
الجمهورية الجزائرية الديمقراطية الشعبية

Ministère de L'Enseignement Supérieur et de la Recherche
Scientifique

UNIVERSITE SAAD DAHLEB DE BLIDA

Faculté des sciences

Département de Mathématiques



MEMOIRE DE MASTER

En Mathématiques

Option : Modèles Stochastiques et Statistique

THÈME :

Caractérisation de la distribution de Lindley et applications dans
l'analyse de survie

Réalisé par

AKABA Widad & OUALI Asma

Soutenu devant le Jury :

TAMI Omar	Université Blida 1	Président
FRIHI Redhouane	Université Blida 1	Examineur
RASSOUL Abdelaziz	ENSH de Blida	Promoteur

Juillet 2022

DÉDICACES

Je dédie ce modeste mémoire :

A mes très chers parents qui m'ont entouré d'amour, m'encourager et qui m'ont guidé durant les moments les plus pénibles de ce long chemin, ma mère qui a été à mes cotés et m'a soutenue durant toute ma vie et mon père qui a sacrifié toute sa vie afin de me voir devenir ce que je suis, merci mes parents et que Dieu prolonge leur vie inchallah

A mes frères :Mohamed, Redouane, Abderraouf et Billel

A mes soeurs : Hadjer et Mariem

A mon fiancé Hamza

A mon grand père et ma grande mère que Dieu les protège et prolonge leur vie inchallah

A toute ma grande famille

A mon Binôme de mémoire Widad

A mes très chères amies

A tous mes collègues de ma promotion

A ceux qui m'ont toujours encouragé pour que je réussisse dans mes études et durant la réalisation de mon mémoire

Finalement à tous ceux qui m'ont aidé de près ou de loin.

OUALI.Asma

DÉDICACES

Je dédie ce modeste mémoire à :

Mes chères parents pour leur soutien et leur encouragement durant toute ma carrière d'études

Ma grande mère et ma seconde mère que Dieu ait pitié d'elle, pour tout le bien que vous m'avez fait, et si dieu le veut, ce sera dans la balance de ses bonnes actions

Mon grand père, que j'aime tant, que Dieu le protège et prenne soin de lui

Mes soeurs et mes frères

Mon binôme ASMA

Mes meilleurs amis

Ainsi à tous ceux qui me connaissent

AKABA.Widad

REMERCIÉMENTS

En tout premier lieu, nous tenons à remercier ALLAH, le tout puissant et miséricordieux, qui m'a donné la volonté, la force et la santé de réaliser ce mémoire.

Nous exprimons notre profonde gratitude et nos sincères remerciements à notre encadreur le professeur RASSOUL Abdelaziz, d'avoir accepté de diriger ce travail, ses conseils, ses corrections, sa gentillesse et sa patience ainsi pour le temps qu'il a consacré à la réalisation de ce mémoire.

Nos sincères remerciements aux membres du jury pour leur présence, pour leur lecture attentive de ce mémoire, ainsi que pour les remarques qu'ils nous ont faites lors de cette soutenance afin d'améliorer notre travail. Ces membres sont : Monsieur TAMI Omar, qui nous a fait l'honneur de présider le jury de soutenance de notre mémoire, et le Monsieur FRIHI Redouane qui examine notre mémoire.

Notre reconnaissance va à nos enseignants qui n'ont pas cessé de donner le mieux tout au long de notre formation.

Enfin, nous tenons à remercier toutes les personnes qui ont aidé de près ou de loin à l'élaboration ce mémoire.

TABLE DES MATIÈRES

Introduction Générale	1
1 Rappels sur les lois de probabilités	3
1.1 Généralités	3
1.2 Variables aléatoires	4
1.3 Variable aléatoire réelle	4
1.4 Variable aléatoire discrète	4
1.4.1 Loi uniforme sur $[1;n]$	5
1.4.2 Loi de Bernoulli	6
1.4.3 Loi Binomiale $\mathcal{B}(n,p)$	6
1.4.4 Loi hypergéométrique	7
1.4.5 Loi géométrique	7
1.4.6 Loi de Poisson	8
1.5 Variables aléatoires continues	9
1.5.1 Loi uniforme	10
1.5.2 Loi exponentielle	10
1.5.3 Loi normale	11
1.5.4 Loi de Gamma	11
1.5.5 Loi de Student	12
1.5.6 Loi du Khi-Deux	13
1.5.7 Loi de Fisher-Snedecor	13
1.6 Fonction de répartition	14
1.7 Fonction caractéristique	14
1.8 Fonction génératrice des moments	15
2 Loi de Lindley	16
2.1 Caractérisation de la distribution de lindley	16

2.1.1	Moments et mesures connexes	17
2.1.2	Fonction de hasard et fonction de survie	20
2.1.3	Ordre stochastique	21
2.1.4	Courbe de Lorenz	21
2.1.5	Statistiques d'ordre extrêmes	22
2.1.6	Fonction Quantile de la distribution Lindley	23
2.1.7	Estimation	24
2.2	Distribution de Lindley à deux paramètres	26
2.2.1	Moments et mesures associées	27
2.2.2	Taux d'échec et durée de vie résiduelle moyenne	28
2.2.3	Ordres stochastiques	29
2.2.4	Estimation des paramètres	30
2.3	Distribution de Lindley Étendue	33
2.3.1	Mesures statistiques de fiabilité	35
2.3.1.1	Fonctions de densité de survie et de taux d'échec	35
2.3.1.2	Quantiles, moments et statistiques d'ordre	36
2.3.1.3	Fonction de durée de vie résiduelle et de taux de dé- faillance inversé	37
2.3.1.4	Courbes de Bonferroni et de Lorenz	39
2.3.1.5	Écartés moyens	41
2.3.2	Estimation et inférence	41
2.4	Distribution de Lindley Généralisée	43
2.4.1	Estimation par la méthode des moments	44
2.4.2	Estimations du maximum de vraisemblance	45
3	Simulations et Applications	47
3.1	Simulation des v. a. de Lindley	47
3.2	Applications sur des données réelles	49
	Conclusion Générale	56

TABLE DES FIGURES

2.1	Présentation graphique de la fonction de densité pour quelques valeurs de θ , noir($\theta=0.5$);rouge($\theta=1$);bleu($\theta=2$)	17
2.2	Représentation graphique du coefficient de variation γ (noir), le coefficient de dissymétrie $\sqrt{\beta}$ (rouge), et le coefficient d'aplatissement β_2 (bleu)	19
2.3	Présentation graphique de la fonction de taux de hasard pour quelques valeurs de θ , noir ($\theta = 0.5$); rouge ($\theta = 1$); bleu ($\theta = 2$).	20
2.4	Présentation graphique de la fonction de survie pour quelques valeurs de θ , noir($\theta=0.5$); rouge ($\theta=1$); bleu($\theta=2$)	21
3.1	L'histogramme, la fonction de répartition, le quantile et la fonction de survie de loi lindley a une seul paramètre	48
3.2	L'histogramme, la fonction de répartition, le quantile et la fonction de survie de loi lindley a deux paramètres	49
3.3	Graphique des paramètres descriptif des données des durées survie de cobays infectés par des bacilles tuberculeux virulents	50
3.4	Histogramme des paramètres descriptif des données des durées survie de cobays infectés par des bacilles tuberculeux virulents	50
3.5	Densités estimées, Densités cumulées estimées, Graphiques p-p pour Lindley à un paramètre ajusté pour l'ensemble de données 1	51
3.6	Densités estimées, Densités cumulées estimées, Graphiques p-p, Graphiques Q-Q pour Lindley à deux paramètre ajusté pour l'ensemble de données 1	51
3.7	Densités estimées, Densités cumulées estimées, Graphiques p-p, Graphiques Q-Q pour Lindley Généralisé ajusté pour l'ensemble de données 1	52
3.8	Densités estimées, Densités cumulées estimées, Graphiques p-p pour Lindley à un paramètre ajusté pour l'ensemble de données 2	54

3.9 Densités estimées des modèles pour l'ensemble de données 2 54

LISTE DES TABLEAUX

3.1	Les estimations des données de durée de survie (en jours) de 72 cobayes infectés par des bacilles tuberculeux virulents	50
3.2	AIC,BIC,K-S,C-V et A-D des modèles basés sur le jeu de données 1 . . .	53
3.3	L'ajustement de la distribution 'Lindley1 , Lindley2'et 'GenLindley par des paramètres de maximum de vraisemblance d'ensemble des données 1	53
3.4	Les temps de rémission (en mois) des patients atteints de cancer de la vessie	53
3.5	Les estimations des données des temps de rémission (en mois) des patients atteints de cancer de la vessie	54
3.6	AIC,BIC,K-S,C-V et A-D des modèles basés sur le jeu des données 2 . .	55
3.7	Ajustement de la distribution 'Lindley1 , Lindley2' par des paramètres de maximum de vraisemblance d'ensemble des données 2	55

ملخص

-يتم تطبيق التوزيعات الاحتمالية بشكل شائع لوصف الظواهر الحقيقية وغالبا ما تستخدم في العديد من المجالات مثل الاقتصاد وعلم الاحياء والطب. هناك العديد من التوزيعات الاحتمالية في هذه المذكرة نميز فيها توزيع لندلي ذو المعلم الواحد ولندلي ذو المعلمين وكذلك توزيع لندلي بثلاث معاملات. بالإضافة الي دراسة خصائص التوزيعات مثل العزوم الدوال المولدة والمميزة و التقدير بطريقة العزوم وطريقة الارجحية (او المعقولة) العظمى. في الأخير تطبيق على بعض البيانات حقيقية التمييز بين التوزيعات.

Résumé :

Les distributions de probabilité sont couramment appliquées pour décrire des phénomènes réels et sont souvent utilisées dans plusieurs domaines tels que l'économie, la biologie et la médecine. Il existe plusieurs distributions de probabilité. Dans ce mémoire nous distinguons la distribution de Lindley à un paramètre et la distribution de Lindley à deux paramètres, ainsi que la distribution de Lindley à trois paramètres en plus d'étudier les caractéristiques des distributions telles que les moments, les fonction génératrices et caractéristiques, l'estimation par la méthode des moments et par la méthode du maximum de vraisemblance. dans la dernière application sur certaines données réelles pour distinguer les distributions.

Abstract :

Probability distributions are commonly applied to describe real phenomena and are often used in several fields such as economics, biology and medicine. There are several probability distributions. In this thesis we distinguish between the Lindley distribution with one parameter and the Lindley distribution with two parameters, as well as the Lindley distribution with three parameters in addition to studying the characteristics of the distributions such as the moments, the generating and characteristic functions, the estimation by the method of moments and by the method of maximum likelihood. in the last application on certain real data to distinguish the distributions.

INTRODUCTION GÉNÉRALE

Dans de nombreuses sciences appliquées telles que la médecine, l'ingénierie et la finance, entre autres, la modélisation et l'analyse de la durée de vie des données sont cruciales. Plusieurs distributions de durée de vie ont été utilisées pour modéliser ces types de données. Par exemple, l'exponentielle, Les distributions de Weibull, gamma, Rayleigh et leurs généralisations (voir, par exemple, Gupta & Kundu, 1999 et Nadarajah & Kotz), Chaque distribution a ses propres caractéristiques dues spécifiquement à la forme de la fonction de taux de défaillance qui peut n'être que monotone décroissante ou croissante ou constante dans son comportement, ainsi que non monotone, étant en forme de baignoire ou même unimodale.

Les distributions de probabilités sont utilisées dans plusieurs domaines de l'activité humaine comme l'économie, l'ingénierie et l'informatique. Dans la littérature, il existe beaucoup de distributions de probabilités différentes.

Parmi ces distributions, La loi de Lindley, proposée par le statisticien britannique Dennis Victor Lindley en 1958[1], est une distribution de probabilité continue. Les propriétés de cette distribution ont été étudiées pour la première fois par Ghitany et al 2008 [2].

Dans ce mémoire on parle de trois parties de la distribution de Lindley. Il est divisé en trois chapitres comme suit :

Chapitre 1 : Dans ce chapitre, nous rappelons certaines notions et certains résultats que nous utiliserons par la suite. Ce rappel comporte la notion d'un variable aléatoire, la notion de variable aléatoire discrete avec quelques exemples de la distribution de probabilités(Loi de Bernoulli, Binomiale, géométrique ...), La notion de variable aléatoire continue avec quelques distributions de probabilités (Loi exponentielle, loi de Gamma ...), Fonction de répartition, Fonction caractéristique et la notion de la fonction génératrice des moment.

Chapitre 2 : Dans ce chapitre, nous allons voir trois parties de la distribution de Lindley. D'abord Les propriétés de la loi de Lindley à un paramètre et ses propriétés sont : Moments et mesures connexes, Fonction de hasard et fonction de survie, Ordre stochastique, Courbe de Lorenz, Statistiques d'ordre extrêmes, Fonction quantile, Estimation du paramètre avec la méthode des moments et la méthode du maximum de vraisemblance. Et voir aussi la Distribution de Lindley à deux paramètres avec ces propriétés : Mesures associées, Taux d'échec et durée de vie résiduelle moyenne, Ordre stochastique et L'estimation des paramètres. Et enfin on voit La distribution de Lindley à trois paramètres.

Chapitre 3 : On va présenter dans ce chapitre la simulation des variables aléatoires de Lindley avec les trois catégories, un seul paramètre, deux paramètres, et trois paramètres. Enfin on va présenter des applications sur des jeux de données réelles et on va essayer d'ajuster ces données par des distributions de Lindley.

CHAPITRE 1

RAPPELS SUR LES LOIS DE PROBABILITÉS

1.1 Généralités

Une loi de probabilité décrit de manière théorique le caractère aléatoire d'une expérience qui est considérée comme aléatoire. la notion "d'expérience aléatoire" est dégagée pour désigner un processus réel de nature expérimentale, ou le hasard intervient avec des issues possibles bien identifiées.

Définition 1.1 Une loi de probabilité est une mesure dont la masse totale vaut 1, en particulier, cette mesure vérifie les trois axiomes des probabilités. Pour (Ω, \mathcal{A}) un espace mesurable, P est une loi de probabilité, mesure de probabilité ou plus simplement probabilité si :

- 1- P est une application de \mathcal{A} dans $[0, 1]$.

2- $P(\Omega) = 1$.

3- P est σ -additive, pour toute famille finie ou dénombrable d'éléments disjoints $(\mathcal{A}_i)_{i \in I}$ de Ω :

$$P(\cup_{i \in I} \mathcal{A}_i) = \sum_{i \in I}^n P(\mathcal{A}_i).$$

- Une conséquence immédiate est $P(\emptyset) = 0$
- Le triple (Ω, \mathcal{A}, P) est appelé espace probabilisé, une loi de probabilité est également appelé distribution de probabilité pour une étude plus appliquée
- Une manière usuelle d'expression d'une loi est l'utilisation d'une variable aléatoire puisque pour toute loi de probabilité P sur (Ω, \mathcal{A}) il existe une variable aléatoire X définie sur un espace probabilisable et de loi P .

1.2 Variables aléatoires

Soit (Ω, \mathcal{A}, P) un espace de probabilité, ou (Ω, \mathcal{A}) , est un espace mesurable et P est une mesure de probabilité sur \mathcal{A} .

Définition 1.2 Une variable aléatoire X est une fonction mesurable $X : (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}, \mathcal{B})$ où \mathcal{B} est la tribue borélienne de \mathbb{R} . On écrit parfois $X = X(\omega)$ pour souligner le fait qu'il s'agit d'une fonction de $\omega \in \Omega$.

Définition 1.3 La fonction de répartition d'une variable aléatoire X est la fonction $F : \mathbb{R} \rightarrow [0, 1]$ définie par

$$F_X(x) = P(X \leq x) = P(X(\omega) \leq x). \quad (1.1)$$

C'est une fonction monotone croissante, continue à droite et telle que $\lim_{x \rightarrow -\infty} F_X(x) = 0$ et $\lim_{x \rightarrow +\infty} F_X(x) = 1$.

La fonction F sera aussi appelé la loi (ou la distribution) de X .

Il existe deux principaux types de variables aléatoires : Les variables aléatoires discrètes et les variables aléatoires continues.

1.3 Variable aléatoire réelle

Définition 1.4 Une variables aléatoire réelle est une variable aléatoire à valeurs dans \mathbb{R} , ou une partie de \mathbb{R} , c'est une fonction définie depuis l'ensemble des résultats possibles d'une expérience aléatoire, dont on doit déterminer la probabilité qu'elle prenne une valeur donnée ou un ensemble donnée de valeurs, les variables aléatoires réelles sont les variable aléatoires les plus couramment étudiées, ce qui conduit les auteurs à omettre l'adjectif réel, et à parler de variable aléatoire tout court.

-Les variables aléatoires sont très utilisées en théorie des probabilités et en statistiques.

Dans les applications, les variables aléatoires sont utilisées pour modéliser le résultat d'un mécanisme non-déterministe ou encore comme le résultat d'une expérience non-déterministe qui génère un résultat aléatoire.

selon sa définition mathématique, la loi de X est la mesure de probabilité P_x définie par la relation

$$P_x = P(X \in A) = P(X^{-1}(A))$$

(jean-Y ves DAUXOIS, 2013)[3]

1.4 Variable aléatoire discrète

Définition 1.5 On dit d'une variable aléatoire est discrète si elle ne peut prendre que des valeurs numériques isolées.

Dans le cas contraire, on dit qu'elle est continue.

Une variable aléatoire discrète finie est donc une application

$\omega \rightarrow X$ ensemble $\{X_1, \dots, X_n\}$ fini des valeurs de F_x ce sera le cas chaque fois que ω est finie, par exemple la somme des numéros des faces lors du lancer de trois dés, le nombre d'enfants d'une famille

-Une variable aléatoire discrète dénombrable est une application de $\omega \in \mathbb{N}$ qui est finie, mais

formé des valeurs bien isolées les unes des autres .

