كلية العلوم و التقنيات فاس
ⵜⴰⵖⵉⵡⵍⴰⵏⵜ ⵏ ⵜⵎⴰⵙⵙⴰⵏⵉⵏ ⴷ ⵜⵙⵏⵉⵇⵉⵏ
Faculté des Sciences et Techniques de Fès

جامعة سيدي محمد بن عبد الله
Université Sidi Mohamed Ben Abdellah

Licence Sciences et Techniques (LST)

CALCUL SCIENTIFIQUE ET APPLICATIONS

MEMOIRE DE FIN D'ETUDES

Pour l'obtention du Diplôme de Licence Sciences et Techniques

Distribution Des Lois Alpha Stable Pour La Caractérisation Des Phonèmes Aléatoires

Présenté par :

- **SOUGHOU RAJAE**

Encadré par :

- **Pr CHAIBI GHIZLANE**

Soutenu Le 12 Juin 2019 devant le jury composé de:

- **Pr EL KHOMSSI Mohammed**
- **Pr EZZAKI FATIMA**

Année Universitaire 2018 / 2019

FACULTE DES SCIENCES ET TECHNIQUES FES – SAISS
B.P. 2202 – Route d'Imouzzer – FES
☎ 212 (0)5 35 61 16 86 – Fax : 212 (0)5 35 60 82 14
Site web : http://www.fst-usmba.ac.ma

Dedicace

A ma très chère maman et mon cher grand père
Qui me trouve en moi la source de leur fierté
A qui je dois tout
A mon frère NIDAL et ma soeur MAROUA
A qui je souhaite un avenir radieux plein de réussite
A mon encadrant qui serai durant cette mémoire une grande soeur
Et a mes amis

Remerciement

Je tiens tout d'abord à remercier Dieu le tout puissant et miséricordieux, qui nous a donné la force et la patience d'accomplir ce modeste travail.
En second lieu, je tiens à remercier mon encadrant madame CHAIBI GHIZLANE, je la remercie pour sa patience, sa disponibilité et surtout ses judicieux conseils, qui ont contribué à alimenter ma réflexion.
Nos vifs remerciements vont également aux membres du jury El KHOUMSSI MOHAMMED et EZZAKI FATIMA pour l'intérêt qu'ils ont porté à notre recherche en acceptant d'examiner notre travail et de l'enrichir par leurs propositions.
Enfin, nous tenons également à remercier toutes les personnes qui ont participé de près ou de loin à la réalisation de ce travail

Table des matières

Introduction **5**

1 Les lois usuelles **7**
1.1 Variables aléatoires. 7
1.2 Les Lois Discrètes. 9
1.2.1 Loi uniforme : 9
1.2.2 Loi de Bernoulli : 10
1.2.3 Loi Binomiale : 11
1.2.4 Loi de Poisson : 12
1.2.5 Loi géométrique : 13
1.3 Variables aléatoires continues. 13
1.4 Les lois continues 15
1.4.1 Loi uniforme continue : 15
1.4.2 Loi de Cauchy : 15
1.4.3 Loi de Gamma : 16
1.4.4 Loi du Khi-deux : 16
1.4.5 Loi exponentielle : 16
1.4.6 Loi de Student : 17
1.4.7 Loi de Fisher : 17
1.4.8 La loi normale 18
1.5 Théorème central limite 18

2 Les Lois Stables **20**
2.1 Les lois stables uni variées : 20
2.1.1 Les inconvénients de la loi stable : 26
2.2 Les lois stables multivariées 27

3 Estimation des paramètres : **29**
3.1 Méthode de McCulloch : 29
3.2 Maximum de vraisemblance : 31
3.3 Méthode des moments 33
3.4 Conclusion 34

Bibliographie **36**

Introduction

En statistique, une des lois les plus utilisées est la loi gaussienne . Elle est célèbre grâce à la représentation en forme de cloche de sa densité de probabilité. De plus, elle est très souvent utilisée car elle possède des propriétés intéressantes :

- la loi gaussienne est stable par combinaison linéaire.
- Le fait que deux paramètres (moyenne et variance) suffisent à le caractériser
- le théorème de la limite centrale établit la convergence en loi de la somme de n variables aléatoires indépendantes et identiquement distribuées vers une loi Gaussienne justifiant son utilisation dans de nombreuses applications.

Cependant, les méthodes fondées sous l'hypothèse gaussienne ne sont plus valables dès lors que cette hypothèse n'est plus vérifiée. L'opérateur va alors se tourner vers d'autres lois de probabilité théoriques utilisées et exploitées dans plusieurs domaines : en mécanique, télécommunication ... dont voici une liste non-exhaustive : Weibull, Rayleigh, log-normal, Rice ...

L'hypothèse que les données soient issues d'une loi de probabilité va être ensuite validée ou rejetée au moyen de tests statistiques comme celui du χ^2 ou de Kolmogorov-Smirnov.

En pratique, il est possible d'observer sur les histogrammes construits à partir des données une asymétrie et/ou une queue lourde. L'asymétrie est caractérisée par le fait que la densité de probabilité n'est pas symétrique par rapport à un mode. Le phénomène de queue lourde se traduit par une décroissance asymptotiquement plus lente de la queue de la distribution considérée par rapport à la queue d'une distribution gaussienne. Ces propriétés obéissent à des lois appelées α-stables. Le concept de distributions α-stables a été introduit par Paul Lévy en 1924 dans l'article intitulé Théorie des erreurs.

La première application des lois stables a été introduite bien avant la parution de l'article de Lévy par l'astronome Danois Holtsmark : la force gravitationnelle exercée par le système stellaire sur un point de l'univers obéit à une loi stable d'exposant caractéristique $\alpha = 1.5$. La particularité de ces distributions est qu'elles sont dites stables par combinaison linéaire, ce qui induit que la loi gaussienne est un cas particulier de lois stables. Cette propriété permet alors de généraliser le théorème de la limite centrale appliqué au cas gaussien au théorème de la limite centrale généralisé appliqué à la loi α-stable.

La définition d'une loi stable est établie à partir de l'expression de la fonction caractéristique dépendant de quatre paramètres :

— l'exposant caractéristique α
— le paramètre de symétrie β
— le paramètre de dispersion γ
— le paramètre de position σ.

Néanmoins, il existe plusieurs paramétrisations de la fonction caractéristique suivant l'application choisie comme celles proposées par Zolotarev ou par Taqqu et Samorodnisky.

L'expression de la densité de probabilité est obtenue par une transformée de Fourier de la fonction caractéristique : la densité de probabilité n'a donc pas d'expression analytique.

À la suite de la parution de l'article de Lévy, les lois stables ont été peu utilisées jusqu'aux travaux de Mandelbrot elles ont eté appliqués à la finance au début des années 60. Le modèle fondamental des variations des prix, introduit par Bachelier en 1900, suivait une loi de Gauss. Cependant, le modèle Gaussien est limité puisqu'il ne prend pas en compte le hasard boursier, représenté par une succession de hausses et de baisses.

Depuis quelques années, les problèmes de communication s'intéressent à modéliser des bruits à faibles probabilité d'apparition mais à fortes amplitudes, qu'on appelle bruit impulsif.De sorte que les α-stables donnent souvent un très bon ajustement par exemple pour certaines bruits de télécommunication.

Ce travail s 'est donné pour but la présentation des lois α-stable . Il comporte trois chapitres :

Dans **le premier chapitre** nous représentons les lois plus utilisées (discrètes et continues) en particulier la loi gaussienne, avec es différentes propriétés (moyenne et variance) , ainsi nous rappelons le théorème de limite centrale .

Le but du **deuxième chapitre** est de présenter les lois stables, et ses propriétés dans le cas uni varié et multivariées .

Le troisième chapitre présente quelques méthodes d'estimation des paramètres des lois stables

1

Les lois usuelles

Nous définissons les notions : événement,tribu,mesure,probabilité à fin de parler d'une variable aléatoire (discrètes et continues) Les variables aléatoires discrètes ne prennent qu'un nombre fini ou dénombrable de valeurs (en général entières). Les variables aléatoires continues peuvent à priori prendre toutes les valeurs dans un intervalle de réels. Cette distinction correspond bien sûr à celle déjà introduite pour les lois de probabilité, parmi ces lois (Poisson, Bernouilli, binomiale, uniforme continues, gamma, khi-2, normale,...) . ce dernier loi nous oblige de initialiser le théorème limite centrale .

1.1 Variables aléatoires.

La théorie mathématique des probabilités a introduite au
16^e siècle par Cardano puis reprise par Fermat et Pascal au 17^e siècle
permettant de résoudre les jeux de hasard.

***Définition* 1.1.1.**
Soit Ω un ensemble quelconque.
On appelle ***algèbre*** *sur Ω (ou algèbre de parties de Ω) un ensemble $\mathcal{A}$ de parties A,B,..., de Ω tel que :*
Si $A, B \in \mathcal{A}$, $\quad A \cap B \in \mathcal{A}$
$A \cup B \in \mathcal{A}$
$\Omega \in \mathcal{A}$
$\emptyset \in (A)$ *($\emptyset$ ensemble vide)*
si $A \in \mathcal{A}$, le complémentaire $\complement A$ appartient aussi à $\mathcal{A}$
De cette définition résulte que $\mathcal{A}$ est stable par intersection e réunion finie.
On dit que $\mathcal{A}$ est un ***tribu****, ou $\sigma-$algébre si la réunion et l'intersection d'une infinité* ***dénombrable*** *d'élélements de $\mathcal{A}$ appartiennent à $\mathcal{A}$.*

***Définition* 1.1.2.**
Les éléments d'une tribu $\mathcal{A}$ sur Ω s'appellent des ***ensembles mesurables****.*
Un ensemble Ω pourvu d'une tribu $\mathcal{A}$ est appelé un ensemble ***probabilisable****. Les éléments de Ω s'appellent* ***événements éléments*** *. Les éléments de $\mathcal{A}$ s'appellent* ***événemnts****.*

Choisir un événemnts A dans une tribu $\mathcal{A}$, c'est faire une épreuve. L'ensemble des événemnts d'une même tribu sur un même ensemble Ω s'appelle une ***catégorie d'épreuves****.*

***Définition* 1.1.3.**
Soit $\mathcal{A}$ une algèbre sur Ω. Soit ϕ une application de $\mathcal{A}$ dans $\mathbb{R}$. A tout sous-ensemble de Ω appartenant à $\mathcal{A}$, ϕ fait corrsepondre un nombre reél. On dit parfois que ϕ est ***fonction d'ensemble****.*
On dit que ϕ est additive si, pour deux ensembles disjoints A et $B \in \mathcal{A}$,

$$\phi(A \cup B) = \phi(A) + \phi(B)$$

On dit que ϕ est $\sigma -$ additive si, pour toute famille ***dénombrable*** *A_n d'ensemble deux à deux* ***disjoints*** *appartenant à $\mathcal{A}$, et dont la réunion appartient à $\mathcal{A}$.*

$$\phi(\cup A_n) = \Sigma\phi(A_n)$$

.
On dit que ϕ est une ***mésure de probabilité*** *(notée p)si :*
p est définie sur une tribu $\mathcal{A}$ de parties de Ω ;
$p \geq 0$
p(Ω)=1 ;
p est $\sigma -$ additive sur $\mathcal{A}$.
La donnée de $\Omega, \mathcal{A}, p$ constitue un espace probabiliste , ou espace de probabilité.

