérifie les assertions suivantes :

1.

$$\forall k \in \mathbb{N}, \quad \mathbb{P}[N(]s, t]) = k] = \frac{[\lambda(t-s)]^k}{k!} e^{-\lambda(t-s)}, \quad]s, t] \subset \mathbb{R}_+$$

2. $\forall n \in \mathbb{N}^*$, si $]s_1, t_1],]s_2, t_2], \dots,]s_n, t_n]$ sont des intervalles de \mathbb{R}_+ deux à deux disjoints, alors $N(]s_1, t_1]), N(]s_2, t_2]), \dots, N(]s_n, t_n])$ sont des variables aléatoires indépendantes.

Résultat : De la proposition ci-dessus, on déduit qu'un HPP(λ) à espace d'état \mathbb{R}_+ est un PPP(μ), où μ est une mesure définie sur $(\mathbb{R}_+, \mathcal{B}_{\mathbb{R}_+})$ dans $[0, +\infty]$ par :

$$\forall A \in \mathcal{B}_{\mathbb{R}_+}, \quad \mu(A) = \lambda|A|, \quad |\cdot| \text{ est la mesure de LEBESGUE sur } (\mathbb{R}_+, \mathcal{B}_{\mathbb{R}_+})$$

B. Représentation d'un HPP(λ)

Dans la proposition qui suit, on va annoncer un résultat qui nous donne la représentation d'un HPP(λ) et nous précise les lois de probabilité de ses réalisations :

Proposition C.3.2 [RESNICK (1992)].

Soit $N = \{N(t), t \in \mathbb{R}_+\}$ un processus ponctuel à espace d'état \mathbb{R}_+ . N est un HPP(λ) ($\lambda > 0$) si et seulement s'il a la représentation suivante :

$$N = \sum_{n=1}^{\infty} \delta_{T_n}, \quad T_n = Y_1 + Y_2 + \dots + Y_n$$

Où : $\{Y_n, n \in \mathbb{N}^*\}$ est une suite de variables aléatoires **indépendantes et identiquement distribuées** selon la loi de probabilité exponentielle de paramètre λ .

Conséquence : De la proposition ci-dessus, on déduit que les réalisations $\{T_n, n \in \mathbb{N}^*\}$ (les instants d'occurrences de l'événement) sont des variables aléatoires qui ont des lois de probabilité gammas telles que :

$$\forall n \in \mathbb{N}^*, \quad T_n \sim \Gamma(n, \lambda)$$

ANNEXE D

INTRODUCTION AU LOGICIEL R

D.1 Présentation du logiciel R

Le logiciel **R** est un logiciel de statistique créé par Ross IHAKA et Robert GENTLEMAN. Il est à la fois un langage informatique et un environnement de travail : les commandes sont exécutées grâce à des instructions codées dans un langage relativement simple, les résultats sont affichés sous forme de texte et les graphiques sont visualisés directement dans une fenêtre qui leur est propre. C'est un clone du logiciel **S-PLUS** qui est fondé sur le langage de programmation orienté objet S, développé par *AT&T Bell Laboratories* en 1988. Ce logiciel sert à manipuler des données, à tracer des graphiques et à faire des analyses statistiques sur ces données.

R est un logiciel **gratuit** et à **code source ouvert** (*open-source*). Il fonctionne sous UNIX (et LINUX), Windows et Macintosh. C'est donc un logiciel **multi-plates-formes**. Il est développé dans la mouvance des logiciels distribués librement, sous les termes de la *GNU General Public Licence*, par une communauté sans cesse plus vaste de bénévoles motivés rassemblés dans le *R Development Core Team*. Tout le monde peut d'ailleurs contribuer à son amélioration en y intégrant de nouvelles fonctionnalités ou méthodes d'analyse non encore implémentées. Cela en fait donc un logiciel en rapide et constante évolution. C'est aussi un outil très puissant et très complet, particulièrement bien adapté pour la mise en œuvre informatique des méthodes statistiques.

Le logiciel **R** est particulièrement performant pour la manipulation de données, le calcul et l'affichage des graphiques. Il possède entre autres choses :

- Un système de documentation intégré très bien conçu (en anglais).
- Des procédures efficaces de traitement des données et des capacités de stockage de ces données.
- Une suite d'opérateurs pour des calculs sur des tableaux et en particulier sur des matrices.
- Une vaste et cohérente collection de procédures statistiques pour l'analyse de données.
- Des capacités graphiques évoluées.
- Un langage de programmation simple et efficace intégrant les conditions, les boucles, la récursivité et des possibilités d'entrée-sortie.

D.2 L'interface graphique de R

L'interface graphique de **R** (c'est-à-dire l'ensemble de ses menus) est très limitée, voire inexistante sur certaines plates-formes, en comparaison avec des autres logiciels standards. Cette minimalité peut dérouter les novices. Toutefois, nous pouvons nuancer cet inconvénient au travers des points suivants :

- Cela offre l'avantage pédagogique d'inciter l'utilisateur à bien maîtriser la procédure statistique qu'il compte appliquer.
- Il existe des outils additionnels qui permettent d'étendre l'interface graphique.