(louis gacogne/Gérard Frugier,1990)[4]

caractéristiques

◆ L'espérance d'une loi discrète est comme suit :

$$E(X) = \sum_{i=1}^n x_i P(X = x_i). \quad (1.2)$$

◆ La variance d'une loi discrète est comme suit :

$$V(X) = E(X^2) - [E(X)]^2 = E((X - E(X))^2). \quad (1.3)$$

◆ La fonction caractéristique d'une variable aléatoire réelle X est la fonction à valeurs complexes définie sur \mathbb{R} par :

$$Q_x(t) = E[e^{itx}] = E[\cos(tx)] + iE[\sin(tx)]. \quad (1.4)$$

◆ La fonction génératrice :

$$G_x(t) = E[e^{tx}] = E[\cos(tx)] + E[\sin(tx)]. \quad (1.5)$$

Exemples

1.4.1 Loi uniforme sur $[1;n]$

X suit la loi uniforme sur $[1;n]$ si :

1. $X(\Omega) = [1;n]$.
2. quelque soit i appartient $[1;n]$,

$$P(X = i) = \frac{1}{n}.$$

On note alors : $X \rightsquigarrow \mathcal{U}([1; n])$.

Il s'agit simplement, comme son nom l'indique, d'une loi dont tous les poids de probabilité sont identiques.

-Son espérance est :

$$E(X) = \frac{n+1}{2}$$

-Sa variance est :

$$V(X) = \frac{n^2-1}{12}$$

1.4.2 Loi de Bernoulli

La variable aléatoire X définie sur un espace de probabilité (Ω, \mathcal{A}, P) , où Ω est l'ensemble fondamental des événements, \mathcal{A} la σ -algèbre des événements et P la probabilité définie sur l'espace probabilisable (Ω, \mathcal{A}) , est dite suivre une loi de Bernoulli de paramètre p si $X(\Omega) = \{0, 1\}$ ces deux valeurs sont dites aussi échec et succès. On a :

$$P(X = 0) = q, \quad P(X = 1) = p \quad \text{avec} \quad p + q = 1$$

-Son espérance est $E(X)=p$

-Sa variance est $V(X)=pq$ est appelée loi de Bernoulli notée $X \rightsquigarrow \mathcal{B}(1;p)$.

1.4.3 Loi Binomiale $\mathcal{B}(n, p)$

Soit la répétition de n épreuves indépendantes de BERNOULLI de paramètre p . la loi de la variable aléatoire X égale au nombre de succès dans la répétition de ces n épreuves est dite Binomiale ;

On a :

$$P(X = k) = C_n^k p^k (1-p)^{n-k}, k = 0, 1, 2, \dots, n. \quad (1.6)$$

-Son espérance est $E(X) = np$

-Sa variance est $V(X) = npq$, avec $p + q = 1$

Loi de la somme de deux variables binomiales :

Soient deux variables aléatoires indépendantes X_1 et X_2 suivant une loi binomiale $\mathcal{B}(n_1, p)$ et $\mathcal{B}(n_2, p)$ respectivement.

La loi de $Z = X_1 + X_2$ est une loi binomiale $\mathcal{B}(n_1 + n_2, p)$.

1.4.4 Loi hypergéométrique

Soit une expérience aléatoire dont le processus est constitué d'un tirage sans remise de n objets parmi N , c-à-d n épreuves de Bernoulli non indépendantes, avec :

$$p = P(\text{succès lors de la 1ère vérification})$$

$$q = P(\text{échec lors de la 1ère vérification}) = 1 - p$$

Soit X une variable aléatoire discrète qui comptabilise le nombre de succès ou d'objets gagnants lorsqu'on tire simultanément n objets dans une urne contenant pN objets gagnants et qN objets perdants (avec $q = 1 - p$, soit un nombre total d'objets valant $pN + qN = N$). X suit une loi hypergéométrique de paramètres n , N et p , notée $X \sim H(n; N; p)$ avec :

$$p(x) = P(X = x) = \frac{C_x^{Np} \times C_{n-x}^{Nq}}{C_n^N}, \text{ avec } x = 0, 1, 2, 3, \dots, \min(n, Np) \quad (1.7)$$

On a :

-Son espérance :

$$E(X) = np$$

-Sa variance :

$$V(X) = npq \left(\frac{N-n}{N-1} \right)$$

1.4.5 Loi géométrique

-La loi géométrique est une loi de probabilité apparaissant dans de nombreuses applications.

-La loi géométrique de paramètre $P(0 < P < 1)$ correspond au modèle suivant :

On considère une épreuve de Bernoulli dont la probabilité de succès est p et celle d'échec $q = 1 - p$.

On renouvelle cette épreuve de manière indépendante jusqu'au premier succès. On appelle X la variable aléatoire donnant le rang du premier succès.

Les valeurs de X sont les entiers naturels non nuls $1, 2, 3, \dots$. La probabilité que $X = k$ est :

$$P(X = k) = q^{k-1} p, k = 1, 2, 3, \dots$$

On dit que X suit une loi géométrique de paramètre p .

-L'espérance d'une variable aléatoire X suivant une loi géométrique de paramètre P est :

$$E(X) = \frac{1}{p} .$$

et sa variance est :

$$V(X) = \frac{(1-p)}{p^2} = \frac{q}{p^2}$$

(ou $q=1-p$ est la probabilité d'échec).

1.4.6 Loi de Poisson

Soit $\lambda > 0$, X suit la loi de Poisson de paramètre λ si :

1- $X(\Omega) = \mathbb{N}$.

2-quelque soit $k \in \mathbb{N}$, $P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$, $k = 0, 1, 2, \dots$ (1.8)

On note alors : $X \rightsquigarrow \mathcal{P}(\lambda)$.

Remarque 1.1 On a facilement $\sum_{k=0}^{\infty} P(X = k) = 1$ (série exponentielle)

Contexte usuel :

c'est la «loi limite» d'une loi binomiale lorsque np reste constant et $n \rightarrow +\infty$.

Soit $\lambda = np$. On considère une v.a.r. X qui suit une loi binomiale $\mathcal{B}(n; \frac{\lambda}{n})$:

$\forall k \in [0; n]$,

$$\begin{aligned} P(X = k) &= \binom{n}{k} \times \left(\frac{\lambda}{n}\right)^k \times \left(1 - \frac{\lambda}{n}\right)^{n-k} \\ &= \frac{\lambda^k}{k!} \frac{n(n-1)\dots(n-k+1)}{n^k} \exp(n-k) \ln\left(1 - \frac{\lambda}{n}\right) \xrightarrow{n \rightarrow +\infty} \frac{\lambda^k}{k!} \exp(-\lambda) \end{aligned} \quad (1.9)$$

- La loi de poisson apparait donc lorsque l'on compte le nombre de succès au cours d'un nombre d'épreuves très élevé, avec une probabilité P du succès très faible : c'est la loi des «événements rares».
- On peut aussi remarquer que

$$\frac{P(X = n)}{P(X = n-1)} = \frac{\lambda}{n}$$

- La suite $(P(X = n))$ des poids de probabilité est donc croissante tant que $n \leq \lambda$, puis elle est décroissante. Les poids «élevés» sont ceux pour lesquels n est proche de λ , et ils sont ensuite rapidement très faibles, ce qui justifie aussi l'appellation «loi des évènements rares».
- L'espérance d'une variable aléatoire de poisson est $E(X) = \lambda$,
- La variance d'une variable aléatoire de poisson est $V(X) = \lambda$.
- Cette loi s'applique pour la détermination d'une probabilité de succès relativement à un évènement réparti dans le temps (par exemple loi du nombre d'appels téléphoniques pendant un intervalle de temps). λ est le nombre moyen de succès correspondant à l'intervalle de temps.

1.5 Variables aléatoires continues

Définition 1.6 Une variable aléatoire est une application de l'univers Ω dans \mathbb{R}

$$X : \begin{cases} \Omega \longrightarrow \mathbb{R} \\ \omega \mapsto X(\omega) \end{cases}$$

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

Caractéristiques

- ◆ L'espérance d'une loi continue est comme suit :

$$E(X) = \int_{-\infty}^{+\infty} xf(x)dx \quad (1.10)$$

- ◆ La variance d'une loi continue est comme suit :

$$V(X) = E(X^2) - [E(X)]^2 = [E(X - E(X))^2] \quad (1.11)$$

- ◆ Fonction caractéristique : si X est une variable aléatoire qui admet pour fonction de densité f_x , sa fonction caractéristique est définie par :

$$\phi_X(t) = E(e^{itx}) = \int_{-\infty}^{+\infty} e^{itx} f_X(x)dx. \quad (1.12)$$

- ◆ Fonction génératrice des moments : si X est une variable aléatoire qui admet pour fonction de densité f_X , sa fonction génératrice des moments est définie par :

$$M_x(t) = E(e^{tx}) = \int_{-\infty}^{+\infty} e^{tx} f_X(x)dx \quad (1.13)$$

Exemples

1.5.1 Loi uniforme

Une variable aléatoire X absolument continue est dite distribuée uniformément sur l'intervalle $[a, b]$ si sa densité de probabilité est :

$$f(x) = \begin{cases} \frac{1}{b-a}, & \text{si } x \in [a, b] \\ 0, & \text{ailleurs} \end{cases} \quad (1.14)$$

♣ Sa fonction de répartition :

$$F_x(x) = \begin{cases} 0, & \text{si } x < a \\ \frac{x-a}{b-a}, & \text{si } x \in [a, b] \\ 1, & \text{si } x \geq b \end{cases} \quad (1.15)$$

♣ L'espérance d'une variable aléatoire uniforme est :

$$\frac{a+b}{2}$$

♣ La variance d'une variable aléatoire uniforme est :

$$\frac{(a-b)^2}{12}$$

♣ La fonction génératrice des moments d'une variable aléatoire continue suivant une loi uniforme est :

$$M_x(t) = E(e^{xt}) = \frac{e^{tb} - e^{ta}}{t(b-a)}.$$

1.5.2 Loi exponentielle

-La loi exponentielle décrit la durée de vie d'un phénomène sans vieillissement (particule radioactive, temps d'attente,....).

-La densité de probabilité d'une variable aléatoire continue suivant une loi exponentielle $\text{Exp}(\alpha)$ est :

$$f(X) = \alpha e^{-\alpha x}, \quad x \in \mathbb{R}, \quad (1.16)$$

Où α est un nombre réel strictement positif, d'espérance :

$$E(X) = \frac{1}{\alpha}$$

et de variance :

$$V(X) = \frac{1}{\alpha^2}.$$

1.5.3 Loi normale

-Une variable aléatoire X absolument continue est dite normale (ou suit une loi normale, ou gaussienne) si elle admet pour densité de probabilité la fonction :

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right), x \in \mathbb{R}, \quad (1.17)$$

où m et σ sont respectivement la moyenne et l'écart-type. On dit aussi que X suit une loi normal de paramètre $\mathcal{N}(m, \sigma)$.

si

$$X^* = \frac{x-m}{\sigma}$$

est la variable aléatoire réduite correspondante, alors

$$E(X^*) = 0 \text{ et } V(X^*) = 1$$

et la densité de X^* s'exprime par :

$$g(u) = \frac{1}{\sqrt{2\pi}} * \exp\left(-\frac{u^2}{2}\right),$$

si X suit une loi normale $\mathcal{N}(m, \sigma)$, alors $E(X) = m$, et $V(X) = \sigma^2$.

On peut ramener tout calcul sur la fonction de répartition d'une variable aléatoire normale

$$\mathcal{N}(m, \sigma)$$

à un calcul sur la fonction de répartition d'une variable aléatoire normal réduite $\mathcal{N}(0, 1)$.

On désigne par Φ cette fonction :

$$P(X \leq a) = P\left(\frac{X-m}{\sigma} \leq \frac{a-m}{\sigma}\right) = P\left(X^* \leq \frac{a-m}{\sigma}\right) = \Phi\left(\frac{a-m}{\sigma}\right)$$

-Les valeurs de Q sont données par les tables. Elles permettent de faire les calculs sur une loi normale de paramètres (m, σ) quelconques.

Cette loi est très importante tant du point de vue théorique que pratique .

1.5.4 Loi de Gamma

On présente la famille de lois Gamma ou d'Euler très utiles pour les propriétés de décroissance rapide de leur fonction de survie. Une variable aléatoire continue suit une loi Gamma, de paramètres $\alpha, \beta \in \mathbb{R}_+^*$,

le premier est appelé paramètre d'échelle alors que β est le paramètre de forme,

si elle admet pour densité de probabilité la fonction :

$$f_X(x) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, & \forall x > 0 \\ 0 & \text{ailleurs} \end{cases},$$

où

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

On note $X \rightsquigarrow GA(\alpha, \beta)$.

♣ L'espérance mathématique de X est :

$$E(X) = \frac{\alpha}{\beta}.$$

♣ La variance de X est :

$$V(X) = \frac{\alpha}{\beta^2}.$$

1.5.5 Loi de Student

-est une loi de probabilité, faisant intervenir le quotient entre une variable suivant une loi normale centrée réduite et la racine carrée d'une variable distribuée suivant la loi du χ^2 .

-Soit Z une variable aléatoire de loi normale centrée et réduite et soit U une variable indépendante de Z et distribuée suivant la loi du χ^2 à K degrés de liberté.

Par définition la variable

$$T = \frac{Z}{\sqrt{\frac{U}{K}}}$$

Suit une loi de Student à K degrés de liberté.

★ La densité de T notée f_T , est donnée par :

$$f_T(t) = \frac{1}{\sqrt{k\pi}} \frac{\Gamma(\frac{K+1}{2})}{\Gamma(\frac{K}{2})} \left(\frac{1}{(1 + \frac{t^2}{K})} \right)^{\frac{K+1}{2}}.$$

Pour $K \geq 1$

Où Γ est la fonction Gamma d'Euler.

La densité f_T associée à la variable T est symétrique, centrée sur 0, en forme de cloche.

★ Son espérance ne peut pas être définie pour $K = 1$, et est nulle pour $K > 1$.

★ Sa variance est infinie pour $K \leq 2$ et vaut $\frac{K}{K-2}$ pour $K > 2$.

1.5.6 Loi du Khi-Deux

-Soient K variables aléatoires X_1, X_2, \dots, X_k indépendantes suivant une loi normale $\mathcal{N}(m_i, \sigma_i), i = 1, \dots, K$.

La variable aléatoire

$$Z = \sum_{i=1}^K \left(\frac{X_i - m_i}{\sigma_i} \right)^2$$

Suit une loi de Khi-deux (χ_k^2) à k degrés de liberté.

La densité de probabilité de la variable aléatoire Z est :

$$f(X) = \begin{cases} \frac{1}{2^{\frac{k}{2}} \Gamma(\frac{k}{2})} e^{-\frac{x}{2}} x^{\frac{k}{2}-1}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

où Γ est la fonction gamma, On a :

$$E(Z) = K, \quad V(Z) = 2K.$$

Cette loi est surtout utile dans les tests statistiques .

1.5.7 Loi de Fisher-Snedecor

Soit X une v.a. suit une loi de Fisher-Snedecor à n_1 et n_2 degrés de liberté si :

1. $X(\Omega) = \mathbb{R}_+$
2. La fonction de densité de probabilité f est définie pour tout x positif par :

$$f(X) = n_1^{\frac{n_1}{2}} \cdot n_2^{\frac{n_2}{2}} \cdot \frac{\Gamma\left(\frac{n_1+n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right) \cdot \Gamma\left(\frac{n_2}{2}\right)} \cdot \frac{x^{\frac{n_2-1}{2}}}{(n_2x + n_1)^{\left(\frac{n_1+n_2}{2}\right)}}$$

avec :

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

Notation : $X \rightsquigarrow F(n_1; n_2)$

-Les moments de la loi de Fisher-Snedecor font l'objet de la proposition suivante ;

Proposition 1.1 Si $X \rightsquigarrow F(n_1; n_2)$, alors

$$E(X) = \frac{n_2}{n_2 - 2} \quad \text{pour } n_2 > 2$$

et

$$V(X) = \frac{2n_2^2(n_1 + n_2 - 2)}{n_1(n_2 - 4)(n_2 - 2)^2} \quad \text{pour } n_2 > 4$$

L'espérance de X ne dépend pas de n_1 et lorsque n_2 tend vers 1 par valeur supérieure, par ailleurs, lorsque n_1 et n_2 tendent vers l'infinie, la variance de X tend vers 0.

1.6 Fonction de répartition

Définition 1.7 Soit X une variable aléatoire réelle définie sur un espace probabilisé (Ω, A, P) . On appelle fonction de répartition de X l'application

$$F : \mathbb{R} \rightarrow [0;1]$$

$$x \rightarrow P(X \leq x)$$

Tout fonction de répartition F possède les propriétés suivantes :

1. F est croissante.
2. $\lim_{x \rightarrow -\infty} F(x) = 0$ et $\lim_{x \rightarrow +\infty} F(x) = 1$
3. F est continue à droite en tout point x de \mathbb{R} .
4. F a une limite à gauche en tout point x de \mathbb{R} , et quelque soit $x \in \mathbb{R}$,

$$F(x) - \lim_{t \rightarrow x^-} F(t) = P(X = x).$$

1.7 Fonction caractéristique

La fonction caractéristique d'une variable aléatoire réelle X est la fonction à valeurs complexes définie sur \mathbb{R} par :

$$\varphi_X(t) = E(e^{ixt}), i^2 = -1$$

Propriétés :

Il y a une relation entre les moments et la fonction caractéristique d'une variable aléatoire.

Lorsque les moments de tout ordre existent et que leur série exponentielle a un rayon de convergence non nul R alors pour tout $t \in]-R, R[$ on a :

$$\varphi_X(t) = E(e^{ixt}) = \sum_{k=0}^{\infty} \frac{i^k E(X^k)}{k!} t^k.$$

Pour tout $k \geq 0$, φ_x est k fois dérivable en 0 et $\varphi_x^k(0) = i^k E(X^k)$. Par exemple : l'espérance et la variance de X : $E(X) = -i\varphi_x'(0)$ et $var(X) = -\varphi_x''(0) + [\varphi_x'(0)]^2$ respectivement.

1.8 Fonction génératrice des moments

Il faut bien garder à l'esprit que ces grandeurs ne suffisent pas à déterminer entièrement la loi de x , même si ce sont de très bons indicateurs. En revanche, la fonction génératrice des moments contient cette information, elle est définie par :

$$M_X(t) = E[e^{tX}]$$

pour un réel t tel que la somme converge (la fonction génératrice n'existe pas pour tout t , ni même pour certaines lois paramétriques comme la loi de Pareto ou la loi log-normale). Notons au passage que l'on retrouve parfois dans la littérature une définition avec une exponentielle complexe, mais l'idée sous-jacente est la même. On peut en effet écrire pour tout $n \in \mathbb{N}$:

$$\begin{aligned} \frac{d^n M_X(t)}{dt^n} \Big|_{t=0} &= \frac{d^n E[e^{tX}]}{dt^n} \Big|_{t=0} \\ &= E \left[\frac{d^n e^{tX}}{dt^n} \right] \Big|_{t=0} \\ &= E[X^n e^{tX}] \Big|_{t=0} \\ &= E[X^n] \end{aligned}$$

On comprend dès lors l'origine du nom de cette fonction, puisque ses dérivées donnent accès à tous les moments d'ordre n de X .