***Définition* 1.1.4.**
Soient Ω, Ω' deux ensembles. Choisissons dans Ω une tribu $\mathcal{A}$, dans Ω' une tribu $\mathcal{A}'$. Soit X une application de Ω dans Ω'. On dit que X est ***mesurable*** *si l'image réciproque par X de tout ensemble mésurable de Ω' est un élement de $\mathcal{A}$, c'est-à-dire un ensemble mesurable dans Ω. Cela sighifie que l'ensemble des éléments ω de Ω dont l'image par X est un ensemble de $\mathcal{A}'$ constitue un ensemble appartenant à la tribu $\mathcal{A}$.*
On dit que la fonction mesurable X est une ***variable aléatoire****, à valeurs dans Ω', si on a probabilisé Ω, c'est-à-dire choisi sur $\mathcal{A}$ une mesure de probabilité p.*

Cas particuliers :

1. Ω' est $\mathbb{R}$.$\mathcal{A}'$ est le tribu borélienne sur $\mathbb{R}$.
   X est alors une variable aléatoire numérique (réelle).
2. Ω' est le plan complexe C. X est une variable aléatoire complexe.
3. Ω' est $\mathbb{R}^n$. X est une variable aléatoire (vectorielle).

***Définition* 1.1.5.**
On appelle fonction de répartition de la variable aléatoire X la fonction F telle que :

$$P[X(\omega) \leq x] = F(x)$$

.
Elle est définie par
F(x)=mesure dans Ω de l'ensemble des ω tels que $X(\omega) \leq x$.
F(x) définit la ***loi de probabilité*** *de X.*

***Définition* 1.1.6.**
On dit qu'une ***variable aléatoire réelle*** *est discrète si elle ne prend qu'un nombre fini ou dénombrable de valeurs :*

$$X(\Omega) \in \{x_k, k \in K \subset \mathbf{N}\}$$

Dans ce cas, la loi de la variable aléatoire X est la loi de probabilité sur l'ensemble des valeurs possibles de X qui affecte la probabilité $\mathbf{P}[X = x_k]$ au singleton x_k.

Remarque :

- En pratique, l'ensemble des valeurs que peut prendre X est $\mathbf{N}$ ou une partie de $\mathbf{N}$.
- Déterminer la loi d'une **variable aléatoire discrète** c'est :

— -Déterminer l'ensemble des valeurs que peut prendre X.
— -Calculer $\mathbf{P}[\mathrm{X} = x_k]$ pour chacune de ces valeurs x_k .

• Rappelons que le seul sens pratique que l'on puisse donner à la notion de probabilité est celui d'une limite de fréquences expérimentales. C'est aussi le sens qu'il faut donner à la notion de loi discrète

• Répétons n fois indépendamment l'expérience aléatoire à l'issue de laquelle X est mesurée. On obtient ainsi un n-uplet $(X_1, \ldots, X_n)$ de variables aléatoires indépendantes de même loi que X (cela s'appelle un échantillon). On peut sur ce n-uplet calculer les fréquences expérimentales des événements "$\mathrm{X} = x_k$".

$$f_n\left(x_k\right) = \frac{1}{n}\left(1_{x_k}\left(\mathrm{X}_1\right) + \cdots + 1_{x_k}\left(\mathrm{X}_n\right)\right)$$

D'après la loi des grands nombres [1] cette fréquence doit converger vers $P[\mathrm{X} = x_k]$. Pour tout n les fréquences expérimentales $f_n(x_k)$ définissent une loi de probabilité discrète sur l'ensemble des x_i.

On représente graphiquement une loi discrète par un diagramme en bâtons : il consiste à tracer au dessus de l'abscisse x_i un segment vertical de longueur proportionnelle à $P[\mathrm{X} = x_k]$.

1.2 Les Lois Discrètes.

***Définition* 1.2.1.**
On appelle loi d'une v.a discrète la donnée de tous les $\mathbf{P}(X = x_i)$ lorsque x_i prend toutes les valeurs possibles dans $X(\Omega)$.

1.2.1 Loi uniforme :

***Définition* 1.2.2.** *On dit qu'une variable aléatoire X suit une loi uniforme sur l'ensemble 1 . . . n avec des probabilités élémentaires identiques si :*

$$\mathbf{P}(X = k) = \frac{1}{n} \qquad \forall k = 1 ... n$$

1. Soit (X_n) une suite de variables aléatoires réelles indépendantes et de même loi admettant une espérance μ . La moyenne empirique $\overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k$ converge en probabilité vers l'espérance : pour tout $\varepsilon > 0$, on a $\lim_{n\to+\infty} \mathbf{P}\left(\left|\overline{X}_n - \mu\right| > \varepsilon\right) = 0$

Exemples :

La distribution des chiffres obtenus au lancer de dé (si ce dernier est non pipé) suit une loi uniforme dont la loi de probabilité est la suivante :

X	1	2	3	4	5	6
$\mathbf{P}(X = x_i)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

avec pour espérance :

$$\mathbf{E}(X) = \frac{1}{6}\sum i = 3,5$$

et pour variance

$$\mathbf{V}(X) = \frac{1}{6}\sum i^2 \mathbf{E}(X^2) = 2,92$$

où les valeurs x_i correspondent au rang i de la variable X dans la série

Dans le cas particulier d'une loi discrète uniforme ou les valeurs de la variable aléatoire X correspondent au rang $x_i = i \quad (\forall i \in [1, n])$

$$\mathbf{E}(X) = \tfrac{n+1}{2} \qquad \mathbf{V}(X) = \tfrac{n^2-1}{12}$$

1.2.2 Loi de Bernoulli :

***Définition* 1.2.3.**
Soit p un nombre réel appartenant à [0 ;1]. On appelle épreuve de Bernoulli de paramètre p toute expérience aléatoire n'admettant que deux issues $\mathcal{A}$ et $\overline{\mathcal{A}}$ de probabilités respectives p et q=1-p.
Soit $\Omega = \{0, 1\}$ l'univers associé à une expérience aléatoire, et soit p un nombre réel appartenant à [0 ;1]. On appelle loi de Bernoulli de paramètre p la loi de probabilité définie sur Ω

$$\begin{cases} \mathbf{P}(1) &= p \\ \mathbf{P}(0) &= 1 - p \end{cases}$$

***Théorème* 1.2.1.**
Soit $p \in [0; 1]$ et B_p la loi de Bernoulli de paramètre p. On pose q=1-p.
— *L'espérance mathématique de B_p est $\mathbf{E(X)}$=p .*
— *La variance de B_p est $\mathbf{V(X)}$=pq.*

Exemples :

On compte là, par exemple, le nombre N de réponses "oui" dans un échantillon de population, lors d'un sondage, afin d'en déduire la proportion de "oui". On effectue une série de n tirages au hasard dans une population. On pose la même question à chacun des n individus. Le but est d'estimer la proportion p d'individus de la population totale qui auraient répondu "oui" (si on leur avait posé la question) à l'aide du nombre N . On remarque que N peut s'écrire :

$$N = \sum_{k=1}^{n} \mathrm{X}_k,$$

où $X_1, X_2, \ldots, X_n$ sont définies par

$$X_k = 1_{\text{la réponse du } k-\text{ème individu est } oui},$$

, i.e. X_k vaut 1 ou 0 selon que la réponse du k-ème individu est "oui" ou "non". Étant une fonction indicatrice, X_k est donc une variable de Bernoulli. Son paramètre est la probabilité de répondre " oui", à savoir la proportion de "oui" dans la population totale, c'est-a-dire p. On a donc

$$\mathbf{E}[N] = \sum_{1\leq i\leq n} \mathbf{P}(X_i = 1) = np, \qquad et \qquad \mathbf{E}[N/n] = p.$$

D'où l'idée, proposée par Bernoulli dans son ouvrage fondateur' Ars Conjectandi', d'estimer cette proportion p a priori inconnue a l'aide de la proportion N/n de "oui" dans l'échantillon, qui est, elle, connue.

En effet la variable $Z = X_iX_j$ vaut 0 ou 1 est donc une variable de Bernoulli. Si T est la taille totale de la population, pour $i < j$, on a alors

$$\mathbf{P}(X_iX_j = 1) = \frac{pT}{T}\frac{pT-1}{T-1} \quad et \quad Cov(X_i, X_j) = -\frac{p(1-p)}{T-1},$$

puis, a l'aide des propriétés de la variance,

$$\mathbf{V}(N) = \frac{p(1-p)n(T-n)}{T-1}.$$

Dans les deux cas considérés ci-dessus, la loi de N est connue explicitement. Cependant, le calcul de l'espérance de N utilisant la décomposition de N en somme de variables de Bernoulli, présenté ci-dessus, est plus simple que le calcul de l'espérance de N utilisant le théorème de transfert : [2]

$$\mathbf{E}[N] = \sum_{1\leq k\leq n} k\mathbf{P}(N = k).$$

1.2.3 Loi Binomiale :

***Définition* 1.2.4.**
On appelle ***épreuve de Bernoulli*** *une épreuve n'ayant que deux issues : Succès (S) et Échec(E).*
On appelle ***schéma de Bernoulli*** *, la répétition n fois, de manière indépendante d'une épreuve de Bernoulli.*
Soit X la variable aléatoire correspondant au nombre de succès a l'issue de schéma de Bernoulli,

2. Pour une variable aléatoire discrète X

$$\mathbf{E}(\varphi(X)) = \sum_{n\geq 0} \varphi(n)\mathbf{P}(X = n)$$

Pour une variable aléatoire continue X

$$\mathbf{E}(\varphi(X)) = \int_{\mathbf{R}} \varphi(x)f(x)dx$$

donc la loi de probabilité de X est la loi Binomiale de paramètres n et p , notée B(n ;p)
Soit n $\in \mathbf{N}$ *et p un nombre réel tel que* $p \in [0;1]$. *Soit E une épreuve de Bernoulli a deux issues* $\mathcal{A}$ *et* $\overline{\mathcal{A}}$ *de probabilités respectives p et q=1-p. Pour tout entier naturel k tel que* $0 \leq k \leq n$*, la probabilité* $\mathbf{P}_k$ *que l'événement A soit réalisé exactement k fois a l'issue de n épreuves indépendantes E est donnée par :*

$$\mathbf{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

***Théorème* 1.2.2.**
Soit N un entier naturel p un nombre réel tel que p$\in [0;1]$ *et B(n,p) la loi binomiale de paramètres n et p. On pose q=1-p.*
L'espérance mathématique de B(n,p) est $\mathbf{E}(X)$*=np*
La variance de B(n,p) est $\mathbf{V}(X)$*=npq.*

Exemples :
Revenons à l'exemple de la loi de Bernouilli. Si les tirages ont eu lieu avec remise (i.e. il est possible que la même personne soit interrogée plusieurs fois), ce qui implique l'indépendance des X_k, et donne

$$\mathrm{V}(N) = \sum_{1 \leq i \leq n} \mathrm{V}(X_i) = np(1-p)$$

1.2.4 Loi de Poisson :

***Définition* 1.2.5.**
Si le nombre moyen d'occurrences dans un intervalle de temps fixé est λ*, alors la probabilité qu'il existe exactement k occurrences (k étant un entier naturel,* $k = 0, 1, 2, \ldots$*) est*

$$p(k) = P(X = k) = \frac{\lambda^k}{k!} \mathrm{e}^{-\lambda}$$

On dit alors que X suit la loi de Poisson de paramètre λ.

