وزارة التعليم العالي والبحث العلمي

*Ministry of Higher Education and Scientific Research*


**BADJI MOKHTAR-ANNABA UNIVERSITY**

**UNIVERSITE BADJI MOKHTAR ANNABA**

جامعة باجي مختار

- عنابة-

**Faculty of Sciences**

**Department of Mathematics**


**Year: 2024/2025**

# THESIS

Presented with a view to obtaining the doctorate degree

## Bayesian Statistical Studies In The Presence Of Censored Data

**Stream**

**Applied Mathematics**

**Speciality**

**Probability and Statistics**

**By**


**BENATALLAH ISSRA NADA**

| | | |
|---|---|---|
| **SUPERVISOR: TALHI HAMIDA** | **MCA.** | **U.B.M.- Annaba** |
| **CO-SUPERVISOR: GOUAL HAFIDA** | **MCA.** | **U.B.M.- Annaba** |

**In front of the jury**

| | | |
|---|---|---|
| **PRESIDENT: CHADLI ASSIA** | **Prof.** | **U.B.M.- Annaba** |
| **EXAMINER: LEULMI SARRA** | **MCA.** | **U.F.M.C- Constantine 1** |
| **EXAMINER: MEZHOUD KENZA ASSIA** | **MCA.** | **U.F.M.C- Constantine 1** |

وزارة التعليم العالي والبحث العلمي

*Ministère de l'Enseignement Supérieur et de la Recherche Scientifique*

**BADJI MOKHTAR-ANNABA**

**UNIVERSITY**

**UNIVERSITE BADJI MOKHTAR ANNABA**

جامعة باجي مختار

-عنابة-

**Faculté des Sciences**

**Département de Mathématiques**

**Année : 2024/2025**

# THÈSE

Présentée en vue de l'obtention du diplôme de Doctorat

## Études statistiques bayésiennes en présence de données censurée

**Filière**
**Mathématiques Appliquées**

**Spécialité**
**Probabilités et Statistique**
**Par**

**BENATALLAH ISSRA NADA**

**DIRECTEUR DE THÈSE: TALHI HAMIDA** **MCA. U.B.M. - Annaba**

**Devant le jury**

**PRESIDENT: CHADLI ASSIA** **Prof. U.B.M. - Annaba**

**EXAMINATEUR** : **LEULMI SARRA** **MCA. U.F.M.C- Constantine 1**

**EXAMINATEUR** : **MEZHOUD KENZA ASSIA** **MCA. U.F.M.C- Constantine 1**

# Dedication

All praise is due to Allah, who granted me strength, patience, and perseverance to reach this moment. I am forever grateful for His countless blessings.

To myself, for believing in my dreams despite all difficulties, for standing firm in the face of challenges—this achievement is the result of my hard work and determination.

To my beloved mother, who stayed up countless nights raising and educating me, who planted this dream in my heart since childhood, and who has been my guiding light every step of the way.

To my dear father and my cherished brothers, Riad, Moin, and Anas, thank you for your unwavering support and encouragement, which have always been a source of motivation for me.

To all my dear friends, with special gratitude to those who have been my pillars of strength throughout my academic journey:

Hadeer, the first to share my joy of success at 2 AM and who stood by my side for three years.

Djihane, who celebrated every achievement of mine as if it were her own.

Sabrina, who always took pride in me in every gathering and prayed for my success in every prayer.

A heartfelt thank you to everyone who encouraged and supported me, and to all my friends at Badji Mokhtar University – Annaba and Brothers

Mentouri University – Constantine. Your presence has been a beautiful part of this journey.

This achievement is not mine alone; it is the result of the unwavering belief and support of those who stood by my side.

# ACKNOWLEDGMENTS

I am very grateful to my thesis supervisor, Ms. Talhi Hamida, and my co-supervisor, Ms. Goual Hafida, at Badji Mokhtar University in Annaba, and members of the LaPS laboratory, for placing their trust in me by proposing this thesis. I sincerely thank them for their availability, kindness, and wise advice.

I would like to express my sincere gratitude to Ms. Chadli Assia, Professor at Badji Mokhtar University in Annaba, for agreeing to review my work and chair my thesis jury. I am very honored by the presence on my thesis jury, as examiners, of:

- Sarra Leulmi, Dr.UFMC. Constantine.

- Kenza Assia Mezhoud, Dr. UFMC. Constantine.

May they find here the expression of my deep gratitude. Finally, I would also like to thank all my professors, colleagues, and friends at the Faculty of Mathematics at UBM, the members of the Laps laboratory, and the Faculty of Sciences at the University of Constantine

# ABSTRACT

This thesis considers Bayesian inference for some recent survival time models in the presence of censoring. The Xlindley model, Zeghdoudi model, and Xexponential model are the intriguing models. With type II right-censored data, the parameters for these three models were calculated using a Bayesian technique under various loss functions, both symmetric (quadratic loss) and asymmetric (LINEX loss and entropy) and balanced. Numerical approaches were used in the classical approach, where the estimators are solutions of a nonlinear system whose solutions are not analytically explicit.

Estimators are provided as a ratio of integrals in the Bayesian technique. Simulations and data analysis were conducted using Markov Chain Monte Carlo (MCMC) techniques, namely the Metropolis-Hastings algorithm, to demonstrate the outcomes.

Finally, we used the Pitman criterion and the integrated mean square error (IMSE) criterion to compare Bayesian estimators and maximum likelihood estimators.

Keywords and expressions: The X-exponential distribution, the X-lindley distribution, the Zeghdoudi distribution, transcription, type II censored data, maximum likelihood estimation, Bayesian estimations, MCMC methods

# RÉSUMÉ

On considère dans cette thèse, l'inférence Bayésienne pour certains modèles récents de durées de survie en présence de censures. Les modèles auquel on s'est intéressé sont le modèle de Xlindley , le modèle de Zeghdoudi et enfin le modèle Xexponentiel , Pour ces trois modèles, on a procédé à l'estimation des paramètres, à l'aide d'une approche Bayésienne sous différentes fonctions de perte aussi bien symétriques (perte quadratique) puis asymétriques (perte LINEX et entropie) et Balanced , avec des données censurées à droite de type II . Dans le cas de l'approche classique, les estimateurs sont solutions d'un système non linéaire dont les solutions ne sont pas explicites analytiquement ; des méthodes numériques ont été adoptées.

Dans l'approche Bayésienne, les estimateurs sont donnés sous forme d'un rapport d'intégrales, les méthodes de Monte-Carlo par Chaine de Markov (MCMC) et en particulier l'algorithme Metropolis -Hastings ont été utilisées pour procéder à des simulations et à une analyse de données permettant d'illustrer les résultats obtenus.

Finalement, Nous avons utilisé le critère de Pitman et le critère de erreur quadratique moyenne intégrée (IMSE) pour comparer les estimateurs bayésiens et les estimateurs du maximum de vraisemblance.

Mots-clés et expressions : La distribution de Xexponentiel, La distribution de Xlindley , la distribution de Zeghdoudi, trancature, données censurées de type II, estimation de maximum de vraisemblance, estimations bayésiennes,

méthodes MCMC.

# الملخص

في هذه الأطروحة، ندرس الاستدلال البايزي لبعض النماذج الحديثة لأوقات البقاء في وجود الرقابة .
كانت النماذج التي اهتممنا بها هي نموذجXlindley ، ونموذجZeghdoudi ، وأخيرًا نموذج X الأسّي. بالنسبة لهذه النماذج الثلاثة، شرعنا في تقدير المعاملات باستخدام نهج بايزي باستخدام دوال خسارة مختلفة، متماثلة (خسارة تربيعية) وغير متماثلة (خسارة LINEX وانتروبيا) ومتوازنة، مع بيانات من النوع الثاني مُراقَبة يمينًا .

في حالة النهج الكلاسيكي، فإن المقدرين هم حلول لنظام غير خطي لا تكون حلوله واضحة تحليليًا؛ تم اعتماد الأساليب العددية

.في النهج البايزي، يتم إعطاء المقدرين في شكل نسبة من التكاملات، وتم استخدام طرق سلسلة ماركوف مونت كارلو (MCMC) وعلى وجه الخصوص خوارزمية متروبوليس- هاستينجز لإجراء عمليات المحاكاة وتحليل البيانات لتوضيح النتائج التي تم الحصول عليها .

وأخيرًا، استخدمنا معيار بيتمان ومعيار خطأ المتوسط المربع المتكامل (IMSE) لمقارنة مقدرين بايز ومقدري الاحتمالية القصوى .

الكلمات المفتاحية والتعابير: التوزيع الأسيX ، توزيعXlindley ، توزيعZeghdoudi ، النسخ، البيانات الخاضعة للرقابة من النوع الثاني، تقدير الاحتمالية القصوى، التقديرات البايزية، طرق MCMC.

# Contents

# General Introduction

Modeling in survival analysis is a very important aspect in many diverse and diverse fields, such as industry, medicine, finance, etc. After modelling, the statistical study of "lifespan" or "reliability" has gained significant attention from statisticians.

The primary issues in this discipline are the estimate of parameters or other dependent quantities within a parametric framework, and the modelling of survival-related phenomena using random distributions.However, direct estimation of the distribution might be taken into consideration as a nonparametric or semiparametric approach. Data that is complete or right-censored is frequently used to draw conclusions. The example where the data are gradually suppressed, as presented by Balakrishnan [49], is the main emphasis of this thesis.

A few distributions stand out among the models because of their proven applicability in many contexts. Among these are the Pareto distribution (Raqab et al., 2010), the Weibull distribution (Tse et al., 2003; El Sagheer, 2018), exponential models, and their various variations. The XLINDLEY (2021), Zeghdoudi (2018), and Xexponential models are the main subjects of this thesis.

The XLindley distribution (XLD) was introduced by Chouia and Zeghdoudi (2021)[30]. It is generated by a special mixture of two distributions: exponential and Lindley, hence its name. It also exhibits statistical properties such as stochastic order, quantile function, maximum likelihood method, and method of moments.The application of the model to a real data set is presented and compared to the fitting. It shows that the XLD is more flexible than other one-parameter distributions.

The Zeghdoudi distribution was introduced by Messaadia and Zeghdoudi (2018)

[36]. They suggested a family of one-parameter exponential distributions based on mixtures of gamma distributions.

the Xexponential model, which is generated by a special mixture of exponential and gamma models. The study explores the mathematical properties of this new model, including the moment generation function, moments, mode, and quantile function. The research evaluates various estimators to estimate the unknown parameter of the distribution using a Monte Carlo simulation. Also, the goodness-of-fit test is provied for this model. Additionally, the flexibility of the Xexponential model is compared to that of other commonly used models, such as gamma, exponential, Lindley, Shanker, Akash, Zeghdoudi, and Chris-Jerry, by applying the model to a real data set.

The results indicate that the proposed model provides more flexibility and a better fit than other models. These findings suggest that the Xexponential model could be useful in modeling real-life data and may warrant further exploration in future research. We use two approaches, the classical maximum likelihood approach and the Bayesian approach, to estimate the distribution parameters. The Bayesian estimators and the corresponding posterior risks (PR) were derived using the Balanced loss function, both symmetric (quadratic loss) and asymmetric (Linex and entropy). The estimators cannot be obtained explicitly, so we use the Monte Carlo method (MCMC) to approximate the solutions in the case of the Bayesian approach. In the case of the maximum likelihood approach, the estimators are solutions to nonlinear systems and were approximated using the EM algorithm. We also use the Pitman proximity criterion to compare the results of the two methods.

**Organization of the thesis**

Chapter one is devoted to the main mathematical tools used in this thesis. We cite the Bayesian approach and its various loss functions; survival analysis and its main characteristics; and finally, the different types of censoring.

In chapter two, we focused on parameter estimation for the truncated xlindley model. The data are assumed to be type II right-censored, the prior distributions on the parameters are gamma distributions, and this is the case under different loss functions

In Chapter 3, we perform a Bayesian analysis of the Zeghdoudi distribution using Type II censored data. Using two types of loss functions: balanced and unbalanced, we use three different loss functions. This estimation includes three cases of prior information.

Chapter four discusses a new probability model, the Xeponential model, by studying its mathematical and probabilistic properties, comparing it to other models, and demonstrating its flexibility using real-world data. We also estimate the parameters of this model using Bayesian estimation.

# Generalities and Mathematical Tools

This chapter deals with parametric statistics in a Bayesian framework. The statistics Bayesian is a rival theory to classical statistics in the sense that each of them proposes opposite, to the same problem; a different approach and resolution. We let's start by recalling some definitions and properties of the theoretical elements of classical analysis, then we present the notions and tools on which it is based Bayesian analysis.

## 1.1 Survival analysis

The objective of our work is the estimation in a censorship model, for this we begin in this chapter by introducing the survival data and the censoring with its different types as well as the assessment of survival function in a sensed model. Survival analysis is the modeling of the time of occurrence of a specific event, the event studied is the passage between two specific states. This time is called lifespan, survival time, or simply a duration that is a positive or zero variable, often assumed to be limited. Sometimes for some reasons it happens that the events of interest are not produced during the observation period, this phenomenon is called the censorship of the data series, in other words for some data you have only a lower or higher border and not the exact value. Although this theory has been developed in biomedical research, censored observations can be found in certain areas of research. For example, in medicine, survival analysis is used to evaluate the effectiveness of a treatment, i.e. to estimate the likely survival time. of a patient. For this, we take a sample of patients, we know for each of them :

- the exact value of survival time (complete uncensored data).
- a lower limit of this duration (censored data)

The latter occurs when a patient cannot be followed until the end of the study For a specific reason, for example, he left the country or when he died from a cause independent of the disease. In engineering, survival analysis is used to estimate the reliability of machines, electronic components, etc. In demography, survival analysis is used to construct mortality tables. These are used to determine the amount of life insurance, among others.

### 1.1.1 The 5 equivalent definitions of reliability

Five equivalent functions define the law of duration: Suppose the survival time X is a positive or zero variable, and absolutely continuous. So its law of probability can be defined by one of the following functions:

**1- The Survival Function (Reliability Function)**

By definition:

$$S(t) = P(X \geq t); t \geq 0$$

For fixed t, it is the probability of survival until time t.

**2- The cumulative distribution function F**

for t fixed, the probability of dying before the moment t; i.e.:

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

**3- The Probability density f**

This is a function $f(t) \geq 0$ , such that for any $t \geq 0$ ,

$$F(t) = \int_0^t f(s)ds$$

If the cumulative distribution function has a derivative at the point t then:

$$f(t) = \lim_{dt \to 0} \frac{P(t \leq X \leq t + dt)}{dt} = F'(t) = -S'(t)$$

For t fixed, probability density characterizes the probability of dying within a short time interval after the moment t

**4- Instant risk (the hazard rate) h**

Called "the hazard rate" is defined as:

$$h(t) = \lim_{dt \to 0} \frac{P(t \leq X \leq t + dt | X \geq t)}{dt} = \frac{f(t)}{S(t)}$$

For fixed t, $h(t)$ characterizes the probability of dying in a small time interval after time t, conditional on having survived until time t. Also this means the risk of instant death for those who survived.

**5- The cumulative hazard rate H**

This is the integral of the hazard rate h:

$$H(t) = \int_0^t h(v)dv = -\ln(S(t))$$

We can deduce the cumulative hazard rate function using the relation:

$$S(t) = \exp(-H(t)) = \exp\left(-\int_0^t h(v)dv\right)$$

The definition of the probability distribution of X is based on one of the following five data, which are equivalent:

$$S(t), F(t), f(t), h(t), H(t)$$

### 1.1.2 Quantities associated with survival distribution

#### Mean and variance of survival time

The Mean survival time $\mathrm{E}(X)$ and the variance of the survival time $V(X)$ are defined by the following quantities:

$$\mathrm{E}(X) = \int_{-\infty}^{+\infty} tf(t)dt = \int_0^{\infty} S(t)dt$$

$$V(X) = 2\int_0^{\infty} tS(t)dt - (\mathrm{E}(X))^2$$

So we can deduce hope and variance from any of the functions: S, F, f, h, and H (but not the other way around)

#### Quantiles of survival time

The median of the survival time is the time t for which the probability of survival $S(t)$ is equal to0.5 , that is to say, the $m$ value which satisfies $S(m) = 0.5$ In the case where the estimator is a step function, there may be a time interval verifying $S(m) = 0.5$. We

must then be careful in the interpretation, if both events framing the median time are distant. It is possible to obtain an interval median time confidence
Consider a $[C_i, C_s]$ confidence interval of level $\alpha$ of the time median $m$ is:

$$\left[S^{-1}(C_s), S^{-1}(C_i)\right]$$

- The quantile function of survival time is defined by:

$$q(p) = \inf\{t : F(t) \geq p, 0 < p < 1\}$$
$$= \inf\{t : S(t) \leq 1 - p\}$$

When the cumulative distribution function F is strictly increasing and continuous SO

$$q(p) = F^{-1}(p) = S^{-1}(p), 0 < p < 1$$

The quantile $q(p)$ is the time when a proposition p from the population has disappeared.

### 1.1.3 Censorship and truncation

One of the characteristics of survival data is the existence of incomplete observations. Indeed, data is often collected partially, in particular, because of the processes censorship and truncation. Censored or truncated data comes from the fact that we have not no access to all the information: instead of observing independent and identically distributed realizations (i.i.d) of duration various disturbances independent or not of the phenomenon studied.

#### censored data

Censorship is the most common phenomenon encountered during data collection survival:

For individual i, consider:

- Its survival time $X_i$.
- Its time of censorship $C_i$
- The duration actually observed $T_i$

**Right censorship**

The lifespan is said to be right censored if the individual did not experience the event at his last observation. In the presence of right censoring, not all lifetimes are observed; for some of them, we only know that they are greater than a certain value known.

**1- Type I censorship**

Let C be a fixed value, instead of observing the variables$X_1,...,X_n$ , we observe $X_i$ only when $X_i \leq C$, otherwise we only know that $X_i \geq C$. We use the following notation:

$$T_i = X_i \wedge C = \min(X_i, C)$$

This censorship mechanism is frequently encountered in industrial applications. By example, we can test the lifespan of n identical objects (light bulbs) over an interval of observations fixed $[0, u]$. In biology, we can test the effectiveness of a molecule on a batch of mice (the mice are sacrificed after a time)

**2- Type II censorship**

It is presented when we decide to observe the survival times of n patients until that k of them have died and to stop the study at that point. Let $X(i)$ and $T(i)$ the order statistics of variables $X_i$ and $T_i$ . The censorship date is therefore $X(k)$ and we observe the following variables:

$$\begin{cases} T_{(1)} = X_{(1)} \\ T_{(k)} = X_{(k)} \\ T_{(k+1)} = X_{(k)} \\ . \\ . \\ . \\ T_{(n)} = X_{(n)} \end{cases}$$

**3- Type III censorship (or random type I censorship)**

Let $C_1, ..., C_n$ be random variables i.i.d. We observe the variables

$$T_i = X_i \wedge C_i$$

The available information can be summarized by:

- The duration actually observed $T_i$.
- An individual $\delta_i = 1_{\{X_i \leq C_i\}}$
- The duration actually observed $T_i$.
- $\delta_i = 1$if the event is observed (hence $T_i = X_i$ ). We observe the "real" durations or the durations
- $\delta_i = 0$ if the individual is censored (hence$T_i = C_i$ ). We observe incomplete durations (censored).

Random censorship is the most common. For example, during a therapeutic trial she can be caused by:

(a)- Loss of sight: the patient leaves the current study and is not seen again (because of a move, the patient decides to seek treatment elsewhere). These are patients "lost to follow-up".

(b)- Stopping or changing treatment: side effects or ineffectiveness of treatment may lead to a change or stoppage of treatment. These patients are excluded from the study.

(c)- End of the study: the study ends while some patients are still alive (they did not experience the event). These are "excluded-living" patients. The "lost to sight" and the "living excluded" correspond to censored observations but the two mechanisms are of a different nature (censoring can be informative in those "lost to follow-up").

**Censorship on the left**

Left censoring corresponds to the case where the individual has already suffered the event before the individual is observed. We only know the date of the event less than a certain date known. For each individual, we can associate a pair of random variables $(T, \delta)$:

$$\begin{cases} T = X \vee C = \max(X, C) \\ \delta_i = 1_{\{X \geq C\}} \end{cases}$$

**Interval censorship**

A date is censored by interval, if instead of observing with certainty the time of the event, the only information available is that it took place between two known dates

**Troncature**

Truncations differ from censorships in the sense that they concern sampling even. Thus, a variable X is truncated by a possibly random subset A of R+ if instead of X, we observe X only if $X \in A$ The points of the sample "truncated" all belong to A, and therefore follow the law of T conditioned by belonging to A. We must not confuse censorship and truncation. If there is truncation, part individuals (therefore $X_i$) are not observable and we only study a sub-sample (sampling problem). Selection bias is a special case of truncation.

**1- Left truncation**

Let Z be a random variable independent of X is only observable if $X > Z$. We observe the couple $(X,Z)$, with $X > Z$.

**2- Right truncation**

Likewise, there is right truncation when X is only observable if $X < Z$.

**3- Truncation by interval**

When a duration is truncated on the right and left, we say that it is interval truncated.

### 1.1.4 Usual survival models

Below we only list the most common models; in a way In general, all distributions used to model positive variables (log-normal, Pareto, logistic, etc.) can be used in survival models, the basic distribution of parametric models of duration is the exponential distribution. The choice of model determines in particular the form of the function of chance ; we will notably distinguish models with a monotone chance function from models making it possible to obtain chance functions in the form of bathtub ; these latter models are little used in insurance, the situation of reference Being an increasing chance rate (in the broad sense) with time, The two most used simple laws: Exponential Law, Weibull Law, gamma law .

#### The exponential law

This law has numerous applications in several areas. It is a simple law, very used in reliability whose failure rate is constant. It describes the life of the materials which suffer sudden failures.

The probability density of an exponential law with parameter $\lambda$ is written:

$$f(t) = \lambda e^{-\lambda t}$$

The reliability function:

$$S(t) = e^{-\lambda t}$$

The failure rate is constant over time :

$$h(t) = \lambda$$

**Memoryless properties of the exponential law:**

A main property of the exponential law is to be without memory or "Memory less property" ((Bon, 1995), (Leemis, 1994)):

$$P(X \geq t + \Delta t | X \geq t) = \frac{e^{-\lambda(t+\Delta t)}}{e^{-\lambda t}} = e^{-\lambda \Delta t} = P(X \geq \Delta t), t \geq 0, \Delta \geq 0$$

This result shows that the conditional law of the lifespan of a device that has worked without breaking down until time t is identical to the law of the lifespan of a new device is considered new (or "as good as new" ), with a duration of exponential life of parameter $\lambda$.