Introduction

Une distribution de Lindley à un seul paramètre a attiré les chercheurs pour son usage en modélant des données de vie, et on a observé dans plusieurs articles que cette distribution a exécuté excellemment. Cette distribution est introduite par Lindley en 1958 [1] comme mélange d'Exp(θ) et de Gamma ($2, \theta$).

2.1 Caractérisation de la distribution de lindley

Soit X une variable aléatoire suivant la distribution de Lindley à un paramètre avec la fonction de densité :

$$f(x, \theta) = \begin{cases} \frac{\theta^2(1+x)e^{-\theta x}}{1+\theta}, & x, \theta > 0 \\ 0, & \text{Sinon} \end{cases} \quad (2.1)$$

La distribution de Lindley est importante pour étudier la modélisation de la fiabilité.

La fonction de répartition correspondante est :

$$F(x, \theta) = 1 - \frac{\theta + 1 + \theta x}{1 + \theta} e^{-\theta x}; x > 0, \theta > 0 \quad (2.2)$$

La première dérivée de (2.1) est :

$$\frac{d}{dx} f(x) = \frac{\theta^2}{1 + \theta} (1 - \theta - \theta x) e^{-\theta x}.$$

Il en résulte que

(i) : Pour $\theta < 1$, $\frac{d}{dx}f(x) = 0$ implique que $x_0 = \frac{1-\theta}{\theta}$ est le point critique unique à laquelle $f(x)$ est maximisée.

(ii) : Pour $\theta \geq 1$, $\frac{d}{dx}f(x) \leq 0$, c'est à dire . $f(x)$ diminue en x , par conséquent, le mode de cette distribution est :

$$mode(x) = \begin{cases} \frac{1-\theta}{\theta}, & 0 < \theta < 1 \\ 0, & \text{ailleurs} \end{cases}$$

La figure (2.1) représente la fonction de densité. de la distribution de Lindley pour quelques valeurs de θ .

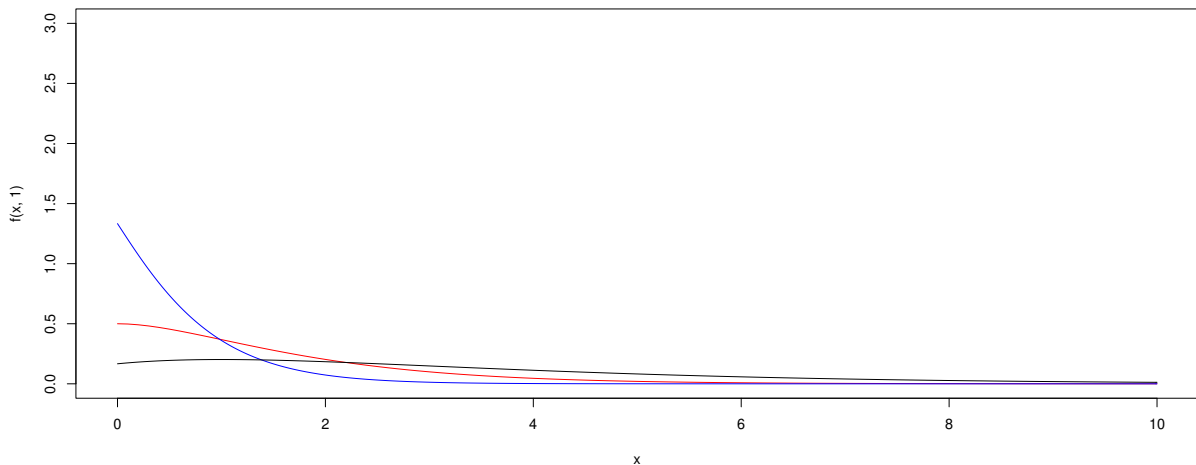


FIGURE 2.1 – Présentation graphique de la fonction de densité pour quelques valeurs de θ , noir($\theta=0.5$);rouge($\theta=1$);bleu($\theta=2$)

2.1.1 Moments et mesures connexes

Le moment d'ordre K de la distribution de Lindley est :

$$\mu'_k = E(X^k) = \frac{k!(\theta + k + 1)}{\theta^k(\theta + 1)}, k = 1, 2, 3, \dots$$

D'où ,on a :

$$\begin{aligned}\mu'_1 &= \frac{\theta + 2}{\theta(\theta + 1)}, \\ \mu'_2 &= \frac{2(\theta + 3)}{\theta^2(\theta + 1)}, \\ \mu'_3 &= \frac{6(\theta + 4)}{\theta^3(\theta + 1)}, \\ \mu'_4 &= \frac{24(\theta + 5)}{\theta^4(\theta + 1)}.\end{aligned}$$

Le moment centré d'ordre k de la distribution de Lindley est définie par :

$$\mu_k = E(X - \mu)^r = \sum_{r=0}^k \binom{k}{r} \mu'_k (-\mu)^{k-r}.$$

D'où, on a

$$\begin{aligned}\mu_2 &= \frac{\theta^2 + 4\theta + 2}{\theta^2(\theta + 1)} = \sigma^2, \\ \mu_3 &= \frac{2(\theta^3 + 6\theta^2 + 6\theta + 2)}{\theta^3(\theta + 1)^3}, \\ \mu_4 &= \frac{3(3\theta^4 + 24\theta^3 + 44\theta^2 + 32\theta + 8)}{\theta^4(\theta + 1)^4}.\end{aligned}$$

Le coefficient de variation (γ), le coefficient de dissymétrie ($\sqrt{\beta_1}$) et

le coefficient d'aplatissement (β_2) sont :

$$\gamma = \frac{\sqrt{\theta^2 + 4\theta + 2}}{\theta + 2},$$

$$\sqrt{\beta_1} = \frac{2(\theta^3 + 6\theta^2 + 6\theta + 2)}{(\theta^2 + 4\theta + 2)^{\frac{3}{2}}}$$

$$\beta_2 = \frac{3(3\theta^4 + 24\theta^3 + 44\theta^2 + 32\theta + 8)}{(\theta^2 + 4\theta + 2)^2}.$$

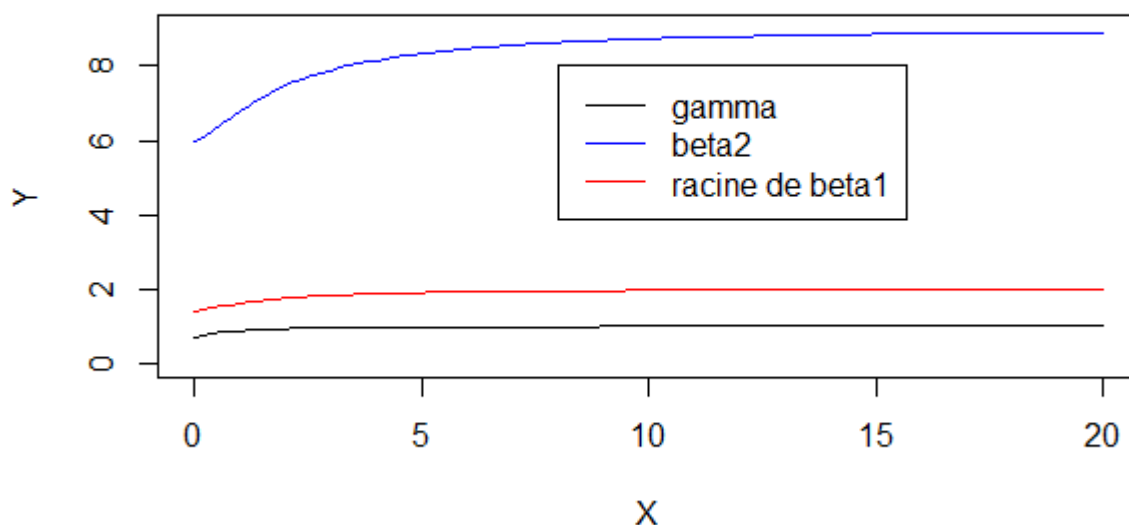


FIGURE 2.2 – Représentation graphique du coefficient de variation γ (noir), le coefficient de dissymétrie $\sqrt{\beta_1}$ (rouge), et le coefficient d'aplatissement β_2 (bleu)

Remarque 2.1 (i) γ est une fonction croissante en θ et $(\frac{1}{\sqrt{2}} < \gamma < 1)$, voir Fig (2.2).
(ii) $\sqrt{\beta_1}$ est une fonction croissante en θ et $\sqrt{2} < \sqrt{\beta_1} < 2$, voir Fig (2.2).
(iii) β_2 est une fonction croissante en θ et $6 < \beta_2 < 9$, voir Fig (2.2).

Théorème 2.1 Soit $X \sim \text{Lindley}(\theta)$. Alors

$$\text{Mode}(X) < \text{Median}(X) < E(X).$$

Preuve. Soit $M = \text{Mode}(X)$, $m = \text{Median}(X)$ et $\mu = E(X) = \frac{\theta+2}{\theta(\theta+1)}$. Depuis la fonction de répartition de la distribution de Lindley, il en résulte que :

$$F(M) = \begin{cases} 1 - \frac{2}{1+\theta} e^{-(1-\theta)}, & 0 < \theta < 1 \\ 0 & \text{ailleurs} \end{cases}$$

$$F(m) = \frac{1}{2}$$

et

$$F(\mu) = 1 - \frac{\theta^2 + 3\theta + 3}{(1 + \theta)^2} \times e^{-\frac{\theta+2}{\theta+1}}.$$

Notons que $F(M)$ est une fonction décroissante en $\theta \in (0, 1)$ et pour tout $\theta > 0$, on a

$$0 \leq F(M) < 1 - 2e^{-1} < (1/2).$$

De même, $F(\mu)$ est une fonction décroissante en $\theta > 0$ et

$$\frac{1}{2} < 1 - 3e^{-2} < F(\mu) < 1.$$

Enfin, étant donné que $F(x)$ est une fonction croissante en $x > 0$. Il est facile de vérifier que $F(M) < F(me) = \frac{1}{2} < F(\mu)$, alors on a $M < m < \mu$. ■

2.1.2 Fonction de hasard et fonction de survie

Soit

$$h(x) = \lim_{\Delta x \rightarrow 0} \frac{P(X < x + \Delta x | X > x)}{\Delta x} = \frac{f(x)}{1 - F(x)} = \frac{\theta^2(1+x)}{\theta + 1 + \theta x}.$$

et

$$S(x) = 1 - F(x) = \frac{\theta + 1 + \theta x}{1 + \theta} e^{-\theta x}.$$

La fonction de taux de hasard et la fonction de survie de la distribution de Lindley, respectivement .

Remarque 2.2 (i)

$$h(0) = f(0) = \frac{\theta^2}{\theta + 1}.$$

(ii) Comme

$$\frac{d}{dx} h(x) = \frac{\theta^2(1+x)}{\theta + 1 + \theta x},$$

$h(x)$ est une fonction croissante en x et θ en plus $\frac{\theta^2}{\theta+1} < h(x) < \theta$.

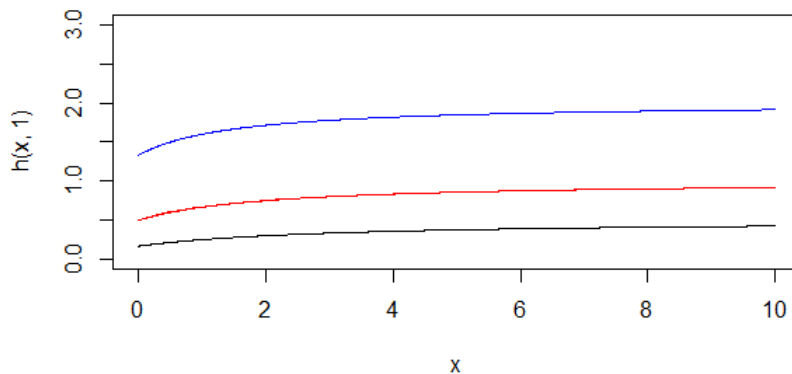


FIGURE 2.3 – Présentation graphique de la fonction de taux de hasard pour quelques valeurs de θ , noir ($\theta = 0.5$); rouge ($\theta = 1$); bleu ($\theta = 2$).

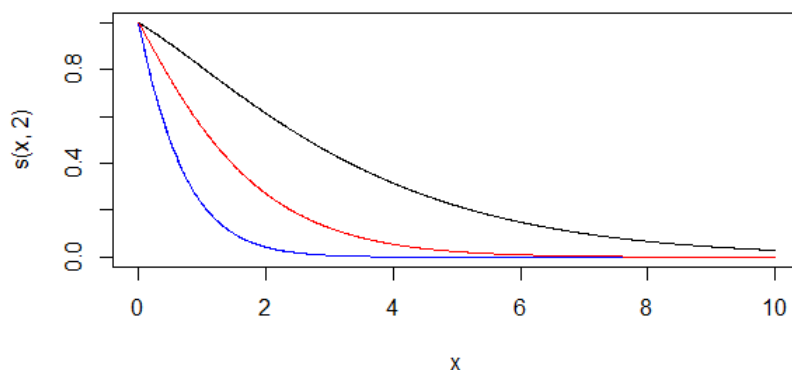


FIGURE 2.4 – Présentation graphique de la fonction de survie pour quelques valeurs de θ , noir($\theta=0.5$); rouge ($\theta=1$); bleu($\theta=2$)

2.1.3 Ordre stochastique

Théorème 2.2 Soient $X \rightsquigarrow \text{Lindely}(\theta_1)$ et $Y \rightsquigarrow \text{Lindely}(\theta_2)$: si $\theta_1 > \theta_2$. Alors : $X \leq_{lr} Y$ et donc $X \leq_{hr} Y$ et $X \leq_s Y$.

Preuve. Notons d'abord que

$$\frac{f_X(t)}{f_Y(t)} = \frac{\theta_1^2(1 + \theta_2)}{\theta_2^2(1 + \theta_1)} e^{-(\theta_1 - \theta_2)t}, t > 0$$

On a, pour $\theta_1 > \theta_2$,

$$\frac{d}{dt} \frac{f_X(t)}{f_Y(t)} = (\theta_2 - \theta_1) \frac{f_X(t)}{f_Y(t)} < 0,$$

$\frac{f_X(t)}{f_Y(t)}$ est décroissante en X . Alors $X \leq_{lr} Y$.

■

2.1.4 Courbe de Lorenz

La courbe de Lorenz pour une variable aléatoire X positif est définie comme le graphe du rapport

$$L(F(x)) = \frac{E(X | X \leq x)F(x)}{E(X)} \quad (2.3)$$

Pour la distribution de Lindley (2.1) on a,

$$E(X | X \leq x)F(x) = \frac{2 + \theta}{\theta(1 + \theta)} - \frac{e^{(-\theta x)}}{1 + \theta} \left[\frac{2}{\theta + 1 + \theta x^{2+2x+x}} \theta \right].$$

Ainsi, à partir de (2.3), on obtient la courbe de Lorenz pour la distribution de Lindley comme suit :

$$L(p) = 1 - \frac{\theta(1+\theta)(1-p)}{(2+\theta)(1+\theta+\theta x)} \left[\frac{2}{\theta} + 1 + \theta x^2 + 2x + x\theta \right].$$

Où $x = F^{-1}(p)$ avec $F(\cdot)$ donnée par (2.2).

2.1.5 Statistiques d'ordre extrêmes

Si X_1, \dots, X_n un échantillon de n variables aléatoires qui suivent la distribution de Lindley et si

$$\bar{X} = \frac{1}{n}(X_1 + \dots + X_n)$$

représente la moyenne d'échantillon alors par le théorème central limite

$$\frac{\sqrt{n}(\bar{X} - E(X))}{\sqrt{\text{var}(X)}} \longrightarrow \mathcal{N}(0, 1)$$

se rapproche de la distribution normale standard quand $n \rightarrow \infty$.

Théorème 2.3 *Parfois, on serait intéressé à étudier la loi asymptotique des valeurs extrêmes $M_n = \max(X_1, \dots, X_n)$ et $m_n = \min(X_1, \dots, X_n)$,*

Pour la fonction de répartition définie dans (2.2), on constate que :

$$\lim_{t \rightarrow \infty} \frac{1 - F(t+x)}{1 - F(t)} = \exp(-\theta x).$$

$$\lim_{t \rightarrow 0} \frac{F(tx)}{F(t)} = x.$$

Ainsi, il en résulte du Théorème 1.4 (Leadbetter et al. [5]) [6] qu'il doit y avoir les constantes de normalisation $a_n > 0; b_n; c_n > 0$ et d_n de telle sorte que :

$$\Pr(a_n(M_n - b_n) \leq x) \longrightarrow \exp(-\theta x).$$

et

$$\Pr(c_n(m_n - d_n) \leq x) \longrightarrow 1 - \exp(-x), \tag{2.4}$$

comme $n \rightarrow \infty$. La forme des constantes de normalisation peut également être déterminée. Par exemple, en utilisant le Corollaire 1.1 (Leadbetter et al. [5], on peut voir que $a_n = 1$ et $b_n = F^{-1}(1 - \frac{1}{n})$ avec $F(\cdot)$ donnée par (2.2).

2.1.6 Fonction Quantile de la distribution Lindley

D'après la fonction de répartition de la distribution Lindley définie en (2.2). Il convient de noter qu'est continue et strictement croissante de sorte que la fonction de quantile X est

$$Q_X(\mu) = F_X^{-1}(\mu), \quad 0 < \mu < 1.$$

Dans le résultat suivant, on donne une expression explicite de Q_X en fonction de la fonction W de Lambert.

Théorème 2.4 *Pour tout $\theta > 0$, la fonction quantile de la distribution Lindley X est :*

$$Q_X(\mu) = -1 - \frac{1}{\theta} - \frac{1}{\theta} W_{-1} \left(\frac{\theta + 1}{\exp(\theta + 1)} (\mu - 1) \right), \quad 0 < \mu < 1, \quad (2.5)$$

Où W_{-1} désigne la branche négative de la fonction W de Lambert.

