HAL open science

Contributions à l'estimation de quantiles extrêmes. Applications à des données environnementales

Jonathan El Methni

To cite this version:

Jonathan El Methni. Contributions à l'estimation de quantiles extrêmes. Applications à des données environnementales. Mathématiques générales [math.GM]. Université de Grenoble, 2013. Français. NNT : 2013GRENM035 . tel-00924293

HAL Id: tel-00924293

https://theses.hal.science/tel-00924293

Submitted on 6 Jan 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ DE GRENOBLE

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Mathématiques Appliquées**

Arrêté ministériel : 7 août 2006

Présentée par

Jonathan EL METHNI

Thèse dirigée par **Stéphane GIRARD**
et codirigée par **Laurent GARDES**

préparée au sein **du Laboratoire Jean Kuntzmann**
et de **l'École Doctorale Mathématiques, Sciences et Technologies de l'Information, Informatique**

Contributions à l'estimation de quantiles extrêmes. Applications à des données environnementales

Thèse soutenue publiquement le **lundi 7 Octobre 2013**,
devant le jury composé de :

Mme. Clémentine PRIEUR
Professeur, Université Joseph Fourier, Présidente
M. Clément DOMBRY
Professeur, Université de Franche-Comté, Rapporteur
Mme. Irène GIJBELS
Professeur, Université Catholique de Louvain, Rapporteur
Mme. Véronique MAUME-DESCHAMPS
Professeur, ISFA-Université Lyon 1, Examinatrice
M. Stéphane GIRARD
Chargé de Recherche, Inria Rhône-Alpes, Directeur de thèse
M. Laurent GARDES
Professeur, Université de Strasbourg, Directeur de thèse

A mes parents Rose-Marie et Mohamed El Methni
et à ma grand-mère Antonia Buffa Valès

"Where there is a will, there is a way."

Proverbe Anglais

Remerciements

Cette thèse n'aurait jamais pu voir le jour sans le soutien d'un grand nombre de personnes que je tiens à remercier pour leurs encouragements, leur générosité et leur gentillesse.

En premier lieu, je tiens à exprimer ma profonde gratitude à l'égard de mon directeur de thèse Stéphane Girard pour la confiance qu'il m'a accordé en acceptant d'encadrer ma thèse et de l'avoir si bien fait. J'aimerai également lui dire à quel point j'ai apprécié ses compétences, sa rigueur et sa clairvoyance. J'ai été extrêmement sensible à ses qualités humaines d'écoute, de disponibilité et de bienveillance. Sans compter son humour et sa simplicité. J'ai pris un très grand plaisir à travailler avec lui.

Je souhaite également remercier profondément mon co-directeur de thèse Laurent Gardes qui malgré la distance a toujours été présent. Par ses conseils avisés, ses multiples relectures et remarques pertinentes il m'a offert le meilleur encadrement possible afin de mener mes travaux de recherche. J'ai beaucoup apprécié sa gentillesse, sa disponibilité et sa patience.

Merci à tous les deux de m'avoir tant apporté aussi bien humainement que scientifiquement. Merci pour nos dicussions si enrichissantes et d'avoir toujours été disponibles pour moi. J'espère avoir été digne de la confiance que vous m'avez accordé et que nous continueront à travailler ensemble. J'ai beaucoup appris à vos côtés et je suis très honoré de vous avoir eu pour encadrants.

J'adresse aussi mes remerciements aux deux rapporteurs de ma thèse Irène Gijbels et Clément Dombry. Merci pour vos relectures attentives, vos remarques, vos conseils et pour l'intérêt que vous avez porté à mon travail. Je tiens également à remercier Clémentine Prieur d'avoir accepté de présider mon jury de thèse. Ma gratitude s'adresse également à Véronique Maume-Deschamps d'avoir accepté d'examiner ma thèse. Un grand merci à tous les membres de mon jury pour leurs remarques et les perspectives scientifiques qu'elles offrent à mes travaux. Merci à vous tous de m'avoir fait l'honneur d'avoir pris part à mon jury de thèse.

Je tiens à remercier très chaleuresement Armelle Guillou pour notre collaboration et pour m'avoir permis de présenter mes travaux.

Merci à Inria Rhône-Alpes qui m'a offert les meilleures conditions de travail qui soient pour mon doctorat. Je tiens à remercier Florence Forbes qui m'a accueilli dans cette superbe équipe de recherche qu'est Mistis et de m'avoir permis de présenter mes travaux lors de plusieurs manifestations scientifiques dans divers pays.

Mes remerciements vont également aux autres membres de l'équipe et plus particulièrement à celles et ceux avec qui j'ai partagé mon bureau : Angelika qui m'a tant aidé, El Hadji pour nos discussions mathématiques et nos travaux futurs, Marc pour ses conseils filmographiques, Gildas avec qui j'ai partagé plus qu'un bureau durant deux ans mais aussi Julie, Hessam, Eugen, Mathieu F et Vasil.

J'aimerai également remercier pour leur gentillesse et leurs encouragements : Christine avec qui j'ai commencé ma thèse, Flor, Marie-José, Laure, Farida, Senan, Nourou, Jean-Baptiste, Ludovic, Lotfi, Thomas V, Trung, Kai, Giao et Darren. Je souhaite également remercier nos assistantes Imma et Françoise pour m'avoir tant facilité la vie durant ces trois années.

Merci Juliette Blanchet pour nos discussions scientifiques sur mes travaux et tes conseils sur l'après-thèse. Merci également Charles Bouveyron de m'avoir invité à présenter mes travaux. Un immense merci à Alexandre Lekina pour ses travaux de thèse et ses nombreux documents qui m'ont beaucoup aidé. Merci Lamiae Azizi pour ton accueil à l'Inria et toute l'aide que tu as pu m'apporter.

A nos parties de babyfoot et plus encore, merci Audrey, Yohan, Fabien, Willy, Manh Toan, Éric, Benoit, Hugo, Dimitar et Thibaud. Pour ton amitié merci David L.

J'aimerai adresser un remerciement à Shakila pour sa gentillesse et son soutien quotidien.

Durant ma thèse j'ai eu l'opportunité d'intervenir dans un collège afin d'initier des élèves de sixième à la recherche en mathématiques. Ce fut l'occasion de travailler dans la classe de Françoise Lemoal que je tiens à remercier. L'intervention s'est faite en compagnie d'un autre doctorant David Salinas aujourd'hui docteur et ami.

Les conférences auxquelles j'ai pu prendre part furent l'occasion de faire des rencontres enrichissantes, merci Anne-Laure Fougères, Cécile Mercadier, Anne Sabourin, Anne-Catherine Favre, Marianne Clausel, Jean-Noël Bacro, Thomas Opitz, Philippe Naveau, Abdelaati Daouia, Philippe Garat et Mathieu Ribatet. Un grand merci à Gilles Stupfler pour ses conseils et sa sympathie.

J'adresse une pensée toute particulière à mes anciens professeurs : Françoise Truc, Julie Dugdale, Corinne Berzin, Catherine et Gérard d'Aubigny, Karim Benhenni, Abdel Abdali, Jean-Michel Billiot, Alain Latour et Azzam Hassan.

Je souhaiterais exprimer ma profonde gratitude envers mes anciens professeurs qui furent également mes collègues durant mon monitorat et qui sont aujourd'hui des amis Rémi Drouilhet et Jean-François Coeurjolly.

Mes remerciements à Franck Corset et à Éric Fontenas pour m'avoir soutenu et conseillé sur l'orientation à donner à ma carrière. Je remercie Alexandre et Meryam pour ces bons moments passés ensemble à Bruxelles.

A mes amis connus en classes préparatoires : Thibault S, Arnold, Fabrice, Baptiste, Tarik et Gauthier. A mes amis de Licence et de Master : Aurore, Alexandra, Claire Thibaut J, et Akrame. Merci Stéphanie pour tes nombreux conseils sur l'après-thèse.

Merci Jérôme L pour ces trois années d'études passées ensemble et nos si nombreux fous rires.

Je remercie Olivier Renaud, Paolo Ghisletta et Julien Chanal qui m'ont donné la chance de rejoindre leur équipe Mad en tant que Post-doctorant. Une pensée à mes nouvelles partenaires de bureau : Emilie pour sa bonne humeur quotidienne et Eda pour sa gentillesse et pour avoir été si conciliante pour les Tds pendant la préparation de ma soutenance de thèse.

Merci à mon ami Thomas A pour nos nombreuses sorties à vélo et nos montées des quatre seigneurs et pour toutes nos discussions. A mes amis de toujours : Robin, Fif, Mc, Kevin, Mathieu B et Rémi B-M merci.

Merci Mathieu R d'avoir été là avec ta joie de vivre tous les jours à l'Inria du début de mon stage jusqu'à ma soutenance de thèse. Merci du fond du coeur pour ton soutien et pour ton aide mon ami.

A nos nombreuses randonnées et pour m'avoir fait découvrir et aimer la montagne, merci Florian et Camille d'être mes amis.

Merci Vincent d'avoir rythmé mes étés de thèse par nos sorties de VTT et en particulier pour nos traversées de la Chartreuse et du Vercors. Même si ma thèse fut rude par moment merci d'avoir été présent Rémi L.

A mes amies, amis et proches : Marie-Joe et Armand, Nahid et Taghi, Léon et Tina, Susana, Jacqueline, Charles et Michel, Rodney et Francie, Michelle et Jean-Pierre, Jean-Pierre et Dane, Christian et Pascal, Bérénice, Constance, Marie, Charline, Géraud, Giuseppe, Jérôme M, Damien, Olga et Pierrick, Ezequiel, Cyrille, Nicolas, Colin, Antoine, Guilhem, Sébastien, Pierre, Micka, Long et Selim merci.

J'aimerai remercier ma famille de Tunisie et d'Espagne qui m'a apporté son soutien. A mes oncles et tantes et en particulier à Azzedine et Taoufik, à mes cousines et cousins. A Lola, Pilar et Josep. A Volga.

Aux deux personnes sans qui ce travail n'aurait pu se réaliser et à qui je dois tout, mes parents. Merci Papa pour m'avoir montré la voie à suivre en m'initiant à la beauté des mathématiques et pour ta patience. Merci Maman pour ton éducation, ton aide de tous les instants et ton soutien inconditionnel. Merci pour votre amour et pour l'exemple que vous êtes pour moi.

Je dédie cette thèse à mes anciens et futurs étudiants et plus particulièrement à Alexia et Lucas.

Résumé

Cette thèse s'inscrit dans le contexte de la statistique des valeurs extrêmes. Elle y apporte deux contributions principales.

Dans la littérature récente en statistique des valeurs extrêmes, un modèle de queues de distributions a été introduit afin d'englober aussi bien les lois de type Pareto que les lois à queue de type Weibull. Les deux principaux types de décroissance de la fonction de survie sont ainsi modélisés. Un estimateur des quantiles extrêmes a été déduit de ce modèle mais il dépend de deux paramètres inconnus, le rendant inutile dans des situations pratiques. La première contribution de cette thèse est de proposer des estimateurs de ces paramètres. Insérer nos estimateurs dans l'estimateur des quantiles extrêmes précédent permet alors d'estimer des quantiles extrêmes pour des lois de type Pareto aussi bien que pour des lois à queue de type Weibull d'une façon unifiée. Les lois asymptotiques de nos trois nouveaux estimateurs sont établies et leur efficacité est illustrée sur des données simulées et sur un jeu de données réelles de débits de la rivière Nidd se situant dans le Yorkshire en Angleterre.

La seconde contribution de cette thèse consiste à introduire et estimer une nouvelle mesure de risque appelé Conditional Tail Moment. Elle est définie comme le moment d'ordre $a \geq 0$ de la loi des pertes au-delà du quantile d'ordre $\alpha \in]0, 1[$ de la fonction de survie. Estimer le Conditional Tail Moment permet d'estimer toutes les mesures de risque basées sur les moments conditionnels telles que la Value-at-Risk, la Conditional Tail Expectation, la Conditional Value-at-Risk, la Conditional Tail Variance ou la Conditional Tail Skewness. Ici, on s'intéresse à l'estimation de ces mesures de risque dans le cas de pertes extrêmes c'est-à-dire lorsque α tend vers 0 lorsque la taille de l'échantillon augmente. On suppose également que la loi des pertes est à queue lourde et qu'elle dépend d'une covariable. Les estimateurs proposés combinent des méthodes d'estimation non-paramétrique à noyau avec des méthodes issues de la statistique des valeurs extrêmes. Le comportement asymptotique de nos estimateurs est établi et illustré aussi bien sur des données simulées que sur des données réelles de pluviométrie provenant de la région Cévennes-Vivarais.

Abstract

This thesis can be viewed within the context of extreme value statistics. It provides two main contributions to this subject area.

In the recent literature on extreme value statistics, a model on tail distributions which encompasses Pareto-type distributions as well as Weibull tail-distributions has been introduced. The two main types of decreasing of the survival function are thus modeled. An estimator of extreme quantiles has been deduced from this model, but it depends on two unknown parameters, making it useless in practical situations. The first contribution of this thesis is to propose estimators of these parameters. Plugging our estimators in the previous extreme quantiles estimator allows us to estimate extreme quantiles from Pareto-type and Weibull tail-distributions in an unified way. The asymptotic distributions of our three new estimators are established and their efficiency is illustrated on a simulation study and on a real data set of exceedances of the Nidd river in the Yorkshire (England).

The second contribution of this thesis is the introduction and the estimation of a new risk measure, the so-called Conditional Tail Moment. It is defined as the moment of order $a \geq 0$ of the loss distribution above the quantile of order $\alpha \in (0, 1)$ of the survival function. Estimating the Conditional Tail Moment permits to estimate all risk measures based on conditional moments such as the Value-at-Risk, the Conditional Tail Expectation, the Conditional Value-at-Risk, the Conditional Tail Variance or the Conditional Tail Skewness. Here, we focus on the estimation of these risk measures in case of extreme losses *i.e.* when α converges to 0 when the size of the sample increases. It is moreover assumed that the loss distribution is heavy-tailed and depends on a covariate. The estimation method thus combines nonparametric kernel methods with extreme-value statistics. The asymptotic distribution of the estimators is established and their finite sample behavior is illustrated both on simulated data and on a real data set of daily rainfalls in the Cévennes-Vivarais region (France).

Table des matières

Remerciements **v**

Résumé **ix**

Abstract **xi**

Introduction générale **1**

1 Présentation de la théorie des valeurs extrêmes univariées **7**
1.1 Introduction au comportement du maximum d'un échantillon 9
1.2 La loi des valeurs extrêmes 14
1.3 La loi des excès 19
1.4 Caractérisation des domaines d'attraction 22
1.4.1 Fonctions à variations régulières 22
1.4.2 Domaine d'attraction de Fréchet 24
1.4.3 Domaine d'attraction de Weibull 25
1.4.4 Domaine d'attraction de Gumbel 26
1.5 Estimation de quantiles extrêmes 29
1.5.1 L'approche par la loi des valeurs extrêmes 30
1.5.2 L'approche par la loi des excès 34
1.5.3 L'approche semi-paramétrique 40
1.5.3.1 Lois à queues lourdes 40
1.5.3.2 Lois à queue de type Weibull 45
1.6 Estimation de quantiles conditionnels extrêmes 50
1.6.1 L'approche paramétrique 52
1.6.2 L'approche semi-paramétrique 52
1.6.3 L'approche non-paramétrique 52

2 Estimation de quantiles extrêmes pour des lois à queues lourdes et légères **55**
2.1 Introduction 57
2.2 Cadre d'étude et présentation du modèle 59
2.2.1 Modèle 59
2.2.2 Définition des estimateurs et leurs lois asymptotiques 61

2.2.3 Application aux lois à queue de type Weibull 64
2.2.4 Application au domaine d'attraction de Fréchet 65
2.3 Contributions théoriques 66
2.3.1 Définition des estimateurs 67
2.3.2 Propriétés asymptotiques 68
2.4 Simulations et illustrations sur données réelles 70
2.4.1 Simulations 70
2.4.2 Application aux crues d'une rivière 80
2.5 Perspectives 83
2.6 Démonstrations 84
2.6.1 Résultats préliminaires 84
2.6.2 Preuves des principaux résultats 86
2.6.3 Preuves des résultats préliminaires 91

3 Estimation non-paramétrique de mesures de risque extrêmes pour des lois conditionnelles à queues lourdes 99
3.1 Introduction 101
3.2 Le Regression Conditional Tail Moment : definition et estimation 107
3.2.1 Une nouvelle mesure de risque 107
3.2.2 Pertes extrêmes et regression 108
3.2.3 Inférence 109
3.3 Principaux résultats 110
3.4 Mesures de risque pour des extrêmes pluviométriques 116
3.4.1 Description du problème et du jeu de données réelles 116
3.4.2 Choix des paramètres de réglage 120
3.4.3 Validation sur simulations 121
3.4.4 Mesures de risque estimées sur des extrêmes pluviométriques . . . 131
3.5 Perspectives 136
3.6 Démonstrations 137
3.6.1 Résultats préliminaires et leurs preuves 137
3.6.2 Preuves des principaux résultats 147

Conclusion et perspectives 153

Bibliographie 157

Introduction générale

"La loi des grands nombres et la distribution gaussienne, fondements de l'étude statistique des grandeurs moyennes, échouent à rendre compte des événements rares ou extrêmes. Pour ce faire, des outils statistiques plus adaptés existent ... mais ne sont pas toujours utilisés !"

Rama Cont

Dans la nuit du 31 janvier au 1er février 1953, de très fortes tempêtes traversant la mer du Nord balayèrent les côtes flamandes et néerlandaises d'ouest en est. Après que plusieurs digues eurent cédé, les provinces néerlandaises de la Hollande et de la Zélande furent particulièrement touchées. Les conséquences de ce raz-de-marée furent désastreuses. On dénombra plus de 2 500 morts, 47 000 habitations inondées et 10 000 détruites, 200 000 hectares de terres inondées, 30 000 têtes de bétail noyées, environ 9% des fermes des Pays-Bas inondées et plus de 400 brèches dans les digues.

A la suite de cette tempête, le gouvernement néerlandais décida la mise en place du "plan Delta" pour se prémunir contre une nouvelle inondation. Il fallut construire un nouveau réseau de digues renforcées de plus de 500 km le long de la côte de la mer du Nord.

Un comité composé de nombreux scientifiques, parmi lesquels des statisticiens, se réunit afin d'étudier le phénomène et proposer ainsi des recommandations sur les hauteurs des digues. Il fallut prendre en compte des facteurs économiques (coût de construction, coût des inondations,...), des facteurs physiques (rôle du vent sur la marée,...), mais aussi les hauteurs de marées enregistrées lors des précédentes inondations. En 1953, les inondations avaient entraîné une montée des eaux à 3.85 mètres au-dessus du niveau de la mer, soit largement en-deçà des 4 mètres atteints le 1er novembre 1570, soit 382 ans auparavant.

Le but était de construire des digues assez grandes, de telle sorte qu'aucune vague ne les dépasse dans un horizon de 10 000 ans. Autrement dit, il convenait de déterminer quelle serait la hauteur de la plus grande vague dans un horizon de temps de 10 000 ans.

La difficulté résidait dans le fait que, pour calculer la hauteur des digues et donc la hauteur maximale d'une vague qui n'a jamais eu lieu, le comité d'experts devait se baser sur les informations des années précédentes ; or il n'y avait que très peu de données disponibles en particulier pour des événements de cette ampleur.

FIGURE 1: Vues aériennes d'Oude Tonge, au nord de la Zélande, le 3 février 1953, après le raz-de-marée. Source : Wikipédia.

Le 4 octobre 1986 marqua l'achèvement du plan Delta aux Pays-Bas. Il s'agit du plus grand chantier de génie civil de tous les temps. Il permit de relier toutes les îles côtières de la province de Zélande par des digues.

Après calcul, le comité d'expert estima que ces digues devraient mesurer au moins 5 mètres, estimation fondée sur l'utilisation de techniques statistiques empruntées à la théorie des valeurs extrêmes qui constitue le sujet de la présente thèse.

De la nécessité d'une théorie

Il est d'un grand intérêt de se prémunir contre les risques extrêmes, qu'ils résultent d'une crise financière, d'un accident nucléaire ou d'une catastrophe naturelle, compte tenu des répercussions humaines, économiques et financières que ces derniers peuvent avoir.

Les inondations de 1570 et de 1953 peuvent nous conduire à nous interroger sur une possible récurrence de ces événements. Le but serait d'être en mesure de prédire l'apparition de tels phénomènes, leurs impacts et le retour de ces derniers. Pour cela, il convient de définir précisément ce qu'est un événement extrême.

Un événement extrême est un événement qui a une faible probabilité de se produire mais qui, lorsqu'il se produit, prend de très petites ou de très grandes valeurs et a un grand impact. On notera la différence avec un événement rare qui, par définition, est un événement dont la probabilité d'occurence est faible. Le fait qu'un événement soit rare n'implique pas qu'il soit extrême ; il est dépourvu de la notion de quantifiabilité (petites ou grandes valeurs). A l'inverse, tout événement extrême est rare au sens où il a une faible probabilité de se produire.

De la difficulté d'estimer des événements extrêmes

Les mathématiques, et plus particulièrement la théorie probabiliste, nous offre des outils très puissants permettant de prédire le hasard et donc de le contrôler. La discipline reliant cette théorie aux événements de tous les jours est la statistique. Cette dernière a pour but d'utiliser les données disponibles afin d'en retirer le maximum d'informations et de modéliser au mieux la loi du hasard régissant le phénomène.

De fait, les événements extrêmes se prêtent difficilement à la modélisation puisque par définition les données sont peu nombreuses et peuvent même être inexistantes. L'information la plus précise est celle contenue dans les valeurs les plus extrêmes observées. A l'inverse, on dispose d'un grand nombre d'observations pour les événements plus fréquents. La difficulté réside dans le fait de prédire des événements extrêmes à l'aide d'événements fréquents. On doit donc, à partir de peu de données, construire un modèle nous permettant d'extrapoler et de prédire un événement sans commune mesure.

La théorie des valeurs extrêmes fournit une base mathématique probabiliste rigoureuse sur laquelle il est possible de construire des modèles statistiques permettant de prévoir l'intensité et la fréquence de ces événements extrêmes. Selon Kotz et Nadarajah [125] et Reiss [145], les premiers développements de cette théorie sont attribués à Nicolas Bernoulli en 1709. Fuller fut probablement le premier à en faire application en 1914.

De nombreux domaines d'applications

La théorie des valeurs extrêmes trouve à s'appliquer dans de nombreux domaines tels qu'en fiabilité [64], en métallurgie [16] et en astrophysique [44]. Elle intéresse également les sciences de l'environnement, avec la modélisation de grands feux de forêts [2] ainsi que la climatologie [153] et la météorologie [31, 35, 85].

Dans un premier article, Einmahl et Magnus [66] proposèrent une application aux temps limites de records en athlétisme ; dans un second, ils s'intéressèrent plus particulièrement à l'estimation du temps minimal possible sur 100m [67].

Aarssen et de Haan [1] proposèrent des résultats afin de calculer l'âge limite possible de l'être humain. Pour d'autres exemples d'applications, se référer au livre de Reiss et Thomas [146]. Au sujet de la mise en garde d'une mauvaise utilisation de la théorie des valeurs extrêmes, on citera Bouleau [23].

Le domaine d'application historique reste l'hydrologie [49, 107, 156], notamment suite aux travaux de Jules Emile Gumbel en 1954 [108] et son ouvrage [109] en 1958. Plus récemment l'utilisation de la théorie des valeurs extrêmes a été vivement recommandée par le rapport Flood Study Report NERC [137] afin de modéliser les lois de probabilités des maxima annuels de précipitations et de crues. Ce rapport, établi en 1975, utilise et teste de nombreux outils statistiques issus de la théorie des valeurs extrêmes sur un très grand nombre de jeux de données de crues de rivières et de pluviométrie en Grande-Bretagne.

Deux autres principaux domaines d'applications sont l'actuariat afin de se prémunir contre des sinistres ayant un grand impact [18, 26, 148, 152] et la finance [70–73, 133].

"With globalisation increasing, you will see more crises. Our whole focus is on extremes now–what is the worst that can happen to you in any situation– because we never want to go through that again."

John Meriweather

Il existe également quelques articles de vulgarisation scientifique visant à introduire la problématique et les applications de la théorie des valeurs extrêmes. Citons notamment Matthews [131] qui donne une très bonne vue d'ensemble. En finance, se référer notamment aux articles de Cont [37, 38], mais également au livre de Taleb [160] et à son article [161].

Enfin, à l'usage de statisticiens intéressés par la théorie des valeurs extrêmes, on conseillera vivement la lecture des ouvrages de Coles [34] et Beirlant *et al.* [16].

Organisation de la thèse Cette thèse s'organise en trois chapitres.

- Le chapitre 1 présente un état de l'art en théorie des valeurs extrêmes, et plus particulièrement sur l'estimation des quantiles extrêmes qui est la problématique de cette thèse. Tout d'abord, on se limitera au cas univarié réel, où l'on présentera les principaux résultats et définitions de la théorie des valeurs extrêmes utiles dans nos travaux. Après avoir introduit le comportement du maximum d'un échantillon dans la Partie 1.1, on présentera les deux principaux outils servant à modéliser le comportement des valeurs extrêmes : la loi des valeurs extrêmes Partie 1.2 et la loi des excès Partie 1.3. On s'intéressera ensuite à la caractérisation des domaines d'attraction dans la Partie 1.4. Enfin on rappellera les différentes méthodes d'estimation de quantiles extrêmes Partie 1.5. Dans la dernière Partie 1.6 on se penchera sur le cas univarié réel en présence d'une covariable. On passera en revue la littérature sur l'estimation des quantiles extrêmes dans ce contexte de régression avec covariable qui a été peu abordé.

- Le chapitre 2, s'intéresse à une nouvelle famille de lois introduite par Gardes *et al.* [88]. Un estimateur des quantiles extrêmes y est proposé, mais il dépend d'un paramètre inconnu, le rendant inutile dans des situations pratiques. On commencera dans la Partie 2.1 par une introduction motivant nos travaux. On rappellera ensuite dans la Partie 2.2 le cadre d'étude et le modèle de Gardes *et al.* [88]. Puis l'on présentera nos contributions théoriques dans la Partie 2.3. Ces contributions consistent dans un premier temps à proposer un estimateur de ce paramètre inconnu. A la suite de quoi cela nous permettra d'estimer des quantiles extrêmes. Les lois asymptotiques de nos estimateurs sont établies et leur efficacité est illustrée dans la Partie 2.4 sur des données simulées et sur un jeu de données réelles. Ce dernier provient du rapport Flood Study Report NERC [137] et consiste en des crues de la rivière Nidd en Grande Bretagne. On finira par donner des perspectives dans la Partie 2.5. En dernier lieu, la Partie 2.6 exposera les preuves de nos résultats théoriques. Les contributions de ce chapitre ont fait l'objet de l'article El Methni *et al.* [69] paru en 2012.

- Au chapitre 3, on commencera dans la Partie 3.1 par une introduction sur la notion de mesure de risque où l'on motive nos travaux. On introduira ensuite dans la Partie 3.2 une nouvelle mesure de risque que l'on appellera le Conditional Tail Moment. Estimer le Conditional Tail Moment permet d'estimer toutes les mesures de risque basées sur les moments conditionnels telles que la Value-at-Risk, la Conditional Tail Expectation, la Conditional Value-at-Risk, la Conditional Tail Variance, ou la Conditional Tail Skewness. Ici, on s'intéressera à l'estimation de ces mesures de risque dans le cas de pertes extrêmes. On supposera aussi que la loi des pertes est à queue lourde et qu'elle dépend d'une covariable. Nos méthodes d'estimation combinent des méthodes

d'estimation non-paramétrique à noyau avec des méthodes issues de la statistique des valeurs extrêmes. Dans la Partie 3.3 on établira le comportement asymptotique de nos estimateurs. On illustrera dans la Partie 3.4 ces comportements sur des données simulées où l'on mettra en place une procédure nous permettant de choisir les paramètres de nos estimateurs. Enfin on appliquera cette procédure et nos estimateurs à un jeu de données réelles pluviométriques provenant de la région Cévennes-Vivarais se situant dans le sud de la France. On donnera quelques perspectives dans la Partie 3.5. Les preuves de nos résultats théroriques se trouvent dans la dernière Partie 3.6. Les contributions de ce chapitre ont fait l'objet du rapport de recherche El Methni *et al.* [68] soumis pour publication.

La conclusion de ce mémoire exposera les perspectives à donner à nos travaux.

Chapitre 1

Présentation de la théorie des valeurs extrêmes univariées

Résumé

Dans ce chapitre, on regroupe des définitions et des résultats sur la théorie des valeurs extrêmes dans le cas univarié réel. Après avoir introduit le comportement du maximum dans la Partie 1.1, on présentera les deux principaux outils servant à modéliser le comportement des valeurs extrêmes d'un échantillon : la loi des valeurs extrêmes Partie 1.2 et la loi des excès Partie 1.3. On s'intéressera ensuite à la caractérisation des domaines d'attraction dans la Partie 1.4. Enfin, on rappelera les différentes méthodes d'estimation de quantiles extrêmes Partie 1.5. Dans la dernière Partie 1.6, on présentera la théorie des valeurs extrêmes univariées en présence d'une covariable et en particulier l'estimation de quantiles extrêmes conditionnels.

Sommaire

1.1 Introduction au comportement du maximum d'un échantillon . 9
1.2 La loi des valeurs extrêmes 14
1.3 La loi des excès 19
1.4 Caractérisation des domaines d'attraction 22
1.4.1 Fonctions à variations régulières 22
1.4.2 Domaine d'attraction de Fréchet 24
1.4.3 Domaine d'attraction de Weibull 25
1.4.4 Domaine d'attraction de Gumbel 26
1.5 Estimation de quantiles extrêmes 29
1.5.1 L'approche par la loi des valeurs extrêmes 30
1.5.2 L'approche par la loi des excès 34
1.5.3 L'approche semi-paramétrique 40
1.6 Estimation de quantiles conditionnels extrêmes 50
1.6.1 L'approche paramétrique 52
1.6.2 L'approche semi-paramétrique 52
1.6.3 L'approche non-paramétrique 52

1.1 Introduction au comportement du maximum d'un échantillon

"There is always going to be an element of doubt, as one is extrapolating into areas one doesn't know about. But what extreme value theory is doing is making the best use of whatever data you do have about extreme phenomena."

Richard Smith

La théorie des valeurs extrêmes a pour but d'étudier et de caractériser le comportement des valeurs extrêmes d'un échantillon de variables aléatoires. On souhaite estimer des petites probabilités ou des quantités dont la probabilité d'observation est très faible, c'est-à-dire proche de zéro. Ces quantités sont appelées quantiles ; on parle de quantile extrême lorsque l'ordre du quantile (probabilité d'observation) converge vers zéro quand la taille de l'échantillon tend vers l'infini.

Ces quantiles extrêmes se situent dans les queues de distributions des lois de probabilité. Ainsi, à l'inverse de l'approche statistique classique qui s'intéresse au comportement moyen et à la variabilité des phénomènes autour de la moyenne, on s'intéresse ici au comportement des queues de distributions. Que pouvons-nous dire du comportement du maximum ou du minimum d'un phénomène si la loi de probabilité qui le régit n'est pas connue ?

Comme évoqué dans l'introduction générale, un exemple d'application historique est l'hydrologie. En effet, si l'on dispose d'un échantillon de hauteurs d'eau annuelles d'un cours d'eau, on peut se poser plusieurs questions intéressantes sur les événements extrêmes. En particulier :

(1) Quelle est la hauteur d'eau qui est atteinte ou dépassée pour une faible probabilité donnée ?
(2) Pour une "grande" hauteur d'eau fixée, qu'elle est la probabilité d'observer une hauteur d'eau qui lui sera supérieure ?

La théorie des valeurs extrêmes permet de répondre à ce type de questions.

Plaçons-nous dans un cadre statistique et considérons n variables aléatoires réelles $Y_1, \ldots, Y_n$ indépendantes et identiquement distribuées (*i.i.d.*) de fonction de répartition $F(y) = \mathbb{P}(Y \leq y)$. Rangeons ces variables aléatoires par ordre croissant, on notera dans la suite l'échantillon ordonné :

$$Y_{1,n} \leq \cdots \leq Y_{n,n}.$$

Remarquons que les statistiques d'ordre ne sont pas indépendantes par définition. Deux statistiques d'ordre sont particulièrement intéressantes pour l'étude des événements extrêmes : ce sont les deux statistiques extrêmes, autrement dit le minimum de l'échantillon $Y_{1,n}$ et le maximum $Y_{n,n}$. On notera qu'il est très facile de passer de l'un à l'autre à l'aide de la relation :

$$Y_{1,n} = -\max(-Y_1, \ldots, -Y_n).$$

Ainsi, la suite des développements du présent chapitre se concentrera sur l'étude du maximum. A partir de l'information apportée par ces variables aléatoires, on souhaite estimer des quantiles extrêmes. Soit $\overline{F}(y) = 1 - F(y)$ la fonction de survie, rappelons la définition d'un quantile.

Définition 1.1. *Le quantile d'ordre* $1-\alpha$ *de la fonction de répartition* F *est défini par :*

$$q(\alpha) := \overline{F}^{\leftarrow}(\alpha) = \inf\{y : \overline{F}(y) \leq \alpha\} \quad \textit{avec} \quad \alpha \in]0, 1[,$$

où $\overline{F}^{\leftarrow}$ est l'inverse généralisée de $\overline{F}$. Par convention $\inf\{\varnothing\} = \infty$. Notons que l'inverse généralisée d'une fonction coïncide avec l'inverse classique lorsque la fonction est continue.

Définition 1.2. *Le quantile extrême d'ordre* $1 - \alpha_n$ *de la fonction de répartition* F *est défini par :*

$$q(\alpha_n) := \overline{F}^{\leftarrow}(\alpha_n) \quad \textit{avec} \quad \alpha_n \to 0 \quad \textit{quand} \quad n \to \infty.$$

En résumé, un quantile sera dit extrême si l'on remplace son ordre α par une suite $\alpha_n \to 0$ quand $n \to \infty$. Le fait que l'ordre $\alpha_n \to 0$ quand $n \to \infty$ indique que l'information la plus importante pour estimer des quantiles extrêmes est contenue dans la queue de distribution (voir Figure 1.1).

Statistiquement parlant, la question (1) se rapporte donc à l'estimation d'un quantile extrême ou niveau de retour en hydrologie. Quant à la question (2), elle se rapporte à l'estimation d'une "petite probabilité", ou de façon équivalente en hydrologie à une période de retour.

Afin d'expliquer les notions de niveau de retour et de période de retour, considérons tout simplement que l'on dispose d'un dé à 6 faces parfaitement équilibré. Cet exemple banal est un cas classique en théorie des probabilités d'une variable aléatoire discrète suivant une loi Uniforme sur [1 : 6]. Si le dé n'est pas pipé, alors la probabilité de tomber sur une des 6 faces est de 1/6. Prenons par exemple la face relative au chiffre 2. On peut alors s'attendre à obtenir cette face au bout d'un certain temps (certain nombre de lancers) ou plus précisément on peut s'attendre à l'obtenir en moyenne tous les $T = 6$ lancers. T est donc la période de retour.

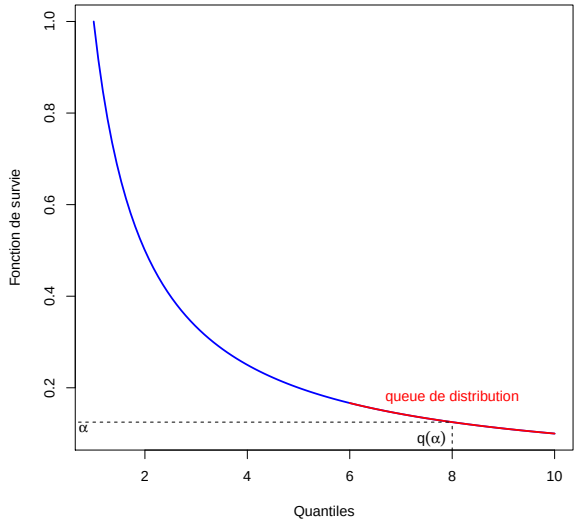


FIGURE 1.1: Quantiles extrêmes et queue de distribution.

Dans le cas d'une variable aléatoire continue Y, la probabilité que $Y = y$ est nulle : $\mathbb{P}(Y = y) = 0$. Cependant il y a une probabilité $\overline{F}(y) = \mathbb{P}(Y \geq y)$ que la variable aléatoire Y soit égale ou supérieure à y. On a alors la fonction appelée période de retour donnée par :

$$T(y) = \frac{1}{\overline{F}(y)}.$$

Elle représente le nombre d'observations tel que, en moyenne, il y ait une observation égale ou supérieure à y. Il est évident que la période de retour augmente lorsque y augmente. On peut alors définir la fonction niveau de retour comme l'inverse de la période de retour :

$$y(T) = \overline{F}^{\leftarrow}\left(\frac{1}{T}\right) = q\left(\frac{1}{T}\right).$$

Le niveau de retour représente le niveau (d'eau par exemple) qui sera atteint ou dépassé pour une certaine période de retour (de probabilité : $1/T$). Un niveau de retour à 1 an revient à calculer le quantile d'ordre $1/365.25$. Il est donc plus simple de s'exprimer en termes de niveau de retour à un an que de parler de quantile d'ordre 0.2737851%. Ainsi on peut mieux caractériser ce qu'est un événement extrême :

Un événement extrême est un événement dont le temps de retour est grand.

Quelle est alors la probabilité d'observer un événement extrême ayant une valeur plus grande que le maximum de l'échantillon ? Autrement dit, quelle est la probabilité que le quantile extrême soit plus grand que le maximum de l'échantillon ?

Afin de simplifier on supposera F continue. Comme les variables aléatoires sont indépendantes et identiquement distribuées, alors lorsque $\alpha_n \to 0$ on a :

$$\begin{aligned}
\mathbb{P}(Y_{n,n} < q(\alpha_n)) &= \mathbb{P}\left(\bigcap_{i=1}^{n}\{Y_i \leq q(\alpha_n)\}\right) \\
&= \prod_{i=1}^{n}\mathbb{P}(Y_i \leq q(\alpha_n)) \\
&= F^n(q(\alpha_n)) \qquad (1.1) \\
&= (1-\alpha_n)^n \\
&= \exp(n\log(1-\alpha_n)) \\
&= \exp(-n\alpha_n(1+o(1)).
\end{aligned}$$

La probabilité que le quantile extrême soit plus grand que le maximum de l'échantillon dépend donc du comportement asymptotique de $n\alpha_n$. Ainsi, lorsque l'on souhaite estimer des quantiles extrêmes, on doit faire la distinction entre deux cas qui sont fonction de la vitesse de convergence de α_n vers 0.

Premier cas : Si $n\alpha_n \to \infty$ alors, $\mathbb{P}(Y_{n,n} < q(\alpha_n)) \to 0$.

Dans ce cas, on cherche à estimer un quantile qui se trouve avec une grande probabilité dans l'échantillon. On est dans le cas où α_n converge lentement vers 0, autrement dit cela revient à supposer que le quantile $q(\alpha_n)$ ne tend pas trop vite vers l'infini quand $n \to \infty$. Dans une telle situation, l'estimation du quantile extrême requiert une interpolation à l'intérieur de l'échantillon. Un estimateur naturel de ce quantile est la $\lfloor n\alpha_n \rfloor$ ième plus grande observation de l'échantillon, autrement dit la statistique d'ordre supérieur $Y_{n-\lfloor n\alpha_n \rfloor+1,n}$ (où $\lfloor . \rfloor$ est la fonction partie entière). Cet estimateur est asymptotiquement gaussien (voir [50]). Remarquons que la condition $n\alpha_n \to \infty$ correspond à un temps de retour $T_n = 1/\alpha_n$ petit devant n

Second cas : Si $n\alpha_n \to 0$ alors, $\mathbb{P}(Y_{n,n} < q(\alpha_n)) \to 1$.

On cherche alors à estimer un quantile qui se trouve avec une grande probabilité en dehors de l'échantillon. Dans ce cas, l'estimateur de $q(\alpha_n)$ ne peut être obtenu en inversant simplement la fonction de répartition empirique :

$$\widehat{F}_n(y) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}\{Y_i \leq y\},$$

car $\widehat{F}_n(y) = 1$ pour $y \geq Y_{n,n}$. On est dans le cas où α_n converge rapidement vers 0, autrement dit cela revient à supposer que le quantile $q(\alpha_n)$ tend suffisament vite vers

l'infini quand $n \to \infty$. Remarquons que la condition $n\alpha_n \to 0$ correspond à un temps de retour T_n grand devant n. Cela nécessite d'extrapoler les résultats de l'échantillon là où il n'y a pas de données observées ; c'est ce qui se produit dans de nombreux domaines d'application, parmi lesquels citons la fiabilité (Ditlevsen [64]), l'hydrologie (Smith [156]), la finance (Embrechts *et al.* [71]), les assurances (Beirlant et Teugels [18] et Brodin et Rootzén [26]), ainsi que la climatologie (Rootzén et Tajvidi [153]).

"*Many real life questions require extrapolation, but since no data or only few has been observed – as by definition extreme events are rare – essential estimations are more often based on feeling than on fact. Extreme value theory is a branch of statistics that deals with such rare situations and that gives a scientific alternative to pure guesswork.*

Chavez–Demoulin et Armin Roehrl.

Le résultat (1.1) nous indique que pour d'estimer des quantiles extrêmes, on a besoin de connaître le comportement du maximum de l'échantillon. Ce comportement est caractérisé par la fonction de répartition du maximum que l'on notera $F_{Y_{n,n}}$. Toujours d'après le résultat (1.1), on a $F_{Y_{n,n}}(y) = F^n(y)$. La loi F du phénomène étant inconnue en pratique, le comportement de F^n sera encore plus difficile à étudier.

On peut cependant remarquer que :

$$\lim_{n\to\infty} F_{Y_{n,n}}(y) = \lim_{n\to\infty} [F(y)]^n = \mathbb{I}\{y \geq y_F\} = \begin{cases} 1 & \text{si} \quad y \geq y_F \\ 0 & \text{si} \quad y < y_F \end{cases} \tag{1.2}$$

où

$$y_F = \sup\{y \in \mathbb{R}, F(y) < 1\},$$

est le point terminal de la loi F avec la convention $\sup\{\varnothing\} = \infty$. Le point terminal y_F représente la borne supérieure du support de la loi. Le résultat (1.2) nous indique que la distribution du maximum $Y_{n,n}$ est une loi dégénérée. Ce résultat fournit très peu d'informations sur le comportement de $Y_{n,n}$. On aimerait obtenir une loi non dégénérée pour le maximum.

L'idée est de procéder à une transformation. La plus connue en statistique est la normalisation illustrée à travers l'exemple du théorème central limite qui, après normalisation, donne la loi asymptotique (non dégénérée) de la moyenne de n variables aléatoires.

Dans la Partie 1.2 on commencera par énoncer le résultat fondamental de la théorie des valeurs extrêmes univariées connu sous le nom de théorème de Fisher-Tippet-Gnedenko. Il établit la loi asymptotique du maximum de l'échantillon $Y_{n,n}$ convenablement renormalisé.

1.2 La loi des valeurs extrêmes

Historiquement, l'étude de la loi de probabilité du maximum d'un échantillon de n variables a été la première approche pour décrire les événements extrêmes. Fisher et Tippet [77] en 1928 ont, les premiers, déduit de manière heuristique les lois limites possibles pour le maximum d'une suite de variables aléatoires indépendantes et de même loi, avant que Gnedenko [97] en 1943 n'obtienne rigoureusement la convergence, dont la preuve fut simplifiée par de Haan [48] en 1976. Les travaux de Von Mises [164] en 1936 et Jenkinson [122] en 1955 ont permis de donner une forme unifiée à ce résultat. Les applications ont commencé suite aux travaux de Gumbel [108] en 1954, en particulier en hydrologie.

Le théorème ci-dessous est fondamental en théorie des valeurs extrêmes car il établit la loi asymptotique du maximum $Y_{n,n}$ convenablement normalisé d'un échantillon.

Théorème 1.3. *Soit $(Y_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes et identiquement distribuées de fonction de répartition F. S'il existe deux suites normalisantes réelles $(a_n)_{n\geq 1} > 0$ et $(b_n)_{n\geq 1} \in \mathbb{R}$ et une loi non-dégénérée $\mathcal{H}_\gamma$ telle que :*

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{Y_{n,n} - b_n}{a_n} \leq y\right) = \lim_{n\to\infty} F^n(a_n y + b_n) = \mathcal{H}_\gamma(y), \tag{1.3}$$

alors à une translation et un changement d'échelle près on a

$$\mathcal{H}_\gamma(y) = \exp\left(-\left(1+\gamma y\right)_+^{-1/\gamma}\right), \tag{1.4}$$

où $\gamma \in \mathbb{R}$ et $z_+ = \max(0, z)$.

Le cas $\gamma = 0$ dans l'équation (1.4) peut être vu comme le cas limite lorsque $\gamma \to 0$. On retrouve alors la loi de Gumbel ayant pour fonction de répartition :

$$\mathcal{H}_0(y) = \exp\left(-\exp\left(-y\right)\right).$$

Le comportement limite du maximum normalisé est ainsi décrit par la fonction de répartition $\mathcal{H}_\gamma$ pour la plus grande partie des lois usuelles. $\mathcal{H}_\gamma$ est appelée fonction de répartition de la loi des valeurs extrêmes, en anglais "Generalized Extreme Value Distribution" notée GEV. Une preuve détaillée de ce théorème peut être trouvée dans Resnick [147] et dans Embrechts *et al.* [71].

Si l'on souhaite faire un parallèle avec le théorème central limite, la suite (a_n) jouerait le rôle d'un paramètre d'échelle ou de dispersion et la suite (b_n) celui d'un paramètre de position. Ces suites ne sont pas uniques. Pour plus d'exemples de suites de normalisation de chaque loi se référer à Embrechts *et al.* [71] page 145.

L'unification du comportement du maximum en une seule fonction de répartition facilite grandement l'étude du comportement du maximum. Cette loi dépend du seul paramètre de forme γ appelé indice des valeurs extrêmes ou indice de queue. Comme on le verra tout au long de cette thèse, γ est le paramètre clé de toute la théorie des valeurs extrêmes. L'estimation de γ nous fournira le comportement de la queue de distribution. En effet, selon son signe, on distingue trois domaines d'attraction dont quelques densités ont été représentées pour γ fixé sur la Figure 1.2.

- si $\gamma > 0$, on dit que F appartient au domaine d'attraction de Fréchet [79], que l'on notera $\mathcal{D}$(Fréchet). Il contient les lois dont la fonction de survie est à décroissance polynomiale, *i.e.* les lois à queues lourdes ou lois de type Pareto. Les lois de ce domaine ont un point terminal y_F infini.
- si $\gamma < 0$, on dit que F appartient au domaine d'attraction de Weibull, que l'on notera $\mathcal{D}$(Weibull). Toutes les lois de ce domaine d'attraction ont un point terminal y_F fini.
- si $\gamma = 0$, on dit que F est dans le domaine d'attraction de Gumbel, que l'on notera $\mathcal{D}$(Gumbel). Il contient les lois dont la fonction de survie est à décroissance exponentielle, *i.e.* les lois à queues légères.

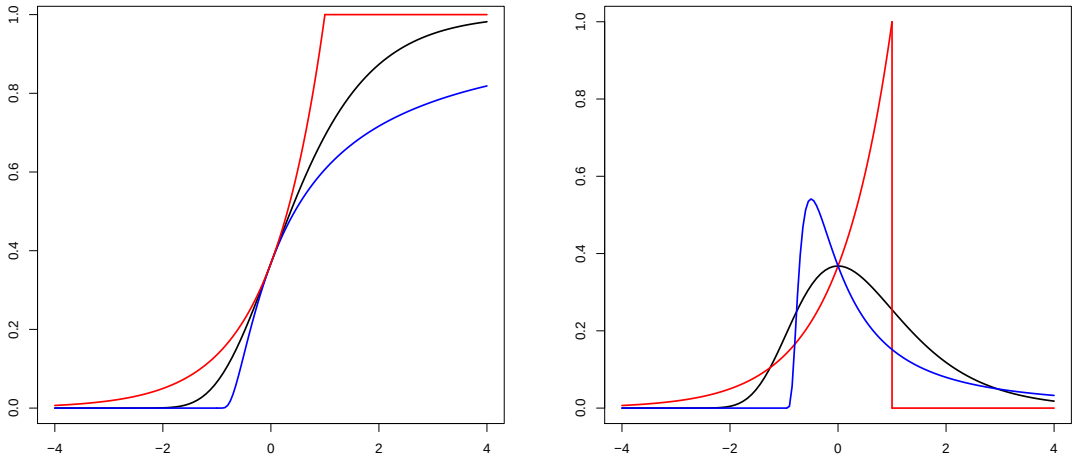


FIGURE 1.2: A gauche : $\mathcal{H}_\gamma$. A droite : les densités associées à la loi des valeurs extrêmes (noir : $\gamma = 0$, bleu : $\gamma = 1$ et rouge : $\gamma = -1$).

Un classement de nombreuses lois par domaine d'attraction est disponible dans Embrechts *et al.* [71] page 145. On donne une liste non exhaustive de l'appartenance des lois à leur domaine d'attraction dans le Tableau 1.1.

Fréchet ($\gamma > 0$)	Gumbel ($\gamma = 0$)	Weibull ($\gamma < 0$)
Pareto	Normale	Uniforme
Student	Exponentielle	Beta
Burr	Log-normale	ReverseBurr
Chi-deux	Gamma	
Fréchet	Weibull	
Log-gamma	Benktander-type-I	
Log-logistique	Benktander-type-II	
Cauchy	Logistique	
	Gumbel	

TABLE 1.1: Quelques lois et leurs domaines d'attraction

Illustrons le Théorème 1.3 à travers l'exemple d'une variable aléatoire Y suivant une loi exponentielle de paramètre 1. Sa fonction de répartition est donnée par :

$$F(y) = \begin{cases} 1 - e^{-y} & \text{si} \quad y \geq 0, \\ 0 & \text{si} \quad y < 0. \end{cases}$$

Le support de la loi étant $\mathbb{R}$ on a $y_F = +\infty$ d'où $Y_{n,n} \xrightarrow{\mathbb{P}} \infty$. Effectuons la normalisation suivante :

$$\begin{aligned} \mathbb{P}\left(\frac{Y_{n,n} - \log(n)}{1} \leq y\right) &= \mathbb{P}(Y_{n,n} \leq y + \log(n)) \\ &= (F(y + \log(n))^n \\ &= (1 - \exp(-y - \log(n)))^n \\ &= \left(1 - \frac{\exp(-y)}{n}\right)^n \\ &\to \exp(-\exp(-y)) \end{aligned}$$

quand $n \to \infty$.

Ainsi convenablement normalisé il apparait que $Y_{n,n}$ converge en loi vers une loi de Gumbel. La convergence du maximum normalisé d'une loi exponentielle vers une loi de Gumbel est illustrée dans la suite de figures données ci-après.

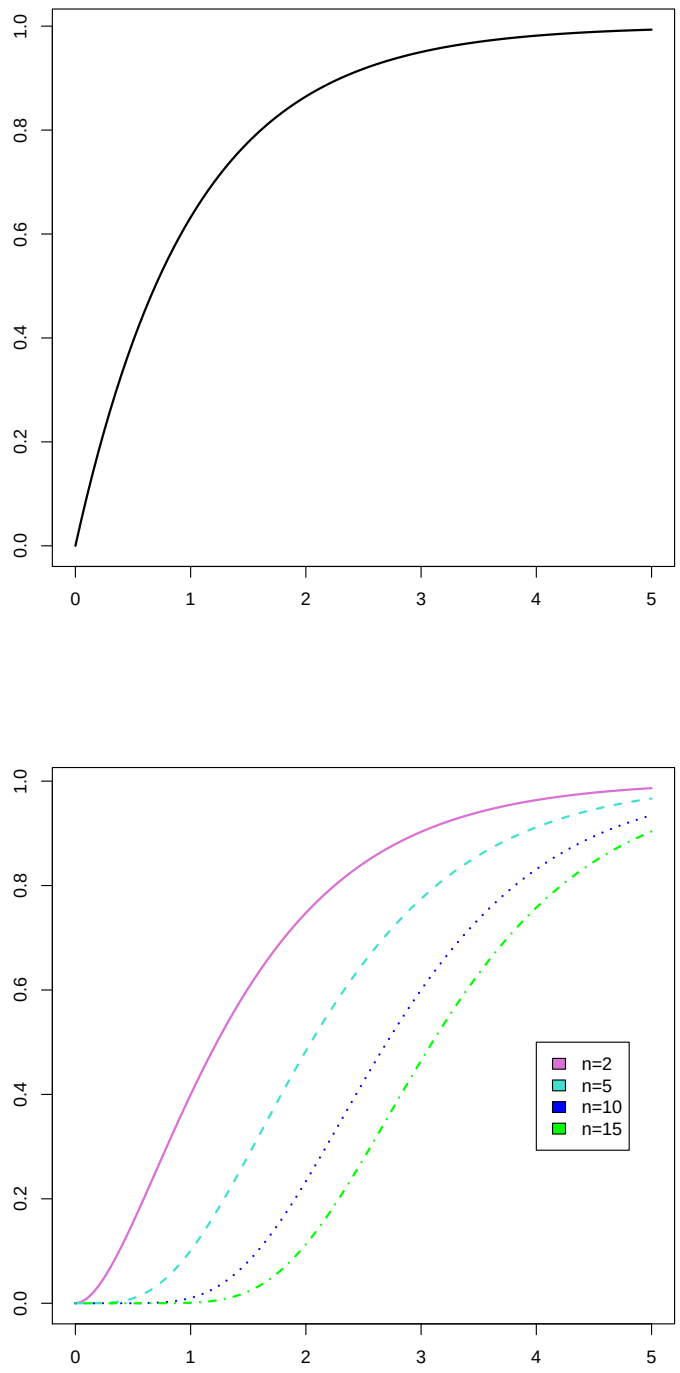


FIGURE 1.3: En haut : fonction de répartition de la loi exponentielle de moyenne 1. En bas : différentes puissances de fonctions de répartitions de la loi exponentielle de moyenne 1.

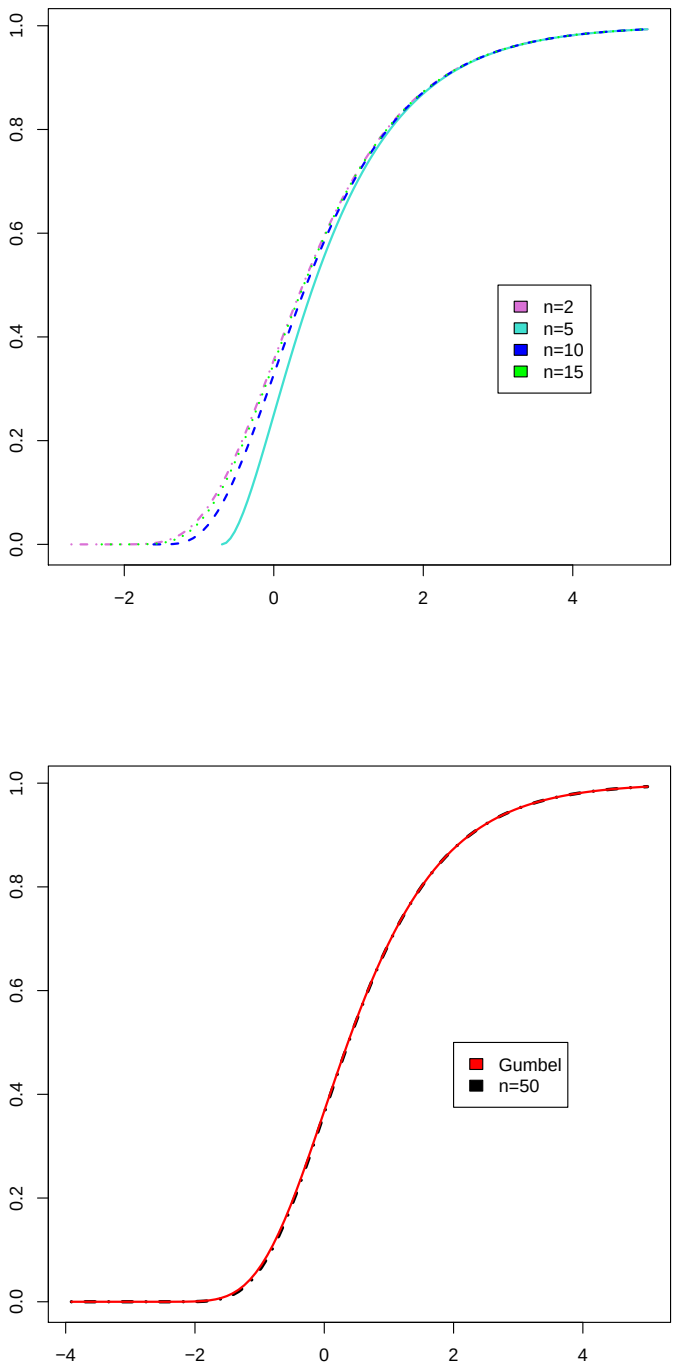


FIGURE 1.4: En haut : différentes puissances de fonctions de répartitions normalisées de la loi exponentielle de moyenne 1. En bas : convergence de la loi du maximum normalisé d'une loi exponentielle de moyenne 1 vers une loi de Gumbel.

Il est irréaliste de croire que seul le maximum de l'échantillon permet de modéliser le comportement des valeurs extrêmes. Les autres grandes valeurs de l'échantillon contiennent elles aussi de l'information sur la queue de distribution. L'approche par dépassements de seuil est une alternative à la loi GEV dans la modélisation du comportement du maximum d'un échantillon se basant sur les "grandes valeurs" de l'échantillon.

Elle se base sur le résultat mathématique (que nous détaillerons dans la partie suivante), affirmant qu'il existe une équivalence entre la loi GEV et la loi des dépassements de seuil aussi appelée loi des excès. La partie 1.3 est dédié au comportement limite des excès au-delà d'un seuil assez grand.

1.3 La loi des excès

L'approche par dépassements de seuil, en anglais "Peaks-Over-Threshold approach" notée POT, repose sur l'utilisation des statistiques d'ordre supérieur de l'échantillon. Elle consiste à ne conserver que les observations dépassant un certain seuil. L'excès au-delà du seuil est défini comme l'écart entre l'observation et le seuil.

Plus précisément, soit un échantillon de variables aléatoires *i.i.d.* $Y_1, \ldots, Y_n$. Soit u un seuil fixé (non aléatoire) tel que $u < y_F$. Considérons les N_u observations $Y_{i_1}, \ldots, Y_{i_{N_u}}$ dépassant le seuil u. On appelle excès au-delà du seuil u les $Z_j := Y_{i_j} - u$, où $j = 1, \ldots, N_u$ voir Figure 1.5 ci-dessous :

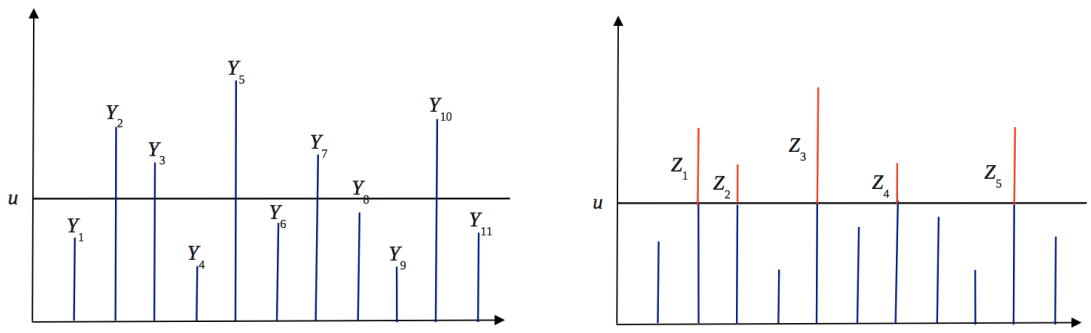


FIGURE 1.5: Excès au delà du seuil u.

On notera F_u la fonction de répartition de l'excès Z au delà du seuil u. La loi des excès est celle de variables aléatoires *i.i.d.* admettant pour fonction de répartition $F_u(y) = \mathbb{P}(Z \leq y | Y > u)$ représentant la probabilité que la variable aléatoire Y ne dépasse pas le seuil u de au moins une quantité y sachant qu'elle dépasse u. F_u décrit ainsi la loi de Z sachant que $Y > u$. On peut la réécrire en fonction de F à l'aide du résultat suivant.