### Weibull's law

It is the most popular law used in several fields (electronics, mechanics, etc.). In particular, it makes it possible to model numerous equipment wear situations. It characterizes the behavior of the system in the three phases of life, period of youth, useful life period and wear period or aging depends on the three parameters following: $\beta$, $\eta$and $\gamma$ .

The probability density of a Weibull distribution has the expression:

$$f(t) = \frac{\beta}{\eta}\left(\frac{t-\gamma}{\eta}\right)^{\beta-1} e^{-\left(\frac{t-\gamma}{\eta}\right)}; t \geq \gamma$$

Or :

- $\beta$ is the shape parameter $\beta > 0$.
- $\eta$ is the scale parameter $\eta \geq 0$ .
- $\gamma$is the position parameter $\gamma \geq 0$

The reliability function is written:

$$S(t) = e^{-\left(\frac{t-\gamma}{\eta}\right)^{\beta}}$$

the hazard rate is given by:

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

According to the values of $\beta$ , the failure rate is either decreasing $\beta < 1$ or constant $\beta = 1$, or rising $\beta > 1$

The Weibull distribution therefore makes it possible to represent the three periods of the life of a device described by the curve in bathtub. The case $\gamma > 0$ corresponds to devices whose probability of failure is zero until a certain age $\gamma$

**The gamma law**

The gamma law is the law of the instant of occurrence of $\alpha^{me}$ event in a process of Poison.

Let $T_i$ the vector representing the inter-event durations (the times between failures successive of a system). If these durations are independent random variables and identically distributed according to an exponential law with parameter $\beta$, then the cumulative rate of appearance of $\alpha$failures follows a gamma law with parameters $(\alpha, \beta)$. Its probability density is written

$$f(t) = \frac{\beta^{\alpha} t^{\alpha-1} e^{-\beta t}}{\Gamma(\alpha)} t \geq 0, \alpha \geq 1, \beta \geq 0$$

the hazard rate is given by:

$$h(t) = \frac{\beta^{\alpha} t^{\alpha-1} e^{-\beta t}}{\int_t^{\infty} \Gamma(\alpha) f(u) du}$$

The gamma law is widely used in the Bayesian approach, it is the natural conjugate of the exponential law with parameter $\lambda$

## 1.2 classic statistics

Descriptive and inferential statistics, as well as exploratory statistics, are the main areas of statistics. Descriptive statistics provides tools to describe a sample. Starting from the sample, inferential statistics can be used to give an indication of the population. One of the main areas of statistics is to provide an indication of a population. In most cases, it is not possible to obtain all the data from the population, so a sample is taken. This sample can then be described using descriptive statistics, for example, by reporting the mean value and dispersion of the sample.

But this does not yet give information about the population, this is the task of inferential statistics. Inferential statistics takes a sample of the population in order to make inferences about the population from that sample. Thus, the goal of inferential statistics is to infer the unknown parameters of the population from the known parameters of a sample.

Therefore, inferential statistics attempts to infer conclusions that go beyond the immediate data, unlike descriptive statistics. To do this, hypothesis tests such as the t-test or analysis of variance are used in inferential statistics.

**The population**

for a statistician, is the almost exhaustive set of individuals having something in common allowing one to define membership in the population and for which we study a or several variables (for example: the size of the French adults) .

**The sample**

for a statistician, is a subset of the population studied for which we carry out a series of measurements on the variable(s) studied .


### 1.2.1 theory of sampling

In statistics, sampling is important for a number of reasons. First, it makes data collection possible in situations where a population's entire size is either unreasonably large or completely unreachable. Second, it offers substantial time and cost reductions without sacrificing the accuracy of the data. It is possible to perform statistical analysis more effectively and with fewer errors by concentrating on a representative sample. Sampling is also quite useful for forecasting future patterns and behaviours. Statisticians can draw conclusions about a wider population from a sample's characteristics by using appropriate sampling strategies. In many disciplines, including political science, public health, and marketing, where knowledge of huge populations is crucial, this extrapolation is basic.

The study of sampling theory in statistics aims to choose a subset of a population that fairly represents the entire population. It is essential for conducting tests or surveys for which it is impractical or even impossible to observe the entire population. By using sampling theory, conclusions can be drawn from samples that are far smaller than the entire population while yet yielding insightful and trustworthy information.

How Sampling Theory Operates ?

Sampling theory essentially suggests methods for choosing subsets of a population so that the resulting data precisely represents the attributes of the complete population. This method, which enables researchers to make significant findings without having to survey the entire population, is fundamental to statistical analysis in many different domains. The efficacy of sampling is dependent on a number of methods and ideas, each of which is adapted to certain situations and study goals.

**Basic Principles of Sampling Theory**

The following fundamental ideas form the basis of sampling theory and guarantee the efficacy of sampling as a technique for gathering and analyzing data:

- Representativeness: The sample needs to represent the population in small proportion, incorporating a range of the set's features.
- Randomness: To prevent prejudice, every member of the population must have an equal probability of being chosen.
- Sample size: Enough samples should be taken to adequately represent faithfully the populace. The rule of diminishing returns holds true even if results from larger samples are typically more reliable.

Precise implementation of these guidelines guarantees the correctness and dependability of the statistical inferences made in addition to facilitating effective data collecting. Using the right sampling strategies reduces bias and provides an accurate representation of the population studied.

**Technique of Sampling Theory**

**Empirical methods:**

The majority of procedures used by polling companies are empirical. Their success is dependent on the investigators' level of skill and their accuracy cannot be estimated.

- Judgment-based sampling: Selection of a sample based on the advice of specialists who have a thorough understanding of the population and can identify representative entities.

Pbme: Expert judgment is arbitrary.

- Sampling using the quota method: Samples are taken at will as long as they adhere to predetermined composition (e.g., sex, age, CSP, etc.).

**Random methods:**

The random sampling of samples and the computation of probabilities are the foundations of random methods.

- Simple random sampling: People are selected at random from the population, and they are sampled independently of one another with no replacement and an equal chance of being selected.
- The stratified random sampling technique operates on the assumption that the population is composed of homogeneous subpopulations known as strata, such as age groups. A straightforward random sampling is conducted in each stratum, with the size of the sample being proportionate to the stratum's size in the population (representative sample). Not every member of the population has the same likelihood to be dismissed. requires that the strata be homogeneous. improves the accuracy of approximations.
- In cluster sampling, groups or families of people are selected at random, and each member of the group is investigated (for example, we choose a building and interview every occupant). When clusters are similar to one another and individuals within a cluster differ from strata, the approach works even better.

### 1.2.2 Classical estimation methods

In this section, the data $X_1, ..., X_n$ are assumed to be n independent realizations of the same underlying random variable X. It is equivalent to assuming $X_1, ..., X_n$ independent and of the same law. Here we will adopt the second formula, which is the most convenient to manipulate.

We frequently seek to determine the value of a quantity that enables us to define a statistical population for a certain variable. If size is the variable under study, for instance, we may try to describe the University of Burgundy student body by average height (the "order of magnitude" of the size) and by variance (a measure of the variability of the size). Parameters in the population are its variance and mean size. We are interested in these values. However, we don't usually know them. We use the size's computed mean and variance, which come from a sample, to approximate these values. Estimates are these values that are computed using the sample data.

There are two types of estimates:

- point estimates
- interval estimates

**(A) Estimation by confidence interval**

We can qualify our point estimate by computing the likelihood that the mean lies within a given interval around our estimate using the sampling distribution of the mean.

**The confidence interval**

- confidence interval: interval calculated from an estimator in order to quantify the precision of the estimate.
- Significance threshold $\alpha$: probability set for a situation that an erroneous result is obtained by chance. In the case of a confidence interval, it is the probability of obtaining a sample with a very low or very high mean.
- Confidence level$1 - \alpha$: this is the complement of the significance threshold. With$\alpha = 5$ , by calculating the confidence intervals of all possible samples, 95 of these would contain the parameter to be estimated.
- Margin of error: half of the confidence interval.
- Standard error: standard deviation of the estimator

**Analysis**

- There is no 95 probability that a confidence interval containing the true parameter exists at a 95 confidence level. The interval is not random, and the real parameter is fixed.
- The parameter will be present in 95 of the computed intervals, and there is a 95 chance of selecting one of the "good" ones.
- It is also known that there was a low probability (less than $\alpha$) of achieving this interval if the estimate was poor (the parameter was outside the interval).
- Ninety-five percent of the surveys—let's say 100—that we conduct will genuinely include the parameter.

**(B) Point estimate**

Finding an approximate value for a parameter based on sample findings is known as parameter estimation. The point estimate of a parameter is the value that is obtained from sample results when it is estimated by a single number. The estimator, a random variable of the sample, is used to calculate the point estimate. The random variable's estimate is the value it takes in the sample that was observed.

**PROPERTIES OF POINT ESTIMATORS**

We want point estimators to have specific qualities when we utilize them regularly. These characteristics are crucial for selecting the associated parameter's best estimator, or the one that approaches the parameter to be estimated as closely as feasible. Multiple estimators can be used for an unknown parameter. For instance, we may use the arithmetic mean of the median or the mode to estimate the parameter m, or mean, of a population. The following lists the characteristics an estimator has to have in order to produce accurate estimates.

We will note :

- $\theta$ the parameter of unknown value .
- $\overset{\wedge}{\theta}$ the estimator of$\theta$

**(1) Unbiased estimator**

An estimator $\overset{\wedge}{\theta}$ of unknown parameter $\theta$ is unbiased if the mean of its sampling distribution is equal to the value $\theta$ of the population parameter to be estimated, that is to say if :

$$\mathrm{E}\left(\overset{\wedge}{\theta}\right) = \theta$$

If the estimator is biased, its bias is measured by the following difference:

$$BIAS = \mathrm{E}\left(\overset{\wedge}{\theta}\right) - \theta$$

**(2) Efficient estimator**

An unbiased estimator is efficient if its variance is the lowest among the variances of other unbiased estimators. More precisely, an unbiased estimator is said to be efficient if its variance reaches the minimum possible value, known as the *Cramér–Rao lower bound (CRLB)*:

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

where $I(\theta)$ denotes the Fisher information.

So, if $\hat{\theta}_1$ and $\hat{\theta}_2$ are two unbiased estimators of the parameter $\theta$, the estimator $\hat{\theta}_1$ is efficient if:

$$\begin{cases} V(\hat{\theta}_1) \leq V(\hat{\theta}_2), \\ Var(\hat{\theta}_1) = \frac{1}{I(\theta)}, \\ E(\hat{\theta}_1) = E(\hat{\theta}_2) = \theta. \end{cases}$$

**(3) Convergent estimator**

An estimator $\overset{\wedge}{\theta}$ is convergent if its distribution tends to concentrate around of the unknown value to be estimated, $\theta$, as the sample size increases, that is to say if :

$$\lim_{n\to\infty} V\left(\overset{\wedge}{\theta}\right) = 0$$

**Note :** An unbiased and convergent estimator is said to be absolutely correct

- The primary characteristics we search for in an estimator are these three attributes. We won't be strict about the estimators' mathematical characteristics.

There are various techniques for estimating points, including the greatest the method of moments, the method of least squares , and likelihood method . The latter is what we'll showcase:

### 1.2.3 Maximum likelihood estimation

In this part, we will explain the maximum likelihood method which is considered most commonly used to find an estimate of an unknown parameter $\theta$ from a set of given samples and in two cases . they have good convergence and efficiency properties when the number of samples is very large.

**(1) Maximum likelihood method in the case complete**

**The likelihood function**

We call the likelihood function of the n-sample, the density joint probability of the n random variables$X_1,...,X_n$ ; $L(x_1,...,x_n,\theta)$ , we considered as a function of $\theta$ only, we note by$L(x,\theta)$ , assuming that $X_1,...,X_n$ is independent and of law $f(x,\theta)$ , then

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

The maximum likelihood method is based on finding the values $\overset{\wedge}{\theta}$ which maximize the likelihood $L(x,\theta)$ :

$$\overset{\wedge}{\theta}=\arg\underset{\theta\in\Theta}{Max}\,L(x,\theta)$$

then we refer to $\overset{\wedge}{\theta}$ as the likelihood estimator (EMV) and observe$\overset{\wedge}{\theta}_{MV}$ .

The advantage of using the logarithm function's monotonicity is that it enables us to identify the maximum $\overset{\wedge}{\theta}_{MV}$ when maximizing the logarithm of the likelihood function, also known as the log-likelihood function, denoted $l(\theta)$.

This allows us to calculate the sum rather than the product:

$$l(\theta)=\log L(x,\theta)=\sum_{i=1}^{n}\log(f(x_i,\theta))$$

When the parameter $\theta$ has a finite dimension of k, meaning that $\theta=(\theta_1,\theta_2,...,\theta_k)^t$ , the solution to $\nabla_\theta=0$ is the estimator $\overset{\wedge}{\theta}_{MV}$ , where $\nabla_\theta$ represents the gradient operator of $l(\theta)$ :

$$\nabla_\theta = \begin{pmatrix} \frac{\partial l(\theta)}{\partial \theta_1} \\ \frac{\partial l(\theta)}{\partial \theta_2} \\ . \\ . \\ . \\ \frac{\partial l(\theta)}{\partial \theta_k} \end{pmatrix}$$

To make sure that each solution is of a true maximum, it must be verified, or the nature of the obtained optimum must be confirmed by calculating the second derivatives.

(2) **Maximum likelihood estimator with censoring**

The goal of this subsection is to find reasonable estimates of reliability, which depend on the parameters of the distribution under study. Since the true value of this parameter is typically unknown in practical situations, we will attempt to use the maximum likelihood method to estimate the parameter's value for the various types of right-censored data.

The likelihood function is computed using the following formula when we have right-censored data:

**Type I censorship:**

censorship that is fixed Considering a fixed $C > 0$ and a sample of survival times $(X_1, ..., X_n)$ , the probability blank of the model corresponding to the observations $(T_1, D_1), ..., (T_n, D_n)$ with:

$$T_i = X_i \wedge C$$

And

$$D_i = \begin{cases} 1 si X_i < C \\ 0 si X_i \geq C \end{cases}$$

is written with both a continuous and a discrete component :

$$L(\theta) = \prod_{i=1}^{n} f(T_i, \theta)^{D_i} S(T_i, \theta)^{1-D_i}$$

Stated differently, the density term influences the likelihood when the output is observed prior to censoring, and the discrete term is found when the survival function on the censoring date is the value in the opposite scenario. As a result, the distribution is discrete when compared to $D_i$ and continuous when compared to $T_i$ .

we can also write the likelihood in the form (up to a multiplicative constant):

$$L(\theta) = \prod_{i=1}^{n} S(T_i, \theta)^{D_i} h(T_i, \theta)^{1-D_i}$$

The loglikelihood, up to an additive constant, is typically used at:

$$l(\theta) = \sum_{i=1}^{n} [D_i \ln(S(T_i, \theta)) + (1 - D_i) \ln(h(T_i, \theta))]$$

**type II censorship (stopping at the kth death) :**

We now assume the scenario in which we agree to terminate the observation upon the occurrence of the first exit, but the end date of the observation is not specified in advance. As a result, the experiment's end date is random and equal to $X(k)$ . Formally, we state that type II censorship exists for a sample of survival times $(X_1, ..., X_n)$ with k > 0 fixed if, rather than directly observing $(X_1, ..., X_n)$ , we observe $(T_1, D_1), ..., (T_n, D_n)$ with:

$$T_i = X_i \wedge X_{(k)}$$

And

$$D_i = \begin{cases} 1 si X_i < X_{(k)} \\ 0 si X_i \geq X_{(k)} \end{cases}$$

The likelihood is formed :

$$L(\theta) = \frac{n!}{k!\,(n-k)!}\left[\prod_{i=1}^{k} f(X_i,\theta)\right] S\Big(X_{(k)},\theta\Big)^{n-k}$$

**Type III censorship(random censorship) :**

An instance of a iid sample When the censorship date is random, type III censoring generalizes type I censoring. Specifically, we have a sample of survival durations $(X_1,...,X_n)$ and another independent sample made up of positive variables $(C_1,...,C_n)$. We say that type III censorship is present for this sample if, rather than directly observing $(X_1,...,X_n)$ , we observe $(T_1,D_1),...,(T_n,D_n)$ with:

$$T_i = X_i \wedge C_i$$

and

$$D_i = \begin{cases} 1 si X_i < C_i \\ 0 si X_i \geq C_i \end{cases}$$

The likelihood of the sample $(T_1,D_1),...,(T_n,D_n)$ is written, with obvious notations:

$$L(\theta) = \prod_{i=1}^{n} S(T_i,\theta)h(T,\theta)^{D_i}$$

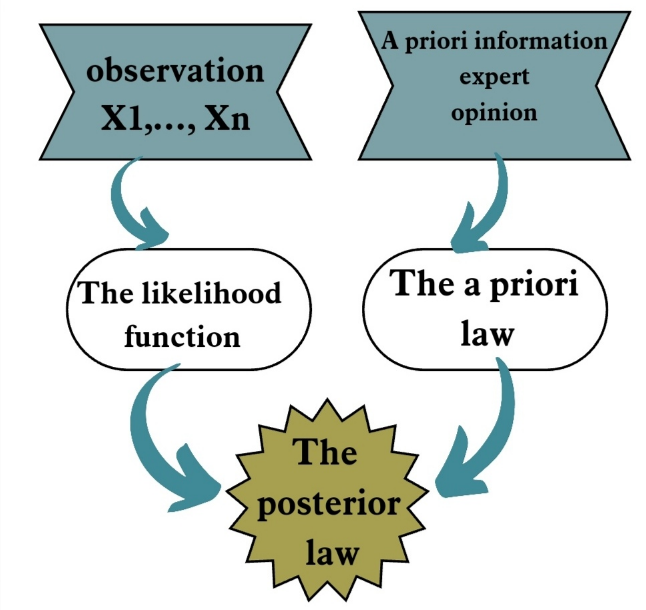
Here, we merely note that fixed censorship is a particular instance of non-informative random censorship, where the censoring law is a Dirac law at point C. It is consequently simple to generalize the statement that was established in the specific instance of fixed censoring.

## 1.3 The Bayesian approach

When estimating a parameter $\theta$ for a given model, the classical statistician views $\theta$ as a fixed parameter that belongs to a space $\Theta$ and we estimate it based on the observed sample. In contrast, the Bayesian approach views$\theta$ as a random variable that is estimated based on information preventing the sample and, if relevant, expert opinion. In a parametric observation model $X \sim f(X/\theta)$ , where $\theta \in \Theta$ limited dimensional space and together they are termed states, and f is the density function that depends on an unknown parameter $\theta$ that needs to be estimated.

The objective of the Bayesian perspective is to build inference methods by making the most use of the information that X provides about the parameter $\theta$. Even if X is merely an arbitrary realization of a law controlled by $\theta$, it offers an update to the data that the experimenter had previously gathered. The foundation of Bayesian statistical analysis consists of two laws: the first is the a priori law, represented by the symbol $\pi(.)$ , which serves as the approach's main engine and summarizes the data on $\theta$ independently of the knowledge revealed by the experimental findings. The second law is the a posteriori law, represented by the symbol $\pi(.|x)$ , which encompasses all of the available data for the parameter $\theta$.

The following diagram summarizes the Bayesian paradigm:




### 1.3.1 Basic of Bayesian Decision Theory

A section on decision theory is highly helpful before discussing estimating, as we shall show that figuring out a Bayes estimator is equivalent to figuring out a decision rule. A decision problem typically relies on the following components in addition to the observation space and the parameter space:

- $\chi$ :The observation space.
- $\Theta$ : The space of natural states (or, in the case of a statistical problem, the space of parameters)
- A : The area of actions or choices, wherein the events are snapshots of observations made by an application $\delta$ known as a decision rule (a statistical function in the case of an observation problem).
- D : the set of decision rules $\delta$ , applications of X in A (the possible estimators). We observe a behaviour. $a = \delta(x)$ is what we have. Selecting a decision rule $\delta \in D$ regarding $\theta \in \Theta$ based on an observation $x \in \chi$ , where x and $\theta$ are connected by the law $f(x|\theta)$, constitutes inference.

In statistics, an estimator serves as the decision rule, and the value of the estimator at the observation point serves as the action. Create a preference connection by taking into account a measure of the loss or cost experienced when making a decision, $\delta$ , and the natural state, $\delta(x)$, in order to make a choice. In order to accomplish this, we first introduce function L, often known as the cost (or loss) function, which is defined as follows: We refer to any function L of $\Theta \times D$ in R as the cost function. When the parameter is $\Theta$, the cost of a decision an is evaluated using $L(\theta, a)$. As a result, it makes it possible to measure the loss brought about by a poor choice or poor assessment of $\Theta$. This is dependent on $\Theta$. A benefit is correlated with a negative cost.

**1-The Average Bayesian risk (frequency):**

Let $L(\theta, \delta)$ be a loss function, we call the average Bayesian risk associated with this function, denoted $R(\theta, \delta)$ the average of the loss function :

$$R(\theta, \delta) = \mathrm{E}[L(\theta, \delta(x))] = \int_{\chi} L(\theta, \delta(x)) f(x|\theta) dx$$

**2-The Bayes risk:**

means associated with the loss function l compared to the a priori law: $\mathrm{E}^{\pi}[R(\theta, \delta)]$. It's about of Bayesian risk or Bayes risk which we note $r(\pi, \delta)$ .

We have:

$$
\begin{aligned}
r(\pi,\delta) &= \mathrm{E}^{\pi}\left[R(\theta,\delta)\right] \\
&= \int_{\Theta} R(\theta,\delta)\,\pi(\theta)d\theta \\
&= \int_{\Theta}\int_{\chi} L(\theta,\delta(x))\,f(x|\theta)\,dx\pi(\theta)\,d\theta \\
&= \int_{\Theta}\int_{\chi} L(\theta,\delta(x))\,\pi(\theta)\,f(x|\theta)\,dxd\theta
\end{aligned}
$$

**3-The posterior risks:**

we define the a posteriori risk denoted $\rho(\pi,\delta|x)$ as being the average of the function of loss compared to the a posteriori law:

$$\rho(\pi,\delta|x) = \mathrm{E}^{\pi(.|x)}\left[L(\theta,\delta(x))\right] = \int_{\Theta} L(\theta,\delta(x))\,\pi(\theta|x)d\theta$$

We have the following result by setting $f_X$ the marginal density of X:

$$
\begin{aligned}
r(\pi,\delta) &= \int_{\chi}\int_{\Theta} L(\theta,\delta(x))\,\pi(\theta|x)d\theta f(x|\theta)\,dx \\
&= \int_{\chi} \rho(\pi,\delta|x)\,f(x|\theta)\,dx
\end{aligned}
$$

the Bayes risk $r(\pi,\delta)$ is the average of the a posteriori risk $\rho(\pi,\delta|x)$ following the marginal law $f(x)$.