Preuve. Pour tout θ fixe, $\theta > 0$, soit $\mu \in (0, 1)$. On doit résoudre l'équation $F_X(x) = \mu$ par rapport à x , pour tous $X > 0$ comme suit :

$$(\theta + 1 + \theta x)e^{-\theta x} = (\theta + 1)(1 - \mu) \quad (2.6)$$

En multipliant par

$$-e^{(-\theta-1)}$$

l'équation (2.6), on obtient :

$$-(\theta + 1 + \theta x)e^{(-\theta-1-\theta x)} = (\theta + 1)(\mu - 1)e^{(-\theta-1)}. \quad (2.7)$$

D'après l'équation (2.7), conjointement avec l'équation :

$$W(z)e^{W(z)} = z$$

, on voit que $-(\theta + 1 + \theta x)$ est la fonction W de Lambert de l'argument réel $(\theta + 1)(\mu - 1)e^{(-\theta-1)}$

Alors, on a

$$W \left(\frac{\theta + 1}{\exp(\theta + 1)} (\mu - 1) \right) = -(\theta + 1 + \theta x), \quad 0 < \mu < 1. \quad (2.8)$$

Toujours, pour tout $\theta > 0$ et $x > 0$ il est immédiat que

$$(\theta + 1 + \theta x) > 1$$

et il peut également être vérifié que puisque $\mu \in (0, 1)$. En prenant en compte les propriétés de la branche négative de la fonction W de Lambert nous l'avons mentionné

plus tôt, l'équation (2.8) devient

$$W_{-1}\left(\frac{\theta+1}{\exp(\theta+1)}(\mu-1)\right) = -(\theta+1+\theta x). \quad (2.9)$$

Ce qui implique le résultat. la preuve du théorème est terminée

■

2.1.7 Estimation

Estimation par la méthode des moments (MoM)

Étant donné un échantillon aléatoire X_1, \dots, X_n , de la distribution de Lindley (2.1), l'estimateur des moments (MoM) de θ est :

$$\hat{\theta}_{MoM} = \frac{-(\bar{X}-1) + \sqrt{(\bar{X}-1)^2 + 8\bar{X}}}{2\bar{X}}, \quad \bar{X} > 0. \quad (2.10)$$

Le théorème suivant montre que l'estimateur θ_{MoM} de θ est biaisé.

Théorème 2.5 *L'estimateur θ_{MoM} de θ est positivement biaisée, i.e.,*

$$E(\theta) - \theta > 0$$

Preuve. Soient $\hat{\theta}_{MoM} = g(\bar{X})$ et $g(t) = \frac{-(t-1) + \sqrt{(t-1)^2 + 8t}}{2t}, \forall t > 0$.

Comme

$$g^{(2)}(t) = \frac{1}{t^3} \left[1 + \frac{3t^3 + 15t^2 + 9t + 1}{[(t-1)^2 + 8t]} \right] > 0,$$

$g(t)$ est strictement convexe. Ainsi, par l'inégalité de Jensen, on a

$$E(g(\bar{X})) > g[E(\bar{X})].$$

Enfin, étant donné que

$$E(g(\bar{X})) = g(\mu) = g\left(\frac{\theta+2}{\theta(\theta+1)}\right) = \theta$$

On obtient $(\hat{\theta}_{MoM}) > \theta$

Le théorème suivant donne la loi limite de $\hat{\theta}_{MoM}$.

■

Théorème 2.6 *L'estimateur $\hat{\theta}_{MoM}$ de θ est convergent et asymptotiquement normal :*

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{D} \mathcal{N}\left(0, \frac{1}{\sigma^2}\right)$$

L'intervalle de confiance de θ pour un seuil de confiance $100(1-\alpha)$ est donnée par :

$$\hat{\theta} \pm Z_{\frac{\alpha}{2}} \frac{1}{\sqrt{n\hat{\sigma}^2}}.$$

Où $Z_{\frac{\alpha}{2}}$ est le $(1 - \frac{\alpha}{2})$ percentile de la distribution normale standard.

Preuve. Étant donné μ est finie, $\bar{X} \xrightarrow{p} \mu$.

$g(t)$ est une fonction continue à $t = \mu$, $g(\bar{X}) \xrightarrow{p} g(\mu)$ c'est-à-dire

$$\hat{\theta} \xrightarrow{p} \theta.$$

Comme

$$\sigma^2 < \infty,$$

par le théorème central limite, on a

$$\sqrt{n}(\bar{X} - \mu) \xrightarrow{D} \mathcal{N}(0, \sigma^2).$$

En outre, puisque $g(\mu)$ est différentiable et $g'(\mu) \neq 0$, par la méthode Delta, on a :

$$\sqrt{n}(g(\bar{X}) - g(\mu)) \xrightarrow{D} \mathcal{N}(0, [g'(\mu)]^2 \sigma^2)$$

Enfin, étant donné que

$$g(\bar{X}) = \hat{\theta}_{MoM}, g(\mu) = \theta$$

et

$$g'(\mu) = \frac{-1}{2\mu^2} \left[1 + \frac{1 + 3\mu}{\sqrt{(\mu - 1)^2 + 8\mu}} \right] = -\frac{1}{\sigma^2}$$

■

Estimation par la méthode du maximum de vraisemblance

Soient $X_i \rightsquigarrow LD(\theta)$, $i = 1, \bar{n}$ n variables aléatoires. La fonction de log-vraisemblance est :

$$\ln(x_i; \theta) = 2n \ln \theta - n \ln(\theta + 1) + \sum_{i=1}^n \ln(x_i + 1) - n\theta \bar{X}.$$

L'estimateur de la méthode du maximum de vraisemblance $\hat{\theta}_{MV}$ de θ est une solution de l'équation :

$$\frac{\partial \ln l(x_i; \beta; \theta)}{\partial \theta} = \frac{2n}{\theta} - \bar{X} - \frac{n}{(\theta + 1)} = 0.$$

On obtient

$$\hat{\theta}_{MV} = \frac{-(\bar{X} - 1) + \sqrt{(\bar{X} - 1)^2 + 8\bar{X}}}{2\bar{X}}, \bar{X} > 0.$$

Avec

$$\frac{\partial^2 \ln l(x_i; \beta; \theta)}{\partial \theta^2} = -\frac{2n}{\theta^2} - \sum_{i=1}^n \frac{x_i^2}{(\theta x_i + \theta)^2} < 0.$$

Remarque 2.3 L'estimateur de la méthode des moments $\hat{\theta}_{MoM}$ et l'estimateur du maximum de vraisemblance $\hat{\theta}_{MV}$ du paramètre θ sont les mêmes.

2.2 Distribution de Lindley à deux paramètres

Une distribution de Lindley à deux paramètres (deux paramètres LD) de paramètres α et θ est défini par sa fonction de densité de probabilité

$$f(x; \alpha; \theta) = \frac{\theta^2}{\theta + \alpha} (1 + \alpha x) \exp\{-\theta x\}, \quad x > 0, \quad \theta > 0, \alpha > -\theta. \quad (2.11)$$

On constate facilement que pour $\alpha = 1$, la distribution (2.11) se réduit au seul paramètre LD (2.1) et pour $\alpha = 0$, il réduit à la distribution exponentielle avec des paramètres θ . (2.11) peut être représentée comme un mélange de distributions exponentielles (θ) et $gamma(2, \theta)$ comme suit :

$$f(x; \alpha, \theta) = p f_1(x) + (1 - p) f_2(x) \quad (2.12)$$

Où

$$p = \frac{\theta}{\theta + \alpha},$$

$$f_1(x) = \theta e^{-\theta x},$$

et

$$f_2(x) = \theta^2 x e^{-\theta x}.$$

La dérivée première de (2.11) est obtenue comme

$$f'(x) = \frac{\theta^2}{\theta + \alpha} (\alpha - \theta - \alpha \theta x) e^{-\theta x}.$$

et donc

$$f'(x) = 0, \text{ donne } x = \frac{\alpha - \theta}{\alpha \theta}.$$

Il en résulte que,

1. Pour $\theta < \alpha$, $x_0 = \frac{\alpha - \theta}{\alpha \theta}$ est l'unique point critique où $f'(x)$ est maximum
2. Pour $\theta \geq \alpha$, $f'(x) \leq 0$ i.e, $f(x)$ diminue en x .

Par conséquent, le mode de distribution est donné par

$$Mode(x) = \begin{cases} \frac{\alpha - \theta}{\alpha \theta}, & \theta < \alpha \\ 0 & \text{ailleurs} \end{cases} \quad (2.13)$$

La fonction de distribution cumulative (c.d.f) des deux paramètres LD est donnée par :

$$F(x) = 1 - \frac{\theta + \alpha + \alpha \theta x}{\theta + \alpha} e^{-\theta x}; x > 0, \theta > 0, \alpha > -\theta \quad (2.14)$$

2.2.1 Moments et mesures associées

Le r ème moment autour de l'origine du LD à deux paramètres a été obtenu comme

$$\mu'_r = \frac{\Gamma(r+1)(\theta + \alpha + \alpha r)}{\theta^r(\theta + \alpha)}; r = 1, 2, \dots \quad (2.15)$$

En prenant $r = 1, 2, 3$, et 4 dans (2.15), les quatre premiers moments autour de l'origine sont obtenues comme

$$\begin{aligned} \mu'_1 &= \frac{\theta + 2\alpha}{\theta(\theta + \alpha)}, \\ \mu'_2 &= \frac{2(\theta + 3\alpha)}{\theta^2(\theta + \alpha)}, \\ \mu'_3 &= \frac{6(\theta + 4\alpha)}{\theta^3(\theta + \alpha)}, \\ \mu'_4 &= \frac{24(\theta + 5\alpha)}{\theta^4(\theta + \alpha)} \end{aligned} \quad (2.16)$$

On vérifie facilement que pour $\alpha=1$, les moments sur l'origine du LD à deux paramètres se réduisent aux moments respectifs du LD à un paramètre. Plus loin, puisque la moyenne de la distribution est toujours supérieure à le mode, la distribution est positivement asymétrique. Les moments centraux de cette distribution ont ainsi été obtenus comme

$$\mu_2 = \frac{\theta^2 + 4\theta\alpha + 2\alpha^2}{\theta^2(\theta + \alpha)^2} \quad (2.17)$$

$$\mu_3 = \frac{2(\theta^3 + 6\theta^2\alpha + 6\theta\alpha^2 + 2\alpha^3)}{\theta^3(\theta + \alpha)^3} \quad (2.18)$$

$$\mu_4 = \frac{3(3\theta^4 + 24\theta^3\alpha + 24\theta^2\alpha^2 + 20\theta^2\alpha^2 - 88\theta\alpha^3 + 120\theta\alpha^2 + 8\alpha^4)}{\theta^4(\theta + \alpha)^4} \quad (2.19)$$

On peut facilement vérifier que pour $\alpha=1$, la central les moments du LD à deux paramètres se réduisent aux moments respectifs du LD à un paramètre. Les coefficients de variation (γ), asymétrie ($\sqrt{\beta_1}$) et d'aplatissement (β_2) du LD à deux paramètres sont donné par

$$\gamma = \frac{\sigma}{\mu_1} = \frac{\sqrt{\theta^2 + 4\theta\alpha + 2\alpha^2}}{\theta + 2\alpha} \quad (2.20)$$

$$\sqrt{\beta_1} = \frac{\mu_3}{\mu_2^{3/2}} = \frac{2(\theta^3 + 6\theta^2\alpha + 6\theta\alpha^2 + 2\alpha^3)}{(\theta^2 + 4\theta\alpha + 2\alpha^2)^{3/2}} \quad (2.21)$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{3(3\theta^4 + 24\theta^3\alpha + 24\theta^2\alpha^2 + 20\theta^2\alpha^2 - 88\theta\alpha^3 + 120\theta\alpha^2 + 8\alpha^4)}{(\theta^2 + 4\theta\alpha + 2\alpha^2)^2} \quad (2.22)$$

On vérifie facilement que pour $\alpha = 1$, les coefficients de variation (γ), d'asymétrie ($\sqrt{\beta_1}$), et d'aplatissement (β_2) de la LD à deux paramètres réduit à la valeur respective coefficients du paramètre LD

2.2.2 Taux d'échec et durée de vie résiduelle moyenne

Pour une distribution continue avec p.d.f. $f(x)$ et c.d.f. $F(x)$, la fonction du taux de défaillance (également connue sous le nom de fonction de taux) et la fonction de durée de vie résiduelle moyenne sont respectivement définies comme

$$h(x) = \lim_{\Delta x \rightarrow 0} \frac{P(X < x + \Delta x | X > x)}{\Delta x} = \frac{f(x)}{1 - F(x)} \quad (2.23)$$

et

$$m(x) = E[X - x | X > x] = \frac{1}{1 - F(x)} \int_x^\infty [1 - F(t)] dt \quad (2.24)$$

La fonction de taux de défaillance correspondante, $h(x)$ et la fonction de durée de vie résiduelle moyenne, $m(x)$ à deux paramètres LD sont donc donnés par

$$h(x) = \frac{\theta^2(1 + \alpha x)}{\theta + \alpha + \theta\alpha x} \quad (2.25)$$

et

$$\begin{aligned} m(x) &= \frac{1}{(\theta + \alpha + \theta \alpha x)e^{-\theta x}} \int_x^\infty (\theta + \alpha + \theta \alpha t)e^{-\theta t} dt \\ &= \frac{\theta + 2\alpha + \theta \alpha x}{\theta(\theta + \alpha + \theta \alpha x)} \end{aligned}$$

On peut facilement vérifier que

$$h(0) = \frac{\theta^2}{\theta + \alpha} = f(0)$$

et

$$m(0) = \frac{\theta + 2\alpha}{\theta(\theta + \alpha)} = \mu'_1.$$

Il est également évident que $h(x)$ est une fonction croissante de x, α et θ , alors que $m(x)$ est une fonction décroissante de x et α , et croissante fonction de θ . Pour $\alpha=1$, (2.25) et (2.26) se réduisent aux mesures correspondantes d'un paramètre LD. La fonction de taux de défaillance et fonction de durée de vie résiduelle moyenne de la distribution montre sa flexibilité sur un paramètre LD et distribution exponentielle.

2.2.3 Ordres stochastiques

L'ordre stochastique des variables aléatoires continues positives est un outil important pour juger de leur comportement. Une variable aléatoire X est dite plus petite qu'une variable aléatoire Y dans

1. Ordre stochastique ($X \leq_{st} Y$) si $F_X(x) \geq F_Y(x)$, pour tout x ;
2. Ordre du taux de risque ($X \leq_{hr} Y$) si $h_X(x) \geq h_Y(x)$, pour tout x ;
3. Ordre de vie résiduel moyen ($X \leq_{mrl} Y$) si $m_X(x) \leq m_Y(x)$, pour tout x ;
4. ordre du rapport de vraisemblance ($X \leq_{lr} Y$) si $\frac{f_X(x)}{f_Y(x)}$ diminue en x .

Les résultats suivants dus à M. Shaked et j.G. Shanthikumar [7] est bien connu pour établir un ordre stochastique des distributions.

$$X \leq_{lr} Y \longrightarrow X \leq_{hr} Y \longrightarrow X \leq_{mrl} Y \Downarrow X \leq_{st} Y \quad (2.26)$$

Le DL à deux paramètres est ordonné par rapport au l'ordre du "rapport de vraisemblance" le plus fort, comme indiqué dans le théorème suivant :

Théorème 2.7 Soit $X \sim LD$ à deux paramètres (α_1, θ_1) et $Y \sim LD$ à deux paramètres (α_2, θ_2) .

Si $\alpha_1 = \alpha_2$ et $\theta_1 \geq \theta_2$ (Ou si $\theta_1 = \theta_2$ et $\alpha_1 \geq \alpha_2$). Alors $X \leq_{lr} Y$ et Par conséquent $X \leq_{hr} Y$, $X \leq_{mrl} Y$ et $X \leq_{st} Y$.

Preuve. Nous avons

$$\frac{f_X(x)}{f_Y(x)} = \left(\frac{\theta_1}{\theta_2}\right)^2 \left(\frac{\theta_2 + \alpha_2}{\theta_1 + \alpha_1}\right) \left(\frac{1 + \alpha_1 x}{1 + \alpha_2 x}\right) e^{-(\theta_1 - \theta_2)x}; x > 0$$

A présent

$$\log \frac{f_X(x)}{f_Y(x)} = 2 \log\left(\frac{\theta_1}{\theta_2}\right) + \log\left(\frac{\theta_2 + \alpha_2}{\theta_1 + \alpha_1}\right) + \log(1 + \alpha_1 x) - \log(1 + \alpha_2 x) - (\theta_1 - \theta_2)x.$$

Ainsi

$$\begin{aligned} \frac{\partial}{\partial x} \log \frac{f_X(x)}{f_Y(x)} &= \frac{1}{1 + \alpha_1 x} - \frac{1}{1 + \alpha_2 x} + (\theta_2 - \theta_1) \\ &= \frac{(\alpha_2 - \alpha_1)x}{(1 + \alpha_1 x)(1 + \alpha_2 x)} + (\theta_2 - \theta_1) \end{aligned}$$

Cas (i) : Si $\alpha_1 = \alpha_2$ et $\theta_1 \geq \theta_2$; Alors

$$\frac{\partial}{\partial x} \log \frac{f_X(x)}{f_Y(x)} < 0$$

Cela signifie que $X \leq_{lr} Y$ et donc $X \leq_{hr} Y$, $X \leq_{mrl} Y$ et $X \leq_{st} Y$.

Cas (ii) : Si $\theta_1 = \theta_2$ et $\alpha_1 \geq \alpha_2$, Alors

$$\frac{\partial}{\partial x} \log \frac{f_X(x)}{f_Y(x)} < 0$$

Cela signifie que $X \leq_{lr} Y$ et donc $X \leq_{hr} Y$, $X \leq_{mrl} Y$ et $X \leq_{st} Y$.

Ce théorème montre la flexibilité de la distribution de Lindely à deux paramètres LD par rapport à la distribution de Lindley à un seul paramètre LD et de la distributions exponentielles.

