République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
UNIVERSITÉ KASDI MERBAH OUARGLA

FACULTÉ DE MATHÉMATIQUES ET SCIENCES DE LA MATIÈRE

DÉPARTEMENT DE MATHÉMATIQUES

Mémoire présenté En Vue De L'obtention Du

DIPLÔME DE MASTER

EN MATHÉMATIQUES

Option : Probabilité et Statistique

Par

Chaima BEKHDIDJA

Intitulé

Réduction de la variance de l'indice des valeurs extrêmes par la méthode de Monte-Carlo

Membres du jury

Abdelkader AMARA	M. C. B	UKMO	Président
Aissa BAHADI	Professeur	UKMO	Examinateur
Fatima MEDDI	M. C. B	UKMO	Rapporteur

Septembre 2018

Dédicaces

Je dédie ce modeste travail

A ma très cher Mère et mon très cher Père
A mes chers sœurs et frères
A mon fiancé

A ceux qui ont veillé pour mon bien être

A ceux qui m'ont toujours encouragé pour que je réussisse dans mes études
A tout ce qui m'ont encouragé lors de la réalisation de mon travail

Sans oublier de dédier ce mémoire à mes très chères amies intimes

Finalement à tous ceux qui m'on aider de proche ou de loin

REMERCIEMENTS

Je remercie Dieu le tout-puissant de m'avoir donné la volonté, la force et le courage pour bien mener et finir mon travail de thèse.

Je voudrais d'abord et avant tout remercier mon encadreur MEDDI FATIMA pour tous les efforts en vue d'établir ce mémoire, elle a eu le rôle fondamental et essentiel et la grande mérite dans tout ce qui a été réalisé, comme cela avait toujours été, près de moi et me guider et corriger mes erreurs et me donner de précieux conseils et les conseils appropriés et les alertes considérés, je la répète mes remerciements pour tout ce qu'elle a fait pour moi à travers la mise en plan à toutes les étapes de la préparation de ce mémoire depuis le début jusqu'à la fin, étape par étape, était que le premier et dernier facteur pour le succès de ce travail, et je lui dis encore une fois merci beaucoup à pour votre appréciation profonde.

Je tient a remercier les membres de jury Monsieur le Professeur Aissa BAHADI pour l'intérêt qui l'apporter mon travail ainsi, Monsieur Abdekader AMARA d'avoir accepter de présider le jury et donner de son temps pour en discuter.

J'exprime ma gratitude à ma famille qui m'a toujours soutenue et encouragée dans la voie que je m'étais fixée. Je remercie particulièrement mes parents qui m'ont stimulée et encouragé pendant mes études. qui étaient toujours prêts à fournir tous les moyens physique et morale pour la réussite de ce projet.

Table des matières

0.1 Abréviations et Notations

TVE	Théorie des valeurs extrêmes.
γ	Indice des valeurs extrêmes.
$\hat{\gamma}^H_{k,n}$	Estimateur de Hill.
iid	Indépendantes et identiquement distribuées.
$v.a$	Variable aléatoire.
$\mathbb{E}$	Espérance mathématique.
Var	Variance mathématique.
$cov(X,Y)$	La covariance entre X et Y.
$\mathbb{P}(A)$	La probabilité de l'évènement A.
s.o	Statistique d'ordre.
$X_{(1)},...X_{(n)}$	Statistiques d'ordre associées à $X_1,...X_n$.
x_F	Point terminal.
$M_n = X_{(n)}$	Maximum de $X_1,...X_n$.
F	Fonction de répartition.
F_n	Fonction de répartition empirique.
$F^{\leftarrow}$	Inverse généralisée de F.
$\bar{F}$	$(1-F)$ Fonction de survie.
G_γ	Famille de la loi des valeurs extrêmes.
Φ	Loi de Gumbel.
Λ	Loi de Fréchet.
Ψ	Loi de Weibull.
$\mathcal{RV}_\alpha$	Variation régulière à l'infini avec l'indice α.
$\mathcal{RV}_0$	Variation lente.
$l(x)$	Fonction à variation lente.
$\mathcal{D}(Frchet)$	Domaine d'attraction de Fréchet.
$\mathcal{D}(Gumbel)$	Domaine d'attraction de Gumbel.
$\mathcal{D}(Weibll)$	Domaine d'attraction de Weibull.
$\mathcal{N}(\mu,\sigma^2)$	Loi normal de moyenne μ et de variance σ^2.
$\mathbb{R}$	Ensemble des réels et $\mathbb{R}^p = \underbrace{\mathbb{R}\times ... \times\mathbb{R}}_{\text{p fois}}$, $p\in\mathbb{N}^*$.
$\mathbb{I}_A$	Fonction indicatrice de l'événement A.

0.2 Introduction

Les méthodes de Monte-carlo ont été utilisées pendant des siècles, mais seulement au cours des dernières décennies, la technique a acquis le statut d'une méthode numérique à part entière capable de traiter les applications les plus complexes. **Stanislaw Ulam** a inventé la méthode en 1946 tout en réfléchissant aux probabilités de gagner un jeu de cartes de solitaire, cette méthode inspirée son nom allant l'intérêt de Ulam par le jeu de Poker ; la capitale de Monaco «Monte-carlo», était une référence connue pour le jeu.
la méthode de Monte-carlo contient quatre techniques principales pour la réduction de la variance :

- variable antithétique.
- stratification.
- échantillonnage préférentiel.
- variable de contrôle.

Le but de ces techniques est de trouver un estimateur d'une statistique telle que son espérance est le même que la statistique, mais sa variance est très faible que la variance de celle-ci. La différence entre les quatre technique est la forme de l'estimateur qu'on cherche.

Les extrêmes sont les évènements dont la probabilité d'apparition est trop faible, ils apparaissent en général dans les contextes physiques nombreux et variés en particulier les catastrophes naturelles. La théorie des valeurs extrêmes est une branche de la statistique qui essaie d'amener une solution face à ces phénomènes. Elle se repose principalement sur des distributions limites des extrêmes et leurs domaines d'attraction. Fisher et Tippet (1928) ils étudiaient la résistance des fils de coton puis plus tard Gnedenko (1943) s'est intéressé à ces distributions. Ils ont énoncé un théorème fondamental avec la création de trois domaines d'attraction, domaine d'attraction de Fréchet, Gumbel et Weibull. Ce théorème intéressant fait référence à un paramètre appelé l'indice de queue qui donne la forme de la queue de distribution. Von Mises (1954) puis Jenkinson (1955) ont rassemblé les distributions de ces trois domaines en une seule écriture. C'est en ce moment que plusieurs auteurs se sont focalises aux estimations de l'indice des valeurs extrêmes. Nous pouvons citer Hill (1975) dans le cas où l'indice est positif.

Dans ce mémoire, on utilise la méthode de Monte-carlo pour réduire la variance de l'estimateur de Hill(1975).

Nous présentons dans le premier chapitre la méthode de Monte-carlo ; en débutant par l'historique de la méthode. Ensuite, nous détaillons les quatre techniques de la réduction de la variance.

Notre principal contribution est proposée au deuxième chapitre, ou nous appliquerons la méthode de Monte-carlo pour la réduction de la variance de l'estimateur de Hill (1975) $\gamma_{k,n}^{\hat{H}}$.

Au préalablement, nous rappellerons quelques notions essentielles sur la théorie des valeurs extrêmes, nous définissons rapidement les notions de domaine d'attraction, fonctions à variations régulières puis nous présentons l'estimateur classique de l'indice de queue de Hill (1975) $\gamma_{k,n}^{\hat{H}}$, ainsi que ces propriétés asymptotiques.

Dans le troisième chapitre, nous présenterons des algorithmes correspondantes a chaque méthode.

Chapitre 1

Présentation de la Méthode de Monte-Carlo

1.1 Historique

Les méthodes numériques connues sous le nom de méthodes de Monte Carlo peuvent être décrites de manière approximative comme des méthodes de simulation statistique, où la simulation statistique est définie en termes assez généraux pour être une méthode qui utilise des séquences de nombres aléatoires pour effectuer la simulation. Les méthodes de Monte Carlo ont été utilisées pendant des siècles, mais seulement au cours des dernières décennies, la technique a acquis le statut d'une méthode numérique à part entière capable de traiter les applications les plus complexes. Stanislaw Ulam, un mathématicien né en Pologne qui a travaillé pour John von Neumann sur Manhattan Project aux Etats-Unis pendant la Seconde Guerre mondiale, Il a inventé la méthode Monte Carlo en 1946 tout en réfléchissant aux probabilités de gagner un jeu de cartes de solitaire. Cité dans Eckhardt (1987), Ulam décrit l'incident comme suit : «Les premières pensées et tentatives que j'ai faites pour pratiquer la Méthode Monte Carlo ont été suggérées par une question qui m'est venue à l'esprit en 1946 alors que j'étais en convalescence d'une maladie et que je jouais des solitaires. Après avoir passé beaucoup de temps à essayer de les estimer par de purs calculs combinatoires, je me suis demandé si une méthode plus pratique que la «pensée abstraite» ne serait pas de dire cent fois et simplement observer et compter le nombre de jeux réussis. C'était déjà possible d'envisager avec le début de la nouvelle ère des ordinateurs rapides, et j'ai immédiatement pensé aux problèmes de diffusion des neutrons et d'autres questions de physique mathématique, et plus généralement comment changer les processus décrits par certaines équations différentielles en une forme équivalente interprétable comme une succession d'opérations aléatoires Plus tard ... [en 1946, I] Desc a renvoyé l'idée à John von Neumann, et nous avons commencé à planifier des calculs réels. » [1]. Le nom "Monte Carlo" a été inventé par Nicholas Metropolis

1. Eckhardt, Stan Ulam, John Von Neumann et la méthode de Monte-carlo, Los Alamos science,1987

inspiré dans l'intérêt de Stanislaw Ulam, son collègue de Manhattan Project à Los Alamos, dans le jeu de poker Monte Carlo, la capitale de Monaco, était une référence connue pour le jeu. Le premier article de Monte Carlo, "The Monte Carlo Method" de Metropolis Ulam, a été publié en 1949 dans le Journal of the American Statistical Association. (pour plus détails Voir : Eckhardt, Roger (1987)." Stan Ulam, John von Neumann, et la méthode de Monte Carlo ", Los Alamos Science, numéro spécial (15), 131-137).

La méthode de Monte Carlo est maintenant utilisé couramment dans de nombreux domaines, de la simulation de phénomènes physiques complexes tels que le transport de rayonnement dans l'atmosphère terrestre, dans la simulation des processus inucléaires ésotériques, dans les expériences de physique des hautes énergies...

1.2 Principe de la méthode

Soient $X_1, ..., X_n$ une suite de variables aléatoires de même loi que X et g une fonction de v.a mesurable et doublement intégrable. On définit I par :

$$I = \mathbb{E}(g(X)),$$

l'estimateur empirique de I est :

$$\hat{I}_n = \frac{1}{n} \sum_{i=1}^{n} g(X_i),$$

est appelé l'estimateur de Monte-carlo, est un estimateur sans biais : $\mathbb{E}(\hat{I}_n) = I$, et si n assez grand $\hat{I}_n$ converge vert I :

$$\hat{I}_n \xrightarrow[n \to +\infty]{} I$$

Théorème 1.2.1 (loi des grands nombres) *soit* $(X_i)_{i \in \mathbb{N}}$ *suite de variables aléatoires iid et* $\mathbb{E}(\mid X \mid) < +\infty$ *alors, on a presque surement :*

$$\lim_{n \longrightarrow +\infty} \bar{X}_n = \lim_{n \longrightarrow +\infty} \frac{1}{n}(x_1 + x_2 + ... + x_n) = \mathbb{E}(X)$$

Autrement dit, si n assez grand on a :

$$\frac{1}{n} \sum_{i=1}^{n} x_i \approx \mathbb{E}(X).$$

Nous avons

$$\begin{aligned}
\mathbb{E}[(\hat{I}_n - I)^2] &= \mathbb{E}[(\hat{I}_n - \mathbb{E}(\hat{I}))^2] \\
&= Var(\hat{I}_n) \\
&= Var\left(\frac{1}{n}\sum_{i=1}^{n} g(X_i)\right) \\
&= \frac{1}{n^2}\sum_{i=1}^{n} Var(g(X_i)) \\
&= \frac{1}{n^2}\left(nVar(g(X))\right) \\
&= \frac{1}{n}Var(g(X)) \qquad (1.1)
\end{aligned}$$

représente l'erreur commise. La méthode de Monte-carlo consiste à améliorer cet estimateur, i.e. réduction d'erreur d'estimation. Cette méthode consiste à réduire la variance $\hat{\sigma}^2$ de $\hat{I}$. Le principe de la méthode de Monte Carlo consiste à trouver une statistique noté Y, telle que :

$$\mathbb{E}(Y) = \mathbb{E}(g(X)) \qquad (1.2)$$

avec

$$Var(Y) \leq Var(g(X)) \qquad (1.3)$$

ce qui une statistique (Y) de même espérance que $g(X)$ mais avec une variance réduite.

