N°d'ordre: ....

People's Democratic Republic of Algeria
Ministry of Higher Education and Scientific Research
Mouloud Mammeri University of Tizi-Ouzou

Faculty of Sciences
Department of Mathematics
Laboratory LMPA

Dissertation Submitted in Partial Fulfilment of the Requirements for the Degree of Master in Mathematics

Field : Mathematics
Specialty : Probabilities & Statistics

Presented by:

NATURINDA EVALYNE

On The Lindley Distribution and its Applications

Defended on June 29, 2024 before the jury composed of :

Mr.	GOUBI Mouloud	MCA	UMMTO	President of the jury
Mrs.	BEDOUHENE Kahina	MCB	UMMTO	Examiner
Mrs.	BOUZIANE Houria	MAA	UMMTO	Examiner
Mrs.	LADJIMI Fetima	MCB	UMMTO	Supervisor

Academic Year : 2023/2024

I dedicate this Master's Thesis:

To my parents Mr. BANGIRANA YERONIMO and Mrs. KENGYEYA TOPISTA whose support, sacrifices and encouragement have been my driving force behind my success.Thank you so much may God protect you always so that you enjoy the fruits of our sweat.

To my brothers and sisters (Judith, Godfrey, Nabaasa and Ninsiima) with whom we have always really worked together since childhood . Thank you so much and may God bless you in whichever you do.

To this surpassing individual, Mr. Lwanga Charles thank you so much for all the sacrifices you made for me right away from missing your exam to go and look where Biskra was located I can never take that for granted. Thank you for standing by my side and for cheering me up. You have really been an integral part of my journey.

To one of the directors of Rwentojo secondary school, Mr. Muhwezi John I cannot really thank you enough for what you have done for us in as far as our education is concerned may the Almighty keep blessing you always.

Lastly, to all my friends that have always wished me well thank you I am really grateful for you all. . . .

Acknowledgements

First and foremost, i would like to thank God (Exalted is He above all) who helped me in the times of weakness, reassurance in fear, and hope in the times of despair, guides me throughout my life and gave me the power and determination to accomplish this research work.

I sincerely express my sincere gratitude to my respectable supervisor Dr. LADJIMI FETIMA, for her continuous encouragement throughout the period of my research work. She was always motivating me to work harder and helped me in different possible ways. Without her active guidance and great support, it would not have been possible to complete this work. May ALLAH reward her immensely.

In addition, great thanks go to the jury members Mr. GOUBI Mouloud the president, Mrs. BEDOUHENE Kahina and Mrs. BOUZIANE Houria the examiners for their invaluable contributions to this thesis. I am grateful for the time and effort each member dedicated to reviewing my work.

I am also thankful to all the professors of my speciality (probability and statistics) for your encouragement and mentorship throughout this journey. Your dedication to my academic development have been truly motivating.

Finally, special thanks to my parents for their endless love, support, prayers and encouragement throughout this journey. Words can not express how grateful I am to these two. And to my siblings and friends I appreciate you all.

Tizi-Ouzou, June 29, 2024.

Table of Contents

List of Figures

List of Tables

Introduction

Statistics equips us with essential tools to analyze data, make informed decisions and predict future outcomes. The quality of the procedures used in statistical analysis depends strongly on the assumed probability model or distribution. For this purpose, considerable efforts have been made in the development of large classes of probability distributions as well as relevant statistical methodologies. Different probability distributions such as exponential, Weibull, Gamma among others have been used to model life time and survival data. However, these distributions seem to be insufficient to capture variability observed in empirical data.

The Gamma and Weibull models have major disadvantage. For the Gamma model, it is known that the cumulative distribution function (or the survival function) does not have an explicit form when the shape parameter is not a positive integer. Since it is in terms of an incomplete gamma function, Therefore, numerical methods (integral calculations) are necessary to approximate the cumulative distribution function, the survival function, or the hazard function. This makes gamma distribution little bit unpopular compared to the Weibull distribution, which has a nice distribution function, survival function and hazard function. In the case of the Weibull model, the major disadvantage is that the convergence in distribution of the maximum likelihood estimators to the normal law is very slow (see the references therein Gupta and Kundu [11]). As a result, confidence interval estimators are generally poor. For these reasons, other survival models have been proposed recently such as generalized exponential model proposed by Gupta and Kundu [11], Lindley distribution, his extensions and generalizations which include Ghitany et al. [9], Asgharzadeha et al. [3], Bakouch et al. [4], Chesneau et al. [8], Hamed and Alzaghal [12], Shanker et al. [26], Nedjar and Zeghdoudi [20] , Ramos and Louzada [22]....

In this work, we are interested in the Lindley distribution introduced by a British statistician Dennis Victor Lindley in 1958 and studied by Ghitany et al. [9]. Which has been widely applied across different fields such as finance, engineering, etc. The first objective of this work is the theoretical study and parameter estimation of the Lindley distribution, as well as the investigation of some extensions of this distribution, including the two-parameter Lindley distribution and the two-parameter weighted Lindley distribution. The second objective is to apply this distributions to real and simulated data. This thesis is structured into four main parts.

The first chapter is devoted to a detailed review of certain definitions and certain result that is going to be used from now onwards that include the notion of probability, random variables, generalities on some probability distributions, survival, hazard and quantile function, moments, the moment generating function and the characteristic function.

In the second chapter we focus on the one-parameter Lindley distribution. We explore its properties such as its probability density function together with its cumulative distribution function, survival, hazard rate and residual life function, moment and its related measures, the characteristic function, the moment generating function, its Lorenz curve, estimation of the parameter by method of moments, maximum likelihood and Bayesian approach and finally simulation.

In the third chapter we examines Lindley with two parameters Lindley distributions i.e the two-parameter Lindley distribution and the two-parameter weighted Lindley distribution. Properties and estimations of parameters of these distributions are also studied.

Finally, the fourth chapter focuses on the applications of Lindley distributions where in the first part we discuss the discrimination procedure between random variables, followed by application to real data and simulated data. We finish this thesis with a conclusion.

Generalities and some probability distributions 1

1.1 Introduction

The purpose of this chapter is to recall the fundamental concepts and tools required in this work, based on the classical literature Balakrishnan et al.[5], Casella and Berger [7], Jacod and Protter [13], Lehmann and Casella [17] and Ross [23].

1.2 Generalities

1.2.1 Basic Definitions

Definition 1.1 ($\sigma-$**field**) *A family $\mathcal{A}$ of subsets of a nonempty set Ω is a $\sigma-$field ($\sigma-$algebra) of Ω if it satisfies the following three axioms:*

i) $\Omega \in \mathcal{A}$.

ii) *If $A \in \mathcal{A}$ then $\bar{A} \in \mathcal{A}$.*

iii) *For any sequence $(A_n)_{n\in\mathbb{N}}$ of elements of $\mathcal{A}$, , we have $\bigcup\limits_{n\in\mathbb{N}} A_n \in \mathcal{A}$.*

Definition 1.2 (Probability) *Given $(\Omega, \mathcal{A})$, a measurable space, where Ω is the set of all possible outcomes of the experiment and $\mathcal{A}$ is a $\sigma-$ field of subsets of Ω. We call a probability an application*

$$P : \mathcal{A} \to [0, 1]$$

such that;

- $P(\Omega) = 1$
- $P\left(\bigcup\limits_{i=1}^{\infty} A_i\right) = \sum\limits_{i=1}^{\infty} P(A_i)$, *for any sequence of mutually exclusive events $A_1, A_2, \ldots$ (that is, events for which $A_i \cap A_j = \varnothing$ when $i \neq j$).*

The triplet $(\Omega, \mathcal{A}, P)$ is called a probability space.

Proposition 1.1 (Properties of a probability) *Let $(\Omega, \mathcal{A}, P)$ be a probability space, the probability P satisfies the following properties.*

i) $\forall A \in \mathcal{A},\ 0 \leq P(A) \leq 1.$

ii) $P(\varnothing) = 0.$

iii) *For any finite sequence $(A_i)_{1 \leq i \leq n}$ of pairwise disjoint elements of $\mathcal{A}$, we have* $P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i).$

iv) *For each event $A \in \mathcal{A}$, $P(A) = 1 - P(\bar{A})$.*

v) $\forall A, B \in \mathcal{A},$ *if $A \subset B$ then $P(A) \leq P(B)$ and $P(B \setminus A) = P(B) - P(A)$.*

vi) $\forall A, B \in \mathcal{A},\quad P(A \cup B) = P(A) + P(B) - P(A \cap B).$

Definition 1.3 (Conditional Probability) *Given $(\Omega, \mathcal{A}, P)$ a probability space and A and B two events of $\mathcal{A}$. The conditional probability of event B give that event A is realised is given by;*

$$P(B|A) = \frac{P(B \cap A)}{P(A)},\ \text{ defined for P(A)} \neq 0.$$

From the conditional probability formula, we deduct that

$$P(A \cap B) = P(A)P(B|A) = P(B)P(A|B).$$

Theorem 1.1 (Bayes Rule) *Let $(H_n)_n$ be a disjoint sequence of events such that $P(H_n) > 0,\ n = 1, 2, ...,$ and $\bigcup_{n=1}^{\infty} H_n = \Omega$. Let $B \in \mathcal{A}$ with $P(B) > 0$ then*

$$P(H_j|B) = \frac{P(H_j)P(B|H_j)}{\sum_{i=1}^{\infty} P(H_i)P(B|H_i)},\ j = 1, 2, ...$$

Proof. From the conditional probability formula,

$$P(H_j|B) = \frac{P(H_j \cap B)}{P(B)} \Rightarrow P(H_j \cap B) = P(B)P(H_j|B)$$

Also

$$P(B|H_j) = \frac{P(B \cap H_j)}{P(H_j)} \Rightarrow P(H_j \cap B) = P(H_j)P(B|H_j)$$
$$\Rightarrow P(B)P(H_j|B) = P(H_j)P(B|H_j)$$
$$\Rightarrow P(H_j|B) = \frac{P(H_j)P(B|H_j)}{P(B)},$$

where $P(B) = \sum_{i=1}^{\infty} P(H_i)P(B|H_i)$ is the total probability

Thus

$$P(H_j|B) = \frac{P(H_j)P(B|H_j)}{\sum_{i=1}^{\infty} P(H_i)P(B|H_i)}, \; j = 1, 2, \ldots$$

□

1.2.2 Random variables

Definition 1.4 *Given $(E, \mathcal{E})$ as a measurable space, a random variable (r.v.) defined on the probability space $(\Omega, \mathcal{A}, P)$, with values in $(E, \mathcal{E})$ is any measurable function $X : \Omega \to E$ i.e ,*

$$\forall B \in \mathcal{E}, \; X^{-1}(B) = \{\omega : X(\omega) \in B\} \in \mathcal{A}$$

A random variable X is real-valued if X is a measurable function from $(\Omega, \mathcal{A})$ to $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, where $\mathcal{B}_{\mathbb{R}}$ is Borel sets.

Remark 1.1 *The set $D_X = X(\Omega) = \{X(w), w \in \Omega\}$ is called a range of Omega or support of X.*

If the support of the random variable X is a discrete set, then X is called a discrete random variable, but if it is a continuous set, then X is called a continuous random variable.

In the remainder of this work, we will only discuss real-valued random variables.

Definition 1.5 (Probability distribution) *Let X be a real random variable defined on the probability space $(\Omega, \mathcal{A}, P)$. The distribution of X is the positive measure denoted P_X with total mass 1, defined on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, by*

$$\forall B \in \mathcal{B}_{\mathbb{R}}, \quad P_X(B) = P[X \in B]$$

1.2.3 Cumulative distribution function

Definition 1.6 *Let X be a random variable defined on the sample space Ω. The cumulative distribution function(c.d.f), or simply the distribution function of X is the real-valued function*

$$F : \mathbb{R} \to [0,1]$$

defined by

$$F(x) = P(X \leq x) = P_X(]-\infty, x]) = P(\omega \in \Omega : X(\omega) \leq x).$$

Proposition 1.2 (Properties of a cumulative distribution function)

Let X be a random variable, and let F be its cumulative distribution function.

i) *If $x \leq y$, then $F(x) \leq F(y)$, so that F is an increasing function;*

ii) *We have*

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

iii) *Let $(x_n)_{n \geq 1}$ be a decreasing sequence of real numbers such that $\lim_{n \to \infty} x_n = x$. Then we have*

$$\lim_{n \to \infty} F(x_n) = F(x).$$

We mention that Property (iii) of Proposition 1.2 means that the function F is a right-continuous function.

Proposition 1.3 *Let a and b be two real numbers such that $a < b$. Then we have*

$$P(a < X \leq b) = F(b) - F(a).$$

Definition 1.7 (Mass function) *A random variable X is discrete if it takes countably many values $\{x_1, x_2, \ldots\}$. We define the probability function or probability mass function for X by $p(x) = P(X = x)$.*

Definition 1.8 (Density function) *A random variable X is continuous if there exists a function f such that $f(x) \geq 0$ for all x, $\int_{-\infty}^{\infty} f(x)dx = 1$ and for every $a \leq b$,*

$$P(a < X < b) = \int_a^b f(x)dx.$$

The function f is called the probability density function (P.D.F). We have that

$$F(x) = \int_{-\infty}^{x} f(t)dt$$

and $f(x) = F'(x)$ at all points x at which F is differentiable.

1.2.4 Convolution product and distribution of sum of random variables

Convolution is one of the basic operations performed in image and signal processing. From a purely mathematical standpoint, convolution is an integral. Imagine two functions f and g. The purpose of the operation is to shift g over f. The resulting overlap that occurs when g is shifted over f is the convolution of f and g.

Definition 1.9 (Convolution product) *Given two real functions f and g, the convolution product of these two functions is given by*

$$f * g(t) = \int_{-\infty}^{+\infty} f(\tau)g(t-\tau)d\tau \tag{1.1}$$

Remark 1.2 *For functions f, g supported on only $[0, \infty[$ (i.e , zero for negative arguments), the integrated limits can be truncated, resulting in*

$$f * g(t) = \int_{0}^{t} f(\tau)g(t-\tau)d\tau \quad \text{for} \quad f, g : [0, \infty[\to \mathbb{R}.$$

Some properties of a convolution product

i) Commutativity: $f * g = g * f$.

ii) Distributivity: $f * (g + h) = f * g + f * h$.

iii) Associativity: $[f(t) * g(t)] * h(t) = f(t) * [g(t) * h(t)]$.

The convolution product is applicable where knowing the distribution of the sum of random variables is needed. Let X and Y are two independent random variables. The distribution of the sum $Z = X + Y$ of two independent integer-valued (and hence discrete) random variables is defined by

$$P(Z = z) = \sum_{k=-\infty}^{+\infty} P(X = k)P(Y = z - k)$$

For independent, continuous random variables with probability density functions (P.D.F) f, g and cumulative distribution functions F , G respectively, we have that the cumulative distribution function of the sum is defined by

$$P(Z \leq z) = \int_{-\infty}^{+\infty} F(z-t)g(t)dt = \int_{-\infty}^{\infty} G(t)f(z-t)dt = F * G(z)$$

1.2.5 Survival, hazard and quantile function

1. Survival function

This gives the probability that a patient , or other object will survive past a certain time. It is also known as the survivor function or reliability function.

Definition 1.10 *Let X be a non negative random variable with cumulative distribution function F and t be the time to the event of interest (survival times, failure rate etc). The survival function of X is the real-valued function noted S given by*

$$S(t) = P(X > t) = 1 - F(t) \tag{1.2}$$

If t is the time to death, then $S(t)$ is the probability that a subject can survive beyond time t.

2. Hazard function

It is also known as failure rate function. It is defined as the conditional probability that the phenomenon stops after a duration X given that we have attained this duration. It is always given by;

$$h(x) = \frac{f(x)}{S(x)} \tag{1.3}$$

Remark 1.3 *Let H be the Cumulative hazard function , i.e,*

$$H(x) = \int_{-\infty}^{x} h(t)dt = \int_{-\infty}^{x} \frac{f(t)}{S(t)}dt$$

since

$$S(t) = 1 - F(t) \Rightarrow f(t) = -S^{'}(t).$$

Thus

$$H(x) = \int_{-\infty}^{x} \frac{-S^{'}(t)}{S(t)}dt = -\ln(S(x))$$

So,

$$S(x) = e^{-H(x)} \text{ and } f(x) = h(x)e^{-H(x)}$$

3. Quantile function

Definition 1.11 *Let X be a random variable with cumulative distribution function F. The quantile function of X is the function*

$$Q_X : [0,1] \to \mathbb{R}$$

defined by

$$Q_X(u) = \inf\{x : F(x) \geq u\}$$

If F is strictly increasing and continuous then $Q_X(u) = F^{-1}(u)$, where F^{-1} is the inverse function of F.

The u^{th} quantile of X is values x_u of X such that $F(x_u) = u$, i.e, $x_u = Q_X(u)$.

Proposition 1.4 (Properties of a quantile function) *Suppose $X : \Omega \mapsto \mathbb{R}$ is a random variable with cumulative density function F and quantile function Q_X then*

a) *Q_X is a non-decreasing function.*

b) $Q_X(F(x)) \leq x, \forall x \in \mathbb{R}$.

c) $F(Q_X(u)) \geq u, \forall u \in (0,1)$

d) *$Q_X(u) \leq x$ if and only if $F(x) \geq u$.*

Proof. **a)** Let $A(u) := \{y : F(y) \geq u\}$ and let $u_1, u_2 \in (0,1)$ such that $u_1 \leq u_2$, then since F is a non-decreasing function, we have $A(u_1) \supseteq A(u_2)$ and hence $Q_X(u_1) \leq Q_X(u_2)$.

b) We have $Q_X(F(x)) = \inf A(F(x))$, but clearly $x \in A(F(x))$ thus

$$Q_X(F(x)) \leq x, \quad \forall x \in \mathbb{R}.$$

The proof of the properties c) and d) is evident. □

1.2.6 Moments of a random Variable

The moments of random variable are mathematical quantities that caracterize the distribution of that variable. In the following, let $(\Omega, \mathcal{A}, P)$ be a probability space and X a random variable.

1. Mathematical expectation of a random variable

Definition 1.12 *The mathematical expectation (or mean) of the random variable X is denoted by $\mathbb{E}(X)$ (or μ) and is defined as*

$$\mathbb{E}(X) = \sum_{x \in D_X} xP(X = x), \text{ if } X \text{ is discrete}$$

and

$$\mathbb{E}(X) = \int_{\mathbb{R}} xf(x)dx, \text{ if } X \text{ is continuous,}$$

where $f(x)$ being the p.d.f of X and D_X is its range.

Proposition 1.5 (Balakrishnan et al. [5]) *Let X be a discrete random variable with probability function p and range D_X. Let $g : \mathbb{R} \mapsto \mathbb{R}$ be a real function. Then the expectation of the random variable $Y = g(X)$ is given by*

$$\mathbb{E}[g(X)] = \sum_{x \in D_x} g(x)p(x).$$

Proposition 1.6 (Balakrishnan et al.[5]) *Let X be a continuous random variable with density function f and $g : \mathbb{R} \mapsto \mathbb{R}$ be a real function. Then the expectation of the random variable $Y = g(X)$ is given by*

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)\mathrm{d}x.$$

Proposition 1.7 (Balakrishnan et al.[5]) *Let X be a random variable and a and b be two real numbers. Then,*

$$\mathbb{E}(aX + b) = a\mathbb{E}(X) + b.$$

2. Moments of any order

Definition 1.13 *The r th moment of a random variable X is defined by*

$$\mathbb{E}(X^r) = \sum_{x_i \in D_X} x^r P(X = x_i), \text{ if } X \text{ is discrete}$$

,

$$\mathbb{E}(X^r) = \int_{-\infty}^{+\infty} x^r f(x)dx, \text{ if } X \text{ is continuous,}$$

Remark 1.4 *The first-order moment is the expectation of the random variable X.*

3. Variance of a random variable

The mean of a random variable doesn't give the information about dispersion of the values of the variable around its mean. We therefore define two parameters of dispersion i.e variance and standard deviation which measure the dispersion of the values of X around the mean $\mathbb{E}(X)$.

Definition 1.14 *Let X be a random variable for which the expectation $\mathbb{E}(X)$ exists. The variance of the random variable X is the quantity noted by $Var(X)$ (or σ^2) defined as*

$$Var(X) = \sigma^2 = \mathbb{E}\left[(X - \mathbb{E}(X))^2\right]$$

Remark 1.5 *Variance can also be found by the following expression;*

$$Var(X) = \sigma^2 = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$$

Indeed,

$$\begin{aligned} Var(X) &= \mathbb{E}((X - \mathbb{E}(X))^2) \\ &= \mathbb{E}(X^2 - 2X\mathbb{E}(X) + \mathbb{E}^2(X) \\ &= \mathbb{E}(X^2) - 2\mathbb{E}(X(\mathbb{E}X)) + \mathbb{E}(\mathbb{E}^2(X)) \\ &= \mathbb{E}(X^2 - 2\mathbb{E}^2(X) + \mathbb{E}^2(X) \\ &= \mathbb{E}(X^2) - \mathbb{E}^2(X) \end{aligned}$$

Proposition 1.8 (Properties of the variance)
Let X be a random variable for which the expectation $\mathbb{E}(X)$ exists. Let a and b be two real numbers.

i) $Var(aX + b) = a^2 Var(X)$;

ii) $Var(aX) = a^2 Var(X)$.

iii) $Var(X) \geq 0$.

Remark 1.6 *If $Var(X) = 0$, it means that the random variable X is a constant equal to its mean.*

Definition 1.15 (Standard deviation) *The standard deviation denoted σ is the square root of variance. i.e*

$$\sigma = \sqrt{Var(X)}$$

4. The central moment of order r

Definition 1.16 *The central moment of order r of a random variable X noted by μ_r is defined as*

$$\mu_r = \mathbb{E}((X - \mathbb{E}(X))^r) = \sum_{x \in D_X} (x - \mathbb{E}(X))^r P(X = x),$$

if X is discrete and if the sum exists.

$$\mu_r = \mathbb{E}((X - \mathbb{E}(X))^r) = \int_{\mathbb{R}} (x - \mathbb{E}(X))^r f(x) dx,$$

if X is continue, where f is the probability density function of the random variable X.

Remark 1.7 *The second-order central moment is the variance of the random variable X.*

5. The coefficient of variation

The Coefficient of variation noted γ is used to compare the variability of two or more sets of data even if they are on different scales or have different units of measurement and is often in statistics to compare the dispersion of data sets with different units or scales of measurements.

The coefficient of variation γ is calculated by

$$\gamma = \frac{\text{standard deviation}}{\text{mean}} = \frac{\sigma}{\mathbb{E}(X)}$$

6. Skewness

The skewness noted $\sqrt{\beta_1}$ tells us the direction of dispersion about the center of the distribution. Skewness refers to the lack of symmetry in distribution. Symmetry signifies that the value of variables are equidistant from the average on both sides. In other words, a balanced pattern of a distribution is called symmetrical distribution; whereas unbalanced pattern of distribution is called asymmetrical distribution. A distribution is said to be positive and negative skewed respectively when the mean $>$ median $>$ mode and mean $<$ median $<$ mode. For symmetrical distribution, mean $=$ median $=$ mode.

