

Ministère de L'Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITÉ SAAD DAHLAB DE BLIDA

Faculté des sciences

Département de Mathématiques



Mémoire de Master

Spécialité : Modélisation Stochastique et Statistique

Par

MOUSSAOUI Meriem & REMIDENE Assia

**Caractérisation des nouvelles distributions flexibles
pour analyser la survie et la fiabilité, avec applications
sur des données réelles**

Devant le jury composé de :

TAMI Omar	Président	MCB.	Univ.Blida 1
FRIHI Redouane	Examineur	MCB.	Univ.Blida 1
RASSOUL Abdelaziz	Directeur de thèse	Prof.	ENSH, Blida

Juin 2025

REMERCIEMENTS

Nous rendons tout d'abord grâce à Dieu, qui nous a donné la force, la volonté et le courage pour accomplir ce modeste travail.

Nous remercions notre promoteur, Monsieur Abdelaziz RASSOUL, pour son soutien et ses conseils continus, ainsi que les membres du jury de soutenance, Monsieur TAMI Omar et Monsieur FRIHI Redouane, pour leurs remarques précieuses. Nous exprimons également notre gratitude aux professeurs de la Faculté de Mathématiques qui ont contribué à élargir nos connaissances scientifiques.

Nous dédions ce travail à nos chères mères, en témoignage de notre profonde reconnaissance pour leur patience, leurs prières et leur soutien indéfectible, ainsi qu'à la mémoire de nos pères, que Dieu ait leurs âmes, pour nous avoir transmis l'amour du savoir, le goût de l'effort et les valeurs du travail. Nos pensées vont aussi à nos sœurs bien-aimées, qui ont toujours été une source de soutien et d'encouragement à chaque étape de notre parcours.

Enfin, nous remercions tous ceux qui nous ont soutenus, collègues et amis, ainsi que toutes les personnes qui ont contribué, de près ou de loin, à la réalisation de ce projet, en espérant que cet effort, bien que modeste, soit une contribution utile, à la hauteur des attentes de ceux qui ont cru en nous.

Au nom de Dieu, j'ai commencé, et par Sa louange j'ai achevé, et par Sa grâce j'ai complété ce parcours.

Je dédie ce succès, fruit de plusieurs années d'efforts et de persévérance, à :

*À celle dont les prières ont été le secret de ma réussite et mon soutien à chaque instant, source d'affection et de tendresse, **ma chère mère**, que Dieu lui accorde longue vie.*

*Ce travail est une dette que je ne pourrai jamais rembourser,
il est pour toi avant d'être pour moi.*

*À l'âme de **mon cher père**, que Dieu ait son âme en paix,
qui a semé en moi l'amour du savoir et s'est toujours efforcé pour moi. . .
J'espère t'avoir honoré ici-bas et dans l'au-delà.*

*À **mes chères sœurs**, Sarah et Yasmine, flammes de ma joie et compagnes de mon chemin,
pour leur sincère bienveillance et leur soutien constant.*

*Merci pour votre patience et votre amour, qui ont été mon moteur.
À vous, toute ma gratitude et mon affection.*

*Et à tous ceux qui m'ont aidée par une parole, un conseil ou une prière
durant mon parcours, parents, voisins et amis.
Je n'oublierai jamais votre bienveillance.*

À vous tous, mes plus sincères remerciements et ma profonde reconnaissance.

Et ma dernière invocation : « Louange à Dieu, Seigneur des mondes ».

Moussaoui Meriem

DÉDICACE 2

À ceux qui, par leur amour, leur soutien et leurs prières, ont illuminé mon chemin et rendu ce rêve possible :

*À **mon père** bien-aimé, que Dieu ait son âme en paix. Pour tout ce qu'il m'a légué de courage, de sagesse et de valeurs inestimables. Sa mémoire est pour moi une lumière qui éclaire mes pas et un refuge dans les moments de doute.*

*À ma **chère mère**, force silencieuse et source intarissable d'amour et de tendresse. Merci pour tes sacrifices, tes prières et ton infinie patience qui ont été mon roc tout au long de ce parcours.*

*À ma sœur **Maria**, confidente et alliée précieuse, pour sa présence réconfortante et son soutien sans faille. Que chaque réussite à venir soit aussi la tienne, car tu fais partie de chacune de mes victoires.*

*À **ma famille**, pour leur bienveillance, leurs encouragements et leur confiance en mes capacités, qui m'ont permis de croire en moi-même et de persévérer malgré les obstacles.*

*À **mes amis**, compagnons de route et artisans discrets de mes sourires, pour leur écoute, leurs mots d'encouragement et la joie qu'ils ont su apporter à mon quotidien.*

À vous tous, qui avez contribué à faire de ce mémoire une réalité, j'adresse ma gratitude la plus profonde et mes remerciements les plus sincères. Puissiez-vous trouver dans ces lignes le reflet de l'amour et de la reconnaissance que je vous porte.

Remidene Assia

TABLE DES MATIÈRES

Introduction Générale	1
1 Notions de base sur les distributions de probabilités	3
1.1 Définition d'une variable aléatoire	3
1.2 Différents types de variables aléatoires	3
1.2.1 Variable aléatoire discrète	3
1.2.1.1 Fonction de masse de probabilité	3
1.2.1.2 Fonction de répartition	4
1.2.1.3 Fonction caractéristique	4
1.2.1.4 Fonction génératrice	4
1.2.1.5 Fonction génératrice des moments	4
1.2.1.6 Valeurs caractéristiques	5
1.2.2 Variable aléatoire continue	5
1.2.2.1 Fonction de densité	5
1.2.2.2 Fonction de répartition	6
1.2.2.3 Fonction caractéristique	6
1.2.2.4 Fonction génératrice des moments	6
1.2.2.5 Valeurs caractéristiques	7
1.3 Lois de probabilité usuelles	7
1.3.1 Cas discret	7
1.3.1.1 Loi uniforme discrète	7
1.3.1.2 Loi Bernoulli	8
1.3.1.3 Loi Binomiale	8
1.3.1.4 Loi de Poisson	8
1.3.1.5 Loi géométrique	9
1.3.2 Cas continue	9
1.3.2.1 Loi uniforme	9

1.3.2.2	Loi normale	9
1.3.2.3	Loi exponentielle	10
1.3.2.4	Loi de gamma	11
1.3.2.5	Loi bêta	11
1.3.2.6	Loi Weibull	12
1.3.2.7	Loi de Pareto	12
1.3.2.8	Loi du Khi-deux	12
1.3.2.9	Loi Student	13
1.3.2.10	Loi de Fisher-Snedecor	13
1.4	Méthodes de Simulation	14
1.4.1	Méthode de l'inversion	14
1.4.2	Méthode d'acceptation-rejet	14
2	Méthodes de génération de nouvelles distributions de probabilités.	15
2.1	Familles de distribution continues	15
2.1.1	Distribution de Pearson	15
2.1.2	Distribution de Burr	16
2.1.3	Distribution de Johnson	17
2.1.4	Distributions lambda généralisées	18
2.1.5	Distributions normales asymétriques	18
2.1.6	Distributions générées par la loi bêta	19
2.1.7	Méthodes de génération de distributions asymétriques	20
2.1.8	Méthode de transformation bêta	20
2.2	Méthode T-X familles de distributions	21
2.2.1	Propriétés de la famille T-X	23
2.2.2	Certaines familles de distributions T-X avec différentes distribu- tions T	25
2.2.2.1	Famille Gamma-X	26
2.2.2.2	Famille Bêta-exponentielle-X	28
2.2.2.3	Famille Weibull-X	31
2.3	Résumé et conclusion	33
3	Étude d'une nouvelle distribution.	36
3.1	Distribution Lomax	36
3.2	Distribution de Lomax transformée par une loi bêta	37
3.3	Propriétés mathématiques de base	40
3.3.1	Fonction quantile	40
3.3.2	Statistiques d'ordre	41
3.4	Estimation des paramètres	42
3.4.1	Méthode de Cramér-von Mises	42

3.4.2	Méthode du maximum de vraisemblance	43
4	Simulations et Applications	46
4.1	Simulation du BTL	48
4.2	Application sur des données réelles	51
CONCLUSION		1

TABLE DES FIGURES

2.1	Graphiques de la distribution gamma-Pareto pour différentes valeurs de paramètres.	28
2.2	Graphiques de la distribution exponentielle-exponentielle-logistique pour différentes valeurs de paramètres.	31
2.3	Graphiques de la distribution Weibull-logistique pour différentes valeurs de paramètres	33
2.4	Sous-familles de la famille de distributions T-X	35
4.1	PDF de la distribution BTL pour différentes valeurs des paramètres . .	47
4.2	CDF de la distribution BTL pour différentes valeurs des paramètres . .	47
4.3	Histogramme des valeurs simulées et densité théorique de la distribution Beta-Lomax (BTL) obtenues par la méthode de l'inverse	48
4.4	Graphiques statistiques pour l'analyse des données	53
4.5	Histogramme de la distribution ajustée et courbe de densité obtenus à l'aide des méthodes d'estimation CVM, ML pour le jeu de données . . .	55
4.6	Graphiques statistiques pour l'analyse des données	56
4.7	Histogramme de la distribution ajustée et courbe de densité obtenus à l'aide des méthodes d'estimation CVM, ML pour le jeu de données . . .	58

LISTE DES TABLEAUX

2.1	Fonctions de densité de probabilité de certaines familles T-X basées sur différentes fonctions $W(.)$	23
2.2	Familles de distributions généralisées dérivées de différentes distributions T	25
4.1	Comparaison des méthodes d'estimation des paramètres par CVM et MLE pour différentes tailles d'échantillons.	49
4.2	Comparaison des méthodes d'estimation des paramètres par CVM et MLE pour différentes tailles d'échantillons.	50
4.3	Données de survie en mois	53
4.4	Statistiques descriptives de l'ensemble de données	53
4.5	Estimations des paramètres et qualité de l'ajustement pour différents modèles ajustés aux données	54
4.6	Valeurs statistiques et leurs valeurs p- associées pour l'ensemble de données	54
4.7	Données sur les temps de défaillance (en heures)	55
4.8	Statistiques descriptives de l'ensemble de données	56
4.9	Estimations des paramètres et qualité de l'ajustement pour différents modèles ajustés aux données	57
4.10	Valeurs statistiques et leurs valeurs p- associées pour l'ensemble de données	57

ملخص :

يهدف هذا المشروع إلى دراسة توزيعات احتمالية جديدة تتميز بمرونة محسنة، مصممة خصيصًا لتعزيز دقة وفعالية تحليل بيانات البقاء على قيد الحياة والموثوقية. انطلاقًا من إطار عام لتحويل التوزيعات الأساسية، يركز البحث على اشتقاق وتوسيع توزيع لوماكس من خلال تطبيق تحويل بيتا عليه، مما يثري خصائصه الرياضية ويمنحه قدرات نمذجة متقدمة. هذا التوزيع الجديد، الذي تُشكل خصائصه الرياضية جوهر المشروع، يخضع لتقييم دقيق لأدائه مقابل توزيعات أخرى في معالجة بيانات حقيقية، حيث تُظهر النتائج فعاليته المتفوقة في توفير ملائمة دقيقة للبيانات التجريبية المعقدة.

الكلمات المفتاحية : توزيع لوماكس ، تحويل بيتا ، تحليل البقاء والموثوقية، النماذج ، بيانات حقيقية ، الملاءمة الإحصائية

Résumé :

Ce projet vise à étudier de nouvelles distributions de probabilité dotées d'une flexibilité améliorée, spécifiquement conçues pour renforcer la précision et l'efficacité de l'analyse des données de survie et de fiabilité. Partant d'un cadre général de transformation des distributions de base, la recherche se concentre sur la dérivation et l'extension de la distribution de Lomax via l'application d'une transformation Bêta, enrichissant ainsi ses propriétés mathématiques et lui conférant des capacités de modélisation avancées. Cette distribution transformée, dont les propriétés mathématiques constituent le cœur du projet, est soumise à une évaluation rigoureuse de ses performances par rapport à d'autres distributions lors du traitement de données réelles, où les résultats démontrent son efficacité supérieure à fournir un ajustement précis aux données empiriques complexes.

Mots-clés : Distribution de Lomax, Transformation Bêta, Analyse de survie et de fiabilité, Modélisation, Données réelles – Ajustement statistique

Abstract :

This project aims to study new, more flexible probability distributions specifically designed to enhance the accuracy and effectiveness of survival and reliability data analysis. Starting from a general framework for transforming basic distributions, the research focuses on deriving and extending the Lomax distribution by applying a Beta transformation to it. This enriches its mathematical properties and grants it advanced modeling capabilities. This transformed distribution, whose mathematical properties form the core of the project, undergoes a rigorous evaluation of its performance against other distributions when applied to real-world data. The results consistently demonstrate its superior effectiveness in providing a precise fit for complex empirical data.

Keywords : Lomax distribution, Beta transformation, Survival and reliability analysis, Modeling – Real data, Statistical fitting

ABRÉVIATIONS ET NOTATIONS

Abréviations

PDF	Fonction de densité de probabilité
CDF	Fonction de répartition cumulative
BTL	Distribution Beta Transformée de Lomax (ou Beta-Lomax Transformé)
GL	Distribution Gamma-Lomax
EL	Distribution Exponentielle-Lomax
MOEL	Distribution Marshall-Olkin Exponentielle-Lomax
BL	Distribution Beta-Lomax
L	Distribution Lomax
GLD	Distributions lambda généralisées
KW-G	La densité des distributions généralisées de Kumaraswamy
T-X	Méthode de transformée
MLE	Méthode du Maximum de Vraisemblance (Maximum Likelihood)
CVM	Méthode de Cramér-von Mises
EMV	Estimateur du Maximum de Vraisemblance
EQM	L'erreur quadratique moyenne
K-S	Statistique de Kolmogorov-Smirnov
C-V	Statistique de Cramér-von Mises
A-D	Statistique d'Anderson-Darling

Notations

f	Fonction de densité
F	Fonction répartition
F^{-1}	Inverse de la fonction de répartition
$v.a$	Variable aléatoire
$(X_1, \dots, X_n)$	Échantillon de taille n de X
$(X_{(1)}, \dots, X_{(n)})$	Échantillon ordonné
$\mathbb{R}$	ensemble de nombre réels
$X_{(i)}$	La $i^{\text{ème}}$ statistique d'ordre

INTRODUCTION GÉNÉRALE

L'étude des données de survie et de fiabilité est un sujet d'importance croissante dans de nombreux domaines appliqués, notamment en ingénierie, médecine et statistiques industrielles. Le défi principal réside dans le choix de modèles probabilistes appropriés capables de représenter avec flexibilité et précision le comportement des données. Dans ce contexte, ce projet vise à développer de nouvelles distributions probabilistes plus flexibles, utilisées pour l'analyse des données de survie, avec des applications pratiques sur des données réelles.

L'étude approfondie des méthodes existantes permet de créer de nouvelles distributions plus flexibles par rapport à les distributions anciennes. Le mémoire est composé de quatre chapitres :

Le premier chapitre présente les concepts fondamentaux liés aux distributions probabilistes. Il introduit les notions de variables aléatoires discrètes et continues, et passe en revue les principales fonctions qui leur sont associées (fonction de répartition, densité, fonctions génératrices), ainsi qu'un rappel des distributions classiques les plus utilisées en statistique.

Dans le deuxième chapitre, nous abordons différentes méthodes de transformation des distributions, notamment la méthode T-X et la transformation par la distribution bêta. Ces approches permettent de construire de nouvelles distributions à partir de distributions de base, offrant ainsi une plus grande flexibilité pour la modélisation des données.

Le troisième chapitre représente le cœur de ce travail. Nous y développons et analysons la distribution Lomax transformée par la méthode bêta, connue sous le nom de distribution BTL. Nous étudions ses propriétés mathématiques et proposons des méthodes d'estimation efficaces pour ses paramètres, permettant ainsi des modèles statistiques plus précis et flexibles.

Enfin, le quatrième chapitre est consacré pour les simulations, estimation et application de la nouvelle distribution. L'application sur des jeux de données réelles issues des domaines de la survie et de la fiabilité. La performance de cette distribution est évaluée avec précision et comparée avec d'autres distributions existantes. Les résultats expérimentaux montrent systématiquement que la nouvelle distribution, en particulier la Lomax transformée par bêta, offre un ajustement supérieur aux données, confirmant ainsi son efficacité et sa valeur en tant qu'outil statistique avancé pour les praticiens et les chercheurs dans ce domaine.

CHAPITRE 1

NOTIONS DE BASE SUR LES DISTRIBUTIONS DE PROBABILITÉS

1.1 Définition d'une variable aléatoire

Définition 1.1 Une variable aléatoire (v.a.) X est une fonction mesurable qui associe à chaque issue ω d'un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ un nombre réel $X(\omega)$. Elle permet ainsi de traduire les résultats d'une expérience aléatoire en valeurs numériques, facilitant leur analyse statistique.

1.2 Différents types de variables aléatoires

Une variable aléatoire (v.a.) est une application $X : \Omega \rightarrow \mathbb{R}$. Si l'ensemble des valeurs prises par $X(\omega)$ est au plus dénombrable, on dit que X est une v.a. discrète ; sinon, on dit qu'elle est continue.

1.2.1 Variable aléatoire discrète

Définition 1.2 Si une variable aléatoire X prend un nombre de valeurs fini ou dénombrable (son ensemble de définition est inclus dans $\mathbb{N}$), on parle de **variable discrète**. On s'intéresse alors à définir l'ensemble des valeurs possibles ainsi que leurs probabilités associées.

1.2.1.1 Fonction de masse de probabilité

Soient P une probabilité sur Ω , et X une variable aléatoire sur Ω . On définit la loi de probabilité P_X de la variable aléatoire X comme suit :

$$\mathbb{P}_X(x_i) = P(\{\omega \in \Omega \mid X(\omega) = x_i\}) = P(X = x_i)$$

telle que $P_X(x_i) \geq 0$, $\sum_{i=1}^n P_X(x_i) = 1$

1.2.1.2 Fonction de répartition

La fonction de répartition d'une variable aléatoire X indique, pour chaque valeur réelle x , la probabilité que X prenne une valeur inférieure ou égale à x . C'est la somme des probabilités des valeurs de X jusqu'à x . On la note F , et on a :

$$F(x) = \sum_{x_i \leq x} P(X = x_i) = \mathbb{P}(X \leq x)$$

La fonction de répartition est toujours croissante, comprise entre 0 et 1, et constitue un outil clé des probabilités théoriques.

1.2.1.3 Fonction caractéristique

Pour une variable aléatoire discrète X avec $P(X = x_k) = p_k$, la fonction caractéristique est :

$$\varphi_X(t) = \mathbb{E}[e^{itX}] = \sum_k e^{itx_k} p_k$$

1.2.1.4 Fonction génératrice

La fonction génératrice $G_X(t)$ de X d'une variable aléatoire discrète X utilisée pour faciliter le calcul des moments (l'espérance et la variance) est définie par :

$$G_X(t) = \mathbb{E}(t^X) = \sum_{i=1}^n t^{x_i} P_X(x_i)$$

Où :

t est un paramètre réel,

$E(t^X)$ représente l'espérance de t^X , c'est-à-dire la somme des puissances de t pondérées par les probabilités associées aux valeurs de X .

1.2.1.5 Fonction génératrice des moments

La fonction génératrice des moments $M_X(t)$ d'une variable aléatoire discrète X permet de calculer les moments (espérance, variance,...) et est définie par :

$$M_X(t) = \mathbb{E}(e^{tX}) = \sum_{i=1}^n e^{tx_i} P_X(x_i)$$

1.2.1.6 Valeurs caractéristiques

- **L'espérance mathématique** d'une variable aléatoire discrète X à valeurs dans $\{x_1, x_2, \dots, x_n\}$, et P_X la loi de X . est définie par :

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

$E(X)$ est aussi appelée la *valeur moyenne* de X .

