



جامعة باجي مختار -
عنابة

BADJI MOKHTAR - ANNABA UNIVERSITY
UNIVERSITÉ BADJI MOKHTAR - ANNABA



كلية العلوم

قسم الرياضيات

DEPARTMENT OF MATHEMATICS
DEPARTEMENT DE MATHEMATIQUES

THESE

Présentée en vue de l'obtention du diplôme de
Doctorat en Sciences
Option : Statistique

Sur la distribution de Poisson-Quasi Lindley et ses applications

Par:
GRINE RAZIKA

Sous la direction de

ZEGHDOUDI Halim

M.C.A

U.B.M. Annaba

Devant le jury

PRESIDENT :

CHADLI Assia

Prof

U.B.M. Annaba

EXAMINATEUR :

BOUZEBDA Salim

Prof.

U TC Compiègne

EXAMINATEUR :

BRAHIMI Brahim

Prof

U. Biskra

EXAMINATEUR :

YAHIA Djabrane

MCA

U. Biskra

Année 2018

Remerciements

*Je tiens à adresser mes remerciements les plus chaleureux et ma profonde gratitude à mon encadreur Monsieur **Zeghdoudi.Halim** de l'université d'Annaba, pour m'avoir proposé le sujet et pour l'encadrement, me poussant toujours à faire plus d'efforts*

*Mes remerciements vont également à Madame **Chadli Assia** Professeure de L'université d'Annaba, qui a bien voulu présider le jury.*

*De même je remercie vivement Monsieur **Bouzebda.Salim** professeur de L'université de Compiègne et Monsieur **Brahimi Brahim**, Monsieur **DjabraneYahia** de l'université Biskra, qui ont acceptées d'examiner ce travail.*

Enfin, je n'oublié pas de remercier toutes les personnes qui mont facilité la tâche

Dédicace

A la mémoire de mes parents

A toute ma famille

A tous mes amis

Table des matières

Table des figures	3
0.1 Introduction générale	7
1 Notions et quelques lois de probabilités	10
1.1 Définitions	10
1.1.1 Variable aléatoire	10
1.1.2 Loi de probabilité	11
1.1.3 Moment ordinaire	12
1.1.4 Moment centré	12
1.1.5 Fonction génératrice des moments	12
1.2 Estimation	13
1.2.1 Définition	13
1.2.2 Propriétés d'un estimateur	14
1.2.3 Normalité asymptotique	15
1.2.4 Construction d'estimateurs	15
1.3 Fonction de survie	17
1.4 Taux de hasard	17
1.5 Fonction W de Lambert	18
1.6 Fonction Quantile	20
1.7 Lois de probabilités usuelles	20
1.7.1 Loi de Poisson	20
1.7.2 Loi Exponentielle	21
2 Distribution de Lindley et ses applications	23
2.1 Distribution de Lindley	23
2.1.1 Moments et mesures connexes	25
2.1.2 Fonction de hasard et fonction de survie	26
2.1.3 Fonction Quantile de la distribution Lindley	27
2.1.4 Estimation	29
2.1.5 Simulation	31

2.2	Distribution Poisson–Lindley discrète	34
2.2.1	Moments et Mesures Connexes	35
2.2.2	Fonction Quantile de la distribution de Poisson Lindley	35
2.2.3	Estimation	37
2.2.4	Simulation	40
2.3	Distribution Quasi Lindley	42
2.3.1	Fonction de hasard et fonction de survie	44
2.3.2	Fonction Quantile de la distribution Quasi Lindley	45
2.3.3	Estimation	46
2.4	Distribution de Deux-Paramètres de Lindley	47
3	Poisson–Quasi Lindley distribution et ses applications	50
3.1	Distribution poisson quasi-Lindley	50
3.2	La Fonction de survie et la Fonction de hasard	54
3.3	Estimations	54
3.3.1	Estimation par le maximum de vraisemblance	54
3.3.2	Estimations par la méthode des moments	56
3.4	Fonction quantile de la distribution poisson quasi-Lindley	57
3.5	Illustrations et exemple	59
3.6	Conclusion et Perspectives	64
	Bibliographie	65

Table des figures

1.1	Les deux branches de la fonction de Lambert[3]	19
3.1	Représentation graphique de la fonction de masse de PQL.	53
3.2	Présentation graphique de la fonction de répartition de la distribution PQL.	53
3.3	Comparaison des différents modèles	63

Abstract

In this thesis, we give a treatment of the mathematical properties for new distribution named poisson quasi Lindley (PQLD), by mixing Poisson and quasi Lindley distributions. The properties studied include: moments, the quantile function, maximum likelihood estimation. Simulation studies and data driven applications.

Keywords: Lindley distribution, Poisson distribution, Poisson-Lindley distribution, Quasi- Lindley distribution, maximum-likelihood estimation.

Résumé

Dans cette thèse on se propose de présenter une nouvelle distribution (discrètes) appelée distribution poisson quasi Lindley, par le mélange des distributions poisson et quasi Lindley. Des propriétés ont été étudiées comme l'estimation par la méthode du moment et maximum de vraisemblance, fonction quantile, des simulations ont été appliquées.

Mots clé : distribution Lindley, distribution Poisson, distribution Poisson-Lindley , distribution Quasi- Lindley , estimation de maximum de vraisemblance

ملخص

في هذه الأطروحة، نقترح توزيع جديدة متقطع يسمى ببواسون تنائي ليندلي و ذلك بالمزج بين توزيع بواسون و توزيع تنائي ليندلي. بعض الخصائص تمت دراستها مثل التباين، الربيعيات والوسط الحسابي بطريقة العزوم و طريقة الارجحية العظمى، في الاخير تطبيق على بعض المعطيات .

الكلمات المفتاحية, توزيع ليندلي, توزيع بواسون, توزيع بواسون ليندلي, توزيع تنائي ليندلي, التقدير بطريقة الارجحية العظمى.

0.1 Introduction générale

les distributions statistiques sont couramment appliquées pour décrire des phénomènes réels et sont les plus souvent utilisées dans différents domaines tels que la médecine, la finance, les sciences du génie biologique et la science actuarielle. La distribution de Lindley à un paramètre est utilisée pour modéliser les données de durée de vie..

Soit X une variable aléatoire qui suit la distribution d'un seul paramètre donné par la fonction de densité suivante :

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

Introduite par **Lindley (1958)**. Sankaran (**1970**) composé (1) avec la distribution poisson pour trouver une nouvelle distribution nommée Poisson- Lindley discrète. Récemment, Asgharzadeh et al. (**2013**), Ghitany et al. (**2008a**) et (**2008b**) ont introduit et étudié une nouvelle distribution ressemble autre à la distribution Lindley, dont sa dérivé est connue comme Zero-truncated Poisson- Lindley et Pareto Poisson-Lindley distributions.

Mahmoudi et Zakerzadeh (**2010**) propose une nouvelle version de mélange de la distribution poisson nommée poisson Lindley généralisée, qui est obtenue par mélange de la distribution poisson avec la distribution Lindley généralisée qui a été obtenue et analysé par **Zakerzadeh et Dolati (2009)** qui ont utilisé gamma (α, θ) et gamma $(\alpha + 1, \theta)$ pour créer une distribution de **Lindley** généralisée.

Shanker et Mishra (**2013**), ont introduit une nouvelle distribution (quasi Lindley distribution) basées sur le mélanges de la distribution gamma $(2, \theta)$ et exponentiel (θ) , où la fonction de densité pour la variable aléatoire X est donnée par :

$$f_{QL}(x; \alpha, \theta) = \begin{cases} \frac{\theta(\alpha + \theta x)e^{-\theta x}}{\alpha + 1}, & x, \theta > 0, \alpha \succ -1 \\ 0, & \text{ailleurs.} \end{cases} \quad (2)$$

Récemment, Zeghdoudi et Nedjar (**2016a, 2016b**) propose de nouvelles version pour la mixture de la distribution poisson nommées distribution poisson gamma Lindley (PGaL) et poisson pseudo Lindley (PPL) et pseudo Lindley distribution (PL) (**2017**), par mélange de la distribution Poisson et gamma Lindley.

Zeghdoudi et Lazri (2016,2017) ont introduit une autre forme de mélange à savoir :

- On lindley-Pareto Distributions ;
- Composite Gamma(2, λ)-Pareto Distribution.

Dans cette thèse, on introduit une nouvelle distribution de durée de vie en utilisant les modèles de mélange et en combinant la distribution de Poisson et la distribution quasi Lindley à deux paramètres qui apportera un plus à la littérature existante sur la modélisation des données de survie, des sciences biologiques et des sciences actuarielles.

Dans le premier chapitre nous rappelons certaines définitions et certains résultats que nous utiliserons par la suite. Ce rappel comporte des généralités sur quelques distributions de probabilités, fonction W de Lambert, fonction quantile, estimation MM et MV.

Dans le chapitre II, nous faisons une collecte des résultats obtenus sur les distributions de Lindley et Poisson Lindley et quasi Lindley en nous inspirant des travaux de Lindley (1958), Ghitany et al. (2008,a) et Sankaran (1970). Shanker, R., & Mishra, A. (2013). Le chapitre III comporte la nouvelle distribution à deux paramètres (Poisson quasi- Lindley) dont on donne quelques propriétés à savoir : la fonction quantile, méthode des moments, estimation du maximum de vraisemblance. Plusieurs simulations sont établies pour examiner le biais et l'erreur quadratique moyenne des estimateurs des paramètres obtenus par la méthode du maximum de vraisemblance et leurs applications dans l'analyse de survie. Une conclusion générale et une bibliographie viendront finir ce travail.

Chapitre 1

Notions et quelques lois de probabilités

1.1 Définitions

Soit X une application définie de Ω dans \mathbb{R} . Soit B un sous ensemble de \mathbb{R} . On appelle image inverse de B par X , la partie de Ω noté $X^{-1}(B)$, notée par :

$$X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\}.$$

1.1.1 Variable aléatoire

La notion la plus utilisée dans la théorie des probabilités et de la statistique c'est la variable aléatoire.

Variable aléatoire réelle

Soit (Ω, \mathcal{F}, P) un espace probabilisé. l'application X définie de Ω dans \mathbb{R} est une variable aléatoire vérifiée :

$$X : \omega \in \Omega \rightarrow X(\omega) \in \mathbb{R}.$$

$$\forall x \in \mathbb{R} : \{X \leq x\} \in \mathcal{F}.$$

\mathcal{F} étant une tribu.

La variable aléatoire réelle X est dite discrète si l'espace probabilisé est dénombrable, sinon la variable aléatoire réelle X est dite continue.

1.1.2 Loi de probabilité

Définition 1.1 On définit la loi de probabilité de la v.a X , par l'application P_X telle que :

$$P_X(X = x) = P(\{\omega \in \Omega : X(\omega) = x\}).$$

Définition 1.2 La fonction de répartition d'une variable aléatoire X , est la fonction F définie pour toute réel x par :

$$F_X : \mathbb{R} \rightarrow [0, 1].$$

$$x \rightarrow F_X(x) = P(X \leq x) = P\{\omega \in \Omega / X(\omega) \leq x\}.$$

Loi de probabilité d'une variable aléatoire discrète

La loi de probabilité d'une variable aléatoire discrète est entièrement déterminée par les probabilités P_i des événements $\{X = x_i\}$, x_i parcourant l'univers image $X(\Omega)$. La loi de probabilité est donnée par les (x_i, p_i) .