The Skewness is calculated by

$$\sqrt{\beta_1} = \frac{\mathbb{E}\left[(X - \mathbb{E}(X))^3\right]}{\sigma^3} = \frac{\mu_3}{\sigma^3}$$

7. Kurtosis

The kurtosis β_2 is basically used to measure the extent how the distribution is more peaked or less peaked than the normal distribution curve. The degree of kurtosis of a distribution is measured relative to that of a normal curve. curves with greater peakedness than the normal curve are called Leptokurtic . Those which are more flat than the normal curve are called Platykurtic and the normal curve is called Mesokurtic.
Kurtosis is calculated by

$$\beta_2 = \frac{\mu_4}{\mu_2^2},$$

where μ_4 is the fourth order central moment of the distribution and μ_2 is the second order central moment of the distribution of a random variable X.

Remark 1.8

- *If $\beta_2 = 3$, the curve is said to be mesokurtic.*
- *If $\beta_2 < 3$, the curve is said to be platykurtic.*
- *If $\beta_2 > 3$, the curve is said to be leptokurtic.*

1.2.7 The moment generating function

Given X a random variable. The moment generating function of X is given by

$$M_X(t) = \mathbb{E}(e^{tX})$$

Remark 1.9 *If X is a continuous random variable then*

$$M_X(t) = \mathbb{E}(e^{tX}) = \int_{\mathbb{R}} e^{tx} f(x)dx,$$

with f being the probability density function of the random variable X.

When the moments of all orders exist and their exponential generating series has a non zero radius of convergence R, then for all $t \in]-R, R[$, we have:

$$M_X(t) = \mathbb{E}\left(e^{tX}\right) = \sum_{k=0}^{\infty} \frac{\mathbb{E}\left(X^k\right)}{k!} t^k.$$

We prove this property if X is a continuous random variable with density function f; the case where X is discrete is proved similarly.

$$\begin{aligned} M_X(t) &= \int_{\mathbb{R}} \left[1 + tx + \frac{(tx)^2}{2!} + \frac{(tx)^3}{3!} + ...\right] f(x)dx \\ &= \int_{\mathbb{R}} f(x)dx + t\int_{\mathbb{R}} xf(x)dx + \frac{t^2}{2!}\int_{\mathbb{R}} x^2 f(x)dx + ... \\ &= \sum_{k\geq 0} \frac{t^k}{k!}\int_{\mathbb{R}} x^k f(x)dx \\ &= \sum_{k\geq 0} \mathbb{E}(X^k)\frac{t^k}{k!} \end{aligned}$$

With this function, we can be able to find any moment of any order by finding its derivative at a point $t = 0$ .i.e

$$M_X^{(r)}(0) = \mathbb{E}(X^r).$$

The moment generating function for the exponential and standard normal distributions are $M_X(t) = \dfrac{\lambda}{\lambda - t}$ and $M_X(t) = e^{\frac{t^2}{2}}$ respectively.

1.2.8 Characteristic functions

Definition 1.17 *A characteristic function of a real random variable X denoted φ_X is the function $\varphi_X : \mathbb{R} \to \mathbb{C}$ defined by*

$$\varphi_X(t) = \mathbb{E}(e^{itX})$$

In other words ,

$$\varphi_X(t) = \mathbb{E}(\cos tX + i\sin tX) = \mathbb{E}(\cos tX) + i\mathbb{E}(\sin tX)$$

Properties of the characteristic functions

Let φ_X be a characteristic function of a random variable X.

i) $\varphi_X(0) = 1$.

ii) $|\varphi_X(t)| \leq 1$.

iii) $\varphi_X(-t) = \overline{\varphi_X(t)} = \mathbb{E}(e^{-itX}) = \mathbb{E}(\cos tX - i \sin tX)$.

iv) φ_X is uniformly continuous. That is for all $\epsilon > 0$, there exists $\delta > 0$ such that $|\varphi(t) - \varphi(s)| \leq \epsilon$ whenever $|t - s| \leq \delta$.

v) $\varphi_{aX+b} = \varphi_{aX} e^{itb} = \varphi_X(at) e^{itb}$, $a, b \in \mathbb{R}$,

vi) Let X and Y are two independent random variables with characteristic functions φ_X and φ_Y, then the characteristic function of $X + Y$ is defined by

$$\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t),$$

vii) The characteristic function of $-X$ is the complex conjugate $\overline{\varphi_X}(t)$ i.e, $\varphi_{-X}(t) = \varphi_X(-t) = \overline{\varphi_X(t)}$.

viii) If a random variable X has moments up to k-th order, then the characteristic function φ_X is k times continuously differentiable on the entire real line. In this case

$$\mathrm{E}\left[X^k\right] = i^{-k}\varphi_X^{(k)}(0).$$

1.2.9 Central limit theorem

Theorem 1.2 (CLT) *Let $X_1, X_2, \ldots.., X_n$ be a sequence of independent random variables having mean 0 and variance σ^2 and the common distribution F and moment generating function M defined in the neighbourhood of zero. Let $S_n = \sum_{i=1}^{n} X_i$ then*

$$\lim_{n\to\infty} P\left(\frac{S_n}{\sigma\sqrt{n}} \leq x\right) = \phi(x), -\infty < x < \infty,$$

where ϕ is the cumulative distribution function of the standard normal distribution.

Proof. Let $Z_n = \dfrac{S_n}{\sigma\sqrt{n}}$. We will show that the moment generating function of Z_n tends to the moment generating function of the standard normal distribution. Since S_n is the sum of independent random variables, $M_{S_n}(t) = [M(t)]^n$ and $M_{Z_n}(t) = \left[M\left(\dfrac{t}{\sigma\sqrt{n}}\right)\right]^n$.

Using Taylor series expansion about Zero:

$$M(s) = M(0) + sM^{'}(0) + \frac{1}{2}s^2M^{''}(0) + \varepsilon_n.$$

Since $\mathbb{E}(X) = 0,\ M^{'}(0) = 0$ and $M^{''}(0) = \sigma^2$. As $n \to \infty$, $\left(\frac{t}{\sigma\sqrt{n}}\right) \to 0$, and

$$M\left(\frac{t}{\sigma\sqrt{n}}\right) = 1 + \frac{1}{2}\sigma^2\left(\frac{t}{\sigma\sqrt{n}}\right)^2 + \varepsilon_n$$

$$M_{Z_n}(t) = \left(1 + \frac{t^2}{2n} + \varepsilon_n\right)^n$$

It can be shown that if $a_n \to a$, then

$$\lim_{n\to\infty}\left(1 + \frac{a_n}{n}\right)^n = e^a$$

From this result, it follows that $M_{Z_n}(t) \to e^{\frac{t^2}{2}}$ as $n \to \infty$ where $e^{\frac{t^2}{2}}$ is the moment generating function of the standard normal distribution, as was to be shown. □

1.3 Some probability distributions

Probability distributions are mathematical models that describe the likelihood of different outcomes in a random experiment.They are divided into two categories i.e discrete distributions and continuous distributions. Some of these distributions include;

1.3.1 Poisson distribution.

This distribution allows modelling of observation of the phenomenon that produces events at a known rate. We suppose that only one event occurs at a time. Considering X a random variable that gives the number of events observed in a time. We have then a Poisson phenomenon and the random variable that gives the number of events in a time has Poisson distribution noted as;

$$X \sim P(\lambda),$$

where λ is the average rate of occurrence.
Its probability law is written as;

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

Its expectation and variance are the same equal to the parameter λ.

Its characteristic function is

$$\varphi_X(t) = e^{\lambda(e^{2t}-1)}.$$

1.3.2 Exponential distribution

Given $\lambda > 0$. A random variable X has exponential distribution of parameter λ if its probability density function f is given by;

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

We note $X \sim exp(\lambda)$
Its cumulative distribution function is given by;

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

This distribution is often used to model the inter-arrival times between successive arrivals of customers in a queueing system. Its expectation and variance are given by: $\mathbb{E}(X) = \dfrac{1}{\lambda}$ and $Var(X) = \dfrac{1}{\lambda^2}$.

Its characteristic function is given by

$$\varphi_X(t) = \frac{\lambda}{\lambda - it}$$

1.3.3 The two-parameter Weibull distribution

The two-parameter Weibull distribution is a widely used probability distribution in reliability engineering and failure analysis. It is characterized by its flexibility in modeling various types of data, particularly for the life data of products and materials. Although it was first identified by Fréchet in 1927, it is named after Waloddi Weibull and is a cousin to both the Fréchet and Gumbel distributions. Waloddi Weibull was the first to promote the usefulness of this distribution by modeling data sets from various disciplines.

Definition 1.18 *A random variable X is said to have a two-parameter Weibull distribution with parameters (α, β), $\alpha > 0$, $\beta > 0$, if its density function f is given by*

$$f(x) = \begin{cases} \frac{\beta}{\alpha^{\beta}} x^{\beta-1} e^{-\left(\frac{x}{\alpha}\right)^{\beta}} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

In this density function, our α represents the scale parameter, and our β represents the shape parameter.

When $\beta = 1$, the Weibull distribution simplifies to the exponential distribution.

The distribution function of a two-parameter Weibull distribution can also be derived and is defined as:

$$F(x) = \begin{cases} 1 - e^{-(t/\alpha)^{\beta}} & x > 0 \\ 0 & x \leq 0 \end{cases}.$$

1.3.4 Generalized exponential distribution

The generalized exponential distribution is an extension of the standard exponential distribution has been proposed and studied by Gupta and Kundu[11] (see also Kundu and Gupta [15]). It includes an additional shape parameter that allows it to model a wide range of data compared to simple exponential distribution.
For a random variable X following a generalized exponential distribution, its probability density function is given by

$$f(x) = \alpha\lambda(1 - e^{-\lambda x})^{\alpha-1}, \quad x > 0 \tag{1.4}$$

where α and λ are shape and scale parameters.

The cumulative distribution function of the generalized exponential distribution is

$$F(x) = (1 - e^{-\lambda x})^{\alpha}, \quad x > 0, \quad \alpha > 0, \quad \lambda > 0 \tag{1.5}$$

The density function of the generalized exponential distribution can take different shapes. For $\alpha \leq 1$, it is a decreasing function and for $\lambda > 1$, it is unimodal, skewed, right tailed similar to that of Weibull.

1.3.5 Uniform distribution

With this probability distribution, all outcomes are equally likely. Within a given range, each value has the same probability of occurring. The probability distribution function (p.d.f) of uniform distribution $(u([a,b]))$ is given by

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{if } x \in [a,b] \\ 0 & \text{otherwise} \end{cases}$$

where a and b are the lower and upper bounds of the distribution. Its cumulative distribution function is given by;

$$F(x) = \begin{cases} 0 & \text{if } x \leq a \\ \dfrac{x-a}{b-a} & \text{if } a \leq x \leq b \\ 1 & \text{if } x \geq b \end{cases}$$

Its expectation and variance are given by

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

Its characteristic function is given by

$$\varphi_X(t) = \frac{1}{b-a}\left(\frac{e^{itb} - e^{ita}}{it}\right)$$

This distribution is commonly used in various fields including statistics, physics, particularly in scenarios where each outcome is equally likely within a specified range.

1.3.6 Normal distribution

It is also known as Gaussian distribution. It is characterised by a bell- shape curve symmetrically centred around its mean. The mean, mode and median of a normal distribution are equal that's why it isn't skewed. The normal distribution is described by two parameters i.e the mean (μ) and standard deviation (σ) which determine its location and spread, respectively.we note it as $N(\mu, \sigma)$

Definition 1.19 *A random variable $X \sim N(\mu,\sigma)$ if its probability density function is given by*

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}},\ x \in \mathbb{R}$$

The analytic expression for the cumulative distribution function of $N(\mu,\sigma)$ doesn't exist. This calls for the need to standardise this distribution.

Standard normal distribution: The standard normal distribution is a particular case of the normal distribution when $\mu = 0$ and $\sigma = 1$ noted by $N(0,1)$, its probability density function is given by

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}, x \in \mathbb{R}.$$

Its characteristic function is given by

$$\varphi_X(t) = \int_{-\infty}^{+\infty} e^{itx} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = e^{-\frac{1}{2}t^2}$$

The cumulative distribution function of $N(\mu,\sigma)$ distribution does not exist and this calls a need to standardise this distribution since only $N(0,1)$ is tabulated.

For a random variable $X \sim N(\mu,\sigma)$, by using change of variable $Z = \dfrac{X-\mu}{\sigma}$, we have

$$P(X \le a) = P\left(\frac{X-\mu}{\sigma} \le \frac{a-\mu}{\sigma}\right) = P\left(Z \le \frac{a-\mu}{\sigma}\right) = \phi\left(\frac{a-\mu}{\sigma}\right)$$

where ϕ is the cumulative distribution function of $N(0,1)$.

For more details on the uni-dimensional and multidimensional normal distribution see Ladjimi[16].

1.3.7 Gamma distribution

The Gamma distribution is a continuous probability distribution that is widely used in statistics to model positive-valued variables that are skewed to the right. It is often used to model waiting times and survival times.

Definition 1.20 (Gamma function) *The Gamma function at a point a is defined as;*

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

Properties 1.1
$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$,
$\Gamma(a+1) = a\Gamma(a)$, *for all* $a > 0$,
$\Gamma(n) = (n-1)!, \;\; n \in \mathbb{N}$.

Definition 1.21 *A random variable X is said to have gamma distribution with parameters* $t > 0, \; \lambda > 0$, *denoted* $\Gamma(t, \; \lambda))$, *if its probability density function is given by;*

$$f(x) = \begin{cases} \dfrac{\lambda e^{-\lambda x}(\lambda x)^{t-1}}{\Gamma(t)} & \textit{if } x \geq 0 \\ 0 & \textit{otherwise} \end{cases}$$

where t is the shape parameter that controls the shape of the distribution. Higher values result in distributions with heavier tails and λ is the rate parameter that controls the rate at which the distribution decays. The exponential probability distribution is the particular case of the gamma distribution we replace t by 1. Its expectation and variance are given by
$\mathbb{E}(X) = \frac{t}{\lambda}$ and $Var(X) = \frac{t}{\lambda^2}$

1.3.8 Beta distribution

Beta distribution is a probability distribution defined on the interval $[0, 1]$. The shape of the distribution is controlled by two parameters usually denoted as a and b. These parameters determine the shape of the distribution curve, with higher values of a leading to a distribution skewed towards the left, and higher values of b leading to a distribution skewed towards the right.

Definition 1.22 (Beta function) *Beta function at a point* $a, \; b > 0$ *is defined as;*

$$\beta(a, \; b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx$$

Properties 1.2
$\beta(a, \; b) = \beta(b, \; a)$
$\beta(a, \; b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$

Definition 1.23 *A random variable X has beta distribution with a and b parameters*

$(X \sim \beta(a,\ b))$ *with* $a,\ b \geq 0$ *if its probability density function is written as;*

$$f(x) = \begin{cases} \dfrac{1}{\beta(a,b)} x^{a-1}(1-x)^{b-1} & \text{if} \quad 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

Its expectation and variance are given by;

$$\mathbb{E}(X) = \frac{a}{a+b} \quad \text{and} \quad \mathrm{Var}(X) = \frac{\mathrm{ab}}{(\mathrm{a}+\mathrm{b})^2(\mathrm{a}+\mathrm{b}+1)}$$

Remark 1.10 *If* $a = b = 1$*, Beta distribution becomes uniform distribution.*

1.4 Estimation of Parameters

All random phenomena are governed by laws of probability, which involve one or more parameters. The goal is always to estimate these parameters from a random sample of the population considered.
Let $X_1, X_2, ..., X_n$ be a sample of random variable X with size n drawn from a population with unknown parameter θ. An estimator $\hat{\theta}$ of θ is a function of the given sample

$$\hat{\theta} = \tau(X_1, X_2, ..., X_n) \tag{1.6}$$

which is a random variable. Replacing $X_1, X_2, ..., X_n$ by their realisations in the equation (1.6) we obtain an estimate of θ.

1.4.1 Quality of estimators

Definition 1.24 (**Bias**) *The bias of an estimator is* $Bias(\hat{\theta}) = \mathbb{E}(\hat{\theta}) - \theta$*. An estimator* $\hat{\theta}$ *is said to be unbiased if* $Bias(\theta) = 0$ *and if* $Bias(\theta) > 0$*, the estimator is said to be positively biased.*

Definition 1.25 (**Convergence of the estimator**) *An estimator* $\hat{\theta}$ *for* θ *is said to be convergent if it converges in probability to* θ*. i.e*

$$\hat{\theta} \xrightarrow{P} \theta \Leftrightarrow \ \forall \epsilon > 0 \quad \lim_{n \to \infty} P(|\hat{\theta} - \theta| > \epsilon) = 0$$

The quality of the estimator $\hat{\theta}$ is measured with the help of mean square error. It is calculated by

$$MSE = \mathbb{E}[(\hat{\theta} - \theta)^2]$$

Definition 1.26 (**Efficiency of the estimator**) *Given two estimators $\hat{\theta}_1$ and $\hat{\theta}_2$ of θ. One estimator is said to be more efficient than the other if its MSE is less than that of the other. Measure of efficiency is equal to* $\dfrac{MSE(\hat{\theta}_1)}{MSE(\hat{\theta}_2)}$.

Definition 1.27 *An estimator is said to be asymptotically unbiased if* $\lim_{n\to\infty} \mathbb{E}(\hat{\theta}) = \theta$

Theorem 1.3 (**Casella and Berger[7]**) *Any unbiased estimator whose variance tends to 0 is convergent.* $(\mathbb{E}(\hat{\theta}) = \theta$ *and* $Var(\hat{\theta}) \to 0) \Leftrightarrow \hat{\theta} \xrightarrow{P} \theta$

Definition 1.28 *Given $\hat{\theta}$ an estimator of the parameter θ of a distribution of a random variable X. Assume that there exist two sequences of strictly positive real functions, $a = a_n(\theta)$ and $b = b_n(\theta)$ such that*

$$\frac{\hat{\theta} - a}{b} \xrightarrow{D} N(0,1).$$

We say that $\hat{\theta}$ is asymptotically normal.

1.4.2 Methods of Estimation

1. Method of Moments

This is a very simple but useful approach of finding an estimators. This approach bases on equating sample moments with theoretical moments of the distribution. Let $X_1, X_2, ..., X_n$ be a random variable from a probability distribution with unknown parameters $\theta_1, \theta_2, ..., \theta_k$. The Estimators $\hat{\theta_1}, \hat{\theta_2}, ..., \hat{\theta_k}$ of these parameters are obtained by solving the following system of equations

$$(S) : \begin{cases} \mathbb{E}(X) = \dfrac{1}{n}\displaystyle\sum_{i=1}^{i=n} X_i \\ \\ \mathbb{E}(X^2) = \dfrac{1}{n}\displaystyle\sum_{i=1}^{i=n} X_i^2 \\ \vdots \\ \mathbb{E}(X^k) = \dfrac{1}{n}\displaystyle\sum_{i=1}^{i=n} X_i^k \end{cases}$$

2. Method of Maximum likelihood

This method involves finding the value of θ that maximizes the likelihood function. Given a random sample $X_1, X_2, ..., X_n$. The likelihood function of this random sample noted $L_n(\theta)$ is given by;

$$L_n(\theta) = L(\theta|X_1, X_2, ..., X_n) = \prod_{i=1}^{n} L(\theta|X_i) = \prod_{i=1}^{n} p(X_i;\theta).$$

where $p(X_i;\theta)$ is the mass function or the density . In many cases, instead of using the likelihood function, we often work with the log-likelihood function noted $\ell_n(\theta)$.

$$\ell_n(\theta) = \ln L_n(\theta) = \sum_{i=1}^{n} \ln p(X_i;\theta)$$

Example 1.1 *Let $x_1, x_2, ..., x_n$ be observations of random variable X which has exponential distribution of parameter $\lambda > 0$, its likelihood function is given by*

$$L_n(\lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i}$$

its log likelihood function $\ell_n(\lambda) = n \log \lambda - \lambda \sum_{i=1}^{n} x_i$

$$\frac{\partial \ell_n(\lambda)}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i = 0$$

$$\Rightarrow \lambda = \frac{n}{\sum_{i=1}^{n}} x_i = \frac{1}{\bar{X}}$$

Also its clear that

$$\frac{\partial^2 \ell_n(\lambda)}{\partial \lambda^2} = \frac{-n}{\theta^2} < 0.$$

Thus

$$\hat{\lambda} = \frac{1}{\bar{X}}.$$

By using the method of moments we get the some estimator of λ.
Since, $\mathbb{E}(X) = \int_0^{\infty} \lambda x e^{-\lambda x} = \frac{1}{\lambda}$ *and by using the equality* $\mathbb{E}(X) = \frac{1}{n}\sum_{i=1}^{n} x_i$ *it follows that*

$$\hat{\lambda} = \frac{1}{\bar{X}}.$$

We can easily show that $\hat{\lambda}$ is an unbiased estimator with variance $V(\hat{\lambda}) = n\lambda^2$.

Example 1.2 *Given a random sample $X_1, X_2, ..., X_n$ that has two- parameter Weibull* *distribution. Its probability density function is given by;*

$$f(x) = \beta\alpha^{-\beta}x^{\beta-1}e^{-(\frac{x}{\alpha})^{\beta}}, \; x > 0, \; \alpha > 0, \; \beta > 0 \tag{1.7}$$

where α is the scale parameter and β is the shape parameter that defines the form of the distribution. Its Likelihood and log-likelihood functions respectively are

$$L_n(\alpha, \beta) = \frac{\beta^n}{\alpha^n}e^{-\sum_{i=1}^{n}(\frac{x_i}{\alpha})^{\beta}}\prod_{i=1}^{n}(\frac{x_i}{\alpha})^{\beta-1}, \tag{1.8}$$

and

$$\ell_n(\alpha, \beta) = nlog\beta - nlog\alpha + (\beta - 1)\sum_{i=1}^{n} log(\frac{x_i}{\alpha}) - \sum_{i=1}^{n}(\frac{x_i}{\alpha})^{\beta} \tag{1.9}$$

MLE for the Weibull distribution can be obtained by solving the equations resulting from setting the two partial derivatives of the log-likelihood function and when the shape parameter β is known.

$\hat{\beta}$ is the solution of $\frac{1}{\hat{\beta}} + \frac{1}{n}\sum_{i=1}^{n}\ln x_i - \frac{\sum_{i=1}^{n} x_i^{\beta}\ln x_i}{\sum_{i=1}^{n} x_i^{\beta}} = 0$ *So,* $\hat{\alpha} = \left(\frac{\sum_{i=1}^{n} x_i^{\beta}}{n}\right)^{\frac{1}{\beta}}$ *and*

$$\hat{\beta} = \left[\left(\sum_{i=1}^{n} x_i^{\hat{\beta}}\ln x_i\right)\left(\sum_{i=1}^{n} x_i^{\hat{\beta}}\right)^{-1} - \frac{1}{n}\sum_{i=1}^{n}\ln x_i\right]^{-1}.$$

In order to solve above equation, one can apply a suitable iterative procedure such as the Newton Raphson method see Almazah and Ismail [2].

3. Bayesian estimation

There are many situations where maximum likelihood estimator doesn't converge especially with higher dimension models. In such cases, the use of Bayesian methods is sought. Bayesian estimation is a statistical approach that combines information collected from experimental data with the knowledge one has prior to performing the experiment. The cornerstone of Bayesian methodology is the Bayes theorem which helps to update our beliefs inform of the probability statements about the parameters after the sample has been taken.