Le package `Rcmdr`, à installer et à charger (activer) via le menu `Packages`, permet d'effectuer des analyses graphiques et statistiques usuelles au moyen d'une interface plus conviviale à l'aide des menus déroulants. Par ailleurs, les instructions **R** permettant de réaliser l'analyse choisie dans les menus `Rcmdr` s'affichent dans une portion de fenêtre dédiée. Cela peut par exemple être utile si l'on ne connaît pas ou que l'on a oublié une instruction **R** nécessaire à la réalisation d'une tâche particulière.

D.3 R et les statistiques

R est un logiciel dans lequel de nombreuses techniques statistiques modernes et classiques ont été implémentées. Les méthodes les plus courantes permettant de réaliser une analyse statistique telles que :

- Statistique descriptive.
- Tests d'hypothèses.
- Analyse de la variance.
- Méthodes de régression linéaire (simple et multiple).

sont enchâssées directement dans le cœur du système. Notez également que la plupart des méthodes avancées de statistique sont aussi disponibles au travers des modules externes appelés **packages**. Ceux-ci sont faciles à installer directement à partir d'un menu du logiciel. Ils sont tous regroupés sur le site internet du *Comprehensive R Archive Network* (*CRAN*) (<http://cran.r-project.org>) sur lequel vous pouvez les consulter. Ce site fournit aussi, pour certains grands domaines d'étude, une liste commentée des packages associés à ces thèmes (appelée *Task View*), ce qui facilite ainsi la recherche d'une méthode statistique particulière. Par ailleurs une documentation détaillée en anglais de chaque package est disponible sur le *CRAN*.

N.B : Il est par ailleurs utile de noter que les méthodes statistiques les plus récentes y sont régulièrement ajoutées par la communauté statistique elle-même.

D.4 R et les graphiques

Une des grandes forces de **R** réside dans ses capacités, bien supérieures à celles des autres logiciels courants du marché, à combiner un langage de programmation avec la possibilité de réaliser des graphiques de qualité. Les graphiques usuels s'obtiennent aisément au moyen des fonctions prédéfinies. Ces dernières possèdent de très nombreux paramètres permettant par exemple d'ajouter

des titres, des légendes, des couleurs, etc. Mais il est également possible d'effectuer des graphiques plus sophistiqués permettant de représenter des données complexes telles que les courbes de surface ou de niveau, des volumes affichés avec un effet 3D, des courbes de densité, et bien d'autres choses encore. Il vous est également possible d'y ajouter des formules mathématiques. Vous pouvez aussi agencer ou superposer plusieurs graphiques sur la même fenêtre, et utiliser de nombreuses palettes de couleurs.

D.5 R et les packages

Les **packages** (**librairies**) sont des bibliothèques externes qui contiennent des collections de fonctions utilisables sous **R** souvent centré sur un sujet particulier. Ils permettent d'étendre les fonctionnalités de **R**. Ils sont disponibles sur le site WEB : <http://cran.r-project.org/bin/windows/contrib/>, dans le dossier correspondant au numéro de votre version de **R**.

D.5.1 Installation des packages

Il existe plusieurs moyens pour installer un nouveau package. Nous présentons, ci-dessous, deux de ces moyens :

I Installation à partir d'un fichier zip situé sur le disque dur

Pour installer un package à partir d'un fichier zip situé sur le disque dur, on suit les étapes suivantes :

1. Télécharger, depuis le site mentionné ci-dessus, le fichier zip correspondant au package que vous voulez utiliser, et enregistrez le sur le disque dur de votre ordinateur.
2. Lancer le logiciel **R** en double-cliquant sur son icône.
3. Allez dans le menu **Packages**, puis dans le sous-menu **Installer le(s) package(s) depuis des fichiers zip ...**
4. Sélectionnez alors le fichier zip situé sur le disque dur, puis cliquez sur « Ouvrir ».

II Installation directe depuis l'Internet

Pour installer un package depuis l'Internet, qui est une méthode plus simple, on suit les étapes suivantes :

1. Lancer le logiciel **R** en double-cliquant sur son icône située sur le Bureau.
2. Allez dans le menu **Packages**, puis dans le sous-menu **Installer le(s) package(s) ...**
3. Sélectionnez un miroir (*CRAN mirror*) proche de votre situation géographique et cliquez sur « OK ».
4. Sélectionnez l'entrée du package que vous voulez utiliser.
5. Cliquez sur « OK ».

D.5.2 Chargement des packages installés

Installer les packages signifie que les fichiers qu'il contient sont « écrits » physiquement sur le disque dur. Lorsque l'on éteint et rallume l'ordinateur, ces fichiers seront toujours présents à l'endroit où ils auront été copiés. Vous n'aurez donc plus besoin de réinstaller ce package, sauf pour en avoir une version plus récente.

Au contraire, charger un package (en mémoire de **R**) signifie qu'il est temporairement mis à disposition de l'utilisateur dans **R**. Mais si l'on ferme et rouvre **R**, ce package ne sera plus disponible sous **R**. Il faudra donc le charger de nouveau.