1.1.3 Moment ordinaire

Le moment ordinaire d'ordre $r \in \mathbb{N}$ de X est défini, s'il existe, par :

$$\mu_r' = E(X^r) = \begin{pmatrix} \sum_{X \in I} x^r P[X = x] & \text{si } X \text{ est discrète} \\ \int_{x \in I} x^r f(x) dx & \text{si } X \text{ est continue} \end{pmatrix}$$

1.1.4 Moment centré

Le moment centré d'ordre $r \in \mathbb{N}$ de X est défini, s'il existe, par :

$$\mu_r = E([X - E(X)]^r) = \begin{pmatrix} \sum_{X \in I} [x - E(X)]^r P[X = x] & \text{si } X \text{ est discrète} \\ \int_{x \in I} [X - E(X)]^r f(x) dx & \text{si } X \text{ est continue} \end{pmatrix}$$

1.1.5 Fonction génératrice des moments

Pour la variable aléatoire X la fonction génératrice des moments est définie par :

$$H_X(u) = E(e^{Xu}), u \in \mathbb{R}.$$

La notion de fonction génératrice peut être utile parfois pour calculer plus facilement les moments de certaines lois de probabilité telle que :

l'espérance et la variance de X :

$$E(X) = H'_X(0).$$

$$Var(X) = E(X^2) - E^2(X) = H''_X(0) - \left[H'_X(0) \right]^2.$$

Le coefficient de variation (γ), le coefficient d'asymétrie ($\sqrt{\beta_1}$) et le coefficient d'aplatissement (β_2) sont :

$$\begin{aligned} \gamma &= \frac{\sqrt{Var(X)}}{\mathbb{E}(X)}. \\ \sqrt{\beta_1} &= \frac{\mathbb{E}(X^3)}{(Var(X))^{\frac{3}{2}}}. \\ \beta_2 &= \frac{\mathbb{E}(X^4)}{(Var(X))^2}. \end{aligned}$$

1.2 Estimation

1.2.1 Définition

Un estimateur de θ est une application T_n de E^n dans F qui à chaque échantillon (x_1, x_2, \dots, x_n) de la loi P_θ associe une variable aléatoire réelle dont on peut déterminer la loi de probabilité, on appelle la valeur $\hat{\theta}$ estimateur ou estimation :

$$\hat{\theta} = T_n(x_1, x_2, \dots, x_n).$$

1.2.2 Propriétés d'un estimateur

Biais

Pour pouvoir considérer T_n comme une valeur approchée de θ , il faut que les valeurs prises par la variable T_n ne s'écartent pas trop de la valeur fixe de θ .

$$B_n(\theta) = E(T_n) - \theta$$

Un estimateur T_n de θ est dit sans biais si pour tout θ , $B_n(\theta) = 0$, et si $B_n(\theta) > 0$, l'estimateur est dit positivement biaisé.

Convergence

Un estimateur T_n est convergent si la suite de v.a. (T_n) converge en probabilité vers θ , soit :

$$\begin{aligned} T_n \xrightarrow{P} \theta &\iff \lim_{n \rightarrow \infty} P(|T_n - \theta| < \varepsilon) \rightarrow 1, n \rightarrow \infty; \forall \varepsilon > 0 \\ &\iff \lim_{n \rightarrow \infty} P(|T_n - \theta| > \varepsilon) \rightarrow 0. \end{aligned}$$

Théorème 1.3 *Théorème 1.4* Tout estimateur sans biais dont la variance tend vers 0 est convergent :

$$(E_\theta(T_n) = \theta \text{ et } Var_\theta(T_n) \rightarrow 0) \Rightarrow T_n \xrightarrow{P} \theta; n \rightarrow \infty.$$

Erreur quadratique moyenne

La qualité d'un estimateur est mesurée à l'aide d'une distance moyenne au paramètre qu'on appelle l'erreur quadratique moyenne (*Mean Squared Error en anglais*)

appelée aussi risque quadratique est l'espérance du carré de l'erreur entre la vraie valeur et sa valeur estimée.

$$EQM(T_n) = E_\theta((T_n - \theta)^2).$$

Si le risque est faible, l'estimateur T_n est proche de θ .

1.2.3 Normalité asymptotique

On applique le théorème central limite suivant

Théorème 1.5 (Théorème central limite). *Soient une suite de variables aléatoires X_1, X_2, \dots, X_n indépendantes qui suit la même loi (donc de même espérance m et de même écart-type σ).*

$$Z_n = \frac{\bar{X}_n - m}{\frac{\sigma}{\sqrt{n}}} \rightsquigarrow N(0, 1).$$

La variable aléatoire Z_n converge en loi vers la loi normale centrée réduite.

1.2.4 Construction d'estimateurs

Méthode du maximum de vraisemblance

Définition 1.6 *On appelle vraisemblance (likelihood) de l'échantillon (X_1, X_2, \dots, X_n)*

la loi de probabilité de ce n-uple, notée $L(x_1, \dots, x_n; \theta)$ est définie par

$$L(x_1, \dots, x_n; \theta) = \prod_{i=1}^n P(X_i = x_i \mid \theta)$$

Si X est une variable discrète, et par :

$$L(x_1, \dots, x_n; \theta) = \prod_{i=1}^n f(x_i; \theta)$$

Si X est une variable aléatoire continue de densité $f(x; \theta)$

Définition 1.7 Un estimateur de maximum de vraisemblance (emv) de θ est la valeur $\hat{\theta}_n$ de θ qui rend maximise la fonction de vraisemblance $L(x_1, \dots, x_n; \theta)$. Donc $\hat{\theta}_n$ sera en général calculé en maximisant la ln-vraisemblance :

$$\hat{\theta}_n = \max \ln L(\theta; x_1, \dots, x_n).$$

La recherche de l'estimateur de vraisemblance peut se faire directement par la recherche du maximum de L , ou dans le cas particulier où la fonction L est deux fois dérivable par rapport à θ comme solution de l'équation suivante :

$$\begin{aligned} \frac{\partial}{\partial \theta} \log L(\theta; x_1, \dots, x_n) &= 0. \\ \text{où } \frac{\partial^2}{\partial \theta^2} \log L(\theta; x_1, \dots, x_n) &< 0. \end{aligned}$$

Dans ce cas, on les résout numériquement.

Méthode des Moments

Dans le cas où le paramètre à estimer est $\theta = E_\theta(X)$ moyenne théorique de la loi, l'estimateur naturel est la moyenne empirique

$$\hat{\theta}_n = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i.$$

de même pour estimer le paramètre $\theta = V_\theta(X)$ variance de X , l'estimateur naturelle serait la variance empirique

$$S_n'^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2.$$

Plus généralement, pour $\theta \in \Theta$, si $E(X) = \varphi(\theta)$, où φ est une fonction inversible, alors l'estimateur de θ par la méthode des moments est :

$$\hat{\theta}_n = \varphi^{-1}(\theta).$$

1.3 Fonction de survie

La fonction de survie est la probabilité pour que l'événement d'intérêt (le décès par exemple) T intervienne après un délai supérieur à t , autrement dit, que l'événement d'intérêt T ne survienne pas avant la date t , c'est-à-dire

$$S(t) = P(T > t) = 1 - F_T(t), \quad t \geq 0. \quad (1.1)$$

1.4 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) :

$$h(t) = \frac{f(t)}{S(t)} = \frac{f(t)}{1 - F_T(t)} . \quad (1.2)$$

1.5 Fonction W de Lambert

Lambert en 1758 définit une fonction complexe multivaluée appelé fonction de Lambert (W) qui est la solution de l'équation suivante

$$W(z) \exp(W(z)) = z. \quad (1.3)$$

Où z est un nombre complexe. Si z est un nombre réel sachant que $z \geq -1/e$ alors $W(z)$ devient une fonction réelle et elle a deux branches réelles possibles.

La première branche réelle qui a des valeurs dans l'intervalle $]-\infty, -1]$ c'est la branche négative et notée par W_{-1} . et la deuxième branche réelle qui a des valeurs dans l'intervalle $[-1, \infty[$ est appelée la branche principale et notée W_0 . Les deux réelles branches de W sont représentés par la figure.1.1. Lémeray [3] a montré qu'il ya d'autres équations qui peuvent être résolues en termes de la fonction (1.3) W de Lambert. À cet effet, le lemme suivant, sera essentiel dans notre travail.

Lemme 1.8 (Lémeray et al.[3]) *Soient a, b et c des nombres complexes fixés. Par rapport à $z \in \mathbb{C}$ la solution de l'équation $z + ab^z = c$ est :*

$$z = c - \frac{1}{\log(b)} W(ab^c \log(b)).$$

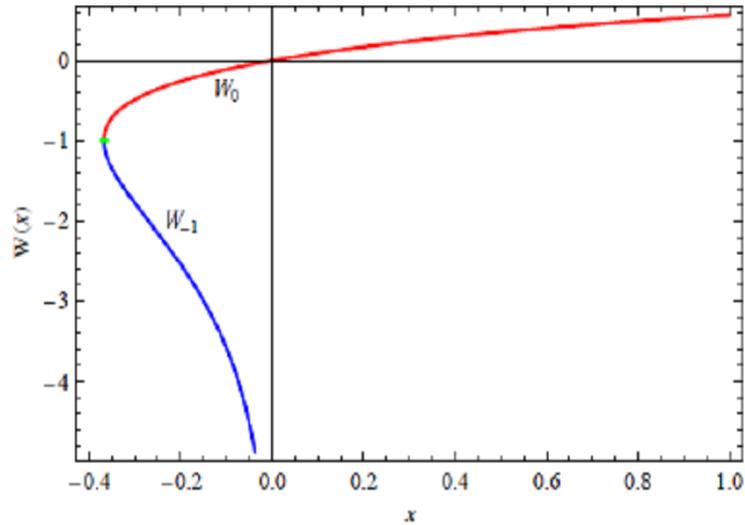


FIG. 1.1: Les deux branches de la fonction de Lambert[3]

Où W est la fonction de Lambert.

Démonstration. Pour tous nombres complexes a, b et c fixés, par rapport à la variable complexe z on doit résoudre l'équation $z + ab^z = c$. Pour cela il faut multiplier les deux côtés de cette équation par $b^c \log(b)$, alors l'équation résultante peut être écrite sous cette forme :

$$(c - z) \log(b) \exp((c - z) \log(b)) = ab^c \log(b). \quad (1.4)$$

Comparent l'équation (1.4) par $b^c \log(b)$ avec l'équation (1.3). Il est clair que $(c - z) \log(b)$ est la fonction W de Lambert de l'argument complexe $ab^c \log(b)$. Par conséquent, on a :

$$W(ab^c \log(b)) = (c - z) \log(b).$$

Ce qui implique le résultat souhaité. Ceci termine la preuve du lemme 1.1. ■

1.6 Fonction Quantile

La fonction quantile d'une variable aléatoire (ou d'une loi de probabilité) est l'inverse de sa fonction de répartition. On appelle fonction quantile de X la fonction, notée Q_X , de $]0, 1[$ dans \mathbb{R} , qui à $u \in]0, 1[$ associe :

$$Q_X(u) = F^{-1}(u) \quad 0 < u < 1. \quad (1.5)$$

1.7 Lois de probabilités usuelles

Dans cette partie on définit quelques lois usuelles les plus utilisées dans notre travail.