On a pour $y \geq 0$:

$$F_u(y) = \mathbb{P}(Z \leq y | Y > u) = \mathbb{P}(Y - u \leq y | Y > u) = \frac{F(u+y) - F(u)}{1 - F(u)}, \tag{1.5}$$

ou de manière équivalente pour la fonction de survie :

$$\overline{F}_u(y) := \mathbb{P}(Z > y | Y > u) = 1 - F_u(y) = \frac{\overline{F}(u+y)}{\overline{F}(u)}. \tag{1.6}$$

L'idée de l'existence d'une équivalence entre la loi GEV et la loi des excès est la suivante. A l'aide du résultat donné dans le Théorème 1.3 pour n assez grand on a,

$$F^n(u) \approx \exp\left(-\left(1 + \gamma\left(\frac{u - b_n}{a_n}\right)\right)^{-1/\gamma}\right),$$

avec $a_n > 0$ et $(b_n, \gamma) \in \mathbb{R}^2$. Ainsi,

$$n \log(F(u)) \approx -\left(1 + \gamma\left(\frac{u - b_n}{a_n}\right)\right)^{-1/\gamma}. \tag{1.7}$$

Si u est assez grand alors un développement limité donne :

$$\log(F(u)) \approx -(1 - F(u)).$$

En remplaçant dans l'expression (1.7) on obtient pour u assez grand

$$1 - F(u) \approx \frac{1}{n}\left(1 + \gamma\left(\frac{u - b_n}{a_n}\right)\right)^{-1/\gamma}.$$

De même pour $y > 0$ on a,

$$1 - F(u+y) \approx \frac{1}{n}\left(1 + \gamma\left(\frac{u + y - b_n}{a_n}\right)\right)^{-1/\gamma}.$$

En remplaçant dans l'expression (1.6) on obtient :

$$F_u(y) \approx 1 - \left(1 + \gamma\frac{y}{\sigma}\right)^{-1/\gamma},$$

avec

$$\sigma = a_n + \gamma(u - b_n). \tag{1.8}$$

Les travaux de Balkema et de Haan [8] et de Pickands [141] donnent un résultat très précis sur l'approximation de cette fonction de survie (ou de manière équivalente de la fonction de répartition) lorsque le seuil u est proche du point terminal y_F.

Théorème 1.4. *(Balkema et de Haan [8], Pickands [141]) F appartient au domaine d'attraction de $\mathcal{H}_\gamma$ si et seulement si il existe $\sigma > 0$ et $\gamma \in \mathbb{R}$ tels que la loi des excès F_u peut être uniformément approchée par une loi de Pareto généralisée notée $\mathcal{G}_{\gamma,\sigma}$, i.e.*

$$\lim_{u \to y_F} \sup_{y \in]0, y_F - u[} |F_u(y) - \mathcal{G}_{\gamma,\sigma}(y)| = 0$$

où

$$\mathcal{G}_{\gamma,\sigma}(y) = 1 - \left(1 + \gamma \frac{y}{\sigma}\right)^{-1/\gamma} \tag{1.9}$$

définie pour $y \in \mathbb{R}^+$ si $\gamma \geq 0$ ou $[0, -\sigma/\gamma[$ si $\gamma < 0$.

Le cas $\gamma = 0$ dans l'expression (1.9) peut être vu comme le cas limite lorsque que $\gamma \to 0$. On a alors :

$$\mathcal{G}_{0,\sigma}(y) = 1 - \exp\left(-\frac{y}{\sigma}\right), \quad y \geq 0.$$

On retrouve une loi exponentielle de paramètre $1/\sigma$. Notons aussi que $\mathcal{G}_{-1,\sigma}(.)$ correspond à la loi uniforme sur $[0, \sigma]$. L'existence des moments de la GPD est liée à la valeur de γ. En effet, le moment d'ordre m existe si et seulement si $\gamma < 1/m$.

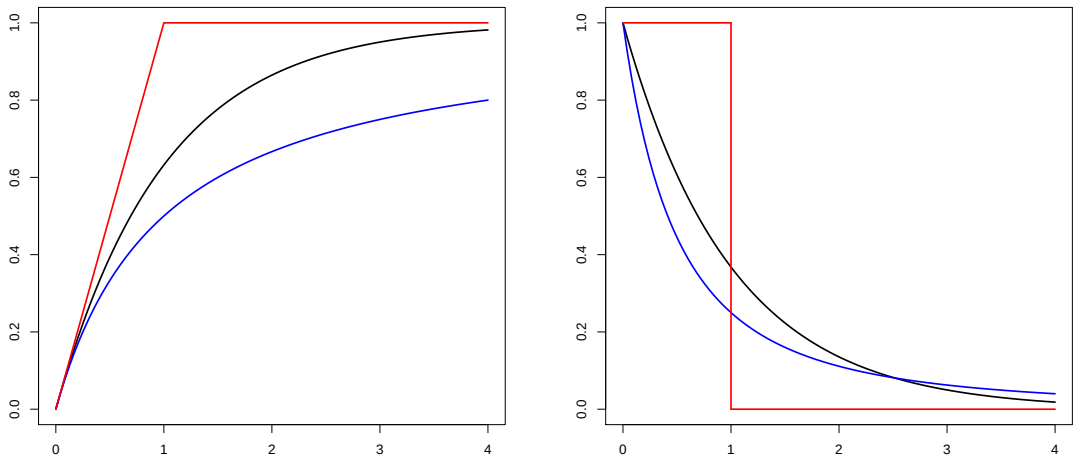


FIGURE 1.6: A gauche : $\mathcal{G}_{\gamma,1}$. A droite : les densités associées à la loi de Pareto généralisée (noir : $\gamma = 0$, bleu : $\gamma = 1$ et rouge : $\gamma = -1$).

Le Théorème 1.4 établit l'équivalence en loi du maximum (convenablement normalisé) d'un échantillon vers une loi des valeurs extrêmes $\mathcal{H}_\gamma$ et la convergence en loi des excès au-delà d'un seuil vers une loi de Pareto généralisée $\mathcal{G}_{\gamma,\sigma}$, lorsque le seuil tend vers la limite supérieure du support de F.

Ainsi les paramètres de la loi de Pareto généralisée, en anglais "Generalized Pareto Distribution" notée GPD (voir expression (1.9)), ont un lien avec ceux de la GEV (voir

expression (1.4)). Le paramètre d'échelle σ de la GPD est tel que $\sigma = a_n + \gamma(u - b_n)$ (voir expression (1.8)). Le paramètre de forme γ est le même que dans la loi GEV, il caractérise donc la lourdeur de la queue de la loi. La dualité qu'il existe entre la GPD et la GEV met en évidence le rôle prépondérant de γ dans le comportement des valeurs extrêmes.

En résumé le Théorème 1.4 suggère donc d'approcher la loi des excès de l'échantillon par une distribution GPD lorsque le seuil u est choisi suffisament grand.

L'idée est résumée dans l'article Mandelbrot's Extremism de Beirlant *et al.* [17]

Excess postulate : *Up to scale, the distribution of excesses over a high threshold stabilizes as the threshold u grows indefinitely large.*

Une application de la GPD à la prédiction de vents extrêmes est réalisée dans l'article de Brodin et Rootzén [26].

Dans la partie suivante on donnera les conditions pour que F appartienne à l'un des trois domaines d'attraction ainsi que ses suites de normalisation (a_n) et (b_n). On donnera de plus une caractérisation de la fonction quantile dans chacun des domaines. Ces conditions étant basées sur la notion de fonctions à variations régulières, on commencera par rappeler la définition de ces fonctions et quelques unes de leurs propriétés.

1.4 Caractérisation des domaines d'attraction

On définit les notions de fonctions à variations régulières et de fonctions à variations lentes qui nous seront utiles par la suite. Pour plus de détails, se référer à Bingham *et al.* [21] où de nombreux résultats sur les fonctions à variations régulières sont donnés.

1.4.1 Fonctions à variations régulières

Commençons par rappeler la définition d'une fonction à variations régulières.

Définition 1.5. *Une fonction $U : \mathbb{R}_+ \to \mathbb{R}_+$ mesurable au sens de Lebesgue, est à variations régulières à l'infini si et seulement si, il existe un réel δ tel que pour tout $\lambda > 0$ on a :*

$$\lim_{y\to\infty} \frac{U(\lambda y)}{U(y)} = \lambda^{\delta}. \tag{1.10}$$

On dit que la fonction U est à variations régulières d'indice $\delta \in \mathbb{R}$ à l'infini, et on notera par la suite $U \in \mathcal{RV}_{\delta}$.

Le lemme suivant nous renseigne sur l'inverse d'une fonction à variations régulières.

Lemme 1.6. *Inverse d'une fonction à variations régulières :*
- *Si U est à variations régulières d'indice $\delta > 0$, alors $U^{\leftarrow}(y)$ est à variations régulières d'indice $1/\delta$.*
- *Si U est à variations régulières d'indice $\delta < 0$, alors $U^{\leftarrow}(1/y)$ est à variations régulières d'indice $-1/\delta$.*

Pour une preuve du Lemme 1.6, on pourra se référer au Théorème 1.5.12 de l'ouvrage de Bingham *et al.* [21] ou à la Proposition 2.6 du livre de Resnick [147].

La notion de fonctions à variations régulières à l'infini est liée à celle des fonctions à variations lentes à l'infini.

Fonctions à variations lentes

Commençons par rappeler la définition d'une fonction à variations lentes à l'infini.

Définition 1.7. *Une fonction $\ell : \mathbb{R}_+ \to \mathbb{R}_+$ mesurable au sens de Lebesgue, est à variations lentes à l'infini si et seulement si, pour tout $\lambda > 0$ on a :*

$$\lim_{y\to\infty} \frac{\ell(\lambda y)}{\ell(y)} = 1. \tag{1.11}$$

Remarquons qu'une fonction à variations régulières à l'infini d'indice $\delta = 0$, est une fonction à variations lentes à l'infini (voir expression (1.10)), et on notera par la suite $\ell \in \mathcal{RV}_0$. On peut facilement montrer que toute fonction à variations régulières d'indice $\delta \in \mathbb{R}$ peut s'écrire sous la forme :

$$U(y) = y^{\delta}\ell(y), \quad \text{avec} \quad \ell \in \mathcal{RV}_0.$$

Ce résultat montre que l'étude des fonctions à variations régulières à l'infini se ramène à l'étude des fonctions à variations lentes à l'infini. On peut citer quelques exemples de fonctions à variations lentes à l'infini, par exemple :

- les fonctions possédant une limite strictement positive à l'infini,
- les fonctions de la forme $\ell : y \to |\log(y)|^d$ où $d \in \mathbb{R}$,
- les fonctions de la forme $\ell : y \to \exp(\log(y)^{\eta})$ où $\eta < 1$,
- les fonctions ℓ telles que :

$$\exists M > 0, \quad \forall y \geq M, \quad \ell(y) = a + by^{-c}(1 + o(1)),$$

  avec $(a, c) \in \mathbb{R}_+^2$ et $b \in \mathbb{R}$. L'ensemble de ces fonctions ℓ est appelé classe de Hall [111].

Un résultat intéressant à signaler et utile dans la suite est la proposition suivante.

Proposition 1.8. *(voir [21, Propriété 1.3.6]) Pour toute fonction à variations lentes à l'infini ℓ on a :*

$$\lim_{y\to\infty} \frac{\log(\ell(y))}{\log(y)} = 0.$$

On peut caractériser les fonctions à variations lentes de manière plus précise à l'aide du théorème de Karamata.

Théorème 1.9. *Toute fonction à variations lentes à l'infini ℓ peut s'écrire sous la forme :*

$$\ell(y) = c(y) \exp\left\{\int_1^y \frac{\varepsilon(u)}{u} du\right\}, \tag{1.12}$$

où $c > 0$ et ε sont deux fonctions mesurables telles que :

$$\lim_{y\to\infty} c(y) = c > 0 \quad \textit{et} \quad \lim_{y\to\infty} \varepsilon(y) = 0.$$

Cette représentation des fonctions à variations lentes est connue sous le nom de *représentation de Karamata* (voir Bingham *et al.* [21, Théorème 1.3.1]).

De plus, si la fonction c est constante, la fonction ℓ est dite normalisée. Si ℓ est normalisée alors d'après l'expression (1.12), ℓ est dérivable de dérivée ℓ' avec pour tout $y > 0$,

$$\ell'(y) = \frac{\varepsilon(y)\ell(y)}{y}.$$

En particulier on a :

$$\lim_{y\to\infty} \frac{y\ell'(y)}{\ell(y)} = 0.$$

A l'aide des différents résultats présentés sur les fonctions à variations régulières et variations lentes à l'infini, on va pouvoir caractériser les différents domaines d'attraction. Sachant la distribution F, on voudrait connaître son domaine d'attraction et ses constantes de normalisation. On rappelle les conditions nécessaires et suffisantes sur une fonction de répartition pour qu'elle appartienne à un domaine d'attraction.

1.4.2 Domaine d'attraction de Fréchet

Rappelons que le domaine d'attraction de Fréchet contient les lois dont la fonction de survie est à décroissance polynomiale, *i.e.* les lois à queues lourdes ou lois de type Pareto. Les lois de ce domaine ont un point terminal y_F infini. En effet, le résultat ci-dessous énoncé par Gnedenko [97] et dont on trouvera une démonstration simple dans le livre de Resnick [147, Proposition 1.11] assure que toute fonction appartenant au domaine d'attraction de Fréchet est une fonction à variations régulières (voir Définition 1.5) et inversement.

Théorème 1.10. *Une fonction de répartition F appartient au domaine d'attraction de Fréchet (avec un indice des valeurs extrêmes $\gamma > 0$) si et seulement si $y_F = \infty$ et si la fonction de survie $\overline{F}$ est à variations régulières d'indice $-1/\gamma$ qui s'écrit $\overline{F} \in \mathcal{RV}_{-1/\gamma}$ i.e.*

$$\forall y > 0, \quad \lim_{t\to\infty} \frac{\overline{F}(ty)}{\overline{F}(t)} = y^{-1/\gamma}.$$

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

$$F(y) = 1 - y^{-1/\gamma}\ell(y) \quad \text{avec} \quad \ell \in \mathcal{RV}_0. \tag{1.13}$$

Les suites de normalisation (a_n) et (b_n) sont données dans ce cas pour tout $n > 0$ par :

$$a_n = \overline{F}^{\leftarrow}(1/n) \quad \text{et} \quad b_n = 0,$$

(voir Resnick [147, Proposition 1.11]). On peut facilement obtenir une caractérisation de la fonction quantile dans le domaine d'attraction de Fréchet. Pour cela on a besoin, par définition du quantile (voir Définition 1.1), d'inverser la fonction de survie. On peut alors à l'aide du Lemme 1.6 montrer que l'expression (1.13) est équivalente à :

$$q(\alpha) = \alpha^{-\gamma}L(\alpha^{-1}) \quad \text{avec} \quad L \in \mathcal{RV}_0, \tag{1.14}$$

où $\alpha \in]0, 1[$.

Le domaine d'attraction de Fréchet regroupe une grande diversité de lois comptant parmi elles des lois usuelles (loi de Student, loi du Chi-deux, loi Log-gamma, loi de Fréchet) se reporter au Tableau 1.1 pour plus d'exemples de lois. Il est donc sujet à de nombreuses applications, citons en particulier Gardes *et al.* [89], Gardes et Girard [85] et Daouia *et al.* [44].

1.4.3 Domaine d'attraction de Weibull

Toutes les lois appartenant au domaine d'attraction de Weibull ont un point terminal y_F fini. Le résultat suivant (voir Gnedenko [97], Resnick [147, Proposition 1.13]) montre que l'on passe du domaine d'attraction de Fréchet (voir sous-Partie 1.4.2) à celui de Weibull par un simple changement de variable dans la fonction de répartition.

Théorème 1.11. *Une fonction de répartition F appartient au domaine d'attraction de Weibull (avec un indice des valeurs extrêmes $\gamma < 0$) si et seulement si $y_F < \infty$ et si la fonction de répartition F_* définie par :*

$$F_*(y) = \begin{cases} 0 & \text{si } y \leq 0 \\ F(y_F - 1/y) & \text{si } y > 0 \end{cases}$$

appartient au domaine d'attraction de Fréchet avec un indice des valeurs extrêmes $-\gamma > 0$ i.e.

$$\forall y > 0, \quad \lim_{t\to\infty} \frac{\overline{F}\left(y_F - 1/(ty)\right)}{\overline{F}\left(y_F - 1/t\right)} = \lim_{t\to\infty} \frac{\overline{F}_*\left(ty\right)}{\overline{F}_*\left(t\right)} = y^{+1/\gamma}.$$

Ainsi, une fonction de répartition F du domaine d'attraction de Weibull s'écrit pour $y \leq y_F$ sous la forme :

$$F(y) = 1 - (y_F - y)^{-1/\gamma}\ell((y_F - y)^{-1}) \quad \text{avec} \quad \ell \in \mathcal{RV}_0. \tag{1.15}$$

Comme on peut facilement passer du domaine d'attraction de Fréchet à celui de Weibull on a les suites de normalisation (a_n) et (b_n) données pour tout $n > 0$ par :

$$a_n = y_F - \overline{F}^{\leftarrow}(1/n) \quad \text{et} \quad b_n = y_F.$$

De même, il est facile d'obtenir une caractérisation de la fonction quantile associée

$$q(\alpha) = y_F - \alpha^{-\gamma}L(1/\alpha), \quad \text{avec} \quad L \in \mathcal{RV}_0, \tag{1.16}$$

où $\alpha \in]0, 1[$.

Ce domaine d'attraction a été considéré notamment par Falk [74], Girard *et al.* [95] et Hall [112] pour estimer le point terminal d'une distribution.

1.4.4 Domaine d'attraction de Gumbel

Le domaine d'attraction de Gumbel contient les lois dont la fonction de survie est à décroissance exponentielle, *i.e.* les lois à queues légères. A contrario des deux autres domaines, il n'y a pas de représentation simple pour les lois appartenant au domaine d'attraction de Gumbel. Il peut être décrit à partir des fonctions de type Von Mises [164].

Le résultat ci-dessous démontré notamment dans Resnick [147, Proposition 1.4] donne une caractérisation plus précise.

Théorème 1.12. *Une fonction de répartition F appartient au domaine d'attraction de Gumbel si et seulement si il existe $t < y_F \leq \infty$ tel que :*

$$\overline{F}(y) = c(y) \exp\left\{ - \int_t^y \frac{1}{\delta(u)} du \right\} \quad \text{où} \quad t < y < y_F, \tag{1.17}$$

$c(y) \to c > 0$ lorsque $y \to y_F$ et δ est une fonction positive et dérivable de dérivée δ' telle que $\delta'(y) \to 0$ lorsque $y \to y_F$.

Dans ce cas, un choix possible pour les suites (a_n) et (b_n) pour tout $n > 0$ est :

$$a_n = q(1/n) \quad \text{et} \quad b_n = \frac{1}{\overline{F}(a_n)} \int_{a_n}^{y_F} \overline{F}(z) dz.$$

Il est difficile de caractériser la fonction quantile de lois appartenant au domaine d'attraction de Gumbel. Du fait que ce domaine d'attraction regroupe une grande diversité de lois, comptant parmi elles la plupart des lois usuelles (voir Tableau 1.1), il est sujet à de nombreuses applications particulièrement en hydrologie voir de Haan [49] et Gumbel [108, 109].

Le domaine d'attraction de Gumbel étant difficile à étudier dans toute sa généralité, de nombreux auteurs se sont concentrés sur une sous-famille facilement caractérisable, la famille de lois à queue de type Weibull.

Lois à queue de type Weibull

Cette famille englobe une grande variété de lois à queues légères ; par exemple la loi de Weibull, la loi Normale, la loi Gamma et les lois Logistiques sont des lois à queue de type Weibull, voir Galambos [80] pour plus de détails.

Ce type de lois intervient dans de nombreuses applications, en hydrologie notamment, mais aussi dans la modélisation de sinistres importants en assurance vie voir Beirlant et Teugels [18]. Comme ces lois appartiennent au domaine d'attraction de Gumbel, elles ont des fonctions de survie qui décroissent vers zéro à une vitesse exponentielle. De quelle manière plus précisément décroissent-elles vers zéro ?

Pour cela donnons la définition de lois à queue de type Weibull.

Définition 1.13. *Une fonction de répartition F a une queue de type Weibull si sa fonction de survie associée $\overline{F}$ satisfait la propriété suivante : Il existe $\beta > 0$ tel que pour tout $\lambda > 0$:*

$$\lim_{t \to \infty} \frac{-\log \overline{F}(\lambda t)}{-\log \overline{F}(t)} = \lambda^{1/\beta} \quad \textit{autrement dit} \quad -\log(\overline{F}) \in \mathcal{RV}_{1/\beta}.$$

La queue de ces lois est contrôlée par le paramètre de forme $\beta > 0$, appelé coefficient à queue de type Weibull. On se réfèrera à Beirlant *et al.* [19] pour un résumé plus général sur les lois à queue de type Weibull. De nombreuses publications récentes leurs sont consacrées, citons par exemple Goegebeur *et al.* [98] qui proposent des tests d'adéquation pour ce type de lois.

Les lois à queue de type Weibull correspondent au cas particulier où l'on suppose dans l'équation (1.17) que la dérivée de la fonction δ est à variations régulières d'indice strictement négatif. Plus présicément, si l'on suppose que $\delta' \in \mathcal{RV}_{-1/\beta}$ où β est l'indice de queue de Weibull, on montre facilement que l'équation (1.17) s'écrit :

$$\overline{F}(y) = \exp\left(-y^{1/\beta}\ell(y)\right) \quad \text{avec} \quad \ell \in \mathcal{RV}_0. \tag{1.18}$$

Toute fonction de répartition pouvant s'écrire sous la forme donnée dans l'expression (1.18) est dite à queue de type Weibull d'indice $\beta > 0$. Ainsi pour les lois à queue de type Weibull, il est facile d'obtenir une caractérisation de la fonction quantile associée :

$$q(\alpha) = (-\log(\alpha))^{\beta} L(-\log(\alpha)) \quad \text{avec} \quad L \in \mathcal{RV}_0 \tag{1.19}$$

où $\alpha \in]0, 1[$.

On remarque que la loi Log-normale n'est pas une loi à queue de type Weibull. Elle appartient à une autre sous famille intéressante de lois, les lois à queue de type Log-weibull. Une variable aléatoire Y suit une loi à queue de type Log-weibull si $\log(Y)$ suit une loi à queue de type Weibull. La décroissance de la fonction de survie des lois à queue de type Log-weibull est plus lente que celle des lois à queue de type Weibull.

Depuis le début du chapitre, on a vu comment modéliser le comportement des valeurs extrêmes d'un échantillon à travers la loi des valeurs extrêmes dans la Partie 1.2 puis la loi des excès dans la Partie 1.3 et la caractérisation des domaines d'attraction dans la Partie 1.4. A l'aide de ces résultats, on va pouvoir s'intéresser à notre but premier qui est l'estimation de quantiles extrêmes. Dans la Partie 1.5 on présentera les différentes approches. La première est basée sur la loi des valeurs extrêmes (sous-Partie 1.2), la deuxième sur la loi des excès (sous-Partie 1.3) et la troisième sur une approche semi-paramétrique faisant appel à la caractérisation des différents domaines d'attraction (sous-Partie 1.4).

1.5 Estimation de quantiles extrêmes

Dans cette partie, on supposera que F appartient à l'un des domaines d'attraction défini précédemment. L'estimation de quantiles extrêmes en dehors de l'échantillon pose problème. En effet, on a vu que supposer que $n\alpha_n \to 0$ quand $n \to \infty$ implique que $\mathbb{P}(Y_{n,n} < q_{\alpha_n}) \to 1$. Par conséquent, comme évoqué dans la Partie 1.1 au début du chapitre (voir Second cas page 12) on ne peut pas estimer le quantile de manière empirique. Dans une telle situation, l'estimation du quantile extrême requiert une extrapolation au-delà de l'échantillon. Il existe en théorie des valeurs extrêmes différentes approches, nous présenterons les trois suivantes :

1. L'approche des maxima par bloc consiste à séparer l'échantillon en sous-blocs choisis arbitrairement assez grands puis à ajuster la loi des maxima de ces blocs à une loi des valeurs extrêmes. D'après le Théorème 1.3 on a $F^n(a_n y + b_n) \approx \mathcal{H}_\gamma(y)$. A partir de cette approximation on pourra estimer des quantiles extrêmes. Cette estimation dépendra de l'ajustement de la loi des valeurs extrêmes, on présentera donc les différentes méthodes d'estimation des paramètres de celle-ci.

2. L'approche par dépassements de seuil consiste à ajuster la loi des excès à une loi de Pareto généralisée. D'après le Théorème 1.4 on a $F_u(y) \approx \mathcal{G}_{\gamma,\sigma}(y)$. A partir de cette approximation on pourra estimer des quantiles extrêmes. Cette estimation dépendra de l'ajustement de F_u à $\mathcal{G}_{\gamma,\sigma}$ donc du choix du seuil et de l'estimation des paramètres γ et $\sigma > 0$. On présentera donc les différentes méthodes d'estimation de ces derniers.

3. L'approche semi-paramétrique :
   - si l'on se place dans le domaine d'attraction de Fréchet, on a la caractérisation de la fonction quantile (voir expression (1.14)). A partir de cette expression, Weissman [167] proposa un estimateur semi-paramétrique du quantile extrême dépendant de l'estimation du paramètre γ. On présentera donc des estimateurs de $\gamma > 0$ dont le plus célèbre est dû à Hill [116] en 1975.
   - si l'on se place dans le domaine d'attraction de Gumbel, et plus particulièrement dans le cas d'une loi à queue de type Weibull, on a la caractérisation de la fonction quantile (voir expression (1.19)). A partir de cette expression, Beirlant *et al.* [19] proposa un estimateur semi-paramétrique du quantile extrême dépendant de l'estimation du paramètre β. On présentera donc des estimateurs de $\beta > 0$ dont le plus célèbre est dû à Beirlant *et al.* [19] en 1996.

Il s'avère qu'il existe une analogie entre toutes ces méthodes d'estimation de quantiles extrêmes. Ces méthodes sont basées sur des estimateurs des paramètres des lois GEV et GPD et en particulier sur des estimateurs de l'indice des valeurs extrêmes. Dans la suite

de cette partie, on présentera les estimateurs les plus importants et on fera un rappel de la littérature.

1.5.1 L'approche par la loi des valeurs extrêmes

Guida et Longo [106] se sont basés sur cette approche pour estimer des quantiles extrêmes. On a, d'après le Théorème 1.3, l'approximation suivante :

$$\mathbb{P}\left(\frac{Y_{n,n}-b_n}{a_n} \leq y\right) = F^n(a_n y + b_n) \approx \mathcal{H}_\gamma(y), \tag{1.20}$$

ou de façon équivalente :

$$\mathbb{P}\left(Y_{n,n} \leq z\right) \approx \mathcal{H}_\gamma\left(\frac{z-b_n}{a_n}\right). \tag{1.21}$$

Dans la suite on introduit :

$$\mathcal{H}_{\gamma,a_n,b_n} = \mathcal{H}_\gamma\left(\frac{z-b_n}{a_n}\right).$$

L'équation (1.21) montre que la loi du maximum $Y_{n,n}$ peut être approchée par $\mathcal{H}_{\gamma,a_n,b_n}$. On peut réécrire l'équation (1.20) comme suit :

$$\lim_{n\to\infty} n\log(F(a_n y + b_n)) = \lim_{n\to\infty} n\log(1-\overline{F}(a_n y + b_n)) = \log(\mathcal{H}_\gamma(y)).$$

On peut montrer que $a_n y + b_n \to y_F$ quand $n \to \infty$ donc $\overline{F}(a_n y + b_n) \to 0$. On peut alors faire un développement limité de $\log(1+u)$ au premier ordre ce qui donne :

$$\overline{F}(a_n y + b_n) \simeq -\frac{1}{n}\log(\mathcal{H}_\gamma(y)).$$

Ce qui équivaut à :

$$\overline{F}(y) \simeq -\frac{1}{n}\log\left(\mathcal{H}_\gamma\left(\frac{y-b_n}{a_n}\right)\right) = -\frac{1}{n}\log\left(\mathcal{H}_{\gamma,a_n,b_n}(y)\right)$$

A l'aide de l'expression de $\mathcal{H}_\gamma$ donnée dans l'expression (1.4), on obtient ainsi une approximation de la fonction de survie en queue :

$$\overline{F}(y) \simeq \frac{1}{n}\left(1+\gamma\left(\frac{y-b_n}{a_n}\right)\right)^{-1/\gamma}, \tag{1.22}$$

que l'on peut prolonger à $\gamma = 0$ en faisant tendre $\gamma \to 0$ dans l'équation (1.22) ce qui donne :

$$\overline{F}(y) \simeq \frac{1}{n}\exp\left(-\frac{y-b_n}{a_n}\right).$$

On souhaite estimer des quantiles, or par définition du quantile, (voir Définition 1.1) il nous faut inverser la fonction de survie de l'équation (1.22) ce qui nous permet d'approcher le quantile $q(\alpha)$ par :

$$q(\alpha) \simeq b_n + \frac{a_n}{\gamma}\left(\left(\frac{1}{n\alpha}\right)^{\gamma} - 1\right). \tag{1.23}$$

De même que précédemment, le cas $\gamma = 0$ dans l'expression (1.23) peut être vu comme le cas limite lorsque $\gamma \to 0$, on a alors :

$$q(\alpha) \simeq b_n - a_n \log(n\alpha).$$

On obtient ainsi un estimateur du quantile extrême donné dans la définition suivante.

Définition 1.14. *L'estimateur du quantile extrême de la loi GEV est défini par :*

$$\hat{q}_n^{gev}(\alpha_n) = \hat{b}_n + \frac{\hat{a}_n}{\hat{\gamma}_n}\left(\left(\frac{1}{n\alpha_n}\right)^{\hat{\gamma}_n} - 1\right),$$

où $(\hat{a}_n, \hat{b}_n, \hat{\gamma}_n)$ *sont respectivement des estimateurs des paramètres* (a_n, b_n, γ).

De même qu'auparavant, comme le cas $\gamma = 0$ peut être vu comme le cas limite lorsque $\gamma \to 0$. On a d'après l'expression (1.24) :

$$\hat{q}_n^{gev}(\alpha) = \hat{b}_n - \hat{a}_n \log(n\alpha).$$

Pour mettre en pratique cette approche basée sur la convergence en loi du maximum (convenablement normalisé) d'un échantillon vers une loi GEV, Jules Emile Gumbel a introduit en 1958 [109] l'approche des maxima par bloc, en anglais "Block maxima approach".

Ainsi si l'on dispose d'un échantillon *i.i.d.* $Y_1, \ldots, Y_n$, il nous faut tout d'abord obtenir des maxima. L'approche des maxima par bloc consiste à séparer l'échantillon en m sous-échantillons (blocs) disjoints, choisis arbitrairement assez grands. On extraira ainsi le maximum de chaque bloc nous donnant un échantillon de maxima noté $Z_1, \ldots, Z_m$. La loi de ces maxima est alors approchée, pour une taille de chaque bloc assez grande, par une loi des valeurs extrêmes (Définition 1.4).

L'idée est résumée dans l'article Mandelbrot's Extremism de Beirlant *et al.* [17].

Block-maximum postulate : *Up to location and scale, the distribution of block maxima stabilizes as the block size grows indefinitely large.*

Une fois cet échantillon de maxima obtenu, on peut alors s'en servir pour estimer de diverses façons les paramètres (γ, a_n, b_n) de la loi GEV. On présentera dans le paragraphe suivant la méthode du maximum de vraisemblance et la méthode des moments pondérés.

Pour cela considérons un échantillon $Z_1, \ldots, Z_m$ de m maxima *i.i.d.* de loi $\mathcal{H}_{\gamma,a_n,b_n}$.

Estimation des paramètres par la méthode du maximum de vraisemblance

Prescott et Walden [142, 143] étudièrent au début des années 1980 l'estimation des paramètres par la méthode du maximum de vraisemblance.

La fonction de log-vraisemblance obtenue à partir de la Définition 1.4 s'écrit :

$$\begin{aligned}\log\left(\mathcal{L}(\gamma, a_n, b_n)\right) &= -m\log(a_n) - \left(1+\frac{1}{\gamma}\right)\sum_{i=1}^{m}\log\left(1+\gamma\left(\frac{Z_i-b_n}{a_n}\right)\right)\\ &- \sum_{i=1}^{m}\left(1+\gamma\left(\frac{Z_i-b_n}{a_n}\right)\right)^{-1/\gamma},\end{aligned}$$

avec

$$1+\gamma\left(\frac{Z_i-b_n}{a_n}\right) > 0 \quad \forall i \in \{1, \ldots, m\}.$$

L'estimateur du maximum de vraisemblance n'est pas explicite. Jenkinson a proposé un algorithme itératif de maximisation de la fonction de vraisemblance. L'algorithme de Newton-Raphson correspondant est donné dans Hosking [118] et amélioré dans Macleod [129].

Les propriétés asymptotiques de l'estimateur du maximum de vraisemblance furent étudiées par Smith [157] en 1985. Il montre que si $\gamma > -1/2$, on a la consistance, l'efficacité et la normalité asymptotique de ces estimateurs.

Zhou [168] et Dombry [65] ont prouvé que l'estimateur du maximum de vraisemblance existe et est consistant pour $\gamma > -1$. Puis en 2010, Zhou [169] a également obtenu la normalité asymptotique pour $-1 < \gamma < -1/2$ et prouvé qu'il n'est pas consistant pour $\gamma < -1$.

Estimation des paramètres par la méthode des moments pondérés

En 1985, Hosking *et al.* [120] proposèrent d'utiliser la méthode des moments pondérés pour estimer les paramètres de la loi GEV.

On définit le moment pondéré d'ordre r par :

$$\mu_r = \mathbb{E}\left(Z\mathcal{H}^r_{\gamma,a_n,b_n}(Z)\right).$$

Cette quantité existe pour $\gamma < 1$ et est donnée par :

$$\mu_r = \frac{1}{r+1}\left(b_n - \frac{a_n}{\gamma}(1-(r+1)^{\gamma}\Gamma(1-\gamma))\right),$$

où Γ est la fonction gamma d'Euler définie pour tout $t > 0$ par :

$$\Gamma(t) = \int_0^{\infty} u^{t-1}\exp(-u)du.$$

En utilisant la formule précédente, trois moments pondérés suffisent pour calculer a_n, b_n et γ. En effet, on a :

$$\begin{aligned}
\mu_0 &= b - \frac{a_n}{\gamma}(1-\Gamma(1-\gamma)),\\
2\mu_1 - \mu_0 &= -\frac{a_n}{\gamma}(1-2^{\gamma})\Gamma(1-\gamma),\\
\frac{3\mu_2 - \mu_0}{2\mu_1 - \mu_0} &= \frac{3^{\gamma}-1}{2^{\gamma}-1}.
\end{aligned}$$

Ainsi en remplaçant respectivement μ_r, $r \in \{0,1,2\}$ par son estimateur empirique

$$\hat{\mu}_{r,n} = \frac{1}{m}\sum_{i=1}^{m} Z_{i,m}\left(\frac{i-1}{m}\right)^r.$$

où les $Z_{1,m},\ldots,Z_{m,m}$ sont les statistiques ordonnées associées à l'échantillon $Z_1,\ldots,Z_m$ et en résolvant le système précédent, on obtient les estimateurs des moments pondérés des paramètres a_n, b_n et γ.

Dans le cas d'échantillons de petite ou de moyenne taille, la méthode des moments pondérés donne de meilleurs résultats que la méthode du maximum de vraisemblance voir Hosking [118]. De plus les estimateurs des moments pondérés sont plus simples à calculer.

La mise en pratique de l'approche des maxima par bloc fait resurgir les défauts de cette dernière. La principale critique faite à cette méthode provient de la perte d'informations qui peut être conséquente lorsqu'on l'applique à un jeu de données. En effet, pour chaque bloc on ne retient qu'une seule observation, la plus grande, faisant ainsi fi de l'information contenue dans les autres grandes valeurs du bloc, autrement dit dans la queue de distribution. Ainsi, selon le découpage des blocs plusieurs valeurs extrêmes peuvent être contenues dans un même bloc, alors qu'un autre n'en contiendra pas, ou tout du moins pas du même ordre de grandeur.

De plus, l'approche maxima par bloc n'est facilement utilisable en pratique que sur des données horaires, journalières, annuelles ou sur un jeu de données structuré. Elle est adaptée lorsque les données sont recueillies de sorte qu'il soit facile d'en extraire des

maxima. Ce qui est par exemple le cas pour des données issues de la climatologie (précipitations, températures). De ce constat surgit un autre problème de mise en pratique. Si l'on fait des blocs annuels par exemple, on aura alors besoin de décennies de données. Cette approche est très coûteuse. Que faire lorsque l'on dispose de "peu" de données ?

"These results seemed to imply that the theory only worked if you had around a million data points which appeared to rule out any real applications, but more recent research shows that it can work pretty well with just a few hundred. It took a long time for the theory to go beyond that assumption."

Richard Smith

Afin de résoudre cette principale critique Bakelma, de Haan et Pickands ont proposé l'approche POT ou l'approche par la loi des excès (voir Partie 1.3). On se fixe un seuil u arbitrairement grand, ce qui coupe notre échantillon en deux parties. On ne travaille plus que sur les plus grandes valeurs que l'on ajuste à une loi de Pareto généralisée. Inverser la fonction de survie de cette loi nous permet d'estimer des quantiles extrêmes.

1.5.2 L'approche par la loi des excès

L'approche par la loi des excès est basée sur l'idée suivante. On a d'après l'équation (1.6), pour tout $y \geq 0$ la relation :

$$\overline{F}(u+y) = \overline{F}(u)\overline{F}_u(y).$$

Si on effectue le changement de variable $z = u + y$, alors l'approximation de la queue de distribution donne :

$$\overline{F}(z) = \overline{F}(u)\overline{F}_u(z-u) \approx \overline{F}(u)\overline{\mathcal{G}}_{\gamma,\sigma}(z-u),$$

où $\overline{\mathcal{G}}_{\gamma,\sigma}$ est la fonction de survie de la loi de Pareto généralisée $\mathcal{G}_{\gamma,\sigma}$ qui nous est donnée dans le Théorème 1.4. On introduit alors la probabilité p que Y dépasse le seuil u, $p = \mathbb{P}(Y > u) = \overline{F}(u)$, d'où :

$$\overline{F}(z) \simeq p\overline{\mathcal{G}}_{\gamma,\sigma}(z - \overline{F}^{\leftarrow}(p)).$$

On obtient ainsi pour $\gamma \in \mathbb{R}$ une approximation de la fonction de survie en queue :

$$\overline{F}(z) \simeq p\left(1 + \gamma\left(\frac{z - \overline{F}^{\leftarrow}(p)}{\sigma}\right)\right)^{-1/\gamma}. \tag{1.24}$$

Le cas $\gamma = 0$ peut être vu comme le cas limite $\gamma \to 0$ dans l'équation (1.24) :

$$\overline{F}(z) \simeq p \exp\left(-\frac{z - \overline{F}^{\leftarrow}(p)}{\sigma}\right).$$

On souhaite estimer des quantiles, or par définition du quantile (voir Définition 1.1), il nous faut inverser la fonction de survie de l'équation (1.24) ce qui donne :

$$q(\alpha) \simeq \overline{F}^{\leftarrow}(p) + \frac{\sigma}{\gamma}\left(\left(\frac{\alpha}{p}\right)^{-\gamma} - 1\right), \tag{1.25}$$

et dans le cas $\gamma = 0$ on fait tendre $\gamma \to 0$ dans l'équation (1.25) ce qui donne :

$$q(\alpha) \simeq \overline{F}^{\leftarrow}(p) - \sigma \log\left(\frac{\alpha}{p}\right). \tag{1.26}$$

Lorsque l'on suppose $\gamma = 0$, deux variantes de cette méthode ont été présentées par Breiman *et al.* [25] en 1990 sous les appellations Exponential Tail (ET) et Quadratique Tail (QT).

On notera la similitude entre l'expression du quantile de la loi GEV (voir l'expression (1.23)) et celle du quantile de la loi GPD (voir l'équation (1.25)). Il y a trois paramètres inconnus dans chacune d'entre elles :

- l'indice des valeurs extrêmes γ qui est le même dans les deux expressions soulignant de fait son importance dans le comportement de la queue de distribution et donc celui des valeurs extrêmes,
- Le paramètre d'échelle σ joue le rôle de a_n dans l'approche GEV,
- Le seuil $u = \overline{F}^{\leftarrow}(p)$ joue le rôle de b_n dans l'approche GEV.

Pour pouvoir estimer des quantiles extrêmes, il nous faut donc estimer ces paramètres.

Le seuil u donné par $\overline{F}^{\leftarrow}(p)$ est un quantile se trouvant dans l'échantillon, facile à estimer par inversion de la fonction de survie empirique (voir Premier cas page 12). Ainsi si $p_n = k_n/n$, où k_n est le nombre d'excès (auparavant noté N_u) on estime $\overline{F}^{\leftarrow}(k_n/n)$ par $Y_{n-k_n+1,n}$. Une fois le seuil u choisi, il nous reste à estimer les paramètres γ et σ afin d'obtenir un estimateur du quantile extrême $q(\alpha_n)$.

Définition 1.15. *L'estimateur du quantile extrême de la loi GPD est défini par :*

$$\hat{q}_n^{gpd}(\alpha_n) = Y_{n-k_n+1,n} + \frac{\hat{\sigma}_n}{\hat{\gamma}_n}\left(\left(\frac{k_n}{n\alpha_n}\right)^{\hat{\gamma}_n} - 1\right),$$

où $\hat{\gamma}_n$ et $\hat{\sigma}_n$ sont des estimateurs des paramètres de forme et d'échelle.

Différentes valeurs du seuil u donneront différents échantillons d'excès plus ou moins grands, ce qui influencera l'estimation des quantiles extrêmes. La question étant jusqu'à

quel point peut-on s'éloigner de l'échantillon, tout en restant dans la queue de distribution là où se situe l'information. Ce seuil doit être suffisament grand (k_n petit) pour que l'on puisse appliquer le Théorème 1.4, mais pas trop car sinon on utilisera peu d'excès donc peu d'informations.

Choix du seuil en pratique

Comme le font remarquer Davison et Smith [47], on ne dispose pas d'outils théoriques permettant de choisir de manière optimale le seuil. Sa détermination reste donc empirique. Généralement, on le détermine graphiquement.

Supposons que la loi des excès $Y - u_0$ au-delà d'un seuil u_0 puisse être approchée par une loi GPD $\mathcal{G}_{\gamma,\sigma_{u_0}}$, alors si $\gamma < 1$ on a :

$$e(u_0) = \mathbb{E}(Y - u_0 | Y > u_0) = \frac{\sigma_{u_0}}{1-\gamma}.$$

Si la loi GPD est valide pour les excès au-delà d'un seuil u_0, elle l'est également pour tout seuil $u > u_0$ avec le même paramètre de forme γ et d'après l'équation (1.8), avec un paramètre d'échelle $\sigma_u = \sigma_{u_0} + \gamma(u - u_0)$. On a alors :

$$e(u) = \mathbb{E}(Y - u | Y > u) = \frac{\sigma_u}{1-\gamma} = \frac{\sigma_{u_0} + \gamma(u - u_0)}{1-\gamma}.$$

Ainsi $e(u)$ est une fonction linéaire en u et peut être estimée par la moyenne empirique des dépassements observés du seuil u soit :

$$\hat{e}_n(u) = \frac{1}{n_u} \sum_{i=1}^{n_u} (Y_i - u),$$

où $Y_1, \ldots, Y_{n_u}$ sont les n_u excès au-delà du seuil u. La recherche du seuil u_0 à partir duquel la loi GPD est valide se fait à l'aide du graphique :

$$\{(u, \hat{e}_n(u)) \,;\, u \leq Y_{n_u,n_u}\} \quad \text{avec} \quad Y_{n_u,n_u} = \max_{i=\{1,\ldots,n_u\}} (Y_i).$$

Ce graphique est appelé "Mean residual life plot". Le seuil u_0 choisi est celui à partir duquel le Mean residual life plot est linéaire en u.

Il reste à estimer γ et σ, pour cela on présentera quatre grandes méthodes très utilisées en pratique :

- la méthode du maximum de vraisemblance,
- la méthode des moments,
- la méthode des moments pondérés,
- l'estimateur DEdH.

Pour cela considérons un échantillon $Z_1, \ldots, Z_{k_n}$ *i.i.d* de loi GPD $\mathcal{G}_{\gamma,\sigma}$.

Estimation des paramètres par la méthode du maximum de vraisemblance

La fonction de log-vraisemblance est obtenue à partir de la loi GPD $\mathcal{G}_{\gamma,\sigma}$ (voir Théorème 1.4) ce qui donne :

$$\log\left(\mathcal{L}(\gamma,\sigma)\right) = -k_n \log(\sigma) - \left(1+\frac{1}{\gamma}\right)\sum_{i=1}^{k_n} \log\left(1+\frac{\gamma}{\sigma}Z_i\right),$$

avec $1+\gamma Z_i/\sigma > 0$ pour $i = 1, \ldots, k_n$ sinon $\mathcal{L}(\gamma,\sigma) = -\infty$.

On ne peut avoir d'expression explicite des estimateurs solutions des équations de vraisemblance. Dans la pratique on utilisera des méthodes de résolution numériques telles que l'algorithme de Newton-Raphson (voir [119]).

Sous l'hypothèse $\gamma > -1/2$, les estimateurs du maximum de vraisemblance sont asymptotiquement gaussiens et efficaces, voir Smith [158]. Pour un exemple d'estimation des paramètres par maximum de vraisemblance voir les travaux de Smith [157] en 1985.

Comme le font remarquer Davison et Smith [47] l'estimateur du maximum de vraisemblance est cependant peu utilisé en pratique car il pose des problèmes numériques et il est peu performant sur des échantillons de petite taille ($n < 500$). Il est préférable dans ce cas d'estimer les paramètres par l'une des trois méthodes suivantes.

Estimation des paramètres par la méthode des moments

La méthode des moments fut introduite par Hosking et Wallis [119] en 1987 afin d'estimer les paramètres de la loi GPD. Comme évoqué précédemment, pour que l'espérance et la variance d'une variable aléatoire Z de loi GPD $\mathcal{G}_{\gamma,\sigma}$ existent il faut que $\gamma < 1/2$. Dans ce cas on a :

$$\mathbb{E}(Z) = \frac{\sigma}{1-\gamma} \quad \text{et} \quad \mathrm{Var}(Z) = \frac{\sigma^2}{(1-\gamma)^2(1-2\gamma)}.$$

On peut alors facilement exprimer les paramètres de la loi GPD γ et σ en fonction de l'espérance et de la variance de Z, soit :

$$\gamma = \frac{1}{2}\left(1-\frac{\mathbb{E}(Z)^2}{\mathrm{Var}(Z)}\right) \quad \text{et} \quad \sigma = \frac{\mathbb{E}(Z)}{2}\left(1+\frac{\mathbb{E}(Z)^2}{\mathrm{Var}(Z)}\right).$$

Ainsi en remplaçant $\mathbb{E}(Z)$ et $\mathrm{Var}(Z)$ par leurs estimateurs empiriques :

$$\overline{Z} := \frac{1}{k_n}\sum_{i=1}^{k_n} Z_i \quad \text{et} \quad s^2(Z) := \frac{1}{k_n - 1}\sum_{i=1}^{k_n}(Z_i - \overline{Z})^2.$$

On obtient les estimateurs des moments de γ et σ, soit :

$$\hat{\gamma}_n^m = \frac{1}{2}\left(1 - \frac{\overline{Z}^2}{s^2(Z)}\right) \quad \text{et} \quad \hat{\sigma}_n^m = \frac{\overline{Z}}{2}\left(1 + \frac{\overline{Z}^2}{s^2(Z)}\right).$$

Ces estimateurs sont asymptotiquement gaussiens si $\gamma < 1/4$ (voir [119]). Leur principal défaut est que leur domaine de validité est très limité.

Estimation des paramètres par la méthode des moments pondérés

Hosking et Wallis [119] en 1987 proposèrent aussi d'utiliser une méthode basée sur les moments pondérés pour estimer les paramètres de la loi GPD. Soit μ_r le moment pondéré d'ordre r de la loi GPD $\mathcal{G}_{\gamma,\sigma}$ définit par :

$$\mu_r := \mathbb{E}(Z(1 - \mathcal{G}_{\gamma,\sigma}(Z))^r) = \frac{\sigma}{(1+r)(1+r-\gamma)} \quad \text{avec} \quad \gamma < 1/r.$$

On peut facilement exprimer les paramètres de la loi GPD γ et σ en fonction de μ_0 et μ_1, soit

$$\gamma = \frac{\mu_0 - 4\mu_1}{\mu_0 - 2\mu_1} \quad \text{et} \quad \sigma = \frac{2\mu_0\mu_1}{\mu_0 - 2\mu_1}.$$

Ainsi en remplaçant respectivement μ_r, $r \in \{0, 1\}$ par son estimateur empirique :

$$\hat{\mu}_{r,n} := \frac{1}{k_n}\sum_{i=1}^{k_n}\left(1 - \frac{i}{1+k_n}\right)^r Z_{i,k_n}$$

où $Z_{1,k_n}, \ldots, Z_{k_n,k_n}$ sont les statistiques ordonnées associées à $Z_1, \ldots, Z_{k_n}$. On obtient alors les estimateurs des moments pondérés de γ et σ, soit :

$$\hat{\gamma}_n^{mp} = \frac{\hat{\mu}_{0,n} - 4\hat{\mu}_{1,n}}{\hat{\mu}_{0,n} - 2\hat{\mu}_{1,n}} \quad \text{et} \quad \hat{\sigma}_n^{mp} = \frac{2\hat{\mu}_{0,n}\hat{\mu}_{1,n}}{\hat{\mu}_{0,n} - 2\hat{\mu}_{1,n}}.$$

Hosking et Wallis [119] ont prouvé que les estimateurs des moments pondérés sont asymptotiquement gaussiens pour $-1 < \gamma < 1/2$ et ont étudié leurs performances sur simulations. Pour des applications pratiques leur domaine de validité reste limité. Cependant Diebolt *et al.* [61, 62] ont introduit une méthode basée sur les moments pondérés généralisés permettant d'agrandir le domaine de validité à $-1 < \gamma < 3/2$.

Estimation des paramètres par l'estimateur DEdH

Dekkers, Einmahl et de Haan (voir [55]) proposent d'estimer les paramètres de la loi GPD comme suit. Soit $(k_n)_{n\geq 1}$ une suite d'entiers avec $1 < k_n \leq n$, on a comme estimateur de $\gamma \in \mathbb{R}$:

$$\hat{\gamma}_n^M = \mathcal{M}_{k_n}^{(1)} + 1 - \frac{1}{2}\left(1 - \frac{\left(\mathcal{M}_{k_n}^{(1)}\right)^2}{\mathcal{M}_{k_n}^{(2)}}\right)^{-1}, \tag{1.27}$$

où

$$\mathcal{M}_{k_n}^{(r)} = \frac{1}{k_n - 1}\sum_{i=1}^{k_n-1}\left(\log(Z_i)\right)^r, \quad r \in \mathbb{R}.$$

L'estimateur $\hat{\gamma}_n^M$ a été introduit par Dekkers, Einmahl et de Haan [55] comme une généralisation de l'estimateur de Hill [116] à tous les domaines d'attraction permettant ainsi d'estimer $\gamma \in \mathbb{R}$. On notera que $\mathcal{M}_{k_n}^{(1)}$ correspond exactement à l'estimateur de Hill. Dans la sous-Partie suivante 1.5.3 on consacrera une grande partie du Paragraphe 1.5.3.1 à l'estimateur de Hill.

Il proposent également d'estimer $\sigma > 0$ par :

$$\hat{\sigma}_n = Y_{n-k_n+1,n}\mathcal{M}_n^{(1)}(1 - \hat{\gamma}_n^M + \mathcal{M}_n^{(1)}).$$

Si on insère ces deux estimateurs dans l'estimateur du quantile extrême de la loi GPD (voir Définition 1.15) on obtient l'estimateur DEdH donné par :

Définition 1.16. *Soit $(k_n)_{n\geq 1}$ une suite d'entiers avec $1 < k_n \leq n$, l'estimateur DEdH est défini par :*

$$\hat{q}_n^M(\alpha_n) = Y_{n-k_n+1,n} + Y_{n-k_n+1,n}\frac{\mathcal{M}_n^{(1)}(1 - \hat{\gamma}_n^M + \mathcal{M}_n^{(1)})}{\hat{\gamma}_n^M}\left(\left(\frac{k_n}{np_n}\right)^{\hat{\gamma}_n^M} - 1\right). \tag{1.28}$$

L'estimateur DEdH a été introduit en 1989 par Dekkers, Einmahl et de Haan [55] d'où son abréviation. Les auteurs établirent sa consistance forte, faible et sa normalité asymptotique. On l'appelle également estimateur des Moments car les $\mathcal{M}_{k_n}^{(r)}$ peuvent être interprétés comme des moments empiriques.

Quelle que soit l'approche choisie, l'élément clé réside dans l'estimation des paramètres de la loi modélisant le comportement des valeurs extrêmes, et en particulier de l'indice des valeurs extrêmes γ. Les estimateurs semi-paramétriques des quantiles extrêmes présentés dans la sous-partie suivante dépendent principalement de l'estimation de γ dans le cas du domaine d'attraction de Fréchet et de β dans le cas de loi à queue de type Weibull. On présentera alors différents estimateurs de ces derniers.

1.5.3 L'approche semi-paramétrique

Le plus connu des estimateurs de quantiles extrêmes basé sur une approche semi-paramétrique est l'estimateur de Weissman [167]. On verra qu'il repose sur l'estimation de $\gamma > 0$. Comme estimateur de l'indice des valeurs extrêmes $\gamma > 0$ on citera le plus connu d'entre eux qui est l'estimateur de Hill [116] ainsi que l'estimateur de Zipf [126, 155] qui est basé sur une approche par moindres carrés.

1.5.3.1 Lois à queues lourdes

Dans le domaine d'attraction de Fréchet, l'approche semi-paramétrique est basée sur la définition de la fonction quantile (voir expression 1.14). Pour tout $\gamma > 0$, on a

$$q(\alpha_n) = \alpha_n^{-\gamma}\ell(\alpha_n^{-1}) \quad \text{avec} \quad \ell \in \mathcal{RV}_0, \tag{1.29}$$

$$q(\beta_n) = \beta_n^{-\gamma}\ell(\beta_n^{-1}) \quad \text{avec} \quad \ell \in \mathcal{RV}_0. \tag{1.30}$$

En divisant (1.30) par (1.29) et à l'aide de la définition d'une fonction à variations lentes (voir Définition 1.7), si α_n est suffisament petit et $\beta_n < \alpha_n$ on obtient l'approximation suivante :

$$q(\beta_n) \simeq q(\alpha_n)\left(\frac{\alpha_n}{\beta_n}\right)^{\gamma}, \tag{1.31}$$

où $q(\alpha_n)$ est un quantile choisi dans l'échantillon, facile à estimer par inversion de la fonction de survie empirique (voir Premier cas page 12) et $q(\beta_n)$ un quantile choisi en dehors de l'échantillon (voir Second cas page 12). On va donc chercher à estimer $q(\beta_n)$ en extrapolant à partir de $q(\alpha_n)$.

Ces approximations sont des cas particuliers de l'approche GPD avec $\sigma = \gamma q(\alpha_n)$. On peut alors remplacer $q(\alpha_n)$ par son estimateur naturel $Y_{n-\lfloor n\alpha_n\rfloor+1,n}$ et γ par un estimateur quelconque $\hat{\gamma}_n$ estimant $\gamma > 0$. Ce qui nous permet d'obtenir l'estimateur de Weissman [167] introduit en 1978.

Définition 1.17. *L'estimateur de Weissman est défini par :*

$$\hat{q}_n^W(\beta_n) = Y_{n-\lfloor n\alpha_n\rfloor+1,n}\left(\frac{\alpha_n}{\beta_n}\right)^{\hat{\gamma}_n}. \tag{1.32}$$

Le terme $(\alpha_n/\beta_n)^{\hat{\gamma}_n}$ est le terme servant à l'extrapolation. Plus de détails sur les propriétés de cet estimateur sont disponibles dans Embrechts *et al.* [71], Weissman [167] et de Haan et Ferreira [50]. L'estimateur de Weissman repose sur l'estimation de $\gamma > 0$. Il

existera ainsi autant d'estimateurs des quantiles extrêmes dans le domaine d'attraction de Fréchet qu'il existera d'estimateurs de $\gamma > 0$.

Il est donc normal qu'une grande partie de la littérature sur les valeurs extrêmes soit dédiée à l'estimation de $\gamma > 0$. Le plus connu de ces estimateurs étant l'estimateur de Hill [116].

Estimateur de Hill

Voici une idée possible pour la construction de l'estimateur de Hill. Passons au logarithme dans l'équation (1.31) ce qui donne :

$$\log(q(\beta_n) - \log(q(\alpha_n))) \simeq \gamma \log(\alpha_n/\beta_n).$$

On choisit comme précédemment, $\alpha_n = k_n/n$ et on considère plusieurs valeurs pour β_n. $\beta_n = i/n$ avec $i = 1, \ldots, k_n - 1$ tout en ayant $\beta_n < \alpha_n$. On obtient alors :

$$\log(q(i/n)) - \log(q(k_n/n)) \simeq \gamma \log(k_n/i).$$

Ainsi en estimant les quantiles par leurs équivalents empiriques on obtient :

$$\log(Y_{n-i+1,n}) - \log(Y_{n-k_n+1,n}) \simeq \gamma \log(k_n/i). \tag{1.33}$$

En sommant de part et d'autre sur $i = 1, \ldots, k_n - 1$, on obtient :

$$\gamma \simeq \frac{\sum_{i=1}^{k_n-1} \log(Y_{n-i+1,n}) - \log(Y_{n-k+1,n})}{\sum_{i=1}^{k_n-1} \log(k_n/i)}.$$

Le dénominateur se réécrit $\log(k_n^{k_n-1}/(k_n-1)!)$. En utilisant la formule de Stirling, il est équivalent à k_n au voisinage de l'infini. On obtient alors l'estimateur de Hill [116].

Définition 1.18. *Soit $(k_n)_{n\geq 1}$ une suites d'entiers avec $1 < k_n \leq n$, l'estimateur de Hill est défini par :*

$$H_n(k_n) = \frac{1}{k_n - 1} \sum_{i=1}^{k_n-1} \log\left(Y_{n-i+1,n}\right) - \log\left(Y_{n-k_n+1,n}\right).$$

Il fut introduit en 1975 afin d'estimer non-paramétriquement $\gamma > 0$. Historiquement, la construction de l'estimateur de Hill a été d'appliquer la méthode du maximum de vraisemblance sur l'ensemble des k_n plus grandes observations de l'échantillon, autrement dit les k_n statistiques d'ordre supérieur. D'autres méthodes pour la construction de l'estimateur de Hill sont données dans le livre de de Haan et Ferreira [50] et dans le livre de Beirlant *et al.* [16].

Le premier résultat sur les propriétés asymptotiques de cet estimateur est dû à Mason [130] en 1982 qui en a démontré la consistance faible. La consistance forte fut établie en 1988 par Deheuvels *et al.* [54].

Théorème 1.19. *Propriétés asymptotiques de l'estimateur de Hill : Soit* $(k_n)_{n\geq 1}$ *une suites d'entiers telle que* $1 < k_n \leq n$, $k_n \to \infty$ *et* $k_n/n \to 0$ *quand* $n \to \infty$ *:*
- *alors* $H_n(k_n)$ *converge en probabilité vers* γ *(consistance faible),*
- *et si de plus* $k_n/\log\log n \to \infty$ *quand* $n \to \infty$*, alors* $H_n(k_n)$ *converge presque sûrement vers* γ *(consistance forte).*

La normalité asymptotique de l'estimateur de Hill a fait l'objet de nombreux travaux citons entre autre Csörgő et Viharos [40], Csörgo et Mason [42], Davis et Resnick [45], De Haan et Peng [51], Hall [113], Häusler et Teugels [115], Smith [158] et de Haan et Resnick [53]. En effet, une hypothèse sur la fonction à variations lentes ℓ est nécessaire. Cette condition porte sur la vitesse de convergence du rapport des fonctions à variations lentes vers 1 (voir Définition 1.7).