The conditional distribution of the parameters after observing the data is the posterior distribution that integrates the prior and the sample information. Suppose we have two discrete random variables, X and Y. The Bayes rule for the conditional probability is

$$P(y|x) = \frac{P(x,y)}{P_X(x)} = \frac{P(x|y)P_Y(y)}{P_X(x)} = \frac{P(x|y)P_Y(y)}{\sum_y P(x|y)P_Y(y)dy} \tag{1.10}$$

The denominator in this expression is a fixed normalising factor that ensures that $\sum_y P(y|x) = 1$.

If Y is continuous, the Bayes theorem can be stated as

$$P(y|x) = \frac{P(x,y)}{P_X(x)} = \frac{P(x|y)P_Y(y)}{P_X(x)} = \frac{P(x|y)P_Y(y)}{\int_y P(x|y)P_Y(y)dy} \tag{1.11}$$

In Bayesian terminology, $P_Y(y)$ represents the probability statement of prior beliefs, $P(x|y)$ is the probability of data x given prior beliefs which is the likelihood and the updated probability $P(y|x)$ is the posterior. We can express the posterior distribution as proportional to the product of likelihood and prior distribution ($P(y|x) \propto P(x|y)P(y)$)

Definition 1.29 *The probability of θ given data $X_1, X_2, ..., X_n$ is the posterior distribution which is given by*

$$\pi(\theta|X) = \frac{L_n(\theta)\pi(\theta)}{g(x)} \tag{1.12}$$

where $g(x)$ is the marginal distribution of X which can be calculated as

$$g(x) = \begin{cases} \sum_{\Theta} L_n(\theta)\pi(\theta), & \text{in the discrete case} \\ \\ \int_{\Theta} L_n(\theta)\pi(\theta)d\theta, & \text{in the continuous case} \end{cases}$$

where Θ is a set of the unknown parameters of the distribution.

Definition 1.30 (Loss function) *We call loss function, all function*

$$\begin{array}{rcrcl} L & : & \Theta \times \mathbb{A} & \to & \mathbb{R} \\ & & (\theta, \delta) & \mapsto & L(\theta, \delta) \end{array}$$

where $\mathbb{A}$ is the space for the actions (decisions) and δ is a decision taken. In summery, loss function measures the loss that happens when a decision is taken. $L(\theta, \delta)$ is negative means gain. A decision is said to be the best decision if it brings us zero loss. There exist different loss functions such as loss function absolute (LAF), squared error loss function(SELF), the Linex (linear-exponential) loss function (LLF), the loss function $0-1$ (LF), the weighted squared error Loss function (WSELF), the precautionary loss function (PLF), the modified (quadratic) squared error Loss function (M/QSELF) and many others. However, the most used ones are squared error loss function(SELF) and the Linux (linear-exponential) loss function (LLF).

Definition 1.31 (Loss a posterior) *Loss a posterior noted $\rho(\pi,\delta)$ is the mean of the loss function with respect to the a posterior distribution. i.e*

$$\rho(\pi,\delta) = \mathbb{E}\left[L(\theta,\delta)\right] \tag{1.13}$$

It is also called the risk a posterior.

Bayes estimator

Bayes estimators are evaluated using loss functions.

Proposition 1.9 (Lehmann and Casella[17]) *Bayes estimator δ^{π}, associated with the prior distribution $\pi(\theta)$ and the squared error loss function(SELF) is the a posterior mean of θ.*

Proof. By definition, Bayes estimator σ^{π} minimizes the loss posterior function.

$$\rho(\pi,\delta) = \mathbb{E}[(\theta-\delta)^2] = \mathbb{E}[\theta^2+\delta^2-2\theta\delta]$$

$$= \mathbb{E}(\theta^2)+\delta^2-2\delta\mathbb{E}(\theta)$$

Finding the derivative with respect to δ and equating it to zero we get

$$\frac{d\rho(\pi,\delta)}{d\delta} = 2\delta - 2\mathbb{E}(\theta) = 0 \Rightarrow \delta^{\pi} = \mathbb{E}(\theta)$$

$\frac{d^2\rho(\pi,\delta)}{d\delta^2} = 2 > 0 \Rightarrow \delta^{\pi}$ is Bayes estimator. Bayes estimators under the remaining loss functions are given in the table below

□

Table 1.1 – *Bayes estimators under different loss functions*

Loss function	Bayes estimator
LLF=$L(\hat{\theta},\theta) = \exp(a(\hat{\theta}-\theta)) - a(\hat{\theta}-\theta) - 1$	$-\frac{1}{a}\ln \mathbb{E}[\exp(-a\theta)]$
LAF = $\mid \theta - \delta \mid = \begin{cases} \theta-\delta & \text{if } \theta > \delta \\ \delta - \theta & \text{if } \theta \leq \delta \end{cases}$	$Median(\theta\|X)$
$LF = \begin{cases} 0 & \text{if } \mid \theta-\delta \mid < \epsilon \\ 1 & \text{if } \mid \theta-\delta \mid \geq \epsilon \end{cases}$	$Mode(\theta\|X)$
WSELF=$\frac{(\theta-\delta)^2}{\theta}$	$(E(\theta^{-1}\|X))^{-1}$
PLF=$\frac{(\theta-\delta)^2}{\delta}$	$\sqrt{E(\theta^2\|X)}$
M/QSELF=$(1-\frac{\delta}{\theta})^2$	$\mathbb{E}(\theta^{-1}\|X)$

Example 1.3 *It is believed that cross-fertilized plants produce taller offsprings than self-fertilized plants. To obtain an estimate of the proportion θ of cross fertilized plants that are taller, an experimenter observes a random sample of 15 pairs of plants that are exactly the same age. Each pair is grown under the same conditions, with some cross fertilized. Based on previous experiment, the experimenter believes that the following are possible values of θ and that the prior probability for each value of θ is $\pi(\theta)$*

Table 1.2 – *Binomial prior distribution*

θ	0.80	0.82	0.84	0.86	0.88	0.90
$\pi(\theta)$	0.13	0.15	0.22	0.25	0.15	0.10

From the experiment, it is observed that in 13 of 15 pairs, the cross-fertilized plant is taller. The Likelihood of obtaining 13 taller cross-fertilized plants in 15 pairs compared with the different prior values of π is given using the binomial distribution i.e $\pi(\theta) = \binom{15}{13}\theta^{13}(1-\theta)^2$. The summary of prior and posterior probabilities is given in the table below.

Table 1.3 – *Summary of posterior distribution using binomial prior*

Prior value of θ	$\pi(\theta)$	$L_n(\theta)$	$\pi(\theta) \times L_n(\theta)$	$\pi(\theta\|X)$
0.80	0.13	0.2309	3.0017×10^{-2}	0.11029
0.82	0.15	0.2578	3.867×10^{-2}	0.14208
0.84	0.22	0.2787	6.1314×10^{-2}	0.22528
0.86	0.25	0.2897	7.2425×10^{-2}	0.2661
0.88	0.15	0.2870	4.305×10^{-2}	0.15817
0.90	0.10	0.2669	2.669×10^{-2}	0.098064
		Total	0.27217	$0.999984 \approx 1.0$

Bayesian estimate of θ under squared error loss function is

$$\hat{\theta} = 0.84879 \approx 0.85.$$

The priors used in this case are informative priors, because they favoured certain values of θ. For example if $\theta = 0.86$, the prior value of θ is $\pi(\theta) = 0.25$ which is higher than all the rest of the values. If there were no information or no strong prior opinions, then we could select a non-informative prior, which would have assigned equal prior probability of $1/6$ to each of the possible values of θ.

Table 1.4 – *Summary of prior and posterior probabilities using non informative prior.*

Prior value of θ	$\pi(\theta)$	$L_n(\theta)$	$\pi(\theta) \times L_n(\theta)$	$\pi(\theta\|X)$
0.80	1/6	0.2309	3.8483×10^{-2}	0.14333
0.82	1/6	0.2578	4.2967×10^{-2}	0.16003
0.84	1/6	0.2787	0.04645	0.173
0.86	1/6	0.2897	4.8283×10^{-2}	0.17982
0.88	1/6	0.2870	4.7833×10^{-2}	0.17815
0.90	1/6	0.2669	4.4483×10^{-2}	0.16567
		Total	0.2685	1.0

The Bayesian estimate for the non-informative prior $\hat{\theta} = 0.85173$

Example 1.4 *Suppose that a random sample $X_1, X_1, ..., X_n$ has a distribution with probability density function(p.d.f)*

$$f(x) = \begin{cases} \theta x^{\theta-1} & \text{if } x \in [0,1] \\ 0 & \text{otherwise} \end{cases}$$

Suppose that the prior distribution $\pi(\theta) = \dfrac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\beta\theta}$, $\theta > 0$. Bayes estimator is found as follows

Likelihood function $L_n(\theta) = \theta^n \prod_{i=1}^{n} x_i^{\theta-1}$

$$\begin{aligned}\pi(\theta|X) &= \frac{L_n(\theta)\pi(\theta)}{\int_0^{+\infty} L_n(\theta)\pi(\theta)d\theta} \\ &= \frac{\theta^n \prod_{i=1}^{n} x_i^{\theta-1} \dfrac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\beta\theta}}{\displaystyle\int_0^{+\infty} \theta^n \prod_{i=1}^{n} x_i^{\theta-1} \frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\beta\theta}d\theta}\end{aligned}$$

First,

$$\int_0^{+\infty} \theta^n \prod_{i=1}^{n} x_i^{\theta-1} \frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\beta\theta}d\theta = \frac{\beta^\alpha}{\Gamma(\alpha)} e^{-\sum_{i=1}^{n} \ln x_i} \int_0^{+\infty} \theta^{n+\alpha-1}e^{-(\beta-\sum_{i=1}^{n} \ln x_i)\theta}d\theta$$

Let $\alpha_1 = n + \alpha$ and $\beta_1 = \beta - \sum_{i=1}^{n} \ln x_i$

It means that

$$\int_0^{+\infty} \theta^n \prod_{i=1}^{n} x_i^{\theta-1} \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta} d\theta = \frac{\beta^\alpha}{\Gamma(\alpha)} e^{-\sum_{i=1}^n \ln x_i} \frac{\Gamma(\alpha_1)}{\beta_1^{\alpha_1}} \int_0^{+\infty} \frac{\beta_1^{\alpha_1}}{\Gamma(\alpha_1)} \theta^{\alpha_1-1} e^{-\beta_1\theta} d\theta$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)} e^{-\sum_{i=1}^n} \frac{\Gamma(\alpha_1)}{\beta_1^{\alpha_1}}$$

because $\int_0^{+\infty} \frac{\beta_1^{\alpha_1}}{\Gamma(\alpha_1)} \theta^{\alpha_1-1} e^{-\beta_1\theta} d\theta = 1.$

Thus

$$\pi(\theta|X) = \frac{\frac{\beta^\alpha}{\Gamma()\alpha} e^{-\sum_{i=1}^n \ln x_i} \theta^{\alpha_1-1} e^{-\beta_1\theta}}{\frac{\beta^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha_1)}{\beta_1^{\alpha_1}} e^{-\sum_{i=1}^n \ln x_i}}.$$

So,

$$\pi(\theta|X) = \frac{\beta_1^{\alpha_1}}{\Gamma(\alpha_1)} \theta^{\alpha_1-1} e^{-\beta_1\theta}$$

which is a $\Gamma(\alpha_1, \beta_1)$ *density, where* $\alpha_1 = n + \alpha$ *and* $\beta_1 = \beta - \sum_{i=1}^{n} \ln x_i$.

Using the squared error loss function, Bayes estimator

$$\hat{\theta} = \mathbb{E}(\theta|X) = \frac{\alpha_1}{\beta_1} = \frac{n+\alpha}{\beta - \sum_{i=1}^n \ln x_i}.$$

A One-Parameter Lindley Distribution 2

The one-parameter Lindley distribution was introduced by a British statistician Dennis Victor Lindley in the context of Bayesian inference as a counter example of fiducial statistics (see Lindley[18]). This distribution is a mixture of $\exp(\theta)$ and gamma$(2, \theta)$.

In the recent years, researchers have given Lindley a special attention for its importance in modelling complex real life time and survival data. The detailed study about its mathematical properties, estimation of parameter and application showing the superiority of Lindley distribution over exponential distribution for the waiting times before service of the bank customers has been done by Ghitany et al. [9]. Other researchers have introduced more flexible generalization of Lindley by compounding with other well-known distributions(see for example Karuppusamy et al. [14], Hamed and Alzaghal [12] and the references therein.) In this chapter we study the properties and characteristics of the one-parameter Lindley distribution.

2.1 Definitions and Properties

In this section, we give the definitions and properties of one-parameter Lindley distribution.

Definition 2.1 (Density function) *Let $\theta > 0$, the Lindley distribution with parameter θ, denoted Lindley (θ), is defined by the density function (p.d.f):*

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

A random variable which has a Lindley distribution is noted by

$$X \sim Lindley(\theta).$$

Proposition 2.1 *The cumulative distribution function (c.d.f) of a random variable X that has Lindley distribution is given by ;*

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

Proof. Let $X \sim Lindley(\theta)$

$$F(x) = \int_{-\infty}^{x} f(t)dt \qquad = \int_{0}^{x} \frac{\theta^2}{\theta + 1}(1+t)e^{-\theta t}dt = \frac{\theta^2}{\theta + 1}\int_{0}^{x}(1+t)e^{-\theta t}dt$$

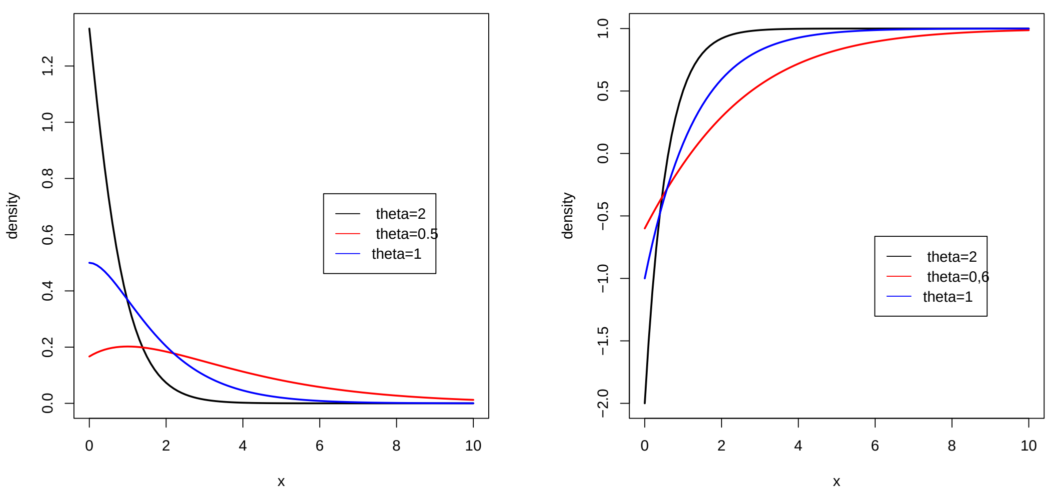
By applying integration by parts, we have

$$\begin{aligned} \int_{0}^{x}(1+t)e^{-\theta t}dt &= -\frac{1}{\theta}(1+t)e^{-\theta t}\Big]_{0}^{x} - \int_{0}^{x} -\frac{1}{\theta}e^{-\theta t}dt \\ &= -\frac{1}{\theta}(1+x)e^{-\theta x} + \frac{1}{\theta} - \frac{1}{\theta^2}(e^{-\theta x} - 1) \end{aligned}$$

Thus

$$\begin{aligned} F(x) &= \frac{\theta^2}{\theta + 1}\left[\theta + 1 - \theta(1+x)e^{-\theta x} - e^{-\theta x}\right] \\ &= 1 - \frac{\theta + 1 + \theta x}{\theta + 1}e^{-\theta x}, \quad x > 0, \; \theta > 0 \end{aligned}$$

□



(a) *Density function* (b) *Cumulative distribution function*

Figure 2.1 – *Probability density and cumulative density function of one parameter Lindley distribution for different values of θ*

2.1.1 Survival, hazard rate and mean residual life function

1. Survival function

We recall that the survival function of a random variable X is given by;

$$S(x) = P(X > x) = 1 - F(x)$$

Proposition 2.2 *Given that $X \sim Lindley(\theta)$. Its survival function is*

$$S(x) = \left(\frac{\theta + 1 + \theta x}{\theta + 1}\right) e^{-\theta x},\ \theta > 0$$

Proof.

$$\begin{aligned} S(x) &= P(X > x) = 1 - F(x) \\ &= 1 - \left(1 - \left(\frac{\theta + 1 + \theta x}{\theta + 1}\right) e^{-\theta x}\right) \\ &= \left(\frac{\theta + 1 + \theta x}{\theta + 1}\right) e^{-\theta x} \end{aligned}$$

□

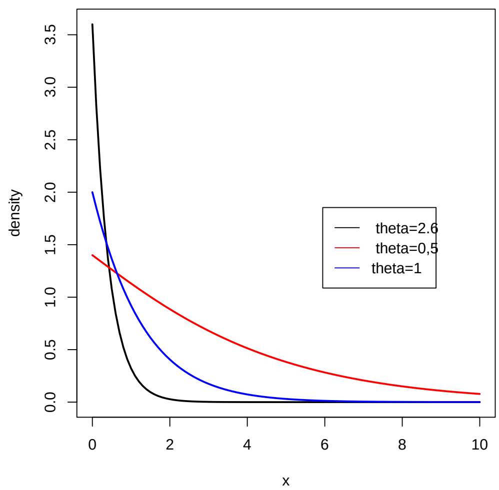


Figure 2.2 – *Graphical representation of survival function of a one-parameter Lindley distribution for different values of θ*

2. Hazard rate function

The hazard rate function of a one-parameter Lindley distribution is the function h defined by

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

Remark 2.1 $h(0) = f(0) = \dfrac{\theta^2}{\theta+1}$ *and since* $\dfrac{d}{dx}(h(x)) = \dfrac{\theta^2(1+x)}{\theta+1+\theta x}$, $h(x)$ *is an increasing function in* x *and* θ *and* $\dfrac{\theta^2}{\theta+1} < h(x) < \theta$

3. Mean residual life

For a continuous distribution with p.d.f $f(x)$ and cumulative distribution function $F(x)$, the mean residual life function is defined as

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

The mean residual life function of a one parameter Lindley distribution is given by

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

Remark 2.2 *For the exponential distribution of parameter* θ, $m(x) = \dfrac{1}{\theta}$ *and so the* $m(x)$ *of Lindley distribution shows its flexibility in fitting a wide range of data patterns over the exponential distribution.*

2.1.2 Mode

The mode of a one parameter Lindley distribution is given by;

$$Mode(X) = \begin{cases} \dfrac{1-\theta}{\theta} & \text{if } 0 < \theta < 1 \\ 0 & \text{otherwise} \end{cases}$$

Indeed, let $X \sim Lindley(\theta)$ with the probability density function f given by (2.1) that is,

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

The first derivative of f is defined

$$\frac{d}{dx}f(x) = \frac{\theta^2}{\theta+1}e^{-\theta x}(1-\theta-\theta x) = 0$$

$$\Rightarrow 1 - 1 - \theta - \theta x = 0$$

$$\Rightarrow x = \frac{1-\theta}{\theta}$$

$$\frac{d^2}{dx^2}f(x) = \frac{-\theta^3}{\theta+1}e^{-\theta x}(1-\theta-\theta x+1) < 0$$

This is because the exponential function is always positive and $1-\theta-\theta x+1$ at a point where $x = \frac{1-\theta}{\theta} = 1 > 0$ and $\theta > 0$ which means that $\frac{d^2}{dx^2}f(x) < 0$. Thus

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

2.1.3 Moment and its related measures

The r th moment of a one-parameter Lindley distribution.

Proposition 2.3 (**Ghitany et al. [9]**) *The r th moment about the origin of the Lindley distribution is given by the expression below.*

$$\mu'_r = \mathbb{E}\left(X^r\right) = \frac{r!(\theta+r+1)}{\theta^r(\theta+1)}, \quad r = 1, 2, \ldots \tag{2.2}$$

Proof. For $X \sim Lindley(\theta)$, $f(x) = \frac{\theta^2}{\theta+1}(1+x)e^{-\theta x}$ from the definition of mathematical expectation, we have

$$\begin{aligned}
\mathbb{E}\left(x^r\right) &= \int_0^{+\infty} x^r f(x)dx \\
&= \int_0^{+\infty} x^r \frac{\theta^2}{\theta+1}(1+x)e^{-\theta x}dx, \quad \theta > 0, \quad x > 0. \\
&= \frac{\theta^2}{\theta+1}\int_0^{+\infty} x^r(1+x)e^{-\theta x}dx \\
&= \frac{\theta^2}{\theta+1}\left[\int_0^{+\infty} x^r e^{-\theta x}dx + \int_0^{+\infty} x^{r+1}e^{-\theta x}\right]dx
\end{aligned}$$

Firstly, we have

$$\begin{aligned}\int_0^{+\infty} x^r e^{-\theta x} dx &= \frac{1}{\theta^{r+1}} \int_0^{+\infty} y^r e^{-y} dy \\ &= \frac{\Gamma(r+1)}{\theta^{r+1}} \\ &= \frac{r\Gamma(r)}{\theta^{r+1}}\end{aligned}$$

so,

$$\int_0^{+\infty} x^r e^{-\theta x} dx = \frac{r!}{\theta^{r+1}}$$

Secondly,

$$\begin{aligned}\int_0^{+\infty} x^{r+1} e^{-\theta x} dx &= \frac{1}{\theta^{r+2}} \int_0^{+\infty} y^{r+1} e^{-y} dy \\ &= \frac{\Gamma(r+2)}{\theta^{r+2}} \\ &= \frac{(r+1)\Gamma(r+1)}{\theta^{r+2}}\end{aligned}$$

so,

$$\int_0^{+\infty} x^{r+1} e^{-\theta x} dx = \frac{(r+1)!}{\theta^{r+2}}$$

$$\begin{aligned}\mathbb{E}(X^r) &= \frac{\theta^2}{\theta+1}\left[\frac{r!}{\theta^{r+1}} + \frac{(r+1)!}{\theta^{r+2}}\right] \\ &= \frac{\theta^2}{\theta+1}\left[\frac{\theta^{r+2} r! + \theta^{r+1}(r+1)r!}{\theta^{r+1}\theta^{r+2}}\right] \\ &= \frac{\theta^{r+3} r!(\theta+r+1)}{(\theta+1)\theta^{r+1}(\theta^{r+2})} \\ &= \frac{r!(\theta+r+1)}{\theta^r(\theta+1)}\end{aligned}$$

Then,

$$\forall r \in \mathbb{N}^\star, \quad \mathbb{E}(X^r) = \frac{r!(\theta+r+1)}{\theta^r(\theta+1)}$$

□

With this expression, we can be able to find all the moments of different orders of a one parameter Lindley distribution. Particularly, we have

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

The central moments of a one-parameter Lindley distribution are;

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

In particular, we have

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

The coefficient of variation, Skewness and Kurtosis are :

$$\gamma = \frac{\sqrt{\theta^2+4\theta+2}}{\theta+2}$$
$$\sqrt{\beta_1} = \frac{2(\theta^3+6\theta^2+6\theta+2)}{(\theta^2+4\theta+2)^{3/2}}$$
$$\beta_2 = \frac{3(3\theta^4+24\theta^3+44\theta^2+32\theta+8)}{(\theta^2+4\theta+2)^2}$$

2.1.4 The characteristic function.

The characteristic function $\phi_X(t)$ of the one parameter Lindley distribution is given in the following result.

