Bayesian Inference of the Weibull-Pareto Distribution

By

Reuben Kimutai Kipchumba

Reg No. I56/37629/2020

DEPARTMENT OF MATHEMATICS

FACULTY OF SCIENCE AND TECHNOLOGY

UNIVERSITY OF NAIROBI

Project submitted in Partial Fulfillment of the Requirements for the award of the Degree of Master of Science in Mathematical Statistics

June 2022

Bayesian Inference of the Weibull-Pareto Distribution

Research Report in Mathematics, Number 2, 2022

Reuben Kimutai Kipchumba
School of Mathematics
College of Biological and Physical sciences
Chiromo, off Riverside Drive
30197-00100 Nairobi, Kenya

Master Thesis

Submitted to the School of Mathematics in partial fulfilment for a degree in Master of Science in Mathematical Statistics

Submitted to: The Graduate School, University of Nairobi, Kenya

Declaration and Approval

I, the undersigned, declare that this dissertation is my original work and to the best of my knowledge, has not been submitted in support of an award of a degree in any other university or institution of learning.

25-08-2022

Signature Date

Reuben Kimutai Kipchumba
Reg No. I56/37629/2020

In my capacity as a supervisor of the candidate's dissertation, I certify that this dissertation has my approval for submission.

6-September-2022

Signature Date

Professor Ivivi J. Mwaniki,
School of Mathematics,
University of Nairobi,
Box 30197, 00100 Nairobi, Kenya.
E-mail: jimwaniki@uonbi.ac.ke

Abstract

There are a wide range of industries that make use of the Weibull distribution, including industrial engineering, general insurance, survival analysis, electrical engineering, and reliability engineering, to name just a few. The Weibull distribution is expanded to become the Weibull-Pareto distribution, also known simply as the Weibull distribution. A notable use of the Weibull-Pareto distribution is in the modeling of asymmetrical data, which is also one of its most important applications. During the course of this inquiry, a hierarchical Bayesian model will be created with the use of a Weibull-Pareto distribution as a reference.

Words to note: MCMC, survival model, right censoring,WPD, Heavy-tailed Distribution, hierarchical Bayesian model,

Dedication

I dedicate this work to my family and friends, as well as those pursuing higher education – don't give up no matter what!

Acknowledgement

I'd want to express my heartfelt appreciation to my supervisor, Professor Joseph Ivivi Mwaniki, for his mentoring during this study. I appreciate my parents', Joyce and Moris' help in maintaining calm throughout the last two years.

Timothy, Elias, and Philemon are my brothers; Linah, Felisters, and Mitchell are my sisters; and Collins, Francis, Gitau, and Justus are my pals. I'd also want to thank Carren Chepkemoi for his invaluable assistance during this project.

Finally, I'd like to thank me for always working so hard to ensure that you're reading this!

Contents

Chapter 1

Introduction

1.1 Study Background

In the field of modeling survival analysis and reliability, the Weibull distribution is applied rather frequently. Because of the value of the shape parameter, it is adaptable and can take on the characteristics of a variety of other distributions. Because of this, it is extremely well-liked among quality control engineers and analysts, particularly those who have experience working with data modeling and reliability. A huge number of academics have come up with many versions of the Weibull distribution; one of these variants is called the EPD(Exponential-Weibull distribution). (Mudholkar et al.,1995).

Both the generalized Weibull distribution that Kollia and Mudholder (1994) produced, as well as the Beta-Weibull distribution that Famoye et al.(2005) constructed, have certain similarities. A prior discussion introduced the idea of the New generalized Weibull distribution. This distribution can also be referred to as the Weibull Pareto distribution. It was generated from a family of probability density models known as "Transformed-Transformer" (Alzaatreh et al., 2013). This distribution has a severely skewed nature when compared to the Weibull distribution, which is the standard. As a consequence of this, the tactic that is suggested is modeling extremely skewed data, which is something that frequently occurs in survival analysis and dependability.

According to Alzaatreh et al. (2013), the value of the shape parameter of the WPD is smaller than one. This conclusion was reached in 2013. There is no such thing as a Maximum Likelihood Estimator (MLE), and this applies to both the scale and the shape parameter. After that, they mandated two different parameter estimation methodologies, which were the modified maximum likelihood estimation (MMLE) and the alternative maximum likelihood estimation (AMLE). Despite the fact that this was the case, the AMLE resulted in a significant bias, and using the MMLE is costly. In the following part, 2.1.1, we will talk about the two different approaches.

The Bayesian Weibull Pareto Model will be explained in detail during the course of this article's discussion. Bayes' methodology differs from the frequentist approach in that it operates under the assumption that the parameters are subject to random variation and adhere to the probability that was previously stated. The researcher's prior credence regarding the distribution parameters is defined by the random distribution, which is captured as the prior distribution.

Estimating Bayesian parameters has been the subject of multiple publications authored by a number of researchers over the past few years. Estimating the generalized lognormal distribution was the topic of an article that Perez and Martin (2009) wrote. With the use of Type-I censoring, Aslam and Noor (2013) were able to estimate the parameters of the inverse Weibull distribution. When conducting data analysis, the Weibull distribution has proven to be extremely helpful, particularly when used to censored data, which is the kind of study that is typically performed in survival analysis. For instance, survival time is the end result that is of interest in cancer research since it is the event of interest. In spite of this, it is sometimes known as the time between complete remission to the beginning of a recurrence. In addition, there is the possibility that some people will not experience the event of interest by the time the actual time to occurrence has passed. As a result, a censored observation takes place. As a consequence of this, in the event that it takes place, we will truncate the final findings due to random factors.

1.2 Problem Statement

Because of the extreme skewness of many real-world data sets, the most recent iteration of the Weibull Distribution is unable to effectively simulate the distributions of a number of these collections of data. Because of this, it is essential to broaden the application of the Weibull-Pareto distribution by incorporating a fresh component into the model. Because of this, the current Weibull-Pareto distribution will have greater flexibility, and the resulting distribution will provide a better fit than the Weibull-Pareto Distribution. The Weibull-Pareto distribution needs to be extended in order to account for this necessity.

1.3 Study Objectives

The purpose of this study is to propose a novel distribution that will be referred to as the generalized Weibull-Pareto Distribution (Generalised WPD), and to deduce its features using the transformation.

The following is a list of the specific goals:
i. Construct the Generalized Weibull-Pareto Distribution, analyze the reliability of the data, and derive a variety of structural features.

While the General Objectives are

i. Make an educated guess as to the values of the modified distribution's parameters.
ii. Evaluate how well the WPD is working.

1.4 Importance of the Study

The incorporation of a generalized parameter into the Weibull-Pareto Distribution will result in a significant increase in both the sensitivity and the efficacy of the statistical tests associated with the distribution. This will be the case because the sensitivity of the tests will increase while the efficacy of the tests will remain the same. This is going to be the case due to the fact that the test's level of sensitivity is going to improve, although the test's level of accuracy will remain unchanged. As a result of this, it will be possible to model and carry out flexible analyses of skewed data sets based on real-world examples in a wide variety of application domains. These studies can be conducted in a wide variety of application domains. There are some real-world data sets that do not follow the Weibull or Pareto distributions as expected. These data sets can be found all around the world. As a direct consequence of this, the Weibull-Pareto distribution needs to be extended by adding a new parameter to it in order to boost its adaptability and make it applicable to a wider number of scenarios. This can be accomplished by making the distribution more general.

Chapter 2

LITERATURE REVIEW

PARAMETER ESTIMATES-BAYESIAN MODEL

2.1 WP Distribution

Lifespans, on the whole, are an excellent illustration of the kind of favorably skewed data that frequently occurs in analyses of survival and dependability. For instance, the data on breast cancer that Khan et al. (2014) investigated is skewed to the right. When modeling this kind of data, the Weibull and distributions are typically the ones that are utilized. These distributions are able to describe right-skewed data because of the parameter choices; but, in the ecent that the data is substantially skewed, As a result of their more rapid decline, the models are unable to accurately represent the right-hand heavy tail.

Alzaatreh et al.(2013) were the ones who came up with the Weibull-Pareto distribution). This distribution is a member of a family of distributions that is commonly referred to as the "Transformed-Transformers" family (T-X family).

2.1.1 Estimation

Estimating the model parameters is our primary objective whenever we work with a probability distribution. The approach of maximum likelihood estimation (MLE) is a good illustration of a choice that should be evident. However, Alzaatreh et al. (2013) emphasized that applying this strategy requires taking into consideration two primary concerns.