Pour résumer, une fois que les packages souhaités ont été installés sur le disque dur de votre ordinateur, il faut les charger dans la mémoire de **R** pour pouvoir les utiliser.

Pour charger un package installé en mémoire de **R**, on peut procéder de deux manières différentes qui suivent :

1. Taper `require(nom_du_package)` dans la console de **R**.
2. Aller dans le menu **Packages** puis dans le sous-menu **Charger le package...**, alors une fenêtre contenant la liste des packages installés sur votre ordinateur apparaîtra, sélectionnez le package que vous souhaitez utiliser, puis cliquez sur « OK ».

BIBLIOGRAPHIE

- [1] **B. ABDOUS, A.L. FOUGÈRES, K. GHOUDI & P. SOULIER.** « Estimation of Bivariate Excess Probabilities for Elliptical Models. ». *Bernoulli*, 14(4). (2008). 1065-1088.
- [2] **D. ALLARD.** *Statistiques spatiales : introduction à la géostatistique.* Cours donné aux universités Montpellier I, Montpellier II et Montpellier SupAgro, Promotion Master Biostatistique.
- [3] **D. APPLEBAUM.** *LÉVY Processes and Stochastic Calculs.* Cambridge Studies in Advanced Mathematics, Cambridge University Press, Volume 93. Cambridge. (2004).
- [4] **B.C. ARNOLD, N. BALAKRISHNAN & H.N. NAGARAJA.** *A First Course in Order Statistics.* Classics In Applied Mathematics, Society for Industrial and Applied Mathematics. Philadelphia. (2008).
- [5] **R.B. ASH & C. A. DOLÉANS -DADE.** *Probability and Measure Theory.* A Harcourt Science and Technology Company, Harcourt Academic Press. San Diego. (2000).
- [6] **K.B. ATHREYA & S.N. LAHIRI.** *Measure Theory and Probability Theory.* Springer Science and Business Media, LLC, Springer. (2006).
- [7] **S. AVOUYI-DOVI & D. GUEGAN.** « Valeurs extrêmes et séries temporelles : application à la finance ». Document de recherche : *Centre d'Etude des Politiques Economiques de l'Université d'Evrey.* Mars (2001).
- [8] **J.N. BACRO & C. GAETAN.** « A Review on Spatial Extreme Modelling ». *Advances and Challenges in Space-time Modelling of Natural Events* (Edited by E. PORCU, J.M. MONTERO & M. SCHLATHER). Lecture Notes in Statistics 207, Springer-Verlag. Berlin. (2012). 103-124.
- [9] **N. BALAKRISHNAN & V.B. NEVZOROV.** *A Primer on Statistical Distributions.* John Wiley & Sons, Inc. New Jersey. (2003).
- [10] **P. BARBE & M. LEDOUX.** *Probabilité.* Collection Enseignements Supérieurs des Mathématiques, EDP Sciences. (2007).
- [11] **D. BARRO.** *Contribution à la modélisation statistique des valeurs extrêmes multivariées.* Thèse de l'Université de Ouagadougou. Décembre (2010).
- [12] **J. BEIRLANT, Y. GOEGEBEUR & J. TEUGELS.** *Statistics of Extremes : Theory and Applications.* John Wiley & Sons, Ltd. (2004).
- [13] **R. BHATTACHARYA & E.C. WAYMIRE.** *A Basic Course in Probability Theory.* Springer Science and Business Media, Inc, Springer. (2007).

