People's Democratic Republic of Algeria
Ministry of Higher Education and Scientific Research

University of Saida Dr Moulay Tahar

Thesis submitted for the Academic

Master's degree

Faculty: Mathematics, Computer Science and Telecommunication

Sector : Mathematics

Specialty: Stochastic Analysis, Process Statistics and Applications

Presented by

Moghorbi Asmaa [1]

Supervised by

Dr. Rajaa Hazeb

Theme:

Statistical Inference Based On Progressive Type-II Censoring Data

Defended on 29/06/2025 in feont of jury composed of:

Dr. F. Benziadi	University of Saida Dr. Moulay Tahar	President
Dr. R. Hazeb	University of Saida Dr. Moulay Tahar	Supervisor
Dr. I. Mekkaoui	University of Saida Dr. Moulay Tahar	Examiner

Academic Year: 2024/2025

[1] e-mail: bouananiasmaa044@gmail.com

Dedication

On this occasion, I dedicate my graduation, my success, and all that is beautiful in my life:

*To the source of strength, dignity, and silent sacrifice... (**My beloved father**), whose wisdom guided my steps and whose prayers protected my journey.*

*To the heart that overflowed with unconditional love, patience, and unwavering support... (**My precious mother**), beneath whose feet lies Paradise, and whose presence was my greatest blessing,*

*To my life partner and steadfast supporter... (**My dear husband**), whose love was my safe haven and whose faith in me gave me strength in moments of doubt,*

*To the ones who have always been my safe haven, my companions in joy and sorrow, and the closest to my heart...(**My brothers**),*

*To the graceful heart who treated me like a sister, stood by me with kindness and encouragement, and shared my joys with sincerity... (**My brothers wife**).*

*To a family that welcomed me with warmth and kindness... (**My husbands family**), especially to his sister, whose affection and support I deeply cherish.*

*To **My family, friends and colleagues**... To everyone who has contributed even a letter in my college life..... To them all: I dedicate this work, which I sincerely ask Almighty God to accept....*

Acknowledgments

First and foremost, I thank ALLAH who helped me and gave me the strength, patience, and courage during my years of study.

I want to start by thanking my parents, my husband, and all my family for their love and support all the way.

Above all, I extend my deepest gratitude to my supervisor, Dr. R. Hazeb, for her time, her kind guidance, and the valuable knowledge she shared with me during the writing of this thesis.

Also, I would like to thank the committee members, Dr. F. Benziadi and Dr. I. Mekkaoui, for examining my work.

In addition, I would like to express my deep appreciation to all the professors from whom I have had the honor to learn throughout my academic journey. Each of you has left an indelible mark on my learning and personal growth.

Finally, I would like to thank my classmates for their encouragement and support throughout this master thesis.

Abstract

This study investigates the estimation and prediction of parameters for the Weibull and Alpha Power Weibull (APW) distributions using progressively Type-II censored data. The APW model provides greater flexibility in modeling different hazard rate shapes. Maximum likelihood and Bayesian methods are used to estimate distribution parameters, reliability, and hazard functions. Bayesian estimates are obtained via MCMC under Squared Error and LINEX loss functions. Monte Carlo simulations evaluate the performance of the methods, and real engineering data confirm their effectiveness. Results show that Bayesian inference with the APW model offers accurate and reliable performance under progressive censoring.
Keywords: Weibull distribution, Alpha Power Weibull distribution (APW), Progressive Type-II censoring, Bayesian estimation, Maximum likelihood, Monte Carlo simulation, MCMC techniques.

Resumé

Cette étude porte sur l'estimation et la prédiction des paramètres des lois de Weibull et de Weibull puissance alpha (APW) en présence de données censurées de type II progressive. Le modèle APW offre une plus grande flexibilité pour modéliser différentes formes de taux de défaillance. Les méthodes du maximum de vraisemblance et bayésienne sont utilisées pour estimer les paramètres de la distribution, ainsi que les fonctions de fiabilité et de risque. Les estimateurs bayésiens sont obtenus à l'aide de la méthode MCMC sous les fonctions de perte quadratique et LINEX. Des simulations de Monte Carlo sont utilisées pour évaluer les performances des méthodes, et des données réelles du domaine de l'ingénierie confirment leur efficacité. Les résultats montrent que l'inférence bayésienne basée sur le modèle APW fournit des estimations précises et fiables en présence de censure progressive.
Mots-clés : distribution de Weibull, distribution de Weibull puissance alpha (APW), censure progressive de type II, estimation bayésienne, maximum de vraisemblance, Simulation de monte carlo , Techniques MCMC.

Contents

List of Tables

List of Figures

List Of Notations And Symbols

MLE :	Maximum Likelihood Estimation
LINEX :	Linear Exponential
WE :	Weibull Distribution
APW :	Alpha Power Weibull Distribution
pmf :	probability mass function
pdf :	probability density function
cdf :	cumulative distribution function
HRF :	hazard rate function
RF :	reliability function
HPD :	highest posterior density
MC :	Monte Carlo
MCMC :	Monte Carlo Markov Chain
M – H :	Metropolis Hastings
K – S :	Kolomogrov Smirnov
MLEs :	Maximum likelihood Estimators
BEs :	Bayesian Estimators
SE :	Standard Error
SEL :	Squared Error Loss
MSE :	Mean Squared Error
LL :	LINEX Loss
CIs :	Confidence intervals
ACIs :	Asymptotic Confidence intervals
ACLs :	Average Confidence lengths
CPs :	Coverage probabilities
RMSE :	root mean squared error
MRAB :	mean relative absolute bias
rv :	random variable
rs :	random sample
log :	Logarithm function
ν :	The shape parameter of LINEX loss function

$\mathbb{P}$:	Probability
$Var(.)$:	variance
$\mathbb{E}(.)$:	Expected value
G :	Gamma Didtribution
$\mathcal{N}$:	Normal Distribution
$\mathcal{B}$:	Beta Distribution
$\mathcal{P}$:	Poisson Distribution
$\mathcal{B}$:	Binomial Distribution
$\mathcal{E}xp$:	Exponential Distribution
$I(.,.)$:	Observed information matrix
$\pi(.)$:	Joint posterior density
$L(.)$:	Likelihood function
$H(.)$:	Cumulative hazard rate function
$h(.)$:	Hazard rate function

Introduction

Statistical models must adapt to the challenge of extracting meaningful inferences from partially observed data. In such contexts, statistical inference plays a crucial role in estimating model parameters, evaluating system performance, and making predictions despite the presence of censored observations.

Statistical inference under censoring mechanisms enables analysts to make valid conclusions using only a subset of the data, typically observed until a certain point or stage. Among various censoring types, progressive Type-II censoring has gained significant importance for its flexibility and realism. Unlike classical Type-I or Type-II censoring, progressive schemes allow for the systematic removal of surviving units at multiple stages during the test. This not only saves resources but also captures more detailed failure behavior [4].

In progressive Type-II censoring, items are placed under test, and failure times are recorded. At each failure time, a pre-determined number of remaining items are removed (censored), leading to a multi-stage observation structure. This design generalizes other censoring schemes and has proven effective in applications where time and cost efficiency are critical.

To analyze data collected under such schemes, various estimation methods are employed. The maximum likelihood estimation (MLE) method is widely used due to its asymptotic properties and interpretability [6]. On the other hand, Bayesian estimation offers a probabilistic framework that incorporates prior information and allows for flexible modeling using sampling-based methods such as the Metropolis-Hastings algorithm [7, 14]. Both approaches are valuable tools in handling progressively censored data.

The Weibull distribution is among the most prominent models in reliability and survival analysis due to its ability to characterize different hazard rate behaviors-decreasing, constant, or increasing depending on its shape parameter. Its versatility makes it a standard choice in industrial applications and life-data analysis [16, 13].

Although the Weibull distribution is powerful, it may lack the flexibility needed to capture complex real-world data patterns. This has motivated the development of more general families of lifetime distributions. Several extensions have been proposed in the literature, including the equi-transponentiated Weibull, modified Weibull, and generalized Weibull distributions. These variants aim to enhance the flexibility of the modeling by introducing additional parameters to control skewness, tail behavior, or hazard shape.

Among these generalizations, the Alpha Power Weibull (APW) distribution has recently attracted attention due to its ability to provide an even wider range of hazard shapes by including an additional parameter, α . The APW distribution not only encompasses the classical Weibull distribution as a special case but also offers an improved fit for complex reliability data. Its flexibility in modeling both monotonic and nonmonotonic hazard functions makes it a suitable candidate for real-world applications involving progressive censoring schemes.

In recent statistical literature, increasing attention has been given to the problem of designing optimal censoring schemes. For a fixed number of units n and observed failures m, the goal is to determine the best progressive censoring scheme $(R_1, R_2, ..., R_m)$ such that $m + \sum_{i=1}^{m} R_i = n$, and the chosen scheme yields the most information about the unknown parameters. This involves defining appropriate information measures and comparing competing schemes .

The main objective of this study is to explore and explain the methods for conducting robust statistical inference, such as parameter estimation and hypothesis testing, when data are collected under this specific censoring scheme, highlighting its practical importance.

The rest of my dissertation work is organized as follows : In Chapter 1, we present the theoretical background necessary for understanding classical and Bayesian estimation methods, along with a discussion on censoring mechanisms, with emphasis on progressive Type-II censoring.
In Chapter 2 focuses on parameter estimation for the Weibull distribution using classical techniques such as maximum likelihood and confidence intervals, as well as Bayesian estimation under different loss functions.
In Chapter 3, the analysis is extended to the Alpha Power Weibull distribution, where both estimation methods are studied. The methodology of Monte Carlo simulation is outlined to support the theoretical procedures, and a real data example is included to illustrate the application of the proposed estimators.

Chapter 1

Backgrounds

This chapter is divided into three important sections. The first one contains some definitions of maximum likelihood estimation. Then, some of definitions and properties on bayesian estimation are also provided. Finally, the concept of censoring is introduced, with a particular focus on progressive Type-II censoring, which plays a central role in the data structure and estimation procedures considered in this study.

1.1 Maximum Likelihood Estimation

Maximum likelihood estimation is a probabilistic approach to solving the density estimation problem. It involves finding the probability distribution and its parameters that best describe the observed data by maximizing a likelihood function.

Likelihood and Log-Likelihood Functions

Let $X = x$ be a realization of a random variable (rv) or vector X with a known probability mass or density function (*pmf* or *pdf*) $f(x;\theta)$, which depends on the observed value x and an unknown parameter θ . This function is typically defined by a statistical model.

The parameter θ, which may be scalar or vector, lies in the parameter space Θ, while the set of all possible values of X is the sample space Ω.

Definition 1.1.1. The likelihood function $L(\theta)$ is the *pmf* or *pdf* of the observed data x, viewed as a function of the unknown parameter θ. That is,

$$L(\theta; x) = f(x;\theta)$$

When X is a random sample (rs), we assume that $x_1, \ldots, x_n$ are observations of the random vector $X = (X_1, \ldots, X_n)$, where $X_i \overset{iid}{\sim} f(x;\theta)$, for $i = 1, \ldots, n$.

due to the assumed independence of the components of X.

Therefore, the likelihood function based on a rs is:

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

The log-likelihood is hence the sum of the individual log-likelihood contributions as

$$logL\left(\theta;X\right)=log\left(\prod_{i=1}^{n}\mathfrak{f}\left(x_i;\theta\right)\right)=\sum_{i=1}^{n}logf\left(x_i;\theta\right).$$

1.1.1 Maximum Likelihood Estimate

The maximum likelihood provides a reliable and general method for estimating parameters. This approach can be used in many different estimation problems.For example,it is useful in reliability analysis, especially when dealing with censored data from different censoring schemes.

Definition 1.1.2. The likelihood function is maximized to produce the maximum likelihood estimate (MLE) $\hat{\theta}_{ML}$ of a parameter θ:

$$\hat{\theta}_{ML}=\arg\max_{\theta\in\Theta}L\left(\theta\right).$$

Example 1.1.1. Let X denote a rs from an exponential distribution $\mathcal{E}xp\left(\theta\right)$.Then:

$$\begin{aligned}L\left(\theta\right) &= \prod_{i=1}^{n}\{\theta e^{(-\theta x_i)\}} \\ &= \theta^n e^{\left(-\theta\sum_{i=1}^{n}x_i\right)}\end{aligned}$$

is the likelihood function of $\theta\in\mathbb{R}^+$. The log-likelihood function is therefore:

$$logL\left(\theta\right) = nlog\left(\theta\right)-\theta\sum_{i=1}^{n}x_i$$

with derivative:

$$\frac{\partial logL\left(\theta\right)}{\partial\theta} = \frac{n}{\theta} - \sum_{i=1}^{n} x_i$$

Setting the derivative to zero, we easily obtain the MLE $\hat{\theta}_{ML} = \frac{n}{\sum\limits_{i=1}^{n} x_i}$ is the mean observed survival time.

1.1.2 Score Function and Fisher Information

Definition 1.1.3. The first derivative of the log-likelihood function

$$S\left(\theta\right) = \frac{\partial logL\left(\theta\right)}{\partial\theta}$$

is called the **score function**.

Remark. Computation of the MLE is typically done by solving the score equation $S\left(\theta\right) = 0$.

Definition 1.1.4. The negative second derivative of the log-likelihood function

$$I\left(\theta\right) = -\frac{\partial^2 logL\left(\theta\right)}{\partial\theta^2} \tag{1.1}$$

is called the **Fisher information**. The value of the Fisher information at the MLE $\hat{\theta}_{ML}$ $I\left(\hat{\theta}_{ML}\right)$, is the observed Fisher information.

Example 1.1.2. Suppose we have observations $x_1, x_2, \ldots, x_n$ of a rs from a normal distribution $\mathcal{N}(\mu, \sigma^2)$ with unknown mean μ and known variance σ^2.

The *pdf* is:

$$f(x_i \mid \mu) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right)$$

The likeliohood is:

$$L(\mu) = \prod_{i=1}^{n} f(x_i \mid \mu) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right)$$

The log-likelihood function and score function are then

$$logL(\mu) = -\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x-\mu)^2$$

and

$$S(\mu) = \frac{d}{d\mu} \, log \, L(\mu) = \frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - \mu)$$

respectively.The solution of score equation $S(\mu) = 0$ is the MLE

$$\hat{\mu}_{ML} = \bar{x}$$

We take the second derivative of the log-likelihood:

$$\frac{d^2}{d\mu^2} \, log \, L(\mu) = -\frac{n}{\sigma^2}$$

so The Fisher information is:

$$I(\mu) = \mathbb{E}\left[-\frac{d^2}{d\mu^2} \, log \, L(\mu) \right] = \frac{n}{\sigma^2}$$

Since this value does not depend on μ, we have:

$$I(\hat{\mu}_{ML}) = I(\mu) = \frac{n}{\sigma^2}$$

Now, assume μ is Known and σ^2 is Unknown The log-likelihood is:

$$log \, L(\sigma^2) = -\frac{n}{2} \, log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2$$

and The MLE and Fisher information of σ^2 are then

$$\hat{\sigma}^2_{ML} = \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^2$$

and

$$I(\sigma^2) = \frac{1}{\sigma^6} \sum_{i=1}^{n} (x_i - \mu)^2 - \frac{n}{2\sigma^4}$$

In the MLE, the observed Fisher information becomes

$$I(\hat{\sigma}^2_{ML}) = \frac{n}{2\hat{\sigma}^4_{ML}}$$

Remark. From a frequentist point of view, the MLE $\hat{\mu}_{ML} = \bar{X}$ is a random variable. Its variance is:

$$\text{Var}(\hat{\mu}_{ML}) = \frac{\sigma^2}{n}$$

This is exactly equal to the inverse of the Fisher information:

$$\text{Var}(\hat{\mu}_{ML}) = \frac{1}{I(\hat{\mu}_{ML})}$$

In general, under regularity conditions, the variance of the MLE $\hat{\theta}_{ML}$ is approximately equal to the inverse observed Fisher information:

$$\text{Var}(\hat{\theta}_{ML}) \approx \frac{1}{I(\hat{\theta}_{ML})}$$

This approximation improves as the sample size increases. Example (1.1.1) is a special case where the equality holds exactly for any sample size.