Condition du second ordre : Il existe $\rho < 0$ et $b(y) \underset{y\to\infty}{\longrightarrow} 0$ tels que uniformément localement pour $\lambda > 1$:

$$\log\left(\frac{\ell(\lambda y)}{\ell(y)}\right) \sim b(y)K_\rho(\lambda), \quad \text{lorsque} \quad y \to \infty,$$

où

$$K_\rho(x) = \int_1^y u^{\rho-1}du.$$

On remarque que nécessairement la fonction b (appelée aussi fonction de biais) est à variations régulières d'indice $\rho < 0$ (voir par exemple Geluk et de Haan [92]). Le paramètre de second ordre $\rho < 0$ contrôle la vitesse de convergence du ratio $\ell(\lambda x)/\ell(x)$ vers 1. Plus ρ sera proche de 0, plus la convergence sera lente et donc plus l'estimation de γ sera difficile.

De ce fait beaucoup d'auteurs se sont intéressés à l'estimation de ρ afin de réduire le biais asymptotique qui apparaît lorsque l'on estime $\gamma > 0$. Citons entre autre Alves *et al.* [3, 4], Gomes et Martins [104], Gomes *et al.* [105], Peng [140]. D'autres estimateurs de cet indice ont été proposés notamment par Beirlant *et al.* [12, 13] qui utilisent un modèle de régression exponentielle pour débiaiser l'estimateur de Hill et par Feuerverger et Hall [76] qui introduisent un estimateur des moindres carrés (voir le livre de Beirlant *et al.* [16] pour plus détails). Deme *et al.* [56] ont par la suite unifié ces travaux.

Cette condition du second ordre est la clef de voûte de toutes les preuves de normalité asymptotique en théorie des valeurs extrêmes. Elle est très souvent utilisée pour étudier

le comportement asymptotique d'estimateurs d'indice ou de quantiles extrêmes. Beirlant *et al.* [12], Häusler et Teugels [115] et Hill [116] s'en servent pour prouver la normalité asymptotique de nombreux estimateurs de l'indice des valeurs extrêmes.

Théorème 1.20. *Normalité asymptotique de l'estimateur de Hill (voir [16]).*

Soit $(k_n)_{n\geq 1}$ *une suite d'entiers telle que* $1 < k_n \leq n$, $k_n \to \infty$ *et* $k_n/n \to 0$ *quand* $n \to \infty$. *Si de plus la condition de second ordre est satisfaite avec* $\sqrt{k_n}b(n/k_n) \to 0$ *quand* $n \to \infty$ *alors :*

$$\sqrt{k_n}\,(H_n(k_n) - \gamma) \xrightarrow{d} \mathcal{N}(0, \gamma^2).$$

Dans les Théorèmes 1.19 et 1.20, (k_n) est une suite d'entiers telle que :

$$\lim_{n\to\infty} k_n = \infty \quad \text{et} \quad \lim_{n\to\infty} k_n/n = 0, \tag{1.34}$$

Une suite vérifiant les deux conditions de l'équation (1.34) sera appelée suite intermédiaire d'entiers.

La première condition $k_n \to \infty$ nous assure que le nombre de statistiques d'ordre utilisé k_n est assez grand afin d'obtenir des estimateurs stables. Cependant si k_n est trop grand, le seuil estimé $Y_{n-k_n+1,n}$ est petit, on sort alors de "la zone" où la fonction de survie est approximativement une fonction polynomiale. Notre approximation n'est plus valable et dans ce cas l'estimateur de Hill a alors un grand biais.

A l'inverse, la seconde condition $k_n/n \to 0$ permet de rester dans la queue de distribution. Cependant si k_n est trop petit, le seuil estimé sera trop grand, on aura alors peu d'excès ce qui engendrera une grande variance dans l'estimation. On doit donc s'assurer de ne pas trop s'éloigner de la queue de distribution là où se trouve l'information importante.

Comme il est dit dans Embrechts *et al.* [71] : "*Let the tail speak for itself.*"

Ces conditions se retrouvent dans toutes les approches basées sur un seuil, en particulier dans l'approche par la loi des excès (voir Partie 1.3). Dans cette dernière le seuil ne doit pas être trop grand (condition similaire à $k_n/n \to 0$) mais suffisamment grand pour avoir le Théorème 1.4 (condition similaire à $k_n \to \infty$).

Pour illustrer l'importance de k_n et rendre compte de la qualité de l'estimation, on trace l'estimateur de Hill en fonction de k_n, autrement dit la fonction $k_n \to H_n(k_n)$ (voir Figure 1.7).

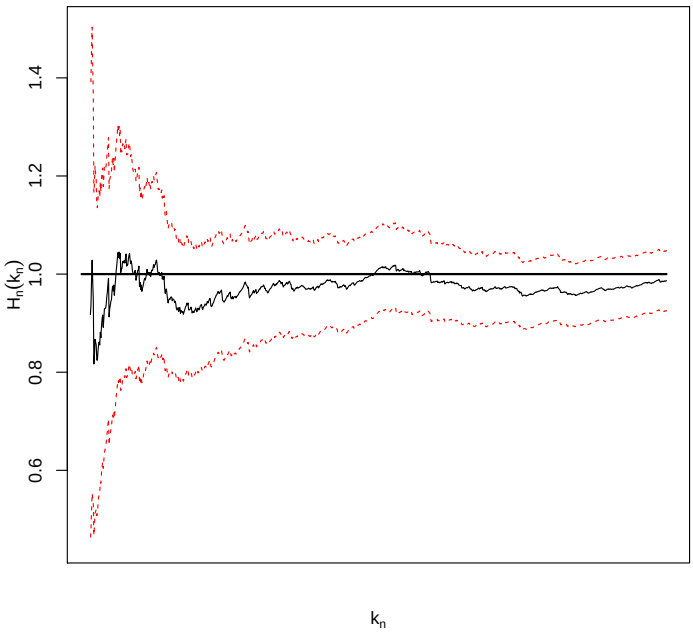


FIGURE 1.7: Graphique de l'estimateur de Hill $H_n(k_n)$, en fonction de k_n (trait noir) avec son intervalle de confiance à 95% (lignes rouge), $n = 1000$ réalisations d'une loi de Pareto standard pour laquelle $\gamma = 1$. La ligne horizontale représente la vraie valeur de γ. Dans ce cas la fonction à variations lentes est constante.

L'utilisation d'un noyau dans l'estimateur de Hill a été étudiée par Csorgo *et al.* [41]. Un estimateur efficace de l'indice des valeurs extrêmes a été proposé par Falk et Marohn [75]. Une liste plus détaillée des différents travaux sur l'estimation de l'indice des valeurs extrêmes est dressée par Csörgő et Viharos [40].

Une interprétation graphique de l'estimateur de Hill est possible (voir expression (1.33)), on l'appelle le "Pareto quantile plot". C'est le graphe des points :

$$(-\log(i/(n+1)), \log(Y_{n-i+1,n})) \quad \text{pour} \quad i = 1, \ldots, k_n. \tag{1.35}$$

Cette représentation est très utile car elle permet de visualiser facilement si les observations sont distribuées selon une loi appartenant au domaine d'attraction de Fréchet. En effet, si un échantillon provient d'une loi appartenant à ce domaine, alors le Pareto quantile plot doit être approximativement linéaire avec une pente γ, pour les petites valeurs de i, autrement dit pour les valeurs extrêmes. Il devient linéaire à partir du point $(-\log(k_n/(n+1)), \log(Y_{n-k_n+1,n}))$. Si l'on utilise la méthode des moindres carrés en forçant la droite de regression à passer par ce point on retrouve l'estimateur de Hill.

Estimateur de Zipf

En 1996 deux travaux indépendants, Schultze et Steinebach [155] et Kratz et Resnick [126] conduirent à l'introduction du même estimateur de l'indice des valeurs extrêmes, l'estimateur de Zipf. Il consiste à utiliser la méthode des moindres carrés sur le précédent graphe des points (voir expression (1.35)).

Définition 1.21. *Soit* $(k_n)_{n\geq 1}$ *une suite d'entiers avec* $1 < k_n \leq n$*, l'estimateur de Zipf est défini par :*

$$\hat{\gamma}_n^Z = \frac{\frac{1}{k_n}\sum_{i=1}^{k_n}\left(\log\left(\frac{k_n+1}{i}\right) - \frac{1}{k_n}\sum_{i=1}^{k_n}\log\left(\frac{k_n+1}{i}\right)\right)\log\left(Y_{n-i+1,n}\right)}{\frac{1}{k_n}\sum_{i=1}^{k_n}\left(\log\left(\frac{k_n+1}{i}\right)\right)^2 - \left(\frac{1}{k_n}\sum_{i=1}^{k_n}\log\left(\frac{k_n+1}{i}\right)\right)^2} \tag{1.36}$$

Il a été introduit dans le but d'obtenir un biais asymptotique plus petit que les estimateurs précédemment introduits, notamment l'estimateur de Hill. Cependant, sa variance asymptotique est deux fois plus grande que celle de l'estimateur de Hill (voir Théorème 1.20).

Le paragraphe suivant est une courte synthèse bibliographique des méthodes d'estimation des quantiles extrêmes pour des lois à queue de type Weibull. Nous montrons comment cette problématique s'inscrit plus largement dans la théorie des valeurs extrêmes et les similitudes qui existent entre les estimateurs du coefficient à queue de type Weibull et l'estimateur de Hill.

1.5.3.2 Lois à queue de type Weibull

Le plus connu des estimateurs des quantiles extrêmes basé sur une approche semi-paramétrique est l'estimateur de Beirlant [19]. On verra qu'il repose sur l'estimation de $\beta > 0$.

Comme les lois à queue de type Weibull appartiennent au domaine d'attraction de Gumbel, elles ont un indice des valeurs extrêmes $\gamma = 0$. L'indice des valeurs extrêmes ne fournit donc aucune information sur la vitesse de décroissance de la fonction de survie à l'intérieur de cette famille de lois. C'est l'indice de queue de Weibull $\beta > 0$ qui donne cette information. Une valeur du paramètre de forme β proche de zéro correspond à une décroissance rapide de la queue de distribution. Réciproquement une valeur de β proche de l'infini correspond à une décroissance lente de la queue de distribution. L'estimation des quantiles extrêmes pour des lois à queue de type Weibull passe alors par l'estimation de β.

Afin d'estimer des quantiles extrêmes, commençons par remarquer que pour des lois à queue de type Weibull on a d'après la définition de la fonction quantile (voir expression (1.19)) pour tout $\beta > 0$:

$$q(\alpha) = (-\log(\alpha))^{\beta}\ell(-\log(\alpha)) \quad \text{avec} \quad \ell \in \mathcal{RV}_0.$$

Ce qui donne en passant au logarithme :

$$\frac{\log q(\alpha)}{\log_2(1/\alpha)} = \beta + \frac{\log \ell(\log(1/\alpha))}{\log_2(1/\alpha)},$$

où $\log_2(.) = \log(\log(.))$.

Ainsi, en utilisant la Proposition 1.8 qui dit que $\log(\ell(x))/\log(x) \to 0$ lorsque $x \to \infty$, on en déduit que lorsque $\alpha \to 0$:

$$\log q(\alpha) \approx \beta \log_2(1/\alpha). \tag{1.37}$$

L'idée est la même que pour la construction de l'estimateur proposé par Weissman [167] (voir Définition 1.17). Elle est basée sur l'approximation donnée par l'équation (1.37) qui assure que pour n assez grand, si on a (k_n) une suite intermédiaire d'entiers (voir l'expression (1.34)) alors on a :

$$\log(q(\alpha_n)) \approx \beta \log_2(1/\alpha_n) \tag{1.38}$$

$$\log(q(k_n/n)) \approx \beta \log_2(n/k_n) \tag{1.39}$$

En soustrayant l'équation (1.39) à l'équation (1.38) on obtient :

$$\log(q(\alpha_n)) \approx \log(q(k_n/n)) + \beta(\log_2(1/\alpha_n) - \log_2(n/k_n)).$$

On applique alors la fonction exponentielle, ce qui donne :

$$q(\alpha_n) \approx q(k_n/n)\left(\frac{\log(1/\alpha_n)}{\log(n/k_n)}\right)^{\beta}. \tag{1.40}$$

Le quantile $q(k_n/n)$ se trouve dans l'échantillon (voir Premier cas page 12.), on peut alors l'estimer par son estimateur empirique soit la statistique d'ordre $Y_{n-k_n+1,n}$. Puis on remplace β par un estimateur quelconque $\hat{\beta}_n$ estimant $\beta > 0$. Ce qui nous permet d'obtenir l'estimateur de Beirlant [19] introduit en 1996.

Définition 1.22. *L'estimateur de Beirlant est défini par :*

$$\hat{q}_n^B(\alpha_n) = Y_{n-k_n+1,n}\left(\frac{\log(1/\alpha_n)}{\log(n/k_n)}\right)^{\hat{\beta}_n}. \tag{1.41}$$

Pour estimer des quantiles extrêmes on procède de la même façon que dans le cas de lois appartenant au domaine d'attraction de Fréchet (voir Définition 1.17), sauf que cette fois on remplace le coefficient à queue de Weibull β par l'un de ces estimateurs. On peut alors remplacer β dans l'expression (1.41) par n'importe quel estimateur de la littérature et ainsi obtenir autant d'estimateurs des quantiles extrêmes qu'il existe d'estimateurs de β.

Estimation de l'indice de queue de type Weibull

Il existe dans la littérature de nombreux estimateurs dédiés à l'estimation du coefficient de queue de type Weibull. Historiquement, la première approche est due à Berred [20] qui proposa un estimateur basé sur des valeurs records. A la suite de ces travaux deux grandes familles d'approches se sont développées.

- Une première famille d'approches (Beirlant *et al.* [9, 10], Broniatowski [27], Dierckx *et al.* [63]) est basée sur les logarithmes des excès au-dessus d'un seuil (log-excesses) des k_n statistiques d'ordre supérieur :

$$\log(Y_{n-i+1,n}) - \log(Y_{n-k_n+1,n}).$$

- Une seconde famille d'approches (Beirlant *et al.* [19], Diebolt *et al.* [59], Gardes et Girard [82, 83], Girard [93], Goegebeur *et al.* [98], Goegebeur et Guillou [101]) se base quant à elle sur les espacements entre les logarithmes (log-spacings) des k_n statistiques d'ordre supérieur :

$$\log(Y_{n-i+1,n}) - \log(Y_{n-i,n}).$$

Le plus simple de ces estimateurs a été proposé dans Beirlant *et al.* [19]. Il peut être construit en remarquant que pour des lois à queue de type Weibull on a d'après l'expression (1.19) de la fonction quantile :

$$\begin{aligned} q(t) &= (\log(1/t))^{\beta}\ell(\log(1/t)), \\ q(s) &= (\log(1/s))^{\beta}\ell(\log(1/s)), \end{aligned}$$

ainsi pour s et t proches de 0 on a :

$$\begin{aligned} \log(q(t)) - \log(q(s)) &= \beta(\log_2(1/t) - \log_2(1/s)) + \log\left(\frac{\ell(\log(1/t))}{\ell(\log(1/s))}\right), \\ &\simeq \beta(\log_2(1/t) - \log_2(1/s)). \end{aligned}$$

L'approximation est justifiée par le fait que ℓ est une fonction à variations lentes à l'infini (voir Définition 1.7). On peut alors considérer l'estimateur de β suivant.

Définition 1.23. *Soit $(k_n)_{n\geq 1}$ une suite d'entiers avec $1 < k_n \leq n$, l'estimateur de Beirlant est défini par :*

$$\hat{\beta}_n^B = \frac{1}{\sum_{i=1}^{k_n-1}(\log_2(n/i) - \log_2(n/k_n))} \sum_{i=1}^{k_n-1} (\log(Y_{n-i+1,n}) - \log(Y_{n-k_n+1,n})). \tag{1.42}$$

La forme de cet estimateur est très proche de celle de l'estimateur de l'indice des valeurs extrêmes proposé par Hill [116] (voir définition (1.18)). En effet, seul le terme de normalisation (dénominateur) diffère.

Une interprétation graphique de cet estimateur est possible. En effet, rappelons que pour des lois à queue de type Weibull on a $\log q(\alpha) \sim \beta \log_2(1/\alpha)$ d'après l'équation (1.37). Ainsi, les points $(\log_2(n/i), \log(Y_{n-i+1,n}))$, $i = 1, \ldots, k_n - 1$ sont approximativement répartis sur une droite de pente β.

Comme dans le domaine d'attraction de Fréchet, il est alors aussi possible d'estimer β par l'estimateur des moindres carrés. L'estimateur $\hat{\beta}_n^Z$ ainsi obtenu est similaire à l'estimateur de l'indice des valeurs extrêmes de Zipf $\hat{\gamma}_n^Z$ dans le cas de lois à queue lourde (voir Définition 1.21).

Un autre estimateur est proposé dans Beirlant *et al.* [10], où β est estimé par :

$$\hat{\beta}_n^{B*} = \frac{\log(n/k_n)}{Y_{n-k_n+1,n}} \frac{1}{k_n - 1} \sum_{i=1}^{k_n-1} (Y_{n-i+1,n} - Y_{n-k_n+1,n}).$$

En remplaçant β par $\hat{\beta}_n^B$ ou $\hat{\beta}_n^{B*}$ dans l'équation (1.41) on obtient deux estimateurs de quantiles extrêmes. L'étude du comportement asymptotique de ces deux estimateurs a été unifiée par Gardes et Girard [81].

Il existe de nombreux estimateurs de β dans la littérature ; détaillons plus précisément deux familles d'estimateurs englobant $\hat{\beta}_n^B$.

Famille d'estimateurs utilisant des poids

La première famille qui nous intéressera est basée sur l'incorporation de poids dans l'estimateur $\hat{\beta}_n^B$ (voir définition 1.23). Elle fut introduite par Gardes et Girard [83]. Les estimateurs appartenant à cette famille sont des combinaisons linéaires de statistiques

d'ordre c'est-à-dire des L-estimateurs. Plus précisément, on considère la famille d'estimateurs $\mathcal{B}_1 = \{\hat{\beta}_n(\zeta),\ \zeta = (\zeta_{1,n}, \ldots, \zeta_{k_n-1,n})\}$ avec :

$$\hat{\beta}_n(\zeta) = \sum_{i=1}^{k_n-1} \zeta_{i,n}(\log(Y_{n-i+1,n}) - \log(Y_{n-k_n+1,n})) \Big/ \sum_{i=1}^{k_n-1} \zeta_{i,n}(\log_2(n/i) - \log_2(n/k_n)) \ , \tag{1.43}$$

où les $\zeta_{i,n}$ sont des poids.

En prenant $\zeta_{i,n} = 1$ pour tout $i = 1, \ldots, k_n - 1$ on retrouve l'estimateur $\hat{\beta}_n^B$ proposé par Beirlant *et al.* [19] (voir l'équation (1.42)).

L'estimateur des moindres carrés $\hat{\beta}_n^Z$ appartient également à la famille $\mathcal{B}_1$ (voir [83, Corollaire 2]) avec les poids :

$$\zeta_{i,n}^Z = \log_2(n/i) - \frac{1}{k_n - 1} \sum_{i=1}^{k_n-1} \log_2(n/i).$$

Famille d'estimateurs utilisant des suites de normalisation

Dans l'estimateur de Beirlant $\hat{\beta}_n^B$ (voir l'équation (1.42)), la somme des écarts entre les logarithmes des statistiques d'ordre est normalisée par la suite :

$$S_n^{(B)} = \sum_{i=1}^{k_n-1} (\log_2(n/i) - \log_2(n/k_n)). \tag{1.44}$$

On peut remplacer cette suite $(S_n^{(B)})$ par une suite positive quelconque (S_n) vérifiant :

$$S_n k_n \log(n/k_n) \to 1. \tag{1.45}$$

Cela nous permet de définir une nouvelle famille d'estimateurs :

$$\mathcal{B}_2 = \{\hat{\beta}_n(S_n),\ S_n > 0 \text{ vérifiant } (1.45)\}$$

avec

$$\hat{\beta}_n(S_n) = \frac{1}{S_n} \sum_{i=1}^{k_n-1} (\log(Y_{n-i+1,n}) - \log(Y_{n-k_n+1,n})). \tag{1.46}$$

Cette famille a été introduite par Gardes et Girard [82] où il est aussi donné un résultat de normalité asymptotique.

Le terme de variance asymptotique est identique pour tous les estimateurs de la famille. Le biais dépend lui par contre du choix de la suite (S_n). Il n'existe cependant pas de

choix annulant ce dernier. Il n'y a pas d'estimateur qui soit préférable dans toutes les situations envisageables. L'efficacité des différents estimateurs est illustrée sur simulations dans Gardes et Girard [87].

Un estimateur de β débiaisé a été proposé dans Diebolt *et al.* [59]. Il est basé sur un modèle de régression exponentielle inspiré de ceux proposés par Beirlant *et al.* [12, 13] et Feuerverger et Hall [76] afin d'estimer l'indice des valeurs extrêmes pour des lois du domaine d'attraction de Fréchet. Diebolt *et al.* [60] proposent alors de remplacer β par cet estimateur débiaisé [59] et obtiennent ainsi un estimateur de quantiles extrêmes débiaisé.

Pour une synthèse bibliographique plus poussée sur l'estimation de quantiles extrêmes pour les lois à queue de type Weibull voir Gardes et Girard [87].

1.6 Estimation de quantiles conditionnels extrêmes

Dans cette dernière partie, nous nous intéressons au cas où la variable aléatoire d'intérêt Y est mesurée conjointement avec une covariable x. Contrairement au problème d'estimation des quantiles de régression classique [124] qui a été largement considéré, peu d'attention a été accordée aux quantiles conditionnels extrêmes.

Ce problème n'a été considéré que récemment [46], et ce malgré de nombreuses applications notamment en finance et en hydrologie. En effet, l'ajout d'une covariable permet de modéliser plus de phénomènes. Par exemple en hydrologie, la variable Y peut représenter le niveau de pluie journalier en millimètres tombée en un point géographique caractérisé par sa position x=(latitude, longitude, altitude) [89]. Un autre exemple d'application possible est l'astrophysique, où Y représente la quantité d'un certain paramètre physique et x une courbe hyperspectrale [44].

Lorsque l'on travaille en présence d'une covariable on distingue deux cas selon la nature de cette dernière :

- Le modèle dit "à plan aléatoire" ou "design aléatoire" pour lequel les données sont des couples $\{(X_i, Y_i), i = 1, \ldots, n\}$ de variables aléatoires réelles indépendantes et de même loi. Dans ce cas, on peut définir la fonction de répartition conditionnelle $F(.|x)$ par :
$$F(y|x) = \mathbb{P}(Y < y|X = x).$$
- Le modèle dit "à plan fixe" ou "design fixe" dont les données sont des couples $\{(x_i, Y_i), i = 1, \ldots, n\}$ où les observations Y_i sont des variables aléatoires réelles indépendantes et les x_i sont des points d'observations non aléatoires. Dans ce cas, la

fonction de répartition de la variable d'intérêt dépend de x et est notée également $F(.|x)$ par analogie avec le cas précédent.

Définition 1.24. *Le quantile conditionnel extrême d'ordre* $1-\alpha_n$ *de* $F(.|x)$ *est défini par :*

$$q(\alpha_n|x) := \overline{F}^{\leftarrow}(\alpha_n|x) \quad \textit{avec} \quad \alpha_n \to 0 \quad \textit{quand} \quad n \to \infty.$$

Par analogie avec le cas sans covariable il est difficile d'estimer $q(\alpha_n|x)$ au-delà du point maximal du sous-échantillon des observations de Y prises dans un voisinage de x.

En présence d'une covariable, les paramètres des lois GEV et GPD dépendent de x. On a pu voir tout au long de ce chapitre la similitude qui existe entre les nombreux estimateurs des quantiles extrêmes dans le cas sans covariable. Dans cette partie, on se concentre sur l'approche POT (voir Partie 1.3) où dans ce cas la loi GPD dépendra des trois fonctions $u(x)$, $\sigma(x)$ et $\gamma(x)$. La fonction $\gamma(x)$ est appelée indice des valeurs extrêmes conditionnel ou indice de queue conditionnel.

Par analogie avec l'estimateur des quantiles extrêmes pour la loi GPD (voir la Définition 1.15) on a l'estimateur du quantile conditionnel extrême donné dans la définition suivante.

Définition 1.25. *L'estimateur du quantile conditionnel extrême est défini par :*

$$\hat{q}_n^{gpd}(\alpha_n|x) = \hat{u}_n(x) + \frac{\hat{\sigma}_n(x)}{\hat{\gamma}_n(x)}\left(\left(\frac{k_n}{n\alpha_n}\right)^{\hat{\gamma}_n(x)} - 1\right),$$

où $\hat{\gamma}_n(x)$, $\hat{u}_n(x)$*et* $\hat{\sigma}_n(x)$ *sont des estimateurs respectifs de* $\gamma(x)$, $u(x)$ *et de* $\sigma(x)$.

Lorsque l'on souhaite estimer des quantiles extrêmes dans le cas particulier de lois à queues lourdes conditionnelles, on retrouve l'estimateur de Weissman (voir Définition 1.17) adapté au cas conditionnel.

Définition 1.26. *L'estimateur de Weissman adapté au cas conditionnel est défini par :*

$$\hat{q}_n^W(\beta_n|x) = \hat{q}_n(\alpha_n|x)\left(\frac{\alpha_n}{\beta_n}\right)^{\hat{\gamma}_n(x)},$$

où dans ce cas $\hat{q}_n(\alpha_n|x)$ correspond à l'estimation d'un quantile conditionnel choisi dans un sous-échantillon des observations prises dans un voisinage de x.

Afin d'estimer des quantiles extrêmes conditionnels le but est alors comme précédemment d'estimer $\gamma(x), \sigma(x)$ et $u(x)$.

1.6.1 L'approche paramétrique

L'ajout d'une covariable est relativement récent dans la littérature. Les premiers travaux sont dus à Smith [159] qui propose de modéliser les maxima par une loi des valeurs extrêmes dont les paramètres sont des fonctions de la covariable. Une forme paramétrique est supposée sur ces fonctions et l'estimation est effectuée par maximum de vraisemblance ou par moindres carrés. Ce fut ensuite au tour de Davison et Smith [47] de proposer de modéliser les excès par une loi de Pareto généralisée dont les paramètres sont des fonctions de la covariable, les auteurs estimant ces derniers par maximum de vraisemblance.

1.6.2 L'approche semi-paramétrique

Beirlant et Goegebeur [14] ont été les premiers à utiliser une approche semi-paramétrique. Ils proposent de transformer tout d'abord les données dans le but d'obtenir des résidus suivant une loi de type Pareto puis de les utiliser dans un modèle de régression exponentielle où les paramètres du dit modèle sont estimés par la méthode du maximum de vraisemblance.

Dans le cas d'une covariable unidimensionnelle, Beirlant et Goegebeur [15] se proposent d'adapter les estimateurs de quantiles proposés par Matthys et Beirlant [132] au cas conditionnel en remplaçant les statistiques d'ordres par les quantiles estimés par la méthode des polynômes locaux, voir [124]. Ils proposent également une procédure pour choisir leur paramètre de lissage et leur seuil.

Wang et Tsai [166] estiment l'indice de queue conditionnel dans le cas d'une loi à queue lourde en utilisant une fonction de lien logarithme afin de relier l'indice de queue conditionnel au prédicteur linéraire induit par la covariable, ce qui constitue le modèle de régression pour l'indice de queue conditionnel. Ils estiment ensuite les paramètres de regression par maximum de vraisemblance. A la suite de ces travaux, Wang *et al.* [165] ont estimé des quantiles conditionnels extrêmes dans le cas de lois issues du modèle de Hall [111]. Dans ce cadre d'étude, Hall [110] a aussi proposé des estimateurs semi-paramétriques des quantiles conditionnels extrêmes.

1.6.3 L'approche non-paramétrique

Les premiers à proposer une approche non-paramétrique sont Davison et Ramesh [46]. Leur approche est basée sur des polynômes locaux. Beirlant *et al.* [16] ont étendu ces résultats aux covariables multidimensionnelles.

Ce fut ensuite Chavez-Demoulin et Davison [32] qui, en utilisant la méthode du maximum de vraisemblance pénalisé, proposèrent des estimateurs splines de quantiles extrêmes conditionnels dans le cas d'une covariable unidimensionnelle.

Dans le cas de lois à queues lourdes, on dénombre de nombreux travaux :

Tout d'abord Beirlant *et al.* [11] proposèrent d'estimer non-paramétriquement des quantiles conditionnels puis de les utiliser comme statistiques d'ordre supérieur pour estimer des quantiles conditionnels extrêmes.

Lorsque la covariable est déterministe fonctionnelle, Gardes et Girard [86] et Gardes *et al.* [89] ont proposé des estimateurs à noyau des quantiles conditionnels extrêmes. Tandis que Gardes et Girard [85] se sont penchés sur une approche par plus proches voisins dans le cas où la covariable est déterministe.

Dans le cas où la variable d'intérêt dépend d'une covariable aléatoire de dimension finie, Daouia *et al.* [44] proposent une estimation par noyau. Cette méthode a été généralisée à tous les domaines d'attraction par Daouia *et al.* [43] qui se sont basés sur des conditions dues à Von Mises (voir [50]) sur la fonction de survie conditionnelle.

Différents travaux ont eu pour but d'estimer l'indice de queue conditionnel de manière non-paramétrique. Citons Gardes *et al.* [90] qui proposent des estimateurs à noyau dépendant de paramètres de réglage. Dans le cas où la covariable est déterministe, une famille d'estimateurs non-paramétriques a été proposée par Gardes et Girard [84]. Ces estimateurs sont basés sur une moyenne pondérée des log-spacings entre les observations sélectionnées à l'aide d'une approche par fenêtre mobile pour la covariable et d'un seuil aléatoire pour la variable d'intérêt. Plus récemment, Gardes et Stupfler [91] se sont intéressés à l'estimation de l'indice de queue conditionnel en présence d'une covariable aléatoire de dimension finie. Ils proposent un estimateur de Hill local lissé adapté à la présence d'une covariable. Cet estimateur est basé sur un raffinement de l'approche par fenêtre mobile. On peut également citer les travaux de Goegebeur *et al.* [102] qui ont introduit une famille d'estimateurs non-paramétriques de l'indice de queue conditionnel en présence d'une covariable aléatoire.

Dans le but de réduire le biais asymptotique (voir condition du second ordre Paragraphe 1.5.3.1), Goegebeur et de Wet [100] et Goegebeur *et al.* [99] ont estimé le paramètre du second ordre non-paramétriquement en utilisant des estimateurs à noyau.

Afin d'estimer non paramétriquement des quantiles conditionnels extrêmes pour des lois à queue de type Weibull (voir Définition 1.13), Goegebeur *et al.* [98] ont introduit une nouvelle famille d'estimateurs du coefficient à queue de type Weibull dépendant d'une

covariable. Cette famille d'estimateurs à noyau est obtenue en faisant une moyenne pondérée des excès au-delà d'un seuil aléatoire assez grand.

Chapitre 2

Estimation de quantiles extrêmes pour des lois à queues lourdes et légères

Résumé

Dans Gardes et al. *[88], une nouvelle famille de lois est introduite, dépendant de deux paramètres τ et θ, qui englobe aussi bien les lois de type Pareto que les lois à queue de type Weibull. Des estimateurs pour θ et pour les quantiles extrêmes sont aussi proposés, mais ils dépendent tous deux d'un paramètre inconnu τ, les rendant inutiles dans des situations pratiques. On commencera dans la Partie 2.1 par une introduction motivant nos travaux. On rappellera ensuite dans la Partie 2.2 le cadre d'étude et le modèle de Gardes* et al. *[88]. Puis l'on présentera nos contributions théoriques dans la Partie 2.3. Ces contributions consistent dans un premier temps à proposer un estimateur de τ qui est indépendant de θ. A la suite de quoi insérer notre estimateur de τ dans les deux estimateurs précédents nous permet alors d'estimer des quantiles extrêmes pour des lois de type Pareto aussi bien que pour des lois à queue de type Weibull d'une même façon. Les lois asymptotiques de nos trois nouveaux estimateurs sont établies et leur efficacité est illustrée dans la Partie 2.4 sur des données simulées et sur un jeu de données réelles. On finira par des perspectives dans la Partie 2.5. En dernier lieu se trouve la Partie 2.6 contenant les preuves de nos résultats théoriques.*

Sommaire

2.1 Introduction **57**
2.2 Cadre d'étude et présentation du modèle **59**
2.2.1 Modèle 59
2.2.2 Définition des estimateurs et leurs lois asymptotiques 61
2.2.3 Application aux lois à queue de type Weibull 64
2.2.4 Application au domaine d'attraction de Fréchet 65
2.3 Contributions théoriques **66**
2.3.1 Définition des estimateurs 67
2.3.2 Propriétés asymptotiques 68
2.4 Simulations et illustrations sur données réelles **70**
2.4.1 Simulations 70
2.4.2 Application aux crues d'une rivière 80
2.5 Perspectives **83**
2.6 Démonstrations **84**
2.6.1 Résultats préliminaires 84
2.6.2 Preuves des principaux résultats 86
2.6.3 Preuves des résultats préliminaires 91

2.1 Introduction

"In applications the three different families give quite different representations of extreme value behaviour. In early applications of extreme value theory, it was usual to adopt one of the three families, and then to estimate the relevant parameters of that distribution. But there are two weakness : first, a technique is required to choose which of the three families is most appropriate for the data at hand ; second, once such a decision is made, subsequent inferences presume this choice to be correct, and do not allow for the uncertainty such a selection involves, even though this uncertainty may be substantial"

Stuart Coles

Soit $Y_1, \ldots, Y_n$ une suite de variables aléatoires *i.i.d* ayant pour fonction de répartition F et notons $Y_{1,n} \leq \cdots \leq Y_{n,n}$ les statistiques ordonnées associées à cet échantillon. Rappelons que le Théorème 1.4 nous assure que pour une grande partie des fonctions de répartition, le maximum $Y_{n,n}$ (après renormalisation) converge en loi vers une loi des valeurs extrêmes ayant pour paramètre de forme γ. Selon son signe, trois domaines d'attraction sont possibles pour F : Fréchet ($\gamma > 0$), Gumbel ($\gamma = 0$) et Weibull ($\gamma < 0$). Rappelons également que les lois appartenant au domaine d'attraction de Weibull ont une queue droite finie (point terminal fini), ce domaine d'attraction n'est de ce fait pas très utilisé dans la plupart des applications.

Dans ce chapitre, on se concentre sur les domaines d'attraction de Fréchet et Gumbel qui peuvent quant à eux avoir une queue droite infinie (point terminal infini). Rappelons qu'une sous famille intéressante du domaine d'attraction de Gumbel est la famille de lois à queue de type Weibull (voir sous-Partie 1.4.4 et plus particulièrement la Définition 1.13).

Comment choisir entre les domaines d'attraction que sont celui de Gumbel et celui de Fréchet pour modéliser le maximum d'un phénomène ? A-t-on une décroissance de la fonction de survie exponentielle ou polynômiale ? Si on se trouve dans le domaine d'attraction de Fréchet quelle est la valeur de l'indice des valeurs extrêmes γ ? Si on se trouve dans le domaine d'attraction de Gumbel, et plus particulièrement dans le cas d'une loi à queue de type Weibull, quelle est la valeur du coefficient β de queue de type Weibull ?

Gardes *et al.* [88] ont étudié plusieurs estimateurs de l'indice de queue et de quantiles extrêmes. Ils ont ainsi pu remarquer la grande similarité qui existe entre la plupart des estimateurs du coefficient à queue de type Weibull (par exemple l'estimateur de Beirlant voir Définition 1.42) et l'estimateur de Hill $H_n(k_n)$ (voir Définition 1.18). Cette similarité

peut sembler surprenante puisque $H_n(k_n)$ a été construit dans le but d'estimer l'indice de queue γ pour des lois appartenant au domaine d'attraction de Fréchet.

Dans l'optique de faire de l'inférence sur la queue de la loi, la plupart des méthodes ou approches consistent à utiliser les k_n statistiques d'ordre supérieures $Y_{n-k_n+1,n} \leq \cdots \leq Y_{n,n}$, c'est le cas de l'estimateur de Hill. Cela semble naturel car l'information de la queue est seulement contenue dans la partie extrême supérieure de l'échantillon. Rappelons que (k_n) est une suite intermédiaire d'entiers (voir l'expression (1.34)).

En ce qui concerne l'estimation du coefficient à queue de type Weibull, des méthodes dédiées furent proposées. Historiquement, la première approche est due à Berred [20] qui proposa un estimateur basé sur des valeurs records. Rappelons que deux grandes familles d'approches se sont développées.

– Une première famille d'approches (Beirlant *et al.* [9, 10], Broniatowski [27], Dierckx *et al.* [63]) est basée sur les logarithmes des excès au-dessus d'un seuil (log-excesses) des k_n statistiques d'ordre supérieur :

$$\log(Y_{n-i+1,n}) - \log(Y_{n-k_n+1,n}).$$

– Une seconde famille d'approches (Beirlant *et al.* [19], Diebolt *et al.* [59], Gardes et Girard [82, 83], Girard [93], Goegebeur *et al.* [98], Goegebeur et Guillou [101]) se base quant à elle sur les espacements entre les logarithmes (log-spacings) des k_n statistiques d'ordre supérieur :

$$\log(Y_{n-i+1,n}) - \log(Y_{n-i,n}).$$

Tous ces estimateurs sont similaires à l'estimateur de Hill. En effet $H_n(k_n)$ (voir Définition 1.18) peut peut être vu aussi bien en termes de log-spacings :

$$H_n(k_n) = \frac{1}{k_n - 1} \sum_{i=1}^{k_n-1} \log(Y_{n-i+1,n}) - \log(Y_{n-k_n+1,n}),$$

qu'en termes de log-excesses :

$$H_n(k_n) = \frac{1}{k_n - 1} \sum_{i=1}^{k_n-1} i \left(\log(Y_{n-i+1,n}) - \log(Y_{n-i,n})\right).$$

Comme exemple, on citera les trois estimateurs proposés dans Gardes et Girard [82] qui sont proportionnels à $H_n(k_n)$.

Pour expliquer ce résultat, Gardes *et al.* [88] introduirent en 2011 un modèle contenant une nouvelle famille de lois englobant notamment aussi bien tout le domaine d'attraction de Fréchet que les lois à queue de type Weibull. Ainsi, ils proposent d'unifier les domaines

d'attraction de Fréchet et de Gumbel à travers une nouvelle famille de lois et d'inclure ainsi les lois à queue de type Weibull dans une famille de lois plus générale.

Dans la partie suivante, on exposera le cadre d'étude et le modèle de Gardes *et al.* [88]. On présentera ensuite nos contributions théoriques dans la Partie 2.3. Les lois asymptotiques de nos estimateurs sont établies et leur efficacité est illustrée dans la Partie 2.4 sur des données simulées et sur un jeu de données réelles. On finira par des perspectives dans la Partie 2.5. En dernier lieu, se trouve la Partie 2.6 contenant les preuves de nos résultats théoriques.

2.2 Cadre d'étude et présentation du modèle

Cette Partie 2.2 est une présentation du modèle de Gardes *et al.* [88] ; le but principal de cette dernière étant d'expliquer pourquoi les statistiques basées sur les log-spacings peuvent être utiles pour estimer les indices de queue aussi bien pour les lois à queue de type Weibull que pour les lois de type Pareto.

2.2.1 Modèle

Dans le but de comprendre la similarité qui existe entre la plupart des estimateurs du coefficient à queue de type Weibull et l'estimateur de Hill, une nouvelle famille de lois a été proposée dans Gardes *et al.* [88]. Cette famille dépend de deux paramètres $\tau \in [0,1]$ et $\theta > 0$. Plus précisément, la famille de fonctions de survie considérée est donnée par :

$$(\mathbf{A}_1(\tau,\theta)) \quad \overline{F}(y) = \exp(-K_\tau^{\leftarrow}(\log H(y))) \text{ avec } y \geq y_* > 0$$

où

$$K_\tau(y) = \int_1^y u^{\tau-1} du \quad \text{avec} \quad \tau \in [0,1].$$

Ici, H est une fonction croissante telle que $H^{\leftarrow} \in \mathcal{RV}_\theta$ où $\theta > 0$. Le paramètre τ nous permet quant à lui de représenter un large panel de lois allant des lois à queue de type Weibull aux lois appartenant au domaine d'attraction de Fréchet, en cela il est proche du paramètre de l'indice des valeurs extrêmes γ.

Mettons tout d'abord en avant le fait que la lourdeur de la queue de la fonction de survie $\overline{F}$ est principalement contrôlée par $\tau \in [0,1]$ et ensuite par $\theta > 0$.

Proposition 2.1. *Soient $\overline{F}_{\tau_1,\theta_1}$ et $\overline{F}_{\tau_2,\theta_2}$ deux fonctions de survie satisfaisant respectivement $(\mathbf{A}_1(\tau_1,\theta_1))$ et $(\mathbf{A}_1(\tau_2,\theta_2))$.*

(i) Si $\tau_1 < \tau_2$ *alors* $\overline{F}_{\tau_1,\theta_1}(y)/\overline{F}_{\tau_2,\theta_2}(y) \to 0$ *quand* $y \to \infty$ *pour tout* $(\theta_1, \theta_2) \in]0, \infty[^2$.

(ii) Si $\tau_1 = \tau_2 = \tau$ *et* $\theta_1 < \theta_2$ *alors* $\overline{F}_{\tau,\theta_1}(y)/\overline{F}_{\tau,\theta_2}(y) \to 0$ *quand* $y \to \infty$.

Ainsi, plus grand sera τ, plus lourde sera la queue. En effet, considérons les deux cas extrêmes $\tau = 0$ et $\tau = 1$.

- Clairement, sous $(\mathbf{A}_1(0, \theta))$, on a $K_0(y) = \log(y)$ ainsi $K_0^{\leftarrow}(y) = \exp(y)$ d'où :

$$\overline{F}(y) = \exp(-H(y)) = \exp(-y^{1/\theta} L(y)),$$

  où L est une fonction à variations lentes à l'infini. On est dans le cas d'une fonction de survie d'une loi à queue de type Weibull où θ coïncide avec l'indice de queue de type Weibull β (voir Définition 1.13).

- A l'opposé, $(\mathbf{A}_1(1, \theta))$ implique $K_1(y) = y - 1$ ainsi $K_0^{\leftarrow}(y) = y + 1$ d'où :

$$\overline{F}(y) = e^1/H(y) = y^{-1/\theta} \tilde{\ell}(y),$$

  où $\tilde{\ell}$ est une fonction à variations lentes à l'infini. Par conséquent, F appartient au domaine d'attraction de Fréchet et θ coïncide avec l'indice de queue $\gamma > 0$ (voir l'équation (1.13)).

Au vu des remarques précédentes, les valeurs intermédiaires de $\tau \in]0, 1[$ correspondent à des lois ayant des queues plus légères que des queues de type Pareto mais plus lourdes que des queues de type Weibull. En effet, on a :

$$\overline{F}(y) = \exp(-h(y)) \quad \text{avec} \quad h(y) \sim ((\tau/\theta)\log(y))^{1/\tau},$$

et ainsi $h(y)/y^a \to 0$ pour tout $a > 0$ tandis que $h(y)/\log(y) \to \infty$ quand $y \to \infty$, cette propriété caractérise une loi de "type exponentielle", voir Harris [114].

La proposition suivante nous donne une caractérisation plus précise et des exemples sont donnés.

Proposition 2.2.

(i) F *vérifie* $(\mathbf{A}_1(0, \theta))$ *si et seulement si* F *est une loi à queue de type Weibull avec comme coefficient de queue de Weibull* θ.

(ii) Si F *vérifie* $(\mathbf{A}_1(\tau, \theta))$, $\tau \in [0, 1[$ *et si* H *est deux fois dérivable alors* F *appartient au domaine d'attraction de Gumbel.*

(iii) F *vérifie* $(\mathbf{A}_1(1, \theta))$ *si et seulement si* F *est dans le domaine d'attraction de Fréchet avec un indice de queue égal à* θ.

Au vu de la Proposition 2.2(i), les lois : Normale, Gamma, Weibull, Benktander II, Logistique et la loi des valeurs extrêmes vérifient toutes $(\mathbf{A}_1(0,\theta))$ puisqu'elles sont des exemples de lois à queue de type Weibull (voir Gardes et Girard [83], Tableau 1).

Des exemples de lois vérifiant $(\mathbf{A}_1(\tau,\theta))$ avec $\tau \in]0,1[$ incluent certaines lois à queue de type Log-weibull. Rappelons qu'une variable aléatoire Y suit une loi à queue de type Log-weibull si $\log(Y)$ suit une loi à queue de type Weibull.

Proposition 2.3. *Supposons que F vérifie $(\mathbf{A}_1(0,\theta))$ avec $\theta \in (0,1]$. Si, de plus, la fonction à variations lentes ℓ est dérivable et $\ell(t) \to \ell_\infty > 0$ quand $t \to \infty$ alors $F(\log .)$ verifie $(\mathbf{A}_1(\theta, \theta\ell_\infty))$.*

Comme exemple, on citera la loi Log-normale standard qui peut être vue comme une loi à queue de type Log-weibull et vérifiant alors $(\mathbf{A}_1(1/2, \sqrt{2}/2))$. De même, la loi Gamma vérifie $(\mathbf{A}_1(0,1))$ et la loi Log-gamma appartient au domaine d'attraction de Fréchet, voir par exemple Embrechts *et al.* [71], Tableau 3.4.2. Enfin, d'autres exemples de lois satisfaisant $(\mathbf{A}_1(1,\theta))$ se trouvent dans le tableau mentionné.

Comme on l'a illustré dans les parties précédentes, le modèle $(\mathbf{A}_1(\tau,\theta))$ nous donne un nouvel outil pour l'analyse des estimateurs de queue basés sur les log-spacings. Cela nous permet d'englober les lois à queue de type Weibull dans un cadre plus général et d'expliquer ainsi pourquoi les estimateurs dédiés à l'estimation de leur indice de queue sont très similaires à l'estimateur de Hill (voir Définition 1.18) et dans le cas de quantiles extrêmes à l'estimateur de Weissman (voir Définition 1.17).

Nous allons maintenant présenter un estimateur $\widehat{\theta}_{n,\tau}(k_n)$ de θ pour cette nouvelle famille de lois et l'estimateur des quantiles extrêmes qui s'en déduit. La normalité asymptotique de ces estimateurs est établie dans la sous-Partie suivante 2.2.2 d'une seule et même façon.

2.2.2 Définition des estimateurs et leurs lois asymptotiques

Dans Gardes *et al.* [88], un estimateur de θ basé sur l'estimateur de Hill est introduit. En notant par (k_n) une suite intermédiaire d'entiers, cet estimateur de θ est donné par :

$$\widehat{\theta}_{n,\tau}(k_n) = \frac{1}{\mu_\tau(\log(n/k_n))} \frac{1}{k_n - 1} \sum_{i=1}^{k_n-1} (\log(Y_{n-i+1,n}) - \log(Y_{n-k_n+1,n})) \tag{2.1}$$

avec, pour tout $t > 0$,

$$\mu_\tau(t) = \int_0^\infty (K_\tau(y+t) - K_\tau(t))\, \mathrm{e}^{-y} dy.$$

Un point crucial est que l'estimateur (2.1) consiste essentiellement à faire la moyenne des log-excesses entre les statistiques ordonnées supérieures. Il est important de noter que $\widehat{\theta}_{n,\tau}(k_n)$ diffère seulement de l'estimateur de Hill (voir Définition 1.18) par une suite normalisée déterministe :

$$\widehat{\theta}_{n,\tau}(k_n) = \frac{H_n(k_n)}{\mu_\tau(\log(n/k_n))}.$$

Cette similarité peut être vue à travers l'étude des log-spacings entre deux quantiles $q(u)$ et $q(v)$ de $\overline{F}$, avec $0 < u < v \leq 1$. Sous $(\mathbf{A}_1(\tau,\theta))$ on a :

$$\log q(u) - \log q(v) = \theta\left(K_\tau(-\log u) - K_\tau(-\log v)\right) + \log\left(\frac{\ell(\exp K_\tau(-\log u))}{\ell(\exp K_\tau(-\log v))}\right). \quad (2.2)$$

Puisque ℓ est une fonction à variations lentes, si les ordres u et v des quantiles sont assez petits, le second terme peut être négligé dans la partie droite de (2.2) pour obtenir :

$$\log q(u) - \log q(v) \simeq \theta\left(K_\tau(-\log u) - K_\tau(-\log v)\right), \quad (2.3)$$

ce qui nous montre que les log-spacings sont approximativement proportionnels à θ. Puisque cette propriété est vérifiée pour tout $\tau \in [0,1]$, elle est ainsi partagée par les lois : de type Pareto, à queue de type Weibull et à queue de type Log-weibull. Notons que cette propriété peut être interprétée graphiquement sur un échantillon en traçant un graphique quantile-quantile. Cela consiste à tracer les points :

$$(K_\tau(\log(n/i)), \log(Y_{n-i+1,n})) \quad \text{pour} \quad i = 1, \ldots, k_n.$$

D'après (2.3), le graphique devrait être approximativement linéaire. De même, un estimateur des quantiles extrêmes (voir Définition 1.2) peut être déduit d'après (2.1) par :

$$\widehat{q}_n\left(\alpha_n, \widehat{\theta}_{n,\tau}(k_n)\right) = Y_{n-k_n+1,n} \exp\left(\widehat{\theta}_{n,\tau}(k_n)\left(K_\tau(\log(1/\alpha_n)) - K_\tau(\log(n/k_n))\right)\right). \quad (2.4)$$

Soulignons le fait que les estimateurs (2.1) et (2.4) sont d'un intérêt seulement théorique puisque dans des situations pratiques τ est inconnu, ce qui ne permet donc pas de les utiliser.

On montre dans la suite que la normalité asymptotique de $\widehat{\theta}_{n,\tau}(k_n)$ et $\widehat{q}_n\left(\alpha_n, \widehat{\theta}_{n,\tau}(k_n)\right)$ peut être établie pour tout $\tau \in [0,1]$ d'une même façon (résultats du travail de Gardes *et al.* [88]). En ce sens, le comportement asymptotique de ces estimateurs est plus une conséquence de la propriété des log-spacings que du comportement de la queue de la loi (qui peut être exponentiel aussi bien que polynômial). Les sous-Parties 2.2.3 et 2.2.4 illustrent nos résultats sur ces deux cas extrêmes $\tau = 0$ et $\tau = 1$.

Pour établir la normalité asymptotique de $\widehat{\theta}_{n,\tau}(k_n)$, on introduit une condition de second ordre sur ℓ.

Condition du second ordre : $(\mathbf{A}_2(\rho))$

Il existe $\rho < 0$ et $b(y) \underset{y\to\infty}{\longrightarrow} 0$ tels que uniformément localement pour $\lambda \geq \lambda_0 > 0$:

$$\log\left(\frac{\ell(\lambda y)}{\ell(y)}\right) \sim b(y)K_\rho(\lambda), \quad \text{lorsque} \quad y \to \infty,$$

avec $|b|$ asymptotiquement décroissante.

Rappelons que nécessairement la fonction b (appelée aussi fonction de biais) est à variations régulières d'indice $\rho < 0$ (voir par exemple Geluk et de Haan [92]). Rappelons également que le paramètre de second ordre $\rho < 0$ contrôle la vitesse de convergence du rapport $\ell(\lambda y)/\ell(y)$ vers 1. Une valeur de ρ proche de 0 implique une faible vitesse de convergence.

La condition $(\mathbf{A}_2(\rho))$ est la clef de voûte de toutes les preuves de normalité asymptotique en théorie des valeurs extrêmes. Elle est utilisée dans Beirlant *et al.* [12], Häusler et Teugels [115] et Hill [116] pour prouver la normalité asymptotique d'un grand nombre d'estimateurs de l'indice des valeurs extrêmes. Notons que la condition du second ordre utilisée pour prouver la normalité asymptotique de Hill diffère quelque peu de celle-ci. En effet, on n'a pas besoin que $|b|$ soit asymptotiquement décroissante (voir la condition du second ordre au Paragraphe 1.5.3.1).

Ainsi les auteurs Gardes *et al.* [88] obtiennent la normalité asymptotique de $\widehat{\theta}_{n,\tau}(k_n)$.

Théorème 2.4. *Supposons que les hypothèses* $(\mathbf{A}_1(\tau,\theta))$ *et* $(\mathbf{A}_2(\rho))$ *soient vérifiées. Soit* (k_n) *une suite intermédiaire d'entiers telle que :*

$$\sqrt{k_n}\, b(\exp K_\tau(\log(n/k_n))) \to \lambda. \tag{2.5}$$

Alors, en introduisant $a_{\tau,\rho} = 1$ *si* $\tau \in [0,1)$ *et* $a_{1,\rho} = 1/(1-\rho)$*, on a :*

$$\sqrt{k_n}\left(\widehat{\theta}_{n,\tau}(k_n) - \theta - a_{\tau,\rho} b(\exp K_\tau(\log(n/k_n)))\right) \xrightarrow{d} \mathcal{N}(0,\theta^2). \tag{2.6}$$

Il apparaît que la variance asymptotique de $\widehat{\theta}_{n,\tau}(k_n)$ donnée par $\mathcal{AV} = \theta^2/k_n$ est indépendante de τ. En particulier, elle reste constante quel que soit le domaine d'attraction de F. Le biais asymptotique au carré est donné par $\mathcal{ASB}(\tau,\rho) = a^2_{\tau,\rho} b^2(\exp K_\tau(\log(n/k_n)))$. De plus, comme b^2 décroit asymptotiquement, alors $\mathcal{ASB}$ est une fonction décroissante de $\tau \in [0,1[$ avec un saut à $\tau = 1$. Ces remarques sont illustrées sur des données simulées dans l'article de Gardes *et al.* [88].

Le résultat suivant établi par Gardes *et al.* [88] permet d'établir la vitesse de convergence de $\widehat{\theta}_{n,\tau}(k_n)$ vers θ dans (2.6).

Proposition 2.5. *La condition (2.5) avec $\lambda \neq 0$ implique :*

$$\log(k_n) = -2\rho a_{\tau,2\rho} K_\tau(\log n)(1 + o(1)).$$

La vitesse de convergence est ainsi d'ordre $\exp(-\rho a_{\tau,2\rho} K_\tau(\log n)(1 + o(1)))$. Une vitesse de convergence géométrique est obtenue seulement dans le domaine d'attraction de Fréchet. En effet, $\tau = 1$ implique $\sqrt{k_n} = n^{-\rho/(1-2\rho)+o(1)}$ qui est consistant avec les conclusions de de Haan et Peng [52]. Les lois à queue de type Weibull donnent des vitesses de convergence logarithmiques. En effet, $\tau = 0$ implique $\sqrt{k_n} = (\log n)^{-\rho+o(1)}$ qui est consistant avec les résultats de Gardes et Girard [82]. Plus généralement, plus lourde sera la queue, meilleure sera la vitesse de convergence.

Gardes *et al.* [88] déduisent alors du Théorème 2.4 la normalité asymptotique de l'estimateur des quantiles extrêmes (2.4).

Théorème 2.6. *Supposons que les hypothèses du Théorème 2.4 soient vérifiées avec $\lambda = 0$. Si, de plus,*

$$(\log(n/k_n))^{1-\tau}(K_\tau(\log(1/\alpha_n)) - K_\tau(\log(n/k_n))) \to \infty \tag{2.7}$$

alors on a :

$$\frac{\sqrt{k_n}}{K_\tau(\log(1/\alpha_n)) - K_\tau(\log(n/k_n))} \left(\frac{\widehat{q}_n\left(\alpha_n, \widehat{\theta}_{n,\tau}(k_n)\right)}{q(\alpha_n)} - 1 \right) \xrightarrow{d} \mathcal{N}(0, \theta^2).$$

Concentrons-nous sur les deux cas particuliers $\tau = 0$ (lois à queue de type Weibull) et $\tau = 1$ (domaine d'attraction de Fréchet).

2.2.3 Application aux lois à queue de type Weibull

Si $\tau = 0$, l'estimateur (2.1) coïncide avec $\widehat{\theta}_n^{(1)}$ introduit dans Gardes et Girard [82]. Notons également que les estimateurs $\widehat{\theta}_n^{(2)}$ et $\widehat{\theta}_n^{(3)}$ dans Gardes et Girard [82] peuvent être respectivement déduits de $\widehat{\theta}_{n,0}(k_n)$ en approchant μ_0 par une somme de Riemann ou en utilisant l'approximation du premier ordre $\mu_0(t) \sim 1/t$ lorsque $t \to \infty$ donnée dans le Lemme 2(i) de Gardes *et al.* [88].

Par ailleurs :

$$\widehat{q}_n\left(\alpha_n, \widehat{\theta}_{n,0}(k_n)\right) = Y_{n-k_n+1,n} \left(\frac{\log(1/\alpha_n)}{\log(n/k_n)} \right)^{\widehat{\theta}_{n,0}(k_n)} \tag{2.8}$$

est l'estimateur proposé dans Gardes et Girard [81]. La forme de l'estimateur (2.8) est similaire à celle de la famille d'estimateurs de quantiles extrêmes dans le cas de loi à queue de type Weibull que l'on a présenté dans le chapitre 1 (voir Définition 1.41). Comme conséquence du Théorème 2.4 et Théorème 2.6, Gardes *et al.* [88] obtiennent le corollaire suivant.

Corollaire 2.7. *Supposons que* $(\mathbf{A}_1(0,\theta))$ *et* $(\mathbf{A}_2(\rho))$ *soient vérifiées. Soit* (k_n) *une suite intermédiaire d'entiers telle que* $\sqrt{k_n}\, b(\log(n/k_n)) \to 0$. *Alors,*

$$\sqrt{k_n}\,(\widehat{\theta}_{n,0}(k_n) - \theta) \xrightarrow{d} \mathcal{N}(0,\theta^2).$$

Si, de plus :

$$\log(n/k_n)\,(\log_2(1/\alpha_n) - \log_2(n/k_n)) \to \infty \tag{2.9}$$

alors,

$$\frac{\sqrt{k_n}}{\log_2(1/\alpha_n) - \log_2(n/k_n)} \left(\frac{\widehat{q}_n\left(\alpha_n, \widehat{\theta}_{n,0}(k_n)\right)}{q(\alpha_n)} - 1 \right) \xrightarrow{d} \mathcal{N}(0,\theta^2).$$

Ces résultats sont très semblables au Corollaire 3.1 dans Gardes et Girard [82] excepté que la condition (2.9) est plus faible que celle utilisée dans l'article mentionné.

2.2.4 Application au domaine d'attraction de Fréchet

Soit $\tau = 1$ remarquons que $\mu_1(t) = 1$ pour tout $t > 0$, l'estimateur (2.1) coïncide alors avec l'estimateur de Hill (voir Définition 1.18) de l'indice de queue. Par ailleurs :

$$\widehat{q}_n\left(\alpha_n, \widehat{\theta}_{n,1}(k_n)\right) = Y_{n-k_n+1,n} \left(\frac{k_n}{n\alpha_n}\right)^{\widehat{\theta}_{n,1}(k_n)} \tag{2.10}$$

est l'estimateur de Weissman (voir Définition 1.17). La forme de l'estimateur (2.10) est similaire à celle de la famille d'estimateurs de quantiles extrêmes dans le cas de loi appartenant au domaine d'attraction de Fréchet que l'on a présenté dans le chapitre 1 (voir Définition 1.17). Une application directe des théorèmes précédents donne les résultats classiques donnés dans le corollaire suivant.

Corollaire 2.8. *Supposons que* $(\mathbf{A}_1(1,\theta))$ *et* $(\mathbf{A}_2(\rho))$ *soient vérifiées. Soit* (k_n) *une suite intermédiaire d'entiers telle que* $\sqrt{k_n}\, b(n/k_n) \to 0$. *Alors :*

$$\sqrt{k_n}(\widehat{\theta}_{n,1}(k_n) - \theta) \xrightarrow{d} \mathcal{N}(0,\theta^2).$$

Si, de plus $k_n/(n\alpha_n) \to \infty$ *alors :*

$$\frac{\sqrt{k_n}}{\log(k_n/(n\alpha_n))}\left(\frac{\widehat{q}_n\left(\alpha_n, \widehat{\theta}_{n,1}(k_n)\right)}{q(\alpha_n)} - 1\right) \xrightarrow{d} \mathcal{N}(0, \theta^2).$$

Comme on l'a illustré dans les parties précédentes, le modèle $(\mathbf{A}_1(\tau, \theta))$ donne un nouvel outil pour l'analyse des estimateurs de l'indice de queue basés sur les log-spacings. Cela permet d'englober les lois à queue de type Weibull dans un cadre plus général et d'expliquer ainsi pourquoi les estimateurs dédiés à l'estimation de β sont très similaires à l'estimateur de Hill et ceux du quantile extrême à celui de Weissman.