Proposition 2.4 (Ghitany et al.[9]) *Let a random variable $X \sim Lindley(\theta)$. Its char-*

acteristic function is given by;

$$\phi_X(t) = \frac{\theta^2(\theta - it + 1)}{(\theta + 1)(\theta - it)^2},$$

where i is the imaginary number.

Proof. We recall that for any $t \in \mathbb{R}$ and $X \sim Lindley(\theta)$,

$$\begin{aligned}\phi_X(t) &= \mathbb{E}(e^{itX}) \\ &= \int_0^{+\infty} e^{itx} \frac{\theta^2}{\theta+1}(1+x)e^{-\theta x} dx\end{aligned}$$

$$\phi_X(t) = \frac{\theta^2}{\theta+1} \int_0^{+\infty} e^{itx}(1+x)e^{-\theta x} dx \tag{2.3}$$

Let $u(x) = 1 + x$ and $v'(x) = e^{(it-\theta)x}$. By applying integration by parts we have

$$\int_0^{+\infty} (1+x)e^{(it-\theta)x} = -\frac{1}{it-\theta} + \frac{1}{(it-\theta)^2}$$

Thus,

$$\begin{aligned}\phi_X(t) &= \frac{\theta^2}{\theta+1}\left(\frac{-(it-\theta)+1}{(it-\theta)^2}\right) \\ &= \frac{\theta^2(\theta - it + 1)}{(\theta+1)(\theta - it)^2}\end{aligned}$$

□

2.1.5 The moment generating function

Proposition 2.5 (**Ghitany et al.[9]**) *The moment generating function of a random variable $X \sim Lindley(\theta)$ is given by*

$$M_X(t) = \frac{\theta}{\theta - t}\left[1 + \frac{t}{(\theta+1)(\theta - t)}\right]$$

Proof.

$$\begin{aligned}M_X(t) &= \int_0^{+\infty} e^{tx} \frac{\theta^2}{\theta+1}(1+x)e^{-\theta x} dx \\ &= \frac{\theta^2}{\theta+1} \int_0^{+\infty} (1+x)e^{(t-\theta)x} dx\end{aligned}$$

Using integration by parts with $u = 1 + x$ and $v^{'} = e^{(t-\theta)x}$,

$$\int_0^{+\infty} (1+x)e^{(t-\theta)x}dx = \frac{1}{t-\theta}(1+x)e^{(t-\theta)x}]_0^{+\infty} - \int_0^{+\infty} \frac{1}{t-\theta}e^{(t-\theta)x}dx$$
$$= \frac{1}{\theta - t} + \frac{1}{(t-\theta)^2},$$

for all $t < \theta$

Thus,

$$\begin{aligned} M_X(t) &= \frac{\theta^2}{\theta+1}\left[\frac{1}{\theta-t} + \frac{1}{(t-\theta)^2}\right] \\ &= \frac{\theta}{\theta-t}\left[\frac{\theta^2(\theta-t)+\theta}{(\theta+1)(\theta-t)}\right] \\ &= \frac{\theta}{\theta-t}\left[\frac{\theta^2-\theta t+\theta}{(\theta+1)(\theta-t)}\right] \\ &= \frac{\theta}{\theta-t}\left[1+\frac{t}{(\theta+1)(\theta-t)}\right] \end{aligned}$$

□

2.1.6 Lorenz curve

A Lorenz curve is one of the measures of the variability of a statistical series. It is a cumulative percentage curve that was first introduced by Max Lorenz. These curves are generally used to measure the variability of the distribution of income and wealth. Hence Lorenz curve is the measure of the deviation of the actual distribution of a statistical series from the line of equal distribution.

Remark 2.3 *If the distance between the Lorenz curve from the line of equal distribution is more, it means that there is more inequality or variability in the series and vice-versa.*

Definition 2.2 *The Lorenz curve of a positive random variable say X is defined as the graph of the ratio*

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

For $X \sim Lindley(\theta)$,

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

Thus from the above definition,

$$L(F(x)) = \frac{2+\theta}{\theta(\theta+1)}\frac{\theta(\theta+1)}{\theta+2} - \frac{\theta(\theta+1)}{(\theta+2)(\theta+1)}e^{-\theta x}\left(\frac{2}{\theta}+1+2x+\theta x+\theta x^2\right)$$
$$= 1 - \frac{\theta}{\theta+2}e^{-\theta x}\left(\frac{2}{\theta}+1+2x+\theta x+\theta x^2\right)$$

2.1.7 Stochastic ordering

Stochastic orders have been used in different areas of probability and statistics.The simplest way of comparing two distribution functions is by comparing their the associated means. However, such a comparison is based on only two single numbers(the means), and therefore it is often not very informative. The most important and common stochastic orders that compare the magnitude of random variables are the stochastic order, hazard rate order, the mean residual life order and the likelihood ratio order.

We recall some basic definitions.

Let X and Y be two random variables. A random variable X is said to be smaller than a random variable Y in the

- stochastic order ($X \leq_{st} Y$) if $F_X(x) \geq F_Y(x)$ for all x.
- Hazard rate order ($X \leq_{hr} Y$) if $h_X(x) \geq h_Y(x)$ for all x.
- Mean residual life order ($X \leq_{mrl} Y$) if $m_X(x) \geq m_Y(x)$ for all x.
- Likelihood ratio order ($X \leq_{lr} Y$) if $(f_X(x))/(f_Y(x))$ decreases in x.

The following implications are well known.

$$\begin{array}{c} (X \leq_{lr} Y) \Rightarrow (X \leq_{hr} Y) \Rightarrow (X \leq_{mrl} Y) \\ \Downarrow \\ (X \leq_{st} Y) \end{array}$$

Theorem 2.1 (**Ghitany et al.[9]**) *Let $X \sim Lindley(\theta_1)$ and $Y \sim Lindley(\theta_2)$. If $\theta_1 > \theta_2$ then $X \leq_{lr} Y$ and hence $X \leq_{hr} Y, X \leq_{mrl} Y$ and $X \leq_{st} Y$.*

Proof. Note that

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

Finding the derivative of this ratio we get

$$\left(\frac{f_X}{f_Y}\right)'(x) = (\theta_2-\theta_1)\frac{\theta_1^2(\theta_2+1)}{\theta_2^2(\theta_1+1)}e^{(\theta_2-\theta_1)x}$$

$$= (\theta_2-\theta_1)\frac{f_X(x)}{f_Y(x)}.$$

Since $\theta_1 > \theta_2$, $\left(\frac{f_X}{f_Y}\right)'(x) \leq 0$ then $\left(\frac{f_X}{f_Y}\right)$ is decreasing in x. That is likelihood ratio order $(X \leq_{lr} Y)$. The remaining statements follow from the implication given in the above theorem.
This result also holds for the exponential distribution. □

2.1.8 Sums, products and ratios

For the distribution of sum, product and ratio of two independent random variables X and Y having Lindley distribution with parameters θ_1 and θ_2 respectively see Ghitany et al.[9]

2.1.9 The quantile function

Let X be a random variable which has Lindley distribution with (c.d.f) F. Its known that F is strictly increasing so the quantile function $Q_X = F^{-1}$ which expression depends on the Lambert W function.

Lambert W function

The Lambert W function has attracted a great deal of attention beginning with Lambert in 1758 and Euler in 1779. It is a multi-valued function complex function defined as the solution of the equation:

$$W(z)\exp(W(z)) = z,$$

where z is a complex number.

If z is a real number such that $z \geq \frac{-1}{e}$ then $W(z)$ becomes a real function and there are two possible real branches. The real branch taking on values in $]-\infty,-1]$ is the negative branch denoted by W_{-1} and the real branch taking on values in $[-1,\infty[$ is the principal branch denoted by W_0. This equation has two possible solutions if $z \in [-\frac{1}{e},0]$ and a unique solution if $z \geq 0$.

The negative branch has the following elementary properties: $W_{-1}(-\frac{1}{e}) = -1$, $W_{-1}(z)$ is decreasing as z increases to 0 and $W_{-1}(z) \to \infty$ as $z \to 0$.

Proposition 2.6 (Ghitany et al.[9]) *Let X be a one parameter Lindley random variable with cumulative distribution function F. The quantile function of X can be written as*

$$Q_X(u) = F^{-1}(u) = -1 - \frac{1}{\theta} - \frac{1}{\theta} W_{-1}\left[\frac{(\theta+1)(u-1)}{e^{(\theta+1)}}\right],$$

where $0 < u < 1$ and W_{-1} is the negative branch of Lambert W function.

Proof. We have

$$\begin{aligned} F_X(x) = u \Rightarrow 1 - \frac{\theta+1+\theta x}{\theta+1} e^{-\theta x} &= u \\ \Rightarrow -(\theta+1+\theta x)e^{-\theta x} &= (u-1)(\theta+1) \\ \Rightarrow -(\theta+1+\theta x)e^{-\theta x}e^{-(\theta+1)} &= e^{-(\theta+1)}(u-1)(\theta+1) \\ \Rightarrow -(\theta+1+\theta x)e^{-\theta x-\theta-1} &= e^{-(\theta+1)}(u-1)(\theta+1) \end{aligned}$$

By use of the equation $W(z)exp(W(z)) = z$, we conclude that $-(\theta+1+\theta x)$ is the value of W at $(u-1)(\theta+1)e^{-(\theta+1)}$ then

$$F_X(x) = u \Rightarrow -(\theta+1+\theta x) = W\left[\frac{(\theta+1)(u-1)}{e^{(\theta+1)}}\right], \; 0 < u < 1$$

since $\forall\theta > 0$, $\forall x > 0$, $\theta+\theta x+1 > 1$ then $-(\theta+1+\theta x) < -1$ this implies $-(\theta+1+\theta x) \in]-\infty,-1]$ and so,

$$F_X(x) = u \Rightarrow -(\theta+1+\theta x) = W_{-1}\left[\frac{(\theta+1)(u-1)}{e^{(\theta+1)}}\right]$$

which implies that , $\theta x = -\theta - 1 - W_{-1}\left[\frac{(\theta+1)(u-1)}{e^{(\theta+1)}}\right]$ then

$$x = F_X^{-1}(u) = Q_X(u) = -1 - \frac{1}{\theta} - \frac{1}{\theta} W_{-1}\left[\frac{\theta+1(u-1)}{e^{(\theta+1)}}\right]$$

□

2.1.10 Exponential family of distributions

An exponential family is a concept in statistics and probability theory. It is defined by probability density functions that can be expressed in the form:

$$f(x,\theta) = D(x)\exp\left[\eta(\theta).\, T(x) - B(\theta)\right]$$

Here the η and B are real-valued functions of the parameters and T is a real-valued statistics (see Lehmann and Casella[17]).

It can be easily seen that the one parameter Lindley distribution belongs to the exponential family by rewriting its probability density function as

$$f(x) = \frac{\theta^2}{1+\theta}(1+x)e^{-\theta x} = D(x)\exp\left[\eta(\theta).\, T(x) - B(\theta)\right]$$

where $\eta(\theta) = -\theta,\ \ T(x) = x,\ \ D(x) = 1 + x$ and $\ B(\theta) = \ln\left(\frac{1+\theta}{\theta^2}\right).$

The exponential family is so essential in different fields such as statistics and can used in Bayesian estimation while choosing the prior distribution.

2.2 Estimation of the parameter

In this section, we shall consider the estimation of the parameter θ of the one-parameter Lindley distribution by using method of moments, maximum likelihood and Bayesian approach.

2.2.1 Method of moments

Given a random sample $X_1, X_2, .., X_n$, of size n from the one parameter Lindley distribution, the estimator of the parameter θ by the method of moments is given by;

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

Since

$\bar{X} = \frac{\theta+2}{\theta(\theta+1)}$ which means that $\theta(\theta+1)\bar{X} = \theta + 2$. Thus

$$\bar{X}\theta^2 + (\bar{X}-1)\theta - 2 = 0$$

Thus

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

$\bar{X} > 0$. The following theorem shows that the estimator of θ is biased.

Theorem 2.2 (Ghitany et al.[9]) *The estimator $\hat{\theta}$ of the parameter θ is positively biased, i.e* $\mathbb{E}(\hat{\theta}) - \theta > 0$

Proof. Let $\hat{\theta} = g(\bar{X})$ where $g(t) = \dfrac{-(t-1)+\sqrt{(t-1)^2+8t}}{2t}$, $t > 0$.

Since

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

$g(t)$ is strictly convex. Thus by Jensen's inequality which states that for a convex function, the function's value of the expectation of a random variable X is less or equal to the expectation of the function's values at points specified by the random variable $(f(\mathbb{E}(X)) \leq \mathbb{E}(f(X)))$. That means we have $\mathbb{E}(g(\bar{X})) > g(E(\bar{X})$. Finally, since $g(\mathbb{E}(\bar{X})) = g(\mu)$ where $\mu = \dfrac{\theta+2}{\theta(\theta+1)}$,

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

thus we obtain $\mathbb{E}(\hat{\theta}) > \theta)$. □

2.2.2 Method of Maximum Likelihood

Given $X_i \sim Lindley(\theta)$, for $i = 1, 2, .., n$ random samples of size n, the log-likelihood function is given by;

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

The maximum likelihood estimator of the parameter θ is the solution of the equation given below

$$\frac{\partial \ell_n(\theta)}{\partial\theta} = \frac{2n}{\theta} - \bar{X} - \frac{n}{(\theta+1)} = 0 \quad (2.4)$$

By solving this equation we obtain

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

with $\frac{\partial^2 \ell_n(\theta)}{\partial\theta^2} = \frac{-2n}{\theta^2} - \sum_{i=1}^{n} \frac{x_i^2}{(\theta x_i + \theta)^2} < 0.$

Remark 2.4 *The method of moments and method of maximum likelihood estimators of the parameter θ are the same.*

Theorem 2.3 (Ghitany et al.[9]) *The estimator $\hat{\theta}$ of the parameter θ is consistent and asymptotically normal.*

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

The large sample $100(1-\alpha)\%$ confidence interval for θ is given by

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

where $z_{\alpha/2}$ is the $1-(\alpha/2)$ percentile of the standard normal distribution.

For the proof see Ghitany et al. [9]

2.2.3 Bayesian Estimation of Lindley distribution

In this section, we use Bayesian procedures to construct the estimators of the unknown parameters of Lindley distribution see (Ali et al. [1] and Metiri et al. [19]). We use Linex loss function to find the Bayes estimator which is defined as

$$L(\hat{\theta}, \theta) = exp(a(\hat{\theta} - \theta)) - a(\hat{\theta} - \theta) - 1, \; a \neq 0 \tag{2.5}$$

Under this function, Bayes estimator

$$\hat{\theta}_{Lin} = -\frac{1}{a} \ln[\mathbb{E}[exp(-a\theta)]] \tag{2.6}$$

when the expectation $\mathbb{E}[exp(-a\theta)]$ exists and is finite.

The posterior expectation of the Linex loss function ρ is got as follows

$$\rho = \mathbb{E}[L(\hat{\theta}, \theta)] \tag{2.7}$$
$$= \mathbb{E}[e^{a(\hat{\theta}-\theta)} - a(\hat{\theta} - \theta) - 1] \tag{2.8}$$
$$= e^{a\hat{\theta}}\mathbb{E}[e^{-a\theta}] - a[\hat{\theta} - \mathbb{E}(\theta)] - 1 \tag{2.9}$$

Posterior distribution using extension of Jeffrey's prior

This prior is assumed as non-informative prior for the parameter θ. It is given by

$$\pi(\theta) = k\frac{n^c}{\theta^{2c}},\ \theta,\ c > 0 \tag{2.10}$$

The posterior distribution for the parameter θ is

$$\pi(\theta|X) = \frac{L_n(\theta)\pi(\theta)}{\int_0^{+\infty} L_n(\theta)\pi(\theta)d\theta} = \frac{\frac{\theta^{2(n-c)}}{(\theta+1)^n}e^{-\theta\sum_{i=1}^n x_i}}{\int_0^{+\infty}\frac{\theta^{2(n-c)}}{(\theta+1)^n}e^{-\theta\sum_{i=1}^n x_i}d\theta} \tag{2.11}$$

Using the Linex loss function, the corresponding Bayes estimator of the parameter θ is

$$\hat{\theta}_{Lin} = -\frac{1}{a}\ln[\mathbb{E}[e^{-a\theta}]]$$

and

$$\begin{aligned}\mathbb{E}[e^{-a\theta}] &= \int_0^{+\infty} e^{-a\theta}\pi(\theta|X)d\theta \\ &= \frac{\int_0^{+\infty}\frac{\theta^{2(n-c)}}{(\theta+1)^n}e^{-\theta(a+\sum_{i=1}^n x_i)}d\theta}{\int_0^{+\infty}\frac{\theta^{2(n-c)}}{(\theta+1)^n}e^{-\theta\sum_{i=0}^n x_i}}d\theta\end{aligned}$$

Posterior distribution using gamma prior

Assuming that the prior information is gamma with parameters α, β,

$$\pi(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\beta\theta}$$

The posterior distribution using gamma prior is given by

$$\pi(\theta|X) = \frac{\frac{\theta^{2n+\alpha-1}}{(\theta+1)^n}e^{-(\beta+\sum_{i=1}^n x_i)\theta}\frac{\beta^\alpha}{\Gamma(\alpha)}\prod_{i=1}^n(1+x_i)}{\int_0^{+\infty}\frac{\theta^{2n+\alpha-1}}{(\theta+1)^n}e^{-(\beta+\sum_{i=1}^n x_i)\theta}\frac{\beta^\alpha}{\Gamma(\alpha)}\prod_{i=1}^n(1+x_i)} \tag{2.12}$$

The corresponding Bayes estimator of the parameter θ under Linex loss function is given by

$$\hat{\theta}_{Lin} = -\frac{1}{a}\ln[\mathbb{E}[e^{-a\theta}]]$$

where

$$\mathbb{E}[e^{-a\theta}] = \int_0^{+\infty} e^{-a\theta}\pi(\theta|X)d\theta = \frac{\int_0^{+\infty}\frac{\theta^{2n+\alpha-1}}{(\theta+1)^n}e^{-(a+\beta+\sum_{i=1}^n x_i)\theta}\frac{\beta^\alpha}{\Gamma(\alpha)}\prod_{i=1}^n(1+x_i)d\theta}{\int_0^{+\infty}\frac{\theta^{2n+\alpha-1}}{(\theta+1)^n}e^{-(\beta+\sum_{i=1}^n x_i)\theta}\frac{\beta^\alpha}{\Gamma(\alpha)}\prod_{i=1}^n(1+x_i)d\theta}$$

Posterior distribution using inverted gamma prior

The inverted gamma prior represents the reciprocal of a variable distributed according to gamma distribution. It is observed that if X has an inverted gamma (α, β) distribution then $\frac{1}{X}$ has a gamma (α, β) distribution. The inverted gamma distribution is given by

$$\pi(\theta) = \frac{\alpha^{\beta}}{\Gamma(\beta)} \frac{1}{\theta^{\beta+1}} e^{-(\frac{\alpha}{\theta})}, \ \alpha > 0, \ \beta > 0, \ \theta > 0 \tag{2.13}$$

The posterior distribution of the parameter θ using the inverted gamma prior is given by;

$$\pi(\theta|X) = \frac{\frac{\theta^{2n-\beta-1}}{(\theta+1)^n} e^{-\frac{\alpha}{\theta} - \theta \sum_{i=1}^{n} x_i}}{\int_0^{+\infty} \frac{\theta^{2n-\beta-1}}{(\theta+1)^n} e^{-\frac{\alpha}{\theta} - \theta \sum_{i=1}^{n} x_i} d\theta}$$

Bayes estimator of the parameter θ under Linex loss function is

$$\hat{\theta}_{Lin} = -\frac{1}{a} \ln[\mathbb{E}(e^{-a\theta})]$$

where

$$\mathbb{E}[e^{-a\theta}] = \frac{\int_0^{+\infty} \frac{\theta^{2n-\beta-1}}{(\theta+1)^n} e^{-\frac{\alpha}{\theta} - \theta(a + \sum_{i=1}^{n} x_i)} d\theta}{\int_0^{+\infty} \frac{\theta^{2n-\beta-1}}{(\theta+1)^n} e^{-\frac{\alpha}{\theta} - \theta \sum_{i=1}^{n} x_i} d\theta}$$

It is observed that of all these prior distributions assumed, there is no simplest form of all these integrals in order to find Bayes estimator. However, there are various methods suggested to approximate the ratio of integrals of which among these methods Lindley(1980) is the simplest.

Lindley approximation

Lindley approximation approaches the ratio of the integrals as a whole and produces a single numerical result.