The first scenario is when c is less than 1, This indicates that the likelihood function approaches infinity when the θ value trends toward the highest value that has been seen, which is represented by the symbol $x_{(1)}$. In the case where c is less than 1 and the estimate θ is approximated by $x_{(1)}$, the β and c MLE does not exist. Smith (1985) conducted study on the topic and presented a different approach known as the AMLE, which makes the assumption θ to be equal $x_{(1)}$. Although employing this method will result in a cheaper cost of computation, it is not successful when c is bigger than 1 since it does not take into account the exponential growth of the data. Figure 2.1 illustrates that when c is less than one, $x(1)$is possible to estimate the most accurate value for θ due to the fact that the probability density function is at its highest point. Despite the fact that this is the case, the WP distribution will become more symmetric as c continues to grow; hence,$x_{(1)}$ will not be able to produce an exact estimate when alternative MLE methods are utilized.

In a subsequent piece of research, Alzaatreh et al. (2013) evaluated the parameters of the WPD using a different technique known as maximum likelihood estimation. This method was utilized. In his computations, he ensured that he took into consideration the probability densities.

The alternate version of greatest likelihood is not quite transparent but consistent. The behavior of it with comparatively smaller samples is not clearly defined at this point. A simulation was carried out by Alzaatreh et al. (2013) with a sample size of two and a total of 36 parameters. They came to the conclusion that the parameters β and θ should each have values of 0,5, 1, and 3, and they came to the conclusion that c should have values of 0,5, 1, 3, and 7. Both a sample size of n = 500 and a sample size of n = 100 were chosen after some deliberation. They applied the transformation that was discussed before for each and every one of the parameters. Within this transformation, the random variable generated are $y1, y2, ...yn$, the random variables belongs to the Weibull distribution with parameters $\frac{1}{\beta}$ and c. Consequently, a random sample taken from the WP has the formula $x_i = \theta exp(y_i)$

When c was more than one, the simulation results revealed that the alternative maximum likelihood estimator possessed a significant amount of bias. In cases where $\hat{\theta} = x(1) > \theta$ by only a narrow margin, there will be a significant amount of bias for the variable c. The conclusion that may be drawn from this is caused by$\hat{c}$.

A significant amount of bias is produced by the phrase $\log(\log(\frac{x_i}{\hat{\theta}}))$. Alzaatreh et al. (2013) explained that if the actual parameter $\theta = 1$, the estimated parameter $\hat{\theta} = 1.3$, and the observed parameter $x_i = 1.3001$, then this indicates that the value $x_i = 1.3001$

is the one that should be used. As a result, the $\log[\log(\frac{x_i}{\hat{\theta}})] = -9.4727$, whereas the initial value is $\log[\log(\frac{x_i}{\theta})] = -1.3377$. In particular, the results of the simulations show that an incorrect estimate of the true value of θ is produced whenever c is made larger and $\hat{\theta} = x_i$ is used.

The other possibility is that c is greater than one, in which case, the estimate θ will be poorly predicted by $x_{(1)}$ and the AMLE will depict a substantial huge bias in its results. Smith (2014) created a modified MMLE,which will be helpful in this circumstance.

2.2 Introduction to Bayesian Statistics

Bayesian statistics is a method of applying the possibility to problems. It gives us mathematical techniques to inform our perceptions about stochastic events in response to new evidence or data about such events (Giovagnoli, 2008). Bayesian inference, in particular, perceives probability as a way of measuring beliefs or trust that a person may have in a specified event. Researchers may well have preconceived ideas for an occasion, but that belief will vary if new evidence is presented (Rossi Allenby, 2003). Bayesian statistics provides a solid theoretical formulation for integrating preconceived beliefs and proof to generate posterior perceptions.

We begin with a definition of conditional probability, which provides a rule for deciding the likelihood of an outcome A given the incidence of some other incident B.The conditional Probability is given by;

$p(A|B) = \frac{P(B|A)P(A)}{P(B)}$ where $P(B) \neq 0$

The parameter P(B) is a simple constant that does not play a big role due to the fact that it is the marginal distribution and does not depend on our previous perspectives. This is due to the fact that the marginal distribution is unaffected by the ideas we have held in the past (Giovagnoli, 2008; Rigollet, 2016). The equation that was just given can, as a direct consequence of this, be rewritten as follow

$P(A|B) \propto P(B|A)P(A)$

If we have a multivariate data distribution, we derive the likelihood function $L(A)$, the posterior distribution will be given by

$P(A|B) \propto L(A)P(A)$

This is the likelihood function that is measured based on an earlier belief (Rigollet, 2016). In spite of the fact that these are not traditional nor frequentist data, it is still possible to reconstruct the result by working under the assumption that there is no prior. This is done by working under the premise that there is no prior variable. In other words, the preceding might be disregarded as relevant by making the assumption that it is comparable to a unit, which will simply serve to emphasize the likelihood (Rigollet, 2016).

It is possible to have invariant priors, which are prior records that contain no additional information at all, but this can only happen under very particular circumstances (Giovagnoli, 2008). Instead, and more generally speaking, some researchers use an objective approach to Bayesian statistics, employing a "non-informative" or "reference" prior. This is an event that led up that carries few (if any) details and allows the data to start driving the results. This makes it possible to have a more straightforward connection between the facts and the inferences that can be taken from them (Giovagnoli, 2008; Van Dongen, 2006).

A prior of this type presumes that the occurrence of any event is comparable to any other (Giovagnoli, 2008). The purpose is to be as objective as is humanly possible while avoiding attempts to affect the likelihood function with previous subjective views, which could cause the results to be inaccurate (Van Dungan, 2006). A non-informative prior would be something like the uniform distribution as an example (Van Dungan, 2006). There are a number of drawbacks associated with using a non-informative prior, and Bayesian statistics in particular is based on the concept that we always have an ex ante anticipation, which can be utilized in a variety of settings. Despite these drawbacks, Bayesian statistics remains one of the most popular approaches to statistical analysis (Giovagnoli, 2008). The following inquiry that requires a response is, "How can we choose the proper prior?"

In reality, although it is true that erroneous priors can occasionally lead to an inaccurate posterior distribution (Taraldsen Lindqvist, 2010), it is also true that sometimes the choice of the prior can have an effect on the posterior distribution. This is the case despite the fact that it is true that erroneous priors can sometimes lead to an inaccurate posterior distribution. This is the case despite the fact that erroneous priors can occasionally lead to an inaccurate posterior distribution. However, this does not change the fact that this is the case. This is true even though it is true that incorrect priors can sometimes cause an incorrect posterior distribution (Rossi Allenby, 2003). You may utilize methods that assess the sensitivity in order to determine whether or not a prior has an effect on the posterior, despite the fact that such methods may also result in responses that are irrelevant (Wasserman, 1996). As a result, locating reliable priors is of the utmost importance, despite the fact that doing so is not always achievable (Wasserman, 1996). For the purposes of this investigation, we might take the approach of a textbook and, for the sake of clarity and conciseness, steer clear of the enormous body of literature around the process of selecting the belief. The use of "conjugate priors," which is a strategy that can be put into practice and refers to a prior that, from a mathematical point of view, has an equal form of distribution that follows the same format as the posterior (Giovagnoli, 2008), is a practical method that may be used.

In the same vein, if a researcher plans to use a normal distribution for her pattern, she must also use a normal distribution as a prior, which can lead to a normal distribution for the results of her study (Giovagnoli, 2008). In point of fact, a beta distribution must be followed as a prior if the pattern was generated by a binomial process. On the other hand, a Gamma distribution must be followed if the pattern was generated by a Poisson process, and the latter is indicated by the fact that the pattern was generated by a Poisson process (Donovan Ruth 2019; Giovagnoli, 2008).

The frequentist method is used to make an estimation in section 2.1.1; this strategy presupposes that the unknown but fixed factors are present. In contrast to this, the Bayesian methodology takes into account random parameters. A probabilistic technique, also known as the prior distribution, is utilized to simulate the stochastic nature of model parameters.

In effort to provide a more tangible example of how Bayesian techniques work, let us investigate the following scenario: If we were to administer medication that had a high rate of success, we would be able to determine, based on the data that was provided, which is binary, whether or not the drug that was administered was successful. This would be the case if we were to administer medication that had a high success rate. Let us assume that the data follow a Bernoulli distribution, in which the probability of being successful is represented by the letter p, and we know from previous experience that this probability is fairly high. If the previous data is defined in the interval (0,1), then the Beta distribution, $Beta(\alpha p, \beta p)$, is the distribution that seems to make the most sense to utilize. If we make some tweaks to the α_p and $\beta_p)$ values, we will be able to cover a wide variety of different form types. For instance, Beta(5,1) can provide us a density that is positively negatively biased in our favor. Given the current state of affairs, selecting this course of action is recommended.