Les estimateurs (2.1) et (2.4) présentés sont d'un intérêt seulement théorique car, dans des situations pratiques τ étant inconnu, ils ne peuvent être utilisés. Ma contribution sera tout d'abord de proposer un estimateur de τ qui est indépendant de θ. Insérer cet estimateur de τ dans les deux estimateurs précédents permet alors d'estimer des quantiles extrêmes pour des lois de type Pareto aussi bien que pour des lois à queue de type Weibull d'une même façon. La fin de ce chapitre est organisée comme suit. Les estimateurs sont définis dans la sous-Partie 2.3.1, leurs propriétés asymptotiques sont établies dans la sous-Partie 2.3.2. Le comportement asymptotique de l'estimateur des quantiles extrêmes est illustré sur des simulations en sous-Partie 2.4.1 et sur un jeu de données réelles dans la sous-Partie 2.4.2. Dans la Partie 2.5 on présentera quelques perspectives à donner à nos travaux. Les preuves de nos résultats sont exposées dans la Partie 2.6.

2.3 Contributions théoriques

La fin du chapitre constitue ma contribution au modèle. Elle a fait l'objet de l'article El Methni *et al.* [69] paru en 2012. On commence par proposer un estimateur $\widehat{\tau}_n$ de τ, indépendant de θ. En remplaçant τ par $\widehat{\tau}_n$ dans (2.1) et (2.4), on obtient deux nouveaux estimateurs qui peuvent être utilisés dans des situations pratiques. On est ainsi capable d'estimer des quantiles extrêmes de lois allant des lois de type Pareto aux lois à queue de type Weibull. Les lois asymptotiques des trois nouveaux estimateurs sont aussi établies.

2.3.1 Définition des estimateurs

Commençons par décrire la construction de l'estimateur de τ. Soient (k_n) et (k'_n) avec $k'_n > k_n$ deux suites d'entiers telles que

$$\begin{aligned} \widehat{\theta}_{n,\tau}(k_n) &\xrightarrow{\mathbb{P}} \theta \\ \widehat{\theta}_{n,\tau}(k'_n) &\xrightarrow{\mathbb{P}} \theta \end{aligned}$$

Il s'ensuit clairement que :

$$\frac{\widehat{\theta}_{n,\tau}(k_n)}{\widehat{\theta}_{n,\tau}(k'_n)} = \frac{H_n(k_n)}{H_n(k'_n)} \frac{\mu_\tau(\log(n/k'_n))}{\mu_\tau(\log(n/k_n))} \xrightarrow{\mathbb{P}} 1.$$

En introduisant pour tout $t > t' > 0$ la fonction définie par :

$$\psi(.;t;t') : \mathbb{R} \to (-\infty, \exp(t-t')) \quad \text{avec} \quad \psi(y;t,t') = \frac{\mu_y(t)}{\mu_y(t')},$$

il s'ensuit que :

$$\frac{H_n(k_n)}{H_n(k'_n)} \overset{\mathbb{P}}{\sim} \psi(\tau; \log(n/k_n), \log(n/k'_n)). \tag{2.11}$$

De plus, on peut remarquer (voir Lemme 2.15 dans la Partie 2.6) que :

$$\psi(.; \log(n/k_n), \log(n/k'_n))$$

est une bijection de $\mathbb{R}$ dans $(-\infty, k'_n/k_n)$. Par conséquent, on peut considérer l'estimateur de τ suivant :

$$\widehat{\tau}_n = \begin{cases} \psi^{-1}\left(\frac{H_n(k_n)}{H_n(k'_n)}; \log(n/k_n), \log(n/k'_n)\right) & \text{si} \quad \frac{H_n(k_n)}{H_n(k'_n)} < \frac{k'_n}{k_n} \\ u & \text{si} \quad \frac{H_n(k_n)}{H_n(k'_n)} \geq \frac{k'_n}{k_n}, \end{cases} \tag{2.12}$$

où u est la réalisation d'une loi uniforme standard. En pratique, seule la première situation doit être considérée, puisque le Lemme 2.17 de la Partie 2.6 nous dit que, pour n assez grand, $H_n(k_n)/H_n(k'_n)$ est presque sûrement plus petit que k'_n/k_n. Il est ainsi possible de remplacer τ par son estimateur $\widehat{\tau}_n$ dans (2.1) pour obtenir un nouvel estimateur de θ donné par :

$$\widehat{\theta}_{n,\widehat{\tau}_n}(k_n) = \frac{H_n(k_n)}{\mu_{\widehat{\tau}_n}(\log(n/k_n))}.$$

De même, en remplaçant τ et $\widehat{\theta}_{n,\tau}(k_n)$ par leurs estimateurs dans (2.4) cela nous donne le nouvel estimateur des quantiles extrêmes suivant :

$$\widehat{q}_n\left(\alpha_n, \widehat{\theta}_{n,\widehat{\tau}_n}(k_n)\right) = Y_{n-k_n+1,n} \exp\left(\widehat{\theta}_{n,\widehat{\tau}_n}(k_n)\left[K_{\widehat{\tau}_n}(\log(1/\alpha_n)) - K_{\widehat{\tau}_n}(\log(n/k_n))\right]\right).$$

Le comportement asymptotique des trois nouveaux estimateurs $\widehat{q}_n\left(\alpha_n, \widehat{\theta}_{n,\widehat{\tau}_n}(k_n)\right)$, $\widehat{\theta}_{n,\widehat{\tau}_n}$ et $\widehat{\tau}_n$ est établi dans la sous-Partie suivante.

2.3.2 Propriétés asymptotiques

Comme premier résultat, on établit la consistance de $\widehat{\tau}_n$ sous l'hypothèse suivante :

$$H^{\leftarrow}(t) = t^{\theta}\ell(t) = c\,t^{\theta}\exp\left(\int_1^y \frac{\varepsilon(u)}{u}du\right) \tag{2.13}$$

avec c une constante positive et $\varepsilon(s) \to 0$ lorsque $s \to \infty$. Mettons en évidence le fait que (2.13) implique $H^{\leftarrow} \in \mathcal{R}_\theta$ et que la fonction à variations lentes à l'infini ℓ est normalisée, voir Bingham *et al.* [21], page 15.

Proposition 2.9. *Supposons que* $(\mathbf{A}_1(\tau,\theta))$ *est vérifiée avec* ℓ *une fonction à variations lentes normalisée. Si* (k_n) *et* (k'_n) *sont deux suites intermédiaires d'entiers telles que* $k_n/k'_n \to 0$*, alors* $\widehat{\tau}_n \xrightarrow{\mathbb{P}} \tau$.

On notera que la consistance de $\widehat{\tau}_n$ est établie pour tout $\theta > 0$ et $\tau \in [0,1]$. C'est en ce sens, que le comportement asymptotique de cet estimateur est plus une conséquence de la propriété des log-spacings que du comportement de la queue (qui peut être exponentielle pour $\tau = 0$ aussi bien que polynômiale pour $\tau = 1$).

Enfin, pour établir la normalité asymptotique des trois estimateurs $\widehat{q}_n\left(\alpha_n, \widehat{\theta}_{n,\widehat{\tau}_n}(k_n)\right)$, $\widehat{\theta}_{n,\widehat{\tau}_n}$ et $\widehat{\tau}_n$ la condition de second d'ordre $(\mathbf{A}_2(\rho))$ sur ℓ est nécessaire.

Le Théorème 2.10 établit la normalité asymptotique de $\widehat{\tau}_n$.

Théorème 2.10. *Supposons que* $(\mathbf{A}_1(\tau,\theta))$ *et* $(\mathbf{A}_2(\rho))$ *soient vérifiées. Si* (k_n) *et* (k'_n) *sont deux suites intermédiaires d'entiers telles que* $k_n/k'_n \to 0$ *et*

$$\sqrt{k'_n}b(\exp K_\tau(\log n/k'_n)) \to 0, \tag{2.14}$$

$$\sqrt{k_n}\left(\log_2(n/k_n) - \log_2(n/k'_n)\right) \to \infty, \tag{2.15}$$

$$\log(n/k'_n)\left(\log_2(n/k_n) - \log_2(n/k'_n)\right) \to \infty, \tag{2.16}$$

alors

$$\sqrt{k_n}\left(\log_2(n/k_n) - \log_2(n/k'_n)\right)(\widehat{\tau}_n - \tau) \xrightarrow{d} \mathcal{N}(0,1).$$

La condition (2.14) est classique en théorie des valeurs extrêmes. Elle impose que le biais induit par la fonction à variations lentes est asymptotiquement négligeable. La condition (2.15) impose à la vitesse de convergence de $\widehat{\tau}_n$ de tendre vers l'infini. Finalement,

la condition (2.16) est de même nature que la condition (2.15), elle impose un certain espacement minimal entre les deux suites (k_n) et (k'_n).

Par ailleurs, si $\tau = 0$, les conditions (2.14) et (2.15) impliquent que $yb(y) \to 0$ lorsque $y \to \infty$. Par conséquent, si $\tau = 0$ et $\rho > -1$, il n'est pas possible de choisir des suites (k_n) et (k'_n) qui satisfassent les hypothèses précédentes. Dans un tel cas, seule la consistance de $\widehat{\tau}_n$ peut être garantie.

Dans le résultat suivant, il est établi que $\widehat{\theta}_{n,\widehat{\tau}_n}$ hérite de la normalité asymptotique de $\widehat{\tau}_n$.

Théorème 2.11. *Supposons que les hypothèses du Théorème 2.10 soient vérifiées. Si, de plus,*

$$\left(\log_2(n/k_n) - \log_2(n/k'_n)\right) / \log_2(n/k_n) \to 0, \tag{2.17}$$

$$\sqrt{k_n}\left(\log_2(n/k_n) - \log_2(n/k'_n)\right) / \log_2(n/k_n) \to \infty, \tag{2.18}$$

alors

$$\frac{\sqrt{k_n}\left(\log_2(n/k_n) - \log_2(n/k'_n)\right)}{\log_2(n/k_n)} \left(\widehat{\theta}_{n,\widehat{\tau}_n}(k_n) - \theta\right) \xrightarrow{d} \mathcal{N}(0, \theta^2).$$

Il apparait que l'estimation de τ a un coût en termes de vitesse de convergence. La condition (2.17) implique que $\widehat{\theta}_{n,\widehat{\tau}_n}(k_n)$ converge moins vite que $\widehat{\theta}_{n,\tau}(k_n)$, voir Lemme 2.19 dans la Partie 2.6. Comme précédemment, la condition (2.18) impose à la vitesse de convergence de $\widehat{\theta}_{n,\widehat{\tau}_n}(k_n)$ de tendre vers l'infini. On notera que cette condition implique la condition (2.15) dans le Théorème 2.10.

De même que pour le Théorème 2.10, on peut voir que, si $\tau = 0$, les conditions (2.14) et (2.18) impliquent $y \log(y)\, b(y) \to 0$ lorsque $y \to \infty$. Ainsi, encore une fois, il n'existe pas de suites (k_n) et (k'_n) dans le cas où $\tau = 0$ et $\rho > -1$.

Si, $\tau \in]0, 1]$ ou si $\tau = 0$ et $\rho < -1$, un choix possible pour les suites d'entiers est

$$\begin{cases} \log(k_n) & = aK_\tau(\log(n)) \\ \log(k'_n) & = a'K_\tau(\log(n)) \end{cases}$$

avec les restrictions suivantes sur $(a, a') \in \mathbb{R}^2$ données par :

$$\begin{cases} 0 < a < a' < 2\rho/(2\rho - 1) & \text{si} \quad \tau = 1 \\ 0 < a < a' < -2\rho & \text{si} \quad 0 < \tau < 1 \\ 2 < a < a' < -2\rho & \text{si} \quad \tau = 0. \end{cases} \tag{2.19}$$

Enfin, dans le cas où $\tau = 0$ et $\rho = -1$, l'existence de suites (k_n) et (k'_n) dépend de la loi sous jacente.

Le dernier théorème est dédié à la loi asymptotique de l'estimateur des quantiles extrêmes.

Théorème 2.12. *Supposons que les hypothèses du Théorème 2.11 soient vérifiées. Si, de plus,*

$$\sqrt{k_n}(\log_2(n/k_n) - \log_2(n/k'_n))\,/\log_2(1/\alpha_n) \to \infty, \tag{2.20}$$

$$(\log(n/k_n))^{1-\tau}[K_\tau(\log(1/\alpha_n)) - K_\tau(\log(n/k_n))] \to \infty, \tag{2.21}$$

$$\log_2(n/k_n)[K_\tau(\log(1/\alpha_n)) - K_\tau(\log(n/k_n))] \Big/ \int_{\log(n/k_n)}^{\log(1/\alpha_n)} \log(u)u^{\tau-1}du \to 0, \tag{2.22}$$

alors

$$\frac{\sqrt{k_n}\,(\log_2(n/k_n) - \log_2(n/k'_n))}{\int_{\log(n/k_n)}^{\log(1/\alpha_n)} \log(u)u^{\tau-1}du}\left(\frac{\widehat{q}_n\left(\alpha_n, \widehat{\theta}_{n,\widehat{\tau}_n}(k_n)\right)}{q(\alpha_n)} - 1\right) \xrightarrow{d} \mathcal{N}(0, \theta^2).$$

Une condition suffisante pour vérifier les conditions (2.21) et (2.22) est :

$$\log_2(1/\alpha_n)/\log_2(n/k_n) \to \infty.$$

Cela impose une borne supérieure sur l'ordre α_n du quantile extrême. A l'opposé, la condition (2.20) impose une borne inférieure sur α_n.

Un choix possible pour l'ordre α_n du quantile extrême est donné par $\log_2(1/\alpha_n) = [\log_2(n)]^\delta$ pour tout $\delta > 1$. Les performances de $\widehat{q}_n\left(\alpha_n, \widehat{\theta}_{n,\widehat{\tau}_n}(k_n)\right)$ sur des échantillons sont illustrées dans la partie suivante.

2.4 Simulations et illustrations sur données réelles

On commencera par illustrer l'efficacité de nos estimateurs à l'aide de simulations dans la sous-Partie 2.4.1, puis on appliquera nos estimateurs à un jeu de données réelles dans la sous-Partie 2.4.2.

2.4.1 Simulations

Dans cette partie, notre estimateur des quantiles extrêmes $\widehat{q}_n\left(\alpha_n, \widehat{\theta}_{n,\widehat{\tau}_n}(k_n)\right)$ sera comparé à l'approche POT (voir Définition 1.15). Rappelons que l'approche POT se base sur une approximation de la distribution des excès, au-desssus d'un seuil "élevé", par une loi de Pareto généralisée notée GPD. Cette approche a été conçue pour estimer des quantiles extrêmes dans tous les domaines d'attraction.

Parmi les nombreuses méthodes disponibles pour estimer les paramètres de la GPD, on se concentrera sur :
– la méthode des moments (voir sous-Partie 1.5.2),
– l'estimateur DEdH (voir Définition 1.16),
car ce sont celles qui donnent les meilleurs résultats sur nos simulations.

Il est important de signaler que l'estimateur des quantiles extrêmes $\widehat{q}_n\left(\alpha_n, \widehat{\theta}_{n,\widehat{\tau}_n}(k_n)\right)$ a besoin de l'inversion numérique de la fonction ψ. Ce calcul est effectué à l'aide d'une procédure par dichotomie puisque le Lemme 2.15(i) nous assure que ψ est croissante. Les simulations ont été réalisées sur le logiciel R Development Core Team [144]. En ce qui concerne l'estimation des paramètres de la GPD par la méthode des moments, on a utilisé le package de Ribatet [149]

Les simulations ont été effectuées sur les douze lois différentes suivantes :

– cinq lois à queue de type Weibull ($\tau = 0$) : la valeur absolue d'une loi Normale centrée réduite ($\theta = 1/2$ et $\rho = -1$), une loi de Weibull ayant un paramètre de forme égal à 2 ($\theta = 2$ et $\rho = -\infty$), une loi Gamma ayant un paramètre de forme égal à 1 ($\theta = 1$ et $\rho = -1$) et deux lois à queue de type Weibull ayant pour fonction quantile $\overline{F}^{\leftarrow}(y) = (-\log y)^{\theta}(1 + (\rho+\theta)(-\log y)^{\rho})$, $y \in]0,1[$ avec $\theta = 1/2$ et $\rho \in \{-1/2, -1/4\}$ (on les notera respectivement W1 et W2),

– deux lois à queue de type Log-weibull ($\tau = 1/2$), voir Gardes *et al.* [88], Paragraphe 2.2 : la loi Lognormale centrée réduite ($\theta = \sqrt{2}/2$ et $\rho = 0$) et la loi ayant pour fonction quantile $\overline{F}^{\leftarrow}(y) = \exp\{\sqrt{2}[(-\log y)^{1/2} - 1]\}$, $y \in]0,1[$ pour laquelle $\theta = \sqrt{2}/2$ et $\rho = -\infty$ que l'on appellera loi Log-weibull.

– cinq lois de type Pareto ($\tau = 1$) : la valeur absolue d'une loi de Cauchy standard ($\theta = 1$ et $\rho = -2$), une loi de Pareto standard ($\theta = 2$ et $\rho = -\infty$), la valeur absolue d'une loi de Student à deux degrés de liberté ($\theta = 1/2$ et $\rho = -2$) et deux lois de Burr ($\theta = 1/2$ et $\rho \in \{-1, -1/2\}$) ayant pour fonction quantile $\overline{F}^{\leftarrow}(y) = (y^{\rho} - 1)^{-\theta\rho}$, $y \in]0,1[$.

Ci-après se trouve le Tableau 2.1 résumant les différentes lois utilisées, et les valeurs de τ, θ et ρ. Ces lois représentent diverses situations :

$$\tau \in \{0, 1/2, 1\}, \quad \theta \in \{1/2, \sqrt{2}/2, 1, 2\} \quad \text{et} \quad \rho \in \{-\infty, -2, -1, -1/2, -1/4, 0\}$$

dans lesquelles la normalité asymptotique de nos estimateurs n'est pas toujours établie.

Dans ce qui suit, on prend $\alpha_n = 10^{-3}$ et on simule $N = 100$ échantillons $(\mathcal{Y}_{n,j})_{j=1,\dots,N}$ de taille $n = 500$. Pour chaque échantillon $\mathcal{Y}_{n,j}$, l'estimateur $\widehat{q}_n\left(\alpha_n, \widehat{\theta}_{n,\widehat{\tau}_n}(k_n)\right)$ est calculé

$\mathcal{DA}$	Type de loi	Exemples	τ	θ	ρ
Gumbel	Lois à queue de type Weibull	Gamma	0	1	-1
		W1	0	1/2	-1/2
		W2	0	1/2	-1/4
		Weibull	0	2	$-\infty$
		Normale	0	1/2	-1
	Lois à queue de type Log-weibull	Log-normale	1/2	$\sqrt{2}/2$	0
		Log-weibull	1/2	$\sqrt{2}/2$	$-\infty$
Fréchet	Lois de type Pareto	Pareto	1	2	$-\infty$
		Student	1	1/2	-2
		Cauchy	1	1	-2
		Burr	1	1/2	-1
		Burr	1	1/2	-1/2

TABLE 2.1: Tableau résumant les différentes lois utilisées pour les simulations.

pour $k'_n = 3, \ldots, 500$ et $k_n = 2, \ldots, 499$ avec $k_n \leq k'_n$. Pour toutes les lois considérées, les erreurs quadratiques moyennes associées aux estimateurs sont calculées comme des fonctions de k'_n et de k_n. Comme notre estimateur des quantiles extrêmes dépend de ces deux suites et que l'approche POT ne dépend que d'une suite, on a décidé de réduire notre estimateur à une suite en pratique.

Pour cela on a representé l'erreur quadratique moyenne associée à $\widehat{q}_n\left(\alpha_n, \widehat{\theta}_{n,\widehat{\tau}_n}(k_n)\right)$ sous la forme d'une matrice avec k'_n en indice de ligne et k_n en indice de colonne. On a ensuite extrait la valeur $c \in [0,1]$ telle que lorsque l'on somme les éléments de la matrice se situant sur la droite $k_n = \lfloor ck'_n \rfloor$ de pente c on obtienne la plus petite valeur. A la suite d'un grand nombre de simulations Monte-Carlo sur toutes les lois considérées, on a choisi la valeur $c = 0.1$.

Sur la Figure 2.1, on a représenté les déciles (voir échelle de couleur) des matrices des erreurs quadratiques moyennes associées à $\widehat{q}_n\left(\alpha_n, \widehat{\theta}_{n,\widehat{\tau}_n}(k_n)\right)$ pour la loi de Pareto et la loi Gamma évoquées précédemment (voir Tableau 2.1).

Avec un tel choix, notre estimateur et l'approche POT dépendent uniquement d'une seule suite d'entiers (k'_n). Pour toutes les lois considérées, les erreurs quadratiques moyennes associées aux estimateurs sont calculées comme des fonctions de k'_n et sont reportées sur les Figures 2.2 à 2.7.

Il apparaît que notre estimateur est plus performant que l'approche POT (que l'on estime les paramètres de la GPD avec l'estimateur des moments ou l'estimateur DEdH) pour presque toutes les valeurs de k'_n. Remarquons également que pour $k'_n \geq 50$, l'erreur quadratique moyenne vue comme une fonction de k'_n associée à notre estimateur est presque constante, quelle que soit la loi.

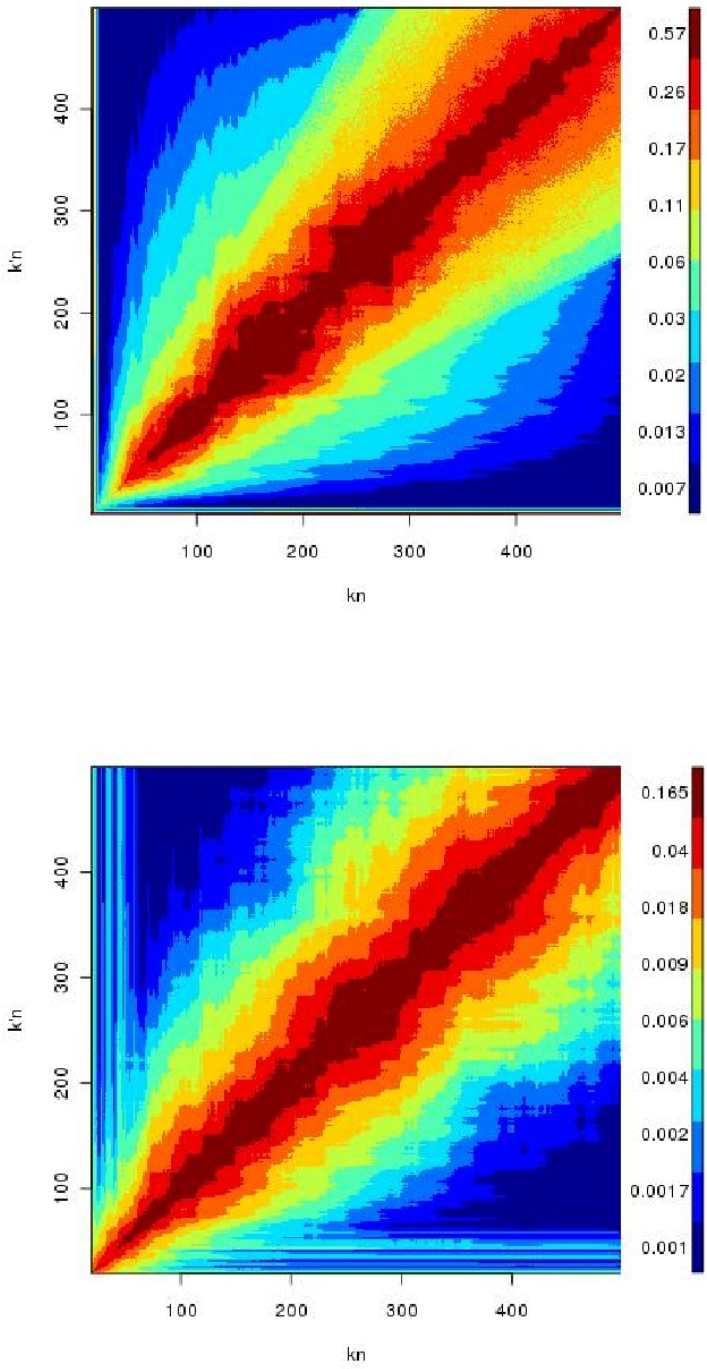


FIGURE 2.1: Déciles de l'erreur quadratique moyenne associée à $\widehat{q}_n\left(\alpha_n, \widehat{\theta}_{n,\widehat{\tau}_n}(k_n)\right)$ vue comme une fonction de k'_n et de k_n. Simulations réalisées sur 100 échantillons de taille 500. Les matrices sont obtenues par symétrie. En haut : loi de Pareto. En bas : loi Gamma.

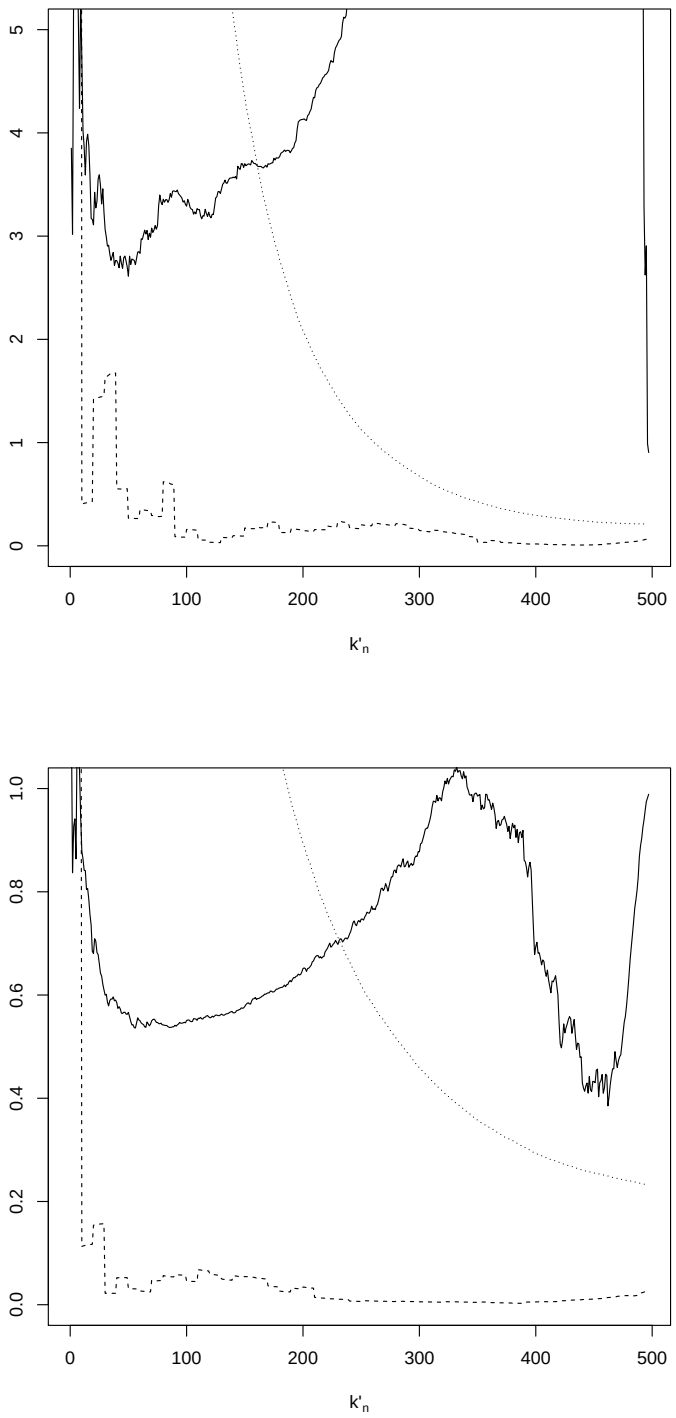


FIGURE 2.2: Erreur quadratique moyenne vue comme une fonction de k'_n associée à $\widehat{q}_n\left(\alpha_n, \widehat{\theta}_{n,\widehat{\tau}_n}(k_n)\right)$ (traits en pointillés), à l'approche POT : estimateur des Moments (traits en gras) et estimateur DEdH (traits pleins). Simulations réalisées sur 100 échantillons de taille 500. En haut : valeur absolue d'une loi de Cauchy. En bas : valeur absolue d'une loi de Student.

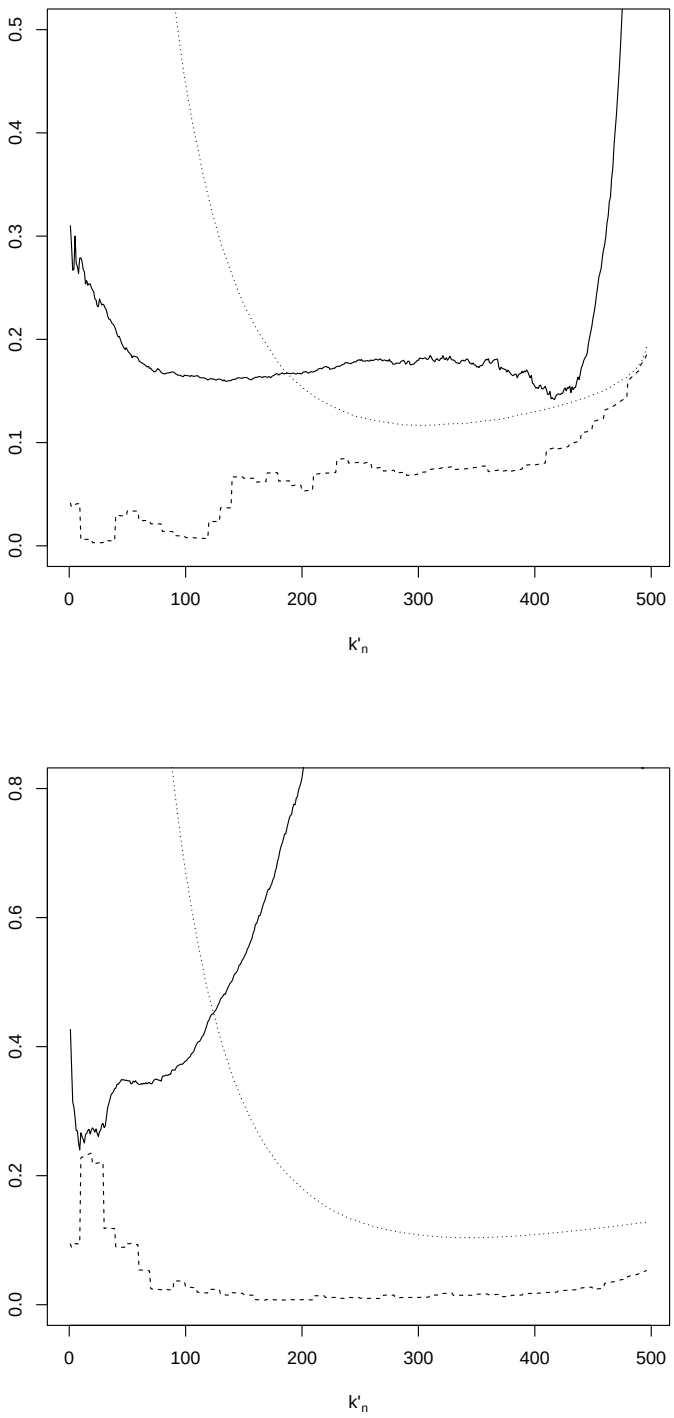


FIGURE 2.3: Erreur quadratique moyenne vue comme une fonction de k'_n associée à $\widehat{q}_n\left(\alpha_n, \widehat{\theta}_{n,\widehat{\tau}_n}(k_n)\right)$ (traits en pointillés), à l'approche POT : estimateur des Moments (traits en gras) et estimateur DEdH (traits pleins). Simulations réalisées sur 100 échantillons de taille 500. En haut : loi de Burr avec $\rho = -1$. En bas : loi de Burr avec $\rho = -1/2$.

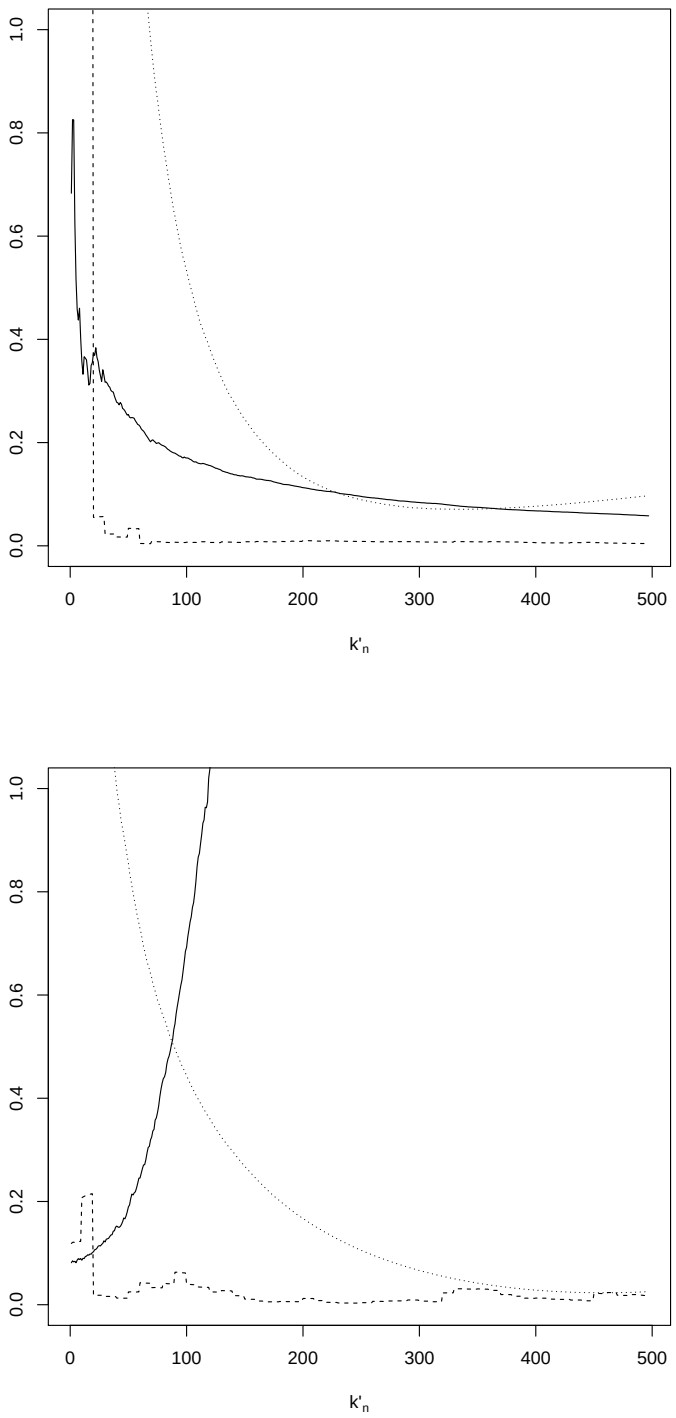


FIGURE 2.4: Erreur quadratique moyenne vue comme une fonction de k'_n associée à $\widehat{q}_n\left(\alpha_n, \widehat{\theta}_{n,\widehat{\tau}_n}(k_n)\right)$ (traits en pointillés), à l'approche POT : estimateur des Moments (traits en gras) et estimateur DEdH (traits pleins). Simulations réalisées sur 100 échantillons de taille 500. En haut : loi de Pareto. En bas : loi de Weibull.

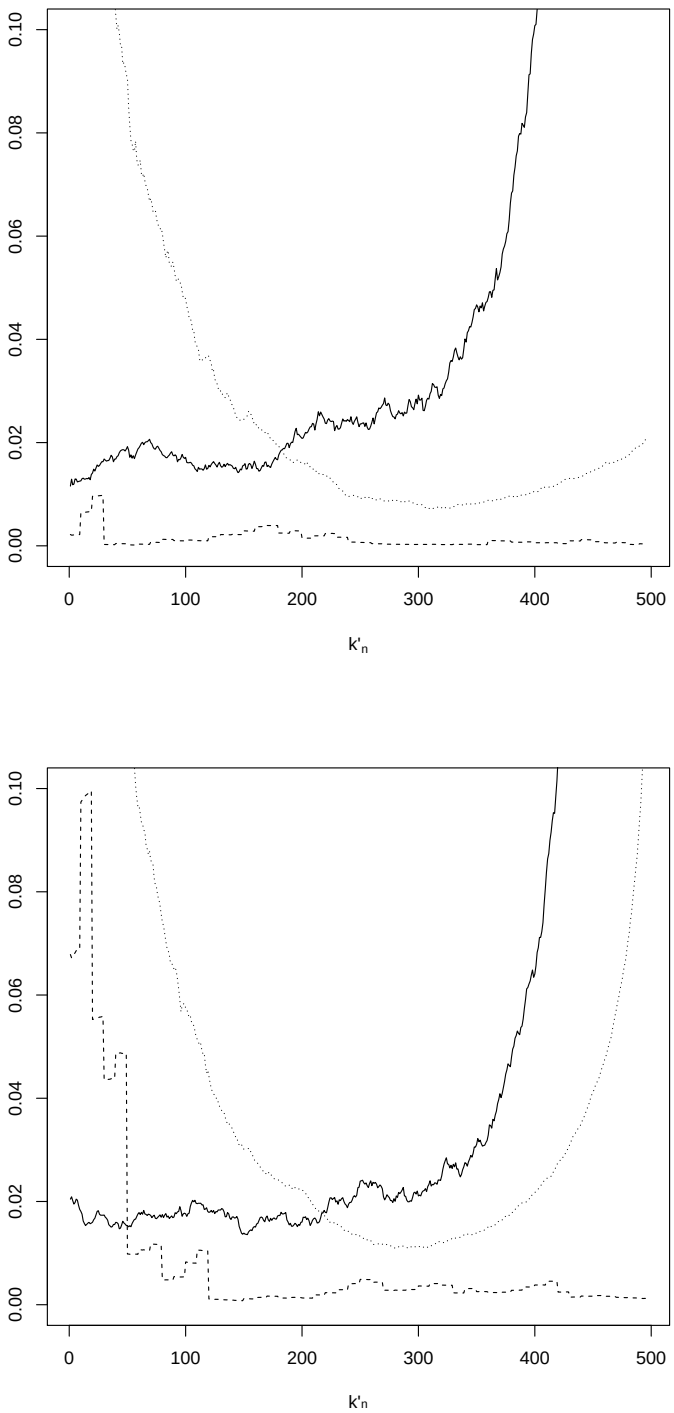


FIGURE 2.5: Erreur quadratique moyenne vue comme une fonction de k'_n associée à $\widehat{q}_n\left(\alpha_n, \widehat{\theta}_{n,\widehat{\tau}_n}(k_n)\right)$ (traits en pointillés), à l'approche POT : estimateur des Moments (traits en gras) et estimateur DEdH (traits pleins). Simulations réalisées sur 100 échantillons de taille 500. En haut : valeur absolue d'une loi Normale. En bas : loi Gamma.

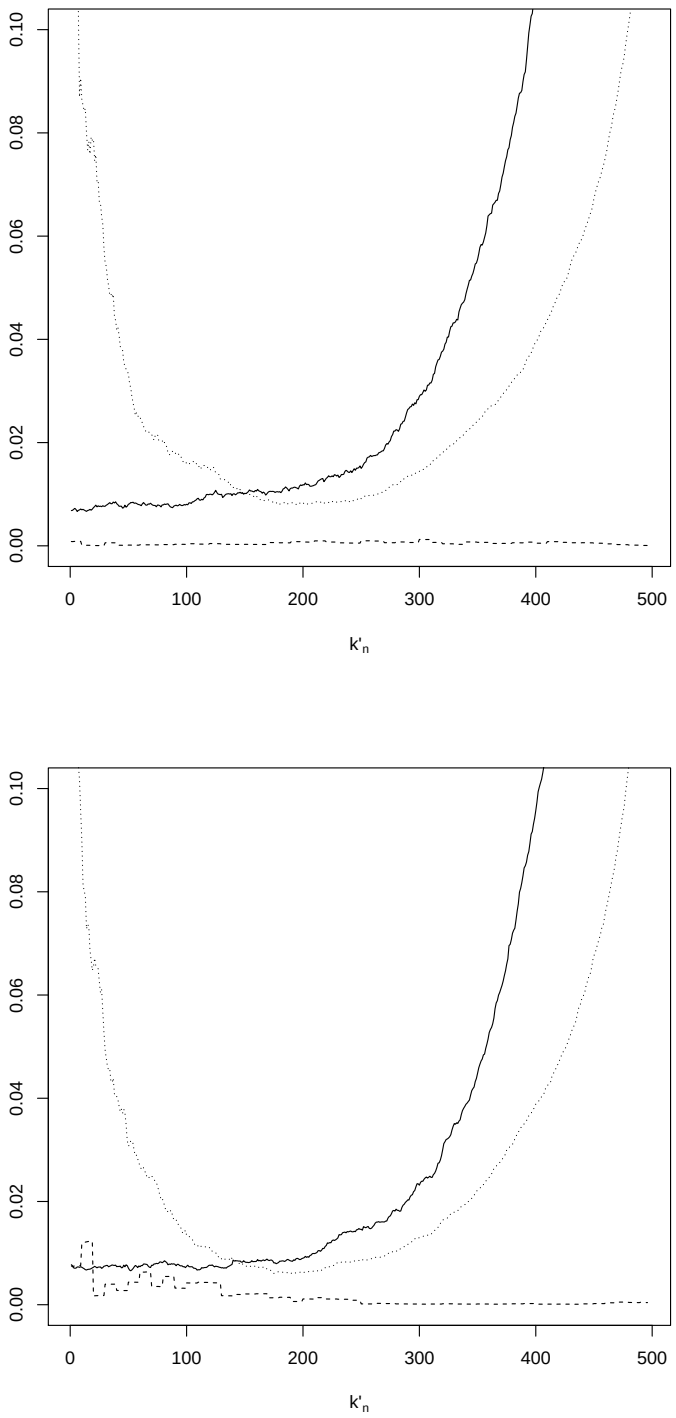


FIGURE 2.6: Erreur quadratique moyenne vue comme une fonction de k'_n associée à $\widehat{q}_n\left(\alpha_n, \widehat{\theta}_{n,\widehat{\tau}_n}(k_n)\right)$ (traits en pointillés), à l'approche POT : estimateur des Moments (traits en gras) et estimateur DEdH (traits pleins). Simulations réalisées sur 100 échantillons de taille 500. En haut : loi à queue de type Weibull W1 avec $\rho = -1/2$. En bas : loi à queue de type Weibull W2 avec $\rho = -1/4$.

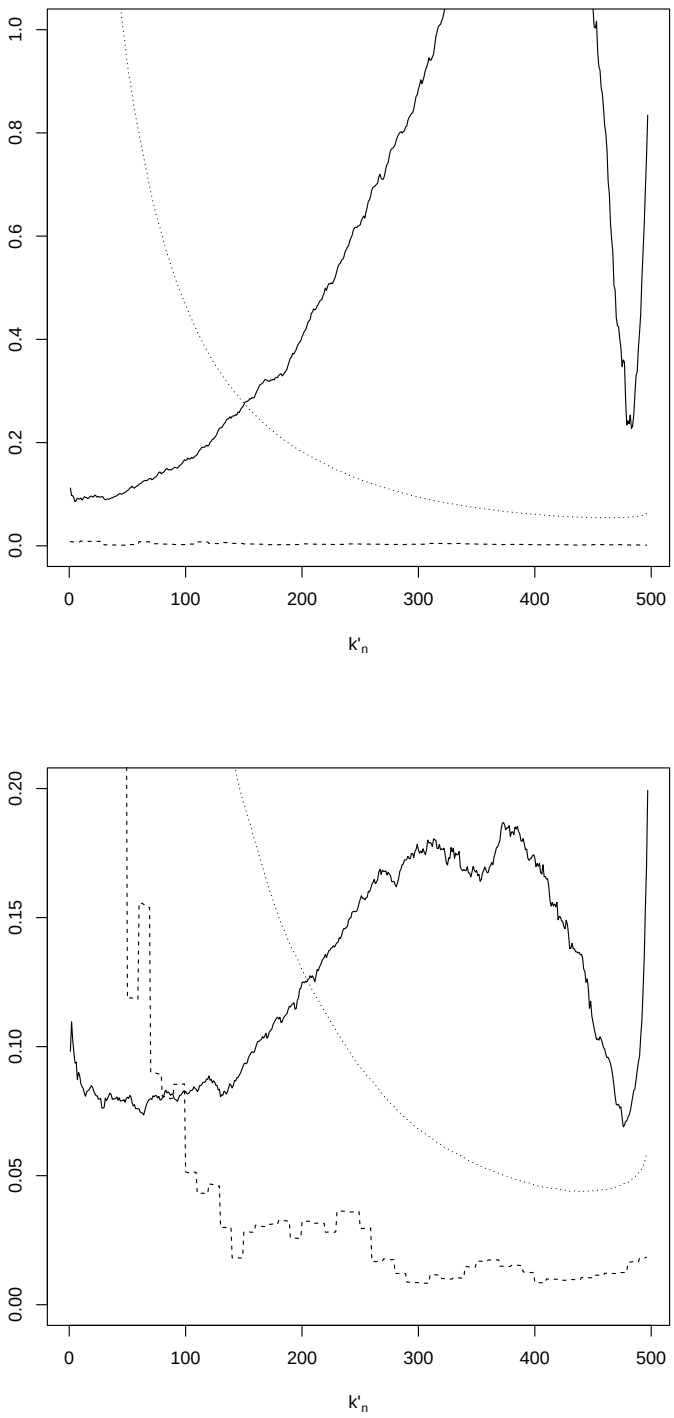


FIGURE 2.7: Erreur quadratique moyenne vue comme une fonction de k'_n associée à $\widehat{q}_n\left(\alpha_n, \widehat{\theta}_{n,\widehat{\tau}_n}(k_n)\right)$ (traits en pointillés), à l'approche POT : estimateur des Moments (traits en gras) et estimateur DEdH (traits pleins). Simulations réalisées sur 100 échantillons de taille 500. En haut : loi Log-normale. En bas : loi à queue de type Log-weibull avec $\rho = -1/4$.

2.4.2 Application aux crues d'une rivière

Les performances de nos estimateurs sont illustrées à travers l'analyse d'événements extrêmes d'un jeu de données de crues de la rivière Nidd provenant du rapport NERC [137]. Ce jeu de données est classique dans la littérature des valeurs extrêmes voir par exemple Davison et Smith [47] ou Hosking et Wallis [119]. Il consiste en 154 excès (crues) au-dessus du niveau $65m^3s^{-1}$durant la période 1934-1969 (35 ans) de la rivière Nidd se situant dans le Yorkshire en Angleterre.

Il n'y a pas de consensus sur le domaine d'attraction de ce jeu de données. Dans Diebolt *et al.* [58], il est supposé qu'il appartient au domaine d'attraction de Fréchet, les lois à queues lourdes sont considérées comme un modèle possible pour ce type de données. Cependant, selon Hosking et Wallis [119], le jeu de données de la rivière Nidd peut raisonnablement être supposé provenir d'une loi appartenant au domaine d'attraction de Gumbel. Ce résultat est en accord avec Diebolt *et al.* [60] où il a été montré qu'une loi à queue de type Weibull peut être considérée pour modéliser ces données. L'estimation de τ est donc d'un grand intérêt.

Les estimations de τ et θ sont données sur la Figure 2.8 comme fonctions de k'_n (rappelons que $k_n = \lfloor 0.1k'_n \rfloor$). On voit que les estimateurs se stabilisent pour $k'_n \geq 80$ avec $\widehat{\tau}_n \simeq 1$ et $\widehat{\theta}_{n,\widehat{\tau}_n}(k_n) \simeq 0.3$. Ces résultats nous indiquent que les données peuvent être supposées provenir d'une loi dans le domaine d'attraction de Fréchet. Ils sont en accord avec ceux obtenus par des méthodes Bayésiennes dans Diebolt *et al.* [58], Figure 7 où l'indice de queue est aussi estimé avec une valeur de 0.3.

Rappelons que la quantité standard d'intérêt dans les études environnementales est le niveau de retour d'ordre N, défini comme étant le niveau qui sera dépassé en moyenne une fois toutes les N années. Ici, on s'intéresse à l'estimation des niveaux de retour à 50 et 100 ans.

Sur la Figure 2.9, notre estimateur des quantiles extrêmes est comparé avec l'approche POT. On a représenté le niveau de retour d'ordre N comme une fonction de k'_n pour $N = 50$ et $N = 100$. On voit que notre estimateur et l'approche POT dont les paramètres de la GPD ont été estimés avec l'estimateur DEdH donnent, des courbes similaires. L'estimation des paramètres de la GPD par l'estimateur des moments donne des résultats quelque peu différents. Cependant, en choisissant $k'_n \simeq 60$ pour toutes les méthodes, on obtient une estimation du niveau de retour à 50 ans qui appartient approximativement à l'intervalle $[340m^3s^{-1}, 375m^3s^{-1}]$, et une estimation du niveau de retour à 100 ans appartenant à l'intervalle $[400m^3s^{-1}, 470m^3s^{-1}]$. Encore une fois, ces résultats sont en accord avec les intervalles de crédibilité obtenus dans Diebolt *et al.* [58], Tableau 1.

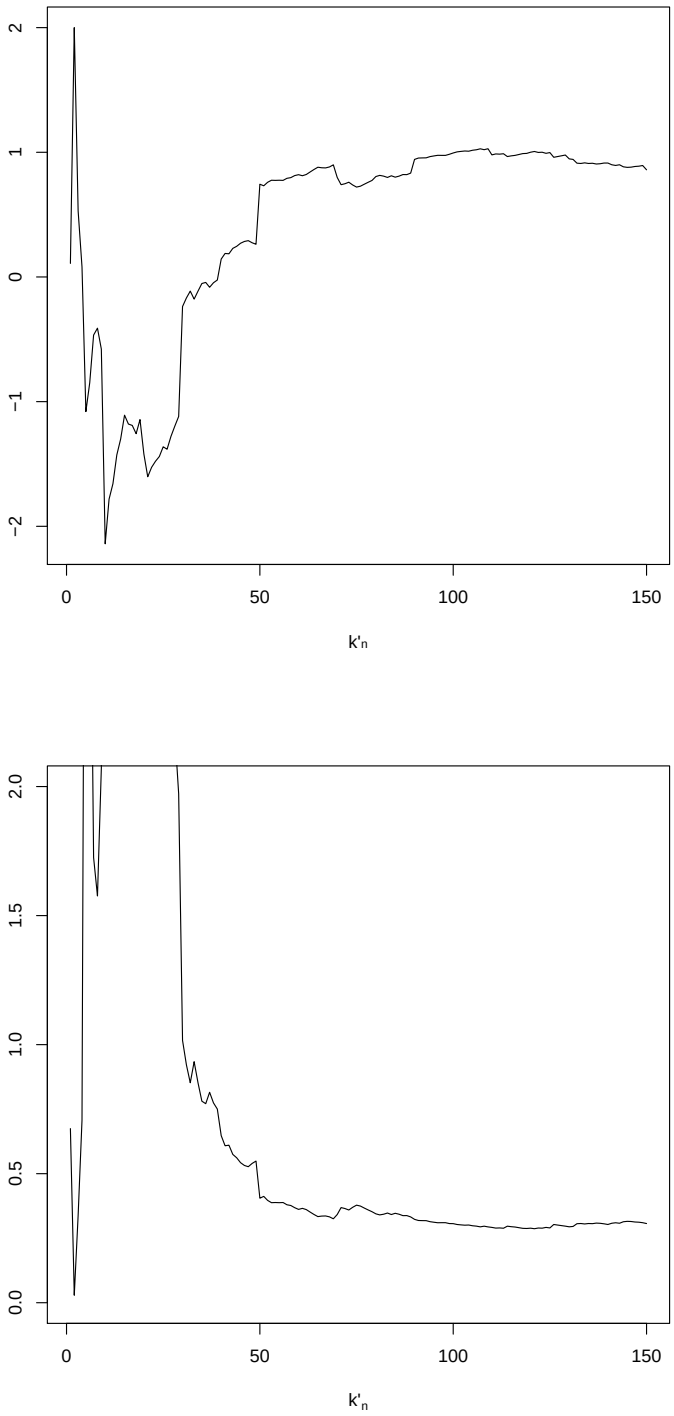


FIGURE 2.8: Résultats obtenus sur le jeu de données de la rivière Nidd En haut : estimation de τ comme fonction de k'_n. En bas : estimation de θ comme fonction de k'_n.

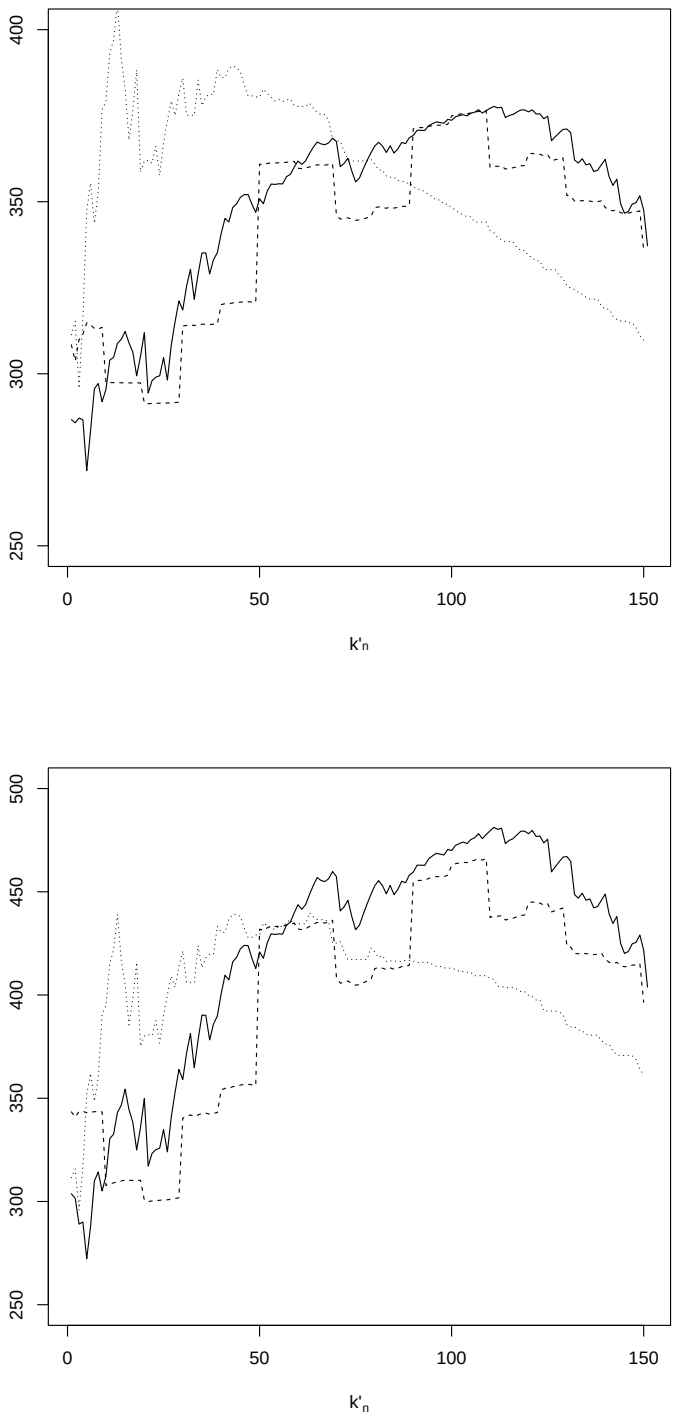


FIGURE 2.9: Résultats obtenus sur le jeu de données de la rivière Nidd. Le niveau de retour comme fonction de k'_n obtenu avec $\widehat{q}_n\left(\alpha_n, \widehat{\theta}_{n,\widehat{\tau}_n}(k_n)\right)$ (traits en pointillés), avec l'approche POT : estimateur des Moments (traits en gras) et estimateur DEdH (traits pleins). En haut : estimation du niveau de retour à 50 ans comme fonction de k'_n. En bas : estimation du niveau de retour à 100 ans comme fonction de k'_n.

2.5 Perspectives

Plusieurs perpectives à donner à nos travaux sont envisageables.

Parmi les suites explorables, on pourra envisager celle de la correction de biais basée sur l'estimation du paramètre du second ordre (Caeiro et Gomes [28], Gomes *et al.* [103]). Dans ce but, un modèle de régression exponentielle similaire à ceux de Beirlant *et al.* [12], Diebolt *et al.* [59, 60], Feuerverger et Hall [76] étendu aux lois appartenant à la famille de fonctions de survie $(\mathbf{A}_1(\tau,\theta))$ serait d'un grand intérêt.

Il serait aussi intéressant d'adapter nos résultats au cas $\tau > 1$ et de chercher des liens possibles avec des lois à queues "super-lourdes" (Fraga Alves *et al.* [78]).

Comme on a pu le voir dans la sous-Partie Simulations 2.4.1, nos estimateurs sont très performants, même pour des lois ayant des valeurs de τ où la normalité asymptotique de nos estimateurs n'est pas établie. En ce qui concerne le choix des suites (k_n) et (k'_n) il serait intéressant d'étudier le cas $\tau = 0$ et $\rho \geq -1$ de façon plus précise.

Ce travail peut également s'étendre à des variables aléatoires $Z = \psi(Y)$ où Y a une loi associée satisfaisant $(\mathbf{A}_1(\tau,\theta))$. Par exemple, en choisissant $\psi(y) = y_F - 1/y$, cela nous permettra de considérer des lois (ayant un point terminal fini y_F) dans le domaine d'attraction de Weibull (voir le passage du domaine d'attraction de Fréchet à celui de Weibull dans le Théorème 1.11). Cela pourrait ainsi permettre d'inclure l'estimateur de Hill négatif (voir par exemple Falk [74] ou de Haan et Ferreira [50], Paragraphe 3.6.2) dans notre modèle.

Finalement, nos travaux ouvrent la porte à la construction de tests d'hypothèses pour les queues de distributions. Pour un jeu de données réelles, on pourrait notamment décider s'il est issu d'une loi à queue lourde ou d'une loi à queue de type Weibull.

2.6 Démonstrations

Cette partie se subdivise en trois sous-Parties. Dans la première sous-Partie 2.6.1, on donne des résultats préliminaires. La seconde sous-Partie 2.6.2 est dédiée aux preuves de nos principaux résultats. Enfin, dans la troisième et dernière sous-Partie 2.6.3, on expose les preuves de nos résultats préliminaires.

2.6.1 Résultats préliminaires

Dans la suite, C est un sous ensemble compact tel que $[0,1] \subset C \subset]-\infty, 2[$. Le premier lemme est un résultat classique sur le comportement à l'infini des transformées de Laplace.

Lemme 2.13. *Soit $y \in C$ et $h_y \in C^\infty(\mathbb{R}^+)$. Soit $i(y) = \min\{j \in \mathbb{N}/h_y^{(j)}(0) \neq 0\}$. Si*

$$\sup_{\substack{y\in C\\ w\geq 0}} \left| h_y^{(i(y)+1)}(w) \right| < \infty,$$

alors :

$$\lim_{t\to\infty} \sup_{y\in C} \left| t^{i(y)+1} \widetilde{h}_y(t) - h_y^{(i(y))}(0) \right| = 0,$$

avec $\widetilde{h}_y(t) = \int_0^{+\infty} \exp(-tu) h_y(u) du$ la transformée de Laplace de h_y.

Le Lemme 2.13 est l'outil clé pour établir les développements asymptotiques uniformes de μ_y et $\partial\mu_y/\partial y$ donnés dans le Lemme 2.14.

Lemme 2.14. *Pour tout $y \in C$ et $t > 0$, on a :*

$$(i) \quad \lim_{t\to\infty} \sup_{y\in C} \left| \frac{\mu_y(t)}{t^{y-1}} - 1 \right| = 0,$$

$$(ii) \quad \lim_{t\to\infty} \sup_{y\in C} \left| \frac{\frac{\partial}{\partial y}\mu_y(t) - \log(t)\mu_y(t)}{t^{y-2}} - 1 \right| = 0.$$

Comme conséquence des développements du Lemme 2.14, certaines propriétés importantes de ψ peuvent être obtenues dans les deux lemmes suivants.

