

Université Mohamed Khider – Biskra



Faculté des Sciences et Sciences de l'Ingénieur
Département de Mathématiques

THESE

Présentée pour obtenir le Diplôme de Doctorat en Sciences en

Mathématiques

Option: Statistique

Par

Djamel MERAGHNI

Thème

Modelling Distribution Tails

Soutenue le

devant la commission d'examen

B. MEZERDI	Professeur	U. de Biskra	<i>Président</i>
A. NECIR	Professeur	U. de Biskra	<i>Rapporteur</i>
H. FELLAG	Professeur	U. de Tizi Ouzou	<i>Examineur</i>
Z. MOHDEB	Professeur	U. de Constantine	<i>Examineur</i>
E. OULD SAID	Maître de Conférences HDR	U. du Littoral, France	<i>Examineur</i>
S. BAHLALI	Maître de Conférences	U. de Biskra	<i>Examineur</i>
L. ABBAOUI	Professeur	U. de Sétif	<i>Invité</i>

Contents

List of Figures	v
List of Tables	ix
Introduction	1
1 Extreme Values	7
1.1 Sums of iid rv's	7
1.1.1 Laws of Large Numbers	8
1.1.2 Central Limit Theorem	10
1.2 Order Statistics	11
1.3 Fluctuations of Maxima	14
1.3.1 GEVD and Limit Distributions of $X_{n,n}$	14
1.3.2 Regular Variation	18
1.3.3 Domains of Attraction	22
1.3.4 GPD Approximation	28
1.4 Stochastic Processes	32
1.4.1 Introduction	32
1.4.2 Special Processes	34
2 Tail Index and High Quantile Estimation	37
2.1 Exploratory Data Analysis	38
2.1.1 Probability and Quantile Plots	38
2.1.2 Mean Excess Function Plot	39
2.1.3 Pareto Quantile Plot	41
2.1.4 Gumbel's Method of Excesses	41
2.1.5 Return Period and Return Level	42
2.2 EVT Procedure	43
2.2.1 Annual Maxima	43
2.2.2 ML Method	44
2.2.3 PWM Method	44
2.3 Semi-parametric Estimation	46
2.3.1 Pickands' Estimator	47
2.3.2 Hill's Estimator	50

2.3.3	Moment Estimator	55
2.3.4	Kernel Type Estimators	57
2.3.5	Bias Correction	58
2.3.6	Comparison of Estimators	60
2.4	Optimal Sample Fraction Selection	60
2.4.1	Graphical Method	61
2.4.2	Mean Squared Error	63
2.4.3	Adaptive Procedures	64
2.4.4	Illustrative Examples	69
2.5	POT Procedure	69
2.5.1	Fitting the GPD	71
2.5.2	Estimating GPD Parameters	71
2.5.3	Threshold Selection	74
2.6	Estimating High Quantiles	76
2.6.1	EVT-based Estimators	77
2.6.2	POT-based Estimator	80
3	Stable Law	81
3.1	Definition and Characterization	81
3.1.1	Stable rv's	81
3.1.2	Parameterization	83
3.1.3	Stable pdf and df	85
3.1.4	Heavy Tails	88
3.1.5	Moments	88
3.2	Generating Stable rv's	90
3.3	Parameter Estimation	94
3.3.1	Sample Quantile Methods	94
3.3.2	Sample cf Methods	96
3.3.3	ML Methods	98
3.3.4	Goodness of Fit	99
3.4	Modelling Financial Asset Returns	100
3.4.1	Asset Returns	102
3.4.2	Stable Fit vs. Normal Fit	104
3.5	Stable Processes	106
4	Estimating Some Actuarial Tools via the Extreme Value Approach	107
4.1	Risk Measures	108
4.1.1	Definitions	108
4.1.2	Estimations	110
4.2	Insurance Premium	111

4.2.1	PHP Principle	112
4.2.2	Estimating $\Pi_{\eta, R_{opt}}$	112
4.2.3	Confidence Bounds for $\Pi_{\rho, R_{opt}}$	121
4.2.4	Comparing $\widehat{\Pi}_{\eta, R_{opt}}^{(spar)}$ with $\widehat{\Pi}_{\eta, R_{opt}}^{(npar)}$	121
4.3	Stable Parameters	122
4.3.1	Estimating the Stability Index α	124
4.3.2	Estimating the Location Parameter μ	125
4.3.3	Estimating the Scale Parameter σ [98]	130
4.4	Concluding Remarks	140
	References	143
A	The Statistical Software <i>R</i>	157
B	Tables of McCulloch	159
C	Abbreviations and Notations	165