- [14] **P. BILLINGSLEY.** *Probability and Measures.* Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, Inc, Troisième édition. New York. (1995).
- [15] **P. BILLINGSLEY.** *Convergence of Probability Measures.* Wiley Series in Probability and Statistics, John Wiley & Sons, Inc, Seconde édition. New York. (1999).
- [16] **N.H. BINGHAM, C.M. GOLDIE & J.L. TEUGELS.** *Regular Variation.* Encyclopedia of Mathematics and its Applications, Cambridge University Press, Volume 27. Cambridge. (1987).
- [17] **J. BLANCHET & A.C. DAVISON.** « Spatial Modeling of Extreme Snow Depth ». *The Annals of Applied Statistics*, Volume 5, N° 3. (2011). 1699-1725.
- [18] **S. BORAK, W. HÄRDLE & R. WERON.** « Stable Distributions ». Discussion Paper 008 : SFB 649, Humboldt-Universität zu Berlin. Février (2005).
- [19] **A. BORCHANI.** « Statistiques des valeurs extrêmes dans le cas de lois discrètes ». Rapport de stage de fin d'étude : ESSEC. Decembre (2010).
- [20] **H. BOUZIANE.** *Calcul et estimation d'une probabilité de ruine.* Mémoire de Magister de l'Université Mouloud MAMMERI de Tizi Ouzou. (2011).
- [21] **L. BREIMAN.** *Probability.* Classics In Applied Mathematics, Society for Industrial and Applied Mathematics. Philadelphia. (1992).
- [22] **P. BRÉMAUD.** *Initiation aux probabilités et aux chaînes de MARKOV.* Springer-Verlag, Seconde édition. Berlin. (2009).
- [23] **P.J. BROCKWELL & R.A. DAVIS.** *Introduction to Time Series and Forecasting.* Springer Texts in Statistics, Springer-Verlag, Seconde édition. New York. (2002).
- [24] **P.J. BROCKWELL & R.A. DAVIS.** *Time series : Theory and Methods.* Springer Series in Statistics, Springer Science and Business Media, LLC, Springer, Seconde édition. New York. (2006).
- [25] **B.M. BROWN & S.I. RESNICK.** « Extreme Values of Independent Stochastic Processes ». *Journal of Applied Probability* 14(4). (1977). 732-739.
- [26] **T.A. BUSHAND, L. DE HAAN & C. ZHOU.** « On Spatial Extremes with Application to a Rainfall Problem ». *The Annals of Applied Statistics*, Volume 2, N° 2. (2008). 624-642.
- [27] **B. CAMINADE & S. NICOLAS.** *Mathématiques, tout-en-un : Cours et exercices corrigés.* J'intègre, Dunod. Paris. (2008).
- [28] **E. CASTILLO, A.S. HADI, N. BALAKRISHNAN & J.M. SARABIA.** *Extreme Value and Related Models with Applications in Engineering and Science.* John Wiley & Sons, Inc. New Jersey. (2005).
- [29] **J.M. CHAMBERS, C.L. MALLOWS & B.W. STUCK.** « A Method for Simulating Stable Random Variables ». *Journal of American Statistical Association*, Volume 71, N° 354. (1976). 340-344.
- [30] **A. CHARPENTIER & M. DENUIT.** *Mathématiques de l'assurance non-vie, tome 2 : Tarification et provisionnement.* Économie et Statistiques Avancées, Economica. (2005).
- [31] **S. COLES.** *An Introduction to Statistical Modeling of Extreme Values.* Springer Series in Statistics, Springer-Verlag. Londres, (2001).
- [32] **S. COLES & J. TAWN.** « Modelling Extreme Multivariate Events ». *Journal of the Royal Statistical Society Series, B* 53(2). (1991). 377-392.

- [33] **S. COLES & J. TAWN.** « Modelling Extremes of the Real Rainfall Process ». *Journal of the Royal Statistical Society Series, B* 58(2). (1996). 329-347.
- [34] **S. COLES & D. WALSHAW.** « Directional Modeling of Extreme wind Speeds ». *Applied Statistics*, 43. (1994). 139-157.
- [35] **D. COOLEY, P. NAVEAU & P. PONCET.** « Variograms for Spatial Max-Stable Random Fields ». *Dependence in Probability and Statistics* (Edited by P. BERTAIL, P. DOUKHAN & P. SOULIER). Lecture Notes in Statistics 187, Springer Science and Business Media, LLC, Springer. (2006). 373-390.
- [36] **D. COOLEY, D. NYCHKA & P. NAVEAU.** « Bayesian Spatial Modeling of Extreme Precipitation Return Levels ». *Journal of American Statistical Association*, 104(479). (2007). 824-840.
- [37] **D.R. COX & N. REID.** « A Note on Pseudolikelihood Constructed from Marginal Densities ». *Biometrika*, 91, 3. (2004). 729-737.
- [38] **N.A.C. CRESSIE.** *Statistics for Spatial Data*. John Wiley & Sons, Inc. New York. (1991).
- [39] **L. D'ESTAMPES.** *Traitement statistique des processus alpha-stables : Mesures de dépendance, identification des AR stables et tests séquentiels tronqués*. Thèse de l'Institut National Polytechnique de Toulouse. Octobre (2003).
- [40] **J. DANIELSSON & C.G. DE VRIES.** « Tail Index Estimation with Very High Frequency Data ». *Journal of Empirical Finance*, 4. (1997). 241-257.
- [41] **A.C. DAVISON, S.A. PADOAN & M. RIBATET.** « Statistical Modeling of Spatial Extremes ». *Statistical Science*, Volume 27, N° 2. (2012). 161-186.
- [42] **A.C. DAVISON & R.L. SMITH.** « Models for Exceedances over High Thresholds ». *Journal of the Royal Statistical Society Series, B* 52. (1990). 393-442.
- [43] **L. DE HAAN.** « A Characterization of Multidimensional Extreme Value Distributions ». *Sankhya Series, A* 40(1). (1978). 85-88.
- [44] **L. DE HAAN.** « A Spectral Representation for Max-Stable Processes ». *The Annals of Probability*, Volume 12, N° 4. (1984). 1194-1204.
- [45] **L. DE HAAN & A. FERREIRA.** *Extreme Value Theory : An Introduction*. Springer Series in Operations Research and Financial Engineering, Springer Science and Business Media, LLC, Springer. (2006).
- [46] **L. DE HAAN & T.T. PEREIRA.** « Spatial Extremes : Models for the Stationary Case ». *The Annals of Statistics*, Volume 34, N° 1. (2006). 146-168.
- [47] **L. DE HAAN & J. PICKANDS.** « Stationary Min-Stable Stochastic Processes ». *Probability Theory and Related Fields*, 72. (1986). 477-492.
- [48] **L. DE HAAN & S.I. RESNICK.** « Limit Theory for Multivariate Sample Extremes ». *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 40. (1977). 317-337.
- [49] **L. DE HAAN & C. ZHOU.** « On Extreme Value Analysis of a Spatial Process ». *REVSTAT - Statistical Journal*, Volume 6, N° 1. (2008). 71-81.
- [50] **P.L. DE MICHEAUX, R. DROUILHET & B. LIQUET.** *Le logiciel R : Maitriser le langage et effectuer des analyses statistiques*. Collection Statistique et Probabilités Avancées, Springer-Verlag France. Paris. (2011).

