

الجمهورية الجزائرية الديمقراطية الشعبية
وزارة التعليم العالي والبحث العلمي

UNIVERSITE BADJI MOKHTAR - ANNABA
BADJI MOKHTAR – ANNABA UNIVERSITY



جامعة باجي مختار – عنابة

Faculté : Sciences

Département : Mathématiques

Domaine : Mathématiques - Informatique

Filière : Mathématiques appliquées

Spécialité : Probabilités et Statistique.

Mémoire

Présenté en vue de l'obtention du Diplôme de Master

Thème:

**Estimation dans un modèle Lomax inverse tronqué en
présence de censure à droite**

Présenté par : Chaïben Nehla

Encadrant : *AIDI Khaoula*

MCA

Université Badji Mokhtar - Annaba

Jury de Soutenance :

Benmostefa F.Z.	Prof	UBM-Annaba	Président
AIDI Khaoula	MCA	UBM-Annaba	Encadrant
Stihi Nadjet	MCB	UBM-Annaba	Examineur

Année Universitaire : 2022/2023

TABLE DES MATIÈRES

1	Préliminaires Statistiques	8
1.1	Définitions	8
1.1.1	Statistique	8
1.1.2	Variable aléatoire	8
1.1.3	Loi de probabilité	8
1.2	Analyse de la durée de survie	9
1.2.1	Analyse de survie	9
1.2.2	Distributions de la durée de survie	9
1.3	Quelques distributions usuelles	10
1.3.1	La distribution uniforme	10
1.3.2	La distribution Exponentielle	11
1.3.3	La distribution Weibull	12
1.3.4	La distribution Rayleigh	13
2	L'estimation de maximum de vraisemblance	14
2.1	Introduction	14
2.2	Propriétés d'un estimateur (qualités)	14
2.2.1	Biais d'un estimateur	15
2.2.2	Convergence d'un estimateur	15
2.2.3	Erreur quadratique moyenne	15
2.2.4	Efficacité	15
2.3	Méthodes d'estimation	16
2.3.1	La méthode des moments	16
2.3.2	La méthode du maximum de vraisemblance	17
2.3.3	La vraisemblance pour les différents types de données	19
3	La famille Lomax Inverse Tronquée Généralisée	21
3.1	Définition	21
3.2	La distribution Lomax Inverse Tronquée - Exponentielle	22
3.2.1	Fonction de répartition et densité	22
3.2.2	La fonction de survie et le taux de hasard	23
3.2.3	Fonction de hasard cumulé	23
3.2.4	L'espérance et la variance	23

3.2.5	Estimation du maximum de vraisemblance pour les données complètes	24
3.2.6	Estimation du maximum de vraisemblance pour les données censurées	25
3.3	La distribution Lomax Inverse Tronquée - Weibull	26
3.3.1	Fonction de répartition et densité	26
3.3.2	La fonction de survie et le taux de hasard	27
3.3.3	Fonction de hasard cumulé	28
3.3.4	L'espérance et la variance	28
3.3.5	Estimation des paramètres dans le cas de donnée complète	28
3.3.6	Estimation des paramètres dans le cas de censure droite	30
4	Simulations	31
4.1	Simulation de la distribution TIL-E	31
4.2	Simulation de la distribution TIL-W	32

Remerciements

Je remercie tout d'abord ALLAH le tout puissant de m'avoir donné la force, la patience et la volonté pour réaliser ce travail dans meilleures circonstances et en bon état.

Mes vifs remerciements sont adressés à mon encadreur Mme AIDI Khaoula, qui s'est toujours très disponible tout au long de la réalisation de ce mémoire, ainsi pour l'inspiration, la patience et ces précieux conseils.

Je remercie également les honorables présidents et membres de jury d'avoir accepté d'examiner ce modeste travail en lui apportant de l'intérêt.

Je remercie l'ensemble des enseignants du département de Mathématiques qui nous ont guidés et suivis tout au long de notre cursus universitaire et à leur tête Monsieur le directeur.

Je tiens à remercier profondément :

Mes chers parents, je ne pourrai pas les remercier et remplir leur droit pour ce qu'ils ont fait et font pour moi. Que Dieu fasse d'eux une couronne et une fierté sur ma tête.

Ma chère soeur et son mari et mes frère pour leurs soutien et encouragements tout au long de la réalisation de ce travail.

Mes amies proches et aussi mes collègues qui m'ont apporté leur soutien moral et intellectuel tout au long de notre démarche.

Enfin, à toutes les personnes qui nous ont aidés dans la réalisation de ce modeste travail.
"Merci à tous"

C'est avec profonde gratitude et sincères mots, que je dédie ce modeste travail :

A qui je prends son nom avec une grande fierté, à qui m'a poussé obtenir à ma première école dans la vie, à qui m'a toujours encadré avec beaucoup d'amour et d'attention : "Mon père".

A mon ange dans la vie, à le sens de l'amour et de tendresse, à celle qui sa prière était le secret de mon succès, qui sacrifiée pour mon chemin et m'illumine de douceur et d'amour à : "Ma mère".

Nous espérons qu'un jour nous pourrons leur rendre un peu de ce qu'ils ont fait pour nous, que Dieu leur prête bonheur et long vie.

A ma chère sœur Sara et son mari et mes frères, que dieu les garde et les protège.

A ma famille.

A toutes mes amies, particulièrement Djemoui Nour El Houda.

L'un des problèmes importants de la modélisation des données de fiabilité est celui des données manquantes ou données censurées. Divers auteurs se sont intéressés aux différents types de censure : censure gauche, censure droite, double censure, etc... . La censure progressive, où l'expérimentateur retire, aléatoirement, un nombre déterminé d'unités après chaque défaillance, a attiré l'attention de beaucoup d'auteurs ces dernières années. Dans ce travail, nous nous intéressons à l'estimation paramétrique dans quelques modèles en présence de censure à droite.

Mots-clés : Analyse de survie, La famille lomax inverse tronquée généralisée, Estimateurs du maximum de vraisemblance,

Le terme "fiabilité" est un néologisme introduit dans les années 1960 pour traduire le terme anglo-saxon "reliability" et si l'on accepte de la considérer comme une science, la fiabilité est la science des défaillances.

Elle s'est développée très rapidement, le domaine potentiel des études de fiabilité comprend toutes les activités humaines (économique, scientifique, technique et industriel, etc.). Cette science est basée sur l'expérience et l'une des composantes essentielles de la qualité d'un produit et elle est retenue en tant que critère fondamental dès le stade de la conception.

La fiabilité est l'aptitude d'une entité à accomplir les fonctions requises dans des conditions données d'utilisation pendant une durée de temps donnée.

Le but de notre travail est de présenter la méthode d'estimation par le maximum de vraisemblance et de l'appliquer sur la famille Lomax inverse Tronquée généralisé

Les distributions tronquées ont été largement appliquées, essentiellement dans les tests de durée de vie et études de fiabilité. La forme tronquée d'une distribution résulte de l'application d'une limite sur la plage de la distribution afin qu'elle soit définie correctement sur un sous-ensemble de la plage d'origine. Ainsi, tronqué les distributions sont utilisées de manière à limiter les occurrences à des valeurs situées au-dessus ou en dessous d'un seuil donné ou dans une fourchette spécifiée. La distribution tronquée inférieure (gauche) est obtenu si les occurrences sont limitées à des valeurs inférieures à un seuil donné. De l'autre D'autre part, la distribution supérieure (droite) tronquée apparaît si les occurrences sont limitées à des valeurs qui se situer au-dessus d'un seuil donné.

Récemment, les distributions tronquées généralisées ont attiré l'attention de plusieurs auteurs. parmi lesquels : M. Ali et al,(A truncated Pareto distribution, 2006), L. Zaninetti et al (The truncated Pareto distribution with applications, 2008), Nadarajah (Some truncated distributions, 2009), Teamah et al. (Random sum of truncated and random truncated Lindley distribution, 2010) M.M. Mohie El-Din. et al(Random sum of mid truncated Lindley distribution, 2010), S.H. Alkarni (A class of truncated binomial lifetime distributions, 2013), Zaninetti (A right and left truncated gamma distribution with application to the stars.L, 2014), Chattopadhyay et al (Fitting truncated geometric distributions in large scale real world networksS, 2014), customers J. Pender (The truncated normal distribution : Applications to queues with impatient 2015), Abid, SH et al (A Truncated Fréchet-G generator of distributions 2017), Amal. S. Hassan et al (ATruncated Lomax distributions, 2020), Hassan A.S et al (A truncated Power Lomax distributions, 2020), SH. ENAMI et al (A truncated Lomax-Exponential distributions, 2022).

En effet, le principe de cette approche consiste à maximiser la fonction de vraisemblance en considérant que celle-ci est une fonction des paramètres, ce qui revient dans la plus part des modèles complexes à résoudre des systèmes non linéaires. Des solutions numériques sont alors nécessaires pour trouver des estimateurs pour les paramètres du modèle et d'en déduire des estimateurs des caractéristiques de la fiabilité, en particulier, la fonction de fiabilité et le taux de panne.

Les estimateurs du maximum de vraisemblance, permettent de déterminer les paramètres de la loi de fiabilité recherchée. On en déduit les indicateurs tels que le taux de défaillance, la fonction de fiabilité, . . .

On a structuré ce mémoire au tour de quatre chapitres.

Dans le premier chapitre nous donnons un rappel sur des notions et définitions de base de la théorie de la fiabilité utilisées dans les autres chapitres.

Dans le deuxième chapitre, nous nous intéressons à l'estimation, nous aborderons deux méthodes d'estimation paramétrique, du maximum de vraisemblance et des moments.

Dans le troisième chapitre, on s'intéresse à l'étude de la famille lomax inverse tronquée généralisé ainsi que ses caractéristiques : la fonction de densité de probabilité, la fonction de répartition, la fonction de survie, les fonctions du taux de hasard et du taux de hasard cumulé.

Enfin, on termine par la simulation des données trouvées dans ce mémoire, nous utilisons le logiciel R.