La méthode de Monte-carlo contient quatre techniques principales pour la réduction de la variance :

1. variable antithétique.
2. stratification.
3. échantillonnage préférentiel.
4. variable de contrôle.

1.2.1 Variable antithétique

Soit X une variable aléatoire de densité f et g une fonction de variable aléatoire mesurable et doublement intégrable. On définit la $v.a$ Y comme suit :

$$Y = \frac{g(X) + g \circ A(X)}{2}$$

et $Y_1, Y_2, ..$ suite de $v.a$ de même loi que Y avec A une transformation mesurable $A : \mathbb{R}^n \mapsto \mathbb{R}^n$ telle que $A(X)$ est de même loi que X.

les variables $g(X)$ et Y sont de même espérance :

$$\begin{aligned}
\mathbb{E}(Y) &= \mathbb{E}\left[\frac{g(X) + g \circ A(X)}{2}\right] \\
&= \frac{\mathbb{E}(g(X)) + \mathbb{E}(g \circ A(X))}{2} \\
&= \frac{2\mathbb{E}(g(X))}{2} \\
&= \mathbb{E}(g(X))
\end{aligned}$$

On va estimer $I = \mathbb{E}(g(X))$ et $I_a = \mathbb{E}(Y)$ pour un échantillon de taille n ; on trouve :

$$\hat{I}_n = \frac{1}{n}\sum_{i=1}^{n} g(X_i)$$

et

$$\begin{aligned}
\hat{I}_{n,a} &= \frac{1}{n}\sum_{i=1}^{n} Y_i \\
&= \frac{1}{n}\sum_{i=1}^{n} \frac{g(X_i) + g \circ A(X_i)}{2}
\end{aligned}$$

Nous avons d'après (1.1) :

$$Var(\hat{I}_n) = \frac{1}{n} Var(g(X)) \tag{1.4}$$

calcule de la variance de Y :

$$\begin{aligned}
Var(\hat{I}_{n,a}) &= Var\left(\frac{1}{n}\sum_{i=1}^{n} \frac{g(X_i) + g \circ A(X_i)}{2}\right) \\
&= \frac{1}{n^2}\sum_{i=1}^{n} \frac{Var(g(X_i) + g \circ A(X_i))}{4} \\
&= \frac{1}{4n^2}\sum_{i=1}^{n} Var(g(X_i)) + Var(g \circ A(X_i)) + 2cov(g(X), g \circ A(X)) \\
&= \frac{1}{4n^2}\sum_{i=1}^{n} 2Var(g(X)) + 2cov(g(X), g \circ A(X)) \\
&= \frac{1}{4n^2}\left[2nVar(g(X)) + 2ncov(g(x), g \circ A(X))\right] \\
&= \frac{1}{2n} Var(g(X)) + \frac{1}{2n} cov(g(x), g \circ A(X)),
\end{aligned}$$

D'après le principe de Monte-Carlo on a :

$$Var(\hat{I}_{n,a}) = \frac{1}{n} Var(Y),$$

alors,

$$\frac{1}{n} Var(Y) = \frac{1}{2n} Var(g(X)) + \frac{1}{2n} cov\left(g(x), g \circ A(X)\right)$$

$$Var(Y) = \frac{1}{2}\left[Var(g(X)) + cov(g(x), g \circ A(X))\right] \tag{1.5}$$

On en déduit les résultats suivants :

1. Condition suffisante

$$cov(g(X), g \circ A(X)) \leq 0 \Longrightarrow Var(Y) < Var(g(X)).$$

2. Condition nécessaire et suffisante

$$cov(g(X), g \circ A(X)) \leq 0 \Longleftrightarrow Var(Y) < Var(g(X)).$$

Proposition 1.2.1 *soient X une variable aléatoire réelle de loi p et A une transformation décroissante de $\mathbb{R}$ telle que $A(X)$ soit de loi p. si la fonction $g : \mathbb{R} \to \mathbb{R}$ est monotone, alors $cov(g(X), g \circ A(X)) \geq 0$, avec égalité uniquement si g est presque sûrement constante.* [2]

1.2.2 Stratification

Soit X une $v.a$ de densité f, $g : \mathbb{R}^n \longmapsto \mathbb{R}$ une fonction de variable aléatoire mesurable et doublement intégrable, D l'ensemble des valeurs prises par X et $(X_i)_{1\leq i\leq n}$ une suite de $v.a$ iid de même loi que X. On définit I par :

$$I = \int_{\mathbb{R}^n} g(X) f(X) dX = \mathbb{E}(g(X))$$

Remarque : dans cet technique, la variable Y n'est pas définit par une expression explicite. On va décomposer l'expression de $\mathbb{E}(g(X))$ et on cherche a minimiser la variance de $g(X)$.

Supposons que D est partitionné en $D_1, ..., D_m$ avec $(D = \cup_{1\leq i\leq m} D_i)$, on définit I_i et P_i comme suit :

$$I_i = \mathbb{E}(g(X) \mid X \in D_i)$$

$$P_i = P(X \in D_i)\ ,\ \sum_{i=1}^{m} P_i = 1$$

2. Julien STOEHR, Méthodes deMonte Carlo (cours), université DAUPHINE, Paris 2017-2018

alors, I est décomposé sous la forme suivante :

$$I = \sum_{i=1}^{m} P_i I_i.$$

Nous pouvons approcher I_i par une méthode de Monte-Carlo a n_i tirages indépendants :

$$I_i \approx \hat{I}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} g(X_j^{(i)}) \qquad \text{avec,} \qquad \sum_{i=1}^{m} n_i = n.$$

Chacune de ces approximations a un coût numérique n_i (puisqu'on fait n_i boucles pour calculer la somme ci-dessus). Nous en déduisons une deuxième approximation :

$$I \approx \hat{I} = \sum_{i=1}^{m} P_i \hat{I}_i.$$

L'erreur commise se décompose dans ce cas en une somme d'erreur :

$$I - \sum_{i=1}^{m} P_i \hat{I}_i \quad = \quad \sum_{i=1}^{m} P_i (I_i - \hat{I}_i).$$

Dans cette somme, chacun des termes est (approximativement) de loi $\mathcal{N}(0, p_i^2 \sigma_i^2 / n_i)$ avec :

$$\begin{aligned}
\sigma_i^2 &= Var(g(X)|X_i \in D_i) \\
&= \mathbb{E}\left[(g(X) - \mathbb{E}(g(X)|X \in D_i))^2 | X \in D_i\right] \\
&= \frac{\mathbb{E}\left[(g(X) - \mathbb{E}(g(X)|X \in D_i))^2\right] \mathbb{I}_{\{X \in D_i\}}}{P_i} \\
&= P_i^{-1}\mathbb{E}[g(X)^2 \mathbb{I}_{\{X \in D_i\}}] + P_i^{-1}\mathbb{E}(g(X)|X \in D_i)^2 \mathbb{E}(\mathbb{I}_{\{X \in D_i\}}) \\
&\quad -2P_i^{-1}\mathbb{E}(g(X)|X \in D_i)\mathbb{E}(g(X)\mathbb{I}_{\{X \in D_i\}}) \\
&= P_i^{-1}\mathbb{E}[g(X)^2 \mathbb{I}_{\{X \in D_i\}}] - P_i^{-2}\mathbb{E}(g(X)\mathbb{I}_{\{X \in D_i\}})^2 \qquad (1.6)
\end{aligned}$$

Nous voulons que cet erreur soit la plus petite possible donc nous voulons minimiser sa variance :

$$\begin{aligned}
Var(I - \hat{I}) &= Var\left[\sum_{i=1}^{m} P_i (I_i - \hat{I}_i)\right] \\
&= \sum_{i=1}^{m} Var\left[P_i (I_i - \hat{I}_i)\right] \\
&= \sum_{i=1}^{m} \frac{P_i^2 \sigma_i^2}{n_i},
\end{aligned}$$

Si nous prenons $n_i = n\frac{P_i\sigma_i}{\sum_{i=1}^m P_j\sigma_j}$ alors,

$$\begin{aligned}
\sum_{i=1}^m \frac{P_i^2\sigma_i^2}{n_i} &\geq \sum_{i=1}^m P_i^2\sigma_i^2 \frac{\sum_{j=1}^m P_j\sigma_j}{nP_i\sigma_i} \\
&\geq \frac{1}{n}(\sum_{j=1}^m P_i\sigma_i)^2 \\
Var(I-\hat{I}) &\geq \frac{1}{n}(\sum_{j=1}^m P_i\sigma_i)^2.
\end{aligned}$$

On calcule maintenant la variance de $g(X)$:

$$\begin{aligned}
Vra(g(X)) &= \mathbb{E}\left[\left(g(X)\mathbb{I}_{\{X\in D_i\}}\right)^2\right] - \mathbb{E}\left[\sum_{i=1}^m g(X)\right]^2 \\
&= \mathbb{E}\left[\sum_{i=1}^m g(X)^2\mathbb{I}_{\{X\in D_i\}}\right] - \left[\sum_{i=1}^m P_i\mathbb{E}(g(X)|X\in D_i)\right]^2 \\
&= \sum_{i=1}^m \{P_i\mathbb{E}(g(X)^2|X\in D_i) - P_i\mathbb{E}(g(X)|X\in D_i)^2\} \\
&\quad + \sum_{i=1}^m P_i\mathbb{E}(g(X)|X\in D_i)^2 - \mathbb{E}\left[\sum_{i=1}^m P_i\mathbb{E}(g(X)|X\in D_i)\right]^2
\end{aligned}$$

D'après l'inégalité de convexité pour la fonction carré (pour $x_1,...,x_m \in \mathbb{R}$ et $\lambda_1,...,\lambda_m \in \mathbb{R}_+$ tel que $\sum_{i=1}^m \lambda_i = 1$ alors $\sum_{i=1}^m \lambda_i x_i^2 \geq (\sum_{i=1}^m \lambda_i x_i)^2$) [3], on a :

$$\sum_{i=1}^m P_i\mathbb{E}(g(X)|X\in D_i)^2 \geq \mathbb{E}\left[\sum_{i=1}^m P_i\mathbb{E}(g(X)|X\in D_i)\right]^2$$

$$\sum_{i=1}^m P_i\mathbb{E}(g(X)|X\in D_i)^2 - \mathbb{E}\left[\sum_{i=1}^m P_i\mathbb{E}(g(X)|X\in D_i)\right]^2 \geq 0$$

3. Sylvain Rubenthaler, Méthodes de Monte-Carlo (Cours et exercices), Université Nice Sophia Antipolis 2017-2018

alors,

$$\begin{aligned}
Var(g(X)) &\geq \sum_{i=1}^{m}\{P_i\mathbb{E}(g(X)^2|X\in D_i)-P_i\mathbb{E}(g(X)|X\in D_i)^2\} \\
&\geq \left(\sum_{i=1}^{m}P_i\left[\mathbb{E}(g(X)^2|X\in D_i)-\mathbb{E}(g(X)|X\in D_i)^2\right]^{1/2}\right)^2 \\
&= \left(\sum_{i=1}^{m}P_i\sigma_i\right)^2.
\end{aligned}$$

On conclue que la variance de $g(X)$ est bien réduite.