Estimating the likelihood of various outcomes is the function of the prior distribution, which is a distribution that models the researcher's belief concerning the model parameter. The interval serves as the foundation for the definition of the uniform prior, which is yet another outstanding example of a prior distribution (0,1). Within the context of this particular case, the preceding contained parameter p can take on any value within the supplied range (0,1). In spite of the fact that this is the case, the Beta(5,1) prior is thought to be informative, in contrast to the Uniform(0,1) prior, which is not thought to be useful. When it comes to the opinions that people have concerning the parameter, the non-informative prior offers very little to no information. A prior that is informative offers a more precise representation of the parameter that is being investigated, such as the Beta distribution. You can see some examples of previous decisions that people have made by looking at Figure 2.1.

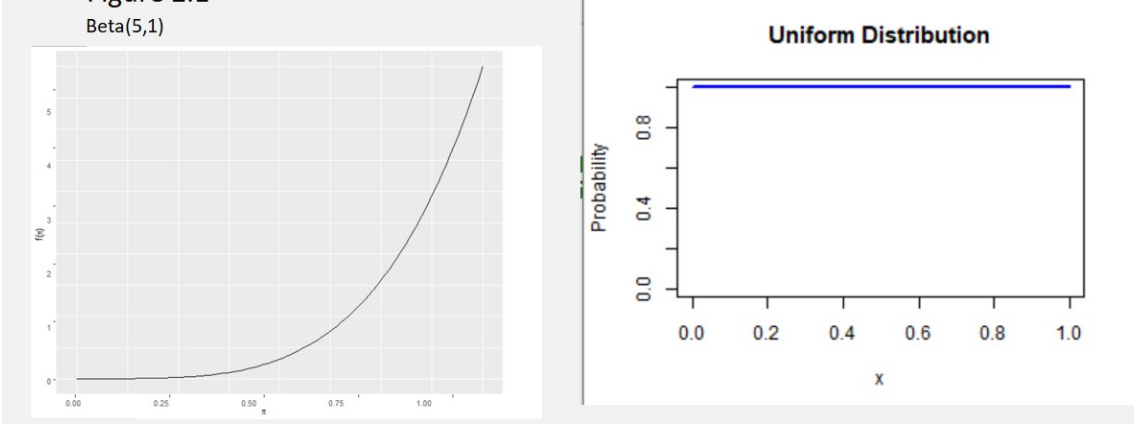


Figure 2.1

The model can be stated hierarchically as follows, using the aforementioned example as an illustration.

Likelihood: $Y_i \sim Bernoulli(n, p)\ i = 1, ..., n$

Prior: $p \sim Beta(\alpha_p, \beta_p)$

In which the hyper-parameters αp and βp can be adjusted to effectively capture the prior information, respectively. Therefore, the following holds true:
$\prod(p|data) = \frac{f(data|p)g(\alpha_p,\beta_p)}{h(data|\alpha_p,\beta_p)}$

Here, $g(p|\alpha_p, \beta_p)$ and $f(data|p)$ correspond, respectively, to the prior and the likelihood. Utilizing the information and observations $Y_1, Y_2....Y_N$, in particular, $h(data|\alpha_p, \beta_p)$ is the marginal data and can be derived as below.

$h(data|\alpha_p, \beta_p) = \int_0^1 f(data|p)g(p|\alpha_p, \beta_p)dp$

We shall refer to the marginal data as h(data) rather than h(data—alpha p,beta p) for the purpose of convenience. $h(data|\alpha_p, \beta_p)$ would read as follows: In day-to-day practice, deriving h(data) might be difficult if the model is more complex because doing so requires integrating all of the model's parameters in their totality. This can be a demanding task. As a result, it is possible to get at the posterior through the application of an approximation;

$\prod(p|data) \propto f(data|p)g(p||\alpha_p, \beta_p)$, normalizing the constant.

The probability distribution, according to Bayes's theorem, takes the following form:

$\prod(\eta|data) = \frac{f(data|\eta)g(\eta)}{h(data)}$

The posterior for η is $\prod(\eta|data)$, the likelihood is $f(data|\eta)$ and the posterior is $g(\eta)$, while prior distribution is $h(data)$. Therefore, the derived likelihood is arrived as;

$L(\eta|x_i) = \prod f(x_i|\eta)$

Within this context, marginal probability functions admirably as a normalizing constant actor. When there is continuous parametric space, the following applies:

$h(data) = \int_\eta f(data|\eta)g(\eta)d\eta$

The following are the steps that Lynch (2007) outlined in order to make generic Bayesian inferences:

1. Perform a check on the model and define the parameters of the model.
2. Indicate the sample size for the Posterior distribution.
3. Using the likelihood, revise the prior in order to obtain a revised posterior in order to produce a summary of the parameters.

We would like for our priors to be conjugates, as this will cause the posterior distribution to belong to the same family of distributions as the previous assumption.Because of the way the model is constructed, we haven't been able to find any earlier study that is applicable to our model that is useful. We shall take samples using the MCMC method in the section that comes after this one, section 2.2.1. The Markov Chain Monte Carlo algorithm gives us a wide variety of choices for dealing with the multivariate component by sampling densities over several dimensions.To provide additional detail, the Gibbs Sampler is the method that works the best for this situation. The following procedures are described by Lynch(2007) in reference to the Gibbs Sampler.

0. provide the parameter vector with the base values, which are denoted by S.

$\eta^{j=0} = \eta_1, \eta_2, ...\eta_k = S,$

v indexes will be available throughout this iteration of the process.

1.Make $v = v + 1$

2.$Sample(\eta_1^v | \eta_2^{v-1}, \eta_3^{v-1}, ...\eta_n^{v-1})$

.

.

.

n.$Sample(\eta_n^v | \eta_1^{jv}, \eta_2^v, ...\eta_{n-1}^v)$

n + 1. Return to number one

2.2.1 Bayesian WPD

The Weibull-Pareto Distribution has been decided to be used as the likelihood for the model. The likelihood is based on three parameters. η is the parameter vector that is being used (c, β, θ). In this particular situation, the gamma and exponential models are being used as priors to the analysis.

Likelihood: $X|c, \theta, \beta \sim WPD(c, \theta, \beta)$

Prior: $c|\alpha_c, \gamma_c \sim \Gamma(\alpha_c, \gamma_c)$, $\beta|\alpha_\beta, \gamma_\beta \sim \Gamma(\alpha_\beta, \gamma_\beta)$

In this scenario, it is not necessary for θ to have a prior distribution in order for the Bayesian form of the AMLE to be applied. But at the other side, we make an exception for the equation $\theta = x_1$, where x_1 refers to the number that represents the bare minimum of the data that is highlighted in Section 2.1. In this particular situation, we are working with two priors, which correspond to two parameters that the Gibbs Sampler is tasked with estimating. Both of these parameters are being estimated by the Gibbs Sampler. In addition to this, we include the modified form of the maximum likelihood estimator in our calculations (MMLE). In this particular situation, the previous value of θ has been abbreviated to $x_{(1)}$.

$$\theta|\alpha_\theta, \gamma_\theta \sim \Gamma(\alpha_\theta, \gamma_\theta) I_{0,x_{(1)}}(\theta)$$

It is possible for the Priors that are going to be selected for β, c, and $theta$ to take over any distribution that is stated in the $\Re_+$ space.Take into account the following: for instance, if we want to make a guess as to what the value of the parameter c is, we have:

$$\prod((c|x) = L(WPD)\Gamma(\alpha_c, \gamma_c)\Gamma(\alpha_\beta, \gamma_\beta)\Gamma(\alpha_\theta, \gamma_\theta)$$

$$= \frac{1}{\Gamma(\alpha_c)\gamma^{\alpha_\beta}} c^{(\alpha_c - 1)} \exp(-\frac{c}{\gamma_c}) \int_\beta \int_\theta \frac{\beta c}{x} \left(\beta \log(\frac{x}{\theta})\right)^{(c-1)} \exp\left(-(\beta \log \frac{x}{\theta})^c\right) \frac{1}{\Gamma(\alpha\beta)\gamma^{\alpha_\beta}} \beta^{(\alpha_\beta - 1)} \exp(-\frac{\beta}{\gamma_\beta}) \frac{1}{\Gamma(\alpha_\theta)\gamma^{\alpha_\theta}} \theta^{(\alpha_\theta - 1)} \exp(-\frac{\theta}{\gamma_\theta}) I_{(0,x_{(1)})}(\theta) d\beta d\theta$$

Both θ and β can have their values determined by us. As can be seen in the preceding equation, it is difficult to get accurate estimates for the parameters. As a result of this, we make use of the Gibbs Sampler, which was initially discussed in section 2.2.

In the following part 4.2, we will make use of the melanoma data, which is a prior test that evaluates how well the exponential and gamma functions. Both the informative and the non-informative were utilized in this study. The first 50,000 iterations generated by the Gibbs Sampler will be discarded, and the 10^{th} iteration will be used to select the iteration from which the estimates will be derived. The previously discussed graphs can be seen in Figures 2.2 and 2.3.