CHAPITRE 1

Préliminaires Statistiques

Dans ce chapitre, nous avons énoncé quelques définitions et propriétés qui seront utilisées plus tard.

1.1 Définitions

1.1.1 Statistique

La statistique est une science qui s'intéresse à la collecte, l'analyse, l'interprétation, la présentation et la prise de décision à partir de données. Elle permet de décrire et de comprendre les phénomènes observés dans la nature ou dans la société, en utilisant des méthodes mathématiques et probabilistes pour évaluer les incertitudes liées aux données.

1.1.2 Variable aléatoire

Soit (Ω, \mathcal{F}, P) un espace probabilisé. Une variable aléatoire (v.a), que l'on note par X , est toute application définie sur l'espace fondamental Ω et à valeur dans l'ensemble des réelles \mathbb{R} telle que :

$$\begin{aligned} X : \Omega &\rightarrow \mathbb{R} \\ \omega &\rightarrow X(\omega) \end{aligned}$$

On distingue deux types de variables aléatoires :

- Lorsque l'ensemble des résultats X d'une variable aléatoire est un ensemble fini ou infini dénombrable, alors cette variable aléatoire est dite discrète.
- Lorsque l'ensemble des résultats X d'une variable aléatoire est un intervalle de nombres réels, alors cette variable aléatoire est dite continue.

1.1.3 Loi de probabilité

La loi de probabilité est un modèle mathématique qui décrit la probabilité d'occurrence de différents événements aléatoires. En termes plus simples, c'est une fonction qui décrit la

probabilité associée à un ensemble des résultats possibles d'une expérience aléatoire, les lois de probabilité permettent de décrire les variables aléatoires sous la forme d'une "expérience type" puis d'analyser cette expérience en détail pour pouvoir déduire les principales caractéristiques de toutes les expériences aléatoires qui sont du même type.

1.2 Analyse de la durée de survie

1.2.1 Analyse de survie

L'analyse de survie désigne un ensemble de techniques statistiques permettant de traiter des données soumises à une censure, qui cherche à modéliser le temps restant avant la mort pour des organismes biologiques (l'espérance de vie) ou le temps restant avant l'échec ou la panne dans les systèmes artificiels, ce que l'on représente graphiquement sous la forme d'une courbe de survie.

1.2.2 Distributions de la durée de survie

Supposons que la durée de survie X soit une variable positive ou nulle, et absolument continue, alors sa loi de probabilité peut être définie par l'une des cinq fonctions équivalentes suivantes (chacune des fonctions ci-dessous peut être obtenue à partir de l'une des autres fonctions) :

Fonction de survie S

La fonction de survie est pour t fixé, la probabilité de survivre jusqu'à l'instant t , c'est-à-dire

$$S(t) = P(X \succ t), t \geq 0$$

Fonction de répartition F

La fonction de répartition (ou c.d.f. pour "cumulative distribution fonction") représente, pour t fixés, la probabilité de mourir avant l'instant t , c'est-à-dire

$$F(t) = P(X \leq t) = 1 - S(t)$$

Densité de probabilité f

C'est la fonction $f(t) \geq 0$ telle que pour tout $t \geq 0$

$$F(t) = \int_0^t f(u) du$$

Si la fonction de répartition F admet une dérivée au point t alors

$$f(x) = \lim_{h \rightarrow 0} \frac{P(t < X < t + h)}{h} = F'(t) = -S'(t)$$

Pour t fixé, la densité de probabilité représente la probabilité de mourir dans un petit intervalle de temps après l'instant t .

Risque instantané λ (ou taux de hasard)

Le risque instantané (ou taux d'incidence), pour t fixé caractérise la probabilité de mourir dans un petit intervalle de temps après t , conditionnellement au fait d'avoir survécu jusqu'au temps t (c'est-à-dire le risque de mort instantané pour ceux qui ont survécu) :

$$\lambda(t) = \lim_{h \rightarrow 0} \frac{p(t \leq X < t+h \mid X \geq t)}{h} = \frac{f(t)}{S(t)} = -\ln(S(t))'$$

Taux de hasard cumulé Λ

Le taux de hasard cumulé est l'intégrale du risque instantané λ :

$$\Lambda(t) = \int_0^t \lambda(u) du = -\ln(S(t))$$

On peut déduire de cette équation une expression de la fonction de survie en fonction du taux de hasard cumulé (ou du risque instantané) :

$$S(t) = \exp(-\Lambda(t)) = \exp\left(-\int_0^t \lambda(u) du\right)$$

On en déduit que

$$f(t) = \lambda(t) \exp\left(-\int_0^t \lambda(u) du\right)$$

1.3 Quelques distributions usuelles

1.3.1 La distribution uniforme

La loi uniforme continue est une généralisation de la fonction rectangle à cause de la forme de sa fonction de densité de probabilité. Elle est paramétrée par les plus petites et plus grandes valeurs a et b que la variable aléatoire uniforme peut prendre. Cette loi continue est souvent notée $U(a, b)$.

La distribution uniforme est utile pour l'échantillonnage à partir des distributions arbitraires. La distribution uniforme est souvent utilisée dans la simulation des échantillons, de nombreux langages de programmation ont la capacité de générer des nombres pseudo-aléatoires qui sont distribués efficacement selon la distribution uniforme standard.

La fonction de densité de probabilité de cette distribution est définie par :

$$f(t) = \frac{1}{\lambda}, 0 < t < \lambda < \infty$$

La fonction de répartition s'écrit comme suit :

$$F(t) = \frac{t}{\lambda}, 0 < t < \lambda < \infty$$

La fonction de survie est :

$$S(t) = 1 - F(t) = 1 - \frac{t}{\lambda}, 0 < t < \lambda < \infty$$

La fonction du taux de hasard est donnée par :

$$h(t) = \frac{f(t)}{S(t)} = (\lambda - t)^{-1}, 0 < t < \lambda < \infty$$

La fonction du taux de hasard cumulé est :

$$H(t) = -\log[S(t)] = -\log\left(1 - \frac{t}{\lambda}\right), 0 < t < \lambda < \infty$$

La moyenne de la distribution uniforme sur $[a, b]$ est

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

Et la variance est

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

1.3.2 La distribution Exponentielle

Une loi exponentielle modélise la durée de vie d'un phénomène sans mémoire, où sans vieillissement, ou sans usure. En d'autres termes, le fait que le phénomène ait duré pendant t heures ne change rien à son espérance de vie à partir du temps t .

Une variable aléatoire continue X suit une loi exponentielle de paramètre (d'intensité ou inverse de l'échelle) $\lambda > 0$ si elle admet pour **densité de probabilité la fonction**

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \forall x \in \mathbb{R}^+ \\ 0 & \text{ailleurs} \end{cases}$$

On note $X \rightsquigarrow \exp(\lambda)$

La fonction de répartition de X

$$F_X(x) = \begin{cases} 1 - \exp(-\lambda x) & \forall x \in \mathbb{R}^+ \\ 0 & \text{ailleurs} \end{cases}$$

Alors la fonction de survie est :

$$S(x) = e^{-\lambda x}, \quad x \geq 0$$

Par conséquent, **la fonction de taux du hasard est :**

$$h(x) = \lambda, \quad \lambda \geq 0$$

On définit l'espérance mathématique, la variance et l'écart type de X

$$E(x) = \frac{1}{\lambda}$$

$$Var(x) = \frac{1}{\lambda^2}$$

$$\sigma(x) = \frac{1}{\lambda}$$

1.3.3 La distribution Weibull

En théorie des probabilités, la loi de Weibull, nommée d'après *Waloddi Weibull* en 1951, est une loi de probabilité continue. La loi de Weibull est un cas spécial de la loi d'extremum généralisée au même titre que la loi de Gumbel ou la loi de Fréchet.

Le domaine d'application de la distribution de Weibull est très vaste, nous citons par exemple : l'analyse de survie, ingénierie de fiabilité et analyse des défaillances, théorie de la valeur extrême et elle est aussi utilisée pour modéliser la taille des sinistres de réassurance et le développement cumulatif des pertes d'asbestose en assurance générale. La distribution de Weibull est aussi utile en génie électrique, ingénierie industrielle, prévisions météorologiques, industrie éolienne pour décrire les distributions de vitesse du vent et en ingénierie des systèmes de communication.

La distribution de Weibull a aussi un rôle dans les systèmes radar pour modéliser la dispersion du niveau des signaux reçus produit par certains types de fouillis.

La fonction de densité de probabilité et la fonction de répartition de la loi Weibull sont, respectivement, données par :

$$f(t) = \frac{\beta}{\eta} \left(\frac{t}{\eta}\right)^{\beta-1} \exp\left\{-\left(\frac{t}{\eta}\right)^\beta\right\}, \quad t \geq 0, \quad \eta, \beta > 0$$

$$F(t) = 1 - \exp\left\{-\left(\frac{t}{\eta}\right)^\beta\right\}$$

La fonction de survie est donnée par :

$$S(t) = 1 - F(t) = \exp\left\{-\left(\frac{t}{\eta}\right)^\beta\right\}$$

La fonction du taux de hasard et du taux de hasard cumulé est :

$$h(t) = \frac{f(t)}{S(t)} = \frac{\beta}{\eta} \left(\frac{t}{\eta}\right)^{\beta-1}$$

$$H(t) = -\log[S(t)] = \left(\frac{t}{\eta}\right)^\beta$$

Où $\beta > 0$ représente le paramètre d'échelle et $\eta > 0$ représente le paramètre de forme de la distribution Weibull.

La moyenne de la distribution de Weibull est :

$$\mu = \frac{\Gamma(1 + 1/\beta)}{\eta}$$

et la variance est :

$$\sigma^2 = \frac{1}{\eta^2} \left[\Gamma\left(1 + \frac{2}{\beta}\right) - \Gamma^2\left(1 + \frac{1}{\beta}\right) \right]$$

où $\Gamma(\eta)$ est la fonction gamma bien connue, définie comme

$$\Gamma(\beta) = (\beta - 1)!$$

1.3.4 La distribution Rayleigh

La distribution de Rayleigh est une distribution de probabilité continue nommée d'après le mathématicien anglais *Rayleigh*.