### 1.3.2 Specification of the prior distribution

the choice of the a priori law is crucial for Bayesian analysis because it affects a directly the posterior distributions. So how do you transfer the information that is known beforehand into law that is known beforehand?, Due to the rarity of sufficiently precise a priori information to determine this law precisely, this question is essential in practice. Schematically, two ways of thinking stand out. The first is based on subjective assumptions that the prior distribution represents knowledge gained from work-related experiences and sound intuitions prior to data observation. A so-called "informative" statute expresses this information.

The second way of thinking is more impartial. When there is minimal information, it is employed. The ability to maintain a Bayesian model in the absence of previous knowledge is thus at issue. We are therefore looking for "uninformative" (or vague) prior distributions expressing a priori ignorance but treating the parameters as random

(1) **The informative a priori law :**

It is based on the information about the parameter $\Theta$ that has been gathered from the views of specialists and researchers. Rarely is the a priori knowledge interpreted with enough accuracy to produce a single a priori law. Conjugate laws, on the other hand, are laws that are calibrated based on the distribution of observations.

While there are several ways to acquire conjugate laws, we will focus on one of the most fascinating here conjugate families.

**The conjugate prior distribution :**

Let $L(\theta) = f(x|\theta)$ be the maximum likelihood function based on the observation $X = x$ .

The class G of distributions is called conjugate with respect to $L(\theta)$ , if the posterior distribution $f(x|\theta)$ is in G for all x whenever the distribution $\pi(\theta)$ is in G

For every probability function, the family G = all distributions is trivially conjugate. In actuality, we search for small sets G unique to the likelihood $L(\theta)$ .

A few examples of prior distributions and the matching probability function are shown in the following table:

| **The likelihood** | **The conjugate prior distribution** | **The posterior distribution** |
|---|---|---|
| $X\|\theta \sim \text{Bin}(n,\theta)$ | $\theta \sim \text{Be}(\alpha,\beta)$ | $\theta\|X \sim \text{Be}(\alpha+x,\beta+n-x)$ |
| $X\|\theta \sim \text{Geom}(\theta)$ | $\theta \sim \text{Be}(\alpha,\beta)$ | $\theta\|X \sim \text{Be}(\alpha+1,\beta+x-1)$ |
| $X\|\pi \sim \text{Po}(e\lambda)$ | $\lambda \propto \text{G}(\alpha,\beta)$ | $\lambda\|X \sim \text{G}(\alpha+x,\beta+e)$ |
| $X\|\pi \sim \exp(\lambda)$ | $\lambda \sim \text{G}(\alpha,\beta)$ | $\lambda\|X \sim \text{G}(\alpha+1,\beta+x)$ |
| $X\|\mu \sim \mathcal{N}(\mu,\sigma^2 \text{ known})$ | $\mu \sim \mathcal{N}(v,\tau^2)$ | $\mu\|X \sim \mathcal{N}\left(\left(\frac{1}{\sigma^2}+\frac{1}{\tau^2}\right)^{-1}\left(\frac{x}{\sigma^2}+\frac{v}{\tau^2}\right),\left(\frac{1}{\sigma^2}+\frac{1}{\tau^2}\right)^{-1}\right)$ |
| $X\|\sigma^2 \sim \mathcal{N}(\mu \text{ known},\sigma^2)$ | $\sigma^2 \sim \text{IG}(\alpha,\beta)$ | $\sigma^2\|X \sim \text{IG}\left(\alpha+\frac{1}{2},\beta+\frac{1}{2}(x-\mu)^2\right)$ |

Table 1.1: Examples of conjugate prior and posterior distributions

(2) **The non-informative a priori law :**

When there is no prior information or when it is challenging to convert the available parameter information into a prior distribution, Bayesian theory can still be used. In these circumstances, it is possible to construct a prior distribution based on subjective considerations using procedures that incorporate our ignorance about the relevant parameters; these approaches are referred to as non-informative.

Occasionally, selecting a non-informative prior distribution results in the definition of a measure rather than a probability, or an inappropriate distribution (a measurable function that is not integrable and is defined on the set R of measure $\pi$ : $\int_R \pi(\theta)\, d\theta$)

Here, we provide one of the most widely used methods for creating non-informative priors.

**Jeffrey's prior law :**

Jeffrey (1961) presented a method that makes it possible to generate non-informative priors. Fisher information is used in this method: $I(\theta)$ , the following could be the argument: $I(\theta)$ is a measure of how much of the observation's $\theta$ information is present. More information is provided by the observation the larger $I(\theta)$. The values of $\theta$ for which $I(\theta)$ is big seem thus to be naturally favoured ( made more probable in accordance with $I(\theta)$); this reduces the prior's effect in favour of the observation. Thus, Jeffrey's rule is associating the preceding measuring to a sampling model defined by its likelihood $L(x|\theta)$ :

$$\pi(\theta) = |\det I(\theta)|^{\frac{1}{2}}$$

where $\theta$is a vector of unknown parameters and $I(\theta)$ is the Fisher information matrix defined by:

$$I(\theta) = -\left\{ \mathrm{E}\left[ \frac{\partial^2 \log L(\theta)}{\partial\theta_i \partial\theta_j} \right] \right\}_{1\le i,j\le n}$$

**Exemple (Berger and Yang 1995) :**

Let $X \sim \mathrm{N}\left(\mu, \sigma^2\right)$ where the mean $\mu$ and the variance $\sigma^2$ are unknown parameters. The density function is written as:

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

The parameter of interest is the vector $\theta = (\mu, \sigma)$.

Following is the Fisher information matrix:

$$I(\theta) = \begin{pmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{1}{2\sigma^4} \end{pmatrix}$$

Jeffrey's prior law is written as:

$$\pi(\theta) = \frac{1}{\sqrt{2\sigma^3}}$$

Although there are alternative methods, such the invariant and reference prior, for creating non-informative priors, Jeffrey's law is still the most frequently applied method as of right now.

### 1.3.3 The posteriori distribution

Once the a priori law is found, we can now construct the a posteriori law, this is done using Bayes' formula, which we recall.

**Bayes' formula:**

In statistical inference, the Bayes theorem is used to update and renew estimates of a parameter, or probability, based on observations and their probability distributions.

First, the simplest version of Bayes' theorem is provided by

$$P(A|B) = \frac{P(A|B)\,P(A)}{P(B)}$$

where A and B are measurable sets with $P(B) \neq 0$.

**The law a posteriori:**

The posterior distribution is the most important quantity in Bayesian inference. It contains all the information available about the unknown parameter $\theta$ after observing the data $X = x$ . (posterior distribution)

Let $X = x$ denote the observed realization of a (possibly multivariate) random variable X with the density function $f(x|\theta)$ .

Let $\pi(\theta)$ be the prior density that allows us to calculate the density function $\pi(\theta|x)$ of the posterior distribution using Bayes' theorem:

$$\pi(\theta|x) = \frac{f(x|\theta)\,\pi(\theta)}{\int f(x|\theta)\,\pi(\theta)\,d\theta}$$

We refer to the likelihood function as $f(x|\theta)$. Since $\theta$ is random, we expressly condition $L(\theta) = f(x|\theta)$ on a particular value of $\theta$.

The denominator can be written as:

$$\int f(x|\theta)\pi(\theta)\,d\theta = \int f(x,\theta)d\theta = f(x)$$

which emphasizes the fact that it is independent of $\theta$. The marginal probability, or $f(x)$, is a significant quantity that influences the choice of Bayesian model.

The density of the posterior distribution is therefore proportional to the product of the likelihood and the posterior density of the distribution with a constant of proportionality $\frac{1}{f(x)}$, This is usually denoted :

$$\pi(\theta|x) \propto f(x|\theta)\pi(\theta)$$

### 1.3.4 The Bayesian Estimators

As already mentioned, any estimate involves a cost that is measured using the loss function. Intuitively, we seek a decision that minimizes the loss function on average.

minimizing $r(\pi,\delta)$ for any value of x is equivalent to minimizing the posteriori risk function. This minimization can be done analytically as it can be approached numerically according to the complexity of the loss function l and the posterior distribution $\pi(.|x)$.

**Bayes estimator of classical loss functions :**

For classical loss functions (Quadratic, Linex,...) the corresponding Bayes estimators are usual characteristics of the posterior distribution .

**(1) Generalized Quadratic loss function :**

A generalized quadratic loss function is a function $L : \Theta \times D \to R^+$ given by:

$$L(\theta,\delta) = \tau(\theta)(\theta-\delta)^2$$

where $\tau(\theta)$ is a non-negative function and the corresponding Bayes estimator is :

$$\overset{\wedge}{\delta}_{GQ} = \frac{\mathrm{E}^{\pi}[\tau(\theta)\theta]}{\mathrm{E}^{\pi}[\tau(\theta)]}$$

The posterior risk for the loss function GQ is :

$$PR\left(\overset{\wedge}{\delta}_{GQ}\right) = \mathrm{E}^{\pi}\left[\tau(\theta)\left(\theta-\overset{\wedge}{\delta}_{GQ}\right)^2\right]$$

where $\tau(\theta)$ is a positive function, for example, $\tau(\theta) = \theta^{\alpha-1}$ and $\alpha$ are assumed to be an integer.

(2) **Linex Loss function :**

A Linex loss function is a function $L : \Theta \times D \to R^{+}$ given by:

$$L(\theta, \delta) = \exp(r(\delta - \theta)) - r(\delta - \theta) - 1$$

The Bayes estimator associated with L (the estimator that minimizes the posterior expectation) :

$$\overset{\wedge}{\delta}_L = -\frac{1}{r}\log\left(\mathrm{E}^{\pi}\left[e^{-r\theta}\right]\right)$$

The posterior risk for the loss function L is :

$$PR\left(\overset{\wedge}{\delta}_L\right) = r\left(\overset{\wedge}{\delta}_{GQ} - \overset{\wedge}{\delta}_L\right)$$

The sign of r representing respectively the direction and the degree of symmetry ($r > 0$: overestimation is more serious than underestimation and vice versa). For a close to zero, the Linex loss is approximately the quadratic loss function:

(3) **Entropy Loss function :**

A Entropy loss function is a function $L : \Theta \times D \to R^{+}$ given by:

$$L(\theta, \delta) = \left(\frac{\delta}{\theta}\right)^{p} - p\log\left(\frac{\delta}{\theta}\right) - 1$$

The Bayes estimator of parameter $\theta$ under this loss function is :

$$\overset{\wedge}{\delta}_E = \left(\mathrm{E}^{\pi}\left[\theta^{-p}\right]\right)^{-\frac{1}{p}}$$

The posterior risk for the loss function E is :

$$PR\left(\overset{\wedge}{\delta_p}\right) = p\left(\mathrm{E}^{\pi}\left[\log(\theta) - \log\left(\overset{\wedge}{\delta_p}\right)\right]\right)$$

- When $p = 1$ , the Bayes estimator coincides with the Bayes estimator under the weighted quadratic loss function $\frac{(\delta-\theta)^2}{\theta}$
- When $p = -1$ , the Bayes estimator coincides with the Bayes estimator under the quadratic loss function.

The following table summarizes the different loss functions used :

| Loss function Expresion | Bayes estimators | posterior risk |
|---|---|---|
| Entropy : $L(\theta,\delta) = \left(\frac{\delta}{\theta}\right)^p - p\log\left(\frac{\delta}{\theta}\right) - 1$ | $\hat{\delta}_E = E_\pi(\theta^{-p})^{\frac{-1}{p}}$ | $p[E_\pi(\log(\theta - \log(\hat{\delta}_E)))]$ |
| Generalized quadratic : $L(\theta,\delta) = \tau(\theta)(\theta-\delta)^2$ | $\hat{\delta}_{GQ} = \frac{E_\pi(\tau(\theta)\theta)}{E_\pi(\tau(\theta)}$ | $E_\pi(\tau(\theta)(\theta-\delta)^2$ |
| Linex : $L(\theta,\delta) = \exp(r(\delta-\theta)) - r(\delta-\theta) - 1$ | $\hat{\delta}_L = \frac{-1}{r}\log(E_\pi(\exp(-r\theta))$ | $r(\hat{\delta}_{GQ} - \hat{\delta}_L)$ |

(4) **Balanced Loss function :**

The symmetry criterion previously discussed is not the only standard used to classify loss functions; Zellner (1994) proposed a more thorough standard known as the balanced criterion, sometimes known as the equilibrium criterion. Enhancing precision and conformance in the estimating process is the goal of reaching equilibrium in the loss function. An uneven loss function is one of the loss functions that were previously addressed. In addition to meeting the symmetry requirement, Zellner's formula allows the loss function to be balanced as follows:

$$L_{\rho,\omega,\delta_0}(\theta,\delta) = \omega\rho(\delta_0,\delta) + (1-\omega)\rho(\theta,\delta)$$

where

- $L_{\rho,\omega,\delta_0}$ Balanced loss function .
- $\rho$ is an arbitrary loss function (Unbalanced loss function).

- $\delta_0$ is a chosen a priori "target" estimator of $\theta$, obtained for instance using the criterion of maximum likelihood, least-squares, or unbiasedness.
- $\theta$ the unknown parameter
- $\omega$ weighted coefficient $\omega \in (0,1)$

**Theorem 1 :**

Under the balanced loss function $L_{\rho,\omega,\delta_0}$ and prior $\pi$ , the Bayes estimator $\delta_{\omega,\pi}(X)$ of $\theta$ is given by the Bayes solution $\delta^*(X)$ with respect to $(\pi_x^*, L_0)$ for all x, where

$$\pi_x^*(\theta) = \omega 1_{\{\delta_0(X)\}}(\theta) + (1-\omega)\pi_x(\theta)$$

i.e., a mixture of a point mass at $\delta_0(X)$ and the posterior $\pi_x(\theta)$ .

We now use specializations of Theorem 1 to pursue several options of $\rho$, and we also provide more observations and examples to support our claims :

**(A) Generalized Quadratic Balanced loss :**

A generalization of squared error loss. In Balanced Loss function the choice :

$$\rho(\theta,\delta)=\tau(\theta)(\theta-\delta)^2$$

So

$$L_{\rho,\omega,\delta_0}(\theta,\delta)=\omega\tau(\delta_0)(\delta-\delta_0)^2+(1-\omega)\tau(\theta)(\theta-\delta)^2$$

The Bayes estimator under balanced loss function :

$$\delta_{\omega,\pi(x)}=\frac{\mathrm{E}^{\pi_x^*}[\theta\tau(\theta)]}{\mathrm{E}^{\pi_x^*}[\tau(\theta)]}=\frac{\omega\delta_0(x)\tau(\delta_0(x))+(1-\omega)\mathrm{E}^{\pi}[\theta\tau(\theta)]}{\omega\tau(\delta_0(x))+(1-\omega)\mathrm{E}^{\pi}[\tau(\theta)]}$$

**(B) Linex Balanced loss :**

A Linex loss. In Balanced Loss function the choice :

$$\rho(\theta,\delta)=\exp(r(\delta-\theta))-r(\delta-\theta)-1$$

So

$$L_{\rho,\omega,\delta_0}(\theta,\delta)=\omega[\exp(r(\delta-\delta_0))-r(\delta-\delta_0)-1]+(1-\omega)[\exp(r(\delta-\theta))-r(\delta-\theta)-1]$$

The Bayes estimator under Linex balanced loss function :

$$\delta_{\omega,\pi(x)}=-\frac{1}{r}\log(-\omega\exp(-r\delta_0(x))+(1-\omega)\mathrm{E}^{\pi}[\exp(-r\theta)])$$

**(C) Entropy Balanced loss :**

A Entropy loss. In Balanced Loss function the choice :

$$\rho(\theta,\delta)=\left(\frac{\delta}{\theta}\right)^p-p\log\left(\frac{\delta}{\theta}\right)-1$$

Where :

$$L_{\rho,\omega,\delta_0}(\theta,\delta) = \omega\left[\left(\frac{\delta_0}{\delta}\right)^p - p\log\left(\frac{\delta_0}{\delta}\right) - 1\right] + (1-\omega)\left[\left(\frac{\theta}{\delta}\right)^p - p\log\left(\frac{\theta}{\delta}\right) - 1\right]$$

The Bayes estimator under Linex balanced loss function :

$$\delta_{\omega,\pi(x)} = \left\{\frac{\omega}{(\delta_0(x))^p} + (1-\omega)\,\mathrm{E}^{\pi}\left[\frac{1}{\theta^p}\right]\right\}^{-\frac{1}{p}}$$

The following table summarizes the balanced loss function :

Table 1.2: The Balanced loss function

| Balanced Loss function | | |
|---|---|---|
| **Unbalanced loss function** | **Expresion** | **Bayes estimators** |
| Generalized Quadratic (GQ) | $L_{\rho,\omega,\delta_0}(\theta,\delta) = \omega\tau(\delta_0)(\delta-\delta_0)^2 + (1-\omega)\tau(\theta)(\theta-\delta)^2$ | $\delta_{\omega,\pi(x)} = \frac{\mathrm{E}^{\pi_x^*}[\theta\tau(\theta)]}{\mathrm{E}^{\pi_x^*}[\tau(\theta)]} = \frac{\omega\delta_0(x)\tau(\delta_0(x))+(1-\omega)\mathrm{E}^{\pi}[\theta\tau(\theta)]}{\omega\tau(\delta_0(x))+(1-\omega)\mathrm{E}^{\pi}[\tau(\theta)]}$ |
| Linex (L) | $L_{\rho,\omega,\delta_0}(\theta,\delta) = \omega[\exp(r(\delta-\delta_0)) - r(\delta-\delta_0) - 1] + (1-\omega)[\exp(r(\delta-\theta)) - r(\delta-\theta) - 1]$ | $\delta_{\omega,\pi(x)} = -\frac{1}{r}\log(-\omega\exp(-r\delta_0(x)) + (1-\omega)\mathrm{E}^{\pi}[\exp(-r\theta)])$ |
| Entropy (E) | $L_{\rho,\omega,\delta_0}(\theta,\delta) = \omega\left[\left(\frac{\delta_0}{\delta}\right)^p - p\log\left(\frac{\delta_0}{\delta}\right) - 1\right] + (1-\omega)\left[\left(\frac{\theta}{\delta}\right)^p - p\log\left(\frac{\theta}{\delta}\right) - 1\right]$ | $\delta_{\omega,\pi(x)} = \left\{\frac{\omega}{(\delta_0(x))^p} + (1-\omega)\mathrm{E}^{\pi}\left[\frac{1}{\theta^p}\right]\right\}^{-\frac{1}{p}}$ |

### 1.3.5 The Bayesian approach's interest

The following are the primary distinctions between the Bayesian technique and the classical or frequency approach that make it intriguing and justify its application:

- The subjective understanding of probabilities that forms the foundation of the Bayesian method makes more sense than frequentist theories, which see probability as a finite amount determined by an endless series of experiments.

- The Bayesian interpretation of probability is associated with a notion of rational betting: the probability assigned to an event is defined by the conditions under which a rational individual is willing to bet on the occurrence of such an event. In the face of uncertainty, the individual's rationality is required to prevent the notion of probability from being arbitrary and to be characterized by certain behavioural constraints.

- Bayes' formula is the instrument that allows to combine the subjective information (or vision) of the modeler and the evidence of the experimental results.

- Results from the Bayesian analysis can be interpreted more directly than those from the frequency technique. For Bayesians, a credibility interval with a confidence rate of $1-\alpha$ serves as the estimator of the posterior probability. Because of its established limits and inclusion of the parameter with a predetermined probability, it is known as a credibility interval. The 95 posterior credibility interval, for instance, is usually the one that is bounded below by the order 2.5 percentile and above by the order 97.5 percentile. The frequency confidence interval in classical inference has arbitrary bounds.

- In classical inference this assertion is no longer true because the (unknown) parameter of the model is not a random variable but a constant quantity. If we consider the set of random samples that can be acquired from the model, parameterized by $\theta$ ,95 of the confidence intervals calculated (based on the various samples) contain the real value of the parameter. This is the accurate interpretation of the confidence interval. Most practitioners who use Bayesian inference without realizing it also use the much more natural Bayesian interpretation.

- Compared to the estimators offered by traditional inference methods, the outcomes of Bayesian inference are more comprehensive. By obtaining the joint distribution of the model parameters, Bayesian techniques allow one to simultaneously consider the impact of global uncertainty on all unknown parameters on future predictions of the system's studied behaviour and on the decisions that are suggested by this behaviour.

### 1.3.6 Difficulty of the Bayesian approach

In the 1930s, the Bayesian technique was defined and reintroduced in the 20th century. The advent of integral calculus, followed by specialized software and computer use, has made the algorithmic challenges obsolete. However, the following main challenges with the Bayesian approach still exist:

first knowledge is a complicated technique that might lead to a subjective evaluation for which it is challenging to define an exact a priori law that can be used to model them. Although there is flexibility in selecting the a priori law, we would like to attach the following characteristics to it:

- It must be as easy as feasible to calculate the joint posterior density from the product of the prior distributions and the observation sample's distribution .
- The posterior distribution must preferably be of the same type as the prior distribution in order to allow an iterative updating calculation .
- The prior distribution needs to be parameterizable, able to be physically interpreted, and able to represent a wide range of physical circumstances or occurrences.

The posterior density (results) depends on how informative past knowledge is. Intuitively, the posterior probability will be drawn to the distribution that will yield the most information and, consequently, the most precise in relation to its average value. The posterior probability is a weighted average between the information provided a priori and the observations of the feedback (the likelihood function).

Therefore, the posterior probability density will vary depending on the distribution type selected to represent the prior information, which may or may not be useful in relation to the observations.

# Bayesian estimation of distributions in the presence of censorship

In this chapter, We provide a brand-new distribution based on the model of Lindley, with an emphasis on the estimation of its unknown parameters. Following the introduction of the new distribution, we discuss two methods for estimating its parameters: the Bayesian approach and the conventional method, the maximum likelihood technique, in the presence of a censored scheme. The Bayesian estimators are derived using a Monte Carlo Markov chains (MCMC) process, while the censored maximum likelihood estimators are derived using the Barzilai-Brown algorithm. The Bayesian estimators are derived from three loss functions: the Linex, the extended quadratic, and the entropy functions. The Bayesian estimations and the maximum likelihood are contrasted using Pitman's closeness criteria. Every estimation method offered has been assessed through simulated experiments. Lastly, we take into account two sample Bayes predictions for future order statistics prediction.

## 2.1 The Bayesian and traditional inference of the truncated Xlindley distribution with censored data

Numerous disciplines, including statistics, engineering sciences, medicine, and finance, employ practical numerical techniques. We observe that in our practical applications, statistics play a crucial role. The presumptive probability distributions are frequently crucial to statistical analysis. But not every statistical problem adheres to the traditional or conventional probability distributions. Selecting a suitable basic model for reliability and survival analysis is becoming increasingly crucial for new data analysis. The results could be significantly affected by even a small departure from the basic model.