- **La variance** d'une variable aléatoire discrète X à valeurs dans $\{x_1, x_2, \dots, x_n\}$ est définie par :

$$\mathbb{V}ar(X) = E\left((X - E(X))^2\right) = E(X^2) - (E(X))^2$$

avec

$$\mathbb{E}(X^2) = \sum_{i=1}^n x_i^2 P_X(x_i)$$

- **L'écart type** d'une variable aléatoire X discrète la quantité :

$$\sigma_X = \sqrt{\mathbb{V}ar(X)}$$

1.2.2 Variable aléatoire continue

Définition 1.3 Une variable aléatoire est dite continue si elle peut prendre toutes les valeurs dans un intervalle donné (borné ou non borné). En règle générale, toutes les variables qui résultent d'une mesure sont de type continue.

1.2.2.1 Fonction de densité

La fonction de densité f_X d'une variable aléatoire continue X est une fonction telle que :

1. $f_X(x) \geq 0$ pour tout $x \in \mathbb{R}$
2. $\int_{-\infty}^{+\infty} f_X(x) dx = 1$
3. Pour tout intervalle $[a, b]$, on a : $P(a \leq X \leq b) = \int_a^b f_X(x) dx$.

1.2.2.2 Fonction de répartition

La fonction de répartition $F_X(x)$ d'une variable aléatoire continue X est définie par :

$$F_X(x) = P(X \leq x)$$

1. F est continue et croissante sur $\mathbb{R}$.
2. $\lim_{x \rightarrow -\infty} F(x) = 0$ et $\lim_{x \rightarrow +\infty} F(x) = 1$.
3. $P(a \leq X \leq b) = F(b) - F(a)$.

Remarque : La fonction de répartition $F_X(x)$ est reliée à la densité par :

$$F_X(x) = \int_{-\infty}^x f_X(t) dt$$

et

$$f_X(x) = F'_X(x) \quad \text{si } F_X \text{ est dérivable en } x$$

1.2.2.3 Fonction caractéristique

Soit une variable aléatoire continue X de densité $f_X(x)$, la fonction caractéristique est :

$$\varphi_X(t) = \mathbb{E}[e^{itX}] = \int_{-\infty}^{+\infty} e^{itx} f_X(x) dx$$

1.2.2.4 Fonction génératrice des moments

Soit X une variable aléatoire continue de densité $f_X(x)$. La fonction génératrice des moments est définie par :

$$M_X(t) = \mathbb{E}[e^{tX}] = \int_{-\infty}^{+\infty} e^{tx} f_X(x) dx$$

Elle existe si $\mathbb{E}[e^{tX}] < +\infty$ pour t proche de 0

1. $M_X(0) = 1$
2. Si X et Y sont indépendantes : $M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$

Remarque : La fonction génératrice $\mathbb{E}[t^X]$ est remplacée par $\mathbb{E}[e^{tX}]$ pour les variables continues en raison de problèmes de convergence.

1.2.2.5 Valeurs caractéristiques

- **L'espérance mathématique** d'une variable aléatoire continue X de densité $f_X(x)$ est définie par :

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} x \cdot f_X(x) dx$$

1. si a, b sont des constantes , alors $E(aX + b) = aE(X) + b$
2. Si X et Y sont indépendantes, alors $E[XY] = E[X] \cdot E[Y]$.
3. L'espérance d'une constante est égale à la constante : $E[c] = c$

- **La variance** mesure la dispersion autour de l'espérance :

$$\text{Var}(X) = E[(X - E(X))^2] = \int_{-\infty}^{+\infty} (x - E(X))^2 f_X(x) dx$$

1. si a, b sont des constantes alors , $\text{Var}(aX + b) = a^2 \text{Var}(X)$
2. Si X et Y sont indépendantes, alors $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$
3. Variance d'une constante est un nulle : $\text{Var}(c) = 0$

- **L'écart type** est donné par :

$$\sigma_X = \sqrt{\text{Var}(X)}$$

1.3 Lois de probabilité usuelles

1.3.1 Cas discret

1.3.1.1 Loi uniforme discrète

On appelle loi uniforme sur $\{1, 2, \dots, n\}$ la loi dont le support est $\{1, 2, \dots, n\}$, notée $\mathcal{U}(\{1, 2, \dots, n\})$ et où tous les événements élémentaires ont la même probabilité, telle qu'une variable aléatoire suivant cette loi vérifie :

$$\mathbb{P}[X = k] = \frac{1}{n} \quad \text{pour } k \in \{1, 2, \dots, n\}$$

Si $X \sim \mathcal{U}(\{1, 2, \dots, n\})$, alors :

$$\mathbb{E}[X] = \frac{n+1}{2} \quad , \quad \text{Var}(X) = \frac{n^2-1}{12}$$

1.3.1.2 Loi Bernoulli

On dit qu'une variable aléatoire X suit la loi de Bernoulli de paramètre $p \in]0, 1[$, notée $X \sim \mathcal{B}(p)$. Si $X(\Omega) = \{0, 1\}$, c'est-à-dire si l'ensemble des réalisations associé à cette expérience aléatoire est $\{0, 1\}$, où :

- 0 correspond à un échec
- 1 correspond à un succès

avec :

$$\mathbb{P}(X = 0) = q, \quad \mathbb{P}(X = 1) = p, \quad \text{et} \quad q = 1 - p$$

Si $X \sim \mathcal{B}(p)$, alors :

$$\mathbb{E}(X) = p \quad \text{et} \quad \mathbb{V}\text{ar}(X) = pq$$

1.3.1.3 Loi Binomiale

Le nombre de succès dans n répétitions indépendantes d'une épreuve de Bernoulli est la somme des variables de Bernoulli correspondantes.

On dit qu'une variable aléatoire X suit une loi binomiale de paramètres $n \in \mathbb{N}$, notée $X \sim \mathcal{Bin}(n, p)$ et $p \in [0, 1]$ si X s'écrit sous la forme :

$$X = \sum_{k=1}^n X_k$$

où $X_1, \dots, X_n$ sont indépendantes et identiquement distribuées (i.i.d.) suivant une loi de Bernoulli de paramètre p .

Si $X \sim \mathcal{Bin}(n, p)$, alors :

$$\mathbb{E}[X] = np, \quad \mathbb{V}\text{ar}(X) = np(1 - p)$$

1.3.1.4 Loi de Poisson

On dit qu'une variable aléatoire X suit la loi de Poisson de paramètre $\lambda > 0$, notée $X \sim \mathcal{P}(\lambda)$. Dans ce cas, la probabilité que X prenne la valeur $k \in \mathbb{N}$ est donnée par :

$$\mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad \text{avec } \lambda > 0$$

Si $X \sim \mathcal{P}(\lambda)$, alors :

$$\mathbb{E}[X] = \lambda, \quad \mathbb{V}\text{ar}(X) = \lambda$$

1.3.1.5 Loi géométrique

Loi géométrique donne la probabilité que le premier succès survienne à l'essai k dans une suite d'épreuves de Bernoulli indépendantes. La variable X suit une loi géométrique de paramètre $p \in]0, 1[$ avec support $\mathbb{N}^*$, notée $X \sim G(p)$. Pour tout $k \in \mathbb{N}^*$, la fonction de masse est donnée par :

$$\mathbb{P}[X = k] = (1 - p)^{k-1} p, \quad \text{avec } p \in]0, 1[$$

Si $X \sim G(p)$, alors :

$$\mathbb{E}[X] = \frac{1}{p}, \quad \text{Var}(X) = \frac{1-p}{p^2}$$

1.3.2 Cas continue

1.3.2.1 Loi uniforme

On dit qu'une variable aléatoire X suit la loi uniforme sur l'intervalle $[a, b] \in \mathbb{R}$, notée $X \sim \mathcal{U}([a, b])$, si sa densité de probabilité est donnée par :

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{si } a \leq x \leq b, \\ 0 & \text{sinon.} \end{cases}$$

La fonction de répartition $F_X(x)$ de X est donnée par :

$$F_X(x) = \begin{cases} 0 & \text{si } x \leq a, \\ \frac{x-a}{b-a} & \text{si } a < x < b \\ 1 & \text{si } x \geq b \end{cases}$$

De plus

$$\mathbb{E}[X] = \frac{a+b}{2}, \quad \text{Var}(X) = \frac{(b-a)^2}{12}$$

1.3.2.2 Loi normale

- *Cas général*

On dit qu'une variable aléatoire X suit la loi normale de paramètres $m \in \mathbb{R}$ et σ^2 , notée $X \sim \mathcal{N}(m, \sigma^2)$, si elle est continue et admet pour densité :

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right).$$

La fonction de répartition $F_X(x)$ est :

$$F_X(x) = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(t-m)^2}{2\sigma^2}\right) dt$$

De plus

$$\mathbb{E}[X] = m \quad \text{et} \quad \mathbb{V}ar(X) = \sigma^2$$

• **Loi normale centrée réduite**

On dit qu'une variable aléatoire X suit la loi normale centrée réduite , notée $X \sim \mathcal{N}(\mathbf{0},1)$, si elle est continue et admet pour densité :

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$$

La fonction de répartition $F_X(x)$ est :

$$F_X(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt$$

et on a

$$\mathbb{E}(X) = 0 \quad \text{et} \quad \mathbb{V}(X) = 1$$

1.3.2.3 Loi exponentielle

Une variable aléatoire X suit une loi exponentielle de paramètre ($\lambda > 0$) , notée $X \sim \mathcal{E}(\lambda)$, si sa densité est donnée par

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases}$$

La fonction de repartition F_X de X est donnée par :

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x} & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases}$$

De plus

$$\mathbb{E}(X) = \frac{1}{\lambda} \quad , \quad \mathbb{V}ar(X) = \frac{1}{\lambda^2}$$

1.3.2.4 Loi de gamma

On dit qu'une variable aléatoire X suit la loi gamma de paramètres $\alpha > 0$ et $\lambda > 0$, notée $X \sim \Gamma(\alpha, \lambda)$, si sa densité est donnée par :

$$f_X(x) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases}$$

où $\Gamma(\alpha) = \int_0^{+\infty} t^{\alpha-1} e^{-t} dt$ (désigne la fonction gamma d'Euler)

Sa fonction de répartition est donnée par :

$$F_X(x) = \frac{1}{\Gamma(\alpha)} \int_0^{\beta x} t^{\alpha-1} e^{-t} dt = \frac{\gamma(\alpha, \lambda x)}{\Gamma(\alpha)}, \quad x \geq 0$$

où $\gamma(\alpha, \beta x) = \int_0^{\beta x} t^{\alpha-1} e^{-t} dt$. (la fonction Gamma incomplète)

De plus

$$\mathbb{E}[X] = \frac{\alpha}{\beta}, \quad \text{Var}(X) = \frac{\alpha}{\beta^2}$$

1.3.2.5 Loi bêta

On dit qu'une variable aléatoire X suit la loi Bêta de paramètres $\alpha, \beta > 0$, notée $X \sim \beta(\alpha, \beta)$, si sa densité est donnée par :

$$f_X(x) = \begin{cases} \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)} & \text{si } 0 < x < 1 \\ 0 & \text{sinon} \end{cases} \quad \text{avec } x \in \mathbb{R}$$

Sa fonction de répartition est donnée par :

$$F_X(x) = \frac{1}{B(\alpha, \beta)} \int_0^x t^{\alpha-1} (1-t)^{\beta-1} dt \quad \text{pour } 0 < x < 1$$

où $B(\alpha, \beta) = \int_0^1 t^{\alpha-1} (1-t)^{\beta-1} dt = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$ (désigne l'intégrale bêta)

De plus

$$\mathbb{E}[X] = \frac{\alpha}{\alpha+\beta}, \quad \text{Var}(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

1.3.2.6 Loi Weibull

On dit qu'une variable aléatoire X suit la loi de Weibull de paramètres $c > 0$ et $\beta > 0$, notée $X \sim \mathcal{W}(c, \beta)$, si sa densité est donnée par :

$$f_X(x) = \beta c x^{\beta-1} e^{-cx^\beta} \cdot \mathbf{1}_{[0,+\infty[}(x), \quad x \in \mathbb{R}$$

Sa fonction de répartition est donnée par :

$$F_X(x) = (1 - e^{-cx^\beta}) \cdot \mathbf{1}_{[0,+\infty[}(x)$$

Son espérance et sa variance sont données par :

$$\mathbb{E}[X] = \frac{\Gamma\left(1 + \frac{1}{(\beta)}\right)}{c^{1/(\beta)}} \quad , \quad \mathbb{V}ar(X) = \frac{\Gamma\left(1 + \frac{2}{(\beta)}\right) - \left(\Gamma\left(1 + \frac{1}{(\beta)}\right)\right)^2}{c^{2/(\beta)}}$$

Si $Y \sim \mathcal{E}(c)$, alors : $Y^{1/(\beta)} \sim \mathcal{W}(c, \beta)$

1.3.2.7 Loi de Pareto

On dit qu'une variable aléatoire X suit une loi de Pareto de paramètres $k > 0$ et $\theta > 0$, notée $X \sim \text{Par}(k, \theta)$, si sa densité est donnée par :

$$f_X(x) = \frac{\theta}{k} \left(\frac{k}{x}\right)^{\theta+1} \quad , \quad x \in \mathbb{R}$$

La fonction de répartition F_X de X est donnée par :

$$F_X(x) = \left(1 - \left(\frac{k}{x}\right)^\theta\right)$$

De plus

$$\mathbb{E}[X] = \frac{k\theta}{\theta-1} \quad \text{pour } \theta > 1$$

$$\mathbb{V}ar[X] = \frac{k^2\theta}{(\theta-1)^2(\theta-2)} \quad \text{pour } \theta > 2$$

1.3.2.8 Loi du Khi-deux

Une variable aléatoire X suit une loi du Khi-deux à n degrés de liberté, notée $X \sim \chi_n^2$ de la forme :

$$X = Z_1^2 + Z_2^2 + \cdots + Z_n^2 \quad \text{avec } Z_i \sim \mathcal{N}(0, 1)$$

La fonction densité est donnée par :

$$f_X(x) = \begin{cases} \frac{1}{2^{n/2}\Gamma(n/2)} x^{(n/2)-1} e^{-x/2} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$$

De plus

$$\mathbb{E}[X] = n, \quad \text{Var}(X) = 2n$$

Si $X_1 \sim \chi_{n_1}^2$ et $X_2 \sim \chi_{n_2}^2$ sont indépendantes, alors $X_1 + X_2 \sim \chi_{n_1+n_2}^2$.

1.3.2.9 Loi Student

Une variable aléatoire X suit une loi de Student à ν degrés de liberté, notée $T \sim \mathcal{T}(\nu)$ de la forme :

$$T = \frac{Z}{\sqrt{Y/\nu}} \quad Z \sim \mathcal{N}(0,1) \quad \text{et} \quad Y \sim \chi_\nu^2$$

avec $Z \sim \mathcal{N}(0,1)$ et $Y \sim \chi^2(\nu)$ deux variables aléatoires indépendantes.

La fonction densité est donnée par :

$$f_T(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \quad x \in \mathbb{R}$$

De plus

$$\mathbb{E}[T] = 0 \quad \text{pour } \nu \geq 2$$

$$\text{Var}[T] = \frac{\nu}{\nu-2} \quad \text{pour } \nu \geq 3$$

1.3.2.10 Loi de Fisher-Snedecor

Une variable aléatoire X suit une Fisher-snedecor à ν_1 degrés de liberté au numérateur et ν_2 degrés de liberté au dénominateur, notée $F_{(\nu_1, \nu_2)}$ de la forme :

$$F = \frac{\nu_2 X_1}{\nu_1 X_2}$$

avec $X_1 \sim \chi_{\nu_1}^2$ et $X_2 \sim \chi_{\nu_2}^2$ deux variables aléatoires indépendantes.

De plus

$$E[X] = \frac{\nu_2}{\nu_2 - 2} \quad \text{pour } \nu_2 > 2$$

$$\text{Var}[X] = \frac{2\nu_2^2(\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2 - 2)^2(\nu_2 - 4)} \quad \text{pour } \nu_2 > 4$$

1.4 Méthodes de Simulation

1.4.1 Méthode de l'inversion

Soit X une variable aléatoire continue de fonction de répartition F_X strictement croissante et continue. La méthode de l'inversion consiste à simuler X en générant une variable aléatoire $U \sim \mathcal{U}([0, 1])$ et en appliquant l'inverse généralisé de F_X , noté F_X^{-1} , tel que :

$$X = F_X^{-1}(U)$$

La fonction F_X^{-1} est définie comme :

$$F_X^{-1}(u) = \inf \{x \in \mathbb{R} : F_X(x) \geq u\}$$

1.4.2 Méthode d'acceptation-rejet

Soit X une variable aléatoire à simuler, de densité f_X connue. On choisit une densité auxiliaire g (facile à simuler) et une constante $M > 0$ telle que :

$$f(x) \leq Mg(x) \quad , \quad \forall x$$

L'algorithme est le suivant :

1. Générer $Y \sim g$ et $U \sim \mathcal{U}(0, 1)$ indépendamment.
2. Accepter Y comme réalisation de X si :

$$U \leq \frac{f(Y)}{Mg(Y)}$$

3. Sinon, rejeter Y et recommencer.

Ces deux méthodes sont fondamentales en simulation de variables aléatoires, la méthode d'inversion étant directe mais nécessitant la fonction inverse de la répartition, tandis que la méthode d'acceptation-rejet est plus générale et permet de simuler des lois complexes à partir d'une densité instrumentale simple.

CHAPITRE 2

MÉTHODES DE GÉNÉRATION DE NOUVELLES DISTRIBUTIONS DE PROBABILITÉS.

Introduction

Les distributions statistiques sont couramment utilisées pour modéliser des phénomènes réels. De nombreuses distributions généralisées ont été développées, souvent avec plusieurs paramètres. Johnson et les autres estiment qu'une distribution à quatre paramètres est généralement suffisante pour les applications pratiques, bien qu'ils doutent de l'utilité d'ajouter un cinquième ou un sixième paramètre.

2.1 Familles de distribution continues

2.1.1 Distribution de Pearson

La distribution de Pearson a été développée par Pearson [1], est une distribution pour lequel chaque fonction de densité de probabilité (p.d.f.) $f(x)$ satisfait une équation différentielle de la forme :

$$\frac{1}{f(x)} \frac{df(x)}{dx} = \frac{a+x}{b_0 + b_1x + b_2x^2} \quad (1.1)$$

où a , b_0 , b_1 et b_2 sont des paramètres

La forme de la solution de l'équation (1.1) dépend du type de racines de $b_0 + b_1x + b_2x^2 = 0$.

1. La forme de $f(x)$ dépend des paramètres a , b_0 , b_1 , b_2 , et si $-a$ n'est pas une racine de $b_0 + b_1x + b_2x^2 = 0$, alors $f(x)$ est finie et $f'(-a) = 0$.

2. $f'(x) = 0$ lorsque $f(x) = 0$. Cependant, si $x \neq -a$ et $f(x) \neq 0$, alors $f'(x) \neq 0$. Pour respecter les conditions $f(x) \geq 0$ et $\int_{-\infty}^{\infty} f(x) dx = 1$, il faut que :

$$\lim_{x \rightarrow \pm\infty} f(x) = 0 \quad \text{et} \quad \lim_{x \rightarrow \pm\infty} f'(x) = 0,$$

ce qui n'est pas toujours le cas pour toutes les solutions. Dans certains cas, la condition $f(x) \geq 0$ n'est pas vérifiée. Il devient alors nécessaire de limiter l'intervalle où $f(x) > 0$ et de poser $f(x) = 0$ en dehors de cet intervalle.