Cette distribution est largement utilisée : En théorie des communications, pour modéliser des trajets multiples de signaux diffusés denses atteignant un récepteur, en sciences physiques (pour modéliser la vitesse du vent, la hauteur des vagues et le rayonnement sonore lumineux), en ingénierie (pour mesurer la durée de vie d'un objet, où la durée de vie dépend de l'âge de l'objet ; comme : résistances, transformateurs et condensateurs dans les radars d'avions) et aussi en imagerie médicale (modéliser la variance du bruit en imagerie par résonance magnétique).

la fonction de densité de probabilité est :

$$f(t) = \frac{t}{\lambda^2} \exp\left(-\frac{t^2}{2\lambda^2}\right), \quad t > 0$$

la fonction de répartition

$$F(t) = 1 - \exp\left(-\frac{t^2}{2\lambda^2}\right), \quad t > 0$$

Les fonctions de survie, taux de hasard et du taux de hasard cumulé sont, respectivement, données par :

$$S(t) = \exp\left(-\frac{t^2}{2\lambda^2}\right), \quad t > 0$$

$$h(t) = \frac{t}{\lambda^2}, \quad t > 0$$

et

$$H(t) = -\log[S(t)] = \frac{t^2}{2\lambda^2}, \quad t > 0$$

L'espérance et la variance d'une variable aléatoire suit une loi de Rayleigh sont les suivantes :

$$\mu = \lambda \sqrt{\frac{\pi}{2}}$$

et

$$\sigma^2 = \frac{4 - \pi}{2} \lambda^2$$

CHAPITRE 2

L'estimation de maximum de vraisemblance

2.1 Introduction

Le problème de l'estimation est l'impossibilité de connaître exactement la valeur d'un paramètre inconnu.

On utilise une statistique d'un échantillon prélevé au hasard ou par d'autres méthodes aléatoires.

On peut donner une estimation ponctuelle la plus proche possible et par intervalle de confiance : dans ce cas, on détermine un intervalle dans lequel se trouve le paramètre à estimer avec une probabilité fixée.

Ayant fait choix d'un estimateur, on se pose la question de savoir s'il s'agit d'un bon estimateur. Il va donc falloir se donner des critères de qualité des estimateurs qui s'exprime par leur convergence, leur biais et leur efficacité. Diverses méthodes permettant d'obtenir des estimateurs de qualités différentes. Dans ce chapitre, nous ne nous intéressons qu'aux deux méthodes d'estimation les plus usuelles, la méthode des moments et la méthode du maximum de vraisemblance.

Mais il faut d'abord définir précisément ce qui est un estimateur.

Définition 2.1 Soit un échantillon de taille n fini, issu d'une population mère et soient x_1, \dots, x_n des réalisations indépendantes de la variable aléatoire X dont la densité $f(x, \theta)$ dépend d'un paramètre inconnu θ . On appelle estimateur de θ d'une variable aléatoire réelle dont $T_n = T(x_1, \dots, x_n)$ ou T une application de $\mathbb{R}^n \rightarrow \mathbb{R}$ qui aux valeurs observées (x_1, \dots, x_n) en fait correspondre un nombre réel noté par :

$$\hat{\theta} = T_n(x_1, \dots, x_n)$$

est appelé estimation ponctuelle du paramètre θ .

2.2 Propriétés d'un estimateur (qualités)

Le choix d'un estimateur va reposer sur ses qualités. Le premier défaut possible concerne la possibilité de comporter un biais.

2.2.1 Biais d'un estimateur

Le biais d'un estimateur $\hat{\theta}$ du paramètre θ est :

$$Biais(\hat{\theta}) = E(\hat{\theta}) - \theta$$

On dit que $\hat{\theta}$ est sans biais ou non-biaisé si

$$Biais(\hat{\theta}) = 0.$$

Le biais est une mesure de l'erreur systématique faite en approximant θ par $\hat{\theta}$

2.2.2 Convergence d'un estimateur

Dénotons par $\hat{\theta}_n$, un estimateur du paramètre θ calculé à partir d'un échantillon de taille n .

Définition

Un estimateur $\hat{\theta}_n$, est convergent si pour tout $\varepsilon > 0$.

$$\lim_{n \rightarrow \infty} P \left(\left[(|\hat{\theta}_n - \theta|) \right] < \varepsilon \right) = 1$$

Ceci signifie : si la taille de l'échantillon est assez grande alors on est (presque) certain que l'estimateur $\hat{\theta}_n$, est très proche de θ .

2.2.3 Erreur quadratique moyenne

L'erreur quadratique moyenne (EQM) d'un estimateur erreur $\hat{\theta}$ du paramètre θ est

$$EQM(\hat{\theta}) = E \left[(\hat{\theta} - \theta)^2 \right]$$

L'EQM est une mesure de la précision d'un estimateur

2.2.4 Efficacité

On dit qu'un estimateur $\hat{\theta}_n$ sans biais de θ est efficace si sa variance atteint la borne la de *Fréchet-Cramer-Rao*

$$Var(\hat{\theta}_n) = \frac{1}{I_n(\theta)}$$

où $I_n(\theta) = nI(\theta)$ est l'information de *Fisher* donnée par l'échantillon avec

$$I(\theta) = -E \left(\frac{\partial^2 \ln f(x, \theta)}{\partial \theta^2} \right)$$

Remarque : On dit qu'un estimateur $\hat{\theta}_n$ sans biais de θ est asymptotiquement efficace si et seulement si et seulement si $Var(\hat{\theta}_n)I_n(\theta)$ tend vers 1 quand n tend vers l'infini.

Proposition : Un estimateur sans biais efficace est convergent.

Preuve. Puisque $\hat{\theta}_n$ est un estimateur sans biais efficace, on a

$$\text{Var}(\hat{\theta}_n) = \frac{1}{I_n(\theta)} = \frac{1}{nI(\theta)}$$

Alors

$$\lim_{n \rightarrow \infty} \text{Var}(\hat{\theta}_n) = \lim_{n \rightarrow \infty} \frac{1}{nI(\theta)} = 0$$

Donc $\hat{\theta}_n$ convergent.

Théorème

Si $\hat{\theta}$ est un estimateur du paramètre θ alors :

$$EQM(\hat{\theta}) = v(\hat{\theta}) + \left[\text{Biais}(\hat{\theta}) \right]^2$$

Le meilleur de deux estimateur θ_1 et θ_2 c'est-à-dire le plus efficace, est celui qui a la plus petite EQM : θ_1 est plus efficace que θ_2 si :

$$EQM(\theta_1) < EQM(\theta_2) \iff \frac{EQM(\theta_1)}{EQM(\theta_2)} < 1$$

Lorsque deux estimateurs son non biaisés, ceci revient à dire que le plus efficace est celui dont la variance est la plus petite.

2.3 Méthodes d'estimation

Dans cette section, nous nous intéressons qu'aux deux méthodes d'estimation le plus usuelles, la méthode du maximum de vraisemblance et la méthode des moments.

2.3.1 La méthode des moments

La méthode des moments est un outil d'estimation intuitif qui date du début des statistiques. Elle consiste à déterminer les paramètres en égalant les estimations empiriques et théoriques des moments.

Dans le cas où le paramètre à estimer est $\theta = E_\theta(X)$, moyenne théorique de la loi, nous avons vu que l'estimateur naturel était la moyenne empirique, ou moyenne de l'échantillon, \bar{X}_n . De même, pour estimer le paramètre $\theta = V_\theta(X)$, variance de la loi, nous retenons logiquement comme estimateur la variance empirique S_n^2 .

Plus généralement, si l'un des moments d'ordre $k \in \mathbb{N}^*$, non centré $m_k = E_\theta(X^k) = m_k(\theta)$, ou centré $\mu_k = E_\theta(X - m_1)^k = \mu_k(\theta)$, dépend de θ , nous allons chercher un estimateur par résolution de l'équation en θ obtenue en égalent moment théorique et moment empirique correspondant, soit :

$$m_{kn} = \frac{1}{n} \sum_{i=1}^n X_i^k = m_k(\theta) \quad \text{ou} \quad \mu_{kn} = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^k = \mu_k(\theta)$$

La solution de l'équation, si elle existe et est unique, sera appelée estimateur obtenu par la méthode des moments. Dans les exemples introductif où $\theta = E_\theta(X)$ et $\theta = V_\theta(X)$, les équations à résoudre s'écrivaient sous forme résolue :

$$\theta = \overline{X_n} \quad \text{et} \quad \theta = S_n^2 .$$

Remarque :

Dans certains cas, comme celui de la loi Gamma, le calcul de la fonction de vraisemblance peut poser des problèmes (l'utilisation de l'ordinateur et d'algorithmes numériques est indispensable) tandis que l'estimation des moments est très facilement accessible.

2.3.2 La méthode du maximum de vraisemblance

La fonction de vraisemblance

Définition : Quand les observations sont toutes discrètes ou toutes continues, on appelle fonction de vraisemblance (ou plus simplement vraisemblance) pour l'échantillon x_1, \dots, x_n , la fonction du paramètre θ :

$$L(\theta; x_1, \dots, x_n) = \begin{cases} P(X_1 = x_1, \dots, X_n = x_n; \theta) & \text{si les } X_i \text{ sont discrètes} \\ f_{x_1, \dots, x_n}(x_1, \dots, x_n; \theta) & \text{si les } X_i \text{ sont continues} \end{cases}$$

Dans tous les exemples que nous traiterons ici, les X_i sont indépendantes et de même loi. Dans ce cas, la fonction de vraisemblance s'écrit :

$$L(\theta; x_1, \dots, x_n) = \begin{cases} \prod_{i=1}^n P(X_i = x_i; \theta) = \prod_{i=1}^n P(X = x_i; \theta) & \text{si les } X_i \text{ sont discrètes} \\ \prod_{i=1}^n f_{x_i}(x_i; \theta) = \prod_{i=1}^n f(x_i; \theta) & \text{si les } X_i \text{ sont continues} \end{cases} \quad (2.1)$$

Remarque : La probabilité et la densité utilisées dans cette définition sont des fonctions des observations x_1, \dots, x_n , dépendant du paramètre θ . À l'inverse, la fonction de vraisemblance est considérée comme une fonction de θ dépendant des observations x_1, \dots, x_n , ce qui permet, par exemple, de dériver cette fonction par rapport à θ .