Theorem 1.1.1. (Asymptotic Normality of MLE)[12]

Let $\hat{\theta}$ be the MLE for an unknown parameter θ. Then, we have

$$\sqrt{n}(\hat{\theta} - \theta) \rightarrow \mathcal{N}\left(0, \frac{1}{I(\theta)}\right)$$

As we can see, the asymptotic variance dispersion of the estimate around true parameter will be smaller when Fisher information is larger.

1.1.3 Delta Method

The delta method is a result concerning the approximate probability distribution for a function of an asymptotically normal statistical estimator from knowledge of the limiting variance of that estimator. The delta method generalizes easily to a multivariate setting, careful motivation of the technique is more easily demonstrated in univariate terms.

Definition 1.1.5. if there is a sequence of random variables X_n satisfying

$$\sqrt{n}\,[X_n - \theta] \xrightarrow{d} \mathcal{N}(0, \sigma^2)$$

where θ and σ^2 are finite valued constants and $\xrightarrow{d}$ denotes convergence in distribution, then

$$\sqrt{n}\,[g(X_n) - g(\theta)] \xrightarrow{d} \mathcal{N}\left(0, \sigma^2 . \left[g'(\theta)\right]^2\right)$$

for any function g satisfying the property that $g'(\theta)$ exists and is non-zero valued.

1.2 Confidence Interval

A confidence interval (CI) provides a range of plausible values for an unknown parameter, such as the mean or variance. Instead of giving a single estimate, it offers an interval that is likely to contain the true parameter with a certain level of confidence, typically 95%. CI are important because they reflect the uncertainty in estimation due to sampling variability, and they help us make informed decisions by quantifying the precision of our estimates.

Definition 1.2.1. For fixed $\gamma \in (0,1)$, a $\gamma.100\%$ confidence interval for θ is defined by two statistics $T_l = h_l(X)$ and $T_u = h_u(X)$ based on a rs X, which fulfill

$$\mathbb{P}(T_l \leq \theta \leq T_u) = \gamma \tag{1.2}$$

for all $\theta \in \Theta$. The statistics T_l and T_u are the limits of the confidence interval, and we assume $T_l \leq T_u$ throughout. The confidence level γ is also called coverage probability.

1.3 Bayesian Estimation

In frequentist inference, the data X are considered random, and point estimates of the unknown parameter θ are treated as functions of the data. The parameter θ is fixed but unknown. The properties of these estimates are studied by looking at how they behave over many possible samples of the data.

In Bayesian inference, named after Thomas Bayes, the unknown parameter θ is treated as a rv with a prior distribution $f(\theta)$. After observing the data $X = x$, Bayes theorem is used to update this prior and obtain the posterior distribution $f(\theta|x)$, which reflects what we know about θ given the data. Unlike frequentist inference, Bayesian inference is based on the observed data $X = x$.

Definition 1.3.1. [12] Let A and B denote two events A,B with $0 < \mathbb{P}(A) < 1$ and $\mathbb{P}(B) > 0$.then

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\,\mathbb{P}(A)}{\mathbb{P}(B)} \tag{1.3}$$

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\,\mathbb{P}(A)}{\mathbb{P}(B|A)\,\mathbb{P}(A) + \mathbb{P}(B|A^c)\,\mathbb{P}(A^c)} \tag{1.4}$$

For a general partition $A_1, A_2, ..., A_n$ with $\mathbb{P}(A_i) > 0$ for all $i = 1, ..., n$ we have that

$$\mathbb{P}\left(A_j|B\right)=\frac{\mathbb{P}\left(B|A_j\right)\mathbb{P}\left(A_j\right)}{\sum\limits_{i=1}^{n}\mathbb{P}\left(B|A_i\right)\mathbb{P}\left(A_i\right)} \tag{1.5}$$

for each $j=1,...,n$.

1.3.1 Posterior Distribution

The posterior distribution is the most important quantity in Bayesian inference. It contains all the information available about the unknown parameter θ after having observed the data $X=x$. Certain characteristics of the posterior distribution can be used to derive Bayesian point estimate.

Definition 1.3.2. Let $X=x$ be the observed realization of a (possibly multivariate) *rv* X with density function $f\left(x|\theta\right)$. Specifying a prior distribution with density function $f\left(\theta\right)$ allows us to compute the density function $f\left(\theta|x\right)$ of the posterior distribution using Bayes theorem.

$$f\left(\theta|x\right)=\frac{f\left(x|\theta\right)f\left(\theta\right)}{\int f\left(x|\theta\right)f\left(\theta\right)d\theta} \tag{1.6}$$

For discrete parameter θ the integral in the denominator has to be replaced with a sum.

Remark. • The term $f(x\mid\theta)$ is the likelihood, denoted by $L(\theta)$ previously denoted by $f(x,\theta)$. Since θ is now random, we write $L(\theta)=f(x\mid\theta)$. The marginal likelihood can also be written as

$$\int f\left(x|\theta\right)f\left(\theta\right)d\theta=\int f\left(x,\theta\right)d\theta=f\left(x\right).$$

which does not depend on θ.

• The posterior density is proportional to the product of the likelihood and the prior:

$$f(\theta\mid x)\propto f(x\mid\theta)f(\theta)\quad\text{or}\quad f(\theta\mid x)\propto L(\theta)f(\theta),$$

with the constant of proportionality ensuring normalization.

Example 1.3.1. The number $X=x$ of events observed in a fixed time interval can be reasonably modeled by the Poisson distribution. That is, we assume

$$X\sim\mathcal{P}(\lambda)$$

where $\lambda>0$ is the unknown rate parameter representing the average number of occurrences per interval.

It is tempting to select a **Gamma distribution** as a prior for λ, because the support of the Gamma distribution matches the parameter space $(0, \infty)$. Let the prior be

$$\lambda \sim G(\alpha, \beta), \quad \text{with } \alpha, \beta > 0,$$

which has the probability density function

$$f(\lambda) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda}, \quad \lambda > 0.$$

The likelihood function based on a single observation x is

$$f(x|\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2, \ldots$$

Hence, the posterior distribution is proportional to the product of the likelihood and the prior:

$$f(\lambda|x) \propto f(x|\lambda) \cdot f(\lambda) \propto \lambda^x e^{-\lambda} \cdot \lambda^{\alpha-1} e^{-\beta\lambda} = \lambda^{\alpha+x-1} e^{-(\beta+1)\lambda}$$

This is recognized as the kernel of a Gamma distribution with updated parameters

$$\lambda|x \sim G(\alpha + x, \beta + 1)$$

More generally, if we observe a random sample $X_1, X_2, \ldots, X_n \sim \mathcal{P}(\lambda)$, the sufficient statistic is $T = \sum_{i=1}^{n} X_i$. Then the posterior becomes

$$\lambda|\mathbf{x} \sim G(\alpha + T, \beta + n)$$

The Bayes estimator under squared error loss is the posterior mean:

$$\hat{\lambda}_{\text{Bayes}} = \frac{\alpha + T}{\beta + n}$$

1.3.2 Choice of the Prior Distribution

Bayesian inference allows the probabilistic specification of prior beliefs through a prior distribution. It is often useful and justified to restrict the range of possible prior distributions to a specific family with one or two parameters. The choice of this family can depend on the type of likelihood function used. We now discuss such a choice.

Conjugate Prior Distributions

A practical way to choose a prior distribution is to select one from a family of distributions that leads to a posterior distribution in the same family. This type of prior is called a *conjugate prior distribution.*

Definition 1.3.3. Let $L(\theta) = f(x|\theta)$ denote a likelihood function based on the observation $X = x$.A class $\mathcal{G}$ of distributions is called conjugate with respect to $L(\theta)$ if the posterior distribution $f(\theta|x)$ is in $\mathcal{G}$ for all x whenever the prior distribution $f(\theta)$ is in $\mathcal{G}$.

- The family $\mathcal{G} = \{\textsf{all distributions}\}$ is trivially conjugate with respect to any likelihood function. In practice one tries to find smaller sets $\mathcal{G}$ that are specific to the likelihood $L(\theta)$.

Table 1.1: Summary of conjugate prior distributions for different likelihood functions

Likelihood	Conjugate prior distribution	Posterior distribution
$X\|\pi \sim Bin(n,\pi)$	$\pi \sim \beta(\alpha,\beta)$	$\pi\|x \sim \beta(\alpha+x,\beta+n-x)$
$X\|\pi \sim Geom(\pi)$	$\pi \sim \beta(\alpha,\beta)$	$\pi\|x \sim \beta(\alpha+1,\beta+x-1)$
$X\|\lambda \sim \mathcal{P}(e.\lambda)$	$\lambda \sim G(\alpha,\beta)$	$\lambda\|x \sim G(\alpha+x,\beta+e)$
$X\|\lambda \sim \mathcal{E}xp(\lambda)$	$\lambda \sim G(\alpha,\beta)$	$\lambda\|x \sim G(\alpha+1,\beta+x)$
$X\|\mu \sim \mathcal{N}(\mu,\sigma^2)$	$\mu \sim \mathcal{N}(\nu,\varsigma^2)$	$\mu\|x \sim \mathcal{N}\left(\left(\frac{1}{\sigma^2}+\frac{1}{\varsigma^2}\right)^{-1}\cdot\left(\frac{x}{\sigma^2}+\frac{\nu}{\varsigma^2}\right),\left(\frac{1}{\sigma^2}+\frac{1}{\varsigma^2}\right)^{-1}\right)$
$X\|\sigma^2 \sim \mathcal{N}(\mu,\sigma^2)$	$\sigma^2 \sim IG(\alpha,\beta)$	$\sigma^2\|x \sim IG\left(\alpha+\frac{1}{2},\beta+\frac{1}{2}(x-\mu)^2\right)$

Improper Prior Distributions

The prior distribution affects the posterior distribution. To reduce this influence, a vague prior is often used such as one with a very large variance. In extreme cases, this may result in an *improper prior*, meaning a prior whose "density" does not integrate to one. Since it lacks a normalizing constant, it is usually written using the proportionality symbol "$\propto$". When using improper priors, it is important to check that the resulting posterior distribution is proper (i.e., it integrates to one). If the posterior is proper, then the use of improper priors is acceptable in Bayesian analysis. We now give a formal definition of an improper prior distribution.

Definition 1.3.4. A prior distribution with density function $f(\theta) \geq 0$ is called improper if

$$\int_\Theta f(\theta)\, d\theta = \infty \quad or \quad \sum_{\theta\in\Theta} f(\theta) = \infty \tag{1.7}$$

for continuous or discrete parameters θ, respectively.

Jeffreys Prior Distributions

It turns out that a particular choice of prior distribution is invariant under reparametrisation. This is Jeffreys prior (after Sir Harold Jeffreys, 1891-1989).

Definition 1.3.5. Let X be a rv with likelihood function $f(x|\theta)$ where θ is an unknown scalar parameter. Jeffreys prior is defined as

$$f(\theta) \propto \sqrt{J(\theta)} \tag{1.8}$$

where $J(\theta)$ is the expected Fisher information of θ. Equation (1.8) is also known as Jeffreys rule.

1.3.3 Bayesian Point Estimate

Statistical inference about θ is based solely on the posterior distribution. Suitable point estimates are location parameters, such as the mean, median or mode, of the posterior distribution. We will formally define those now for a scalar parameter θ.

Definition 1.3.6. • The posterior mean $\mathbb{E}(\theta|x)$ is the expectation of the posterior distribution:

$$\mathbb{E}(\theta|x) = \int \theta f(\theta|x)\, d\theta$$

• The posterior mode Mod$(\theta|x)$ is the mode of the posterior distribution:

$$Mod(\theta|x) = \arg\max_{\theta} f(\theta|x)$$

• The posterior median Med$(\theta|x)$ is the median of the posterior distribution, i.e. any number a that satisfies

$$\int_{-\infty}^{a} f(\theta|x)\, d\theta = 0.5 \qquad and \qquad \int_{a}^{\infty} f(\theta|x)\, d\theta = 0.5$$

Properties of Bayesian Point Estimate

To estimate an unknown parameter θ, there are at least three possible Bayesian point estimate available: the posterior mean, mode, and median.Which one should we choose in a specific application? To answer this question, we take a decision-theoretic view and first introduce the notion of a **loss function**.

Definition 1.3.7. Loss function

A loss function $l(a, \theta) \in \mathbb{R}$ quantifies the loss encountered when estimating the true parameter θ by a. If $a = \theta$, the associated loss is typically set to zero: $l(\theta, \theta) = 0$.

Common loss functions

- Quadratic loss: $l(a, \theta) = (a - \theta)^2$
- Linear loss: $l(a, \theta) = |a - \theta|$
- Zero-one loss:

$$l_\varepsilon(a, \theta) = \begin{cases} 0, & \text{if } |a - \theta| \leq \varepsilon \\ 1, & \text{if } |a - \theta| > \varepsilon \end{cases}, \quad \text{where } \varepsilon > 0$$

We now choose the point estimate a that minimises the posterior expected loss with respect to $f(\theta \mid x)$. Such a point estimate is called a **Bayes estimate**.

Definition 1.3.8. (Bayes estimate)

A Bayes estimate of θ with respect to a loss function $l(a, \theta)$ minimises the expected posterior loss:

$$E[l(a, \theta) \mid x] = \int_\Theta l(a, \theta) f(\theta \mid x)\, d\theta$$

It turns out that commonly used Bayesian point estimates correspond to Bayes estimates under specific loss functions.

Result 1.3.1. [12]

- The posterior mean is the Bayes estimate under quadratic loss.
- The Bayes estimate to linear loss is the posterior median.
- The posterior mode is the Bayes estimate under zero loss as $\varepsilon \to 0$.

To prove (1.3.1) we need the next integral rule.

Definition 1.3.9. (Leibniz Integral Rule)

Let a, b and f be real-valued functions that are continuously differentiable in t .Then the Leibniz integral rule is

$$\frac{\partial}{\partial t} \int_{a(t)}^{b(t)} f(x, t)\, dx = \int_{a(t)}^{b(t)} \frac{\partial}{\partial t} f(x, t)\, dx - f\{a(t), t\}.\frac{d}{dt} a(t) + f\{b(t), t\}.\frac{d}{dt} b(t) \quad (1.9)$$

This rule is also known as differentiation under the integral sign.