1.7.1 Loi de Poisson

La loi de Poisson est une distribution discrète très utile dans l'étude de la survenue dans le temps d'événements homogènes (le nombre d'absents par jour dans une entreprise, le nombre de clients dans une file d'attente durant des laps de temps de même durée).

Une variable aléatoire suit une loi de Poisson de paramètre $\lambda \in \mathbb{R}_+^*$ (qui est à la fois

la moyenne et la variance) si :

$$P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}; \quad k \in \mathbb{N}.$$

L'espérance mathématique et la variance de X :

$$E[X] = var[X] = \lambda.$$

1.7.2 Loi Exponentielle

Une loi exponentielle modélise la durée de vie d'un phénomène sans mémoire, ou 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 :

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

On note $X \rightsquigarrow EXP(\lambda)$.

La fonction de répartition de X :

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

L'espérance mathématique (ou durée de vie moyenne) et la variance de X :

$$E[X] = \frac{1}{\lambda}, var[X] = \frac{1}{\lambda^2}.$$

L'estimateur $\hat{\lambda}_{MoM}$ de paramètre λ obtenu par la méthode des moments est :

$$\hat{\lambda}_{MoM} = \frac{1}{E[X]}.$$

Le logarithme de la vraisemblance d'un échantillon issu d'une loi exponentielle est donné par :

$$\log L(x_i, \alpha) = n \log \lambda - \lambda \sum_{i=1}^n \log x_i.$$

Chapitre 2

Distribution de Lindley et ses applications

La distribution de Lindley d'un seul paramètre est introduit par Lindley en 1958 comme mélange des deux distribution $\text{Exp}(\theta)$ et $\text{Gamma}(2, \theta)$. Plus de détails sur la distribution de Lindley peut être trouvés en Ghitany et autres [15]. Cette distribution traitée par plusieurs chercheurs pour son usage en modélisent des données de vie, et on l'a observé en plusieurs articles que cette distribution a bien étudié .

2.1 Distribution de Lindley

Soient Y_1 et Y_2 deux variables aléatoires indépendantes. Pour $\theta > 0$, on considère la variable aléatoire $X = Y_1$ et $X = Y_2$ avec les probabilités respectivement $P_1 = \frac{\theta}{1+\theta}$ et $P_2 = \frac{1}{1+\theta}$. Où $Y_1 \sim \text{exp}(\theta)$ et $Y_2 \sim \text{Gamma}(2, \theta)$.

La distribution lindley est spécifiée par la fonction densité suivante :

$$f(x; \theta) = \frac{\theta^2(1+x)e^{-\theta x}}{1+\theta}; \quad x, \theta > 0 \quad (2.1)$$

Où la fonction de répartition est donné par :

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

La dérivée première 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-à-d . $f(x)$ diminue en x .

Remarque 2.1 *Le mode de cette distribution est :*

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

Théorème 2.2 *Soit $X \sim \text{Lindley}(\theta)$. Alors,*

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

Démonstration. Soient $M = Mode(X)$, $m = Median(X)$ et $\mu = E(X) = \frac{\theta+2}{\theta(\theta+1)}$. A partir 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

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

Notons que $F(M)$ est une fonction décroissante en $\theta \in (0, 1)$ et, pour tout $\theta > 0$, $0 \leq F(M) < 1 - 2e^{-1} < (1/2)$.

De même, $F(\mu)$ est une fonction croissante en $\theta > 0$ et $(1/2) < 1 - 3e^{-2} < F(\mu) < 1$.

Enfin, étant donné que $F(x)$ est une fonction croissante en $x > 0$ pour tout $\theta > 0$. on

a $M < m < \mu$. ■

2.1.1 Moments et mesures connexes

Le r^{ieme} moment de la distribution de Lindley est :

$$\mu_r' = E(X^r) = \frac{r!(\theta + r + 1)}{\theta^r(\theta + 1)}, r = 1, 2, \dots$$

En particulier 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 de la distribution de Lindley est donné par :

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

En particulier on a

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

Le coefficient de variation (γ), le coefficient de d'asymétrie ($\sqrt{\beta_1}$) et le coefficient d'aplatissement (β_2) sont :

$$\begin{aligned}\gamma &= \frac{\sqrt{\theta^2 + 4\theta + 2}}{\theta + 2}, \\ \sqrt{\beta_1} &= \frac{2(\theta^3 + 6\theta^2 + 6\theta + 2)}{(\theta^2 + 4\theta + 2)^{\frac{3}{2}}}, \\ \beta_2 &= \frac{3(3\theta^4 + 24\theta^3 + 44\theta^2 + 32\theta + 8)}{(\theta^2 + 4\theta + 2)^2}.\end{aligned}$$

2.1.2 Fonction de hasard et fonction de survie

Remarque 2.3 Pour la distribution Lindley la fonction de taux de hasard est :

$$h(x) = \lim_{\Delta x \rightarrow 0} \frac{P(X < x + \Delta x \mid X > x)}{\Delta x} = \frac{f(x)}{1 - F(x)} = \frac{\theta^2(1 + x)}{\theta + 1 + \theta x}. \quad (2.3)$$

Et la fonction de survie est :

$$S(x) = 1 - F(x) = \frac{\theta + 1 + \theta x}{1 + \theta} e^{-\theta x}. \quad (2.4)$$

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

$$\frac{\theta^2}{\theta+1} < h(x) < \theta.$$

2.1.3 Fonction Quantile de la distribution Lindley

La distribution Lindley de paramètre θ est spécifiée par sa fonction de répartition définie en (2.2). Il convient de noter qu'elle est continue et strictement croissante de sorte que la fonction de quantile X est $Q_X(u) = F_X^{-1}(u)$, $0 < u < 1$. 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 :*

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

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

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

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

En multipliant par $-\exp(-\theta - 1)$ l'équation (2.6), on obtient :

$$-(\theta + 1 + \theta x) \exp(-\theta - 1 - \theta x) = (\theta + 1)(u - 1) \exp(-\theta - 1). \quad (2.7)$$

D'après l'équation (2.7), conjointement avec l'équation (1.3), on voit que $-(\theta + 1 + \theta x)$ est la fonction W de Lambert de l'argument réel $(\theta + 1)(u - 1) \exp(-\theta - 1)$.

Alors, on a :

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

■

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 $u \in (0, 1)$, en prenant en compte les propriétés de la branche négative de la fonction W de Lambert a présenté en première chapitre, l'équation (2.8) devient :

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

Ce qui implique le résultat. La preuve du théorème est terminée ■.

2.1.4 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 :

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

Le théorème suivant montre que l'estimateur $\hat{\theta}_{MoM}$ de θ est biaisé.

Théorème 2.5 *L'estimateur $\hat{\theta}_{MoM}$ de θ est positivement biaisée, i.e. $E(\hat{\theta}) - \theta > 0$.*

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

Comme $g''(t) = \frac{1}{t^3} \left[1 + \frac{3t^3 + 15t^2 + 9t + 1}{[(t-1)^2 + 8t]^{\frac{3}{2}}} \right] > 0$, $g(t)$ 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

$$E(\hat{\theta}_{MoM}) > \theta.$$

■ ■

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 :*

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

L'intervalle de confiance de θ pour un seuil de confiance $100(1 - \alpha)\%$ est donné 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.

Démonstration. Étant donné que 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 :

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

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

$$\sqrt{n} (g(\bar{X}) - g(\mu)) \xrightarrow{P} N \left(0, [g'(\mu)]^2 \sigma^2 \right).$$

Enfin, étant donné que :

$$g(\bar{X}) = \hat{\theta}_{MoM}, g(\mu) = \theta, \quad \text{et} \quad g'(\mu) = \frac{-1}{2\mu^2} \left[1 + \frac{1 + 3\mu}{\sqrt{(\mu - 1)^2 + 8\mu}} \right] = -\frac{1}{\sigma^2}.$$

■

Estimation par la méthode du maximum de vraisemblance

Soient $X_i \sim LD(\theta)$, $i = \overline{1, n}$ n variables aléatoires. La fonction de logvraisemblance est :

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

L'estimateurs 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

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

Remarque2.3 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.

2.1.5 Simulation

Cette section étudie le comportement des estimateurs du maximum de vraisemblance $\hat{\theta}_{MV}$ pour un échantillon de taille finie (n). La simulation est réalisée pour chaque couple $(\theta; n)$, où $\theta = 0.1, 1, 9$ et $n = 20, 40, 60, 80, 100$. Alors on a l'algorithme suivant :

- Choisir les valeurs initiales de θ_0 pour spécifier la distribution de Lindley ;

- Choisir la taille de l'échantillon n ;
- Générer N échantillons indépendants de taille n de $LD(\theta)$;
- Calculer les estimations $\hat{\theta}_{MV}$ de θ pour chacun des N échantillons ;
- Calculer :

(i) La moyenne des estimateurs obtenus sur tous les N échantillons

$$\text{biais moyen}(\theta) = \frac{1}{N} \sum_{i=1}^N (\hat{\Theta}_i - \Theta_0) .$$

(ii) L'erreur quadratique moyenne EQM des estimations simulées

$$EQM(\theta) = \frac{1}{N} \sum_{i=1}^N (\hat{\Theta}_i - \Theta_0)^2 .$$

Tableau 2.1 : les biais moyens de l'estimateur $\hat{\theta}$

n	$\theta = 0.1$	$\theta = 1$	$\theta = 9$
20	0.00265	0.03068	0.41459
40	0.00118	0.01699	0.18019
60	0.00090	0.01060	0.13665
80	0.00061	0.00746	0.09073
100	0.00058	0.00570	0.07728

Tableau 2.2 : la moyenne EQM de l'estimateur $\hat{\theta}$

n	$\theta = 0.1$	$\theta = 1$	$\theta = 9$
20	0.00028	0.03335	4.25425
40	0.00013	0.01543	1.91203
60	0.00009	0.01024	0.22017
80	0.00007	0.00752	0.90044
100	0.00005	0.00590	0.70490

Remarque 2.7 (i) *Le tableau 2.1 présente un biais positif, comme indiqué dans le théorème 2.3. Le tableau montre également que le biais diminue (augmente) quand $n(\theta)$ augmente respectivement.*

(ii) *Le tableau 2.2 montre que l'erreur quadratique moyenne diminue (augmente) lorsque $n(\theta)$ augmente respectivement.*

2.2 Distribution Poisson–Lindley discrète

Une distribution composée de Poisson peut être obtenue en composant la distribution de Poisson et une distribution due à Lindley. Cette distribution a été introduit par Sankaran [14] pour modéliser des données de comptage.