3. Si $b_1 = b_2 = 0$, alors l'équation (1.1) devient plus simple :

$$\frac{d \log f(x)}{dx} = \frac{x + a}{b_0}$$

d'où

$$f(x) = K \exp \left[\frac{(x + a)^2}{2b_0} \right]$$

où K est une constante

4. Lorsque $b_1 = b_2 = 0$, la solution de l'équation (1.1) devient une distribution normale d'espérance $-a$ et d'écart-type $\sqrt{b_0}$ (avec $b_0 > 0$), où $K = \sqrt{2\pi b_0}$.
5. La forme de la distribution de Pearson varie selon les paramètres a, b_0, b_1 et b_2 , conduisant à une classification en plusieurs types basée sur les solutions de son équation différentielle caractéristique.

2.1.2 Distribution de Burr

Burr [2] a proposé un système de distributions continues caractérisé par l'équation différentielle :

$$dF = F(1 - F)g(x)dx \tag{1.2}$$

où $0 \leq F \leq 1$ et $g(x)$ est une fonction appropriée, non négative sur le domaine de x . La solution à cette équation différentielle, pour une fonction $g(x)$ donnée, est obtenue comme suit :

$$F(x) = \frac{1}{1 + e^{-G(x)}}$$

où

$$G(x) = \int_{-\infty}^x g(u) du$$

Burr a proposé 12 solutions à cette équation différentielle (correspondant à différents choix de $g(x)$), que nous listons ci-dessous :

1. **Type I** : $F(x) = x$ pour $0 < x < 1$
2. **Type II** : $F(x) = (1 + e^{-x})^{-k}$
3. **Type III** : $F(x) = (1 + x^c)^{-k}$
4. **Type IV** : $F(x) = \left[1 + \left(\frac{c-x}{x}\right)^{1/c}\right]^{-k}$
5. **Type V** : $F(x) = (1 + ce^{-\tan(x)})^{-k}$
6. **Type VI** : $F(x) = (1 + e^{-k \sinh(x)})^{-k}$
7. **Type VII** : $F(x) = 2^{-k}(1 + \tanh(x))^k$
8. **Type VIII** : $F(x) = 1 - e^{-x}$
9. **Type IX** : $F(x) = 1 - \frac{2}{2+c[(1+e^x)^k-1]}$
10. **Type X** : $F(x) = (1 - e^{-x^2})^k$
11. **Type XI** : $F(x) = [K - 2\sin(2\pi x)]^k$ pour $0 < x < 1$
12. **Type XII** : $F(x) = 1 - (1 + x^c)^{-k}$

où c et k sont des paramètres positifs et $-\infty < x < \infty$ sauf indication contraire. L'introduction de $x = (z-\mu)/\sigma$ permettrait d'ajouter des paramètres de position et d'échelle, si nécessaire. Nous avons choisi de ne pas les inclure ici.

Ces douze distributions définies par ces fonctions de distribution peuvent être appelées les distributions de Burr de Type I à XII .

2.1.3 Distribution de Johnson

Johnson [3] a proposé un système pour générer des distributions en utilisant une transformation de normalisation :

$$Z = \gamma + \delta f\left(\frac{x - \xi}{\lambda}\right) \quad (1.3)$$

où $f(\cdot)$ est la fonction de transformation, Z est une variable aléatoire normale standardisée, γ et δ sont des paramètres de forme, λ est un paramètre d'échelle et ξ est un paramètre de position.

Johnson a introduit trois transformations ($\delta, \lambda > 0$) qui unifient plusieurs lois classiques. La loi log-normale vérifie : $\log(X - \xi) \sim \mathcal{N}$. D'autres lois, comme la loi de Weibull, apparaissent via des transformations telles que $(X - \xi)^c$ ou $e^{-(X-\xi)}$.

2.1.4 Distributions lambda généralisées

Cette famille de distributions est définie en termes de fonction quantile :

$$Q(y) = Q(y; \lambda_1, \lambda_2, \lambda_3, \lambda_4) = \lambda_1 + \frac{y^{\lambda_3} - (1-y)^{\lambda_4}}{\lambda_2} \quad \text{où } 0 \leq y \leq 1. \quad (1.4)$$

Les paramètres λ_1 et λ_2 sont respectivement les paramètres de position et d'échelle, tandis que λ_3 et λ_4 déterminent l'asymétrie et le kurtosis.[4]

Soit y une variable prenant des valeurs dans l'intervalle $[0, 1]$. La fonction $x = Q(y)$ transforme y en une valeur sur l'axe des x . La densité de probabilité est définie par :

$$f(Q(y)) = \frac{1}{Q'(y)} \quad \text{où } Q'(y) = \frac{dQ(y)}{dy}$$

La condition pour que $Q(y)$ soit une fonction de distribution inverse légitime est que $Q'(y) > 0$ pour tout $y \in [0, 1]$. La densité de probabilité correspondante est également donnée par :

$$f(x) = \frac{\lambda_2}{\lambda_3 y^{\lambda_3-1} + \lambda_4 (1-y)^{\lambda_4-1}} \quad \text{avec } x = Q(y) \quad (1.5)$$

L'existence d'une densité de probabilité valide nécessite que $\lambda_3 y^{\lambda_3-1} + \lambda_4 (1-y)^{\lambda_4-1}$ ait le même signe pour tout $y \in [0, 1]$ et que λ_2 prenne également ce même signe.

La GLD (Generalized Lambda Distribution) a été proposée, englobant à la fois la GLD et la distribution bêta généralisée, définie par :

$$f(x) = \begin{cases} \frac{(x-\beta_1)^{\beta_3} (\beta_1+\beta_2-x)^{\beta_4}}{B(\beta_3+1, \beta_4+1) \beta_2^{\beta_3+\beta_4+1}} & \text{pour } \beta_1 \leq x \leq \beta_1 + \beta_2 \\ 0 & \text{sinon} \end{cases} \quad (1.6)$$

où $B(.,.)$ est la fonction bêta complète.

2.1.5 Distributions normales asymétriques

Azzalini [5] a introduit la famille des distributions normales asymétriques. Supposons que X et Y sont des variables aléatoires indépendantes, chacune ayant une densité symétrique autour de zéro. Pour tout λ , nous avons :

$$0.5 = P(X - \lambda Y < 0) = \int_{-\infty}^{\infty} f_Y(y) F_X(\lambda y) dy \quad (1.7)$$

Ainsi, la fonction de densité de probabilité est donnée par :

$$2f_Y(y)F_X(\lambda y)$$

Si X et Y suivent une loi normale standard $N(0, 1)$, alors la famille des distributions normales asymétriques admet la densité suivante :

$$2\phi(x)\Phi(\lambda x) \quad (1.8)$$

où $\phi(x)$ et $\Phi(x)$ sont respectivement la densité et la fonction de répartition de la loi normale $N(0, 1)$.

2.1.6 Distributions générées par la loi bêta

Eugene et al [6] ont utilisé la distribution bêta comme générateur pour développer la famille des distributions bêta-générées. La fonction de répartition cumulative (c.d.f) d'une variable aléatoire bêta-générée X est définie par :

$$G(x) = \int_0^{F(x)} b(t)dt \quad (1.9)$$

où $b(t)$ est la densité de la variable bêta et $F(x)$ est la fonction de répartition de toute variable aléatoire. La densité de probabilité associée est :

$$g(x) = \frac{1}{B(\alpha, \beta)} f(x) F^{\alpha-1}(x) (1 - F(x))^{\beta-1} \quad (1.10)$$

Cette famille de distributions généralise les distributions des statistiques d'ordre pour la variable aléatoire X avec fonction de répartition $F(x)$

Depuis l'article d'Eugene et al, de nombreuses distributions bêta-générées ont été étudiées, notamment la distribution bêta-Gumbel, la distribution bêta-exponentielle, la distribution bêta-Weibull et la distribution bêta-gamma

Plus récemment, Jones [7] et Cordeiro et de Castro [8] ont étendu cette famille en remplaçant la distribution bêta dans (1.9) par la distribution de Kumaraswamy.

$$b(x) = \alpha\beta x^{\alpha-1} (1 - x^\alpha)^{\beta-1}, \quad x \in (0, 1)$$

La densité des distributions généralisées de Kumaraswamy (KW-G) est alors donnée par :

$$g(x) = \alpha\beta f(x) F^{\alpha-1}(x) (1 - F^\alpha(x))^{\beta-1} \quad (1.11)$$

Plusieurs distributions généralisées à partir de (1.11) ont été étudiées dans la littérature, y compris la distribution Kumaraswamy-Weibull, la distribution gamma généralisée de Kumaraswamy et la distribution demi-normale généralisée de Kumaraswamy

2.1.7 Méthodes de génération de distributions asymétriques

Ferreira et Steel [9] ont introduit une méthode pour générer des distributions asymétriques à travers des transformations inverses de probabilité intégrale, une distribution G est une version asymétrique d'une distribution symétrique F , générée par un mécanisme de déformation P , si sa densité est de la forme :

$$g(y|F, P) = f(y)p(F(y)) \quad (1.12)$$

La densité ci-dessus est une fonction pondérée de $f(y)$ avec le poids $p(F(y))$

- S est égal à F uniquement dans le cas où P est une distribution uniforme sur $(0, 1)$.
- Si P est fixé et que F varie, obtenir une distribution S symétrique pour tout F équivaut à la symétrie de P autour de $\frac{1}{2}$.

Comme indiqué dans l'équation (1.12), l'énergie est proportionnelle à la masse.

Remarque

- La famille normale asymétrique en (1.8) est un cas particulier de cette famille.
- En relâchant l'hypothèse que $F(\cdot)$ est symétrique, la famille générée par la bêta (1.10) est un cas particulier de (1.12)

2.1.8 Méthode de transformation bêta

La **transformation bêta (BT)**, proposée par FASNA (2022)[10], définit une nouvelle famille de distributions continues à partir d'une fonction de répartition de base $F(x)$. Pour une variable aléatoire continue X , la fonction de répartition transformée $G(x)$ est donnée par :

$$G(x) = \begin{cases} \frac{\beta}{\beta - 1} [1 - \beta^{-F(x)}] & \text{si } \beta > 0, \beta \neq 1 \\ F(x) & \text{si } \beta = 1 \end{cases}$$

où β est un paramètre de transformation. Cette formulation permet d'ajuster la forme de la distribution.

La densité de probabilité associée, obtenue par dérivation de $G(x)$, est :

$$g(x) = \begin{cases} \frac{\beta \log \beta}{\beta - 1} f(x) \beta^{-F(x)} & \text{si } \beta > 0, \beta \neq 1 \\ f(x) & \text{si } \beta = 1 \end{cases}$$

Cette transformation conserve les propriétés fondamentales des distributions de probabilité tout en offrant une plus grande flexibilité dans la modélisation.

2.2 Méthode T-X familles de distributions

La loi bêta et la loi Kummer-Weber-Gamma sont des familles de distributions définies sur l'intervalle $[0, 1]$ ce qui les rend adaptées pour modéliser des variables aléatoires représentant des proportions ou des probabilités. L'idée est de généraliser l'utilisation de générateurs de distributions à des intervalles plus larges que $[0, 1]$

Soit $r(t)$ la densité de probabilité (p.d.f) d'une variable aléatoire T définie sur $[a, b]$, où $-\infty \leq a < b \leq \infty$. Soit $W(F(x))$ une fonction de la fonction de répartition $F(x)$ de toute variable aléatoire X , telle que $W(F(x))$ satisfasse les conditions suivantes :

$$\begin{cases} W(F(x)) \in [a, b] \\ W(F(x)) \text{ est différentiable et monotone non décroissante} \\ W(F(x)) \rightarrow a \text{ lorsque } x \rightarrow -\infty \text{ et } W(F(x)) \rightarrow b \text{ lorsque } x \rightarrow \infty \end{cases} \quad (2.1)$$

Définition 2.1 Soit X une variable aléatoire avec densité de probabilité $f(x)$ et fonction de répartition $F(x)$. Soit T une variable aléatoire continue avec densité de probabilité $r(t)$ définie sur un intervalle $[a, b]$. La fonction de répartition $G(x)$ d'une nouvelle famille de distributions est définie par :

$$G(x) = \int_a^{W(F(x))} r(t) dt \quad (2.2)$$

où $W(F(x))$ est une fonction de $F(x)$ qui satisfait les conditions précédentes.(2.1)

On peut également exprimer $G(x)$ sous la forme $G(x) = R(W(F(x)))$ où $R(t)$ est la fonction de répartition de la variable aléatoire T . La densité de probabilité $g(x)$ associée à cette nouvelle famille de distributions est donnée par :

$$g(x) = f(x) \cdot W'(F(x)) \cdot r(W(F(x))) \quad (2.3)$$

Remarque

- La fonction de répartition dans (2.2) est une fonction composée de $(R \circ W \circ F)(x)$.
- La distribution $g(x)$ dans (2.2) est appelée « T-X » ou « transformée-transforme » car elle résulte de la transformation de la fonction de densité de probabilité $r(t)$ via la fonction $W(Fx)$
- Si X est une variable aléatoire discrète, alors $G(x)$ représente la fonction de répartition d'une famille de distributions discrètes.
- la distribution (1.12) est un cas particulier de (2.3) lorsque $W(F(x)) = F(x)$, avec $r(\cdot)$ jouant le rôle d'une fonction de pondération.

Différents choix de $W(Fx)$ génèrent de nouvelles familles de distributions, et sa définition dépend du support de la variable aléatoire T

Exemples de choix de $W(F(x))$ selon le support de T

1. **Support de T est $[0, 1]$:** On peut choisir

$$W(F(x)) = F(x) \quad \text{ou} \quad W(F(x)) = F^\alpha(x)$$

Cela correspond à la famille de distributions générée par la bêta.

2. **Support de T est $[a, \infty)$ avec $a \geq 0$:** On peut choisir

$$\left\{ \begin{array}{l} W(F(x)) = -\log(1 - F(x)) \\ W(F(x)) = F(x)/(1 - F(x)) \\ W(F(x)) = -\log(1 - F^\alpha(x)) \\ W(F(x)) = F^\alpha(x)/(1 - F^\alpha(x)) \end{array} \right. \quad \text{avec } \alpha > 0$$

3. **Support de T est $(-\infty, \infty)$:** On peut choisir

$$\left\{ \begin{array}{l} W(F(x)) = \log(-\log(1 - F(x))) \\ W(F(x)) = \log[F(x)/(1 - F(x))] \\ W(F(x)) = \log(-\log(1 - F^\alpha(x))) \\ W(F(x)) = \log[F^\alpha(x)/(1 - F^\alpha(x))] \end{array} \right.$$

En utilisant $W(F(x)) = -\log(1 - F(x))$ dans le deuxième exemple, la fonction de répartition $G(x)$ dans (2.2) est donnée par :

$$G(x) = \int_0^{-\log(1-F(x))} r(t) dt = R\{-\log(1 - F(x))\} \quad (2.4)$$

où $R(t)$ est la fonction de répartition de la variable aléatoire T .

La fonction de densité de probabilité associée à (2.4) est :

$$g(x) = \frac{f(x)}{1 - F(x)} \cdot r(-\log(1 - F(x))) = h(x) \cdot r(-\log(1 - F(x))) \quad (2.5)$$

où $h(x)$ est la fonction de risque pour la variable aléatoire X avec la fonction de répartition $F(x)$.

Les familles de distributions générées à partir des autres fonctions $W(\cdot)$ mentionnées dans les exemples 2 et 3 sont données dans le Tableau 2.1.

Dans la suite de cet article, nous nous concentrerons sur le cas où T a un support $[0, \infty)$ et $W(F(x)) = -\log(1 - F(x))$. Pour simplifier, nous utiliserons le nom de famille de distributions « T-X » pour la nouvelle famille de distributions dans (2.5).

Support of T	$W(F(x))$	$g(x)$
$[0, \infty)$	$\frac{F(x)}{1-F(x)}$	$\frac{f(x)}{(1-F(x))^2} r\left\{\frac{F(x)}{1-F(x)}\right\}$
$[0, \infty)$	$-\log(1 - F^\alpha(x))$	$\frac{\alpha f(x)F^{\alpha-1}(x)}{1-F^\alpha(x)} r\{-\log(1 - F^\alpha(x))\}$
$[0, \infty)$	$\frac{F^\alpha(x)}{1-F^\alpha(x)}$	$\frac{\alpha f(x)F^{\alpha-1}(x)}{(1-F^\alpha(x))^2} r\left\{\frac{F^\alpha(x)}{1-F^\alpha(x)}\right\}$
$(-\infty, \infty)$	$\log(-\log(1 - F(x)))$	$\frac{f(x)r\{\log(-\log(1-F(x)))\}}{(F(x)-1)\log(1-F(x))}$
$(-\infty, \infty)$	$\log\left(\frac{F(x)}{1-F(x)}\right)$	$\frac{f(x)}{F(x)(1-F(x))} r\left\{\log\left(\frac{F(x)}{1-F(x)}\right)\right\}$
$(-\infty, \infty)$	$\log(-\log(1 - F^\alpha(x)))$	$\frac{\alpha f(x)F^{\alpha-1}(x)}{(F^\alpha(x)-1)\log(1-F^\alpha(x))} r\{\log(-\log(1 - F^\alpha(x)))\}$
$(-\infty, \infty)$	$\log\left(\frac{F^\alpha(x)}{1-F^\alpha(x)}\right)$	$\frac{\alpha f(x)}{F(x)(1-F(x))} r\left\{\log\left(\frac{F^\alpha(x)}{1-F^\alpha(x)}\right)\right\}$

TABLE 2.1 – Fonctions de densité de probabilité de certaines familles T-X basées sur différentes fonctions $W(\cdot)$

2.2.1 Propriétés de la famille T-X

- (a) La densité peut être écrite comme $g(x) = h(x)r(H(x))$, où $H(x)$ est la fonction de hasard cumulée.
- (b) La relation entre les variables aléatoires X et T est donnée par : $X = F^{-1}(1 - e^{-T})$
 La relation $G(x) = R(-\log(1 - F(x)))$ permet d'exprimer X en fonction de T comme $X = F^{-1}(1 - e^{-T})$. Cela facilite la simulation de X en générant d'abord T selon $r(t)$, puis en appliquant cette transformation. De plus, l'espérance de X est donnée par $E(X) = E\{F^{-1}(1 - e^{-T})\}$.

La fonction quantile $Q(\lambda)$, pour $0 < \lambda < 1$, de la famille de distributions T-X peut être calculée en utilisant la formule :

$$Q(\lambda) = F^{-1}\{1 - e^{-R^{-1}(\lambda)}\} \quad (2.6)$$

L'entropie de Shannon [11] mesure l'incertitude d'une variable aléatoire X , définie par
 Pour une variable discrète X :

$$H(X) = - \sum_x p(x) \log p(x)$$

Pour une variable continue avec densité $g(x)$:

$$H(X) = - \int_{-\infty}^{\infty} g(x) \log g(x) dx = E\{-\log(g(X))\}$$

Le théorème 2.1 relie l'entropie de la distribution $g(x)$ à celle du générateur $r(t)$.

Théorème 2.1 *Si X suit la famille de distributions*

$$g(x) = \frac{f(x)}{1-F(x)} \cdot r(-\log(1-F(x)))$$

l'entropie de Shannon de X , notée η_X , est donnée par :

$$\eta_X = -E\{\log f(F^{-1}(1-e^{-T}))\} - \mu_T + \eta_T \quad (2.7)$$

où μ_T et η_T sont respectivement la moyenne et l'entropie de Shannon de T avec la f.d.p. $r(t)$.