- [51] **P. DEHEUVELS**. « Caractérisation complète des lois extrêmes multivariées et de la convergence aux types extrêmes ». *Publications de l'Institut de Statistique de l'Université de Paris*, 23. (1978). 1-36.
- [52] **P. DEHEUVELS**. « Probabilistic Aspects of Multivariate Extremes ». *Statistical Extremes and Applications* (Edited by J. TIAGO DE OLIVEIRA). D. Reidel Publishing Company. Dordrecht. (1984). 117-130.
- [53] **P.J. DIGGLE & P.J. RIBEIRO JR.** *Model-based Geostatistics*. Springer Series in Statistics, Springer Science and Business Media, LLC, Springer. New York. (2007).
- [54] **C. DOMBRY**. *Théorie spatiale des extrêmes et propriétés des processus max-stables*. Document de synthèse en vue de l'habilitation à diriger des recherches, Université de Poitiers, UFR Sciences Fondamentales et Appliquées. Novembre (2012).
- [55] **P. EMBRECHTS, C. KLÜPPELBERG & T. MIKOSCH**. *Modelling Extremal Events for Insurance and Finance*. Springer-Verlag. Berlin. (1997).
- [56] **P. EMBRECHTS, S.I. RESNICK & G. SAMORODNITSKY**. « Extreme Value Theory as a Risk Management Tool ». *North American Actuarial Journal*, 26. (1999). 30-41.
- [57] **M. FALK, J. HÜSLER & R. REISS**. *Laws of Small Numbers : Extremes and Rare Events*. Springer, Troisième édition. Basel. (2011).
- [58] **W. FELLER**. *An Introduction to Probability Theory and its Applications*, Volume 1. John Wiley & Son, Inc, Troisième édition. New York. (1968).
- [59] **W. FELLER**. *An Introduction to Probability Theory and its Applications*, Volume 2. John Wiley & Son, Inc, Seconde édition. New York. (1971).
- [60] **B. FINKENSTÄDT & H. ROOTZÉN**. *Extreme Values in Finance, Telecommunications and the Environment*. Monographs on Statistics and Applied Probability 99, Chapman & Hall/CRC, CRC Press LLC. Boca Raton. (2004).
- [61] **R.A. FISHER & L.H.C. TIPPETT**. « Limiting Forms of the Frequency of the Largest or Smallest Member of a Sample ». *Proceedings of the Cambridge Philosophical Society*, 24. (1928). 180-190.
- [62] **D. FOATA & A. FUCHS**. *Calcul des probabilités : Cours, exercices et problèmes corrigés*. Dunod, Seconde édition. Paris. (1998).
- [63] **C. FONSECA, L. PEREIRA, H. FERREIRA & A.P. MARTINS**. « Generalized Madogram and Pairwise Dependence of Maxima over two Regions of a Random Field ». arXiv : 1104.2637v2 [math.ST] 24 Jan 2012. (2012).
- [64] **A.L. FOUGÈRES & P. SOULIER**. « Limit Conditional Distributions for Bivariate Vectors with Polar Representation ». *Stochastic Models*, 26(1). (2010). 54-77.
- [65] **M. FRÉCHET**. « Sur la loi de probabilité de l'écart maximum ». *Annales de la société Polonaise de Mathématiques*, 6. (1927). 93-116.
- [66] **C. GAETAN & X. GUYON**. *Modélisation et statistique spatiales*. Mathématiques et Applications 63, Springer-Verlag. Berlin. (2008).
- [67] **J. GALAMBOS**. *The Asymptotic Theory of Extreme Order Statistics*. Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, Inc. New York. (1987).
- [68] **A.E. GELFAND, P.J. DIGGLE, M. FUENTES & P. GUTTORP**. *Handbook of Spatial Statistics*. Chapman & Hall/CRC, Handbooks of Modern Statistical Methods, Taylor and Francis Group, LLC. New York. (2010).