1.2.3 Échantillonnage préférentiel

Soit X une $v.a$ de densité f, $g:\mathbb{R}^n\longmapsto\mathbb{R}$ une fonction de variable aléatoire mesurable et doublement intégrable, $h>0$ une autre densité équivalente a f. La densité h est appelée loi instrumentale (ou loi d'importance) et le rapport $\dfrac{f(X)}{h(X)}$ est appelé poids d'importance. Nous cherchons un estimateur appelé Y de $g(X)$ tel que

$$\mathbb{E}(Y)=\mathbb{E}(g(X))$$

et

$$Var(Y)<Var(g(X)).$$

Supposons que Y est définit par :

$$Y=\frac{g(X)f(X)}{h(X)}$$

on vérifier que $\mathbb{E}(Y)=\mathbb{E}(g(X))$:

$$\begin{aligned}
\mathbb{E}(g(X)) &= \int_{\mathbb{R}^n}g(x)f(x)dx \\
&= \int_{\mathbb{R}^n}\frac{g(x)f(x)}{h(x)}h(x)dx \\
&= \mathbb{E}\left(\frac{g(X)f(X)}{h(X)}\right) \\
&= \mathbb{E}(Y)
\end{aligned}$$

Comparaison des variances :
On a

$$\begin{aligned} Var(g(X)) &= \mathbb{E}(g(X)^2) - \mathbb{E}(g(X))^2 \\ &= \int g^2(X)f(X)dX - \mathbb{E}_f\left[g(X)\right]^2 \end{aligned}$$

$$\begin{aligned} Var(Y) &= \mathbb{E}(Y^2) - \mathbb{E}(Y)^2 \\ &= \mathbb{E}_h\left[\frac{g^2(X)f^2(X)}{h^2(X)}\right] - \mathbb{E}_h\left[\frac{g(X)f(X)}{h(X)}\right]^2 \\ &= \int g^2(X)\frac{f^2(X)}{h^2(X)}h(X)dX - \left(\int \frac{g(X)f(X)}{h(X)}h(X)\right)^2 \\ &= \int g^2(X)\frac{f^2(X)}{h(X)}dX - \left(\int g(X)f(X)\right)^2 \\ &= \mathbb{E}_f\left[\frac{g^2(X)f(X)}{h(X)}\right] - \mathbb{E}_f\left[g(X)\right]^2 \end{aligned}$$

$$Var(Y) = \mathbb{E}_f\left[\frac{g^2(X)f(X)}{h(X)}\right] - \mathbb{E}_f\left[g(X)\right]^2$$

alors

$$\begin{aligned} Var(g(X)) - Var(Y) &= \mathbb{E}(g(X)^2) - \mathbb{E}_f\left[\frac{g^2(X)f(X)}{h(X)}\right] \\ &= \int g^2(X)f(X)dX - \int g^2(X)\frac{f^2(X)}{h(X)}dX \\ &= \int g^2\left(1 - \frac{f(X)}{h(X)}\right)f(X)dX. \end{aligned}$$

$$Var(g(X)) - Var(Y) = \int g^2\left(1 - \frac{f(X)}{h(X)}\right)f(X)dX$$

Conclusion

$$\int g^2\left(1 - \frac{f(X)}{h(X)}\right)f(X)dX > 0 \Longrightarrow Var(Y) < Var(g(X)).$$

On peut minimiser la variance de Y à l'aide du Proposition suivante :

Proposition 1.2.2 *L'estimateur d'échantillonnage préférentiel de variance minimale de* Y *,est obtenu pour la densité instrumentale*[4] *:*

$$h^*(X) = \frac{1}{\int_{\mathbb{R}^n} \mid g(X) \mid f(X)dX} \mid g(X) \mid f(X) \tag{1.7}$$

1.2.4 Variable de contrôle

Soit $(X_i)_{i\leq n}$ une suite de $v.a$'s iid de même loi que X, g une fonction de variable aléatoire mesurable et doublement intégrable, Z une autre $v.a$ tel que $\mathbb{E}(Z)$ connue ou facilement calculable, appelé variable de contrôle, $Z_1, Z_2, ...$ une suite de $v.a$ de même loi que Z. On définit la variable Y sous la forme suivante :

$$Y = g(X) - c(Z - \mathbb{E}(Z)) \text{ (c : paramètre constant fixé),}$$

Vérifiant facilement que $\mathbb{E}(Y) = \mathbb{E}(g(X))$,
l'estimateur empirique de $I_c = \mathbb{E}(Y)$ est donné par :

$$\begin{aligned} \hat{I}_c &= \frac{1}{n}\sum_{i=1}^{n} Y_i \\ &= \frac{1}{n}\sum_{i=1}^{n} g(X_i) - c(Z_i - \mathbb{E}(Z_i)), \end{aligned}$$

et l'estimateur empirique de $I = \mathbb{E}(g(X))$ est

$$\hat{I} = \frac{1}{n}\sum_{i=1}^{n} g(X_i).$$

La variance de $\hat{I}$ est donné par :

$$Var(\hat{I}) = \frac{1}{n} Var(g(X)) \tag{1.8}$$

4. Julien STOEHR, Méthodes deMonte Carlo (cours), université DAUPHINE, Paris 2017-2018

Calcul de la variance de Y :

$$\begin{aligned}
Var(\hat{I}_c) &= Var[\frac{1}{n}\sum_{i=1}^{n} Y_i] \\
&= \frac{1}{n} Var(Y) \\
Var(Y) &= nVar(\hat{I}_c) \\
&= nVar\left[\frac{1}{n}\sum_{i=1}^{n} g(X_i) - c(Z_i - \mathbb{E}(Z_i))\right] \\
&= \frac{1}{n} Var\left[\sum_{i=1}^{n} g(X_i) - c(Z_i - \mathbb{E}(Z_i))\right] \\
&= \frac{1}{n}\left[\sum_{i=1}^{n} Var(g(X)) + c^2 \times Var(Z) - 2c \times cov(g(X), Z)\right] \\
&= \left[Var(g(X)) + c^2 \times Var(Z) - 2c \times cov(g(X), Z)\right]
\end{aligned}$$

$$Var(Y) = Var(g(X)) + c^2 \times Var(Z) - 2c \times cov(g(X), Z) \tag{1.9}$$

On cherche maintenant la valeur optimal de c pour réduire $Var(Y)$:
D'après (1.9), on conclue que la variance de Y devient plus faible que la variance de $g(X)$ si et seulement si on choisit c et Z telles que

$$c^2 \times Var(Z) - 2c \times cov(Y, Z) < 0.$$

La variance de Y étant une fonction quadratique et convexe en c, elle admet un unique minimum c^* :
On pose

$$\begin{aligned}
k(c) &= Var(g(X)) + c^2 \times Var(Z) - 2c \times cov(g(X), Z) \\
k'(c) &= 2c \times Var(Z) - 2cov(g(X), Z) \\
k'(c) = 0 &\implies c = \frac{cov(g(X), Z)}{Var(Z)},
\end{aligned}$$

donc

$$c^* = \frac{cov(g(X), Z)}{Var(Z)}. \tag{1.10}$$

Nous remplacerons (1.10) dans (1.9) :

$$\begin{aligned}
Var(Y) &= \left[Var(g(X)) + (c)^2 \times Var(Z) - 2c \times cov(g(X), Z)\right] \\
&= \left[Var(g(X)) + \frac{cov(g(X), Z)^2}{Var(Z)^2} \times Var(Z) - 2 \times \frac{cov(g(X), Z)}{Var(Z)} \times cov(g(X), Z)\right] \\
&= \left[Var(g(X)) - \frac{cov(g(X), Z)^2}{Var(Z)}\right]
\end{aligned}$$

$$Var(Y) = \left[Var(g(X)) - \frac{cov(g(X), Z)^2}{Var(Z)}\right] \tag{1.11}$$

On a,

$$\frac{cov(g(X), Z)^2}{Var(Z)} > 0.$$

Alors,

$$Var(Y) < Var(g(X)).$$

Chapitre 2

Réduction de la variance de l'estimateur d'indice des valeurs extrêmes γ

Plusieurs chercheurs se sont intéressés à l'estimation de l'indice des valeurs extrêmes γ et son quantile extrême correspondant telle que Hill(1975), Pickands(1975), De Haan et Resnick(1980), Hall(1982), Mason(1982), Hall et Welsh(1985), Dekkers et al(1989), Davis et Resnick(1989), Grubel et Wolf(1994). l'estimateur de l'indice le plus célèbre est L'estimateur de Hill (1975). [1]

On a vu dans le premier chapitre la méthode de Monte-carlo pour la réduction de la variance. Par cette dernier, dans ce chapitre, on va réduire la variance de l'estimateur d'indice des valeurs extrêmes. Nous avons représentons quelque notions essentielles sur la théorie des valeurs extrêmes, nous définissons les statistiques d'ordres, les domaines d'attraction et ses caractéristiques et quelques fonctions telle que la fonction de répartition et la fonction de quantile.

2.1 Théorie des valeurs extrêmes

La théorie des valeurs extrêmes communément appelée « Extreme Value Theory » (EVT) en anglais, est une vaste théorie dont le but est d'étudier les événements rares c'est-à-dire les événements dont la probabilité d'apparition est faible. Autrement dit elle essaie d'amener des éléments de réponses aux intempéries, aux inondations, aux catastrophes naturelles, aux problèmes financiers, etc. en prédisant leurs occurrences dans les années à venir. En d'autres termes on veut 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 extrêmes car l'ordre de ces quantiles tend vers zéro lorsque la taille de l'échantillon, n tend vers l'infini.

1. Jonathan EL METHNI, Contribution à l'estimateur de quantiles extrêmes application à des données environnementales, Université de GRENOBLE.

2.1.1 Statistique d'ordre

Les statistiques d'ordre fournissent des informations sur la distribution de queue ; pour ça ils sont très importantes dans la théorie des valeurs extrêmes. en effet, de façon naturelle et depuis long-temps, dans les problèmes de données censurées ou tronquées.

Définition 2.1.1 *Soit $X_1, X_2, ..., X_n$ suite de variables aléatoires iid de fonction de répartition commune F. et soit les va's $X_{1,n}, X_{2,n}, ..., X_{n,n}$ tel que :*

$$X_{1,n} \leq X_{2,n} \leq ... \leq X_{n,n}$$

sont appelées les statistique d'ordre de l'échantillon $X_1, X_2, ..., X_n$.
les deux statistiques d'ordre suivantes :

$$X_{1,n} = min(X_1, X_2, ..., X_n) \text{ et } X_{n,n} = max(X_1, X_2, ..., X_n)$$

sont appelées les statistiques d'ordre extrêmes. ils sont intéressantes pour l'étude des évènements extrêmes.

il est facile d'écrire l'un en fonction de l'autre :

$$min(X_1, X_2, ..., X_n) = -max(-X_1, -X_2, ..., -X_n)$$

Définition 2.1.2 *(la fonction de répartition empirique) La fonction de répartition empirique de l'échantillon $(X_1, X_2, ..., X_n)$ notée F_n est donnée par :*

$$F_n(x) = \frac{1}{n} \sum \mathbb{I}_{]-\infty,x[}(X_i), \ x \in \mathbb{R}$$

par une autre formule, en utilisant les (s.o) comme suit :

$$F_n(x) = \begin{cases} 0 & si, \ x \leq X_{1,n} \\ \frac{i-1}{n} & si, \ X_{i-1,n} < x \leq X_{i,n} \ , \ 2 \leq i < n \\ 1 & si, \ x > X_{n,n} \end{cases}$$

Définition 2.1.3 *(les fonction de quantile et de quantile de queue) On définit la fonction des quantiles q comme suit :*

$$q(t) = F^{\leftarrow}(t) = inf\{x \in \mathbb{R} : F(x) \geq t\}, \ 0 < t < 1$$

où $F^{\leftarrow}$ est l'inverse généralisée de F. dans la théorie des valeurs extrêmes on utilise souvent la fonction de quantile de queue, elle est définit par :

$$U(t) = q(1 - 1/t) = (1/\bar{F})^{\leftarrow}, \ 1 < t < \infty,$$

Fonction à variation régulière

Définition 2.1.4 *Une fonction mesurable $f : \mathbb{R}_+ \longmapsto \mathbb{R}_+$ est dite à variation régulière à l'infini avec l'indice $\alpha \in \mathbb{R}$, et on écrit $f \in \mathcal{RV}_\alpha$ si et seulement si, pour tout $t > 0$,*

$$\lim_{x \longrightarrow \infty} \frac{f(tx)}{f(x)} = t^\alpha$$

Dans le cas où $\alpha = 0$, la fonction f est dite à variation lente à l'infini, et la notation l est habituellement utilisée pour une telle fonction.