Supposons que le paramètre λ de la distribution de Poisson a une distribution appartenant à la famille exponentielle de distribution donnée par :

$$dF(\lambda) = e^{\lambda\Phi} h(\lambda) B(\Phi) d\lambda, \text{ où } h(\lambda) = 1 + \lambda \text{ et } B(\Phi) = \frac{[-\Phi]^2}{(1-\Phi)}.$$

Alors la distribution de Poisson composée est :

$$\begin{aligned} P_x(\Phi) &= \int_0^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} dF(\lambda) \\ &= \frac{B(\Phi)}{x!} \left[\int_0^{\infty} e^{(\Phi-1)\lambda} \lambda^x d\lambda + \beta \int_0^{\infty} e^{(\Phi-1)\lambda} \lambda^{x+1} d\lambda \right] \\ &= \frac{\Phi^2}{(1-\Phi)} \left(\frac{1-\Phi+x+1}{(1-\Phi)^{x+2}} \right). \end{aligned}$$

Alors on remplace Φ par $-\theta$ on trouve :

$$P_x(\theta) = \theta^2 \frac{(x+2+\theta)}{(\theta+1)^{x+3}}.$$

La fonction de densité de Poisson-Lindley (*PLD*) est :

$$f_{PLD}(x; \theta) = P_x(\theta) = \theta^2 \frac{(x+2+\theta)}{(\theta+1)^{x+3}}, \quad x = 0, 1, \dots, \theta > 0. \quad (2.11)$$

La fonction de répartition correspondante est :

$$F_{PLD}(x) = 1 - \frac{\theta^2 + 3\theta + 1 - \theta x}{(\theta+1)^{x+3}}, \quad x = 0, 1, \dots, \theta > 0. \quad (2.12)$$

La fonction génératrice de Poisson-Lindley (*PLD*) est :

$$M_X(s) = E(e^{sX}) = \frac{\theta^2}{\theta + 1} \frac{2 + \theta - s}{(\theta + 1 - s)^2}.$$

2.2.1 Moments et Mesures Connexes

Soit $X \rightsquigarrow PLD(\theta)$, La moyenne et la variance de X sont :

$$E(X) = \frac{2 + \theta}{\theta(\theta + 1)}, \quad (2.13)$$

$$E(X^2) = \frac{\theta^2 + 4\theta + 6}{\theta^2(\theta + 1)^2}, \quad (2.14)$$

$$Var(X) = \frac{\theta^3 + 4\theta^2 + 6\theta + 2}{\theta^2(\theta + 1)^2}. \quad (2.15)$$

Le coefficient d'asymétrie($\sqrt{\beta_1}$) et le coefficient d'aplatissement(β_2) sont :

$$\sqrt{\beta_1} = \frac{2(\theta + 1)^4(\theta + 2) - \theta^3(\theta + 2)(\theta + 3)}{[2(\theta + 1)^3 - \theta^2(\theta + 2)]^{\frac{3}{2}}},$$

$$\beta_2 = 3 + \frac{2(\theta + 1)^5[(\theta + 3)^2 - 3] - \theta^4(\theta + 2)[(\theta + 4)^2 - 3]}{[2(\theta + 1)^3 - \theta^2(\theta + 2)]^2}.$$

2.2.2 Fonction Quantile de la distribution de Poisson Lindley

Notons Q_X la fonction quantile de X défini selon la formule (1.5).(voir[24])

Désormais, $\log(\cdot)$ désigne le logarithme naturel et $[t]$ représente le plafond d'un nombre réel t , qui est, $t := \inf\{k \in Z : k \geq t\}$. Avec la précédente notation, on est en mesure de déclarer ce qui suit.

Théorème 2.8 *Pour tout $\theta > 0$, la fonction quantile de la distribution de Poisson Lindley X est/*

$$Q_X(u) = -\frac{\theta^2 + 3\theta + 1}{\theta} - \frac{1}{\log(\theta + 1)} W_{-1} \left(\frac{\log(\theta + 1)}{\theta(\theta + 1)^{\frac{\theta^2+1}{\theta}}} (u - 1) \right), \quad 0 < u < 1. \quad (2.16)$$

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

Démonstration. *Pour tout $\theta > 0$, soit $u \in (0, 1)$ Afin d'obtenir l'expression de Q_X , on a pour résoudre l'équation $F_X(x) = u$ par rapport à k , pour tout $k \geq 0$, comme suit : ■*

$$\frac{\theta^2 + 3\theta + 1 + \theta k}{(\theta + 1)^k} = 1 - u, k \geq 0. \quad (2.17)$$

l'équation (2.17) peut être écrit comme suit :

$$k + \frac{(\theta + 1)^3}{\theta} (u - 1) (\theta + 1)^k = - \left(\frac{\theta^2 + 3\theta + 1}{\theta} \right). \quad (2.18)$$

Maintenant, on applique le lemme 1.1 pour résoudre l'équation (2.18) par rapport à k . Par conséquent, l'égalité suivante est réalisée :

$$W \left(\frac{\log(\theta + 1)}{\theta(\theta + 1)^{\frac{\theta^2+1}{\theta}}} (u - 1) \right) = - \left(k + \frac{\theta^2 + 3\theta + 1}{\theta} \right) \log(\theta + 1). \quad (2.19)$$

Voyant maintenant l'équation. (2.19). Pour tout $\theta > 0, k \geq 0$ et $u \in (0, 1)$ les inégalités suivantes sont vérifiées :

$$(i) \quad \frac{-1}{e} < \frac{\log(\theta + 1)}{\theta(\theta + 1)^{\frac{\theta^2+1}{\theta}}} (u - 1) < 0,$$

et

$$(ii) \left(k + \frac{\theta^2 + 3\theta + 1}{\theta}\right) \log(\theta + 1) > 1.$$

En vertu des inégalités (i) et (ii) ci-dessus ainsi que les propriétés de la fonction W de Lambert, la branche réelle de W impliqué dans l'équation (2.19) est précisément la branche W_{-1} , négative, ce qui conduit au résultat souhaité.

2.2.3 Estimation

Maximum de vraisemblance

Étant donné un échantillon aléatoire X_1, \dots, X_n , de la distribution de Poisson-Lindley discrète (2.11), la fonction de logvraisemblance est :

$$\log l(x_i; \beta, \theta) = 2n \log \theta - n(\bar{x} + 3) \log(1 + \theta) + \sum_{i=1}^n \log[x_i + \theta + 2].$$

Sankaran [14] a montré que l'estimateur de la méthode du maximum de vraisemblance $\hat{\theta}$ de θ est une solution de l'équation :

$$\frac{\partial \ln l(x_i, \theta)}{\partial \theta} = \frac{2n}{\theta} - n\left(\frac{\bar{x} + 3}{1 + \theta}\right) + \sum_{i=1}^n \frac{1}{x_i + \theta + 2} = 0. \quad (2.20)$$

Sankaran [14] a déclaré que la résolution de l'équation (2.20) est équivalente à la résolution d'un polynôme de degré $(n + 1)$, et l'équation (2.20) peuvent avoir plusieurs solutions. Dans ce qui suit, on montre que l'équation (2.20) a une solution unique pour tout n .

Théorème 2.9 *l'équation (2.20) est équivalente à :*

$$\varphi(\theta) = 2n - n\bar{x}\theta - \sum_{i=1}^n \frac{(x_i + 1)\theta}{x_i + \theta + 2} = 0.$$

La fonction $\varphi(\theta)$ est strictement décroissante en θ , puisque

$$\varphi'(\theta) = -n\bar{x} - \sum_{i=1}^n \frac{(x_i + 1)(x_i + 2)}{(x_i + \theta + 2)^2} < 0.$$

Comme $\varphi(0) = 2n$ et $\varphi(\infty) = -\infty$, il en résulte que $\varphi(\theta)$ traversera l'axe θ qu'une seule fois, c'est-à-dire il existe un unique $\hat{\theta}$ tel que $\varphi(\theta) = 0$.

Remarque 2.10 *Bien qu'efficace de l'estimateur du maximum de vraisemblance $\hat{\theta}$ mais ce dernier n'est pas simplement obtenu et il peut être suffisant pour des raisons pratiques d'utiliser l'estimation des moments.*

Méthode des moments

Étant donné un échantillon aléatoire X_1, \dots, X_n , de la distribution de Poisson-Lindley discrète (2.11), l'estimateur des moments (*MoM*) de θ est :

$$\hat{\theta}_{MoM} = \frac{-(\bar{X} - 1) + \sqrt{(\bar{X} - 1)^2 + 8\bar{X}}}{2\bar{X}}, \bar{X} > 0. \quad (2.21)$$

Le théorème suivant montre que l'estimateur $\hat{\theta}_{MoM}$ de θ est positivement biaisé.

Théorème 2.11 *L'estimateur $\hat{\theta}_{MoM}$ de θ est positivement biaisée, c'est-à-dire*

$$E(\hat{\theta}) - \theta > 0$$

Démonstration. Soit $\hat{\theta} = g(\bar{X})$ et $g(t) = \frac{-(t-1) + \sqrt{(t-1)^2 + 8t}}{2t} \quad \forall t > 0$.

Comme $g''(t) = \frac{1}{t^3} \left[1 + \frac{3t^3 + 15t^2 + 9t + 1}{[(t-1)^2 + 8t]^{\frac{3}{2}}} \right] > 0$, $g(t)$ 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 :

$$E(\hat{\theta}_{MoM}) > \theta.$$

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

Théorème 2.12 *L'estimateur $\hat{\theta}_{MoM}$ de θ est convergent et asymptotiquement normal :*

$$\sqrt{n}(\hat{\theta}_{MoM} - \theta) \xrightarrow{D} N(0, \vartheta^2(\theta)).$$

Où

$$\vartheta^2(\theta) = \frac{\theta^2(\theta + 1)^2(\theta^3 + 4\theta^2 + 6\theta + 2)}{(\theta^2 + 4\theta + 2)^2}.$$

Démonstration. Etant donné $\mu < \infty$, $\bar{X} \xrightarrow{P} \mu$, $g(t)$ est une fonction continue à $t = \mu$, $g(\bar{X}) \xrightarrow{P} g(\mu)$, c'est-à-dire $\hat{\theta}_{MoM} \xrightarrow{P} \theta$. Comme $\sigma^2 < \infty$, par le théorème central limite, on a :

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

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

$$\sqrt{n}(g(\bar{X}) - g(\mu)) \xrightarrow{P} N\left(0, [g'(\mu)]^2 \sigma^2\right).$$

Enfin, étant donné que :

$$g(\bar{X}) = \hat{\theta}_{MOM}, g(\mu) = \theta, \text{ et } g'(\mu) = \frac{-1}{2\mu^2} \left[1 + \frac{1+3\mu}{\sqrt{(\mu-1)^2 + 8\mu}} \right] = -\frac{\theta^2(\theta+1)^2}{(\theta^2+4\theta+2)}.$$

■

En conséquence du théorème 2.6, L'intervalle de confiance de θ pour un seuil de confiance $100(1-\alpha)\%$ est donné par :

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

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

2.2.4 Simulation

Dans cette section, nous étudions le comportement des estimateurs MOM pour un échantillon de taille (n) finie. Une étude de simulation comporte les étapes suivantes est effectuée pour chaque couple $(\theta; n)$, où $\theta = \frac{1}{3}, 1, 3$ et $n = 20, 40, 60, 80, 100$.