- [69] **R. GENÇAY & F. SELÇUK.** « Extreme Value Theory and Value-at-Risk : Relative Performance in Emerging Markets ». *International Journal of Forecasting*, 20. (2004). 287-303.
- [70] **M.M. GHOLAM REZAEI.** *Geostatistics of Extremes : A Composite Likelihood Approach*. Thèse de l'École Polytechnique Fédérale de Lausanne. Décembre (2010).
- [71] **E. GINÉ, M.G. HAHN & P. VATAN.** « Max-Infinitely Divisible and Max-Stable Sample Continuous Processes ». *Probability Theory and Related Fields*, 87(2). (1990). 139-165.
- [72] **B.V. GNEDENKO.** « Sur la distribution limite du terme maximum d'une série aléatoire ». *Annales de Mathématiques*, 44. (1943). 423-453.
- [73] **B.V. GNEDENKO & A.N. KOLMOGOROV.** *Limit Distributions for Sums of Independent Random Variables*. Addison-Wesley Series in Statistics, Addison-Wesley Publishing Company, Inc. (1954).
- [74] **J. GUMBEL.** *Statistics of Extremes*. Columbia University Press. New York. (1958).
- [75] **Z. HAN.** « Living to 100 and Beyond : An Extreme Value Study ». *Living to 100 and Beyond Symposium Monograph*. (2005).
- [76] **J.E. HEFFERNAN & S.I. RESNICK.** « Limit Laws for Random Vectors with an Extreme Component ». *The Annals of Applied Probability*, 17(2). (2007). 537-571.
- [77] **J.E. HEFFERNAN & A.G. STEPHENSON.** *An Introduction to Statistical Modeling of Extreme Values*. R Package ismev, Version 1.39. (2012).
- [78] **J.E. HEFFERNAN & J. TAWN.** « A Conditional Approach for Multivariate Extreme Values ». *Journal of the Royal Statistical Society Series, B* 66(3). (2004). 497-546.
- [79] **C.C. HEYDE.** « A Contribution to the Theory of Large Deviations for Sums of Independent Random Variables ». *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 7. (1967). 303-308.
- [80] **B.M. HILL.** « A Simple General Approach to Inference about the Tail of a Distribution ». *Annals of Statistics*, 3(5). (1975). 1163-1174.
- [81] **A. JARRAH.** *Développement de méthodes statistiques et probabilistes en corrosion par piqûres pour l'estimation de la profondeur maximale : Application à l'aluminium A5*. Thèse de l'École Nationale Supérieure d'Arts et Métiers. Décembre (2009).
- [82] **A.G. JOURNEL & C.J. HUIJBREGTS.** *Mining Geostatistics*. Academic Press. London. (1978).
- [83] **Z. KABLUCHKO.** « Spectral Representations of Sum - and Max-stable Processes ». *Extremes*, 12(4). (2009). 401-424.
- [84] **Z. KABLUCHKO.** « Extremes of Independent Gaussian Processes ». *Extremes*, 14. (2011). 285-310.
- [85] **Z. KABLUCHKO & M. SCHLATHER.** « Ergodic Properties of Max-Infinitely Divisible Processes ». *Stochastic Processes and their Applications*, 120(3). (2010). 281-295.
- [86] **Z. KABLUCHKO, M. SCHLATHER & L. DE HAAN.** « Stationary Max-Stable Fields Associated to Negative Definite Functions ». *The Annals of Probability*, Volume 37, N° 5. (2009). 2042-2065.
- [87] **D. KANNAN.** *An Introduction to Stochastic Processes*. Elsevier North Holland, Inc. New York. (1979).

- [88] **R.W. KATZ, M.B. PARLANGE & P. NAVEAU.** « Statistics of Extremes in Hydrology ». *Advances in Water Resources*, 25. (2002). 1287-1304.
- [89] **B. KERESZTES.** *Segmentation de zones d'intérêt par processus objets : Application à la détection des réseaux de failles sismiques.* Thèse de l'Université de Bordeaux I. Décembre (2009).
- [90] **H. KLAJNMIĆ.** « Niveaux de retour pour les vitesses extrêmes des vents ». XXXV^{èmes} journées de Statistique. Lyon. (2003).
- [91] **S. KOTZ & S. NADARAJAH.** *Extreme Value Distributions : Theory and Applications.* Imperial College Press. (2000).
- [92] **B. LAPEYRE, É. PARDOUX & R. SENTIS.** *Méthodes de Monte-Carlo pour les équations de transport et de diffusion.* Mathématiques et Applications 29, Springer-Verlag. Berlin. (1998).
- [93] **N.D. LE & J.V. ZIDEK.** *Statistical Analysis of Environmental Space-Time Processes.* Springer Series in Statistics, Springer Science and Business Media, Inc, Springer. New York. (2006).
- [94] **M.R. LEADBETTER, G. LINDGREN & H. ROOTZÉN.** *Extremes and Related Properties of Random Sequences and Processes.* Springer-Verlag. New York. (1983).
- [95] **M. LEJEUNE.** *Statistique : La théorie et ses applications.* Collection Statistique et Probabilités Appliquées, Springer-Verlag France, Seconde édition. Paris. (2010).
- [96] **A. LEKINA.** *Estimation non-paramétrique des quantiles extrêmes conditionnels.* Thèse de l'Université de Grenoble. Octobre (2010).
- [97] **B.G. LINDSAY.** « Composite Likelihood Methods ». *American Mathematical Society, Contemporary Mathematics*, Volume 80. (1988). 221-239.
- [98] **F. LONGIN.** « La théorie des valeurs extrêmes : présentation et premières applications en finance ». *Journal de la Société de Statistique de Paris*, Tome 136, N° 1. (1995). 77-95.
- [99] **F. LONGIN.** « Value at Risk : Une nouvelle approche fondée sur les valeurs extrêmes ». *Annales d'Economie et de Statistique*, 52. (1998). 23-51.
- [100] **F. LONGIN.** « From Value at Risk to Stress Testing : The Extreme Value Approach ». *Journal of Banking and Finance*, 24. (2000). 1097-1130.
- [101] **A. MANSOUR.** *Probabilités et statistiques pour les ingénieurs : Cours, exercices et programmation.* Lavoisier. (2007).
- [102] **V. MARIĆ.** *Regular Variation and Differential Equations.* Lecture Notes in Mathematics 1726, Springer-Verlag. Berlin. (2000).
- [103] **A.J. MC NEIL.** *Calculating Quantile Risk Measures for Financial Time Series using Extreme Value Theory.* Department of Mathematics, ETH Swiss, Federal Technical University E-Collection. (1998).
- [104] **A.J. MC NEIL & T. SALADIN.** « The Peaks Over Thresholds Method for Estimating High Quantiles of Loss Laws ». Department of Mathematic, ETH Zurich. (1997).
- [105] **D. MERAGHNI.** *Modelling Distributions Tails.* Thèse de l'Université de Biskra. Janvier (2008).
- [106] **M. OESTING, Z. KABLUCHKO & M. SCHLATHER.** « Simulation of BROWN-RESNICK Processes ». *Extremes* DOI 10.1007/s10687-011-0128-8. (2011).