L'estimateur de maximum de vraisemblance(EMV)

En suivant le raisonnement précédent, pour n quelconque, il est logique de dire que la valeur la plus vraisemblable de θ est la valeur pour laquelle la probabilité d'observer x_1, \dots, x_n est la plus forte possible. Cela revient à faire comme si c'était l'éventualité la plus probable qui s'était produite au cours de l'expérience.

Définition

L'estimation de maximum de vraisemblance de θ est la valeur $\hat{\theta}_n$ de θ qui rend maximale la fonction de vraisemblance $L(\theta; x_1, \dots, x_n)$. L'estimateur de maximum de vraisemblance (EMV) de θ est la variable aléatoire correspondante.

Comme dans l'exemple, dans la plupart des cas, la fonction de vraisemblance s'exprime comme un produit.

Donc $\hat{\theta}_n$ sera en général calculé en maximisant la log-vraisemblance :

$$\theta^n = \arg \max_{\theta} \ln \mathcal{L}(\theta; x_1, \dots, x_n)$$

Quand $\theta = (\theta_1, \dots, \theta_d) \in \mathbb{R}^d$ et que toutes les dérivées partielles ci-dessous existent, $\hat{\theta}_n$ est solution du système d'équations appelées équations de vraisemblance :

$$\forall j \in \{1, \dots, d\}, \frac{\partial}{\partial \theta_j} \ln L(\theta; x_1, \dots, x_n) = 0$$

Information de Fisher et efficacité

Considérons une variable aléatoire X dont la loi dépend de s paramètres notés comme un vecteur $\theta \in T \subset \mathbb{R}^s$. Nous considérons soit ses probabilités $f(x, \theta) = P(X = x, \theta)$ (**cas discret**), soit sa densité $f(x, \theta)$ (**cas continu**). Nous utiliserons dans les énoncés la fonction de densité, mais le cas discret est tout à fait analogue. Nous étudions d'abord le cas $s = 1$, puis le cas général. La notion suivante a été introduite par Ronald Aylmer Fisher qui quantifie l'information relative à un paramètre contenue dans une distribution.

Score de l'échantillon

On définit la fonction score par :

$$S_n(\theta) = \frac{\partial}{\partial \theta} \ln L(X_1, \dots, X_n; \theta) = \frac{L'(X_1, \dots, X_n; \theta)}{L(X_1, \dots, X_n; \theta)}$$

- Pour tout θ , Le score est une variable aléatoire.
- Le score s'annule à un optimum en θ de la fonction de vraisemblance.
- Le score est centré : $E(S_n(\theta)) = 0$.
- La variance du score (si elle existe) s'appelle l'information de Fisher apportée par l'échantillon sur θ ,

$$\begin{aligned} I_n(\theta) &= V(S_n(\theta)) = E((S_n(\theta))^2) \\ &= E\left(\left(\frac{\partial \ln L(X_1, \dots, X_n; \theta)}{\partial \theta}\right)^2\right) \end{aligned}$$

Cette quantité mesure l'information apportée par un échantillon sur le paramètre.

Propriétés de l'information de Fisher La quantité d'information de Fisher vérifie les propriétés suivantes :

- La positivité : on a

$$I_n(\theta) = V(S_n(\theta)) \geq 0, \forall \theta.$$

- L'additivité : si le domaine de définition de la v.a X ne dépend pas de paramètre θ , chaque observation apporte la même information,

$$I_n(\theta) = nI_1(\theta)$$

- Autre formulation : si le domaine de définition de X ne dépend pas de θ alors l'information de Fisher :

$$I_n(\theta) = -E\left(\frac{\partial^2 \ln L(X_1, \dots, X_n; \theta)}{\partial \theta^2}\right)$$

si cette quantité existe.

Borne de Cramer-Rao et efficacité **Théorème 2.4** Sous les hypothèse de Cramer-Rao, en particulier si $E = X(\Omega)$ est indépendant du paramètre à estimer θ , pour tout estimateurs sans biais T_n de θ on a :

$$V(T_n) \geq \frac{1}{I_n(\theta)} = B_F(\theta)$$

La quantité $B_F(\theta)$ est la borne inférieure de Fréchet-Darmois-Cramer-Rao (FDCR en abrégé).

Définition 2.4

Un estimateur sans biais $T_n(\theta)$ est dit efficace si sa variance est égale à la borne inférieure de FDCR :

$$V(T_n) = \frac{1}{I_n(\theta)}$$

2.3.3 La vraisemblance pour les différents types de données

Le but de notre travail est de trouver des estimations raisonnables de la fiabilité qui dépend des paramètres de la distribution étudiée, et dans la plus part des situations pratiques la vraie valeur n'est pas connue, donc on va essayer d'estimer la valeur de cette fonction pour les différents types des données en utilisons la méthode du maximum de vraisemblance.

Données complètes

L'outil de base de l'estimation est la notion de vraisemblance. Plaçons-nous pour l'instant dans le cas de données complètes. On a observé n réels x_1, x_2, \dots, x_n , que l'on considère comme des réalisations d'un n -échantillon (x_1, x_2, \dots, x_n)

Les x_i sont des variables aléatoires indépendantes et de même loi. Cette loi inconnue a pour densité f , par définitions, la vraisemblance dans notre cas est définie par

$$L = \prod_{i=1}^n f(x_i)$$

Données censurées de différents types

Lors de l'analyser des données de survie, on est souvent confronté à des données incomplètes dues à deux phénomènes différents : la censure et la troncature. Dans cette partie, nous n'étudierons que le phénomène de la censure

La censure est le phénomène le plus courant rencontré lors de la collecte de données de survie. Il existe trois types de censure, censure à droite, censure à gauche et censure par intervalle.

Pour l'individu i , considérons

- son temps de survie X_i .
- son temps de censure C_i .
- la durée réellement observée T_i .

1. **Censure à droite** : Si l'individu n'a pas vécu l'évènement, la durée de vie est censurée à droite lors de sa dernière apparition. Autrement dit, si $X_i \geq C_i$. Il existe trois types de censure à droite :

1-1 Censure à droite de type I (Censure non aléatoire)

Dans ce cas, les variables X_i , avec $i = 1, \dots, n$, ne sont pas toutes observées. On observe X_i pendant une durée de censure inférieure ou égale à C et $C_i = C$, sinon on sait juste $X_i > C$.

Nous utilisons la notation suivante : $T_i = X_i \wedge C = \min(X_i, C)$.

Par exemple, on peut tester la durée de vie de n objets identiques (par exemple des ampoules) sur un intervalle d'observation fixé $[0, u]$

1-2 Censure à droite de type II

Il survient lorsque nous décidons d'observer la durée de survie de n patients jusqu'à ce que k d'entre eux meurent, moment auquel l'étude est interrompue. Soient $X_{(i)}$ et $T_{(i)}$ les statistiques d'ordre des variables X_i et T_i . Par conséquent, La date de censure est donc $X_{(k)}$ et nous observons les variables suivantes

$$\begin{aligned} T_{(1)} &= X_{(1)} \\ &\dots\dots \\ &\dots\dots \\ &\dots\dots \\ T_{(k)} &= X_{(k)} \\ T_{(k+1)} &= X_{(k)} \\ &\dots\dots \\ &\dots\dots \\ &\dots\dots \\ T_{(n)} &= X_{(k)} \end{aligned}$$

1-3 Censure à droite de type III (Censure aléatoire)

Supposons que les C_i sont indépendantes et identiquement distribuées(i.i.d). Dans ce cas, on observe la variable $T_i = X_i \wedge C_i$.

Lorsqu'il existe un examen des droits de type III, les informations disponibles pour chaque individu est $\{T_i, d_i\}$ avec

$$\begin{cases} T_i = X_i \wedge C_i = \min(X_i, C_i) \\ \Delta_i = \begin{cases} 1 & \text{Si } X_i \leq C_i \\ 0 & \text{Sinon} \end{cases} \end{cases}$$

Elle est la plus courante.

2 Censure à gauche :

La durée de vie est censurée à gauche si la valeur exacte n'est pas connue avant C_i . Autrement dit, la durée de vie est censurée à gauche si l'individu a subi l'évènement avant qu'il soit observé. En présence de censure à gauche on a :

$$\begin{cases} T_i = X_i \vee C_i = \max(X_i, C_i) \\ \Delta_i = \begin{cases} 1 & \text{Si } X_i \geq C_i \\ 0 & \text{Si } X_i \leq C_i \end{cases} \end{cases}$$

3 Censure par intervalle :

Si l'évènement n'a pas été observé, mais qu'il s'est produit entre deux dates d'observation. On dit censuré par intervalle. La seule information disponible est que cela s'est passé entre deux dates connues.

CHAPITRE 3

La famille Lomax Inverse Tronquée Généralisée

Dans ce chapitre, nous commençons par présenter la famille Lomax Inverse Tronquée généralisée et ses différentes formes de fonction de risque, ensuite on calcul les estimateurs du maximum de vraisemblance des paramètres inconnus à partir des données initiales dans les deux cas.

3.1 Définition

La famille Lomax Inverse Tronquée généralisée est une distribution continue définie par la CDF suivante

$$F(x) = 1 - 2^\alpha [1 + (1 - G(x))^{-1}]^{-\alpha}, \quad x \in R \quad (1)$$

où $\alpha > 0$ est un paramètre de forme et $G(x) = G(x, \zeta)$ est un cdf d'une distribution mère, dépendant éventuellement d'un ou plusieurs paramètres symbolisés par ζ (dans la suite, nous expliquerons seul si nécessaire).