■

2.2.4 Estimation des paramètres

Estimations du maximum de vraisemblance

Soit (x_1, x_2, \dots, x_n) un échantillon aléatoire de taille n d'une LD à deux paramètres (2.11) et soit f_x la valeur observée fréquence dans l'échantillon correspondant à $X = x(x = 1, 2, \dots, k)$, tel que

$$\sum_{x=1}^k f_x = n;$$

où k est le plus grand valeur observé n'ayant pas zéro fréquence. La fonction de vraisemblance L des deux paramètres LD (2.11) est donnée par

$$L = \left(\frac{\theta^2}{\theta + \alpha} \right)^n \prod_{x=1}^K (1 + \alpha x)^{f_x} e^{-n\theta \bar{x}} \quad (2.27)$$

et donc la fonction log de vraisemblance est obtenue comme

$$\log L = n \log \theta_n^2 \log(\theta + \alpha) + \sum_{x=1}^k f_x \log(1 + \alpha x) - n\theta \bar{X} \quad (2.28)$$

Les deux équations de log de vraisemblance sont ainsi obtenues comme

$$\frac{\partial \log L}{\partial \theta} = \frac{2n}{\theta} - \frac{n}{\theta + \alpha} - n\bar{X} = 0 \quad (2.29)$$

$$\frac{\partial \log L}{\partial \alpha} = -\frac{n}{\theta + \alpha} + \sum_{x=1}^k \frac{x f_x}{1 + \alpha x} = 0 \quad (2.30)$$

Équation (2.29) donne

$$\bar{X} = \frac{\theta + 2\alpha}{\theta(\theta + \alpha)},$$

qui est le moyenne de la DL à deux paramètres. Les deux équations (2.29) et (2.30) ne semblent pas résolus directement . Cependant, la méthode de notation de Fisher peut être appliquée pour résoudre ces équations. Car, nous avons

$$\frac{\partial^2 \log L}{\partial \theta^2} = -\frac{2n}{\theta^2} + \frac{n}{(\theta + \alpha)^2} \quad (2.31)$$

$$\frac{\partial^2 \log L}{\partial \theta \partial \alpha} = \frac{n}{(\theta + \alpha)^2} \quad (2.32)$$

$$\frac{\partial^2 \log L}{\partial \alpha^2} = \frac{n}{(\theta + \alpha)^2} - \sum_{x=1}^k \frac{x f_x}{(1 + \alpha x)^2} \quad (2.33)$$

Les équations suivantes pour $\hat{\theta}$ et $\hat{\alpha}$ peuvent être résolues

$$\begin{bmatrix} \frac{\partial^2 \log L}{\partial \theta^2} & \frac{\partial^2 \log L}{\partial \theta \partial \alpha} \\ \frac{\partial^2 \log L}{\partial \theta \partial \alpha} & \frac{\partial^2 \log L}{\partial \alpha^2} \end{bmatrix}_{\hat{\theta}=\theta_0, \hat{\alpha}=\alpha_0} \times \begin{bmatrix} \hat{\theta} - \theta_0 \\ \hat{\alpha} - \alpha_0 \end{bmatrix} = \begin{bmatrix} \frac{\partial \log L}{\partial \theta} \\ \frac{\partial \log L}{\partial \alpha} \end{bmatrix}_{\hat{\theta}=\theta_0, \hat{\alpha}=\alpha_0} \quad (2.34)$$

Où θ_0 et α_0 sont les valeurs initiales de θ et α respectivement. Ces équations sont résolues itérative-ment jusqu'à des estimations suffisamment proches de $\hat{\theta}$ et $\hat{\alpha}$ sont obtenues.

Estimations de Moments

En utilisant les deux premiers moments autour de l'origine de la distribution de Lindley à deux paramètres, nous avons

$$\frac{\mu_2}{\mu_1^2} = k(say) = \frac{2(\theta + 3\alpha)(\theta + \alpha)}{(\theta + 2\alpha)^2} \quad (2.35)$$

Posons $\theta = b\alpha$, on a

$$\frac{\mu_2}{\mu_1^2} = \frac{2(b+3)(b+1)}{(b+2)^2} = \frac{2b^2 + 8b + 6}{b^2 + 4b + 4} = k$$

Cela donne

$$(2-k)b^2 + 4(2-k)b + 2(3-2k) = 0 \quad (2.36)$$

Qui est une équation quadratique dans b . Remplacement du premier et les seconds instants μ'_1 et μ'_2 par le respectif exemples de moments \bar{X} et m'_2 une estimation de k peut être obtenu, à l'aide de laquelle, l'équation (2.36) peut être résolue et une estimation de b obtenue. Encore une fois, en remplaçant $\theta = b\alpha$ dans l'expression de la moyenne de la LD à deux paramètres, on obtient

$$\bar{X} = \frac{b+2}{\alpha b(b+1)}$$

et donc une estimation de α est donnée par

$$\hat{\alpha} = \left(\frac{b+2}{b(b+1)} \right) \frac{1}{\bar{X}} \quad (2.37)$$

Enfin, une estimation de θ est obtenue comme

$$\hat{\theta} = \left(\frac{b+2}{b+1} \right) \frac{1}{\bar{X}} \quad (2.38)$$

2.3 Distribution de Lindley Étendue

Introduction

Nous considérons la distribution de Lindley qui a été introduite par Lindley (1958)[1] avec la fonction de densité donnée par

$$f(x, \theta) = \frac{\theta^2}{1 + \theta} (1 + x) \exp -\theta x, x > 0, \theta > 0,$$

et la fonction de distribution

$$F(x, \theta) = 1 - \frac{1 + \theta + \theta x}{1 + \theta} \exp -\theta x. \quad (2.39)$$

Bien que la distribution de Lindley ait attiré peu d'attention dans la littérature statistique en raison de la grande popularité de la distribution exponentielle bien connue, elle a une forme fermée et, selon Hussain (2006)[8], la distribution de Lindley est importante pour étudier la modélisation de la fiabilité contrainte-résistance. Par ailleurs, certains chercheurs ont proposé de nouvelles classes de distributions basées sur des modifications de la distribution de Lindley, incluant également propriétés. L'idée principale est toujours orientés en imbriquant les anciennes distributions dans des structures plus souples. Sankaran (1970)[9] a introduit le discret Distribution Poisson-Lindley en combinant les distributions Poisson et Lindley. Ghitany, Atieh et Nadarajah (2008)[2] étudié la plupart des propriétés statistiques de la distribution de Lindley, montrant que cette distribution peut fournir une meilleure ajustement que la distribution exponentielle. Mahmoudi et Zakerzadeh (2010)[10] ont proposé une version étendue du composé Distribution de Poisson qui a été obtenue en combinant la distribution de Poisson avec la distribution de Lindley généralisée. Louzada, Roman et Cancho (2011)[11] ont proposé la distribution géométrique exponentielle complémentaire en combinant les les distributions géométriques et exponentielles. Le but de cet article est d'introduire une extension de la distribution de Lindley qui offre une distribution plus flexible pour modéliser les données de durée de vie, notamment en fiabilité, en termes de formes de taux de défaillance. La nouvelle distribution peut accueillir à la fois des taux d'échec décroissants et croissants en tant que ses prédécesseurs, ainsi que des taux d'échec unimodaux et en forme de baignoire. Nous étendons la distribution de Lindley par exponentiation. Plusieurs auteurs ont envisagé des extensions de la survie habituelle distributions en suivant cette idée. Par exemple, Mudholkar, Srivastava et Freimer (1995)[12] ont considéré l'exponentielle Distribution de Weibull comme généralisation de la distribution de Weibull, Gupta et Kundu (1999)[13] ont introduit la distribution exponentielle distribution exponentielle comme généralisation de la distribution exponentielle usuelle, Nadarajah et Kotz (2006)[14] ont proposé une

distributions de type exponentielle prolongeant les distributions de Fréchet, gamma, Gumbel et Weibull et Barriga, Louzada-Neto et Cancho (2011)[15] ont proposé la distribution de puissance exponentielle complémentaire en exponentielle l'exponentielle distribution de puissance proposée par Smith et Bain (1975)[16]. Par conséquent, nous étendons la distribution de Lindley via sa fonction de distribution, en considérant particulièrement de (2.39) donné par

$$F(x) = 1 - \left(\frac{1 + \theta + \theta x}{1 + \theta} \right)^\alpha \exp(-(\theta x)^\beta), \quad (2.40)$$

Où $\alpha \in \mathbb{R}^- \cup \{0, 1\}$, $\theta > 0$ et $\beta > 0$. Ci-après, l'extension de la distribution de Lindley sera désignée par la distribution Lindley (EL) étendue. La motivation de la distribution EL proposée donnée par (2.40) découle de sa capacité à modéliser les données de durée de vie avec l'augmentation, taux de défaillance décroissants, uni-modaux et en forme de baignoire. Ceci se justifie en introduisant les paramètres α et β , comme nous le verrons voir dans la section suivante. De ce fait, le modèle représente une bonne alternative aux modèles gamma et Weibull les plus populaires. Distributions de durée de vie qui souffrent de ne pas présenter de taux de défaillance uni-modaux et en forme de baignoire. Aussi, nous verrons plus loin que le paramètre θ peut être interprété comme une borne supérieure de la fonction de taux de défaillance, qui est une caractéristique importante pour modèles de durée de vie. Peu de distributions de durée de vie ont leurs paramètres directement interprétantes en termes de taux de risque fonctions, comme le Weibull. Une autre caractéristique intéressante de la distribution EL est qu'elle a des expressions de forme fermée pour ses fonction de distribution cumulative et fonction de taux de défaillance, ce qui n'est pas le cas, par exemple, pour la distribution gamma. De plus, la distribution EL a plusieurs cas particuliers

- Pour $\alpha = 0$ et $\beta = 1$ La distribution EL est réduite à la distribution de Lindley (Lindley, 1958)[1], et
- Pour $\alpha = 0$, elle se réduit à la distribution de Weibull. Aussi,
- Pour $\beta = 0$ et $\alpha \in \mathbb{R}^-$, Le (EL) est réduite à une distribution de Pareto donnée par

$$F(x) = 1 - \left(\frac{c}{c+x} \right)^\delta,$$

où $x > 0$, $c = 1 + \frac{1}{\theta}$ et $\delta > 0$.. De plus,

- Pour $\beta = 1$ et $\alpha \in \mathbb{R}^-$, La fonction de distribution définie par (2.40) se réduit à la distribution de Pareto conique (ou généralisée) avec

$$F(x) = 1 - \left(\frac{c}{c+x} \right)^\delta e^{-\lambda x},$$

où $x > 0$, $c = 1 + \frac{1}{\lambda}$ et $\delta > 0$. Cette distribution a été discutée par Kagan et Schoen-

berg (2001)[17] et utilisé pour modéliser la taille des tremblements de terre. D'ailleurs, l'éq. (2.40) représente le produit des fonctions de survie $(1 - F(x))$ des distributions Lomax (Pareto II) et Weibull[18], respectivement, pour tout β et $\alpha \in \mathbb{R}_+$. c'est-à-dire que la distribution EL peut être considérée comme un mélange d'une distribution Lomax et d'une distribution de Weibull pour les anciens α et β . Tout au long du mémoire, nous utiliserons les équations suivantes :

$$\int_b^\infty x^c (1 + \theta + \theta x)^\alpha (\theta x)^a e^{-(\theta x)^\beta} dx = \frac{1}{\theta^{1+c} \beta} \sum_{j=0}^{\infty} \binom{\alpha}{j} (1 + \theta)^{\alpha-j} \Gamma\left(\frac{1+a+c+j}{\beta}, (\theta b)^\beta\right), \quad (2.41)$$

où $\gamma(\alpha, b)$ est la fonction gamma incomplète supérieure donnée par $\gamma(s, b) = \int_b^\infty t^{s-1} e^{-t} dt$.

$$\int_0^b (1 + \theta + \theta x)^\alpha (\theta x)^a e^{-(\theta x)^\beta} dx = \frac{1}{\theta \beta} \sum_{i=0}^{\infty} \binom{\alpha}{i} (1 + \theta)^{\alpha-i} \sigma\left(\frac{1+i+a}{\beta}, (\theta b)^\beta\right), \quad (2.42)$$

où $\gamma(\alpha, b)$ est la fonction gamma incomplète inférieure donnée par

$$\gamma(\alpha, b) = \int_0^b t^{\alpha-1} e^{-t} dt = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \frac{b^{\alpha+k}}{\alpha+k}$$

2.3.1 Mesures statistiques de fiabilité

Dans cette section, nous donnons quelques mesures statistiques et de fiabilité importantes pour la distribution EL.

2.3.1.1 Fonctions de densité de survie et de taux d'échec

Les fonctions de densité et de survie associées à (2.40) sont données par

$$f(x) = \frac{\theta(1 + \theta + \theta x)^{\alpha-1}}{(1 + \theta)^\alpha} (\beta(1 + \theta + \theta x)(\theta x)^{\beta-1} - \alpha) e^{-(\theta x)^\beta}, x > 0 \quad (2.43)$$

$$\bar{F}(x) = \left(\frac{1 + \theta + \theta x}{1 + \theta}\right)^\alpha \exp\{-\theta x\}^\beta, x > 0 \quad (2.44)$$

D'après (2.43) et (2.44), la fonction de taux de défaillance (ou d'aléa) est donnée par

$$r(x) = \frac{\beta(1 + \theta + \theta x)\theta^\beta x^{\beta-1} - \theta\alpha}{1 + \theta + \theta x} \quad (2.45)$$

Nous pouvons obtenir la dérivée première de $r(x)$ comme

$$r'(x) = \beta(\beta - 1)\theta^\beta x^{\beta-2} + \frac{\theta^2 \alpha}{(1 + \theta + \theta x)^2}.$$

Il est évident que $r'(x) \leq 0$, pour $\beta \leq 1$ et $\alpha \leq 0$. Plus généralement la fonction $r(x)$ est croissante pour $\alpha > k$ et décroissante pour $\alpha < k$ où

$$k = -\beta(\beta - 1)(\theta x)^{\beta-2}(1 + \theta + \theta x)^2.$$

De plus, pour certaines combinaisons aux valeurs des paramètres, nous avons fonction de taux d'échec uni-modal ou en forme de baignoire. Par conséquent, On conclut que les paramètres de forme α et β influencent la forme de la fonction de taux de défaillance.

- Pour $\beta > 1$,

$$r(0) = f(0) = \frac{-\theta\alpha}{1 + \theta}.$$

Donc, à l'origine $r(x)$ varie continûment avec les paramètres. Ceci est en contraste avec les familles de distribution de Weibull et gamma, où $r(0) = 0$ ou $r(0) = \infty$ pour les deux familles et donc $r(0)$ est discontinues dans les paramètres de telles familles.

- Pour $\beta = 1$,

$$\lim_{x \rightarrow \infty} r(x) = \theta,$$

la fonction $r(x)$ est majorée par θ et continue dans les paramètres du EL distribution, comme la distribution Gamma mais contrairement à celle de Weibull. De plus, dans ce cas, la distribution EL est réduite à un distribution exponentielle.

2.3.1.2 Quantiles, moments et statistiques d'ordre

Le pième quantile x_p de la distribution EL, l'inverse de la fonction de distribution $F(x_p) = p$, est donné par

$$x_p = \left(\frac{\alpha}{\theta^\beta} \ln \left(\frac{1 + \theta + \theta x_p}{(1 + \theta)(1 - p)^{\frac{1}{\alpha}}} \right) \right)^{\frac{1}{\beta}}. \quad (2.46)$$

Beaucoup de caractéristiques intéressantes d'une distribution peuvent être étudiées à travers ses moments.

Soit X une variable aléatoire suivant la loi EL avec des paramètres α , θ et β . Les expressions de l'espérance mathématique, de la variance et des moment centrés à l'origine de X peuvent être obtenues à l'aide de la formule bien connue

$$E[X^r] = r \int_0^\infty x^{r-1} \bar{F}(x) dx$$

Par conséquent, il suit :

$$E[X^r] = \frac{r}{\beta\theta^r} \sum_{j=0}^{\infty} \binom{\alpha}{j} \frac{\Gamma\left(\frac{r+j}{\beta}\right)}{(1+\theta)^j}$$

En particulier, $\alpha = 1$ implique

$$E[X] = \mu = \frac{(1+\theta)\Gamma(1/\beta) + \Gamma(2/\beta)}{\theta\beta(1+\theta)}$$

De plus, pour $\alpha = \beta = 1$ on obtient

$$\mu = \frac{2+\theta}{\theta(1+\theta)},$$

qui est la moyenne de la distribution originale de Lindley. Maintenant, nous discutons de certaines propriétés des statistiques d'ordre de la distribution EL. Les statistiques de commandes sont parmi les outils les plus fondamentaux des statistiques non paramétriques et de l'inférence.

Soit X_1, \dots, X_n un échantillon aléatoire prélevé sur la distribution EL et $X_{1:n}, \dots, X_{n:n}$ désigne les statistiques d'ordre de l'échantillon. Ensuite, la densité de probabilité de la i ème statistique d'ordre $X_{i:n}$ est donné par

$$\begin{aligned} f_{i:n}(x) &= \frac{n!\theta(\beta(1+\theta+\theta x)(\theta x)^{\beta-1} - \alpha)(1+\theta+\theta x)^{(n-i+1)\alpha-1}}{(n-i)!(i-1)!(1+\theta)^{n\alpha}} \\ &\times \left[(1+\theta)^\alpha - (1+\theta+\theta x)^\alpha e^{-(\theta x)^\beta} \right]^{i-1} e^{-(n-i+1)(\theta x)^\beta} \end{aligned}$$

L'évolution de la statistique d'ordre $x_{i:n}$ peut être obtenue à partir du résultat connu

$$E[X_{i:n}^r] = r \sum_{k=n-i+1}^n (-1)^{k-n+i-1} \binom{k-1}{n-i} \binom{n}{k} \int_0^\infty x^{r-1} [\bar{F}(x)]^k dx$$

En substituant $\bar{F}(x)$ donné par (2.44), on obtient

$$E[X_{i:n}^r] = \frac{r}{\beta\theta^r} \sum_{k=n-i+1}^n \frac{(-1)^{k-n+i-1}}{K^{\frac{r}{\beta}}} \binom{k-1}{n-i} \binom{n}{k} \sum_{j=0}^{\infty} \binom{k\alpha}{j} \frac{\Gamma\left(\frac{r+j}{\beta}\right)}{k^{\frac{j}{\beta}}(1+\theta)^j}$$

2.3.1.3 Fonction de durée de vie résiduelle et de taux de défaillance inversé

Étant donné qu'un composant survit jusqu'au temps $t \geq 0$, la durée de vie résiduelle est la période au-delà de t jusqu'au moment de la défaillance et définie par le vide conditionnel variable $x - \frac{t}{x} > t$.

En terme de fiabilité, on sait bien que la fonction de vie différente et la proportion de plusieurs évolutions de la vie individuelle déterminent la répartition (GuptaGupta,1983)[19].