Proposition 2.1.1 *Soient $\alpha \in \mathbb{R}$ et $f \in \mathcal{RV}_\alpha$ alors, il existe une fonction à variation lente l à l'infini telle que :*

$$\forall x > 0,\ f(x) = x^\alpha l(x).$$

Définition 2.1.5 *Une fonction de répartition F sur $\mathbb{R}$ appartient à une classe à variation régulière s'il existe $\alpha \geq 0$ tel que $1 - F \in \mathcal{RV}_{-\alpha}$ sur $\mathbb{R}$, ou d'une manière équivalente :*

$$1 - F(x) \sim x^{-\alpha} l(x),\ \textit{quand}\ x \longrightarrow \infty$$

pour certaines $l \in \mathcal{RV}_0$.

2.1.2 Distribution des valeurs extrêmes

Soit $(X_i)_{i\geq 1}$ une suite de variables aléatoires réelles iid de fonction de répartition F_X. La théorie des valeurs extrêmes a pour but d'étudier le maximum

$$M_n = \max_{i\leq n}(X_i)$$

Lorsque n tend vers l'infini, le maximum converge vers $x^* = sup\{x \in \mathbb{R}, F_X(x) < 1\}$ et par conséquent vers une limite dégénérée. Il convient de renormaliser M_n par un seuil adéquat, c'est- 'a-dire de déterminer un seuil u_n s'écrivant sous la forme $u_n = u_n(t) = a_n t + b_n$ où $a_n > 0$, $b_n \in \mathbb{R}$ et $t \in \mathbb{R}$ est un paramètre, de sorte que $\mathbb{P}(M_n \leq u_n(t)) = F_X^n(u_n(t))$ converge vers une limite non dégénérée c'est-a-dire une fonction de répartition qui n'est pas une fonction saut. En d'autres termes, il s'agit d'obtenir une convergence en loi du type

$$\frac{M_n - b_n}{a_n} \xrightarrow[n\to\infty]{d} Y$$

Théorème 2.1.1 *(Fisher et Tippett, 1928, Gnedenkon, 1943) Soit $X_1, ..., X_n$ une suite de n variables aléatoires réelles iid de loi continue P et $M_n = max(X_1, ..., X_n)$. s'il existe deux suites réelles $(a_n)_{n\geq 1}$ et $(b_n)_{n\geq 1}$, avec $b_n > 0$, et une fonction de répartition non-dégénérée G_γ telle que,*

$$\lim_{n \longrightarrow \infty} \mathbb{P}[\frac{M_n - b_n}{a_n} \leq x] = \lim_{n \longrightarrow \infty} [F(a_n x + b_n)]^n = G_\gamma(x),\ x \in \mathbb{R}$$

alors G_γ est nécessairement l'un des trois types suivantes :

$$\begin{aligned}
\textit{loi de Gumbel : } \Phi_\gamma(x) &= exp(-exp(-x)) \ -\infty < x < +\infty, \\
\textit{loi de Fréchet : } \Lambda_\gamma(x) &= \begin{cases} 0 & x < 0 \\ exp(-x^{-1/\gamma}) & x \geq 0, \gamma > 0, \end{cases} \\
\textit{loi de Weibull : } \Psi_\gamma(x) &= \begin{cases} exp(-(-x)^{-1/\gamma}) & x < 0, \gamma < 0, \\ 1 & x \geq 0 \end{cases}
\end{aligned}$$

avec, G_γ est la loi des valeurs extrêmes, γ est l'indice des valeurs extrêmes et les suites $(a_n)_{n\geq 1}$ et $(b_n)_{n\geq 1}$ sont appelées suites de normalisation(ou paramètres de normalisation) [2].

2.1.3 Domaines d'attraction

Le théorème précédant découle immédiatement que le comportement de la queue de distribution d'une fonction est complètement caractérisé par un unique paramètre noté γ, et appelé indice des valeurs extrêmes. Le signe de ce paramètre est un indicateur essentiel sur la forme de la queue de distribution. il faut donc distinguer les troix cas possibles :

- ❒ Si $\gamma > 0$, F appartient au domaine d'attraction de Fréchet, et l'on note $F \in \mathcal{D}(Frechet)$, Il contient toutes les lois dont la fonction de survie décroit comme une fonction puissance. Ce sont les lois à « queue lourde » ou lois de type Pareto. ces lois ont un point terminal x_F infini. ex : Burr, Pareto strict,..
- ❒ Si $\gamma = 0$, F appartient au domaine d'attraction de Gumbel, et l'on note $F \in \mathcal{D}(Gumbel)$, les queues des lois appartenant à ce domaine décroissent de manière exponentielle(i.e. les lois à queue légères) et le point terminal x_F peut être fini ou non. ex : Gamma, Logestic,..
- ❒ Si $\gamma < 0$, F appartient au domaine d'attraction de Weibull, et l'on note $F \in \mathcal{D}(Weibll)$, les lois de ce domaine d'attraction ont un point terminal x_F fini. ex : Reverse Burr, uniforme,.. [3]

Domaine d'attraction de Fréchet

Théorème 2.1.2 *Une fonction de répartition F appartient au domaine d'attraction de Fréchet avec un indice des valeurs extrêmes $\gamma > 0$ si et seulement si $x_F = sup\{x : F(x) < 1\} =$*

2. Fatima MEDDI, Théorie des valeurs extrêmes(cours), université Kasdi Merbah, Ouargla, Alger, 2017-2018

3. Hayat RAMDANI, Mémoire présenté En Vue De l'obtention Du diplôme de master en mathématiques, Université Kasdi Merbah OUARGLA, Alger,juin 2017

$+\infty$ *et sa fonction de survie* $\bar{F} \in \mathcal{RV}_{-1/\gamma}$ *:*[4]

$$1 - F(x) = x^{-1/\gamma} l(x)$$

Dans ce cas un choix possible des suites de normalisation a_n et b_n est :

$$a_n = F^{\leftarrow}(1 - \frac{1}{n}) = \bar{F}^{\leftarrow}(\frac{1}{n}) \text{ et } b_n = 0$$

Domaine d'attraction de Gumbel

Théorème 2.1.3 *une fonction de répartition F appartient au domaine d'attraction de Gumbel si et seulement si*

$$x_F \leq \infty \ \textit{et} \ \bar{F}(x) = c(x)\text{exp}\{-\int_y^x \frac{g(t)}{a(t)}dt\},\ y < x < x_F,$$

où $c(x) \to c > 0$*,* $g(x) \to 1$ *et* $a'(x) \to 0$ *si* $x \to x_F$*.*
un choix possible pour la fonction a est la fonction moyenne des excès définie par[5] *:*

$$a(x) = \frac{1}{\bar{F}(x)} \int_x^{x_F} F(t)dt, x < x_F.$$

Dans ce cas les suites a_n et b_n sont ainsi définies :

$$b_n = F^{\leftarrow}(1 - \frac{1}{n}) \text{ et } a_n = a(b_n).$$

Domaine d'attraction de Weibull

Théorème 2.1.4 *Une fonction de répartition F appartient au domaine d'attraction de Weibull avec* $\gamma < 0$ *si et seulement si* $x_F < +\infty$ *et en plus* $1 - F^*$ *est une fonction à variation régulière d'indice* $-1/\gamma$ *c'est-à-dire*

$$F^* = 1 - (x_F - x)^{-1/\gamma} l((x_F - x)^{-1})$$

avec,

$$F^*(x) = \begin{cases} 0 & si\ x \leq 0, \\ F(x_F - x^{-1}) & sinon. \end{cases}$$

où l est une fonction à variation lente à l'infini ($l \in \mathcal{RV}_0$*)*[6].

4. Hayat RAMDANI, Mémoire présenté En Vue De l'obtention Du diplôme de master en mathématiques, Université Kasdi Merbah OUARGLA, Alger,juin 2017
5. Hayat RAMDANI, Mémoire présenté En Vue De l'obtention Du diplôme de master en mathématiques, Université Kasdi Merbah OUARGLA, Alger,juin 2017
6. Hayat RAMDANI, Mémoire présenté En Vue De l'obtention Du diplôme de master en mathématiques, Université Kasdi Merbah OUARGLA, Alger,juin 2017

Dans ce domaine d'attraction les suites de normalisation sont déterminées comme suit :

$$a_n = x_F - F^{\leftarrow}(1 - \frac{1}{n}) \text{ et } b_n = x_F.$$

2.2 Estimateur de Hill(1975) $\gamma^H_{k_n}$

L'estimateur de Hill de l'indice de queue (ou indice des valeurs extrêmes), uniquement définit pour les indices positifs $\gamma > 0$. La construction de l'estimateur de Hill est basée sur la méthode de Maximum de vraisemblance où on se sert des statistiques d'ordres supérieur à certain seuil u, pour ne garder que les observations les plus grandes, de façon à ce quelles suivent approximativement une distribution Pareto. il est définit par la statistique suivante :

$$\begin{aligned} \hat{\gamma}^H_{k_n} &= \frac{1}{k}\sum_{i=1}^{k} \log X_{(n-i+1)} - \log X_{(n-k)} \\ &= \frac{1}{k}\sum_{i=1}^{k} i(\log X_{(n-i+1)} - \log X_{(n-i)}) \end{aligned} \tag{2.1}$$

où $X_{(1)} \leq ... \leq X_{(n)}$ est l'échantillon ordonné associé aux variables aléatoires $X_1, .., X_n$

2.2.1 Propriétés asymptotiques

Plusieurs chercheurs ont essayé de déterminer les propriétés asymptotiques de l'estimateur de Hill. Mason (1982) a prouvé la consistance faible de l'estimateur de Hill pour toute suite $k = k(n)$ satisfaisant $k \longrightarrow \infty$ et $k/n \longrightarrow 0$ quand $n \longrightarrow \infty$ appelée suite intermédiaire d'entiers. La condition $k \longrightarrow \infty$ assure que la taille de statistiques d'ordre k est assez grande afin d'obtenir des estimateurs stables. Par contre, la condition $k/n \longrightarrow 0$ permet de rester dans la queue de distribution. Davis et Resnick (1984) ont proposé sa normalité asymptotique sous les conditions de Von Mises ; Csörgö et Mason (1985) ont présenté sa normalité asymptotique en introduisant l'approximation des processus empiriques par les ponts browniens. Dans cette même lancée Resnick et de Haan (1998) ont montré cette propriété asymptotique. Rappelons à présent les propriétés asymptotiques de l'estimateur. Pour cela, nous allons commencer par les conditions du premier et du second ordre avec les fonctions quantiles définies ainsi [7] :

$$q(t) = F^{\leftarrow}(t) = inf\{x \in \mathbb{R} : F(x) \geq t\},\ 0 < t < 1$$

et

$$U(s) = q(1 - 1/s) = (1/\bar{F})^{\leftarrow},\ 1 < s < \infty,$$

7. Hadjer DJOUHRI, Bootstrap de l'indice des valeurs extrêmes en présence de censure aléatoire à droite, Université Kasdi Merbah ouargla, ALGER, juin 2016