Figure 2.2

Gamma(2,2) **Gamma(1,1)**

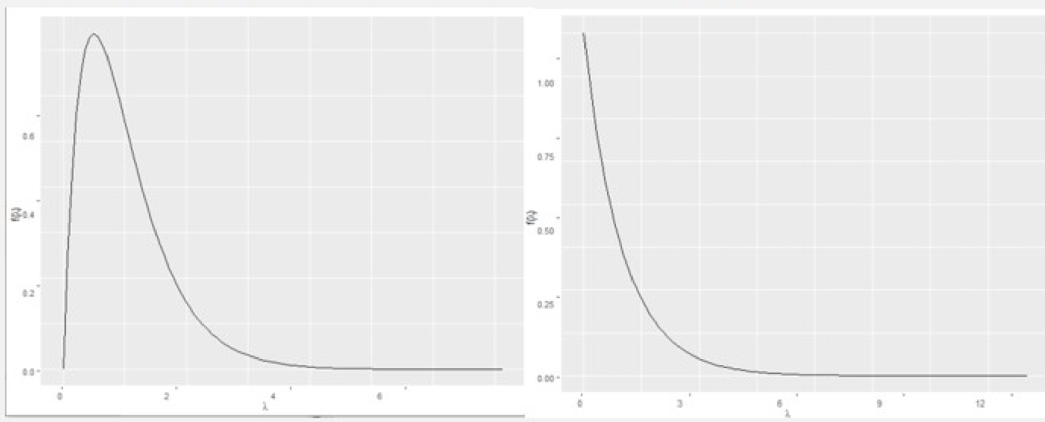
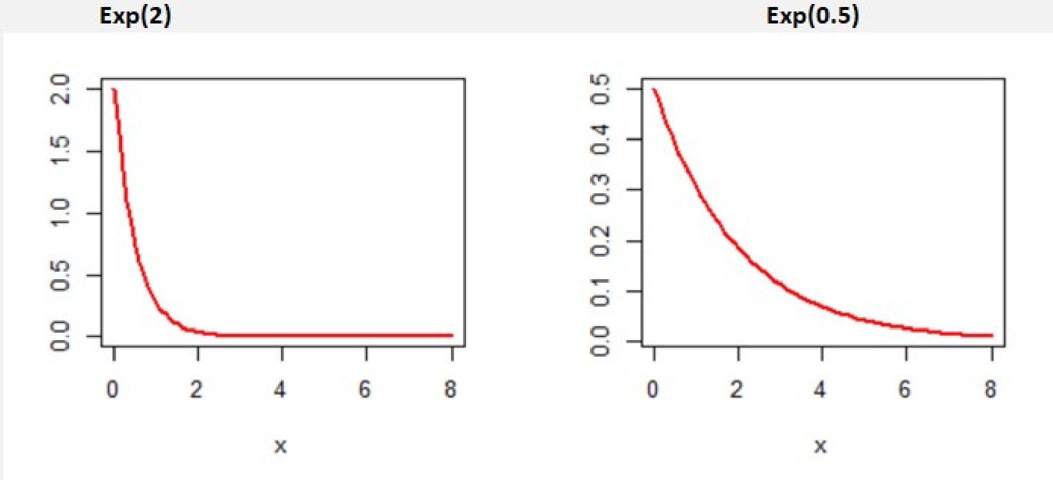


Figure 2.3

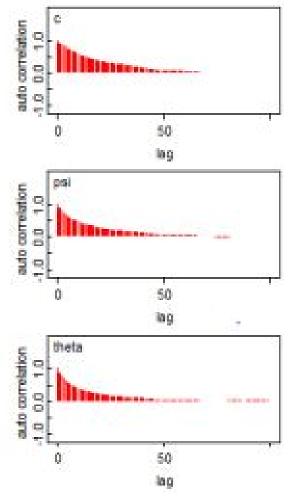
Exp(2) **Exp(0.5)**



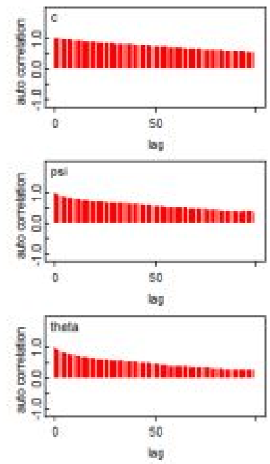
The below plots indicates the autocorrelation and the history of the prior choices are depicted in Figures 2.4 and 2.5 respectively. In order to determine if the Markov Chain is functioning properly, the history plots and the autocorrelation of the parameters can be investigated. R version 4.0.3 is what we're working with to carry out the implementation of the convergence checks. In order to facilitate a deeper comprehension of the data, the parameter values can be displayed in the form of a function of the sampler number when utilizing the history plots. In the event that the model is shown to converge, there will be oscillations of the plots to the right and left of the median of the posterior distribution. If we look back through the history of the data and find a trend, that will be a clear evidence that the data are not converging. The autocorrelation plot, on the other hand, will show you whether or not there is a correlation between the data that are subsequent to one another. If there is correlation, the samples won't be as successful in moving completely on the posterior distribution. As a result, it may be necessary to perform a significant number of iterations or to reconstruct the entire Bayesian model.

The non-informative priors of the exponential and the gamma display substantial autocorrelation, as shown in Figures 2.4 and 2.5, and the history plot depicts a trend. Therefore, it can be concluded that the non-informative priors demonstrate non-convergence. The informative priors point to relatively few connections, and the history plots are completely random in nature. This suggests that the informative priors have reached a consensus.

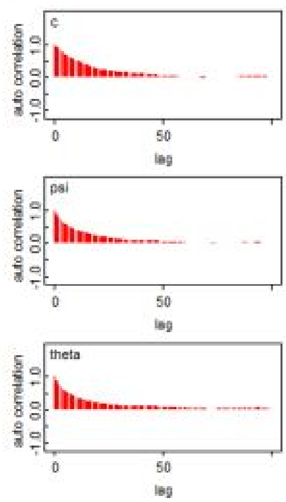
Figure 2.4: Auto-correlation Plots



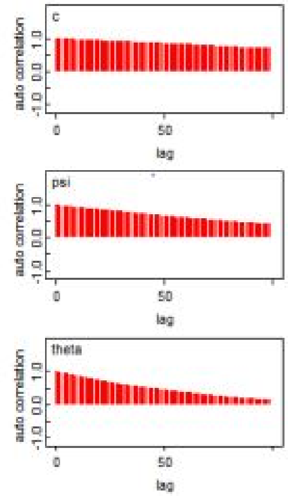
(a) Exponential Informative



(b) Exponential Non-Informative

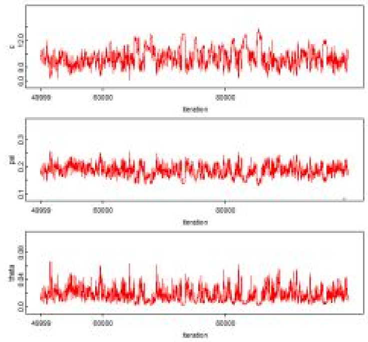


(c) Gamma Informative

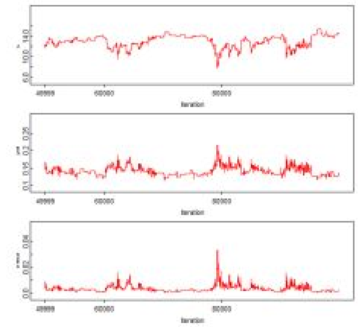


(d) Gamma Non-Informative

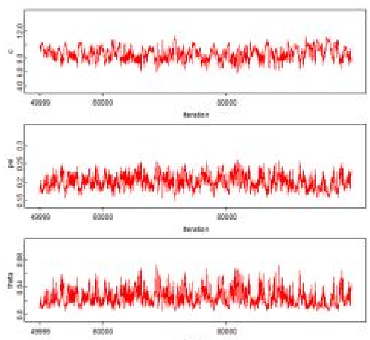
Figure 2.5: History



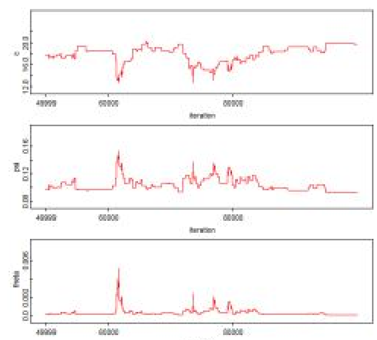
(a) Exponential Informative



(b) Exponential Non-Informative



(c) Gamma Informative



(d) Gamma Non-Informative

Chapter 3

RESEARCH METHODOLOGY

3.1 Weibull-Pareto Parameters

Let's say we're going to use the notation F(x) to represent the cumulative distribution function (CDF) of any random variable X, and let's say we're going to use the notation r(t) to represent the probability density function (PDF) of any stochastic variable T defined on the interval $[0, \infty]$. Suppose we're going to do this so that we can compare the two.
In the event that the random variable follows the parameters of the Weibull distribution with the values of c and γ,

$$r(t) = \frac{c}{\gamma}\frac{f(x)}{1-F(x)}\left(\frac{\log(1-F(x))}{\gamma}\right)^{c-1}\exp(-\left(\frac{-\log(1-F(x))}{\gamma}\right)^{c})$$