If n is sufficiently large, according to Lindley (1980), any ratio of the integral of the form

$$I(x) = \mathbb{E}[\mu(\theta)] = \frac{\int_{\theta} \mu(\theta) exp[\ell_n(\theta) + g(\theta)] d\theta}{\int_{\theta} exp[\ell_n(\theta) + g(\theta)] d\theta} \tag{2.14}$$

where $\mu(\theta)$ is a function of θ only, $\ell_n(\theta)$ is log of likelihood function of Lindley distribution and $g(\theta)$ is log of prior of θ

$$I(x) = \mu(\hat{\theta}) + 0.5[(\hat{\mu}_{\theta\theta} + 2\hat{\mu}_{\theta}\hat{p}_{\theta})\hat{\sigma}_{\theta\theta}] + 0.5[(\hat{\mu}_{\theta}\hat{\sigma}_{\theta\theta})(\hat{\ell}_{n\theta\theta\theta}\hat{\sigma}_{\theta\theta})]$$

where
$\hat{\theta}$ is the maximum likelihood estimator of the parameter θ. we have

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

$\hat{\mu}_{\theta} = \dfrac{\partial\mu(\hat{\theta})}{\partial\hat{\theta}}$, $\hat{\mu}_{\theta\theta} = \dfrac{\partial^2\mu(\hat{\theta})}{\partial\hat{\theta}^2}$, $\hat{p}_{\theta} = \dfrac{\partial g(\hat{\theta})}{\partial\hat{\theta}}$, $\hat{\ell}_{n\theta\theta} = \dfrac{\partial^2\ell_n(\hat{\theta})}{\partial\hat{\theta}^2}$, $\hat{\sigma}_{\theta\theta} = -\dfrac{1}{\hat{\ell}_{n\theta\theta}}$, and $\hat{\ell}_{n\theta\theta\theta} = \dfrac{\partial^3\ell(\hat{\theta})}{\partial\hat{\theta}^3}$. For the Bayes estimates of Lindley distribution,

$$I(x) = \mathbb{E}[\mu(\theta)] = \frac{\int_0^{+\infty}\mu(\theta)exp[\ell_n(\theta) + g(\theta)]d\theta}{\int_0^{+\infty}exp[\ell_n(\theta) + g(\theta)]d\theta}.$$

Estimation of the Parameter

Case of the Extension of Jeffrey's prior: Using the extension of Jeffreys prior and following the steps explained above we have $\mu(\theta) = e^{-a\theta}$, $\ell_n(\theta) = 2n\ln\theta - n\ln(1+\theta) - \theta\sum_{i=1}^{n}x_i + \sum_{i=1}^{n}\ln(1+x_i)$, $g(\theta) = c\ln(n) - 2c\ln\theta$, $\hat{\mu}_{\theta} = -ae^{-a\hat{\theta}}$, $\hat{\mu}_{\theta\theta} = -a^2e^{-a\hat{\theta}}$, $\hat{p}_{\theta} = -\dfrac{2c}{\theta}$, $\hat{\ell}_{n\theta\theta} = \dfrac{-2n}{\theta^2} + \dfrac{n}{(1+\theta)^2}$, $\hat{\sigma_{\theta\theta}} = \dfrac{\theta^2(1+\theta)^2}{2n(1+\theta)^2 - n\theta^2}$, and $\hat{\ell}_{n\theta\theta\theta} = \dfrac{4n}{\theta^3} - \dfrac{2n}{(1+\theta)^3}$ Thus

$$\begin{aligned}\mathbb{E}[\exp(-a\theta)|X] = e^{-a\hat{\theta}} &+ 0.5\left[ae^{-a\hat{\theta}}\left(a + \frac{4c}{\hat{\theta}}\right)\frac{\hat{\theta}^2\left(1+\hat{\theta}\right)^2}{2n(1+\hat{\theta})^2 - n\hat{\theta}^2}\right] \\ &- 0.5\left[ae^{-a\hat{\theta}}\left(\frac{4n}{\hat{\theta}^3} - \frac{2n}{(1+\hat{\theta})^3}\right)\left(\frac{\hat{\theta}^2(1+\hat{\theta})^2}{2n(1+\hat{\theta})^2 - n\hat{\theta}^2}\right)^2\right]\end{aligned}$$

Thus Bayes estimator under Linex function and using extension of Jeffreys prior is

$$\begin{aligned}\hat{\theta}_{Lin} = -\frac{1}{a}\ln&\left[e^{-a\hat{\theta}} + 0.5\left[ae^{-a\hat{\theta}}\left(a + \frac{4c}{\hat{\theta}}\right)\frac{\hat{\theta}^2(1+\hat{\theta})^2}{2n(1+\hat{\theta})^2 - n\hat{\theta}^2}\right]\right. \\ &\left.-0.5\left[ae^{-a\hat{\theta}}\left(\frac{4n}{\hat{\theta}^3} - \frac{2n}{(1+\hat{\theta})^3}\right)\left(\frac{\hat{\theta}^2(1+\hat{\theta})^2}{2n(1+\hat{\theta})^2 - n\hat{\theta}^2}\right)^2\right]\right] \qquad (2.15)\end{aligned}$$

Case of the gamma prior : Using the same steps mentioned above, for the gamma prior we find

$\mu(\theta) = e^{-a\theta}$, $\quad \ell_n(\theta) = 2n\ln\theta - n\ln(1+\theta) - \theta\sum_{i=1}^{n} x_i + \sum_{i=1}^{n}\ln(1+x_i)$,
$g(\theta) = a\ln\beta - \ln\Gamma(a) + (\alpha-1)\ln\theta - \beta\theta$, $\hat{\mu}_\theta = -ae^{-a\hat{\theta}}$, $\hat{\mu}_{\theta\theta} = -a^2e^{-a\hat{\theta}}$,
$\hat{p}_\theta = \frac{\alpha-1}{\hat{\theta}} - \beta$, $\hat{\ell}_{n\theta\theta} = \frac{-2n}{\hat{\theta}^2} + \frac{n}{(1+\hat{\theta})^2}$, $\sigma_{\hat{\theta}\theta} = -\frac{\hat{\theta}^2(1+\hat{\theta})^2}{n\hat{\theta}^2 - 2n(1+\hat{\theta})^2}$ and
$\hat{\ell}_{n\theta\theta\theta} = \frac{4n}{\hat{\theta}^3} - \frac{2n}{(1+\hat{\theta})^3}$

$$\mathbb{E}[\exp(-a\theta)|x] = e^{-a\hat{\theta}} + 0.5\left[\frac{2na\hat{\theta}^2e^{-a\hat{\theta}}(1+\hat{\theta})^4(-2(1+\hat{\theta})^3+\hat{\theta}^5)}{n^2\hat{\theta}^3(1+\hat{\theta})^3(\hat{\theta}^4-2\hat{\theta}^2(1+\hat{\theta})^2+4(1+\hat{\theta}^4))}\right] + 0.5\left[\frac{e^{-a\hat{\theta}}(1+\hat{\theta})^2(a^2\hat{\theta}^3+2a\hat{\theta}^3(\alpha-1))+\hat{\theta}^3\beta(1+\hat{\theta})^2}{n\hat{\theta}}(\hat{\theta}^2-2(1+\hat{\theta})^2)\right] \quad (2.16)$$

Substituting this equation in equation (2.6), we can be able to find Bayes estimator using gamma prior.

Case of the Extension of inverted gamma prior: The same way, Bayes estimator using inverse gamma prior

$$\hat{\theta}_{Lin} = -\frac{1}{a}\ln\left[e^{-a\hat{\theta}} + 0.5\left[a(a-2)e^{-a\hat{\theta}}\left(\frac{\alpha}{\hat{\theta}^2} - \frac{\beta+1}{\theta}\right)\frac{\hat{\theta}^2(1+\hat{\theta})^2}{2n(1+\hat{\theta})^2 - n\hat{\theta}^2}\right] - 0.5\left[ae^{-a\hat{\theta}}\left(\frac{4n}{\hat{\theta}^3} - \frac{2n}{(1+\hat{\theta})^3}\right)\left(\frac{\hat{\theta}^2(1+\hat{\theta})^2}{2n(1+\hat{\theta})^2 - n\hat{\theta}^2}\right)^2\right]\right] \quad (2.17)$$

2.2.4 Simulation

In this section, we embark on a journey to explore the simulation of a one parameter Lindley distribution. By simulating data from this distribution, we aim to gain insights that can enhance its practical ability and inform decision- making processes. Throughout this exploration, we will use simulation techniques to generate data sets that follow the one parameter Lindley distribution. There exist different methods for generating random data from different probability distributions such as the inversion method, rejection method, Monte Carlo Markov chain method (mcmc) such as metropolis-Hastings and Gibbs sampling among others (see Ross [24] or Ortgiese et al.[21]). Since $F(x) = u$, where u is an observation from the uniform distribution $(0,1)$, can not be solved explicitly in x, the inversion method for generating random data from the one parameter Lindley distribution fails. However, we can use the fact that the Lindley distribution is a special mixture of Exponential (θ) and gamma $(2,\theta)$ distributions:

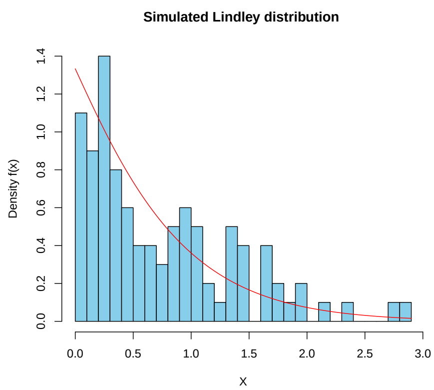
$$f(x) = pf_1(x) + (1-p)f_2(x), \quad x > 0, \quad \theta > 0,$$

where $p = \dfrac{\theta}{\theta+1}$, $f_1(x) = \theta e^{-\theta x}$ and $f_2(x) = \theta^2 x e^{-\theta x}$.

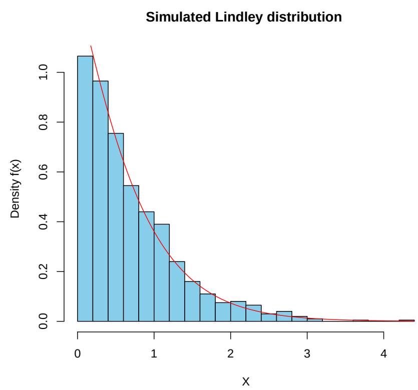
To generate random data $X_i, i = 1, ..., n$ from the one parameter Lindley distribution with parameter θ we have the following algorithm:

1. Generate $u_i \sim uniform(0,1),\ i = 1, ..., n$.
2. Generate $v_i \sim exponential(\theta), i = 1, ..., n$.
3. Generate $w_i = gamma(2,\theta), i = 1, .., n$.
4. if $u_i \leq p = \dfrac{\theta}{\theta+1}$, then set $X_i = w_i, i = 1, ..., n$.

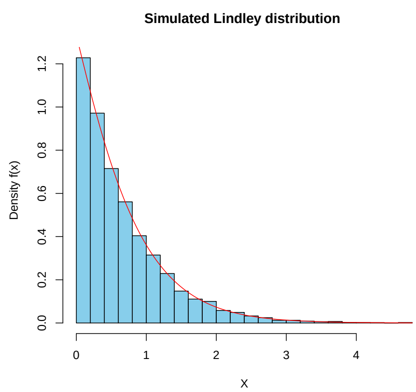
For the program of this simulation see the annex.
We can also be able to generate random data from the one parameter Lindley distribution using metropolis-Hastings while using $\exp(\theta)$ as the proposal distribution.



(a)

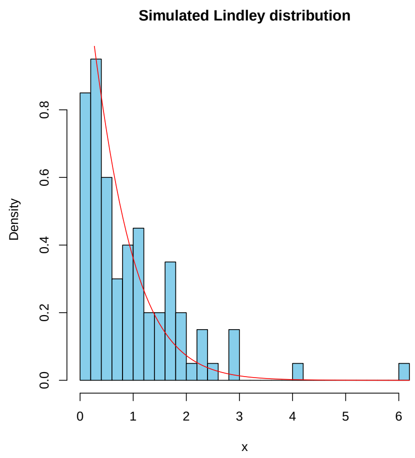


(b)

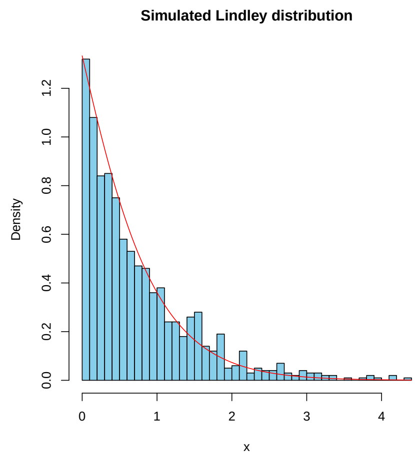


(c)

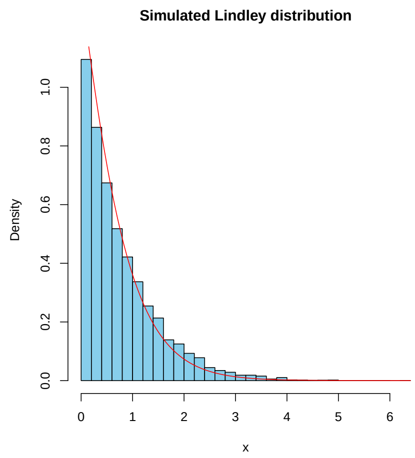
Figure 2.3 – *Histograms of simulated values of one parameter Lindley distribution for different values of n using the fact that Lindley is a mixture of Exponential and gamma distributions for* $\theta = 2.0$: *(a) n=100, (b) n=1000 and (c) n=10000*



(a)



(b)



(c)

Figure 2.4 – *Histograms of simulated values of one parameter Lindley distribution for different values of n using Metropolis-Hastings algorithm for* $\theta = 2.0$ *: (a) n=100, (b) n=1000 and (c) n=10000*

To test the performance of this simulation method, we generate datasets of sizes 40, 100, and 500 of Lindley distribution with the parameter $\theta = 2$, we will fit these data to the Lindley distribution using maximum likelihood estimation (MLE) to estimate the parameter. To verify the quality of this fit, we will use the Kolmogorov-Smirnov (K-S) test. The p-value from the K.S test will also be calculated to assess the statistical significance of the fit. The results are reported in the table bellow.

n	Estimate of Parameter θ	$K-S$	p-value
40	2.0438	0.087195	0.8954
100	1.98846	0.071565	0.685
500	1.990628	0.03106	0.7202

Across all sample sizes, the estimated θ values are very close to the true parameter value 2, indicating that the maximum likelihood estimation (MLE) method accurately captures the parameter of the Lindley distribution. The K-S test statistics are consistently low, and the p-values are all well above the typical significance level (e.g., 0.05), suggesting that we consider the null hypothesis that the data follow the Lindley distribution.

These results collectively demonstrate that the Lindley distribution provides an excellent fit to the simulated data across various sample sizes. The high p-values indicate that the observed differences between the empirical distribution and the fitted Lindley distribution are not statistically significant, affirming the robustness and reliability of the MLE method for estimating the parameter of the Lindley distribution.

The Two-Parameter Lindley Distributions 3

3.1 A two-parameter Lindley distribution

The one-parameter Lindley distribution does not provide enough flexibility for analyzing different types of lifetime data because of having only one parameter. To increase the flexibility for modelling purposes it will be useful to consider further alternatives of this distribution. In this section, we are interested in a two-parameter Lindley distribution studied in Shanker et al. [27], of which the Lindley distribution (2.1) is a particular case. For other two-parameter Lindley distribution see Shanker et al. [26].

A two-parameter Lindley distribution with parameters α and θ is defined by its probability density function(p.d.f)

$$f(x) = \frac{\theta^2}{\theta+\alpha}(1+\alpha x)e^{-\theta x}, \quad x > 0,\ \theta > 0,\ \alpha > -\theta \tag{3.1}$$

At $\alpha = 1,$ the two-parameter distribution reduces to one parameter distribution and at $\alpha = 0$, it reduces to an exponential distribution with parameter θ. This probability density function of two-parameter Lindley distribution can be shown as a mixture of exp (θ) and gamma $(2,\ \theta)$ distributions as follows:

$$f(x) = pf_1(x) + (1-p)f_2(x) \tag{3.2}$$

where $p = \dfrac{\theta^2}{\theta + \alpha}$, $f_1(x) = \theta e^{-\theta x}$ and $f_2(x) = \theta^2 x e^{-\theta x}$.
The first derivative of p.d.f is obtained as $f^{'}(x) = \dfrac{\theta^2}{\theta + \alpha}(\alpha - \theta - \alpha\theta x)e^{-\theta x}$ and so $f^{'}(x) = 0$ gives $x = \dfrac{\alpha - \theta}{\alpha\theta}$. From this it follows that for $\theta < \alpha$, $x_0 = \dfrac{\alpha - \theta}{\alpha\theta}$ is the unique critical point at which $f(x)$ is maximum. Also for $\theta \geq \alpha, f^{'}(x; \alpha, \theta) \leq 0$ i.e $f(x)$ is decreasing in x. Therefore, the mode of the distribution is given by

$$Mode(X) = \begin{cases} \dfrac{\alpha - \theta}{\alpha\theta} & \text{if } \theta < \alpha \\ 0 & \text{otherwise} \end{cases}$$

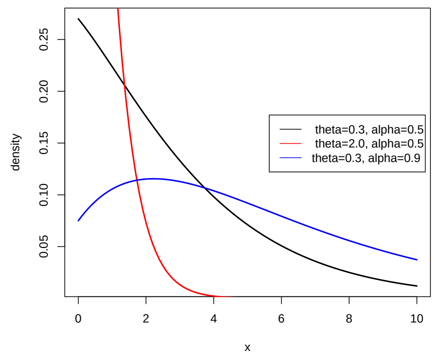


Figure 3.1 – *Probability density function of a two-parameter Lindley distribution for different values of θ and α*

The cumulative distribution function (c.d.f) of the two-parameter Lindley distribution is given by

$$F(x) = 1 - \frac{\theta + \alpha + \alpha\theta x}{\theta + \alpha} e^{-\theta x}, \quad x > 0, \ \theta > 0, \ \alpha > -\theta \tag{3.3}$$

3.1.1 Survival, failure rate and mean residual life functions

The survival function of the distribution can be expressed as

$$S(x) = \frac{\theta + \alpha + \alpha\theta x}{\theta + \alpha} e^{-\theta x}, \quad x > 0, \ \theta > 0, \ \alpha > -\theta \tag{3.4}$$

The corresponding failure rate (hazard) function $h(x)$ and the mean residual life function $m(x)$ are given by

$$h(x) = \frac{\theta^2(1+\alpha x)}{\theta+\alpha+\alpha\theta x}, \quad x > 0,\ \theta > 0,\ \alpha > -\theta \tag{3.5}$$

$$m(x) = \frac{\theta+2\alpha+\theta\alpha x}{\theta(\theta+\alpha+\theta\alpha x)} \tag{3.6}$$

It is obvious that $h(x)$ is an increasing function of x, α and θ, whereas $m(x)$ is a decreasing function of x, α and increasing function of θ. The failure rate function and mean residual life function of the two parameter Lindley distribution show its flexibility over one parameter Lindley distribution.

3.1.2 Statistical properties

1. Moments

The r th moment $\mu_r^{'}$ of a two parameter Lindley distribution can be obtained as

$$\mu_r^{'} = \mathbb{E}(X^r) = \frac{\Gamma(r+1)(\theta+\alpha+\alpha r)}{\theta^r(\theta+\alpha)}, r = 1, 2, ... \tag{3.7}$$

Indeed, Taking $r = 1, 2, 3,$ and 4 in (please write here the number of the equation of the expression of mean given above), the first four moments are obtained as

$$\mu_1^{'} = \frac{\theta+2\alpha}{\theta(\theta+\alpha)}, \quad \mu_2^{'} = \frac{2(\theta+3\alpha)}{\theta^2(\theta+\alpha)}, \quad \mu_3^{'} = \frac{6(\theta+4\alpha)}{\theta^3(\theta+\alpha)}, \quad \mu_4^{'} = \frac{24(\theta+5\alpha)}{\theta^4(\theta+\alpha)} \tag{3.8}$$

We verify easily that for $\alpha = 1$, the moments about the origin of the two parameter Lindley distribution reduces to the moments of the one parameter Lindley distribution. The central moments of this distribution are

$$\mu_2 = \frac{\theta^2+4\theta\alpha+2\alpha^2}{\theta^2(\theta+\alpha)^2}, \quad \mu_3 = \frac{2(\theta^3+6\theta^2\alpha+6\theta\alpha^2+2\alpha^3)}{\theta^3(\theta+\alpha)^3},$$

and

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

It can be easily verified that for $\alpha = 1$, the central moments of the two-parameter Lindley distribution reduces to the central moments of the one-parameter Lindley distribution. Thus the coefficient of variation , skewness and kurtosis noted as γ, $\sqrt{\beta_1}$ and β_2 respectively of the two-parameter Lindley distribution are given by

$$\gamma = \frac{\sqrt{\theta^2 + 4\theta\alpha + 2\alpha^2}}{\theta + 2\alpha} \tag{3.9}$$

$$\sqrt{\beta_1} = \frac{2(\theta^3 + 6\theta^2\alpha + 6\theta\alpha^2 + 2\alpha^3)}{(\theta^2 + 4\theta\alpha + 2\alpha^2)^{3/2}} \tag{3.10}$$

$$\beta_2 = \frac{3(3\theta^4 + 24\theta^3\alpha + 24\theta^2\alpha + 20\theta^2\alpha^2 - 88\theta\alpha^3 + 120\theta\alpha^2 + 8\alpha^4)}{(\theta^2 + 4\theta\alpha + 2\alpha^2)} \tag{3.11}$$

The two-parameter Lindley distribution is positively skewed since its mean is greater than its mode.

2. Stochastic ordering

Theorem 3.1 (Shanker et al.[27]) *Let $X \sim Lindley(\alpha_1, \theta_1)$ and $Y \sim Lindley(\alpha_2, \theta_2)$. If $\alpha_1 = \alpha_2$ and $\theta_1 \geq \theta_2$ (or if $\theta_1 = \theta_2$ and $\alpha_1 \geq \alpha_2$), then $X \leq_{lr} Y$ and hence $X \leq_{hr} Y, X \leq_{mrl} Y$ and $X \leq_{st} Y$.*

Proof.

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

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

Thus

$$\frac{\partial}{\partial x}\ln\frac{f_X(x)}{f_Y(y)} = \frac{1}{1+\alpha_1(x)} - \frac{1}{1+\alpha_2 x} + (\theta_2 - \theta_1)$$

$$\frac{\partial}{\partial x}\ln\frac{f_X(x)}{f_Y(y)} = \frac{(\alpha_2-\alpha_1)x}{(1+\alpha_1 x)(1+\alpha_2 x)} + (\theta_2-\theta_1)$$

Case(i): If $\alpha_1 = \alpha_2$ and $\theta_1 \geq \theta_2$, then

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

which means that $X \leq_{lr} Y$ and hence $X \leq_{hr} Y, X \leq_{mrl} Y$ and $X \leq_{st} Y$.

case(ii): If $\theta_1 = \theta_2$ and $\alpha_1 \geq \alpha_2$, then

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

which means that $X \leq_{lr} Y$ and hence $X \leq_{hr} Y, X \leq_{mrl} Y$ and $X \leq_{st} Y$.