Proposition 2.2.1 *(Conditions du premier ordre, de Haan et Ferreira (2006)). Les assertions suivantes sont équivalentes*[8] *:*

1. *F est à queue lourde*

$$F \in \mathcal{D}(Frchet), \gamma > 0$$

2. *$1 - F$ est une fonction à variation régulière à l'infini d'indice $-1/\gamma$*

$$\lim_{t \longrightarrow \infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-1/\gamma}, x > 0$$

3. *$q(1 - s)$ est une fonction à variation régulière à zéro d'indice $-\gamma$*

$$\lim_{t \longrightarrow 0} \frac{q(1 - tx)}{q(1 - t)} = x^{-\gamma}, x > 0$$

4. *U une fonction à variation régulière à l'infini d'indiceγ*

$$\lim_{s \longrightarrow \infty} \frac{U(sx)}{U(s)} = x^{\gamma}, x > 0$$

Proposition 2.2.2 *(Conditions du second ordre de Haan et Ferreira (2006)). Une fonction de répartition $F \in \mathcal{D}(Frchet)$, $\gamma > 0$, admet une condition du second ordre à l'infini si elle satisfait à l'une des assertions suivantes :*

1. *Il existe un paramètre $\rho \leq 0$, et une fonction A_1 qui tend vers 0 (ne change pas de signe à l'infini) définie par, $\forall x > 0$*

$$\lim_{t \longrightarrow \infty} \frac{(1 - F(tx))/(1 - F(t)) - x^{-1/\gamma}}{A_1(t)} = x^{-1/\gamma} \frac{x^{\rho} - 1}{\rho}$$

2. *S'il existe un paramètre $\rho \leq 0$ et une fonction A_2 qui tend vers 0 (ne change pas de signe à l'infini) définie par, $\forall x > 0$*

$$\lim_{s \longrightarrow 0} \frac{U(1 - sx)/U(1 - s) - x^{-\gamma}}{A_2(t)} = x^{-\gamma} \frac{x^{\rho} - 1}{\rho}$$

3. *S'il existe un paramètre $\rho \leq 0$ et une fonction A qui tend vers 0 (ne change pas de signe à l'infini) définie par, $\forall x > 0$*

$$\lim_{t \longrightarrow \infty} \frac{U(tx)/U(t) - x^{\gamma}}{A(t)} = x^{\gamma} \frac{x^{\rho} - 1}{\rho}$$

8. Hadjer DJOUHRI, Bootstrap de l'indice des valeurs extrêmes en présence de censure aléatoire à droite, Université Kasdi Merbah ouargla, ALGER, juin 2016

si $\rho = 0$ on remplace $(x^\rho - 1)/\rho$ par $\log x$.
Les fonctions $A(.), A_1(.), A_2(.)$ sont à variations régulières à l'infini d'indices respectifs $\rho, \rho/\gamma$ et $-\rho$, avec $A_1(t) = A(1/(1-F(t)))$ et $A_2(s) = A(1/s)$. [9]

Ces deux conditions ont permis de déterminer les propriétés asymptotiques de certains estimateurs de l'indice des valeurs extrêmes.

Théorème 2.2.1 *(Propriétés asymptotiques de l'estimateur de Hill)* [10]*. Soit k_n, $n \geq 1$ une suite d ?entiers telle que $1 \leq k_n \leq n$, $k_n \longrightarrow \infty$ et $kn/n \longrightarrow 0$ quand $n \longrightarrow \infty$.*

1. ***Consistance faible** : $\hat{\gamma}^H_{k_n}$ converge en probabilité vers γ.*
2. ***Consistance forte** : Si de plus $k_n/\log n \log n \longrightarrow \infty$ quand $n \longrightarrow \infty$, alors $\hat{\gamma}^H_{k_n}$ converge presque sûrement vers γ.*
3. ***Normalité asymptotique** : Si la condition*

$$\lim_{t \longrightarrow \infty} \frac{U(tx)/U(t) - x^\gamma}{A(t)} = x^\gamma \frac{x^\rho - 1}{\rho}$$

est satisfaite avec $\sqrt{k_n} A(n/k_n) \longrightarrow \lambda \in \mathbb{R}$ quand $n \longrightarrow \infty$, alors

$$\sqrt{k_n}(\hat{\gamma}^H_{k_n} - \gamma) \xrightarrow{\mathcal{L}} \mathcal{N}(\lambda/(1-\rho), \gamma^2)$$

Remarque : D'après la normalité asymptotique précédente, si $\sqrt{k_n} A(n/k_n) \longrightarrow 0$, on peut écrire $\sqrt{k_n}(\hat{\gamma}^H_{k_n}/\gamma - 1) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$ qu'il me donne le résultat suivante :

$$\hat{\gamma}^H_{k_n} \xrightarrow{\mathcal{D}} \mathcal{N}(\gamma, \gamma^2/K_n) \tag{2.2}$$

2.3 Réduction de la variance de l'estimateur de Hill $\hat{\gamma}^H_{k,n}$

Nous appliquerons la méthode de Monte-carlo pour la réduction de la variance de l'estimateur de Hill(1975) $\hat{\gamma}^H_{k,n}$. Dans chaque technique on va poser une $v.a$ Y tel que les variables Y et $\hat{\gamma}^H_{k,n}$ sont de même espérance, mais la variance de Y est plus faible que la variance de $\hat{\gamma}^H_{k,n}$.

9. Hadjer DJOUHRI, Bootstrap de l'indice des valeurs extrêmes en présence de censure aléatoire à droite, Université Kasdi Merbah ouargla, ALGER, juin 2016
10. Hadjer DJOUHRI, Bootstrap de l'indice des valeurs extrêmes en présence de censure aléatoire à droite, Université Kasdi Merbah ouargla, ALGER, juin 2016

2.3.1 Variable antithétique

Soit X une $v.a$ de distribution de Pareto de paramètre γ. $X_1, .., X_n$ une suite de variables aléatoires iid de même loi que X, $g : \mathbb{R}^n \mapsto \mathbb{R}$ une fonction de $v.a$ mesurable et doublement intégrable, $A : \mathbb{R}^n \mapsto \mathbb{R}^n$ une transformation de même loi que X (la variable centrée réduite de X). Nous appliquerons la technique "variable antithétique" de Monte-carlo pour réduire la variance de l'estimateur de Hill(1975) $\hat{\gamma}^H_{k,n}$:

On définit la fonction g et la transformation A comme suit :

$$\begin{aligned} g(X_1, .., X_n) &= \hat{\gamma}^H_{k_n}(X_1, .., X_n) \\ &= \frac{1}{k}\sum_{i=1}^{k} \log X_{(n-i+1)} - \log X_{(n-k)} \quad (2.3) \\ A(X) &= \frac{X - \mathbb{E}(X)}{\sqrt{Var(X)}} \\ &= V \quad (2.4) \end{aligned}$$

D'après (2.3) et (2.4), la variable antithétique Y est donnée par :

$$\begin{aligned} Y &= \frac{g(X) + g \circ A(X)}{2} \\ &= \frac{\hat{\gamma}^H_{k_n}(X) + \hat{\gamma}^H_{k_n} \circ A(X)}{2} \\ &= \frac{\hat{\gamma}^H_{k_n}(X) + \hat{\gamma}^H_{k_n}(V)}{2} \end{aligned}$$

$$Y = \frac{\hat{\gamma}^H_{k_n}(X) + \hat{\gamma}^H_{k_n}(V)}{2} \quad (2.5)$$

On vérifier que g et Y sont de même espérance :

$$\begin{aligned}
\mathbb{E}(Y) &= \mathbb{E}\left[\frac{\hat{\gamma}^H_{k_n}(X)+\hat{\gamma}^H_{k_n}(V)}{2}\right] \\
&= \frac{1}{2}\mathbb{E}\left[\hat{\gamma}^H_{k_n}(X)+\hat{\gamma}^H_{k_n}(V)\right] \\
&= \frac{1}{2}\left[\mathbb{E}(\hat{\gamma}^H_{k_n}(X))+\mathbb{E}(\hat{\gamma}^H_{k_n}(V))\right] \\
&= \frac{1}{2}2\mathbb{E}(\hat{\gamma}^H_{k_n}(X)) \\
&= \mathbb{E}(\hat{\gamma}^H_{k_n}(X)) \\
&= \mathbb{E}(g(X))
\end{aligned}$$

maintenant, On va calculer la variance de Y et l'on comparer avec la variance de $\hat{\gamma}^H_{k_n}$:

$$\begin{aligned}
Var(Y) &= Var\left[\frac{\hat{\gamma}^H_{k_n}+\hat{\gamma}^H_{k_n}\circ A}{2}\right] \\
&= \frac{1}{4}\left[Var(\hat{\gamma}^H_{k_n})+Var(\hat{\gamma}^H_{k_n}\circ A(X))+2cov(\hat{\gamma}^H_{k_n},\hat{\gamma}^H_{k_n}\circ A)\right] \\
&= \frac{1}{4}\left[2Var(\hat{\gamma}^H_{k_n})+2cov(\hat{\gamma}^H_{k_n},\hat{\gamma}^H_{k_n}\circ A)\right] \\
&= \frac{1}{2}Var(\hat{\gamma}^H_{k_n})+\frac{1}{2}cov(\hat{\gamma}^H_{k_n},\hat{\gamma}^H_{k_n}\circ A) \\
&= \frac{1}{2}Var(\hat{\gamma}^H_{k,n})+\frac{1}{2}cov(\hat{\gamma}^H_{k,n},\hat{\gamma}^H_{k,n}\circ A) \\
&= \frac{1}{2}\left[Var(\hat{\gamma}^H_{k,n})+cov(\hat{\gamma}^H_{k,n},\hat{\gamma}^H_{k,n}\circ A)\right] \\
&= \frac{1}{2}\left[\gamma^2/K_n+cov(\hat{\gamma}^H_{k,n},\hat{\gamma}^H_{k,n}\circ A)\right]
\end{aligned}$$

$$Var(Y)=\frac{1}{2}\left[\gamma^2/K_n+cov(\hat{\gamma}^H_{k_n},\hat{\gamma}^H_{k,n}\circ A)\right] \tag{2.6}$$

(2.6) est de même résultat que (1.5) malgré que nous avons travaillé sans l'estimation d'espérance.