Proof. We first derive the posterior mean $\mathbb{E}(\theta|x)$ as the Bayes estimate with respect to quadratic loss. The expected quadratic loss is

$$\begin{aligned} \mathbb{E}\{l(a,\theta)\,|x\} &= \int l(a,\theta)\, f(\theta|x)\, d\theta \\ &= \int (a-\theta)^2 f(\theta|x)\, d\theta \end{aligned}$$

Setting the derivative with respect to a to zero leads to

$$2\int (a-\theta)\, f(\theta|x)\, d\theta = 0 \Leftrightarrow a - \int \theta f(\theta|x)\, d\theta = 0 \tag{1.10}$$

It immediately follows that $a = \int \theta f(\theta|x)\, d\theta = \mathbb{E}(\theta|x)\,.$

- Consider now the expected linear loss

$$\begin{aligned} \mathbb{E}\{l(a,\theta)\,|x\} &= \int l(a,\theta)\, f(\theta|x)\, d\theta = \int |a-\theta| f(\theta|x)\, d\theta \\ &= \int_{\theta\le a} (a-\theta)\, f(\theta|x)\, d\theta + \int_{\theta>a} (\theta-a)\, f(\theta|x)\, d\theta \end{aligned}$$

The derivative with respect to a can be calculated using Leibniz′s integral rule (1.9):

$$\begin{aligned} \frac{\partial}{\partial a}\mathbb{E}\{l(a,\theta)\,|x\} &= \frac{\partial}{\partial a}\int_{-\infty}^{a} (a-\theta)\, f(\theta|x)\, d\theta + \frac{\partial}{\partial a}\int_{a}^{\infty} (\theta-a)\, f(\theta|x)\, d\theta \\ &= \int_{-\infty}^{a} f(\theta|x)\, d\theta - (a-(-\infty))\, f(-\infty|x)\,.0 + (a-a)\, f(a|x)\,.1 \\ &- \int_{a}^{\infty} f(\theta|x)\, d\theta - (a-a)\, f(a|x)\,.1 + (\infty-a)\, f(\infty|x)\,.0 \\ &= \int_{-\infty}^{a} f(\theta|x)\, d\theta - \int_{a}^{\infty} f(\theta|x)\, d\theta \end{aligned}$$

Setting this equal to zero yields the posterior median $a = Med(\theta|x)$ as the solution for the estimate.

- Finally, the expected zero-one loss is

$$\begin{aligned} \mathbb{E}\{l(a,\theta)\,|x\} &= \int l_\varepsilon(a,\theta)\, f(\theta|x)\, d\theta \\ &= \int_{-\infty}^{a-\varepsilon} f(\theta|x)\, d\theta + \int_{a+\varepsilon}^{+\infty} f(\theta|x)\, d\theta \\ &= 1 - \int_{a-\varepsilon}^{a+\varepsilon} f(\theta|x)\, d\theta \end{aligned}$$

This will be minimised if the integral $\int_{a-\varepsilon}^{a+\varepsilon} f(\theta|x)\, d\theta$ is maximised. For small ε the integral is approximately $2\varepsilon f(a|x)$, which is maximised through the posterior mode $a = Mod(\theta|x)$. □

1.4 Censoring and Progressive Censoring

In survival analysis and reliability studies, data are often subject to censoring due to time or cost limitations. Understanding the types of censoring is crucial before applying estimation techniques. This section provides a brief overview of censoring mechanisms, with a special focus on progressive censoring, which is particularly relevant for this study [5].

1.4.1 Censoring

Life-testing and reliability studies have recently gained significant attention. It is understood that collecting complete lifetime data can be inefficient, expensive, and time-consuming. In addition, in some experiments, not all failure times can be observed. For these reasons, censored sampling is used in life-testing experiments.

Types Of Censoring

- **Right**: is the most common type, where the event has not occurred by the end of the observation period.
- **Left**: happens when the event occurs before the observation period begins.
- **Interval**: occurs when the event occurs but is only known to fall within a specific time frame.

A more advanced form of censoring, called **progressive censoring**, was developed by reliability practitioners to enhance data collection flexibility.

1.4.2 Progressive Censoring

In today's competitive market, product reliability has become crucial. Consumers now expect high-quality products with long useful life. To meet these expectations, manufacturers conduct reliability and life-testing experiments to understand product failure patterns and design effective warranties. Progressive censoring is one of these methods that is used in these experiments. It helps gather more accurate data, especially when some items are removed during testing. This technique improves the precision of statistical inference compared to conventional methods, making it valuable for product improvement and quality assessment.

- **Genesis**

Issues relating to progressive censoring can be dated back more than 40 years. To give a glimpse of its history, here is a query made in 1966 (Query 18, *Technometrics*, August 1966):

> "It is not uncommon in our life-testing for items to fail for reasons quite unrelated to the normal failure mechanism. For example, consider a number of lamps placed simultaneously on life-test. One of the lamps might be accidently broken after the start of the test but before all the lamps had burned out. If all lamps but one had burned out and the last were accidently broken, the population parameters are easily estimated by techniques designed to deal with censoring on the right. Breakage of any lamp but the longest lived one in the sample, however, introduces the problem of how to utilize the information that this lamp burned the observed number of hours before it was destroyed. What procedures can be recommended?"

The response to this enquiry was given by Dr. A. Clifford Cohen of University of Georgia in a subsequent issue of *Technometrics*. Yet, the above presented passage, being the first documented practical enquiry about the loss or removal of industrial units from experimentation prior to the termination of the experiment and due to causes other than failure, may be viewed as the genesis of "real-life" problem-based research on the topic of progressive censoring [5].

Now, you are probably wondering, *what exactly is progressive censoring?*

- **The Need for Progressive Censoring**

In reliability experiments, it's common for test units to be removed or lost before failure, either unintentionally (like breakage or dropout) or deliberately (to save time, cost, or resources). Traditional censoring assumes a single point of removal, but this doesn't always reflect real-world situations. Progressive censoring addresses this gap by allowing multiple, planned removals at various points in time. This flexibility makes it especially useful in experiments where early removal can save costs or where continuous monitoring is necessary. It also helps improve the accuracy and efficiency of statistical analysis.

- **Types of Progressive Censoring**

1. **Progressive Type-I Censoring :** In this type of censoring, some elements of the test are removed at different times during the study. The test continues until a fixed time is chosen before the test starts. This method is often used in reliability studies to check how long items last before they fail.

2. **Progressive Type-II Censoring :** In this method, the test continues until a prefixed number of failures occur. After each failure, a certain number of remaining units are randomly withdrawn from the test.

   This approach provides more control over the number of failures observed compared to traditional censoring methods,and allows for efficient data collection with limited resources.

Definition 1.4.1. The joint *pdf* of all m progressively Type-II censored rs

$X_{1:m:n}, X_{2:m:n}, \ldots, X_{m:m:n}$ is given by:

$$f_{X_{1:m:n},X_{2:m:n},\ldots,X_{m:m:n}}(x_1, x_2, \ldots, x_m) = c\prod_{i=1}^{m} f(x_i)\,[1-F(x_i)]^{R_i}\,, \tag{1.11}$$

where $X_1 < X_2 < ... < X_m$. and the constant c is:

$c = n(n-R_1-1)(n-R_1-R_2-2)\ldots(n-R_1-R_2-\cdots-R_{m-1}-m+1)$.

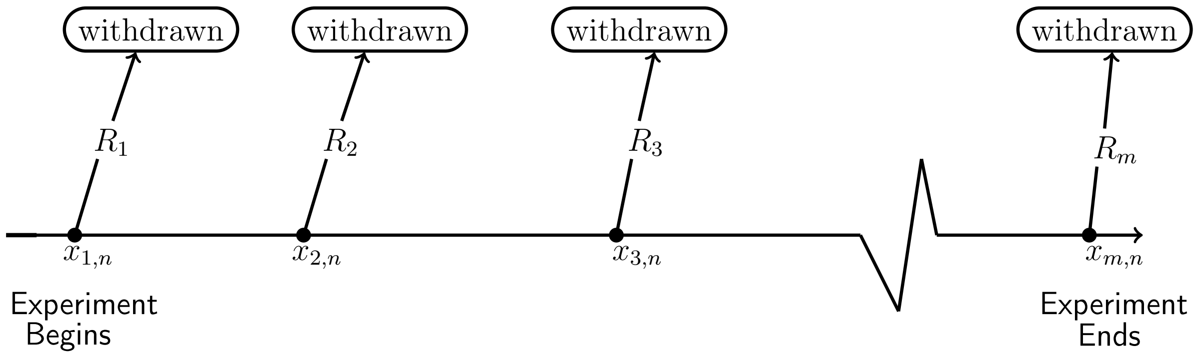

Figure 1.1: Progressive censoring of a life test with censoring scheme $\mathcal{R} = (R_1, ..., R_m)$

Chapter 2

Statistical Inference Based On Progressive Type-II Censoring from Weibull Distribution

In this chapter, We considerd classical methods including MLE and asymptotic confidence intervals, as well as Bayesian estimation. A brief numerical comparison is also given.

2.1 Classical Estimation

Classical estimation is a statistical method used to infer unknown population parameters from sample data through objective procedures such as the method of moments, MLE, and least squares, aiming to produce estimators that are unbiased, consistent, efficient, and sufficient, and it is widely applied in statistical inference based on progressive Type-II censoring from the Weibull distribution.

Weibull Distribution

The Weibull distribution (WE) is one of the most widely used distributions in reliability and survival studies. It plays an important role in analyzing skewed data and it is quite useful in diverse fields ranging from engineering to medical scopes (Lawless, 1982)[16]. A detailed discussion of the WE distribution has been provided by Johnson et al (1995) [13].
Definition 2.1.1. The WE distribution with shape α and scale λ parameters has the cumulative distribution function (*cdf*)

$$F(x|\alpha, \lambda) = 1 - e^{-\lambda x^{\alpha}}, \quad x > 0, \alpha, \lambda > 0, \tag{2.1}$$

and *pdf*

$$f(x|\alpha,\lambda) = \begin{cases} \alpha\lambda x^{\alpha-1}e^{-\lambda x^{\alpha}}, & \text{if } x > 0, \\ 0, & \text{if } x \le 0, \end{cases} \tag{2.2}$$

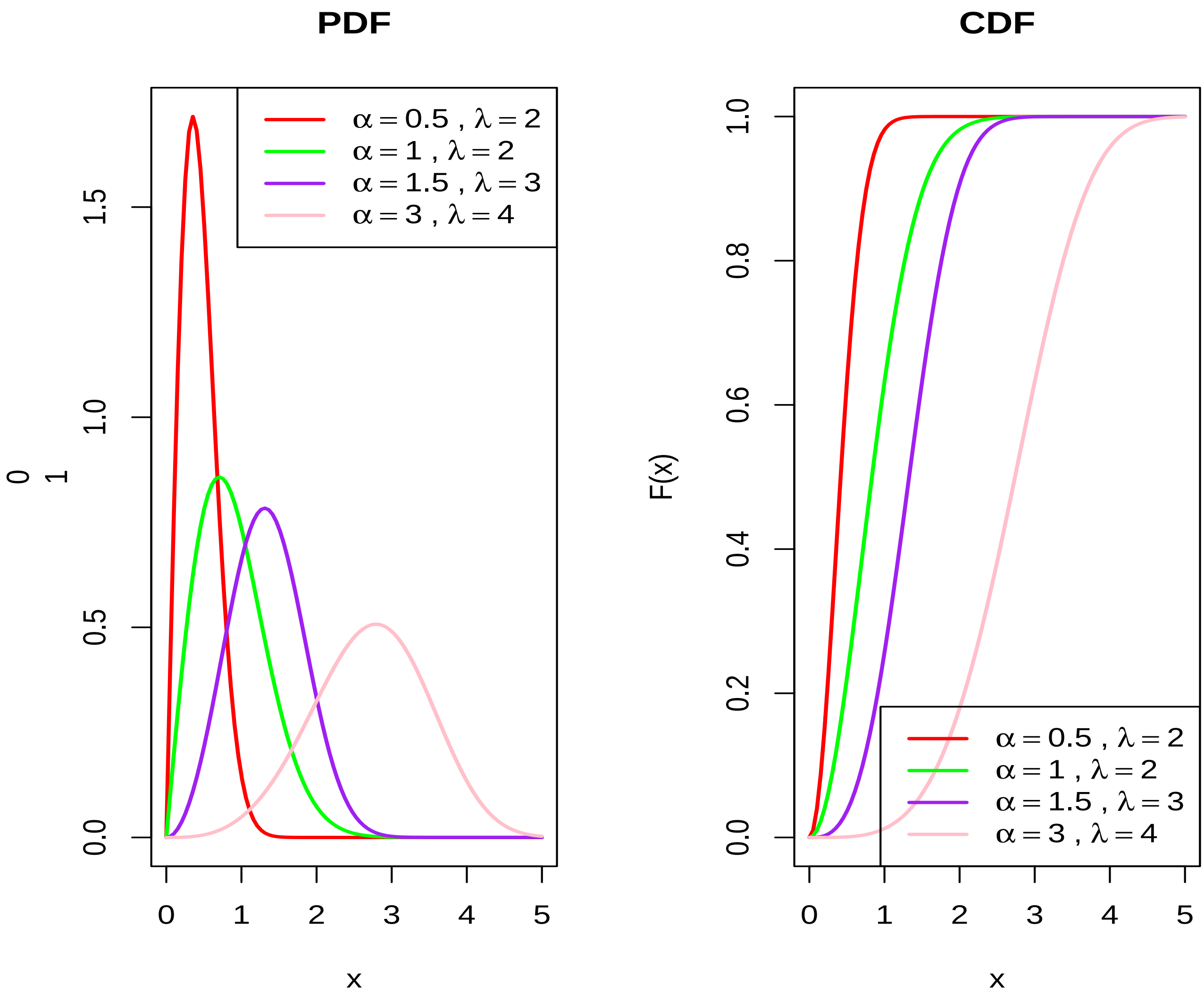

Figure 2.1: The PDF and CDF of the WE distribution using some specified values.

2.1.1 Maximum Likelihood Estimation

Let $X = (X_{1:m:n}, X_{2:m:n}, \ldots, X_{m:m:n})$ be a progressively Type-II censored sample of size m from a total of n units from WE distribution based on the censoring scheme $R = (R_1, R_2, \ldots, R_m)$, where: $n = m + \sum_{j=1}^{m} R_j$

The likelihood function under progressive Type-II censoring is:

$$L(\alpha,\lambda|x) = c\prod_{i=1}^{m} f(x_{i:m:n}|\alpha,\lambda)[1-F(x_{i:m:n}|\alpha,\lambda)]^{R_i}. \tag{2.3}$$

Substituting (2.1), (2.2) in (2.3), we obtain

$$L(\alpha,\lambda|x) \propto \alpha^m\lambda^m\left(\prod_{i=1}^{m} x_{i:m:n}^{\alpha-1}\right)\exp\left(-\lambda\sum_{i=1}^{m}(1+R_i)x_{i:m:n}^{\alpha}\right). \tag{2.4}$$

Then the log-likelihood function is:

$$\begin{aligned}
log\,L(\alpha,\lambda|x) &= log\left(\alpha^m\lambda^m\prod_{i=1}^{m} x_{i:m:n}^{\alpha-1}e^{-\lambda\sum_{i=1}^{m}(1+R_i)x_{i:m:n}^{\alpha}}\right)\\
&= log\,(\alpha^m) + log\,(\lambda^m) + \sum_{i=1}^{m} log(x_{i:m:n}^{\alpha-1}) + log(e^{-\lambda\sum_{i=1}^{m}(1+R_i)x_{i:m:n}^{\alpha}})\\
&= mlog\alpha + mlog\lambda + (\alpha-1)\sum_{i=1}^{m} logx_{i:m:n} - \lambda\sum_{i=1}^{m}(1+R_i)x_{i:m:n}^{\alpha}
\end{aligned}$$

We use the partial derivatives of the log-likelihood function with respect to α and λ, we get:

$$\frac{\partial\,log\,L(\alpha,\lambda|x)}{\partial\lambda} = \frac{m}{\lambda} - \sum_{i=1}^{m}(1+R_i)x_{i:m:n}^{\alpha} = 0$$

and

$$\frac{\partial\,log\,L(\alpha,\lambda|x)}{\partial\alpha} = \frac{m}{\alpha} + \sum_{i=1}^{m} log\,x_{i:m:n} - \lambda\sum_{i=1}^{m}(1+R_i)x_{i:m:n}^{\alpha}\,log\,x_{i:m:n} = 0 \tag{2.5}$$

The MLEs of λ is

$$\hat{\lambda}_{\text{ML}} = \frac{m}{\sum_{i=1}^{m}(1+R_i)x_{i:m:n}^{\alpha}} \tag{2.6}$$

Substituting (2.6) into (2.5) for λ:

$$\frac{m}{\alpha} + \sum_{i=1}^{m} log\,x_{i:m:n} - \left(\frac{m}{\sum_{k=1}^{m}(1+R_i)x_{i:m:n}^{\alpha}}\right)\left(\sum_{i=1}^{m}(1+R_i)x_{i:m:n}^{\alpha}\,log\,x_{i:m:n}\right) = 0 \tag{2.7}$$

Since we can not obtain the solution explicitly for the shape parameter α from equation (2.7), it is typically estimated numerically using iterative algorithms such as Newton-

Raphson, secant, bisection, or fixed-point iteration.For more details about the existence and uniqueness of these MLEs and uniqueness, see Balakrishnan and Kateri (2008) [6].