There are many other types of probability distributions. In this work, we are interested in the XL distribution, a one-parameter distribution that combines the Lindley and exponential distributions. The exponential distribution is one of the frequently used continuous probability distributions. The Lindley distribution, a probability distribution used to characterize the lifespan of a processor or other device, is also used to simulate the time between events.

Assuming that $Y$ is a random variable with the previously stated one-parameter distribution XL, its density function may be found using:

$$f_{XL}(y,\alpha) = \frac{\alpha^2(1+y)e^{-\alpha y}}{1+\alpha} \qquad y,\alpha > 0 \tag{2.1}$$

Its cumulative function is :

$$F_{XL}(y,\alpha) = 1 - \left[1 + \frac{\alpha y}{(1+\alpha)^2}\right]e^{-\alpha y} \tag{2.2}$$

This work's concept is focused on creating a new distribution from the XL distribution using upper truncated data. The aforementioned statement serves as the inspiration for this work. In general, it is appropriate to experiment with simpler distributions rather than more complex ones. We found that the XLindley distribution (XL) is straightforward and easy to use, and that it provides adequate fits to a variety of real-world data sets, including those related to the corona, ebola, and Nipah viruses.

### 2.1.1 The UXL distribution

A statistical experiment's "truncation" is the process of eliminating all values that deviate from predefined ranges. whereby the remaining data points fall inside these constraints, giving us shortened data. Stated differently,

If only the values of $X$ for which $X \leq c$ $(Y > c)$ are taken into consideration, that is, we omit all other values of $Y$ for which $Y > c$ $(Y_c)$, then we say that $Y$ is upper (lower) truncated at a certain point level $c$.

The upper truncated XLindly distribution's probability density function (PDF) at the location $\beta > 0$ is provided by

$$f_{UXL}(y,\alpha,\beta) = \frac{f_{XL}(y,\alpha)}{F_{XL}(y,\beta)} \qquad y,\alpha,\beta > 0$$

replacing by (2.1) and (2.2):

$$f_{UXL}(y,\alpha,\beta) = \frac{\alpha^2(1+y)(1+\beta)^2 e^{-\alpha y+\beta y}}{(1+\alpha)A_\beta(y)} \tag{2.3}$$

where

$$A_\beta(y) = (1+\beta)^2(e^{\beta y}-1)+\beta y$$

The cumulative function (CDF) can then be calculated using the same formula, yielding:

$$\begin{aligned} F_{UXL}(y,\alpha,\beta) &= \frac{F_{XL}(y,\alpha)}{F_{XL}(y,\beta)} \\ &= \frac{1-[1+\dfrac{\alpha y}{(1+\alpha)^2}]e^{-\alpha y}}{1-[1+\dfrac{\beta y}{(1+\beta)^2}]e^{-\beta y}} \end{aligned}$$

$$F_{UXL}(y,\alpha,\beta) = \frac{A_\alpha(y)(1+\beta)^2 e^{-\alpha y+\beta y}}{(1+\alpha)^2 A_\beta(y)} \tag{2.4}$$

where

$$A_\alpha(y) = (1+\alpha)^2(e^{\alpha y}-1)+\alpha y$$

The following provides the corresponding survival function:

$$S(y) = 1-F(y;\alpha,\beta) = \frac{(1+\theta)^2 A_\beta(y) - A_\alpha(y)(1+\beta)^2 e^{\alpha y+\beta y}}{(1+\alpha)^2 A_\beta(y)} \tag{2.5}$$

At this time, $t$, the failure rate is represented as follows:

$$H(t) = \frac{f(t)}{S(t)} = \frac{(1+\alpha)\alpha^2(1+t)(1+\beta)^2}{(1+\alpha)^2 A_\beta(t)e^{\alpha t-\beta t} - A_\alpha(t)(1+\beta)^2} \tag{2.6}$$

The joint probability density function is as follows, assuming that the random variables $Y_1, Y_2, ..., Y_n$ are independent and do follow the UXL distribution:

$$f_{UXL}(y_1, y_2, ..., y_n, \alpha, \beta) = \frac{\alpha^{2n}(1+\beta)^{2n} e^{\sum_{i=1}^{n}(-\alpha y_i + \beta y_i)}}{(1+\alpha)^n} \prod_{i=1}^{n} \frac{1+y_i}{A_\beta(y_i)} \tag{2.7}$$

### 2.1.2 Estimation of maximum likelihood

In traditional statistical conclusions, we employ the Maximum Likelihood Estimation (MLE) method, which is widely utilized due to its flexible and straightforward reasoning. A statistical method for estimating the parameters of a probability distribution that has been assumed given some observed data is called maximum likelihood estimation, or MLE. This is achieved by making the observed data as likely as possible given the assumed statistical model by maximizing a probability function. But in most situations, figuring out the maximum of the probability function will require the application of numerical methods.

The joint probability density function (2.7) is the likelihood function when the data is complete. To estimate the parameters in this case, type II censored data is of interest. We suppose that the $m$-sample $(Y_1, Y_2, ..., Y_m)$ is created from the UXL distribution, taking into account the $n$-sample $(Y_1, Y_2, ..., Y_n)$ and a fixed constant $m$. This sample's likelihood function is:

for $n, m \in \mathbb{N}$

$$L(\theta, \beta, Y) = N \prod_{i=1}^{m} f_{UXL}(y, \theta, \beta)[1 - F_{UXL}(y_m, \theta, \beta)]^{n-m}$$

where $N = \dfrac{n!}{(n-m)!}$

We have the following in place of both (2.3) and (2.2):

$$L(\alpha,\beta,Y) = NB^m e^{\sum_{i=1}^{m}(-\alpha y_i+\beta y_i)} C^{n-m}(x_m)\prod_{i=1}^{m}\frac{1+y_i}{A_\beta(y_i)}. \tag{2.8}$$

where

$$\begin{cases} B = \dfrac{\alpha^2(1+\beta)^2}{1+\alpha}, \\ C(y_m) = \dfrac{(1+\alpha)^2 A_\beta(y_m) - A_\alpha(y_m)(1+\beta)^2 e^{-\alpha y_m+\beta y_m}}{(1+\alpha)^2 A_\beta(y_m)} \end{cases}.$$

Looking at the logarithm, we discover:

$$\begin{aligned} l(y,\alpha,\beta) &= \ln L(\alpha,\beta,y) \\ &= \ln N + m\ln B + \sum_{i=1}^{m}(-\alpha y_i+\beta y_i) + (n-m)\ln C(y_m) + \sum_{i=1}^{m}\ln\left(\frac{1+y_i}{A_\beta(y_i)}\right) \\ &= \ln N + m\ln B - \sum_{i=1}^{m}(-\alpha y_i + \sum_{i=1}^{m}\beta y_i) + (n-m)\ln C(y_m) \\ &+ \sum_{i=1}^{m}\ln(1+y_i) - \sum_{i=1}^{m}\ln(A_\beta(y_i)). \end{aligned}$$

By noting :

$$\frac{\partial B}{\partial\alpha} = (1+\beta)^2\left(\frac{2\alpha+\alpha^2}{(1+\alpha)^2}\right)$$

$$\frac{\partial B}{\partial\beta} = 2(1+\beta)\left(\frac{\alpha^2}{1+\alpha}\right)$$

$$\begin{aligned}\frac{\partial C(y_m)}{\partial\beta} &= \frac{(1+\alpha)^2\frac{\partial A_\beta}{\partial\beta} - (A_\alpha(2+2\beta)e^{-\alpha y_m+\beta y_m} + A_\alpha(y_m)\beta e^{-\alpha y_m+\beta y_m})(1+\theta)^2 A_\beta}{((1+\theta)^2 A_\beta(y_m))^2} \\ &- \frac{(1+\alpha)^2\frac{\partial A_\beta}{\partial\beta}[(1+\theta)^2 A_\beta + A_\alpha(1+\beta)e^{-\alpha y_m+\beta y_m}}{((1+\alpha)^2 A_\beta(y_m))^2}\end{aligned}$$

$$\frac{\partial A_\beta(y)}{\partial\beta} = (2+2\beta)(e^{\beta y}-1) + y(1+\beta)^2 e^{\beta y} + y$$

$$\frac{\partial A_\alpha(y)}{\partial\alpha} = (2+2\alpha)(e^{\alpha y}-1) + y(1+\alpha)^2 e^{\alpha y} + y$$

The following non-linear system is solved to get the maximum likelihood estimators $\hat{\alpha}_{ML}$ and $\hat{\beta}_{ML}$:

$$(S)\begin{cases} \hat{\alpha}_{ML} = \frac{\partial l}{\partial \alpha} = m.\frac{\frac{\partial B}{\partial \alpha}}{B} - \sum_{i=1}^{m} y_i + (n-m)\frac{\frac{\partial C(y_m)}{\partial \alpha}}{C(y_m)} & = 0 \\ \hat{\beta}_{ML} = \frac{\partial l}{\partial \beta} = m.\frac{\frac{\partial B}{\partial \beta}}{B} + \sum_{i=1}^{m} y_i + (n-m)\frac{\frac{\partial C(y_m)}{\partial \beta}}{C} - \frac{\frac{\partial A_\beta(y)}{\partial \beta}}{A_\beta(y)} & = 0 \end{cases} \tag{2.9}$$

Since the estimators' formulation is obviously hard to calculate analytically, we employ numerical methods to get approximations of the estimators using the $\mathcal{R}$ programming language. The following section discusses the findings.

#### Simulation research

We want to conduct a simulation study using a range of sample sizes, with $n = 20, 50, 200$ and $m = 12, 30, 120$, which correspond to various effective sample sizes. After $M = 10000$ sample generations, we take $\alpha = 2$ and $\beta = 1$ to get the findings.

After using the R programming language, the results are shown. Specifically, we utilized the package BB and the command BBsolve, which stands for Barzilai-Brown, which is renowned for its high capacity for solving large-scale nonlinear systems; for further information, see Varadhan and Gilbert. [14]

| N = 5000 | n = 20 | n = 50 | n = 200 |
|---|---|---|---|
| m | 12 | 30 | 120 |
| $\alpha$ | 2,0502 (0,0152) | 1,9234 (0,0217) | 1,9872 (0,0078) |
| $\beta$ | 0,6135 (0,0078) | 0,7397 (0,0054) | 0,9573 (0,0045) |

Table 2.3: The quadratic error of the parameters' MLE (in brackets).

As we can see from the above table, all of the estimated values for $\alpha$ are rather near to the actual value of $\alpha$, and the associated error is minimal. However, for $\beta$, large values of $n$ and $m$ result in the best estimation value and the smallest quadratic error.

### 2.1.3 The Bayesian estimating method

This section discusses the Bayesian technique, which treats the unknown parameters as random variables; in other words, we start with a prior distribution of the parameters to be estimated based on a piece of prior information.

We make use of both the informative and non-informative forms of prior distribution.

We assume that we have a gamma distribution as an informative prior for the first parameter ($\alpha$).

$$\pi(\alpha) = \frac{a^b}{\Gamma(b)}\alpha^{b-1}e^{-a\alpha} \qquad \alpha > 0, a, b > 0$$

Because of its adaptability and ability to provide conjugate prior distributions, the gamma distribution is frequently employed as a prior distribution.

A non-informative prior distribution is employed for the second parameter ($\beta$);

$$\pi(\beta) = \frac{1}{\beta}$$

The joint prior distribution, taking into account that the parameters are independent, is:

$$\pi(\alpha, \beta) = \frac{a^b}{\beta\Gamma(b)}\alpha^{b-1}e^{-a\alpha} \qquad \alpha, \beta > 0, a, b > 0$$

For type II censored data, Bayesian estimation is also performed. The posterior distribution is then read using (2.8) as follows:

$$\pi(\alpha, \beta, Y) = k\beta^m e^{\sum_{i=1}^{m}(-\alpha y_i + \beta y_i)} C_{y_n}^{n-m}\beta^{-1}\alpha^{b-1}e^{-a\alpha}\prod_{i=1}^{m}\frac{1+y_i}{A_\beta(y_i)} \tag{2.10}$$

where $K = \int_0^{+\infty}\int_0^{+\infty}\beta^m e^{\sum_{i=1}^{m}(-\theta y_i + \beta y_i)} C_{y_n}^{n-m}\beta^{-1}\theta^{b-1}e^{-a\theta}\prod_{i=1}^{m}\frac{1+y_i}{A_\beta(y_i)}d\theta d\beta$

**Estimators and their corresponding risks**

1. We obtain the estimators and associated risks under the entropy loss function ($p$ is an integer):

$$\hat{\alpha}_(EN) = [K\int\int B^m e^{\sum_{i=1}^{m}(-\alpha y_i+\beta y_i)-a\alpha} C^{n-m}(y_m)\beta^{-1}\alpha^{b-1-p}\prod_{i=1}^{m}\frac{1+y_i}{A_\beta(y_i)}d\alpha d\beta]^{-\frac{1}{p}}$$
$$\hat{\beta}_(EN) = [K\int\int B^m e^{\sum_{i=1}^{m}(-\alpha y_i+\beta y_i)-a\alpha} C^{n-m}(y_m)\beta^{-p-1}\alpha^{b-1}\prod_{i=1}^{m}\frac{1+y_i}{A_\beta(y_i)}d\alpha d\beta]^{-\frac{1}{p}} \tag{2.11}$$

$$PR(\hat{\alpha}_{EN}) = pE_\pi(\ln\alpha - \ln\hat{\alpha}_(EN))$$
$$PR(\hat{\beta}_{EN}) = pE_\pi(\ln\beta - \ln\hat{\beta}_(EN))$$

2. According to the generalized quadratic loss function, we obtain the estimators and the associated risks:

$$\hat{\alpha}_Q = \frac{\int_0^{+\infty}\int_0^{+\infty}\beta^m e^{\sum_{i=1}^{m}(-\alpha y_i+\beta y_i)} C^{n-m}(y_m)\beta^{-1}\alpha^{\gamma+b-1}\prod_{i=1}^{m}\frac{1+y_i}{A_\beta(y_i)}d\alpha d\beta}{\int_0^{+\infty}\int_0^{+\infty}\beta^m e^{\sum_{i=1}^{m}(-\alpha y_i+\beta y_i)} C^{n-m}(y_m)\beta^{-1}\alpha^{\gamma+b-2}\prod_{i=1}^{m}\frac{1+y_i}{A_\beta(y_i)}d\alpha d\beta}$$
$$\hat{\beta}_Q = \frac{\int_0^{+\infty}\int_0^{+\infty}B^m e^{\sum_{i=1}^{m}(-\alpha y_i+\beta y_i)-a\alpha} C^{n-m}(y_m)\beta^{\gamma-1}\alpha^{b-1}\prod_{i=1}^{m}\frac{1+y_i}{A_\beta(y_i)}d\alpha d\beta}{\int_0^{+\infty}\int_0^{+\infty}B^m e^{\sum_{i=1}^{m}(-\alpha y_i+\beta y_i)-a\alpha} C^{n-m}(y_m)\beta^{\gamma-2}\theta^{b-1}\prod_{i=1}^{m}\frac{1+y_i}{A_\beta(y_i)}d\alpha d\beta} \tag{2.12}$$

$$PR(\hat{\theta}_Q) = E_\pi(\alpha^{\gamma+1}) - 2\hat{\alpha}_Q E_\pi(\alpha^\gamma) + \hat{\alpha}_Q E_\pi(\alpha^{\gamma-1})$$
$$PR(\hat{\beta}_Q) = E_\pi(\beta^{\gamma+1}) - 2\hat{\beta}_Q E_\pi(\beta^\gamma) + \hat{\beta}_Q E_\pi(\beta^{\gamma-1})$$

3. We obtain the estimators and associated risks ($r$ is an integer) under the Linex loss function. :

$$\hat{\alpha}_L = \frac{-k}{r}\ln\Big[\int\int B^m e^{\sum_{i=1}^m(-\alpha y_i+\beta y_i)-a\alpha-r\alpha} C^{n-m}(y_m)\beta^{-1}\alpha^{b-1}\prod_{i=1}^m \frac{1+y_i}{A_\beta(y_i)} d\alpha d\beta\Big]$$
$$\hat{\beta}_L = \frac{-k}{r}\ln\Big[\int\int B^m e^{\sum_{i=1}^m(-\alpha y_i+\beta y_i)-a\alpha-r\beta} C^{n-m}(y_m)\beta^{-1}\alpha^{b-1}\prod_{i=1}^m \frac{1+y_i}{A_\beta(y_i)} d\alpha d\beta\Big] \tag{2.13}$$

$$PR(\hat{\alpha}_L) = r(\hat{\alpha}_Q - \hat{\alpha}_L)$$
$$PR(\hat{\beta}_L) = r(\hat{\beta}_Q - \hat{\beta}_L)$$

Since it is evident that the Bayesian estimators in 2.11, **??**, and 2.13 cannot be calculated analytically, we suggest using an MCMC approach to approximate them in order to achieve the estimation values.

**Simulation research**

Using the MCMC procedure we described, we select $a = 2$ and $b = 1$ for the hyper-parameters of the prior distribution. This way, the prior mean becomes the expected value of the parameter. The Bayesian estimation under the entropy, generalized quadratic, and Linex loss functions is shown in the following tables.

For the integers $p, \gamma, r \in \{-2, -1.5, -1, -0.5, 2, 1.5, 1, 0.5\}$, we selected various options.

| | N = 5000 | n = 20 | n = 50 | n = 200 |
|---|---|---|---|---|
| $p$ | m | 12 | 30 | 120 |
| -2 | $\alpha$ | 2.0942 (0.0008) | 2.3990 (0.1644) | 2.2144 (0.0019) |
| | $\beta$ | 1.3188 (0.0699) | 1.2839 (0.0090) | 0,.7034 (0.0110) |
| -1.5 | $\alpha$ | 2.1067 (0.0091) | 1.7188 (0.1443) | 2.2179 (0.0017) |
| | $\beta$ | 0.4407 (0.0611) | 0.4077 (0.0661) | 0.7060 (0.0012) |
| -1 | $\alpha$ | 2.1041 (0.0009) | 1.6205 (0.0171) | 2.2167 (0.0001) |
| | $\beta$ | 1.4177 (0.0072) | 1.3633 (0.0073) | 0.7051 (0.0003) |
| -0.5 | $\alpha$ | 1.7981 (0.0038) | 1.7830 (0.0733) | 2.2148 (0.0009) |
| | $\beta$ | 0.6493 (0.0308) | 0.8755 (0.319) | 0.7037 (0.0009) |
| 0.5 | $\alpha$ | 1.8998 (0.0008) | 1.8895 (0.0729) | 1.9814 (0.0001) |
| | $\beta$ | 0.7638 (0.0071) | 0.9856 (0.0065) | 1.0024 (0.0002) |
| 1 | $\alpha$ | 1.6981 (0.0038) | 2.4830 (0.0733) | 1.2148 (0.0009) |
| | $\beta$ | 0.5491 (0.0308) | 1.3055 (0.0319) | 0.6037 (0.0009) |
| 1.5 | $\alpha$ | 1.7053 (0.0035) | 1.6701 (0.0667) | 2.2169 (0.0009) |
| | $\beta$ | 1.4239 (0.0199) | 1.3881 (0.0303) | 0.7059 (0.0003) |
| 2 | $\alpha$ | 1.7697 (0.0099) | 1.7644 (0.1173) | 2.2188 (0.0031) |
| | $\beta$ | 1.4579 (0.0997) | 1.4259 (0.0944) | 0.7071 (0.0014) |

Table 2.4: PR and Bayes estimators under the entropy loss function (in brackets).

| | N = 5000 | n = 20 | n = 50 | n = 200 |
|---|---|---|---|---|
| $\gamma$ | m | 12 | 30 | 120 |
| -2 | $\alpha$ | 1.6490 (0.0089) | 1.6825 (0.0041) | 1.6432 (0.0016) |
| | $\beta$ | 0.6657 (0.1491) | 0.5033 (0.0611) | 0.8113 (0.0008) |
| -1.5 | $\alpha$ | 1.7990 (0.0087) | 2.0825 (0.0061) | 2.2127 (0.0016) |
| | $\beta$ | 0.8657 (0.7091) | 0.7039 (0.0633) | 0.7120 (0.0008) |
| -1 | $\alpha$ | 1.9282 (0.0004) | 1.9841 (0.0001) | 2.0015 (0.0001) |
| | $\beta$ | 0.9296 (0.0003) | 0.9789 (0.0009) | 0.9998 (0.0001) |
| -0.5 | $\alpha$ | 2.0994 (0.0089) | 2.0888 (0.0070) | 2.2138 (0.0018) |
| | $\beta$ | 1.2999 (0.0825) | 1.2701 (0.711) | 0.7131 (0.0012) |
| 0.5 | $\alpha$ | 1.7510 (0.0095) | 1.7926 (0.0077) | 2.1839 (0.0020) |
| | $\beta$ | 0.6891 (0.0909) | 1.3591 (0.995) | 1.7139 (0.0019) |
| 1 | $\alpha$ | 1.7575 (0.0091) | 2.0977 (0.0078) | 2.1841 (0.0031) |
| | $\beta$ | 1.4228 (0.1094) | 1.3803 (0.1071) | 1.7149 (0.0025) |
| 1.5 | $\alpha$ | 1.6743 (0.0098) | 1.5632 (0.0081) | 2.1232 (0.0042) |
| | $\beta$ | 1.4768 (0.1241) | 0.6754 (0.1181) | 0.7903 (0.0033) |
| 2 | $\alpha$ | 2.1099 (0.0098) | 2.0990 (0.0081) | 2.1841 (0.0042) |
| | $\beta$ | 1.4768 (0.1241) | 1.4191 (0.1181) | 0.7158 (0.0033) |

Table 2.5: PR and Bayes estimators under generalized quadratic loss function (in brackets)

| | N = 5000 | n = 10 | n = 50 | n = 200 |
|---|---|---|---|---|
| $r$ | m | 12 | 30 | 120 |
| -2 | $\alpha$ | 1.7021 (0.0039) | 1.6549 (0.0066) | 2.1015 (0.0006) |
| | $\beta$ | 1.4647 (0.1041) | 1.1058 (0.0147) | 1.4315 (0.0481) |
| -1.5 | $\alpha$ | 1.5409 (0.1666) | 1.5861 (0.0009) | 1.7174 (0.0003) |
| | $\beta$ | 1.4721 (0.1884) | 1.4183 (0.0131) | 0.7145 (0.0004) |
| -1 | $\alpha$ | 1.7201 (0.0039) | 1.6815 (0.0038) | 1.8179 (0.0012) |
| | $\beta$ | 1.4806 (0.0411) | 1.4455 (0.0519) | 0.7054 (0.0013) |
| -0.5 | $\alpha$ | 1.5555 (0.0519) | 1.5815 (0.0183) | 2.0070 (0.0057) |
| | $\beta$ | 0.2191 (0.0049) | 0.1251 (0.0195) | 0.7094 (0.0057) |
| 0.5 | $\alpha$ | 2.2041 (0.0107) | 1.9813 (0.0031) | 2.0019 (0.0015) |
| | $\beta$ | 0.8070 (0.0020) | 1.1619 (0.0105) | 0.8021 (0.0012) |
| 1 | $\alpha$ | 1.7228 (0.0105) | 1.7819 (0.0081) | 1.8153 (0.0004) |
| | $\beta$ | 0.7117 (0.1033) | 1.4639 (0.0581) | 0.7059 (0.0013) |
| 1.5 | $\alpha$ | 2.0182 (0.0013) | 2.0501 (0.0007) | 1.9731 (0.0003) |
| | $\beta$ | 0.9692 (0.0014) | 0.9462 (0.0099) | 0.9859 (0.0004) |
| 2 | $\alpha$ | 1.6991 (0.0007) | 1.8058 (0.0147) | 2.2061 (0.0015) |
| | $\beta$ | 1.3815 (0.0183) | 1.1251 (0.0195) | 0.7091(0.0032) |

Table 2.6: PR (in brackets) and Bayes estimators under the Linex loss function.