Preuve. : L'entropie de Shannon de X est donnée par :

$$\eta_X = E(-\log[g(X)]) = -E(\log f(X)) + E(\log(1-F(X))) + E(-\log r(-\log(1-F(X))))$$

En utilisant la variable aléatoire $T = -\log(1-F(X))$ avec la f.d.p. $r(t)$, on obtient les résultats suivants :

$$E(\log f(X)) = E\{\log f(F^{-1}(1-e^{-T}))\}$$

$$E(\log(1-F(X))) = -E(T) = -\mu_T$$

$$E(-\log r(-\log(1-F(X)))) = \eta_T$$

Ainsi, l'entropie de Shannon de X est :

$$\eta_X = -E\{\log f(F^{-1}(1-e^{-T}))\} - \mu_T + \eta_T$$

■

Pour l'asymétrie $\alpha_3 = \mu_3/\sigma^3$ et la kurtosis $\alpha_4 = \mu_4/\sigma^4$, qui peuvent être difficiles à calculer pour certaines distributions (comme Cauchy, Lévy, et Pareto). La mesure d'asymétrie S définie par Galton [12] et la mesure de kurtosis K définie par Moors [13] sont basées sur des fonctions quantiles et sont définies comme suit :

$$S = \frac{Q(6/8) - 2Q(4/8) + Q(2/8)}{Q(6/8) - Q(2/8)} \quad (2.8)$$

$$K = \frac{Q(7/8) - Q(5/8) + Q(3/8) - Q(1/8)}{Q(6/8) - Q(2/8)} \quad (2.9)$$

L'asymétrie S mesure l'orientation de la queue (gauche ou droite), et la kurtosis K mesure la lourdeur de la queue.

Pour la famille $T-X$, ces mesures peuvent être calculées à partir de la fonction quantile (2.6) et des distributions de T et X .

2.2.2 Certaines familles de distributions T-X avec différentes distributions T

La famille de distributions T-X peut être classée en deux sous-familles :

- Une sous-famille ayant la même distribution X mais différentes distributions T .
- Une sous-famille ayant la même distribution T mais différentes distributions X .

Cette section étudie des sous-familles de distributions obtenues en faisant varier la loi de T tout en gardant X fixe. Le tableau 2.2 présente plusieurs de ces sous-familles pour différentes lois de T .

Nom	Densité $r(t)$	Densité de la famille $g(x)$
Exponentielle	$\theta e^{-\theta t}$	$\theta f(x)(1-F(x))^{\theta-1}$
Beta-exponentielle	$\frac{\lambda e^{-\lambda x}(1-e^{-\lambda x})^{\alpha-1}}{B(\alpha, \beta)}$	$\frac{\lambda f(x)}{B(\alpha, \beta)}(1-F(x))^{\lambda\beta-1}(1-(1-F(x))^{\lambda})^{\alpha-1}$
Expo-exponentielle	$\frac{\alpha \lambda (1-e^{-\lambda x})^{\alpha-1}}{e^{-\lambda x}}$	$\alpha f(x)\{1-(1-F(x))^{\lambda}\}^{\alpha-1}(1-F(x))^{\lambda-1}$
Gamma	$\frac{1}{\Gamma(\alpha)\beta^{\alpha}} t^{\alpha-1} e^{-t/\beta}$	$\frac{f(x)}{\Gamma(\alpha)\beta^{\alpha}} (-\log(1-F(x)))^{\alpha-1} (1-F(x))^{\frac{1}{\beta}-1}$
Demi-normale	$\frac{1}{\sigma} \left(\frac{2}{\pi}\right)^{1/2} e^{-t^2/2\sigma^2}$	$\frac{1}{\sigma} \left(\frac{2}{\pi}\right)^{1/2} \frac{f(x)}{1-F(x)} \exp(-(\log(1-F(x)))^2/2\sigma^2)$
Lévy	$\left(\frac{c}{2\pi}\right)^{1/2} \frac{e^{-c/2t}}{t^{3/2}}$	$\left(\frac{c}{2\pi}\right)^{1/2} \frac{f(x)}{F(x)} e^{-c/2(\log(1-F(x)))} (\log(1-F(x)))^{-3/2}$
Log-logistique	$\frac{\beta(t/\alpha)^{\beta-1}}{\alpha^{\beta}(1+(t/\alpha)^{\beta})^2}$	$\frac{\beta}{\alpha} \frac{f(x)}{F(x)(1-F(x))} (-\log(1-F(x)))^{\beta-1}$
Rayleigh	$\frac{t}{\sigma^2} e^{-t^2/2\sigma^2}$	$\frac{f(x)\log(1-F(x))}{\sigma^2(1-F(x))} \exp(-(\log(1-F(x)))^2/2\sigma^2)$
Gumbel Type-2	$\alpha \beta t^{-\alpha-1} e^{-\beta t^{-\alpha}}$	$\frac{\alpha \beta f(x)}{1-F(x)} (-\log(1-F(x)))^{-\alpha-1} \times \exp(-\beta\{-\log(1-F(x))\}^{-\alpha})$
Lomax	$\frac{\lambda k}{(1+\lambda x)^{k+1}}$	$\frac{f(x)}{(1-F(x))} \frac{\lambda k}{((1-\lambda \log(1-F(x))))^{k+1}}$
Beta inversée	$\frac{t^{\beta-1}(1+t)^{-\beta-\gamma}}{B(\beta, \gamma)}$	$\frac{f(x)}{B(\alpha, \gamma)(1-F(x))} \{(1-\log(1-F(x)))\}^{\beta-1} \times \{1-\log(1-F(x))\}^{-\beta-\gamma}$
Gaussienne inverse	$\sqrt{\frac{\lambda}{2\pi t^3}} e^{-\frac{\lambda}{2\mu^2 t}(t-\mu)^2}$	$\frac{f(x)}{1-F(x)} \left(\frac{\lambda}{2\pi(-\log(1-F(x)))^3}\right)^{1/2} \times \exp\left(\frac{-\lambda\{-\log(1-F(x))-\mu\}^2}{2\mu^2(-\log(1-F(x)))}\right)$
Weibull	$\frac{c}{\gamma} \left(\frac{t}{\gamma}\right)^{c-1} e^{-(t/\gamma)^c}$	$\frac{c}{\gamma} \frac{f(x)}{1-F(x)} \{-\log(1-F(x))\}^{c-1} \times \exp(-\{(1/\gamma)\log(1-F(x))\}^c)$

TABLE 2.2 – Familles de distributions généralisées dérivées de différentes distributions T

2.2.2.1 Famille Gamma-X

Soit T une variable aléatoire suivant une distribution gamma avec les paramètres α et β . Sa fonction de densité est donnée par :

$$r(t) = \frac{1}{\Gamma(\alpha)\beta^\alpha} t^{\alpha-1} e^{-t/\beta} \quad , \quad t > 0$$

La densité de probabilité de la famille gamma-X est définie par :

$$g(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} f(x) (-\log(1 - F(x)))^{\alpha-1} (1 - F(x))^{\frac{1}{\beta}-1} \quad (2.10)$$

En utilisant l'équation (2.4) et la fonction de répartition de la distribution gamma exprimée en termes de la fonction gamma incomplète, on obtient :

$$R(t) = \frac{1}{\Gamma(\alpha)} \gamma(\alpha, t/\beta)$$

où

$$\gamma(\alpha, t) = \int_0^t u^{\alpha-1} e^{-u} du$$

Ainsi, la fonction de répartition de la famille gamma-X est donnée par :

$$G(x) = \frac{\gamma(\alpha, -\log(1 - F(x)))}{\Gamma(\alpha)}$$

Les distributions de la forme $(F(x))^c$ et $(1 - F(x))^c$ sont respectivement appelées les familles Exp(F) et Exp(1-F).

Lemme 2.1 *L'entropie de Shannon de la famille gamma-X de distributions est donnée par :*

$$\eta_X = -\mathbb{E} \left[\log f \left(F^{-1}(1 - e^{-T}) \right) \right] + \alpha(1 - \beta) + \log \beta + \log \Gamma(\alpha) + (1 - \alpha)\psi(\alpha)$$

où ψ est la fonction digamma.

$$\psi(x) = \frac{d}{dx} \ln \Gamma(x)$$

La fonction digamma est la dérivée du logarithme de la fonction gamma. Elle intervient souvent dans le calcul d'espérances et d'entropies de distributions comme la gamma et la beta.

Preuve. Il s'ensuit du Théorème 2.1 en utilisant $\mu_T = \alpha\beta$ et l'entropie de Shannon pour la distribution gamma, qui est donnée par :

$$H(X) = \alpha + \ln(\beta) + \ln \Gamma(\alpha) + (1 - \alpha)\psi(\alpha)$$

où $\Gamma(\alpha)$ est la fonction gamma et $\psi(\alpha)$ est la fonction digamma.

$$\eta_T = \alpha + \log \beta + \log \Gamma(\alpha) + (1 - \alpha)\psi(\alpha)$$

■

Lorsque $\alpha = 1$, la famille gamma- X devient une distribution exponentielle. Pour $\alpha = n$ et $\beta = 1$, elle correspond à la densité de la n -ième valeur record supérieure d'une suite de variables aléatoires indépendantes et identiquement distribuées

La distribution gamma généralisée est un cas particulier où X suit une loi de Weibull, donnant :

$$g(x) = \frac{c}{\gamma^{\alpha c} \Gamma(\alpha) \beta^{\alpha}} x^{\alpha c - 1} e^{-\frac{1}{\beta}(x/\gamma)^c}, \quad c > 0 \quad (2.11)$$

En posant $\delta = \beta \gamma^c$ dans l'équation (2.11), la distribution se réduit à la distribution gamma généralisée d'Amoroso [14]. Lorsque $c = \gamma = 1$, elle devient la distribution gamma.

Si $f(x)$ est la densité de probabilité de la distribution de Pareto, alors à partir de (3.1), on obtient :

$$g(x) = \frac{k^{\alpha}}{x \Gamma(\alpha) \beta^{\alpha}} \left(\frac{\theta}{x}\right)^{k/\beta} \left(\log\left(\frac{x}{\theta}\right)\right)^{\alpha-1}, \quad x > \theta$$

En posant $\beta/k = c$, cela donne :

$$g(x) = \frac{1}{x \Gamma(\alpha) c^{\alpha}} \left(\frac{\theta}{x}\right)^{1/c} \left(\log\left(\frac{x}{\theta}\right)\right)^{\alpha-1}, \quad x > \theta \quad (2.12)$$

Cette distribution est appelée distribution gamma-Pareto. Lorsque $\alpha = 1$, elle se réduit à la distribution de Pareto, ce qui en fait une généralisation de cette dernière. La Figure 2.1 montre que le paramètre α rend la densité gamma-Pareto plus flexible, variant de la forme J inversé à une forme concave.

La fonction de répartition est donnée par :

$$G(x) = \frac{\gamma(\alpha, c^{-1} \log(x/\theta))}{\Gamma(\alpha)}$$

La fonction quantile est obtenue en résolvant l'équation $G(x) = p$, pour $0 \leq p \leq 1$.

L'analyse de l'asymétrie de Galton et de la kurtosis de Moors montre que :

- L'asymétrie de Galton est toujours positive, indiquant une distribution asymétrique à droite.
- Pour $c \geq 1$ fixé, elle est une fonction croissante de α .

- Pour $c < 1$ fixé, elle est décroissante en α , et pour α fixé, elle est croissante en c .
- La kurtosis de Moors est une fonction croissante de α et c .

Ces propriétés montrent la flexibilité de la distribution gamma-Pareto pour modéliser différentes formes de données.

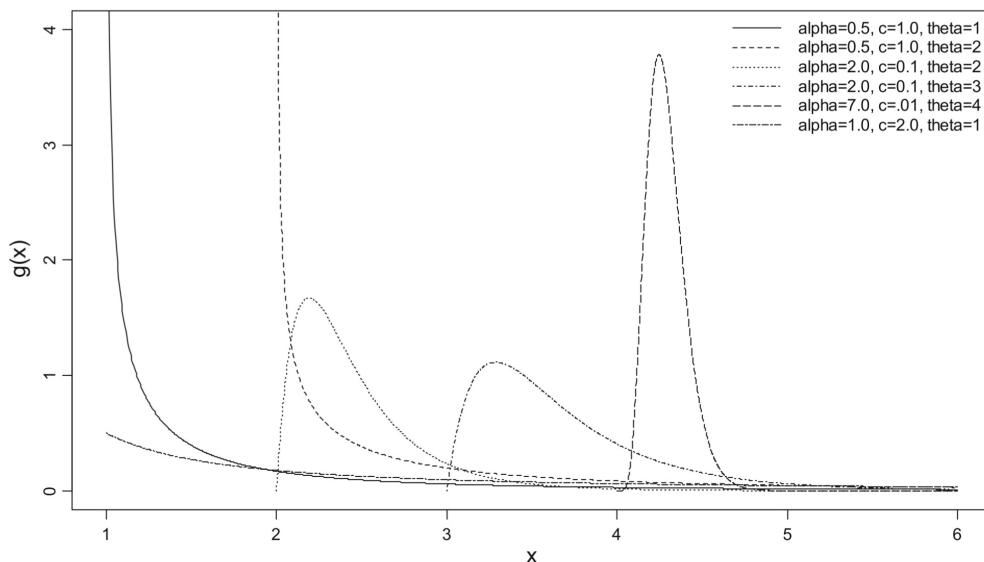


FIGURE 2.1 – Graphiques de la distribution gamma-Pareto pour différentes valeurs de paramètres.

2.2.2.2 Famille Bêta-exponentielle-X

Si une variable aléatoire T suit la distribution bêta-exponentielle selon Nadarajah and Kotz[15], alors la fonction de risque est donnée par :

$$r(t) = \lambda(B(\alpha, \beta))^{-1} e^{-\lambda\beta t} (1 - e^{-\lambda t})^{\alpha-1}$$

D'après (2.5), la fonction de densité de probabilité (f.d.p) de la famille bêta-exponentielle-X est définie par :

$$g(x) = \lambda B(\alpha, \beta)^{-1} f(x) (1 - F(x))^{\lambda\beta-1} \{1 - (1 - F(x))^\lambda\}^{\alpha-1} \quad (2.13)$$

La fonction de répartition de (2.13) peut être exprimée en termes de la fonction bêta incomplète $I_x(a, b)$. La fonction de répartition de la famille bêta-exponentielle-X est donnée par :

$$G(x) = 1 - I_{(1-F(x))^\lambda}(\lambda(\beta-1) + 1, \alpha)$$

Lemme 2.2 La mesure d'entropie de Shannon de la famille de distributions bêta-exponentielle-

X est donnée par :

$$\begin{aligned}\eta_X &= -E\left\{\log f\left(F^{-1}(1 - e^{-T})\right)\right\} + \log(\lambda^{-1}B(\alpha, \beta)) \\ &+ (\alpha + \beta - 1)\psi(\alpha + \beta) - (\alpha - 1)\psi(\alpha) - \beta\psi(\beta) - [\psi(\alpha + \beta) - \psi(\beta)]/\lambda\end{aligned}$$

Preuve. Cela découle du Théorème 2.1 en utilisant la moyenne

$$\mu_T = [\psi(\alpha + \beta) - \psi(\beta)]/\lambda$$

et l'entropie de Shannon

$$\eta_T = \log(\lambda^{-1}B(\alpha, \beta)) + (\alpha + \beta - 1)\psi(\alpha + \beta) - (\alpha - 1)\psi(\alpha) - \beta\psi(\beta)$$

pour la distribution bêta-exponentielle

■

Cas particuliers de la famille bêta-exponentielle- X :

La famille bêta-exponentielle- X généralise plusieurs distributions connues. Ses cas particuliers incluent :

- (1) La famille bêta-générée de (1.10) est un cas particulier de (2.13) lorsque $\lambda = 1$. Ainsi, cette famille permet de générer toutes les distributions bêta-générées.
- (2) La famille bêta-exponentielle- X se réduit à
Lorsque $\alpha = 1$, elle devient la distribution $\text{Exp}(1 - F(x))$.
Lorsque $\beta = 1$ et $\lambda = 1$, elle devient la distribution $\text{Exp}(F(x))$.
- (3) Lorsque $\beta = 1$, (2.13) se réduit à la famille exponentielle-exponentielle- X avec la f.d.p. :

$$g(x) = \alpha \lambda f(x) \left[1 - (1 - F(x))^\lambda\right]^{\alpha-1} (1 - F(x))^{\lambda-1} \quad (2.14)$$

La fonction de répartition de (2.14) peut s'écrire comme :

$$G(x) = \left[1 - (1 - F(x))^\lambda\right]^\alpha$$

En utilisant $D(x) = 1 - F(x)$ dans (2.14), la famille exponentielle-exponentielle- X se réduit à la famille K W-G.

Si X est une variable aléatoire uniforme, alors à partir de (3.4), la distribution bêta-exponentielle-uniforme est définie comme :

$$g(x) = \frac{\lambda}{B(\alpha, \beta)} \frac{1}{b-a} \left(\frac{b-x}{b-a}\right)^{\lambda\beta-1} \left\{1 - \left(\frac{b-x}{b-a}\right)^\lambda\right\}^{\alpha-1}, \quad a < x < b. \quad (2.15)$$

Si nous utilisons la transformation $y = 1 - x$ dans (2.15), alors la distribution se réduit à :

- (i) la distribution bêta généralisée de première espèce lorsque $b = 1$.
- (ii) la distribution bêta lorsque $a = 0$ et $b = \lambda = 1$.
- (iii) la distribution à double limite de Kumaraswamy lorsque $a = 0$ et $b = \beta = 1$.

La distribution exponentielle-Weibull est un membre de la famille exponentielle - exponentielle- X lorsque X est une variable aléatoire Weibull . Sa fonction de densité de probabilité (f.d.p.) de la distribution de Weibull est donnée par :

$$g(x) = \frac{c\lambda\alpha}{\gamma} \left(\frac{x}{\gamma}\right)^{c-1} \left(1 - e^{-\lambda(x/\gamma)^c}\right)^{\alpha-1} e^{-\lambda(x/\gamma)^c}, \quad x > 0; c, \gamma, \alpha, \lambda > 0 \quad (2.16)$$

En posant $\delta = \lambda\gamma^c$, la formule (2.16) devient celle de la distribution exponentielle-Weibull de Mudholkar et al [16].

- Si $\gamma = c = 1$, on obtient la distribution exponentielle-exponentielle
- Si $\lambda = c = 1$, on retrouve la distribution de Weibull.
- Si $\lambda = \gamma = c = 1$, la distribution devient exponentielle.

La distribution logistique généralisée de type I est un cas particulier de la distribution exponentielle-exponentielle-logistique. Lorsque $f(x)$ est la densité logistique standard, la f.d.p. devient :

$$g(x) = \frac{\alpha\lambda e^{-\lambda x}}{(1 + e^{-x})^{\lambda+1}} \left(1 - \frac{e^{-\lambda x}}{(1 + e^{-x})^\lambda}\right)^{\alpha-1}, \quad -\infty < x < \infty \quad (2.17)$$

La distribution exponentielle-exponentielle-logistique (2.17) devient :

- logistique généralisée de type I si $\lambda = 1$.
- logistique standard si $\alpha = \lambda = 1$.

La Figure 2.2 montre les graphiques des fonctions de densité exponentielles-exponentielles-logistiques pour différentes valeurs de paramètres, y compris les cas particuliers.