- [107] **J. OUVRARD**. *Probabilités 2 : Master-Agrégation*. Cassini. Paris. (2000).
- [108] **S.A. PADOAN, M. RIBATET & S.A. SISSON**. « Likelihood-based Inference for Max-Stable Processes ». *Journal of the American Statistical Association* 105. (2010). 263-277.
- [109] **E. PARADIS**. « R pour débutants ». *Institut des Sciences de l'Évolution, Université Montpellier II*. Septembre (2005).
- [110] **J. PICKANDS**. « Statistical Inference using Extreme Order Statistics ». *Annals of Statistics*, 3. (1975). 119-131.
- [111] **B. REGGAD**. « Fondements de la théorie des valeurs extrêmes, ses principales applications et son apport à la gestion des risques du marché pétrolier ». *Math. Sci. Hum, Mathematics and Social Sciences*, 47^{ème} Année, N° 186. (2009)2. 29-63.
- [112] **R.D. REISS & M. THOMAS**. *Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields*. Birkhäuser Verlag, Troisième édition. Basel. (2007).
- [113] **S.I. RESNICK**. « Point Processes, Regular Variation and Weak Convergence ». *Advances in Applied Probability*, Volume 18, N° 1. (1986). 66-138.
- [114] **S.I. RESNICK**. *Extreme Values, Regular Variation, and Point Processes*. Applied Probability Trust, Springer-Verlag. New York, (1987).
- [115] **S.I. RESNICK**. *Adventures in Stochastic Processes*. Birkhäuser. Boston. (1992).
- [116] **S.I. RESNICK**. *A Probability Path*. Birkhäuser. Boston. (1998).
- [117] **S.I. RESNICK**. *Heavy-Tail Phenomena : Probabilistic and Statistical Modeling*. Springer Series in Operations Research and Financial Engineering, Springer Science and Business Media, LLC, Springer, (2006).
- [118] **S.I. RESNICK & R. ROY**. « Random usc Functions, Max-Stable Processes and Continuous Choice ». *The Annals of Applied Probability*, 1(2). (1991). 267-292.
- [119] **M. RIBATET**. *A User's Guide to the SpatialExtremes Package*. Chair of Statistics, École Polytechnique Fédérale de Lausanne. Switzerland. (2009).
- [120] **M. RIBATET**. « An Introduction to Max-Stable Processes ». *Institut de Mathématiques et de Modélisation de Montpellier*. June 10 (2011).
- [121] **M. RIBATET**. *Modelling Spatial Extremes*. **R** Package **SpatialExtremes**, Version 1.9-0. (2012).
- [122] **T. RONCALLI**. « Théorie des valeurs extrêmes ou modélisation des événements rares pour la gestion des risques ». DESS 203 : *Université de Paris IX Dauphine*. Janvier (2002).
- [123] **H. ROOTZÉN & N. TAJVIDI**. « Extreme Value Statistics and Wind Storm Losses : A Case Study ». *Scandinavian Actuarial Journal*. (1997). 70-94.
- [124] **M. SCHLATHER**. « Models for Stationary Max-Stable Random Fields ». *Extremes*, 5(1). (2002). 33-44.
- [125] **M. SCHLATHER & J. TAWN**. « Inequalities for the Extremal Coefficients of Multivariate Extreme Value Distributions ». *Extremes*, 5(1). (2002). 87-102.
- [126] **M. SCHLATHER & J. TAWN**. « A Dependence Measure for Multivariate and Spatial Extreme Values : Properties and Inference ». *Biometrika*, 90(1). (2003). 139-156.