We see that:

- The optimal posterior risk under the entropy loss function is $p = 0.5$.
- The optimal posterior risk under the generalized loss function is $\gamma = -1$.
- The optimal posterior risk under the Linex loss function is $r = 1.5$ a.
- When $n$ and $m$ are huge, our posterior risk is at its lowest.

The following table then shows the three loss functions at their peak performance, when $p = 0.5$, $\gamma = -1$, and $r = 1.5$, so that they can be compared based on the corresponding posterior risk.

| | N = 5000 | n = 20 | n = 50 | n = 200 |
|---|---|---|---|---|
| Loss functions | m | 12 | 30 | 120 |
| Entropy (p=0.5) | $\alpha$ | 1.8998 (0.0008) | 1.8895 (0.0729) | 1.9814 (0.0001) |
| | $\beta$ | 0.7638 (0.0071) | 0.9856 (0.0065) | 1.0024 (0.0002) |
| (GQ)$\gamma$= -1 | $\alpha$ | 1.9282 (0.0004) | 1.9841 (0.0001) | 2.0015 (0.0001) |
| | $\beta$ | 0.9296 (0.0003) | 0.9789 (0.0009) | 0.9998 (0.0001) |
| Linex (r = 1.5) | $\alpha$ | 2.0182 (0.0013) | 2.0501 (0.0007) | 1.9731 (0.0003) |
| | $\beta$ | 0.9692 (0.0014) | 0.9462 (0.0099) | 0.9859 (0.0004) |

Table 2.7: PR and Bayes estimators under the three loss functions (in brackets).

Among the three loss functions, it is evident that the generalized quadratic loss function provides the lowest posterior risk.

### 2.1.4 Comparison of the estimation methods

There are numerous ways to compare various estimate techniques. The Pitman Criterion is a straightforward and rational process that may be used to compare two estimators. It is defined as follows: (see Jozani [11])

An estimate $\delta_1$ of a parameter $\theta$ performs better than another estimator termed $\delta_2$ if (under the Pitman closeness criterion)

$$P_\theta[|\delta_1 - \theta| < |\delta_2 - \theta|] > \frac{1}{2}.$$

| | N = 5000 | n = 20 | n = 50 | n = 200 |
|---|---|---|---|---|
| Loss functions | m | 12 | 30 | 120 |
| EN (p=0.5) | $\alpha$ | 0.579 | 0.682 | 0.625 |
| | $\beta$ | 0.543 | 0.542 | 0.599 |
| (Q)$\gamma$= -1 | $\alpha$ | 0.789 | 0.602 | 0.668 |
| | $\beta$ | 0.753 | 0.559 | 0.643 |
| L(r = 1.5) | $\alpha$ | 0.697 | 0.634 | 0.5779 |
| | $\beta$ | 0.632 | 0.579 | 0.5623 |

Table 2.8: comparison of the estimators using Pitman.

The values of the Pitman probabilities, which enable us to compare the Bayesian estimators with the MLE estimator under the three loss functions for $p = 0.5$, $\gamma = -1$, and $r = 1.5$, are shown in Table 2.8. The aforementioned definition states that the Bayesian estimators outperform the MLE estimators when the probability is greater than 0.5. Then, we observe that the Bayesian estimators of the two parameters outperform the MLE based on this criterion. Additionally, when compared to the other two loss functions, the generalized quadratic loss function has the best values with $0.789_{|n=20,m=12|}$ and $0.753_{|n=20,m=12|}$.

### 2.1.5 Bayesian prediction for future order statistics

In the event that $Z_1, \cdots, The dataset Z_N$ is a random selection of size $N$ of future undetected observations from a comparable distribution. we assume that both samples are independent. Alternatively, we assume that $Y_{1,m,n}, \cdots, Y_{m,m,n}$ is a type II censored future sample of size $m$, selected from a two parameters UX Lindley distribution. In a future sample of size $N$, we aim to construct a Bayesian forecast for the $k^{th}, 1 \leq k \leq N$, ordered lifespan.

The density function that is calculated for the $k^{th}$, ordered lifespan $Z_{(k)}$ in the next collection (of size $N$) is as follows:

$$H_{(k)}(z_{(k)};\alpha,\beta) = k\binom{N}{k}\left[S(z_{(k)})\right]^{N-k}\left[F_{UXL}(z_{(k)},\alpha,\beta)\right]^{k-1} f_{UXL}(z_{(k)},\alpha,\beta).$$

The survival function and the functions provided in (2.3), (2.4), and (2.5) are used to obtain

$$H_{(k)}(z_{(k)};\alpha,\beta) = \alpha^2(1+z_{(k)})(1+\beta)^2 e^{-\alpha z_{(k)}+\beta z_{(k)}} \frac{D^{N-k}E^{k-1}}{(1+\alpha)^{2N-1}A_\beta(z_{(k)})^N} \tag{2.14}$$

where

$$D(z_{(k)}) = (1+\theta)^2 A_\beta(z_{(k)}) - A_\alpha(z_{(k)})(1+\beta)^2 e^{\alpha z_{(k)}+\beta z_{(k)}} F(z_{(k)}) = A_\alpha(z_{(k)})(1+\beta)^2 e^{-\alpha z_{(k)}+\beta z_{(k)}}$$

Consequently

$$H_{(k)}(Z_{(k)}|Y) = \int_0^{+\infty}\int_0^{+\infty}\int_0^{+\infty} H_{(k)}(z_{(k)};\alpha,,\beta)\pi(\alpha,\beta|Y)\mathrm{d}\alpha\mathrm{d}\beta,$$

The joint posterior density (2.10) is $\pi(\alpha,\beta|Y)$. It is evident that this Bayesian predictive density function cannot be computed or represented analytically.

employing the MCMC sampling process outlined in the previous sections, we suggest employing the natural predictor $\hat{Z}_k = E(Z_{(k)})$ in order to achieve a simulation-based prediction.

# Bayesian analysis under balanced loss function

Using type II censored data, we conduct a Bayesian study of the Zeghdoudi distribution in this chapter. We employ three distinct loss functions using two different types of loss functions: balanced and unbalanced. We provide Bayes estimators and the associated posterior risks for this estimation, which takes into account three scenarios of prior knowledge: availability and lack of main information. Since these estimators' analytical forms are unattainable, we suggest the Markov chain Monte-Carlo (MCMC) method. Furthermore, we derive maximum likelihood estimators given initial values for the model's parameters. Additionally, we use balanced and unbalanced loss functions to compare their performance with that of the Bayesian estimators.

## 3.1 The Zeghdoudi distribution

According to Messaadia and Zeghdoudi (2018), the Zeghdoudi distribution's probability density function is

$$f_{ZD}(Y,\theta) = \frac{\theta^3 y(1+y)e^{-\theta y}}{2+\theta}, \quad y,\theta > 0, \tag{3.15}$$

as well as its cumulative function is

$$F_{ZD}(Y) = \frac{1-(y^2\theta^2+\theta(\theta+2)y+\theta+2)}{(\theta+2)e^{-\theta y}}, \quad y,\theta > 0. \tag{3.16}$$

## 3.2 The unknown parameters' estimation

### 3.2.1 Function of maximum likelihood

Suppose that the data is type II censored, meaning that for a given m $\in (1,2,,m)$, we only observe $(Y_1, Y_2, \ldots, Y_m)$. Let the sample $(Y_1, Y_2, \ldots, Y_n)$ be created using the Zeghdoudi model. The function of likelihood is

$$L(Y,\theta) = N \times \left[1 - F_{y_i}(y_m)\right]^{n-m} \times \prod_{i=1}^{m} f_{y_i}(y_i)$$

where $N = \frac{n!}{(n-m)!}$

The following provides the likelihood function:

$$L(Y,\theta) = \frac{N\theta^{3m}}{(\theta+2)^n} B^{n-m} \times \prod_{i=1}^{m} A_i e^{-\theta y_i} \tag{3.17}$$

Where

$$\begin{cases} A_i = y_i(1+y_i) \\ B = \left[\theta^2 y_m^2 + \theta(\theta+2) y_m + \theta + 2\right] \mathrm{e}^{-\theta y_m} \end{cases}$$

The logarithm that corresponds to this is

$$l(Y,\theta) = \ln L(y,\theta) = \ln \mathrm{N} + 3m.ln\theta + n.\ln(\theta+2) + (n-m)\ln B + \sum_{i=1}^{m} [\ln A_i - \theta y_i] \tag{3.18}$$

The maximum likelihood estimators $\theta_{ML}$ of the parameter $\theta$ are obtained from the solution of the subsequent non-linear system.

$$\frac{\partial}{\partial\theta} l(y,\theta) = \frac{3m}{\theta} + \frac{n}{\theta+2} + (n-m)\frac{B_1}{B} - \sum_{i=1}^{m} y_i = 0 \tag{3.19}$$

where

$$B_1 = \left[\theta^2\left(y_m^3 - y_m^2\right) + \theta y_m + 1\right] \mathrm{e}^{-\theta y_m}$$

It appears that the equation (4.26) cannot be solved analytically. To get approximations, we'll use numerical techniques. To determine the estimated value of the maximum likelihood estimator $\theta_{ML}$ of the parameter $\theta$, we will utilize the $R$ package $BB$. Non-linear systems of equations can be successfully solved with the R package $BB$; see Varadhan and Gilbert (2010).

### 3.2.2 Prior and posterior distributions

The posterior distribution for the parameters ($\theta$) varies depending on the prior distribution used; it can be evaluated using the three previously mentioned priors as follows: **Informative prior and non-informative prior**. Prior can be classified into two types based on the abundance of primary information:

**(i) Informative Prior**

Here, we suppose that the Zeghdoudi distribution's parameter $\theta$ follows an independent gamma distribution:

$$\pi_1(\theta) = \frac{a^b}{\Gamma(b)}\theta^{b-1}\exp(-a\theta), \quad \theta > 0, a, b > 0,$$

where the constants $a, b$ are called hyper-parameters The posterior distribution of $\theta$ reads

$$\pi_1(\theta|Y) = K\frac{\theta^{3m+b-1}B^{n-m}}{(\theta+2)^n}\mathrm{e}^{-a\theta}\prod_{i=1}^{m}A_i e^{-\theta y_i} \tag{3.20}$$

where the normalizing constant is $K$.

**(ii) Non Informative Prior**

$$\pi_2(\theta) = \frac{1}{\theta}, \quad \theta > 0, \tag{3.21}$$

The posterior distribution of $\theta$ reads

$$\pi_2(\theta|Y) = K\frac{\theta^{3m-1}B^{n-m}}{(\theta+2)^n}\prod_{i=1}^{m} A_i e^{-\theta y_i} \tag{3.22}$$

**(iii) Fisher information Prior**

The definition of Fisher information is:

$$\pi_3(\theta) = I(\theta) = -\mathrm{E}\left[\frac{\partial^2 l(y,\theta)}{\partial\theta^2}\right], \tag{3.23}$$

$$\pi_3(\theta) = \mathrm{E}_Y\left[\frac{3m}{\theta^2} + \frac{n}{(\theta+2)^2} + (n-m)\frac{B_2B - B_1^2}{B^2}\right] = E_\theta \tag{3.24}$$

where

$$B_2 = \left[\theta^2\left(x_m^3 - x_m^4\right) + \theta\left(2x_m^3 - x_m^2\right)\right]e^{-\theta x_m}$$

The following is the posterior distribution of $\theta$:

$$\pi_3(\theta|Y) = KE_\theta\frac{\theta^{3m}B^{n-m}}{(\theta+2)^n}\prod_{i=1}^{m} A_i e^{-\theta y_i} \tag{3.25}$$

### 3.2.3 Using Bayesian estimation when different loss functions are not balanced:

We will now proceed in the following order to find Bayes estimators under an unbalanced loss function:

**(A) The generalized quadratic losss function**

**1- Informative Prior**

The following formulas provide the Bayes estimators for the generalized quadratic loss function:

$$\hat{\theta}_{Q1} = \frac{\int_0^\infty \frac{\theta^{3m+b+\alpha-1} B^{n-m}}{(\theta+2)^n} \mathrm{e}^{-a\theta} \prod_{i=1}^{m} A_i e^{-\theta y_i} d\theta}{\int_0^\infty \frac{\theta^{3m+b+\alpha-2} B^{n-m}}{(\theta+2)^n} \mathrm{e}^{-a\theta} \prod_{i=1}^{m} A_i e^{-\theta y_i} d\theta}$$

Following that, the associated posterior risk are

$$PR\left(\hat{\theta}_{Q1}\right) = \mathrm{E}_{\pi_1}\left(\theta^{\alpha+1}\right) + \hat{\theta}_{Q1}^2 \mathrm{E}_{\pi_1}\left(\theta^{\alpha-1}\right) - 2\hat{\theta}_{Q1} \mathrm{E}_{\pi_1}\left(\theta^{\alpha}\right)$$

**2- Non Informative Prior**

The following formulas provide the Bayes estimators for the generalized quadratic loss function:

$$\hat{\theta}_{Q2} = \frac{\int_0^\infty \frac{\theta^{3m+\alpha-1} B^{n-m}}{(\theta+2)^n} \prod_{i=1}^{m} A_i e^{-\theta y_i} d\theta}{\int_0^\infty \frac{\theta^{3m+\alpha-2} B^{n-m}}{(\theta+2)^n} \prod_{i=1}^{m} A_i e^{-\theta y_i} d\theta}$$

Following that, the associated posterior risk are

$$PR\left(\hat{\theta}_{Q2}\right) = \mathrm{E}_{\pi_2}\left(\theta^{\alpha+1}\right) + \hat{\theta}_{Q2}^2 \mathrm{E}_{\pi_2}\left(\theta^{\alpha-1}\right) - 2\hat{\theta}_{Q2} \mathrm{E}_{\pi_2}\left(\theta^{\alpha}\right)$$

**3- Fisher information Prior**

The following formulas provide the Bayes estimators for the generalized quadratic loss function:

$$\hat{\theta}_{Q3} = \frac{\int_0^\infty E_\theta \frac{\theta^{3m+\alpha} B^{n-m}}{(\theta+2)^n} \prod_{i=1}^{m} A_i e^{-\theta y_i} d\theta}{\int_0^\infty E_\theta \frac{\theta^{3m+\alpha-1} B^{n-m}}{(\theta+2)^n} \prod_{i=1}^{m} A_i e^{-\theta y_i} d\theta}$$

Following that, the associated posterior risk are

$$PR\left(\hat{\theta}_{Q3}\right) = \mathrm{E}_{\pi_3}\left(\theta^{\alpha+1}\right) + \hat{\theta}_{Q3}^{2}\mathrm{E}_{\pi_3}\left(\theta^{\alpha-1}\right) - 2\hat{\theta}_{Q3}\mathrm{E}_{\pi_3}\left(\theta^{\alpha}\right)$$

**(B) The Entropy loss function**

**1- Informative Prior**

We derive the following estimators under the entropy loss function:

$$\hat{\theta}_{EN1} = \left[K\int_{0}^{\infty}\frac{\theta^{3m+b-p-1}B^{n-m}}{(\theta+2)^{n}}\mathrm{e}^{-a\theta}\prod_{i=1}^{m}A_i e^{-\theta y_i}\,d\theta\right]^{-\frac{1}{p}}$$

Following that, the associated posterior risk are

$$PR\left(\hat{\theta}_{EN1}\right) = p\left[\mathrm{E}_{\pi_1}\left(\ln\theta - \ln\hat{\theta}_{EN1}\right)\right]$$

**2- Non Informative Prior**

The following estimators are obtained under the entropy loss function:

$$\hat{\theta}_{EN2} = \left[K\int_{0}^{\infty}\frac{\theta^{3m-p-1}B^{n-m}}{(\theta+2)^{n}}\prod_{i=1}^{m}A_i e^{-\theta y_i}\,d\theta\right]^{-\frac{1}{p}}$$

Following that, the associated posterior risk are

$$PR\left(\hat{\theta}_{EN2}\right) = p\left[\mathrm{E}_{\pi_2}\left(\ln\theta - \ln\hat{\theta}_{EN2}\right)\right]$$

**3- Fisher information Prior**

We derive the following estimators under the entropy loss function:

$$\hat{\theta}_{EN3} = \left[ K \int_{0}^{\infty} E_{\theta} \frac{\theta^{3m-p} B^{n-m}}{(\theta+2)^{n}} \prod_{i=1}^{m} A_{i} e^{-\theta y_{i}} d\theta \right]^{-\frac{1}{p}}$$

Following that, the associated posterior risk are

$$PR\left(\hat{\theta}_{EN3}\right) = p\left[\mathrm{E}_{\pi_3}\left(\ln\theta - \ln\hat{\theta}_{EN3}\right)\right]$$

**(C) The Linex loss function.**

**1- Informative Prior**

The following estimators are obtained under the Linex loss function:

$$\hat{\theta}_{LI1} = -\frac{1}{r} \ln \left[ K \int_{0}^{\infty} \frac{\theta^{3m+b-1} B^{n-m}}{(\theta+2)^{n}} \mathrm{e}^{-\theta(a+r)} \prod_{i=1}^{m} A_{i} e^{-\theta y_{i}} d\theta \right]$$

and the associated risk to the posterior include

$$PR\left(\hat{\theta}_{LI1}\right) = r\left[\hat{\theta}_{Q1} - \hat{\theta}_{LI1}\right]$$

**2- Non Informative Prior**

The estimators that we obtain under the Linex loss function are as follows:

$$\hat{\theta}_{LI2} = -\frac{1}{r} \ln \left[ K \int_{0}^{\infty} \frac{\theta^{3m-1} B^{n-m}}{(\theta+2)^{n}} \mathrm{e}^{-\theta r} \prod_{i=1}^{m} A_{i} e^{-\theta y_{i}} d\theta \right]$$

and the associated risk to the posterior include

$$PR\left(\hat{\theta}_{LI2}\right) = r\left[\hat{\theta}_{Q2} - \hat{\theta}_{LI2}\right]$$

**3- Fisher information Prior**

The estimators that we obtain under the Linex loss function are as follows:

$$\hat{\theta}_{LI3} = -\frac{1}{r}\ln\left[K\int_{0}^{\infty} E_{\theta}\frac{\theta^{3m}B^{n-m}}{(\theta+2)^{n}}\mathrm{e}^{-\theta r}\prod_{i=1}^{m}A_i e^{-\theta y_i}d\theta\right]$$

and the associated risk to the posterior include

$$PR\left(\hat{\theta}_{LI3}\right) = r\left[\hat{\theta}_{Q3} - \hat{\theta}_{LI3}\right]$$

### 3.2.4 Estimation using Bayesian under balanced loss functions:

**Balanced loss function**

We employ the balanced criterion, often known as the equilibrium criterion, which was established by Zellner in 1994. Increasing precision and conformance in the estimating process is the goal of reaching equilibrium in the loss function. The aforementioned loss functions are regarded as unbalanced loss functions.

Zellner's formula allows the loss function to be balanced in addition to the symmetry criterion in the following ways:

$$L_{L,\omega,\hat{\theta}_Q}\left(\hat{\theta}_{QB},\theta\right) = \omega L\left(\hat{\theta}_Q,\hat{\theta}_{QB}\right) + (1-\omega)L\left(\hat{\theta}_{QB},\theta\right)$$

$L_{L,\omega,\hat{\theta}_Q}\left(\hat{\theta}_{QB},\theta\right)$ Balanced loss function

$\omega$ weighted coefficient, $w \in (0,1)$

$\theta_0$ Primary estimator for the parameter $\theta$ depends on the observations

$L\left(\hat{\theta}_Q,\hat{\theta}_{QB}\right)$Unbalanced loss function.

$L\left(\hat{\theta}_Q,\hat{\theta}_{QB}\right)$Unbalanced loss function for the likelihood function..

It is evident that the initial estimator $\theta_0$ and the weighted coefficient $(\theta)$ have a significant influence on the balanced loss function.