Par conséquent, nous obtenons le mouvement de l'ordre de la vie différente par le biais de la formule générale.

$$\mu_r(t) = E((X - t)^r | X > t) = \frac{1}{\bar{F}(t)} \int_t^\infty r(x - t)^{r-1} \bar{F}(x) dx$$

Postuler (2.41), Le développement binomial à $(x - t)^{r-1}$ et en remplaçant $\bar{F}(\cdot)$ donné par (2.44) dans la formule ci-dessus impliquent

$$\mu_r(t) = \frac{r(1 + \theta)^\alpha e^{(\theta t)^\beta}}{\beta \theta^r (1 + \theta + \theta t)^\alpha} \sum_{j=0}^{r-1} \binom{r-1}{j} (-1)^j (\theta t)^j \sum_{k=0}^{\infty} \binom{\alpha}{k} \frac{\Gamma\left(\frac{r-j+k}{\beta}, (\theta t)^\beta\right)}{(1 + \theta)^k}, r \geq 1.$$

Ainsi, nous obtenons la durée de vie résiduelle de la répartition de la distribution EL comme

$$\mu_1(t) = \mu(t) = \frac{e^{(\theta t)^\beta}}{\beta \theta (1 + \theta + \theta t)^\alpha} \sum_{k=0}^{\infty} \binom{\alpha}{k} (1 + \theta)^{\alpha-k} \Gamma\left(\frac{1+k}{\beta}, (\theta t)^\beta\right) \quad (2.47)$$

En particulier, pour $\alpha = 1$, nous obtenons

$$\mu(0) = E(X) = \frac{(1 + \theta)\Gamma(1/\beta) + \Gamma(2/\beta)}{\theta\beta(1 + \theta)}$$

De plus, si $\alpha = \beta = 1$, alors

$$\mu(t) = \frac{2 + \theta + \theta t}{\theta(1 + \theta + \theta t)}$$

qui est la fonction de durée de vie résiduelle moyenne de la distribution de Lindley d'origine. En outre, le deuxième moment de la durée de vie résiduelle de la distribution EL est

$$\mu_2(t) = \frac{2e^{(\theta t)^\beta}}{\beta \theta^2 (1 + \theta + \theta t)^\alpha} \sum_{k=0}^{\infty} \binom{\alpha}{k} (1 + \theta)^{\alpha-k} \left\{ \Gamma\left(\frac{2+k}{\beta}, (\theta t)^\beta\right) - \theta t \Gamma\left(\frac{1+k}{\beta}, (\theta t)^\beta\right) \right\}.$$

La variance de la durée de vie résiduelle de la distribution EL peut être obtenue facilement en utilisant $\mu_2(t)$ et $\mu(t)$, et par conséquent son coefficient de variation.

D'autre part, nous discutons par analogie de la durée de vie résiduelle inversée et de certaines de ses propriétés. La durée de vie résiduelle inversée peut être définie comme la variable aléatoire conditionnelle $t - X | X \leq t$ qui désigne le temps écoulé à la suite de la défaillance d'un composant, car sa durée de vie est inférieure ou égale à t . Ce phénomène aléatoire peut également s'appeler le temps d'inactivité (ou le temps écoulé depuis l'échec); pour plus de détails que vous pouvez voir (Kundu Nanda, 2010 [20]; Nanda, Singh, Misra et Paul, 2003 [21]). En outre, en termes de fiabilité, les thèmes de la vie différente et la vie inverse et les motivations négatives de la vie différente caractérisent la répartition de manière unique.

En utilisant (2.40) et (2.43), La fonction de taux de défaillance inversée (ou de risque inversé) est donnée par

$$h(x) = \frac{f(x)}{F(x)} = \frac{\theta(1 + \theta + \theta x)^{\alpha-1}(\beta(1 + \theta + \theta x)(\theta x)^{\beta-1} - \alpha)}{(1 + \theta)^\alpha e^{(\theta x)^\beta} - (1 + \theta + \theta x)^\alpha},$$

Il n'est pas établi que $h(0) = \infty$, $h(0)$ est discontinu dans le paramètre de la distribution de EL.

Le mouvement d'ordre de la vie différente peut être obtenu par la forme bien connue

$$m_r(t) = E((t - X)^r | X \leq t) = \frac{1}{F(t)} \int_0^t r(t - x)^{r-1} F(x) dx,$$

Par conséquent

$$m_r(t) = \frac{rt^{r-1}(1 + \theta)^\alpha e^{(\theta t)^\beta}}{(1 + \theta)^\alpha e^{(\theta t)^\beta} - (1 + \theta + \theta t)^\alpha} \times \sum_{k=0}^{r-1} \binom{r-1}{k} \left(\frac{-1}{t}\right)^k \left\{ \frac{t^{1+k}}{1+k} - \frac{1}{\beta \theta^{1+k}} \sum_{i=0}^{\infty} \binom{\alpha}{i} \frac{\gamma\left(\frac{1+i+k}{\beta}, (\theta t)^\beta\right)}{(1 + \theta)^i} \right\}, r \geq 1.$$

Ainsi, la moyenne de la durée de vie résiduelle inversée de la distribution EL est donnée par

$$m_1(t) = \frac{(1 + \theta)^\alpha e^{(\theta t)^\beta}}{(1 + \theta)^\alpha e^{(\theta t)^\beta} - (1 + \theta + \theta t)^\alpha} \left\{ t - \frac{1}{\theta \beta} \sum_{i=0}^{\infty} \binom{\alpha}{i} \frac{\gamma\left(\frac{1+i}{\beta}, (\theta t)^\beta\right)}{(1 + \theta)^i} \right\}$$

et le deuxième mouvement de la vie différente de la répartition de EL est donné par

$$m_2(t) = \frac{t(1 + \theta)^\alpha e^{(\theta t)^\beta}}{(1 + \theta)^\alpha e^{(\theta t)^\beta} - (1 + \theta + \theta t)^\alpha} \times \left\{ t + \frac{2}{\theta \beta} \sum_{i=0}^{\infty} \binom{\alpha}{i} \frac{1}{(1 + \theta)^i} \left[\frac{\gamma\left(\frac{2+i}{\beta}, (\theta t)^\beta\right)}{\theta t} - \gamma\left(\frac{1+i}{\beta}, (\theta t)^\beta\right) \right] \right\}.$$

En utilisant $m_1(t)$ et $m_2(t)$, nous avons obtenu la variation de la durée de vie différente de la distribution de EL, et donc le coefficient de variation de la durée de vie résiduelle inversée de la distribution EL peut être facilement obtenue.

2.3.1.4 Courbes de Bonferroni et de Lorenz

Les courbes de Bonferroni et de Lorenz et l'indice de Gini ont de nombreuses applications non seulement en économie pour étudier le revenu et la pauvreté, mais aussi dans d'autres domaines tels que la fiabilité, la médecine et l'assurance. La courbe de

Bonferroni $B_F[F(x)]$ est donné par

$$B_F[F(x)] = \frac{1}{\mu F(x)} \int_0^x u f(u) du,$$

ou de manière équivalente donnée par

$$B_F(p) = \frac{1}{\mu_p} \int_0^p F^{-1}(t) dt,$$

où $p = F(x)$ et

$$F^{-1}(t) = \inf\{x : F(x) \geq t\}.$$

De la relation entre l'aire de visite et la durée de vie résiduelle donnée par le théorème de (fonction de densité, de survie et de taux d'échec) de Pundir, Arora et Jain (2005) [22], la courbe de Bonferroni de la fonction de distribution F de la distribution EL est donnée par

$$B_F[F(x)] = \frac{1}{1 - \left(\frac{1+\theta+\theta x}{1+\theta}\right)^\alpha e^{-(\theta x)^\beta}} \times \left\{ 1 - \frac{1}{\theta \mu \beta} \sum_{j=0}^{\infty} \binom{\alpha}{j} \frac{\Gamma\left(\frac{1+j}{\beta}, (\theta x)^\beta\right)}{(1+\theta)^j} - \frac{x}{\mu} \left(\frac{1+\theta+\theta x}{1+\theta}\right)^\alpha e^{-(\theta x)^\beta} \right\}.$$

De plus, la courbe de Lorenz de F qui suit la distribution EL peut être obtenue par l'expression suivante :

$$L_F[F(x)] = B_F[F(x)]F(x).$$

Le temps total mis à l'échelle et le temps total cumulé sur la transformation d'essai d'une fonction de distribution F (Pundir et al., 2005 [22]) sont définis par

$$S_F[F(t)] = \frac{1}{\mu} \int_0^t \bar{F}(u) du$$

et

$$C_F = \int_0^1 S_F[F(t)] f(t) dt,$$

respectivement. Si $F(t)$ est la fonction de distribution EL spécifiée par (2.40) alors en utilisant la formule (2.42),

$$S_F[F(t)] = \frac{1}{\theta \mu} \sum_{j=0}^{\infty} \binom{\alpha}{j} \frac{1}{(1+\theta)^j} \sum_{k=0}^{\infty} \frac{(-1)^k (\theta t)^{1+k\beta+j}}{k! (1+k\beta+j)}$$

et

$$C_F = \frac{1}{\theta\mu} \sum_{j=0}^{\infty} \binom{\alpha}{j} \frac{1}{(1+\theta)^j} \sum_{k=0}^{\infty} \frac{(-1)^k}{k!(1+k\beta+j)}$$

$$\times \left\{ \sum_{i=0}^{\infty} \binom{\alpha}{i} \frac{\gamma\left(\frac{1+i+j+\beta+k\beta}{\beta}, \theta^\beta\right)}{(1+\theta)^i} - \frac{\alpha}{\beta} \sum_{i=0}^{\infty} \binom{\alpha-1}{i} \frac{\gamma\left(\frac{2+i+j+k\beta}{\beta}, \theta^\beta\right)}{(1+\theta)^i} \right\}.$$

L'indice de Gin peut être obtenu à partir de la relation $G = 1 - C_F$.

2.3.1.5 Écart moyen

La quantité de dispersion dans une population peut être mesurée par la totalité des écarts par rapport à la moyenne et à la médiane.

Pour une variable aléatoire X , $f(x)$, de fonction de distribution $F(x)$, de moyenne $\mu = E(X)$ et $M = Median(X)$, l'écart moyen par rapport au moyen et à l'éloignement par rapport au médian, respectivement, est défini par

$$\delta_1(X) = \int_0^{\infty} |x - \mu| f(x) dx = 2\mu F(\mu) - 2\mu + 2 \int_{\mu}^{\infty} x f(x) dx$$

et

$$\delta_2(X) = \int_0^{\infty} |x - M| f(x) dx = 2MF(M) - M - \mu + 2 \int_M^{\infty} x f(x) dx.$$

Si X est une variable aléatoire EL spécifiée par (2.43) puis en utilisant la formule (2.41),

$$\delta_1(X) = 2\mu F(\mu) - 2\mu + 2L(\mu)$$

et

$$\delta_2(X) = 2MF(M) - M - \mu + 2L(M).$$

Où

$$L(b) = \frac{1}{\theta} \sum_{j=0}^{\infty} \binom{\alpha}{j} \frac{\Gamma\left(\frac{1+\beta+j}{\beta}, (\theta b)^\beta\right)}{(1+\theta)^j} - \frac{\alpha}{\theta\beta} \sum_{j=0}^{\infty} \binom{\alpha-1}{j} \frac{\Gamma\left(\frac{2+j}{\beta}, (\theta b)^\beta\right)}{(1+\theta)^{1+j}}.$$

2.3.2 Estimation et inférence

Dans cette section, nous considérons l'estimation du maximum de vraisemblance et fournissons des expressions pour le Fisher observé associé matrice d'informations. En supposant que les durées de vie sont distribuées indépendamment, Les estimations du maximum de vraisemblance (MLE) des paramètres sont obtenus par maximisation

directe de la fonction log-vraisemblance donnée par,

$$\begin{aligned} \log L(\alpha, \beta, \theta) &= n(\log(\theta) - \alpha \log(1 + \theta)) + \sum_{i=1}^n \alpha \log(1 + \theta + \theta x_i) \\ &+ \sum_{i=1}^n \log(\beta(1 + \theta + \theta x_i)(\theta x_i)^{\beta-1} - \alpha) - \sum_{i=1}^n (\theta x_i)^\beta. \end{aligned} \quad (2.48)$$

Il s'ensuit que les estimâtes du maximum de vraisemblance (MLE), disons $\hat{\alpha}$, $\hat{\beta}$, et $\hat{\lambda}$, sont les solutions simultanées des équations,

$$\begin{aligned} \sum_{i=1}^n (\beta(1 + \theta + \theta x_i)(\theta x_i)^{\beta-1} - \alpha)^{-1} &= -n \log(1 + \theta) + \sum_{i=1}^n \log(1 + \theta + \theta x_i), \\ \sum_{i=1}^n (\theta x_i)^\beta \log(\theta x_i) &= \sum_{i=1}^n \frac{(1 + \theta + \theta x_i)(\theta x_i)^{\beta-1} (1 + \beta \log(\theta x_i))}{\beta(1 + \theta + \theta x_i)(\theta x_i)^{\beta-1} - \alpha} \end{aligned}$$

et

$$\begin{aligned} \beta \theta^{-1} \sum_{i=1}^n (\theta x_i)^\beta &= n \left(\frac{1}{\theta} - \frac{\alpha}{(1 + \theta)} \right) + \sum_{i=1}^n \frac{\alpha(1 + x_i)}{(1 + \theta + \theta x_i)} \\ &+ \sum_{i=1}^n \frac{\beta(\theta x_i)^{\beta-1} (1 + x_i + (1 + \theta + \theta x_i)\theta^{-1}(\beta - 1))}{\beta(1 + \theta + \theta x_i)(\theta x_i)^{\beta-1} - \alpha}. \end{aligned}$$

Pour l'estimation d'intervalle de (α, β, θ) , nous considérons la matrice d'information de Fisher observée donnée par,

$$I_F(\alpha, \beta, \theta) = - \begin{pmatrix} I_{\alpha\alpha} & I_{\alpha\beta} & I_{\alpha\theta} \\ I_{\beta\alpha} & I_{\beta\beta} & I_{\beta\theta} \\ I_{\theta\alpha} & I_{\theta\beta} & I_{\theta\theta} \end{pmatrix}_{(\alpha, \beta, \theta) = (\hat{\alpha}, \hat{\beta}, \hat{\theta})}$$

où l'élément de la matrice $I_F(\alpha, \beta, \theta)$ sont donnés. Dans les conditions qui sont remplies pour les paramètres α, β et θ à l'intérieur de l'espace des paramètres, la distribution asymptotique de $(\hat{\alpha} - \alpha, \hat{\beta} - \beta, \hat{\theta} - \theta)$ comme $n \rightarrow \infty$, est une 3-variable normale avec une moyenne nulle et une matrice de covariance $I_F^{-1}(\alpha, \beta, \theta)$

$$\begin{aligned} I_{\alpha\alpha} &= - \sum_{i=1}^n (\beta T_{lbi} - \alpha)^{-2} \\ I_{\alpha\beta} = I_{\beta\alpha} &= \sum_{i=1}^n \frac{T_{lbi} + \beta T_{lbi} \log(\theta x_i)}{(\beta T_{lbi} - \alpha)^2}, \end{aligned}$$

$$I_{\alpha\theta} = I_{\theta\alpha} = -\frac{n}{1+\theta} + \sum_{i=1}^n \frac{1+x_i}{T_{li}} + \sum_{i=1}^n \frac{\beta T_{bi}(1+x_i+T_{li}(\beta-1)\theta^{-1})}{(\beta T_{lbi}-\alpha)^2}$$

$$I_{\beta\beta} = \sum_{i=1}^n \left\{ \frac{T_{lbi}(2\log(\theta x_i) + \beta(\log(\theta x_i))^2)}{\beta T_{lbi}-\alpha} - \frac{(T_{lbi} + \beta T_{lbi}\log(\theta x_i))^2}{(\beta T_{lbi}-\alpha)^2} \right\} - \sum_{i=1}^n (\theta x_i)^\beta (\log(\theta x_i))^2,$$

$$I_{\beta\theta} = I_{\theta\beta} = \sum_{i=1}^n \left\{ ((T_{bi}(1+x_i) + T_{lbi}(\beta-1)\theta^{-1})(1 + \beta\log(\theta x_i)) + \beta T_{lbi}\theta^{-1})(\beta T_{lbi}-\alpha)^{-1} - T_{lbi}(1 + \beta\log(\theta x_i)) \right. \\ \left. - \sum_{i=1}^n \frac{(\theta x_i)^\beta (1 + \beta\log(\theta x_i))}{\theta} \right\},$$

et

$$I_{\theta\theta} = n \left(\frac{-1}{\theta^2} + \frac{\alpha}{(1+\theta)^2} \right) - \sum_{i=1}^n \frac{\alpha(1+x_i)^2}{T_{li}^2} \\ + \sum_{i=1}^n \left\{ (2\beta(1+x_i)T_{bi}(\beta-1)\theta^{-1} + \beta T_{lbi}(\beta-1)\theta^{-2}(\beta-2))(\beta T_{lbi}-\alpha)^{-1} \right. \\ \left. - (\beta T_{bi})^2(1+x_i + \beta T_{li}(\beta-1)\theta^{-1})^2(\beta T_{lbi}-\alpha)^{-2} - \sum_{i=1}^n ((\theta x_i)^\beta \beta^2 \theta^{-2} - \beta(\theta x_i)^\beta \theta^{-2}) \right\},$$

(2.50)

Où

$$T_{li} = (1 + \theta + \theta x_i), T_{bi} = (\theta x_i)^{\beta-1}$$

et

$$T_{lbi} = T_{li} T_{bi}.$$

2.4 Distribution de Lindley Généralisée

La distribution de Lindley Généralisée (Généralized Lindley Distribution (GLD)) avec les paramètres α et θ est définie par sa fonction de densité de probabilité

$$f(x; \alpha; \theta; \beta) = \frac{\theta^2(\theta x)^{\alpha-1}(\alpha + \beta x)e^{-\theta x}}{(\theta + \beta)\Gamma(\alpha + 1)}; x, \theta, \alpha, \beta > 0. \quad (2.51)$$

Il est facile de voir que si $\alpha = \beta = 1$, la fonction (2.51) de GLD se réduit à la fonction de distribution de LD (2.1), si $\beta = 0$, elle se réduit à la distribution Gamma (α, θ) et si $(\alpha, \beta) = (1, 0)$, elle se réduit à la distribution Exponentielle (θ) . La fonction de densité (2.51) de GLD peut être montrée sous forme de mélange des distributions Gamma (α, θ) et Gamma $(\alpha + 1, \theta)$ comme suit :

$$f(x; \alpha; \theta; \beta) = p_1 f_1(x) + p_2 f_2(x)$$

Où $p_1 = \frac{\theta}{\theta + \beta}$, $p_2 = \frac{\beta}{\theta + \beta}$, $f_1(x) = \frac{\theta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\theta x}$ et $f_2(x) = \frac{\theta^{\alpha+1}}{\Gamma(\alpha+1)} x^\alpha e^{-\theta x}$.