Lemme 2.15. *Pour tout $t > t' > 0$ et $y \in \mathbb{R}$, on a :*

(i) $y \to \psi(y; t, t')$ est une fonction croissante,

(ii) $\lim_{y\to\infty} \psi(y; t, t') = \exp(t - t')$.

Lemme 2.16. *Pour tout $y \in C$ et $t > t' > 0$, on a :*

$$\frac{\partial}{\partial y}\psi(y;t,t') \quad = \quad \log(t/t')\psi(y;t,t')\left(1 + O\left(\frac{1}{t'\log(t/t')}\right)\right), \quad \textit{lorsque} \quad t' \to \infty.$$

Le lemme suivant nous assure que l'estimateur $\widehat{\tau}_n$ est bien défini. Dans (2.12), le cas où $H_n(k_n)/H_n(k'_n) \geq k'_n/k_n$ est finalement de probabilité asymptotiquement nulle.

Lemme 2.17. *Soient (k_n) et (k'_n) deux suites intermédiaires d'entiers telles que $k_n/k'_n \to 0$. Si $\widehat{\theta}_{n,\tau}(k_n)/\widehat{\theta}_{n,\tau}(k'_n) \xrightarrow{\mathbb{P}} 1$ alors :*

$$\mathbb{P}\left(\frac{H_n(k_n)}{H_n(k'_n)} \geq \frac{k'_n}{k_n}\right) \to 0 \quad \textit{lorsque} \quad n \to \infty.$$

Gardes *et al.* [88] n'ont pas montré que $\widehat{\theta}_{n,\tau}(k_n)$ est un estimateur consistant de θ quand τ est connu et sous des hypothèses générales. Prouvons le dans le lemme suivant.

Lemme 2.18. *Supposons que $(\mathbf{A}_1(\tau,\theta))$ est vérifiée avec une fonction à variations lentes ℓ normalisée. Si (k_n) est une suite intermédiaire d'entiers, alors : $\widehat{\theta}_{n,\tau}(k_n) \xrightarrow{\mathbb{P}} \theta$.*

La normalité asymptotique de $\widehat{\theta}_{n,\tau}(k_n)$ et $\widehat{q}_n\left(\alpha_n, \widehat{\theta}_{n,\tau}(k_n)\right)$ a déjà été établie dans le cas particulier où τ est connu par Gardes *et al.* [88]. Citons deux résultats de ce travail.

Lemme 2.19. *(Théorème 1, Gardes* et al. *[88]).*
Supposons que $(\mathbf{A}_1(\tau,\theta))$ et $(\mathbf{A}_2(\rho))$ soient vérifiées. Si (k_n) est une suite intermédiaire d'entiers telle que $\sqrt{k_n}\, b(\exp K_\tau(\log(n/k_n))) \to 0$, alors :

$$\sqrt{k_n}\left(\frac{\widehat{\theta}_{n,\tau}(k_n)}{\theta} - 1\right) \xrightarrow{d} \mathcal{N}(0,1). \tag{2.23}$$

Lemme 2.20. *(Théorème 2, Gardes* et al. *[88]). Sous les hypothèses du Lemme 2.19 et si, de plus,*

$$(\log(n/k_n))^{1-\tau}(K_\tau(\log(1/\alpha_n)) - K_\tau(\log(n/k_n))) \to \infty,$$

alors :

$$\frac{\sqrt{k_n}}{K_\tau(\log(1/\alpha_n)) - K_\tau(\log(n/k_n))} \log\left(\frac{\widehat{q}_n\left(\alpha_n, \widehat{\theta}_{n,\tau}(k_n)\right)}{q(\alpha_n)}\right) \xrightarrow{d} \mathcal{N}(0,\theta^2).$$

Le lemme suivant établit que θ peut être remplacé par $\widehat{\theta}_{n,\tau}(k'_n)$ dans (2.23) sans modifier la loi asymptotique.

Lemme 2.21. *Supposons que* $(\mathbf{A}_1(\tau,\theta))$ *et* $(\mathbf{A}_2(\rho))$ *soient vérifiées. Soient* (k_n) *et* (k'_n) *deux suites intermédiaires d'entiers telles que :*

$$\sqrt{k'_n}b(\exp K_\tau(\log n/k'_n)) \to 0 \quad et \quad k_n/k'_n \to 0.$$

On a :

$$\sqrt{k_n}\left(\frac{\widehat{\theta}_{n,\tau}(k_n)}{\widehat{\theta}_{n,\tau}(k'_n)} - 1\right) \xrightarrow{d} \mathcal{N}(0,1).$$

Le dernier lemme quantifie l'effet d'estimer τ dans $\widehat{\theta}_{n,\tau}(k_n)$.

Lemme 2.22. *Supposons que* $(\mathbf{A}_1(\tau,\theta))$ *et* $(\mathbf{A}_2(\rho))$ *soient vérifiées. Soient* (k_n) *et* (k'_n) *deux suites intermédiaires d'entiers telles que :*

$$k_n/k'_n \to 0, \quad \sqrt{k_n}\left(\log_2(n/k_n) - \log_2(n/k'_n)\right)/\log_2(n/k_n) \to \infty,$$

$$\sqrt{k'_n}b(\exp K_\tau(\log n/k'_n)) \to 0 \quad et \quad \log(n/k'_n)\left(\log_2(n/k_n) - \log_2(n/k'_n)\right) \to \infty.$$

On a :

$$\frac{\sqrt{k_n}(\log_2(n/k_n) - \log_2(n/k'_n))}{\log_2(n/k_n)}\left(\frac{\widehat{\theta}_{n,\widehat{\tau}_n}(k_n)}{\widehat{\theta}_{n,\tau}(k_n)} - 1\right) \xrightarrow{d} \mathcal{N}(0,1).$$

2.6.2 Preuves des principaux résultats

Preuve de la Proposition 2.9 – Pour tout $\varepsilon > 0$, écrivons :

$$\mathbb{P}\left(|\widehat{\tau}_n - \tau| > \varepsilon\right) = \mathbb{P}\left(\widehat{\tau}_n > \tau + \varepsilon\right) + \mathbb{P}\left(\widehat{\tau}_n < \tau - \varepsilon\right)$$

et considérons les deux termes séparément. Clairement, on a :

$$\begin{aligned}
\mathbb{P}(\widehat{\tau}_n > \tau + \varepsilon) &= \mathbb{P}\left(\{\widehat{\tau}_n > \tau + \varepsilon\} \middle| \left\{\frac{H_n(k_n)}{H_n(k'_n)} \geq \frac{k'_n}{k_n}\right\}\right)\mathbb{P}\left(\frac{H_n(k_n)}{H_n(k'_n)} \geq \frac{k'_n}{k_n}\right) \\
&+ \mathbb{P}\left(\{\widehat{\tau}_n > \tau + \varepsilon\} \cap \left\{\frac{H_n(k_n)}{H_n(k'_n)} < \frac{k'_n}{k_n}\right\}\right) \\
&\leq \mathbb{P}\left(\frac{H_n(k_n)}{H_n(k'_n)} \geq \frac{k'_n}{k_n}\right) \qquad (2.24) \\
&+ \mathbb{P}\left(\frac{H_n(k_n)}{H_n(k'_n)} \geq \psi(\tau + \varepsilon; \log(n/k_n), \log(n/k'_n))\right). \qquad (2.25)
\end{aligned}$$

Concentrons nous sur la probabilité (2.24), notons que, $\widehat{\theta}_{n,\tau}(k_n)$ et $\widehat{\theta}_{n,\tau}(k'_n)$ sont des estimateurs consistants de θ en probabilité par le Lemme 2.18. Alors $\widehat{\theta}_{n,\tau}(k_n)/\widehat{\theta}_{n,\tau}(k'_n) \xrightarrow{\mathbb{P}} 1$ et le Lemme 2.17 impliquent :

$$\mathbb{P}\left(\frac{H_n(k_n)}{H_n(k'_n)} \geq \frac{k'_n}{k_n}\right) \to 0.$$

La seconde probabilité (2.25) tend vers 0 par (2.11) et le Lemme 2.15(i). En procédant de la même façon, on peut prouver que $\mathbb{P}\left(\widehat{\tau}_n < \tau - \varepsilon\right) \to 0$ ce qui achève la preuve. ■

Preuve du Théorème 2.10 – Soit $v_n = \sqrt{k_n}\left(\log_2(n/k_n) - \log_2(n/k'_n)\right)$, notre but est de prouver que

$$\mathbb{P}\left(v_n(\widehat{\tau}_n - \tau) \leq s\right) \to \Phi(s),$$

pour tout $s \in \mathbb{R}$ où Φ est la fonction de répartition d'une loi Normale centrée réduite. Pour cela, premièrement remarquons que le Lemme 2.21 implique

$$\widehat{\theta}_{n,\tau}(k_n)/\widehat{\theta}_{n,\tau}(k'_n) \xrightarrow{\mathbb{P}} 1.$$

Ainsi, en introduisant $E_n(s) = \{v_n(\widehat{\tau}_n - \tau) \leq s\}$, on a :

$$\begin{aligned}
\mathbb{P}(E_n(s)) &= \mathbb{P}\left(E_n(s) \cap \left\{\frac{H_n(k_n)}{H_n(k'_n)} < \frac{k'_n}{k_n}\right\}\right) \\
&+ \mathbb{P}\left(E_n(s) \middle| \left\{\frac{H_n(k_n)}{H_n(k'_n)} \geq \frac{k'_n}{k_n}\right\}\right) \mathbb{P}\left(\frac{H_n(k_n)}{H_n(k'_n)} \geq \frac{k'_n}{k_n}\right) \\
&=: T_n^{(1)}(s) + o(1),
\end{aligned}$$

par le Lemme 2.17. D'après la définition de $\widehat{\tau}_n$ donnée dans (2.12) et en rappelant que, d'après le Lemme 2.15(i), $\psi(.; \log(n/k_n), \log(n/k'_n))$ est une fonction croissante, on obtient :

$$\begin{aligned}
T_n^{(1)}(s) &= \mathbb{P}\left(\left\{\frac{H_n(k_n)}{H_n(k'_n)} \leq \psi(\tau + s/v_n; \log(n/k_n), \log(n/k'_n))\right\} \cap \left\{\frac{H_n(k_n)}{H_n(k'_n)} < \frac{k'_n}{k_n}\right\}\right) \\
&= \mathbb{P}\left(\frac{H_n(k_n)}{H_n(k'_n)} \leq \min\left(\psi(\tau + s/v_n; \log(n/k_n), \log(n/k'_n)), \frac{k'_n}{k_n}\right)\right).
\end{aligned}$$

Le Lemme 2.15(ii) implique alors :

$$\begin{aligned}
T_n^{(1)}(s) &= \mathbb{P}\left(\frac{H_n(k_n)}{H_n(k'_n)} \leq \psi(\tau + s/v_n; \log(n/k_n), \log(n/k'_n))\right) \\
&= \mathbb{P}\left(\frac{H_n(k_n)}{H_n(k'_n)} \leq \frac{\mu_{\tau+s/v_n}(\log(n/k_n))}{\mu_{\tau+s/v_n}(\log(n/k'_n))}\right),
\end{aligned}$$

et, d'après (2.1) et le fait que :

$$\zeta_n := \sqrt{k_n}\left(\frac{\widehat{\theta}_{n,\tau}(k_n)}{\widehat{\theta}_{n,\tau}(k'_n)} - 1\right) \xrightarrow{d} \mathcal{N}(0,1),$$

(voir Lemme 2.21), on a :

$$\begin{aligned}
T_n^{(1)}(s) &= \mathbb{P}\left(\zeta_n \leq \sqrt{k_n}\left(\frac{\mu_\tau(\log(n/k_n'))}{\mu_\tau(\log(n/k_n))}\frac{\mu_{\tau+s/v_n}(\log(n/k_n))}{\mu_{\tau+s/v_n}(\log(n/k_n'))} - 1\right)\right) \\
&= \mathbb{P}\left(\zeta_n \leq \sqrt{k_n}\frac{\mu_\tau(\log(n/k_n'))}{\mu_\tau(\log(n/k_n))}\left[\psi(\tau+s/v_n;\log(n/k_n),\log(n/k_n'))\right.\right. \\
&- \left.\left.\psi(\tau;\log(n/k_n),\log(n/k_n'))\right]\right).
\end{aligned}$$

Un développement de Taylor au premier ordre nous donne :

$$\begin{aligned}
T_n^{(1)}(s) &= \mathbb{P}\left(\zeta_n \leq \frac{s}{(\log_2(n/k_n) - \log_2(n/k_n'))}\frac{\mu_\tau(\log(n/k_n'))}{\mu_\tau(\log(n/k_n))}\right. \\
&\times \left.\frac{\partial}{\partial y}\psi(\tau_0;\log(n/k_n),\log(n/k_n'))\right),
\end{aligned}$$

où $\tau_0 = \tau + s\eta_0/v_n$ avec $\eta_0 \in]0,1[$. Comme $\tau_0 \to \tau$, pour n assez grand, $\tau_0 < 2$ et le Lemme 2.16 implique :

$$\begin{aligned}
T_n^{(1)}(s) &= \mathbb{P}\left(\zeta_n \leq s\frac{\mu_{\tau_0}(\log(n/k_n))}{\mu_{\tau_0}(\log(n/k_n'))}\frac{\mu_\tau(\log(n/k_n'))}{\mu_\tau(\log(n/k_n))}\left(1\right.\right. \\
&+ \left.\left. O\left(\frac{1}{\log(n/k_n')\,(\log_2(n/k_n) - \log_2(n/k_n'))}\right)\right)\right) \\
&= \mathbb{P}\left(\zeta_n \leq s\frac{\mu_{\tau_0}(\log(n/k_n))}{\mu_{\tau_0}(\log(n/k_n'))}\frac{\mu_\tau(\log(n/k_n'))}{\mu_\tau(\log(n/k_n))}(1+o(1))\right).
\end{aligned}$$

Par ailleurs, le Lemme 2.14(i) implique :

$$\begin{aligned}
s\frac{\mu_{\tau_0}(\log(n/k_n))}{\mu_{\tau_0}(\log(n/k_n'))}\frac{\mu_\tau(\log(n/k_n'))}{\mu_\tau(\log(n/k_n))} &= s\left(\frac{\log(n/k_n)}{\log(n/k_n')}\right)^{\tau_0-\tau}(1+o(1)), \\
&= s\exp\left(\frac{s\eta_0}{\sqrt{k_n}}\right)(1+o(1)) \underset{n\to\infty}{\longrightarrow} s,
\end{aligned}$$

et ainsi on a :

$$\begin{aligned}
|T_n^{(1)}(s) - \Phi(s)| &= |\mathbb{P}(\zeta_n \leq s(1+o(1))) - \Phi(s)| \\
&\leq |\mathbb{P}(\zeta_n \leq s(1+o(1))) - \Phi(s(1+o(1)))| + |\Phi(s(1+o(1))) - \Phi(s)| \\
&\leq \sup_{y\in\mathbb{R}}|\Phi(y) - \mathbb{P}(\zeta_n \leq y)| + |\Phi(s(1+o(1))) - \Phi(s)|
\end{aligned}$$

par Embrechts *et al.* [71], page 552. Ceci achève la preuve du Théorème 2.10. ■

Preuve du Théorème 2.11 – Dans un premier temps, le Lemme 2.19 montre que :

$$\widehat{\theta}_{n,\tau}(k_n) - \theta = \frac{\zeta_n\theta}{\sqrt{k_n}} \quad \text{où} \quad \zeta_n \xrightarrow{d} \mathcal{N}(0,1),$$

dans un second temps, le Lemme 2.22 implique :

$$\frac{\widehat{\theta}_{n,\widehat{\tau}_n}(k_n)}{\widehat{\theta}_{n,\tau}(k_n)} - 1 = \frac{\delta_n \log_2(n/k_n)}{\sqrt{k_n}(\log_2(n/k_n) - \log_2(n/k'_n))} \quad \text{où} \quad \delta_n \xrightarrow{d} \mathcal{N}(0,1). \tag{2.26}$$

En rassemblant ces résultats, il s'ensuit que :

$$\begin{aligned}
\widehat{\theta}_{n,\widehat{\tau}_n}(k_n) - \theta &= \widehat{\theta}_{n,\tau}(k_n)\left(\frac{\widehat{\theta}_{n,\widehat{\tau}_n}(k_n)}{\widehat{\theta}_{n,\tau}(k_n)} - 1\right) + \left(\widehat{\theta}_{n,\tau}(k_n) - \theta\right) \\
&= \widehat{\theta}_{n,\tau}(k_n)\frac{\delta_n \log_2(n/k_n)}{\sqrt{k_n}(\log_2(n/k_n) - \log_2(n/k'_n))} + \frac{\zeta_n \theta}{\sqrt{k_n}}.
\end{aligned}$$

Par conséquent, on a immédiatement :

$$\begin{aligned}
\frac{\sqrt{k_n}(\log_2(n/k_n) - \log_2(n/k'_n))}{\log_2(n/k_n)}\left(\widehat{\theta}_{n,\widehat{\tau}_n}(k_n) - \theta\right) &= \widehat{\theta}_{n,\tau}(k_n)\delta_n \\
&+ \zeta_n \theta \frac{\log_2(n/k_n) - \log_2(n/k'_n)}{\log_2(n/k_n)}
\end{aligned}$$

et le Théorème 2.11 en découle. ■

Preuve du Théorème 2.12 – Considérons le développement suivant :

$$\begin{aligned}
\log\left(\frac{\widehat{q}_n\left(\alpha_n, \widehat{\theta}_{n,\widehat{\tau}_n}(k_n)\right)}{q(\alpha_n)}\right) &= \log\left(\frac{\widehat{q}_n\left(\alpha_n, \widehat{\theta}_{n,\widehat{\tau}_n}(k_n)\right)}{\widehat{q}_n\left(\alpha_n, \widehat{\theta}_{n,\tau}(k_n)\right)}\right) + \log\left(\frac{\widehat{q}_n\left(\alpha_n, \widehat{\theta}_{n,\tau}(k_n)\right)}{q(\alpha_n)}\right) \\
&=: T_n^{(2)} + T_n^{(3)}.
\end{aligned}$$

Par (2.4), on a :

$$\begin{aligned}
T_n^{(2)} &= \widehat{\theta}_{n,\widehat{\tau}_n}(k_n)[K_{\widehat{\tau}_n}(\log(1/\alpha_n)) - K_{\widehat{\tau}_n}(\log(n/k_n))] \\
&- \widehat{\theta}_{n,\tau}(k_n)[K_\tau(\log(1/\alpha_n)) - K_\tau(\log(n/k_n))] \\
&= [K_\tau(\log(1/\alpha_n)) - K_\tau(\log(n/k_n))]\left(\widehat{\theta}_{n,\widehat{\tau}_n}(k_n) - \widehat{\theta}_{n,\tau}(k_n)\right) \\
&\quad + \widehat{\theta}_{n,\widehat{\tau}_n}(k_n)\left[\left(K_{\widehat{\tau}_n}(\log(1/\alpha_n)) - K_{\widehat{\tau}_n}(\log(n/k_n))\right)\right. \\
&- \left.(K_\tau(\log(1/\alpha_n)) - K_\tau(\log(n/k_n)))\right] \\
&=: T_n^{(2,1)} + T_n^{(2,2)}.
\end{aligned}$$

Concentrons nous sur le premier terme $T_n^{(2,1)}$. L'équation (2.26) nous donne :

$$T_n^{(2,1)} = \frac{\log_2(n/k_n)[K_\tau(\log(1/\alpha_n)) - K_\tau(\log(n/k_n))]}{\sqrt{k_n}\,(\log_2(n/k_n) - \log_2(n/k'_n))}\widehat{\theta}_{n,\tau}(k_n)\delta_n$$

ce qui implique :

$$\begin{aligned}\frac{\sqrt{k_n}(\log_2(n/k_n)-\log_2(n/k'_n))}{\int_{\log(n/k_n)}^{\log(1/\alpha_n)} u^{\tau-1}\log(u)du}T_n^{(2,1)} &= \frac{\delta_n \log_2(n/k_n)\widehat{\theta}_{n,\tau}(k_n)}{\int_{\log(n/k_n)}^{\log(1/\alpha_n)} u^{\tau-1}\log(u)du} \\ &\times [K_\tau(\log(1/\alpha_n))-K_\tau(\log(n/k_n))] \\ &= o_{\mathbb{P}}(1).\end{aligned}$$

Considérons alors $T_n^{(2,2)}$. Le Théorème 2.10 affirme que :

$$\widehat{\tau}_n = \tau + \frac{\xi_n}{\sqrt{k_n}\,(\log_2(n/k_n)-\log_2(n/k'_n))} =: \tau + \xi_n\sigma_n \quad \text{où} \quad \xi_n \xrightarrow{d} \mathcal{N}(0,1).$$

En remplaçant dans $T_n^{(2,2)}$, on obtient :

$$\begin{aligned}T_n^{(2,2)} &= \widehat{\theta}_{n,\widehat{\tau}_n}(k_n)[(K_{\tau+\sigma_n\xi_n}(\log(1/\alpha_n))-K_{\tau+\sigma_n\xi_n}(\log(n/k_n))) \\ &- (K_\tau(\log(1/\alpha_n))-K_\tau(\log(n/k_n)))].\end{aligned}$$

Par definition :

$$K_\tau(\log(1/\alpha_n))-K_\tau(\log(n/k_n)) = \int_{\log(n/k_n)}^{\log(1/\alpha_n)} u^{\tau-1}du,$$

et donc il en découle immédiatement que :

$$\begin{aligned}T_n^{(2,2)} &= \widehat{\theta}_{n,\widehat{\tau}_n}(k_n)\int_{\log(n/k_n)}^{\log(1/\alpha_n)}\left(u^{\tau+\sigma_n\xi_n-1}-u^{\tau-1}\right)du \\ &= \widehat{\theta}_{n,\widehat{\tau}_n}(k_n)\int_{\log(n/k_n)}^{\log(1/\alpha_n)} u^{\tau-1}\left(u^{\sigma_n\xi_n}-1\right)du.\end{aligned}$$

Soit $\varphi(y) := \exp(y)-1-y$, on en déduit que :

$$\begin{aligned}T_n^{(2,2)} &= \widehat{\theta}_{n,\widehat{\tau}_n}(k_n)\sigma_n\xi_n\int_{\log(n/k_n)}^{\log(1/\alpha_n)} u^{\tau-1}\log(u)du \\ &+ \widehat{\theta}_{n,\widehat{\tau}_n}(k_n)\int_{\log(n/k_n)}^{\log(1/\alpha_n)} u^{\tau-1}\varphi(\sigma_n\xi_n\log(u))du.\end{aligned}$$

Il existe alors $c>0$ tel que $y<\log(c)$ cela implique $|\varphi(y)|<\frac{c}{2}y^2$. Par conséquent, comme :

$$\sigma_n\log_2(1/\alpha_n)\to 0 \quad \text{et} \quad \sigma_n\log_2(n/k_n)\to 0,$$

pour n assez grand, on a :

$$\begin{aligned}\left|\int_{\log(n/k_n)}^{\log(1/\alpha_n)} u^{\tau-1}\varphi(\sigma_n\xi_n\log(u))du\right| &\le \int_{\log(n/k_n)}^{\log(1/\alpha_n)} u^{\tau-1}|\varphi(\sigma_n\xi_n\log(u))|du \\ &\le \frac{c}{2}\sigma_n^2\xi_n^2\int_{\log(n/k_n)}^{\log(1/\alpha_n)} u^{\tau-1}[\log(u)]^2du.\end{aligned}$$

Ainsi,

$$\begin{aligned}\frac{\left|\int_{\log(n/k_n)}^{\log(1/\alpha_n)} u^{\tau-1}\varphi(\sigma_n\xi_n\log(u))du\right|}{\sigma_n\int_{\log(n/k_n)}^{\log(1/\alpha_n)} u^{\tau-1}\log(u)du} &\leq \frac{c\sigma_n\xi_n^2\int_{\log(n/k_n)}^{\log(1/\alpha_n)} u^{\tau-1}[\log(u)]^2du}{2\int_{\log(n/k_n)}^{\log(1/\alpha_n)} u^{\tau-1}\log(u)du}\\ &\leq \frac{c\xi_n^2\log_2(1/\alpha_n)}{2\sqrt{k_n}(\log_2(n/k_n)-\log_2(n/k_n'))} = o_{\mathbb{P}}(1),\end{aligned}$$

et, en remplaçant dans $T_n^{(2,2)}$, on obtient :

$$\frac{\sqrt{k_n}\,(\log_2(n/k_n)-\log_2(n/k_n'))}{\int_{\log(n/k_n)}^{\log(1/\alpha_n)}\log(u)u^{\tau-1}du}T_n^{(2,2)} = \xi_n\widehat{\theta}_{n,\widehat{\tau}_n}(k_n)+o_{\mathbb{P}}(1)\xrightarrow{d}\mathcal{N}(0,\theta^2).$$

Finalement, le Lemme 2.20 affirme que :

$$T_n^{(3)} = \frac{\varrho_n[K_\tau(\log(1/\alpha_n))-K_\tau(\log(n/k_n))]}{\sqrt{k_n}} \quad \text{where} \quad \varrho_n\xrightarrow{d}\mathcal{N}(0,\theta^2),$$

et ainsi, sous nos hypothèses,

$$\frac{\sqrt{k_n}(\log_2(n/k_n)-\log_2(n/k_n'))}{\int_{\log(n/k_n)}^{\log(1/\alpha_n)} u^{\tau-1}\log(u)du}T_n^{(3)} = o_{\mathbb{P}}(1).$$

En combinant les résultats précédents et en utilisant la delta méthode, le Théorème 2.12 est prouvé. ■

2.6.3 Preuves des résultats préliminaires

Preuve du Lemme 2.13 – Pour plus de simplicité dans la preuve on remplacera $i(y)$ par i. Un développement de Taylor à l'ordre $i+1$ nous donne :

$$h_y(u) = \frac{u^i}{i!}h_y^{(i)}(0)+\frac{u^{i+1}}{(i+1)!}h_y^{(i+1)}(\eta u),$$

avec $\eta\in\,]0,1[$ et par conséquent,

$$\begin{aligned}\widetilde{h}_y(t) &= \int_0^\infty \exp(-tu)\frac{u^i}{i!}h_y^{(i)}(0)du+\int_0^\infty\exp(-tu)\frac{u^{i+1}}{(i+1)!}h_y^{(i+1)}(\eta u)du,\\ &=: T_y^{(1)}(t)+T_y^{(2)}(t).\end{aligned}$$

Il s'ensuit que :

$$T_y^{(1)}(t) = \frac{h_y^{(i)}(0)}{i!}\int_0^\infty u^i\exp(-tu)du = \frac{h_y^{(i)}(0)}{t^{i+1}},$$

et le changement de variable $v = tu$ implique :

$$T_y^{(2)}(t) = \frac{1}{(i+1)!}\frac{1}{t^{i+2}}\int_0^\infty \exp(-v)v^{i+1}h_y^{(i+1)}\left(\eta\frac{v}{t}\right)dv.$$

Donc, on a, uniformément pour $y \in C$:

$$\begin{aligned}|T_y^{(2)}(t)| &\leq \sup_{\substack{z\in C\\ w\geq 0}}\left|h_z^{(i+1)}(w)\right|\frac{1}{(i+1)!}\frac{1}{t^{i+2}}\int_0^\infty \exp(-v)v^{i+1}dv\\ &= \frac{1}{t^{i+2}}\sup_{\substack{z\in C\\ w\geq 0}}\left|h_z^{(i+1)}(w)\right| = O\left(\frac{1}{t^{i+2}}\right),\end{aligned}$$

et ainsi :

$$\widetilde{h}_y(t) = \frac{h_y^{(i)}(0)}{t^{i+1}} + O\left(\frac{1}{t^{i+2}}\right) = \frac{1}{t^{i+1}}\left(h_y^{(i)}(0) + O\left(\frac{1}{t}\right)\right),$$

ce qui conclut la preuve. ■

Preuve du Lemme 2.14 – (i) Par definition, pour tout $y \in \mathbb{R}$ et $t > 0$, on a :

$$\mu_y(t) = \int_0^\infty (K_y(u+t) - K_y(t))\exp(-u)du = \int_0^\infty \left(\int_t^{u+t} w^{y-1}dw\right)\exp(-u)du.$$

En utilisant le théorème de Fubini, il s'ensuit :

$$\mu_y(t) = \int_t^\infty w^{y-1}\left(\int_{w-t}^\infty \exp(-u)du\right)dw = \exp(t)\int_t^\infty w^{y-1}\exp(-w)dw, \tag{2.27}$$

et le changement de variable $u = w/t - 1$ implique :

$$\mu_y(t) = t^y\int_0^\infty \exp(-tu)(u+1)^{y-1}du = t^y\widetilde{h}_y(t)$$

avec $h_y(t) = (t+1)^{y-1}$. En appliquant le Lemme 2.13 avec $i = 0$ on conclut la preuve du point (i).

(ii) D'après (2.27) on obtient, pour $y \in \mathbb{R}$ et $t > 0$,

$$\mu_y(t) = \exp(t)\int_t^\infty w^{y-1}\exp(-w)dw = \exp(t)\Gamma(y,t), \tag{2.28}$$

où $\Gamma(y,t)$ est la fonction Gamma incomplète supérieure. On a ainsi (voir par exemple Artin [5]) :

$$\frac{\partial}{\partial y}\mu_y(t) = \exp(t)\int_t^\infty \exp(-w)\log(w)w^{y-1}dw, \tag{2.29}$$

et le changement de variable $u = w/t - 1$ implique :

$$\begin{aligned}
\frac{\partial}{\partial y}\mu_y(t) &= \exp(t)\int_0^\infty \exp(-tu)\exp(-t)t^{y-1}(u+1)^{y-1}\log(t(u+1))tdu \\
&= t^y\int_0^\infty \exp(-tu)(u+1)^{y-1}\log(u+1)du \\
&+ \log(t)t^y\int_0^\infty \exp(-tu)(u+1)^{y-1}du \\
&=: t^y\int_0^\infty \exp(-tu)g_y(u)du + \log(t)\mu_y(t),
\end{aligned}$$

avec $g_y(u) := (u+1)^{y-1}\log(u+1)$. En appliquant le Lemme 2.13 avec $i = 1$, on conclut la preuve. ■

Preuve du Lemme 2.15 – (i) Premièrement, remarquons que $\psi(y;t,t') > 0$ pour tout $y \in \mathbb{R}$ et $t > t' > 0$. Par ailleurs, un calcul classique nous montre que :

$$\begin{aligned}
\frac{\partial}{\partial y}\psi(y;t,t') &= \psi(y;t,t')\left(\frac{\partial/\partial y(\mu_y(t))}{\mu_y(t)} - \frac{\partial/\partial y(\mu_y(t'))}{\mu_y(t')}\right) \\
&=: \psi(y;t,t')\left(Q_y(t) - Q_y(t')\right),
\end{aligned} \tag{2.30}$$

où, par (2.27) et (2.29),

$$Q_y(z) = \frac{\int_z^\infty \log(w)w^{y-1}\exp(-w)dw}{\int_z^\infty w^{y-1}\exp(-w)dw}.$$

Remarquons que Q_y est une fonction croissante sur $(0,\infty)$ car :

$$Q_y'(z) = \frac{z^{y-1}\exp(-z)\int_z^\infty w^{y-1}\exp(-w)\log(w/z)dw}{\left(\int_z^\infty w^{y-1}\exp(-w)dw\right)^2} > 0.$$

Par conséquent $t > t'$ implique :

$$\partial/\partial y(\psi(y;t,t')) = \psi(y;t,t')(Q_y(t) - Q_y(t')) > 0$$

et conclut la première partie de la preuve.
(ii) D'après (2.28), on a

$$\frac{1}{\psi(y;t,t')}\frac{\exp(-t')}{\exp(-t)} - 1 = \frac{\mu_y(t')\exp(-t')}{\mu_y(t)\exp(-t)} - 1 = \frac{\int_{t'}^t w^{y-1}\exp(-w)dw}{\int_t^\infty w^{y-1}\exp(-w)dw} =: M_y(t,t').$$

Par aileurs, en considérant les inégalités suivantes :

$$0 < \int_{t'}^t w^{y-1}\exp(-w)dw < t^{y-1}\left(\exp(-t') - \exp(-t)\right),$$

et

$$\int_t^\infty w^{y-1}\exp(-w)dw > \int_{2t}^\infty w^{y-1}\exp(-w)dw > (2t)^{y-1}\exp(-2t),$$

il s'ensuit que :

$$M_y(t,t') < 2^{1-y}\frac{\exp(-t')-\exp(-t)}{\exp(-2t)},$$

et ainsi $M_y(t,t') \to 0$ quand $y \to \infty$. La conclusion en découle. ■

Preuve du Lemme 2.16 – Rappelons que C est un sous ensemble compact tel que $[0,1] \subset C \subset (-\infty, 2)$. D'après (2.30) et le Lemme 2.14(ii), on a :

$$\begin{aligned}\frac{\partial}{\partial y}\psi(y;t,t') &= \log(t/t')\psi(y;t,t')\left(1+\frac{1}{\log(t/t')}\frac{t^{y-2}}{\mu_y(t)}(1+o(1))\right.\\ &- \left.\frac{1}{\log(t/t')}\frac{t'^{y-2}}{\mu_y(t')}(1+o(1))\right),\end{aligned}$$

pour tout $y \in C$ lorsque $t' \to \infty$. De plus, une application directe du Lemme 2.14(i) implique $t^{y-2}/\mu_y(t) = (1/t)(1+o(1))$. En remplaçant dans l'expression précédente et en rappelant que $t > t'$, alors :

$$\frac{\partial}{\partial y}\psi(y;t,t') = \log(t/t')\psi(y;t,t')\left(1+O\left(\frac{1}{t'\log(t/t')}\right)\right).$$

Le résultat est prouvé. ■

Preuve du Lemme 2.17 – Remarquons que :

$$\begin{aligned}\mathbb{P}\left(\frac{H_n(k_n)}{H_n(k_n')} \geq \frac{k_n'}{k_n}\right) &= \mathbb{P}\left(\frac{\widehat{\theta}_{n,\tau}(k_n)}{\widehat{\theta}_{n,\tau}(k_n')} \geq \frac{k_n'}{k_n}\frac{\mu_\tau(\log(n/k_n'))}{\mu_\tau(\log(n/k_n))}\right)\\ &= \mathbb{P}\left(\frac{\widehat{\theta}_{n,\tau}(k_n)}{\widehat{\theta}_{n,\tau}(k_n')} \geq \frac{k_n'}{k_n}\left(\frac{\log(n/k_n)}{\log(n/k_n')}\right)^{1-\tau}(1+o(1))\right)\end{aligned}$$

par le Lemme 2.14(i). On a $k_n'/k_n \to \infty$ qui implique $(\log(n/k_n)/\log(n/k_n'))^{1-\tau} \geq 1$ et comme par hypothèse $\widehat{\theta}_{n,\tau}(k_n)/\widehat{\theta}_{n,\tau}(k_n') \xrightarrow{\mathbb{P}} 1$ alors la conclusion en découle. ■

Preuve du Lemme 2.18 – Selon le Lemme 3 dans Gardes *et al.* [88], on a :

$$\begin{aligned}\widehat{\theta}_{n,\tau}(k_n) &= \frac{1}{\mu_\tau(\log(n/k_n))}\left\{\theta\frac{1}{k_n-1}\sum_{i=1}^{k_n-1}[K_\tau(F_i+E_{n-k_n+1,n})-K_\tau(E_{n-k_n+1,n})]\right.\\ &\left.+\frac{1}{k_n-1}\sum_{i=1}^{k_n-1}\log\left(\frac{\ell(\exp K_\tau(F_i+E_{n-k_n+1,n}))}{\ell(\exp K_\tau(E_{n-k_n+1,n}))}\right)\right\}. \quad (2.31)\end{aligned}$$

De plus, le Lemme 2(ii) et le Lemme 5 dans Gardes *et al.* [88] impliquent :

$$\frac{1}{\mu_\tau(E_{n-k_n+1,n})}\frac{1}{k_n-1}\sum_{i=1}^{k_n-1}\left[K_\tau(F_i+E_{n-k_n+1,n})-K_\tau(E_{n-k_n+1,n})\right]\xrightarrow{\mathbb{P}}1. \tag{2.32}$$

De plus, comme la fonction à variations lentes ℓ est supposée normalisée, il s'ensuit que :

$$\log\left(\frac{\ell(\exp(K_\tau(F_i+E_{n-k_n+1,n})))}{\ell(\exp(K_\tau(E_{n-k_n+1,n})))}\right)=\int_{\exp(K_\tau(E_{n-k_n+1,n}))}^{\exp(K_\tau(F_i+E_{n-k_n+1,n}))}\varepsilon(u)\frac{du}{u},$$

avec $\varepsilon(s)\to 0$ lorsque $s\to\infty$. Ainsi pour tout $i=1,\dots,k_n-1$, on a :

$$\begin{aligned}\left|\int_{\exp(K_\tau(E_{n-k_n+1,n}))}^{\exp(K_\tau(F_i+E_{n-k_n+1,n}))}\varepsilon(u)\frac{du}{u}\right| &\leq \sup_{u\geq\exp(K_\tau(E_{n-k_n+1,n}))}|\varepsilon(u)|\\ &\times \int_{\exp(K_\tau(E_{n-k_n+1,n}))}^{\exp(K_\tau(F_i+E_{n-k_n+1,n}))}\frac{du}{u}\\ &= o_{\mathbb{P}}(1)\left[K_\tau(F_i+E_{n-k_n+1,n})-K_\tau(E_{n-k_n+1,n})\right].\end{aligned}$$

Ainsi, (2.32) implique :

$$\frac{1}{\mu_\tau(\log(n/k_n))}\frac{1}{k_n-1}\sum_{i=1}^{k_n-1}\log\left(\frac{\ell(\exp K_\tau(F_i+E_{n-k_n+1,n}))}{\ell(\exp K_\tau(E_{n-k_n+1,n}))}\right)=o_{\mathbb{P}}(1). \tag{2.33}$$

En rassemblant (2.31)-(2.33), le Lemme 2.18 est prouvé. ■

Preuve du Lemme 2.21 – Premièrement, notons que les deux conditions :

$$\sqrt{k'_n}b(\exp K_\tau(\log n/k'_n))\to 0 \text{ et } k_n/k'_n\to 0 \text{ impliquent } \sqrt{k_n}b(\exp K_\tau(\log n/k_n))\to 0$$

du fait que $|b|$ est asymptotiquement décroissante. Le Lemme 2.19 affirme que :

$$\widehat{\theta}_{n,\tau}(k_n)=\theta+\theta\zeta_n/\sqrt{k_n}\quad\text{où}\quad \zeta_n\xrightarrow{d}\mathcal{N}(0,1)$$

et

$$\widehat{\theta}_{n,\tau}(k'_n)=\theta+\theta\zeta'_n/\sqrt{k'_n}\quad\text{où}\quad \zeta'_n\xrightarrow{d}\mathcal{N}(0,1).$$

Par conséquent, on a :

$$\begin{aligned}\frac{\widehat{\theta}_{n,\tau}(k_n)}{\widehat{\theta}_{n,\tau}(k'_n)} &= \frac{1+\zeta_n/\sqrt{k_n}}{1+\zeta'_n/\sqrt{k'_n}}\\ &= \left(1+\frac{\zeta_n}{\sqrt{k_n}}\right)\left(1-\frac{\zeta'_n}{\sqrt{k'_n}}+o_{\mathbb{P}}\left(\frac{1}{\sqrt{k'_n}}\right)\right)\\ &= 1+\frac{\zeta_n}{\sqrt{k_n}}+o_{\mathbb{P}}\left(\frac{1}{\sqrt{k_n}}\right),\end{aligned}$$

et donc :

$$\sqrt{k_n}\left(\frac{\widehat{\theta}_{n,\tau}(k_n)}{\widehat{\theta}_{n,\tau}(k_n')}-1\right)=\zeta_n+o_{\mathbb{P}}(1).$$

Cela conclut la preuve. ■

Preuve du Lemme 2.22 – Un développement de Taylor à l'ordre 1 implique :

$$\frac{\widehat{\theta}_{n,\widehat{\tau}_n}(k_n)}{\widehat{\theta}_{n,\tau}(k_n)}-1=\frac{\mu_\tau(\log(n/k_n))}{\mu_{\widehat{\tau}_n}(\log(n/k_n))}-1=(\tau-\widehat{\tau}_n)\frac{\partial/\partial y(\mu_y(\log(n/k_n)))\big|_{y=\tau_0}}{\mu_{\widehat{\tau}_n}(\log(n/k_n))},$$

avec $\tau_0=\widehat{\tau}_n+\eta(\tau-\widehat{\tau}_n)$ et $\eta\in\,]0,1[$. Soit :

$$w_n:=\frac{\sqrt{k_n}\,(\log_2(n/k_n)-\log_2(n/k_n'))}{\log_2(n/k_n)}\quad\text{et}\quad F_n(s):=\left\{w_n\left(\frac{\widehat{\theta}_{n,\widehat{\tau}_n}(k_n)}{\widehat{\theta}_{n,\tau}(k_n)}-1\right)\le s\right\}.$$

Rappelons que Φ est la fonction de répartition d'une loi Normale centrée réduite. Notre but est de prouver que $\mathbb{P}\left(F_n(s)\right)\to\Phi(s)$, pour tout $s\in\mathbb{R}$. En gardant à l'esprit que C est un sous ensemble compact tel que $[0,1]\subset C\subset(-\infty,2)$ et $A_n=\{\tau_0\in C\}\cap\{\widehat{\tau}_n\in C\}$, on a :

$$\mathbb{P}(F_n(s))=\mathbb{P}(F_n(s)\cap A_n)+\mathbb{P}\left(F_n(s)\,|A_n^c\right)\mathbb{P}\left(A_n^c\right)=:T_n^{(4)}(s)+T_n^{(5)}(s),$$

où A_n^c est l'événement complémentaire de A_n. Premièrement considérons $T_n^{(5)}(s)$. Le Théorème 2.10 affirme que :

$$\sqrt{k_n}\left(\log_2(n/k_n)-\log_2(n/k_n')\right)(\widehat{\tau}_n-\tau)=\xi_n\xrightarrow{d}\mathcal{N}(0,1).$$

Par conséquent, sous les hypothèses du Théorème 2.10, $\widehat{\tau}_n\xrightarrow{\mathbb{P}}\tau\in[0,1]$ et $\tau_0\xrightarrow{\mathbb{P}}\tau\in[0,1]$. Ainsi, $\mathbb{P}(A_n^c)\to 0$ ce qui implique $T_n^{(5)}(s)\to 0$. Concentrons-nous sur $T_n^{(4)}(s)$. En se plaçant sous l'événement A_n, et en combinant le Lemme 2.14(ii) avec le Théorème 2.10, il s'ensuit que :

$$\begin{aligned}
w_n\left(\frac{\widehat{\theta}_{n,\widehat{\tau}_n}(k_n)}{\widehat{\theta}_{n,\tau}(k_n)}-1\right) &= w_n(\tau-\widehat{\tau}_n)\left(\frac{\partial/\partial y(\mu_y(\log(n/k_n)))\big|_{y=\tau_0}}{\mu_{\widehat{\tau}_n}(\log(n/k_n))}\right)\\
&= \xi_n\frac{\log_2(n/k_n)\mu_{\tau_0}(\log(n/k_n))+\log(n/k_n)^{\tau_0-2}(1+o_{\mathbb{P}}(1))}{\log_2(n/k_n)\mu_{\widehat{\tau}_n}(\log(n/k_n))}\\
&= \xi_n\frac{\mu_{\tau_0}(\log(n/k_n))}{\mu_{\widehat{\tau}_n}(\log(n/k_n))}\left(1+\frac{1+o_{\mathbb{P}}(1)}{\log(n/k_n)\log_2(n/k_n)}\right),
\end{aligned}$$

d'après le Lemme 2.14(i). De plus, en se plaçant sous l'événement A_n, et en utilisant le Lemme 2.14(i) et le Théorème 2.10, on a :

$$\begin{aligned}
\frac{\mu_{\tau_0}(\log(n/k_n))}{\mu_{\widehat{\tau}_n}(\log(n/k_n))} &= (\log(n/k_n))^{\tau_0-\widehat{\tau}_n}(1+o_{\mathbb{P}}(1)) \\
&= (\log(n/k_n))^{\eta(\tau-\widehat{\tau}_n)}(1+o_{\mathbb{P}}(1)) \\
&= \exp(\eta(\tau-\widehat{\tau}_n)\log_2(n/k_n))(1+o_{\mathbb{P}}(1)) \\
&= \exp\left(-\frac{\eta\xi_n\log_2(n/k_n)}{\sqrt{k_n}(\log_2(n/k_n)-\log_2(n/k'_n))}\right)(1+o_{\mathbb{P}}(1)) \\
&\xrightarrow{\mathbb{P}} 1.
\end{aligned}$$

On obtient alors :

$$w_n\left(\frac{\widehat{\theta}_{n,\widehat{\tau}_n}(k_n)}{\widehat{\theta}_{n,\tau}(k_n)}-1\right)=\xi_n(1+o_{\mathbb{P}}(1)),$$

et, en remplaçant dans $T_n^{(4)}(s)$, on obtient finalement :

$$\begin{aligned}
T_n^{(4)}(s) &= \mathbb{P}(\{\xi_n \le s(1+o_{\mathbb{P}}(1))\}\cap A_n) \\
&= \mathbb{P}(\xi_n \le s(1+o_{\mathbb{P}}(1))) - \mathbb{P}(\{\xi_n \le s(1+o_{\mathbb{P}}(1))\}\,|A_n^c)\mathbb{P}(A_n^c).
\end{aligned}$$

En utilisant les mêmes arguments que dans la preuve du Théorème 2.10, on voit facilement que $\mathbb{P}(\xi_n \le s(1+o_{\mathbb{P}}(1))) \to \Phi(s)$ et ainsi $T_n^{(4)}(s) \to \Phi(s)$ puisque $\mathbb{P}(A_n^c) \to 0$. En combinant les résultats précédents la preuve du Lemme 2.22 est achevée. ■

Chapitre 3

Estimation non-paramétrique de mesures de risque extrêmes pour des lois conditionnelles à queues lourdes

Résumé

On commence dans la Partie 3.1 par une introduction sur la notion de mesure de risque où l'on motive nos travaux. On introduit ensuite dans la Partie 3.2 une nouvelle mesure de risque que l'on appelera le Conditional Tail Moment. Elle est définie comme le moment d'ordre $a \geq 0$ de la loi des pertes sachant que l'on se trouve au-dessus du quantile supérieur d'ordre α où $\alpha \in]0,1[$. Estimer le Conditional Tail Moment permet d'estimer toutes les mesures de risque basées sur les moments conditionnels telles que la Value-at-Risk, la Conditional Tail Expectation, la Conditional Value-at-Risk, la Conditional Tail Variance ou la Conditional Tail Skewness. Ici, on s'intéressera à l'estimation de ces mesures de risque dans le cas de pertes extrêmes (où $\alpha \to 0$ n'est plus fixé). On suppose aussi que la loi des pertes est à queue lourde et qu'elle dépend d'une covariable. Nos méthodes d'estimation combinent des méthodes d'estimation non-paramétrique à noyau avec des méthodes issues de la statistique des valeurs extrêmes. Dans la Partie 3.3, on établira le comportement asymptotique de nos estimateurs. On illustrera dans la Partie 3.4 ces comportements aussi bien sur des données simulées que sur des données réelles. Le jeu de données réelles est un jeu de données pluviométriques provenant de la région Cévennes-Vivarais se situant dans le sud de la France. Enfin, on donnera quelques perspectives dans la Partie 3.5. Les preuves de nos résultats théoriques se trouvent dans la Partie 3.6.

Sommaire

3.1 Introduction **101**
3.2 Le Regression Conditional Tail Moment : definition et estimation **107**
3.2.1 Une nouvelle mesure de risque 107
3.2.2 Pertes extrêmes et regression 108
3.2.3 Inférence 109
3.3 Principaux résultats **110**
3.4 Mesures de risque pour des extrêmes pluviométriques **116**
3.4.1 Description du problème et du jeu de données réelles 116
3.4.2 Choix des paramètres de réglage 120
3.4.3 Validation sur simulations 121
3.4.4 Mesures de risque estimées sur des extrêmes pluviométriques . . . 131
3.5 Perspectives **136**
3.6 Démonstrations **137**
3.6.1 Résultats préliminaires et leurs preuves 137
3.6.2 Preuves des principaux résultats 147

3.1 Introduction

"In the last fifty years, the ten most extreme days in the financial markets represent half the returns."

Nassim Nicholas Taleb

La maîtrise des risques est un sujet de préoccupation aussi bien en hydrologie qu'en finance et en actuariat. Se prémunir contre les risques en finance et en actuariat est primordial afin d'anticiper respectivement des crises financières ou d'importants sinistres coûteux pour les assurances. En hydrologie, l'étude des pluies et débits extrêmes est d'un grand intérêt afin de se prémunir contre des catastrophes telles que des inondations et ainsi permettre l'aménagement du territoire.

Pour cela, il existe des outils permettant de quantifier et de prédire le risque : les mesures de risque. Elles permettent d'évaluer un niveau de dangerosité d'un risque, mais également de comparer différents risques entre eux. La quantification, la prévention, la comparaison et la prédiction du risque sont des éléments essentiels de notre société.

Dans la suite, la variable aléatoire d'intérêt dont l'on souhaite quantifier le risque sera notée Y. Elle désignera également la variable aléatoire du montant de perte. Donnons à présent la définition précise d'une mesure de risque.

Définition 3.1. *On appelle mesure de risque une fonction $\mathcal{R}$ associant à Y une valeur positive ou nulle telle que :*

$$\mathcal{R} : Y \longrightarrow \mathbb{R}^+$$

De grandes valeurs de $\mathcal{R}(Y)$ indiqueront que Y est "dangereux". $\mathcal{R}(Y)$ peut être vu comme le capital à détenir pour faire face aux pertes Y.

Une des mesures de risque les plus connues, si ce n'est la plus connue, est la Value-at-Risk (notée VaR) introduite dans les années 90 par Morgan [135] (voir Jorion [123] pour plus de détails). C'est historiquement la plus utilisée des mesures de risque.

L'idée derrière le concept de la VaR est la suivante : on fixe un seuil probabiliste α et on calcule une valeur $\mathrm{VaR}(\alpha)$ qui sera telle que la probabilité que la catastrophe survienne soit plus petite que α. On peut voir la $\mathrm{VaR}(\alpha)$ comme le montant d'extra capital dont une entreprise a besoin afin de réduire à α la probabilité de faire faillite.

En termes statistiques, la VaR de niveau $\alpha \in]0, 1[$ correspond au quantile de niveau $1 - \alpha$ de la fonction de répartition (voir Définition 1.1) de la loi des pertes.

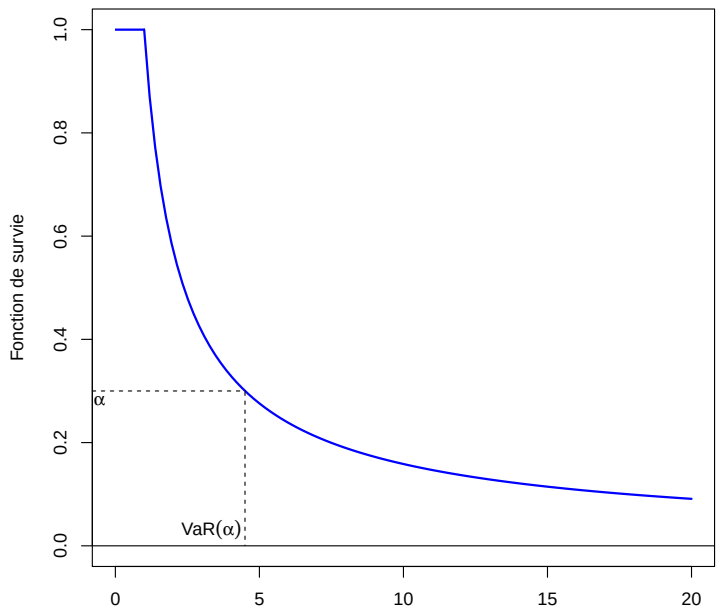


FIGURE 3.1: Graphique d'une fonction de survie modélisant la distribution des pertes. On a représenté une Value-at-Risk d'un certain niveau α.

Cette mesure de risque présente cependant plusieurs défauts.

En premier lieu, elle ne donne qu'une information ponctuelle au quantile $q(\alpha)$ et aucune information au-delà de ce point ; elle ne prend ainsi pas en compte l'importance du sinistre lorsqu'il survient mais seulement sa fréquence. De ce fait, comme elle ne prend pas en considération ce que serait la perte au-delà de ce quantile, elle ne fait donc pas la différence entre une catastrophe qui coûtera 1 million et celle qui coûtera 1 milliard d'euros. La VaR sous estime les pertes.

Un des principaux reproche fait à la VaR et que des variables aléatoires à queues légères et à queues lourdes (voir Tasche [162]) peuvent avoir la même VaR(α).

"Value at Risk is to finance what body temperature is to a patient, an indication of bad health but not an instrument telling us what is wrong and for less a clue on how to get the patient system healthy again".

Paul Embrechts

Pour répondre à la nécessité de principes théoriques et pratiques, Artzner *et al.* [6, 7] ont cherché à caractériser ce qui ferait qu'une mesure de risque soit "performante". Pour cela ils ont introduit la notion de mesure de risque cohérente.

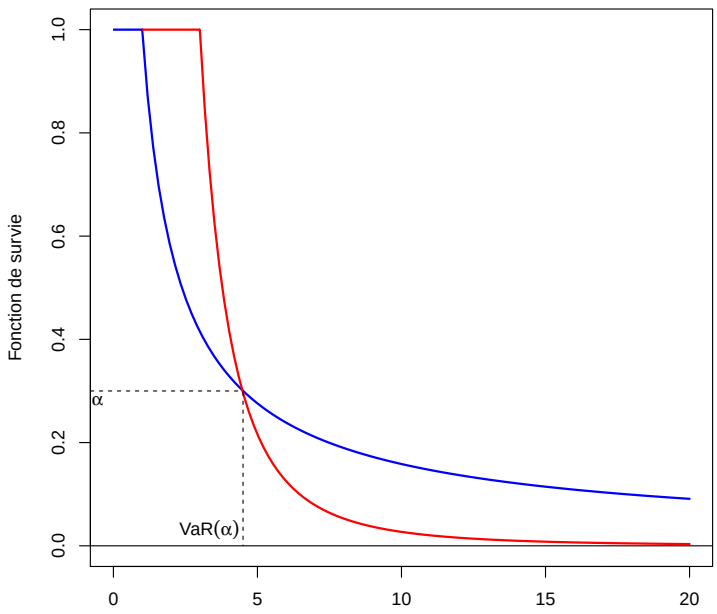


FIGURE 3.2: Illustration graphique de deux fonctions de survies distinctes ayant la même VaR(α).

Définition 3.2. *Une mesure de risque $\mathcal{R}$ est dite cohérente si, pour deux variables aléatoires Z et T, elle satisfait les 4 propriétés suivantes :*

1. *Monotonie : si $\mathbb{P}(Z \leq T) = 1$, alors $\mathcal{R}(Z) \leq \mathcal{R}(T)$,*
2. *Homogénéité positive : $\mathcal{R}(aZ) = a\mathcal{R}(Z)$, $\forall a > 0$,*
3. *Invariance par translation : $\mathcal{R}(Z + b) = \mathcal{R}(Z) + b$, $\forall b > 0$,*
4. *Sous-additivité : $\mathcal{R}(Z + T) \leq \mathcal{R}(Z) + \mathcal{R}(T)$.*

La monotonie permet de s'assurer que si la variable aléatoire Z est presque sûrement plus petite que T alors le risque associé à Z est plus petit que celui associé à T.

L'homogénéité positive signifie que le fait de mesurer une proportion d'un risque revient à considérer la proportion de la mesure du risque seul.

De même, l'invariance par translation implique qu'ajouter une valeur positive au risque implique l'ajout de cette même valeur à la mesure du risque.

L'homogénéité et l'invariance par translation impliquent qu'une mesure de risque cohérente apprécie la nature aléatoire du risque.

La sous-additivité traduit le fait que considérer deux risques simultanément est moins risqué que traiter les risques séparément. Cela intégre donc l'idée de diversification du risque.

Le principal défaut de la Value-at-Risk est qu'elle n'est pas une mesure de risque cohérente car elle ne vérifie pas la condition de sous-additivité. Une mesure de risque cohérente est la Conditional Tail Expectation (notée CTE) [7], aussi appelée dans la littérature sur les mesures de risque Tail-Value-at-Risk, Average-Value-at-Risk, Tail Conditional Expectation ou Expected Shortfall dans le cas d'une loi continue des pertes.

La VaR s'intéresse aux probabilités d'événements extrêmes alors que la CTE s'intéresse à ce qui se passe en moyenne lorsque ces événements surviennnent. En effet, la CTE au niveau α est définie comme la perte moyenne attendue sachant que la VaR(α) est dépassée. Elle représente donc la moyenne des $(1-\alpha)100\%$ sinistres les plus élevés et peut être vue comme le montant d'extra capital nécessaire afin de surmonter en moyenne les pires exercices.

Cette mesure de risque donne des informations sur la distribution de Y au-delà de la VaR(α). Ainsi, contrairement à la VaR(α), elle prend en compte l'information contenue dans la queue de la loi, ce qui est primordial aussi bien dans un contexte financier, actuariel ou hydrologique. L'épaisseur de la queue de la distribution des sinistres est un élément fondamental dans l'évaluation du niveau de danger d'un risque ayant une faible probabilité d'apparition. La Conditional Tail Expectation a été étudiée par Artzner *et al.* [7], Cai et Tan [29], Cai *et al.* [30], Tasche [162] et est fréquemment utilisée dans les investissements financiers ou dans l'actuariat voir Brazaukas *et al.* [24] et Landsman et Valdez [128].

A la suite de la VaR et de la CTE, Rockafellar et Uryasev [150] ont introduit une nouvelle mesure de risque cohérente, la Conditional-Value-at-Risk (notée CVaR) [151] définie comme la moyenne pondérée de la VaR et de la CTE.

Une autre mesure de risque basée sur la VaR et sur la CTE est la Stop-loss Premium réassurance (notée SP) avec un niveau de rétention égal à la VaR [29], voir la Partie 3.2 pour une définition plus précise.

Il existe de nombreuses autres mesures de risque, citons également, la Conditional Tail Variance (notée CTV) [163]. Elle représente la variance de la loi des pertes sachant que les pertes sont supérieures à la VaR. De même, la Conditional Tail Skewness (notée CTS) [117] est définie comme le coefficient d'asymétrie de la loi des pertes sachant que les pertes sont supérieures à la VaR. Le choix de la mesure la plus appropriée pour quantifier le risque en assurance comme en finance reste sujet à débat.

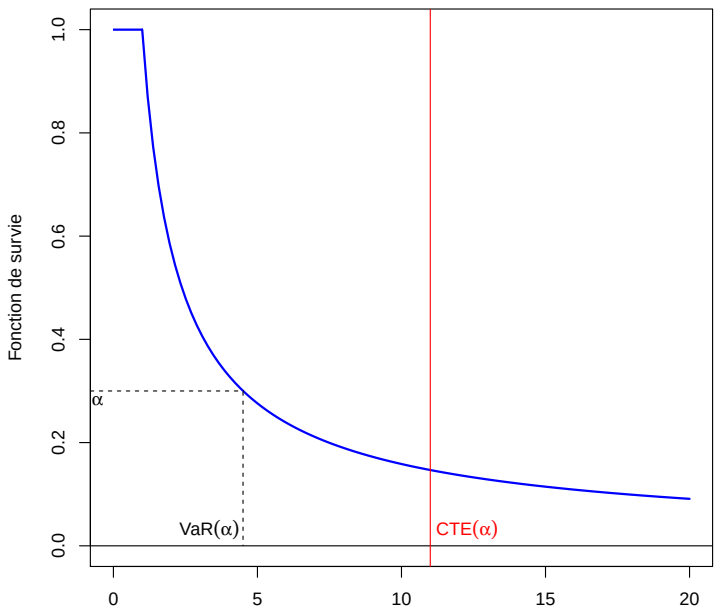


FIGURE 3.3: Graphique d'une fonction de survie modélisant la distribution des pertes. On a représenté une Value-at-Risk d'un certain niveau α et la CTE(α) correspondante à cette dernière.

Dans ce chapitre, on commencera par introduire un nouvel outil permettant d'unifier l'estimation des mesures de risque mentionnées juste avant. Ce nouvel outil sera appelé le Conditional Tail Moment (noté CTM). Il est défini comme le moment d'ordre $a \geq 0$ de la loi des pertes sachant que l'on se trouve au-dessus de la VaR de niveau α.