The next step will be to progressively find the Bayes estimators under the balanced loss function as follows:

**(A) The generalized quadratic losss function**

The general formula of the balanced generalized quadratic loss function :
$L_{L,\omega,\hat{\theta}_{GQ}}\left(\hat{\theta}_{GQB}-\theta\right)=\omega L\left(\hat{\theta}_{GQ},\hat{\theta}_{GQB}\right)+(1-\omega)L\left(\hat{\theta}_{GQB}-\theta\right)$

**1- Informative Prior**

The following formula provides the Bayes estimator under the balanced generalized quadratic loss function:

$$\hat{\theta}_{QB1}=\frac{\omega\left[\hat{\theta}_{Q1}\right]^{\alpha}+(1-\omega)\,\mathrm{E}_{\pi_1}\left(\theta^{\alpha}\right)}{\omega\left[\hat{\theta}_{Q1}\right]^{\alpha-1}+(1-\omega)\,\mathrm{E}_{\pi_1}\left(\theta^{\alpha-1}\right)}$$

and the associated risk to the posterior include

$$PR\left(\hat{\theta}_{QB1}\right)=\mathrm{E}^{*}_{\pi_1}\left(\tau(\theta)\left(\theta-\hat{\theta}QB1\right)\right)$$

**2- Non Informative Prior**

The following formula provides the Bayes estimator under the balanced generalized quadratic loss function:

$$\hat{\theta}_{QB2}=\frac{\omega\left[\hat{\theta}_{Q2}\right]^{\alpha}+(1-\omega)\,\mathrm{E}_{\pi_2}\left(\theta^{\alpha}\right)}{\omega\left[\hat{\theta}_{Q2}\right]^{\alpha-1}+(1-\omega)\,\mathrm{E}_{\pi_2}\left(\theta^{\alpha-1}\right)}$$

and the associated risk to the posterior include

$$PR\left(\hat{\theta}_{QB2}\right)=\mathrm{E}^{*}_{\pi_2}\left(\tau(\theta)\left(\theta-\hat{\theta}QB2\right)\right)$$

**3- Fisher information Prior**

The following formula provides the Bayes estimator under the balanced generalized quadratic loss function:

$$\hat{\theta}_{QB3}=\frac{\omega\left[\hat{\theta}_{Q3}\right]^{\alpha}+(1-\omega)\,\mathrm{E}_{\pi_3}\left(\theta^{\alpha}\right)}{\omega\left[\hat{\theta}_{Q3}\right]^{\alpha-1}+(1-\omega)\,\mathrm{E}_{\pi_3}\left(\theta^{\alpha-1}\right)}$$

and the associated risk to the posterior include

$$PR\left(\hat{\theta}_{QB3}\right) = \mathrm{E}^{*}_{\pi_3}\left(\tau(\theta)\left(\theta - \hat{\theta}QB3\right)\right)$$

**(B) The Entropy loss function**

The balanced entropy loss function's general formula is:

$L_{L,\omega,\hat{\theta}(EN)}\left(\hat{\theta}_{ENB},\theta\right) = \omega L\left(\hat{\theta}_{E},\hat{\theta}_{ENB}\right) + (1-\omega)\,L\left(\hat{\theta}_{ENB},\theta\right)$

**1- Informative Prior**

The following formula provides the Bayes estimator under the balanced entropy loss function:

$$\hat{\theta}_{ENB1} = \left[\frac{\omega}{\left(\hat{\theta}_{EN1}\right)}^{p} + (1-\omega)\,\mathrm{EN}_{\pi_1}\left(\frac{1}{\theta^{p}}\right)\right]^{-\frac{1}{p}}$$

and the associated risk to the posterior include

$$PR\left(\hat{\theta}_{ENB1}\right) = p\left[\mathrm{EN}^{*}_{\pi_1}\left(\ln\theta - \ln\hat{\theta}_{ENB1}\right)\right]$$

**2- Non Informative Prior**

The following formula provides the Bayes estimator under the balanced entropy loss function:

$$\hat{\theta}_{ENB2} = \left[\frac{\omega}{\left(\hat{\theta}_{EN2}\right)}^{p} + (1-\omega)\,\mathrm{EN}_{\pi_2}\left(\frac{1}{\theta^{p}}\right)\right]^{-\frac{1}{p}}$$

and the associated risk to the posterior include

$$PR\left(\hat{\theta}_{ENB2}\right) = p\left[\mathrm{EN}^{*}_{\pi_2}\left(\ln\theta - \ln\hat{\theta}_{ENB2}\right)\right]$$

**3- Fisher information Prior**

The following formula provides the Bayes estimator under the balanced entropy loss function:

$$\hat{\theta}_{ENB3} = \left[ \frac{\omega}{\left(\hat{\theta}_{EN3}\right)^p} + (1-\omega)\,\mathrm{E}_{\pi_3}\left(\frac{1}{\theta^p}\right) \right]^{-\frac{1}{p}}$$

and the associated risk to the posterior include

$$PR\left(\hat{\theta}_{ENB3}\right) = p\left[\mathrm{EN}^*_{\pi_3}\left(\ln\theta - \ln\hat{\theta}_{ENB3}\right)\right]$$

(C) **The Linex loss function.**

The balanced Linex loss function's general formula is:
$L_{L,\omega,\hat{\theta}_L I}\left(\hat{\theta}_{LIB},\theta\right) = \omega L\left(\hat{\theta}_L,\hat{\theta}_{LIB}\right) + (1-\omega)\,L\left(\hat{\theta}_{LIB},\theta\right)$

**1- Informative Prior**

The following formula provides the Bayes estimator under the balanced Linex loss function:

$$\hat{\theta}_{LIB1} = -\frac{1}{r}\ln\left[\omega\exp\left(-r\hat{\theta}_{LI1}\right) + (1-\omega)\,\mathrm{E}_{\pi_1}(\exp(-r\theta))\right]$$

and the associated risk to the posterior include

$$PR\left(\hat{\theta}_{LIB1}\right) = r\left[\hat{\theta}_{QB1} - \hat{\theta}_{LIB1}\right]$$

**2- Non Informative Prior**

$$\hat{\theta}_{LIB2} = -\frac{1}{r}\ln\left[\omega\exp\left(-r\hat{\theta}_{LI2}\right) + (1-\omega)\,\mathrm{E}_{\pi_2}(\exp(-r\theta))\right]$$

and the associated risk to the posterior include

$$PR\left(\hat{\theta}_{LIB2}\right) = r\left[\hat{\theta}_{QB2} - \hat{\theta}_{LIB2}\right]$$

**3- Fisher information Prior**

$$\hat{\theta}_{LIB3} = -\frac{1}{r}\ln\left[\omega \exp\left(-r\hat{\theta}_{L3}\right) + (1-\omega)\,\mathrm{E}_{\pi_3}\left(\exp\left(\,-r\theta\right)\right)\right]$$

and the associated risk to the posterior include

$$PR\left(\hat{\theta}_{LIB3}\right) = r\left[\hat{\theta}_{QB3} - \hat{\theta}_{LIB3}\right]$$

## 3.3 Simulation and results

Sample sizes ($n = 10, 25, 50, 100$), a number of parameter values ($\theta = 0.2, 0.5, 1, 5$), as well as ($a = b = 2$) and ($\alpha = p = r = 2$), were selected for this simulation study. Additionally, ($w = 0.5$) was chosen to override the aligned in the estimation process, meaning that this weighted value would give the initial estimator and Bayes estimator in their customized formulas the same loss. ($K = 1000$) was the number of replications used. The (R3.5.1) program was used to write the simulation program. following the parameter's estimation.

Table 3.9: Estimated Zeghdoudi distribution parameters

| $\theta_0$ | $n$ | $\hat{\theta}_{Q2}$ | $\hat{\theta}_{QB2}$ | $\hat{\theta}_{EN2}$ | $\hat{\theta}_{ENB2}$ | $\hat{\theta}_{Q1}$ | $\hat{\theta}_{QB1}$ | $\hat{\theta}_{EN1}$ | $\hat{\theta}_{ENB1}$ |
|---|---|---|---|---|---|---|---|---|---|
| 0.2 | 10 | 0.2217216 | 0.21563552 | 0.236936799 | 0.22324312 | 0.299624533 | 0.204586986 | 0.208681309 | 0.209115373 |
| | 25 | 0.224128161 | 0.221645598 | 0.229525038 | 0.224344036 | 0.214039847 | 0.216601441 | 0.218425995 | 0.218794515 |
| | 50 | 0.20735369 | 0.206280153 | 0.209590225 | 0.20739842 | 0.203083284 | 0.20414495 | 0.205104525 | 0.20515557 |
| | 100 | 0.207274759 | 0.206738385 | 0.2083694 | 0.207285706 | 0.205100011 | 0.205651012 | 0.206140606 | 0.206171309 |
| 0.5 | 10 | 0.634151472 | 0.602443898 | 0.713420406 | 0.642078365 | 0.517280271 | 0.544008298 | 0.56430575 | 0.567521037 |
| | 25 | 0.60144412 | 0.589415328 | 0.62759396 | 0.602490202 | 0.553135596 | 0.56526102 | 0.574410042 | 0.575898243 |
| | 50 | 0.590509727 | 0.58460463 | 0.602812013 | 0.590755773 | 0.566057243 | 0.572378 | 0.5771156 | 0.57798 |
| | 100 | 0.51623197 | 0.51365081 | 0.521499643 | 0.516284647 | 0.505951 | 0.50851 | 0.51096 | 0.511015 |
| 1 | 10 | 1.508920699 | 1.433474664 | 1.697535786 | 1.527782207 | 1.215023857 | 1.286526243 | 1.325480572 | 1.3417546 |
| | 25 | 1.149113 | 1.12613036 | 1.19907403 | 1.151111069 | 1.05847047 | 1.080809289 | 1.099180873 | 1.10116449 |
| | 50 | 1.059284969 | 1.048692119 | 1.081353406 | 1.059726338 | 1.017403144 | 1.027751207 | 1.037352 | 1.037726 |
| | 100 | 1.075467413 | 1.070090076 | 1.08644157 | 1.075577154 | 1.05363994 | 1.059176339 | 1.064072019 | 1.064392379 |
| 5 | 10 | 5.682602878 | 5.398472734 | 6.392928238 | 5.753635414 | 5.095285492 | 5.104814041 | 5.103947809 | 5.1091452 |
| | 25 | 5.314219249 | 5.207935 | 5.54527226 | 5.323461 | 5.02004674 | 5.06084861 | 5.097740846 | 5.099695662 |
| | 50 | 5.556690528 | 5.501123623 | 5.672454914 | 5.559005815 | 5.332266074 | 5.388911396 | 5.436820311 | 5.441188514 |
| | 100 | 5.380825631 | 5.353921503 | 5.435732015 | 5.381374694 | 5.271585661 | 5.299301518 | 5.323779579 | 5.325398476 |

# Bayesian inference for the new Xexponential distribution and its applications

The Xexponential model, a novel probability model created by combining gamma and exponential models, is presented in this chapter. The study investigates this novel model's mathematical characteristics, such as the quantile function, mode, moments, and moment generation function. Using a Monte Carlo simulation, the study assesses many estimators to determine the distribution's unknown parameter. Additionally, this model's goodness-of-fit test is provided. Additionally, by using the model on an actual data set, the flexibility of the Xexponential model is contrasted with that of other widely used models, including gamma, exponential, Lindley, Shanker, Akash, Zeghdoudi, and Chris-Jerry; The findings show that compared to previous models, the suggested model offers greater flexibility and a better fit. These results imply that the Xexponential model might be valuable for simulating real-world data and might merit more investigation in subsequent studies.

## 4.1 THE NEW DISTRIBUTION'S FORMATION

**Theorem 1**

$$\mathbb{E}(Y) = \int_{\mathfrak{Re}} t f(y) dy \Leftrightarrow \mathbb{E}(Y) = \int_{\mathfrak{Re}} S(y) dy$$

**Proof:** We said that : $\mathbb{E}(Y)=\int\limits_{0}^{\infty} y f(y)\,dy$

$$\mathbb{E}(Y)=\int\limits_{0}^{\infty}\left(\int\limits_{0}^{y} dt\right) f(y)dy=\int\limits_{t=0}^{\infty}\int\limits_{y=t}^{\infty} f(y)dtdy=\int\limits_{t=0}^{\infty}\int\limits_{y=t}^{\infty} f(y)dydt$$
$$=\int\limits_{t=0}^{\infty}[F(y)]_{y=t}^{\infty}dt=\int\limits_{t=0}^{\infty}[F(\infty)-F(t)]dt=\int\limits_{t=0}^{\infty}[1-F(t)]dt$$
$$=\int\limits_{0}^{\infty} S(y)dy$$

**(B)THE Xexponential DISTRIBUTION**

A new statistical family with the probability density function known as the new XLindley distribution (NXL) was recently presented by Khodja et al. [1].

$$g(Y)=\frac{\theta}{2}(1+\theta y)e^{-\theta y},\quad y,\theta>0, \tag{4.26}$$

Our suggested model, known as Xexponential, is derived as (4.26) , when $f(Y)=\frac{S(Y)}{\mathbb{E}(Y)}$

**Theorem 2.** $f(Y)$ is density

- $f(y)\succ 0$
- $\int\limits_{0}^{\infty} f(y)\,dy=1$

**Proof**

- we have $f(y)=\frac{S(y)}{\mathbb{E}(Y)}$ and $S(Y)=\left(\frac{\theta y}{2}+1\right)e^{-\theta y}\succ 0$ ; $\mathbb{E}(Y)=\frac{3}{2\theta}\succ 0$ so $f(y)\succ 0$
- $$I=\int\limits_{0}^{\infty} f(y)\,dy=\int\limits_{0}^{\infty}\frac{S(y)}{\mathbb{E}(y)}dy=\frac{1}{\mathbb{E}(Y)}\int\limits_{0}^{\infty} S(y)\,dy.$$

  using Theorem 1 we can find

  $$I=\frac{1}{\mathbb{E}(Y)}\mathbb{E}(Y)=1$$

SO the Xexponential distribution formula is :

$$f(Y) = \frac{\theta}{3}(\theta y + 2)e^{-\theta y}, \quad y, \theta > 0, \tag{4.27}$$

and it can be obtained by mixture of $f_1(y) \sim Exp(\theta)$ and $f_2(y) \sim g(2, \theta)$ with $p_1 = \frac{2}{3}; p_2 = \frac{1}{3}$

## 4.2 STATISTICAL PROPERTIES

**(A) MODE**

The following two limits, respectively, address the (4.27) behavior at zero and infinity:

$$\lim_{y \to 0} f(y) = \frac{2\theta}{3} and \lim_{y \to \infty} f(y) = 0$$

The first derivative of $f(y)$ is

$$\frac{df(y)}{dy} = -\frac{\theta^2}{3}(1 + \theta y)e^{-\theta y} < 0 \tag{4.28}$$

and its second derivative is

$$\frac{d^2 f(y)}{dy^2} = \frac{\theta^4 y}{3} e^{-\theta y}$$

With $\frac{d^2 f(y)}{dy^2} > 0; y > 0$ and $\frac{d^2 f(y)}{dy^2} < 0; y < 0$ ; where $\left(0, \frac{2\theta}{3}\right)$ is inflection point and the mode of Y is $\frac{2\theta}{3}$

**(B) SURVIVAL AND HAZARD RATE FUNCTION**

The following is the definition of the Y's cumulative distribution function.

$$F_X(x, \theta) = 1 - \left(\frac{\theta x}{3} + 1\right)e^{-\theta x} \tag{4.29}$$

then, the survival function $S_Y(Y,\theta)$ and hazard rate function $h_Y(Y,\theta)$ are defined as follows, respectively, for the Xexponential distribution:

$$S_Y(y,\theta) = 1 - F_Y(y,\theta) = \left(\frac{\theta y}{3} + 1\right)e^{-\theta y} \tag{4.30}$$

$$h_Y(y,\theta) = \frac{f_Y(y,\theta)}{S_Y(y,\theta)} = \frac{y\theta^2 + 2\theta}{\theta y + 3} \tag{4.31}$$

**Proposition 1.**

The $h_Y(y,\theta)$ (4.31) is an increasing function.

**Proof.** we have

$$\frac{dh_Y(y,\theta)}{dy} = \frac{\theta^2}{(\theta y + 3)} \succ 0 \tag{4.32}$$

then, the $h_Y(y,\theta)$ is an increasing function

**(C) MOMENTS**

The following is the definition of the rth moments of the Y.

$$\mu_r = \mathbb{E}(Y^r) = \int_0^\infty y^r f(y)\,dy = \int_0^\infty y^r \tfrac{\theta}{3}(\theta y + 2)\,e^{-\theta y} = \tfrac{1}{3\theta^r}\left[\Gamma(r+2) + 2\Gamma(r+1)\right] \tag{4.33}$$

**Proposition 2.** The following definitions apply to Y's mean, variance, coefficients of variation, skewness, and kurtosis, respectively:

$$\mathbb{E}(Y) = \frac{4}{3\theta}; Var(Y) = \frac{14}{9\theta^2}$$

$$Skewness = \frac{\mathbb{E}(Y^3)}{(Var(Y))^{\frac{3}{2}}} = \frac{\frac{12}{\theta^3}}{\left(\frac{14}{9\theta^2}\right)^{\frac{3}{2}}} = 6,1852 \succ 0$$

$$Kurtosis = \beta_2 = \frac{\mathbb{E}(Y^4)}{(Var(Y))^2} = \frac{\frac{56}{\theta^4}}{\left(\frac{14}{9\theta^2}\right)^2} = 23,1429$$

$$C.V = \gamma = \frac{\sqrt{Var(Y)}}{\mathbb{E}(Y)} = \frac{\sqrt{\frac{14}{9\theta^2}}}{\frac{4}{3\theta}} = \frac{\sqrt{14}}{4}$$

Based on the skewness and kurtosis, the new distribution is right skewed and leptokurtic.

**Theorem 3.** Let $Y \sim XeXp(\theta)$. Then the $median(Y) < \mathbb{E}(Y)$.

**Proof** Let$m = Median(Y)$ and $\mu = \mathbb{E}(Y) = \frac{4}{3\theta}$ Since the c.d.f. is given by (4.29) , it follows that $F(m) = \frac{1}{2}$ and $F(\mu) = 1 - \frac{13}{9}e^{-\frac{4}{3}}$ Note that $\frac{1}{2} \prec 1 - \frac{13}{9}e^{-\frac{4}{3}}$ . Finally, since F(Y) is an increasing function in $y \succ 0$ and $\theta \succ 0$ , we have $m \prec \mu$

**(D) ENTROPY**

It is widely accepted that the degree of uncertainty in a probability distribution can be determined using entropy and information. However, the properties of entropy have led to the creation of several correlations.

The random variable's entropy Y quantifies the fluctuation in the uncertainty. The following is the definition of Rényi's entropy.

$$I_R(s) = \frac{1}{1-s}\log\left\{\int_0^\infty f^s(y)\,dy\right\}$$

where $s(integer) \succ 0$ and $s \neq 1$ For the now distribution , we have

$$I_R(s) = \tfrac{1}{1-s}\log\left\{\int_0^\infty \left(\tfrac{\theta}{3}(\theta y+2)e^{-\theta y}\right)^s dy\right\} = \tfrac{1}{1-s}\log\left\{\int_0^\infty \tfrac{\theta^s}{3^s}(\theta y+2)^s e^{-\theta y s}dy\right\}$$

where

$$\int_0^\infty \tfrac{\theta^s}{3^s}(\theta y+2)^s e^{-\theta y s}dy = \tfrac{\theta^s}{3^s}\sum_{k=0}^{s}\tfrac{s!}{(s-k)!}\int_0^\infty (\theta y)^{s-k}2^k e^{-\theta y s}dy = \tfrac{\theta^s}{3^s}\sum_{k=0}^{s}2^k\tfrac{s!\theta^{k-s}\Gamma(s-k+1)}{(s-k)!(\theta s)^{s-k+1}}$$

The following is the calculation of the Rényi entropy for the Now distribution.

$$I_R(m) = \frac{1}{1-m}\log\left\{\frac{\theta^m}{3^m}\sum_{k=0}^{m}2^k\frac{m!\theta^{k-m}\Gamma(m-k+1)}{(m-k)!(\theta m)^{m-k+1}}\right\}$$

**(E) STRESS-STRENGTH RELIABILITY**

The longevity of a component subjected to random stress (represented by variable Y) and random strength (represented by variable Z) is referred to as stress-strength reliability. The component will fail right away if the stress is greater than its strength. If not, the component will continue to operate as intended until the stress surpasses its strength.

The stress-strength parameter, represented statistically as $R = P[Y \prec Z]$, quantifies component reliability. This idea is widely applicable in many domains, but it is especially employed in engineering to investigate the aging of concrete pressure vessels, the static fatigue of ceramic components, the degradation of rocket motors, and other related problems. For a component having independent strength and stress random variables Z and Y, one method of calculating the stress-strength reliability R is as follows: The following is the distribution with parameters $\theta_1$ and $\theta_2$, respectively.

$$R = P[Y \prec Z] = \int_0^\infty P(Y \prec Z|Z=z)f_Z(z,\theta)\,dz$$
$$= \int_0^\infty f_Z(z,\theta_1)F_Z(z,\theta_2)\,dz$$
$$= \int_0^\infty \left(\tfrac{\theta_1}{3}(\theta_1 z+2)e^{-\theta_1 z}\right)\left(1-\left(\tfrac{\theta_2 z}{3}+1\right)e^{-\theta_2 z}\right)dz$$

$$R = \frac{27\theta_1^3 + 18\theta_2^3 + 49\theta_1^2\theta_2 + 3\theta_2^2\theta_1 + 35\theta_2^2}{9(\theta_1+\theta_2)^3}$$

**(F) STOCHASTIC ORDERING**

**Definition 1.** Let Z and Y be two random variables. The following therefore states that Z is smaller than Y.

- Stochastic order $(Z <_s Y)$ , if $F_Z(t) < F_Y(t)$ ,$\forall t$
- Convex order $(Z \leq_{cz} Y)$ , if for all convex functions $\phi$ and .provided expectation exist $\mathrm{E}(\phi(Z)) \leq \mathrm{E}(\phi(Y))$
- Hazard rate order $(Z \leq_{hr} Y)$ , if $h_Z(t) \geq h_Y(t)$ ; $\forall t$
- Likelihood ratio order $(Z <_{lr} Y)$ ; if $\frac{f_Z(t)}{f_Y(t)}$ is decreasing in t

**Remark 1.** Likelihood ratio order $\Rightarrow$ Hazard rate order $\Rightarrow$ Stochastic order If $\mathbb{E}(X) = \mathbb{E}(Y)$ then Convex order $\Leftrightarrow$ Stochastic order.

**Theorem 4.** Let $Z_i \sim XeXp(\theta_i)$ ; $i = 1;2$ be two random variables. If $\theta_1 \geq \theta_2$ , then : $Z <_{lr} Y$ ; $Z \leq_{hr} Y$ ; $Z <_s Y$ and $Z \leq_{cz} Y$ .

**Proof.** we have

$$\frac{f_Z(t)}{f_Y(t)} = \frac{\theta_1(\theta_1 t + 2)}{\theta_2(\theta_2 t + 2)} e^{-(\theta_1 - \theta_2)t}$$

For simplification, we use $\ln\left(\frac{f_Z(t)}{f_Y(t)}\right)$ Now, we can find:

$$\frac{d \ln\left(\frac{f_Z(t)}{f_Y(t)}\right)}{dt} = \frac{\theta_2\left(\theta_1\theta_2 t^2 + (2\theta_2 + \theta_1)t + 2\right) - \theta_1\left(\theta_1\theta_2 t^2 + (2\theta_1 + \theta_2)t + 2\right)}{(\theta_1 t + 2)(\theta_2 t + 2)}$$

where $z \succ 0$ and $\theta_1, \theta_2 \succ 0$ , and if $\theta_1 > \theta_2$ , we have $\frac{d \ln\left(\frac{f_Z(t)}{f_Y(t)}\right)}{dt} \leq 0$ . This means that $Z <_{lr} Y$ Also, according to Remark 1 the theorem is proved.