Conclusion : La variance de Y obtenue est plus petite que la variance de $\hat{\gamma}^H_{k,n}$ si est seulement si $cov(\hat{\gamma}^H_{k,n},\hat{\gamma}^H_{k,n}\circ A)$ est négatif

$$Var(Y)<Var(\hat{\gamma}^H_{k,n}) \Longleftrightarrow cov(\hat{\gamma}^H_{k,n},\hat{\gamma}^H_{k,n}\circ A)\leq 0 \tag{2.7}$$

2.3.2 Stratification

Soit X une $v.a$ de densité f, D l'ensemble des valeurs prises par X et $(X_i)_{1\leq i\leq n}$ une suite de $v.a$ iid de même loi que X. On définie I comme suit :

$$\begin{aligned} I &= \mathbb{E}(\hat{\gamma}^H_{k,n}(X)) \\ &= \int \hat{\gamma}^H_{k,n}(X)f(x)dx. \end{aligned}$$

Supposons que D est partitionné en $D_1, ..., D_m$ avec ($D = \cup_{1\leq i\leq m} D_i$), on définit I_i et P_i comme suit :

$$\begin{aligned} I_i &= \mathbb{E}(\hat{\gamma}^H_{k,n} \mid X \in D_i) \\ P_i &= P(X \in D_i), \end{aligned}$$

alors I est décomposé sous la forme :

$$\begin{aligned} I &= \sum_{i=1}^{m} P_i I_i \\ &= \sum_{i=1}^{m} P_i \mathbb{E}(\hat{\gamma}^H_{k,n} \mid X \in D_i) \end{aligned}$$

Nous pouvons approcher I_i par une méthode de Monte-Carlo a n_i tirages indépendants :

$$I_i \approx \hat{I}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} \hat{\gamma}^H_{k,n}(X^{(i)}), \text{ avec } \sum_{i=1}^{m} n_i = n.$$

Nous en déduisons une deuxième approximation :

$$\begin{aligned} I \approx \hat{I} &= \sum_{i=1}^{m} p_i \hat{I}_i \\ &= \sum_{i=1}^{m} p_i \left[\frac{1}{n_i} \sum_{j=1}^{n_i} \hat{\gamma}^H_{k,n}(X^{(i)}) \right]. \end{aligned}$$

L'erreur commise se décompose dans ce cas en une somme d'erreur :

$$I - \sum_{i=1}^{m} P_i \hat{I}_i = \sum_{i=1}^{m} P_i (I_i - \hat{I}_i)$$

Dans cette somme, chacun des termes est (approximativement) de loi $\mathcal{N}(0, p_i^2\sigma_i^2/n_i)$ avec :

$$\begin{aligned} \sigma_i^2 &= Var(\hat{\gamma}^H_{k,n}(X) | X_i \in D_i) \\ &= P_i^{-1}\mathbb{E}[(\hat{\gamma}^H_{k,n}(X))^2 \mathbb{I}_{\{X\in D_i\}}] - P_i^{-2}\mathbb{E}(\hat{\gamma}^H_{k,n}(X)\mathbb{I}_{\{X\in D_i\}})^2 \quad \text{(voir (1.6))} \end{aligned}$$

Nous voulons que cette erreur soit la plus petite possible, donc nous voulons minimiser sa variance :

$$\begin{aligned} Var(I-\hat{I}) &= Var\left[\sum_{i=1}^{m} P_i(I_i-\hat{I}_i)\right] \\ &= \sum_{i=1}^{m} Var\left[P_i(I_i-\hat{I}_i)\right] \\ &= \sum_{i=1}^{m} \frac{P_i^2\sigma_i^2}{n_i}, \end{aligned}$$

on prend $n_i = n\frac{P_i\sigma_i}{\sum_{i=1}^{m} P_j\sigma_j}$ alors,

$$\begin{aligned} \sum_{i=1}^{m} \frac{P_i^2\sigma_i^2}{n_i} &\geq \sum_{i=1}^{m} P_i^2\sigma_i^2 \frac{\sum_{j=1}^{m} P_j\sigma_j}{nP_i\sigma_i} \\ &\geq \frac{1}{n}(\sum_{j=1}^{m} P_i\sigma_i)^2 \\ Var(I-\hat{I}) &\geq \frac{1}{n}(\sum_{j=1}^{m} P_i\sigma_i)^2 \end{aligned}$$

Calcule de la variance de $\hat{\gamma}_{k,n}^H(X)$:

$$\begin{aligned} Vra(\hat{\gamma}_{k,n}^H(X)) &= \mathbb{E}\left[\left(\hat{\gamma}_{k,n}^H(X)\mathbb{I}_{\{X\in D_i\}}\right)^2\right] - \mathbb{E}\left[\sum_{i=1}^{m} \hat{\gamma}_{k,n}^H(X)\right]^2 \\ &= \mathbb{E}\left[\sum_{i=1}^{m} \hat{\gamma}_{k,n}^H(X)^2 \mathbb{I}_{\{X\in D_i\}}\right] - \left[\sum_{i=1}^{m} P_i\mathbb{E}(\hat{\gamma}_{k,n}^H(X)|X\in D_i)\right]^2 \\ &= \sum_{i=1}^{m}\{P_i\mathbb{E}((\hat{\gamma}_{k,n}^H(X))^2|X\in D_i) - P_i\mathbb{E}(\hat{\gamma}_{k,n}^H(X)|X\in D_i)^2\} \\ &\quad + \sum_{i=1}^{m} P_i\mathbb{E}(\hat{\gamma}_{k,n}^H(X)|X\in D_i)^2 - \mathbb{E}\left[\sum_{i=1}^{m} P_i\mathbb{E}(\hat{\gamma}_{k,n}^H(X)|X\in D_i)\right]^2 \end{aligned}$$

Nous appliquerons l'inégalité de convexité pour la fonction carré (pour $x_1, ..., x_m \in \mathbb{R}$ et $\lambda_1, ..., \lambda_m \in \mathbb{R}_+$ tel que $\sum_{i=1}^{m} \lambda_i = 1$ alors $\sum_{i=1}^{m} \lambda_i x_i^2 \geq (\sum_{i=1}^{m} \lambda_i x_i)^2$) [11], on trouve :

$$\sum_{i=1}^{m} P_i\mathbb{E}(\hat{\gamma}_{k,n}^H(X)|X\in D_i)^2 \geq \mathbb{E}\left[\sum_{i=1}^{m} P_i\mathbb{E}(\hat{\gamma}_{k,n}^H(X)|X\in D_i)\right]^2$$

11. Sylvain Rubenthaler, Méthodes de Monte-Carlo (Cours et exercices), Université Nice Sophia Antipolis 2017-2018

$$\sum_{i=1}^{m} P_i \mathbb{E}(\hat{\gamma}_{k,n}^{H}(X)|X \in D_i)^2 - \mathbb{E}\left[\sum_{i=1}^{m} P_i \mathbb{E}(\hat{\gamma}_{k,n}^{H}(X)|X \in D_i)\right]^2 \geq 0$$

alors,

$$\begin{aligned}
Var(\hat{\gamma}_{k,n}^{H}(X)) &\geq \sum_{i=1}^{m}\{P_i \mathbb{E}((\hat{\gamma}_{k,n}^{H}(X))^2|X \in D_i) - P_i \mathbb{E}(\hat{\gamma}_{k,n}^{H}(X)|X \in D_i)^2\} \\
&\geq \left(\sum_{i=1}^{m} P_i \left[\mathbb{E}((\hat{\gamma}_{k,n}^{H}(X))^2|X \in D_i) - \mathbb{E}((\hat{\gamma}_{k,n}^{H}(X))|X \in D_i)^2\right]^{1/2}\right)^2 \\
&= \left(\sum_{i=1}^{m} P_i \sigma_i\right)^2
\end{aligned}$$

On conclue que la variance de $\hat{\gamma}_{k,n}^{H}(X)$ est bien réduite.

2.3.3 Échantillonnage préférentiel

Soit X une $v.a$ de densité f, h une autre densité équivalente a f, $g : \mathbb{R}^n \longmapsto \mathbb{R}$ une fonction de variables aléatoires mesurable et doublement intégrable définit par :

$$\begin{aligned}
g : (X_1, ..., X_n) &\longmapsto \hat{\gamma}_{k,n}^{H}(X) \\
(X_1, ..., X_n) &\longmapsto \frac{1}{k}\sum_{i=1}^{k} \log X_{(n-i+1)} - \log X_{(n-k)}
\end{aligned}$$

avec $X_{(1)}, ..., X_{(n)}$ la statistique d'ordre correspondante a la suite des $v.a$'s iid $X_1, ..., X_n$ de même loi que X. On définit la variable Y sous la forme :

$$\begin{aligned}
Y &= \frac{g(X)f(X)}{h(X)} \\
&= \frac{\hat{\gamma}_{k,n}^{H}(X)f(X)}{h(X)}.
\end{aligned}$$

Vérifiant que $\mathbb{E}(Y) = \mathbb{E}(\hat{\gamma}_{k,n}^{H})$:

$$\begin{aligned}
\mathbb{E}_h(Y) &= \mathbb{E}_h\left(\frac{\hat{\gamma}_{k,n}^{H}(X)f(X)}{h(X)}\right) \\
&= \int_{\mathbb{R}^n} \frac{\hat{\gamma}_{k,n}^{H}(X)f(X)}{h(X)} h(X)dX \\
&= \int_{\mathbb{R}^n} \hat{\gamma}_{k,n}^{H}(X)f(X)dX \\
&= \mathbb{E}_f\left(\hat{\gamma}_{k,n}^{H}(X)\right),
\end{aligned}$$

maintenant on va comparer les variance de Y et $\hat{\gamma}^H_{k,n}$:
On a

$$\begin{aligned}
Var(\hat{\gamma}^H_{k,n}) &= \mathbb{E}((\hat{\gamma}^H_{k,n}(X))^2) - \mathbb{E}(\hat{\gamma}^H_{k,n}(X))^2 \\
&= \int (\hat{\gamma}^H_{k,n})^2(X) f(X) dX - \mathbb{E}_f \left[\hat{\gamma}^H_{k,n}(X)\right]^2
\end{aligned}$$

et

$$\begin{aligned}
Var(Y) &= \mathbb{E}(Y^2) - \mathbb{E}(Y)^2 \\
&= \mathbb{E}_h \left[\frac{(\hat{\gamma}^H_{k,n})^2(X) f^2(X)}{h^2(X)}\right] - \mathbb{E}_h \left[\frac{\hat{\gamma}^H_{k,n}(X) f(X)}{h(X)}\right]^2 \\
&= \int (\hat{\gamma}^H_{k,n})^2(X) \frac{f^2(X)}{h^2(X)} h(X) dX - \left(\int \frac{\hat{\gamma}^H_{k,n}(X) f(X)}{h(X)} h(X)\right)^2 \\
&= \int (\hat{\gamma}^H_{k,n})^2(X) \frac{f^2(X)}{h(X)} dX - \left(\int \hat{\gamma}^H_{k,n}(X) f(X)\right)^2 \\
&= \mathbb{E}_f \left[\frac{(\hat{\gamma}^H_{k,n})^2(X) f(X)}{h(X)}\right] - \mathbb{E}_f \left[\hat{\gamma}^H_{k,n}(X)\right]^2
\end{aligned}$$

alors

$$\begin{aligned}
Var(\hat{\gamma}^H_{k,n}) - Var(Y) &= \mathbb{E}((\hat{\gamma}^H_{k,n}(X))^2) - \mathbb{E}_f \left[\frac{(\hat{\gamma}^H_{k,n})^2(X) f(X)}{h(X)}\right] \\
&= \int (\hat{\gamma}^H_{k,n})^2(X) f(X) dX - \int (\hat{\gamma}^H_{k,n})^2(X) \frac{f^2(X)}{h(X)} dX \\
&= \int (\hat{\gamma}^H_{k,n})^2 \left(1 - \frac{f(X)}{h(X)}\right) f(X) dX. \qquad (2.8)
\end{aligned}$$

D'après la proposition (1.2.2), on remplacerons (1.7) dans (2.8) :

$$\begin{aligned}
\int (\hat{\gamma}^H_{k,n})^2 \left(1 - \frac{f(X)}{h(X)}\right) f(X)dX &= \int (\hat{\gamma}^H_{k,n})^2 \left(1 - \frac{f(X)}{h^*(X)}\right) f(X)dX \\
&= \int (\hat{\gamma}^H_{k,n})^2 \left(1 - \frac{\int |\hat{\gamma}^H_{k,n}(X)| f(X)dX}{|\hat{\gamma}^H_{k,n}(X)| f(X)} f(X)\right) f(X)dX \\
&= \int (\hat{\gamma}^H_{k,n})^2 \left(1 - \frac{\int |\hat{\gamma}^H_{k,n}(X)| f(X)dX}{|\hat{\gamma}^H_{k,n}(X)|}\right) f(X)dX \\
&= \int (\hat{\gamma}^H_{k,n})^2 \left(1 - \frac{\mathbb{E}\left[|\hat{\gamma}^H_{k,n}(X)|\right]}{|\hat{\gamma}^H_{k,n}(X)|}\right) f(X)dX \ > 0
\end{aligned}$$

alors

$$Var(\hat{\gamma}^H_{k,n}(X)) - Var(Y) \ > \ 0$$

on conclue que :

$$Var\left(\frac{\hat{\gamma}^H_{k,n}(X) f(X)}{h(X)}\right) < Var(\hat{\gamma}^H_{k,n})$$