La fonction de densité de probabilité (pdf) de la famille TIL-G est donné par

$$f(x) = 2^\alpha \alpha g(x) (1 - G(x))^{-2} [1 + (1 - G(x))^{-1}]^{-\alpha-1}, \quad x \in R \quad (2)$$

avec $g(x) = g(x, \zeta)$ est le pdf de la distribution mère. Ci-après, par la suite, une variable aléatoire X ayant le cdf donné comme Eq. (1) (ou, de manière équivalente, le pdf donné comme équation (2)) sera désigné par $X \sim TIL - G$.

La fonction de survie, disons $S(x) = S(x, \zeta)$, et le fonction de hasard, disons $h(x) = h(x, \zeta)$, sont donnés par

$$S(x) = 2^\alpha [1 + (1 - G(x))^{-1}]^{-\alpha}, \quad x \in R \quad (3)$$

Et

$$h(x) = \frac{\alpha g(x) (1 - G(x))^{-2}}{1 + (1 - G(x))^{-1}}, \quad x \in R \quad (4)$$

La fonction quantile de la famille $TIL - G$, disons $Q(u) = Q(u, \zeta)$, est obtenu en inversant l'Eq. (1) comme

$$Q(u) = G^{-1} \left[1 - \left(2(1-u)^{-1/\alpha} - 1 \right)^{-1} \right], \quad u \in (0, 1) \quad (5)$$

avec $G^{-1}(u) = G^{-1}(u; \zeta)$ est la fonction quantile du distribution mère . La fonction $Q(u)$ est à la base de plusieurs probabilités mesurent et peuvent servir pour générer des valeurs à partir de toute distribution spéciale de la famille.

3.2 La distribution Lomax Inverse Tronquée - Exponentielle

3.2.1 Fonction de répartition et densité

Soit x une variable aléatoire, la fonction de distribution cumulative (cdf) de la distribution de base (Exponentielle) avec paramètres λ et donnée par :

$$G(x) = 1 - \exp(-\lambda x) \quad (6)$$

La fonction de densité de loi exponentielle donnée par

$$g(x) = \lambda e^{-\lambda x} \quad (7)$$

Pour développer le CDF de la distribution Lomax Inverse Tronquée - Exponentielle ($TIL - E$), l'expressions de l'équation (6) est remplacée dans l'équation (1) comme suit :

$$F(x) = 1 - 2^\alpha (1 + e^{\lambda x})^{-\alpha} \quad (8)$$

La fonction de densité de probabilité est (remplaçant (6) et (7) dans (2)) :

$$f(x) = \alpha \lambda 2^\alpha e^{\lambda x} (1 + e^{\lambda x})^{-\alpha}$$

Pour simuler une variable provenant de la distribution $TIL - E$, nous utilisons l'équation non linéaire :

$$x = \frac{1}{\lambda} \ln \left[-1 + \left[2^{-\alpha} (1 - u) \right]^{\frac{-1}{\alpha}} \right]$$

Où u suit la distribution uniforme $U(0, 1)$.

Un exemple de la densité $TIL - E$ est présenté dans la figure (1.1).

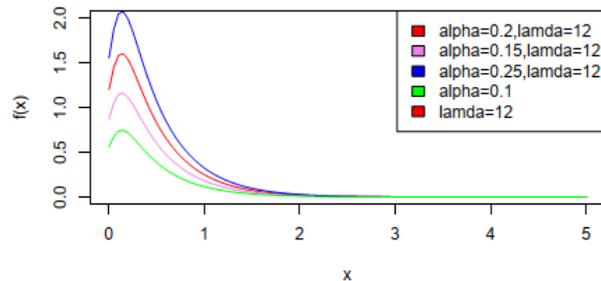


Fig. 1.1. Densité de la distribution $TIL - E$ pour le paramètre de forme $\alpha = (0.2; 0.15; 0.25; 0.1)$ avec $\lambda = 1.2$.

3.2.2 La fonction de survie et le taux de hasard

La fonction de survie du modèle $TIL - E$ est donnée par

$$S(x) = 2^\alpha (1 + e^{\lambda x})^{-\alpha}$$

Sa fonction de hasard est donnée par :

$$h(x) = \frac{\alpha \lambda e^{\lambda x}}{1 + e^{\lambda x}}$$

Le graphe du taux de hasard de cette distribution $TIL - E$ est présenté dans la figure (1.2).

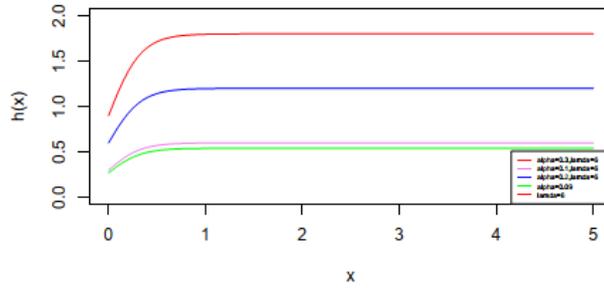


Fig. 1.2. Taux de hasard de la distribution $TIL - E$ pour $\alpha = (0.3; 0.1; 0.2; 0.09)$ avec $\lambda = 6$.

3.2.3 Fonction de hasard cumulé

La fonction de hasard cumulé du modèle $TIL - E$ est donné par :

$$\begin{aligned} H(t) &= -\ln\{S(t)\} \\ H(t) &= -\alpha \ln 2 + \alpha \ln(1 + e^{\lambda t}) \end{aligned}$$

3.2.4 L'espérance et la variance

L'espérance est définie comme :

$$E(T) = \int_0^{+\infty} t dF(t) = \int_0^{+\infty} S(t) dt = \int_0^{+\infty} 2^\alpha (1 + e^{\lambda t})^{-\alpha} dt$$

Comme pour la distribution de $TIL - E$, le calcul ci-dessus nécessite l'intégration numérique.

La variance peut être obtenue par les formules suivantes :

$$\begin{aligned}
 V(T) &= E(T^2) - E(T)^2 \\
 V(T) &= \int_0^{+\infty} t^2 dF(t) - E(T)^2 = 2 \int_0^{+\infty} tS(t)dt - E(T)^2 \\
 V(T) &= 2 \int_0^{+\infty} t(2^\alpha(1 + e^{\lambda t})^{-\alpha})dt - E(T)^2
 \end{aligned}$$

3.2.5 Estimation du maximum de vraisemblance pour les données complètes

Estimation du maximum de vraisemblance

Soit une variable aléatoire $x \sim TIL - E(\theta)$ de vecteur de paramètres $\theta = (\alpha, \lambda)$. Afin de calculer les estimateurs du maximum de vraisemblance des paramètres inconnus de cette distribution, on a besoin de la fonction de vraisemblance donnée par l'équation suivante :

$$\begin{aligned}
 L(x, \theta) &= \prod_{i=1}^n f(x_i, \theta) \\
 &= \prod_{i=1}^n (\alpha \lambda 2^\alpha e^{\lambda x_i} (1 + e^{\lambda x_i})^{-\alpha-1})
 \end{aligned}$$

Et l'expression de la fonction log-vraisemblance est :

$$\begin{aligned}
 l(x, \theta) &= \sum_{i=1}^n \ln f(x_i, \theta) \\
 &= n \ln(\alpha \lambda) + n \alpha \ln(2) + \lambda \sum_{i=1}^n x_i - (\alpha + 1) \sum_{i=1}^n \ln(1 + e^{\lambda x_i})
 \end{aligned}$$

Pour notre modèle $TIL - E$, les fonctions de score sont obtenues comme suit :

$$\begin{aligned}
 \frac{\partial L}{\partial \alpha} &= \frac{n}{\alpha} + n \ln(2) - \sum_{i=1}^n \ln(1 + e^{\lambda x_i}) \\
 \frac{\partial L}{\partial \lambda} &= \frac{n}{\lambda} + \sum_{i=1}^n x_i - (\alpha + 1) \sum_{i=1}^n \frac{x_i e^{\lambda x_i}}{1 + e^{\lambda x_i}}
 \end{aligned}$$

Nous ne pouvons pas trouver les formules explicites des estimateurs du maximum de vraisemblance des paramètres, c'est pourquoi on utilise des approches numériques telles que la méthode de *Newton Raphson*, *l'algorithme BB* ou autres.

La matrice d'information de Fisher

Soit $X = (x_1, \dots, x_n)$ un échantillon aléatoire, et soit $f(x, \theta)$ la fonction densité de probabilité pour un certain modèle de données, de vecteur de paramètres $\theta = (\theta_1, \dots, \theta_n)$, alors la matrice d'information de Fisher $I_n(\theta)$ sur l'échantillon de taille n est donnée par la matrice symétrique $k \times k$

$$I_n(\theta)_{k,l} = \text{cov} \left[\frac{\partial \ln f(x, \theta)}{\partial \theta_k}, \frac{\partial \ln f(x, \theta)}{\partial \theta_l} \right]$$

ou bien

$$I_n(\theta)_{k,l} = -E \left[\frac{\partial^2 \ln f(x, \theta)}{\partial \theta_k \partial \theta_l} \right]$$

Après plusieurs simplifications, on obtient tous les éléments :

$$\begin{aligned} \frac{\partial^2 l(\theta)}{\partial \alpha^2} &= -\frac{n}{\alpha^2} \\ \frac{\partial^2 l(\theta)}{\partial \lambda^2} &= -\frac{n}{\lambda^2} - (\alpha - 1) \sum_{i=1}^n \frac{x_i^2 e^{\lambda x_i}}{(1 + e^{\lambda x_i})^2} \\ \frac{\partial^2 l(\theta)}{\partial \alpha \partial \lambda} &= \frac{\partial^2 l(\theta)}{\partial \lambda \partial \alpha} = -\sum_{i=1}^n \frac{x_i e^{\lambda x_i}}{1 + e^{\lambda x_i}} \end{aligned}$$

Les estimateurs du maximum de vraisemblance $\hat{\theta}_n$ des paramètres inconnus de la distribution $TIL - E$ sont calculés sur les données initiales.