In the event that X follows a Pareto distribution together with a pdf

$$f(x) = \frac{k\theta}{x^{(k+1)}}\ x > \theta$$

The generalized family family CDF can be calculated as follows:

G(x) $= \int_0^{-\log(1-F(x))} r(t)dt.$

Therefore, the GWPD $g(x, c, \theta, \beta)$ reduces to

$$g(x) = \frac{kc}{\gamma x}\left(\frac{k}{\gamma}\log(\frac{x}{\theta})\right)^{(c-1)}\exp(-\left(\frac{k}{\gamma}\log(\frac{x}{\gamma})\right)^{c}),\ x > \theta$$

Assuming that $\beta = \frac{k}{\gamma}$, we have the pdf of the WPD, which has the parameters c, β, and θ

g(x) $= \frac{\beta c}{x}(\beta \log \frac{x}{\theta})^{c-1} \exp\left(-(\beta \log(\frac{x}{\theta}))^c\right)$, $x > \theta, c, \beta, \theta > 0$

The $WPD(c, \beta, \theta)$ has a CDF

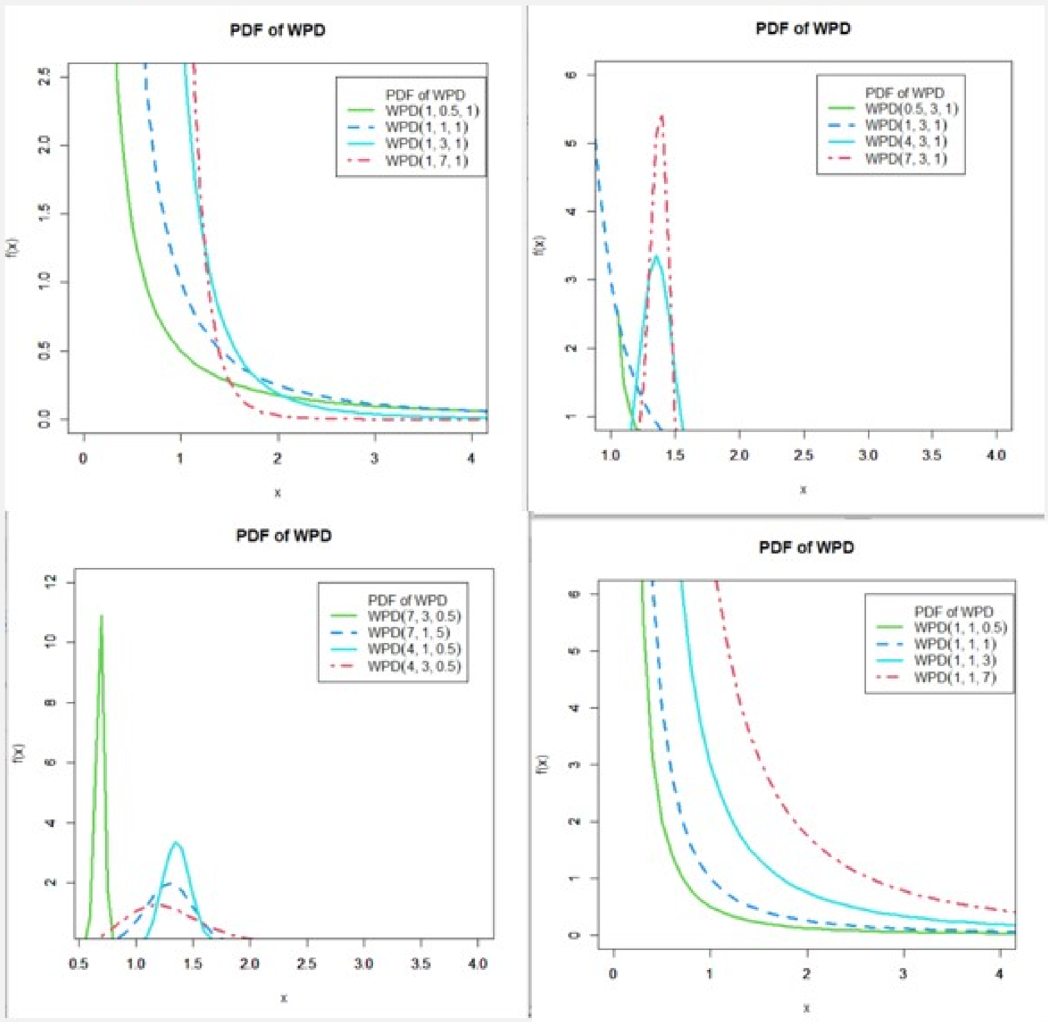
$G(x) = 1 - \exp\left(-(\beta \log(\frac{x}{\theta}))^c\right)$

As a result of the connection between the Weibull distribution and the WPD. It is possible to get at the transform by producing some random samples based on the WPD. Therefore, we can get started by producing random observations based on the Weibull distribution, and then we can use this transformation to get results based on the Weibull-Pareto distribution.

If we have a variable, let's call it Y, and that variable follows the Weibull distribution, then the parameters of that distribution are $\frac{1}{\beta}, c$. The equation for the variable X is: $X = \theta e^Y$

In light of this, we can deduce that the distribution is $WPD(\beta, c, \theta)$

Figure 3.1 PDF OF Weibull-Pareto Disribution



The pdf is given by;

$f(x,\theta,\phi) = (x-\theta)^{(c-1)}\, q\,(x-\theta;\Phi)\,$, for $x>\theta$

A situation in which the parameter vectors θ and Φ are not known. For the alternative maximum likelihood estimator $\Phi=(c,\beta)$, it is proposed that $x_{(1)}=\theta$. Additionally, the observed sample minimum would need to be removed from the data. The alternative maximum likelihood estimator function for the WPD provided by.

$L_* = \sum_{x\neq x_{(1)}} \log\, \mathsf{g}\Big(x_i; c, \beta, x_{(1)}\Big)$

$$= \sum_{x\neq x_{(1)}} \left(c\log\beta + \log c - \log x_i + (c-1)\log((\log(\frac{x_i}{x_{(1)}})) - (\beta\log(\frac{x_i}{x_{(1)}}))^c \right)$$

When we take the derivatives of c and β, we get.

$\frac{\partial L_*}{\partial c} = \sum_{x_i\neq x_{(1)}} \left(\frac{c}{\beta} - c\beta^{(c-1)}(\log(\frac{x_i}{x_{(1)}}))^c\right)$
$= \sum_{x_{(i)}\neq x_{(1)}} \left(c + \log\beta + \log(\log(\frac{x_i}{x_{(1)}})) - \log\beta(\beta\log(\frac{x_i}{x_{(1)}}))^c - (\beta\log(\frac{x_i}{x_{(1)}}))^c \log(\log(\frac{x_i}{x_{(1)}})\right)$

Bringing this equation down to zero gives us the following:

$\beta = \left(\frac{(n-n^{'})}{\sum_{x_i\neq x_{(1)}} \log(\frac{x_i}{x_{(1)}})}\right)^{\frac{1}{c}}$

where the count of the lowest possible occurrence is denoted by $n^{'}$.