- [127] **R. SERFOZO.** *Basics of Applied Stochastic Processes*. Probability and Its Applications, Springer-Verlag. Berlin. (2009).
- [128] **A.N. SHIRYAEV.** *Probability*. Graduate Texts in Mathematics, Springer, Seconde édition. (1984).
- [129] **M. SIBUYA.** « Bivariate Extreme Statistics I ». *Annals of the Institute of Statistical Mathematics*, Tokyo, 11. (1960). 195-210.
- [130] **R.L. SMITH.** « Max-Stable Processes and Spatial Extremes ». *Unpublished manuscript*. (1990).
- [131] **R.L. SMITH.** « Extreme Value Statistics in Meteorology and the Environment ». *environmental Statistics*. (2001). 300-357.
- [132] **S.A. STOEV.** « Max-Stable Processes : Representations, Ergodic, Properties and Statistical Applications ». *Dependence in Probability and Statistics* (Edited by P. DOUKHAN, G. LANG, D. SURGAILIS & G. TEYSSIÈRE). Lecture Notes in Statistics 200, Springer-Verlag. Berlin. (2010). 21 - 42.
- [133] **C. SUQUET.** *Intégration, analyse de FOURIER et probabilités*. Université des Sciences et Technologies de Lille, U.F.R de Mathématiques Pures et Appliquées. (2003-2004).
- [134] **N. TAJVIDI.** *Characterisation and Some Statistical Aspects of Univariate and Multivariate Generalised PARÉTO Distributions*. Chalmers Tekniska Högskola, Department of Mathematics, Göteborg. (1996).
- [135] **R.A. THATCHER.** « The Long-Term Pattern of Adult Mortality and the Highest Attained Age ». *Journal of the Royal Statistical Society Series, A* 162. (1999). 5-43.
- [136] **J. TIAGO DE OLIVEIRA.** « La représentation des distributions extrémales bivariées ». *Bulletins de l'Institut International de Statistique*, 39(2). (1962). 477 - 480.
- [137] **S.N. TOKPAVI.** *Essais sur la Value-at-Risk : mesures de risque intra-journalières et tests de validation*. Thèse de l'Université d'Orléans. Décembre (2008).
- [138] **C. VARIN.** « On Composite Marginal Likelihoods ». *Advances in Statistical Analysis*, 92. (2008). 1-28.
- [139] **C. VARIN, N. REID & D. FIRTH.** « An Overview of Composite Likelihood Methods ». *Statistica Sinica*, 21. (2011). 5-42.
- [140] **C. VARIN & P. VIDONI.** « A Note on Composite Likelihood Inference and Model Selection ». *Biometrika*, 92(3). (2005). 519-528.
- [141] **R. VON MISES.** « La distribution de la plus grande des n valeurs ». *Selected Papers II, American Mathematical Society*. (1936). 271-294.
- [142] **Y. WANG & S.A. STOEV.** « On the Structure and Representations of Max-Stable Processes ». *Advances in Applied Probability* 42. (2010). 855-877.

RÉSUMÉ

Les événements climatiques extrêmes ont été étudiés par de nombreux chercheurs durant les dernières décennies et les statisticiens ont également développé des outils statistiques capables de traiter ce type de données. Bien que la théorie des valeurs extrêmes aie été largement développée et utilisée dans la modélisation des événements tels que les précipitations extrêmes et les canicules, la nature spatiale des données climatiques requiert de nouveaux outils pour tenir compte au mieux des extrêmes spatiaux. Ce mémoire fait d'abord une revue de la théorie des valeurs extrêmes univariée et de la théorie des champs aléatoires ainsi qu'une revue des modèles et des méthodes existantes pour la modélisation des événements extrêmes. Puis combine ces deux théories afin de donner de nouveaux modèles statistiques (**champs aléatoires max-stables**) qui décrivent les événements extrêmes tenant compte de la composante spatiale. Les questions principales qui se posent sont la structure de la dépendance spatiale des extrêmes (**coefficient extrémal**) et les méthodes qui permettent d'estimer les paramètres des champs aléatoires max-stables (**vraisemblance composite**).

Mots clés : Théorie des valeurs extrêmes, Distribution des valeurs extrêmes généralisée, Extrêmes spatiaux, Champs aléatoires max-stables, Coefficient extrémal, Vraisemblance composite.

ABSTRACT

Extreme climate events have been investigated by many researchers in recent decades, and statisticians too have developed statistical tools capable of dealing with them. Although extreme value theory has been extensively developed and used in modelling events such as extreme rainfall and heat waves, the spatial nature of climate data requires new tools to deal with spatial extremes. This work makes first a review of the univariate extreme value theory and the random fields theory as well as a review of the models and the existing methods for the modelling of the extreme events. Then combines these two theories in order to give new statistical models (**Max-Stable Random Fields**) that describe the extreme events taking into account the spatial component. The main questions that arise are the structure of the spatial dependence of the extremes (**Extremal Coefficient**) and the methods that permit to estimate the parameters of the max-stable random fields (**Composite Likelihood**).

Key words : Extreme Value Theory, Generalised Extreme Value Distribution, Spatial Extremes, Max-Stable Random Fields, Extremal Coefficient, Composite Likelihood.