On montrera que l'estimation du CTM permet d'estimer toutes les mesures de risque basées sur les moments conditionnels d'ordre arbitraire sachant que les pertes sont supérieures à la VaR. Il est par exemple clair que le Conditional Tail Moment d'ordre un est la Conditional Tail Expectation.

Notre seconde contribution sera d'estimer le CTM dans le cas de pertes extrêmes, en se servant pour cela de la théorie des valeurs extrêmes. Bien que de nombreux liens entre la théorie des valeurs extrêmes et les mesures de risque aient fait l'objet de recherches citons notamment Embrechts [70], Embrechts *et al.* [71, 72, 73], McNeil *et al.* [133], l'estimation de mesures de risque est habituellement réalisée dans la littérature pour des valeurs fixées de α voir par exemple Deme *et al.* [57] et Necir *et al.* [136].

L'estimateur de mesures de risque extrêmes que l'on propose englobe également la présence d'une covariable. Dans ce contexte, notre nouvelle mesure de risque sera appelée le Regression Conditional Tail Moment (notée RCTM). En effet, en finance, la loi des

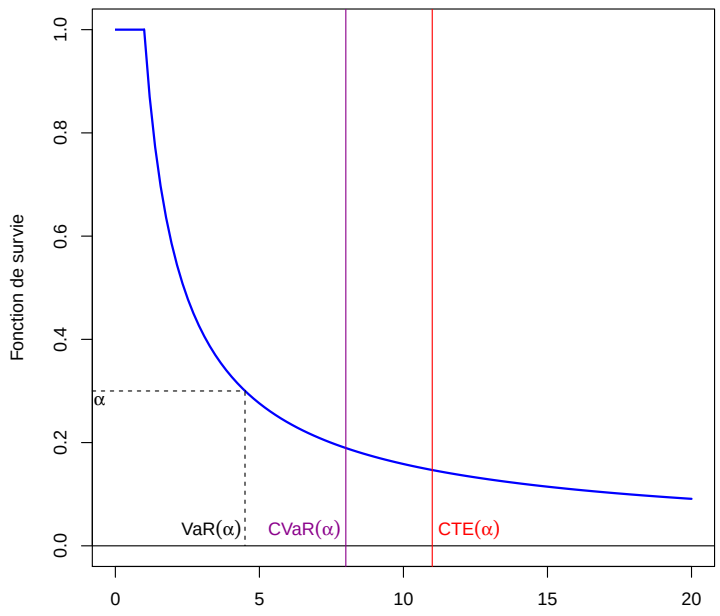


FIGURE 3.4: Graphique d'une fonction de survie modélisant la distribution des pertes. On a représenté une Value-at-Risk d'un certain niveau α, la CTE(α) associée ainsi que la CVaR(α) correnspondante à ces dernières.

pertes peut être affectée par de nombreux facteurs, tels que les taux d'intérêts ou l'inflation. De même les météorologues s'intéressent aux pluies extrêmes comme fonction de la position géographique [44, 85].

Le chapitre est organisé comme suit. La définition du RCTM et de ses liens avec les mesures de risque classiques sont donnés dans la Partie 3.2. Les propriétés asymptotiques de nos estimateurs sont établies dans la Partie 3.3. On illustrera l'efficacité de ces derniers sur des simulations et sur un jeu de données réelles dans la Partie 3.4. Enfin on donnera quelques perspectives dans la dernière Partie 3.5. Les preuves de nos résultats théoriques se trouvent dans la Partie 3.6.

Les contributions de ce chapitre ont fait l'objet du rapport de recherche El Methni *et al.* [68] soumis pour publication.

3.2 Le Regression Conditional Tail Moment : definition et estimation

3.2.1 Une nouvelle mesure de risque

Soit $Y \in \mathbb{R}$ une variable aléatoire de pertes. Pour $\alpha \in]0,1[$, la Value-at-Risk de niveau α est la quantité $\mathrm{VaR}(\alpha)$ satisfaisant :

$$\mathbb{P}(Y > \mathrm{VaR}(\alpha)) = \alpha.$$

Comme expliqué dans l'introduction, la Value at Risk est la plus connue des mesures de risque [135] mais de nombreuses autres existent dans la littérature :

- La Conditional Tail Expectation [7] est définie par :

$$\mathrm{CTE}(\alpha) := \mathbb{E}(Y|Y > \mathrm{VaR}(\alpha)).$$

- La Conditional Tail Variance a été introduite par Valdez [163] et est définie par :

$$\mathrm{CTV}(\alpha) := \mathbb{E}\left((Y - \mathrm{CTE}(\alpha))^2 |Y > \mathrm{VaR}(\alpha)\right).$$

Elle mesure la variabilité conditionnelle de Y sachant que $Y > \mathrm{VaR}(\alpha)$ et nous dit à quel point les événements s'éloignent de la $\mathrm{CTE}(\alpha)$.

- La Conditional Tail Skewness a été définie dans [117] par :

$$\mathrm{CTS}(\alpha) := \mathbb{E}\left(Y^3 |Y > \mathrm{VaR}(\alpha)\right) \Big/ (\mathrm{CTV}(\alpha))^{3/2}.$$

- La Conditional-Value-at-Risk est définie par :

$$\mathrm{CVaR}_\lambda(\alpha) := \lambda \mathrm{VaR}(\alpha) + (1-\lambda)\mathrm{CTE}(\alpha),$$

avec $0 \leq \lambda \leq 1$. Il est clair que $\mathrm{CVaR}_1(\alpha) = \mathrm{VaR}(\alpha)$ et $\mathrm{CVaR}_0(\alpha) = \mathrm{CTE}(\alpha)$. Cette mesure de risque permet de quantifier des risques au-delà de la $\mathrm{VaR}(\alpha)$ et elle est de plus cohérente (voir Définition 3.2) si $0 \leq \lambda < 1$ (voir Rockafellar et Uryasev [151]). D'autres propriétés fondamentales de cette mesure de risque se trouvent dans Rockafellar et Uryasev [150].

- La mesure de rique Stop-loss Premium réassurance avec un niveau de rétention égal à $\mathrm{VaR}(\alpha)$ [29] est définie par :

$$\mathrm{SP}(\alpha) := \mathbb{E}((Y - \mathrm{VaR}(\alpha))_+) = \alpha\left(\mathrm{CTE}(\alpha) - \mathrm{VaR}(\alpha)\right).$$

Elle est proportionnelle à la différence entre la CTE(α) et la VaR(α). Cette mesure de risque permet ainsi de mettre en évidence les cas dangereux.

Le but premier de ce chapitre est d'unifier les définitions des mesures de rique énoncées juste auparavant. Pour cela, on introduit une nouvelle mesure de risque, le Conditional Tail Moment (noté CTM) définie par :

$$\mathrm{CTM}_a(\alpha) := \mathbb{E}(Y^a | Y > \mathrm{VaR}(\alpha)),$$

où $a \geq 0$ est tel que le moment d'ordre a de Y existe. Il est facile de voir que toutes les mesures de risque précédentes de niveau α peuvent être réécrites :

$$\Phi(\mathrm{VaR}(\alpha), \mathrm{CTM}_1(\alpha), \mathrm{CTM}_2(\alpha), \mathrm{CTM}_3(\alpha)),$$

où la fonction $\Phi : \mathbb{R}^4 \mapsto \mathbb{R}$ est prise dans le Tableau 3.1 ci-dessous :

Mesure de risque	$\Phi(t_0, t_1, t_2, t_3)$
CTE(α)	t_1
CTV(α)	$t_2 - t_1^2$
CTS(α)	$t_3/(t_2 - t_1^2)^{3/2}$
CVaR(α)	$\lambda t_0 + (1-\lambda)t_1,\ \lambda \in [0,1]$
SP(α)	$\alpha(t_1 - t_0)$

Table 3.1: Liens entre le CTM et les mesures de risque classiques.

De manière plus générale, le CTM peut être utilisé dans le but de définir n'importe quelle mesure de risque basée sur les moments conditionnels de la variable aléatoire des pertes sachant que les pertes sont supérieures à la VaR de niveau α. En effet, on peut par exemple introduire le Conditional Tail Kurtosis à l'aide de la fonction :

$$\Phi(t_0, t_1, t_2, t_3, t_4) = \frac{t_4}{(t_2 - t_1^2)^2}.$$

3.2.2 Pertes extrêmes et regression

Comme évoqué dans l'introduction, notre second but est d'adapter les mesures de risque classiques dans le cas de pertes extrêmes en présence d'une covariable $X \in \mathbb{R}^p$. Pour cela, le niveau fixé $\alpha \in]0,1[$ est remplacé par une suite $(\alpha_n) \in]0,1[$, telle que $\alpha_n \to 0$. De plus, si l'on note par $\overline{F}(.|x)$ la fonction de survie conditionnelle de Y sachant $X = x$, on peut alors définir la Regression Value-at Risk par :

$$\mathrm{RVaR}(\alpha_n|x) := \overline{F}^{\leftarrow}(\alpha_n|x) = \inf\{t, \overline{F}(t|x) \leq \alpha_n\},$$

et le Regression Conditional Tail Moment d'ordre a par :

$$\mathrm{RCTM}_a(\alpha_n|x) := \mathbb{E}(Y^a|Y > \mathrm{RVaR}(\alpha_n|x), X = x),$$

où $a \geq 0$ est tel que le moment d'ordre a de Y existe. Notons que dans ce cadre d'étude, la $\mathrm{RVaR}(\alpha_n|x)$ est le quantile conditionnel extrême de niveau $\alpha_n \in]0,1[$ (voir Définition 1.24 au chapitre 1).

Il est alors assez facile d'adapter les mesures de risque classiques dans le cas de pertes extrêmes en présence d'une covariable en appliquant la fonction adéquate (voir le Tableau 3.1) au vecteur :

$$(\mathrm{RVaR}(\alpha_n|x), \mathrm{RCTM}_1(\alpha_n|x), \mathrm{RCTM}_2(\alpha_n|x), \mathrm{RCTM}_3(\alpha_n|x)).$$

Cela nous donne les mesures de risque suivantes :

$$\mathrm{RCTE}(\alpha_n|x),\ \mathrm{RCTV}(\alpha_n|x),\ \mathrm{RCTS}(\alpha_n|x),\ \mathrm{RCVaR}(\alpha_n|x) \text{ et } \mathrm{RSP}(\alpha_n|x).$$

3.2.3 Inférence

Soient (X_i, Y_i), $i = 1, \ldots, n$, des copies indépendantes du couple de variables aléatoires (X, Y). Afin d'estimer le RCTM, on commence par donner l'égalité suivante :

$$\mathrm{RCTM}_a(\alpha_n|x) = \frac{1}{\alpha_n}\varphi_a(\varphi_0^{\leftarrow}(\alpha_n|x)|x), \tag{3.1}$$

où pour $y > 0$, le moment conditionnel d'ordre $a \geq 0$ est donné par :

$$\varphi_a(y|x) = \mathbb{E}\left(Y^a \mathbb{I}\{Y > y\}|X = x\right). \tag{3.2}$$

L'estimation du RCTM dépend ainsi de l'estimation du moment conditionnel. On propose d'utiliser un estimateur à noyau classique (voir Parzen [139] et Rosenblatt [154]) donné par :

$$\hat{\varphi}_{a,n}(y|x) = \sum_{i=1}^{n} K_h(x - X_i) Y_i^a \mathbb{I}\{Y_i > y\} \Big/ \sum_{i=1}^{n} K_h(x - X_i), \tag{3.3}$$

où $\mathbb{I}\{.\}$ est la fonction indicatrice et $h = h_n$ est une suite non aléatoire telle que $h \to 0$ lorsque $n \to \infty$. Dans le cas où $a = 0$ dans l'équation (3.3), on retrouve l'estimateur à noyau de la fonction de survie conditionnelle introduit par Collomb [36]. On a aussi introduit $K_h(t) = K(t/h)/h^p$ où K est une densité sur $\mathbb{R}^p$. Dans ce contexte d'étude, h est appelé paramètre de lissage.

Puisque $\hat{\varphi}_{a,n}(.|x)$ est une fonction décroissante, on peut définir un estimateur de $\varphi_a^{\leftarrow}(\alpha|x)$ pour $\alpha \in]0,1[$ par :

$$\hat{\varphi}_{a,n}^{\leftarrow}(\alpha|x) = \inf\{t, \hat{\varphi}_{a,n}(t|x) < \alpha\}. \tag{3.4}$$

En remarquant que $\varphi_0(y|x) = \overline{F}(y|x)$, la RVaR de niveau α_n est alors estimée par :

$$\widehat{\mathrm{RVaR}}_n(\alpha_n|x) = \hat{\varphi}_{0,n}^{\leftarrow}(\alpha_n|x).$$

On retrouve ainsi l'estimateur du quantile conditionnel extrême étudié par Daouia *et al.* [43, 44]. Pour plus d'exemples d'estimateurs du quantile conditionnel extrême, on pourra se reporter à la sous-Partie 1.6 du chapitre 1. Le RCTM d'ordre a est alors estimé par :

$$\widehat{\mathrm{RCTM}}_{a,n}(\alpha_n|x) = \frac{1}{\alpha_n}\hat{\varphi}_{a,n}(\hat{\varphi}_{0,n}^{\leftarrow}(\alpha_n|x)|x). \tag{3.5}$$

Un estimateur de chaque mesure de risque mentionnée auparavant est alors donné par :

$$\Phi(\widehat{\mathrm{RVaR}}_n(\alpha_n|x), \widehat{\mathrm{RCTM}}_{1,n}(\alpha_n|x), \widehat{\mathrm{RCTM}}_{2,n}(\alpha_n|x), \widehat{\mathrm{RCTM}}_{3,n}(\alpha_n|x)), \tag{3.6}$$

où la fonction Φ est choisie dans le Tableau 3.1. Les estimateurs obtenus seront notés :

$$\widehat{\mathrm{RCTE}}_n(\alpha_n|x),\ \widehat{\mathrm{RCTV}}_n(\alpha_n|x),\ \widehat{\mathrm{RCTS}}_n(\alpha_n|x),\ \widehat{\mathrm{RCVaR}}_{\lambda,n}(\alpha_n|x) \text{ et } \widehat{\mathrm{RSP}}_n(\alpha_n|x).$$

Comme exemple, on a l'estimateur de la RCTE qui est simplement donné par :

$$\widehat{\mathrm{RCTE}}_n(\alpha_n|x) = \frac{1}{\alpha_n}\hat{\varphi}_{1,n}(\hat{\varphi}_{0,n}^{\leftarrow}(\alpha_n|x)|x).$$

Comme on peut le voir dans l'équation (3.6), les mesures de risque classiques dépendent de la RVaR et du RCTM. En conséquence, on établira la loi jointe asymptotique des estimateurs du RCTM et de la RVaR dans la partie suivante. La loi asymptotique de toutes les mesures de risque classiques sera une conséquence directe de ce résultat.

3.3 Principaux résultats

Notre hypothèse principale est la suivante :

(F.1) On suppose que la fonction de survie conditionnelle de Y sachant $X = x$ est à queue lourde et admet une fonction de densité.

En résumé, **(F.1)** revient à supposer que la loi conditionnelle de Y sachant $X = x$ appartient au domaine d'attraction de Fréchet. L'hypothèse **(F.1)** équivaut à supposer que pour tout $y > 0$, $\overline{F}(y|x) = \mathbb{P}(Y > y|X = x)$ est à variations régulières à l'infini

(voir Bingham *et al.* [21]) avec un indice $-1/\gamma(x)$ que l'on notera $\overline{F}(.|x) \in \mathcal{RV}_{-1/\gamma(x)}$ *i.e* pour tout $\lambda > 0$ on a :

$$\lim_{y\to\infty} \frac{\overline{F}(\lambda y|x)}{\overline{F}(y|x)} = \lambda^{-1/\gamma(x)}.$$

Dans ce contexte, $\gamma(.)$ est une fonction positive de la covariable x et sera appelée indice de queue conditionnel puisqu'elle contrôle la lourdeur de la queue de la loi conditionnelle de Y sachant $X = x$.

On peut alors remarquer que sous l'hypothèse **(F.1)**, une condition suffisante pour l'existence du $\mathrm{RCTM}_a(1/.|x)$ est $a < 1/\gamma(x)$. Comme établi dans le Lemme 3.9, la condition **(F.1)** implique également que pour tout $a \in [0, 1/\gamma(x)[$, $\varphi_a(.|x) \in \mathcal{RV}_{a-1/\gamma(x)}$. Puisque de plus, $\overline{F}(.|x) \in \mathcal{RV}_{-1/\gamma(x)}$, on a alors : $\mathrm{RCTM}_a(1/.|x) \in \mathcal{RV}_{a\gamma(x)}$. Cela équivaut à supposer que pour $a \in [0, 1/\gamma(x)[$ et pour tout $y > 0$,

$$\mathrm{RCTM}_a(1/y|x) = y^{a\gamma(x)}\ell_a(y|x), \tag{3.7}$$

où pour x fixé, $\ell_a(.|x)$ est une fonction à variations lentes à l'infini, *i.e* pour tout $\lambda > 0$,

$$\lim_{y\to\infty} \frac{\ell_a(\lambda y|x)}{\ell_a(y|x)} = 1. \tag{3.8}$$

Afin d'établir la normalité asymptotique de (3.5), les conditions suivantes sont nécessaires. Premièrement, comme il est remarqué dans [21] p.15, puisque les fonctions à variations lentes sont d'un intérêt seulement asymptotique, on peut ainsi (sans perdre en généralité) supposer que dans (3.7) :

(F.2) $\ell_a(.|x)$ est normalisée pour tout $a \in [0, 1/\gamma(x)[$.

Dans ce cas, la représentation de Karamata (voir le Théorème 1.9) d'une fonction à variations lentes nous donne :

$$\ell_a(y|x) = c_a(x) \exp\left(\int_1^y \frac{\varepsilon_a(u|x)}{u} du\right), \tag{3.9}$$

où $c_a(.)$ est une fonction positive et $\varepsilon_a(y|x) \to 0$ lorsque $y \to \infty$. Ainsi, $\ell_a(.|x)$ est différentiable et la fonction auxiliaire est donnée par $\varepsilon_a(y|x) = y\ell_a'(y|x)/\ell_a(y|x)$. Cette fonction joue un rôle important en théorie des valeurs extrêmes car elle contrôle la vitesse de convergence dans l'équation (3.8), et plus généralement le biais des estimateurs en théorie des valeurs extrêmes. Comme estimateur à noyau de cette fonction, on pourra citer les travaux de Goegebeur *et al.* [99].

Goegebeur et de Wet [100] supposent que la fonction auxiliaire est à variations régulières à l'infini et s'intéressent à l'estimation de l'indice conditionnel de variations régulières. Ici, on se limite à faire l'hypothèse que pour tout $a \in [0, 1/\gamma(x)[$,

(F.3) $|\varepsilon_a(.|x)|$ est continue et décroissante à l'infini.

Une condition de type Lipschitz sur la fonction densité g de X est aussi nécessaire. Pour tout $(x, x') \in \mathbb{R}^p \times \mathbb{R}^p$, la distance entre x et x' sera notée $d(x, x')$. On peut introduire l'hypothèse suivante :

(L) Il existe une constante $c_g > 0$ telle que $|g(x) - g(x')| \leq c_g d(x, x')$.

L'hypothèse suivante est classique lorsque l'on fait de l'estimation par noyau.

(K) K est une densité bornée sur $\mathbb{R}^p$, de support S inclu dans la boule unité de $\mathbb{R}^p$.

Pour $\xi > 0$, la plus grande oscillation au point $(x, y) \in \mathbb{R}^p \times \mathbb{R}_*^+$ du moment conditionnel d'ordre $a \in [0, 1/\gamma(x)[$ est donnée par :

$$\omega(x, y, a, \xi, h) = \sup \left\{ \left| \frac{\varphi_a(z|x)}{\varphi_a(z|x')} - 1 \right|, \ z \in [(1-\xi)y, (1+\xi)y] \text{ and } \ x' \in B(x, h) \right\},$$

où $B(x, h)$ est la boule de centre x et de rayon h. Enfin, pour tout ensemble fini E, l'ensemble $\mathcal{L}(E)$ est défini par :

$$\mathcal{L}(E) = \{e_i + e_j, (e_i, e_j) \in E \times E\} \cup E.$$

On peut dès lors établir notre résultat principal.

Théorème 3.3. *Supposons que* **(F.1)**, **(F.2)**, **(L)** *et* **(K)** *soient vérifiées. Introduisons* $0 \leq a_1 < a_2 < \cdots < a_J$ *où* J *est un entier positif. Pour tout* $x \in \mathbb{R}^p$ *tel que* $g(x) > 0$ *et* $\gamma(x) < 1/(2a_J)$, *on introduit une suite* (α_n) *avec* $\alpha_n \to 0$ *et* $nh^p\alpha_n \to \infty$ *quand* $n \to \infty$. *S'il existe* $\xi > 0$ *tel que*

$$nh^p\alpha_n \left(h \vee \max_{a \in \mathcal{L}(\{0, a_1, \ldots, a_J\})} \omega(x, \varphi_0^{\leftarrow}(\alpha_n|x), a, \xi, h) \right)^2 \to 0,$$

alors, le vecteur aléatoire :

$$\sqrt{nh^p\alpha_n} \left\{ \left(\frac{\widehat{\mathrm{RCTM}}_{a_j,n}(\alpha_n|x)}{\mathrm{RCTM}_{a_j}(\alpha_n|x)} - 1 \right)_{j \in \{1, \ldots, J\}}, \left(\frac{\widehat{\mathrm{RVaR}}_n(\alpha_n|x)}{\mathrm{RVaR}(\alpha_n|x)} - 1 \right) \right\}$$

est asymptotiquement Gaussien, centré, avec une matrice de covariance de taille $(J+1)\times(J+1)$ *donnée par* $\|K\|_2^2\gamma^2(x)\Sigma(x)/g(x)$ *où pour* $(i,j)\in\{1,\dots,J\}^2$ *on a :*

$$\Sigma_{i,j}(x)=\frac{a_ia_j(2-(a_i+a_j)\gamma(x))}{(1-(a_i+a_j)\gamma(x))},$$

$$\Sigma_{J+1,j}(x)=a_j,\quad \Sigma_{i,J+1}(x)=a_i,\quad \textit{et}\quad \Sigma_{J+1,J+1}(x)=1.$$

Autrement écrit, la matrice de variance-covariance asymptotique est

$$\frac{\|K\|_2^2\gamma^2(x)}{g(x)}\left(\begin{array}{c|c} \frac{a_ia_j(2-(a_i+a_j)\gamma(x))}{(1-(a_i+a_j)\gamma(x))} & \begin{matrix} a_1\\ \vdots\\ a_J\end{matrix} \\ \hline a_1\cdots a_J & 1\end{array}\right)$$

Elle est proportionnelle à $\gamma^2(x)$ ainsi si $\gamma(x)$ augmente (*i.e.* plus la queue est lourde) plus la variance de nos estimateurs augmente.

La densité $g(x)$ de la covariable est au dénominateur de la matrice de variance-covariance asymptotique ainsi moins il y a de point (*i.e.* plus la densité est faible) plus la variance des estimateurs sera grande.

La condition $nh^p\alpha_n\to\infty$ est nécéssaire et suffisante pour qu'il y ait presque sûrement au moins une observation dans la région $B(x,h)\times[\varphi_0^{\leftarrow}(\alpha_n|x),+\infty[$ de $\mathbb{R}^p\times\mathbb{R}$.

La condition

$$\sqrt{nh^p\alpha_n}\left(h\vee\max_{a\in\mathcal{L}(\{0,a_1,\dots,a_J\})}\omega(x,\varphi_0^{\leftarrow}(\alpha_n|x),a,\xi,h)\right)\to 0$$

implique que le biais induit par le lissage est négligeable en comparaison de l'écart type.

Le Théorème 3.3 permet d'établir la normalité asymptotique de n'importe quel estimateur d'une mesure de risque basée sur un moment arbitraire au-dessus d'un quantile conditionnel extrême.

En particulier, les estimateurs $\widehat{\mathrm{RCTE}}_n(\alpha_n|x)$, $\widehat{\mathrm{RCVaR}}_{\lambda,n}(\alpha_n|x)$ et $\widehat{\mathrm{RSP}}_n(\alpha_n|x)$ reposent sur le moment d'ordre un, leur normalité asymptotique peut donc être obtenue sous l'hypothèse $\gamma(x)<1/2$.

Si $\alpha_n=\alpha$ fixé on retrouve la condition de normalité asymptotique classique : $nh^p\to\infty$. En l'absence de covariable on retrouve la condition de normalité asymptotique classique en théorie des valeurs extrêmes : $n\alpha_n\to\infty$.

Corollaire 3.4. *Supposons que* **(F.1)**, **(F.2)**, **(L)** *et* **(K)** *soient vérifiées. Pour tout* $x \in \mathbb{R}^p$ *tel que* $g(x) > 0$ *et* $\gamma(x) < 1/2$, *introduisons une suite* (α_n) *avec* $\alpha_n \to 0$ *et* $nh^p\alpha_n \to \infty$ *as* $n \to \infty$. *S'il existe* $\xi > 0$ *tel que*

$$nh^p\alpha_n \left(h \vee \max_{a\in\{0,1,2\}} \omega(x, \varphi_0^{\leftarrow}(\alpha_n|x), a, \xi, h) \right)^2 \to 0,$$

alors on a :

$$\begin{aligned}
\sqrt{nh^p\alpha_n}\left(\frac{\widehat{\mathrm{RCTE}}_n(\alpha_n|x)}{\mathrm{RCTE}(\alpha_n|x)} - 1\right) &\xrightarrow{d} \mathcal{N}\left(0, \frac{2\gamma^2(x)(1-\gamma(x))}{1-2\gamma(x)}\frac{\|K\|_2^2}{g(x)}\right),\\
\sqrt{nh^p\alpha_n}\left(\frac{\widehat{\mathrm{RCVaR}}_{\lambda,n}(\alpha_n|x)}{\mathrm{RCVaR}_\lambda(\alpha_n|x)} - 1\right) &\xrightarrow{d} \mathcal{N}\left(0, \frac{\gamma^2(x)(\lambda^2+2-2\lambda-2\gamma(x))}{1-2\gamma(x)}\frac{\|K\|_2^2}{g(x)}\right),\\
\sqrt{nh^p\alpha_n}\left(\frac{\widehat{\mathrm{RSP}}_n(\alpha_n|x)}{\mathrm{RSP}(\alpha_n|x)} - 1\right) &\xrightarrow{d} \mathcal{N}\left(0, \frac{\gamma^2(x)}{1-2\gamma(x)}\frac{\|K\|_2^2}{g(x)}\right).
\end{aligned}$$

L'estimateur de la $\mathrm{RCTV}(\alpha_n|x)$ repose sur le moment d'ordre deux. Afin d'obtenir sa normalité asymptotique, on a besoin d'une condition plus forte sur $\gamma(x)$ qui est $\gamma(x) < 1/4$.

Corollaire 3.5. *Supposons que* **(F.1)**, **(F.2)**, **(L)** *et* **(K)** *soient vérifiées. Pour tout* $x \in \mathbb{R}^p$ *tel que* $g(x) > 0$ *et* $\gamma(x) < 1/4$, *introduisons une suite* (α_n) *avec* $\alpha_n \to 0$ *et* $nh^p\alpha_n \to \infty$ *quand* $n \to \infty$. *S'il existe* $\xi > 0$ *tel que*

$$nh^p\alpha_n \left(h \vee \max_{a\in\{0,\dots,4\}} \omega(x, \varphi_0^{\leftarrow}(\alpha_n|x), a, \xi, h) \right)^2 \to 0,$$

alors on a :

$$\sqrt{nh^p\alpha_n}\left(\frac{\widehat{\mathrm{RCTV}}_n(\alpha_n|x)}{\mathrm{RCTV}(\alpha_n|x)} - 1\right) \xrightarrow{d} \mathcal{N}\left(0, V(x)\frac{\|K\|_2^2}{g(x)}\right),$$

où

$$V(x) = \frac{8(1-\gamma(x))(1-2\gamma(x))(1+2\gamma(x)+3\gamma^2(x))}{(1-3\gamma(x))(1-4\gamma(x))}.$$

De même, l'estimateur de la $\mathrm{RCTS}(\alpha_n|x)$ repose sur le moment d'ordre trois. Afin d'obtenir sa normalité asymptotique, on a besoin d'une condition encore plus forte sur $\gamma(x)$ qui est $\gamma(x) < 1/6$.

Corollaire 3.6. *Supposons que* **(F.1)**, **(F.2)**, **(L)** *et* **(K)** *soient vérifiées. Pour tout* $x \in \mathbb{R}^p$ *tel que* $g(x) > 0$ *et* $\gamma(x) < 1/6$, *introduisons une suite* (α_n) *avec* $\alpha_n \to 0$ *et* $nh^p\alpha_n \to \infty$ *quand* $n \to \infty$. *S'il existe* $\xi > 0$ *tel que*

$$nh^p\alpha_n \left(h \vee \max_{a\in\{0,\dots,6\}} \omega(x, \varphi_0^{\leftarrow}(\alpha_n|x), a, \xi, h) \right)^2 \to 0,$$

alors on a :

$$\sqrt{nh^p\alpha_n} \left(\frac{\widehat{\mathrm{RCTS}}_n(\alpha_n|x)}{\mathrm{RCTS}(\alpha_n|x)} - 1 \right) \xrightarrow{d} \mathcal{N}\left(0, \nu(x)\frac{\|K\|_2^2}{g(x)}\right),$$

où

$$\nu(x) = \frac{18(1 - 13\gamma(x) + 50\gamma^2(x) - 44\gamma^3(x) - 23\gamma^4(x) - 3\gamma^5(x))}{(1-3\gamma(x))(1-4\gamma(x))(1-5\gamma(x))(1-6\gamma(x))}.$$

Dans le Théorème 3.3, la condition $nh^p\alpha_n \to 0$ nous donne une borne inférieure sur le niveau de la mesure de risque que l'on souhaite estimer. Cette restriction est une conséquence de l'utilisation de l'estimateur à noyau (3.3) qui ne permet pas d'extrapoler au-delà du maximum de l'échantillon dans la boule $B(x, h)$. En conséquence α_n doit être un ordre de quantile extrême se trouvant dans l'échantillon.

Afin de passer outre cette restriction, on propose d'adapter l'estimateur de Weissman (initialement construit pour l'estimation de quantile non conditionnel, voir Définition 1.17) à l'estimation du RCTM :

$$\widehat{\mathrm{RCTM}}^W_{a,n}(\beta_n|x) = \widehat{\mathrm{RCTM}}_{a,n}(\alpha_n|x) \left(\frac{\alpha_n}{\beta_n}\right)^{a\hat{\gamma}_n(x)}, \tag{3.10}$$

où a est une valeur fixée, $0 < \beta_n < \alpha_n$ et $\hat{\gamma}_n(x)$ est un estimateur de l'indice de queue conditionnel $\gamma(x)$. Pour plus d'exemples d'estimateurs de l'indice de queue conditionnel, on pourra se reporter à la sous-Partie 1.6 du chapitre 1.

Comme illustré dans le théorème suivant, le facteur d'extrapolation $(\alpha_n/\beta_n)^{a\hat{\gamma}_n(x)}$ nous permet d'estimer le RCTM pour des niveaux β_n arbitrairement petits.

Théorème 3.7. *Supposons que les hypothèses du Théorème 3.3 soient vérifiées ainsi que la condition* **(F.3)**. *Considérons* $\hat{\gamma}_n(x)$ *un estimateur de l'indice de queue tel que*

$$\sqrt{nh^p\alpha_n}(\hat{\gamma}_n(x) - \gamma(x)) \xrightarrow{d} \mathcal{N}\left(0, v^2(x)\right),$$

avec $v(x) > 0$. *Si, de plus,* $(\beta_n)_{n\geq 1}$ *est une suite positive telle que* $\beta_n \to 0$, $\beta_n/\alpha_n \to 0$ *et* $\sqrt{nh^p\alpha_n}\varepsilon_a(1/\alpha_n|x) \to 0$ *quand* $n \to \infty$, *on a alors :*

$$\frac{\sqrt{nh^p\alpha_n}}{\log(\alpha_n/\beta_n)} \left(\frac{\widehat{\mathrm{RCTM}}^W_{a,n}(\beta_n|x)}{\mathrm{RCTM}_a(\beta_n|x)} - 1 \right) \xrightarrow{d} \mathcal{N}\left(0, (av(x))^2\right).$$

La condition $\sqrt{nh^p\alpha_n}\varepsilon_a(1/\alpha_n|x) \to 0$ implique que le biais induit par la fonction à variations lentes normalisée est négligeable en comparaison de l'écart type.

La condition $\beta_n/\alpha_n \to 0$ permet d'extrapoler et donc de choisir un ordre β_n arbitrairement petit.

Notons que la normalité asymptotique de :

$$\widehat{\mathrm{RVaR}}_n^W(\beta_n|x) = \widehat{\mathrm{RVaR}}_n(\alpha_n|x)\left(\frac{\alpha_n}{\beta_n}\right)^{\hat{\gamma}_n(x)}$$

a été établie par Daouia *et al.* [44]. Remarquons de plus que l'on retrouve exactement l'estimateur de Weissman adapté au cas conditionnel (voir Définition 1.26). Par conséquent, en remplaçant $\widehat{\mathrm{RVaR}}_n$ par $\widehat{\mathrm{RVaR}}_n^W$ et $\widehat{\mathrm{RCTM}}_{a,n}$ par $\widehat{\mathrm{RCTM}}_{a,n}^W$ dans (3.6) cela nous founit des estimateurs pour toutes les mesures de risque considérées dans ce chapitre pour des niveaux arbitrairement petits. Leurs normalités asymptotiques sont une simple conséquence du Théorème 3.7.

Dans la partie suivante, on appliquera nos estimateurs à un jeu de données pluviométriques afin d'estimer des mesures de risque associées à des pluies extrêmes.

3.4 Mesures de risque pour des extrêmes pluviométriques

Le jeu de données pluviométriques est décrit dans la sous-Partie 3.4.1. L'utilisation de nos estimateurs de mesures de risque nécessite de choisir nos deux paramètres de réglage h et α. Une procédure automatique est proposée dans la sous-Partie 3.4.2. L'efficacité de cette procédure est validée sur des simulations dans la sous-Partie 3.4.3. Finalement, dans la dernière sous-Partie 3.4.4, l'ensemble de nos résultats (procédure et estimateurs de mesures de risque) sont appliqués à ce jeu de données réelles.

3.4.1 Description du problème et du jeu de données réelles

Le comportement et l'efficacité de nos estimateurs sont illustrés sur un jeu de données pluviométriques provenant de la région Cévennes-Vivarais se situant au sud de la France. Ce jeu de données nous provient de Météo-France et consiste en des relevés journaliers de pluie mesurés en $N = 523$ stations pluviométriques de 1958 à 2000.

Dans ce contexte d'étude, la variable d'intérêt Y est la hauteur de pluie journalière mesurée en millimètres (mm). Le nombre de mesures varie en chaque station $t \in \{1, \ldots, N\}$, on le notera n_t. Le nombre total d'observations étant $n = \sum_{t=1}^N n_t = 5513734$. Pour X, on prendra les coordonnées géographiques des stations en trois dimensions (longitude,

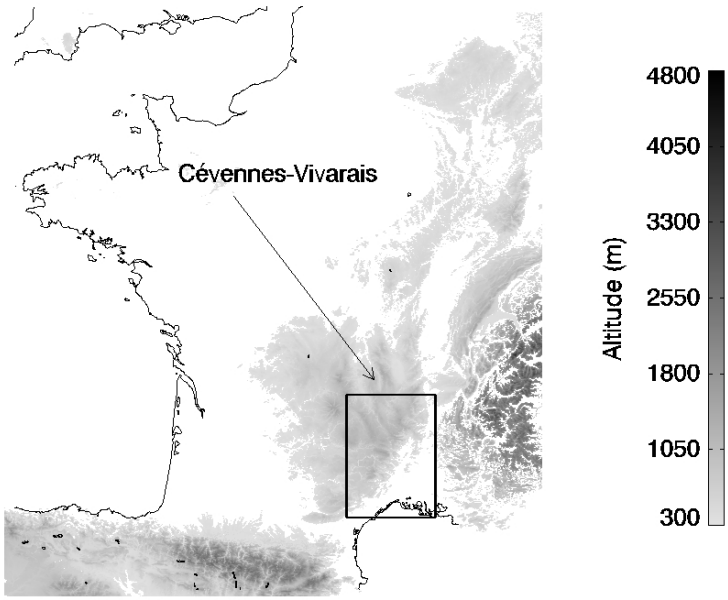


FIGURE 3.5: Carte situant la région Cévennes-Vivarais en France avec en échelle de couleur l'altitude.

latitude et altitude). L'ensemble des stations est représenté sur la Figure 3.6. On notera les coordonnées des stations :

$$S = \{x_t = (x_{1,t}, x_{2,t}, x_{3,t}); t = 1, \ldots, N\},$$

où $(x_{1,t}, x_{2,t}, x_{3,t})$ représenteront respectivement la longitude, la latitude et l'altitude à la station x_t.

Des statistiques sur les pluies extrêmes sont très souvent utilisées lorsqu'une inondation survient pour quantifier la rareté d'un événement de cette ampleur. Un problème typique est d'estimer la quantité d'eau qu'il pleuvra un jour exceptionnel de fortes précipitations, qui est attendu en moyenne toutes les T années. Les hydrologues ont pour habitude de s'intéresser à la valeur $T = 100$, correspondant à un événement centennal. Rappelons qu'en termes statistiques, le problème est d'estimer un niveau de retour à T années ce qui revient à estimer un quantile des pluies journalières de niveau $\beta = 1/(365.25 \times T)$.

L'étude statistique de précipitations extrêmes a intéressé de nombreux auteurs, on citera Coles et Tawn [35] et Cooley *et al.* [39] qui ont modélisé les précipitations par une loi de Pareto Généralisée GPD (voir Partie 1.3). Les précipitations extrêmes dans la

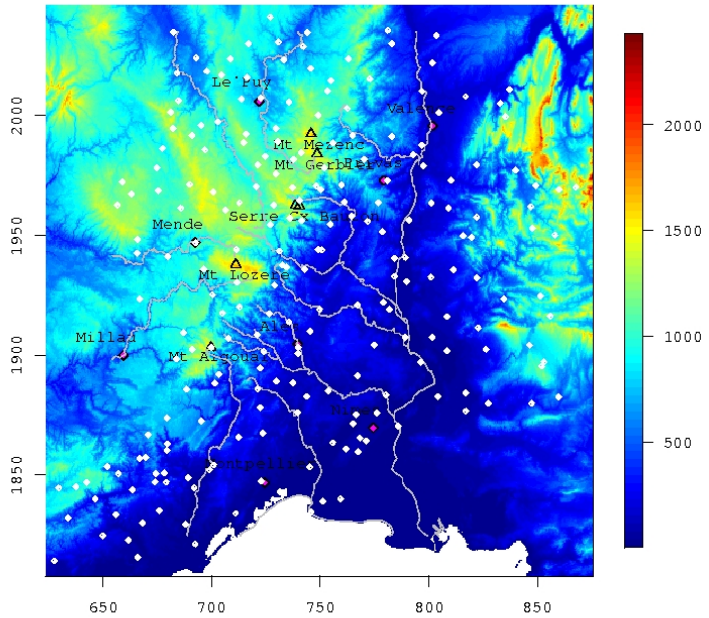
région Cévennes-Vivarais ont déjà fait l'objet de diverses études dont entre autres Bois *et al.* [22] et Gardes et Girard [85].

Les premiers contributeurs ont utilisé un jeu de données de pluies horaires mesurées entre les années 1948 et 1991 sur 48 stations. Ils se sont intéressés au niveau de retour à 10 ans des précipitations qu'ils ont estimé en ajustant une loi de Gumbel aux données et en utilisant une méthode de krigeage [127].

Les seconds contributeurs ont supposé que les pluies extrêmes pouvaient être modélisées par une loi à queue lourde. Partant d'un jeu de données de pluies horaires mesurées entre les années 1993 et 2000 sur 142 stations, ils ont estimé le niveau de retour à 10 ans des précipitations par la méthode des plus proches voisins. Les résultats obtenus permirent à ces auteurs de remettre en cause l'hypothèse selon laquelle la loi des précipitations dans la région Cévennes-Vivarais est une loi à queue légère.

Compte tenu de ce travail, nous avons alors supposé que la loi des pluies journalières dans cette région était à queue lourde. Le but de nos travaux est d'aller plus loin et d'estimer la moyenne des pluies se trouvant au-dessus du niveau de retour à T années, autrement dit la RCTE de niveau $\beta = 1/(365.25 \times T)$. Pour cela, on devra d'abord estimer $\gamma(x)$ pour pouvoir extrapoler (voir le résultat (3.10)) et ensuite calculer le niveau de retour à 100 ans.

Notre résultat sera l'estimation de mesures de risque pour des extrêmes pluviométriques en des sites où il n'y a pas de station.



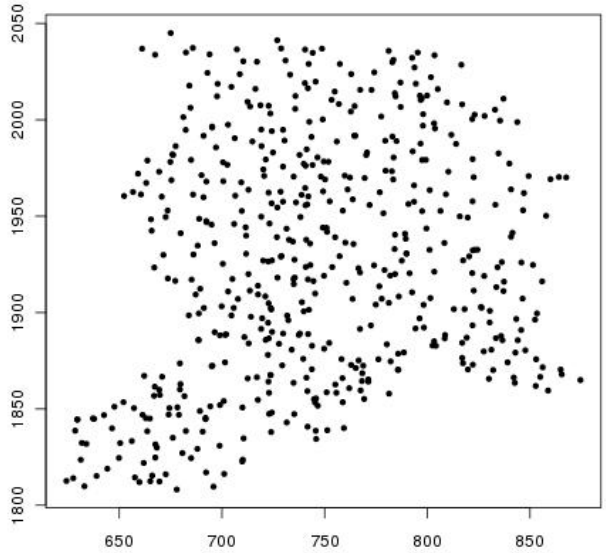


FIGURE 3.6: En haut : quelques stations d'observations (losanges blancs) représentées en fonction de leurs coordonnées géographiques. Horizontalement on a la longitude (en kilomètres), verticalement la latitude (en kilomètres) et en échelle de couleurs l'altitude (en mètres). Sur la carte les montagnes sont représentées par des triangles, les cours d'eaux par des lignes grises et les principales villes par des losanges roses. En bas : l'ensemble des 523 stations d'intérêt.

3.4.2 Choix des paramètres de réglage

Nos estimateurs de mesures de risque dépendent principalement des deux paramètres de réglage h et α qui sont difficiles à choisir dans la pratique.

Le choix de la fenêtre h, qui contrôle le degré de lissage, est un problème récurrent en statistique non-paramétrique. De même, rappelons qu'en théorie des valeurs extrêmes, le choix du nombre de statistique d'ordre ou de manière équivalente le choix de α, est un problème classique d'une grande importance puisqu'il fait apparaître un compromis entre biais et variance. Choisir une grande valeur de α donne un biais important (puisque l'on sort de la queue de la distribution), alors que choisir une petite valeur de α nous engendrera une grande variance (car on dispose de peu d'observations), voir à ce propos le Théorème 3.3.

On se propose de mettre en place une procédure de type validation croisée afin de choisir de manière simultanée h et α. Plus précisément, notre procédure sera du type "leave-one-out cross validation". Elle a pour but de permettre la mise en application de nos estimateurs de manière efficace. Dans ce but, on considère deux ensembles de valeurs.

Le premier :

$$\mathcal{A} = \{\alpha_1 \leq \cdots \leq \alpha_R\}$$

est tel que $\alpha_1 > 1/\min(n_j)$ pour $j \in \{1, \ldots, N\}$ de sorte qu'il y ait au moins une observation utilisée dans l'estimation et $\alpha_R < 0.1$ pour s'assurer que l'on reste dans la queue de distribution.

Le second ensemble de valeurs :

$$\mathcal{H} = \{h_1 \leq \cdots \leq h_M\},$$

est tel qu'il y ait au moins une observation dans la boule $B(x, h_1)$ pour tout x.

Notre procédure est basée sur l'estimation du paramètre $\gamma(x)$ pour plusieurs raisons. Ce paramètre contrôle la lourdeur de la queue de la loi (voir hypothèse **(F.1)**) ainsi que la lourdeur du RCTM (voir l'équation (3.7)) et enfin l'extrapolation (voir l'équation (3.10)). De plus, il existe de nombreux estimateurs de ce paramètre dans la littérature des valeurs extrêmes auxquels se comparer.

Le principe de notre procédure est de choisir le couple empirique $(h_{emp}, \alpha_{emp}) \in \mathcal{H} \times \mathcal{A}$ pour lequel deux estimations différentes de l'indice $\gamma(x_t)$ en chaque station t coïncident approximativement. Pour les hydrologues il est primordial de combiner des informations locales et régionales pour cela nous avons choisis les deux estimateurs suivants.

Le premier estimateur, noté $\hat{\gamma}_{n,t}$, est l'estimateur de Hill (voir Définition 1.18), il dépend seulement de α et il est uniquement basé sur les mesures de pluies obtenues en chaque station t.

Le second estimateur, noté $\hat{\gamma}_n(x_t)$ est l'estimateur de l'indice de queue conditionnel introduit par Daouia *et al.* [44]. Rappelons sa définition.

Définition 3.8. *Soit $(\alpha_n)_{n\geq 1}$ une suite positive telle que $\alpha_n \to 0$ et $nh^p\alpha_n \to \infty$ quand $n \to \infty$. Un estimateur à noyau de type Hill est donné pour tout $J > 1$ par :*

$$\hat{\gamma}_n(x) = \sum_{j=1}^{J}[\log \widehat{\mathrm{RVaR}}_n(\tau_j\alpha_n|x) - \log \widehat{\mathrm{RVaR}}_n(\tau_1\alpha_n|x)] \Big/ \sum_{j=1}^{J} \log(\tau_1/\tau_j) \ ,$$

où $(\tau_j)_{j\geq 1}$ est une suite de poids positive et décroissante.

Il dépend à la fois de α, h et de la suite $(\tau_j)_{j\geq 1}$. Il utilise toutes les mesures comprises dans la boule $B(x_t, h)$ à l'exception des mesures de la station courante t. On verra dans la sous-Partie suivante 3.4.3 les choix possibles pour la suite $(\tau_j)_{j\geq 1}$ en pratique.

En résumé, l'idée principale est de choisir le couple (h_{emp}, α_{emp}) pour lequel les estimations locales $\gamma(x_t)$ et celles prédites en utilisant les stations voisines $\hat{\gamma}_n(x_t)$ sont cohérentes. Pour être plus précis, l'algorithme mis en place est le suivant :

1. Début de la boucle sur toutes les paires $(h_i, \alpha_j) \in \mathcal{H} \times \mathcal{A}$ et sur toutes les stations $t \in \{1, \ldots, N\}$.
2. Appliquer l'estimateur de Hill à la station t pour un niveau α_j pour obtenir $\hat{\gamma}_{n,t,j}$.
3. Appliquer l'estimateur de l'indice de queue conditionnel à l'aide des mesures contenues dans la boule $B(x_t, h_i) \setminus \{x_t\}$ pour un niveau α_j pour obtenir $\hat{\gamma}_{n,i,j}(x_t)$.
4. Calculer la distance : $W_{h_i,\alpha_j}(x_t) = (\hat{\gamma}_{n,t,j} - \hat{\gamma}_{n,i,j}(x_t))^2$.
5. Fin de la boucle.
6. La paire optimale est donnée par :

$$(h_{emp}, \alpha_{emp}) = \underset{(h_i,\alpha_j)\in\mathcal{H}\times\mathcal{A}}{\arg\min} \ \text{mediane} \ \{W_{h_i,\alpha_j}(x_t) \ , t \in \{1, \ldots, N\}\}.$$

3.4.3 Validation sur simulations

Dans cette partie, on teste la procédure précédente sur deux lois à queues lourdes conditionnelles : une loi de Fréchet et une loi de Burr. La fonction de survie conditionnelle de la loi de Fréchet est donnée par :

$$\overline{F}(y|x) = 1 - \exp(-y^{-1/\gamma(x)}), \quad \text{pour} \quad y \geq 0.$$

La RCTE associée peut s'écrire :

$$\mathrm{RCTE}(\beta|x) = \frac{1}{\beta}\int_0^{\mathrm{RVaR}(\beta|x)^{-1/\gamma(x)}} t^{\gamma(x)}\exp(-t)dt,$$

avec

$$\mathrm{RVaR}(\beta|x) = (-\log(1-\beta))^{-\gamma(x)}.$$

La fonction de survie conditionnelle de la loi de Burr choisie est donnée par :

$$\overline{F}(y|x) = \left(1 + y^{1/\gamma(x)}\right)^{-1}, \quad \text{pour} \quad y \geq 0.$$

La RCTE associée peut s'écrire :

$$\mathrm{RCTE}(\beta|x) = \mathcal{I}\left(\frac{\mathrm{RVaR}(\beta|x)^{-1/\gamma(x)}}{1+\mathrm{RVaR}(\beta|x)^{-1/\gamma(x)}}, 1-\gamma, 1+\gamma\right),$$

avec

$$\mathrm{RVaR}(\beta|x) = \beta^{-\gamma(x)}(1-\beta)^{\gamma(x)},$$

et où

$$\mathcal{I}(r,p,q) = \frac{B(r,p,q)}{B(p,q)} \quad \text{avec} \quad B(r,p,q) = \int_0^r w^{p-1}(1-w)^{q-1}dw$$

étant la fonction Beta incomplète et $B(p,q)$ la fonction Beta.

Pour notre validation sur simulation, on a choisi l'indice de queue conditionnel suivant :

$$\gamma : x \in [0,1] \to \gamma(x) = \frac{1}{2}\left(\frac{1}{10} + \sin(\pi x)\right)\left(\frac{11}{10} - \frac{1}{2}\exp\left(-64(x-1/2)^2\right)\right), \tag{3.11}$$

dont on donne la représentation graphique dans la Figure 3.7 ci-après.

Si l'on garde en mémoire les restrictions sur $\gamma(x)$ dues à nos résultats asymptotiques (voir Théorème 3.3), ce choix de fonction rend l'estimation plus difficile. En effet, on peut remarquer que $\gamma(x)$ est proche de $1/2$ lorsque $x = 0.3$ or $x = 0.7$.

La fonction $\gamma(x)$ choisie étant définie sur l'intervalle unité [0,1] (voir équation (3.11)) il nous faut normaliser les vecteurs que l'on utilisera comme covariable. La normalisation d'un vecteur positif $v = (v_i)_{i\in\{1,\ldots,n\}} > 0$ en un vecteur $w = (w_i)_{i\in\{1,\ldots,n\}} \in [0,1]$, est obtenue à l'aide de la transformation :

$$w_i = \frac{v_i - \min(v_1,\ldots,v_n)}{\max(v_1,\ldots,v_n) - \min(v_1,\ldots,v_n)} \quad \text{pour} \quad i \in \{1,\ldots,n\}.$$

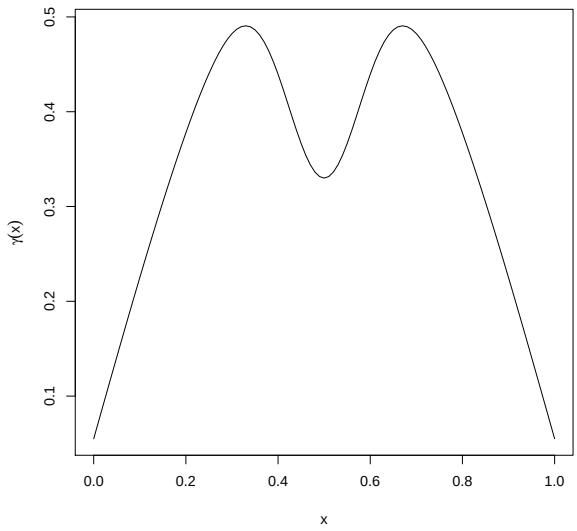


FIGURE 3.7: Représentation de la fonction $\gamma(x)$.

Soient z_1, z_2 et z_3 respectivement la latitude, la longitude et l'altitude normalisées dans l'intervalle unité [0,1]. Deux choix de covariables x ont été utilisés pour $\gamma(x)$:

$$x_{euc} := \sqrt{(z_1^2 + z_2^2)/2} \quad \text{et} \quad x_{alt} := z_3.$$

Les paramètres de réglage sont choisis dans les ensembles :

$$\mathcal{A} = \{1/(6\times 365.25), 1/(5\times 365.25), \ldots, 1/365.25, 4.10^{-3}, 6.10^{-3}, \ldots, 10^{-2}, 2.10^{-2}, \ldots, 10^{-1}\}$$

et

$$\mathcal{H} = \{14, 15, \ldots, 30\}.$$

Sous certaines hypothèses, les auteurs Daouia *et al.* [44] (Corollaire 2) ont obtenu la normalité asymptotique de $\hat{\gamma}_n(x)$ (voir Définition 3.8). La variance asymptotique de l'estimateur vaut :

$$\frac{\gamma^2(x)\|K\|_2^2}{nh^p\alpha_n g(x)} V_j, \tag{3.12}$$

où

$$V_J = \left(\sum_{j=1}^{J} \frac{2(J-j)+1}{\tau_j} - J^2\right) \Bigg/ \left(\sum_{j=1}^{J} \log(\tau_1/\tau_j)\right)^2.$$

On peut remarquer que la variance asymptotique de l'estimateur $\hat{\gamma}_n(x)$ est, à un facteur d'échelle $V_j\|K\|_2^2/g(x)$ près, identique à celle de l'estimateur classique de Hill dans le cas non conditionnel (voir Théorème 1.20). Ce facteur d'échelle fait intervenir les suites de poids $(\tau_j)_{j\geq 1}$. Ainsi, un mauvais choix de ces suites aura forcément une influence sur la

volatilité de l'estimateur.

Comme exemple de suites de poids nous proposons d'utiliser :

- la suite de poids harmonique définie pour tout $j = 1, \ldots, J$ par $\tau_j^{Ha} = 1/j$. Dans ce cas, $V_J^{Ha} = J(J-1)(2J-1)/(6\log^2(J!))$ est une fonction convexe de J. Son minimum qui est atteint en $J_{opt}^{Ha} = 9$ vaut $V_9^{Ha} \simeq 1.245$.
- la suite de poids géométrique définie par $\tau_j^G = (1/j)^{(j/J)}$ pour tout $j = 1, \ldots, J$. La fonction V_J^G est convexe et son minimum est atteint pour $J_{opt}^G = 15$ et $V_{15}^G \simeq 1.117$.

Dans les deux cas, le nombre de termes J a été choisi afin de minimiser la variance asymptotique de $\hat{\gamma}_n(x)$.

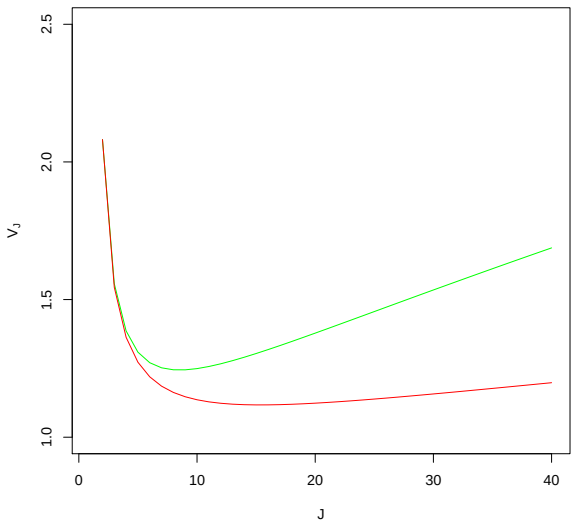


FIGURE 3.8: Représentation de V_J^{Ha} et V_J^G en fonction de J. En ordonnée on a la fonction V_j et en abscisse J. On remarquera que V_J^G croit moins vite que V_J^{Ha}.

Enfin, on a choisi pour noyau, un noyau bi-quadratique :

$$K(x) := K(z_1, z_2) = \frac{15}{16}\left[1 - (z_1^2 + z_2^2)\right]^2 \mathbb{I}\{z_1^2 + z_2^2 \leq 1\},$$

dont on donne la représentation graphique dans la Figure 3.9 ci-après.

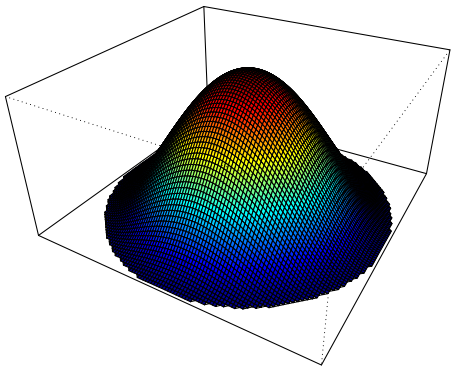


FIGURE 3.9: Représentation du noyau $K(z_1, z_2)$.

Pour s'assurer des performances de notre procédure, on la compare à une stratégie Oracle (choix optimal) (h_{opt}, α_{opt}) qui est basée sur la connaissance de la vraie fonction indice de queue conditionnel.

$$(h_{opt}, \alpha_{opt}) = \underset{(h_i,\alpha_j)\in\mathcal{H}\times\mathcal{A}}{\arg\min} \text{ mediane}\{V_{h_i,\alpha_j}(x_t)\ , t \in \{1, \ldots, N\}\},$$

où

$$V_{h_i,\alpha_j}(x_t) = (\gamma(x_t) - \hat{\gamma}_{n,i,j}(x_t))^2$$

est la distance à la vraie fonction indice de queue conditionnel. Les paramètres choisis par la procédure sont donnés dans le Tableau 3.2 ci-dessous.

Loi de Burr		**Loi de Fréchet**	
x_{euc} et τ_j^{Ha}	x_{alt} et τ_j^{G}	x_{euc} et τ_j^{G}	x_{alt} et τ_j^{Ha}
$h_{opt} = 22$	$h_{opt} = 24$	$h_{opt} = 22$	$h_{opt} = 26$
$h_{emp} = 24$	$h_{emp} = 22$	$h_{emp} = 20$	$h_{emp} = 24$
$\alpha_{opt} = 1/365.25$	$\alpha_{opt} = 1/(365.25 \times 2)$	$\alpha_{opt} = 1/(365.25 \times 3)$	$\alpha_{opt} = 1/365.25$
$\alpha_{emp} = 0.001$	$\alpha_{emp} = 1/365.25$	$\alpha_{emp} = 1/365.25$	$\alpha_{emp} = 0.004$

TABLE 3.2: Résultats de la procédure de sélection.

On peut voir que notre procédure de validation croisée choisit approximativement les mêmes paramètres de réglage que la stratégie Oracle pour tous les choix de lois, de covariable et de poids. Une remarque encore plus importante à faire est que l'on peut observer sur les Figures 3.10 et 3.11 que les lois des erreurs sur l'indice de queue sont

très proches. Ces résultats nous indiquent que notre procédure de type validation croisée est presque aussi performante que la stratégie Oracle, qui elle connait la vraie valeur de l'indice de queue.

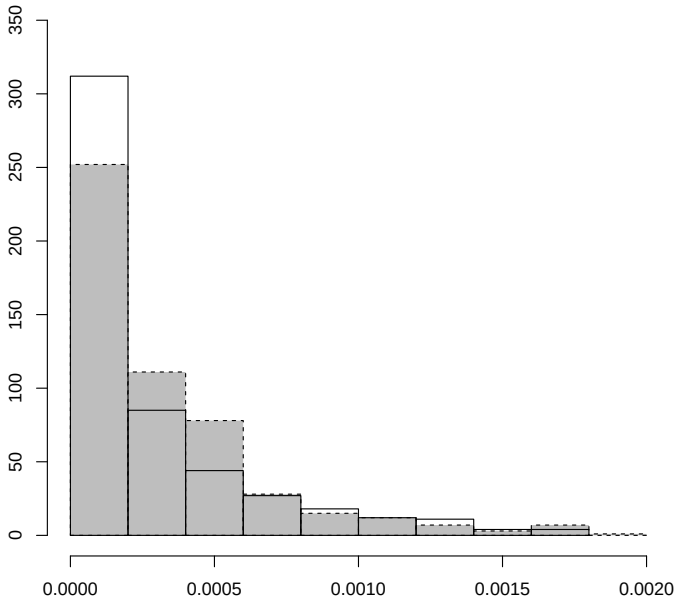
Des résultats similaires peuvent être observés sur la RCTE extrapolée définie par :

$$\widehat{\mathrm{RCTE}}_n^W(\beta|x) = \widehat{\mathrm{RCTE}}_n(\alpha|x)\left(\frac{\alpha}{\beta}\right)^{\hat{\gamma}_n(x)},$$

que l'on a calculé pour un niveau $\beta = 1/(365.25 \times 100)$ correspondant à une pluie centennale. Dans ce cas, la qualité de l'estimation est assurée grâce à l'erreur relative :

$$Q_n(x) = \left(\frac{\widehat{\mathrm{RCTE}}_n^W(\beta|x)}{\mathrm{RCTE}(\beta|x)} - 1\right)^2.$$

Les deux histogrammes de $Q_n(x_t)$, $t \in \{1, \ldots, N\}$ obtenus avec (h_{emp}, α_{emp}) et (h_{opt}, α_{opt}) sont représentés sur les Figures 3.12 et 3.13. Les deux ensembles de paramètres donnent approximativement les mêmes lois d'erreurs.