$c^{-1} + \sum_{x_i\neq x_{(1)}} \log(\frac{x_i}{x_{(1)}}) - \frac{\sum_{x_i\neq x_{(1)}(\log(\frac{x_i}{x_{(1)}}))^c \log(\log(\frac{x_i}{x_{(1)}}))}}{\sum_{x_i\neq x_{(1)}(\log(\frac{x_i}{x_{(1)}}))^c}} = 0$

Take a look at the $\log$ likelihood function, for example.

$$L_n(c,\theta,\beta) = \sum_{i=1}^{n} \log \mathsf{g}\Big(x_{(i)}; c, \beta, \theta\Big)$$

In the event that this is defined for θ is less than $x_{(1)}$. We have used the conventional MLE method, and for each parameter of interest, we have taken the derivatives. It is intended to demonstrate that;

$\frac{\partial L_n(c,\theta,\beta)}{\partial\theta}$ exist and where $\theta < x_{(1)}$, it reduces to,

$$\frac{\partial L_n(c,\theta,\beta)}{\partial\theta} = \frac{1}{\theta}\sum_{i=1}^{n} \frac{1-c}{\log(\frac{x_i}{\theta})} + \frac{c\beta^c}{\theta}\sum_{i=1}^{n}(\log(\frac{x_i}{\theta}))^{(c-1)}$$

As can be seen in the following example, the solution that is found on the right-hand side of the equation is continuous on the interval that ranges from $0 < \theta < x_{(1)}$

$\frac{\partial L_n(c,\theta,\beta)}{\partial\theta}$ exist. if we set $\frac{\partial L_n(c,\theta\beta)}{\partial\theta}$ equal to zero, we have;

$$\sum_{i=1}^{n} \frac{1-c}{\log(\frac{x_{(i)}}{\theta}}) + \frac{c\beta^c}{\theta}\sum_{i=1}^{n}(\log(\frac{x_{(i)}}{\theta}))^{(c-1)} = 0,$$

$$\beta = \frac{n}{\sum_{i=1}^{n}(\log(\frac{x_i}{\theta}))^c}$$

$$c^{-1} + \sum_{i=1}^{n}\log(\frac{x_i}{\theta}) - \frac{\sum_{i=1}^{n}(\log(\frac{x_i}{\theta}))^c\log(\frac{x_i}{\theta})}{\sum_{i=1}^{n}(\log(\frac{x_i}{\theta}))^c} = 0$$

The employment of numerical methods is required in order to find solutions to the aforementioned three equations. In this case, there is a possibility of gradual convergence, despite the fact that the starting assumption is that the procedure will not converge. As a direct consequence of this, parameter estimation is challenging. To yet, a concept for a strategy using a hierarchical Bayesian model has proven to be excellent.

3.2 The Simulation Technique

In the following part of the article, we will present two distinct simulation situations for you to view. The Bayesian WPD model is going to be displayed in the first stage in terms of the bias and mean square errors (MSE) of the model parameters. This is going to be the primary focus of this step. This is going to be the main point of attention. The outcomes of this research were reevaluated with a variety of other parameter configurations. The generation of one hundred informational collections, each of which contains one hundred observations, and one hundred informational indexes, each of which contains five hundred observations, is performed for each parameter decision. The accompanying recreation study was carried out in order to examine how well this model can continue to function even after some of the observed values were omitted from the data. During the course of this inquiry, one of our primary objectives is to locate the true parameter value for a number of different censoring percentiles by employing the 95 percent credible interval. Tables 3.3 and 3.4 contain outcomes that are reliant on a total of 500 pieces of pseudoinformation, each of which has 100 observations. These results are provided as a dependency on the tables. Because the Weibull-Pareto distribution model is so complicated, we limit the number of simulations that we run to either 100 or 500 unique data sets that have been developed. This decision was made because of the difficulty of the model. No matter what the circumstances are, at a later stage we are going to talk about a very big number of the different iterations that could occur. We came to the conclusion that the exponential prior would work best for both β and c. In addition, as was mentioned in Section 2.2.1, we considered adopting a prior that was exponentially truncated on θ. Our prior choices are,

$\theta \sim \exp(0.05)T(0, x_{(1)})$, $\beta \sim \exp(0.05)$, $c \sim \exp(0.05)$ where $x_{(1)}$ is the sample minimum.

3.2.1 The Generation of Data

In this part of the presentation, we will demonstrate the algorithm that can be used to extract pseudo information from the WP circulation. We discovered that the T-X family is quite useful since it ties the Weibull-Pareto distribution to the Weibull distribution. This was very helpful to us. In order to generate data, we follow the steps that are listed below.

1. Generate n observations using the Weibull distribution by adjusting $\frac{1}{\beta}$, θ and c.
2. In order to generate data of size from the WPD, you can use simulation and Step 1.
3. Repeat steps 1 and 2 from the previous section using significant iterations.
4. Repeat steps 1 through 3 while selecting various values for the parameters each time to generate new sets of data.

3.3 Performance of Bayesian WPD mode:Simulation Study 1

The goal of the research is to determine how accurately the Weibull-Pareto model can estimate the model's myriad parameters, and the study's purpose is to do so. We set n equal to 100 in Tables 3.1 and 3.2; however, in Table 3.2, we set n equal to 500. We conduct experiments with a variety of sample sizes to investigate the impact that increasing the number of observations has on the mean squared error (MSE) as well as the bias. The purpose of this study is to evaluate how well the Weibull-Pareto distribution works when applied to data that has a positive skewness to it. When c is given extremely small values, the Weibull-Pareto distribution displays a positive skew as a consequence of this. Figure 2.1 presented this information earlier on.

Table 3.1: Simulation Results for n=100

True value			Bias			MSE		
c	β	θ	$\hat{c}$	$\hat{\beta}$	$\hat{\theta}$	$\hat{c}$	$\hat{\beta}$	$\hat{\theta}$
0.5	0.5	0.5	0.002129811	0.073102786	-0.000002164	0.001761814	0.01895894	2.002612e-07
0.5	0.5	1	0.0099150100	0.0333320080	0.0001145717	0.001593829	0.01199942	2.988987e-06
0.5	0.5	3	0.0091700225	0.0772137927	-0.0002322325	0.002082500	0.01922018	5.380738e-06
1	0.5	0.5	0.006265250	0.013903652	-0.001628796	0.0071278183	0.0027844764	0.0001129739
1	0.5	1	0.015711468	0.010521510	-0.003227082	0.0074881016	0.0030497129	0.0004371173
1	0.5	3	0.0126477950	-0.0001631812	-0.0029130300	0.008015383	0.003231458	0.004849308
1	1	1	0.0181986190	0.0136460478	-0.0008130798	0.0083963957	0.0137898502	0.0001659547
1	1	3	0.01291250	0.01263585	-0.00881712	0.006798370	0.014594735	0.001290123
1	3	1	0.02482298	0.01923111	-0.00134432	0.009365873	0.1196966	1.080989e-05
1	3	3	0.009365939	0.062604910	-0.002626022	0.0077594085	0.1097467409	0.0001132329
1	7	1	0.0183454520	0.0439909200	-0.0006295968	0.008634790	0.6293276	2.560957e-06
4	0.5	0.5	0.15924688	0.02263368	0.03699240	0.786479570	0.008670353	0.026776585
4	0.5	1	0.05652267	0.02837874	0.09484234	0.612100354	0.007922381	0.103516462
4	0.5	3	0.14240777	0.02181913	0.21100991	0.736452697	0.007194806	0.804046570
4	1	0.5	0.4241059525	0.0006509685	-0.0163274190	1.292133717	0.033847635	0.007633421
4	3	1	0.86961504	-0.12338393	-0.04923811	2.702958484	0.403874856	0.009796301
4	3	3	1.1247682	-0.1955078	-0.1776196	4.3029700	0.4265535	0.1068641
4	7	1	1.24921152	-0.53722243	-0.03582712	5.096688635	2.582148315	0.003720049
7	0.5	0.5	-0.1472447	0.1036607	0.1803744	3.84196003	0.03695746	0.10242634
7	0.5	1	-0.30082735	0.09647315	0.33992266	2.77187225	0.02778827	0.33596013
7	0.5	3	-0.3593476	0.1062062	1.1067052	2.68143224	0.03265477	3.49608159
7	1	0.5	0.67829400	0.10219010	0.01170808	5.46684825	0.08569965	0.01383927
7	1	1	0.615419020	0.088010525	0.009819396	5.00692051	0.08421026	0.05392994
7	1	3	0.3153701	0.1493958	0.1614523	4.6895699	0.1284626	0.6313439
7	3	0.5	2.12247436	-0.04913447	-0.03192561	12.255698323	0.673126395	0.003397868
7	3	1	1.22950478	0.30888657	-0.03063851	9.42094087	1.22883087	0.01234657
7	3	3	1.7346162	0.1015389	-0.1577227	12.9595889	0.9568738	0.1449145

The bias can be calculated by subtracting the mean parameter value from the actual value of the parameter, which is denoted by the notation $c - \hat{c}$.. When calculating the Mean Square error, the variation on addition of the squared bias is used as a starting point. To put it another way, that would be $Var(\hat{c}) - bias^2 c$. In this particular situation, $c \leq 1$. The model's mean squared error (MSE) and bias are remarkably low across the board for all of the parameters that were selected. This demonstrates that the model is accurate even when applied to data with a positive skew. When c is greater than one, there is an increase in both the MSE and the bias. When c is greater than one and β is greater than one, the MSE and bias for β both go up.