ملخص

إن ظهور قيم قصوى (عليا أو دنيا) في سلسلة من القياسات الخاصة بظاهرة ما، يعبر عن وقوع أحداث نادرة والتي بالرغم من احتمالها الضعيف، غالبا ما يكون لها انعكاسات سلبية على أصحاب القرار من أشخاص أو مؤسسات.

في هذه الرسالة، المتكونة من مقدمة و أربعة فصول، نهتم بإيجاد نموذجا إحصائيا لأذيال التوزيعات باستعمال نظرية القيم القصوى.

مقدمة الرسالة خصصت لنبذة تاريخية حول هذه النظرية و سرد لمختلف تطبيقاتها في مجالات المال والأعمال، الضمانات، الري، الأحوال الجوية، الاتصالات، ... في الفصل الأول نتطرق لبعض التعاريف و النتائج الخاصة بإحصائيات الترتيب و تغيرات الحدود القصوى وفي الفصل الثاني نقدم عرضا مفصلا لمختلف الطرق المقترحة لتقدير مؤشر القيم القصوى. أما الفصل الثالث فخصص لدراسة تفصيلية لإحدى التوزيعات ذات الأذيال الثقيلة المسماة التوزيعات المستقرة والتي تستعمل كنموذج بديل لتوزيعات "غوص" عند وصف المعطيات ذات القيم القصوى خاصة المعطيات المالية. أخيرا و في الفصل الرابع نستعمل نظرية القيم القصوى لاقتراح مقدرات لبعض الأدوات الخاصة بمجالي المالية و الضمانات مع الإشارة إلى أن بعض نتائج هذا الفصل ستنتشر قريبا في مجلات علمية ذات شهرة عالمية.

RESUME

L'apparition de valeurs extrêmes (aussi bien supérieures qu'inférieures) dans une série d'observations relatives à un certain phénomène témoigne de l'occurrence d'événements rares, qui malgré leur faible probabilité ont des répercussions (souvent négatives) sur les décideurs (individus ou institutions).

Cette thèse, composée d'une introduction et de quatre chapitres, porte sur le comportement asymptotique des distributions statistiques. On y propose la modélisation des queues de distributions par le biais de la théorie des valeurs extrêmes qui permet d'extrapoler le comportement des queues à partir des plus grandes (ou plus petites) valeurs observées.

Dans l'introduction, un historique sur cette théorie est donné ainsi qu'un inventaire de ses applications dans des domaines aussi variés que la finance, assurance, hydrologie, météorologie, télécommunications, physique,...

Le premier chapitre est une sorte de rappel des définitions de base et des résultats les plus importants sur les statistiques d'ordre et les fluctuations des maxima.

Au deuxième chapitre, on passe en revue les différentes méthodes d'estimation de l'indice des valeurs extrêmes et des quantiles extrêmes. On distingue deux approches d'estimation : la première, appelée « EVT procedure », est basée sur la distribution généralisée des valeurs extrêmes (GEVD) et la deuxième, appelée « POT procedure », est basée sur la distribution généralisée de Pareto (GPD).

Une étude détaillée d'une distribution à queues lourdes spéciale est présentée au troisième chapitre. Il s'agit de la distribution Lévy-stable qui joue un rôle de plus en plus important en finance et en télécommunications (entre autres). Fort du théorème central limite généralisé, le modèle stable est présenté comme alternative à celui de Gauss (utilisé à tort) pour la description des données dans plusieurs domaines de la vie socio-économique.

Enfin, le quatrième chapitre est consacré à l'estimation de quelques outils d'actuariat par l'approche des valeurs extrêmes. Quelques résultats du chapitre (estimation semi-paramétrique de la prime d'assurance et l'estimation du paramètre d'échelle de la loi Lévy-stable) sont originaux et vont apparaître très prochainement dans des revues internationales de renommée bien établie.