La fonction de répartition (c.d.f) de cette distribution, donnée par l'équation (2.17), est

$$G(x) = \left(1 - (1 + e^x)^{-\lambda}\right)^\alpha$$

Ainsi, la fonction quantile associée peut s'écrire comme suit :

$$Q(p) = \log\left(\left(1 - p^{1/\alpha}\right)^{-1/\lambda} - 1\right), \quad 0 \leq p \leq 1 \quad (2.18)$$

Les équations (2.18), (2.8) et (2.9) peuvent être utilisées pour obtenir l'asymétrie de Galton et la kurtose de Moors pour la distribution exponentielle-exponentielle-logistique. Elle peut être asymétrique à gauche, à droite ou symétrique selon α et λ

Pour $\lambda > 1$ fixé :

→ L'asymétrie de Galton augmente avec α .

→ La kurtose de Moors diminue avec λ .

Pour $\alpha > 1$ fixé :

→ L'asymétrie de Galton diminue avec λ .

→ La kurtose de Moors diminue avec α .

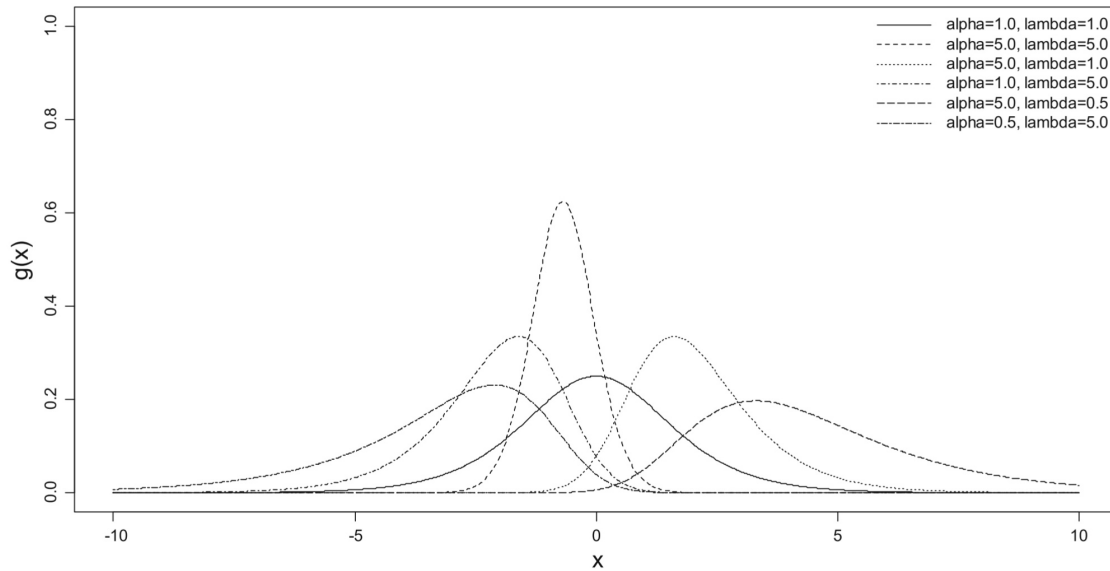


FIGURE 2.2 – Graphiques de la distribution exponentielle-exponentielle-logistique pour différentes valeurs de paramètres.

2.2.2.3 Famille Weibull-X

Si une variable aléatoire T suit une loi de Weibull avec les paramètres c et β , alors la densité de probabilité est donnée par :

$$r(t) = \frac{c}{\beta} \left(\frac{t}{\beta} \right)^{c-1} e^{-\left(\frac{t}{\beta} \right)^c}, \quad t \geq 0$$

D'après l'équation (2.5), la famille Weibull-X est donnée par :

$$g(x) = \frac{c}{\beta} \frac{f(x)}{1-F(x)} \cdot \left\{ \frac{-\log(1-F(x))}{\beta} \right\}^{c-1} \cdot \exp \left\{ - \left(\frac{-\log(1-F(x))}{\beta} \right)^c \right\} \quad (2.19)$$

La fonction de répartition (c.d.f) de la loi de Weibull est $R(t) = 1 - e^{-\left(\frac{t}{\beta} \right)^c}$ et donc, d'après l'équation (2.4), la fonction de répartition de la famille Weibull-X est :

$$G(x) = 1 - \exp \left\{ - \left[\frac{-\log(1-F(x))}{\beta} \right]^c \right\} \quad (2.20)$$

Lemme 2.3 *L'entropie de Shannon de la famille Weibull-X de distributions est donnée par :*

$$\eta_X = -E \left\{ \log f \left(F^{-1}(1 - e^{-T}) \right) \right\} - \beta \Gamma \left(1 + \frac{1}{c} \right) + \gamma \left(1 - \frac{1}{c} \right) - \log \left(\frac{c}{\beta} \right) + 1$$

où γ est la constante d'Euler.

Preuve. Cela découle du Théorème 2.1 en utilisant la moyenne :

$$\mu_T = \beta \Gamma \left(1 + \frac{1}{c} \right)$$

et l'entropie de Shannon :

$$\eta_T = \gamma \left(1 - \frac{1}{c} \right) - \log \left(\frac{c}{\beta} \right) + 1$$

pour la distribution de Weibull, comme indiqué par Song [17]

■

Cas particulier de la famille Weibull-X :

- (1) Lorsque $c = 1$, la famille Weibull-X se réduit à une distribution exponentielle de forme $\text{Exp}(1 - F(x))$.
- (2) La distribution logistique généralisée de type II apparaît comme un cas particulier de la distribution logistique-Weibull.
- (3) Si $F(x)$ représente la fonction de répartition (c.d.f) de la distribution logistique standard, alors l'équation (2.20) se simplifie en :

$$G(x) = 1 - \exp \left\{ - \left[\frac{\log(1 + e^x)}{\beta} \right]^c \right\} , \quad -\infty < x < \infty \quad (2.21)$$

- (4) Lorsque $c = 1$, cette distribution (2.21) devient précisément la distribution logistique généralisée de type II.

La Figure 2.3 montre les densités de la distribution Weibull-logistique pour différentes valeurs de paramètres. y compris le cas particulier.

À partir de (2.21), la fonction quantile de la distribution Weibull-logistique peut être écrite comme suit :

$$Q(p) = \log \left\{ \exp \left[\beta (-\log(1 - p))^{1/c} \right] - 1 \right\} , \quad 0 \leq p \leq 1 \quad (2.22)$$

Les équations (2.22) , (2.8) et (2.9) peuvent être utilisées pour obtenir l'asymétrie de Galton et l'aplatissement de Moors.

Pour un β fixé :

→ L'asymétrie de Galton est une fonction décroissante de c .

→ Pour un c fixé, l'asymétrie de Galton est une fonction croissante de β .

Pour un c fixé :

→ L'aplatissement de Moors est une fonction croissante de β lorsque $c \leq 1$.

→ L'aplatissement de Moors est une fonction décroissante de β lorsque $c > 1$.

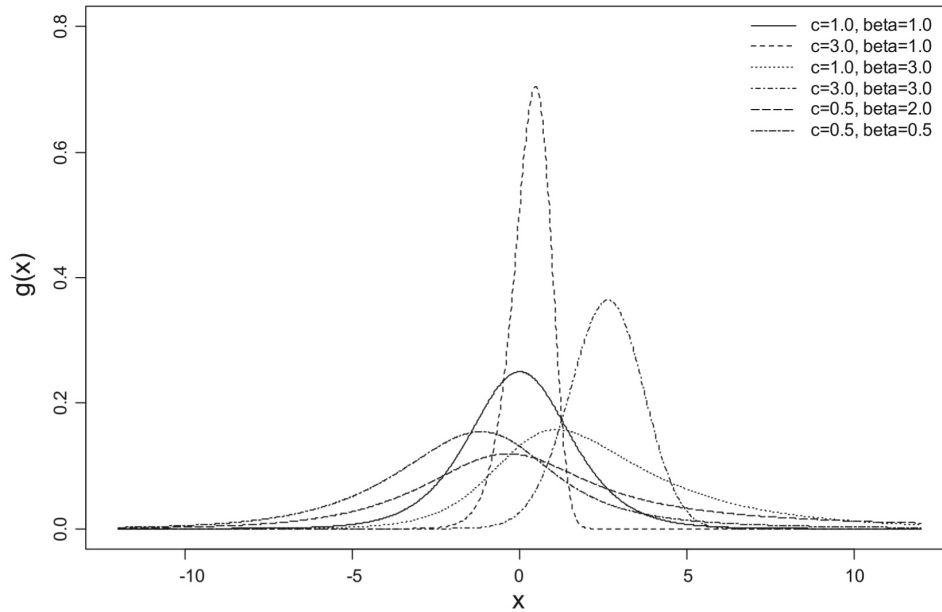


FIGURE 2.3 – Graphiques de la distribution Weibull-logistique pour différentes valeurs de paramètres

2.3 Résumé et conclusion

Cette méthode génère de nouvelles familles de distributions via la fonction composée $(R \circ W \circ F)(x)$, où $W(\cdot)$ relie T et X . Elle inclut des distributions existantes et est illustrée dans le tableau 2.1.

Cet article étudie la fonction $W(F(x)) = -\log(1 - F(x))$ avec T sur $[0, \infty)$. Il propose des mesures de l'asymétrie et de la kurtose (Galton et Moors). Trois sous-familles de la famille $T-X$ (gamma- X , beta-exponentielle- X , Weibull- X) sont explorées, montrant la diversité des distributions dans cette famille.

Le tableau 2.2 présente diverses sous-familles de distributions T et X , avec des relations possibles pour certaines combinaisons. Dans la distribution gamma, α est de forme et β d'échelle et pour Pareto, θ est d'échelle et k de forme. Donc dans la gamma-Pareto, θ reste d'échelle et $\frac{\beta}{k} = c$ devient de forme. Les paramètres de ces distributions

sont en cours d'étude, avec des applications montrant que la distribution gamma-Pareto est adaptée à divers types de données.

La figure 2.4 montre la relation en arbre de la famille $T - X$, avec de nombreuses sous-familles et des cas particuliers de distributions existantes.

Le tableau 2.1 présente des variantes de la famille T-X générant de nouvelles distributions nécessitant des recherches approfondies. Les études futures devront :

- (i) Analyser les propriétés des distributions créées via différentes fonctions $W(.)$
- (ii) Explorer de nouvelles distributions spécifiques
- (iii) Développer des méthodes d'estimation adaptées (au-delà de 3 paramètres)

La méthode ML est difficile pour plus de trois paramètres.

De nouvelles distributions plus générales et flexibles ont été développées, notamment dans la famille $T-X$, adaptées aux distributions asymétriques, à queues lourdes ou bimodales. La distribution bêta-normale a été utilisée avec succès pour ajuster des données réelles par Eugene et al [6]. et Famoye et al [18]. La famille $T-X$ peut ajuster des distributions unimodales, bimodales et multimodales.

L'article traite des distributions $T-X$, où T et X sont continues ou T est continue et X est discrète, nécessitant des fonctions $W(.)$ spécifiques.

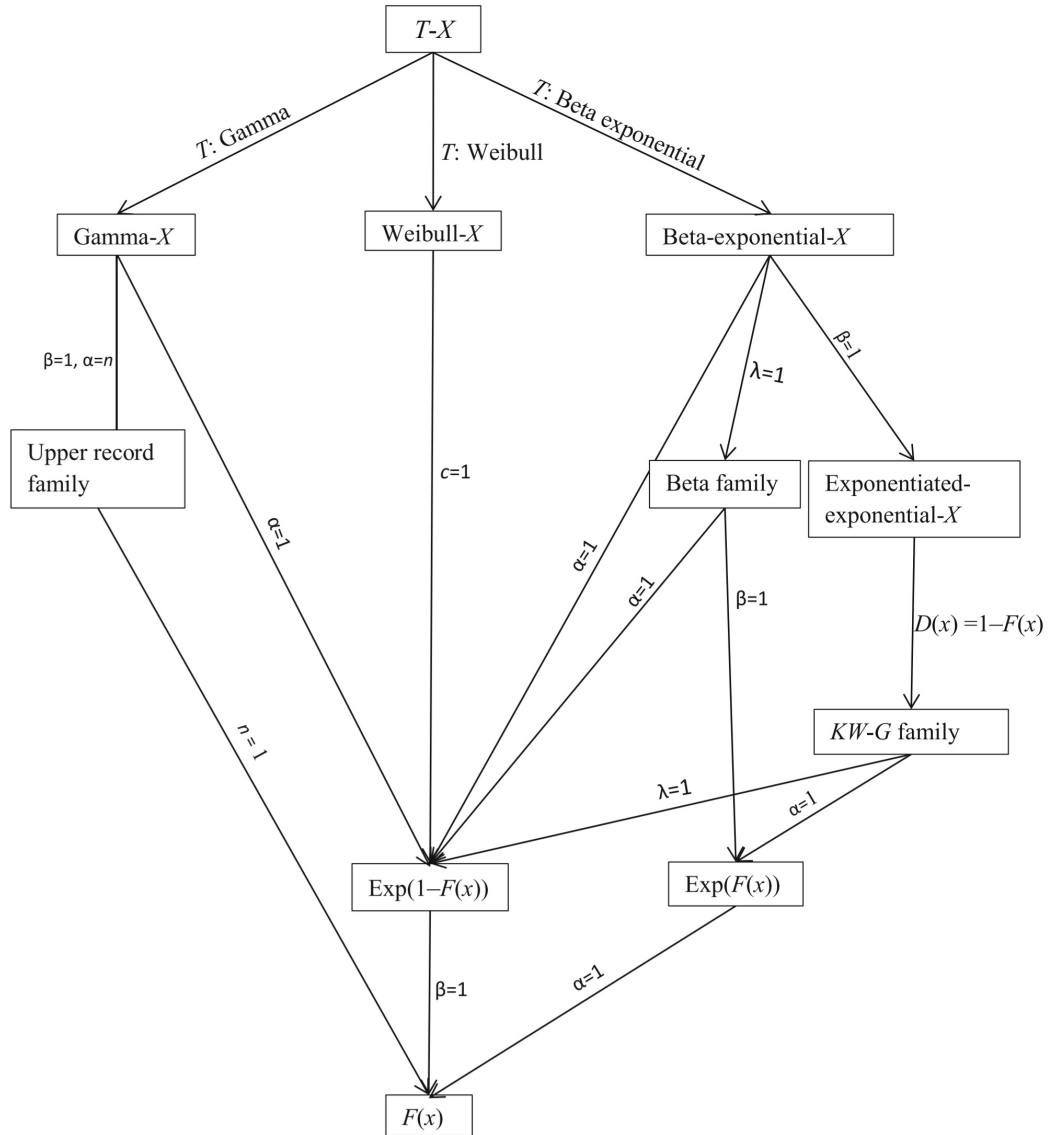


FIGURE 2.4 – Sous-familles de la famille de distributions T-X

CHAPITRE 3

ÉTUDE D'UNE NOUVELLE DISTRIBUTION.

Introduction

La distribution Lomax, est une généralisation du modèle de Pareto de type II, est utilisée pour l'analyse de la durée de vie, la fiabilité et l'analyse de survie. Plusieurs extensions de cette distribution ont été proposées. Par exemple, la distribution de Lomax transformée par une loi bêta (BTL) ajoute de la flexibilité en intégrant des paramètres supplémentaires, permettant ainsi de mieux modéliser les données réelles.

3.1 Distribution Lomax

Soit X une variable aléatoire suivant une loi de Lomax de paramètres $\lambda > 0$ (échelle) et $\alpha > 0$ (forme), ce que l'on note $X \sim \text{Lomax}(\lambda, \alpha)$, si sa fonction de répartition est donnée par :

$$F(x) = 1 - \left(1 + \frac{x}{\lambda}\right)^{-\alpha}, \quad x \geq 0 \quad (1.1)$$

La fonction de densité de probabilité est :

$$f(x) = \frac{\alpha}{\lambda} \left(1 + \frac{x}{\lambda}\right)^{-(\alpha+1)}, \quad x \geq 0 \quad (1.2)$$

La fonction de survie du modèle Lomax est donnée par :

$$S(x) = 1 - F(x) = \left(1 + \frac{x}{\lambda}\right)^{-\alpha}, \quad x \geq 0 \quad (1.3)$$

La fonction quantile $Q(u)$ est l'inverse de la fonction de répartition $F(x)$. Pour un quantile d'ordre $u \in (0, 1)$, tel que $F(x) = u$, alors la fonction quantile est définie par :

$$Q(u) = \lambda \left((1 - u)^{-\frac{1}{\alpha}} - 1 \right), \quad u \in (0, 1) \quad (1.4)$$

Pour simuler une variable provenant de la distribution Lomax, nous utilisons l'équation :

$$x = \lambda \left((1 - u)^{-\frac{1}{\alpha}} - 1 \right)$$

où u est une variable aléatoire uniforme sur l'intervalle $[0, 1]$

La fonction de risque est définie par :

$$h(x) = \frac{f(x)}{s(x)} = \frac{\alpha}{x + \lambda} \quad , \quad x \geq 0 \quad (1.5)$$

Son espérance et sa variance sont données par :

$$\mathbb{E}[X] = \frac{\lambda}{\alpha - 1} \quad , \quad \text{pour } \alpha > 1$$

$$\mathbb{V}ar(X) = \frac{\lambda^2 \alpha}{(\alpha - 1)^2 (\alpha - 2)} \quad , \quad \text{pour } \alpha > 2$$

3.2 Distribution de Lomax transformée par une loi bêta

Soit X une variable aléatoire suivant une loi de Lomax classique, dont la fonction de répartition (c.d.f) est :

$$F(x) = 1 - \left(1 + \frac{x}{\lambda} \right)^{-\alpha} \quad , \quad x \geq 0$$

où $\lambda > 0$ et $\alpha > 0$

Par l'utilisation de la **Beta Transformation (BT)** proposée par Fasna (2022) [10] appliquée à la fonction de répartition du modèle Lomax notée $F(x)$, est définie par :

$$G(x) = \begin{cases} \frac{\beta}{\beta - 1} \left[1 - \beta^{-F(x)} \right] & \text{si } \beta \neq 1, \beta > 0 \\ F(x) & \text{si } \beta = 1 \end{cases} \quad (2.1)$$

avec $\beta > 0$

En remplaçant $F(x)$ dans (2.1) par la fonction de répartition de Lomax, nous obtenons explicitement :

$$G(x) = \begin{cases} \frac{\beta}{\beta - 1} \left[1 - \beta^{-\left(1 - \left(1 + \frac{x}{\lambda} \right)^{-\alpha} \right)} \right] & \text{si } \beta \neq 1, \beta > 0 \\ 1 - \left(1 + \frac{x}{\lambda} \right)^{-\alpha} & \text{si } \beta = 1 \end{cases} \quad , \quad x \geq 0 \quad (2.2)$$

La densité BT transformée associée à (2.1) est donnée par :

$$g(x) = \begin{cases} \frac{\beta}{\beta-1} \ln(\beta) \beta^{-F(x)} f(x) & \text{si } \beta \neq 1, \beta > 0 \\ f(x) & \text{si } \beta = 1 \end{cases}$$

où $f(x)$ est la densité classique de Lomax

Maintenant, en remplaçant $F(x)$ et $f(x)$, la densité explicite devient :

$$g(x) = \begin{cases} \frac{\beta}{\beta-1} \ln(\beta) \beta^{-(1-(1+\frac{x}{\lambda})^{-\alpha})} \frac{\alpha}{\lambda} \left(1 + \frac{x}{\lambda}\right)^{-(\alpha+1)} & \text{si } \beta \neq 1, \beta > 0 \\ \frac{\alpha}{\lambda} \left(1 + \frac{x}{\lambda}\right)^{-(\alpha+1)} & \text{si } \beta = 1 \end{cases}, \quad x \geq 0 \quad (2.3)$$

La fonction de survie $s(x)$ est :

$$S(x) = \begin{cases} 1 - \frac{\beta}{\beta-1} (1 - \beta^{-F(x)}) & \text{si } \beta \neq 1, \beta > 0 \\ 1 - F(x) & \text{si } \beta = 1 \end{cases}$$