2.1.2 Asymptotic Confidence Interval

Since it is not easy to derive the exact distribution of the $MLEs$ in Eq's (2.6) and (2.5), we cannot obtain the exact confidence intervals (CIs) for the parameters α and λ Consequently, asymptotic CIs (ACIs) of the parameters are derived using the asymptotic distribution of $MLEs$. To this end, we need to find the variance-covariance matrix of the $MLEs$. The observed information matrix of $\theta = (\alpha, \lambda)$ is given by:

$$I(\theta) = -\begin{pmatrix} \frac{\partial^2 logL(\alpha,\lambda|x)}{\partial\alpha^2} & \frac{\partial^2 logL(\alpha,\lambda|x)}{\partial\alpha\partial\lambda} \\ \frac{\partial^2 logL(\alpha,\lambda|x)}{\partial\lambda\partial\alpha} & \frac{\partial^2 logL(\alpha,\lambda|x)}{\partial\lambda^2} \end{pmatrix} \tag{2.8}$$

where

$$\frac{\partial^2 logL(\alpha,\lambda|x)}{\partial\lambda^2} = -\frac{m}{\lambda^2} \tag{2.9}$$

$$\frac{\partial^2 logL(\alpha,\lambda|x)}{\partial\lambda\partial\alpha} = \frac{\partial^2 logL(\alpha,\lambda|x)}{\partial\alpha\partial\lambda} = -\sum_{i=1}^{m}(1+R_i)x_{i:m:n}^{\alpha}\, log\, x_{i:m:n} \tag{2.10}$$

$$\frac{\partial^2 logL(\alpha,\lambda|x)}{\partial\alpha^2} = -\frac{m}{\alpha^2} - \lambda\sum_{i=1}^{m}(1+R_i)x_{i:m:n}^{\alpha}(log\, x_{i:m:n})^2 \tag{2.11}$$

Hence, the inverse of the observed information matrix is given by:

$$I^{-1}(\theta) = -\begin{pmatrix} \frac{\partial^2 logL(\alpha,\lambda|x)}{\partial\alpha^2} & \frac{\partial^2 logL(\alpha,\lambda|x)}{\partial\alpha\partial\lambda} \\ \frac{\partial^2 logL(\alpha,\lambda|x)}{\partial\lambda\partial\alpha} & \frac{\partial^2 logL(\alpha,\lambda|x)}{\partial\lambda^2} \end{pmatrix}^{-1} = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} \tag{2.12}$$

where

$$V_{11} = \frac{\frac{m^2}{\lambda}}{\left(\frac{m^2}{\alpha} + \lambda S_1\right)\left(\frac{m}{\lambda^2}\right) - S_2^2}$$

$$V_{12} = V_{21} = \frac{-S_2}{\left(\frac{m}{\alpha^2} + \lambda S_1\right)\left(\frac{m}{\lambda^2}\right) - S_2^2}$$

$$V_{22} = \frac{\frac{m}{\alpha^2} + \lambda S_1}{\left(\frac{m}{\alpha^2} + \lambda S_1\right)\left(\frac{m}{\lambda^2}\right) - S_2^2}$$

Let us denote:

$$S_1 = \sum_{i=1}^{m}(1+R_i)x_{i:m:n}^{\hat{\alpha}}(log\, x_{i:m:n})^2, \quad S_2 = \sum_{i=1}^{m}(1+R_i)x_{i:m:n}^{\hat{\alpha}}\, log(x_{i:m:n})$$

The asymptotic joint distribution of the $MLEs$ $\hat{\alpha}$ and $\hat{\lambda}$ is approximated by a bivariate normal distribution, and is given by:

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\lambda} \end{pmatrix} \overset{\text{D}}{\sim} \mathcal{N}\left[\begin{pmatrix} \alpha \\ \lambda \end{pmatrix} \quad , \quad \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}\right] \tag{2.13}$$

The inverse matrix is:

$$I^{-1}(\hat{\theta}) = \frac{1}{\left(\frac{m}{\hat{\alpha}^2} + \hat{\lambda} S_1\right)\left(\frac{m}{\hat{\lambda}^2}\right) - S_2^2}\begin{pmatrix} \frac{m}{\hat{\lambda}^2} & -S_2 \\ -S_2 & \frac{m}{\hat{\alpha}^2} + \hat{\lambda} S_1 \end{pmatrix}$$

By replacing α and λ with their MLEs, we obtain the estimated variance-covariance matrix of the MLEs $\hat{\boldsymbol{\theta}} = (\hat{\alpha}, \hat{\lambda})$:

$$I^{-1}(\hat{\boldsymbol{\theta}}) = \begin{pmatrix} \text{var}(\hat{\alpha}) & \text{cov}(\hat{\alpha}, \hat{\lambda}) \\ \text{cov}(\hat{\lambda}, \hat{\alpha}) & \text{var}(\hat{\lambda}) \end{pmatrix} \tag{2.14}$$

Thus, the approximate $(1-\alpha)100\%$ CIs for the parameters α and λ are respectively given by:

$$(L_\alpha, U_\alpha) = \hat{\alpha} \pm z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{\alpha})}, \tag{2.15}$$

$$(L_\lambda, U_\lambda) = \hat{\lambda} \pm z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{\lambda})}, \tag{2.16}$$

where $z_{1-\frac{\alpha}{2}}$ is the $(1 - \dfrac{\alpha}{2})$ quantile of the standard normal distribution.

However, since the parameters α and λ are strictly positive, the lower bounds of these asymptotic intervals may sometimes be negative. To address this issue, we follow the log-transformation approach using the delta method. The asymptotic distribution of $log(\hat{\theta}_j)$ for $j = 1, 2$ is given by:

$$log(\hat{\theta}_j) - log(\theta_j) \overset{D}{\sim} \mathcal{N}\left(0, \frac{\text{var}(\hat{\theta}_j)}{\hat{\theta}_j^2}\right) \tag{2.17}$$

Therefore, modified asymptotic $(1-\alpha)100\%(0 < \alpha < 1)$ CIs for α and λ can be easily obtained, respectively, as follows:

$$\left(\hat{\alpha}\exp\left(-z_{1-\frac{\alpha}{2}}\frac{\sqrt{\text{var}(\hat{\alpha})}}{\hat{\alpha}}\right), \hat{\alpha}\exp\left(z_{1-\frac{\alpha}{2}}\frac{\sqrt{\text{var}(\hat{\alpha})}}{\hat{\alpha}}\right)\right), \tag{2.18}$$

$$\left(\hat{\lambda}\exp\left(-z_{1-\frac{\alpha}{2}}\frac{\sqrt{\text{var}(\hat{\lambda})}}{\hat{\lambda}}\right), \hat{\lambda}\exp\left(z_{1-\frac{\alpha}{2}}\frac{\sqrt{\text{var}(\hat{\lambda})}}{\hat{\lambda}}\right)\right). \tag{2.19}$$

2.1.3 Bootstrap

The bootstrap method is a useful resampling technique for constructing confidence intervals, especially when the asymptotic distribution is not reliable or the exact distribution is unknown. To construct the bootstrap confidence intervals for the parameters α and λ based on progressively Type-II censored data, we proceed as follows:

1. Generate B bootstrap samples from the estimated Weibull distribution with parameters $\hat{\alpha}$ and $\hat{\lambda}$.

2. For each bootstrap sample, apply the same progressive Type-II censoring scheme used in the original sample and compute the MLEs $\hat{\alpha}^*(b)$ and $\hat{\lambda}^*(b)$ for $b = 1, 2, \ldots, B$.

3. After obtaining the B bootstrap replicates of the MLEs, sort the bootstrap estimates of each parameter in ascending order.

4. The $100(1 - \alpha)\%$ percentile bootstrap confidence intervals for the parameters are

then given by:

$$\left(\hat{\alpha}^*_{(\lfloor \alpha B/2 \rfloor)}, \hat{\alpha}^*_{(\lceil (1-\alpha/2)B \rceil)}\right), \tag{2.20}$$

$$\left(\hat{\lambda}^*_{(\lfloor \alpha B/2 \rfloor)}, \hat{\lambda}^*_{(\lceil (1-\alpha/2)B \rceil)}\right), \tag{2.21}$$

where $\hat{\alpha}^*_{(i)}$ and $\hat{\lambda}^*_{(i)}$ denote the i-th order statistics of the bootstrap replicates of $\hat{\alpha}$ and $\hat{\lambda}$ respectively.

The bootstrap CIs are computationally intensive but often yield more accurate results than asymptotic intervals, especially for small sample sizes or highly censored data.

2.2 Bayesian Estimation

In this section, we develop the Bayesian estimation procedure for the parameters of the Weibull distribution under progressively Type-II censored data. The methodology is based on specifying suitable prior distributions, deriving the posterior distributions, and computing Bayes estimators under various loss functions. Moreover, a Monte Carlo simulation method is applied to obtain the estimates when analytical forms are not available.

2.2.1 Prior Distribution

We assume the parameters α and λ are *a priori* independent:

- The prior for λ is chosen as a Gamma distribution:

  $$\pi_1(\lambda|a_0, b_0) = \frac{b_0^{a_0}}{\Gamma(a_0)}\lambda^{a_0-1}e^{-b_0\lambda}, \quad \lambda > 0,$$

  where $a_0 > 0$ and $b_0 > 0$ are chosen to reflect the prior knowledge about λ.

- The prior for α is $\pi_2(\alpha)$ assumed to be log-concave and defined on $(0, \infty)$.which means that the logarithm of the density function is concave. This property simplifies the implementation of certain sampling algorithms, especially in Bayesian estimation frameworks. Log-concave densities are particularly useful in Adaptive Rejection Sampling (ARS). For more details, see Berger and Sun (1993) [7] and Kundu (2008) [14] and Devroye (1984) [11].

2.2.2 Posteroir Distribution

Joint Posterior Distribution

By combining the likelihood with the prior, we obtain the joint posterior of α and λ as:

$$\begin{aligned}\pi(\alpha,\lambda|\text{data}) &\propto L(\alpha,\lambda)\cdot\pi_1(\lambda|a_0,b_0)\cdot\pi_2(\alpha)\\ &\propto (\alpha^m\lambda^{m+a_0-1}\left(\prod_{j=1}^{m}x_{j:m:n}^{\alpha-1}\right)\exp\left\{-\lambda\left(b_0+\sum_{j=1}^{m}(1+R_j)x_{j:m:n}^{\alpha}\right)\right\}\pi_2(\alpha)\end{aligned}$$

Conditional Posterior Distribution

To perform Bayesian estimation,we next derive the conditional posterior distributions of each parameter.

- *Conditional posterior of λ given α*: Fixing α,the conditional posterior of λ is given by:

$$\pi(\lambda\mid\alpha,\text{data})\propto\lambda^{m+a_0-1}\exp\left\{-\lambda\left(b_0+\sum_{j=1}^{m}(1+R_j)x_{j:m:n}^{\alpha}\right)\right\}.$$

  This is the kernel of a Gamma distribution:

$$\lambda|\alpha,\text{data}\sim G\left(a_0+m,\ b_0+\sum_{j=1}^{m}(1+R_j)x_{j:m:n}^{\alpha}\right).$$

- *Conditional posterior of α given λ and data* : Conversely,fixing λ the conditional posterior of α becomes:

$$\pi(\alpha|\lambda,\text{data})\propto\alpha^m\prod_{j=1}^{m}x_{j:m:n}^{\alpha-1}\exp\left\{-\lambda\sum_{j=1}^{m}(1+R_j)x_{j:m:n}^{\alpha}\right\}\pi_2(\alpha). \tag{2.22}$$

Marginal posterior of α:

To derive the marginal posterior of α,we integrate out λ from the joint posterior:

$$\pi(\alpha\mid\text{data})\propto\pi_2(\alpha)\left(\prod_{j=1}^{m}x_{j:m:n}^{\alpha}\right)\int_0^{\infty}\lambda^{m+a_0-1}e^{-\lambda(b_0+\sum_{j=1}^{m}(1+R_j)x_{j:m:n}^{\alpha})}d\lambda.$$

Using the known identity: $\int_0^{\infty}\lambda^{k-1}e^{-c\lambda}d\lambda=\Gamma(k)c^{-k}$, with: $k = a_0 + m$ and $c = b_0 + \sum_{j=1}^{m}(1+R_j)x_{j:m:n}^{\alpha}$

we get:

$$\pi(\alpha \mid \text{data}) \propto \left(\prod_{j=1}^{m} x_{j:m:n}^{\alpha}\right)\left(b_0 + \sum_{j=1}^{m}(1+R_j)x_{j:m:n}^{\alpha}\right)^{-(a_0+m)} \pi_2(\alpha).$$

Including the factor α^m, the final form is:

$$\pi(\alpha|\text{data}) \propto \alpha^m \prod_{j=1}^{m} x_{j:m:n}^{\alpha-1}\left(b_0 + \sum_{j=1}^{m}(1+R_j)x_{j:m:n}^{\alpha}\right)^{-(a_0+m)} \pi_2(\alpha). \tag{2.23}$$

This distribution is log-concave if $\pi_2(\alpha)$ is log-concave.

2.2.3 Bayes Estimates under Different Loss Functions

The sample-based technique using the conditional posterior distribution can be used to obtain the Bayes estimates (BEs) of $\theta = \alpha$ or λ under three different loss functions. For a parameter θ and a decision rule a.