Exemples :
On a constaté 0,6 accident par jour en moyenne sur 720 jours. Sur 2 jours, nous avons :

$$P(X = 2; 0,6) = \frac{0,6^2 \times e^{-0,6}}{2!} = 0,09878609450$$

Le tableau montre, qu'en moyenne, il y aura 395 jours sur 720 sans accident où on ne déplorera qu'un seul accident au maximum.

QAJ	P	FE
0	$0,5488116$	$395,144378$
1	$0,3292869$	$237,08626$
2	$0,0987860$	$71,125988$
3	$0,0029572$	$14,225197$
4	$0,0229635$	$2,133779$

QAJ : Quantité d'accidents par jour.
P :probabilité.
FE :fréquence éstimé.

1.2.5 Loi géométrique :

***Définition* 1.2.6.**
Soit p un nombre réel tel que $p \in [0;1]$. On pose q=1-p. On appelle loi géométrique de paramètre p la loi de probabilité notée $G(p)$ donnant le temps d'attente du premier succès dans une succession d'épreuves indépendantes de Bernoulli de paramètre p.

$$k \in \mathrm{N}^*, \mathbf{P}(X = k) = pq^{k-1}$$

***Théorème* 1.2.3.**
Soit p un nombre réel tel que $p \in [0;1]$ et $G(p)$ la loi géométrique de paramètre p . On pose q=1-p .
- *L'espérance mathématique de $G(p)$ est $\mathbf{E}(\mathbf{X}) = \frac{1}{p}$;*
- *La variance de $G(p)$ est $\mathbf{V}(\mathbf{X}) = \frac{q}{p^2}$*

Exemples :
si les tirages ont eu lieu sans remise (i.e. en évitant d'interroger 2 fois la même personne), auquel cas les X_k ne sont pas indépendants. Alors

$$\mathbf{V}(N) = \sum_{1\leq i\leq n} \mathbf{V}(X_i) + 2 \sum_{1\leq i<j\leq n} Cov(X_i, X_j).$$

En ce cas N suit la loi hypergéométrique, et les calculs requièrent de connaitre la taille totale de la population, qu'on notera T dans la suite. On a

$$Cov(X_i, X_j) = \mathbf{E}[X_i X_j] - \mathbf{E}[X_i]\mathbf{E}[X_j] = \mathbf{P}(X_i X_j = 1) - p^2$$

1.3 Variables aléatoires continues.

Nous proposons ci-dessous une liste non-exhaustive de lois continues couramment utilisées en statistique. Elles ont la particularité d'être uni modales.

***Définition* 1.3.1.**
*La variable aléatoire est dite **continue** si l'ensemble $X(\Omega)$ est un intervalle (ou une réunion d'intervalles) de $\mathbf{R}$.*

Remarque :

La description d'une loi continue diffère de celles des lois discrètes puisque pour une variable aléatoire continue X, la probabilité que X prenne une valeur bien précise x est nulle.

$$\mathbf{P}[\mathrm{X} = x] = 0$$

. Il y a en effet une infinité de valeurs dans $\mathbf{R}$ ou dans un intervalle, et au regard de toutes ces valeurs précises, le poids de la valeur particulière est tellement insignifiant qu'il en est nul !

En effet Soient a et b deux valeurs infiniment petites, et F la fonction de répartition. On peut écrire :
$0 \leq P(X = x_0) \leq P(x_0 - a \leq X \leq x_0 + b)$ $0 \leq P(X = x_0) \leq F(x_0 + b) - F(x_0 - a)$

$0 \leq P(X = x_0) \leq F(x_0 + b) - F(x_0) + F(x_0) - F(x_0 - a)$

F est continue donc :
$F(x_0 + b) - F(x_0)$ tend vers 0 lorsque b tend vers 0
$F(x_0) - F(x_0 - a)$ tend vers 0 lorsque a tend vers 0
Ce qui permet de conclure que : $0 \leq P(X = x_0) \leq 0$
Ainsi $P(X = x_0) = 0$

***Définition* 1.3.2.**
La fonction de répartition d'une variable aléatoire réelle X, notée F est définie par :

$$F : \begin{cases} \mathbf{R} \longrightarrow [0;1] \\ x \longmapsto \mathbf{P}(X \leq x) \end{cases}$$

continue à droite et vérifiant les conditions aux limites :

$$\lim_{x \to -\infty} F(x) = 0$$
$$\lim_{x \to +\infty} F(x) = 1$$

***Définition* 1.3.3.**
Une variable aléatoire est dite à densité si sa fonction de répartition est dérivable. La fonction de densité appelée plus communément fonction de densité de probabilité d'une variable aléatoire X, notée f vaut :

$$\mathbf{P}[a \leq X \leq b] = \int_a^b f(t)dt$$

et la probabilité de trouver X dans un intervalle [a ; b] donné, apparaître comme l'aire d'une partie du graphique située entre la courbe de la densité f et l'axe des abscisses.

$$f : \begin{cases} \mathbf{R} \longrightarrow \mathbf{R} \\ x \longmapsto F'(x) \end{cases}$$

Lorsque la fonction de densité de probabilité est intégrable, on obtient les propriétés suivantes :

$$\forall x \in \mathbf{R}, f(x) \geq 0$$
$$\int_{-\infty}^{+\infty} f(x)dx = 1$$
$$F(x) = \int_{-\infty}^{x} f(u)du$$

***Définition* 1.3.4.**
L'espérance mathématique d'une variable aléatoire à densité X est définie par

$$\mathbf{E}[X] = \int_{-\infty}^{+\infty} xf(x)dx$$

L'espérance est appelée aussi moment d'ordre 1.

***Définition* 1.3.5.**
Il est possible d'estimer les moments à des ordres supérieurs à 1 . Un moment d'ordre k noté $\mathbf{M}_k$ *est définie par :*

$$\mathbf{M}_k(X) = \mathbf{E}\left[X^k\right] = \int_{-\infty}^{+\infty} x^k f(x)dx$$

***Définition* 1.3.6.**
La fonction caractéristique d'une variable aléatoire X, notée ϕ_X, est obtenue à partir de la relation :

$$\phi_X(u) = \mathbf{E}[e^(iuX)] = \int_{-\infty}^{+\infty} f(x)e^(iux)dx$$

Il est possible d'obtenir la fonction de densité de probabilité en calculant la transformée de Fourier de sa fonction caractéristique ϕ_X si elle est intégrable :

$$f(x) = \frac{1}{2\Pi}\int_{-\infty}^{+\infty} \phi_X(u)e^(- iux)du$$

1.4 Les lois continues

1.4.1 Loi uniforme continue :

***Définition* 1.4.1.**
Une v.a. X suit une loi de ***Uniforme continue*** *si elle admet pour densité de probabilité définie sur [a ; b] par :*

$$f(x) = \begin{cases} \frac{1}{b-a} & si\ x \in]a;b[\\ 0 & sinon \end{cases}$$

Les lois uniformes continues sont très utiles de par la facilité des calculs qu'elles permettent et pour leur utilisation dans les simulations.

1.4.2 Loi de Cauchy :

***Définition* 1.4.2.**
Soit $a \in \mathbf{R}_+^*$. On appelle loi de Cauchy de paramètre a la loi de probabilité absolument continue dont une densité est donnée par :

$$f(x) = \frac{a/\pi}{a^2 + x^2}, \quad x \in \mathbf{R}$$

Cette mesure est identifiée par la notation $C(a)$

Les lois de Cauchy apparaissent naturellement en Analyse (équation de **Laplace** dans le demi-plan supérieur) et en Probabilités. Ce sont des lois dont tous les moments divergent. Leurs fonctions caractéristiques sont évidemment bien définies, un calcul non immédiat montre que :

$$\varphi(\theta) = \int_{-\infty}^{+\infty} \mathrm{e}^{\mathrm{i}\theta x} \times \frac{a/\pi}{a^2 + x^2}\mathrm{d}x = \mathrm{e}^{-|a\theta|}$$

qui est, a un facteur multiplicatif prés, la densité de la loi de Laplace de paramètre $a > 0$. On remarquera que cette transformée de Fourier n'est pas dérivable en 0^- ce qui implique, mais on le sait déjà, que la loi de Cauchy de paramètre $a > 0$ n'admet pas de moyenne.

1.4.3 Loi de Gamma :

***Définition* 1.4.3.**
Une v.a X suit la loi Gamma de paramètres λ et a ($\lambda > 0$,$a > 0$), notée $\gamma(\lambda; a)$ si sa densité est :

$$f(x) = \frac{\lambda^a}{\Gamma(a)} e^{-\lambda x} x^{a-1} \mathbf{1}_{]0,+\infty[}(x), x \in \mathbf{R}$$

avec $\Gamma(a) = \int_0^{+\infty} e^{-x} x^{a-1} dx$

Caractéristiques de la loi Gamma

1. $\mathbf{E}(\mathbf{X}) = \frac{a}{\lambda}$
2. $\mathbf{V}(\mathbf{X}) = \dfrac{a}{\lambda^2}$
3. Fonction caractéristique :

$$\phi_x(u) = \mathbf{E}(\mathrm{e}^{itX} \tag{1.1}$$

$$= \frac{\lambda^a}{\Gamma(a)} \int_x^{\infty} \mathrm{e}^{-(\lambda - u)x} x^{a-1} \mathrm{d}x \tag{1.2}$$

$$= \tfrac{\lambda^a}{(\lambda - iu)^a} \tag{1.3}$$

Proposition* 1.4.1.** ***(Propriétés de la loi Gamma :)

1. $\Gamma(a) = (a-1)\Gamma(a-1), \quad \forall a > 1$
2. $\Gamma(a) = (a-1)!, \quad \forall a \in \mathbf{N}^*$
3. $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

1.4.4 Loi du Khi-deux :

***Définition* 1.4.4.**
Soit $(X_i)_{1 \leq i \leq n}$ une suite de variables indépendants et identiquement distribuée (i.i.d) de loi $\mathcal{N}(0,1)$. Alors la loi de $X = X_1^2 + X_2^2 + \cdots + X_n^2$ est appelée loi du χ^2 à n degrés de libertés de densité :

$$f(x) = \begin{cases} \frac{1}{2^{n/2}\Gamma(n/2)} e^{-x/2} x^{\frac{n}{2}-1} & si\ x > 0 \\ 0 & sinon. \end{cases}$$

***Théorème* 1.4.1.** *Si $X = X_1^2 + X_2^2 + \cdots + X_n^2$ suit la loi χ^2 alors :*

1. *Espérance* $\mathbf{E}(\mathbf{X}) = n$
2. *Variance :*$\mathbf{V}(\mathbf{X}) = 2n$
3. *Fonction de répartition :* $F(x) = \int_{-\infty}^{x} \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} e^{-\frac{t}{2}} x^{\frac{n}{2}-1} \mathbf{1}_{[}10, +\infty[(t)\mathrm{d}t, \quad x \in \mathbf{R}$.