**(G) LORENZ CURVE**

Let Y be a random variable $f_Y(y)$, and the Lorenz curve L is provided by the $F_Y(y)$.

$$L(F(Y)) = \frac{\int_{-\infty}^{y} tf(t)\,dt}{\mathbb{E}(Y)}$$

where the average is indicated by $\mathbb{E}(Y)$. Plotting the Lorenz curve $L(F)$ as a parametric function in $y$ is then possible: $L(y)$ vs. $F(y)$. The number calculated here is referred to in other contexts as the size-biased distribution, and it plays a significant part in renewal theory. It is available for Xexponential.

$$\int_{0}^{y} tf(t)dt = \frac{4}{3\theta} - \frac{(\theta y+2)^2}{3\theta}$$

The following is how we derive the Lorenz curve for the Xexponential:

$$L(p) = 1 - \frac{(1-p)(\theta y+2)^2}{\left(\frac{\theta y}{3}+1\right)}$$

Where $y = F^{-1}(p)$ with $F(.)$ given by (4.29)

**(H) EXTREME ORDER STATISTICS OF Xexponential**

Let $Y_1,...,Y_n$ a sample of n random variables that follow the Xexponential and if $\bar{Y} = \frac{\sum_{i=1}^{n} Y_i}{n}$ represents the sample mean then by the central limit theorem $\frac{\sqrt{n}\left(\bar{Y}-\mathbb{E}(Y)\right)}{\sqrt{Var(Y)}}$ approximates the standard normal distribution when $n \to \infty$

**Theorem 5.** We will study the asymptotic law of extreme values $S_n = \max(Y_1,...,Y_n)$ and $I_n = \min(Y_1,...,Y_n)$ For the distribution function defined in (4.29) , we find that

$$\lim_{t\to\infty} \frac{1-F_Y(t+y)}{1-F_Y(t)} = e^{-\theta y}$$

and

$$\lim_{t\to 0} \frac{F_Y(ty)}{F_Y(t)} = y$$

According to Theorem 1.6.2 in Leadbetter et al. [10] that there must be norming constants $\alpha_n > 0$ , $\beta_n, \gamma_n > 0$ and $\delta_n$ such that

$$P\{\alpha_n (S_n - \beta_n) \leq y\} \rightarrow \exp(-\exp(-\theta y))$$

and

$$P\{\gamma_n (I_n - \delta_n) \leq y\} \rightarrow 1 - \exp(-y)$$

as $n \rightarrow \infty$ . Using Corollary 1.6.3 in Leadbetter et al. [10], we can see that $\alpha_n = 1$ and $\beta_n = F^{-1}\left(1 - \frac{1}{n}\right)$ with $F(.)$ given by (4.29)

## 4.3 Xexponential Parameter Estimation

### 4.3.1 Maximum likelihood function

The suggested model parameter will be estimated in this section using a variety of techniques. As we will show, we will maximize or minimize an objective function to achieve the estimator.

The likelihood function is defined as follows for an n-sample $(Y_1, ..., Y_n)$ produced by the now distribution:

$$L(Y, \theta) = \frac{\theta^n}{3^n}\left[\prod_{i=1}^{n}(\theta y_i + 2)\right] e^{-\theta \sum_{i=1}^{n} y_i}$$

The next value is the log-likelihood function

$$\ln L(y, \theta) = n[\ln(\theta) - \ln(3)] + \sum_{i=1}^{n} \ln(\theta y_i + 2) - \theta \sum_{i=1}^{n} y_i \tag{4.34}$$

With regard to $\theta$, the derivatives of $\ln L(y, \theta)$ are calculated as follows.

$$\frac{d \ln L(y, \theta)}{d\theta} = \frac{n}{\theta} + \sum_{i=1}^{n} \frac{y_i}{\theta y_i + 2} - \sum_{i=1}^{n} y_i$$

The maximum likelihood estimation MLE of$\theta$, $\overset{\wedge}{\theta}_{MLE}$ can be obtained by either solving the non-linear equation $\frac{d\ln L(y,\theta)}{d\theta} = 0$
, or by directly maximizing equation (4.34) with regard to $\theta$. $\frac{d\ln L(y,\theta)}{d\theta} = 0$ appears to be an analytically unsolvable problem. To get close answers, we'll use numerical techniques. To determine the estimated value of the maximum likelihood estimator $\overset{\wedge}{\theta}_{MLE}$ of the parameter $\theta$, we will utilize the R package BB. Varadhan and Gilbert (2010) demonstrate how to solve non-linear systems of equations using the R package BB.

### 4.3.2 Estimation using Bayesian under different loss functions.

One metric used to assess accuracy in the Bayesian estimation process is the loss function, which is the amount of loss that results from a Bayes judgment on an unknown parameter. $\left(\overset{\wedge}{\theta} - \theta\right)$ is a measure of the difference between the estimated and actual values of this parameter. It should have a real, non-negative value and is typically represented by the symbol $L\left(\overset{\wedge}{\theta}, \theta\right)$.

**(A). Prior and posterior distributions.**

The unknown parameter is seen as a random variable (r.v.) rather than a fixed constant in the Bayesian approach. From here, one can assume previous distributions of the unknown parameter in order to integrate the variation in the parameter. We suppose that the parameter $\theta$ follows the Gamma distribution as a previous distribution.

$$\pi(\theta) = \frac{a^b \theta^{b-1}}{\Gamma(b)} e^{-a\theta}, \quad \theta > 0, a, b > 0, \tag{4.35}$$

The posterior density is then :

$$\pi(\theta|Y) = \frac{\pi(\theta) L(y,\theta)}{\int_0^\infty \pi(\theta) L(y,\theta) d\theta}$$

SO

$$\pi(\theta|Y) = K\theta^{n+b-1} \prod_{i=1}^{n} (\theta y_i + 2) e^{-\theta\left(a + \sum_{i=1}^{n} y_i\right)} \tag{4.36}$$

where the normalizing constant is $K$.

**(B).Unbalanced Bayesian estimation using different loss functions and their posterior risks.**

Loss functions are often divided into two primary categories based on symmetry criteria: The first is a symmetric loss function, which assumes that the loss in the positive direction is equal to the loss in the negative direction.One of the most widely used symmetric loss functions is the quadratic loss function.

The entropy loss function is one example of the second class of loss functions, known as asymmetric loss functions, in which we suppose that the amount of loss under Bayes decisions does not necessarily have to be equal in both the positive and negative directions.The next step will be to find Bayes estimators under imbalanced loss functions in the following order:

**1)-The quadratic loss function in general**

The generalized quadratic loss function is defined as $P\left(\hat{\theta},\theta\right) = \tau(\theta)\left(\hat{\theta}-\theta\right)^2$ where $\tau(\theta) = \theta^{\alpha-1}$

The following formulas provide the Bayes estimators for the generalized quadratic loss function:

$$\hat{\theta}_Q = \frac{\int_0^\infty \theta^{\alpha+n+b-1} \prod_{i=1}^{n} (\theta y_i + 2)\, e^{-\theta\left(a+\sum_{i=1}^{n} y_i\right)} d\theta}{\int_0^\infty \theta^{\alpha+n+b-2} \prod_{i=1}^{n} (\theta y_i + 2)\, e^{-\theta\left(a+\sum_{i=1}^{n} y_i\right)} d\theta} \tag{4.37}$$

Next, the associated posterior risks are

$$PR\left(\hat{\theta}_Q\right) = \mathrm{E}_\pi\left(\theta^{\alpha+1}\right) - 2\hat{\theta}_Q \mathrm{E}_\pi\left(\theta^{\alpha}\right) + \hat{\theta}_Q^2 \mathrm{E}_\pi\left(\theta^{\alpha-1}\right)$$

**2)-The Entropy loss function**

$P\left(\hat{\theta},\theta\right) = \left(\frac{\hat{\theta}}{\theta}\right)^p - p\ln\left(\frac{\hat{\theta}}{\theta}\right)^p - 1$ The following estimators are obtained under the entropy loss function:

$$\hat{\theta}_E N = \left[K \int_0^\infty \theta^{n+b-p-1} \prod_{i=1}^{n} (\theta y_i + 2)\, e^{-\theta\left(a+\sum_{i=1}^{n} y_i\right)} d\theta\right] \tag{4.38}$$

Next, the associated posterior risks are

$$PR\left(\hat{\theta}_E N\right) = p\mathrm{E}_\pi\left[\ln(\theta) - \ln\left(\hat{\theta}_E N\right)\right]$$

**(C). Bayesian estimation using balanced loss functions and their posterior dangers.**

The symmetry criterion discussed above is not the only criterion used to classify loss functions; Zellner (1994) established a more thorough standard known as the balanced criterion, or equilibrium criterion. Increasing precision and conformance in the estimating process is the goal of reaching equilibrium in the loss function. The aforementioned loss functions are regarded as unbalanced loss functions.

In addition to the symmetry condition, Zellner's formula allows the loss function to be balanced in the following way:

$$L_{\varphi,\omega,\theta_0}\left(\hat{\theta},\theta\right) = \omega\varphi\left(\hat{\theta},\theta_0\right) + (1-\omega)\,\varphi\left(\hat{\theta},\theta\right) \tag{4.39}$$

where

$L_{\varphi,\omega,\theta_0}$ : Balanced loss function.

$\varphi\left(\hat{\theta},\theta\right)$ : Unbalanced loss function.

$\theta_0$ :Primary estimator for the paramete $\theta$ depends on the observations.

$\omega$ weighted coefficient, $\omega \in (0,1)$

now we will find bayes estimators under unbalanced loss functions sequentially as follows :

**1)-The generalized quadratic losss function**

The following formula provides the Bayes estimators under the balanced generalized quadratic loss function:

$$\hat{\theta}_{QB} = \frac{\omega\left[\hat{\theta}_Q\right]^{\alpha} + (1-\omega)\,\mathrm{E}_\pi\left[\theta^{\alpha}\right]}{\omega\left[\hat{\theta}_Q\right]^{\alpha-1} + (1-\omega)\,\mathrm{E}_\pi\left[\theta^{\alpha-1}\right]} \tag{4.40}$$

and the corresponding posterior risks are

$$PR\left(\overset{\wedge}{\theta}_{QB}\right) = \mathrm{E}_{\pi^*}\left[\theta^{\alpha-1}\left(\theta - \overset{\wedge}{\theta}_{QB}\right)^2\right]$$

**2)-The Entropy loss function**

The balanced entropy loss function's Bayes estimator is provided by the formula :

$$\overset{\wedge}{\theta}_{ENB} = \left[\frac{\omega}{\left(\overset{\wedge}{\theta}_{E}N\right)^p} + (1-\omega)\,\mathrm{E}_{\pi}\left(\frac{1}{\theta^p}\right)\right]^{-\frac{1}{p}} \tag{4.41}$$

and the corresponding posterior risks are

$$PR\left(\overset{\wedge}{\theta}_{ENB}\right) = p\left[\mathrm{E}_{\pi^*}\left(\ln\theta - \ln\overset{\wedge}{\theta}_{ENB}\right)\right]$$

## 4.4 NUMERICAL SIMULATION

We will use an MCMC simulation method to evaluate the performance of the suggested Bayes estimators with the MLE estimators under the balanced loss function. Assuming that $\theta = 1$ and $a = b = 1$, we then suppose that $\theta = 0.1$ and $a = 0.5$ $b = 0.1$. Using varying sample sizes ( $n = 20, 40, 80, 150, 300$ respectively), we get the following outcomes.

### 4.4.1 Likelihood Estimation

The values of the estimators employing the function BB solve are shown in the following table, and the outcomes are as follows:

**(A).** $\theta = 1$ **and** $a = b = 1$

**(B).** $\theta = 0.1$ **and** $a = 0.5$ $b = 0.1$

| $n$ | Parameter | $\overset{\wedge}{\theta}_M$ |
|---|---|---|
| 20 | $\theta$ | 1,1278(0,0061) |
| 40 | $\theta$ | 1,1134(0,0076) |
| 80 | $\theta$ | 1,1930(0,0053) |
| 150 | $\theta$ | 1,2394(0,0065 |
| 300 | $\theta$ | 1,2872(0,0078) |

Table 4.10: the quadratic error MLE of the parameters for $\theta = 1$ and $a = b = 1$ (in brackets).

| $n$ | Parameter | $\overset{\wedge}{\theta}_M$ |
|---|---|---|
| 20 | $\theta$ | 0,08945(0,0021 |
| 40 | $\theta$ | 0,08320(0,0043) |
| 80 | $\theta$ | 0,1951(0,0031) |
| 150 | $\theta$ | 0,1973(0,003 |
| 300 | $\theta$ | 01921(0,0011) |

Table 4.11: the quadratic error MLE of the parameters with $\theta = 0.1$ and $a = 0.5$ $b = 0.1$ (in brackets).

The projected values of $\theta$ do not closely resemble the actual value of $\theta$, we note.

### 4.4.2 Bayesian Estimation

**(A). the balanced generalized quadratic loss function**

utilizing the generalized quadratic loss function that is balanced

| $N$ | Parameter | $\alpha$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | −2,5 | -2 | -1 | −0,5 | 0,5 | 1 | 1,5 | 2 |
| 20 | $\theta$ | 0,6267 (0,0047) | 0,8961 (0,0054) | 0,9233 (0,0021) | 0,9723 (0,0009) | 0,8494 (0,0084) | 0,7609 (0,0086) | 0,8351 (0,0057) | 0,1276 (0,0067) |
| 40 | $\theta$ | 0,6490 (0,0089) | 0,7990 (0,0087) | 0,9181 (0,0005) | 0,9798 (0,0008) | 0,7510 (0,0085) | 0,7575 (0,0087) | 0,6743 (0,0073) | 0,1099 (0,0088) |
| 80 | $\theta$ | 0,8745 (0,0076) | 0,8950 (0,0075) | 0,8199 (0,0015) | 0,9998 (0,0006) | 0,8598 (0,0054) | 0,8976 (0,0043) | 0,8654 (0,0034) | 0,1087 (0,0054) |
| 150 | $\theta$ | 0,7825 (0,0041) | 0,0825 (0,0061) | 0,9739 (0,0001) | 0,9878 (0,0569) | 0,7926 (0,0077) | 0,0977 (0,0078) | 0,5632 (0,0081) | 0,0990 (0,0081) |
| 300 | $\theta$ | 0,6432 (0,0011) | 1,2127 (0,0006) | 2,0018 (0,0031) | 0,9796 (0,0003) | 1,0839 (0,0020) | 1,3240 (0,0021) | 1,1243 (0,0032) | 1,1841 (0,0012) |
| | | | | | | | | | |

Table 4.12: The Bayes estimators with $\theta = 1$ and $a = b = 1$ (in brackets) under the balanced generalized quadratic loss function and PR.

| $N$ | Parameter | $\alpha$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | $-2,5$ | -2 | -1 | $-0,5$ | 0,5 | 1 | 1,5 | 2 |
| 20 | $\theta$ | 0,07231 (0,0047) | 0,05961 (0,0054) | 0,09221 (0,0021) | 0,09823 (0,0019) | 0,08494 (0,0084) | 0,07609 (0,0086) | 0,08351 (0,0057) | 0,01276 (0,0067) |
| 40 | $\theta$ | 0,07490 (0,0089) | 0,0710 (0,0087) | 0,09181 (0,0005) | 0,9998 (0,0008) | 0,07510 (0,0085) | 0,07575 (0,0087) | 0,06743 (0,0073) | 0,01099 (0,0088) |
| 80 | $\theta$ | 0,06792 (0,0076) | 0,0850 (0,0075) | 0,08199 (0,0015) | 0,09991 (0,0006) | 0,08598 (0,0054) | 0,08976 (0,0043) | 0,08654 (0,0034) | 0,01087 (0,0054) |
| 150 | $\theta$ | 0,2825 (0,0041) | 0,0825 (0,0061) | 0,0731 (0,0001) | 0,9869 (0,0729) | 0,7926 (0,0077) | 0,0977 (0,0078) | 0,5632 (0,0081) | 0,0990 (0,0081) |
| 300 | $\theta$ | 0,0453 (0,0016) | 0,0127 (0,0016) | 0,0932 (0,0001) | 0,09878 (0,0001) | 0,1839 (0,0020) | 0,1841 (0,0031) | 0,1232 (0,0042) | 0,0741 (0,0042) |

Table 4.13: Under the balanced generalized quadratic loss function and PR, the Bayes estimators are $\theta = 0.1$ and $a = 0.5$ $b = 0.1$ (in brackets).

We note that the optimal posterior risk is given by the value $\alpha = -0,5$. Additionally, when n is high, we get the minimum appropriate posterior risk.

**(B). The balanced Entropy loss function**

The balanced entropy loss function is used.

| $n$ | Parameter | | $P$ | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | -2 | $-1,5$ | -1 | $-0,5$ | 0,5 | 1 | 1,5 | 2 |
| 20 | $\theta$ | 1,0990 (0,0044) | 1,1181 (0,0003) | 0,7209 (0,0171) | 0,8833 (0,0333) | 0,6834 (0,0433) | 1,1188 (0,0060) | 0,8765 (0,034) | 0,9076 (0,1153) |
| 40 | $\theta$ | 1,0942 (0,008) | 1,1067 (0,0091) | 1,1041 (0,0009) | 1,7981 (0,0038) | 1,6981 (0,0038) | 1,1194 (0,0081) | 1,7053 (0,0035) | 1,7697 (0,0099) |
| 80 | $\theta$ | 1,3892 (0,100) | 1,2387 (0,0033) | 0,7645 (0,0101) | 0,8965 (0,0343) | 0,5623 (0,0653) | 1,0987 (0,0080) | 0,6759 (0,0002) | 0,8765 (0,1003) |
| 150 | $\theta$ | 1,3990 (0,1644) | 1,2188 (0,1443) | 0,6205 (0,0171) | 0,7830 (0,0733) | 0,4830 (0,0733) | 1,1088 (0,0070) | 0,6701 (0,0667) | 0,7654 (0,1173) |
| 300 | $\theta$ | 1,2144 (0,0019) | 1,2179 (0,0017) | 1,2167 (0,0001) | 1,2148 (0,0009) | 1,2148 (0,0009) | 1,1038 (0,0018) | 1,0969 (0,0009) | 1,0886 (0,0031) |

Table 4.14: The Bayes estimators with $\theta = 1$ and $a = b = 1$ (in brackets) under the balanced entropy loss function and PR.

| $n$ | Parameter | | $P$ | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | -2 | $-1,5$ | -1 | $-0,5$ | 0,5 | 1 | 1,5 | 2 |
| 20 | $\theta$ | 0,06911 (0,0004) | 0,4189 (0,0013) | 0,1409 (0,0071) | 0,1830 (0,003) | 0,0831 (0,0033) | 00982 (0,0003) | 0,8465 (0,004) | 0,1076 (0,1003) |
| 40 | $\theta$ | 0,0942 (0,082) | 0,1067 (0,0001) | 0,1111 (0,0019) | 0,1981 (0,0008) | 0,0582 (0,0038) | 0,1194 (0,0001) | 1,7053 (0,0015) | 0,7697 (0,0009) |
| 80 | $\theta$ | 0,3892 (0,100) | 02387 (0,0033) | 0,1045 (0,0101) | 0,8455 (0,0343) | 0,2623 (0,0653) | 0,0994 (0,0009) | 0,6759 (0,0002) | 0,1765 (0,1003) |
| 150 | $\theta$ | 0,1943 (0,1004) | 0,2188 (0,1401) | 0,1005 (0,0101) | 0,7830 (0,070) | 0,4830 (0,0033) | 0,1022 (0,0020) | 0,6701 (0,021) | 0,114 (0,0123) |
| 300 | $\theta$ | 0,2144 (0,0011) | 0,2179 (0,0012) | 0,1167 (0,0009) | 0,2148 (0,0005) | 0,2148 (0,0003) | 0,1002 (0,0011) | 0,0969 (0,0009) | 0,0886 (0,0031) |

Table 4.15: The Bayes estimators with $\theta = 0.1$ and $a = 0.5$ $b = 0.1$ (in brackets) with the balanced entropy loss function and PR.

We can observe that the values $p = 1$, $n = 300$ offer the best posterior risk in the following table.

| $n$ | Parameter | Entropy ($P =$ 1 | 1) $Q(\lambda$ $= 0,5)$ |
|---|---|---|---|
| 20 | | 1,1188 $(0,0060)$ | 0,9723 $(0,0009)$ |
| 40 | $\theta$ | 1,1194 $(0,0081)$ | 0,9798 $(0,0008)$ |
| 80 | | 1,0987 $(0,0080)$ | 0,9898 $(0,0006)$ |
| 150 | $\theta$ | 1,1088 $(0,0070)$ | 0,9895 $(0,0729)$ |
| 300 | $\theta$ | 1,1038 $(0,0018)$ | 0,9899 $(0,0001)$ |

Table 4.16: Under the two loss functions with $\theta = 1$ and $a = b = 1$, Bays estimators and PR (in brackets) .

| $n$ | Parameter | Entropy ($P =$ 1 | 1)$GQ(\lambda$ $= 0,5)$ |
|---|---|---|---|
| 20 | | 00982 $(0,0003)$ | 0,09823 $(0,0019)$ |
| 40 | $\theta$ | 0,1194 $(0,0001)$ | 0,9998 $(0,0008)$ |
| 80 | | 0,0994 $(0,0009)$ | 0,09991 $(0,0006)$ |
| 150 | $\theta$ | 0,1022 $(0,0020)$ | 0,9869 $(0,0729)$ |
| 300 | $\theta$ | 0,1002 $(0,0011)$ | 0,09878 $(0,0001)$ |

Table 4.17: Under the two loss functions with $\theta = 0.1$ and $a = 0.5$ $b = 0.1$, Bays estimators and PR (in brackets) .

Pitman's proximity criterion was used to compare the estimators.