- The BEs under different loss functions are:
- **Squared Error Loss** ($l_1(a,\theta) = (a-\theta)^2$):

$$\hat{\theta}_{B_1} = \mathbb{E}_{\text{posterior}}(\theta|\text{data}) = \int_0^{\infty}\int_0^{\infty} \theta\, \pi(\alpha,\lambda|\text{data})\, d\alpha\, d\lambda.$$

- **Absolute Error Loss** ($l_2(a,\theta) = |a-\theta|$):

$$\hat{\theta}_{B_2} = \text{Med}_{\text{posterior}}(\theta|\text{data}).$$

- **LINEX Loss** ($l_3(a,\theta) = (a/\theta)^{\nu} - \nu\, ln(a/\theta) - 1,\ \nu \neq 0$):

$$\hat{\theta}_{B_3} = \left[\mathbb{E}_{\text{posterior}}(\theta^{-\nu}|\text{data})\right]^{-1/\nu} = \left[\int_0^{\infty}\int_0^{\infty} \theta^{-\nu}\pi(\alpha,\lambda|\text{data})\, d\alpha\, d\lambda\right]^{-1/\nu}.$$

Proof. (Bayes Estimator under LINEX Loss)

The LINEX (Linear Exponential) loss function is defined as:

$$l(a,\theta) = \left(\frac{a}{\theta}\right)^{\nu} - \nu\, ln\left(\frac{a}{\theta}\right) - 1,$$

Find the Bayes estimate a that minimizes the posterior expected loss:

$$\begin{aligned} E[l(a,\theta)|x] &= \int l(a,\theta)f(\theta|x)d\theta \\ &= \int \left[\left(\frac{a}{\theta}\right)^{\nu} - \nu\, ln\left(\frac{a}{\theta}\right) - 1\right] f(\theta|x)d\theta \\ &= a^{\nu}\int \theta^{-\nu}f(\theta|x)d\theta - \nu\, ln(a)(\theta|x)d\theta + \nu\int ln(\theta)f(\theta|x)d\theta - 1 \\ &= a^{\nu}E[\theta^{-\nu}|x] - \nu\, ln(\delta) + \nu E[ln(\theta)|x] - 1 \end{aligned}$$

Setting the derivative with respect to a to zero leads to

$$\nu a^{\nu-1}E[\theta^{-\nu}|x] - \nu/a = 0 \Leftrightarrow a^{\nu}\, E[\theta^{-\nu}|x] = 1 \Leftrightarrow a = \left(E[\theta^{-\nu}|x]\right)^{-1/\nu}$$

□

2.2.4 Monte Carlo Simulation Algorithm

Since explicit forms of Bayesian estimation are often intractable, we use Monte Carlo (MC) sampling:

step (1) Generate α_1 from the log-concave density function $\pi(\alpha|\text{data})$ using the method proposed by Devroye (1984) [11].

step (2) Generate λ_1 from the Gamma conditional posterior given α_1.

step (3) Repeat Steps(1) and (2) M-times and obtain MC samples $(\alpha_i, \lambda_i), i = 1, \dots, M$.

These resulting samples (α_l, λ_l) are used to approximate the BEs of the parameters and also to construct the corresponding simulated confidence intervals. The simulation steps and comparaison criteria in this section are adapted from Abu Awwad et al.(2014)[1]

2.3 Numerical Comparison

To evaluate the performance of the sample-based estimators and predictors, a MC simulation study was conducted. For specified values of n, m, and a given censoring scheme, progressively Type-II censored samples were generated from the WE distribution with fixed parameters $\alpha = 2$ and $\lambda = 1$, following the algorithm developed by Balakrishnan and Aggarwala (2000)[4]. In each simulation run, both the MLEs and BEs of α and λ were computed under various loss functions. This process was repeated 5000 times. The average bias and the mean squared error (MSE) were calculated for each estimator. The results, rounded to four decimal places, are presented in Tables (2.1) and (2.2), where the MSE values are given in parentheses.

Table 2.1: The average biases and MSEs for the Bayes and MLEs of α

n	m	Censoring Scheme	MLE	L_1	L_2	$L_3(\nu = 0.1)$	$L_3(\nu = 1)$	$L_3(\nu = 5)$
20	10	(10, 9*0)	0.2363 (0.3739)	0.0881 (0.2670)	0.0985 (0.2690)	0.0875 (0.2670)	0.0870 (0.2659)	0.0846 (0.2668)
20	10	(2,0,2,0,2,0,2,0,2,0)	0.3260 (0.6513)	0.1038 (0.2452)	0.1130 (0.2475)	0.1033 (0.2451)	0.1030 (0.2450)	0.1012 (0.2450)
20	10	(9*0,10)	0.3578 (0.6763)	0.0962 (0.2763)	0.1068 (0.2791)	0.0978 (0.2762)	0.0975 (0.2762)	0.0958 (0.2761)
20	10	(5*0,5*2)	0.4449 (0.8167)	0.1041 (0.3389)	0.1093 (0.3499)	0.0993 (0.3409)	0.0989 (0.3408)	0.0971 (0.3404)
30	10	(0,0.5,5.5,5,4*0)	0.3542 (0.6002)	0.1183 (0.2292)	0.1270 (0.2196)	0.1179 (0.2278)	0.1175 (0.2280)	0.1159 (0.2280)
30	10	(4*0,10,10,4*0)	0.2963 (0.5171)	0.0978 (0.2943)	1.069 (0.2968)	0.0974 (0.2943)	0.0970 (0.2943)	0.0953 (0.2944)
30	15	(0,0.5,5.5,5,10*0)	0.1386 (0.1542)	0.0305 (0.1371)	0.0364 (0.1374)	0.0304 (0.1371)	0.0302 (0.1371)	0.0295 (0.1372)
30	15	(14*0,15)	0.1751 (0.3213)	0.0368 (0.2288)	0.0421 (0.2293)	0.0366 (0.2288)	0.0365 (0.2288)	0.0359 (0.2289)
30	15	(15,14*0)	0.1737 (0.2012)	0.0758 (0.1679)	0.0815 (0.1691)	0.0756 (0.1679)	0.0754 (0.1679)	0.0747 (0.1680)
40	10	(0,0,0,10,10,10,4*0)	0.1332 (0.1845)	0.0773 (0.1635)	0.0859 (0.1654)	0.0769 (0.1643)	0.0765 (0.1634)	0.0749 (0.1633)
40	10	(10*3)	0.2415 (0.5109)	0.0659 (0.2631)	0.0656 (0.2643)	0.0566 (0.2631)	0.0563 (0.2631)	0.0548 (0.2631)
40	20	(10*2,10*0)	0.0948 (0.1289)	0.0209 (0.0692)	0.0252 (0.0708)	0.0208 (0.0691)	0.0207 (0.0691)	0.0204 (0.0690)
40	20	(10*0,10*2)	0.1228 (0.1368)	0.0085 (0.0913)	0.0123 (0.0912)	0.0092 (0.0904)	0.0090 (0.0902)	0.0081 (0.0913)
40	20	(19*0,20)	0.0894 (0.1343)	0.0092 (0.1211)	0.0093 (0.1220)	0.0092 (0.1220)	0.0090 (0.1219)	0.0095 (0.1220)
40	20	(20,19*0)	0.1278 (0.1361)	0.0097 (0.1154)	0.0140 (0.1157)	0.0096 (0.1154)	0.0095 (0.1154)	0.0092 (0.1155)

Table 2.2: The average biases and MSEs for the Bayes and MLEs of λ

n	m	Censoring Scheme	MLE	L_1	L_2	$L_3(\nu = 0.1)$	$L_3(\nu = 1)$	$L_3(\nu = 5)$
20	10	(10, 9*0)	2.2465 (2.7777)	0.1311 (0.1414)	0.0947 (0.1316)	0.0761 (0.1198)	0.0304 (0.1063)	-0.1685 (0.1020)
20	10	(2,0,2,0,2,0,2,0,2,0)	2.105 (4.4293)	0.1569 (0.1505)	0.1190 (0.1303)	0.1001 (0.1325)	0.0534 (0.1074)	0.1433 (0.0958)
20	10	(9*0,10)	3.3451 (6.7805)	0.1347 (0.1580)	0.1107 (0.1500)	0.0781 (0.1215)	-0.2806 (0.1117)	-1.1756 (0.1076)
20	10	(5*0,5*2)	3.3310 (6.7452)	0.1582 (0.1713)	0.1238 (0.1580)	0.1002 (0.1430)	0.0522 (0.1255)	-1.1524 (0.1115)
30	10	(0,0.5,5.5,5,4*0)	3.3941 (6.5399)	0.1528 (0.1934)	0.1182 (0.1737)	0.0968 (0.1646)	0.0504 (0.1416)	0.1498 (0.1240)
30	10	(4*0,10,10,4*0)	2.6459 (6.5394)	0.1487 (0.1643)	0.1124 (0.1450)	0.0924 (0.1372)	0.0456 (0.1204)	-0.1548 (0.1107)
30	15	(0,0.5,5.5,5,10*0)	1.1004 (1.1231)	0.0958 (0.0931)	0.0755 (0.0900)	0.0685 (0.0814)	0.0277 (0.0690)	-0.0170 (0.0646)
30	15	(14*0,15)	1.196 (2.2662)	0.0887 (0.0979)	0.0566 (0.0910)	0.0613 (0.0864)	0.0205 (0.0792)	-0.1149 (0.0741)
30	15	(15,14*0)	1.1005 (1.1633)	0.0853 (0.0853)	0.0941 (0.0787)	0.0487 (0.0750)	0.0185 (0.0687)	-0.1134 (0.0656)
40	10	(0,0,10,10,10,10,4*0)	1.669 (2.1705)	0.1232 (0.1493)	0.0937 (0.1376)	0.0717 (0.1250)	0.0386 (0.1011)	0.0450 (0.1040)
40	10	(10*3)	2.0673 (5.5623)	0.1075 (0.1689)	0.0753 (0.1536)	0.0582 (0.1422)	0.0039 (0.1047)	-0.1977 (0.1022)
40	20	(10*2,10*0)	0.0722 (0.0823)	0.0681 (0.0637)	0.0590 (0.0616)	0.0517 (0.0574)	-0.1087 (0.0550)	-0.2082 (0.0532)
40	20	(10*0,10*2)	0.0827 (0.0702)	0.0763 (0.0758)	0.0701 (0.0720)	0.0590 (0.0690)	0.0647 (0.0647)	0.0589 (0.0589)
40	20	(19*0,20)	0.0785 (1.1007)	0.0801 (0.0792)	0.0625 (0.0734)	0.0428 (0.0718)	0.1104 (0.0670)	-0.0671 (0.0580)
40	20	(20,19*0)	0.0924 (0.0810)	0.0564 (0.0763)	0.0419 (0.0737)	0.0291 (0.0709)	-0.0969 (0.0676)	-0.0933 (0.0635)

The numerical results from tables (2.1) and (2.2) show that BEs generally outperform MLEs by having lower bias and MSE. The effectiveness of both estimation techniques is sensitive to variations in sample size and the chosen sampling scheme. While the performance of BEs for the parameter α was similar under schemes L_1 and L_3, the BEs for the λ parameter were better under L_2 than L_1. Furthermore, across most scenarios, the estimates for λ calculated using the LINEX loss (LL) function L_3 were superior to all others.

Chapter 3

Statistical Inference Based On Progressive Type-II Censoring from Alpha Power Weibull Distribution

Several generalizations of classical lifetime distributions have been introduced to better fit real data. Commonly used models include the *exponential*, *Weibull*, *log-normal*, *gamma*, and *generalized exponential* distributions. While these models are useful, they may lack sufficient flexibility to capture complex data behaviors.

To address this, the *Alpha Power transformation*, introduced by Mahdavi and Kundu (2017) [17], has been applied to enhance baseline models.

In particular, the *Alpha Power Weibull (APW)* distribution extends the standard Weibull by adding a shape parameter, offering more flexibility in modeling various failure rates.

This chapter studies statistical inference for the parameters of the APW distribution based on *progressive Type-II censored data*. We present the distribution, derive MLE and Bayesian estimators, construct confidence intervals, and provide a numerical comparison.

Alpha Power Weibull Distribution

One of the most flexible distributions is known as the APW distribution which was introduced by Nassar et al. [19] by utilizing the alpha power transformation method introduced by Mahdavi and Kundu [17]. It can be considered to be a flexible extension of the traditional Weibull distribution and can deliver several desirable properties and better flexibility in the form of the hazard and density functions. If X is a rv that follows the APW distribution, then its *pdf* and *cdf* can be expressed as

$$f(x;\alpha,\beta,\lambda) = \frac{\lambda\beta\, log(\alpha) x^{\beta-1} e^{-\lambda x^{\beta}} \alpha^{1-e^{-\lambda x^{\beta}}}}{\alpha - 1}, \quad x > 0, \alpha, \beta, \lambda > 0,\ \alpha \neq 1 \tag{3.1}$$

and

$$F(x;\alpha,\beta,\lambda) = \frac{\alpha^{1-e^{-\lambda x^{\beta}}} - 1}{\alpha - 1} \tag{3.2}$$

where α and β are shape parameters and λ is a scale parameter

• The reliability function (RF) and hazard rate function (HRF) of the APW distribution are given by:

$$R(x;\alpha,\beta,\lambda) = \frac{\alpha}{\alpha - 1}\left(1 - \alpha^{-\exp(-\lambda x^{\beta})}\right) \tag{3.3}$$

$$h(x;\alpha,\beta,\lambda) = \frac{\lambda\beta\, log(\alpha) x^{\beta-1} \exp(-\lambda x^{\beta})}{\alpha^{\exp(-\lambda x^{\beta})} - 1} \tag{3.4}$$

For $\alpha = 1$, the APW distribution reduces to the alpha power exponential distribution proposed by Mahdavi and Kundu [17].

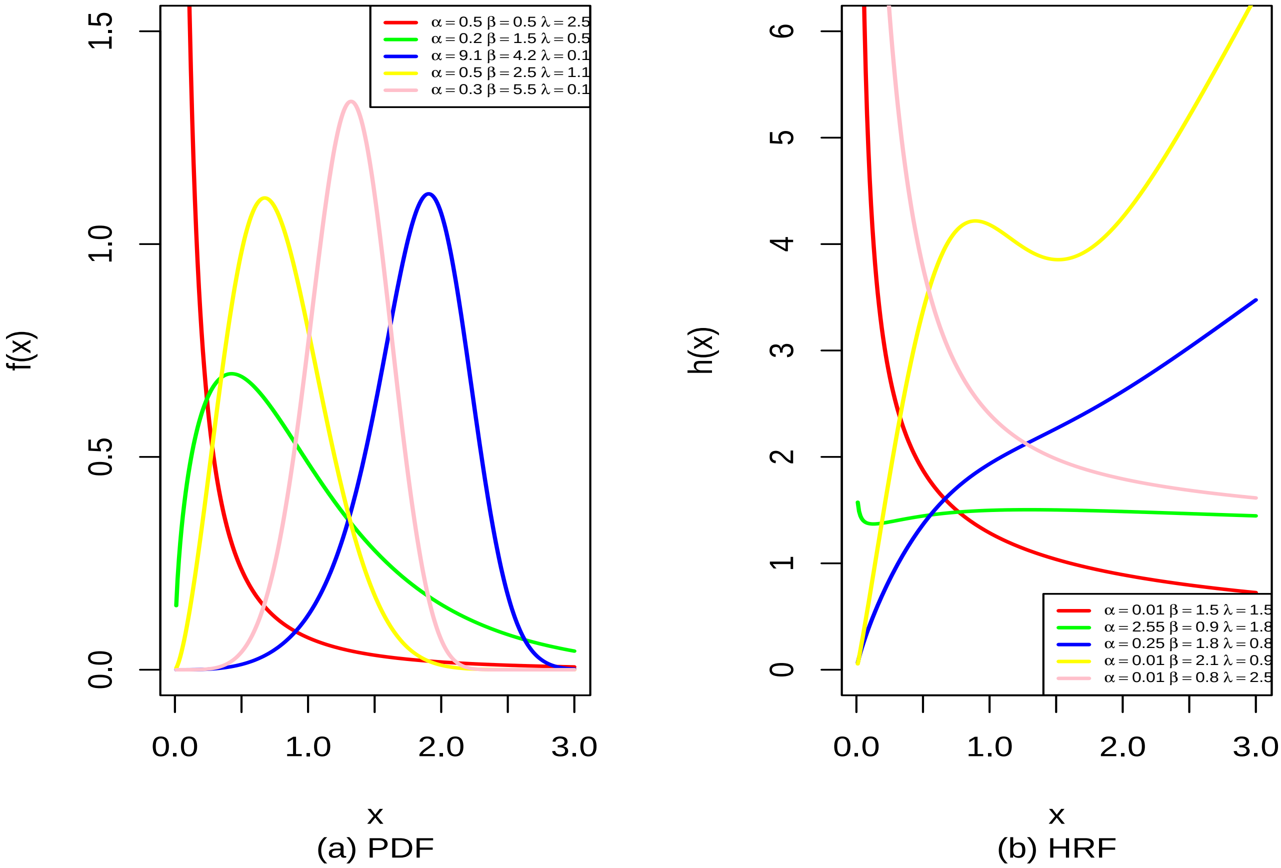

Figure 3.1: The PDFs and HRFs of the APW distribution using some specified values.