Puis, en utilisant l'équation (1.3) pour Lomax classique, on obtient :

$$S(x) = \begin{cases} \frac{1}{\beta-1} (\beta^{-F(x)} - 1) & \text{si } \beta \neq 1, \beta > 0 \\ \left(1 + \frac{x}{\lambda}\right)^{-\alpha} & \text{si } \beta = 1 \end{cases}, \quad x \geq 0 \quad (2.4)$$

La fonction de risque $h(x)$ est donnée par :

$$h(x) = \begin{cases} \frac{\beta}{\beta-1} \ln(\beta) \frac{\beta^{-F(x)} f(x)}{\frac{1}{\beta-1} (\beta^{-F(x)} - 1)} & \text{si } \beta \neq 1, \beta > 0 \\ \frac{f(x)}{\left(1 + \frac{x}{\lambda}\right)^{-\alpha}} & \text{si } \beta = 1 \end{cases}$$

Après simplification, la fonction de risque $h(x)$ s'exprime comme :

$$h(x) = \begin{cases} \frac{\beta \ln(\beta) \beta^{-F(x)} f(x)}{\beta^{-F(x)} - 1} & \text{si } \beta \neq 1, \beta > 0 \\ \frac{\alpha}{x + \lambda} & \text{si } \beta = 1 \end{cases}, \quad x \geq 0 \quad (2.5)$$

L'espérance de la distribution BTL s'écrit sous forme intégrale :

$$\mathbb{E}[X] = \int_0^{+\infty} x g(x) dx$$

En substituant $g(x)$, on obtient :

$$\mathbb{E}[X] = \int_0^{+\infty} x g(x) dx = \begin{cases} \frac{\beta}{\beta-1} \ln(\beta) \int_0^{+\infty} x \beta^{-F(x)} f(x) dx & \text{si } \beta \neq 1, \beta > 1 \\ E_F[X] & \beta = 1 \end{cases}$$

$$\mathbb{E}[X] = \begin{cases} \frac{\beta}{\beta-1} \ln(\beta) \int_0^{+\infty} x \beta^{-(1-(1+\frac{x}{\lambda})^{-\alpha})} \frac{\alpha}{\lambda} \left(1 + \frac{x}{\lambda}\right)^{-(\alpha+1)} dx & \text{si } \beta \neq 1, \beta > 0 \\ E_F[X] & \text{si } \beta = 1 \end{cases}$$

où $E_F[X]$ est l'espérance classique de Lomax est :

$$E_F[X] = \frac{\lambda}{\alpha-1}, \quad \alpha > 1$$

La variance de la distribution BTL est donnée par :

$$\text{Var}[X] = E[X^2] - (E[X])^2$$

avec

$$\mathbb{E}[X^2] = \int_0^{+\infty} x^2 g(x) dx = \begin{cases} \frac{\beta}{\beta-1} \ln(\beta) \int_0^{+\infty} x^2 \beta^{-F(x)} f(x) dx & \text{si } \beta \neq 1, \beta > 1 \\ E_F[X^2] & \text{si } \beta = 1 \end{cases}.$$

En substituant $g(x)$ et $E[X^2]$, on obtient :

$$\text{Var}[X] = \begin{cases} \frac{\beta}{\beta-1} \ln(\beta) \int_0^{+\infty} x^2 \cdot \beta^{-F(x)} f(x) dx - (\mathbb{E}[X])^2 & \text{si } \beta \neq 1, \beta > 0 \\ \text{Var}_F[X] & \text{si } \beta = 1 \end{cases}$$

où $\text{Var}_F[X]$ la variance classique de Lomax est :

$$\text{Var}_F[X] = \frac{\lambda^2 \alpha}{(\alpha-1)^2(\alpha-2)}, \quad \alpha > 2$$

3.3 Propriétés mathématiques de base

3.3.1 Fonction quantile

La fonction quantile $Q(u)$ de la distribution de Lomax transformée par la méthode bêta donne la valeur pour une probabilité $u \in (0, 1)$. Elle est définie par :

$$Q(u) = \begin{cases} \lambda \left[\left(1 + \frac{\ln\left(1 - \frac{(\beta-1)u}{\beta}\right)}{\ln \beta} \right)^{-1/\alpha} - 1 \right] & \text{si } \beta > 0, \beta \neq 1, \beta > 0 \\ \lambda \left[(1-u)^{-1/\alpha} - 1 \right] & \text{si } \beta = 1 \end{cases}$$

où :

$\lambda > 0$ est un paramètre d'échelle.

$\alpha > 0, \beta > 0$ sont des paramètres de forme.

Preuve. Le quantile d'ordre q , noté x_q , de la variable aléatoire BTC est défini par :

$$u = P(X \leq x) = F(x) \quad , \quad x > 0$$

En utilisant la CDF dans (2.2) transformée , nous avons :

- **Cas $\beta = 1$:** Posons $G(Q(u)) = u$, on a :

$$1 - \left(1 + \frac{Q(u)}{\lambda}\right)^{-\alpha} = u$$

$$\left(1 + \frac{Q(u)}{\lambda}\right)^{-\alpha} = 1 - u$$

$$1 + \frac{Q(u)}{\lambda} = (1 - u)^{-1/\alpha}$$

On obtient donc :

$$Q(u) = \lambda \left[(1 - u)^{-1/\alpha} - 1 \right]$$

- **Cas général ($\beta > 0, \beta \neq 1$) :** Posons $F(Q(u)) = u$, on a :

$$\frac{\beta}{\beta - 1} \left[1 - \beta^{-F(Q(u))} \right] = u$$

D'où :

$$\beta^{-F(Q(u))} = 1 - \frac{(\beta - 1)u}{\beta}$$

En prenant le logarithme en base β :

$$\begin{aligned} -F(Q(u)) &= \ln \beta \left(1 - \frac{(\beta-1)u}{\beta} \right) \\ F(Q(u)) &= -\log \beta \left(1 - \frac{\beta-1}{\beta} u \right) \end{aligned}$$

En utilisant la définition de $F(Q(u))$:

$$\begin{aligned} 1 - \left(1 + \frac{Q(u)}{\lambda} \right)^{-\alpha} &= -\ln \beta \left(1 - \frac{\beta-1}{\beta} u \right) \\ \left(1 + \frac{Q(u)}{\lambda} \right)^{-\alpha} &= 1 + \ln \beta \left(1 - \frac{\beta-1}{\beta} u \right) \\ \left(1 + \frac{Q(u)}{\lambda} \right)^{-\alpha} &= 1 + \frac{\ln \left(1 - \frac{\beta-1}{\beta} u \right)}{\ln \beta} \end{aligned}$$

Enfin, on isole x :

$$\begin{aligned} 1 + \frac{Q(u)}{\lambda} &= \left[1 + \frac{\ln \left(1 - \frac{\beta-1}{\beta} u \right)}{\ln \beta} \right]^{-\frac{1}{\alpha}} \\ Q(u) &= \lambda \left\{ \left[1 + \frac{\ln \left(1 - \frac{\beta-1}{\beta} u \right)}{\ln \beta} \right]^{-\frac{1}{\alpha}} - 1 \right\} \end{aligned}$$

■

3.3.2 Statistiques d'ordre

Soit un échantillon de taille n de variables aléatoires indépendantes et identiquement distribuées suivant la distribution transformée G .

Les statistiques d'ordre sont les valeurs triées :

$$X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$$

La fonction de répartition du k -ième ordre est donnée par :

$$\begin{aligned} F_{X_{(k)}}(x) &= P(X_k \leq x) = \sum_{j=k}^n \binom{n}{j} [G(x)]^j [1 - G(x)]^{n-j} \\ F_{X_{(k)}}(x) &= \sum_{i=k}^n \binom{n}{i} \left(\frac{\beta}{\beta-1} \left[1 - \beta^{-(1-(1+x/\lambda)^{-\alpha})} \right] \right)^i \left(1 - \frac{\beta}{\beta-1} \left[1 - \beta^{-(1-(1+x/\lambda)^{-\alpha})} \right] \right)^{n-i} \end{aligned}$$

La fonction de densité du k -ième ordre est :

$$f_{X_{(k)}}(x) = \frac{n!}{(k-1)!(n-k)!} [G(x)]^{k-1} [1-G(x)]^{n-k} g(x)$$

$$f_{X_{(k)}}(x) = \frac{n!}{(k-1)!(n-k)!} \left(\frac{\beta}{\beta-1} \left[1 - \beta^{-(1+(x/\lambda)^{-\alpha})} \right] \right)^{k-1}$$

$$\times \left(1 - \frac{\beta}{\beta-1} \left[1 - \beta^{-(1+(x/\lambda)^{-\alpha})} \right] \right)^{n-k}$$

$$\times \left(\frac{\alpha \beta \log \beta}{\lambda(\beta-1)} (1+x/\lambda)^{-(\alpha+1)} \beta^{-(1+(x/\lambda)^{-\alpha})} \right)$$

où :

$F_{X(k)}$ est la fonction de répartition transformée

$f_{X(k)}$ est la densité de la distribution transformée

Des cas particuliers comme le minimum X_1 ou le maximum X_n s'obtiennent en posant $k = 1$ ou $k = n$.

3.4 Estimation des paramètres

Dans cette section, nous utilisons deux méthodes la méthode du maximum de vraisemblance et la méthode de Cramér-von Mises pour estimer les paramètres des distributions BTL

On suppose que $X_1, X_2, \dots, X_n$ est un échantillon aléatoire de taille n issu d'une distribution dont les paramètres sont inconnus.

3.4.1 Méthode de Cramér-von Mises

Les estimateurs de Cramér-von Mises pour la distribution Lomax-Bêta (BTL) minimisent la distance entre la fonction de répartition empirique et théorique $G(x | \alpha, \lambda, \beta)$

Pour un échantillon ordonné $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$, la statistique W^2 peut être calculée de manière équivalente par la formule :

$$W^2 = \sum_{i=1}^n \left[G(x_{(i)}) - \frac{2i-1}{2n} \right]^2 + \frac{1}{12n}$$

Cette formule est plus pratique pour le calcul lorsque la fonction de répartition théorique $F(x)$ est connue.

- **Cas $\beta = 1$:** Lorsque $\beta = 1$, la transformation de Fasna se réduit à la distribution de base elle-même. Donc, $G(x) = F(x)$, qui est la CDF de la distribution de Lomax :

$$G(x; \alpha, \lambda, \beta) = 1 - \left(1 + \frac{x}{\lambda}\right)^{-\alpha}$$

Dans ce cas, la statistique de Cramér-von Mises à minimiser pour un échantillon ordonné $x_{(1)}, \dots, x_{(n)}$ est :

$$W^2(\alpha, \lambda) = \frac{1}{12n} + \sum_{i=1}^n \left[G(x_{(i)}; \alpha, \lambda, 1) - \frac{2i-1}{2n} \right]^2$$

$$W^2(\alpha, \lambda) = \frac{1}{12n} + \sum_{i=1}^n \left[1 - \left(1 + \frac{x_{(i)}}{\lambda}\right)^{-\alpha} - \frac{2i-1}{2n} \right]^2$$

Les estimations des paramètres $\hat{\alpha}$ et $\hat{\lambda}$ sont obtenues en minimisant cette fonction $W^2(\alpha, \lambda)$ par rapport à $\alpha > 0$ et $\lambda > 0$.

- **Cas $\beta > 0, \beta \neq 1$:** Lorsque β est différent de 1, la CDF transformée est :

$$G(x; \alpha, \lambda, \beta) = \frac{\beta}{\beta-1} \left[1 - \beta^{-(1-(1+\frac{x}{\lambda})^{-\alpha})} \right]$$

Dans ce cas, la statistique de Cramér-von Mises à minimiser pour un échantillon ordonné $x_{(1)}, \dots, x_{(n)}$ est :

$$W^2(\alpha, \lambda, \beta) = \frac{1}{12n} + \sum_{i=1}^n \left[G(x_{(i)}; \alpha, \lambda, \beta) - \frac{2i-1}{2n} \right]^2$$

$$W^2(\alpha, \lambda, \beta) = \frac{1}{12n} + \sum_{i=1}^n \left[\frac{\beta}{\beta-1} \left[1 - \beta^{-(1-(1+\frac{x_i}{\lambda})^{-\alpha})} \right] - \frac{2i-1}{2n} \right]^2$$

Les estimations des paramètres $\hat{\alpha}$, $\hat{\lambda}$ et $\hat{\beta}$ sont obtenues en minimisant cette fonction $W^2(\alpha, \lambda, \beta)$ par rapport à $\alpha > 0$, $\lambda > 0$ et $\beta > 0, \beta \neq 1$.

3.4.2 Méthode du maximum de vraisemblance

Soit un paramètre inconnu α à estimer et $X = (x_1, x_2, \dots, x_n)$. l' échantillon observé, la fonction de vraisemblance est définie par :

$$L(\alpha | X) = P\alpha(X = x_1, x_2, \dots, x_n)$$

ou, dans le cas d'une densité de probabilité,

$$L(\alpha | X) = \prod_{i=1}^n f_{\alpha}(x_i)$$

où f_{α} est la densité ou la fonction de masse associée à la loi paramétrée par θ .

Pour simplifier le calcul, on maximise souvent la log-vraisemblance, définie par :

$$\ell(\alpha | X) = \log L(\alpha | X) = \sum_{i=1}^n \log f_{\alpha}(x_i)$$

L'estimation se fait alors en résolvant l'équation :

$$\frac{\partial \ell(\alpha | X)}{\partial \alpha} = 0$$

et en vérifiant que la solution correspond à un maximum (dérivée seconde négative) .

Pour estimer les paramètres α, β, λ de la BT-Lomax, voici la procédure complète du maximum de vraisemblance :

- **Cas $\beta = 1$:** La fonction de densité de probabilité est :

$$g(x) = \frac{\alpha}{\lambda} \left(1 + \frac{x}{\lambda}\right)^{-(\alpha+1)}$$

La fonction de Vraisemblance s'écrit :

$$\begin{aligned} L(\alpha, \lambda) &= \prod_{i=1}^n g(x_i) \\ L(\alpha, \lambda) &= \prod_{i=1}^n \frac{\alpha}{\lambda} \left(1 + \frac{x_i}{\lambda}\right)^{-(\alpha+1)} \\ L(\alpha, \lambda) &= \left(\frac{\alpha}{\lambda}\right)^n \prod_{i=1}^n \left(1 + \frac{x_i}{\lambda}\right)^{-(\alpha+1)} \end{aligned}$$

La fonction de log-vraisemblance est :

$$\ell(\alpha, \lambda) = n \ln\left(\frac{\alpha}{\lambda}\right) - (\alpha + 1) \sum_{i=1}^n \ln\left(1 + \frac{x_i}{\lambda}\right)$$

Les équations du maximum de vraisemblance sont :

$$\begin{cases} \frac{\partial \ell}{\partial \alpha} = \frac{n}{\alpha} - \sum_{i=1}^n \ln\left(1 + \frac{x_i}{\lambda}\right) = 0 \\ \frac{\partial \ell}{\partial \lambda} = -\frac{n}{\lambda} + (\alpha + 1) \sum_{i=1}^n \frac{x_i}{\lambda^2(1 + \frac{x_i}{\lambda})} = 0 \end{cases}$$

Ces équations pour $\hat{\alpha}$ et $\hat{\lambda}$ doivent généralement être résolues numériquement de manière itérative.

- **Cas $\beta > 0, \beta \neq 1$:** Dans ce cas, la distribution est la distribution de Lomax standard avec fonction de densité :

$$g(x) = \frac{\alpha\beta \ln(\beta)}{\lambda(\beta-1)} \left(1 + \frac{x}{\lambda}\right)^{-(\alpha+1)} \beta^{-(1-(1+\frac{x}{\lambda})^{-\alpha})}$$

La fonction de vraisemblance est donc :

$$\begin{aligned} L(\alpha, \lambda, \beta) &= \prod_{i=1}^n g(x_i) \\ L(\alpha, \lambda, \beta) &= \prod_{i=1}^n \left[\frac{\alpha\beta \ln(\beta)}{\lambda(\beta-1)} \left(1 + \frac{x_i}{\lambda}\right)^{-(\alpha+1)} \beta^{-(1-(1+\frac{x_i}{\lambda})^{-\alpha})} \right] \\ L(\alpha, \lambda, \beta) &= \left(\frac{\alpha\beta \ln(\beta)}{\lambda(\beta-1)} \right)^n \left(\prod_{i=1}^n \left(1 + \frac{x_i}{\lambda}\right)^{-(\alpha+1)} \right) \left(\prod_{i=1}^n \beta^{-(1-(1+\frac{x_i}{\lambda})^{-\alpha})} \right) \end{aligned}$$

On considère la fonction de log-vraisemblance $\ell(\alpha, \lambda, \beta)$:

$$\begin{aligned} \ell(\alpha, \lambda, \beta) &= \ln(L(\alpha, \lambda, \beta)) = \sum_{i=1}^n \ln g(x_i) \\ \ell(\alpha, \lambda, \beta) &= n \ln \left(\frac{\alpha\beta \ln(\beta)}{\lambda(\beta-1)} \right) - (\alpha+1) \sum_{i=1}^n \ln \left(1 + \frac{x_i}{\lambda} \right) - \sum_{i=1}^n \left(1 - \left(1 + \frac{x_i}{\lambda} \right)^{-\alpha} \right) \ln(\beta) \\ \ell(\alpha, \lambda, \beta) &= n \ln \left(\frac{\alpha \ln(\beta)}{\lambda(\beta-1)} \right) + n \ln(\beta) - (\alpha+1) \sum_{i=1}^n \ln \left(1 + \frac{x_i}{\lambda} \right) - n \ln(\beta) + \ln(\beta) \sum_{i=1}^n \left(1 + \frac{x_i}{\lambda} \right)^{-\alpha} \\ \ell(\alpha, \lambda, \beta) &= n \ln \left(\frac{\alpha \ln(\beta)}{\lambda(\beta-1)} \right) - (\alpha+1) \sum_{i=1}^n \ln \left(1 + \frac{x_i}{\lambda} \right) + \ln(\beta) \sum_{i=1}^n \left(1 + \frac{x_i}{\lambda} \right)^{-\alpha} \end{aligned}$$

On résout les dérivées partielles de $\ell(\alpha, \lambda, \beta)$ par rapport à α , λ , et β :

$$\begin{cases} \frac{\partial \ell}{\partial \alpha} = \frac{n}{\alpha} - \sum_{i=1}^n \ln \left(1 + \frac{x_i}{\lambda} \right) - \ln(\beta) \sum_{i=1}^n \left(1 + \frac{x_i}{\lambda} \right)^{-\alpha} \ln \left(1 + \frac{x_i}{\lambda} \right) = 0 \\ \frac{\partial \ell}{\partial \lambda} = -\frac{n}{\lambda} + (\alpha+1) \sum_{i=1}^n \frac{x_i}{\lambda^2(1+\frac{x_i}{\lambda})} - \alpha \ln(\beta) \sum_{i=1}^n \left(1 + \frac{x_i}{\lambda} \right)^{-(\alpha+1)} \frac{x_i}{\lambda^2} = 0 \\ \frac{\partial \ell}{\partial \beta} = \frac{n}{\beta \ln(\beta)} + \frac{n}{\ln(\beta)} - \frac{n}{\beta-1} + \sum_{i=1}^n \left(1 + \frac{x_i}{\lambda} \right)^{-\alpha} \ln \left(1 + \frac{x_i}{\lambda} \right) = 0 \end{cases}$$

La résolution analytique de ce système d'équations est généralement difficile et nécessite souvent des méthodes numériques.