1.4.5 Loi exponentielle :

***Définition* 1.4.5.**
On dit que X suit la loi exponentielle de paramètre $a > 0$, ce que l'on note si elle est absolument continue, et admet pour densité :

$$f(x) = \begin{cases} ae^{-ax} & si\ x\ ¿\ 0 \\ 0 & si\ x\ ¡\ 0 \end{cases}$$

X admet alors une espérance et une variance :

$$\mathbf{E}(X) = \frac{1}{a} \text{ et } \mathbf{V}(X) = \frac{1}{a^2}$$

Exemples :
La durée de vie d'un ordinateur portable exprimée en années est une variable aléatoire X suivant la loi exponentielle de paramètre $\lambda = 0,125$
La probabilité que la durée de vie de cet ordinateur portable dépasse 5 ans est $P(X > 5) = 1 - \int 0,125e^{-0,125t}dt = e^{-0,125t\times5} \approx 0,535$ La probabilité que la durée de vie de cet ordinateur portable soit inférieure à 3 ans est

$$P(X < 3) = \int 0,125e^{-0,125t}dt = 1 - e^{-0,125\times3} \approx 0,313$$

1.4.6 Loi de Student :

***Définition* 1.4.6.**
Soient X et Y deux variables indépendantes telles que $X \sim \mathcal{N}(0,1)$ et $Y \sim \chi_n^2$. Alors la variable aléatoire $T_n = \frac{X}{\sqrt{\frac{Y}{n}}}$ suit la loi de Student à n degrés de libertés, sa densité est

$$f(x) = \frac{1}{\sqrt{n\pi}} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}} , \quad \forall x \in \mathbf{R}$$

***Théorème* 1.4.2.** *Si X suit la loi de student alors : Espérance : $\mathbf{E}(X) = 0$ si $n > 0$*

Variance : $\mathbf{V}(X) = \frac{n}{n-2}$ si $n > 2$

1.4.7 Loi de Fisher :

***Théorème* 1.4.3.**
Une variable aléatoire réelle distribuée selon la loi de Fisher peut être construite comme le quotient de deux variables aléatoires indépendantes, U_1 et U_2, distribuées chacune selon une Loi du χ_2 et ajustées pour leurs nombres de degrés de liberté, respectivement d_1 et d_2 :

$$\mathrm{F}\left(d_1, d_2\right) \sim \frac{U_1/d_1}{U_2/d_2}$$

La densité de probabilité d'une loi de Fisher, $F(d_1, d_2)$, est donnée par :

$$f(x) = \frac{\left(\frac{d_1 x}{d_1 x + d_2}\right)^{d_1/2} \left(1 - \frac{d_1 x}{d_1 x + d_2}\right)^{d_2/2}}{x \mathrm{B}\left(d_1/2, d_2/2\right)}$$

pour tout réel $x \geq 0$, où d_1 et d_2 sont des entiers positifs et B est la fonction bêta[3]

3. En mathématiques, la fonction bêta est un type d'intégrale d'Euler définie pour tous nombres complexes x et y de parties réelles strictement positives par

$$\mathrm{B}(x, y) = \int_0^1 t^{x-1}(1-t)^{y-1} \mathrm{d}t$$

1.4.8 La loi normale

***Définition* 1.4.7.**
La loi normale est une loi de probabilité absolument continue qui dépend de deux paramètres : son espérance, un nombre réel noté μ, et son écart type, un nombre réel positif noté sigma. On dit que X suit une loi normale de paramètres μ et sigma notée $\mathcal{N}(\mu, \sigma^2)$ si la densité de probabilité donnée par :

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

La courbe de cette densité est appelée courbe de Gauss ou courbe en cloche.
Cas particulier : *lorsque $\mu = 0$ et $\sigma = 1$: on parle de la loi normale centrée réduite.*

***Théorème* 1.4.4.** *Si X suit une loi normale $\mathcal{N}(\mu, \sigma^2)$ alors : • la densité de probabilité $\varphi : \mathbf{R} \to \mathbf{R}_+$ est donnée par : $\varphi(t) = \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}\frac{(t-\mu)^2}{\sigma^2}}$, pour tout $t \in \mathbf{R}$*

• la fonction de répartition F :$\mathbf{R} \to \mathbf{R}_+$ est donnée par : $F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} \mathrm{e}^{-\frac{1}{2}\frac{(t-\mu)^2}{\sigma^2}} \mathrm{d}t$, pour tout $x \in \mathbf{R}$

• la fonction caractéristique $\phi : \mathbf{R} \to \mathbf{C}$ est donnée par : $\phi(t) = \mathrm{e}^{\mu i t - \frac{1}{2}\sigma^2 t^2}$, pour tout $t \in \mathbf{R}$

• la fonction génératrice des moments $M : \mathbf{R} \to \mathbf{R}_+$ est donnée par : $M(t) = \mathrm{e}^{\mu t + \frac{1}{2}\sigma^2 t^2}$, pour tout $t \in \mathbf{R}$, où $\mu \in \mathbf{R}$ et $\sigma \in \mathbf{R}_+^\star$

• pour le cas où $\sigma = 0$, les fonctions de densité et de répartition ne sont pas définies.

1.5 Théorème central limite

***Définition* 1.5.1.**
Le théorème central limite (aussi improprement appelé théorème de la limite centrale ou centrée) établit la convergence en loi de la somme d'une suite de variables aléatoires vers la loi normale. Intuitivement, ce résultat affirme que toute somme de variables aléatoires indépendantes tend dans certains cas vers une variable aléatoire gaussienne.

***Théorème* 1.5.1.**

Soit $X_1, X_2, ...$ une suite de variables aléatoires réelles définies sur le même espace de probabilité, indépendantes et identiquement distribuées de carré intégrable suivant la même loi D. Supposons que l'espérance μ et l'écart-type σ de D existent et soient finis avec $\sigma \neq 0$ Considérons la somme $S_n = X_1 + X_2 + ... + X_n$
Alors

- *l'espérance de S_n est n μ*
- *son écart-type vaut $\sigma\sqrt{n}$.*

Afin de formuler mathématiquement cette approximation, nous allons poser

$$\overline{X}_n = \frac{S_n}{n} = \frac{X_1 + X_2 + \ldots + X_n}{n}$$

et

$$Z_n = \frac{\mathrm{S}_n - n\mu}{\sigma\sqrt{n}} = \frac{\overline{X}_n - \mu}{\sigma\sqrt{n}}$$

de sorte que l'espérance et l'écart-type de Z_n valent respectivement 0 et 1 : la variable est ainsi dite centrée et réduite.

Le théorème central limite énonce alors que la suite de variables aléatoires $Z_1, Z_2, \ldots, Z_n, \ldots$ converge en loi vers une variable aléatoire Z, définie sur le même espace probabilisé, et de loi normale centrée réduite $\mathcal{N}(0,1)$ lorsque n tend vers l'infini.

Théorème 1.5.2. Généralisations du théorème central limite
On peut, au prix d'une formulation un peu moins simple, supprimer l'hypothèse selon laquelle les variables X_n sont de même loi. Les variables X_n restent toutefois indépendantes : soit donc $(X_n)_{n\geq 1}$ une suite de variables aléatoires définies sur le même espace de probabilité indépendantes. Supposons que, pour $n \geq 1$, X_n ait une espérance finie μ_n et un écart-type fini σ_n, et posons :

$$s_n^2 = \sum_{i=1}^{n} \sigma_i^2$$

et,

$$Z_n = \frac{1}{s_n} \sum_{i=1}^{n} (X_i - \mu_i)$$

2

Les Lois Stables

Avant de définir les lois α-stables, nous allons introduire une famille de lois plus générale : les lois indéfiniment divisibles. C'est à partir de ces lois que sera précisée la forme de la fonction caractéristique des lois stables.

L'intérêt principal de telles lois réside dans la solution du problème suivant : comment détermine-t-on toutes les distributions qui s'expriment comme limite d'une somme de n variables aléatoires réelles (V.A.R.) indépendantes et identiquement distribuées (i.i.d.) ?

2.1 Les lois stables uni variées :

Intuitivement, si une suite de V.A.R. (T_n) converge en loi vers une V.A.R. T et que ces (T_n) s'expriment comme une somme de n V.A.R. indépendantes de même loi, alors T va aussi s'exprimer de la même manière. Introduisons alors la définition suivante :

***Définition* 2.1.1.**
Une V.A.R. X à une distribution indéfiniment divisible si et seulement si $\forall n, \exists X_1, \dots, X_n$ indépendantes et de même loi telles que $X \stackrel{d}{=}$ [1] $X_1 + \cdots + X_n$

Remarque :

Les V.A.R. X_i n'ont pas même loi que X. En revanche, comme nous verrons dans les exemples suivants, elles appartiennent à la même famille de loi.
Cette classe de V.A.R. permet de résoudre notre problème. En effet, on a le théorème suivant.

***Théorème* 2.1.1.**
Une V.A.R. X est la limite d'une somme de n V.A.R. i.i.d. si et seulement si X est indéfiniment divisible
La démonstration est détaillée dans Shiryayev (1984, pages 336)

Remarque :

Une des caractérisations des lois indéfiniment divisibles est que leur fonction caractéristique peut s'écrire comme puissance nème d'une autre fonction caractéristique.

1. en distribution

***Théorème* 2.1.2.**
Si X a une distribution indéfiniment divisible, alors sa fonction caractéristique s'écrit :

$$\varphi_X(t) = \exp\left\{i\mu t + \int_{-\infty}^{+\infty} \frac{e^{itx} - 1 - it\sin x}{x^2} M(dx)\right\}$$

où μ est un réel et$M$ est une mesure qui attribue une masse finie à tout intervalle fini et telle que les deux intégrales suivantes

$$M^+(x) = \int_x^{+\infty} y^{-2} M(dy) \quad et \quad M^-(-x) = \int_{-\infty}^{-x} y^{-2} M(dy)$$

sont convergentes pour tout $x > 0$ La démonstration est détaillée dans Feller (1971, pages 554–565).