3.1 Classical Estimation

3.1.1 Maximum Likelihood Estimation

The MLEs of the parameters α, β, and λ as well RF and HRF of the APW distribution under progressively Type-II censored data are given by

$$
\begin{aligned}
L(\alpha,\beta,\lambda \mid x) &= c\prod_{i=1}^{m} f(x_{i:m:n})\,[1-F(x_{i:m:n})]^{R_i} \\
&\propto \prod_{i=1}^{m}\left[\lambda\beta\, log(\alpha) x_i^{\beta-1} e^{-\lambda x_i^{\beta}}\cdot \frac{\alpha^{1-e^{-\lambda x_i^{\beta}}}}{\alpha-1}\right]\prod_{i=1}^{m}\left[\frac{\alpha}{\alpha-1}\left(1-\alpha^{-e^{-\lambda x_i^{\beta}}}\right)\right]^{R_i} \\
&\propto (\lambda\beta\, log\,\alpha)^m\left[\prod_{i=1}^{m} x_i^{\beta-1}\right] e^{-\lambda\sum_{i=1}^{m} x_i^{\beta}}\,\alpha^{\sum_{i=1}^{m}(1-e^{-\lambda x_i^{\beta}})}\left(\frac{1}{\alpha-1}\right)^m \\
&\quad \left(\frac{\alpha}{\alpha-1}\right)^{\sum_{i=1}^{m} R_i}\prod_{i=1}^{m}\left(1-\alpha^{-e^{-\lambda x_i^{\beta}}}\right)^{R_i} \\
&\propto (\lambda\beta\, log\,\alpha)^m\left(\frac{\alpha}{\alpha-1}\right)^n\left[\prod_{i=1}^{m} x_i^{\beta-1}\right]\exp\left(-\lambda\sum_{i=1}^{m} x_i^{\beta}\right) \\
&\quad \exp\left(log\,\alpha\sum_{i=1}^{m}(1-e^{-\lambda x_i^{\beta}})\right)\prod_{i=1}^{m}\left(1-\alpha^{-e^{-\lambda x_i^{\beta}}}\right)^{R_i} \\
&\propto [\lambda\beta\, log(\alpha)]^m\left(\frac{\alpha}{\alpha-1}\right)^n\exp\left(-\lambda\sum_{i=1}^{m} x_i^{\beta}-log(\alpha)\sum_{i=1}^{m} e^{-\lambda x_i^{\beta}}\right) \\
&\quad \prod_{i=1}^{m} x_i^{\beta-1}\prod_{i=1}^{m}\left(1-\alpha^{-e^{-\lambda x_i^{\beta}}}\right)^{R_i}
\end{aligned}
$$

Where $x_i = x_{i:m:n}$, for $i = 1,\ldots,m$ and $n = m + \sum_{i=1}^{m} R_i$

Then the log likelihood function can be written as follows:

$$
\begin{aligned}
log\,L(\alpha,\beta,\lambda \mid x) &= m\,log[\lambda\beta\,log(\alpha)] + n\,log\left(\frac{\alpha}{\alpha-1}\right) - \lambda\sum_{i=1}^{m} x_i^{\beta} - log(\alpha)\sum_{i=1}^{m} e^{-\lambda x_i^{\beta}} \\
&\quad + \sum_{i=1}^{m} R_i\,log\left(1-\alpha^{-e^{-\lambda x_i^{\beta}}}\right). \qquad (3.5)
\end{aligned}
$$

The MLEs of α, β, λ, denoted by $\hat{\alpha}, \hat{\beta}, \hat{\lambda}$, respectively, can be obtained by maximizing the log-likelihood function in (3.5). Equivalently, they can be obtained by solving the following system of nonlinear equations:

$$\frac{\partial\, log\, L(\alpha,\beta,\lambda \mid x)}{\partial \alpha} = \frac{m}{\alpha\, log(\alpha)} + n\left(\frac{1}{\alpha} - \frac{1}{\alpha-1}\right) - \frac{1}{\alpha}\sum_{i=1}^{m} e^{-\lambda x_i^{\beta}} + \frac{1}{\alpha}\sum_{i=1}^{m} R_i \frac{e^{-\lambda x_i^{\beta}}}{\psi_i} = 0, \tag{3.6}$$

$$\frac{\partial\, log\, L(\alpha,\beta,\lambda \mid x)}{\partial \beta} = \frac{m}{\beta} - \lambda \sum_{i=1}^{m} x_i^{\beta}\, log(x_i) + \lambda\, log(\alpha) \sum_{i=1}^{m} v_i\, log(x_i) - \lambda\, log(\alpha) \sum_{i=1}^{m} R_i \frac{v_i\, log(x_i)}{\psi_i} = 0, \tag{3.7}$$

$$\frac{\partial\, log\, L(\alpha,\beta,\lambda \mid x)}{\partial \lambda} = \frac{m}{\lambda} - \sum_{i=1}^{m} x_i^{\beta} + log(\alpha) \sum_{i=1}^{m} v_i - log(\alpha) \sum_{i=1}^{m} R_i \frac{v_i}{\psi_i} = 0, \tag{3.8}$$

where $\psi_i = \alpha^{e^{-\lambda x_i^{\beta}}} - 1$ and $v_i = x_i^{\beta} e^{-\lambda x_i^{\beta}}$.

The MLEs of the parameters for APW distribution under progressively Type-II censored data can be obtained by solving equations (3.6)-(3.8) using the Newton-Raphson iterative method. It is worth noting that the MLEs of the APW distribution based on Type-II (non-progressive) censoring can be derived directly from (3.6)-(3.8) by setting $R_1 = R_2 = \cdots = R_{m-1} = 0$. Using the invariance property of MLEs, the estimates of the $R(t)$ and the $h(t)$ at a specific time t are given by:

$$\hat{R}(t) = \frac{\hat{\alpha}}{\hat{\alpha}-1}\left(1 - \hat{\alpha}^{-\exp(-\hat{\lambda} t^{\hat{\beta}})}\right), \quad \hat{h}(t) = \frac{\hat{\lambda}\hat{\beta}\, log(\hat{\alpha}) t^{\hat{\beta}-1} \exp(-\hat{\lambda} t^{\hat{\beta}})}{\hat{\alpha}^{\exp(-\hat{\lambda} t^{\hat{\beta}})} - 1}. \tag{3.9}$$

3.1.2 Asymptotic Confidence Interval

It is important to construct CI for the unknown parameters α, β, and λ, as well as for the $R(t)$ and $h(t)$. To achieve this, we employ the asymptotic properties of the MLEs.

Under standard regularity conditions, the asymptotic distribution of the MLEs $(\hat{\alpha}, \hat{\beta}, \hat{\lambda})$ is a multivariate normal distribution with mean (α, β, λ) and covariance matrix given by the inverse of the Fisher information matrix, denoted by $I^{-1}(\alpha, \beta, \lambda)$. However, due to the complex expressions of the second-order partial derivatives of the log-likelihood function, computing $I^{-1}(\alpha, \beta, \lambda)$ analytically is challenging.

Instead, we use the observed Fisher information matrix evaluated at the MLEs, i.e.,

$I^{-1}(\hat{\alpha}, \hat{\beta}, \hat{\lambda})$, which is computed as follows:

$$I^{-1}(\hat{\alpha}, \hat{\beta}, \hat{\lambda}) = \begin{bmatrix} -\frac{\partial^2 L(\alpha,\beta,\lambda \mid x)}{\partial \alpha^2} & -\frac{\partial^2 L(\alpha,\beta,\lambda \mid x)}{\partial \alpha \partial \beta} & -\frac{\partial^2 L(\alpha,\beta,\lambda \mid x)}{\partial \alpha \partial \lambda} \\ -\frac{\partial^2 L(\alpha,\beta,\lambda \mid x)}{\partial \beta \partial \alpha} & -\frac{\partial^2 L(\alpha,\beta,\lambda \mid x)}{\partial \beta^2} & -\frac{\partial^2 L(\alpha,\beta,\lambda \mid x)}{\partial \beta \partial \lambda} \\ -\frac{\partial^2 L(\alpha,\beta,\lambda \mid x)}{\partial \lambda \partial \alpha} & -\frac{\partial^2 L(\alpha,\beta,\lambda \mid x)}{\partial \lambda \partial \beta} & -\frac{\partial^2 L(\alpha,\beta,\lambda \mid x)}{\partial \lambda^2} \end{bmatrix}^{-1}_{(\hat{\alpha},\hat{\beta},\hat{\lambda})} \tag{3.10}$$

The elements of the Fisher information matrix are obtained from the log-likelihood function as follows:

$$\begin{aligned} \frac{\partial^2 logL(\alpha,\beta,\lambda \mid x)}{\partial \alpha^2} &= \frac{m[1+log(\alpha)]}{[\alpha\, log(\alpha)]^2} n\left[(\alpha-1)^{-2} - \alpha^{-2}\right] + \frac{1}{\alpha^2}\sum_{i=1}^{m} e^{-\lambda x_i^{\beta}} \sum_{i=1}^{m} \alpha^{-2}\psi_i^{-2} R_i e^{-\lambda x_i^{\beta}} \phi_i, \\ \frac{\partial^2 logL(\alpha,\beta,\lambda \mid x)}{\partial \beta^2} &= -\frac{m}{\beta^2} - \lambda \sum_{i=1}^{m} x_i^{\beta}(log\, x_i)^2 + \lambda\, log(\alpha)\sum_{i=1}^{m}\varphi_i - \lambda\, log(\alpha)\sum_{i=1}^{m} R_i\varphi_i\psi_i^{-1} \\ &\quad - \lambda^2 \overset{2}{log}(\alpha)\sum_{i=1}^{m} R_i w_i \psi_i^{-2}, \\ \frac{\partial^2 logL(\alpha,\beta,\lambda \mid x)}{\partial \lambda^2} &= -\frac{m}{\lambda^2} - log(\alpha)\sum_{i=1}^{m} x_i^{\beta} v_i - log(\alpha)\sum_{i=1}^{m} R_i v_i u_i \psi_i^{-2}, \end{aligned}$$

and

$$\begin{aligned} \frac{\partial^2 logL(\alpha,\beta,\lambda \mid x)}{\partial \alpha \partial \beta} &= \frac{\lambda}{\alpha}\sum_{i=1}^{m} v_i\, log(x_i) + \frac{\lambda}{\alpha}\sum_{i=1}^{m} R_i u_i e^{-\lambda x_i^{\beta}}\, log(x_i)\psi_i^{-2}, \\ \frac{\partial^2 logL(\alpha,\beta,\lambda \mid x)}{\partial \alpha \partial \lambda} &= \frac{1}{\alpha}\sum_{i=1}^{m} v_i + \frac{1}{\alpha}\sum_{i=1}^{m} R_i u_i e^{-\lambda x_i^{\beta}} \psi_i^{-2}, \\ \frac{\partial^2 logL(\alpha,\beta,\lambda \mid x)}{\partial \beta \partial \lambda} &= -\sum_{i=1}^{m} x_i^{\beta}\, log(x_i) + log(\alpha)\sum_{i=1}^{m}\varphi_i\, \overset{-1}{log}(x_i) - log(\alpha)\sum_{i=1}^{m} R_i\varphi_i\, \overset{-1}{log}(x_i)\psi_i^{-1} \\ &\quad - \lambda\, \overset{2}{log}(\alpha)\sum_{i=1}^{m} R_i w_i\, \overset{-1}{log}(x_i)\psi_i^{-2}, \end{aligned}$$

where

$$\varphi_i = v_i(log\, x_i)^2(1-\lambda x_i^{\beta}), \quad w_i = v_i^2(log\, x_i)^2\alpha^{e^{-\lambda x_i^{\beta}}}, \quad \phi_i = 1-\alpha^{e^{-\lambda x_i^{\beta}}}(1+e^{-\lambda x_i^{\beta}}),$$

$$\text{and} \quad u_i = x_i^{\beta}[1+\alpha^{e^{-\lambda x_i^{\beta}}}(log(\alpha)e^{-\lambda x_i^{\beta}}-1)].$$

Then, the $100(1-\varepsilon)\%$ ACIs for the unknown parameters α, β, and λ are computed as:

$$\hat{\alpha} \pm z_{\varepsilon/2}\sqrt{\hat{\mathrm{v}}(\hat{\alpha})}, \quad \hat{\beta} \pm z_{\varepsilon/2}\sqrt{\hat{\mathrm{v}}(\hat{\beta})}, \quad \hat{\lambda} \pm z_{\varepsilon/2}\sqrt{\hat{\mathrm{v}}(\hat{\lambda})} \tag{3.11}$$

where $\hat{v}(\hat{\alpha})$, $\hat{v}(\hat{\beta})$, and $\hat{v}(\hat{\lambda})$ are the estimated variances obtained from the diagonal elements of (3.10), and $z_{\varepsilon/2}$ is the upper $\varepsilon/2$ quantile of the standard normal distribution.

To obtain the ACIs for $R(t)$ of the APW distribution, we use the delta method. For this purpose, we need the first-order partial derivatives of $R(t)$ with respect to the unknown parameters α, β, and λ, as follows:

$$\frac{\partial R(t)}{\partial \alpha} = \frac{1 + (\alpha - 1)e^{-\lambda t^{\beta}} - \alpha^{e^{-\lambda t^{\beta}}}}{(\alpha - 1)^2 \alpha^{e^{-\lambda t^{\beta}}}},$$
$$\frac{\partial R(t)}{\partial \beta} = -\frac{\lambda \alpha \, log(\alpha) \, log(t) t^{\beta} e^{-\lambda t^{\beta}}}{(\alpha - 1)\alpha^{e^{-\lambda t^{\beta}}}},$$
$$\frac{\partial R(t)}{\partial \lambda} = \frac{\alpha \, log(\alpha) t^{\beta} e^{-\lambda t^{\beta}}}{(\alpha - 1)\alpha^{e^{-\lambda t^{\beta}}}},$$

Let $Y_R = \left(\frac{\partial R}{\partial \alpha}, \frac{\partial R}{\partial \beta}, \frac{\partial R}{\partial \lambda}\right)^{\top} |_{(\alpha,\beta,\lambda)=(\hat{\alpha},\hat{\beta},\hat{\lambda})}$

Then, the estimated variances of $\hat{R}(t)$ can be approximated by:

$$\hat{v}(\hat{R}) \approx Y_R I^{-1}(\hat{\alpha}, \hat{\beta}, \hat{\lambda}) Y_R^{\top},$$

and the ACIs of $R(t)$ at the confidence level $100(1-\varepsilon)\%$ are given respectively by:

$$\hat{R}(t) \pm z_{\varepsilon/2}\sqrt{\hat{v}(\hat{R})}.$$

the first-order partial derivatives of $h(t)$ with respect to the unknown parameters α, β, and λ To obtain the ACIs for $h(t)$

$$\frac{\partial h(t)}{\partial \alpha} = \frac{\lambda \beta t^{\beta-1} e^{-\lambda t^{\beta}} \left[\alpha^{e^{-\lambda t^{\beta}}}(1 - log(\alpha) e^{-\lambda t^{\beta}}) - 1\right]}{\alpha \left(\alpha^{e^{-\lambda t^{\beta}}} - 1\right)^2},$$
$$\frac{\partial h(t)}{\partial \beta} = \frac{\lambda t^{\beta-1} e^{-\lambda t^{\beta}} \, log(\alpha)[1 + \beta \, log(t) - \lambda \beta t^{\beta} \, log(t)]}{\alpha^{e^{-\lambda t^{\beta}}} - 1} - \frac{\lambda^2 \beta t^{2\beta-1} e^{-2\lambda t^{\beta}} \, log^2(\alpha) \, log(t) \alpha^{e^{-\lambda t^{\beta}}}}{(\alpha^{e^{-\lambda t^{\beta}}} - 1)^2},$$
$$\frac{\partial h(t)}{\partial \lambda} = \frac{\beta t^{\beta-1} e^{-\lambda t^{\beta}} \, log(\alpha)[1 - \lambda t^{\beta}]}{\alpha^{e^{-\lambda t^{\beta}}} - 1} - \frac{\lambda \beta t^{2\beta-1} e^{-2\lambda t^{\beta}} \, log^2(\alpha) \alpha^{e^{-\lambda t^{\beta}}}}{(\alpha^{e^{-\lambda t^{\beta}}} - 1)^2}.$$

Let $Y_h = \left(\frac{\partial h}{\partial \alpha}, \frac{\partial h}{\partial \beta}, \frac{\partial h}{\partial \lambda}\right)^\top |_{(\alpha,\beta,\lambda)=(\hat{\alpha},\hat{\beta},\hat{\lambda})}$

Then, the estimated variances of $\hat{R}(t)$ can be approximated by:

$$\hat{v}(\hat{h}) \approx Y_h I^{-1}(\hat{\alpha}, \hat{\beta}, \hat{\lambda}) Y_h^\top,$$

and the ACIs of $h(t)$ is:

$$\hat{h}(t) \pm z_{\varepsilon/2}\sqrt{\hat{v}(\hat{h})}.$$

3.2 Bayesian Estimation

In this section, we consider the Bayesian estimation of the unknown parameters α, β, and λ, along with the RF and HRF, based on progressively Type-II censored data. The estimation is performed under both the squared error loss (SEL) and the LINEX loss (LL) functions.