3.2.6 Estimation du maximum de vraisemblance pour les données censurées

Soit X_1, X_2, \dots, X_n un échantillon de variable aléatoire *i.i.d* de la loi $TIL - E$, de vecteur de paramètre $\theta = (\alpha, \lambda)^T$. Chaque individu i est supposé avoir une durée de vie x_i et un temps de censure C_i , où X_i et C_i sont des variables aléatoires indépendantes. Supposons que les données sont constituées de n observations indépendantes $x_i = \min(X_i, C_i)$ pour $i = 1, \dots, n$. La distribution de C_i ne dépend pas des paramètres inconnus de X_i .

La fonction de vraisemblance s'écrit :

$$l(x, \theta) = \prod_{i=1}^n f(x_i)^{\delta_i} S(x_i)^{1-\delta_i}, \quad \delta_i = 1_{X_i < C_i}$$

La fonction log-vraisemblance du modèle :

$$L(x, \theta) = \sum_{i=1}^n \delta_i \ln f(x_i) + \sum_{i=1}^n (1 - \delta_i) \ln S(x_i)$$

alors l'équation de la log-vraisemblance devient :

$$\begin{aligned} L(x, \theta) &= \sum_{i=1}^n \delta_i \ln h(x_i) + \sum_{i=1}^n \ln S(x_i) \\ &= \sum_{i=1}^n \delta_i [\ln f(x_i) - \ln S(x_i)] + \sum_{i=1}^n \ln S(x_i) \end{aligned}$$

Ce que donne :

$$L(x, \theta) = \sum_{i=1}^n \delta_i [\ln(\alpha\lambda) + \lambda x_i - \ln(1 + e^{\lambda x_i})] + \sum_{i=1}^n \alpha \ln(2) - \alpha \sum_{i=1}^n \ln(1 + e^{\lambda x_i})$$

Les fonctions de score pour les paramètres α et λ sont données par :

$$\begin{aligned} \frac{\partial L}{\partial \alpha} &= \sum_{i=1}^n \delta_i \left[\frac{1}{\alpha} \right] + n \ln(2) - \sum_{i=1}^n \ln(1 + e^{\lambda x_i}) \\ \frac{\partial L}{\partial \lambda} &= \sum_{i=1}^n \delta_i \left[\frac{1}{\lambda} + x_i - \frac{x_i e^{\lambda x_i}}{1 + e^{\lambda x_i}} \right] - \alpha \sum_{i=1}^n \frac{x e^{\lambda x_i}}{1 + e^{\lambda x_i}} \end{aligned}$$

Pour résoudre ce système de fonctions non linéaires, on fait appel aux méthodes numériques, telles que la méthode de Newton Raphson, l'algorithme Barzilai-Borwein (**BB**) ou d'autres algorithmes.

3.3 La distribution Lomax Inverse Tronquée - Weibull

3.3.1 Fonction de répartition et densité

Soit x une variable aléatoire, la fonction de distribution cumulative (cdf) de la distribution de base (weibull) avec paramètres et est donnée par :

$$G(x) = 1 + e^{-\lambda x^\beta} \quad (9)$$

Pour développer le *CDF* de la distribution Lomax Inverse Tronquée -Weibull (*TIL-W*), l'expressions de l'équation (9) est remplacée dans l'équation (1) comme suit :

$$F(x) = 1 - 2^\alpha \left[1 + e^{\lambda x^\beta} \right]^{-\alpha} \quad (10)$$

La fonction de densité de probabilité est

$$f(x) = \alpha \lambda \beta x^{\beta-1} 2^\alpha e^{\lambda x^\beta} (1 + e^{\lambda x^\beta})^{-1-\alpha} \quad (11)$$

Pour simuler une variable provenant de la distribution *TIL-W*, nous utilisons l'équation non linéaire :

$$x = \left[\frac{1}{\lambda} \ln \left[-1 + \left[2^{-\alpha} (1 - u) \right]^{-\frac{1}{\alpha}} \right] \right]^{\frac{1}{\beta}}$$

Où u suit la distribution uniforme $U(0, 1)$.

Un exemple de la densité $TIL - W$ est présenté dans la figure (2.1).

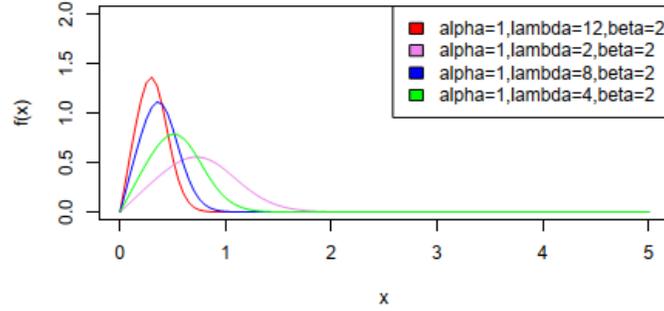


Fig. 2.1. Densité de la distribution $TIL - W$ pour le paramètre de forme $\lambda = (12; 2; 8; 4)$ avec $\alpha = 1$ et $\beta = 2$.

3.3.2 La fonction de survie et le taux de hasard

La fonction de survie du modèle $TIL - W$ est donnée par

$$S(x) = 2^\alpha \left[1 + e^{\lambda x^\beta} \right]^{-\alpha}$$

Sa fonction de hasard est donnée par :

$$h(x) = \frac{\alpha \lambda \beta x^{\beta-1} e^{\lambda x^\beta}}{1 + e^{\lambda x^\beta}}$$

Le graphe du taux de hasard de cette distribution $TIL - W$ est présenté dans le figure (2.2) .

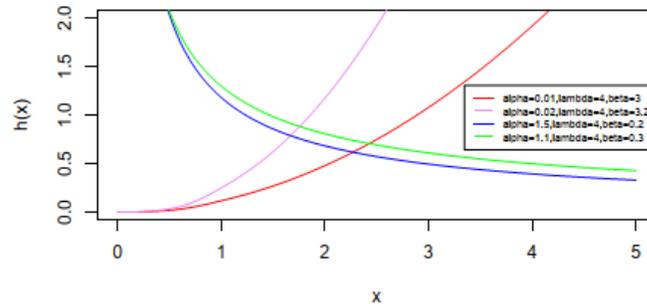


Fig. 2.3. Taux de hasard de la distribution $TIL - W$ pour $\lambda = 4$ avec $\alpha = (0.01; 0.02; 1.5; 1.1)$ et $\beta = (3; 3.2; 0.2; 0.3)$.

3.3.3 Fonction de hasard cumulé

La fonction de hasard cumulé du modèle $TIL - W$ est donné par :

$$\begin{aligned} H(x) &= -\ln\{S(x)\} \\ H(x) &= -\alpha \ln 2 - \alpha \ln \left[1 + e^{\lambda x^\beta}\right] \end{aligned}$$

3.3.4 L'espérance et la variance

L'espérance est définie comme :

$$E(T) = \int_0^{+\infty} t dF(t) = \int_0^{+\infty} S(t) dt = \int_0^{+\infty} 2^\alpha \left[1 + e^{\lambda t^\beta}\right]^{-\alpha} dt$$

Comme pour la distribution de $TIL - W$, le calcul ci-dessus nécessite l'intégration numérique.

La variance peut être obtenue par les formules suivantes :

$$\begin{aligned} V(T) &= E(T^2) - E(T)^2 \\ V(T) &= \int_0^{+\infty} t^2 dF(t) - E(T)^2 = 2 \int_0^{+\infty} t S(t) dt - E(T)^2 \\ V(T) &= 2 \int_0^{+\infty} t (2^\alpha \left[1 + e^{\lambda t^\beta}\right]^{-\alpha}) dt - E(T)^2 \end{aligned}$$

3.3.5 Estimation des paramètres dans le cas de donnée complète

Estimation du maximum de vraisemblance

La fonction de vraisemblance pour un n -échantillon aléatoire (x_1, x_2, \dots, x_n) issu de la distribution(11) est

$$\begin{aligned} L(x, \theta) &= \prod_{i=1}^n f(x_i, \theta) \\ &= \prod_{i=1}^n \left(\alpha \lambda \beta x_i^{\beta-1} 2^\alpha e^{\lambda x_i^\beta} (1 + e^{\lambda x_i^\beta})^{-\alpha-1} \right) \end{aligned}$$

où $\theta = (\alpha, \lambda, \beta)$ vecteur de paramètres

Et l'expression de la fonction log-vraisemblance est :

$$\begin{aligned} l(x, \theta) &= \sum_{i=1}^n \ln f(x_i, \theta) \\ &= n \ln(\alpha) + n \ln(\lambda) + n \ln(\beta) + (\beta - 1) \sum_{i=1}^n \ln(x_i) + n\alpha \ln(2) \\ &\quad + \lambda \sum_{i=1}^n x_i^\beta - (\alpha + 1) \sum_{i=1}^n \ln(1 + e^{\lambda x_i^\beta}) \end{aligned}$$

Pour notre modèle $TIL - W$, les fonctions de score sont obtenues comme suit :

$$\begin{aligned}\frac{\partial L}{\partial \alpha} &= \frac{n}{\alpha} + n \ln(2) - \sum_{i=1}^n \ln(1 + e^{\lambda x_i^\beta}) \\ \frac{\partial L}{\partial \lambda} &= \frac{n}{\lambda} + \sum_{i=1}^n x_i^\beta - (\alpha + 1) \sum_{i=1}^n \frac{x_i^\beta e^{\lambda x_i^\beta}}{1 + e^{\lambda x_i^\beta}} \\ \frac{\partial L}{\partial \beta} &= \frac{n}{\beta} + \lambda \sum_{i=1}^n x_i^\beta \ln(x_i) - (\alpha + 1) \sum_{i=1}^n \frac{\lambda x_i^\beta \ln(x_i) e^{\lambda x_i^\beta}}{1 + e^{\lambda x_i^\beta}} + \sum_{i=1}^n \ln(x_i)\end{aligned}$$

Pour résoudre ce système non linéaire, on utilise n'importe quelle méthode.