### 4.4.3 THE ESTIMATION METHODS COMPARISON

In this subsection, we suggest contrasting the maximum likelihood estimators with the best Bayesian estimators found above. We suggest using the following criteria for this. The integrated mean square error (IMSE) and Pitman closeness (Pitman 1937; Fuller, 1982; and Dozani, 2012) are defined as follows:

**Definition 2.**

An estimator $\theta_1$ of a parameter $\theta$ dominates in the sense of Pitman closeness criterions another estimator $\theta_2$ , if for all $\theta \in \Theta$ :

$$P_\theta[|\theta_1 - \theta| < |\theta_2 - \theta|] > \frac{1}{2}.$$

Consider the estimators $\theta_i$ , $i \in \{1,...,N\}$ obtained with N samples of model. In the following, we present the values of the Pitman probabilities which allow us to compare the Bayesian estimators with the MLE under the tow loss function

The table 5 where should as follow, when the probability is greater than $-0,5$ , the Bayesian estimators is better than the MLE estimators. Then we notice that, according to this criterion:

| $n$ | $EN(p=1)$ | $Q(\alpha=-0,5)$ |
|---|---|---|
| 20 | $0,267$ | $0,345$ |
| 40 | $0,779$ | $0,589$ |
| 80 | $0,654$ | $0,564$ |
| 150 | $0,687$ | $0,603$ |
| 300 | $0,698$ | $0,661$ |

Table 4.18: Comparing the estimators using Pitman with $\theta = 1$ and $a = b = 1$ .

| $n$ | $EN(p=1)$ | $Q(\alpha=-0,5)$ |
|---|---|---|
| 20 | $0,467$ | $0,431$ |
| 40 | $0,682$ | $0,613$ |
| 80 | $0,761$ | $0,534$ |
| 150 | $0,734$ | $0,592$ |
| 300 | $0,709$ | $0,598$ |

Table 4.19: Comparing the estimators using Pitman with $\theta = 0.1$ and $a = 0.5$ $b = 0.1$ .

Pitman's criterion states that the generalized quadratic loss function has the best values when compared to the other loss functions, and that the Bayesian estimators of the parameters are superior to the MLE.

## 4.5 Real data analysis

This section uses voltage data to show how the new one-parameter model can be applied in real-world situations. The failure and operating times of a selection of devices from a broader system field-tracking study are represented by the data set. [?] examined the data; for further in-depth details, see [?]he following is the p.d.f. information on competing models:

$$
\begin{array}{lll}
G_G(y) = \frac{1}{\theta^{\alpha}\Gamma(\alpha)} y^{\alpha-1} \exp\left(-\frac{y}{\theta}\right) & , & y, \alpha, \theta > 0 \\
G_W(y) = \frac{\alpha}{\theta}\left(\frac{y}{\theta}\right)^{\alpha-1} \exp\left(-\left(\frac{y}{\theta}\right)^{\alpha}\right) & , & x, \alpha, \theta > 0 \\
G_E(y) = \theta \exp(-\theta y) & , & y, \theta > 0 \\
G_L(y) = \frac{\theta^2}{(1+\theta)^2}(1+y)\exp(-\theta y) & , & y, \theta > 0 \\
G_S(y) = \frac{\theta^2}{\theta^2+1}(\theta+y)\exp(-\theta y) & , & y, \theta > 0 \\
G_A(y) = \frac{\theta^3}{\theta^2+2}\left(1+y^2\right)\exp(-\theta y) & , & y, \theta > 0 \\
G_Z(y) = \frac{y\theta^3}{2+\theta}(1+y)\exp(-\theta y) & , & y, \theta > 0 \\
G_{CJ}(y) = \frac{\theta^2}{2+\theta}\left(1+\theta y^2\right)\exp(-\theta y) & , & y, \theta > 0 \\
G_{y\exp}(y) = \frac{\theta}{2}(1+\theta y)e^{-\theta y}, y, \theta > 0 & , & y, \theta > 0 \\
\hline
\end{array}
$$

For all models based on voltage data, the parameter estimate process is completed using the MLE method. We calculate the MLEs of the model's parameter along with its standard errors (SE), $\widehat{\ell}$, the Akaike Information Criterion (AIC), the Bayesian information criterion (BIC), the consistent AIC (CAIC), the Hannanâ €"Quinn information criterion (HQIC), the Kolmogorovâ €"Smirnov statistic (KS), the Cramér von Mises statistic (CVM), the Anderson-Darling statistics (AD), and the p-value of these statistics (KS p-value, CVM p-value, and AD p-value).

Table 4.20: The voltage data's goodness of fit findings

| | Xexponential | G | E | L | S | A | Z | CJ | XG |
|---|---|---|---|---|---|---|---|---|---|
| $\widehat{\ell}$ | -184.1358 | -185.0207 | -185.2901 | -187.4624 | -187.9747 | -195.5284 | -195.1676 | -192.0710 | -190.7116 |
| AIC | 370.2716 | 374.0413 | 372.5803 | 376.9248 | 377.9493 | 393.0567 | 392.3352 | 386.1419 | 383.4232 |
| BIC | 371.6728 | 376.8437 | 373.9815 | 378.3260 | 379.3505 | 394.4579 | 393.7364 | 387.5431 | 384.8244 |
| CAIC | 370.4144 | 374.4858 | 372.7231 | 377.0677 | 378.0922 | 393.1996 | 392.4781 | 386.2848 | 383.5661 |
| HQIC | 370.7198 | 374.9379 | 373.0285 | 377.3731 | 378.3976 | 393.5050 | 392.7835 | 386.5902 | 383.8715 |
| KS (KS p-value) | 0.2151(0.1246) | 0.2173(0.1178) | 0.2161(0.1214) | 0.2294(0.0851) | 0.2302(0.0831) | 0.2499(0.0471) | 0.2495(0.0477) | 0.2374(0.0679) | 0.2329(0.0772) |
| AD (AD p-value) | 1.8720(0.1085) | 2.0004(0.0922) | 2.0022(0.0920) | 3.1803(0.0225) | 3.3080(0.0194) | 5.9712(0.0010) | 5.8873(0.0011) | 5.2249(0.0023) | 4.7973(0.0037) |
| CVM (CVM p-value) | 0.3260(0.1140) | 0.3464(0.1001) | 0.3678(0.0875) | 0.3842(0.0790) | 0.3891(0.0766) | 0.5029(0.0384) | 0.5003(0.0390) | 0.4642(0.0485) | 0.4471(0.0537) |
| $\theta$ | 0.0086 | 0.0067 | 0.0056 | 0.0112 | 0.0113 | 0.0169 | 0.0169 | 0.0165 | 0.0163 |
| $\alpha$ | | 1.1896 | | | | | | | |
| SE of $\theta$ | 0.0013 | 0.2669 | 0.0010 | 0.0014 | 0.0014 | 0.0018 | 0.0018 | 0.0018 | 0.0018 |
| SE of $\alpha$ | | 0.0018 | | | | | | | |

Table 4.20 displays the modeling results for voltage data. Table 4.20 makes it clear that, when all factors are considered, the Xexponential distribution produces better outcomes than the other 10 models. This finding suggests that the suggested model

outperformed the two-parameter Weibull and gamma distributions. The Xexponential distribution will become a very competitive model since it performs well over a broad range of one-parameter distributions, including Weibull and gamma distributions.

## 4.6 Goodness-of-Fit Test

The Xexponential distribution is regarded as a lifetime model since $x > 0$. This section develops the modified Chi-squared Goodness-of-Fit Test for the Xexponential distribution using the Nikulin-Rao-Robson (NRR) statistic. To provide a large range of sample sizes and for different values of the parameter $\theta$, a simulation study is carried out.

Suppose that the distribution of $n$ independent identically distributed random variables, $X_1, X_2, \cdots, X_n$, belongs to the family of Xexponential distributions. The issue of verifying this hypothesis is $H_0$.

$$H_0 : \mathbf{P}(Y_i \leq x) = F_{_{NXL}}(y,\theta), \quad y > 0, \quad \theta > 0.$$

Using the points, we split the positive real line into $r$ sub-intervals $I_1, I_2, \cdots, I_r$.

$$0 = a_0 < a_1 < \cdots < a_{r-1} < a_r = +\infty, \quad I_i = ]a_{i-1}, a_i], \quad I_i \cap I_j = \emptyset, i \neq j,$$

$$i, j = 1, \cdots, r, \quad \cup_{i=1}^{r} I_i = \mathbf{R}_+,$$

The vector of frequencies $v = (v_1, v_2, \cdots, v_r)^T$ and the probability vector $p(\theta) = (p_1(\theta), p_2(\theta), \cdots, p_r($

where

$$p_i(\theta) = \int_{a_{i-1}}^{a_i} f_{_{NXL}}(y,\theta)dy = F_{_{NXL}}(a_i,\theta) - F_{_{NXL}}(a_{i-1},\theta), \quad i = 1, 2, \cdots, r.$$

Pearson suggested a test based on the so-called Pearson's $\chi^2$ test of the following form to test the hypothesis $H_0$:

$$X_n^2(\theta) = X_n^T(\theta)X_n(\theta) = \sum_{i=1}^{r} \frac{(v_i - np_i(\theta))^2}{np_i(\theta)}, \tag{4.42}$$

where

$$X_n(\theta) = \left( \frac{v_1 - np_1(\theta)}{\sqrt{np_1(\theta)}}, \frac{v_2 - np_2(\theta)}{\sqrt{np_2(\theta)}}, \cdots, \frac{v_r - np_r(\theta)}{\sqrt{np_r(\theta)}} \right)^T.$$

Under $H_0$, if $\theta$ is known, it was shown by K. Pearson in 1900 (see e.g. [18]) that

$$\lim_{n\to\infty} \mathbf{P}(X_n^2(\theta)) \leq y | H_0) = \mathbf{P}(\chi^2_{r-1} \leq y). \tag{4.43}$$

When $X_n^2(\theta) > C_\alpha$, where $C_\alpha$ is the critical value of the Pearson's test, $C_\alpha = \chi^2_{r-1,\alpha}$ is the upper $\alpha-$ quantile of the $\chi^2$ distribution with $r-1$ degrees of freedom, then the hypothesis $H_0$ must be rejected at a significance level $\alpha$.

The sample

The standard Chi-squared Pearson's test (4.42) for continuous distribution with shift and scale parameters was modified in 1973 by [21, 22]. Similarly, [23] had shown the same outcome for exponential families. Since 1988, the test has been known as the Nikulin-Rao-Robson (NRR) test ([25], [18]), and it can be expressed as follows (see [19]):

$$Y_n^2(\hat{\theta}_n) = X_n^2(\hat{\theta}_n) + X_n^T(\hat{\theta}_n)B(\hat{\theta}_n)(I(\hat{\theta}_n) - J(\hat{\theta}_n))^{-1}B^T(\hat{\theta}_n)X_n(\hat{\theta}_n),$$

with

$$B(\theta) = (b_1, b_2, \cdots, b_r)^T, \;\; b_i(\theta) = \frac{1}{\sqrt{p_i(\theta)}}\frac{\partial p_i(\theta)}{\partial \theta}, \quad i = 1, 2, \cdots, r.$$

And $nJ = nB^T(\theta)B(\theta)$ is the Fisher's information matrix of the vector of frequencies,

and $I$ is the the Fisher's information matrix of $X_i$, such that

$$nI = I_n = -E\left(\frac{\partial \ell^2(\theta)}{\partial \theta^2}\right),$$

where

$$\frac{\partial \ell^2(\theta)}{\partial \theta^2} = -\frac{n}{\theta^2} - \sum_{i=1}^{n} \frac{X_i^2}{(2+\theta X_i)^2}.$$

The following describes the asymptotic behavior of the statistics $Y_n^2(\hat{\theta}_n)$:

**Theorem**[[21, 22]]

$$\lim_{n\to\infty} \mathbf{P}\left(Y_n^2(\hat{\theta}_n) \le x|H_0\right) = \mathbf{P}(\chi^2_{r-1} \le x). \tag{4.44}$$

It should be noted that [24], (also see [27], [26], [20]

## 4.7 Simulation Study

We create samples of Xexponential distribution with $n = 20, 30, 50, 80, 100, 150, 200, 300, 500$ and 1000, and for different values of the parameter $\theta = (0.01, 0.5, 2, 5, 10)$, in order to empirically analyze the behavior of the statistic $Y_n^2$.

The significance level for this study is $\alpha = 0.05$. Examine the equiprobability scenario. Every sample $(n, \theta)$ is replicated five hundred thousand times. The empirical confidence level ($E.L.$), which counts the instances in which $Y_n^2 \le C_\alpha = \chi^2_{r-1,1-\alpha}$ divided by 5000, is determined after computing the NRR statistic $Y_n^2$ for each operation.

Theoretically, $EL$

# Conclusion and Outlook

In this study, we introduced and examined a novel Lindley model variant known as the two parameters Upper truncated XL (UXLE) model, and we conducted Bayesian inference using right-censored data. Under a censored scheme, we present the model, give its failure rate and survival function, and determine the maximum probability. A detailed comparison of the two approaches is carried out using the Pitman criterion, and Bayesian estimation under three loss functions is also built and investigated. Particularly noteworthy are the following findings: the generalized quadratic offers the best estimators out of the three loss functions used. Both approaches provide the minimal error (the posterior risk for the Bayesian technique and the quadratic error for the maximum likelihood approach) when the sample size is large. It is typically stressed that the Bayesian methodology and the Maximum Likelihood methods are recommended since a comparison of the two approaches shows that Bayesian estimation outperforms maximum likelihood estimation, especially when the sample size is small.

Also We draw the conclusion that balanced loss functions offer effective Bayesian estimators when (undefined) knowledge about the phenomenon being studied is lacking, as in the case of an uninformative prior, based on the experimental results for estimating the Zaghdoudi distribution parameter. Unbalanced loss functions, on the other hand, perform better when there is an informative prior since balanced loss functions might not be as effective.

Then, in order to model count data, we proposed a novel one-parameter distribution known as Xexponential, and we looked at a particular instance inside this model. We investigated the parameter estimation techniques and statistical characteristics. The simulation shows that the maximum likelihood approach is the most effective one. Data sets on the yearly maximum flood and survival times of patients with breast cancer are used to illustrate the value of the Xexponential distribution.This novel distribution outperformed the functions employed in this paper to produce optimal estimation, such as Weibull, gamma, Lindley, xLindley, etc.

Also, we compared Bayesian estimators under Different loss functions, for each loss function we obtained a suitable parameter which optimises the Bayesian estimation. Additionally, we use the Nikulin-Rao-Robson statistic for the Xexponential distribution to create a modified Chi-square goodness of fit test. In order to attain optimal estimation for future prospects, we can develop a blend of the loss functions employed in this paper. We think modelling actuarial or dependability data with this new model would be intriguing.

In future work, we will study robust Bayesian analysis under a general class of balanced loss functions.

# BIBLIOGRAPHY


[1] A. Chadli, H. Talhi, H. Fellag : *Comparison of the Maximum Likelihood and Bayes Estimators of the parameters and mean Time Between Failure for bivariate exponential Distribution under different loss functions*, Afr. Stat, **8**,(2013), 499 – 514.

[2] A. S. Yadav, E. Altun, and H. M. Yousof : Burr-Hatke Exponential Distribution: A Decreasing Failure Rate Model, Statistical Inference and Applications. Annals of Data Science, 8(2),(2021), 241-260.

[3] E. Pitman (1937). The closest estimates of statistical parameters : *Mathematical proceeding of the Cambridge philosophical society*, **33**(2).

[4] G. G. Hamedani, M. Rasekhi, S. Najibi, H. M. Yousof and M. Alizadeh : *Type II general exponential class of distributions*, Pakistan Journal of Statistics and Operation Research, **15**(2), (2019), 503 – 523.

[5] H. Aiachi, A. Chadli, H. Talhi: *Bayesian Inference of three parameters Burr XII based on progressively type II censored observations*, Journal of Applied Probability and Statistics, **17**(2),(2022), 071 – 092.

[6] H. Talhi, H. Aiachi : *On Truncated Zeghdoudi Distribution : Posterior Analysis under Different Loss Functions for Type II Censored Data*, Pakistan Journal of Statistics and Operation Research, **17**(2),(2021), 497 – 508.

[7] H. Talhi, H. Aiachi and N. Rahmania : *Bayesian estimation of a competing risk model based on Weibull and exponential distributions under right censored data*, Monte Carlo Methods and Applications, **28**(2), (2022), 163 – 174.

[8] H. Zeghdoudi, S. Chouia : *The XLindley Distribution: Properties and Application*, Journal of Statistical Theory and Application. **1 20**(2), (2021), 318 – 327

[9] M. A. Aldahlan, F. Jamal, C. Chesneau, M. Elgarhy and I. Elbatal: *The Truncated Cauchy Power Family of Distributions with Inference and Applications*, Entropy, **22**(3),(2020), 346.

[10] M.E. Ghitany, D.K. Al-Mutairi, and S. Nadarajah : *Zero-truncated Poisson–Lindley distribution and its application*, Math. Comput. Simulation, **79**, (2008a), 279 – 287.

[11] M. J. Jozani, K. F. Davies, and N. Balakrishnan : *Pitman closeness results concerning ranked set sampling*, Statistics and Probability Letters, **82**(12), (2012), 2260 – 2269.

[12] M. M. Mansour, N. S. Butt, H. M. Yousof, S. I. Ansari, and M. Ibrahim : *A Generalization of Reciprocal Exponential Model: Clayton copula, Statistical Properties and Modelling Skewed and Symmetric Real Data Sets*, Pakistan Journal of Statistics and Operation Research, **16**(2), (2020f) 373 – 386.

[13] R.A.R. Bantan ,F. Jamal, C. Chesneau and Mo. Elgarhy :*Truncated Inverted Kumaraswamy Generated Family Of Distributions with Applications*, Entropy, **21**(11) (2019), 1089, https://doi.org/10.3390/e21111089.

[14] R. Varadhan, and P.Gilbert : *An R package for solving a large system of nonlinear equations and for optimizing a high-dimensional nonlinear objective function*, Journal of statistical software, **32**(4), (2010)

[15] W.A. Fuller: *Closeness of estimators*, Encyclopedia of statistical sciences **2**, Wiley, (1982).

[16] F. Monther Al-Badran: *Bayes Estimation under Balanced Loss Functions*, The Journal of Administration and Economics **42**,119, (2019).

[17] ZELLNER, A. (1994).: *Bayesian and Non-Bayesian estimation using balanced loss functions*, Statistical Decision Theory and Methods V; (J.O. Berger and S.S. Gupta Eds). **New York: Springer-Verilog**,337-390

[18] F. Drost, *Asymptotics for Generalized Chi-squared Goodness-of-fit Tests*, Amsterdam: Centre for Mathematics and Computer Sciences,CWI Tracs, 48. (1988)

[19] Greenwood, P. S. and Nikulin, M. . A guide to Chi-squared Testing, John Wiley and Sons, New York. (1996)

[20] M. Mirvaliev, *An investigation of generalized chi-squared type statistics*, Doctoral thesis, Academy of Science of the Republic of Uzbekistan, Tashkent.(2001)

[21] M. Nikulin, *Chi-square Test For Continuous Distributions with Shift and Scale Parameters*, Teor. Veroyatn.Primen., 18, No. 3, (1973a), 559-568.

[22] M. Nikulin, *On a Chi-square Test For Continuous Distributions*,Thery of Probability and Applications, 18 , (1973b), 638-639.

[23] K.C. Rao and D.S. Robson, *A chi-square statistic for goodness-of-fit tests within the exponential family*. Communications in . Statistics, 3 , (1974) 1139-1153.

[24] T.A. Hsuan and D.S.Robson, *The $\chi^2-$ Goodness-of-fit Tests with Moment Type Estimator*. Communications in Statistics - Theory and Methods, 16 , (1976), 1509-1519.

[25] A.W. Van Der Vaart, *Asymptotic Statistics*. Cambridge Series in Statistics and probabilistic Mathematics. Cambridge: Cambridge University Press. (1998)

[26] V. Voinov, R. Alloyarova and N Pya, *Recent Achievemets in Modified Chi-squared Goodness-of-fit Testing*. In:Statistical Models and Methods for Biomedical and Technical Systems,(Eds. F. Vonta, M. Nikulin, N. Limnios, C. Huber.)Birkhäuser, Boston,(2008) p241-258.

[27] V. Voinov, N. Pya and R. Alloyarova, *A Comparative Study of Some Modified Chi-squared Tests*, Communi. in Stat.-Simul. and Comput., 38:2, (2009), 355-367.

.

[28] Patil, G.P. and Rao, C.R. (1977). The weighted distributions: a survey of their applications, Applications of Statistics, 383–405.

[29] Rao, R.C. (1965), The Theory of Least Squares When the Parameters are Stochastic and Its Application to the Analysis of Growth Curves, Biometrika, 52 (3/4), 447-458.

[30] Chouia, S. and Zeghdoudi, H. (2021), *The xlindley distribution: Properties and Application* .Journal of Statistical Theory and Applications, 20(2):318–327.

[31] Cordeiro, G. M., Ortega, E. M., and Nadarajah, S. (2010). *The kumaraswamy weibull distribution with application to failure data*.Journal of the Franklin Institute, 347(8):1399– 1429.

[32] Ghitany, M. E., Atieh, B., and Nadarajah, S. (2008). *Lindley distribution and its application* . Mathematics and computers in simulation, 78(4):493–506.

[33] K., S. K. (2018). *Pranav distribution with properties and its applications* . Biom Biostat Int J, 7(3):244–254.

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

[35] Meeker, W. Q., Escobar, L. A., and Pascual, F. G. (2022). *Statistical methods for reliability data* . John Wiley and Sons.

[36] Messaadia, H. and Zeghdoudi, H. (2018). *Zeghdoudi distribution and its applications* . International Journal of Computing Science and Mathematics, 9(1):58–65.

[37] Onyekwere, C. K. and Obulezi, O. J. (2022). *Chris-jerry distribution and its applications* . Asian Journal of Probability and Statistics, 20(1):16–30.

[38] Sen, S., Maiti, S. S., and Chandra, N. (2016). *The xgamma distribution: statistical properties and application* .Journal of Modern Applied Statistical Methods, 15(1):38.

[39] Shanker, R. (2015a). *Akash distribution and its applications* .Int J Probab Stat, 4(3):65–75.

[40] Shanker, R. (2015b). *Shanker distribution and its applications* .Int J Stat Appl., 5(6):338–348.

[41] Shanker, R. (2016a). *Amarendra distribution and its applications* .American Journal of Mathematics and Statistics, 6(1):44–56.

[42] Shanker, R. (2016b). *Aradhana distribution and its applications* .Int J Probab Stat, 6(1):23–34.

[43] Shanker, R. (2016c). *Devya distribution and its applications* .International Journal of Statistics and Applications, 6(4):189–202.

[44] Shanker, R. (2016d). *Shambhu distribution and its applications* .International Journal of Statistics and Applications, 5(2):48–63.

[45] Shanker, R. (2016e). *Sujatha distribution and its applications* .Statistics in Transition. New Series, 17(3):391–410.

[46] Shanker, R. (2017a). *Rama distribution and its application* .Int J Stat Appl, 7(1):26–35.

[47] Shanker, R. (2017b). *Rani distribution and its application* .Biometrics and Biostatistics International Journal, 6(1):1–10.

[48] Shanker, R. and Shukla, K. K. (2017). *Ishita distribution and its applications*. Biometrics and Biostatistics International Journal, 5(2):1–9.

[49] Balakrishnan N., and Aggarwala R. (2000). Progressive Censoring : Theory, Methods and Applications, Boston : Birkhauser.