3.2.1 Prior Distribution

There was no conjugate prior to the APW distribution. As a result, we presumptively use gamma priors, which are thought to be more flexible than other priors and adjust to the support of the parameters. Additionally, the independent gamma priors are clear and straightforward, which may avoid many complicated inferential issues, see also in this regard Kundu and Howlader [15], Dey et al.[10] and Nassar et al.[20] Let:

$$\alpha \sim \mathcal{G}(a_1, b_1), \beta \sim \mathcal{G}(a_2, b_2), \lambda \sim \mathcal{G}(a_3, b_3),$$

which leads to the joint prior distribution:

$$g(\alpha, \beta, \lambda) \propto \alpha^{a_1-1}\beta^{a_2-1}\lambda^{a_3-1}\exp\left(-(b_1\alpha + b_2\beta + b_3\lambda)\right), \quad \alpha, \beta, \lambda > 0. \tag{3.12}$$

3.2.2 Posterior Distribution

Combining the likelihood function with the prior in Equation (3.12), the joint posterior density function becomes:

$$\pi(\alpha,\beta,\lambda \mid x) = A^{-1}\frac{\alpha^{n+a_1-1}\beta^{m+a_2-1}\lambda^{m+a_3-1}[log(\alpha)]^m}{(\alpha-1)^n}$$
$$\times \exp\left[-\lambda\left(\sum_{i=1}^{m} x_i^{\beta}+b_3\right)\right]\exp\left(-log(\alpha)\sum_{i=1}^{m} e^{-\lambda x_i^{\beta}}-b_1\alpha-b_2\beta\right)$$
$$\times\prod_{i=1}^{m}\left(1-\alpha^{-e^{-\lambda x_i^{\beta}}}\right)^{R_i}, \tag{3.13}$$

where A is the normalizing constant and given by

$$A=\int_0^{\infty}\int_0^{\infty}\int_0^{\infty}\frac{\alpha^{n+a_1-1}\beta^{m+a_2-1}\lambda^{m+a_3-1}[log(\alpha)]^m}{(\alpha-1)^n}$$
$$\times \exp\left[-\lambda\left(\sum_{i=1}^{m} x_i^{\beta}+b_3\right)\right]\exp\left(-log(\alpha)\sum_{i=1}^{m} e^{-\lambda x_i^{\beta}}-b_1\alpha-b_2\beta\right)$$
$$\times\prod_{i=1}^{m}\left(1-\alpha^{-e^{-\lambda x_i^{\beta}}}\right)^{R_i} d\alpha\, d\beta\, d\lambda.$$

From the posterior density in Equation (3.13), the full conditional distributions are as follows:

For α:

$$\pi_1(\alpha \mid \beta,\lambda,x) \propto \frac{\alpha^{m+a_1-1}[log(\alpha)]^m}{(\alpha-1)^n}\exp\left(-log(\alpha)\sum_{i=1}^{m} e^{-\lambda x_i^{\beta}}-b_1\alpha\right)$$
$$\times\prod_{i=1}^{m}\left(1-\alpha^{-e^{-\lambda x_i^{\beta}}}\right)^{R_i}, \tag{3.14}$$

For β:

$$\pi_2(\beta \mid \alpha,\lambda,x) \propto \beta^{m+a_2-1}\exp\left(-\lambda\sum_{i=1}^{m} x_i^{\beta}-log(\alpha)\sum_{i=1}^{m} e^{-\lambda x_i^{\beta}}-b_2\beta\right)$$
$$\times\prod_{i=1}^{m}\left(1-\alpha^{-e^{-\lambda x_i^{\beta}}}\right)^{R_i}, \tag{3.15}$$

For λ:

$$\pi_3(\lambda \mid \alpha, \beta, x) \propto \lambda^{m+a_3-1} \exp\left(-\lambda \sum_{i=1}^{m} x_i^{\beta} - log(\alpha) \sum_{i=1}^{m} e^{-\lambda x_i^{\beta}} - b_3\lambda\right) \times \prod_{i=1}^{m} \left(1 - \alpha^{-e^{-\lambda x_i^{\beta}}}\right)^{R_i}. \tag{3.16}$$

It is clear that the full conditional distributions of α, β and λ , respectively,can not be reduced to any well-known distributions.

3.2.3 Bayes Estimators Under Squared Error Loss (SEL)

Under the SEL function, the BEs of any function $\phi(\alpha, \beta, \lambda)$ is the posterior mean:

$$\tilde{\phi}_{SEL} = \mathbb{E}[\phi(\alpha, \beta, \lambda)|x] = \frac{\int_0^\infty \int_0^\infty \int_0^\infty \phi(\alpha, \beta, \lambda)\, \pi(\alpha, \beta, \lambda|x)\, d\alpha d\beta d\lambda}{\int\int\int \pi(\alpha, \beta, \lambda|x)\, d\alpha d\beta d\lambda}. \tag{3.17}$$

Since the integrals in Equation (3.17) are analytically intractable, we employ the Markov Chain Monte Carlo (MCMC) method[22], particularly the Metropolis-Hastings (MH) algorithm, to approximate the BEs.

3.3 Numerical Results

3.3.1 Metropolis-Hastings Sampling Procedure

Since direct sampling of α, β, and λ is infeasible, we employ the M-H algorithm with a normal proposal distribution. This allows us to generate samples from the full conditional distributions to obtain Bayes estimates and construct the highest posterior density(HPD) credible intervals for the unknown parameters as well as RF and HRF.

we suggest applying the following steps of the M-H algorithm:

step (1) Set initial values for (α, β, λ), denoted by $(\alpha^{(0)}, \beta^{(0)}, \lambda^{(0)})$.

step (2) Set iteration index $j = 1$.

step (3) Generate α^* from $\mathcal{N}(\alpha^{(j-1)}, \hat{v}(\alpha^{(j-1)})$.

step (4) Compute the acceptance probability:

$$p(\alpha^{(j-1)}|\alpha^*) = \min\left\{1, \frac{\pi_1(\alpha^*|\beta^{(j-1)}, \lambda^{(j-1)})}{\pi_1(\alpha^{(j-1)}|\beta^{(j-1)}, \lambda^{(j-1)})}\right\}.$$

step (**5**) Generate $u \sim U(0,1)$.

step (**6**) If $u \leq p(\alpha^{(j-1)}|\alpha^*)$, set $\alpha^{(j)} = \alpha^*$; otherwise, set $\alpha^{(j)} = \alpha^{(j-1)}$.

step (**7**) Repeat steps 3–6 for β and λ to generate $\beta^{(j)}$ and $\lambda^{(j)}$ from their corresponding posterior densities.

step (**8**) Compute the RF and HRF by substituting $(\alpha^{(j)}, \beta^{(j)}, \lambda^{(j)})$ into their respective functional forms for $t > 0$.

step (**9**) Set $j = j + 1$.

step (**10**) Repeat steps 3–8 for a total of Q iterations, resulting in the samples:

$$\left\{\alpha^{(j)}, \beta^{(j)}, \lambda^{(j)}, R^{(j)}(t), h^{(j)}(t)\right\}, \quad j = 1, 2, \ldots, Q.$$

step (**11**) Discard the first M samples as burn-in. Then, compute the Bayes estimates of α, β, λ, $R(t)$, and $h(t)$ under the SEL function as:

$$\tilde{\phi}_{\mathrm{SEL}} = \frac{1}{Q-M} \sum_{j=M+1}^{Q} \phi^{(j)}.$$

step (**12**) Under the LL function proposed by Varian [23], the Bayes estimates are given by:

$$\tilde{\phi}_{\mathrm{LL}} = -\frac{1}{q}\, log\left(\frac{1}{Q-M} \sum_{j=M+1}^{Q} e^{-q\phi^{(j)}}\right), \quad q \neq 0.$$

step (**13**) Finally, we apply the method proposed by Chen and Shao [8] to compute the HPD credible intervals for α, β, λ, $R(t)$, and $h(t)$ using the posterior samples.

3.3.2 Methodology of the Monte Carlo Simulation

A comprehensive MC simulation study is conducted to compare the behavior of the proposed estimators for the parameters α,β,λ,$R(t)$ and $h(t)$.

Simulation Design and Estimation Procedure

The simulation assumes true parameter values $(\alpha, \beta, \lambda) = (0.5, 1.5, 0.1)$, and reliability characteristics are computed at time point $t = 0.5$, for which the true values are $R(t) = 0.952$ and $h(t) = 0.145$. Various combinations of sample sizes $n = 50$ and $n = 100$, and numbers of observed failures $m = 20, 40, 80$ are used. Three progressive censoring schemes are considered:

- Scheme 1: $R_1 = n - m$, $R_i = 0$ for $i \neq 1$

- Scheme 2: $R_{m/2} = n - m$, $R_i = 0$ for $i \neq m/2$
- Scheme 3: $R_m = n - m$, $R_i = 0$ for $i \neq m$

Two informative gamma priors are adopted for Bayesian estimation:

- Prior 1: $a_i = (2.5, 7.5, 0.5)$, $b_i = 5$
- Prior 2: $a_i = (5, 15, 1)$, $b_i = 10 \quad$ for $i = 1, 2, 3$

The M-H algorithm was used to generate 12,000 MCMC samples. The first 2,000 iterations were discarded as a burn-in period, and the remaining 10,000 samples were used to compute the Bayesian estimates and their 95% HPD credible intervals.

The simulation design and estimation procedure were inspired by the methodology used in Alotaibi et al.(2022)[3]

3.4 Comparison Between Obtained Estimators on Real Engineering Data

This section presents the practical application of the proposed estimators using two real-life datasets from the engineering field to evaluate the performance of APW distribution under progressive Type-II censoring.

Data Description

The first dataset, referred to as **Data-I**, consists of the failure times of 20 mechanical components reported by Murthy and al.[18]. The second dataset, **Data-II**, contains accelerated life test data of metallic specimens subjected to a stress level of 2.6×10^4 psi.These data set hase been reported and analysed by Cheng and Elsayed [9]. The ordered data points of both data sets I and II are provided in (3.1).

Table 3.1: The failure times of mechanical components and metal-coupons.

Data	Failure Times
I	0.067, 0.068, 0.076, 0.081, 0.084, 0.085, 0.085, 0.086, 0.089, 0.098, 0.098, 0.114, 0.114, 0.115, 0.121, 0.125, 0.131, 0.149, 0.160, 0.485
II	2.33, 2.58, 2.68, 2.76, 2.90, 3.10, 3.12, 3.15, 3.18, 3.21, 3.21, 3.29, 3.35, 3.36, 3.38, 3.38, 3.42, 3.42, 3.42, 3.44, 3.49, 3.50, 3.50, 3.51, 3.51, 3.52, 3.52, 3.56, 3.58, 3.58, 3.60, 3.62, 3.63, 3.66, 3.67, 3.70, 3.70, 3.72, 3.72, 3.74, 3.75, 3.76, 3.79, 3.79, 3.80, 3.82, 3.89, 3.89, 3.95, 3.96, 4.00, 4.00, 4.00, 4.03, 4.04, 4.06, 4.08, 4.08, 4.10, 4.12, 4.14, 4.16, 4.16, 4.16, 4.20, 4.22, 4.23, 4.26, 4.28, 4.32, 4.32, 4.33, 4.33, 4.37, 4.38, 4.39, 4.39, 4.43, 4.45, 4.45, 4.52, 4.56, 4.56, 4.60, 4.64, 4.66, 4.68, 4.70, 4.70, 4.73, 4.74, 4.76, 4.76, 4.86, 4.88, 4.89, 4.90, 4.91, 5.03, 5.17, 5.40, 5.60

To assess the suitability of the APW distribution, the Kolmogorov-Smirnov (K-S) statistic and corresponding p values were calculated for each datase, and the model parameters the MLEs of the model parameters (α, β, λ) were estimated using the MLEs, as presented in (3.2). In addition, the algorithm proposed by Balakrishnan and Cramer [5] was applied to generate progressive Type-II censored samples from the complete data sets. The censoring schemes adopted in this analysis are listed in (3.3).

Table 3.2: Summary fit of the APW distribution under real data sets.

Data	$\hat{\alpha}$ (SE)	$\hat{\beta}$ (SE)	$\hat{\lambda}$ (SE)	K-S (p-value)
I	42849.2 (9.2940)	0.92007 (0.1488)	21.2764 (7.1551)	0.181 (0.530)
II	68891.6 (2.0970)	2.72066 (0.2052)	0.00658 (0.0192)	0.045 (0.986)

Table 3.3: Various Type-II progressively censored samples from mechanical components and metal-coupons data sets.

Data (Sample)	m	R	**Type-II Progressive Censored Data**
Data-I (S_1)	15	(5, 0*14)	0.067, 0.085, 0.086, 0.089, 0.098, 0.098, 0.114, 0.114, 0.115, 0.121, 0.125, 0.131, 0.149, 0.160, 0.485
Data-I (S_2)		(0*7, 5, 0*7)	0.067, 0.066, 0.076, 0.081, 0.084, 0.085, 0.085, 0.086, 0.115, 0.121, 0.125, 0.131, 0.149, 0.160, 0.485
Data-I (S_3)		(0*14, 5)	0.067, 0.066, 0.076, 0.081, 0.084, 0.085, 0.085, 0.086, 0.089, 0.098, 0.098, 0.114, 0.114, 0.115, 0.121
Data-II (S_1)	45	(57, 0*44)	2.33, 4.10, 4.12, 4.14, 4.16, 4.16, 4.16, 4.20, 4.22, 4.23, 4.26, 4.28, 4.32, 4.32, 4.33, 4.33, 4.37, 4.38, 4.39, 4.39, 4.43, 4.45, 4.45, 4.52, 4.56, 4.60, 4.64, 4.66, 4.68, 4.73, 4.79, 4.81, 4.84, 4.88, 4.90, 4.95, 4.97, 5.00, 5.05, 5.10, 5.14, 5.18, 5.21, 5.22, 5.25
Data-II (S_2)		(0*22, 57, 0*22)	2.33, 2.68, 2.72, 2.90, 3.14, 3.16, 3.18, 3.21, 3.21, 3.25, 3.35, 3.36, 3.38, 3.39, 3.40, 3.41, 3.44, 3.46, 3.50, 3.51, 3.55, 3.56, 3.58, 3.58, 3.58, 3.60, 3.62, 3.63, 3.63, 3.66, 3.67, 3.70, 3.72, 3.74, 3.75, 3.76, 3.79, 3.79, 3.80
Data-II (S_3)		(0*44, 57)	2.33, 2.58, 2.68, 2.72, 2.90, 3.14, 3.16, 3.18, 3.21, 3.21, 3.25, 3.35, 3.36, 3.38, 3.39, 3.40, 3.41, 3.44, 3.46, 3.50, 3.51, 3.55, 3.56, 3.58, 3.58, 3.58, 3.60, 3.62, 3.63, 3.63, 3.66, 3.67, 3.70, 3.72, 3.74, 3.75, 3.76, 3.79, 3.79, 3.80

In (3.2), the results show that the p-values for both datasets greater than 0.05, indicating that the APW distribution is well-suited to represent the real-world data.
In table (3.3), demonstrates the adaptability of the model to censored and incomplete data settings.