- Choisir les valeurs initiales de θ_0 pour spécifier la distribution de *PLD*;
- Choisir la taille de l'échantillon n ;
- Générer N échantillons indépendants de taille n de *LD*(θ);
- Calculer les estimateurs $\hat{\theta}_{MV}$ de θ pour chacun des N échantillons;
- Calculer :

(i) La moyenne des estimateurs obtenus sur tous les N échantillons

$$\text{average biais}(\theta) = \frac{1}{N} \sum_{i=1}^N (\hat{\Theta}_i - \Theta_0).$$

(ii) L'erreur quadratique moyenne EQM des estimations simulées :

$$EQM(\theta) = \frac{1}{N} \sum_{i=1}^N (\hat{\Theta}_i - \Theta_0)^2.$$

Tableau 2.3 : les biais moyens de l'estimateur $\hat{\theta}$

n	$\theta = \frac{1}{3}$	$\theta = 1$	$\theta = 3$
20	0.011908	0.064915	0.588291
40	0.005658	0.033095	0.230338
60	0.003531	0.021611	0.149065
80	0.002471	0.014849	0.107402
100	0.001887	0.011892	0.083455

Tableau 2.4 : la moyenne EQM de l'estimateur $\hat{\theta}$

n	$\theta = \frac{1}{3}$	$\theta = 1$	$\theta = 3$
20	0.004637	0.080496	4.424323
40	0.002104	0.032186	0.963864
60	0.001349	0.020220	0.538328
80	0.000999	0.014581	0.366871
100	0.00077800	0.011507	0.278360

Remarque 2.13 (i) Le tableau 2.3 présente un biais positif, comme indiqué dans le théorème 1.3. Le tableau montre également que le biais diminue (augmente) quand n (θ) augmente (respectivement).

(ii) Le tableau 2.4 montre que l'erreur quadratique moyenne diminue (augmente) lorsque n (θ) augmente respectivement .

2.3 Distribution Quasi Lindley

La distribution Quasi Lindley (QLD) avec deux paramètres α et θ est définie par sa fonction de densité de probabilité

$$f(x; \alpha, \theta) = \frac{\theta^2(\alpha + x\theta)e^{-\theta x}}{\alpha + 1}; \quad x > 0, \theta > 0, \alpha > -1. \quad (2.22)$$

Il est facile de voir que si $\alpha = \theta$, l'équation (2.22) de QLD se réduit à la fonction de distribution de Lindley (2.1) et si $\alpha = 0$, elle réduit à la distribution $Gamma(2, \theta)$.

La fonction de densité (2.22) de QLD peut être montré sous forme de mélange des distributions Exponentielle (θ) et $Gamma(2, \theta)$ comme suit :

$$f(x; \alpha, \theta) = pf_1(x) + (1 - p) f_2(x)$$

où $p = \frac{\alpha}{\alpha+1}$, $f_1(x) = \theta e^{-\theta x}$ et $f_2(x) = \theta^2 x e^{-\theta x}$. (Pour plus de détails voir. Shanker et al.(2013) [20]).

La dérivée première de (2.22) est :

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

Il en résulte que :

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

(ii) pour $\alpha \geq 1$, $\frac{d}{dx}f(x) \leq 0$, c-à-d. $f(x)$ diminue en x_0 , dont le mode de la distribution QL est donnée par :

$$Mode = \begin{cases} \frac{1-\alpha}{\theta}, & |\alpha| < 1 \\ 0 & \end{cases}$$

La fonction de répartition correspondante est :

$$F(x) = 1 - \frac{1 + \alpha + \theta x}{\alpha + 1} e^{-\theta x}; x > 0, \theta > 0, \alpha > -1. \quad (2.23)$$

Le r^{ieme} moment de la distribution de quasi-Lindley est :

$$\mu_r^l = \frac{\Gamma(r+1)(\alpha+r+1)}{\theta^r(\alpha+1)}, r = 1, 2, \dots$$

En particulier on a :

$$\mu_1^l = \frac{(\alpha+2)}{\theta(\alpha+1)}, \mu_2^l = \frac{2(\alpha+3)}{\theta^2(\alpha+1)}, \mu_3^l = \frac{6(\alpha+4)}{\theta^3(\alpha+1)}, \mu_4^l = \frac{24(\alpha+5)}{\theta^4(\alpha+1)}.$$

Le moment centré de la distribution de Lindley est donné par :

$$\mu_k = E \left\{ (X - \mu)^k \right\} = \sum_{r=0}^k \binom{k}{r} \mu_r^l (-\mu)^{k-r}.$$

En particulier on a

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

Le coefficient de variation (γ), le coefficient de d'asymétrie ($\sqrt{\beta_1}$) et le coefficient d'aplatissement (β_2) sont :

$$\begin{aligned} \gamma &= \frac{\sqrt{\alpha^2 + 4\alpha + 2}}{\alpha + 2}, \\ \sqrt{\beta_1} &= \frac{2(\alpha^3 + 6\alpha^2 + 6\alpha + 2)}{(\alpha^2 + 4\alpha + 2)^{\frac{3}{2}}}, \\ \beta_2 &= \frac{3(3\alpha^4 + 24\alpha^3 + 44\alpha^2 + 32\alpha + 8)}{(\alpha^2 + 4\alpha + 2)^2}. \end{aligned}$$

2.3.1 Fonction de hasard et fonction de survie

Pour la distribution Quasi Lindley la fonction de taux du hasard est :

$$h(x) = \lim_{\Delta x \rightarrow 0} \frac{P(X < x + \Delta x | X > x)}{\Delta x} = \frac{f(x)}{1 - F(x)} = \frac{\theta(\alpha + \theta x)}{\alpha + 1 + \theta x}. \quad (2.24)$$

et la fonction de survie est :

$$S(x) = 1 - F(x) = \frac{\alpha + 1 + \theta x}{1 + \alpha} e^{-\theta x}. \quad (2.25)$$

2.3.2 Fonction Quantile de la distribution Quasi Lindley

La distribution Quasi Lindley du paramètre θ et α est spécifiée par sa fonction de répartition définie en (2.22). Il convient de noter qu'elle est continue et strictement croissante de sorte que la fonction de quantile X est $Q_X(u) = F_X^{-1}(u)$, $0 < u < 1$. Dans le résultat suivant, on donne une expression explicite de Q_X en fonction de la fonction W de Lambert.

Théorème 2.14 *Pour tout $\theta > 0$, la fonction quantile de la distribution Lindley X est :*

$$Q_X(u) = -\frac{1}{\theta} - \frac{\alpha}{\theta} - \frac{1}{\theta} W_{-1} \left(\frac{\alpha + 1}{\exp(\alpha + 1)} (u - 1) \right), \quad 0 < u < 1, \quad (2.26)$$

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

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

$$(\alpha + 1 + \theta x) e^{-\theta x} = (\alpha + 1)(1 - u). \quad (2.27)$$

En multipliant par $-\exp(-\alpha - 1)$ l'équation (2.27), on obtient :

$$-(\alpha + 1 + \theta x) \exp(-\alpha - 1 - \theta x) = (\alpha + 1)(u - 1) \exp(-\alpha - 1). \quad (2.28)$$

D'après l'équation (2.28), conjointement avec l'équation (1.3), on voit que $-(\alpha + 1 + \theta x)$ est la fonction W de Lambert de l'argument réel $(\alpha + 1)(u - 1) \exp(-\alpha - 1)$. Alors,

on a

$$W\left(\frac{\alpha+1}{\exp(\alpha+1)}(u-1)\right) = -(\alpha+1+\theta x), 0 < u < 1. \quad (2.29)$$

Toujours, pour tout $\theta > 0$ et $x > 0$ il est immédiat que $(\alpha+1+\theta x) > 1$ et il peut également être vérifié que $u \in (0, 1)$. en prenant en compte les propriétés de la branche négative de la fonction W de Lambert a présenté en première chapitre, l'équation (2.29) devient

$$W_{-1}\left(\frac{\alpha+1}{\exp(\alpha+1)}(u-1)\right) = -(\alpha+1+\theta x). \quad (2.30)$$

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

2.3.3 Estimation

Estimation par la méthode des moments (MoM)

Étant donné un échantillon aléatoire X_1, \dots, X_n , de QLD , les estimateurs des moments $\hat{\theta}$ de θ et $\hat{\alpha}$ de α peuvent être obtenus comme suit :

$$\hat{\theta} = \left(\frac{\hat{\alpha}+2}{\hat{\alpha}+1}\right) \frac{1}{\bar{X}}, \quad \bar{X} > 0. \quad (2.31)$$

$$\hat{\alpha} = \frac{-1}{k-2} \left[2k + \sqrt{2}\sqrt{2-k} - 4\right] \quad (2.32)$$

Où

$$\frac{\mu_2^!}{\mu_1^!} = k = \frac{2(\alpha+3)(\alpha+1)}{(\alpha+2)^2}.$$

Estimation par la méthode du maximum de vraisemblance

Soient x_1, x_2, \dots, x_n un échantillon aléatoire de taille n de QLD. Et soit f_x la fréquence observée dans l'échantillon correspondant à $X = x$ ($x = 1, 2, \dots, k$) sachant que $\sum_{x=1}^k f_x = n$. La fonction de vraisemblance L de la distribution QL est donnée par :

$$L = \left(\frac{\theta}{\alpha + 1} \right)^n \prod_{x=1}^k (\alpha + x\theta)^{f_x} e^{-n\theta\bar{X}}. \quad (2.33)$$

Alors, le log-vraisemblance est :

$$\log L(x; \alpha, \theta) = n \log \theta - n \log (1 + \alpha) + \sum_{x=1}^k f_x \log (\alpha + \theta x) - n\theta\bar{X}. \quad (2.34)$$

La dérivé de logvraisemblance est

$$\frac{\partial \log L}{\partial \theta} = \frac{n}{\theta} + \sum_{x=1}^k \frac{x f_x}{\alpha + x\theta} - n\bar{X} = 0, \quad (2.35)$$

$$\frac{\partial \log L}{\partial \alpha} = -\frac{n}{\alpha + 1} + \sum_{x=1}^k \frac{f_x}{\alpha + \theta x} = 0. \quad (2.36)$$

Les deux équations (2.35), (2.36) ne semblent pas être résolues directement, cependant la méthode de Fisher soring peut être appliquée pour résoudre ce système.

2.4 Distribution de Deux-Paramètres de Lindley

La distribution de Deux Paramètres de Lindley (*Two Parameter Lindley* (*Two PLD*)) avec les paramètres α et θ est définie par sa fonction de densité de probabilité

$$f(x; \alpha, \theta) = \frac{\theta^2}{\theta + \alpha} (1 + \alpha x) e^{-\theta x}; \quad x > 0, \theta > 0, \alpha > -\theta. \quad (2.37)$$

Il est facile de voir que si $\alpha = 1$, la fonction *TwoPLD* réduit à la fonction de distribution de LD(2.1) et si $\alpha = 0$, elle se réduit à la distribution Exponentielle (θ). La fonction *TwoPLD* peut être montré sous forme de mélange des distributions exponentielle (θ) et gamma ($2, \theta$) comme suit :

$$f(x; \alpha, \theta) = pf_1(x) + (1 - p) f_2(x).$$

Où $p = \frac{\theta}{\theta + \alpha}$, $f_1(x) = \theta e^{-\theta x}$ et $f_2(x) = \theta^2 x e^{-\theta x}$. (pour plus de détails voir. S.Sharma et al.(2013)[21].