Remarque :

La fonction caractéristique caractérise la loi de probabilité d'une variable aléatoire, i.e. :
-si on connait la fonction caractéristique de X, on connait la loi de X,
-si deux variables aléatoires ont même fonction caractéristique, c'est qu'elles ont même loi de probabilité
Si ϕ est la fonction caractéristique d'une variable aléatoire$X$ et si ϕ est intégrable, c'est-à-dire si :

$$\int_{-\infty}^{+\infty} |\varphi(t)| dt < +\infty$$

alors X admet une densité de probabilité f dééfinie sur $\mathbf{R}$ par :

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} \varphi(t) dt \quad \forall x \in \mathbf{R}$$

Exemples :

Beaucoup de lois connues sont indéfiniment divisibles. Si une V.A.R. X suit :
-la loi normale $\mathcal{N}(m, \sigma^2)$: sa fonction caractéristique s'écrit :

$$\varphi_X(t) = \int_{-\infty}^{+\infty} e^{itx} f(x) dx \tag{2.1}$$

$$= \sum_{j=1}^{n} e^{itx_j} \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}\frac{(t-m)^2}{\sigma^2}} \tag{2.2}$$

$$= \exp\left\{imt - \frac{t^2\sigma^2}{2}\right\} \tag{2.3}$$

$$= \left[\exp\left\{i\frac{m}{n}t - \frac{t^2\frac{\sigma^2}{n}}{2}\right\}\right]^n \tag{2.4}$$

comme puissance n ème de la fonction caractéristique d'une loi normale $\mathcal{N}\left(\frac{m}{n}, \frac{\sigma^2}{n}\right)$

- la loi de Cauchy $\mathcal{C}(c)$:sa fonction caractéristique s'écrit :

$$\varphi_X(t) = \int_{-\infty}^{+\infty} e^{itx} f(x) dx \tag{2.5}$$
$$= \mathbb{E}\left(e^{itX}\right) \tag{2.6}$$
$$= \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{e^{itx}}{x^2+1} dx \tag{2.7}$$
$$= \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\cos(tx)}{x^2+1} dx + \frac{i}{\pi}\int_{-\infty}^{\infty} \frac{\sin(tx)}{x^2+1} dx \tag{2.8}$$
$$= \operatorname{ch} t - \frac{2}{\pi}\operatorname{sh} t \int_0^{\infty} \frac{\sin v}{v} dv \tag{2.9}$$

et puisque,

$$\int_0^{\infty} \frac{\sin v}{v} dv = \frac{\pi}{2}$$

pour tout t positif

$$\mathbf{E}\left(e^{itx}\right) = \operatorname{ch} t - \operatorname{sh} t = e^{-t}$$

et, en raison de la parité de la fonction, pour tout t réel,

$$\mathbf{E}\left(e^{itx}\right) = e^{-|t|}$$

$$\begin{aligned}\varphi_X(t) &= \exp(-c|t|)\\ &= \left[\exp\left(-\frac{c}{n}|t|\right)\right]^n\end{aligned}$$

comme puissance n^{eme} de la fonction caractéristique d'une loi de Cauchy $\mathcal{C}\left(\frac{c}{n}\right)$

- la loi de Poisson $\mathcal{P}(\lambda)$:

$$\begin{aligned}\mathbf{E}\left[e^{itX}\right] &= \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} e^{itk}\\ &= e^{-\lambda} \sum_{k=0}^{\infty} \frac{\left(\lambda e^{itk}\right)^k}{k!}\\ &= e^{-\lambda} \exp\left(\lambda e^{it}\right)\\ &= \left[\exp\left\{\frac{\lambda}{n}\left(e^{it}-1\right)\right\}\right]^n\end{aligned}$$

comme puissance n^{eme} de la fonction caractéristique d'une loi de Poisson $\mathcal{P}\left(\frac{\lambda}{n}\right)$

- la loi Gamma $\Gamma(r, \lambda)$:sa fonction caractéristique s'écrit :

$$\varphi_X(t) = \int_{-\infty}^{+\infty} e^{itx} f(x) dx \tag{2.10}$$
$$= \int_{-\infty}^{+\infty} e^{itx} \frac{\lambda^r}{\Gamma(r)} e^{-\lambda x} x^{r-1} \mathbf{1}_{]0,+\infty[}(x) \tag{2.11}$$

telle que

$$\Gamma(r) = \int_0^{+\infty} e^{-x} x^{r-1} dx$$

$$\begin{aligned}
\Phi_X(t) &= \mathrm{E}\left(e^{itx}\right) && (2.12)\\
&= \frac{\lambda^r}{\Gamma(r)} \int_0^{+\infty} e^{-\lambda x} e^{itx} x^{r-1} dx && (2.13)\\
&= \frac{\lambda^r}{\Gamma(r)} \int_0^{+\infty} e^{-(\lambda - it)x} x^{r-1} dx && (2.14)\\
&= \frac{\lambda^r}{(\lambda - it)^r} && (2.15)
\end{aligned}$$

Ainsi

$$\begin{aligned}
\varphi_X(t) &= \frac{1}{\left(1 - \frac{it}{\lambda}\right)^r}\\
&= \left(\frac{1}{\left(1 - \frac{it}{\lambda}\right)^{\frac{r}{n}}}\right)^n
\end{aligned}$$

comme puissance n^{eme} de la fonction caractéristique d'une loi Gamma $\Gamma\left(\frac{r}{n}, \lambda\right)$

***Définition* 2.1.2.**

a) Soient a, b deux nombres réels positifs et X_1, X_2 deux variables aléatoires indépendantes de même loi que X . Il existe $c \in \mathbf{R}^+$ et $d \in \mathbf{R}$ qui satisfont :

$$aX_1 + bX_2 = cX + d \quad \textit{en distribution}$$

b) Soit n un entier positif, et $X_1, X_2, \ldots, X_n$ n copies indépendantes de X . Alors il existe $c_n \in \mathbf{R}^+$ et $d_n \in \mathbf{R}$ tels que $X_1 + X_2 + \ldots + X_n = c_n X + d_n$ en distribution Lorsque $d_n = 0$, on parle de distribution strictement stable.

Remarque :
On peut montrer qu'il existe une constante α, $0 < \alpha \leq 2$, telle que $a_k = k^{1/\alpha}$ pour $k \in \mathbf{N}^*$
La démonstration est détaillée dans Feller (1971, pages 170-171)

***Proposition* 2.1.1.**
Si X est stable, X est indéfiniment divisible.

Preuve
Il suffit de prendre des V . A . R .

$$Y_j = \frac{X_j - \frac{b_n}{n}}{a_n}, j = 1, \ldots, n$$

Comme les X_j sont indépendantes, les Y_j sont aussi indépendantes et

$$Y_1 + \cdots + Y_n = \frac{X_1 + \cdots + X_n}{a_n} - \frac{b_n}{a_n}$$

Or $X_1 + \cdots + X_n \stackrel{d}{=} a_n X + b_n$, d'où $\dot{}\ Y_1 + \cdots + Y_n \stackrel{d}{=} X$

Remarque :
La réciproque est fausse (voir dans le contre exemple suivant, , la loi de Poisson).

En effet, soient X_1 et X_2 deux V . A . R . suivant une loi de Poisson. Supposons que X_1 et X_2 sont stables, alors il existe $a > 0$ et b tels que :

$$X_1 + X_2 \stackrel{d}{=} aX_1 + b$$

Par égalité des moyennes et des variances, nous pouvons voir que

$$\left\{ \begin{array}{l} 2\lambda = a\lambda + b \\ 2\lambda = a^2\lambda \end{array} \right. \Longrightarrow \left\{ \begin{array}{l} b = (2-\sqrt{2})\lambda \\ a = \sqrt{2} \end{array} \right.$$

ce qui entraine une contradiction car $X_1 + X_2$ a ses valeurs uniquement dans $\mathbf{N}$ alors que $\sqrt{2}X_1 + (2-\sqrt{2})\lambda$ n'a pas que des valeurs dans $\mathbf{N}$.

***Théorème* 2.1.3.**
Une V.A.R. X est la limite en distribution des V.A.R. $\frac{(X_1+\cdots+X_n-b_n)}{a_n}, a_n > 0$ *si et seulement si* X *est stable.*
La démonstration est détaillée dans Shiryayev (1984, pages 338–339)

***Corollaire* 2.1.1.**
Si X a une distribution stable, alors sa fonction caractéristique s'écrit :

$$\phi_X(t) = \left\{ \begin{array}{ll} \exp\left(-\sigma^\alpha |t|^\alpha \left(1 - i\beta \operatorname{sgn}(t) \tan \frac{\pi\alpha}{2}\right) + i\mu t\right) & \textit{pour } \alpha \neq 1 \\ \exp\left(-\sigma |t| \left(1 + i\beta \frac{2}{\pi} \operatorname{sgn}(t) \log |t|\right) + i\mu t\right) & \textit{pour } \alpha = 1 \end{array} \right.$$

où les paramètres satisfont les restrictions suivantes : $\alpha \in [0,2], \sigma \in \mathcal{R}_0^+$
$\beta \in [-1,1], \mu \in \mathcal{R}$.

Une variable aléatoire qui a cette expression pour sa FC est notée $X \sim S_\alpha(\sigma, \beta, \mu)$.

Mais cette représentation de la FC , qui s'appelle paramétrisation standard,a le désavantage de ne pas être continue en ses paramètres. En fait, il y a discontinuité en les points où $\alpha = 1$ et $\beta = 0$. Par ailleurs il existe d'autres paramétrisations de la FC plus adaptées aux différents problèmes.

Ici, on fait référence à deux paramétrisations proposées par Zolotarev

$$\phi_X(t) = \left\{ \begin{array}{ll} \exp\left(-\sigma^\alpha |t|^\alpha \left[1 + i\beta \operatorname{sgn}(t) \tan \frac{\pi\alpha}{2} \left(|\sigma t|^{1-\alpha} - 1\right)\right] + i\mu_0 t\right) & \text{pour } \alpha \neq 1 \\ \exp\left(-\sigma |t| \left[1 + i\beta \frac{2}{\pi} \operatorname{sgn}(t) (\log |t| + \log \sigma)\right] + i\mu_0 t\right) & \text{pour } \alpha = 1 \end{array} \right.$$

Cette représentation S^0 se note $X \sim S_\alpha^0(\sigma, \beta, \mu_0)$. Les paramètres α, β et σ de la paramétrisation S^0 sont les mêmes que ceux de la paramétrisation standard, mais μ et μ_0 sont reliés par :

$$\mu = \left\{ \begin{array}{ll} \mu_0 - \beta\sigma \tan \frac{\pi\alpha}{2} & \text{pour } \alpha \neq 1 \\ \mu_0 - \beta \frac{2}{\pi} \sigma \log \sigma & \text{pour } \alpha = 1 \end{array} \right.$$

Cette paramétrisation est très importante parce que la fonction caractéristique, la densité et la fonction cumulative de répartition sont continues par rapport aux quatre paramètres.