Table 3.2: Simulation for results n=500

True value			Bias			MSE		
c	β	θ	$\hat{c}$	$\hat{\beta}$	$\hat{\theta}$	$\hat{c}$	$\hat{\beta}$	$\hat{\theta}$
0.5	0.5	0.5	5.596911e-03	1.384806e-02	-7.328075e-05	1.339933e-03	1.417710e-02	8.451127e-08
0.5	0.5	1	7.531968e-03	7.531968e-03	-4.508175e-05	1.409118e-03	1.573573e-02	8.861059e-07
0.5	0.5	3	0.008797122	0.051014747	-0.000357250	1.929890e-03	1.826296e-02	4.885888e-06
1	0.5	0.5	-0.0012497210	-0.0007515835	0.0002451115	1.223116e-03	6.237983e-04	4.646336e-06
1	0.5	1	0.0028293625	0.0030992490	0.0001162338	1.313297e-03	6.088439e-04	2.161319e-05
1	0.5	3	-0.0017509243	-0.0017509243	-0.0004026875	0.0008745643	0.0006008416	0.0001518166
1	1	1	0.0010709212	-0.0037865287	-0.0000217805	1.377850e-03	2.654785e-03	4.211163e-06
1	1	3	0.006507584	-0.0015396895	-0.0000382000	1.201734e-03	2.535635e-03	4.326752e-05
1	3	3	-0.0030374823	0.0095009975	-0.0001363425	1.187428e-03	1.834478e-02	3.489241e-06
1	3	1	8.154675e-03	-9.874600e-04	1.338901e-03	1.338901e-03	2.154246e-02	1.080989e-05
1	7	1	4.355002e-03	-5.608451e-02	-6.797575e-05	1.184372e-03	1.133858e-01	1.023097e-07
4	0.5	0.5	0.038499715	0.006342355	0.010955393	0.167494824	0.002181525	0.007539611
4	0.5	1	0.077152303	0.077152303	-0.001557034	0.212408305	0.002319323	0.030607729
4	0.5	3	0.030607729	0.007371553	0.071236146	0.170972268	0.001918046	0.243701035
4	1	0.5	0.129761378	-0.012261333	-0.009609649	0.162493926	0.006097347	0.001409241
4	3	1	0.21013405	-0.06566586	-0.01246237	0.33505571	0.09393229	0.00136805
4	3	3	0.15553516	-0.04756925	-0.03447667	0.228879979	0.069509991	0.008803241
4	7	1	0.11262364	-0.07919613	-0.00400735	0.178763771	0.340733447	0.000148442
7	0.5	0.5	0.03657421	0.02086180	0.03820247	1.278143427	0.005969986	0.005969986
7	0.5	1	0.245844248	0.009642056	0.036645576	0.036645576	0.00523846	0.07440450
7	0.5	3	0.11888068	0.01589822	0.17984952	1.612725718	0.006773238	0.849629403
7	1	0.5	0.120854155	0.050013526	0.009339225	1.784286979	0.034474075	0.007168955
7	3	0.5	0.77232087	-0.08332107	-0.01415032	2.807819288	0.207978787	0.001045332
7	1	1	0.67104516	-0.01030971	-0.04057308	3.74455055	0.03429284	0.03429284
7	1	3	0.17691185	0.04321077	0.03092214	2.44122483	0.03696885	0.28205981
7	3	1	0.80810035	-0.04121598	-0.02746845	4.656942717	0.319094556	0.006507584
7	3	3	0.9278334	-0.1150892	-0.1022471	3.7579442	0.2773772	0.0523737

The value of the parameter θ was found to be relatively accurate. When β and c were at their largest value, the maximum was also the point at which the mean squared error and bias were at their highest levels. When there are more observations, both the MSE and the bias decrease, which can be shown by comparing Table 3.2 to Table 3.1. In the event that we carry out a comparison of the last three rows of Tables 3.1 and 3.2 ((WPD(7,3,1), WPD(7,1,3), and WPD(7,3,3)), we will see that both the MSE and the bias reduce considerably as was to be anticipated.

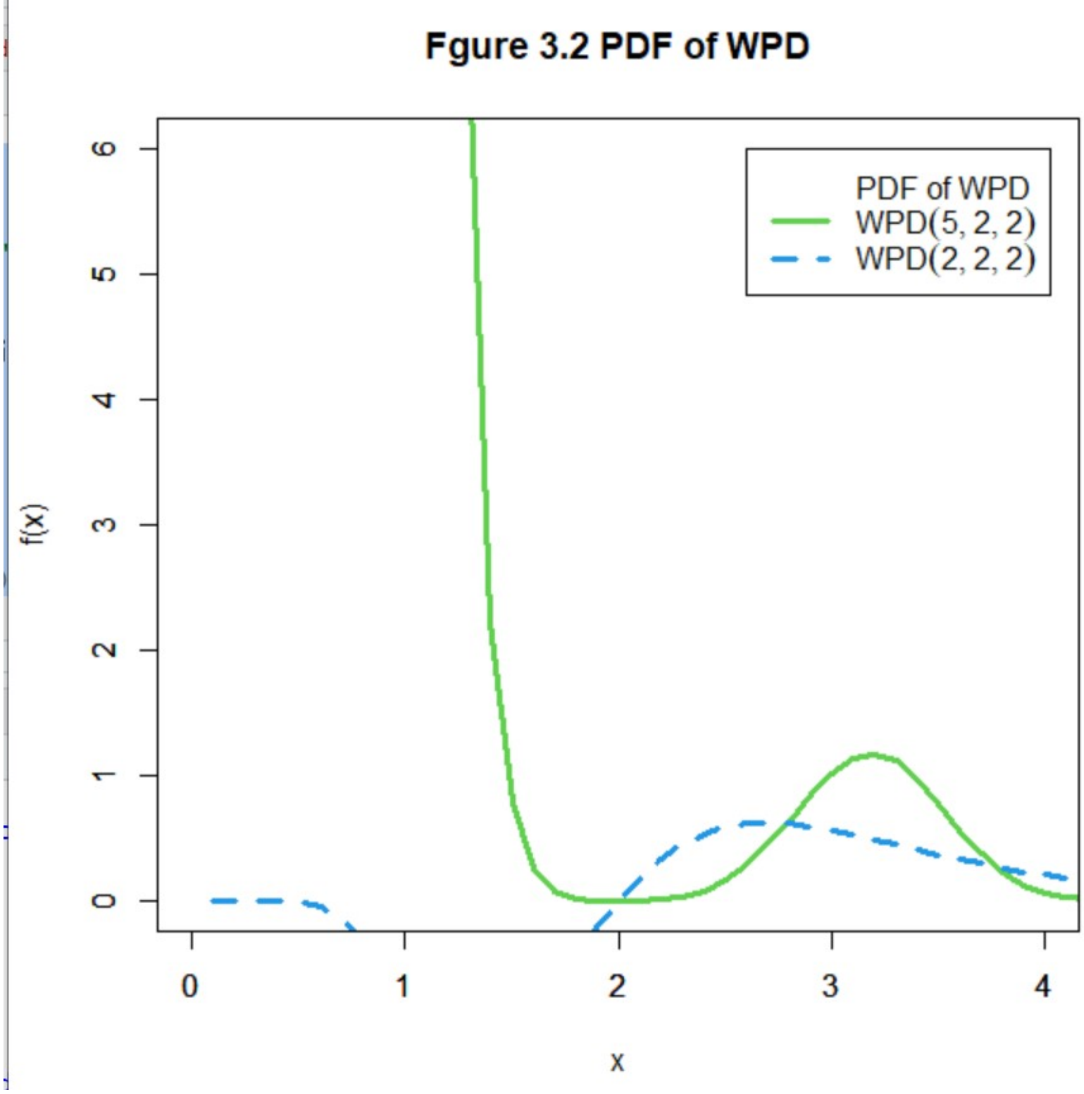
3.4 Bayesian WPD under Censoring: Simulation Study 2

In accordance with the survival model below

$$S(x, t) = [g(x)]^t * [1 - G(x)]^{(t-1)}$$

The equation t = (0,1) represents the censoring variable in the Weibull-Pareto distribution. The probability distribution function g(x) and the cumulative distribution function G(x) are both based on x. The formula for the survival function is written down as $S(x) = 1 - G(x)$. The value is acquired from the survival function if the data that was observed was censored; if this was not the case, then the value was taken from the probability distribution function (pdf), which is designated by the letter g(x). On the simulation, right-censoring was used, and 100 observations were created in accordance with the procedures outlined in Section 3.2. Censorship is performed in the following manner: first, an order is established for the observations; next, beginning with the j^{th} observation and continuing with the censoring of everything else; finally, the results are expressed as a percentage: (100-j). In accordance with what is presented in Figure 3.1, we investigated both the symmetrical and asymmetrical aspects of the model.