□

3.1.3 Order Statistics

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from two-parameter Lindley distribution 3.1. Let $X_{(1)} < X_{(2)} < \ldots < X_{(n)}$ denote the corresponding order statistics. The p.d.f. and the c.d.f. of the k th order statistic, say $Y = X_{(k)}$ are given by

$$\begin{aligned} f_Y(y) &= \frac{n!}{(k-1)!(n-k)!}F^{k-1}(y)\{1-F(y)\}^{n-k}f(y) \\ &= \frac{n!}{(k-1)!(n-k)!}\sum_{l=0}^{n-k}\binom{n-k}{l}(-1)^l F^{k+l-1}(y)f(y) \end{aligned}$$

and

$$\begin{aligned} F_Y(y) &= \sum_{j=k}^{n}\binom{n}{j}F^j(y)\{1-F(y)\}^{n-j} \\ &= \sum_{j=k}^{n}\sum_{l=0}^{n-j}\binom{n}{j}\binom{n-j}{l}(-1)^l F^{j+l}(y), \end{aligned}$$

respectively, for $k = 1, 2, 3, \ldots, n$. Thus, the p.d.f. and the c.d.f of the k th order statistics of the distribution of density (3.1) are obtained as

$$f_Y(y) = \frac{n!\theta^2(1+\alpha x)e^{-\theta x}}{(\alpha+\theta)(k-1)!(n-k)!}\sum_{l=0}^{n-k}\binom{n-k}{l}(-1)^2 \times \left[1 - \frac{\alpha+\theta+\alpha\theta x}{\alpha+\theta}e^{-\theta_x}\right]^{k+l-1}$$

and

$$F_Y(y) = \sum_{j=k}^{n}\sum_{l=0}^{n-j}\binom{n}{j}\binom{n-j}{l}(-1)^l\left[1 - \frac{\alpha+\theta+\alpha\theta x}{\alpha+\theta}e^{-\theta x}\right]^{j+l}$$

3.1.4 Estimation

1. Maximum likelihood estimates

Let $x_1, x_2, ..., x_n$ be observations from a random sample X of size n from a two-parameter Lindley distribution. Its likelihood function is given by

$$L_n(\alpha,\theta) = \left(\frac{\theta^2}{\theta+\alpha}\right)^n e^{-n\theta\bar{X}}\prod_{i=1}^{n}(1+\alpha x_i)$$

Taking natural logarithm of likelihood function we get

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

Finding the derivative of the log likelihood function with respect to θ and then with respect to α we get these two equations

$$\frac{\partial\ell_n(\alpha,\theta)}{\partial\theta} = \frac{2n}{\theta} - \frac{n}{\theta+\alpha} - n\bar{X} = 0$$

and

$$\frac{\partial\ell_n(\alpha,\theta)}{\partial\alpha} = -\frac{n}{\theta+\alpha} + \sum_{i=1}^{k}\frac{x_i}{1+\alpha x_i} = 0$$

From the derivative of log likelihood function with respect to θ, it is clear that

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

which is the mean of the two-parameter Lindley distribution.These two equations do not seem to be solved directly thus opting for fisher's scoring method that iteratively updates the parameter estimates based on the first and second derivatives of the log-likelihood function until convergence is reached.
We have,

$$\frac{\partial^2 \ell_n(\alpha,\theta)}{\partial\theta^2} = -\frac{2n}{\theta^2} + \frac{n}{(\theta+\alpha)^2},$$
$$\frac{\partial^2 \ell_n(\alpha,\theta)}{\partial\theta\partial\alpha} = \frac{\partial^2 \ell_n(\alpha,\theta)}{\partial\alpha\partial\theta} = \frac{n}{(\theta+\alpha)^2}, \quad \frac{\partial^2 \ell_n(\alpha,\theta)}{\partial\alpha^2} = \frac{n}{(\theta+\alpha)^2} - \sum_{i=1}^{n} \frac{x_i^2}{(1+\alpha x_i)^2}.$$

The following equations for $\hat{\theta}$ and $\hat{\alpha}$ can be solved.

$$\begin{bmatrix} \frac{\partial^2 \ell_n(\alpha,\theta)}{\partial\theta^2} & \frac{\partial^2 \ell_n(\alpha,\theta)}{\partial\theta\partial\alpha} \\ \frac{\partial^2 \ell_n(\alpha,\theta)}{\partial\theta\partial\alpha} & \frac{\partial^2 \ell_n(\alpha,\theta)}{\partial\alpha^2} \end{bmatrix}_{\substack{\hat{\theta}=\theta_0 \\ \hat{\alpha}=\alpha_0}} \begin{bmatrix} \hat{\theta}-\theta_0 \\ \hat{\alpha}-\alpha_0 \end{bmatrix} = \begin{bmatrix} \frac{\partial \ell_n(\alpha,\theta)}{\partial\theta} \\ \frac{\partial \ell_n(\alpha,\theta)}{\partial\alpha} \end{bmatrix}_{\substack{\hat{\theta}=\theta_0 \\ \hat{\alpha}=\alpha_0}} \tag{3.12}$$

where θ_0 and α_0 are the initial values of θ and α respectively. These equations can be solved iteratively till close estimates of $\hat{\theta}$ and $\hat{\alpha}$ are obtained.

2. Estimates from moments

Using the first two moments about the origin of the two parameter Lindley distribution, we have

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

taking $\theta = b\alpha$, we get

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

This gives

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

which is the quadratic equation in b. Replacing the first and second moments by their respective sample moments,an estimate of k can be obtained. Again, substituting $\theta = b\alpha$ in the expression for mean of the two parameter Lindley distribution,we get

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

thus an estimate of α is given by

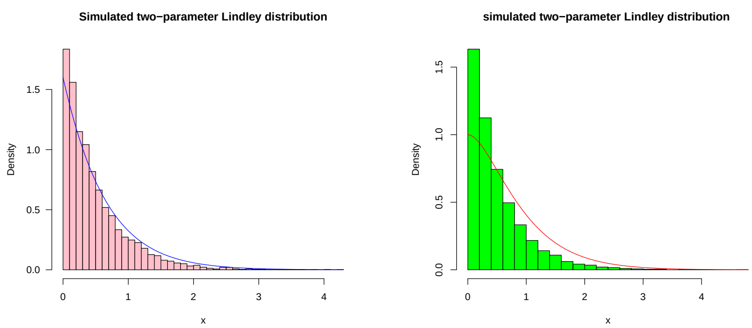
$$\hat{\alpha} = \left(\frac{b+2}{b(b+1)}\right)\frac{1}{\bar{X}}$$

and finally, an estimate of θ is obtained as

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

3.1.5 Simulation of two-parameter Lindley distribution

In this section, we simulate the two-parameter Lindley distribution for $N = 10000$ using Metropolis-Hasting Algorithm and the fact that the density of this distribution is a mixture of exp (θ) and gamma $(2,\ \theta)$ for $\theta = 2$



(a) *Metropolis Hastings method* (b) *Using the sum method*

Figure 3.2 – *Histograms showing two-parameter Lindley distribution: (a)*$N = 10000,\ \theta = 2$ *and (b)* $N = 10000,\ \theta = 2, \alpha = 2$

3.2 A two-parameter weighted Lindley distribution

The one parameter Lindley distribution has been generalized by several authors to increase its flexibility in the survival analysis data . One of these generalizations is the two-parameter weighted Lindley distribution which was proposed by Ghitany et al. [10] by considering a normalising constant and a suitable weighted function.

When an investigator records an observation by nature according to certain stochastic model, the recorded observation will not have the original distribution unless every observation is given an equal chance of being recorded. For example, suppose that the original observation x_0 comes from a distribution with probability density function(p.d.f) $f_0(x_0;\theta_1)$, where θ_1 is the parameter vector, and that observation is re-weighted by a weighted function $w(x;\ \theta_2) > 0, \theta_2$ is a new parameter vector, then x comes from a distribution with (p.d.f)

$$f(x) = Aw(x;\ \theta_2)f_0(x;\theta_1)$$

where A is a normalizing constant. Distributions of this type are called Weighted distributions. The study of weighted distributions is useful as it provides methods of extending distributions for added flexibility in fitting data. In this section we are interested in the origin two-parameter weighted Lindley studied by Ghitany et al. [10], for other models, see Asgharzadeha et al. [3].

We consider a two-parameter Lindley distribution with p.d.f

$$f(x) = Bx^{\alpha-1}f_0(x;\ \theta), x > 0,\ \alpha > 0,\ \theta > 0$$

where $B = \dfrac{\theta^{\alpha-1}}{(\theta+\alpha)\Gamma(\alpha)}(\theta+1)$ is a normalizing constant and

$$f_0(x;\theta) = \frac{\theta^2}{\theta+1}(1+x)e^{-\theta x},\ x > 0,\ \theta > 0.$$

Thus the probability density function of a two-parameter weighted Lindley distribution is given by

$$f(x) = \frac{\theta^{\alpha+1}}{(\theta+\alpha)\Gamma(\alpha)} x^{\alpha-1}(1+x)e^{-\theta x}, \ x > 0, \ \alpha > 0, \ \theta > 0, \tag{3.13}$$

where $\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y} dy, \ \alpha > 0$, is the gamma function. The p.d.f (3.13) can be expressed as a two component mixture.
$f(x) = pf_1(x) + (1-p)f_2(x)$, where $p = \dfrac{\theta}{\theta+\alpha}$ and

$$f_j(x) = \frac{\theta^{\alpha+j-1}}{\Gamma(\alpha+j-1)} x^{\alpha+j-2} e^{-\theta x}, \ x > 0, \ \ \alpha > 0, \ \ \theta > 0, \ \ j = 1, 2$$

is the probability density function of gamma distribution with shape parameter $\alpha + j - 1$ and scale parameter θ, denoted by $\Gamma(\alpha + j - 1, \theta), \ j = 1, 2$.

3.2.1 Structural properties

In this section, we present the structural shape properties of the density, moments, hazard, survival and mean residual life function of a two-parameter weighted Lindley distribution.

1. Probability density function

At $x = 0$ and $x = \infty$, the behaviours of the probability density function of the weighted Lindley distribution $f(x)$ respectively are given by

$$f(0) = \begin{cases} \infty & \text{if } \ \alpha < 1, \\ \dfrac{\theta^2}{\theta+1} & \text{if } \alpha = 1, \\ 0 & \text{if } \ \alpha > 1, \end{cases} \qquad\qquad f(\infty) = 0$$

Theorem 3.2 (Ghitany et al. [10]) *The probability density function of the weighted Lindley distribution is*

(i) *decreasing if* $\{\theta < \alpha < 1 : (\theta+\alpha)^2 - 4\theta \leq 0\}$ *or* $\{\alpha \leq 1, \theta \geq \alpha\}$

(ii) *Unimodal if* $\{\alpha = 1, \theta < 1\}$ *or* $\{\alpha > 1, \theta > 0\}$.

(iii) *Decreasing-increasing-decreasing if* $\{\theta < \alpha < 1 : (\theta+\alpha)^2 - 4\theta > 0\}$.

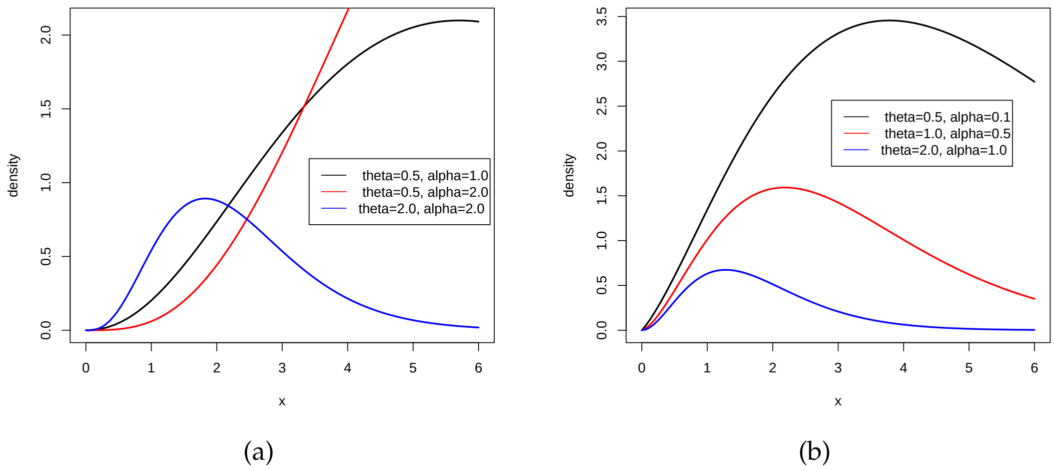


Figure 3.3 – *Probability density function of two-parameter weighted Lindley distribution :(a)* $\theta = 0.5,\ \alpha = 1,\ \theta = 0.5, \alpha = 2,\ \theta = 2,\ \alpha = 2$ *(b)* $\theta = 0.5$, $\alpha = 0.1,\ \theta = 1,\ \alpha = 0.5,\ \theta = 2,\ \alpha = 1$

The r th moment of the weighted Lindley distribution is given by

$$\mu_r^{'} = \mathbb{E}(X^r) = p\int_0^\infty x^r f_1(x)dx + (1-p)\int_0^\infty x^r f_2(x)dx = \frac{(\theta+\alpha+r)\Gamma(\alpha+r)}{\theta^r(\theta+\alpha)\Gamma(\theta)}, r = 1, 2, \ldots$$

Therefore, the mean and variance of the weighted Lindley distribution, respectively are

$$\mu^{'} = \frac{\alpha(\theta+\alpha+1)}{\theta(\theta+\alpha)},\ \mu_2^{'} = \sigma^2 = \frac{(\alpha+1)(\theta+\alpha)^2 - \theta^2}{\theta^2(\theta+\alpha)^2}$$

2. Survival function

The fact that the probability density function of the two-parameter Lindley distribution is a two component mixture, the survival function S is given by

$$\begin{aligned} S(x) = P(X > x) &= p\int_x^\infty f_1(y)dy + (1-p)\int_x^\infty f_2(y)dy \\ &= \frac{(\theta+\alpha)\Gamma(\alpha,\theta x) + (\theta x)^\alpha e^{-\theta x}}{(\theta+\alpha)\Gamma(\alpha)}, \; x > 0, \; \alpha > 0, \; \theta > 0 \end{aligned}$$

where $\Gamma(\alpha,\theta x) = \int_{\theta x}^\infty y^{\alpha-1}e^{-y}dy, \; \alpha > 0, \; \theta x \geq 0$, is the upper incomplete gamma function.

Now, the hazard function which is defined as $h(x) = \dfrac{f(x)}{S(x)}$ of the weighted Lindley distribution is given by

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

At $x = 0$ and $x = \infty$, the behaviours of the hazard function of the weighted Lindley distribution $h(x)$ respectively are given by

$$h(0) = f(0) = \begin{cases} \infty & \text{if } \alpha < 1, \\ \dfrac{\theta^2}{\theta+1} & \text{if } \alpha = 1, \\ 0 & \text{if } \alpha > 1, \end{cases} \qquad h(\infty) = \theta$$

It is difficult to study the shape of $h(x)$ because it includes the upper incomplete gamma function.Fortunately, following lemma is used to determine its shape.

Lemma 3.1 (Ghitany et al. [10]) *Let X be a non-negative continuous random variable with twice differentiable p.d.f $f(x)$ and hazard rate function $h(x)$. Let $\eta(x) = -\dfrac{d\ln f(x)}{dx}$.*

(i) If $\eta(x)$ is increasing (decreasing) in x, then $h(x)$ is increasing (decreasing) in x.

(ii) If $\eta(x)$ has a bathtub (upside-down bathtub) shape, $h(x)$ has also a bathtub (upside-down bathtub) shape.

Using this lemma to determine the shape of hazard rate function we have

$$\begin{aligned}
\eta(x) &= -\frac{d\ln f(x)}{dx} = -\frac{f^{'}(x)}{f(x)} \\
&= -\frac{\left[\frac{\theta^{\alpha+1}}{(\theta+\alpha)\Gamma(\alpha)}(\alpha-1)x^{\alpha-2}(1+x)e^{-\theta x} + \frac{\theta^{\alpha+1}}{(\theta+\alpha)\Gamma(\alpha)}x^{\alpha-1(e^{-\theta x}-\theta(1+x)e^{-\theta x})}\right]}{\frac{\theta^{\alpha+1}}{(\theta+\alpha)\Gamma(\alpha)}(1+x)e^{-\theta x}} \\
&= -\frac{\alpha-1}{x} - \frac{1}{1+x} + \theta
\end{aligned}$$

it follows that $\eta^{'}(x) = \dfrac{\alpha-1}{x^2} + \dfrac{1}{(1+x)^2}$. For $0 < \alpha < 1$, $\eta^{'}(x) = 0$ implies that $\eta(x)$ has a global minimum at $x^* = (-(1-\alpha) + \sqrt{1-\alpha})/\alpha$. $\eta(x)$ is bathtub shaped. Therefore, $h(x)$ is also bathtub shaped. For $\alpha \geq 1$, $\eta^{'}(x) \geq 0$, i.e. $\eta(x)$ is increasing thus $h(x)$ is also increasing. We can conclude that the hazard rate function of a weighted Lindley distribution is bathtub shaped (increasing) if $0 < \alpha < 1 (\alpha \geq 1)$ for all $\theta > 0$

3. Mean residual life function

Using the fact that the density of the weighted Lindley distribution can be expressed as a two-component mixture, its mean residual life function $m(x) = \mathbb{E}(X - x|X > x)$ is given by

$$\frac{[\alpha(\theta+\alpha+1) - (\theta+\alpha)\theta x]\Gamma(\alpha,\theta x) + (\theta+\alpha+1)(\theta x)^{\alpha}e^{-\theta x}}{\theta[(\theta+\alpha)\Gamma(\alpha,\theta x) + (\theta x)^{\alpha}e^{-\theta x}]}$$

At $x = 0$ and $x = \infty$,

$$m(0) = \frac{\alpha(\theta+\alpha+1)}{\theta(\theta+\alpha)}$$

and

$$m(\infty) = \frac{1}{\theta}.$$

3.2.2 Estimation

Let $x = (x_1, ..., x_n)$ be a realization of the random sample $X = (X_1, ..., X_n)$, where $X_1, ..., X_n$ are independent and identically distributed (i.i.d) random variables with parameter vector $\beta = (\theta,\ \alpha),\ \ \theta > 0,\ \ \alpha > 0$. From the density given by equation (3.13) the Likelihood function can be written as:

$$L_n(\beta) = \left(\frac{\theta^{\alpha+1}}{(\theta+\alpha)\Gamma(\alpha)}\right)^n e^{-\theta\Sigma_{i=1}^n x_i} \prod_{i=1}^n x_i^{\alpha-1}(1+x_i)$$

where $\Gamma(\alpha)$ is the complete gamma function. The log-likelihood function for θ and α is

$$\begin{aligned}\ell_n(\beta) &= \ln\left(\frac{\theta^{\alpha+1}}{(\theta+\alpha)\Gamma(\alpha)}\right)^n + \ln(e^{-\theta\Sigma_{i=1}^n x_i}) + \ln\left(\prod_{i=1}^n x_i^{\alpha-1}(1+x_i)\right)\\ &= n[(\alpha+1)\ln(\theta) - \ln(\theta+\alpha) - \ln\Gamma(\alpha)] - \theta\Sigma_{i=1}^n x_i\\ &+ (\alpha-1)\Sigma_{i=1}^n log(x_i) + \Sigma_{i=1}^n \ln(1+x_i)\end{aligned}$$

Differentiating with respect to θ and α we have the vector $u_\beta = (u_\theta,\ u_\alpha)$ with components

$$u_\theta = n\left[\frac{\alpha+1}{\theta} - \frac{1}{\theta+\alpha}\right] - \Sigma_{i=1}^n x_i$$

and

$$u_\alpha = n\left[\ln(\theta) - \frac{1}{\theta+\alpha} - \psi(\alpha)\right] + \Sigma_{i=1}^n \ln(x_i)$$

where $\psi(\alpha) = \dfrac{d}{d\alpha}\ \ln\Gamma(\alpha) = \dfrac{\Gamma'(\alpha)}{\Gamma(\alpha)}$ which represent digamma function. To determine $\hat{\theta}$ and $\hat{\alpha}$, we solve simultaneous equations $u_\theta = 0$ and $u_\alpha = 0$. The equation $u_\theta = 0$ can be solved algebraically for θ giving

$$\hat{\theta}(\alpha) = \frac{\alpha(1-\bar{X}) + \sqrt{\alpha^2(1+\bar{X})^2 + 4\alpha\bar{X}}}{2\bar{X}}$$

which is the maximum likelihood estimate of θ when α is assumed to be known, where $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$. The maximum Likelihood estimate for α can be obtained solving $u_\alpha = 0$ by setting $\theta = \hat{\theta}(\alpha)$.

In this situation we use the asymptotic normality results to obtain the asymptotic confidence interval. We can state the results as

$$\sqrt{n}[(\hat{\alpha} - \alpha), (\hat{\theta} - \theta)] \rightarrow \mathcal{N}_2\left(0, \mathbf{I}^{-1}(\alpha, \lambda)\right)$$

where $\mathbf{I}(\alpha, \theta)$ is the Fisher Information matrix. The elements of Fisher information matrix $\mathbf{I} = [I_{ij}], i, j = 1, 2$ about $(\alpha,\ \theta)$ from a single observation are given by

$$I_{11} = -\mathbb{E}\left[\frac{\partial^2 \ell_n(\beta)}{\partial \alpha^2}\right] = \psi^{'}(\alpha) - \frac{1}{(\theta+\alpha)^2}, \ I_{22} = -\mathbb{E}\left[\frac{\partial^2 \ell_n(\beta)}{\partial \theta^2}\right] = \frac{\alpha+1}{\theta^2} - \frac{1}{(\theta+\alpha)^2},$$

and

$$I_{12} = I_{21} = -\mathbb{E}\left[\frac{\partial^2 \ell_n(\beta)}{\partial \alpha \partial \theta}\right] = -\left[\frac{1}{\theta} + \frac{1}{(\theta+\alpha)^2}\right]$$

$$\mathbf{I} = \begin{pmatrix} \psi^{'}(\alpha) - \dfrac{1}{(\theta+\alpha)^2} & -\left(\dfrac{1}{\theta} + \dfrac{1}{(\theta+\alpha)^2}\right) \\ -\left(\dfrac{1}{\theta} + \dfrac{1}{(\theta+\alpha)^2}\right) & \dfrac{\alpha+1}{\theta^2} - \dfrac{1}{(\theta+\alpha)^2} \end{pmatrix}$$

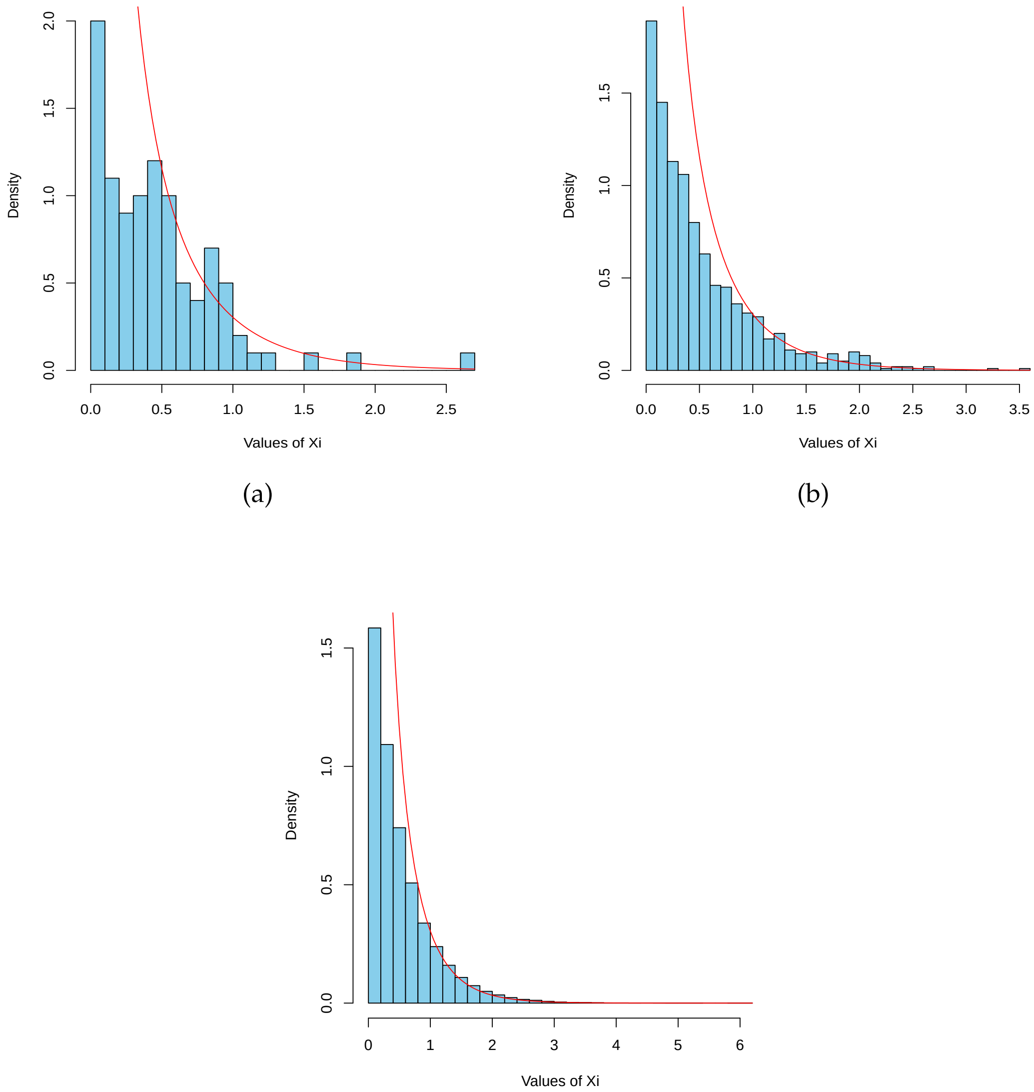
where $\psi^{'}(\alpha) = \frac{d^2}{d\alpha^2}\ln\Gamma(\alpha)$ is trigamma function. The inverse of $\mathbf{I}$ evaluated at $\hat{\theta}$ and $\hat{\alpha}$ provides the asymptotic variance-covariance matrix of the maximum Likelihood estimates. It has been shown in Ghitany et al. [10] that the MLEs $\hat{\mathbf{ff}}$ and $\hat{\theta}$ are positively correlated. The asymptotic $100(1-\alpha)\%$ confidence interval of α and θ, respectively, are given by

$$\hat{\alpha} \pm z_{\alpha/2}\sqrt{\widehat{\text{Var}(\hat{\alpha})}}, \quad \hat{\theta} \pm z_{\alpha/2}\sqrt{\widehat{\text{Var}(\hat{\theta})}},$$

where $\widehat{\text{Var}(\hat{\alpha})}(\widehat{\text{Var}(\hat{\theta})})$ is the MLE of $\text{Var}(\hat{\alpha})(\text{Var}(\hat{\theta}))$ and $z_{\alpha/2}$ is the upper $\alpha/2$ quantile of the standard normal distribution.