Donc, elle est bien conditionnée numériquement pour le calcul. Une autre paramétrisation $\mathcal{S}^1$ est donnée par :

$$\phi_X(t) = \begin{cases} \exp\left(-\sigma_2^\alpha |t|^\alpha \exp\left[-i\beta_2 sgn(t)\frac{\pi}{2}K(\alpha)\right] + i\mu t\right) & \text{pour } \alpha \neq 1 \\ \exp\left(-\sigma_2 |t|^\alpha \exp\left[-i\beta_2 \operatorname{sgn}(t) \log|t|\right) + i\mu t\right) & \text{pour } \alpha = 1 \end{cases}$$

où

$$K(\alpha) = \alpha - 1 + \operatorname{sgn}(1-\alpha) = \begin{cases} \alpha & \text{pour } \alpha < 1 \\ \alpha - 2 & \text{pour } \alpha > 1 \end{cases}$$

Les paramètres α et μ sont les mêmes que pour la paramétrisation standard, les autres paramètres satisfont les relations suivantes :

$$\sigma_2 = \sigma\left(1 + \beta^2 \tan^2 \frac{\pi\alpha}{2}\right)^{\frac{1}{2\alpha}}$$

$$\tan\left(\beta_2 \frac{\pi K(\alpha)}{2}\right) = \beta \tan\frac{\pi\alpha}{2}$$

pour $\alpha \neq 1$ et,

$$\sigma_2 = \frac{2}{\pi}\sigma \text{ et } \beta_2 = \beta$$

pour $\alpha = 1$

Une loi stable est définie par quatre paramètres :

1) Le paramètre α appelé exposant caractéristique ou indice de stabilité décrit la forme de la distribution ou le degré d'épaisseur de la queue de distribution $(0 < \alpha \leq 2)$. Plus α est petit, plus les queues de la distribution sont épaisses. Autrement dit, plus est petit, plus nous constatons l'existence de tr̀es grandes fluctuations. Une distribution gaussienne a la valeur maximum de α soit $\alpha = 2$.

Plus le paramètre α est petit, plus la courbe de la densité est pointue et a des queues de distribution épaisse.

2) Le paramètre β donne une idée de l'asymétrie de la distribution. C'est le paramètre d'asymétrie. Si β est égal à -1 (resp +1) la distribution est totalement asymétrique à gauche (resp. à droite). Lorsque β vaut zéro alors la distribution est symétrique.

Lorsque β est positif (resp. négatif), le mode est à gauche (resp. à droite) de la moyenne. Lorsque β est positif (resp. négatif), la queue de distribution est plus épaisse à droite (resp. à gauche).

3) Le paramètre σ est appelé facteur d'échelle. Plus σ est grand, plus les données sont volatiles. Le paramètre σ permet de cintrer plus ou moins le corps de la distribution.

4) Le paramètre de localisation μ correspond, pour α supérieur à 1, à la moyenne de la loi de distribution si, $\beta = 0$ alors μ est la médiane. Dans les autres cas le paramétre μ ne peut pas être interprété.

2.1.1 Les inconvénients de la loi stable :

Le principal inconvénient est que pour la plupart des lois connues, nous avons une forme explicite de la densité (normale,Cauchy, gamma,...). Pour la loi α-stable, nous n'avons que la forme explicite de la fonction caractéristique. A l'aide de la transformée inverse de la fonction caractéristique, donnée par

$$f_X(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \exp(-itx)\varphi_X(t)dt$$

nous pouvons obtenir f sous la forme d'une intégrale

$$f_X(x) = \frac{1}{\pi}\int_{0}^{+\infty} \exp\left(-t^\alpha\right)\cos\left[xt + \beta t^\alpha w(t,\alpha)\right]dt$$

Et les densités des lois stables sont inconnues sauf dans trois cas :
- La distribution gaussienne $S_2(0,\sigma,\mu)$ où $f(x) = \frac{1}{2\sigma\sqrt{\pi}}\exp\left(-\frac{(x-\mu)^2}{4\sigma^2}\right)$
- La distribution de Cauchy $\mathrm{S}_1(0,\sigma,\mu)$ où $f(x) = \frac{2\sigma}{\pi((x-\mu)^2+4\sigma^2)}$
- La distribution de Lévy $\mathrm{S}_{1/2}(1,\sigma,\mu)$
  $f(x) = \left(\frac{\sigma}{2\pi}\right)^{\frac{1}{2}}(x-\mu)^{-\frac{3}{2}}\exp\left(-\frac{\sigma}{2(x-\mu)}\right)\times \mathrm{I}_{]\mu,x[}(x)$

***Théorème* 2.1.4.**
Soient X_1 et X_2 deux variables aléatoires stables et indépendantes, avec $X_i \sim S_\alpha(\sigma_i,\beta_i,\mu_i)$, pour i=1,2, alors $X_1+X_2 \sim S_\alpha(\sigma,\beta,\mu)$ où $\sigma = (\sigma_1^\alpha+\sigma_2^\alpha)^{\frac{1}{\alpha}}$, $\beta = \frac{\beta_1\sigma_1^\alpha+\beta_2\sigma_2^\alpha}{\sigma_1^\alpha+\sigma_2^\alpha}$, $\mu = \mu_1+\mu_2$

Preuve

$$\begin{aligned}\phi_{X_1+X_2}(t) &= E\left(\exp\left(it\left(X_1+X_2\right)\right)\right)\\ &= \exp\left(-\sigma_1^\alpha|t|^\alpha\left[1-i\beta_1\operatorname{sgn}(t)\tan\tfrac{\pi\alpha}{2}\right]+i\mu_1 t\right)\times\exp\left(-\sigma_2^\alpha|t|^\alpha\left[1-i\beta_2\operatorname{sgn}(t)\tan\tfrac{\pi\alpha}{2}\right]+i\mu_2 t\right)\\ &= \exp\left(-\sigma_1^\alpha|t|^\alpha\left[1-i\beta_1\operatorname{sgn}(t)\tan\tfrac{\pi\alpha}{2}\right]-\left(-\sigma_2^\alpha|t|^\alpha\left[1-i\beta_2\operatorname{sgn}(t)\tan\tfrac{\pi\alpha}{2}\right]\exp\left(i\left(\mu_1+\mu_2\right)t\right)\right.\right.\\ &= \exp\left(\left(-\sigma_1^\alpha-\sigma_2^\alpha\right)|t|^\alpha+i|t|^\alpha\left(\beta_1\sigma_1^\alpha+\beta_2\sigma_2^\alpha\right)\operatorname{sgn}(t)\tan\tfrac{\pi\alpha}{2}\left(\exp\left(i\left(\mu_1+\mu_2\right)t\right)\right.\right.\\ &= \exp((-\sigma_1^\alpha-\sigma_2^\alpha)\left[1-\left(i\tfrac{\beta_1\sigma_1^\alpha+\beta_2\sigma_2^\alpha}{\sigma_1^\alpha+\sigma_2^\alpha}\right)\operatorname{sgn}(t)\tan\tfrac{\pi\alpha}{2}\right]\exp\left(i\left(\mu_1+\mu_2\right)t\right)\end{aligned}$$

***Théorème* 2.1.5.**
Si $X \sim S_\alpha(\sigma,\beta,\mu)$ et $a\in R$, alors : $X+a\sim S_\alpha(\sigma,\beta,\mu+a)$

Preuve

$$\begin{aligned}\phi_{X+a}(t) &= E(\exp(it(X+a))\\ &= \exp\left(-\sigma^\alpha|t|^\alpha\left[1-i\beta\operatorname{sgn}(t)\tan\frac{\pi\alpha}{2}\right]+i\mu t\right)\times\exp(iat)\\ &= \exp\left(-\sigma^\alpha|t|^\alpha\left[1-i\beta\operatorname{sgn}(t)\tan\frac{\pi\alpha}{2}\right]\right)\times\exp(i(\mu+a)t)\end{aligned}$$

***Théorème* 2.1.6.**

Si $X \sim S_\alpha(\sigma,\beta,\mu)$ et $a \in R$, alors :

$$aX \sim \begin{cases} S_\alpha(|a|\sigma,(\mathrm{sgn}(a)\beta,a\mu) & \textit{pour } \alpha \neq 1 \\ S_1\left(|a|\sigma,\mathrm{sgn}(a)\beta,a\mu-\frac{2}{\pi}a(\log|a|)\sigma\beta\right) & \textit{pour } \alpha = 1 \end{cases}$$

***Théorème* 2.1.7.**

$\phi_{aX}(t) = E(\exp(it(aX))$
$=\exp(-a\sigma^\alpha|t|^\alpha[1-i\beta\,\mathrm{sgn}(t)\tan\frac{\pi\alpha}{2}]+ia\mu t)$
$=\exp(-a\sigma^\alpha|t|^\alpha - i\,\mathrm{sgn}(a)\beta\,\mathrm{sgn}(t)\tan\frac{\pi\alpha}{2}+ia\mu t)$

***Théorème* 2.1.8.**

si $X=(X_1,\ldots,X_d)$ est un vecteur α-stable, alors, pour tous réels $b_1,\ldots,b_d$, la variable aléatoire réelle $\sum_{l=1}^d b_l X_l$ est α -stable.

Preuve

$$\phi_{b_1X_1+b_2X_2}(t) = E\left(\exp\left(it\left(b_1X_1+b_2X_2\right)\right)\right)$$

$$= \exp\left(-b_1\sigma_1^\alpha|t|^\alpha\left[1-i\beta_1\,\mathrm{sgn}(t)\tan\tfrac{\pi\alpha}{2}\right]+ib_1\mu_1t\right)\times\exp\left(-b_2\sigma_2^\alpha|t|^\alpha\left[1-i\beta_2\,\mathrm{sgn}(t)\tan\tfrac{\pi\alpha}{2}\right]+ib_2\mu_2t\right)$$

$$= \exp\left(-b_1\sigma_1^\alpha|t|^\alpha\left[1-i\beta_1\,\mathrm{sgn}(t)\tan\tfrac{\pi\alpha}{2}\right]-b_2\sigma_2^\alpha|t|^\alpha\left[1-i\beta_2\,\mathrm{sgn}(t)\tan\tfrac{\pi\alpha}{2}\right]\times\exp\left(i\left(b_1\mu_1+b_2\mu_2\right)t\right)\right.$$

$$= \exp\left(\left(-b_1\sigma_1^\alpha-b_2\sigma_2^\alpha\right)|t|^\alpha+i|t|^\alpha\left(b_1\beta_1\sigma_1^\alpha+b_2\beta_2\sigma_2^\alpha\right)\mathrm{sgn}(t)\tan\tfrac{\pi\alpha}{2}\left(\exp\left(i\left(b_1\mu_1+b_2\mu_2\right)t\right)\right.\right.$$

$$= \exp\left(\left(-b_1\sigma_1^\alpha-b_2\sigma_2^\alpha\right)\right)\left[1-\left(i\frac{b_1\beta_1\sigma_1^\alpha+b_2\beta_2\sigma_2^\alpha}{\sigma_1^\alpha+\sigma_2^\circ}\right)\mathrm{sgn}(t)\tan\tfrac{\pi\alpha}{2}\right]\times\exp\left(i\left(b_1\mu_1+b_2\mu_2\right)t\right)$$

***Théorème* 2.1.9.**

Soit X une V.A.R. $S_\alpha(\mu,\beta,\gamma)$,

$$\begin{cases} \lim_{t\to+\infty} t^\alpha\mathbf{P}(X>t) = \gamma C(\alpha)\frac{1+\beta}{2} \\ \lim_{t\to+\infty} t^\alpha\mathbf{P}(X<-t) = \gamma C(\alpha)\frac{1-\beta}{2} \end{cases}$$

où $C(\alpha) = \left(\int_0^{+\infty} x^{-\alpha}\sin x dx\right)^{-1}$

La démonstration est détaillée dans Samorodnitsky et Taqqu (1994, pages 16–18).