La fonction de répartition correspondante est :

$$F(x) = 1 - \frac{\theta + \alpha + \alpha \theta x}{\theta + \alpha} e^{-\theta x}; x > 0, \theta > 0, \alpha > -\theta. \quad (2.38)$$

Le moment d'ordre k de la distribution de Deux-Paramètres de Lindley est :

$$\mu_k^l = E(X^k) = \frac{\Gamma(k+1)(\theta + \alpha + \alpha k)}{\theta^k(\theta + \alpha)}, k = 1, 2, \dots$$

d'où, on a :

$$\mu_1^l = \frac{(\alpha + 2\alpha)}{\theta(\theta + \alpha)}, \mu_2^l = \frac{2(\alpha + 3\alpha)}{\theta^2(\theta + \alpha)}, \mu_3^l = \frac{6(\alpha + 4\alpha)}{\theta^3(\theta + \alpha)}, \mu_4^l = \frac{24(\alpha + 5\alpha)}{\theta^4(\theta + \alpha)}.$$

La fonction de logvraisemblance de la distribution de deux-Paramètres de Lindley est:

$$\log L(x; \alpha, \theta) = n \log \theta^2 - n \log (\alpha + \theta) + \sum_{i=0}^n \log (1 + \alpha \theta x_i) - n \theta \bar{X}.$$

Estimation par la méthode des moments

Étant donné un échantillon aléatoire X_1, \dots, X_n , de deux-Paramètres de Lindley, les estimateurs des moments $\hat{\theta}_{MoM}$ de θ et $\hat{\alpha}$ de α peuvent être obtenus comme suit :

$$\hat{\theta} = \frac{1}{\mu_2'} \left(2\bar{X} + \sqrt{2} \sqrt{2\bar{X}^2 - \mu_2'} \right), \bar{X} > 0. \quad (2.39)$$

$$\hat{\alpha} = \frac{\hat{\theta} - \bar{X}\hat{\theta}^2}{\bar{X}\hat{\theta} - 2}. \quad (2.40)$$

Chapitre 3

Poisson–Quasi Lindley distribution et ses applications

Dans ce chapitre on introduit une nouvelle distribution à deux paramètres où on va étudier quelques propriétés comme : la fonction quantile, méthode des moments, estimation du maximum de vraisemblance. Des simulations sont établies pour examiner le biais et l'erreur quadratique moyenne des estimateurs des paramètres obtenus par la méthode du maximum de vraisemblance et leur application dans l'analyse de survie.

3.1 Distribution poisson quasi-Lindley

Une distribution composée de Poisson peut être obtenue en composant la distribution de Poisson et une distribution quasi-Lindley.

On a $dF(\lambda) = e^{\lambda\Phi}h(\lambda)B(\Phi)d\lambda$, ou $h(\lambda) = \alpha + \theta\lambda$ et $B(\Phi) = \frac{-\Phi}{\alpha + 1}$, alors la distribution poisson composée est ([14]) :

$$\begin{aligned} P_x(\Phi) &= \int_0^{\infty} \frac{e^{-\lambda}\lambda^x}{x!} dF(\lambda) \\ &= \frac{B(\Phi)}{x!} \left[\alpha \int_0^{\infty} e^{(\Phi-1)\lambda}\lambda^x d\lambda + (-\Phi) \int_0^{\infty} e^{(\Phi-1)\lambda}\lambda^{x+1} d\lambda \right] \\ &= \frac{-\Phi}{(\alpha + 1)} \left(\frac{\alpha(1 - \Phi) - \Phi(x + 1)}{(1 - \Phi)^{x+2}} \right). \end{aligned}$$

On remplace alors Φ par $-\theta$ on trouve :

$$P_x(\theta) = \frac{\theta}{\alpha + 1} \left(\frac{\alpha(1 + \theta) + \theta(x + 1)}{(1 + \theta)^{x+2}} \right)$$

La fonction de densité Poisson-quasi Lindley est donnée alors :

$$f_{PQL}(x; \alpha, \theta) = \frac{\theta(\alpha + \theta + \alpha\theta + \theta x)}{(1 + \alpha)(1 + \theta)^{x+2}}, \quad x = 0, 1, \dots, \quad \theta > 0, \quad \alpha > -1 \quad (3-1)$$

Remarque 3.1 si $\alpha = \theta$, cette distribution est la distribution Poisson-Lindley. La première et la deuxième dérivée de $f_{PQL}(x)$ sont :

$$\frac{d}{dx} f_{PQL}(x) = -\frac{\theta(\theta \ln(\theta + 1) - \theta + \alpha \ln(\theta + 1) + \theta\alpha \ln(\theta + 1) + x\theta \ln(\theta + 1))}{(\theta + 1)^{x+2}(\alpha + 1)}$$

et

$$\frac{d^2}{dx^2} f_{PQL}(x) = \frac{\theta \ln(\theta + 1)(\theta \ln(\theta + 1) - 2\theta + \alpha \ln(\theta + 1) + \theta\alpha \ln(\theta + 1) + x\theta \ln(\theta + 1))}{(\theta + 1)^{x+2}(\alpha + 1)}$$

Où la solution de $\frac{d}{dx} f_{PQL}(x) = 0$, est $x = \frac{1}{\ln(\theta+1)} - \frac{(\theta+\alpha+\theta\alpha)}{\theta}$ et

$$\frac{d^2}{dx^2} f_{PQL}(\hat{x}) = \frac{-\theta^2 (\ln(\theta+1)) (\theta+1)^{\frac{1}{\theta \ln(\theta+1)} (\theta \ln(\theta+1) - \theta + \alpha \ln(\theta+1) + \theta \alpha \ln(\theta+1)) - 2}}{\alpha + 1} < 0.$$

- pour $\theta, \alpha, \hat{x} > 0$, $\hat{x} = \frac{1}{\ln(\theta+1)} - \frac{(\theta+\alpha+\theta\alpha)}{\theta}$ c'est le seul point critique pour la quelle $f_{PQL}(x; \alpha, \theta)$ soit maximisée et $f_{PQL}(x)$ est concave .

- si $\hat{x} < 0$, la fonction de densité $f_{PQL}(x)$ est discrète en \hat{x} .

Par conséquent, le mode de la distribution PQL est donnée par :

$$Mode(X) = \begin{cases} \frac{1}{\ln(\theta+1)} - \frac{(\theta+\alpha+\theta\alpha)}{\theta} & \text{pour } \theta > 0, \alpha > -1 \\ 0 & \text{ailleurs.} \end{cases} \quad (3-2)$$

On peut déterminer facilement la fonction de répartition pour PQL

$$F_{PQL}(x) = 1 - \frac{\alpha + 2\theta + \alpha\theta + \theta x + 1}{(1 + \alpha)(1 + \theta)^{x+2}}, x = 0, 1, \dots, \theta > 0, \alpha > -1 \quad (3-3)$$

Les figures 1 et 2 représentent la fonction de masse et la fonction de répartition de la distribution Poisson quasi- Lindley pour quelques valeurs de ($\alpha; \theta$).

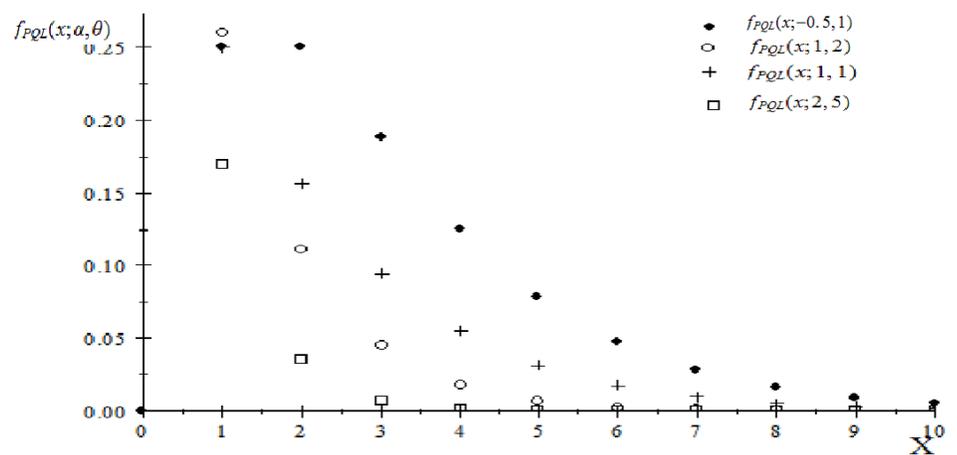


FIG. 3.1: Représentation graphique de la fonction de masse de PQL.

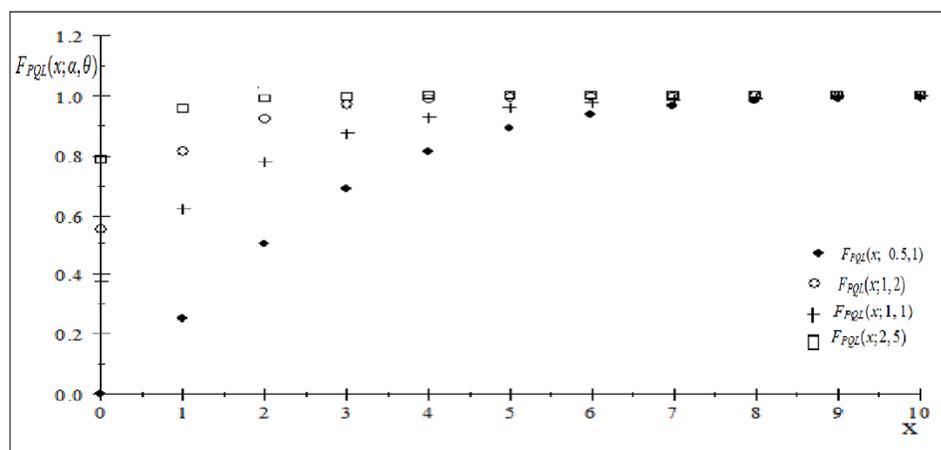


FIG. 3.2: Présentation graphique de la fonction de répartition de la distribution PQL.

3.2 La Fonction de survie et la Fonction de hasard

Soit la fonction de survie et hasard respective

$$S_{PQL}(x) = 1 - F_{PQL}(x) = \frac{\alpha + 2\theta + \alpha\theta + \theta x + 1}{(1 + \alpha)(1 + \theta)^{x+2}}$$

$$h_{PQL}(x) = \frac{f_{PQL}(x)}{1 - F_{PQL}(x)} = \frac{\theta(\alpha + \theta + \alpha\theta + \theta x)}{\alpha + 2\theta + \alpha\theta + \theta x + 1}.$$

Proposition 3.2 *Soit $h_{PQL}(x)$ la fonction hasard de X . Alors $h_{PQL}(x)$ est croissante.*

Démonstration. Selon Glaser (1980) pour la (PQL)

$$\rho(x) = -\frac{f'_{PQL}(x)}{f_{PQL}(x)} = \frac{\theta \ln(\theta + 1) - \theta + \alpha \ln(\theta + 1) + \theta\alpha \ln(\theta + 1) + x\theta \ln(\theta + 1)}{\alpha + \theta + \alpha\theta + \theta x}.$$

Il est claire que $\rho'(x) = \frac{\theta^2}{(\theta + \alpha + x\theta + \theta\alpha)^2} > 0, \forall x, \alpha, \theta$ implique que $h_{PQL}(x)$ est croissante.