CHAPITRE 4

SIMULATIONS ET APPLICATIONS

Introduction

Afin d'analyser le modèle créé et notée Lomax transformé par Beta (BTL), nous proposons des études de simulations. L'approche repose sur deux étapes claires.

La génération de la v. a. suit une distribution BTL avec différentes valeurs des paramètres et de différents taille de l'échantillon.

L'estimation des paramètres α, λ et β en appliquant deux méthodes distinctes : la méthode de Cramér–von Mises et l'estimation par maximum de vraisemblance. Les erreurs quadratiques moyennes associées à chaque estimateur sont calculées pour évaluer leur précision.

La validation du modèle. Les résultats obtenus démontrent clairement l'adéquation du modèle BTL avec les observations empiriques, confirmant ainsi les conclusions des simulations numériques

La distribution BT-Lomax (Lomax transformée par Beta selon Fasna, 2022) est une généralisation flexible de la distribution Lomax classique, permettant une meilleure modélisation des données à queues lourdes avec asymétrie variable. sa fonction de répartition est donnée par :

$$G(x) = \begin{cases} \frac{\beta}{\beta-1} \left[1 - \beta^{-(1+(1+\frac{x}{\lambda})^{-\alpha})} \right] & \text{si } \beta \neq 1, \beta > 0 \\ 1 - \left(1 + \frac{x}{\lambda} \right)^{-\alpha} & \text{si } \beta = 1 \end{cases}, \quad x \geq 0.$$

La fonction de densité BT transformée BT transformée est :

$$g(x) = \begin{cases} \frac{\beta}{\beta-1} \ln(\beta) \beta^{-\left(1+\left(1+\frac{x}{\lambda}\right)^{-\theta}\right)} \frac{\alpha}{\lambda} \left(1+\frac{x}{\lambda}\right)^{-(\alpha+1)} & \text{si } \beta \neq 1, \beta > 0 \\ \frac{\alpha}{\lambda} \left(1+\frac{x}{\lambda}\right)^{-(\alpha+1)} & \text{si } \beta = 1 \end{cases}, \quad x \geq 0$$

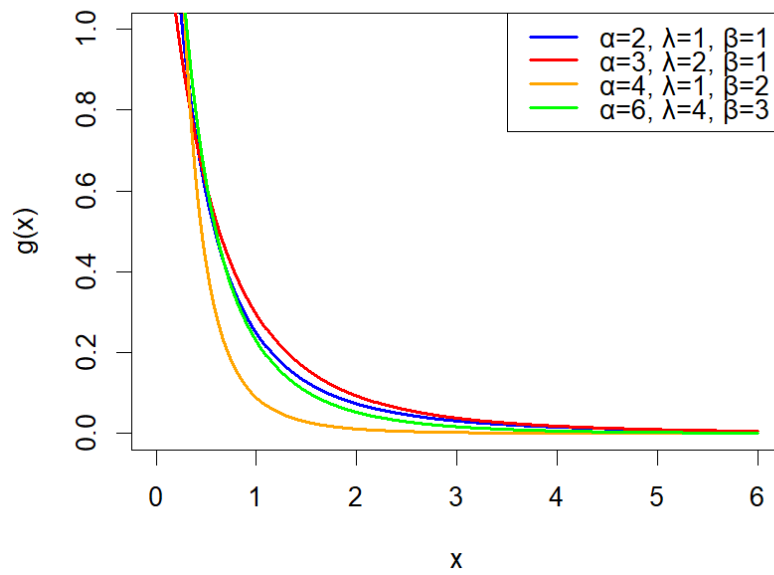


FIGURE 4.1 – PDF de la distribution BTL pour différentes valeurs des paramètres

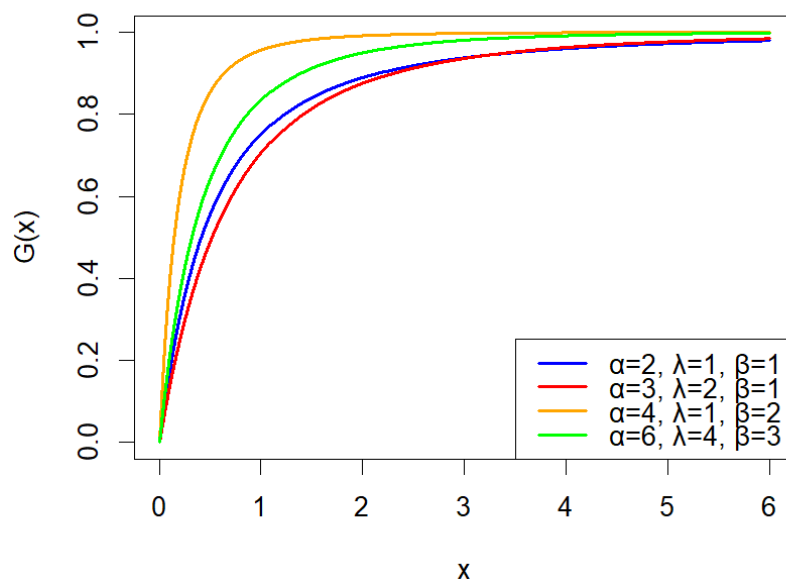


FIGURE 4.2 – CDF de la distribution BTL pour différentes valeurs des paramètres

4.1 Simulation du BTL

Pour simuler une variable aléatoire x suivant une distribution donnée, la méthode d'inversion consiste à :

1. Générer une variable uniforme $u \sim \mathcal{U}(0, 1)$.
2. Calculer $x = Q(u)$ où Q est la fonction quantile (l'inverse de la fonction de répartition).

La fonction quantile $Q(u)$ de la distribution Lomax transformée par beta (BTL) est donnée par :

$$Q(u) = \begin{cases} \lambda \left[\left(1 + \frac{\ln\left(1 - \frac{(\beta-1)u}{\beta}\right)}{\ln \beta} \right)^{-1/\alpha} - 1 \right] & \text{si } \beta \neq 1, \beta > 0 \\ \lambda \left[(1-u)^{-1/\alpha} - 1 \right] & \text{si } \beta = 1 \end{cases}$$

où :

$\lambda > 0$ est un paramètre d'échelle.

$\alpha > 0, \beta > 0$ sont des paramètres de forme.

$u \in (0, 1)$ est une variable uniforme.

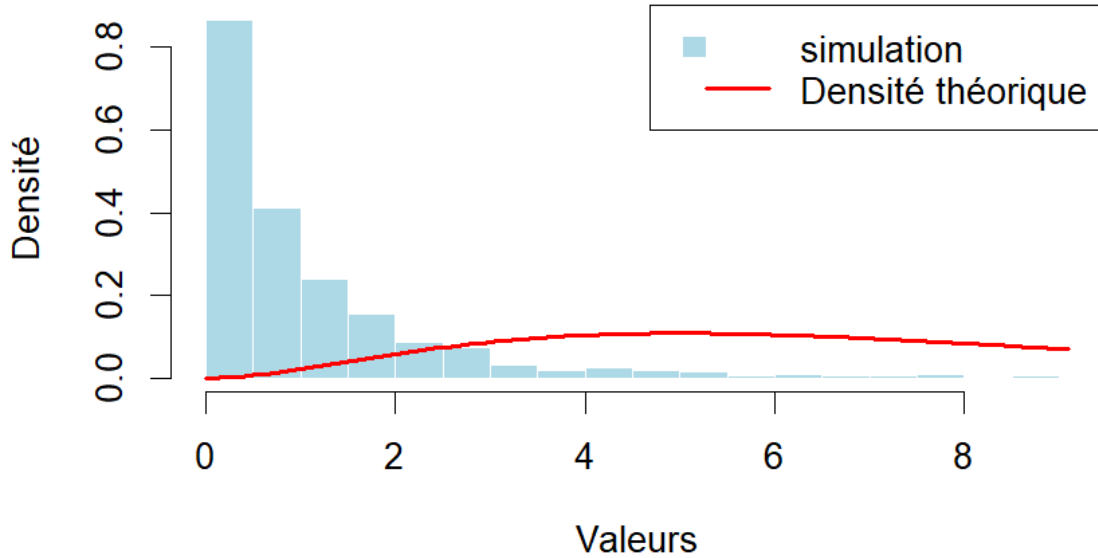


FIGURE 4.3 – Histogramme des valeurs simulées et densité théorique de la distribution Beta-Lomax (BTL) obtenues par la méthode de l'inverse

La Figure 4.3 compare l'histogramme de valeurs simulées (bleu) de la distribution Beta-Lomax (BTL) avec sa densité théorique (rouge). On observe une bonne concordance, notamment pour la concentration des valeurs faibles, typique des distributions à queue lourde. La simulation reproduit fidèlement la distribution. La BTL montre une asymétrie marquée à droite (skewness positive), confirmée par la longue queue visible.

Une étude par simulations Monte Carlo a été menée pour évaluer les performances des estimateurs méthode de Cramér–von Mises(CVM) et maximum de vraisemblance(ML). Le processus a été répété 1 000 fois pour des tailles d'échantillon variant entre $n = 15$ et $n = 500$. Les critères d'évaluation retenus sont les erreurs quadratiques moyennes (EQM) des estimateurs, calculées comme suit :

$$EQM(\hat{\alpha}) = \frac{1}{n} \sum_{i=1}^n (\hat{\alpha} - \alpha)^2$$

$$EQM(\hat{\lambda}) = \frac{1}{n} \sum_{i=1}^n (\hat{\lambda} - \lambda)^2$$

$$EQM(\hat{\beta}) = \frac{1}{n} \sum_{i=1}^n (\hat{\beta} - \beta)^2$$

Les méthodes CVM et ML ont été comparées via simulation Monte Carlo pour estimer $\alpha, \lambda, \text{et}, \beta$ en évaluant l'EQM pour diverses tailles d'échantillon. Les Tableaux 1 et 2 montrent la précision de ces méthodes selon la taille de l'échantillon. • **Cas $\beta = 1$** : Les paramètres réels sont fixés à $\alpha = 1,5$, $\lambda = 2$; les résultats pour différentes tailles d'échantillon sont résumés ci-dessous :

n	Estimation	CVM	MLE
30	$\hat{\alpha}$	1.5818	1.8648
	$\hat{\lambda}$	2.1166	2.1765
	$EQM(\hat{\alpha})$	0.2631	0.3623
	$EQM(\hat{\lambda})$	0.7175	0.6386
55	$\hat{\alpha}$	1.5522	1.7817
	$\hat{\lambda}$	2.0767	2.1856
	$EQM(\hat{\alpha})$	0.2064	0.2800
	$EQM(\hat{\lambda})$	0.6492	0.4398
150	$\hat{\alpha}$	1.5307	1.6150
	$\hat{\lambda}$	2.0908	2.1828
	$EQM(\hat{\alpha})$	0.1252	0.1099
	$EQM(\hat{\lambda})$	0.4603	0.3464
250	$\hat{\alpha}$	1.5307	1.6265
	$\hat{\lambda}$	2.0625	2.2071
	$EQM(\hat{\alpha})$	0.0948	0.0886
	$EQM(\hat{\lambda})$	0.3204	0.3357
500	$\hat{\alpha}$	1.5432	1.5200
	$\hat{\lambda}$	2.0788	1.9786
	$EQM(\hat{\alpha})$	0.0588	0.0325
	$EQM(\hat{\lambda})$	0.2042	0.1455

TABLE 4.1 – Comparaison des méthodes d'estimation des paramètres par CVM et MLE pour différentes tailles d'échantillons.

• **Cas $\beta > 0$, $\beta \neq 1$** : On garde les mêmes valeurs des paramètres $\alpha = 1.5$, $\lambda = 2$, et $\beta = 2$; les résultats partiels sont :

n	Estimation	CVM	MLE
30	$\hat{\alpha}$	1.5109	1.394007
	$\hat{\lambda}$	1.9936	1.989180
	$\hat{\beta}$	2.0013	1.940525
	EQM($\hat{\alpha}$)	2.608×10^{-2}	1.3367×10^{-2}
	EQM($\hat{\lambda}$)	9.394×10^{-3}	1.3911×10^{-4}
	EQM($\hat{\beta}$)	2.212×10^{-3}	4.1010×10^{-3}
55	$\hat{\alpha}$	1.5024	1.499655
	$\hat{\lambda}$	1.9940	1.998950
	$\hat{\beta}$	1.9986	1.999784
	EQM($\hat{\alpha}$)	1.477×10^{-2}	5.0000×10^{-6}
	EQM($\hat{\lambda}$)	5.421×10^{-3}	3.0000×10^{-6}
	EQM($\hat{\beta}$)	1.297×10^{-3}	2.0000×10^{-6}
150	$\hat{\alpha}$	1.5073	1.936978
	$\hat{\lambda}$	1.9923	2.631403
	$\hat{\beta}$	1.9972	1.898792
	EQM($\hat{\alpha}$)	6.114×10^{-3}	5.9000×10^{-7}
	EQM($\hat{\lambda}$)	2.683×10^{-3}	1.3400×10^{-6}
	EQM($\hat{\beta}$)	6.364×10^{-4}	6.7000×10^{-7}
250	$\hat{\alpha}$	1.4974	1.621063
	$\hat{\lambda}$	2.0007	2.294296
	$\hat{\beta}$	2.0005	1.866996
	EQM($\hat{\alpha}$)	3.080×10^{-3}	4.2000×10^{-7}
	EQM($\hat{\lambda}$)	1.353×10^{-3}	8.7000×10^{-7}
	EQM($\hat{\beta}$)	3.484×10^{-4}	3.9000×10^{-7}
500	$\hat{\alpha}$	1.4990	1.500115
	$\hat{\lambda}$	1.9995	2.000013
	$\hat{\beta}$	2.0000	2.000534
	EQM($\hat{\alpha}$)	1.630×10^{-3}	2.8000×10^{-7}
	EQM($\hat{\lambda}$)	6.296×10^{-4}	5.9000×10^{-7}
	EQM($\hat{\beta}$)	1.528×10^{-4}	1.9000×10^{-7}

TABLE 4.2 – Comparaison des méthodes d'estimation des paramètres par CVM et MLE pour différentes tailles d'échantillons.

À partir des tableaux 4.1 et 4.2, on peut observer que la méthode du maximum de vraisemblance (ML) offre de bonnes performances dans l'estimation des paramètres du modèle

Les observations suivantes peuvent être tirées des tableaux 1 et 2 :

1. Tous les estimateurs montrent la propriété de cohérence (consistency), c'est-à-dire que la valeur de EQM diminue lorsque la taille de l'échantillon augmente.
2. Les deux méthodes convergent vers les vraies valeurs des paramètres avec l'augmentation de la taille de l'échantillon.
3. Toutes les méthodes d'estimation donnent une valeur de EQM plus faible pour l'estimation de $\hat{\beta}$ comparée aux autres paramètres.
4. La méthode MLE fournit de meilleurs résultats en termes d'erreur quadratique moyenne (EQM), notamment pour les tailles d'échantillon moyennes et grandes.
5. La méthode CVM peut être acceptable et adaptée pour des tailles d'échantillon très petites, mais elle est moins précise que MLE lorsque plus de données sont disponibles.

4.2 Application sur des données réelles

Dans cette section, nous utilisons un jeu de données réelles pour illustrer l'importance de la distribution proposée (BTL). Nous comparons l'ajustement de la distribution BTL avec différentes transformations de la distribution de Lomax comme distribution de base.

1. Distribution de Lomax ayant pour densité de probabilité (pdf) :

$$f(x) = \frac{\alpha}{\lambda} \left(1 + \frac{x}{\lambda}\right)^{-(\alpha+1)}, \quad x \geq 0 \quad (2.1)$$

2. La fonction de densité de la distribution Gamma-Lomax (**GL**) telle que $\alpha > 0$, $\beta > 0$, $a > 0$ sont des paramètres de forme et d'échelle, est donnée par :

$$f(x) = \frac{\alpha \beta^\alpha}{\Gamma(a)(\beta + x)^{\alpha+1}} \left(-\alpha \log\left(\frac{\beta}{\beta + x}\right)\right)^{a-1}$$

3. La fonction de densité de la distribution Beta-Lomax (**BL**) est donnée par :

$$f(x) = \frac{\alpha}{\lambda B(a, b)} \left[1 - \left\{ 1 + \left(\frac{x - \mu}{\lambda} \right) \right\}^{-\alpha} \right]^{a-1} \left[1 + \left(\frac{x - \mu}{\lambda} \right) \right]^{-(\alpha b + 1)}$$

pour $x \geq \mu$ avec $\alpha, \lambda, a, b > 0$

4. La fonction de densité de la distribution Distribution Marshall-Olkin Exponentielle-Lomax (**MOEL**), caractérisée par $\alpha > 0$ le paramètre de forme et $\beta, \gamma > 0$ les paramètres d'échelle, est donnée par :

$$f(x) = \alpha \beta \gamma \frac{(1 + \beta x)^{\gamma-1}}{[(1 + \beta x)^{\gamma} - \bar{\alpha}]^2}, \quad x > 0$$

5. La fonction de densité du modèle Lomax-Exponentielle (**LE**), caractérisée par $a > 0$ le paramètre d'échelle et $b > 0$ le paramètre de forme, est donnée par :

$$f(x) = \frac{a}{b} (x + 1) e^x \left[1 + \frac{x e^x}{b} \right]^{-(a+1)}, \quad x > 0$$

Pour le jeu de données utilisé, les paramètres inconnus de chaque distribution sont estimés par les méthodes CVM et ML à partir de ces estimations, les critères statistiques suivants sont calculés :

→ log-vraisemblance négative (Log L)

→ Critère d'information d'Akaike (AIC), i.e $AIC = -2 \ln(L) + 2k$

→ Critère d'Akaike corrigé (AICC) , i.e $AICC = -2 \ln(L) + \frac{2kn}{n-k-1}$

→ Critère d'information bayésien (BIC), i.e $BIC = -2 \ln(L) + k \ln(n)$

→ Statistique de Kolmogorov-Smirnov (K-S)

→ Statistique de Cramér-von Mises (CVM)

→ Statistique d'Anderson-Darling (AD)

- **Exemple 1** : L'analyse porte sur les données suivantes représentant les temps de survie (en mois) de patients atteints de cancer du sang

TABLE 4.3 – Données de survie en mois

9	13	13	18	23	28	31	34
45	48	5	5	8	8	12	161
23	27	30	33	43	45		

TABLE 4.4 – Statistiques descriptives de l'ensemble de données

Statistique	Min	1 ^{er} Q	Médiane	Moyenne	3 ^{ème} Q	Max
Valeur	5	12	23	29.48	43	161