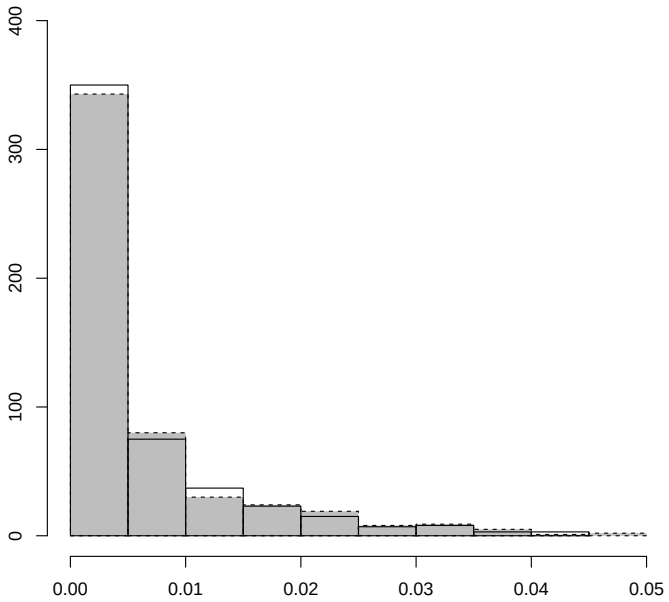


FIGURE 3.10: Histogrammes des erreurs $V_{h_{opt},\alpha_{opt}}(x_t)$ (traits pleins, barres blanches) et $V_{h_{emp},\alpha_{emp}}(x_t)$ (traits en pointillés, barres grises) calculés sur des données simulées pour une loi de Burr. En haut : x_{euc} et τ_j^{Ha}. En bas : x_{alt} and τ_j^{G}.

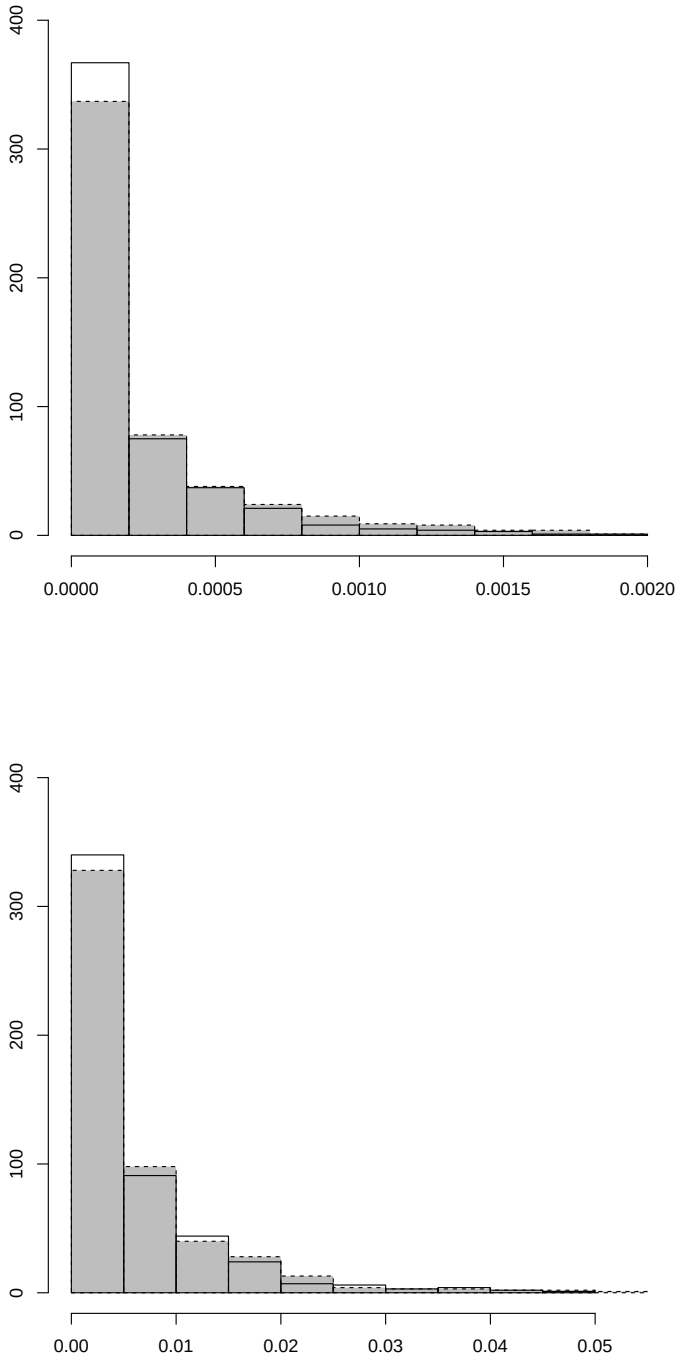


FIGURE 3.11: Histogrammes des erreurs $V_{h_{opt},\alpha_{opt}}(x_t)$ (traits pleins, barres blanches) et $V_{h_{emp},\alpha_{emp}}(x_t)$ (traits en pointillés, barres grises) calculés sur des données simulées pour une loi de Fréchet. En haut : x_{euc} et τ_j^G. En bas : x_{alt} et τ_j^{Ha}.

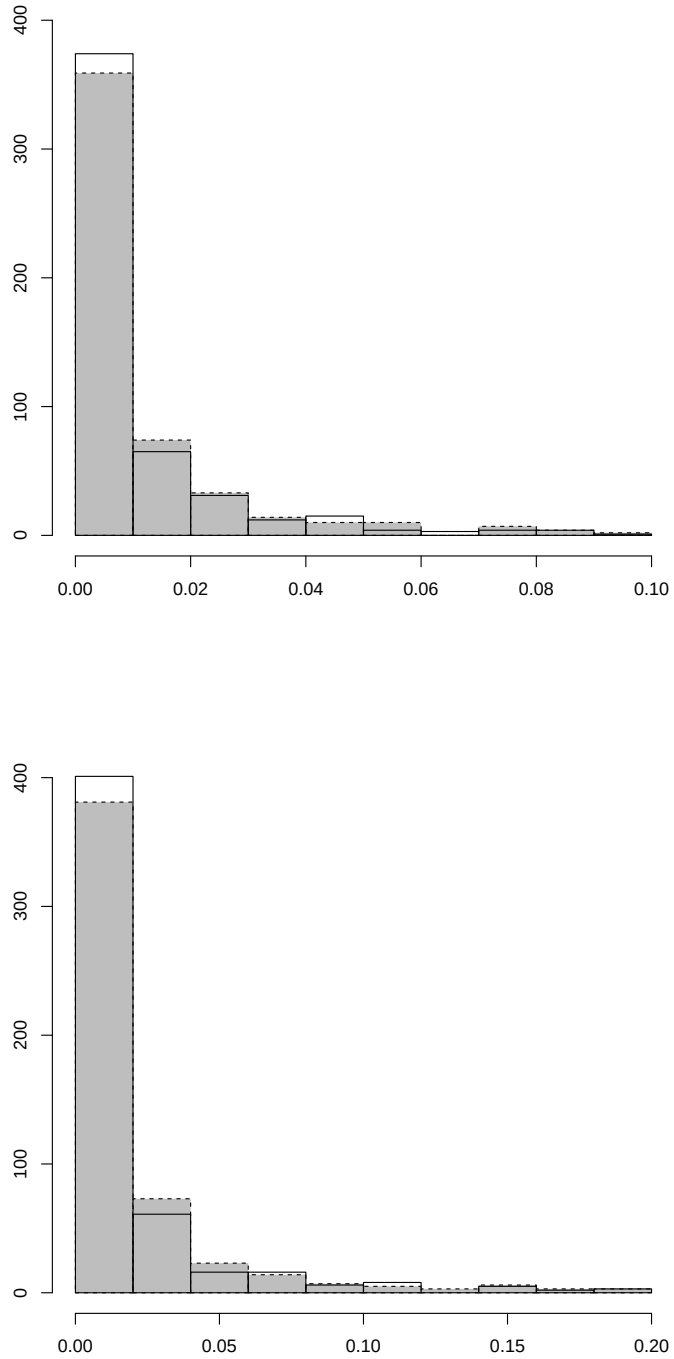


FIGURE 3.12: Histogrammes des erreurs obtenus sur les RCTE extrapolées sur des données simulées pour une loi de Burr. Jeu de paramètre (h_{opt}, α_{opt}) : traits pleins et barres blanches, (h_{emp}, α_{emp}) : traits en pointillés et barres grises. En haut : x_{euc} et τ_j^{Ha}. En bas x_{alt} et τ_j^{G}.

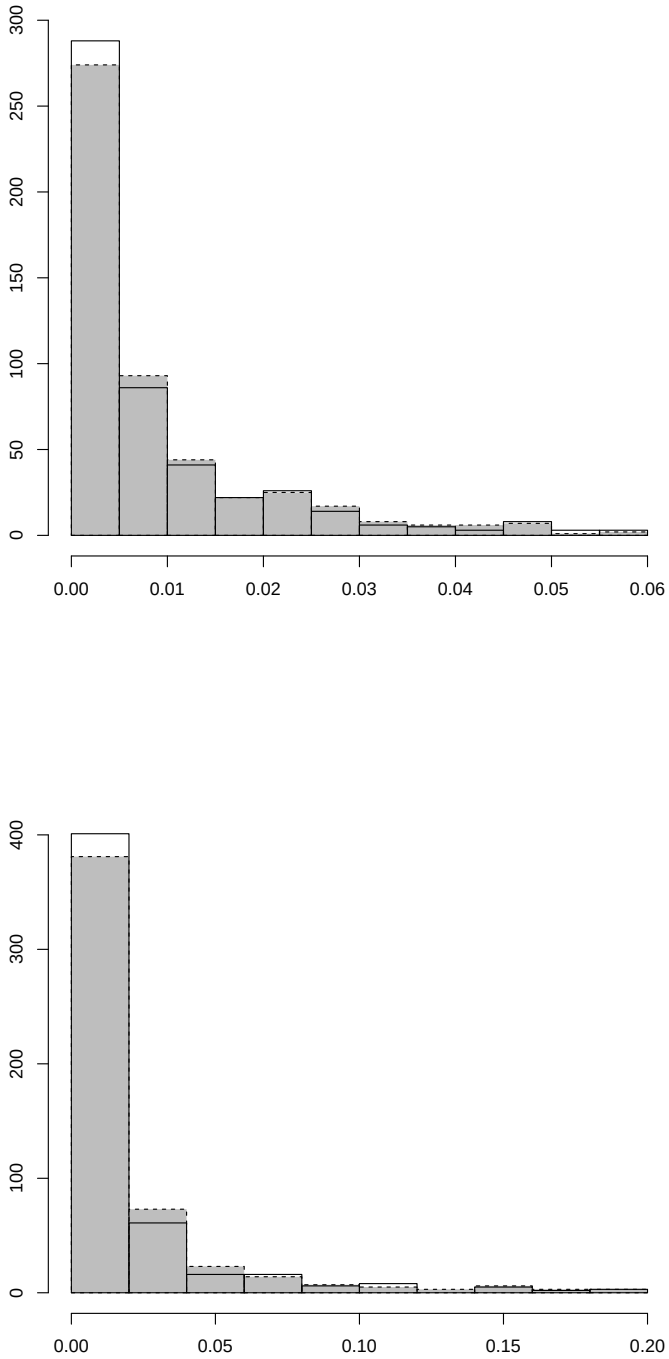


FIGURE 3.13: Histogrammes des erreurs obtenus sur les RCTE extrapolées sur des données simulées pour une loi de Fréchet. Jeu de paramètre (h_{opt}, α_{opt}) : traits pleins et barres blanches, (h_{emp}, α_{emp}) : traits en pointillés et barres grises. En haut : x_{euc} et τ_j^G. En bas x_{alt} et τ_j^{Ha}.

3.4.4 Mesures de risque estimées sur des extrêmes pluviométriques

L'estimation de l'indice de queue conditionel est effectué sur une grille régulière couvrant la région géographique d'intérêt. Cette grille consiste en 200×200 sites, voir Figure 3.14. Dans notre cas on se limite à une covariable X à p=2 dimensions qui dépend de la longitude et de la latitude. La distance d utilisée sera la distance euclidienne.

La procédure par validation croisée appliquée à notre jeu de données réelles avec τ_j^{Ha} donne comme valeurs pour nos paramètres de réglage $h_{emp} = 24$ et $\alpha_{emp} = 1/(365.25 \times 3)$. $h_{emp} = 24$ km correspond au rayon de la boule délimitant les stations utilisées pour lisser. $\alpha_{emp} = 1/(365.25 \times 3)$ correspond à un niveau de retour à 3 ans. Ces deux paramètres obtenus, on peut alors obtenir une estimation de $\gamma(x)$.

Pour illustrer notre estimation et donc la manière avec laquelle on interpole par noyau (voir Figure 3.14) on procède comme suit :

- **Etape 1 :** On se place au point g de la grille qui sera le centre de la boule.
- **Etape 2 :** A l'aide de notre procédure (voir sous-Partie 3.4.2) on fixe le rayon h de la boule de centre g que l'on notera $B(g, h)$.
- **Etape 3 :** On ne sélectionne que les stations x_t pour lesquelles $x_t \in B(g, h)$.

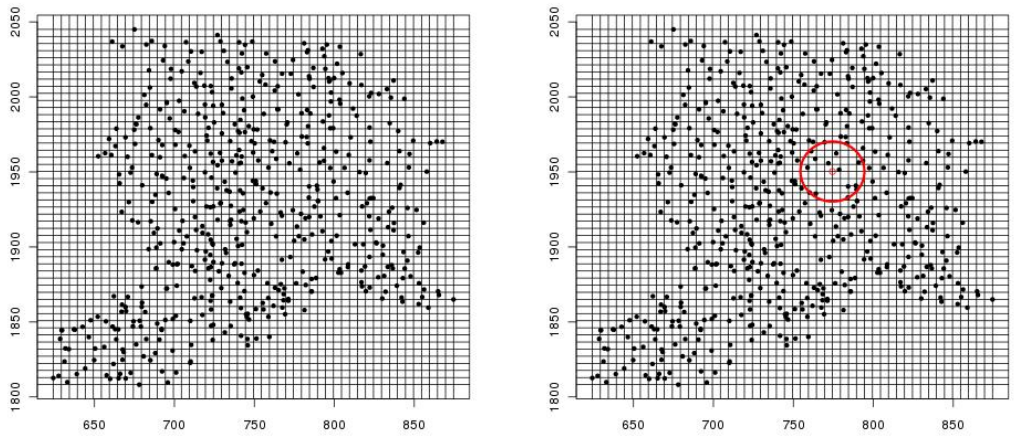


FIGURE 3.14: A gauche : les 523 stations et la grille régulière couvrant la région géographique d'intérêt. A droite : la boule $B(g, h)$ où l'on effectue l'interpolation.

On obtient ainsi une estimation de $\gamma(x)$ sur toute la grille pour les valeurs de nos paramètres dont la représentation est donnée Figure 3.15.

On souhaite vérifier si l'indice de queue conditionnel peut être considéré comme constant sur la région étudiée. A l'aide de la loi asymptotique de $\hat{\gamma}_n(x)$ établie dans [44], Corollaire 2, on peut construire des intervalles de confiance en chaque point de la grille. En effet, à l'aide de la variance asymptotique (voir équation (3.12)) où l'on estime γ par

$\hat{\gamma}_n(x)$, la densité $g(x)$ par $\hat{g}_n(x)$ (voir expression (3.14)) et où l'on remplace V_j par $V_9^{Ha} \simeq 1.245$ (car on a utilisé la suite τ_j^{Ha}) on obtient, l'intervalle de niveau de confiance IC à $(1-\eta)\%$ en chaque point de la grille :

$$\mathrm{IC}_{1-\eta}(x) = \hat{\gamma}_n(x)\left(1 \pm \|K\|_2 \Phi^{-1}\left(1-\frac{\eta}{2}\right)\sqrt{\frac{V_9^{Ha}}{nh^2\alpha_n\hat{g}_n(x)}}\right),$$

où Φ est la fonction de répartition de la loi Normale centrée réduite.

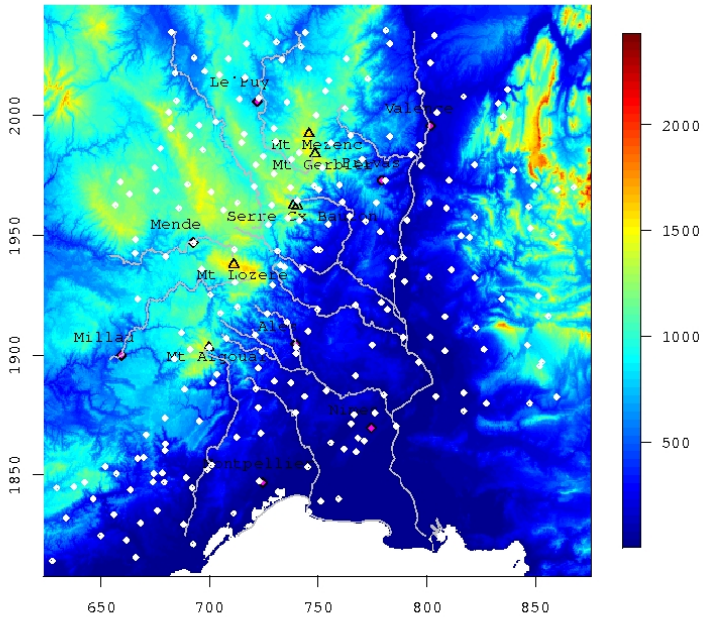
Pour un niveau de confiance de 95%, il s'est avéré qu'il n'y a aucune valeur commune aux intervalles, on voit donc qu'il est raisonnable de ne pas supposer $\gamma(x)$ constant sur toute la grille. On remarque que sur toute la grille, l'estimation de $\gamma(x)$ est significativement inférieure à $1/2$; ainsi on peut utiliser nos estimateurs de mesures de risque.

On peut maintenant passer à l'estimation de mesures de risque associées à une période de retour de 100 ans. On s'intéresse plus particulièrement à la $\mathrm{RVaR}_n(\beta|x)$ et à la $\mathrm{RCTE}_n(\beta|x)$ où $\beta = 1/(365.25 \times 100)$.

Les résultats obtenus à l'aide de l'estimateur associé $\widehat{\mathrm{RVaR}}_n^W(\beta|x)$ sont donnés Figure 3.16. Il apparaît que le niveau de retour augmente globalement avec l'altitude. La dérive du taux de précipitations en fonction de l'altitude est en adéquation avec les statistiques descriptives des précipitations dans la région Cévennes-Vivarais réalisées par Molinié *et al.* [134]. Le niveau de retour à 100 ans estimé $\widehat{\mathrm{RVaR}}_n^W(\beta|x)$ est similaire aux résultats obtenus par Ceresetti *et al.* [31] où les auteurs utilisent des méthodes de krigeage [127].

On peut maintenant calculer la RCTE correspondante au niveau de retour à 100 ans à l'aide de l'estimateur $\widehat{\mathrm{RCTE}}_n^W(\beta|x)$. Elle est un complément au niveau de retour permettant de mieux mesurer la lourdeur de la queue de distribution et ainsi repérer les zones où les pluies sont plus importantes. Elle correspond à la moyenne des pluies se situant au-dessus du niveau de retour à 100 ans.

Tout comme pour le niveau de retour, il apparaît que la RCTE augmente avec l'altitude, voir Figure 3.17. Il est intéressant de remarquer que la $\widehat{\mathrm{RCTE}}_n^W(\beta|x)$ peut être jusqu'à 150 millimètres plus haute que $\widehat{\mathrm{RVaR}}_n^W(\beta|x)$ dans les régions montagneuses. En effet, les amplitudes des précipitations sont très différentes pour les deux mesures de risque. Pour le niveau de retour on a un maximum qui se situe aux alentours de 350 mm alors que pour la RCTE correspondante il est proche de 500 mm. On peut aussi remarquer que les endroits à risque ne sont pas exactement les mêmes. Tous les résultats ont été obtenus sous hypothèse d'indépendance et de stationnarité temporelle.



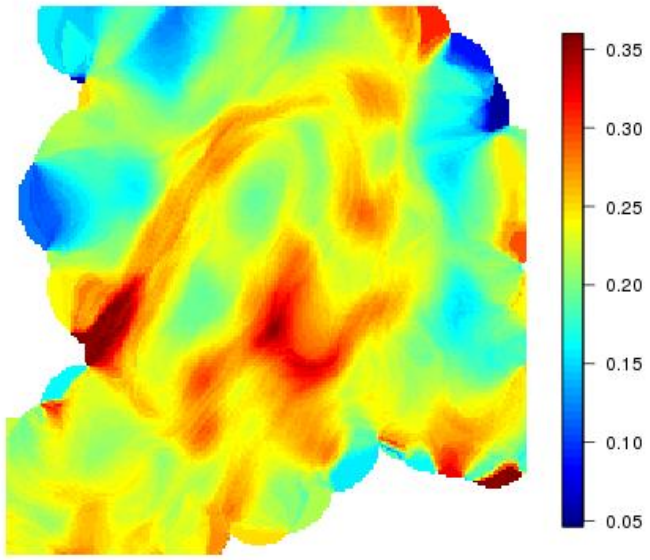
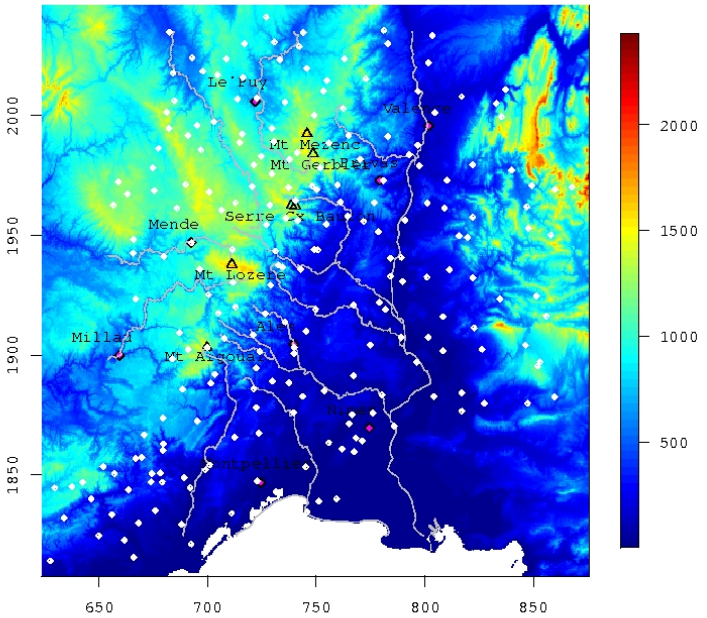


FIGURE 3.15: En haut : carte de la région géographique d'intérêt. En bas : $\hat{\gamma}_n(x)$.



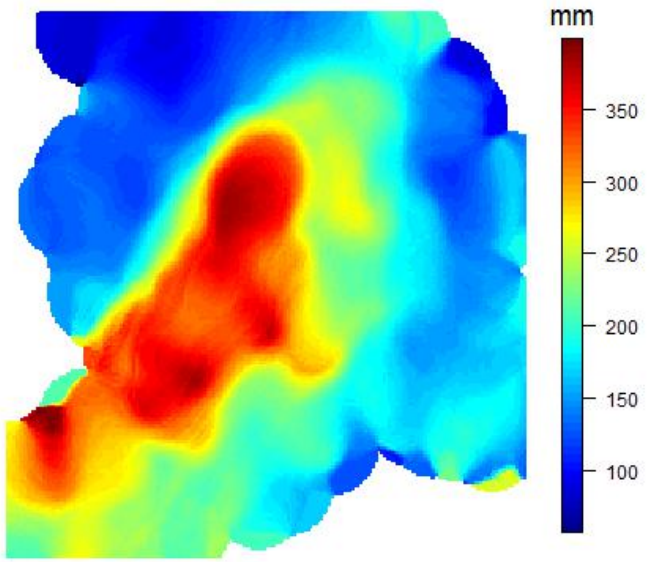
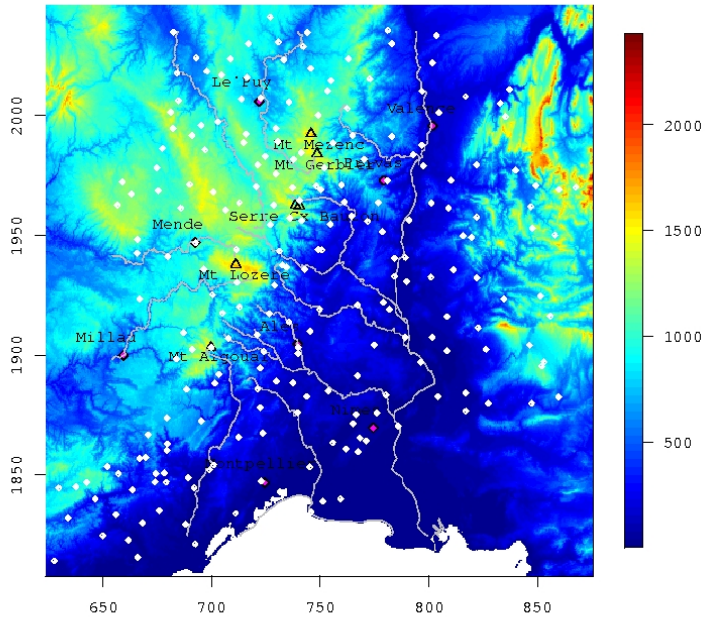


FIGURE 3.16: En haut : carte de la région géographique d'intérêt. En bas : $\widehat{\mathrm{RVaR}}_n^W(1/(100\times 365.25)|x)$. L'échelle de couleur donne la hauteur d'eau en millimètres.



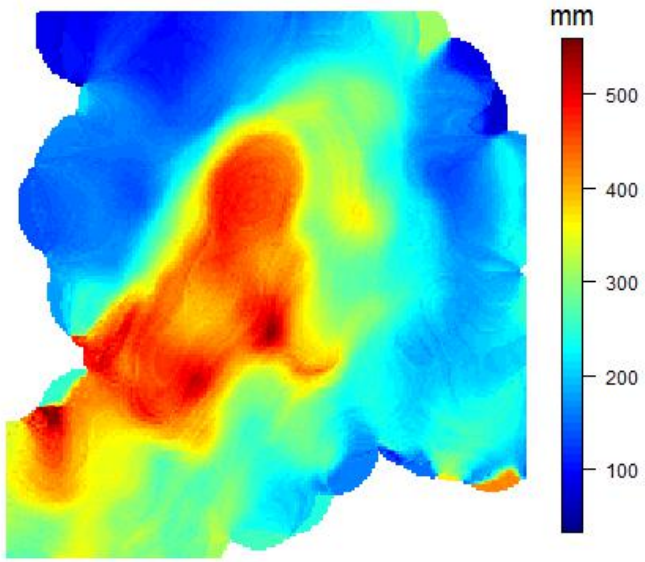


FIGURE 3.17: En haut : carte de la région géographique d'intérêt. En bas : $\widehat{\mathrm{RCTE}}_n^W(1/(100 \times 365.25)|x)$. L'échelle de couleur donne la hauteur d'eau en millimètres.

3.5 Perspectives

Ce travail nous a permis de faire le lien entre les mesures de risque utilisées en finance et en actuariat et le risque en hydrologie. Nous avons ainsi introduit une nouvelle méthode statistique permettant d'estimer des mesures de risque extrêmes en des sites où l'on ne dispose pas de stations de mesure. Plus particulièrement, on a pu présenter une application à la région Cévennes-Vivarais en comparant l'estimation du niveau de retour à 100 ans et de la RCTE correspondante.

Ces mesures de risque sont un nouvel outil dans la prévention des risques en hydrologie. Il serait très intéressant de prendre d'autres covariables en compte. Par exemple on pourrait tenir compte de l'altitude, l'orientation par rapport à la mer ou encore le type de temps. Pour cela il nous faudrait définir des nouvelles distances. Il serait aussi intéressant d'appliquer notre méthode à un nouveau type de données et de pouvoir estimer des mesures de risque dans d'autres domaines d'applications, par exemple en actuariat ou en finance.

Parmi les différentes suites à donner à ce travail, il serait intéressant d'essayer de lever la restriction sur $\gamma(x)$ dans nos théorèmes (voir Partie 3.3). Necir *et al.* [136] ont estimé la CTE pour des lois à queues lourdes (sans la présence d'une covariable) et ont obtenu la restriction $1/2 < \gamma < 1$. On pourrait essayer d'adapter leurs outils pour pallier à nos limitations sur $\gamma(x)$.

Afin d'étendre nos résultats aux autres domaines d'attraction (Gumbel et Weibull), on pourrait s'inspirer des travaux de Daouia *et al.* [43]. Ces derniers ont estimé les quantiles conditionnels extrêmes en présence d'une covariable dans tous les domaines d'attraction en se basant sur des conditions dues à Von Mises (voir [50]) sur la fonction de survie conditionnelle. On pourrait alors estimer pour des pertes extrêmes et en présence d'une covariable le RCTM et ainsi toutes les mesures de risque classiques présentées dans ce chapitre.

Lorsque la covariable est non aléatoire de dimension infinie, Gardes *et al.* [89] ont proposé un estimateur du quantile conditionnel extrême pour des lois à queues lourdes. En se basant sur ces résultats, on pourrait envisager d'étendre nos travaux à ce cas et ainsi estimer toutes les mesures de risque classiques dans le cas d'une covariable non aléatoire de dimension infinie.

Une autre piste très intéressante à explorer serait d'estimer $\gamma(x)$ à l'aide de combinaisons de différents ordres du RCTM. En effet, à l'aide du Lemme 3.9 on peut facilement obtenir un équivalent du RCTM et ainsi à l'instar des travaux de Girard *et al.* [94, 95, 96], on pourrait estimer $\gamma(x)$.

3.6 Démonstrations

3.6.1 Résultats préliminaires et leurs preuves

Ce lemme nous donne un équivalent de $\varphi_a(y|x)$ quand $y \to \infty$. On se réfère à Chen *et al.* [33, Corollaire 3.2] et à Pan *et al.* [138, Proposition 4.1] pour un résultat similaire dans un cas sans covariable.

Lemme 3.9. *Sous l'hypothèse* **(F.1)**, *si* $y \to \infty$, *alors pour* $a \in [0, 1/\gamma(x))$,

$$\varphi_a(y|x) = \frac{1}{1 - a\gamma(x)} y^a \overline{F}(y|x)(1 + o(1)).$$

De plus, si l'on ajoute la condition **(F.2)**, *la dérivée* $\varphi_a'(.|x)$ *de la fonction* $\varphi_a(.|x)$ *existe et est à variations régulières telle que :*

$$\varphi_a'(y|x) = \frac{a\gamma(x) - 1}{\gamma(x)} \frac{\varphi_a(y|x)}{y}(1 + o(1)).$$

Preuve – Premièrement, une intégration par partie nous donne :

$$a \int_y^\infty z^{a-1} \overline{F}(z|x) dz = \varphi_a(y|x) - y^a \overline{F}(y|x). \tag{3.13}$$

En utilisant le résultat de Resnick [147, équation (0.32)] avec :

$$y \to y^{a-1} \overline{F}(y|x) \in \mathcal{RV}_{a-1/\gamma(x)-1} \quad \text{où} \quad a - 1/\gamma(x) - 1 < -1$$

et $y \to \infty$ on obtient :

$$\int_y^\infty z^{a-1} \overline{F}(z|x) dz = \frac{\gamma(x)}{1 - a\gamma(x)} y^a \overline{F}(y|x)(1 + o(1)).$$

En remplaçant dans l'équation (3.13) et en divisant des deux cotés par :

$$\frac{1}{1 - a\gamma(x)} y^a \overline{F}(y|x)$$

on a :

$$\frac{\varphi_a(y|x)}{\frac{1}{1-a\gamma(x)} y^a \overline{F}(y|x)} - 1 + a\gamma(x) = a\gamma(x)(1 + o(1)),$$

ce qui conclut la première partie de la preuve. Deuxièmement, sous la condition **(F.2)**, en dérivant des deux cotés l'équation (3.13) on obtient :

$$\varphi_a'(y|x) = y^a \overline{F}'(y|x) = y^{a-1} \overline{F}(y|x) \frac{y\overline{F}'(y|x)}{\overline{F}(y|x)},$$

et en utilisant le corollaire du Théorème 0.6 dans [147], il s'ensuit que :

$$\frac{y\overline{F}'(y|x)}{\overline{F}(y|x)} = -\frac{1}{\gamma(x)}(1+o(1)),$$

ce qui conclut la preuve. ■

Comme conséquence du Lemme 3.9 et de l'équation (3.1), on obtient que

$$\mathrm{RCTE}(\alpha_n|x)/\mathrm{RVaR}(\alpha_n|x) \sim 1/(1-\gamma(x))$$

qui est une extension du résultat sans covariable de Hua et Joe [121] et de Zhu et Li [170]. Le second lemme est aussi d'une nature analytique. Il nous donne un développement au second ordre du RCTM.

Lemme 3.10. *Supposons que les hypothèses* **(F.1)**, **(F.2)** *et* **(F.3)** *soient vérifiées. Soient* $0 < \beta_n < \alpha_n$ *deux suites telles que* $\alpha_n \to 0$ *lorsque* $n \to \infty$. *Alors,*

$$|\log \mathrm{RCTM}_a(\alpha_n|x) - \log \mathrm{RCTM}_a(\beta_n|x) + a\gamma(x)\log(\alpha_n/\beta_n)| = O\left(\log(\alpha_n/\beta_n)\varepsilon_a(1/\alpha_n)\right).$$

Preuve – En se servant de (3.7) et de l'hypothèse **(F.2)**, on a :

$$\log \mathrm{RCTM}_a(\alpha_n|x) = -a\gamma(x)\log(\alpha_n) + \log(c_a(x)) + \int_1^{1/\alpha_n} \frac{\varepsilon_a(u|x)}{u}du,$$

par conséquent :

$$\Delta_n := \log \mathrm{RCTM}_a(\alpha_n|x) - \log \mathrm{RCTM}_a(\beta_n|x) + a\gamma(x)\log(\alpha_n/\beta_n) = \int_{1/\beta_n}^{1/\alpha_n} \frac{\varepsilon_a(u|x)}{u}du.$$

A l'aide de **(F.3)**, on obtient $|\Delta_n| \leq |\varepsilon_a(1/\alpha_n)|\log(\alpha_n/\beta_n)$, la conclusion en découle. ■

Remarquons que l'estimateur à noyau (3.3) du moment conditionnel peut être réécrit sous la forme :

$$\hat{\varphi}_{a,n}(y_n|x) = \hat{\psi}_{a,n}(y_n|x)/\hat{g}_n(x)$$

où

$$\hat{\psi}_{a,n}(y|x) = \frac{1}{n}\sum_{i=1}^{n} K_h(x-X_i)Y_i^a\mathbb{I}\{Y_i > y\},$$

est un estimateur de $\psi_a(y|x) = g(x)\varphi_a(y|x)$ et $\hat{g}_n(x)$ est l'estimateur à noyau classique de la densité $g(x)$:

$$\hat{g}_n(x) = \frac{1}{n}\sum_{i=1}^{n} K_h(x-X_i). \tag{3.14}$$

Lemme 3.11. *Supposons que les hypothèses* **(F.1)**, **(F.2)**, **(L)** *et* **(K)** *soient vérifiées. Soit* $x \in \mathbb{R}^p$ *tel que* $g(x) > 0$ *et soit* $y_n \to \infty$ *tel que* $nh^p\overline{F}(y_n|x) \to \infty$.

(i) *Soit* $0 \leq a < 1/\gamma(x)$. *Si* $\omega(x, y_n, a, 0, h) \to 0$ *alors*

$$\mathbb{E}(\hat{\psi}_{a,n}(y_n|x)) = \psi_a(y_n|x)(1 + O(h) + O(\omega(x, y_n, a, 0, h))).$$

(ii) *Soient* $0 \leq a_1 < \cdots < a_{J+1} < 1/(2\gamma(x))$ *où* J *est un entier positif. Considérons alors des suites* $(y_{n,j})$, $j = 1, \ldots, J+1$ *telles que :*

$$\max_{j \in \{1,\ldots,J+1\}} \left\{ \left| \frac{y_{n,j}}{y_n} - 1 \right| \right\} \to 0.$$

Si il existe, $\xi > 0$ *tel que* $\max_{a \in \mathcal{L}(\{a_1,\ldots,a_{J+1}\})} \omega(x, y_n, a, \xi, h) \to 0$ *alors, le vecteur aléatoire :*

$$\left\{ \sqrt{nh^p \overline{F}(y_n|x)} \left(\frac{\hat{\psi}_{a_j,n}(y_{n,j}|x) - \mathbb{E}(\hat{\psi}_{a_j,n}(y_{n,j}|x))}{\psi_{a_j}(y_{n,j}|x)} \right) \right\}_{j \in \{1,\ldots,J+1\}}$$

est asymptotiquement Gaussien, centré, avec une matrice de variance-covariance $\|K\|_2^2 \Sigma^{(1)}(x)/g(x)$ *où*

$$\Sigma_{i,j}^{(1)}(x) = \frac{(1 - a_i\gamma(x))(1 - a_j\gamma(x))}{1 - (a_i + a_j)\gamma(x)}, \ (i,j) \in \{1, \ldots, J+1\}^2.$$

Preuve – **(i)** Puisque les copies (X_i, Y_i), $i = 1, \ldots, n$ sont identiquement distribuées, il s'ensuit que :

$$\mathbb{E}(\hat{\psi}_{a,n}(y_n|x)) = \int_{\mathbb{R}^p} K_h(x-t)\varphi_a(y_n|t)g(t)dt = \int_S K(u)\varphi_a(y_n|x-hu)g(x-hu)du,$$

sous l'hypothèse **(K)**. Considérons :

$$|\mathbb{E}(\hat{\psi}_{a,n}(y_n|x)) - \psi_a(y_n|x)| \leq \varphi_a(y_n|x) \int_S K(u)|g(x-hu) - g(x)|du \tag{3.15}$$

$$+ \varphi_a(y_n|x) \int_S K(u) \left| \frac{\varphi_a(y_n|x-hu)}{\varphi_a(y_n|x)} - 1 \right| g(x-hu)du. \tag{3.16}$$

Sous l'hypothèse **(L)**, et puisque $g(x) > 0$, on a :

$$(3.15) \leq \varphi_a(y_n|x)c_g h \int_S d(u,0)K(u)du = \varphi_a(y_n|x)O(h). \tag{3.17}$$

Par ailleurs, au vu de l'équation (3.17) :

$$\begin{aligned} (3.16) &\leq \varphi_a(y_n|x)\omega(x, y_n, a, 0, h) \int_S K(u)g(x-hu)du \\ &\leq \varphi_a(y_n|x)g(x)\omega(x, y_n, a, 0, h)(1 + o(1)) \\ &\leq \psi_a(y_n|x)\omega(x, y_n, a, 0, h)(1 + o(1)). \end{aligned} \tag{3.18}$$

Combiner les équations (3.17) et (3.18) permet de conclure la première partie de la preuve.

(ii) Soient $\beta \neq 0$ un vecteur de $\mathbb{R}^{J+1}$ et $\Lambda_n(x) = (nh^p\psi_0(y_n|x))^{-1/2}$. Considérons alors la variable aléatoire :

$$\begin{aligned}
\Psi_n &= \sum_{j=1}^{J+1} \beta_j \left(\frac{\hat{\psi}_{a_j,n}(y_{n,j}|x) - \mathbb{E}(\hat{\psi}_{a_j,n}(y_{n,j}|x))}{\Lambda_n(x)\psi_{a_j}(y_{n,j}|x)} \right), \\
&= \sum_{i=1}^{n} \frac{1}{n\Lambda_n(x)} \left\{ \sum_{j=1}^{J+1} \frac{\beta_j K_h(x-X_i) Y_i^{a_j} \mathbb{I}\{Y_i \geq y_{n,j}\}}{\psi_{a_j}(y_{n,j}|x)} \right. \\
&- \left. \mathbb{E}\left(\sum_{j=1}^{J+1} \frac{\beta_j K_h(x-X) Y^{a_j} \mathbb{I}\{Y \geq y_{n,j}\}}{\psi_{a_j}(y_{n,j}|x)} \right) \right\}, \\
&=: \sum_{i=1}^{n} Z_{i,n}.
\end{aligned}$$

Il est clair que $\{Z_{i,n},\ i = 1, \ldots, n\}$ est un ensemble de variables aléatoires centrées, indépendantes et identiquement distribuées de variance :

$$\mathrm{var}(Z_{1,n}) = \frac{1}{n^2 h^{2p} \Lambda_n^2(x)} \mathrm{var}\left(\sum_{j=1}^{J+1} \beta_j K\left(\frac{x-X}{h}\right) \frac{Y^{a_j}\mathbb{I}\{Y \geq y_{n,j}\}}{\psi_{a_j}(y_{n,j}|x)} \right) = \frac{1}{n^2 h^p \Lambda_n^2(x)} \beta^t B \beta,$$

où B est la matrice de variance-covariance de taille $(J+1) \times (J+1)$ définie par :

$$B_{j,l} = \frac{A_{j,l}}{\psi_{a_j}(y_{n,j}|x)\psi_{a_l}(y_{n,l}|x)},$$

pour tout $(j,l) \in \{1, \ldots, J+1\}^2$ et

$$\begin{aligned}
A_{j,l} &= \frac{1}{h^p} \mathrm{cov}\left(K\left(\frac{x-X}{h}\right) Y^{a_j}\mathbb{I}\{Y \geq y_{n,j}\}, K\left(\frac{x-X}{h}\right) Y^{a_l}\mathbb{I}\{Y \geq y_{n,l}\} \right), \\
&= \|K\|_2^2 \mathbb{E}\left(\frac{1}{h^p} Q\left(\frac{x-X}{h}\right) Y^{a_j+a_l} \mathbb{I}\{Y \geq y_{n,j} \vee y_{n,l}\} \right) \\
&- h^p \mathbb{E}(K_h(x-X)Y^{a_j}\mathbb{I}\{Y \geq y_{n,j}\})\mathbb{E}(K_h(x-X)Y^{a_l}\mathbb{I}\{Y \geq y_{n,l}\}),
\end{aligned}$$

avec $Q(.) := K^2(.)/\|K\|_2^2$ satisfaisant aussi l'hypothèse **(K)**. On peut alors utiliser la partie **(i)** de la preuve pour obtenir :

$$\begin{aligned}
A_{j,l} &= \|K\|_2^2 \psi_{a_j+a_l}(y_{n,j} \vee y_{n,l}|x)(1 + O(h) + O(\omega(x, y_{n,j} \vee y_{n,l}, a_j + a_l, 0, h))) \\
&- h^p \psi_{a_j}(y_{n,j}|x)\psi_{a_l}(y_{n,l}|x)(1 + O(h) + O(\omega(x, y_{n,j}, a_j, 0, h))) \\
&\times (1 + O(h) + O(\omega(x, y_{n,l}, a_l, 0, h))).
\end{aligned}$$

Soit $\xi > 0$ tel que :

$$\max_{a \in \mathcal{L}(\{a_1,\ldots,a_{J+1}\})} \omega(x, y_n, a, \xi, h) \to 0.$$

En remarquant que $\omega(x, y_{n,j}, a_j, 0, h) \leq \omega(x, y_n, a_j, \xi, h)$ pour n assez grand, on obtient :

$$\begin{aligned} A_{j,l} &= \|K\|_2^2 \psi_{a_j+a_l}(y_{n,j} \vee y_{n,l}|x)(1 + O(h) + O(\omega(x, y_n, a_j + a_l, \xi, h))) \\ &- h^p \psi_{a_j}(y_{n,j}|x)\psi_{a_l}(y_{n,l}|x)(1 + O(h) + O(\omega(x, y_n, a_j, \xi, h))) \\ &\times (1 + O(h) + O(\omega(x, y_n, a_l, \xi, h))). \end{aligned}$$

Or on a,

$$\max\left(\omega(x, y_n, a_j, \xi, h), \omega(x, y_n, a_l, \xi, h), \omega(x, y_n, a_j + a_l, \xi, h)\right) \leq \max_{a \in \mathcal{L}(\{a_1,\ldots,a_{J+1}\})} \omega(x, y_n, a, \xi, h),$$

ce qui donne :

$$\begin{aligned} B_{j,l} &= \frac{\|K\|_2^2 \psi_{a_j+a_l}(y_{n,j} \vee y_{n,l}|x)}{\psi_{a_j}(y_{n,j}|x)\psi_{a_l}(y_{n,l}|x)}(1 + O(h) + O(\omega_n(y_n, \xi))) \\ &- h^p(1 + O(h) + O(\omega_n(y_n, \xi))). \end{aligned}$$

Rappelons que puisque $\psi_a(.|x)$ est à variations régulières, il s'ensuit que :

$$\psi_{a_j}(y_{n,j}|x) \sim \psi_{a_j}(y_n|x) \to 0$$

pour tout $j \in \{1, \ldots, J+1\}$. Le Lemme 3.9 implique ainsi :

$$B_{j,l} = \frac{\|K\|_2^2}{\psi_0(y_n|x)} \frac{(1 - a_j\gamma(x))(1 - a_l\gamma(x))}{1 - (a_j + a_l)\gamma(x)}(1 + o(1)) = \frac{\|K\|_2^2}{\psi_0(y_n|x)} \Sigma_{j,l}^{(1)}(x)(1 + o(1)).$$

De plus,

$$\mathrm{var}(Z_{1,n}) \sim \|K\|_2^2 \beta^t \Sigma^{(1)}(x)\beta/n.$$

Comme conclusion préliminaire on peut affirmer que la variance de Ψ_n converge vers $\|K\|_2^2 \beta^t \Sigma^{(1)}(x)\beta$. Afin d'utiliser le théorème de Lyapounov il reste à prouver qu'il existe $\eta > 0$ tel que :

$$\sum_{i=1}^{n} \mathbb{E}\left|Z_{i,n}\right|^{2+\eta} = n\mathbb{E}\left|Z_{1,n}\right|^{2+\eta} \to 0.$$

Un calcul immédiat nous donne :

$$\begin{aligned} \mathbb{E}\left|Z_{1,n}\right|^{2+\eta} &= \left(\frac{1}{n\Lambda_n(x)}\right)^{2+\eta} \mathbb{E}\left|\sum_{j=1}^{J+1} \frac{\beta_j K_h(x - X) Y^{a_j} \mathbb{I}\{Y \geq y_{n,j}\}}{\psi_{a_j}(y_{n,j}|x)}\right. \\ &- \left.\mathbb{E}\left(\sum_{j=0}^{J+1} \frac{\beta_j K_h(x - X) Y^{a_j} \mathbb{I}\{Y \geq y_{n,j}\}}{\psi_{a_j}(y_{n,j}|x)}\right)\right|^{2+\eta}. \end{aligned}$$

Par ailleurs pour chaque couple de variables aléatoires (T_1, T_2) ayant un moment d'ordre $2+\eta$ fini, on a :

$$\mathbb{E}\left(|T_1+T_2|^{2+\eta}\right) \leq 2^{2+\eta} \max_{i=\{1,2\}} \mathbb{E}\left(|T_i|^{2+\eta}\right),$$

ce qui donne :

$$\mathbb{E}\left|Z_{1,n}\right|^{2+\eta} \leq \left(\frac{2}{n\Lambda_n(x)}\right)^{2+\eta} \mathbb{E}\left|\sum_{j=1}^{J+1} \frac{\beta_j K_h\left(x-X\right) Y^{a_j} \mathbb{I}\{Y \geq y_{n,j}\}}{\psi_{a_j}\left(y_{n,j}|x\right)}\right|^{2+\eta}.$$

En utilisant le Lemme 3.9 et le fait que $y_{n,j} = y_n(1+o(1))$ pour $j \in \{1,\ldots,J+1\}$ on a :

$$\begin{aligned}
\mathbb{E}\left|Z_{1,n}\right|^{2+\eta} &\leq \left(\frac{2}{n\Lambda_n(x)\psi_0(y_n|x)}\right)^{2+\eta} \\
&\times \mathbb{E}\left|\sum_{j=1}^{J+1} \beta_j K_h\left(x-X\right) \mathbb{I}\{Y \geq y_{n,j}\} \left(\frac{Y}{y_{n,j}}\right)^{a_j} (1-a_j\gamma(x))\right|^{2+\eta} (1+o(1)).
\end{aligned}$$

Soient $\tilde{a} = \max\{a_1,\ldots,a_{J+1}\}$ et $\tilde{y}_n = \min\{y_{n,1},\ldots,y_{n,J+1}\}$, il s'ensuit que pour n assez grand,

$$\begin{aligned}
n\mathbb{E}\left|Z_{1,n}\right|^{2+\eta} &\leq 2n\left(\frac{2(1-\tilde{a}\gamma(x))}{nh^p\Lambda_n(x)\psi_0(y_n|x)\tilde{y}_n^{\tilde{a}}}\right)^{2+\eta} \\
&\times \sum_{j=1}^{J+1} |\beta_j|^{2+\eta}\, \mathbb{E}\left(K\left(\frac{x-X}{h}\right) Y^{\tilde{a}} \mathbb{I}\{Y \geq \tilde{y}_n\}\right)^{2+\eta}.
\end{aligned}$$

En choisissant η tel que $0 < \eta < -2 + 1/(\tilde{a}\gamma(x))$, le point **(i)** implique :

$$\begin{aligned}
\mathbb{E}\left(K\left(\frac{x-X}{h}\right) Y^{\tilde{a}} \mathbb{I}\{Y \geq \tilde{y}_n\}\right)^{2+\eta} &= h^p \|K\|_{2+\eta}^{2+\eta} \mathbb{E}\left(N_h\left(x-X\right) Y^{\tilde{a}(2+\eta)} \mathbb{I}\{Y \geq \tilde{y}_n\}\right), \\
&= h^p \|K\|_{2+\eta}^{2+\eta} \psi_{\tilde{a}(2+\eta)}(\tilde{y}_n|x)(1+o(1)),
\end{aligned}$$

puisque $N(.) := K^{2+\eta}(.)/\|K\|_{2+\eta}^{2+\eta}$ vérifie aussi l'hypothèse **(K)**. En utilisant le Lemme 3.9 et le fait que $\tilde{y}_n = y_n(1+o(1))$, on obtient :

$$n\mathbb{E}\left|Z_{1,n}\right|^{2+\eta} = O\left(\Lambda_n^{\eta}(x)\right) \to 0,$$

lorsque $n \to \infty$ ce qui conclut la preuve. ■

Les comportements asymptotiques de nos estimateurs $\hat{\varphi}_{a,n}(.|x)$ et $\hat{\varphi}_{a,n}^{\leftarrow}(.|x)$ sont établis dans les deux propositions suivantes.

Proposition 3.12. *Supposons que les hypothèses* **(F.1)**, **(F.2)**, **(L)** *et* **(K)** *soient vérifiées. Soit $x \in \mathbb{R}^p$ tel que $g(x) > 0$ et soient $0 \leq a_1 < \cdots < a_{J+1} < 1/(2\gamma(x))$ où J est un entier positif. Considérons $y_n \to \infty$ tel que $nh^p\overline{F}(y_n|x) \to \infty$ lorsque $n \to \infty$ et les suites $(y_{n,j})$, $j \in \{1, \ldots, J+1\}$ telles que :*

$$\max_{j\in\{1,\ldots,J+1\}} \left| \frac{y_{n,j}}{y_n} - 1 \right| \to 0.$$

S'il existe $\xi > 0$ tel que

$$nh^p\bar{F}(y_n|x)\left(h \vee \max_{a\in\mathcal{L}(\{a_1,\ldots,a_{J+1}\})} \omega(x, y_n, a, \xi, h)\right)^2 \to 0$$

alors, le vecteur aléatoire :

$$\left\{\sqrt{nh^p\overline{F}(y_n|x)}\left(\frac{\hat{\varphi}_{a_j,n}(y_{n,j}|x)}{\varphi_{a_j}(y_{n,j}|x)} - 1\right)\right\}_{j\in\{1,\ldots,J+1\}}$$

est asymptotiquement Gaussien, centré, avec une matrice de variance-covariance :

$$\frac{\|K\|_2^2\Sigma^{(1)}(x)}{g(x)}.$$

Preuve – En gardant à l'esprit les notations du Lemme 3.11, il s'ensuit les développements suivants :

$$\Lambda_n^{-1}(x)\sum_{j=1}^{J+1}\beta_j\left(\frac{\hat{\varphi}_{a_j,n}(y_{n,j}|x)}{\varphi_{a_j}(y_{n,j}|x)} - 1\right) = \frac{\Delta_{1,n} + \Delta_{2,n} - \Delta_{3,n}}{\hat{g}_n(x)}, \tag{3.19}$$

où

$$\begin{aligned}
\Delta_{1,n} &= g(x)\Lambda_n^{-1}(x)\sum_{j=1}^{J+1}\beta_j\left(\frac{\hat{\psi}_{a_j,n}(y_{n,j}|x) - \mathbb{E}(\hat{\psi}_{a_j,n}(y_{n,j}|x))}{\psi_{a_j}(y_{n,j}|x)}\right),\\
\Delta_{2,n} &= g(x)\Lambda_n^{-1}(x)\sum_{j=1}^{J+1}\beta_j\left(\frac{\mathbb{E}(\hat{\psi}_{a_j,n}(y_{n,j}|x)) - \psi_{a_j}(y_{n,j}|x)}{\psi_{a_j}(y_{n,j}|x)}\right),\\
\Delta_{3,n} &= \left(\sum_{j=1}^{J+1}\beta_j\right)\Lambda_n^{-1}(x)\left(\hat{g}_n(x) - g(x)\right).
\end{aligned}$$

Ainsi, à l'aide du Lemme 3.11(ii), le terme aléatoire $\Delta_{1,n}$ peut se réécrire :

$$\Delta_{1,n} = g(x)\|K\|_2\sqrt{\beta^t\Sigma^{(1)}(x)\beta}\xi_n, \tag{3.20}$$

où ξ_n converge vers une loi Normale centrée réduite. Le terme déterministe $\Delta_{2,n}$ est contrôlé par le Lemme 3.11(i) :

$$\begin{aligned}\Delta_{2,n} &= O\left(h\Lambda_n^{-1}(x)\right) + O\left(\Lambda_n^{-1}(x) \max_{a\in\mathcal{L}(\{a_1,\dots,a_{J+1}\})} \omega(x,y_n,a,\xi,h)\right)\\ &= O\left(nh^{p+2}\bar{F}(y_n|x)\right)^{1/2} + O\left(nh^p\bar{F}(y_n|x) \max_{a\in\mathcal{L}(\{a_1,\dots,a_{J+1}\})} \omega^2(x,y_n,a,\xi,h)\right)^{1/2}\\ &= o(1). \end{aligned} \tag{3.21}$$

Finalement, $\Delta_{3,n}$ est un terme classique lorsque l'on estime une densité par noyau. Il peut être borné par [44], Lemme 4 :

$$\begin{aligned}\Delta_{3,n} &= O(h\Lambda_n^{-1}(x)) + O_P(\Lambda_n^{-1}(x)(nh^p)^{-1/2}),\\ &= O\left(nh^{p+2}\overline{F}(y_n|x)\right)^{1/2} + O_P(\overline{F}(y_n|x))^{1/2} = o_P(1). \end{aligned} \tag{3.22}$$

En rassemblant les équations (3.19)–(3.22), on obtient :

$$\hat{g}_n(x)\Lambda_n^{-1}(x)\sum_{j=1}^{J+1}\beta_j\left(\frac{\hat{\varphi}_{a_j,n}(y_{n,j}|x)}{\varphi_{a_j}(y_{n,j}|x)} - 1\right) = g(x)\|K\|_2\sqrt{\beta^t\Sigma^{(1)}(x)\beta}\xi_n + o_P(1).$$

Le dernier point et $\hat{g}_n(x) \xrightarrow{\mathbb{P}} g(x)$ donnent :

$$\sqrt{nh^p\overline{F}(y_n|x)}\sum_{j=1}^{J+1}\beta_j\left(\frac{\hat{\varphi}_{a_j,n}(y_{n,j}|x)}{\varphi_{a_j}(y_{n,j}|x)} - 1\right) = \|K\|_2\sqrt{\frac{\beta^t\Sigma^{(1)}(x)\beta}{g(x)}}\xi_n + o_P(1),$$

ce qui prouve le résultat énoncé. ■

Proposition 3.13. *Supposons que les hypothèses* **(F.1)**, **(F.2)**, **(L)** *et* **(K)** *soient vérifiées. Soit $x \in \mathbb{R}^p$ tel que $g(x) > 0$ et soient $0 \le a_1 < \cdots < a_{J+1} < 1/(2\gamma(x))$ où J est un entier positif. Considérons $\alpha_n \to 0$ tel que $nh^p\alpha_n \to \infty$ lorsque $n \to \infty$. Soient $(\alpha_{n,j})$, $j = 1,\dots,J+1$ des suites telles que :*

$$\max_{j\in\{1,\dots,J+1\}}\left|\frac{\varphi_{a_j}^{\leftarrow}(\alpha_{n,j}|x)}{\varphi_0^{\leftarrow}(\alpha_n|x)} - 1\right| \to 0,$$

S'il existe $\xi > 0$ tel que

$$nh^p\alpha_n\left(h \vee \max_{a\in\mathcal{L}(\{a_1,\dots,a_{J+1}\})}\omega(x,\varphi_0^{\leftarrow}(\alpha_n|x),a,\xi,h)\right)^2 \to 0$$

alors, le vecteur aléatoire :

$$\left\{\sqrt{nh^p\alpha_n}\left(\frac{\hat{\varphi}_{a_j,n}^{\leftarrow}(\alpha_{n,j}|x)}{\varphi_{a_j}^{\leftarrow}(\alpha_{n,j}|x)} - 1\right)\right\}_{j\in\{1,\dots,J+1\}}$$

est asymptotiquement Gaussien, centré, avec une matrice de variance-covariance :

$$\frac{\|K\|_2^2\Sigma^{(2)}(x)}{g(x)} \quad \text{où} \quad \Sigma_{i,j}^{(2)}(x) = \frac{\gamma^2(x)}{1-(a_i+a_j)\gamma(x)}, \ (i,j) \in \{1,\dots,J+1\}^2.$$

Preuve – Introduisons pour $j \in \{1,\dots,J+1\}$, les termes suivants :

$$\begin{aligned}
\sigma_{n,j}(x) &= \varphi_{a_j}^{\leftarrow}(\alpha_{n,j}|x)(nh^p\alpha_n)^{-1/2}, \\
v_{n,j}(x) &= \alpha_{n,j}^{-1}\frac{\gamma(x)}{1-a_j\gamma(x)}(nh^p\alpha_n)^{1/2}, \\
W_{n,j}(x) &= v_{n,j}(x)\left(\hat{\varphi}_{a_j}(\varphi_{a_j}^{\leftarrow}(\alpha_{n,j}|x)+\sigma_{n,j}(x)z_j)|x) - \varphi_{a_j}(\varphi_{a_j}^{\leftarrow}(\alpha_{n,j}|x)+\sigma_{n,j}(x)z_j)|x\right), \\
t_{n,j}(x) &= v_{n,j}(x)\left(\alpha_{n,j} - \varphi_{a_j}(\varphi_{a_j}^{\leftarrow}(\alpha_{n,j}|x)+\sigma_{n,j}(x)z_j)|x\right),
\end{aligned}$$

où $(z_1,\dots,z_{J+1}) \in \mathbb{R}^{J+1}$. Examinons le comportement asymptotique de la fonction de répartition définie par :

$$\begin{aligned}
\Phi_n(z_1,\dots,z_{J+1}) &= \mathbb{P}\left(\bigcap_{j=1}^{J+1}\left\{\sigma_{n,j}^{-1}(x)(\hat{\varphi}_{a_j,n}^{\leftarrow}(\alpha_{n,j}|x) - \varphi_{a_j}^{\leftarrow}(\alpha_{n,j}|x)) \le z_j\right\}\right), \\
&= \mathbb{P}\left(\bigcap_{j=1}^{J+1}\{W_{n,j}(x) \le t_{n,j}(x)\}\right).
\end{aligned}$$

Concentrons nous d'abord sur les termes déterministes $t_{n,j}(x)$, $j \in \{1,\dots,J+1\}$. D'après le Lemme 3.9, pour tout $a \in [0, 1/(2\gamma(x)))$, la fonction $\varphi_a(.|x)$ est dérivable ainsi, pour chaque $j \in \{1,\dots,J+1\}$ il existe $\theta_{n,j} \in]0,1[$ tel que :

$$\varphi_{a_j}\left(\varphi_{a_j}^{\leftarrow}(\alpha_{n,j}|x)|x\right) - \varphi_{a_j}\left(\varphi_{a_j}^{\leftarrow}(\alpha_{n,j}|x)+\sigma_{n,j}(x)z_j|x\right) = -\sigma_{n,j}(x)z_j\varphi'_{a_j}(r_{n,j}|x), \quad (3.23)$$

où $r_{n,j} = \varphi_{a_j}^{\leftarrow}(\alpha_{n,j}|x) + \theta_{n,j}\sigma_{n,j}(x)z_j$. Il est ainsi évident que $r_{n,j} \sim \varphi_{a_j}^{\leftarrow}(\alpha_{n,j}|x) \to \infty$. De plus le Lemme 3.9 donne

$$\varphi'_{a_j}(r_{n,j}|x) = \frac{(a_j\gamma(x)-1)\alpha_{n,j}}{\gamma(x)\varphi_{a_j}^{\leftarrow}(\alpha_{n,j}|x)}(1+o(1)). \quad (3.24)$$

Au vu des deux équations (3.23) et (3.24), on termine avec :

$$t_{n,j}(x) = \frac{(1-a_j\gamma(x))v_{n,j}(x)\sigma_{n,j}(x)\alpha_{n,j}z_j}{\gamma(x)\varphi_{a_j}^{\leftarrow}(\alpha_{n,j}|x)}(1+o(1)) = z_j(1+o(1)). \quad (3.25)$$

Penchons nous sur les termes aléatoires $W_{n,j}(x)$, $j \in \{1,\dots,J+1\}$. Il est clair que les suites $y_{n,j} := \varphi_{a_j}^{\leftarrow}(\alpha_{n,j}|x) + \sigma_{n,j}(x)z_j$, $j = 1,\dots,J+1$ et $y_n := \varphi_0^{\leftarrow}(\alpha_n|x)$ satisfont les

hypothèses de la Proposition 3.12. Par conséquent on a,

$$W_{n,j}(x) = \frac{\gamma(x)}{1 - a_j\gamma(x)} \frac{\varphi_{a_j}(\varphi_{a_j}^{\leftarrow}(\alpha_{n,j}|x) + \sigma_{n,j}(x)z_j|x)}{\alpha_{n,j}} (nh^p\alpha_n)^{1/2} \left(\frac{\hat{\varphi}_{a_j}(y_{n,j}|x)}{\varphi_{a_j}(y_{n,j}|x)} - 1 \right).$$

De plus, puisque $\varphi_a(.|x)$ est à variations régulières, les équivalences suivantes sont vérifiées :

$$\frac{\varphi_{a_j}(\varphi_{a_j}^{\leftarrow}(\alpha_{n,j}|x) + \sigma_{n,j}(x)z_j|x)}{\alpha_{n,j}} = \frac{\varphi_{a_j}(\varphi_{a_j}^{\leftarrow}(\alpha_{n,j}|x)(1 + o_{\mathbb{P}}(1))|x)}{\alpha_{n,j}} = 1 + o_{\mathbb{P}}(1).$$

Comme conséquence du théorème de Slutsky, le vecteur aléatoire $(W_{n,1}, \ldots, W_{n,J+1})$ est égal à $A(x)\xi_n$ où

$$A(x) = \text{diag}\left(\frac{\gamma(x)}{1 - a_1\gamma(x)}, \ldots, \frac{\gamma(x)}{1 - a_{J+1}\gamma(x)} \right),$$

et ξ_n est un $(J+1)-$ vecteur aléatoire convergeant vers une loi Normale centrée de matrice de variance-covariance $\|K\|_2^2\Sigma^{(1)}(x)/g(x)$. En tenant compte de l'équation (3.25), on obtient que $\Phi_n(z_1, \ldots, z_{J+1})$ converge vers une fonction de répartition d'une loi Normale centrée de matrice de variance-covariance

$$\frac{\|K\|_2^2 A(x)\Sigma^{(1)}(x)A(x)}{g(x)} = \frac{\|K\|_2^2\Sigma^{(2)}(x)}{g(x)}.$$

Ce qui est le résultat attendu. ■

3.6.2 Preuves des principaux résultats

Preuve du Théorème 3.3.