■

3.3 Estimations

3.3.1 Estimation par le maximum de vraisemblance

Dans cette partie on discute les points d'estimation pour les paramètres de $PQL(\theta, \alpha)$. Soit la fonction log-vraisemblance pour une seule observation (dite x_i) pour le vecteur de paramètre (θ, α) peut être écrit comme suit :

$$\ln l(x; \alpha, \theta) = \ln \theta + \ln(\alpha + \theta + \alpha\theta + \theta x) - \ln(1 + \alpha) - 2 \ln(1 + \theta) - x \ln(1 + \theta), \quad (3-4)$$

la dérivation de $\ln l(x; \theta, \alpha)$ par rapport a θ et α est :

$$\begin{aligned} \frac{\partial \ln l(x; \theta, \alpha)}{\partial \theta} &= \frac{1}{\theta} + \left(\frac{1+\alpha+x}{\alpha+\theta+\alpha\theta+\theta x} \right) - \frac{2+x}{1+\theta}, \\ \frac{\partial \ln l(x; \theta, \alpha)}{\partial \alpha} &= \left(\frac{1+\theta}{\alpha+\theta+\alpha\theta+\theta x} \right) - \frac{1}{1+\alpha}. \end{aligned} \quad (3-5)$$

L'estimateur de maximum de vraisemblance $\hat{\theta}$ de θ et $\hat{\alpha}$ de α est obtenu par la résolution de système non linière

$$\begin{cases} \hat{\theta} = \frac{1}{x} \\ \hat{\alpha} = \frac{-x}{1+x} \end{cases} \quad (3-6)$$

et

$$\begin{cases} E(\hat{\theta}) = \frac{\theta(\alpha+\theta+\alpha\theta)}{(1+\alpha)(1+\theta)^3} \\ -1 \leq E(\hat{\alpha}) \leq 0 \end{cases} \quad (3-7)$$

Proposition 3.3 *Soit X_1, X_2, \dots, X_n les n variable aléatoires indépendante de la distribution $PQL(\alpha, \theta)$. Alors la fonction génératrice des moments de $S = \sum_{i=1}^n X_i$, est donnée par :*

$$M_S(t) = \frac{(\alpha\theta)^n e^t + \theta^{2n}}{(1+\alpha)^n (e^t + \theta e^t - e^{2t})^n}$$

et

$$M_X(t) = \mathbb{E}(e^{tX}) = \frac{\alpha\theta e^t + \theta^2}{(1+\alpha)(e^t + \theta e^t - e^{2t})} \quad (3-8)$$

Corollaire 3.4 *Soit $X \sim PQL(\theta, \alpha)$, l'espérance et la variance pour X est :*

$$\begin{aligned} \mathbb{E}(X) &= \frac{2+\alpha}{(1+\alpha)\theta} \\ Var(X) &= \frac{2+4\alpha+\alpha^2+\theta(\alpha+2)(1+\alpha)}{(1+\alpha)^2\theta^2} \end{aligned} \quad (3-9)$$

Démonstration. On a $\mathbb{E}(X) = M'_X(t = 0)$, $\mathbb{E}(X^2) = M''_X(t = 0)$ et $Var(X) = \mathbb{E}(X^2) - \mathbb{E}^2(X)$.

$$\mathbb{E}(X) = \frac{2 + \alpha}{(1 + \alpha)\theta}$$

■

$$\mathbb{E}(X^2) = \frac{6 + 2\theta + \alpha(2 + \theta)}{(1 + \alpha)\theta^2}$$

3.3.2 Estimations par la méthode des moments

Utilisant m et m_2 le premier et le deuxième moment de la distribution PQL, on a :

$$\begin{cases} m = \frac{2+\alpha}{(1+\alpha)\theta} \\ m_2 = \frac{6+2\theta+\alpha(2+\theta)}{(1+\alpha)\theta^2} \end{cases} \quad (3-10)$$

Où $m_2 = S^2 + m^2$ et S^2 est la variance. On résout ce système non linéaire on trouve le couple $(\hat{\theta}, \hat{\alpha})$, où $(\hat{\theta}, \hat{\alpha}) \succ 0$ pour toute $S \succ 0, m \succ 0$. La solution du système non linéaire (3-10) donne :

$$(m_2 - m)\theta^2 - 4m\theta + 2 = 0 \quad \text{et} \quad \alpha = \frac{2 - \theta m}{\theta m - 1}.$$

Il est facile d'avoir que la solution de $(m_2 - m)\theta^2 - 4m\theta + 2 = 0$ est :

$$\theta = -\frac{1}{m - m_2} \left(2m + \sqrt{2} \sqrt{m - m_2 + 2m^2} \right), \quad \text{si } m - m_2 \neq 0,$$

Alors,

$$\hat{\theta} = -\frac{1}{m - m_2} \left(2m + \sqrt{2} \sqrt{m - m_2 + 2m^2} \right) \text{ et } \hat{\alpha} = \frac{2 - \hat{\theta}m}{\hat{\theta}m - 1}. \quad (3-11)$$

3.4 Fonction quantile de la distribution poisson quasi-Lindley

D'après la fonction de répartition de la distribution poisson quasi-Lindley, il convient de noter qu'elle est continue et strictement croissante de sorte que la fonction de quantile X est $Q_X(u) = F_X^{-1}(u)$; où $0 \leq u \leq 1$. Dans le résultat suivant, on donne une expression explicite de Q_X en fonction de la fonction W de Lambert.

Théorème 3.5 *Pour tout $\theta, \alpha > 0$, la fonction quantile de X qui suit la distribution poisson-quasi Lindley est*

$$Q_X(u) = -\frac{\alpha + 2\theta + \alpha\theta + 1}{\theta} - \frac{1}{\ln(1 + \theta)} W_{-1} \left(\frac{\ln(1 + \theta)}{\theta(1 + \theta)^{\frac{\alpha + \theta\alpha + 1}{\theta}}} (u - 1) \right)$$

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

Démonstration.

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

$$\alpha + 1 + 2\theta + \theta\alpha = (\alpha + 1)(1 + \theta)^{x+2}(1 - u). \quad (3.12)$$

En multipliant par $(1 + \theta)^{-\frac{(\alpha+1+2\theta+\theta\alpha)}{\theta}} \log(1 + \theta)$ l'équation (3, 12), on obtient :

$$\frac{(u-1)(\alpha+1)}{\theta(1+\theta)^{\frac{(\alpha+1+2\theta+\theta\alpha)}{\theta}}} (1+\theta)^x \log(1+\theta) = - \left(\frac{(\alpha+1+2\theta+\theta\alpha)}{\theta} + x \right) \log(1+\theta). \quad (3.13)$$

D'après l'équation (3, 13), conjointement avec l'équation (1.4), on a

$$W \left(\frac{(u-1)}{\theta(1+\theta)^{\frac{(\alpha+1+2\theta+\theta\alpha)}{\theta}}} \log(1+\theta) \right) = - \left(\frac{\alpha+1+2\theta+\theta\alpha}{\theta} + x \right) \log(1+\theta), \quad 0 < u < 1. \quad (3.14)$$

Toujours, pour tout $\theta > 0$ et $x > 0$ il est immédiat que $\left(\frac{(\alpha+1+2\theta+\theta\alpha)}{\theta} + x \right) > 1$ et il peut également être vérifié que $u \in (0, 1)$. en prenant en compte les propriétés de la branche négative de la fonction W de Lambert a présenté en première chapitre, l'équation (3, 15) devient

$$W_{-1} \left(\frac{(u-1)}{\theta(1+\theta)^{\frac{(\alpha+1+2\theta+\theta\alpha)}{\theta}}} \log(1+\theta) \right) = - \left(\frac{\alpha+1+2\theta+\theta\alpha}{\theta} + x \right). \quad (3.15)$$

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

3.5 Illustrations et exemple

Exemple 1

Tableau 3.1 présente quelque quantiles de la distribution Poisson Quasi Lindley pour différentes valeurs de α et θ , où on utilise l'expression de $Q_X(u)$ donné dans le théorème 3.1.

Tableau 3.1. Quantiles de la distribution Poisson Quasi-Lindley

u	$\theta = 0.1, \alpha = 0.1$	$\theta = 0.1, \alpha = 0.5$	$\theta = 3, \alpha = 1$	$\theta = 5, \alpha = 1$
0.01	3.2884	7.8757	-0.33136	-0.50238
0.05	4.4242	8.6062	-0.29346	-0.4739
0.1	5.765	9.5437	-0.24395	-0.43667
0.15	7.0627	10.513	-0.19181	-0.39743
0.25	9.6363	12.569	$-7.8345.10^{-2}$	-0.3119
0.3	10.949	13.671	$-1.6190.10^{-2}$	-0.26497
0.35	12.3	14.832	$5.0285.10^{-2}$	-0.21474
0.40	13.705	16.064	0.12176	-0.16067
0.45	15.180	17.381	0.19911	-0.10211
0.5	16.745	18.799	0.28342	$-3.8208.10^{-2}$
0.55	18.425	20.341	0.37617	$3.2156.10^{-2}$
0.60	20.252	22.037	0.47930	0.11049
0.65	22.269	23.928	0.59557	0.19889

0.7	24.539	26.076	0.72901	0.30046
0.75	27.157	28.573	0.88581	0.41995
0.8	30.283	31.577	1.0753	0.56534
0.85	34.211	35.378	1.32	0.7515
0.9	39.596	40.625	1.6601	1.0117
0.95	48.508	49.368	2.2339	1.4516

Le tableau 3.2 représente le mode, la moyenne et la médiane de la distribution poisson quasi-Lindley pour différentes valeurs de θ et α .

Tableau 3.2. Mode, moyenne et médiane de PQL

	$\theta = 0.01, \alpha = 0.1$	$\theta = 0.1, \alpha = 0.5$	$\theta = 0.05, \alpha = 1$
Q_1	104.66	12.569	30.601
$Median = Q_2$	172.92	18.799	41.819
Q_3	272.77	28.573	60.008
$Mode$	89.399	3.9921	-1.5041
$Mean$	190.91	16.667	30

Simulation

Dans cette section, nous étudions le comportement des estimateurs MM pour un échantillon de taille (n) fini. Une étude de simulation comporte les étapes suivantes est effectuée pour chaque triple $(\theta, \alpha; n)$, ou $\theta = 0.01, \alpha = 0.1, 0.01, 1$ et pour $\alpha = 0.5, \theta = 0.05, 1, 5$ et $n = 10, 30, 50$.