Bayesian and MLE with Interval Analysis

The parameters α, β, and λ, along with the $R(t)$ and $h(t)$, were estimated using both the MLE method and the Bayesian approach under the SEL and LINEX loss functions. The results are reported in table 3.4. and to evaluate the precision of the interval estimates, the mean confidence length (ACL) were calculated for both ACI derived from MLE and HPD intervals obtained through Bayesian estimation, as shown in table 3.5.

Table 3.4: MLE and Bayesian point estimates with their (SEs).

Data (Sample)	**Par.**	**MLE**	**SEL**	**LL ($q = -3$)**	**LL ($q = +3$)**
Data-I (S_1)	α	5215.91 (9.9806)	5215.79 (0.0012)	5215.79 (0.1173)	5215.79 (0.1175)
	β	0.9813 (0.1654)	0.97475 (0.0009)	0.97475 (0.0963)	0.97475 (0.1066)
	λ	19.6584 (7.6000)	19.5330 (0.0013)	19.5330 (0.1253)	19.5329 (0.1251)
	$R(0.1)$	0.6668 (0.0897)	0.6547 (0.0004)	0.6547 (0.0121)	0.6546 (0.0125)
	$h(0.1)$	11.0652 (3.0169)	11.3187 (0.0120)	11.3274 (0.2624)	11.3099 (0.2449)
Data-I (S_2)	α	4109.16 (7.3782)	4109.12 (0.0014)	4109.12 (0.0920)	4109.12 (0.0934)
	β	0.9064 (0.1382)	0.90381 (0.0002)	0.90381 (0.0026)	0.90380 (0.0022)
	λ	17.2650 (5.2702)	17.2594 (0.0002)	17.2594 (0.0507)	17.2594 (0.0505)
	$R(0.1)$	0.62372 (0.0906)	0.61496 (0.0004)	0.61497 (0.0887)	0.61497 (0.0887)
	$h(0.1)$	11.4492 (2.9366)	11.6298 (0.0109)	11.6370 (0.1687)	11.6227 (0.1737)
Data-I (S_3)	α	3037.22 (12.1411)	3037.10 (0.0012)	3037.10 (0.0700)	3037.10 (0.0675)
	β	1.87803 (0.0651)	1.87459 (0.0012)	1.87459 (0.0022)	1.87458 (0.0030)
	λ	179.7011 (1.1842)	179.581 (0.0007)	179.581 (0.0537)	179.580 (0.0533)
	$R(0.1)$	0.55294 (0.0957)	0.54048 (0.0006)	0.54048 (0.0103)	0.54048 (0.0104)
	$h(0.1)$	30.1922 (5.9534)	30.7723 (0.0289)	30.8225 (0.6305)	30.7221 (0.5303)
Data-II (S_1)	α	13184.1 (7.5350)	13184.1 (0.0005)	13184.1 (0.0191)	13184.1 (0.0197)
	β	5.0096 (0.0631)	5.0089 (0.0003)	5.0089 (0.0040)	5.0089 (0.0044)
	λ	4.5494 (0.0561)	4.5482 (0.0002)	4.5482 (0.0065)	4.5482 (0.0063)
	$R(4.5)$	0.22086 (0.0366)	0.20115 (0.0002)	0.20116 (0.0093)	0.20115 (0.0092)
	$h(4.5)$	3.0188 (0.8109)	3.0591 (0.0138)	3.0592 (0.0641)	3.0591 (0.0635)
Data-II (S_2)	α	23183.8 (2.2339)	23184.0 (0.0005)	23184.0 (0.0035)	23184.0 (0.0034)
	β	6.0392 (0.0328)	6.0370 (0.0002)	6.0370 (0.0045)	6.0370 (0.0045)
	λ	2.2188 (0.0508)	2.2155 (0.0002)	2.2155 (0.0080)	2.2155 (0.0081)
	$R(4.5)$	0.38192 (0.0537)	0.37624 (0.0003)	0.37624 (0.0051)	0.37624 (0.0052)
	$h(4.5)$	3.0412 (0.6852)	3.0138 (0.0118)	3.0138 (0.0663)	3.0138 (0.0664)
Data-II (S_3)	α	3050.18 (0.9681)	3050.2 (0.0002)	3050.2 (0.0030)	3050.2 (0.0031)
	β	2.9132 (0.0312)	2.9136 (0.0002)	2.9136 (0.0045)	2.9136 (0.0045)
	λ	83.2076 (0.5012)	83.1268 (0.0003)	83.1268 (0.0074)	83.1268 (0.0075)
	$R(4.5)$	0.20862 (0.0385)	0.20175 (0.0003)	0.20175 (0.0085)	0.20175 (0.0086)
	$h(4.5)$	3.17180 (0.8818)	3.1477 (0.0014)	3.1478 (0.0239)	3.1476 (0.0242)

Table 3.5: Two-sided 95% ACI/HPD credible interval estimates with their [lengths].

Data (Sample)	Par.	ACI	HPD
Data-I (S_1)	α	(5196.32, 5235.40) [39.123]	(5215.31, 5216.26)[0.9531]
	β	(0.65693, 1.30542) [0.6485]	(0.87338, 1.07297) [0.1996]
	λ	(6.25054, 39.92530) [27.675]	(10.9529, 20.0888) [9.1359]
	$R(0.1)$	(0.49210, 0.84254) [0.3504]	(0.47944, 0.81822) [0.3388]
	$h(0.1)$	(0.10165, 0.47025) [0.3686]	(0.10238, 0.42260) [0.3202]
Data-I (S_2)	α	(4076.31, 4142.13) [65.768]	(4108.73, 4109.50) [0.7737]
	β	(0.77452, 1.41599) [0.6415]	(1.12760, 1.27295) [0.1454]
	λ	(6.93102, 27.5973) [20.666]	(17.1641, 17.3587) [0.1946]
	$R(0.1)$	(0.54478, 0.91921) [0.3744]	(0.58897, 0.91018) [0.3212]
	$h(0.1)$	(0.25040, 1.19322) [0.9428]	(0.37761, 1.38590) [1.0082]
Data-I (S_3)	α	(3034.20, 3067.24) [33.042]	(3036.14, 3036.86) [0.7218]
	β	(1.76811, 1.98884) [0.2197]	(1.78621, 1.97257) [0.1864]
	λ	(6.88532, 9.20257) [2.3172]	(7.05691, 8.94801) [1.8911]
	$R(0.1)$	(0.68386, 0.92160) [0.2377]	(0.75402, 0.91181) [0.1578]
	$h(0.1)$	(0.18322, 0.51653) [0.3333]	(0.34789, 0.68975) [0.3419]
Data-II (S_1)	α	(13169.3, 13198.8) [29.538]	(13183.8, 13184.3) [0.3942]
	β	(3.81867, 5.78267) [1.9640]	(3.50499, 5.20983) [1.7048]
	λ	(0.00000, 0.00269) [0.0026]	(0.00000, 0.00228) [0.0023]
	$R(4.5)$	(1.28740, 2.75757) [1.4737]	(1.46972, 2.58228) [1.1126]
	$h(4.5)$	(0.21054, 1.65011) [1.4396]	(0.22366, 1.43333) [1.2097]
Data-II (S_2)	α	(1.84368, 2.78397) [0.9403]	(2.24354, 2.38236) [0.1388]
	β	(3.64932, 5.19769) [1.5484]	(3.29496, 5.10868) [1.8137]
	λ	(0.00000, 0.00368) [0.0037]	(0.00000, 0.00329) [0.0033]
	$R(4.5)$	(0.26859, 1.49060) [1.2220]	(0.28783, 1.40897) [1.1211]
	$h(4.5)$	(0.12093, 1.45211) [1.3312]	(0.13882, 1.31335) [1.1745]
Data-II (S_3)	α	(3082.16, 3084.69) [2.5257]	(3044.39, 3044.94) [0.5470]
	β	(1.62004, 1.66970) [0.0497]	(1.62668, 1.66288) [0.0362]
	λ	(0.00000, 0.00014) [0.0001]	(0.00000, 0.00011) [0.0001]
	$R(4.5)$	(1.40594, 2.19600) [0.7901]	(1.45360, 2.10871) [0.6551]
	$h(4.5)$	(1.44349, 4.90012) [3.4566]	(2.62990, 3.70098) [1.0711]

In table 3.4, the results show a high level of agreement between the MLE and Bayesian estimates, indicating stability and consistency in the estimation process,with the Bayesian under SEL has the smallest SEs,reflecting higher precision. The values of $R(t)$ were relatively high, reflecting a strong survival probability, while the values of $h(t)$ remained low, indicating an acceptable failure rate. And in table 3.5, the Bayesian HPD intervals outperform the traditional ACIs by providing generally shorter interval lengths. This supports the preference for Bayesian estimation, particularly under progressive censoring.

Conclusion

This thesis has explored statistical inference methodologies for analyzing data subject to progressive Type II censoring, with a focus on two important lifetime distributions: the Weibull distribution and the Alpha Power Weibull (APW) distribution. Progressive Type II censoring, by allowing systematic removal of surviving units at multiple stages, offers a flexible and resource-efficient framework for reliability and survival studies. The work has demonstrated how classical maximum likelihood estimation (MLE) and Bayesian estimation approaches can be effectively adapted to this censoring scheme.

The theoretical developments presented include detailed derivations of likelihood functions, score functions, and Fisher information matrices under progressive censoring, which enable the computation of MLEs and their asymptotic properties. Bayesian methods were also developed, incorporating prior information and utilizing Monte Carlo simulation techniques such as the Metropolis-Hastings algorithm to approximate posterior distributions and obtain Bayesian estimators under various loss functions.

A comparative numerical study through extensive simulations showed that Bayesian estimators often perform favorably in terms of bias and mean squared error, especially when prior information is available. The application to real engineering data illustrated the practical utility of the proposed methods, with the APW distribution providing a more flexible fit than the classical Weibull distribution due to its additional shape parameter.

Overall, this study contributes to the reliability and survival analysis literature by providing robust inferential tools tailored to progressive Type-II censored data, enhancing the modeling flexibility and inferential accuracy. Future work may extend these methods to other complex censoring schemes and explore their applications in diverse fields such as biomedical studies and industrial quality control.

Bibliography

[1] Abu Awwad, R. R., Raqab, M. Z., and Al-Mudahakha, I. M. (2014). Statistical inference based on progressively type II censored data from Weibull model. *Communications in Statistics - Simulation and Computation.* https://doi.org/10.1080/03610918.2013.842589

[2] Ahmed, E. A. (2014). Bayesian estimation based on progressive Type-II censoring from two parameter bathtub-shaped lifetime model: an Markov chain Monte Carlo approach, *Journal of Applied Statistics*, 41:4, 752-768.

[3] Alotaibi, R., Nassar, M., Rezk, H., and Elshahhat, A. (2022). Inferences and engineering applications of Alpha Power Weibull distribution using progressive Type-II censoring. *Mathematics*, 10(16), 2901. https://doi.org/10.3390/math10162901

[4] Balakrishnan, N. and Aggrawala, R. (2000). *Progessive censoring: Theory, methods and applications*. Birkhauser, Boston.

[5] Balakrishnan, N., Cramer, E. (2014) *The Art of Progressive Censoring.* Springer: Birkhauser, NY, USA.

[6] Balakrishnan, N. and Kateri, M. (2008). On the maximum likelihood estimation of parameters of Weibull distribution based on complete and censored data. *Statistics & Probability Letters*, 78(17), 2971-2975.

[7] Berger, J. O. and Sun, D. (1993). Bayesian analysis for the Poly-Weibull distribution, *Journal of the American Statistical Association*, 88, 1412-1418.

[8] Chen, M. H., Shao, Q. M. (1999). Monte Carlo estimation of Bayesian credible and HPD intervals. *J. Comput. Graph. Stat.* , 8, 69-92.

[9] Cheng, Y., Elsayed, E. (2013). Accelerated Life Testing Model for a Generalized Birnbaum-Saunders Distribution. QUALITA2013. Available online: https://hal.archives-ouvertes.fr/hal-00823134/

[10] Dey, S., Wang, L., Nassar, M.(2021) Inference on Nadarajah-Haghighi distribution with constant stress partially accelerated life tests under progressive type-II censoring. *J. Appl. Stat.* , 49, 1-22.

[11] Devroye, L. (1984). A simple algorithm for generating random variates with a log-concave density.*Computing*, 33, 247-257.

[12] Held, L., Daniel, S. B (2016) .*Applied statistical inference.* Springer, Berlin Heidelberg. 16, 978-3 .

[13] Johnson, N. L, Kotz, S. and Balakrishnan, N. (1995). *Continuous univariate distribution*, 2nd edition, Wiley and Sons, New York.

[14] Kundu, D. (2008). Bayesian inference and life testing plan for the Weibull distribution in presence of progressive censoring, *Technometrics*, 50, 144-154.

[15] Kundu, D., Howlader, H. (2010). Bayesian inference and prediction of the inverse Weibull distribution for Type-II censored data.*Comput. Stat. Data Anal*, 54, 1547-1558.

[16] Lawless, J. F. (1982). *Statistical models and methods for life time data*. John Wiley, New York.

[17] Mahdavi, A., Kundu, D. (2017). A new method for generating distributions with an application to exponential distribution.*Commun-Stat-Theory Methods*, 46, 6543-6557.

[18] Murthy, D. N. P., Xie, M., Jiang, R. (2004).*Weibull Models.* Wiley Series in Probability and Statistics; Wiley: Hoboken, NJ, USA.

[19] Nassar, M. ,Alzaatreh, A. , Mead, M. , Abo-Kasem, O. (2017). Alpha power Weibull distribution: Properties and applications.*Commun.-Stat.-Theory Methods*, 46, 10236-10252.

[20] Nassar, M.; Dey, S., Wang, L. , Elshahhat, A. (2021) Estimation of Lindley constant-stress model via product of spacing with Type-II censored accelerated life data.*Commun.-Stat.-Simul. Comput.*

[21] Runcorn, K. (1989). Sir Harold Jeffreys (1891-1989). Nature, 339(6220), 102-102.

[22] Robert, C. P., & Casella, G. (2004). *Monte Carlo Statistical Methods.* Springer

[23] Varian, H. R. (1975). A Bayesian approach to real state assessment. In E. F. Stephen A. Zellner (Eds.), *In Studies in Bayesian Econometrics and Statis-*

tics in Honor of Leonard J. Savage (pp. 195-208). North-Holland Publishing Co.: Amsterdam, The Netherlands.