La fonction génératrice correspondante est :

$$M(t) = \left(\frac{\theta}{\theta - t} \right)^{\alpha+1} \left(\frac{\theta - t + \beta}{\theta + \beta} \right)$$

Les moments d'ordre 1, 2 et 3 de la distribution de Lindley Généralisée sont :

$$\begin{aligned} \mu'_1 &= -\frac{1}{\theta + \beta} + \frac{\alpha + 1}{\theta}, \\ \mu'_2 &= -\frac{2(\alpha + 1)}{\theta(\theta + \beta)} + \frac{\alpha^2 + 3\alpha + 2}{\theta^2}, \\ \mu'_3 &= \frac{\alpha + 1}{\theta^2}(-3\alpha + 6) + \frac{(\alpha + 1)(\theta + \beta)}{\theta} \left[(\alpha + 1)^2 + \frac{3\alpha + 5}{\theta^2} \right] \end{aligned}$$

La fonction de log-vraisemblance de la distribution de Lindley Généralisée est :

$$\begin{aligned} \ln L(x; \alpha; \theta; \beta) &= n(\alpha + 1) \ln \theta - n \ln(\beta + \theta) - n \ln \Gamma(\alpha + 1) \\ &+ (\alpha - 1) \sum_{i=0}^n \ln(x_i) + \sum_{i=0}^n \ln(\alpha + \beta x_i) - n\theta \bar{X} \end{aligned}$$

Estimation

2.4.1 Estimation par la méthode des moments

Étant donné un échantillon aléatoire X_1, \dots, X_n de taille n de la distribution GLD (2.51), Les estimâtes des moments (MoM), $\hat{\beta}$ de β , $\hat{\theta}$ de θ et $\hat{\alpha}$ de α sont la solution des

équations des moments :

$$\begin{aligned} \alpha(\theta + \beta) + \beta - \bar{X}\theta(\theta + \beta) &= 0 \\ -2(\alpha + 1)\theta + (\alpha^2 + 3\alpha + 2)(\theta + \beta) - n\theta^2(\theta + \beta) &= 0 \\ \frac{\alpha + 1}{\theta^2}(-3\alpha + 6) + \frac{(\alpha + 1)(\theta + \beta)}{\theta} \left[(\alpha + 1)^2 + \frac{3\alpha + 5}{\theta^2} \right] - k &= 0 \end{aligned}$$

Il est possible de calculer $\hat{\beta}, \hat{\theta}$ et $\hat{\alpha}$ numériquement. mais en utilisant \bar{X} , $n = \mu'_2$ et $k = \mu'_3$.

2.4.2 Estimations du maximum de vraisemblance

Nous considérons l'estimation de maximum de vraisemblance (MLE) des paramètres, Si X_1, \dots, X_n , est un échantillon aléatoire de X distribué selon $GL(\alpha, \theta, \beta)$, alors la fonction log-vraisemblance, $l(\alpha, \theta, \beta)$ est :

$$l(\alpha, \theta, \beta) = n(\alpha + 1)\log(\theta) - n\log(\beta + \theta) - n\log\Gamma(\alpha + 1) + (\alpha - 1) \sum_{i=1}^n \log(x_i) + \sum_{i=1}^n \log(\alpha + \beta x_i) - \theta \sum_{i=1}^n x_i.$$

Les dérivés de $l(\alpha, \theta, \beta)$ par rapport à α , θ et β sont :

$$\frac{\partial l}{\partial \alpha} = n\log(\theta) - n\Psi(\alpha + 1) + \sum_{i=1}^n \log(x_i) + \sum_{i=1}^n \frac{1}{\alpha + \beta x_i} \quad (2.52)$$

$$\frac{\partial l}{\partial \theta} = \frac{n(\alpha + 1)}{\theta} - \frac{n}{\beta + \theta} - \sum_{i=1}^n x_i, \quad (2.53)$$

$$\frac{\partial l}{\partial \beta} = \sum_{i=1}^n \frac{x_i}{\alpha + \beta x_i} - \frac{n}{\beta + \theta}, \quad (2.54)$$

Où

$$\Psi(t) = \frac{\Gamma'(t)}{\Gamma(t)}$$

, désigne la fonction digamma. Les équations (2.52) et (2.54) peuvent être résolus simultanément pour trouver les estimâtes de maximum de vraisemblance de α, θ et β . La distribution GL satisfait toutes les conditions de régularité (voir, Bain [1, pp.86-87]) d'une manière similaire à la distribution Gamma, et donc en appliquant l'approximation habituelle sur grand échantillon, les estimâtes $(\hat{\alpha}, \hat{\theta}, \hat{\beta})$ traité comme étant approximativement bivarié normal avec le vecteur moyen (α, θ, β) et matrice de variance-covariance I^{-1} , Où I est la matrice d'information de Fisher, dont les éléments sont

données par

$$\begin{aligned}
 -E\left(\frac{\partial^2 l}{\partial \alpha^2}\right) &= n\Psi'(1 + \alpha) + nJ_0(\alpha, \theta, \beta), \\
 -E\left(\frac{\partial^2 l}{\partial \alpha \partial \theta}\right) &= -\frac{n}{\theta} \\
 -E\left(\frac{\partial^2 l}{\partial \alpha \partial \beta}\right) &= nJ_1(\alpha, \theta, \beta), \\
 -E\left(\frac{\partial^2 l}{\partial \theta^2}\right) &= \frac{n(\alpha + 1)}{\theta^2} - \frac{n}{(\beta + \theta)^2}, \\
 -E\left(\frac{\partial^2 l}{\partial \theta \partial \beta}\right) &= -\frac{n}{(\beta + \theta)^2}, \\
 -E\left(\frac{\partial^2 l}{\partial \beta^2}\right) &= nJ_2(\alpha, \theta, \beta) - \frac{n}{(\beta + \theta)^2}
 \end{aligned}$$

Pour $i=0,1,2$,

$$\begin{aligned}
 J_i(\alpha, \theta, \beta) &= E\left(\frac{X^i}{(\alpha + \beta X)^2}\right) \\
 &= \frac{\theta^{1-i}}{\alpha\Gamma(\alpha + 1)(\beta + \theta)} \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \left(\frac{\beta}{\alpha\theta}\right)^k \Gamma(\alpha + k + i),
 \end{aligned}$$

Où X distribué comme $GL(\alpha, \theta, \beta)$.

(pour plus de détails, voir Hojatollah Zakerzadeh, Dolati Ali (2009)[23])

Introduction

Pour étudier l'utilité de la loi de Lindley, on va présenter dans ce chapitre la simulation des variables aléatoires de Lindley avec les trois catégories, un seul paramètre, deux paramètres, et trois paramètres. En suites on va présenter des applications sur des jeux de données réelles et on va essayez d'ajuster ces données par des distributions de Lindley.

3.1 Simulation des v. a. de Lindley

La simulation des variables aléatoires de Lindley est réalisé avec le package LindleyR du logiciel R. Le package LindleyR calcule la densité de probabilité, la distribution cumulative, le quantile et les fonctions de taux de risque et génère des écarts aléatoires à partir de la distribution de Lindley discrète et continue ainsi que pour 19 de ses modifications. Il génère également des écarts aléatoires censurés à partir de toute distribution de probabilité disponible dans R. Les distributions de Lindley, unimodales et exponentielles peuvent être utilisées comme distributions de censure.

Pour un échantillon de taille $n = 1000$ d'une v.a. de Lindley avec un seul paramètre $\theta = 1.5$, on obtient la figure (3.1) qui représente, l'histogramme, la fonction de répartition, le quantile et la fonction de survie

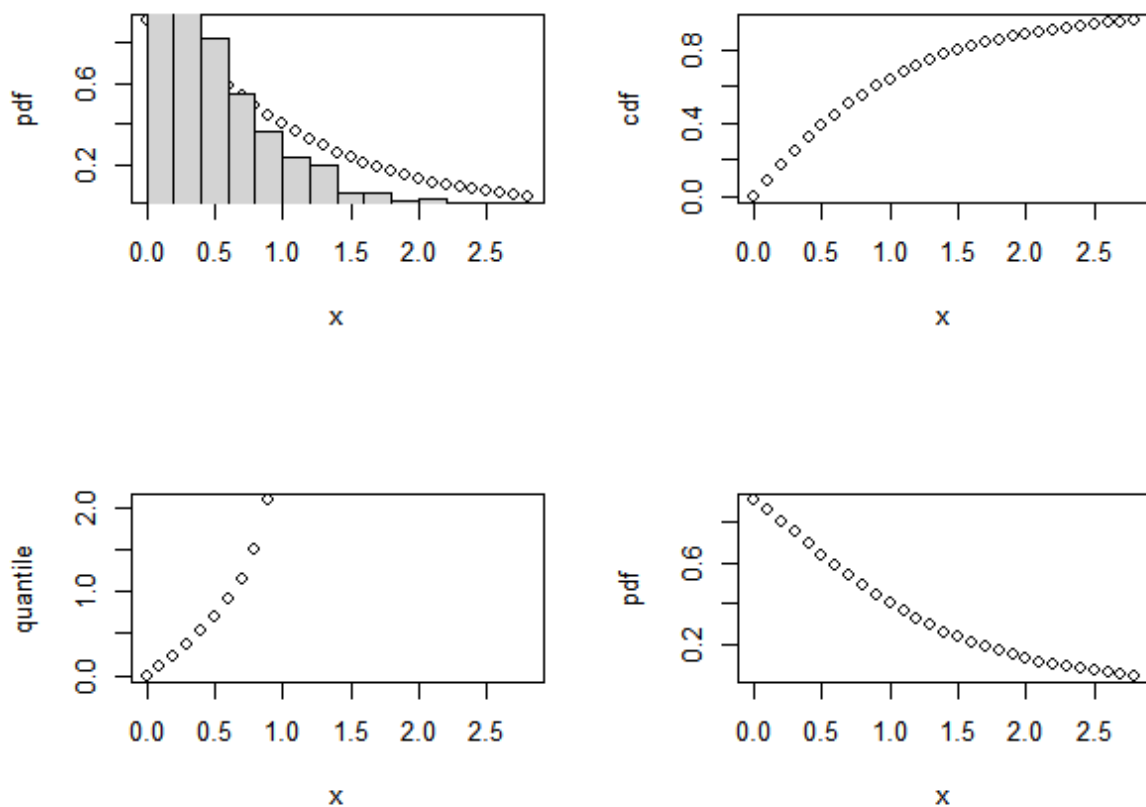


FIGURE 3.1 – L’histogramme, la fonction de répartition, le quantile et la fonction de survie de loi lindley a une seul paramètre

Estimation du paramètre :

L’estimateur des paramètres par la méthode des moments (MoM) de θ :

$$\hat{\theta}_{MoM} = \frac{-(\bar{X} - 1) + \sqrt{(\bar{X} - 1)^2 + 8\bar{X}}}{2\bar{X}}, \quad \bar{X} > 0.$$

alors pour notre échantillon

$$\hat{\theta} = 1.447087$$

Pour un échantillon de taille $n = 1000$ d’une v.a. de Lindley avec deux paramètres, $\theta = 1.5$ et $\alpha = 1.5$, on obtient la figure (3.2) qui représente, l’histogramme, la fonction de répartition, le quantile et la fonction de survie

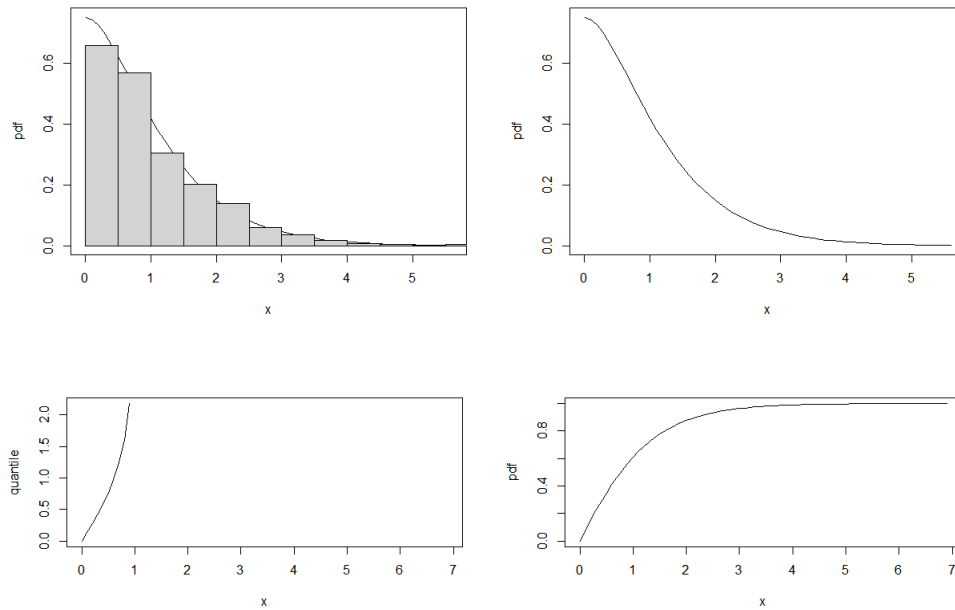


FIGURE 3.2 – L’histogramme, la fonction de répartition, le quantile et la fonction de survie de loi lindley a deux paramètres

Estimation à deux paramètres :

Les estimations de α et θ sont obtenues comme :

$$\hat{\alpha} = \left(\frac{b+2}{b(b+1)} \right) \frac{1}{\bar{X}}$$

$$\hat{\theta} = \left(\frac{b+2}{b+1} \right) \frac{1}{\bar{X}}$$

alors pour notre échantillon, on a :

pour $b = 1$

$$\hat{\alpha} = \hat{\theta} = 1.490856$$

3.2 Applications sur des données réelles

Dans cette section, nous utilisons deux ensembles de données réelles pour montrer que la distribution beta de Lindley peut être un meilleur modèle que celui basé sur la distribution de Lindley.

Ensemble de données 1 : Les données suivantes représentent les durées de survie (en jours) de 72 cobayes infectés par bacilles tuberculeux virulents, observés et rapportés par Bjerkedal [2]. Les données sont les suivantes : 0.1, 0.33, 0.44, 0.56, 0.59,

0.72, 0.74, 0.77, 0.92, 0.93, 0.96, 1, 1, 1.02, 1.05, 1.07, 1.07, 1.08, 1.08, 1.08, 1.09, 1.12, 1.13, 1.15, 1.16, 1.2, 1.21, 1.22, 1.22, 1.24, 1.3, 1.34, 1.36, 1.39, 1.44, 1.46, 1.53, 1.59, 1.6, 1.63, 1.63, 1.68, 1.71, 1.72, 1.76, 1.83, 1.95, 1.96, 1.97, 2.02, 2.13, 2.15, 2.16, 2.22, 2.3, 2.31, 2.4, 2.45, 2.51, 2.53, 2.54, 2.54, 2.78, 2.93, 3.27, 3.42, 3.47, 3.61, 4.02, 4.32, 4.58, 5.55,

Les données (waitingtimes) sont disponibles dans le package (Fildistr) donc on va l'échantillon de paramètre discription

Min	1st Qu	Médiane	Moyenne	3rd Qu	Max
0.100	1.080	1.495	1.768	2.240	5.550

TABLE 3.1 – Les estimations des données de durée de survie (en jours) de 72 cobayes infectés par des bacilles tuberculeux virulents

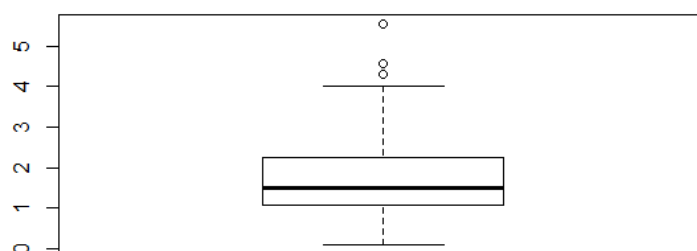


FIGURE 3.3 – Graphique des paramètres descriptif des données des durées survie de cobayes infectés par des bacilles tuberculeux virulents

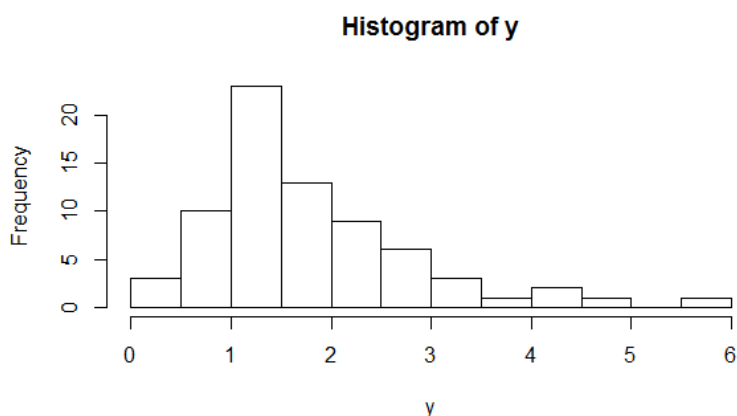


FIGURE 3.4 – Histogramme des paramètres descriptif des données des durées survie de cobayes infectés par des bacilles tuberculeux virulents