- Choisir les valeurs initiale de θ_0, α_0 pour les éléments correspondants du vecteur de $\Theta = (\theta, \alpha)$ pour spécifier la distribution PQL.
- Choisir la taille de l'échantillon n ;
- Générer N échantillon indépendantes de taille n à partir de $PQL(\theta, \alpha)$;
- Calculer l'estimation MM $\hat{\Theta}_n$ de Θ_0 pour chacun de N échantillons;
- Calculer la moyenne des estimateurs obtenus sur tous les N échantillons,

$$\text{bais moyen}(\theta) = \frac{1}{N} \sum_{i=1}^N (\hat{\Theta}_i - \Theta_0)$$

Et l'erreur quadratique moyenne :

$$EQM(\theta) = \frac{1}{N} \sum_{i=1}^N (\hat{\Theta}_i - \Theta_0)^2$$

Tableau 3.3. Biais moyenne des estimations simulées de la distribution PQL

	$\theta = 0.01$	$\alpha = 0.1$	$\theta = 0.01$	$\alpha = 0.01$	$\theta = 0.01$	$\alpha = 1$
	bais(θ)	bais(α)	bais(θ)	bais(α)	bais(θ)	bais(α)
n=10	0.00002556	-0.0056129	0.0000025	-0.00050509	0.000264	-0.088393
n=30	0.00000852	-0.00187096	0.000000833	-0.000168363	0.000088	-0.02946433
n=50	0.000005112	-0.00112258	0.0000005	-0.000101018	0.0000528	-0.0176786

	$\theta = 0.05$	$\alpha = 0.5$	$\theta = 1$	$\alpha = 0.5$	$\theta = 5$	$\alpha = 0.5$
	bais(θ)	bais(α)	bais(θ)	bais(α)	bais(θ)	bais(α)
$n = 10$	0.0006842	-0.038232	0.02743	-0.061022	0.5756	-0.11132
$n = 30$	0.00029806	-0.012744	0.009143333	-0.02034066	0.1918666	-0.0371066
$n = 50$	0.00013684	-0.0076464	0.005486	-0.0122044	0.11512	-0.022264

Tableau 3.4. l'erreur quadratique moyenne

	$\theta = 0.01$	$\alpha = 0.1$	$\theta = 0.01$	$\alpha = 0.01$	$\theta = 0.01$	$\alpha = 1$
	EQM(θ)	EQM(α)	EQM(θ)	EQM(α)	EQM(θ)	EQM(α)
$n = 10$	0.0000000066	0.000315046	$6.3 \cdot 10^{-11}$	0.0000025511	0.000000697	0.078133224
$n = 30$	0.0000000022	0.000105015	$2 \cdot 10^{-11}$	0.0000008504	0.0000002323	0.026044408
$n = 50$	0.0000000013	0.000063009	10^{-11}	0.0000005102	0.0000001394	0.015626644

	$\theta = 0.05$	$\alpha = 0.5$	$\theta = 1$	$\alpha = 0.5$	$\theta = 5$	$\alpha = 0.5$
	EQM(θ)	EQM(α)	EQM(θ)	EQM(α)	EQM(θ)	EQM(α)
$n = 10$	0.0000046813	0.01461685	0.00752404	0.03723684	3.3131536	0.123921424
$n = 30$	0.0000015604	0.00487228	0.002508016	0.0124228	1.10438453	0.041307141
$n = 50$	0.0000009363	0.00292337	0.001504809	0.00744736	0.66263072	0.024784284

Exemple 2

Tableau 3.5 montre quelques distributions de copies de chiffres aléatoires avec des fréquences attendues obtenues en ajuste les distributions de Poisson, Poisson-Lindley et Poisson -quasi Lindley.

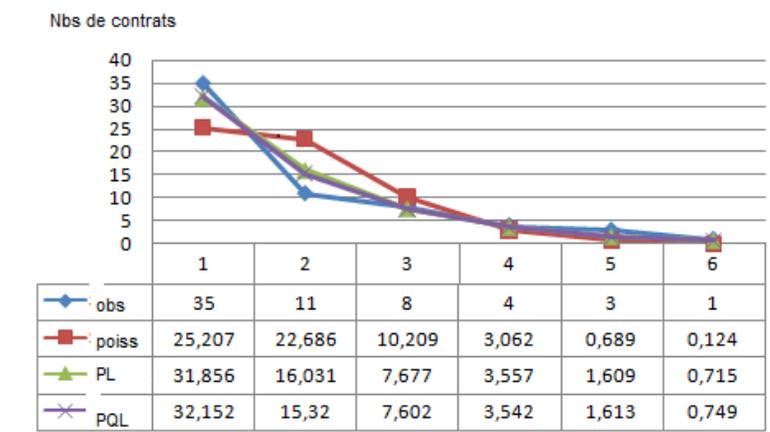


FIG. 3.3: Comparaison des différents modèles

Tableau 3.5. distributions de Poisson, Poisson-Lindley et Poisson -quasi Lindley

Num d'erreurs	Obs freq	Poisson	Poisson-Lindley	Poisson -quasi Lindley
<i>per group</i>	$m=0.9, n=2.451$	$\hat{\theta}=1.08$	$\hat{\theta}=1.547$	$\hat{\theta}=1.398, \hat{\alpha}=0.786$
0	35	25.207	31.856	32.152
1	11	22.686	16.031	15.32
2	8	10.209	7.677	7.602
3	4	3.062	3.557	3.542
4	3	0.689	1.609	1.613
5	1	0.124	0.715	0.749
Totale	62	62	62	62
χ^2	-	24.524	3.271	3.216

3.6 Conclusion et Perspectives

Dans ce travail, nous proposons une nouvelle distribution à deux paramètres, appelée distribution PQL qui contient la distribution poisson Lindley comme cas particulier. Diverses propriétés de cette distribution ont été examinées, notamment la fonction de densité (pdf), la fonction de répartition (cdf), la fonction de survies et le taux de risque, la fonction génératrice de moments (mgf), la moyenne, variance et certains résultats. En outre, les estimations du maximum de vraisemblance et les estimations du moment sont discutées. Finalement nous avons ajusté le modèle PQL à plusieurs ensembles de données réelles pour montrer le potentiel de cette nouvelle distribution.

La distribution PQL a une qualité d'ajustement (goodness of fit) meilleur que celle de poisson et poisson-lindley, dont on peut considérée comme un outil important pour la modélisation des données actuarielles.

Pour les études futures, nous pouvons faire une application actuarielle au modèle de risque collectif en considérant la distribution proposée comme distribution primaire et exponentielle et Erlang comme distributions secondaires.

Bibliographie

- [1] A. Asgharzadeh , Hassan S. Bakouch, L. Esmaeili , Pareto Poisson–Lindley distribution and its application, *Journal of Applied Statistics*, pp. 1–18 (2013).
- [2] D. V. Lindley, Fiducial distributions and Bayes’theorem. *Journal of the Royal Society, series B*, 20, 102-107 (1958).
- [3] E.M.Lémeray,Racines de quelques équations transcendantes. Intégration d’une équation aux différences mêlées. Racines imaginaires, *Nouvelles Ann. Math.* (1897) 540-546.
- [4] J. F. Lawless , 2003. *Statistical models and methods for lifetime data*. Wiley,New York pp 204 - 263.
- [5] H. Zeghdoudi, S. Nedjar (2016a). Gamma Lindley distribution and its application. *Journal of Applied Probability and Statistics* Vol 11, N 1, 129-138.
- [6] H. Zeghdoudi, S. Nedjar (2016b) A Pseudo Lindley distribution and its application., *J. Afrika Statistika* Vol 11 N 1, 923_932.
- [7] H. Zeghdoudi, S. Nedjar (2017). On Pseudo Lindley Distribution : Properties and

- Applications. Journal of New Trends in Mathematical Sciences Vol 5, N 1, 59-65
- [8] H. Zeghdoudi, N. Lazri (2016). On Lindley-Pareto Distribution : Properties and Application GSTF Journal of Mathematics, Statistics, Operations Research (JM-SOR) Vol.3 No.2 pp 1-7.
- [9] H. Zeghdoudi, N. Lazri (2017). Composite Gamma(2, λ)-Pareto Distribution International Journal of Applied Mathematics and Statistics, Int. J. Appl.Math. Stat. ; Vol. 56 ; Issue No. 5, PP 101-107.
- [10] H. Zeghdoudi, N. Lazri (2017). Lindley-Pareto Distribution, Journal of Statistical Distributions and Applications JSDA-D-17-00018.
- [11] H. S. Bakouch, B. M. Al-Zahrani, A. Al-Shomrani, V. A. Marchi and F. Louzada (2012). An extended Lindley distribution, Journal of the Korean Statistical Society, Vol 41, 75-85.
- [12] H. Zakerzadah, A. Dolati (2010). Generalized Lindley distribution. J. Math. Ext.3(2),pp. 13-25.
- [13] Lindley, D. V. (1958). Fiducial distributions and bayes theorem. Journal of the Royal Society, series B, 20, 102-107
- [14] M. Sankaran, The discrete Poisson-Lindley distribution. Biometrics, 26, 145-149 (1970).
- [15] M. E. Ghitany, B. Atieh, S. Nadarajah (2008a). Lindley distribution and its applications. Math. Comput. Simulation , 78, pp. 493-506 .

- [16] M.E.Ghitany,D.K Almutairi(2009) : Estimation methods for the discrete Poisson–Lindley distribution, *Journal of Statistical Computation and Simulation*, 79 :1, 1-9.
- [17] M.E. Ghitany, D.K. Al-Mutairi, and S. Nadarajah, Zero-truncated Poisson–Lindley distribution and its application, *Math. Comput. Simulation* 79, pp. 279–287 (2008).
- [18] P. Jodrá (2010). Computer generation of random variables with Lindley or Poisson–Lindley distribution via the Lambert W function. *Mathematics and Computers in Simulation (MATCOM)*, Vol. **81**, issue **4**, pages 851-859.
- [19] R. E. Glaser, Bathtub and related failure rate characterizations.*J.Amer.Statist. Assoc.*,75,667-672 (1980).
- [20] R.Shanker and A. Mishra, A quasi lindley distribution. *African Journal of Mathematics and computer science research*. vol. 6(4), pp.64-71 (2013).
- [21] R.Shanker, S.Sharma, R.Shanker. (2013). A two-parameter Lindley distribution for modeling waiting and survival times data. *Applied Mathematics*, vol. 4, pp. 363-368.
- [22] R.M. Corless, G.H. Gonnet, D.E.G. Hare, D.J. Jeffrey, D.E. Knuth (1996). On the Lambert W function, *Adv. Comput. Math.* **5** 329–359.
- [23] R.Grine,H.Zeghdoudi, (2017).On Poisson Quasi-Lindley Distribution and its Applications.*journal of modern applied Statistical and Methods*. vol 16, N 2, 403-417.

- [24] S.Nedjar (2017). these doctorat, Poisson Pseudo Lindley Distributions et leurs applications en assurance vie
- [25] S. Nedjar,H. Zeghdoudi (2016c). On Gamma Lindley distribution : proprieties and simulation. Journal of Computational and Applied Mathematics 298, pp167-174.
- [26] S. Nedjar,H. Zeghdoudi (2017). A Poisson pseudo Lindley distribution and its application . Journal of Probability and Statistical Sciences . Vol. 15, N 1, 19-28
- [27] Zakerzadah H, Dolati A. Generalized Lindley distribution. J. Math. Ext. 3(2) :13-25 (2010).