2.3.4 Variable de contrôle

Soit $(X_i)_{i\leq n}$ une suite de $v.a$'s iid de même loi que X, Z variable de contrôle de X est aussi de même loi que X. On définit la fonction de $v.a$ g comme suite :

$$\begin{aligned}
g : (X_1, ..., X_n) &\longmapsto \hat{\gamma}^H_{k,n}(X) \\
(X_1, ..., X_n) &\longmapsto \frac{1}{k}\sum_{i=1}^{k} \log X_{(n-i+1)} - \log X_{(n-k)}
\end{aligned}$$

On définit la variable Y sous la forme :(c : paramètre constant fixé)

$$\begin{aligned}
Y &= g(X) - c(Z - \mathbb{E}(Z)) \\
&= \hat{\gamma}^H_{k,n} - c(Z - \mathbb{E}(Z))
\end{aligned}$$

On a

$$\begin{aligned}
\mathbb{E}[Y] &= \mathbb{E}[\hat{\gamma}^H_{k,n} - c(Z - \mathbb{E}(Z))] \\
&= \mathbb{E}[\hat{\gamma}^H_{k,n}] - c\mathbb{E}[Z - \mathbb{E}(Z)] \\
&= \mathbb{E}[\hat{\gamma}^H_{k,n}] - c[\mathbb{E}[Z] - \mathbb{E}[(Z)]] \\
&= \mathbb{E}[\hat{\gamma}^H_{k,n}]
\end{aligned}$$

Calcul de la variance de Y :

$$\begin{aligned}
Var(Y) &= Var\left[\hat{\gamma}^H_{k,n}(X) - c(Z - \mathbb{E}(Z))\right] \\
&= Var[\hat{\gamma}^H_{k,n}(X)] + c^2 Var[Z - \mathbb{E}(Z)] - 2c \times cov\left(\hat{\gamma}^H_{k,n}(X), Z - \mathbb{E}(Z)\right) \\
&= Var[\hat{\gamma}^H_{k,n}(X)] + c^2 Var(Z) - 2c \times cov[Z, \mathbb{E}(Z)] - 2c \times cov(\hat{\gamma}^H_{k,n}(X), Z) + 2c \times cov[g, \mathbb{E}(Z)] \\
&= Var[\hat{\gamma}^H_{k,n}(X)] + c^2 \times Var(Z) - 2c \times cov(\hat{\gamma}^H_{k,n}(X), Z) \qquad (2.9)
\end{aligned}$$

On pose

$$k(c) = Var(\hat{\gamma}^H_{k,n}(X)) + c^2 \times Var(Z) - 2c \times cov(\hat{\gamma}^H_{k,n}(X), Z)$$

est une fonction quadratique et convexe en c, elle admet un unique minimum c^* :

$$c^* = \frac{cov(g(X), Z)}{Var(Z)},$$

Nous remplacerons la valeur de c^* dans (2.9) :

$$\begin{aligned}
Var(Y) &= Var(\hat{\gamma}^H_{k,n}) + (c^*)^2 \times Var(Z) - 2c^* \times cov(\hat{\gamma}^H_{k,n}, Z) \\
&= Var(\hat{\gamma}^H_{k,n}) + \frac{cov(\hat{\gamma}^H_{k,n}, Z)^2}{Var(Z)^2} \times Var(Z) - 2 \times \frac{cov(\hat{\gamma}^H_{k,n}, Z)}{Var(Z)} \times cov(\hat{\gamma}^H_{k,n}, Z) \\
&= Var(\hat{\gamma}^H_{k,n}) - \frac{cov(\hat{\gamma}^H_{k,n}, Z)^2}{Var(Z)}.
\end{aligned}$$

On a

$$\frac{cov(\hat{\gamma}^H_{k,n}, Z)^2}{Var(Z)} > 0 \qquad (2.10)$$

alors

$$Var(Y) < Var(\hat{\gamma}^H_{k,n}) \qquad (2.11)$$

Chapitre 3

Simulations et algorithmes

3.1 Simulation de loi de probabilité continue par méthode d'inversion

Définition 3.1.1 *Soit F la fonction de répartition de la v.a réelle X*

$$F(t) = P(X \geq t) \text{ et } \alpha \in]0, 1[$$

- *On appelle quantile ou fractile d'ordre α le nombre réel χ_α vérifiant :*

$$F(\chi_\alpha) = \alpha$$

- *Lorsque F est inversible on a :*

$$F^{-1}(\alpha) = \chi_\alpha$$

- *Lorsque F n'est pas inversible, on définit la fonction inverse généralisée de F :*

$$Q(s) = F^{-1}(s) = inf\{t : F(t) \geq s\}, \forall t \in \mathbb{R}; \forall s \in]0, 1[$$

Théorème 3.1.1 *Soit F une fonction de répartition d'inverse généralisée F^{-1}. Alors :*

- *Si U est une variable aléatoire uniforme sur $[0, 1]$, la variable aléatoire $X = F^{-1}(U)$, admet une fonction de répartition égale à F.*
- *Si F est continue sur $\mathbb{R}$ et si X est une v.a de loi F, alors la variable aléatoire $U = F(X)$ est de loi uniforme sur $[0, 1]$.* [1]

1. Fatima MEDDI, Simulation stochastique (cours-TP), université KASDI MERBAH OUARLA, Algeria, 2015/2016.

3.1.1 Algorithme

- Une condition minime pour l'application de cette méthode est de connaitre la forme explicite de F^{-1}.
- Cette méthode suggère que pour générer des échantillon d'une $v.a$ X pour laquelle F^{-1} est connue, on peut générer des nombre aléatoires U uniforme sur $[0, 1]$ et faire $X = F^{-1}(U)$.
- Nous avons alors l'algorithme d'inversion suivant :
  — Générer $U \longrightarrow \mathcal{U}(0, 1)$.
  — Faire $X = F^{-1}(U)$
  — Sortir X

3.1.2 Simulation d'une $v.a$ de Pareto(γ)

Sa fonction de répartition est : pour tout $x \geq 1$

$$F(x) = 1 - x^{-1/\gamma}, \gamma > 0$$

siut U une $v.a$ Uniforme (0,1), on a

$$\begin{aligned} F(x) = 1 - x^{-1/\gamma} &= u \\ x^{-1/\gamma} &= 1 - u \\ x &= (1 - u)^{-\gamma} \end{aligned}$$

Nous faisons

$$X = F^{-1}(U) = (1 - U)^{-\gamma}$$

• Le programme sous R est donnée par :

```
n=1000
gamma<-1.1
a<-0
b<-1
U<-runif(n,a,b)
X<-(1-U)^(-gamma)
```

3.2 Choix du nombre d'extrêmes k

Les résultats concernant les estimateurs de l'indice des valeurs extrêmes sont asymptotiques : ils sont obtenus lorsque $k_n \longrightarrow \infty$ et $k_n/n \longrightarrow 0$. La difficulté en pratique consiste à choisir le nombre d'extrêmes k_n utilisé dans les estimations. Dans ce travail nous utilisons l'estimateur de Hill (1975) ce qui rend difficile l'utilisation de l'estimateur en pratique si aucune indication sur le choix de k_n n'est donnée. Des travaux ont montré qu'en utilisant trop d'observations, dans la procédure d'estimation de γ, on observe un biais substantiel tandis que l'utilisation de peu d'observations conduit à une variance considérable. Ce probème a été longuement abordé dans la littérature, voir par exemple Balkema et de Haan (1974), Hall et Welsh (1985), Dekkers et de Haan (1993), Reiss et Thomas (1997), de Haan et Peng (1998), Drees et Kaufmann (1998), Danielsson et al. (2001), Cheng et Peng (2001) Beirlant et al.(2002),Beirlant et al. (2004) [2], etc.

3.2.1 Méthode graphique

C'est la méthode la plus simple pour la détermination de k_n. Elle consiste à tracer le graphe $\{(k_n, \hat{\gamma}_{k_n}) : 1 \leq k_n < n\}$ dans le but de trouver une valeur optimale de k_n. Ainsi le k_n optimal (notée k_{opt}) est choisi dans la première région où l'estimateur $\hat{\gamma}_{k_n}$ devient stable [3].

Algorithme

— Calcul $\hat{\gamma}^H_{k_n}$ pour chaque $1 \leq k_n < n$
— tracer le graphe $\{(k_n, \hat{\gamma}_{k_n}) : 1 \leq k_n < n\}$

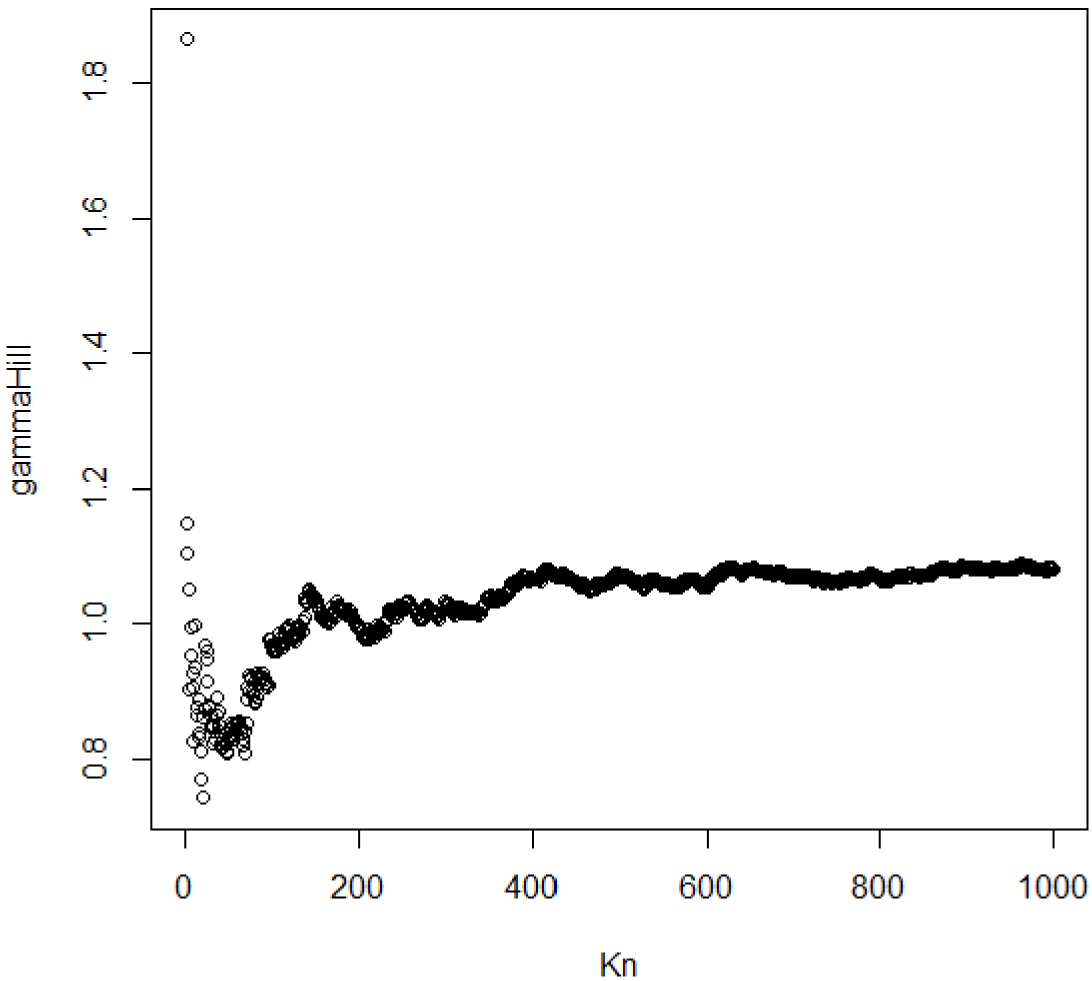
Programme correspondent sous R

```
# tri de X par valeurs croissantes
x<-sort(x)
# l'estimateur de Hill vs K
gammaHill<-numeric(n-1)
Kn<-numeric(n-1)
for(k in 1 :n-1)
{
y<-numeric(k)
for(i in 1 :k)
{
y[i]<-log(x[n-i+1])-log(x[n-k])
}
gammaHill[k]<-(1/k)*sum(y)
```

2. EL HADJI DEME, Quelques contributions à la Théorie univariée des Valeurs Extrêmes et Estimation des mesures de risque actuariel pour des pertes à queues lourdes, UNIVERSITÉ GASTON BERGER DE SAINT-LOUIS, Sénégal, 5 juin 2013

3. EL HADJI DEME, Quelques contributions à la Théorie univariée des Valeurs Extrêmes et Estimation des mesures de risque actuariel pour des pertes à queues lourdes, UNIVERSITÉ GASTON BERGER DE SAINT-LOUIS, Sénégal, 5 juin 2013

Kn[k]<-k
}
# le graphe $\{(k_n, \hat{\gamma}_{k_n}) : 1 \leq k_n < n\}$
plot(Kn,gammaHill)



D'après le graphique ci-dessus $\{(k_n, \hat{\gamma}_{k_n}) : 1 \leq k_n < n\}$, $\hat{\gamma}^H_{k_n}$ devient stable depuis un nombre d'extrêmes égal a 318. Donc $K_{opt} = 318$.