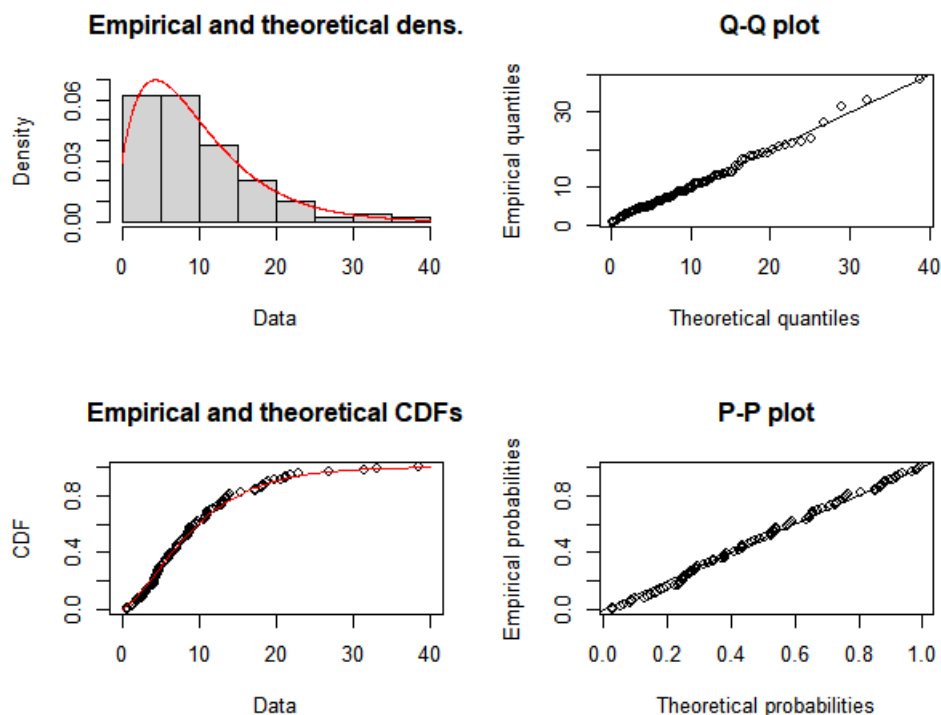


FIGURE 3.5 – Densités estimées, Densités cumulées estimées, Graphiques p-p pour Lindley à un paramètre ajusté pour l'ensemble de données 1

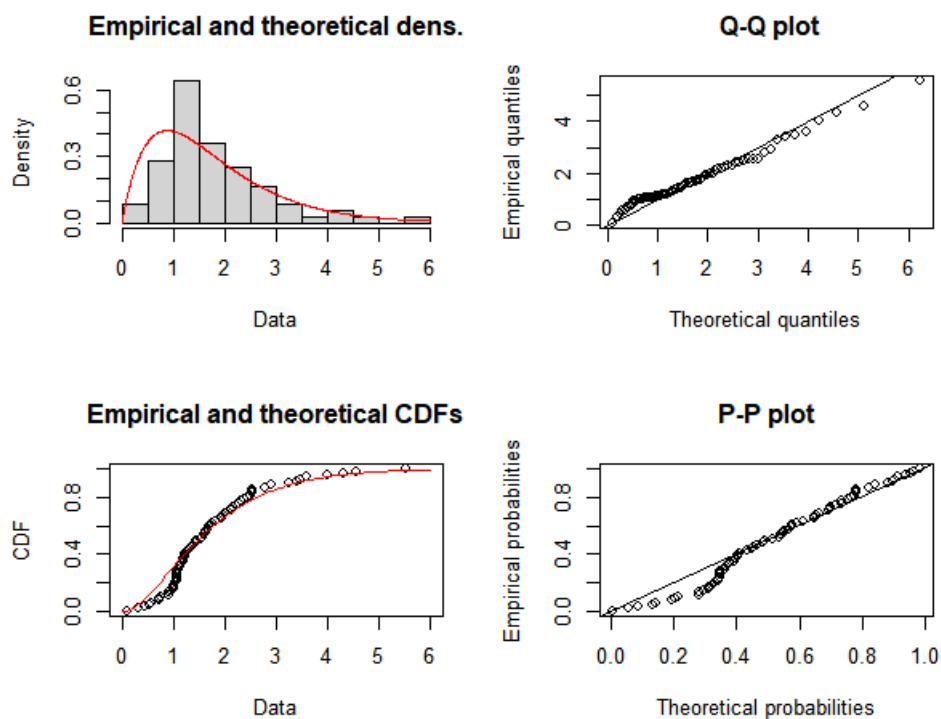


FIGURE 3.6 – Densités estimées, Densités cumulées estimées, Graphiques p-p, Graphiques Q-Q pour Lindley à deux paramètre ajusté pour l'ensemble de données 1

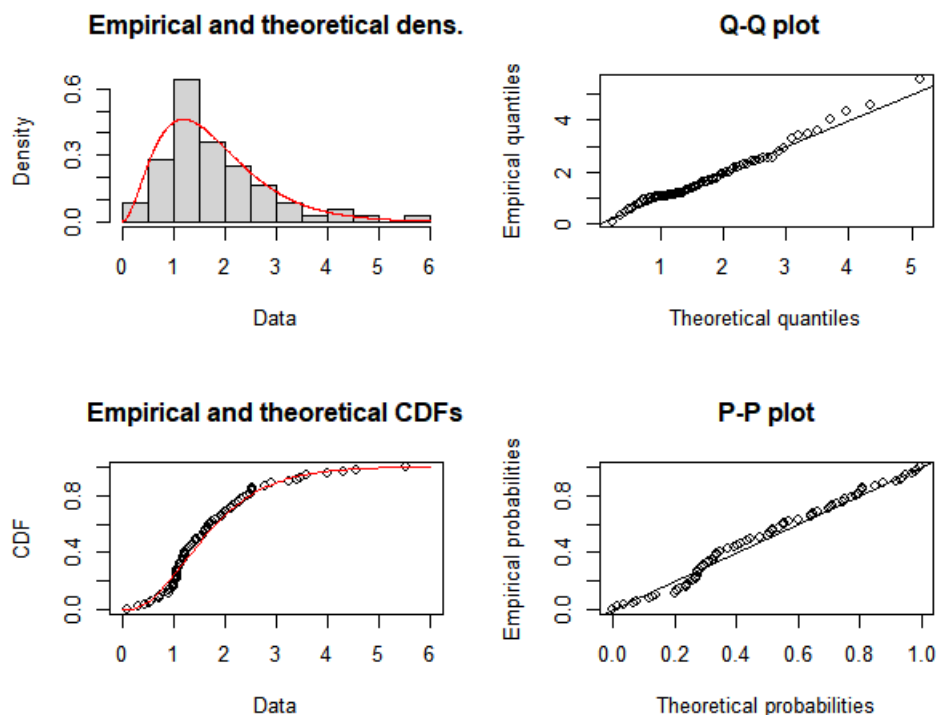


FIGURE 3.7 – Densités estimées, Densités cumulées estimées, Graphiques p-p, Graphiques Q-Q pour Lindley Généralisé ajusté pour l’ensemble de données 1

Afin de comparer les trois modèles de distribution, nous considérons des critères comme, AIC (Akaike critère d’information), BIC (critère d’information bayésien) et K-S (test de Kolmogorov-Smirnov), C-V (cramer-von Mises) et A-D (Anderson-Darling) pour le jeu de données. La meilleure distribution correspond aux petits, AIC, BIC, K-S, C-V, et Valeurs A-D :

$$AIC = 2K - 2l,$$

$$BIC = K * \log(n) - 2l, \quad \text{et} \quad K - S = \sup_x |F_n(x) - F(x)|$$

$$V = \sqrt{\frac{x^2/n}{\min(I,J) - 1}}, \quad \text{et} \quad A^2 = n \int_{-\infty}^{+\infty} \frac{(F_n(x) - F(x))^2}{F(x)(1 - F(x))} dF(x)$$

où

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n I_{x_i \leq x}$$

est la fonction de distribution empirique, $F(x)$ est la fonction de distribution cumulative, K est le nombre de paramètres dans le modèle statistique, n la taille de l’échantillon et l est la valeur maximisée de la fonction log-vraisemblance sous le modèle

considéré.

Modèle	AIC	BIC	K-S	C-V	A-D
Lindley1	640.0748	642.6800	0.6767904	0.5818715	0.48631191
Lindley2	199.0483	203.6016	0.1680613	0.3051425	1.8457649
GenLindley	194.1786	201.0086	0.08884199	0.09709719	0.58927033

TABLE 3.2 – AIC,BIC,K-S,C-V et A-D des modèles basés sur le jeu de données 1

Modèle	ML Estimations	Std.Erreur
Lindley1	$\hat{\theta}=0.1865717$	0.01327453
Lindley2	$\hat{\theta}=1.131085$ $\hat{\alpha}=8.371976$	0.09425701 NaN
GenLindley	$\hat{\theta}=1.824644$ $\hat{\alpha}=2.336106$ $\hat{\beta}=14.766665$	0.3428056 0.7679767 44.1908296

TABLE 3.3 – L’ajustement de la distribution ‘Lindley1 , Lindley2’et ‘GenLindley par des paramètres de maximum de vraisemblance d’ensemble des données 1



D’après le tableau (3.2) le meilleur ajustement est la distribution de Lindley Généralisée à 3 paramètres de l’ensemble des données 1.

Ensemble de données 2 : L’ensemble de données donné dans le tableau 1 représente un ensemble de données non censurées correspondant à la rémission fois (en mois) d’un échantillon aléatoire de 128 patients atteints de cancer de la vessie rapportés dans Lee et wang [12]

TABLE 3.4 – Les temps de rémission (en mois) des patients atteints de cancer de la vessie

0.08	2.09	3.48	4.87	6.94	8.66	13.11	23.63	0.20	2.23
0.52	4.98	6.97	9.02	13.29	0.40	2.26	3.57	5.06	7.09
0.22	13.80	25.74	0.50	2.4	6 3.64	5.09	7.26	9.47	14.24
0.82	0.51	2.54	3.70	5.17	7.28	9.74	14.76	26.31	0.81
0.62	3.82	5.32	7.32	10.06	14.77	32.15	2.64	3.88	5.32
0.39	10.34	14.83	34.26	0.90	2.69	4.18	5.34	7.59	10.66
0.96	36.66	1.05	2.69	4.23	5.41	7.62	10.75	16.62	43.01
0.19	2.75	4.26	5.41	7.63	17.12	46.12	1.26	2.83	4.33
0.66	11.25	17.14	79.05	1.35	2.87	5.62	7.87	11.64	17.36
0.40	3.02	4.34	5.71	7.93	11.79	18.10	1.46	4.40	5.85
0.26	11.98	19.13	1.76	3.25	4.50	6.25	8.37	12.02	2.02
0.31	4.51	6.54	8.53	12.03	20.28	2.02	3.36	6.76	12.07
0.73	2.07	3.36	6.93	8.65	12.63	22.69	5.49		

Les données (waitingtimes) sont disponibles dans le package (Fildistr) donc on va l’échantillon de paramètres discription

Min	1st Qu	Médiane	Moyenne	3rd Qu	Max
0.080	2.615	5.410	9.037	11.348	79.050

TABLE 3.5 – Les estimations des données des temps de rémission (en mois) des patients atteints de cancer de la vessie

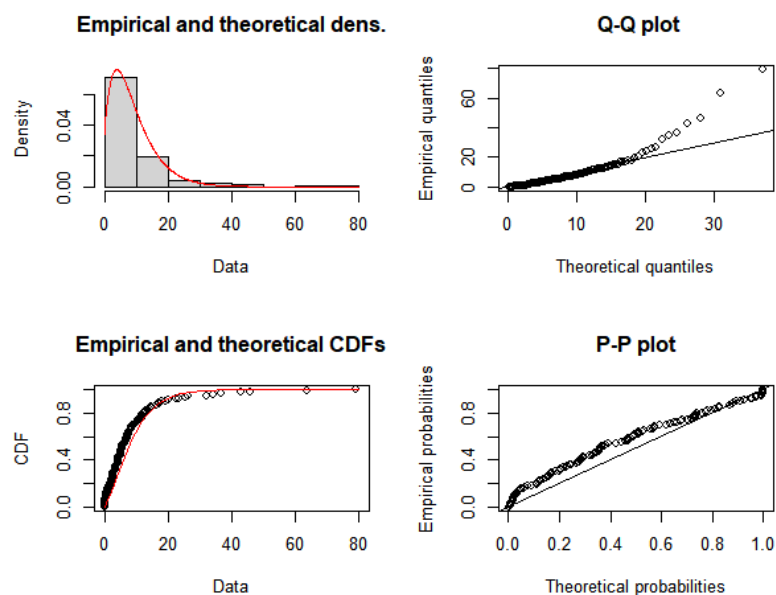


FIGURE 3.8 – Densités estimées, Densités cumulées estimées, Graphiques p-p pour Lindley à un paramètre ajusté pour l'ensemble de données 2

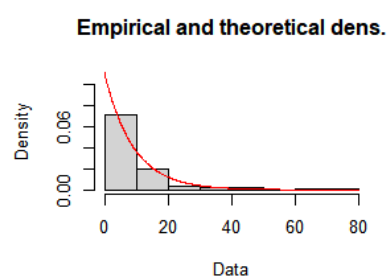


FIGURE 3.9 – Densités estimées des modèles pour l'ensemble de données 2

Modèle	AIC	BIC	K-S	C-V	A-D
Lindley1	857.5144	860.3664	0.1484062	1.0344224	7.3987830
Lindley2	823.5407	829.2447	0.07198795	0.14966378	0.98896423

TABLE 3.6 – AIC,BIC,K-S,C-V et A-D des modèles basés sur le jeu des données 2

Modèle	ML Estimations	Std.Erreur
Lindley1	$\hat{\theta}=0.2026642$	0.01275709
Lindley2	$\hat{\theta}=1.106947$ $\hat{\alpha}=1.368161$	0.01556052 0.01210723

TABLE 3.7 – Ajustement de la distribution 'Lindley1 , Lindley2' par des paramètres de maximum de vraisemblance d'ensemble des données 2

♠

On peut voir dans cet exemple que le LD à deux paramètres donne des ajustements beaucoup plus proches que le LD à un paramètre et offre ainsi une meilleure alternative au LD à un paramètre pour la modélisation données sur les temps d'attente et de survie.

CONCLUSION GÉNÉRALE

Dans ce travail, nous avons étudié trois distributions de Lindley, Lindley à un seul paramètre, Lindley à deux paramètres et Lindley à trois paramètres. On caractérise ces distributions, estimation de ses paramètres par la méthode des moments et la méthode de vraisemblance et on a simulé les variables aléatoires avec un estimation de paramètre et application sur des données réelles ou bien on a ajusté des données réelles par la distribution de Lindley, on a trouver généralement Lindley Généralisé le plus adéquat pour les deux jeux de temps.

BIBLIOGRAPHIE

- [1] Dennis V Lindley. Fiducial distributions and bayes' theorem. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 102–107, 1958.
- [2] Mohamed E Ghitany, Barbra Atieh, and Saralees Nadarajah. Lindley distribution and its application. *Mathematics and computers in simulation*, 78(4) :493–506, 2008.
- [3] Jean-Yves Dauxois, Agathe Guilloux, and Syed NUA Kirmani. Estimation in a competing risks proportional hazards model under length-biased sampling with censoring. *Lifetime data analysis*, 20(2) :276–302, 2014.
- [4] Gérard Frugier and Louis Gacogne. Hachette technique. mathématiques générales. bts. dut. tertiaire.
- [5] François-Xavier Lejeune. *Histogramme, régressogramme et polygone de fréquences en temps continu*. PhD thesis, Université Pierre et Marie Curie-Paris VI, 2007.
- [6] Malcolm R Leadbetter, Georg Lindgren, and Holger Rootzén. *Extremes and related properties of random sequences and processes*. Springer Science & Business Media, 2012.
- [7] Moshe Shaked and J George Shanthikumar. The multivariate hazard construction. *Stochastic Processes and Their Applications*, 24(2) :241–258, 1987.
- [8] Farzana Hussain, Mehdi Hojjati, Masami Okamoto, and Russell E Gorga. Polymer-matrix nanocomposites, processing, manufacturing, and application : an overview. *Journal of composite materials*, 40(17) :1511–1575, 2006.
- [9] Munuswamy Sankaran. 275. note : The discrete poisson-lindley distribution. *Biometrics*, pages 145–149, 1970.
- [10] E Mahmoudi and H Zakerzadeh. Generalized poisson–lindley distribution. *Communications in Statistics—Theory and Methods*, 39(10) :1785–1798, 2010.

- [11] Francisco Louzada, Mari Roman, and Vicente G Cancho. The complementary exponential geometric distribution : Model, properties, and a comparison with its counterpart. *Computational Statistics & Data Analysis*, 55(8) :2516–2524, 2011.
- [12] 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.
- [13] Rameshwar D Gupta and Debasis Kundu. Theory & methods : Generalized exponential distributions. *Australian & New Zealand Journal of Statistics*, 41(2) :173–188, 1999.
- [14] Saralees Nadarajah and Samuel Kotz. The exponentiated type distributions. *Acta Applicandae Mathematica*, 92(2) :97–111, 2006.
- [15] Hassan S Bakouch, Bander M Al-Zahrani, Ali A Al-Shomrani, Vitor AA Marchi, and Francisco Louzada. An extended lindley distribution. *Journal of the Korean Statistical Society*, 41(1) :75–85, 2012.
- [16] Robert M Smith and Lee J Bain. An exponential power life-testing distribution. *Communications in Statistics-Theory and Methods*, 4(5) :469–481, 1975.
- [17] Yan Y Kagan and F Schoenberg. Estimation of the upper cutoff parameter for the tapered pareto distribution. *Journal of Applied Probability*, 38(A) :158–175, 2001.
- [18] Gauss M Cordeiro, Edwin MM Ortega, Božidar V Popović, and Rodrigo R Pescim. The lomax generator of distributions : Properties, minification process and regression model. *Applied Mathematics and Computation*, 247 :465–486, 2014.
- [19] HP Singh and SM Gorey. A remark on gupta, gupta and singh optional randomized response model. *Journal of Applied Mathematics, Statistics and Informatics*, 15(1) :43–73, 2019.
- [20] Chanchal Kundu and Asok K Nanda. Some reliability properties of the inactivity time. *Communications in Statistics—Theory and Methods*, 39(5) :899–911, 2010.
- [21] Asok K Nanda, Harshinder Singh, Neeraj Misra, and Prasanta Paul. Reliability properties of reversed residual lifetime. *Communications in Statistics-Theory and Methods*, 32(10) :2031–2042, 2003.
- [22] Sudesh Pundir, Sangeeta Arora, and Kanchan Jain. Bonferroni curve and the related statistical inference. *Statistics & probability letters*, 75(2) :140–150, 2005.
- [23] Hojatollah Zakerzadeh and Ali Dolati. Generalized lindley distribution. 2009.