***Théorème* 2.1.10.**

Si X suit une loi $S_\alpha(\mu,\beta,\gamma)$

1. *Si* $\alpha=2, \forall p, \mathbf{E}|X|^p<+\infty$
2. *Si* $0<\alpha<2, \begin{cases} \forall 0\leqslant p<\alpha, \mathbb{E}|X|^p<+\infty \\ \forall p\geqslant\alpha, \mathbb{E}|X|^p=+\infty \end{cases}$

2.2 Les lois stables multivariées

***Définition* 2.2.1.**

Le vecteur aléatoire $X=(X_1,\ldots,X_n)$ est alpha-stable si pour tout k et pour tout vecteur $\underline{X}^{(1)},\ldots,\underline{X^{(k)}}$ de même loi que $\underline{X}$, il existe $a_k>0$ et $D^{(k)}$ tels que :

$$\underline{X}^{(1)}+\cdots+X^{(k)} \stackrel{d}{=} a_k\underline{X}+D^{(k)}$$

Remarque :

1. L'égalité en distribution des vecteurs précédents entraine l'égalité en distribution de chaque composante, c'est-à-dire

$$X_j^{(1)} + \cdots + X_j^{(k)} \stackrel{d}{=} a_k X_j + D_j^{(k)}$$

   .

2. Une conséquence de la proposition suivante est que le paramètre a_k joue le même rôle que dans le cas univarié. Il existe alors une constante $\alpha, 0 < \alpha \leqslant 2$, telle que $a_k = k^{1/\alpha}$.

***Proposition* 2.2.1.**
Pour $\underline{X}$ un vecteur α -stable, toute combinaison linéaire des composantes de $\underline{X}$ est une V.A.R. α -stable.

Preuve
Soit $Y = \sum_{j=1}^n l_j X_j$. Prenons alors k réplications $Y_1 = \sum_{j=1}^n l_j X_j^{(1)}, \ldots, Y_k = \sum_{j=1}^n l_j X_j^{(k)}$ et calculons la distribution de la somme des Y_j :

$$\begin{aligned}
Y_1 + \cdots + Y_k &\stackrel{d}{=} \sum_{j=1}^n l_j X_j^{(1)} + \cdots + \sum_{j=1}^n l_j X_j^{(k)} \\
&\stackrel{d}{=} l_j \left(X_j^{(1)} + \cdots + X_j^{(k)} \right) \\
&\stackrel{d}{=} \sum_{j=1}^n l_j \left(a_k X_j + D_j^{(k)} \right) \text{ d'après la remarque précédente} \\
&\stackrel{d}{=} a_j \sum_{j=1}^n l_j X_j + \sum_{j=1}^n l_j D_j^{(k)} \\
&\stackrel{d}{=} a_k Y + b_k
\end{aligned}$$

Donc la V.A.R Y est bien α-stable.

3

Estimation des paramètres :

Un grand nombre de méthodes existent pour estimer les paramères d'une loi stable. **DuMouchel** (1971) a développé un algorithme qui utilise le principe du maximum de vraisemblance. **McCulloch** (1979) a développé un algorithme rapide qui estime les paramètres d'une loi stable symétrique (b = 0) à l' aide du maximum de vraisemblance. . **Zolotarev** (1980) a estimé les paramètres $\alpha, \beta et \gamma$ par la méthode des moments après avoir fixé le paramètre de localisation (μ connu).

Notre objectif est de savoir si les chroniques financières sont mieux modélisé es par une loi normale ou par une loi stable . Pour cela nous allons utiliser la méthode de **McCulloch** (datant de 1986), qui a étendu la méthode de Fama et Roll aux cas $\beta \in [-1; 1]$ et $\alpha \in [0, 6; 2]$.
Cette méthode est efficace en ce qui concerne la précision des estimateurs par rapport au temps de calcul nécessaire à l'estimation. Elle est normalement utilisé e pour fournir des valeurs initiales des paramètres pour démarrer des méthodes plus efficaces mais plus lourdes et plus longues

3.1 Méthode de McCulloch :

Soit F la fonction de répartition d'une variable aléatoire stable $X \sim S_\alpha(\sigma, \beta, \mu)$, et soit x_p le quantile d'ordre p, c'est-à-dire, $F(x_p) = p$, et $\hat{x}_p$ le quantile empirique correspondant. Pour éviter une fausse asymétrie des petits échantillons, une correction est nécessaire : si les x_i sont ordonnées de façon croissante cette correction se fait en posant

$$\hat{x}_{q(i)} = x_i \text{ où } \; q(i) = \frac{2i - 1}{2n}$$

Puis, on effectue une interpolation linéaire pour obtenir $\hat{x}_p$ à partir de $\hat{x}_{q(i)}\hat{x}_{q(i+1)}$
où $q(i) \leqslant p \leqslant q(i+1)$ L'estimateur obtenu x_p est un estimateur convergent de x_p.
McCulloch définit :

$$\nu_\alpha = \frac{x_{0.95} - x_{0.05}}{x_{0.75} - x_{0.25}}$$

$$\nu_\beta = \frac{x_{0.95} + x_{0.05} - 2x_{0.5}}{x_{0.95} - x_{0.05}}$$

et montre que ces indices ne dépendent pas de σ ni de μ . De plus ils sont respectivement des fonctions décroissante et croissante de α et β. Cette relation peut s'inverser, donc les paramètres α et β peuvent être vus comme des fonctions de ν_α et ν_β, soit :

$$\alpha = \phi_1\left(\nu_\alpha, \nu_\beta\right), \beta = \phi_2\left(\nu_\alpha, \nu_\beta\right)$$

McCulloch construit ensuite deux nouveaux indices :

$$\nu_\sigma = \frac{x_{0.75} - x_{0.25}}{\sigma}$$

et,

$$\nu_\zeta = \frac{\zeta - x_{0.5}}{\sigma}$$

où

$$\zeta = \begin{cases} \mu + \beta\sigma \tan\frac{\pi\alpha}{2} & \text{pour } \alpha \neq 1 \\ \mu & \text{pour } \alpha = 1 \end{cases}$$

ν_σ et ν_ζ dépendent seulement de α et β c'est-à-dire

$$\nu_\sigma = \phi_3(\alpha, \beta)$$

et

$$\nu_\zeta = \phi_4(\alpha, \beta)$$

Si les fonctions ϕ_i, pour i=1,2,3,4, sont connues, alors l'algorithme d'estimation des quatre paramètres peut être défini comme ci-dessous.

Procédure concrète de l'algorithme :

- Ordonner l'échantillon.
- Calculer les quantiles empiriques $\hat{x}_{0.05}, \hat{x}_{0.25}, \hat{x}_{0.5}, \hat{x}_{0.75}$ et $\hat{x}_{0.95}$.
- Estimer les indices ν_α et ν_β :

$$\hat{\nu}_\alpha = \frac{\hat{x}_{0.95} - \hat{x}_{0.05}}{\hat{x}_{0.75} - \hat{x}_{0.25}}$$

$$\hat{\nu}_\beta = \frac{\hat{x}_{0.95} + \hat{x}_{0.05} - 2\hat{x}_{0.5}}{\hat{x}_{0.95} - \hat{x}_{0.05}}$$

- Estimer α et β

$$\hat{\alpha} = \phi_1\left(\hat{\nu}_\alpha, \hat{\nu}_\beta\right), \hat{\beta} = \phi_2\left(\hat{\nu}_\alpha, \hat{\nu}_\beta\right)$$

- Estimer σ

$$\hat{\sigma} = \frac{\hat{x}_{0.75} - \hat{x}_{0.25}}{\phi_3(\hat{\alpha}, \hat{\beta})}$$

- Estimer ζ

$$\hat{\zeta} = \hat{x}_{0.5} + \hat{\sigma}\phi_4(\hat{\alpha}, \hat{\beta})$$

- Estimer μ

$$\hat{\boldsymbol{\mu}} = \hat{\zeta} - \hat{\boldsymbol{\beta}}\hat{\sigma}\tan\frac{\pi\hat{\boldsymbol{\alpha}}}{2}$$

Caractéristiques de la méthode
Comme $\hat{x_p}$ est un estimateur convergent et asymptotiquement de loi normale de x_p , et que les fonctions ϕ_i sont continues, alors les estimateurs des paramètres sont convergents et asymptotiquement de loi normale. Le point clé de la méthode est le calcul des fonctions ϕ_i, $i = 1, 2, 3, 4$. On peut construire des tables pour un réseau de points et interpoler bilinéairement, en principe les tables de DuMouchel peuvent être utilisées. Sans considérer le travail préliminaire de génération de ces tables, la méthode est facile à implémenter. La complexité de la méthode est de l'ordre de $O(n \log n)$.

3.2 Maximum de vraisemblance :

***Définition* 3.2.1.**
La méthode de maximum de vraisemblance peut s'appliquer aussi au cas de variables aléatoires stables. Ceci a été étudié depuis longtemps. Le point clé pour appliquer la méthode est le calcul de la densité d'une variable aléatoire stable parce qu'il n'existe pas de formule exacte pour la densité de ces lois.

Énoncé :

Soit une famille paramétrique de distributions de probabilités D_θ dont les élèments sont associés soit à une densité de probabilité connue (distribution continue), soit à une fonction de masse connue (distribution discrète), notée $f(x|\theta)$. On tire un échantillon de n valeurs $x_1, x_2, \ldots, x_n$ de la distribution, et l'on calcule la densité de probabilité associée aux données observées

$$f_\theta\left(x_1, \ldots, x_n; \theta\right) = \prod_{i=1}^{n} f\left(x_i|\theta\right)$$

Ceci étant une fonction de θ avec $x_1, \ldots, x_n$ fixés, c'est une **vraisemblance**.

$$L(\theta) = f_\theta\left(x_1, \ldots, x_n; \theta\right)$$

Lorsque θ n'est pas observable, la méthode du maximum de vraisemblance utilise les valeurs de θ qui maximisent $L(\theta)$ estimateur de θ : c'est l'estimateur du maximum de vraisemblance de θ noté $\hat{\theta}$. Par exemple dans le cas du produit discret, on effectue un tirage de n valeurs, il faut donc trouver le paramètre qui maximise la probabilité d'avoir tiré ce tirage.