La matrice d'information de Fisher

Les éléments de la matrice d'information de Fisher définie par :

$$I_n(\theta)_{k,l} = -E\left[\frac{\partial^2 \ln f(x, \theta)}{\partial \theta_k \partial \theta_l}\right]$$

Après plusieurs simplifications, on obtient tous les éléments :

$$\frac{\partial^2 l(\theta)}{\partial \alpha^2} = -\frac{n}{\alpha^2}$$

$$\frac{\partial^2 l(\theta)}{\partial \lambda^2} = -\frac{n}{\lambda^2} - (\alpha - 1) \sum_{i=1}^n \frac{x_i^{2\beta} e^{\lambda x_i^\beta}}{(1 + e^{\lambda x_i^\beta})^2}$$

$$\frac{\partial^2 l(\theta)}{\partial \beta^2} = -\frac{n}{\lambda^2} + \lambda \sum_{i=1}^n x_i^\beta \ln x_i^2 - (\alpha + 1) \sum_{i=1}^n \frac{\lambda \ln x_i^2 \left(\lambda x_i^{2\beta} e^{\lambda x_i^\beta} + x_i^\beta e^{\lambda x_i^\beta} + x_i^\beta e^{2\lambda x_i^\beta} \right)}{(1 + e^{\lambda x_i^\beta})^2}$$

$$\frac{\partial^2 l(\theta)}{\partial \alpha \partial \lambda} = \frac{\partial^2 l(\theta)}{\partial \lambda \partial \alpha} = -\sum_{i=1}^n \frac{x_i^\beta e^{\lambda x_i^\beta}}{1 + e^{\lambda x_i^\beta}}$$

$$\frac{\partial^2 l(\theta)}{\partial \alpha \partial \beta} = \frac{\partial^2 l(\theta)}{\partial \beta \partial \alpha} = -\lambda \sum_{i=1}^n \frac{x_i^\beta \ln x_i e^{\lambda x_i^\beta}}{1 + e^{\lambda x_i^\beta}}$$

$$\frac{\partial^2 l(\theta)}{\partial \beta \partial \lambda} = \frac{\partial^2 l(\theta)}{\partial \lambda \partial \beta} = \sum_{i=1}^n x_i^\beta \ln x_i - (\alpha + 1) \sum_{i=1}^n \frac{\ln x_i \left(\lambda x_i^{2\beta} e^{\lambda x_i^\beta} + x_i^\beta e^{\lambda x_i^\beta} + x_i^\beta e^{2\lambda x_i^\beta} \right)}{(1 + e^{\lambda x_i^\beta})^2}$$

3.3.6 Estimation des paramètres dans le cas de censure droite

Considérons cette fois-ci le cas où les données peuvent être censurées aléatoirement à droite. Soit x une variable aléatoire distribuée selon une loi $TIL-W$, de vecteur de paramètre $\theta = (\alpha, \lambda, \beta)^T$.

L'équation de la log-vraisemblance est donnée alors par :

$$\begin{aligned} L(x, \theta) &= \sum_{i=1}^n \delta_i \ln h(x_i) + \sum_{i=1}^n \ln S(x_i) \\ &= \sum_{i=1}^n \delta_i [\ln f(x_i) - \ln S(x_i)] + \sum_{i=1}^n \ln S(x_i) \end{aligned}$$

Ce que donne :

$$L(x, \theta) = \sum_{i=1}^n \delta_i \left[\ln(\alpha\lambda\beta) + (\beta - 1) \ln(x_i) + \lambda x_i^\beta - \ln(1 + e^{\lambda x_i^\beta}) \right] + \sum_{i=1}^n \alpha \ln(2) - \alpha \sum_{i=1}^n \ln(1 + e^{\lambda x_i^\beta})$$

Les fonctions de score sont obtenues :

$$\begin{aligned} \frac{\partial L}{\partial \alpha} &= \sum_{i=1}^n \delta_i \left[\frac{1}{\lambda} \right] + n \ln(2) - \sum_{i=1}^n \ln(1 + e^{\lambda x_i^\beta}) \\ \frac{\partial L}{\partial \lambda} &= \sum_{i=1}^n \delta_i \left[\frac{1}{\lambda} + x_i^\beta - \frac{x_i^\beta e^{\lambda x_i^\beta}}{1 + e^{\lambda x_i^\beta}} \right] - \alpha \sum_{i=1}^n \frac{x_i^\beta e^{\lambda x_i^\beta}}{1 + e^{\lambda x_i^\beta}} \\ \frac{\partial L}{\partial \beta} &= \sum_{i=1}^n \delta_i \left[\frac{1}{\beta} + \ln(x_i) + \lambda x_i^\beta \ln x_i - \lambda x_i^\beta \ln(x_i) - \frac{\lambda x_i^\beta \ln(x_i) e^{\lambda x_i^\beta}}{1 + e^{\lambda x_i^\beta}} \right] - \alpha \sum_{i=1}^n \frac{\lambda x_i^\beta \ln(x_i) e^{\lambda x_i^\beta}}{1 + e^{\lambda x_i^\beta}} \end{aligned}$$

Comme les formes analytiques des estimateurs du maximum de vraisemblance des paramètres ne peuvent pas être déduites, alors on passe aux méthodes numériques.

Nous présentons quelques simulations pour différentes tailles d'échantillons afin d'évaluer l'efficacité des estimateurs du maximum de vraisemblance des paramètres inconnus de les deux modèles

4.1 Simulation de la distribution TIL-E

Nous menons une étude par simulations numériques en utilisant le logiciel R.

Différentes méthodes ont été dérivées pour simuler une variable aléatoire, comme la méthode de l'inverse, la méthode du rejet, méthode de composition et bien d'autres encore, à partir de différentes distributions de probabilité dans le domaine des statistiques computationnelles. La technique idéale pour simuler à partir de la distribution $TIL - E$ est la méthode d'inverse.

Nous pouvons simuler la variable aléatoire x donnée par :

$$x = \frac{1}{\lambda} \ln \left[-1 + \left[2^{-\alpha}(1-u) \right]^{\frac{-1}{\alpha}} \right],$$

Où u est une variable aléatoire uniforme dans $[0, 1]$.

Pour cela, nous générons 1000 échantillons de variables aléatoires de différentes tailles $n = 20$, $n = 50$, $n = 150$ et $n = 300$, nous calculons ensuite les valeurs de la moyenne des estimateurs du maximum de vraisemblance simulés α et λ , et leurs erreurs quadratiques moyennes (notées EQM).

Les résultats obtenus sont présentés dans les tableaux suivant :

$N = 1000$	$n = 20$	$n = 50$	$n = 150$	$n = 300$
$\alpha = 0.8$	0.8964	0.8452	0.8267	0.8043
EQM	9.29×10^{-6}	2.05×10^{-6}	7.12×10^{-7}	1.89×10^{-8}
$\lambda = 0.5$	0.5706	0.5283	0.5077	0.4828
EQM	4.99×10^{-6}	8.01×10^{-8}	5.93×10^{-8}	2.92×10^{-7}

Tab.1. La moyenne des estimateurs du maximum de vraisemblance $(\hat{\alpha}, \hat{\lambda})$ des paramètres et leurs erreurs quadratiques moyennes.

$N = 1000$	$n = 20$	$n = 50$	$n = 150$	$n = 300$
$\alpha = 1.2$	0.2234	0.2048	0.1808	0.1680
<i>EQM</i>	5.4823×10^{-7}	2.3698×10^{-8}	3.6714×10^{-7}	1.0217×10^{-6}
$\lambda = 2.8$	1.4543	1.4587	1.4041	1.4006
<i>EQM</i>	2.9576×10^{-6}	3.4536×10^{-6}	1.6878×10^{-8}	4.2982×10^{-10}

Tab.2.Les valeurs moyennes des *EMV* ($\hat{\alpha}, \hat{\lambda}$) des paramètres et leurs erreurs quadratiques moyennes pour les données censurées

On observe que lorsque la taille de l'échantillon augmente, l'erreur quadratique moyenne diminue et elle est inférieure à 0.01(1%) pour un paramètre c'est à dire, la valeur estimée converge toujours vers la vraie valeur des paramètres de la distribution *TIL - E*.

4.2 Simulation de la distribution **TIL-W**

On considère le modèle *TIL - W*. Les données ont été simulées $N = 10,000$ fois (avec les tailles des échantillons $n_1 = 20, n_2 = 50, n_3 = 150$ et $n_4 = 300$). En utilisant le logiciel *R* et l'algorithme Barzilai-Borwein (**BB**) (Ravi, 2009) pour les calculs des moyennes des valeurs simulées des estimateurs du maximum de vraisemblance $\hat{\alpha}, \hat{\lambda}$ et $\hat{\beta}$ des paramètres et leurs erreurs quadratiques moyennes (notées S.ME), nous obtenons les résultats suivants présentés dans la table 3 et 4

$N = 1000$	$n = 20$	$n = 50$	$n = 150$	$n = 300$
$\alpha = 2.9$	3.0801	3.0402	2.9075	2.9068
<i>EQM</i>	3.24×10^{-5}	1.96×10^{-5}	5.74×10^{-8}	4.62×10^{-8}
$\lambda = 1.5$	1.5711	1.5576	1.5275	1.5177
<i>EQM</i>	5.06×10^{-6}	3.32×10^{-6}	7.5812×10^{-7}	3.1478×10^{-7}
$\beta = 4$	4.0344	4.0413	4.0245	4.0145
<i>EQM</i>	1.18×10^{-6}	1.70×10^{-6}	6.01×10^{-7}	2.12×10^{-7}

Tab.3.La moyenne des estimateurs du maximum de vraisemblance ($\hat{\alpha}, \hat{\lambda}, \hat{\beta}$) des paramètres et leurs erreurs quadratiques moyennes.