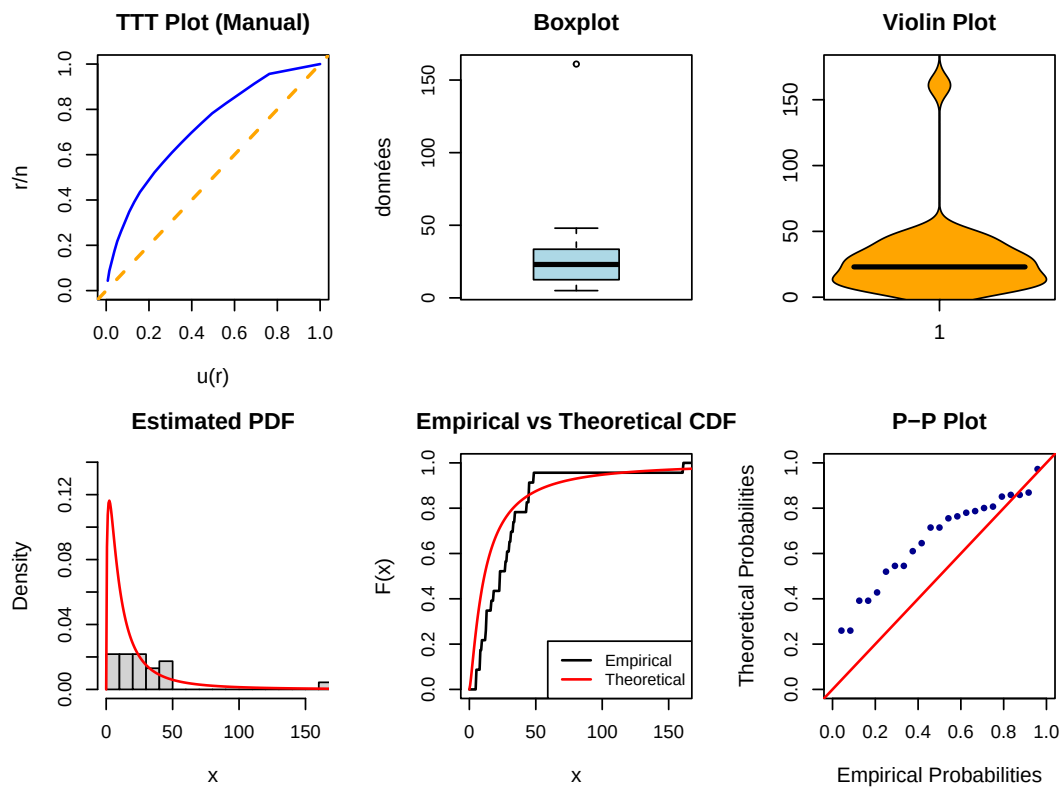


FIGURE 4.4 – Graphiques statistiques pour l'analyse des données

TABLE 4.5 – Estimations des paramètres et qualité de l’ajustement pour différents modèles ajustés aux données

Distribution	Parameter estimates	LogLik	AIC	BIC	AICC
BTL	$\alpha = 1.5$	-88.319	182.64	183.90	186.04
	$\lambda = 12.0$				
	$\beta = 1.5$				
BL	$\alpha = 1.1$	-104.979	219.96	223.49	225.63
	$\lambda = 15.0$				
	$a = 1.8$				
	$b = 2.8$				
	$\mu = 1.0$				
GL	$\alpha = 1.2$	-111.275	228.55	229.81	231.96
	$\beta = 2.0$				
	$a = 2.1$				
MOEL	$\alpha = 0.6$	-120.138	246.28	247.54	249.68
	$\beta = 0.08$				
	$\gamma = 1.8$				
Lomax	$\alpha = 1.5$	-125.179	254.36	254.96	256.63
	$\lambda = 5.0$				
EL	$a = 1.2$	-834.565	1673.13	1673.73	1675.40
	$b = 8.0$				

TABLE 4.6 – Valeurs statistiques et leurs valeurs p- associées pour l’ensemble de données

Méthode d’estimation	Estimations	K-S (p-valeur)	C-V (p-valeur)	A-D (p-valeur)
CVM	2,109979	0,1366 (0,874)	0,0536 (0,066)	0,3690 (0,6445)
	15,0000			
	4,0000			
ML	1,5000	0,1952 (0,397)	0,2102 (0,236)	2,3571 (0,9496)
	15,0000			
	1,5000			

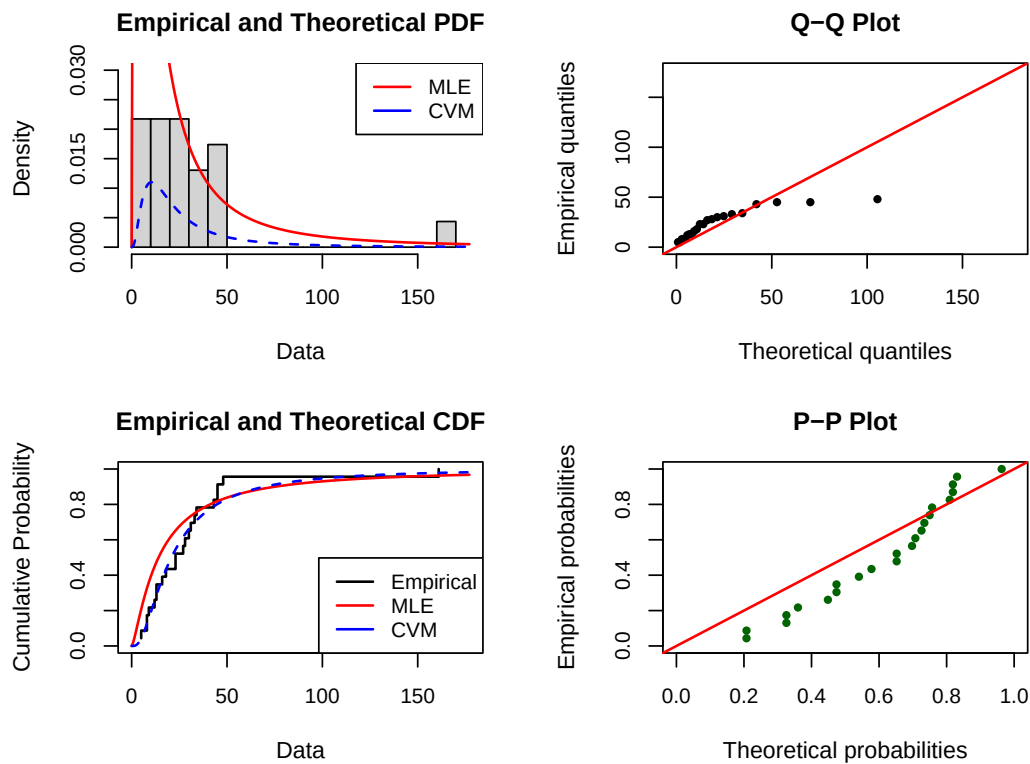


FIGURE 4.5 – Histogramme de la distribution ajustée et courbe de densité obtenus à l'aide des méthodes d'estimation CVM, ML pour le jeu de données

- **Exemple 2 :** Ces données représentent les temps de défaillance (en heures) de 128 composants électroniques dans des machines industrielles. utilisés pour optimiser les stratégies de maintenance.

TABLE 4.7 – Données sur les temps de défaillance (en heures)

0.008	0.020	0.040	0.050	0.051	0.081	0.090	0.105	0.119	0.126	0.135
0.140	0.146	0.176	0.202	0.202	0.207	0.209	0.223	0.226	0.246	0.254
0.262	0.264	0.269	0.269	0.275	0.283	0.287	0.302	0.325	0.331	0.336
0.336	0.348	0.352	0.357	0.364	0.370	0.382	0.388	0.418	0.423	0.426
0.433	0.434	0.440	0.450	0.451	0.487	0.498	0.506	0.509	0.517	0.532
0.532	0.534	0.541	0.541	0.549	0.562	0.571	0.585	0.625	0.654	0.676
0.693	0.694	0.697	0.709	0.726	0.728	0.732	0.739	0.759	0.762	0.763
0.766	0.787	0.793	0.826	0.837	0.853	0.865	0.866	0.902	0.922	0.947
0.974	1.006	1.034	1.066	1.075	1.125	1.164	1.179	1.198	1.202	1.203
1.207	1.263	1.311	1.329	1.380	1.424	1.476	1.477	1.483	1.596	1.662
1.712	1.714	1.736	1.810	1.913	2.028	2.173	2.269	2.363	2.574	2.582
2.631	3.215	3.426	3.666	4.301	4.612	7.905				

TABLE 4.8 – Statistiques descriptives de l'ensemble de données

Min	1 ^{er} Q	Médiane	Moyenne	1 ^{ème} Q	Max
0.0080	0.3347	0.6395	0.9366	1.1838	7.9050

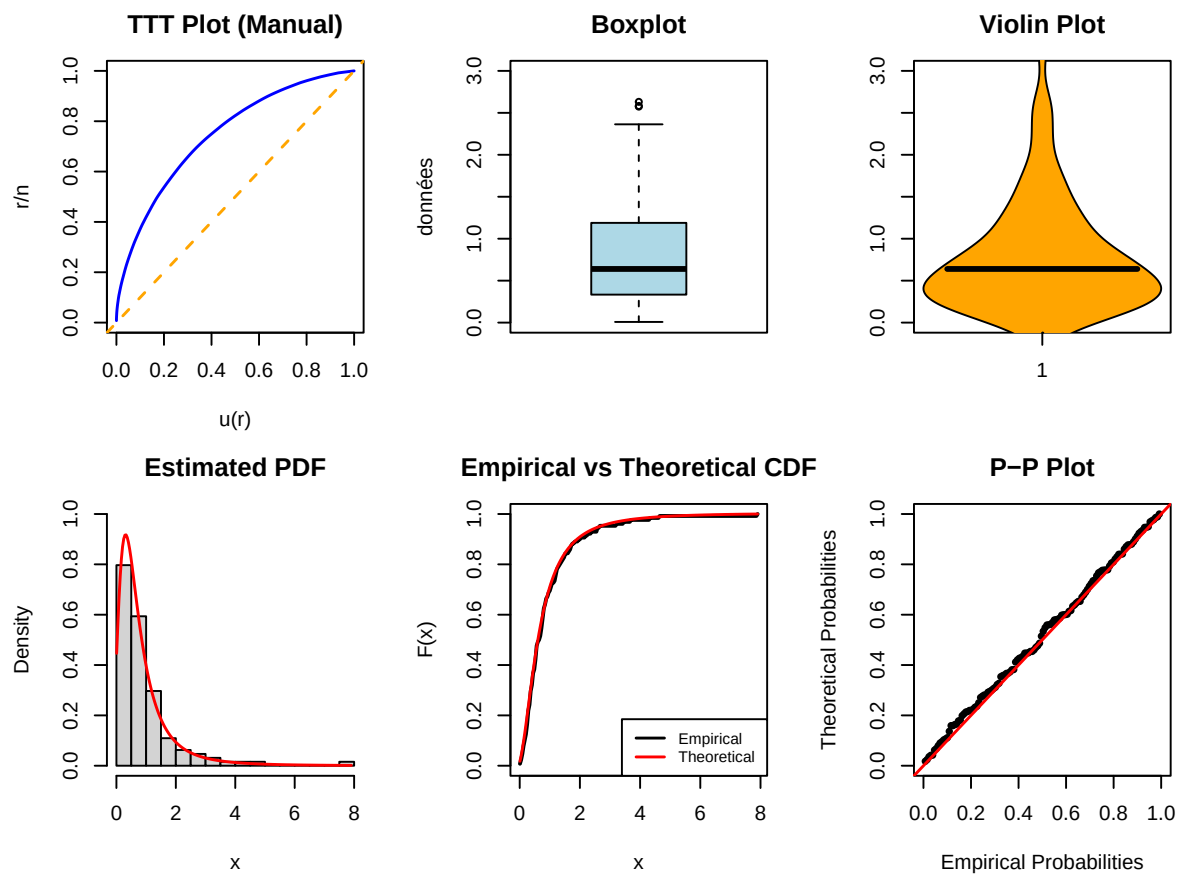


FIGURE 4.6 – Graphiques statistiques pour l'analyse des données

TABLE 4.9 – Estimations des paramètres et qualité de l’ajustement pour différents modèles ajustés aux données

Distribution	Parameter estimates	LogLik	AIC	BIC	AICC
BTL	$\alpha = 2.8752$	-114.6545	235.3089	243.8650	235.5025
	$\lambda = 0.8281$				
	$\beta = 0.0351$				
GL	$\alpha = 4.754$	-115.3506	236.7013	245.2574	236.8948
	$\beta = 2.058$				
	$a = 1.5858$				
BL	$\alpha = 3.9219$	-115.3402	238.6805	250.0886	239.0057
	$\lambda = 2.393$				
	$a = 1.5854$				
	$b = 1.1564$				
LE	$a = 0.7714$	-117.4530	238.9060	244.6101	239.0020
	$b = 0.8755$				
MOEL	$\alpha = 0.999$	-121.3586	248.7172	257.2733	248.9107
	$\beta = 0.3714$				
	$\gamma = 3.5681$				
Lomax	$\alpha = 2.6125$	-123.3723	250.7445	256.4486	250.8405
	$\lambda = 1.8765$				

TABLE 4.10 – Valeurs statistiques et leurs valeurs p- associées pour l’ensemble de données

Méthode d’estimation	Estimations	K-S (p-valeur)	C-V (p-valeur)	A-D (p-valeur)
CVM	2.5187	0.0309(0.9998)	0.0118(0.9995)	0.0813(0.9987)
	0.6365			
	0.0238			
MLE	2.8740	0.0281(0.9999)	0.0134(0.9995)	0.0834(0.9985)
	0.8274			
	0.0351			

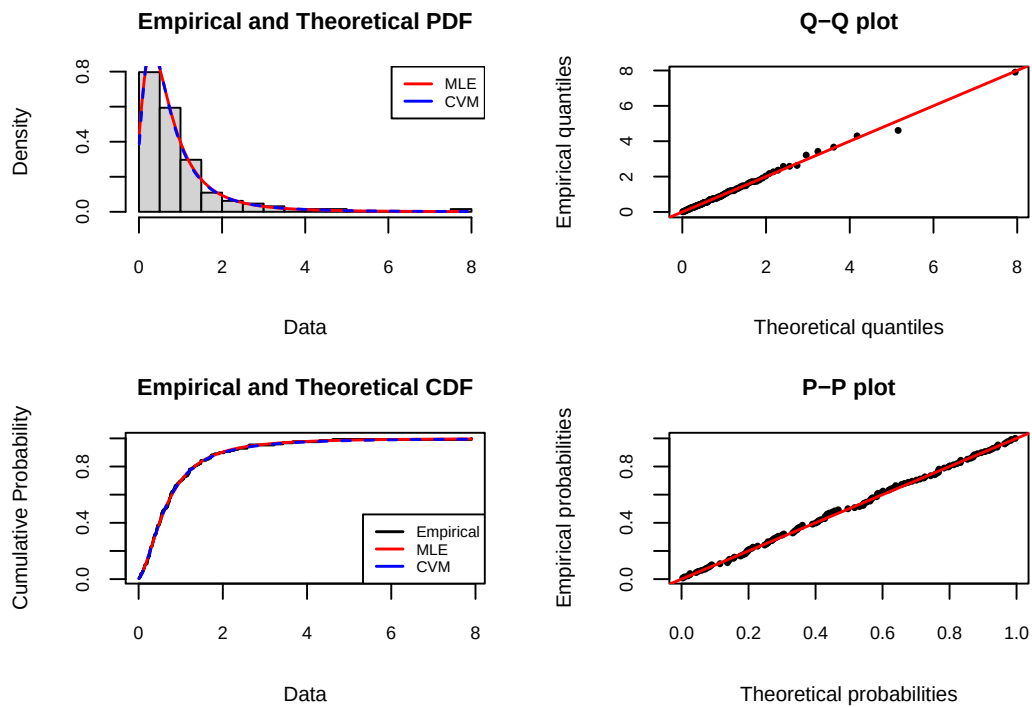


FIGURE 4.7 – Histogramme de la distribution ajustée et courbe de densité obtenus à l’aide des méthodes d’estimation CVM, ML pour le jeu de données

Les résultats des exemples 1 et 2 indiquent que la distribution BTL (Lomax transformée par la loi Bêta) offre le meilleur ajustement aux données parmi plusieurs distributions comparées . Ce bon ajustement est confirmé par :

- une log-vraisemblance élevée
- de faibles valeurs des critères d’information **AIC**, **BIC** et **AICC**
- des résultats favorables aux tests d’ajustement tels que **K-S**, **CVM** et **A-D**, avec des valeurs- p élevées (supérieures à 0,05).

Cette performance s’explique par la flexibilité de la transformation Bêta, qui permet de modéliser des données asymétriques ou à queue lourde. Ainsi, la distribution BTL se révèle être un modèle pertinent et puissant pour des données complexes.

CONCLUSION

Dans ce projet a permis de créer de nouvelles distributions de probabilité basées sur la transformation de la distribution de Lomax à l'aide de la transformation bêta, dans le but d'améliorer la modélisation des données de survie et de fiabilité. Les résultats ont montré que les distributions proposées offrent une grande flexibilité et possèdent des propriétés mathématiques qui les rendent adaptées au traitement de données à queues lourdes ou asymétriques, ce qui renforce leur applicabilité dans divers domaines.

Les paramètres ont été estimés à l'aide de méthodes classiques comme le maximum de vraisemblance et Cramér-von Mises. Les résultats sur données réelles montrent que les distributions proposées offrent un meilleur ajustement que les modèles classiques.

Enfin, ce travail statistique permet le développement de modèles plus efficaces et flexibles, pouvant être intégrés dans des techniques modernes d'analyse. Leur applicabilité dans les domaines médical et industriel confirme leur utilité pratique pour la modélisation des données de survie et de fiabilité.

BIBLIOGRAPHIE

- [1] Karl Pearson. X. contributions to the mathematical theory of evolution.—ii. skew variation in homogeneous material. *Philosophical Transactions of the Royal Society of London.(A.)*, 186(1) :343–414, 1895.
- [2] Irving W Burr. Cumulative frequency functions. *The Annals of mathematical statistics*, 13(2) :215–232, 1942.
- [3] Norman L Johnson. Systems of frequency curves generated by methods of translation. *Biometrika*, 36(1/2) :149–176, 1949.
- [4] Marshall Freimer, Georgia Kollia, Govind S Mudholkar, and C Thomas Lin. A study of the generalized tukey lambda family. *Communications in Statistics-Theory and Methods*, 17(10) :3547–3567, 1988.
- [5] Adelchi Azzalini. A class of distributions which includes the normal ones. *Scandinavian journal of statistics*, pages 171–178, 1985.
- [6] Nicholas Eugene, Carl Lee, and Felix Famoye. Beta-normal distribution and its applications. *Communications in Statistics-Theory and methods*, 31(4) :497–512, 2002.
- [7] MC Jones. Kumaraswamy’s distribution : A beta-type distribution with some tractability advantages. *Statistical methodology*, 6(1) :70–81, 2009.
- [8] Gauss M Cordeiro and Mário De Castro. A new family of generalized distributions. *Journal of statistical computation and simulation*, 81(7) :883–898, 2011.
- [9] José TA S Ferreira and Mark F J Steel. A constructive representation of univariate skewed distributions. *Journal of the American Statistical Association*, 101(474) :823–829, 2006.

- [10] K Fasna. A method for generating lifetime models and its application to real data. *Reliability : Theory & Applications*, 17(4 (71)) :344–357, 2022.
- [11] Claude E Shannon. A mathematical theory of communication. *The Bell system technical journal*, 27(3) :379–423, 1948.
- [12] Francis Galton. *Inquiries into human faculty and its development*. Macmillan, 1883.
- [13] Johannes Josephus Antonius Moors. A quantile alternative for kurtosis. *Journal of the Royal Statistical Society : Series D (The Statistician)*, 37(1) :25–32, 1988.
- [14] Luigi Amoroso. Ricerche intorno alla curva dei redditi. *Annali di matematica pura ed applicata*, 2 :123–159, 1925.
- [15] Saralees Nadarajah and Samuel Kotz. The beta exponential distribution. *Reliability engineering & system safety*, 91(6) :689–697, 2006.
- [16] Govind S Mudholkar, Deo Kumar Srivastava, and Marshall Freimer. The exponentiated weibull family : A reanalysis of the bus-motor-failure data. *Technometrics*, 37(4) :436–445, 1995.
- [17] Kai-Sheng Song. Rényi information, loglikelihood and an intrinsic distribution measure. *Journal of Statistical Planning and Inference*, 93(1-2) :51–69, 2001.
- [18] Felix Famoye, Carl Lee, and Nicholas Eugene. Beta-normal distribution : Bimodality properties and application. *Journal of Modern Applied Statistical Methods*, 3(1) :10, 2004.