3.2.3 Simulation of the two-parameter weighted Lindley distribution

Here are the histograms of simulated weighted Lindley distributions for different values of N using the fact that this distribution is a mixture of $\exp(\theta)$ and gamma $(2, \theta)$



(a) (b)

(c)

Figure 3.4 – *Histograms showing simulated two-parameter weighted Lindley distribution using the fact that it is a mixture of two distributions for* $\alpha = 0.1, \ \theta = 2 : (a) \ N = 100, \quad (b) \ N = 1000, \ (c) \ N = 10000$.

Applications of Lindley distribution 4

4.1 Introduction

This section focuses on the applications of Lindley distribution where in the first part we discuss the discrimination procedure of random variables, followed by application to real data and simulated data.

4.2 Discrimination of random variables

From a statistical modelling perspective, It is an important problem in statistics to test which probability model some given observations follow. Suppose that an experimenter has observed n data points say $x_1, x_1, ..., x_n$ and he wants to use either exponential model with parameter λ or one parameter Lindley model with parameter θ, which one will he prefer?

4.2.1 Discrimination criterion methods

There exist several methods for discriminating random variables such as statistical tests, visualisation techniques, maximum Likelihood Ratio (RML) and methods based on log likelihood function such as

- Akaike Information criterion (AIC). This statistical measure was introduced by Hirotugu Akaike in 1973. It helps in model selection by balancing the goodness of fit and the complexity of the model.
- Hannan-Quinn Information Criterion (HQIC). It offers a compromise between AIC and BIC in terms of model complexity. Th penalty term $2k\ln(\ln(n))$ lies between the lighter penalty of AIC and the heavier penalty of BIC.
- Bayesian Information Criterion (BIC). It is closely related to Akaike informa-

tion criterion. When fitting models, it is possible to increase the likelihood by adding parameters, but by doing so may result in overfitting. The BIC resolves this problem by introducing a penalty term for the number of parameters in the model.

- Corrected Akaike Information Criterion (AICC). It is a modification of AIC. While AIC tends to perform well with larger datasets, it can lead to overfitting when the sample size is small. AICC introduces a correction for finite sample sizes, making it more reliable in such cases.
- K-s statistics (Kolmogorov-Smirnov statistics).Its the greatest distance between the cumulative distribution function of each sample.
- Consistent Akaike Information Criterion (CAIC). It is also used particularly when dealing with large datasets where overfitting is a concern.

These criteria are computed as follows:

$$AIC = -2\ln L_n + 2k, \quad AICC = AIC + \frac{2k(k+1)}{(n-k-1)},$$
$$BIC = -2\ln L_n + k\ln n, \quad HQIC = -2\ln L_n + 2k\ln(\ln(n)),$$
$$CAIC = -2\ln L_n + k(ln(n)+1), \quad D = \sup_x |F_n(x) - F_0(x)|$$

where k is the number of parameters, n is the sample size and F_n is the empirical distribution function.

The best model is the one that provides the minimum values of these criteria.

4.2.2 Discrimination between exponential and Lindley distributions

In this subsection, we study the discrimination procedure based on maximum Likelihood Ratio (RML) to discriminate between exponential and one parameter Lindley distributions (see for instance Vaidyanathan and Varghese [28]). For the discrimination between Lindley and Xgamma distributions (see Sen et al. [25]).
We shall use the following notations throughout the section:
The probability density function (p d f) of a Lindley distribution with parameter θ is denoted by

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

We shall denote it by $X \sim LD(\theta)$. The density function of exponential with parameter λ is denoted by

$$f_{\exp}(x) = \lambda e^{-\lambda x}, \;\; x \geq 0, \; \lambda > 0.$$

We shall denote it by $X \sim \exp(\lambda)$.

Ratio of maximum likelihoods

Let $x_1, x_2, ..., x_n$ be realisations of independent and identically distributed (i.i.d) random variables $X_1, X_2, ..., X_n$ from an exponential distribution or from a one parameter Lindley distribution. The likelihood functions of exponential and one parameter Lindley distributions respectively are given by

$$L_n(\lambda) = \lambda^n e^{-\lambda \sum_{i=1}^{n} x_i} \tag{4.1}$$

$$L_n(\theta) = \left(\frac{\theta^2}{\theta+1}\right)^n e^{-\theta \sum_{i=1}^{n} x_i} \prod_{i=1}^{n} (1+x_i) \tag{4.2}$$

The ratio of maximum likelihood (RML) is defined as

$$RML = \frac{L_n(\hat{\lambda})}{L_n(\hat{\theta})} \tag{4.3}$$

where $\hat{\lambda}$ and $\hat{\theta}$ are estimators of λ and θ, respectively, based on $\{X_1, X_2, ..., X_n\}$. Taking natural logarithm of the RML, we have the log-likelihood ratio denoted by T as

$$T = n\ln(\hat{\lambda}) - \hat{\lambda}\sum_{i=1}^{n} x_i - 2n\ln\hat{\theta} - \sum_{i=1}^{n}\ln(1+x_i) + \hat{\theta}\sum_{i=1}^{n} x_i + n\ln(1+\hat{\theta}), \tag{4.4}$$

where $\hat{\lambda}$ and $\hat{\theta}$ are given by $\frac{1}{\bar{X}}$ and $\frac{-(\bar{X}-1)+\sqrt{(\bar{X}-1)^2+8\bar{X}}}{2\bar{X}}$ respectively. where $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

The proposed procedure is as follows. Select exponential distribution if $T > 0$, otherwise select Lindley distribution. $T > 0$ implies that the numerator of RML is larger than the denominator. This means that the sample observations are more likely to have been generated from the exponential distribution than from Lindley distribution and hence we choose $\exp(\lambda)$ to model the given data. If

$T < 0$, the denominator of RML will be larger than the numerator, which implies that $LD(\theta)$ is a better fit to the given data.

Asymptotic properties of RML

We denote the almost sure convergence by a.s, $\mathbb{E}_{LD}[h(Y)]$ as mean of $h(Y)$ under the assumption that $Y \sim LD(\theta)$ and $\mathbb{E}_{\exp}[h(Y)]$ as mean of $h(Y)$ under the assumption that $Y \sim \exp(\lambda)$.

Case 1: The null hypothesis is that given data is from LD(θ) and the alternative hypothesis is that given data is from $\exp(\lambda)$. The following lemma is needed to find the asymptotic distribution of T.

Lemma 4.1 (**Vaidyanathan and Varghese [28]**) *Under the assumption that data is from Lindley distribution, as $n \to \infty$*

a) $\hat{\theta} \to \theta$ *a.s*

b) $\hat{\lambda} \to \lambda^*$ *a.s where λ^* is the value of λ that maximizes* $\mathbb{E}_{LD}\left[\ln f_{\exp}(x)\right]$.

c) *Let* $T^* = \ln\left(\dfrac{L_n(\lambda^*)}{L_n(\theta)}\right)$ *then* $(T - \mathbb{E}_{LD}(T))/\sqrt{n}$ *is asymptotically equivalent to* $T^* - \mathbb{E}_{LD}\left[\dfrac{T^*}{\sqrt{n}}\right]$.

Proof. **a)** It is clear that $\hat{\theta} \to \theta$ a.s by using the convergence property of Maximum Likelihood estimators.

b) In this case, it is observed that $\hat{\lambda}$ is a pseudo MLE of λ. Hence $\hat{\lambda} \nrightarrow \lambda$. However $\hat{\lambda} \to \lambda^*$ a.s, where λ^* is the value of λ that minimizes Kullback-Leibler information criteria. λ^* is obtained by minimizing

$$\mathbb{E}_{LD}\left[\ln\left(\frac{f_{LD}(x)}{f_{\exp}(x)}\right)\right] \tag{4.5}$$

Minimizing equation (4.5) is the same as maximizing $\mathbb{E}_{LD}\left[\ln f_{exp}(x)\right]$ λ^* is got as follows

$$\begin{aligned}\mathbb{E}_{LD}(\ln f_{\exp}(x)) &= \int_0^\infty \ln(\lambda e^{-\lambda x})\frac{\theta^2}{\theta+1}(1+x)e^{-\theta x}dx \\ &= \int_0^\infty (\ln\lambda - \lambda x)\frac{\theta^2}{\theta+1}(1+x)e^{-\theta x}dx\end{aligned}$$

$$\begin{aligned}\mathbb{E}_{LD}(\ln f_{\exp}(x)) =&(\ln\lambda)\left[\int_0^\infty \frac{\theta^2}{\theta+1}e^{-\theta x}dx + \int_0^\infty \frac{\theta^2}{\theta+1}xe^{-\theta x}dx\right]\\ &-\lambda\left[\int_0^\infty \frac{\theta^2}{\theta+1}xe^{-\theta x}dx + \int_0^\infty \frac{\theta^2}{\theta+1}x^2e^{-\theta x}dx\right]\end{aligned} \tag{4.6}$$

By applying integration by parts we get

$$\begin{aligned}\mathbb{E}_{LD}(\ln f_{\exp}(x)) &= \ln\lambda\left[\frac{\theta}{\theta+1}+\frac{1}{\theta+1}\right]-\lambda\left[\frac{1}{\theta+1}+\frac{2}{\theta(\theta+1)}\right]\\ &= \ln\lambda-\lambda\left(\frac{\theta+2}{\theta(\theta+1)}\right)\end{aligned}$$

let $g(\lambda) := \ln\lambda-\lambda\left(\frac{\theta+2}{\theta(\theta+1)}\right)$

finding the derivative of g with respect to λ we get

$$g^{'}(\lambda) = \frac{1}{\lambda}-\left(\frac{\theta+2}{\theta(\theta+1)}\right) \quad \text{and} \quad g^{''}(\lambda) = -\frac{1}{\lambda^2}.$$

$$g^{'}(\lambda) = 0 \Rightarrow \lambda = \frac{\theta(\theta+1)}{\theta+2}$$

At $\lambda = \frac{\theta(\theta+1)}{\theta+2}$, we have $g^{''}(\lambda) = g^{''}(\frac{\theta(\theta+1)}{\theta+2}) - \frac{(\theta+2)^2}{(\theta(\theta+1))^2} < 0$

thus

$$\lambda^* = \frac{\theta(\theta+1)}{\theta+2}$$

□

Theorem 4.1 (**Vaidyanathan and Varghese [28]**) *Under the assumption that data is from Lindley distribution, T is asymptotically distributed as normal with mean $\mathbb{E}_{LD}(T)$ and variance $V_{LD}(T)$.*

- Suppose that the given data is from exponential distribution

Lemma 4.2 *Under the assumption that data is from* $\exp(\lambda)$, *as* $n \to \infty$

a. $\hat{\lambda} \to \lambda$ *a.s*

b. $\hat{\theta} \to \theta^*$ *a.s where* θ^* *is the value of* θ *that maximizes* $\mathbb{E}_{\exp}(\ln f_{LD}(x))$.

c. *Let* $T^* = \ln(L_n(\lambda)/L_n(\theta^*))$ *then* $(T-\mathbb{E}_{\exp}(T))/\sqrt{n}$ *is asymptotically equivalent to* $(T^*-\mathbb{E}_{\exp}(T^*))/\sqrt{n}$.

The proof to this lemma is similar to that of lemma(put the number of that lemma of case 1 above) with

$$\theta^* = \frac{\lambda - 1 + \sqrt{1 + 6\lambda + \lambda^2}}{2}$$

Theorem 4.2 (**Vaidyanathan and Varghese [28]**) *Under the assumption that the data is from an Exponential distribution, T is asymptotically distributed as normal with mean $\mathbb{E}_{\exp(T)}$ and variance $V_{\exp}(T)$.*

4.2.3 Application

The following table representing the lifetime's data relating to relief times (in minutes) of 20 patients receiving an analgesic and reported by Shanker et al. [26].

1.1	1.4	1.3	1.7	1.9	1.8	1.6	2.2	1.7	2.7
4.1	1.8	1.5	1.2	1.4	3	1.7	2.3	1.6	2

Estimate of parameters for the two models together with different criteria are computed

Table 4.1 – *MLEs, $-2\ln L_n$, AIC, AICC, BIC, HQIC, D and CAIC of the fitted distributions of the above data set*

model	MLEs	$-2\ln L_n$	AIC	AICC	BIC	HQIC	D	CAIC
Exponential	0.5263158	65.7	67.7	67.9	68.7	67.696	0.3895	69.69
Lindley	0.816118	60.50	62.50	62.72	63.49	62.62	0.3411	64.496

From the table of values, we find that the log-likelihood ratio $T = -2.587526$ implying that the given data has been picked from Lindley distribution thus choosing Lindley distribution as a better fit to the given data.

It is also clear that Lindley distribution is better for modelling this data since it corresponds to lower $-2\ln L_n$, AIC, AICC, BIC, HQIC, D and CAIC.

4.3 Application to real data

4.3.1 Application to life time data

In this section, we use a real data set to show that two-parameter Lindley distribution can be a better model than one based on one parameter Lindley, exponential and the two-parameter Weibull distributions. Let's recall that the density of Weibull of parameter α and β is

$$f(x) = \alpha\beta^{\alpha}x^{\alpha-1}\exp\left[-(\beta x)^{\alpha}\right], \quad x > 0, \alpha > 0 \text{ and } \beta > 0.$$

The following table represents waiting times (in minutes) before service of 100 bank customers reported by Ghitany et al. [9].

Table 4.2 – *Waiting times (in minutes) of 100 bank customers.*

0.8	0.8	1.3	1.5	1.8	1.9	1.9	2.1	2.6	2.7
2.9	3.1	3.2	3.3	3.5	3.6	4.0	4.1	4.2	4.2
4.3	4.3	4.4	4.4	4.6	4.7	4.7	4.8	4.9	4.9
5.0	5.3	5.5	5.7	5.7	6.1	6.2	6.2	6.2	6.3
6.7	6.9	7.1	7.1	7.1	7.1	7.4	7.6	7.7	8.0
8.2	8.6	8.6	8.6	8.8	8.8	8.9	8.9	9.5	9.6
9.7	9.8	10.7	10.9	11.0	11.0	11.1	11.2	11.2	11.5
11.9	12.4	12.5	12.9	13.0	13.1	13.3	13.6	13.7	13.9
14.1	15.4	15.4	17.3	17.3	18.1	18.2	18.4	18.9	19.0
19.9	20.6	21.3	21.4	21.9	23.0	27.0	31.6	33.1	38.5

Estimation of the parameters are carried out using method of moments and are recorded in the table below.

Table 4.3 – *Summary fit for the life time data*

Model	Estimate of parameters
Lindley (θ)	$\hat{\theta} = 0.187$
Lindley $(\alpha, \ \theta)$	$\hat{\alpha} = 2.96669, \quad \hat{\theta} = 0.196209$
Exponential (λ)	$\hat{\lambda} = 0.10125$
Weibull $(\alpha, \ \beta)$	$\hat{\alpha} = 3.194, \quad \hat{\beta} = 1.170$

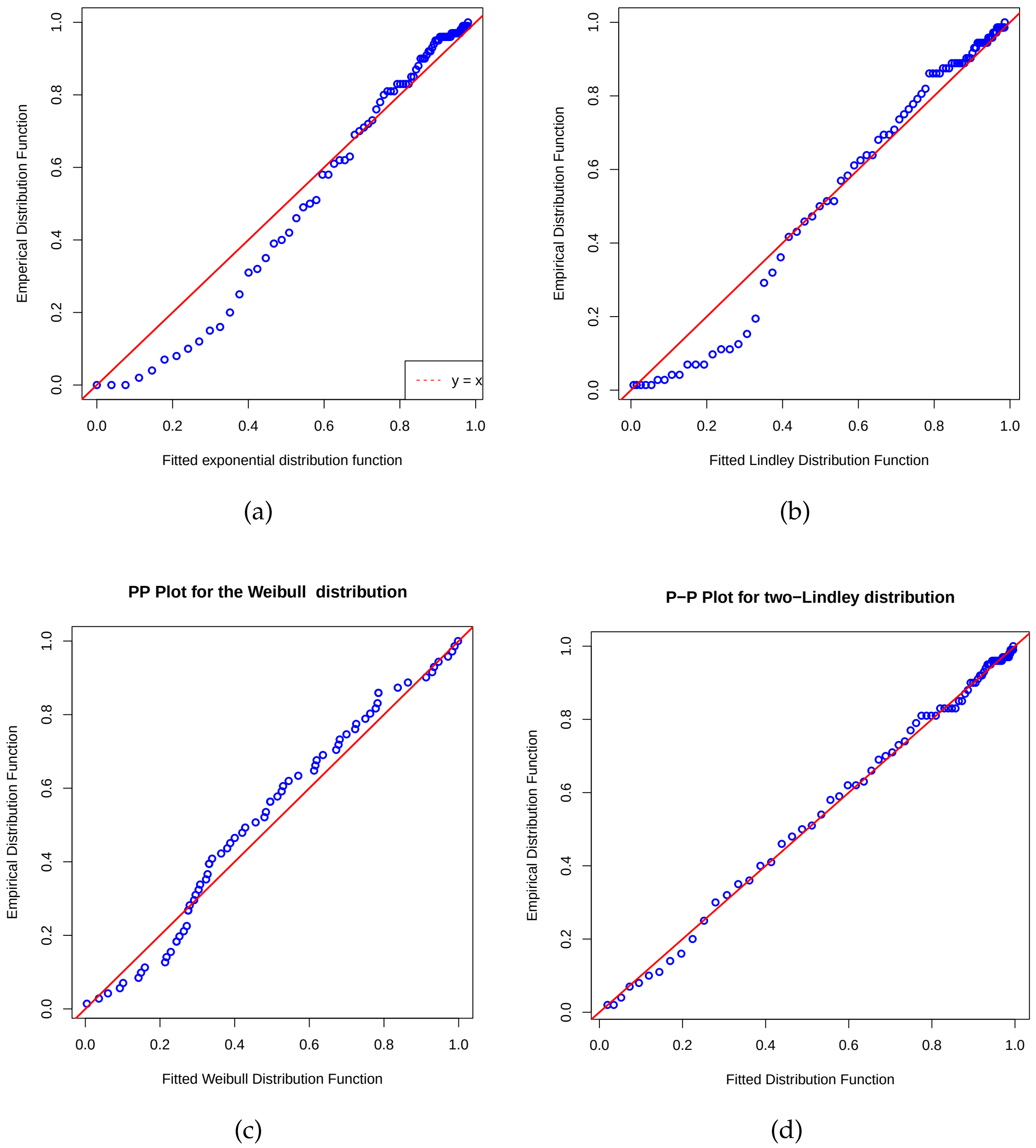


Figure 4.1 – *P-P plots for (a) exponential, (b) one parameter Lindley, (c) Weibull and (d) Two-Lindley*

Observing the p-p plots above, points closely follow the diagonal line in the two-parameter Lindley distribution model compared to Weibull, exponential and one parameter Lindley distributions. This shows that two-parameter Lindley distribution fits data well that the rest. Thus we use two-parameter Lindley distribution to analyse some of characteristics of waiting time data such as

- Expected waiting time data is $\mu = 1.663$.
- Probability of waiting less than 5 minutes is 0.979.
- Probability of waiting more than 5 minutes is 0.021.

By characterising the distribution of waiting times, organisations can gain insights into the efficiency of their service processes, potential bottlenecks, and optimise resource allocation to minimise wait times and improve customer satisfaction.

4.3.2 Application to survival data

Lindley distribution has found significant applications in survival analysis due to its versatility and ability to fit a variety of data patterns. In survival analysis, the primary objective is to model and analyse time-to-event data, which makes the choice of an appropriate statistical distribution crucial. However much we have seen that two-parameter Lindley model gives better fit, in this section we will show that the two-parameter weighted Lindley distribution fits data well compared to one-parameter Lindley, exponential, Weibull models and generalized exponential.

We present an application used in Ghitany et al. [10] to the data set representing the survival times (in days) of 72 guinea pigs infected with virulent tubercle bacilli, observed and reported by Bjerkedal et al.[6]. Its values are

10	33	44	56	59	72	74	77	92	93	96	100
100	102	105	107	107	108	108	108	109	112	113	115
116	120	121	122	122	124	130	134	136	139	144	146
153	159	160	163	163	168	171	172	176	183	195	196
197	202	213	215	216	222	230	231	240	245	251	253
254	254	278	293	327	342	347	361	402	432	458	555

In this application we use the MLE method for estimating the parameters of the considered models. For the models weighted Lindley and generalized- exponential we use the function optim of R software. The results obtained are given in the table below

Table 4.4 – *Summary fit for the survival data*

Model	MLEs	$-2\ln L_n$	AIC	AICC	BIC
Lindley (θ)	$\hat{\theta} = 0.011$	865.636	867.636	867.693	866.174
Exponential (λ)	$\hat{\lambda} = 0.006$	888.191	890.191	890.248	888.729
Weibull (α, β)	$\hat{\alpha} = 199.602,\ \hat{\beta} = 1.825$	4581.514	4585.514	4585.688	4582.590
Weighted Lindley (α, θ)	$\hat{\alpha} = 2.105,\ \hat{\theta} = 0.018$	851.546	855.546	855.720	852.623
Generalized exponential (c, λ)	$\hat{c} = 3.6748,\ \hat{\lambda} = 0.011$	851.6212	855.6212	855.7951	860.1745

A close examination of the table 4.4 with estimators and some of the comparison methods reveals that the two-parameter weighted Lindley model is the best choice among the competing models, since it has the lowest $-2\ln L_n$, AIC, AICC, and BIC. This is also supported by the probability-probability (P-p) plots in figures below.