Soit X une variable aléatoire réelle, de loi discrète ou continue, dont on veut estimer un paramètre θ . On note $\mathcal{D}_\theta$ cette famille de lois paramétriques. Alors on définit une fonction f telle que :

$$f(x; \theta) = \begin{cases} f_\theta(x) & \text{si } X \text{ est une v.a. continue} \\ P_\theta(X = x) & \text{si } X \text{ est ure v.a. discrète} \end{cases}$$

$f_\theta(x)$ représente la densité de X (où θ apparaît)

On appelle vraisemblance de θ au vu des observations $(x_1, \ldots, x_i, \ldots, x_n)$ d'un n-échantillon indépendamment et identiquement distribué selon la loi $\mathcal{D}_\theta$, le nombre :

$$L\left(x_1,\ldots,x_i,\ldots,x_n;\theta\right)=f\left(x_1;\theta\right)\times f\left(x_2;\theta\right)\times\ldots\times f\left(x_n;\theta\right)=\prod_{i=1}^{n}f\left(x_i;\theta\right)$$

On cherche à trouver le maximum de cette vraisemblance pour que les probabilités des réalisations observées soient aussi maximum. Ceci est un problème d'optimisation. On utilise généralement le fait que si L est dérivable (ce qui n'est pas toujours le cas) et si L admet un maximum global en une valeur $\theta=\hat{\theta}$, alors la dérivée première s'annule en $\theta=\hat{\theta}$ et que la dérivée seconde est négative.
Réciproquement, si la dérivée première s'annule en $\theta=\hat{\theta}$ et que la dérivée seconde est strictement négative en $\theta=\hat{\theta}$, alors $\theta=\hat{\theta}$ est un maximum local de $L\left(x_1,\ldots,x_i,\ldots,x_n;\theta\right)$. Il est alors nécessaire de vérifier qu'il s'agit bien d'un maximum global. La vraisemblance étant positive et le logarithme népérien une fonction croissante, il est équivalent et souvent plus simple de maximiser le logarithme népérien de la vraisemblance (le produit se transforme en somme, ce qui est plus simple à dériver).
On peut facilement construire la statistique $Y_n=\Theta$ qui est l'estimateur voulu Ainsi en pratique :
• La condition nécessaire

$$\frac{\partial L\left(x_1,\ldots,x_i,\ldots,x_n;\theta\right)}{\partial\theta}=0$$

ou

$$\frac{\partial \ln L\left(x_1,\ldots,x_i,\ldots,x_n;\theta\right)}{\partial\theta}=0$$

permet de trouver la valeur $\theta=\hat{\theta}$.
$\cdot\theta=\hat{\theta}$ est un maximum local si la condition suffisante est remplie au point critique $\theta=\hat{\theta}$:

$$\frac{\partial^2 L\left(x_1,\ldots,x_i,\ldots,x_n;\theta\right)}{\partial\theta^2}<0$$

ou

$$\frac{\partial^2 \ln L\left(x_1,\ldots,x_i,\ldots,x_n;\theta\right)}{\partial\theta^2}<0$$

Exemple :(Loi de Bernouilli)
L'ensemble des valeurs possibles est $\{0,1\}$. Le paramètre inconnu est p. Si $(x_1,\ldots,x_n)\in\{0,1\}^n$ est un échantillon, la vraisemblance vaut :

$$L\left(x_1,\ldots,x_n,p\right)=p\Sigma x_i(1-p)(n-\Sigma x_i)$$

Son logarithme est :

$$\log\left(L\left(x_1,\ldots,x_n,p\right)\right)=\left(\sum x_i\right)\log p+\left(n-\sum x_i\right)\log(1-p)$$

La dérivée par rapport à p est :

$$\frac{\partial\log\left(L\left(x_1,\ldots,x_n,p\right)\right)}{\partial p}=\left(\sum x_i\right)\frac{1}{p}-\left(n-\sum x_i\right)\frac{1}{1-p}$$

Elle s'annule pour :

$$\widehat{p}=\frac{\sum x_i}{n}$$

La dérivée seconde est :

$$\frac{\partial^2 \log\left(L\left(x_1, \ldots, x_n, p\right)\right)}{\partial p^2} = -\left(\sum x_i\right)\frac{1}{p^2} - \left(n - \sum x_i\right)\frac{1}{(1-p)^2}$$

Elle est strictement négative, la valeur $\widehat{p}$ est bien un maximum. Si $(X_1, \ldots, X_n)$ est un échantillon de la loi de Bernoulli de paramètre p, l'estimateur du maximum de vraisemblance de p est :

$$\frac{\sum X_i}{n}$$

à savoir la fréquence empirique.

3.3 Méthode des moments

Press a proposé une autre méthode basée sur une transformation de la fonction caractéristique. Pour tout α

$$|\phi(t)| = \exp\left(-\sigma^\alpha |t|^\alpha\right)$$

alors,

$$-\log|\phi(t)| = \sigma^\alpha |t|^\alpha$$

Il y a deux cas à considérer, $\alpha \neq 1$ et $\alpha = 1$ Dans le premier cas, on prend deux valeurs non nulles de $t, t_1 \neq t_2$, alors

$$-\log\left|\phi\left(t_k\right)\right| = \sigma^\alpha \left|t_k\right|^\alpha$$

pour $k = 1, 2$. On résout le système d'équations en α et σ, et l'on remplace $\phi(t)$ par $\hat{\phi}(t)$; on arrive à :

$$\hat{\alpha} = \frac{\log \frac{\log\left|\hat{\phi}(t_1)\right|}{\log\left|\hat{\phi}(t_2)\right|}}{\log \frac{|t_1|}{|t_2|}}$$

et,

$$\log\hat{\sigma} = \frac{\log\left(-\log\left|\hat{\phi}\left(t_2\right)\right|\right)\log|t_1| - \log\left(-\log\left|\hat{\phi}\left(t_1\right)\right|\right)\log|t_2|}{\log\frac{\log\hat{\phi}(t_1)}{\log\hat{\phi}(t_2)}}$$

Pour l 'estimation β et μ on considére $u(t) = \operatorname{Im}(\log\phi(t))$.

$$u(t) = \mu t + \sigma^\alpha |t|^\alpha \beta \operatorname{sgn}(t) \tan\frac{\alpha\pi}{2}$$

On prend deux valeurs non nulles de $t, t_3 \neq t_4$ il vient

$$\frac{u\left(t_k\right)}{t_k} = \mu + \left[\sigma^\alpha \left|t_k\right|^{\alpha-1} \tan\frac{\alpha\pi}{2}\right]\beta$$

pour k=3,4 .Comme

$$\hat{\phi}(t) = \left(\frac{1}{n}\sum_{i=1}^{n}\cos\left(t x_i\right)\right) + i\left(\frac{1}{n}\sum_{i=1}^{n}\sin\left(t x_i\right)\right)$$

alors,

$$\tan \hat{u}(t) = \frac{\sum_{i=1}^{n} \cos(tx_i)}{\sum_{i=1}^{n} \sin(tx_i)}$$

Á partir de cette expression on arrive à

$$\hat{\boldsymbol{\beta}} = \frac{\frac{\hat{u}(t_4)}{t_4} - \frac{\hat{u}(t_3)}{t_3}}{\left[|t_4|^{\hat{\alpha}-1} - |t_3|^{\hat{\alpha}-1}\right] \hat{\sigma}^{\hat{\alpha}} \tan \frac{\hat{\alpha}\pi}{2}}$$

et,

$$\hat{\boldsymbol{\mu}} = \frac{|t_4|^{\hat{\alpha}-1} \frac{\hat{u}(t_3)}{t_3} - |t_3|^{\hat{\alpha}-1} \frac{\hat{u}(t_4)}{t_4}}{\left[|t_4|^{\hat{\alpha}-1} - |t_3|^{\hat{\alpha}-1}\right]}$$

Le cas $\alpha = 1$ est plus simple, en suivant le même raisonnement, on arrive aux résultats suivants :

$$\hat{\sigma} = -\frac{\log\left|\hat{\phi}(t_1)\right|}{|t_1|}$$

$$\hat{\boldsymbol{\beta}} = \frac{\frac{\hat{u}(t_3)}{t_3} - \frac{\hat{u}(t_4)}{t_4}}{\frac{2}{\pi}\hat{\sigma} \log\left|\frac{t_4}{t_3}\right|}$$

$$\hat{\mu} = \frac{\log|t_4| \frac{\hat{u}(t_3)}{t_3} - \log|t_3| \frac{\hat{u}(t_4)}{t_4}}{\log\left|\frac{t_4}{t_3}\right|}$$

La méthode des moments est très facile à implémenter et elle est très efficace en temps de calcul mais très imprécise si l'échantillon n'est pas normalisé.

3.4 Conclusion

l'hypothèse de normalité des variations des prix ou des rendements est mise à mal par certains mouvements extrêmes des marchés. La distribution réelle est plus large sur les bords (queue de distribution épaisse), plus cintrée et plus pointue. Les lois stables approchent mieux les lois des rendements financiers que les lois normales utilisées actuellement dans de nombreux modèles.

Parmi les modèles mathématiques qui prennent en compte la structure fractale des marchés sont les lois stables. Nous avons tout de même travaillé avec ces lois car elles possèdent de nombreuses qualités :

- Les distributions stables permettent de prendre en compte les queues de distribution épaisses observées en pratique sur la loi des rendements et intègrent les discontinuités observées sur le marché . Dans le cas où le paramètre de stabilité α est inférieur à 2, leur variance est infinie (i.e. instable).

- Les lois stables sont stables par combinaison linéaire et ce sont les seules qui s'obtiennent comme limites des sommes linéairement normées de variables i.i.d. Nous pouvons donc utilisé le théorème centrale limite généralisé .

Les lois stables sont des distributions fractales du fait de leur propriété de stabilité par addition. En effet, la somme de deux variables α-stables i.i.d. est une variable stable de même exposant caractéristique a. Les variables stables possèdent ainsi des propriétés d'invariance d'échelle. (Nous pouvons généraliser cette proposition à n variables α-stables i.i.d.).

- Elles sont définies par seulement quatre paramètres, ce qui les rend simple à utiliser en pratique.
- Elles sont une généralisation de la loi normale. Nous n' avons donc pas à réfuter les modèles existants basés sur la loi de Gauss mais seulement à les généraliser.

Bibliographie

[1] Shiryayev A. N. ►*Probability, volume 95 of Graduate Texts in Mathematics.Springer-Verlag.(1984)*

Feller W. ►*Une introduction à la théorie des probabilités et à ses applications. Vol. II. John Wiley et Sons Inc., 2 éd*

[2] Samorodnitsky G. • Taqqu M. S. : ►*Stable non-Gaussian random processes. Stochastic Modeling. Chapman et Hall, New York-London (1994).*

[3] MCCULLOCH J.H. ►*Estimateurs simples et cohérents de paramètres de distribution stables. Communications en statistiques. Simulation et Calcul(1986)*

[4] NOLAN J. ►*algorithm for evaluating stable densities in Zolotarev's (M) parametrization. Preprint American University Washington (1996)*

[5] DUMOUCHEL W.H. ►*Stable Distributions in Statistical Inference. PhD. thesis, Dept. of Statistics, Yale University (1971).*

[6] ZOLOTAREV V.M. ► *On représentation of stable laws by intégrais. Selected Translation in Mathematical Statistics and Probability,(1966). 6, 84-88.*

[7] ZOLOTAREV V.M. ►*One-dimensional stable distributions, Trans. of Math. Monographs, AMS Vol. 65 (1986).*

[8] ama, E.F. et Roll, R. ► *Parameter estimates for symmetric stable distributions. Journal of the American Statistical Association.66(334) :pp. 331–338, 1971*

[9] Kotz, S., Johnson, N.L. et Balakrishnan, N. ► *multivariate distributions : models and applications, volume 1. Wiley-Interscience, 2 édition. (2000)*

[10] Koutrouvelis, I.A. ►*Regression-type estimation of the parameters of stable laws. Journal of the American Statistical Association 75(372) :pp. 918–928, December 1980*