$N = 1000$	$n = 20$	$n = 50$	$n = 150$	$n = 300$
$\alpha = 0.6$	0.6505	0.6259	0.6241	0.6097
<i>EQM</i>	2.5558×10^{-6}	6.7405×10^{-7}	5.8310×10^{-7}	9.4091×10^{-8}
$\lambda = 3$	3.1425	3.0481	3.0433	3.0295
<i>EQM</i>	2.0318×10^{-5}	2.3187×10^{-6}	1.8752×10^{-6}	8.7534×10^{-7}
$\beta = 3.2$	3.5362	3.2034	3.5163	3.2034
<i>EQM</i>	0.0001	1.1601×10^{-8}	0.0001	1.1791×10^{-8}

Tab.4.Lesvaleursmoyennesdes *EMV* ($\hat{\alpha}, \hat{\lambda}, \hat{\beta}$) des paramètres et leurs erreurs quadratiques moyennes pour les données censurées

Les résultats des simulations confirment le fait que pour les modèles réguliers, les estimateurs du maximum de vraisemblance sont convergents.

Le domaine de la modélisation est riche de diverses distributions statistiques très utiles pour modéliser les données dans des nombreux domaines tels que l'économie, l'ingénierie et les études biologiques. Cependant, dans de nombreuses situations, les distributions classiques ne permettent pas de s'adapter adéquatement aux données réelles.

Dans ce travail, nous avons réussi à présenter le nouveau modèle « TIL-G », ainsi que les différentes fonctions de cette distribution ; la fonction de densité de probabilité, la fonction de répartition, la fonction de survie, la fonction du taux de hasard et du taux de hasard cumulé. Nous avons étudié certaines de ces propriétés mathématiques. Les paramètres du modèle proposé sont estimés par la méthode du maximum de vraisemblance en donnant la forme explicite des fonctions de scores et les éléments de la matrice d'information de Fisher.

Une étude de simulation est réalisée pour étudier la performance des estimateurs de maximum de vraisemblance.

L'idée de générer de nouveaux modèles étendus à partir des modèles classiques suscite actuellement un grand intérêt parmi les chercheurs, plusieurs extensions de certaines distributions de durée de vie bien connues ont été développées au cours des dernières années. Cette évolution est suivie par des nombreuses approches pour générer des nouvelles familles de distributions dans le but d'augmenter les chances de modéliser un grand nombre des données réelles.

- [1] A.A. Teamah, A. El-Bar and A. Salem, Random sum of truncated and random truncated Lindley distribution, the Int. Journal of Applied Mathematics, 23 (2010), 961-971.
- [2] A.FOLDES et L. REJTO : Strong uniforme consistency for nonparametric survival curve estimators for randomly censored data. The Annals of Statistics, p. 122-129, 1981b.
- [3] Abid SH, Abdulrazak RK. [0,1] truncated Fréchet-G generator of distributions. Appl Math. 2017; 7(3) : 51-66.
- [4] Abid, S.H., & Abdulrazak, R.K. (2017). [0,1] truncated Fréchet-G generator of distributions. Applied Mathematics, 7, 51-66.
- [5] Akaike H. (1973). Information theory and an extension of the maximum likelihood principle. Second International Symposium on Information Theory, 267-281.
- [6] Aldahlan, M.A., 2019. Type II truncated Fréchet generated family of distributions. International Journal of applied Mathematics, 7, pp.221-228.
- [7] Aldahlan, M.A., Jamal, F., Chesneau, C., Elgarhy, M. and Elbatal, I., 2019. The truncated Cauchy power family of distributions with inference and applications. Entropy, 22, p.346.
- [8] Analyse des durées de survie avec R : <http://www.stat.nus.edu.sg/stachen/Rsurv.pdf>
- [9] B. W. TURNBULL : The empirical distribution function with arbitrarily grouped, censored and truncated data. Journal of the Royal Statistical Society. Series B (Methodological), 38(3) :290-295, 1976
- [10] Bantan, R.A.R., Chesneau, C., Jamal, F. and Elgarhy, M., 2020. On the analysis of new Covid-19 cases in Pakistan using an exponentiated version of the M family of distributions. Mathematics, 8(953), pp. 1-20.
- [11] Bantan, R.A.R., Jamal, F., Chesneau, C. and Elgarhy, M., 2019. Truncated inverted Kumaraswamy generated family of distributions with applications. Entropy, 21, p. 1089.
- [12] Brito, C.R., Rêgo, L.C., Oliveira, W.R. and Gomes-Silva, F., 2019. Method for generating distributions and classes of probability distributions : The univariate case. Hacettepe Journal of Mathematics and Statistics, 48, pp.897-930.
- [13] Catherine Huber, Cour de Modélisation Biostatistique en Splus :
- [14] Christian JUTTEN. Détection, Estimation, Information. Univ.Grenoble Alpes-Polytech Grenoble. Juillet 2018.

- [15] Diouf A. (2019) . Modélisation des effets de traitements fongiques sur la dynamique de population d'organismes sentinelles, Thes the doctorat en Mathématiques et Applications, Université Assane Seck de Ziguinchor.
- [16] Domanski, C.,& Jędrzejczak, A. (1998). Maximum likelihood estimators of the dagum model parameters. *International Advance in Economic Research*, 4(3), 243-252.
- [17] Droesbeke, J.-J., Fichet, B. & Tassi, Ph. (rédacteurs; 1989) : Analyse statistique des durées de vie. Modélisation des données censurées; éd. Economica.
- [18] Estimation. S. Le Digabel, Ecole Polytechnique de Montréal. MTH2302D. A 2017.
- [19] Hassan, A.S., Sabry, M.A.H and Elsehetry, A.M., 2020. A new family of upper-truncated distributions : Properties and estimation. *Thailand Statistician*, 18(2), pp.196-214.
- [20] Hougaard P.(2000). *Analysis of Multivariate Survival Data*. New York : Springer-Verlag.
- [21] J. Pender, The truncated normal distribution : Applications to queues with impatient customers, *Operations Research Letters* 43 (2015) 40-45.
- [22] Jamal, F., Bakouch, H.S. and Nasir, M.A., 2019. A truncated general-G class of distributions with application to truncated Burr-G family, *RevStat*, in press, <https://hal.archives-ouvertes.fr/hal-01956892/document>.
- [23] Kleiber, C. and Kotz, S., 2003. *Statistical Size Distributions in Economics and Actuarial Sciences*, Vol. 470, New York, Wiley.
- [24] Klein P., Melvin L. M. (2005) *Survival Analysis. Statistics for Biology and Health*. Springer.
- [25] Kumar, D., Singh, U., Sing, S.K. and Mukherjee, S., 2017. The new probability distribution : An aspect to a life time distribution. *Mathematical Sciences Letters*, 6(1), pp.35-42.
- [26] L. Zaninetti and M. Ferraro, On the truncated Pareto distribution with applications, *Cent. Eur. J. Phys.*, 6(1)(2008) 1-6.
- [27] L. Zaninetti, A right and left truncated gamma distribution with application to the stars. *Advanced Studies in Theoretical Physics*, 23 (2014), 1139-1147.
- [28] Lee, E,T.and Wang, J., 2003. *Statistical Methods for Survival Data Analysis*, Vol. 476, New York, Wiley.
- [29] Li Jialiang, Ma Shuangge. (2013) . *Survival Analysis in Medicine and Genetics*, Chapman and Hall/CRC.
- [30] M. Ali and S. Nadarajah, A truncated Pareto distribution, *Compu. Commun.*,30 (2006)1-4.
- [31] M. Lejeune. *Statistique-La théorie et ses applications*. Springer, 2010.
- [32] M.M. Mohie El-Din. A.A. Teamah, A. El-Bar and A. Salem, Random sum of mid truncated Lindley distribution, *Journal of Advanced Research in Statistics and Probability*, 2(1) (2010), 27-36.
- [33] N. BRESLOW et J.CROWLEY : A large sample study of the life table and product limit estimates under random censorship. *The Annals of Statistics*, 2(3) : 437-453, 1974.
- [34] Probabilité. Echantillonnage, estimation. I.U.T,Département d'infomatique. Année 2008-2009.
- [35] S Nadarajah, Some truncated distributions, *Acta Appl. Math.*, 106 (2009) 105-123.

- [36] S. Chattopadhyay, C.A. Murthy and S. K. Pal, Fitting truncated geometric distributions in large scale real world networks, *Theoretical Computer Science* 551 (2014) 22-38.
- [37] S.H. Alkarni, A class of truncated binomial lifetime distributions, *Open Journal of Statistics* 3 (2013), 305-313.
- [38] Saint Pierre Philippe (2015). Intoduction à l'analyse de survie (analyse des durées de vie). Support de cours (perso.math.univ-toulouse.fr/psaintpi/files/2021/04/Cours_Survie_1.pdf).
- [39] SAS / STAT User's Guide :
- [40] Shao, Q.(2002). Maximum likelihood estimation for generalised logistic distributions.
- [41] V. PATILEA et J.-M. ROLIN : Product-limit estimators of the survival function with left or right censored data. *Rap. tech.*, Institut de Statistique, Université Catholique de Louvain,2004.
- [42] W. STUTE et J-L. WANG : The strong law under random censorship. *The Annals of Statistics*, 21(3) :1591-1607, 1993.
- [43] Yadav, A.S., Shukla, S. and Kumari, A., 2019. Statistical inference for truncated inverse Lomax distribution and its application to survival data. *Annals of Data Science*, to appear, DOI : 10.1007/s40745-019-00235-2.
- [44] Yadav, A.S., Sing, S.K. and Singh, U., 2016b. Reliability estimation for inverse Lomax distribution under type-II censored data using markov chain monte carlo method. *International Journal of Mathimatics and Statistics*, 17(1), pp.128-146.
- [45] Yadav, A.S., Singh, S.K. and Singh, U., 2016a. On hybrid censored inverse Lomax distribution : Application to the survival data. *Statistica*, 76(2), pp.185-203.
- [46] [http : // www.math.wpi.edu / saspdf / stat / pdf idx.htm](http://www.math.wpi.edu/saspdf/stat/pdfidx.htm)
- [47] [http : //www.biommedicale.univ-paris5.fr / survie / enseign / cours_stat_avec_splus.pdf](http://www.biommedicale.univ-paris5.fr/survie/enseign/cours_stat_avec_splus.pdf)