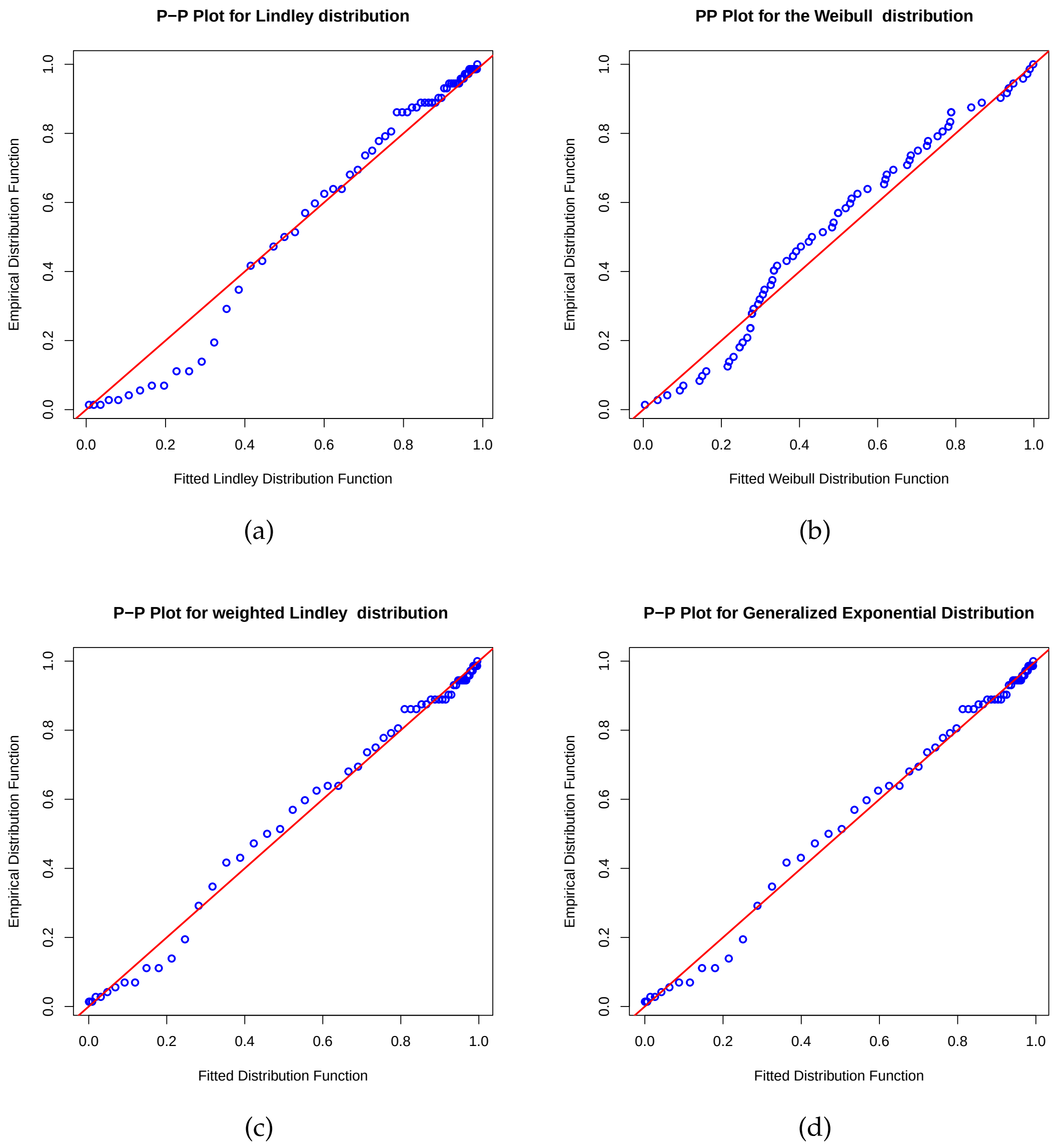
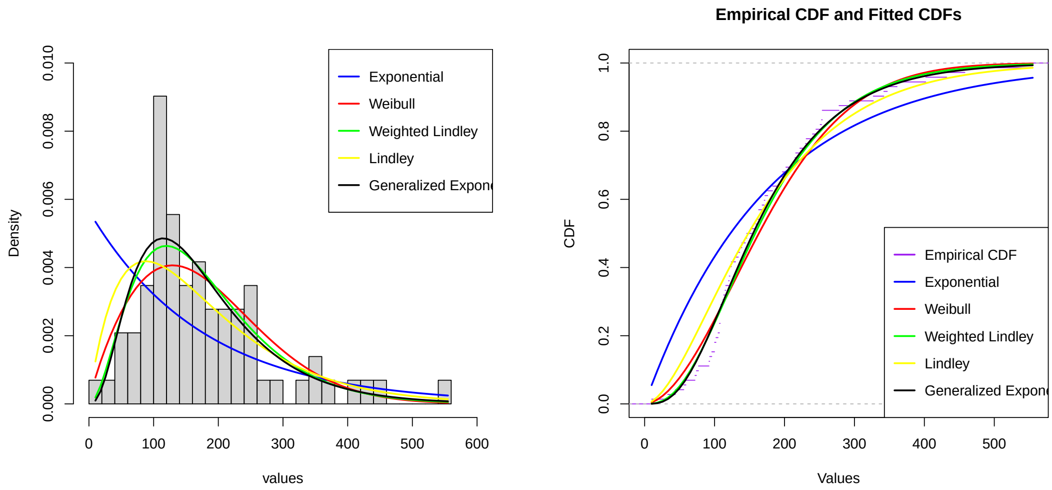


Figure 4.2 – *P-P plot for different models.*



(a) (b)

Figure 4.3 – *Graph of fitted PDFs and CDFs (Survival data set).*

The use of weighted Lindley model enhances the accuracy of survival analysis, offering improved parameter estimation and better fit to empirical data. This approach is valuable where understanding the distribution of survival times is crucial for decision-making and policy development.

4.4 Application to simulated data

In this section, we present the main results obtained from simulations. This part allows us to test the effectiveness of the maximum likelihood method and the Bayesian approach in estimating the one parameter Lindley model on simulated data following this model. It also allows for the comparison of the performance of both approaches. The results were achieved using R software (see program 3 in the Annex). For this, we proceed as follow

We generate 10000 samples of the Lindley distribution, of size $n = 20$, 50, 100, 200 and 500 to represent small, moderate and large sample sizes from Lindley distribution with two value of $\theta(\theta = 0.1,\ \theta = 3$. Maximum likelihood and Bayesian estimators are noted in the tables below together with their corresponding MSEs. Bayesian estimators of parameters are evaluated under Linex loss function and using two priors i.e extension of Jeffreys with $c = 1,\ c = 2$ and the inverted gamma prior; $(\alpha, \beta) \in \{(1, 1.5), (1.5, 2)\}$ with the two values of $a(a = 1,\ a = -1)$.

The following measures were computed:

$$\mathrm{MSE}(\theta) = \frac{1}{N_{sim}} \sum_{i=1}^{N_{sim}} \left(\widehat{\theta}_i - \theta\right)^2$$

Average bias of the simulated estimates $\theta_i,\ i = 1, 2, ..., N_{sim}$

$$\mathrm{Bias}(\theta) = \frac{1}{N_{sim}} \sum_{i=1}^{N_{sim}} \left(\widehat{\theta}_i - \theta\right)$$

$$\bar{\hat{\theta}} = \frac{1}{N_{sim}} \sum_{i=1}^{N_{\mathrm{sim}}} \hat{\theta}_i$$

The results are summarized and tabulated in the following tables.

Table 4.5 – *MLE and Bayes Estimates under linex loss function* ($\alpha = 1,\ b = 1.5,\ a = -1$, $c = 1$)

n	Method	$\bar{\hat{\theta}}$	MSE	Bias
$n = 30$	MLE	0.1264968	0.002415809	0.02649681
	Extension of Jeffrey prior	0.1235499	0.00219517	0.02354986
	Inverted Gamma prior	0.1550891	0.004534634	0.05508908
$n = 50$	MLE	0.1110698	0.0008069258	0.01106976
	Extension of Jeffrey prior	0.1100233	0.0007727713	0.01002332
	Inverted Gamma prior	0.1230098	0.001178395	0.02300978
$n = 100$	MLE	0.1057439	0.0003718375	0.005743872
	Extension of Jeffrey prior	0.1052437	0.0003633185	0.00524372
	Inverted Gamma prior	0.1118018	0.000469239	0.0118018
$n = 200$	MLE	0.1030814	0.0001843936	0.003081359
	Extension of Jeffrey prior	0.1028371	0.0001821637	0.00283708
	Inverted Gamma prior	0.1061325	0.0002101968	0.006132471
$n = 500$	MLE	0.1012363	7.608762e-05	0.001236293
	Extension of Jeffrey prior	0.1011402	7.572517e-05	0.001140197
	Inverted Gamma prior	0.1024627	8.02289e-05	0.002462732

Table 4.6 – *MLE and Bayes Estimates under linex loss function* ($\alpha = 1.5$, $b = 2$, $a = 1$, $c = 2.5$)

n	Method	$\bar{\hat{\theta}}$	MSE	Bias
$n = 30$	MLE	2.944352	0.4906837	-0.05564769
	Extension of Jeffrey prior	2.630747	0.5012555	-0.3692529
	Inverted Gamma prior	2.934057	0.4819162	-0.06594262
$n = 50$	MLE	2.97337	0.3105848	-0.0266298
	Extension of Jeffrey prior	2.770938	0.3080475	-0.2290617
	Inverted Gamma prior	2.967071	0.3071386	-0.03292906
$n = 100$	MLE	2.988489	0.154836	-0.1183531
	Extension of Jeffrey prior	2.881647	0.154836	-0.1183531
	Inverted Gamma prior	2.985319	0.1554709	-0.0146812
$n = 200$	MLE	2.994183	0.08023076	-0.005817243
	Extension of Jeffrey prior	2.939311	0.0796838	-0.06068891
	Inverted Gamma prior	2.992596	0.0800064	-0.00740391
$n = 500$	MLE	2.997204	0.03103324	-0.002795955
	Extension of Jeffrey prior	2.974904	0.03098408	-0.02509574
	Inverted Gamma prior	2.996569	0.03099939	-0.003430606

From the tables above, we observe that all the estimates show the property of consistency i.e the MSEs and the biases decrease as the sample size increases. Also the MSEs of Bayesian estimators using extension of Jeffrey prior are less compared to the rest of the estimators for $n = 20$, 50 , 100, 200, 500.

Conclusion and Perspectives

This work aims to study the theoretical and practical aspects of the one parameter and two parameters Lindley models . The first part of this work is dedicated to a review of the Lindley model. We presented the form of the cumulative distribution function, the probability density function, the survival function, and the hazard function. The method for simulating the random variable X with distribution was also presented.

In the second part, our objective was to present in detail the methods for estimating the parameters of the models. Firstly, we reviewed the principle of the maximum likelihood method for estimating this model. The obtim function in R was used to estimate the parameters for the Lindley models with two parameters because the problem does not have an exact solution. Secondly, we presented the Bayesian approach for the one parameter Lindley distribution.

Next, we focused on applications to simulated and real data. The aim of this part is to apply the Lindley models by estimating his parameters using both the maximum likelihood method and the Bayesian approach. A comparison in terms of the mean squared error between the two methods (maximum likelihood and Bayesian) was conducted.

Annex

Source Code with R

The computer programs were implemented in R using the fitdistrplus and MASS packages. First, we present the programs developed for the calculation of the cumulative distribution functions and density functions. Next, we introduce the method to simulate a random variable X with the $Lindley(\theta)$ distribution. Finally, we provide the R programs that simulate X by Metropolis-Hasting and return the maximum likelihood estimators and Bayesian estimators for the parameters θ.

==

Program 1: Calculation of the functions (cumulative distribution, density)

==

```
plindley   <-  function(x, theta) {
1 -(((1 +theta+ theta * x)* exp(-theta * x))/(theta+1))
}
DLindley = function(x, theta)
{return((theta ^2 / (1+theta)) * exp(-theta * x) * (1 +x))}
```

==

Program 2: Simulation of one-parameter Lindley distribution using the method of the fact that it is a mixture of $\exp(\theta)$ and $\text{gamma}(2, \theta)$

==

```
n <- 10000
theta <- 2
p <- (theta )/ (theta + 1)
Ui <- runif(n, min=0, max=1)
Vi <- rexp(n, rate = theta)
Wi <- rgamma(n, shape = 2, rate = theta)
Xi <- ifelse(Ui <= p, Vi, Wi)
```

==

Program 3: Simulation and estimation of θ by MLE and Bayesian method
==

```
estimer-theta <- function(mean)    {
((1 - mean) + sqrt((mean - 1)\^ \ 2 + 8 * mean)) / (2 * mean)
}
estimer-theta-lin <- function(theta-hat, a = -1, c = 1, n)
{term1 <- exp(-a * theta-hat) +
0.5 * (a * exp(-a * theta-hat)
*(a + (4 * c/theta-hat))*(theta-hat^2
*(1 +theta-hat)^2 / (2*n *(1 +theta-hat)^2-n *theta-hat^2)))
term2 <- 0.5*(a*exp(-a*theta-hat)*((4*n / theta-hat^3)
-(2 * n / (1 + theta-hat)^ 3))
* ((theta-hat ^  2 * (1 + theta-hat) ^2) /
(2 * n * (1 + theta-hat) ^ 2 - n * theta-hat ^ 2)) ^2)
theta-lin <- -(1 / a ) * log(term1 - term2)
return(theta-lin)
}
estimer-theta-lingam <- function(theta-hat, a = -1, alpha = 1,
beta = 1.5, n)
{term3 <- exp(-a * theta-hat)
term4 <- 0.5 * a * (a - 2) * exp(-a * theta-hat) *
((alpha / theta-hat^2)-((beta + 1) / theta-hat))
* (theta-hat ^ 2 * (1 + theta-hat) ^2) /
(2 * n * (1 + theta-hat)^ 2 - n * theta-hat ^2)
term5 <- 0.5 * a * exp(-a * theta-hat) * ((4 * n / theta-hat^3)
-(2 * n / (1 + theta-hat)^3))*((theta-hat^2 * (1 + theta-hat)^2)
/(2 * n * (1 + theta-hat) ^2 - n * theta-hat ^ 2))^2
theta-lingam <- -(1/a)*(log( term3 + term4 -term5))
return(theta-lingam)
}
lambda <- 0.1
theta <- 0.1
f <- function(x) {
(theta ^  2 / (1 + theta)) * exp(-theta * x) * (1 + x)
}
N_{sim} <- 10000
```

```
sample-size <- 500
theta-estimations <- numeric(N_{sim})
theta-estimations-lin <- numeric(N_{sim})
theta-estimations-lingam <- numeric(N_{sim})
generate-sample-and-estimate-theta <-
function(sample-size, lambda, f)
{ x <- numeric(sample-size)
x[1] <- 1
for  (i \ in 2:sample-size) {
y <- rexp(1, rate = lambda)
w <- lambda * exp(-lambda * x[i - 1])
v <- lambda * exp(-lambda * y)
rho <- min(1, (f(y) * w) / (f(x[i - 1]) * v))
u <- runif(1)
if (u < rho) {
x[i] <- y
} else {
x[i] <- x[i - 1]
 }
 }
mean-empirique <- mean(x)
theta-hat <- estimer-theta(mean-empirique)
theta-lin <- estimer-theta-lin (theta-hat, n = sample-size)
theta-lingam <- estimer-theta-lingam (theta-hat, n=sample-size)
return(c(theta-hat, theta-lin, theta-lingam))
}
for (i  in 1:N_{sim}){
thetas <- generate-sample-and-estimate-theta
(sample-size, lambda, f)
theta-estimations[i] <- thetas[1]
theta-estimations-lin[i] <- thetas[2]
theta-estimations-lingam[i] <- thetas[3]
}
mse <- mean (((theta-estimations) - theta )^2 )
mse-lin <- mean((theta-estimations-lin) -theta)^2)
mse-lingam<- mean((theta-estimations)-lingam-theta)^2))
bias-mle<- mean((theta-estimations )-theta)
bias-jef<- mean(theta-estimations-lin -theta)
bias-ingam <- mean((theta-estimations-lingam - theta))
```

```
print(mse)
print(bias-mle)
print(mse-lin)
print(bias-jef)
print(mse-lingam)
print(bias-ingam)
mean-theta-estimations <- mean(theta-estimations)
mean-theta-estimations-lin <- mean(theta-estimations-lin)
mean-theta-estimations-lingam <- mean(theta-estimations-lingam)
print(mean-theta-estimations)
print(mean-theta-estimations-lin)
print(mean-theta-estimations-lingam)
```

Bibliography

[1] Sajid Ali, Muhammad Aslam, and Syed Mohsin Ali Kazmi. A study of the effect of the loss function on bayes estimate, posterior risk and hazard function for lindley distribution. *Applied Mathematical Modelling*, 37(8): 6068–6078, 2013.

[2] Mohammed MA Almazah and Muhammad Ismail. Selection of efficient parameter estimation method for two-parameter weibull distribution. *Mathematical Problems in Engineering*, 2021(1):5806068, 2021.

[3] A Asgharzadeha, Hassan S Bakouch, S Nadarajahd, and F Sharafia. A new weighted lindley distribution with application. *Brazilian Journal of Probability and Statistics*, pages 1–27, 2016.

[4] Hassan S. Bakouch, Sanku Dey, Pedro Luiz Ramos, and Francisco Louzada. Binomial-exponential 2 distribution: Different estimation methods and weather applications. *Trends in Applied and Computational Mathematics*, 18:233–251, 2017. URL https://api.semanticscholar.org/CorpusID:125599260.

[5] Narayanaswamy Balakrishnan, Markos V Koutras, and Konstadinos G Politis. *Introduction to probability: models and applications*. John Wiley & Sons, 2019.

[6] Tor Bjerkedal et al. Acquisition of resistance in guinea pies infected with different doses of virulent tubercle bacilli. *American Journal of Hygiene*, 72 (1):130–48, 1960.

[7] George Casella and Roger Berger. *Statistical inference*. CRC Press, 2024.

[8] Christophe Chesneau, Lishamol Tomy, and Jiju Gillariose. A new modified lindley distribution with properties and applications. *Journal of Statistics and Management Systems*, 24(7):1383–1403, 2021.

[9] Mohamed E Ghitany, Barbra Atieh, and Saralees Nadarajah. Lindley distribution and its application. *Mathematics and computers in simulation*, 78 (4):493–506, 2008.

[10] Mohamed E Ghitany, F Alqallaf, Dhaifalla K Al-Mutairi, and HA Husain. A two-parameter weighted lindley distribution and its applications to survival data. *Mathematics and Computers in simulation*, 81(6):1190–1201, 2011.

[11] Rameshwar D Gupta and Debasis Kundu. Exponentiated exponential family: an alternative to gamma and weibull distributions. *Biometrical Journal: Journal of Mathematical Methods in Biosciences*, 43(1):117–130, 2001.

[12] Duha Hamed and Ahmad Alzaghal. New class of lindley distributions: properties and applications. *Journal of Statistical Distributions and Applications*, 8:1–22, 2021.

[13] Jean Jacod and Philip Protter. *Probability essentials*. Springer Science & Business Media, 2004.

[14] Sadasivan Karuppusamy, Vinoth Balakrishnan, and Keerthana Sadasivan. Modified one-parameter lindley distribution and its applications. *Int. J. Eng. Res. Appl*, 8(1):50–56, 2018.

[15] Debasis Kundu and Rameshwar D Gupta. Generalized exponential distribution: Bayesian estimations. *Computational Statistics & Data Analysis*, 52 (4):1873–1883, 2008.

[16] Fetima Ladjimi. *Contribution à l'étude de l'intégrale stochastique et du mouvement brownien*. PhD thesis, ummto, 2011.

[17] Erich L Lehmann and George Casella. *Theory of point estimation*. Springer Science & Business Media, 2006.

[18] Dennis V Lindley. Fiducial distributions and bayes' theorem. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 102–107, 1958.

[19] Farouk Metiri, Halim zeghdoudi, and Mohamed Riad Remita. On bayes estimates of lindley distribution under linex loss function : Informative and non informative priors. 2016. URL https://api.semanticscholar.org/CorpusID:46201692.

[20] Sihem Nedjar and Halim Zeghdoudi. On gamma lindley distribution: Properties and simulations. *Journal of Computational and Applied Mathematics*, 298:167–174, 2016.

[21] Marcel Ortgiese, Ellen Baake, and Hans-Otto Georgii. *Stochastics: introduction to probability and statistics*. Walter de Gruyter, 2008.

[22] PL Ramos and F Louzada. The generalized weighted lindley distribution: Properties, estimation, and applications. *Cogent Mathematics*, 3(1):1256022, 2016.

[23] Sheldon M Ross. *A first course in probability*, volume 2. Macmillan New York, 1976.

[24] Sheldon M Ross. *Simulation*. academic press, 2022.

[25] Subhradev Sen, Hazem Al-Mofleh, and Sudhansu S Maiti. On discrimination between the lindley and xgamma distributions. *Annals of Data Science*, 8:559–575, 2021.

[26] R Shanker, H Fesshaye, and S Sharma. On two parameter lindley distribution and its applications to model lifetime data. *Biometrics & Biostatistics International Journal*, 3(1):1–8, 2016.

[27] Rama Shanker, Shambhu Sharma, and Ravi Shanker. A two-parameter lindley distribution for modeling waiting and survival times data. 2013.

[28] VS Vaidyanathan and A Sharon Varghese. Discriminating between exponential and lindley distributions. *Journal of Statistical Theory and Applications*, 18(3):295–302, 2019.

Abstract

Probability distributions are commonly applied to describe real phenomena and are often used in several fields such as economics, biology and medicine. There are several probability distributions. In this thesis we study the Lindley distribution with one parameter, Lindley distribution with two parameters as well as the two-parameter weighted Lindley distribution. We also explore the fundamental properties of the Lindley distribution, including its probability density function, cumulative distribution function, survival and hazard rate function, moments, characteristic function, moment generating function, etc and explores parameter estimation methods, including method of moments, maximum likelihood and Bayesian approaches as well. Finally applications of Lindley distribution to different real data sets and simulated data.

Keywords: Random variable, probability distribution, survival function, Lindley distribution, maximum likelihood estimation, Bayesian estimation, simulation.

Résumé

Les distributions de probabilité sont couramment appliquées pour décrire des phénomènes réels et sont souvent utilisées dans plusieurs domaines tels que l'économie, biologie et la médecine. Il existe plusieurs distributions de probabilité. Dans ce mémoire, nous étudions la distribution de Lindley à un paramètre et la distribution de Lindley à deux paramètres, ainsi que la distribution de Lindley pondérée à deux paramètres. Nous explorons également les propriétés fondamentales de la distribution de Lindley, y compris sa fonction de densité de probabilité, sa fonction de répartition, sa fonction de taux de survie et de risque, moments, fonction caractéristique, fonction génératrice des moments, etc. Et voyons ensuite les méthodes d'estimation des paramètres, notamment la méthode des moments, maximum de vraisemblance et les approches Bayésiennes. Enfin, des applications de la distribution de Lindley à différents ensembles de données réels et simulés.

Mots-Clefs: Variable aléatoire, distribution de probabilité, fonction de survie, distribution de Lindley , estimation du maximum de vraisemblance, estimation Bayésienne, simulation.