Considérons pour $j \in \{1, \dots, J\}$ les termes suivants :

$$\begin{aligned}
v_{n,j}(x) &= \frac{(1 - a_j\gamma(x))(nh^p\alpha_n)^{1/2}}{\gamma(x)\varphi_0^{\leftarrow}(\alpha_n|x)}, \\
\sigma_{n,j}(x) &= \varphi_{a_j}(\varphi_0^{\leftarrow}(\alpha_n|x)|x)(nh^p\alpha_n)^{-1/2}, \\
\sigma_{n,0}(x) &= \varphi_0^{\leftarrow}(\alpha_n|x)(nh^p\alpha_n)^{-1/2}, \\
t_{n,j} &= v_{n,j}(x)\left(\varphi_0^{\leftarrow}(\alpha_n|x) - \varphi_{a_j}^{\leftarrow}\left(\varphi_{a_j}(\varphi_0^{\leftarrow}(\alpha_n|x)|x) + \sigma_{n,j}(x)z_j|x\right)\right), \\
W_{n,j}(x) &= v_{n,j}(x)\left(\hat{\varphi}_{a_j,n}^{\leftarrow}\left(\varphi_{a_j}(\varphi_0^{\leftarrow}(\alpha_n|x)|x) + \sigma_{n,j}(x)z_j|x\right)\right. \\
&- \left.\varphi_{a_j}^{\leftarrow}\left(\varphi_{a_j}(\varphi_0^{\leftarrow}(\alpha_n|x)|x) + \sigma_{n,j}(x)z_j|x\right)\right), \\
W_{n,j}^{(0)}(x) &= v_{n,j}(x)\left(\hat{\varphi}_{0,n}^{\leftarrow}(\alpha_n|x) - \varphi_0^{\leftarrow}(\alpha_n|x)\right), \\
W_{n,0}^{(0)}(x) &= \sigma_{n,0}^{-1}(x)\left(\hat{\varphi}_{0,n}^{\leftarrow}(\alpha_n|x) - \varphi_0^{\leftarrow}(\alpha_n|x)\right),
\end{aligned}$$

où $(z_0, z_1, \dots, z_J) \in \mathbb{R}^{J+1}$. Regardons le comportement asymptotique de la fonction de répartition définie par :

$$\begin{aligned}
&\Phi_n(z_0, z_1, \dots, z_J) \\
= \ &\mathbb{P}\left(\left\{\bigcap_{j=1}^{J}\left\{\sigma_{n,j}^{-1}(x)(\hat{\varphi}_{a_j,n}(\hat{\varphi}_{0,n}^{\leftarrow}(\alpha_n|x)|x) - \varphi_{a_j}(\varphi_0^{\leftarrow}(\alpha_n|x)|x)) \le z_j\right\}\right\} \bigcap \left\{W_{n,0}^{(0)}(x) \le z_0\right\}\right), \\
= \ &\mathbb{P}\left(\left\{\bigcap_{j=1}^{J}\left\{W_{n,j}(x) - W_{n,j}^{(0)}(x) \le t_{n,j}\right\}\right\} \bigcap \left\{W_{n,0}^{(0)}(x) \le z_0\right\}\right).
\end{aligned}$$

Concentrons-nous d'abord sur les termes déterministes $t_{n,j}(x)$, $j = 1, \dots, J$. D'après le Lemme 3.9, pour tout $a \in [0, 1/(2\gamma(x)))$, $\varphi_a^{\leftarrow}(.|x)$ est une fonction à variations régulières dérivable telle que :

$$(\varphi_a^{\leftarrow})'(y_n|x) = \frac{1}{\varphi_a'(\varphi_a^{\leftarrow}(y_n|x)|x)} = \frac{\gamma(x)\varphi_a^{\leftarrow}(y_n|x)}{(a\gamma(x) - 1)y_n}(1 + o(1)), \tag{3.26}$$

lorsque $n \to \infty$. Pour tout $j \in \{1, \dots, J\}$, un développement de Taylor à l'ordre un donne :

$$\varphi_{a_j}^{\leftarrow}(\varphi_{a_j}(\varphi_0^{\leftarrow}(\alpha_n|x)|x)|x) - \varphi_{a_j}^{\leftarrow}\left(\varphi_{a_j}(\varphi_0^{\leftarrow}(\alpha_n|x)|x) + \sigma_{n,j}(x)z_j|x\right) = -\sigma_{n,j}(x)z_j q_{n,j}(x),$$

où

$$q_{n,j}(x) = (\varphi_{a_j}^{\leftarrow})'(\varphi_{a_j}(\varphi_0^{\leftarrow}(\alpha_n|x)|x) + \theta_{n,j}\sigma_{n,j}(x)z_j|x)$$

avec $(\theta_{n,1},\dots,\theta_{n,J}) \in]0,1[^J$. Puisque :

$$\sigma_{n,j}(x)/\varphi_{a_j}(\varphi_0^{\leftarrow}(\alpha_n|x)|x) = (nh^p\alpha_n)^{-1/2} \to 0$$

lorsque $n \to \infty$, l'équation (3.26) implique :

$$q_{n,j}(x) = \frac{\gamma(x)\varphi_0^{\leftarrow}(\alpha_n|x)}{(a_j\gamma(x)-1)\varphi_{a_j}(\varphi_0^{\leftarrow}(\alpha_n|x)|x)}(1+o(1)).$$

Ainsi,

$$\varphi_{a_j}^{\leftarrow}(\varphi_{a_j}(\varphi_0^{\leftarrow}(\alpha_n|x)|x)|x) - \varphi_{a_j}^{\leftarrow}\left(\varphi_{a_j}(\varphi_0^{\leftarrow}(\alpha_n|x)|x) + \sigma_{n,j}(x)z_j|x\right) = \frac{z_j}{v_{n,j}(x)}(1+o(1)), \tag{3.27}$$

ce qui montre que pour tout $j \in \{1,\dots,J\}$, $t_{n,j} \to z_j$ lorsque $n \to \infty$. Concentrons nous sur les termes aléatoires $W_{n,j}(x)$, $j = 1,\dots,J$. Il est clair que :

$$W_{n,j}(x) = \frac{1-a_j\gamma(x)}{\gamma(x)}(nh^p\alpha_n)^{1/2}\left(\frac{\hat{\varphi}_{a_j,n}^{\leftarrow}\left(\varphi_{a_j}(\varphi_0^{\leftarrow}(\alpha_n|x)|x) + \sigma_{n,j}(x)z_j|x\right)}{\varphi_{a_j}^{\leftarrow}\left(\varphi_{a_j}(\varphi_0^{\leftarrow}(\alpha_n|x)|x) + \sigma_{n,j}(x)z_j|x\right)} - 1\right)(1+o(1)),$$

puisque d'après l'équation (3.27) :

$$\frac{\varphi_{a_j}^{\leftarrow}\left(\varphi_{a_j}(\varphi_0^{\leftarrow}(\alpha_n|x)|x) + \sigma_{n,j}(x)z_j|x\right)}{\varphi_0^{\leftarrow}(\alpha_n|x)} = 1 + \frac{z_j}{\varphi_0^{\leftarrow}(\alpha_n|x)v_{n,j}(x)}(1+o(1)) = 1+o(1).$$

De plus, on a :

$$W_{n,j}^{(0)}(x) = \frac{1-a_j\gamma(x)}{\gamma(x)}(nh^p\alpha_n)^{1/2}\left(\frac{\hat{\varphi}_{0,n}^{\leftarrow}(\alpha_n|x)}{\varphi_0^{\leftarrow}(\alpha_n|x)} - 1\right).$$

Comme conséquence il en résulte qu'en appliquant la Proposition 3.13 avec $a_{J+1} = 0$, $\alpha_{n,j} = \varphi_{a_j}(\varphi_0^{\leftarrow}(\alpha_n|x)|x) + \sigma_{n,j}(x)z_j$ pour $j = 1,\dots,J$ et $\alpha_{n,J+1} = \alpha_n$ cela implique

$$\left\{\left\{W_{n,j}(x) - W_{n,j}^{(0)}(x)\right\}_{j=1,\dots,J}, W_{n,0}^{(0)}(x)\right\} = M(x)\xi_n,$$

où M est la matrice de taille $(J+1)\times(J+1)$ définie par :

$$M(x) = \left(\begin{array}{c|c} \tilde{A}(x) & \begin{matrix} c_1(x) \\ \vdots \\ c_J(x) \end{matrix} \\ \hline 0\cdots 0 & 1 \end{array}\right)$$

avec

$$\tilde{A}(x) = \operatorname{diag}\left(\frac{1-a_1\gamma(x)}{\gamma(x)},\dots,\frac{1-a_J\gamma(x)}{\gamma(x)}\right) \quad \text{et} \quad c_j = -\frac{1-a_j\gamma(x)}{\gamma(x)}, \quad j \in \{1,\dots,J\}$$

et où ξ_n est un vecteur Gaussien, centré de taille $(J+1)$ avec une matrice de variance-covariance $\|K\|_2^2\Sigma^{(2)}(x)/g(x)$. Puisque pour chaque $j \in \{1,\dots,J\}$, $t_{n,j} \to z_j$ lorsque $n \to \infty$, la fonction de répartition Φ_n converge alors vers une fonction de répartition Gaussienne ayant une matrice de variance-covariance égale à

$$\frac{\|K\|_2^2 M(x)\Sigma^{(2)}(x)M(x)^t}{g(x)} = \frac{\|K\|_2^2\gamma^2(x)\Sigma(x)}{g(x)},$$

ce qui est le résultat escompté. ■

Preuve du Corollaire 3.5. Il est clair que d'après le Théorème 3.3 on a pour $i = 1, 2$,

$$\widehat{\mathrm{RCTM}}_{i,n}(\alpha_n|x) = \mathrm{RCTM}_i(\alpha_n|x)\left(1 + (nh^p\alpha_n)^{-1/2}\xi_{i,n}\right),$$

où le vecteur aléatoire $(\xi_{1,n}, \xi_{2,n})$ est asymptotiquement Gaussien, centré de matrice de variance-covariance $\Sigma^{(3)}$ définie par :

$$\Sigma^{(3)}_{i,j} = ij\gamma^2(x)\frac{2-(i+j)\gamma(x)}{1-(i+j)\gamma(x)}\frac{\|K\|_2^2}{g(x)},\ (i,j) \in \{1,2\}^2.$$

Ainsi,

$$\begin{aligned}
\widehat{\mathrm{RCTV}}_n(\alpha_n|x) - \mathrm{RCTV}(\alpha_n|x) &= (nh^p\alpha_n)^{-1/2}\left(\mathrm{RCTM}_2(\alpha_n|x)\xi_{2,n}\right.\\
&- 2\mathrm{RCTM}_1^2(\alpha_n|x)\xi_{1,n}\\
&- \left.(nh^p\alpha_n)^{-1/2}\mathrm{RCTM}_1^2(\alpha_n|x)\xi_{1,n}^2\right),\\
&= (nh^p\alpha_n)^{-1/2}\left(\mathrm{RCTM}_2(\alpha_n|x)\xi_{2,n}\right.\\
&- \left.2\mathrm{RCTM}_1^2(\alpha_n|x)\xi_{1,n}(1+o(1))\right).
\end{aligned}$$

Puisque pour $a \in [0, 1/(2\gamma(x)))$ et $b > 0$ le Lemme 3.9 implique :

$$\mathrm{RCTM}_a^b(\alpha_n|x) = \frac{(\varphi_0^{\leftarrow}(\alpha_n|x))^{ab}}{(1-a\gamma(x))^b}(1+o(1)),$$

on a alors :

$$\mathrm{RCTV}(\alpha_n|x) = \frac{\gamma^2(x)}{(1-2\gamma(x))(1-\gamma(x))^2}(\varphi_0^{\leftarrow}(\alpha_n|x))^2(1+o(1)).$$

Ainsi, d'après le Théorème 3.3 :

$$(nh^p\alpha_n)^{1/2}\left(\frac{\widehat{\mathrm{RCTV}}_n(\alpha_n|x)}{\mathrm{RCTV}(\alpha_n|x)} - 1\right) = \frac{4\gamma(x)-2}{\gamma^2(x)}\xi_{1,n}(1+o(1)) + \frac{(1-\gamma(x))^2}{\gamma^2(x)}\xi_{2,n}(1+o(1)),$$

est asymptotiquement Gaussien, centré de variance :

$$\left(\frac{4\gamma(x)-2}{\gamma^2(x)}, \frac{(1-\gamma(x))^2}{\gamma^2(x)}\right) \Sigma^{(3)} \left(\frac{4\gamma(x)-2}{\gamma^2(x)}, \frac{(1-\gamma(x))^2}{\gamma^2(x)}\right)^t.$$

La conclusion s'ensuit. ■

Preuve du Corollaire 3.6. Il est clair que d'après le Théorème 3.3 on a pour $i = 1, 2, 3$,

$$\widehat{\mathrm{RCTM}}_{i,n}(\alpha_n|x) = \mathrm{RCTM}_i(\alpha_n|x)\left(1 + (nh^p\alpha_n)^{1/2}\xi_{i,n}\right),$$

où le vecteur aléatoire $(\xi_{1,n}, \xi_{2,n}, \xi_{3,n})$ est asymptotiquement Gaussien, centré de matrice de variance-covariance $\Sigma^{(4)}$ définie par :

$$\Sigma^{(4)}_{i,j} = ij\gamma^2(x)\frac{\|K\|_2^2}{g(x)}\frac{2-(i+j)\gamma(x)}{1-(i+j)\gamma(x)}, \ (i,j) \in \{1,2,3\}^2.$$

D'après la preuve du Corollaire 3.5, on a :

$$\frac{\widehat{\mathrm{RCTV}}_n(\alpha_n|x)}{\mathrm{RCTV}(\alpha_n|x)} = 1+(nh^p\alpha_n)^{-1/2}\left(\frac{4\gamma(x)-2}{\gamma^2(x)}\xi_{1,n}(1+o(1)) + \frac{(1-\gamma(x))^2}{\gamma^2(x)}\xi_{2,n}(1+o(1))\right),$$

et ainsi on obtient :

$$\left(\frac{\widehat{\mathrm{RCTV}}_n(\alpha_n|x)}{\mathrm{RCTV}(\alpha_n|x)}\right)^{3/2} = 1+\frac{3}{2}(nh^p\alpha_n)^{-1/2}\left(\frac{4\gamma(x)-2}{\gamma^2(x)}\xi_{1,n}(1+o(1)) + \frac{(1-\gamma(x))^2}{\gamma^2(x)}\xi_{2,n}(1+o(1))\right).$$

Il est clair que d'après le Théorème 3.3,

$$(nh^p\alpha_n)^{1/2}\left(\frac{\widehat{\mathrm{RCTS}}_n(\alpha_n|x)}{\mathrm{RCTS}(\alpha_n|x)} - 1\right) = \xi_{3,n} - \frac{3}{2}\left(\frac{4\gamma(x)-2}{\gamma^2(x)}\xi_{1,n}(1+o(1)) + \frac{(1-\gamma(x))^2}{\gamma^2(x)}\xi_{2,n}(1+o(1))\right),$$

est asymptotiquement Gaussien, centré de variance :

$$\left(-\frac{3}{2}\frac{4\gamma(x)-2}{\gamma^2(x)}, -\frac{3}{2}\frac{(1-\gamma(x))^2}{\gamma^2(x)}, 1\right) \Sigma^{(4)} \left(-\frac{3}{2}\frac{4\gamma(x)-2}{\gamma^2(x)}, -\frac{3}{2}\frac{(1-\gamma(x))^2}{\gamma^2(x)}, 1\right)^t.$$

Le résultat est ainsi prouvé. ■

Preuve du Théorème 3.7.

La preuve est basée sur le développement suivant :

$$\frac{\sqrt{nh_n^p\alpha_n}}{\log(\alpha_n/\beta_n)}\left(\log \widehat{\mathrm{RCTM}}^W_{a,n}(\beta_n|x) - \log \mathrm{RCTM}_a(\beta_n|x)\right) = \frac{\sqrt{nh_n^p\alpha_n}}{\log(\alpha_n/\beta_n)}(Q_{n,1}+Q_{n,2}+Q_{n,3}),$$

avec :

$$\begin{aligned}
Q_{n,1} &= a\sqrt{nh_n^p\alpha_n}(\hat{\gamma}_n(x)-\gamma(x)),\\
Q_{n,2} &= \frac{\sqrt{nh_n^p\alpha_n}}{\log(\alpha_n/\beta_n)}\log\left(\widehat{\mathrm{RCTM}}_{a,n}(\alpha_n|x)/\mathrm{RCTM}_a(\alpha_n|x)\right),\\
Q_{n,3} &= \frac{\sqrt{nh_n^p\alpha_n}}{\log(\alpha_n/\beta_n)}(\log\mathrm{RCTM}_a(\alpha_n|x)-\log\mathrm{RCTM}_a(\beta_n|x)+a\gamma(x)\log(\alpha_n/\beta_n)).
\end{aligned}$$

Considérons les trois termes séparément. Sous les hypothèses du Théorème 3.7, on a $Q_{n,1}\xrightarrow{d}\mathcal{N}\left(0,(av(x))^2\right)$. Le Théorème 3.3 implique :

$$\widehat{\mathrm{RCTM}}_{a,n}(\alpha_n|x)/\mathrm{RCTM}_a(\alpha_n|x)\xrightarrow{\mathbb{P}}1$$

quand $n\to\infty$ et ainsi :

$$Q_{n,2}=\frac{\sqrt{nh_n^p\alpha_n}}{\log(\alpha_n/\beta_n)}\left(\frac{\widehat{\mathrm{RCTM}}_{a,n}(\alpha_n|x)}{\mathrm{RCTM}_a(\alpha_n|x)}-1\right)(1+o_p(1))=\frac{O_p(1)}{\log(\alpha_n/\beta_n)}.$$

Comme conséquence, $Q_{n,2}\xrightarrow{\mathbb{P}}0$ lorsque $n\to\infty$. Finalement, le Lemme 3.10 implique :

$$Q_{n,3}=O\left(\sqrt{nh_n^p\alpha_n}\varepsilon_a(1/\alpha_n|x)\right)$$

qui converge vers 0 par hypothèse. ■

Conclusion et perspectives

"There are no serious competitor models to those provided by extreme value theory."

Stuart Coles

L'objectif de cette thèse était double.

Dans un premier temps, le but était de compléter le modèle de Gardes *et al.* [88] permettant d'estimer des quantiles extrêmes pour des lois de type Pareto aussi bien que pour des lois à queue de type Weibull d'une même façon. Ce modèle reposait sur un paramètre inconnu. Notre but premier fut de l'estimer, suite à quoi nous avons obtenu un nouvel estimateur des quantiles extrêmes utile dans des situations pratiques. Après avoir établi les lois asymptotiques de nos estimateurs on a pu illustrer leur efficacité sur des données simulées et sur un jeu de données réelles issu de l'hydrologie. Ces contributions ont fait l'objet de l'article El Methni *et al.* [69] paru en 2012.

Dans un second temps, le but était d'estimer les mesures de risque classiques que sont la Value-at-Risk, la Conditional Tail Expectation, la Conditional Value-at-Risk, la Conditional Tail Variance, ou encore la Conditional Tail Skewness dans le cas de pertes extrêmes et où l'on suppose également que la loi des pertes est à queue lourde et dépend d'une covariable aléatoire de dimension finie. Pour cela, on a introduit une nouvelle mesure de risque que l'on a appelé le Conditional Tail Moment. On a vu qu'estimer cette mesure de risque nous permettait d'estimer toutes les précédentes. Pour ce faire, on a introduit de nouvelles méthodes d'estimation, combinant des méthodes d'estimation non-paramétrique à noyau avec des méthodes issues de la statistique des valeurs extrêmes. Après avoir établi les lois asymptotiques de nos estimateurs, on a illustré leurs comportements sur des données simulées et sur un jeu de données réelles issu de l'hydrologie. Ces contributions ont fait l'objet du rapport de recherche El Methni *et al.* [68] soumis pour publication.

Cette thèse offre de nombreuses perspectives, aussi bien d'un point théorique avec les nouveaux axes de recherche qu'elle dessine que d'un point de vue pratique avec ses applications à l'hydrologie.

D'un point de vue théorique, on a vu dans les deux Parties 2.5 et 3.5 les perspectives possibles à donner à nos travaux à court terme.

- En ce qui concerne la première contribution, on pourra considérer la correction de biais basée sur l'estimation du paramètre du second ordre (Caeiro et Gomes [28], Gomes *et al.* [103]) parmi les suites explorables. Dans ce but, un modèle de régression exponentielle similaire à ceux de Beirlant *et al.* [12], Diebolt *et al.* [59, 60] et Feuerverger et Hall [76] étendu aux lois appartenant à la famille de fonctions de survie $(\mathbf{A}_1(\tau,\theta))$ serait d'un grand intérêt.

  Il serait aussi intéressant d'adapter nos résultats au cas $\tau > 1$ et de chercher des liens possibles avec des lois à queues "super-lourdes" (Fraga Alves *et al.* [78]). Comme on a pu le voir dans la sous-Partie Simulations 2.4.1, nos estimateurs sont très performants même pour des lois ayant des valeurs de τ où la normalité asymptotique de nos estimateurs n'est pas établie. En ce qui concerne le choix des suites (k_n) et (k'_n) il serait intéressant d'étudier le cas $\tau = 0$ et $\rho \geq -1$ de façon plus précise.

  Ce travail peut également s'étendre à des variables aléatoires $Z = \psi(Y)$ où Y a une loi associée satisfaisant $(\mathbf{A}_1(\tau,\theta))$. Par exemple, en choisissant $\psi(y) = y_F - 1/y$, cela nous permettra de considérer des lois (ayant un point terminal fini y_F) dans le domaine d'attraction de Weibull (voir le passage du domaine d'attraction de Fréchet à celui de Weibull dans le Théorème 1.11). Cela pourrait ainsi permettre d'inclure l'estimateur de Hill négatif (voir par exemple Falk [74] ou de Haan et Ferreira [50], Paragraphe 3.6.2) dans notre modèle.

  Finalement, nos travaux ouvrent la voie à la construction de tests d'hypothèses pour les queues de distributions. Pour un jeu de données réelles, on pourrait notamment décider s'il est issu d'une loi à queue lourde ou d'une loi à queue de type Weibull.

- Enfin pour la seconde contribution, parmi les différentes suites à donner à ce travail, il serait intéressant d'essayer de lever la restriction sur $\gamma(x)$ dans nos théorèmes (voir Partie 3.3). Necir *et al.* [136] ont estimé la CTE pour des lois à queues lourdes (sans la présence d'une covariable) et ont obtenu la restriction $1/2 < \gamma < 1$. On pourrait essayer d'adapter leurs outils pour pallier nos limitations sur $\gamma(x)$.

Afin d'étendre nos résultats aux autres domaines d'attraction (Gumbel et Weibull), on pourrait s'inspirer des travaux de Daouia *et al.* [43]. Ces derniers ont estimé les quantiles conditionnels extrêmes en présence d'une covariable dans tous les domaines d'attraction en se basant sur des conditions dues à Von Mises (voir [50]) sur la fonction de survie conditionnelle. On pourrait alors estimer pour des pertes extrêmes et en présence d'une covariable le RCTM, et ainsi toutes les mesures de risque classiques présentées dans cette thèse.

Lorsque la covariable est non aléatoire de dimension infinie, Gardes *et al.* [89] ont proposé un estimateur du quantile conditionnel extrême pour des lois à queues lourdes. En se basant sur ces résultats, on pourrait envisager d'étendre nos travaux à ce cas, et ainsi estimer toutes les mesures de risque classiques dans le cas d'une covariable non aléatoire de dimension infinie.

Une autre piste très intéressante à explorer serait d'estimer $\gamma(x)$ à l'aide de combinaisons de différents ordres du RCTM. En effet, à l'aide du Lemme 3.9 on peut facilement obtenir un équivalent du RCTM, et ainsi à l'instar des travaux de Girard *et al.* [94, 95, 96] on pourrait estimer $\gamma(x)$.

En ce qui concerne la suite commune à donner à ces deux contributions, on pourrait envisager de faire le lien entre nos deux contributions en adaptant le modèle de Gardes *et al.* [88] à la présence d'une covariable et ainsi être capable d'estimer des quantiles conditionnels extrêmes pour des lois de type Pareto aussi bien que pour des lois à queue de type Weibull d'une même façon. On pourrait aussi envisager de proposer un estimateur du RCTM dans ce modèle, ce qui nous permettrait d'estimer toutes les mesures de risque classiques pour des lois de type Pareto aussi bien que pour des lois à queue de type Weibull d'une même façon. On vient de voir que l'on peut également étendre le modèle de Gardes *et al.* [88] au domaine d'attraction de Weibull, cela nous permettrait alors d'englober tous les domaines d'attractions.

D'un point de vue applicatif, on pu mettre en pratique nos estimateurs sur des données hydrologiques, soulignant encore une fois le fort lien historique qui unit théorie des valeurs extrêmes et hydrologie. En ce qui concerne la première contribution, on a pu appliquer nos estimateurs à un jeu de données de crues de la rivière Nidd en Grande Bretagne dont on donne les résultats dans la sous-Partie 2.4.2.

Pour la seconde contribution, on a fait le lien entre les mesures de risque utilisées en finance et en actuariat et le risque en hydrologie. On a ainsi introduit une nouvelle méthode statistique permettant d'estimer des mesures de risque extrêmes en des sites où l'on ne dispose pas de stations de mesure. Plus particulièrement, on a pu présenter

une application à la région Cévennes-Vivarais en comparant l'estimation du niveau de retour à 100 ans et de la RCTE correspondante dont on donne les résultats dans la sous-Partie 3.4.4.

Ces mesures de risque sont un nouvel outil dans la prévention des risques en hydrologie. Il serait très intéressant de prendre d'autres covariables en compte. Par exemple on pourrait tenir compte de l'altitude, l'orientation par rapport à la mer ou encore le type de temps. Pour cela il nous faudrait définir des nouvelles distances. Il serait aussi intéressant d'appliquer notre méthode à un nouveau type de données et de pouvoir estimer des mesures de risque dans d'autres domaines d'applications, par exemple en actuariat ou en finance.

Enfin, il est primordial de garder à l'esprit que les résultats obtenus sont basés sur des modèles et qu'ils sont (le plus souvent) une simplification de la réalité ; il convient en conséquence d'adopter une certaine prudence dans l'interprétation des résultats obtenus.

"*The key message is that extreme value theory cannot do magic–but it can do a whole lot better than empirical curvefitting and guesswork. My answer to the sceptics is that people aren't given well-founded methods like extreme value theory, they will just use dubious ones instead.*"

Jonathan Tawn

Bibliographie

[1] Aarssen, K. et de Haan, L. (1994). On the maximal life span of humans. *Mathematical Population Studies*, 4(4):259–281. 4

[2] Alvarado, E., Sandberg, D. et Pickford, S. (1998). Modeling large forest fires as extreme events. *Northwest Science*, 72:66–75. 4

[3] Alves, M., de Haan, L. et Lin, T. (2003a). Estimation of the parameter controlling the speed of convergence in extreme value theory. *Mathematical Methods of Statistics*, 12:155–176. 42

[4] Alves, M., Gomes, M. et de Haan, L. (2003b). A new class of semi-parametric estimators of the second order parameter. *Portugaliae Mathematica*, 60:193–214. 42

[5] Artin, E. (1964). *The gamma function*. Holt, Rinehart and Winston. 92

[6] Artzner, P., Delbaen, F., Eber, J. et Heath, D. (1997). Thinking coherently. *Risk Magazine*, 10:68–71. 102

[7] Artzner, P., Delbaen, F., Eber, J. et Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9:203–228. 102, 104, 107

[8] Balkema, A. et de Haan, L. (1974). Residual life time at a great age. *The Annals of Probability*, 2(5):792–804. 20, 21

[9] Beirlant, J., Bouquiaux, C. et Werker, B. (2006). Semiparametric lower bounds for tail index estimation. *Journal of Statistical Planning and Inference*, 136:705–729. 47, 58

[10] Beirlant, J., Broniatowski, M., Teugels, J. et Vynckier, P. (1995). The mean residual life function at great age : Applications to tail estimation. *Journal of Statistical Planning and Inference*, 45:21–48. 47, 48, 58

[11] Beirlant, J., de Wet, T. et Goegebeur, Y. (2004a). Nonparametric estimation of extreme conditional quantiles. *Journal of statistical computation and simulation*, 74(8):567–580. 53

[12] Beirlant, J., Dierckx, G., Goegebeur, Y. et Matthys, G. (1999). Tail index estimation and an exponential regression model. *Extremes*, 2(2):177–200. 42, 43, 50, 63, 83, 154

[13] Beirlant, J., Dierckx, G., Guillou, A. et Staăricaă, C. (2002). On exponential representations of log-spacings of extreme order statistics. *Extremes*, 5(2):157–180. 42, 50

[14] Beirlant, J. et Goegebeur, Y. (2003). Regression with response distributions of pareto-type. *Computational statistics & data analysis*, 42(4):595–619. 52

[15] Beirlant, J. et Goegebeur, Y. (2004). Local polynomial maximum likelihood estimation for pareto-type distributions. *Journal of Multivariate Analysis*, 89(1):97–118. 52

[16] Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J., de Waal, D. et Ferro, C. (2004b). *Statistics of Extremes : Theory and Applications*. John Wiley & Sons. 4, 41, 42, 43, 52

[17] Beirlant, J., Schoutens, W. et Segers, J. (2004c). Mandelbrot's extremism. *Wilmott Magazine*, March:97–103. 22, 31

[18] Beirlant, J. et Teugels, J. (1992). Modelling large claims in non-life insurance. *Insurance : Mathematics and Economics*, 11(1):17–29. 4, 13, 27

[19] Beirlant, J., Teugels, J. et Vynckier, P. (1996). *Practical analysis of extreme values*. Leuven university press, Leuven. 28, 29, 45, 46, 47, 49, 58

[20] Berred, M. (1991). Record values and the estimation of the Weibull tail-coefficient. *Comptes-rendus de l'Académie des Sciences*, T. 312(Série I):943–946. 47, 58

[21] Bingham, N., Goldie, C. et Teugels, J. (1987). *Regular Variation*. Cambridge University Press. 22, 23, 24, 68, 111

[22] Bois, P., Obled, C., de Saintignon, M. et Mailloux, H. (1997). Atlas expérimental des risques de pluies intenses dans la région des Cévennes-Vivarais. *Pôle Grenoblois d'études et de recherche pour la prévention des risques naturels*, Grenoble 2éme édition. 118

[23] Bouleau, N. (1991). Splendeurs et misères des lois de valeurs extêmes. *Revue Risques*, 4:85–92. 4

[24] Brazaukas, V., Jones, B., Puri, L. et Zitikis, R. (2008). Estimating conditional tail expectation with actuarial applications in view. *Journal of Statistical Planning and Inference*, 128:3590–3604. 104

[25] Breiman, L., Stone, C. et Kooperberg, C. (1990). Robust confidence bounds for extreme upper quantiles. *Journal of Statistical Computation and Simulation*, 37(3-4):127–149. 35

[26] Brodin, E. et Rootzén, H. (2009). Univariate and bivariate GPD methods for predicting extreme wind storm losses. *Insurance : Mathematics and Economics*, 44(3): 345–356. 4, 13, 22

[27] Broniatowski, M. (1993). On the estimation of the Weibull tail coefficient. *Journal of Statistical Planning and Inference*, 35:349–366. 47, 58

[28] Caeiro, F. et Gomes, I. (2002). A class of asymptotically unbiased semi-parametric estimators of the tail index. *Test*, 11(2):345–364. 83, 154

[29] Cai, J. et Tan, K. (2007). Optimal retention for a stop-loss reinsurance under the VaR and CTE risk measures. *Astin Bulletin*, 37(1):93. 104, 107

[30] Cai, J.-J., Einmahl, J., de Haan, L. et Zhou, C. (2012). Estimation of the marginal expected shortfall : the mean when a related variable is extreme. *submitted*. 104

[31] Ceresetti, D., Ursu, E., Carreau, J., Anquetin, S., Creutin, J., Gardes, L., Girard, S. et Molinié, G. (2012). Evaluation of classical spatial-analysis schemes of extreme rainfall. *Natural Hazards and Earth System Sciences*, 12:3229–3240. 4, 132

[32] Chavez-Demoulin, V. et Davison, A. (2005). Generalized additive modelling of sample extremes. *Journal of the Royal Statistical Society : Series C (Applied Statistics)*, 54(1):207–222. 53

[33] Chen, D., Mao, T., Pan, X. et Hu, T. (2012). Extreme value behavior of aggregate dependent risks. *Insurance : Mathematics and Economics*, 50(1):99–108. 137

[34] Coles, S. (2001). *An introduction to statistical modeling of extreme values*. Springer. 4

[35] Coles, S. et Tawn, J. (1996). A bayesian analysis of extreme rainfall data. *Applied Statistics*, pages 463–478. 4, 117

[36] Collomb, G. (1980). Estimation non paramétrique de probabilités conditionnelles. *Comptes-rendus de l'Académie des Sciences*, 291:427–430. 109

[37] Cont, R. (2009a). Risques financiers : quelle modélisation mathématique ? *Pour la Science*, (375):24–27. 4

[38] Cont, R. (2009b). Les statistiques face aux événements extrêmes. *Pour La Science*, pages 24–27. 4

[39] Cooley, D., Nychka, D. et Naveau, P. (2007). Bayesian spatial modeling of extreme precipitation return levels. *Journal of the American Statistical Association*, 102(479):824–840. 117

[40] Csörgő, S. et Viharos, L. (1998). Estimating the tail index. *In* Szyszkowicz, B., éditeur : *Asymptotic Methods in Probability and Statistics*, pages 833–881. Test, North-Holland, Amsterdam. 42, 44

[41] Csorgo, S., Deheuvels, P. et Mason, D. (1985). Kernel estimates of the tail index of a distribution. *The Annals of Statistics*, 13(3):1050–1077. 44

[42] Csörgo, S. et Mason, D. (1985). Central limit theorems for sums of extreme values. *In Mathematical Proceedings of the Cambridge Philosophical Society*, volume 98, pages 547–558. Cambridge Univ Press. 42

[43] Daouia, A., Gardes, L. et Girard, S. (2013). On kernel smoothing for extremal quantile regression. *Bernoulli*. 53, 110, 136, 155

[44] Daouia, A., Gardes, L., Girard, S. et Lekina, A. (2011). Kernel estimators of extreme level curves. *Test*, 20(2):311–333. 4, 25, 50, 53, 106, 110, 116, 121, 123, 131, 144

[45] Davis, R. et Resnick, S. (1984). Tail estimates motivated by extreme value theory. *The Annals of Statistics*, pages 1467–1487. 42

[46] Davison, A. et Ramesh, N. (2000). Local likelihood smoothing of sample extremes. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, 62(1):191–208. 50, 52

[47] Davison, A. et Smith, R. (1990). Models for exceedances over high thresholds. *Journal of the Royal Statistical Society Series B*, 52(3):393–442. 36, 37, 52, 80

[48] de Haan, L. (1976). Sample extremes : an elementary introduction. *Statistica Neerlandica*, 30(4):161–172. 14

[49] de Haan, L. (1990). Fighting the arch–enemy with mathematics. *Statistica Neerlandica*, 44(2):45–68. 4, 27

[50] de Haan, L. et Ferreira, A. (2006). *Extreme Value Theory : An Introduction.* Springer Series in Operations Research and Financial Engineering, New York Inc. 12, 40, 41, 53, 83, 136, 154, 155

[51] De Haan, L. et Peng, L. (1998). Comparison of tail index estimators. *Statistica Neerlandica*, 52(1):60–70. 42

[52] de Haan, L. et Peng, L. (1998). Comparison of tail index estimators. *Statistica Neerlandica*, 52(1):60–70. 64

[53] de Haan, L. et Resnick, S. (1998). On asymptotic normality of the Hill estimator. *Stochastic Models*, 14(4):849–866. 42

[54] Deheuvels, P., Haeusler, E. et Mason, D. (1988). Almost sure convergence of the Hill estimator. *Mathematical Proceedings of the Cambridge Philosophical Society*, 104(2):371–381. 42

[55] Dekkers, A., Einmahl, J. et de Haan, L. (1989). A moment estimator for the index of an extreme-value distribution. *The Annals of Statistics*, 17(4):1833–1855. 39

[56] Deme, E., Gardes, L. et Girard, S. (2013a). On the estimation of the second order parameter for heavy-tailed distributions. *REVSTAT - Statistical Journal.* 42

[57] Deme, E., Girard, S. et Guillou, A. (2013b). Reduced-bias estimator of the conditional tail expectation of heavy-tailed distributions. *http ://hal.inria.fr/hal-00823260.* 105

[58] Diebolt, J., El-Aroui, M., Garrido, M. et Girard, S. (2005). Quasi-conjugate Bayes estimates for GPD parameters and application to heavy tails modeling. *Extremes*, 8:57–78. 80

[59] Diebolt, J., Gardes, L., Girard, S. et Guillou, A. (2008). Bias-reduced estimators of the Weibull tail-coefficient. *Test*, 17(2):311–331. 47, 50, 58, 83, 154

[60] Diebolt, J., Gardes, L., Girard, S. et Guillou, A. (2008b). Bias-reduced extreme quantiles estimators of Weibull tail-distributions. *Journal of Statistical Planning and Inference*, 138:1389–1401. 50, 80, 83, 154

[61] Diebolt, J., Guillou, A. et Rached, I. (2004). A new look at probability-weighted moments estimators. *Comptes Rendus de l'Académie des Sciences*, 338(Série I):629–634. 38

[62] Diebolt, J., Guillou, A. et Rached, I. (2007). Approximation of the distribution of excesses through a generalized probability-weighted moments method. *Journal of Statistical Planning and Inference*, 137(3):841–857. 38

[63] Dierckx, G., Beirlant, J., De Waal, D. et Guillou, A. (2009). A new estimation method for Weibull-type tails based on the mean excess function. *Journal of Statistical Planning and Inference*, 139:1905–1920. 47, 58

[64] Ditlevsen, O. (1994, Balkema, Rotterdam). Distribution arbitrariness in structural reliability. *Structural Safety and Reliability*, pages 1241–1247. 4, 13

[65] Dombry, C. (2013). Maximum likelihood estimators for the extreme value index based on the block maxima method. *http ://hal.archives-ouvertes.fr/hal-00780279*. 32

[66] Einmahl, J. et Magnus, J. (2008). Records in athletics through extreme-value theory. *Journal of the American Statistical Association*, 103(484):1382–1391. 4

[67] Einmahl, J. et Sander, S. (2011). Ultimate 100-m world records through extreme-value theory. *Statistica Neerlandica*, 65(1):32–42. 4

[68] El Methni, J., Gardes, L. et Girard, S. (2013). Nonparametric estimation of extreme risk measures from conditional heavy-tailed distributions. *http ://hal.archives-ouvertes.fr/hal-00830647*. 6, 106, 153

[69] El Methni, J., Gardes, L., Girard, S. et Guillou, A. (2012). Estimation of extreme quantiles from heavy and light tailed distributions. *Journal of Statistical Planning and Inference*, 142(10):2735–2747. 5, 66, 153

[70] Embrechts, P. (2000). *Extremes and integrated risk management*. Risk Books. 4, 105

[71] Embrechts, P., Klüppelberg, C. et Mikosch, T. (1997). *Modelling extremal events for insurance and finance*. Springer Verlag. 13, 14, 15, 40, 43, 61, 88, 105

[72] Embrechts, P., Resnick, S. et Samorodnitsky, G. (1998). Living on the edge. *Risk*, 11(1):96–100. 105

[73] Embrechts, P., Resnick, S. et Samorodnitsky, G. (1999). Extreme value theory as a risk management tool. *North American Actuarial Journal*, 3(2):30–41. 4, 105

[74] Falk, M. (1995). Some best parameter estimates for distributions with finite endpoint. *Statistics : A Journal of Theoretical and Applied Statistics*, 27:115–125. 26, 83, 154

[75] Falk, M. et Marohn, F. (1997). Efficient estimation of the shape parameter in Pareto models with partially known scale. *Statistics and Decisions-International Journal for Stochastic Methods and Models*, 15(3):219–230. 44

[76] Feuerverger, A. et Hall, P. (1999). Estimating a tail exponent by modelling departure from a Pareto distribution. *The Annals of Statistics*, 27(2):760–781. 42, 50, 83, 154

[77] Fisher, R. et Tippet, L. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proceedings of the Cambridge Philosophical Society*, 24:180–190. 14

[78] Fraga Alves, M., de Haan, L. et Neves, C. (2009). A test procedure for detecting super heavy tails. *Journal of Statistical Planning and Inference*, 139:213–227. 83, 154

[79] Fréchet, M. (1927). Sur la loi de probabilité de l'écart maximum. *Annales de la Société Polonaise de Mathématique*, 6:93–116. 15

[80] Galambos, J. (1987). *The asymptotic theory of extreme order statistics*. R.E. Krieger publishing company. 27

[81] Gardes, L. et Girard, S. (2005). Estimating extreme quantiles of Weibull tail-distributions. *Communication in Statistics - Theory and Methods*, 34:1065–1080. 48, 65

[82] Gardes, L. et Girard, S. (2006). Comparison of Weibull tail-coefficient estimators. *REVSTAT - Statistical Journal*, 4(2):163–188. 47, 49, 58, 64, 65

[83] Gardes, L. et Girard, S. (2008a). Estimation of the Weibull tail-coefficient with linear combination of upper order statistics. *Journal of Statistical Planning and Inference*, 138:1416–1427. 47, 48, 49, 58, 61

[84] Gardes, L. et Girard, S. (2008b). A moving window approach for nonparametric estimation of the conditional tail index. *Journal of Multivariate Analysis*, 99:2368–2388. 53

[85] Gardes, L. et Girard, S. (2010). Conditional extremes from heavy-tailed distributions : an application to the estimation of extreme rainfall return levels. *Extremes*, 13(2):177–204. 4, 25, 53, 106, 118

[86] Gardes, L. et Girard, S. (2012). Functional kernel estimators of large conditional quantiles. *Electronic Journal of Statistics*, 6:1715–1744. 53

[87] Gardes, L. et Girard, S. (2013). Estimation de quantiles extrêmes pour les lois à queue de type Weibull : une synthèse bibliographique. *Journal de la Société Française de Statistique*. 50

[88] Gardes, L., Girard, S. et Guillou, A. (2011). Weibull tail-distributions revisited : a new look at some tail estimators. *Journal of Statistical Planning and Inference*, 141(1):429–444. 5, 55, 57, 58, 59, 61, 62, 63, 64, 65, 71, 85, 94, 95, 153, 155

[89] Gardes, L., Girard, S. et Lekina, A. (2010). Functional nonparametric estimation of conditional extreme quantiles. *Journal of Multivariate Analysis*, 101:419–433. 25, 50, 53, 136, 155

[90] Gardes, L., Guillou, A. et Schorgen, A. (2012). Estimating the conditional tail index by integrating a kernel conditional quantile estimator. *Journal of Statistical Planning and Inference*, 142(6):1586–1598. 53

[91] Gardes, L. et Stupfler, G. (2013). Estimation of the conditional tail index using a smoothed local Hill estimator. *Extremes*. 53

[92] Geluk, J. et de Haan, L. (1987). Regular variation, extensions and tauberian theorems. *Center for Mathematics and Computer Science, Amsterdam*, 40:1–131. 42, 63

[93] Girard, S. (2004). A Hill type estimate of the Weibull tail-coefficient. *Communication in Statistics - Theory and Methods*, 33(2):205–234. 47, 58

[94] Girard, S., Guillou, A. et Stupfler, G. (2012a). Estimating an endpoint with high order moments. *Test*, 21:697–729. 136, 155

[95] Girard, S., Guillou, A. et Stupfler, G. (2012b). Estimating an endpoint with high order moments in the Weibull domain of attraction. *Statistics and Probability Letters*, 82:2136–2144. 26, 136, 155

[96] Girard, S., Guillou, A. et Stupfler, G. (2013). Frontier estimation with kernel regression on high order moments. *Journal of Multivariate Analysis*, 116:172–189. 136, 155

[97] Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *The Annals of Mathematics*, 44(3):423–453. 14, 24, 25

[98] Goegebeur, Y., Beirlant, J. et de Wet, T. (2010a). Generalized kernel estimators for the Weibull-tail coefficient. *Communications in Statistics - Theory and Methods*, 39(20):3695–3716. 28, 47, 53, 58

[99] Goegebeur, Y., Beirlant, J. et de Wet, T. (2010b). Kernel estimators for the second order parameter in extreme value statistics. *Journal of Statistical Planning and Inference*, 140:2632–2652. 53, 111

[100] Goegebeur, Y. et de Wet, T. (2012). Local estimation of the second order parameter in extreme value statistics and local unbiased estimation of the tail index. *Communications in Statistics-Theory and Methods*, 41:3575–3607. 53, 112

[101] Goegebeur, Y. et Guillou, A. (2010). Goodness-of-fit testing for Weibull-type behavior. *Journal of Statistical Planning and Inference*, 140:1417–1436. 47, 58

[102] Goegebeur, Y., Guillou, A. et Schorgen, A. (2013). Nonparametric regression estimation of conditional tails - the random covariate case. *Statistics*. 53

[103] Gomes, I., Caeiro, F. et Figueiredo, F. (2004). Bias reduction of a tail index estimator through an external estimation of the second-order parameter. *Statistics*, 38(6):497–510. 83, 154

[104] Gomes, M. et Martins, M. (2001). Generalizations of the Hill estimator–asymptotic versus finite sample behaviour. *Journal of Statistical Planning and Inference*, 93:161–180. 42

[105] Gomes, M., Martins, M. et Neves, M. (2000). Semi-parametric estimation of the second order parameter, asymptotic and finite sample behaviour. *Extremes*, 3:207–229. 42

[106] Guida, M. et Longo, M. (1988). Estimation of probability tails based on generalized extreme value distributions. *Reliability Engineering & System Safety*, 20(3):219–242. 30

[107] Guillou, A. et Willems, P. (2006). Application de la théorie des valeurs extrêmes en hydrologie. *Revue de statistique appliquée*, 54(2):5–32. 4

[108] Gumbel, E. (1954). *Statistical theory of extreme values and some practical applications : a series of lectures*. Numéro Applied Mathematics Series, 33. National Bureau of Standards, Washington. 4, 14, 27

[109] Gumbel, E. (1958). *Statistics of extremes*. Columbia University Press, Columbia. 4, 27, 31

[110] Hall, P.and Tajvidi, N. (2000). Nonparametric analysis of temporal trend when fitting parametric models to extreme-value data. *Statistical Science*, pages 153–167. 52

[111] Hall, P. (1978). Representations and limit theorems for extreme value distributions. *Journal of Applied Probability*, pages 639–644. 23, 52

[112] Hall, P. (1982a). On estimating the endpoint of a distribution. *The Annals of Statistics*, 10(2):556–568. 26

[113] Hall, P. (1982b). On some simple estimates of an exponent of regular variation. *Journal of the Royal Statistical Society. Series B (Methodological)*, 44:37–42. 42

[114] Harris, R. (1996). Gumbel re-visited - a new look at extreme value statistics applied to wind speeds. *Journal of Wind Engineering and Industrial Aerodynamics*, 59:1–22. 60

[115] Häusler, E. et Teugels, J. (1985). On asymptotic normality of Hill's estimator for the exponent of regular. *The Annals of Statistics*, 13:743–756. 42, 43, 63

[116] Hill, B. (1975). A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, 3(5):1163–1174. 29, 39, 40, 41, 43, 48, 63

[117] Hong, J. et Elshahat, A. (2010). Conditional tail variance and conditional tail skewness. *Journal of Financial and Economic Practice*, 10(1):147–156. 104, 107

[118] Hosking, J. (1985). Algorithm as 215 : Maximum-likelihood estimation of the parameters of the generalized extreme-value distribution. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 34(3):301–310. 32, 33

[119] Hosking, J. et Wallis, J. (1987). Parameter and quantile estimation for the generalized Pareto distribution. *Technometrics*, 29(3):339–349. 37, 38, 80

[120] Hosking, J., Wallis, J. et Wood, E. (1985). Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics*, 27(3):251–261. 32

[121] Hua, L. et Joe, H. (2011). Second order regular variation and conditional tail expectation of multiple risks. *Insurance : Mathematics and Economics*, 49(3):537–546. 138

[122] Jenkinson, A. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society*, 81(348):158–171. 14

[123] Jorion, P. (2007). *Value at risk : the new benchmark for managing financial risk.* McGraw-Hill New York. 101

[124] Koenker, R. et Bassett, G. (1978). Regression quantiles. *Econometrica : journal of the Econometric Society*, pages 33–50. 50, 52

[125] Kotz, S. et Nadarajah, S. (2000). *Extreme value distributions : theory and applications.* London : Imperial college press. 3

[126] Kratz, M. et Resnick, S. (1996). The qq-estimator and heavy tails. *Stochastic Models*, 12(4):699–724. 40, 45

[127] Krige, D. (1951). A statistical approach to some basic mine valuation problems on the witwatersrand. *Journal of the Chemical, Metallurgical and Mining Society of South Africa*, 52:119–139. 118, 132

[128] Landsman, Z. et Valdez, E. (2003). Tail conditional expectations for elliptical distributions. *North American Actuarial Journal*, 7:55–71. 104

[129] Macleod, A. (1989). A remark on algorithm as 215 : Maximum-likelihood estimation of the parameters of the generalized extreme-value distribution. *Applied Statistics*, 38(1):198–199. 32

[130] Mason, D. (1982). Laws of large numbers for sums of extreme values. *The Annals of Probability*, pages 754–764. 42

[131] Matthews, R. (1996). Far out forecasting. *The New Scientist*, 2051(October 12):36–40. 4

[132] Matthys, G. et Beirlant, J. (2003). Estimating the extreme value index and high quantiles with exponential regression models. *Statistica Sinica*, 13(3):853–880. 52

[133] McNeil, A., Frey, R. et Embrechts, P. (2005). *Quantitative risk management : concepts, techniques, and tools*. Princeton university press. 4, 105

[134] Molinié, G., Yates, E., Ceresetti, D., Anquetin, S., Boudevillain, B., Creutin, J. et Bois, P. (2012). Rainfall regimes in a mountainous mediterranean region : Statistical analysis at short time steps. *Journal of Applied Meteorology and Climatology*, 51:429–448. 132

[135] Morgan, J. (1997). Creditmetric. Rapport technique, JP Morgan, New York. 101, 107

[136] Necir, A., Rassoul, A. et Zitikis, R. (2010). Estimating the conditional tail expectation in the case of heavy-tailed losses. *Journal of Probability and Statistics*, ID 596839:17 pages. 105, 136, 154

[137] NERC (1975). Flood studies report. Rapport technique, London, Natural Environement Research Council. 4, 5, 80

[138] Pan, X., Leng, X. et Hu, T. (to appear, 2013). Second-order version of karamata's theorem with applications. *Statistics & Probability Letters*. 137

[139] Parzen, E. (1962). On estimation of a probability density function and mode. *The Annals of Mathematical Statistics*, 33:1065–1076. 109

[140] Peng, L. (1998). Asymptotically unbiased estimators for the extreme-value index. *Statistics & Probability Letters*, 38(2):107–115. 42

[141] Pickands, J. (1975). Statistical inference using extreme order statistics. *The Annals of Statistics*, 3(1):119–131. 20, 21

[142] Prescott, P. et Walden, A. (1980). Maximum likelihood estimation of the parameters of the generalized extreme-value distribution. *Biometrika*, 67(3):723–724. 32

[143] Prescott, P. et Walden, A. (1983). Maximum likelihood estimation of the parameters of the three-parameter generalized extreme-value distribution from censored samples. *Journal of Statistical Computation and Simulation*, 16(3-4):241–250. 32

[144] R Development Core Team (2008). *R : A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. 71

[145] Reiss, R.-D. (1989). *Approximate distributions of order statistics : with applications to nonparametric statistics*. Springer-Verlag New York. 3

[146] Reiss, R.-D. et Thomas, M. (2001). *Statistical analysis of extreme values : with applications to insurance, finance, hydrology and other fields*. Birkhäuser Verlag. 4

[147] Resnick, S. (1987). *Extreme Values, Regular Variation, and Point Processes*. Springer Verlag. 14, 23, 24, 25, 26, 137, 138

[148] Resnick, S. (1997). Discussion of the Danish data on large fire insurance losses. *ASTIN Bulletin*, 27:139–151. 4

[149] Ribatet, M. A. (2006). *A User's Guide to the POT Package (Version 1.0)*. 71

[150] Rockafellar, R. et Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk*, 2:21–42. 104, 107

[151] Rockafellar, R. et Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, 26:1443–1471. 104, 107

[152] Rootzén, H. et Tajvidi, N. (1997). Extreme value statistics and wind storm losses : a case study. *Scandinavian Actuarial Journal*, 1997(1):70–94. 4

[153] Rootzén, H. et Tajvidi, N. (2001). Can losses caused by wind storms be predicted from meteorological observations ? *Scandinavian Actuarial Journal*, 2001(2):162–175. 4, 13

[154] Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. *The Annals of Mathematical Statistics*, pages 832–837. 109

[155] Schultze, J. et Steinebach, J. (1996). On least squares estimators of an exponential tail coefficient. *Statistics and Decisions-International Journal for Stochastic Methods and Models*, 14(4):353–372. 40, 45

[156] Smith, J. (1991). Estimating the upper tail of flood frequency distributions. *Water Resources Research*, 23:1657–1666. 4, 13

[157] Smith, R. (1985). Maximum likelihood estimation in a class of nonregular cases. *Biometrika*, 72(1):67–90. 32, 37

[158] Smith, R. (1987). Estimating tails of probability distributions. *The Annals of Statistics*, 15(3):1174–1207. 37, 42

[159] Smith, R. (1989). Extreme value analysis of environmental time series : an application to trend detection in ground-level ozone. *Statistical Science*, 4(4):367–377. 52

[160] Taleb, N. (2007a). *The Black Swan : The Impact of the Highly Improbable.* Penguin Books. 4

[161] Taleb, N. (2007b). Black swans and the domains of statistics. *The American Statistician*, 61(3):198–200. 4

[162] Tasche, D. (2002). Expected shortfall and beyond. *Journal of Banking & Finance*, 26:1519–1533. 102, 104

[163] Valdez, E. (2005). Tail conditional variance for elliptically contoured distributions. *Belgian Actuarial Bulletin*, 5:26–36. 104, 107

[164] Von Mises, R. (1954). La distribution de la plus grande de n valeurs. *American Mathematical Society, Providence, RI*, II(In Selected Papers):271–294. 14, 26

[165] Wang, H., Li, D. et He, X. (2012). Estimation of high conditional quantiles for heavy-tailed distributions. *Journal of the American Statistical Association*, 107:1453–1464. 52

[166] Wang, H. et Tsai, C. (2009). Tail index regression. *Journal of the American Statistical Association*, 104(487):1233–1240. 52

[167] Weissman, I. (1978). Estimation of parameters and large quantiles based on the k largest observations. *Journal of the American Statistical Association*, 73(364):812–815. 29, 40, 46

[168] Zhou, C. (2009). Existence and consistency of the maximum likelihood estimator for the extreme value index. *Journal of Multivariate Analysis*, 100(4):794–815. 32

[169] Zhou, C. (2010). The extent of the maximum likelihood estimator for the extreme value index. *Journal of Multivariate Analysis*, 101(4):971–983. 32

[170] Zhu, L. et Li, H. (2012). Asymptotic analysis of multivariate tail conditional expectations. *North American Actuarial Journal*, 16(3):350–363. 138