3.3 Simulation de la méthode de Monte-Carlo

3.3.1 Variable antithétique

- génération d'une suite de $v.a$ de loi de pareto de paramètre γ.

- Choix du nombre des valeurs extrêmes K.
- calculer $\hat{\gamma}^H_{k,n}(X)$, $\mathbb{E}(\hat{\gamma}^H_{k,n}(X))$ et $Var(\hat{\gamma}^H_{k,n}(X))$ avec :

$\hat{\gamma}^H_{k,n}(X) = \frac{1}{k}\sum_{i=1}^{k} \log X_{(n-i+1)} - \log X_{(n-k)}$

$\mathbb{E}(\hat{\gamma}^H_{k,n}(X)) = \gamma$

$Var(\hat{\gamma}^H_{k,n}(X)) = \gamma^2/K_n$

- générer V, avec :

$$V = \frac{X - \mathbb{E}(X)}{\sqrt{Var(X)}}$$

- générer Y, avec :

$$Y = \frac{\hat{\gamma}^H_{k,n}(X) + \hat{\gamma}^H_{k,n}(V)}{2}$$

- calculer $cov(\hat{\gamma}^H_{k,n}(X), \hat{\gamma}^H_{k,n}(V))$ pour montrer quelle est négatif.
- calculer $Var(Y)$:

$$Var(Y) = \frac{1}{2}\left[Var(\hat{\gamma}^H_{k,n}) + cov(\hat{\gamma}^H_{k,n}, \hat{\gamma}^H_{k,n} \circ A)\right]$$

- comparer $Var(Y)$ avec $Var(\hat{\gamma}^H_{k,n})$

3.3.2 Stratification

- génération d'une suite de $v.a$ de loi de pareto de paramètre γ.
- Choix du nombre des valeurs extrêmes K.
- calculer $\hat{\gamma}^H_{k,n}(X)$, $\mathbb{E}(\hat{\gamma}^H_{k,n}(X))$ et $Var(\hat{\gamma}^H_{k,n}(X))$ avec :

$\hat{\gamma}^H_{k,n}(X) = \frac{1}{k}\sum_{i=1}^{k} \log X_{(n-i+1)} - \log X_{(n-k)}$

$\mathbb{E}(\hat{\gamma}^H_{k,n}(X)) = \gamma$

$Var(\hat{\gamma}^H_{k,n}(X)) = \gamma^2/K_n$

- Décomposition de D de l'ensemble des valeurs prise par X, tel que

$$\cup_{i=1}^{m} D_i = D$$

- Calcul I_i, P_i, σ_i, avec

$$\begin{aligned} I_i &= \mathbb{E}[\hat{\gamma}_k^H \mid X \in D_i] \\ P_i &= P[\hat{\gamma}_k^H \mid X \in D_i] \\ \sigma_i &= \sqrt{P_i^{-1}\mathbb{E}[(\hat{\gamma}_k^H)^2 \mathbb{I}_{\{X\in D_i\}}] - P_i^{-2}\mathbb{E}(g(X)\mathbb{I}_{\{X\in D_i\}})^2} \end{aligned}$$

- Comparer $Var(\hat{\gamma_k^H})$ avec $\left(\sum_{i=1}^m P_i\sigma_i\right)^2$.

3.3.3 Échantillonnage préférentiel

- génération d'une suite de $v.a$ de loi de pareto de paramètre γ.
- Choix du nombre des valeurs extrêmes K.
- calculer $\hat{\gamma}_{k,n}^H(X)$, $\mathbb{E}(\hat{\gamma}_{k,n}^H(X))$ et $Var(\hat{\gamma}_{k,n}^H(X))$ avec :

  $\hat{\gamma}_{k,n}^H(X) = \frac{1}{k}\sum_{i=1}^k \log X_{(n-i+1)} - \log X_{(n-k)}$

  $\mathbb{E}(\hat{\gamma}_{k,n}^H(X)) = \gamma$

  $Var(\hat{\gamma}_{k,n}^H(X)) = \gamma^2/K_n$

- définie les densités f et h :

$$\begin{aligned} f(x) &= x - \frac{\gamma}{\gamma-1}x^{1-1/\gamma} \\ h(x) &= \frac{1}{\int_{\mathbb{R}^n} \mid g(X) \mid f(X)dX} \mid g(X) \mid f(X) \end{aligned}$$

- générer Y :

$$Y = \frac{\hat{\gamma}_{k,n}^H(X)f(x)}{h(x)}$$

- calculer $\int (\hat{\gamma}_{k,n}^H)^2 \left(1 - \frac{f(X)}{h(X)}\right) f(X)dX$ pour montrer quelle est positif.
- calculer $Var(Y)$, avec :

$$Var(Y) = \mathbb{E}_f\left[\frac{(\hat{\gamma}_{k,n}^H)^2(X)f(X)}{h(X)}\right] - \mathbb{E}_f\left[\hat{\gamma}_{k,n}^H(X)\right]^2$$

- comparer $Var(Y)$ avec $Var(\hat{\gamma}_{k,n}^H)$

3.3.4 Variable de contrôle

- génération de deux suites de $v.a$ de loi de pareto de paramètre γ.
- Choix du nombre des valeurs extrêmes K.
- calculer $\hat{\gamma}^H_{k,n}(X)$, $\mathbb{E}(\hat{\gamma}^H_{k,n}(X))$ et $Var(\hat{\gamma}^H_{k,n}(X))$ avec :

  $\hat{\gamma}^H_{k,n}(X) = \frac{1}{k}\sum_{i=1}^{k} \log X_{(n-i+1)} - \log X_{(n-k)}$

  $\mathbb{E}(\hat{\gamma}^H_{k,n}(X)) = \gamma$

  $Var(\hat{\gamma}^H_{k,n}(X)) = \gamma^2/K_n$

- définie la constante c :

$$c = \frac{cov(\hat{\gamma}^H_{k,n}, Z)}{Var(Z)}$$

- générer Y avec :

$$Y = \hat{\gamma}^H_{k,n}(X) - c(Z - \mathbb{E}(Z))$$

- calculer $\frac{cov(\hat{\gamma}^H_{k,n}, Z)^2}{Var(Z)}$
- calculer $Var(Y)$, avec

$$Var(Y) = Var(\hat{\gamma}^H_{k,n}) - \frac{cov(\hat{\gamma}^H_{k,n}, Z)}{Var(Z)}$$

- comparer $Var(Y)$ avec $Var(\hat{\gamma}^H_{k,n})$

Conclusion

Dans ce travail, nous avons utilisé les quatre techniques de la méthode de Monte-Carlo (variable antithétique, stratification, échantillonnage préférentielle et variable de contrôle) pour la réduction de la variance de l'estimateur de Hill (1975). Notamment nous avons donné un algorithme pour chaque méthode présentée en détail.

Dans un premier temps, notre objectif est de présenter ce résultat sous forme d'un programme sous R pour obtenir par chaque méthode une réduction de la variance de l'estimateur de Hill, ainsi une comparaison numérique entre les quatre réductions envisagées qui nous permettra de choisir la plus petite d'entre aux dans le but de l'exploiter dans une étude quelconque de valeurs extrêmes.

Dans un sens analytique en terme de réduction de la variance de l'estimateur de Hill (1975) nous seront amener à démontrer laquelle des variances réduite représentera la valeur minimal de cet dernière, cela en présentant trois lemmes sous forme d inégalités deux à deux pour sélectionner la méthode qui donnera la plus petite valeur de la variance réduite précitée. La démonstration des lemmes sera une bonne perspective pour la suite de notre travail.

Bibliographie

[1] Roger, Eckhardt, Stan Ulam, John Von Neumann et la méthode de Monte-carlo, Los Alamos science,1987

[2] Pathé NDAE, Modélisation des valeurs extrêmes conditionnellement en présence de censure, Université Gaston BERGER de SAINT-LOUIS.

[3] Julien STOEHR, Méthode de Monte-carlo, Université de Paris Dauphine.

[4] Laure ELIE, Bernard LAPEYRE, Introduction aux méthode de Monte-carlo, septembre 2001.

[5] Jonathan EL METHNI, Contribution à l'estimateur de quantiles extrêmes application à des données environnementales, Université de GRENOBLE.

[6] Sylvain RUBENTHALLER, Méthode de Monte-carlo, Université nice Sophia Antipolis 2017/2018.

[7] Fatima MEDDI, Simulation stochastique (cours-TP), université KASDI MERBAH OUARLA, Algeria, 2015/2016.

[8] EL HADJI DEME, Quelques contributions à la Théorie univariée des Valeurs Extrêmes et Estimation des mesures de risque actuariel pour des pertes à queues lourdes, UNIVERSITÉ GASTON BERGER DE SAINT-LOUIS, Sénégal, 5 juin 2013

Résumé :

Nous considérons le problème de la réduction de la variance par la méthode de Monte-carlo de l'estimateur de l'indice des valeurs extrêmes dans le cas domaine d'attraction de Fréchet (γ positif). Pour ceci nous avons pris en considération celui de Hill (1975) pour une éventuelle réduction de sa variance. Nous avons présenté des algorithmes correspondants pour chaque technique présentée dans le but de rendre cette réduction de variance numérique et exploitable dans la pratique.

Mots clés : Monte-carlo, réduction de la variance, théorie des valeurs extrêmes, l'indice des valeurs extrêmes, l'estimateur de Hill (1975).

Abstract :

We consider the problem of variance reduction by the Monte carlo method of the extreme value index estimator in the Fréchet domain of attraction (γ positive). For this we have taken into consideration that of Hill (1975) for a possible reduction of its variance. We presented corresponding algorithms for each technique presented in order to make this variance reduction numerical and exploitable in practice.

Key words: Monte Carlo, variance reduction, extreme value theory, extreme value index, Hill estimator (1975).

ملخـص:

تناولنا مشكلة تصغير التباين من خلال طريقة مونت كارلو لمقدّر مؤشر القيمة القصوى في مجال جاذبية فريشت (γ موجبة). ولهذا فقد اخترنا مقدّر هيل(1975) بهدف تصغير تباينه. قدمنا خوارزميات مطابقة لكل تقنية تم تناولها من أجل جعل تصغير التباين هذا عدديًا واستغلالًا عمليًا.

الكلمات المفتاحية: مونت كارلو ، تصغير التباين ، نظرية القيمة القصوى ، مؤشر القيمة القصوى ، تقدير هيل (1975).