The mean square error and the bias can be determined using Tables 3.3 and 3.4, following the steps outlined in section 3.3. The censoring percentage is denoted by the letter $\%C$." The confidence interval for the percentage of times that the true parameter was recorded by the 95 percent credible interval, divided by the total number of iterations, which in this case was 500. It is clear from looking at Table 3.3 that both the mean squared error (MSE) and the bias for $\hat{c}$are on the higher end of the scale. In the case of β, the mean squared error as well as the bias are on the lower end of the spectrum for the 20 percent and 10 percent censoring, respectively.



Fgure 3.2 PDF of WPD

In spite of the fact that this is the case, there was an increase in self-censorship at both the 35 and the 50 percentage point levels. There is not much of a difference between the bias and the MSE, which is an encouraging indicator for the parameter θ. When the censoring percentage was increased, both the mean square error (MSE) and the bias increased across the board for all of the parameters. In addition, the performance of the credible interval dropped, particularly in terms of its ability to capture a proportion of the actual values.

As we can notice on following Table 3.4, $\hat{c}$ has a high MSE in addition to its bias, while $\hat{\beta}$ has both a large MSE and its bias, regardless of the censoring option used. The MSE and bias of $\hat{\theta}$ are both extremely high.

Table 3.3: Right Censoring for WPD(2,2,2)

	Bias(MSE)			%True Parameter Captured by CI		
%C	$\hat{c}$	$\hat{\beta}$	$\hat{\theta}$	$\hat{c}$	$\hat{\beta}$	$\hat{\theta}$
10	-0.26939507	0.01421504	0.02632264	0.744	0.946	0.87
	(0.11441356)	(0.03343525)	(0.00457954)			
20	-0.49162637	-0.07831612	0.04394210	0.0302	0.946	0.804
	(0.269685739)	(0.028299786)	(0.004787532)			
35	-0.75643641	-0.34672244	0.06169448	0.026	0.502	0.634
	(0.59125417)	(0.13150875)	(0.00665626)			
50	-0.95687543	-0.75339821	0.06558217	0	0.006	0.534
	(0.927656797)	(0.574439426)	0.007724839			

Table 3.4: Right Censoring for WPD(5,2,2)

	Bias(MSE)			%True Parameter Captured by CI		
%C	$\hat{c}$	$\hat{\beta}$	$\hat{\theta}$	$\hat{c}$	$\hat{\beta}$	$\hat{\theta}$
10	-1.3463562	0.4909197	0.1709286	0.662	0.802	0.778
	2.43496554	0.45633039	0.05693664			
20	-1.9829797	0.5566727	0.2347656	0.302	0.696	0.622
	4.29508220	0.48576893	0.07609605			
35	-2.6791479	0.4747181	0.3001660	0.04	0.578	0.38
	7.4101037	0.3472820	0.1083566			
50	-3.13417537	0.09953852	0.31507083	0.008	0.95	0.3
	9.99747159	0.06539004	0.11810655			

Table 3.5: WPD(5,2,2) Parameter Estimate for 35% and 50%

	Mean			Sd		
%C	$\hat{c}$	$\hat{\beta}$	$\hat{\theta}$	$\hat{c}$	$\hat{\beta}$	$\hat{\theta}$
35	2.320852	2.474718	2.300166	0.4540593	0.2550084	0.0914429
50	1.865825	2.099539	2.315071	0.37754154	0.23091182	0.07767302

Chapter 4

APPLICATION AND DISCUSSIONS

4.1 Tribolium Confusum and Tribolium Casteneum

In this section of the study, we apply the Weibull-Pareto to the data from two distinct investigations on the adult number of Tribolium confusum and investigate the outcomes of this analysis. The data have been evaluated in order to have a better knowledge of how well the Weibull-Pareto distributions mount for the various forms of the data that have been applied. This has been done in order to have a better understanding of how well the distributions mount. The data in Table 4.1 are pretty symmetrical, whereas the data in Table 4.2 have a very long tail that is skewed to the left. This indicates that Table 4.2's data are more accurate.

Park (1954) and Park et al. were the ones who kept track of the two separate data points, which were 29 degrees and 24 degrees census (1964). In this portion of the article, the WPD, the generalized Weibull distribution, the Weibull distribution, and the exponentiated-Weibull distribution are all contrasted with one another. The values that were actually observed for the two data points are shown in both Table 4.1 and Table 4.2. Both tables also contain the predicted frequency for the two data points. Figures 4.1 and 4.2 each provide a visual representation of the data that is reported in Tables 4.1 and 4.2, respectively. Data similar to that which is shown in Figure 4.2 may be found in Table 4.2.

The model takes into account the outlying portions of the data that are derived from the primary data. If one examines Table 4.1, they will see that the projected frequency for the Weibull model produces a value of 11.3131. This can be seen by looking at the table.

The Weibull-Pareto distribution has an expected frequency of 5.12015 and is most closely associated with the range of values 30-40 for the x value. This distribution was named after the two statisticians who first developed it. The exponentiated-Weibull(EW) model and the generalized Weibull(GW) model both arrived at the same correct result when calculating the initial value. Even though it would appear that the named models function as anticipated, the Weibull model does not do very well when it comes to capturing the tail characteristics of the data for the three most recent observations. When compared to the other three models, the values dropped at a rate that was significantly faster than average. Both EW and WPD had the same level of performance, which enabled them to successfully catch the tail of the data. Although the drop for the Generalized Weibull occurs more quickly than the drops for the EW and WPD, the drop for the Generalized Weibull occurs more slowly than the drop for the Weibull.

The Weibull distribution produced inaccurate estimates of the x-values in Table 4.2. It would appear that the WPD and the EW are able to produce an approximation of the data's tail values. It would appear that the GW does not take into account the tail values. It can be seen that the GW, EW, and WPD all capture the same thing in the centre for the x-values.

The histograms that were generated combined with the overlay curves for each table provide a clear pictorial picture of how each table fared in comparison to the others. Figure 4.1 presents a breakdown into their individual figures of the numerous approaches that can be utilized in order to make an estimation of the data. As can be seen on the top right, the Weibull tail declines at a rate that is significantly faster than that of any other model. Figure 4.2 makes it quite evident that the curves are superimposed on top of one another. When compared side-by-side, WPD, Generalized Weibull, and Exponentiated Weibull are difficult to differentiate from one another.

Figure 4.3 demonstrates, once more, that the right tail of the Weibull distribution falls off very quickly and does not provide a precise approximation of the data. When compared to WPD and EW, GW offers more accurate predictions of the middle x-values. As can be seen in Figure 4.4, the models all perform similarly well in terms of capturing the tails of the data.

Figure 4.1: Histogram and Curves

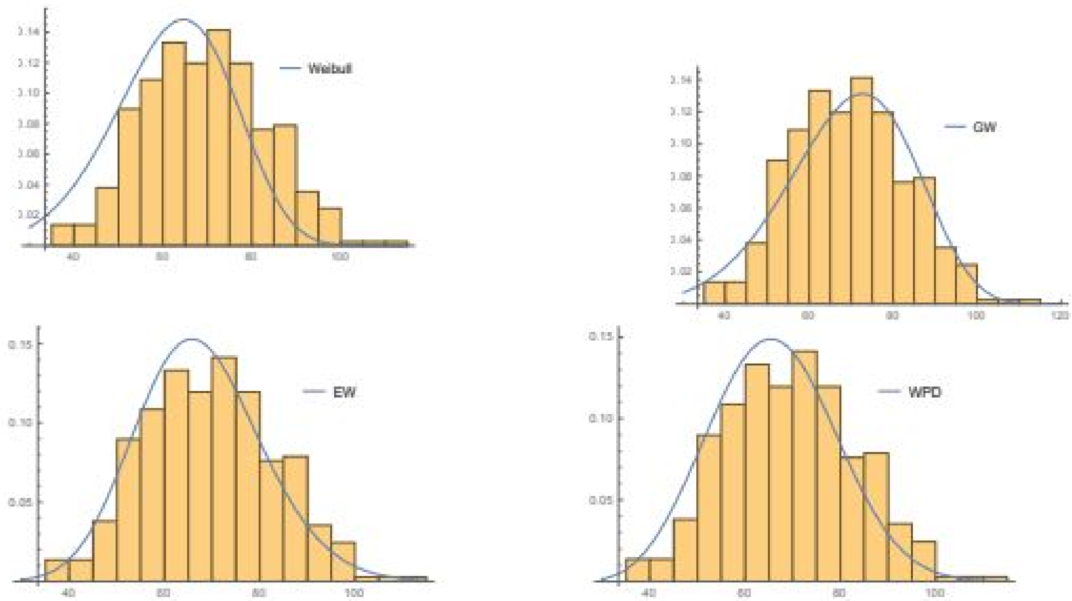


Table 4.1: Observed and expected frequencies of Confusum 29°

x-value	Observed	Weibull	Generalized Weibull	Exponentiated Weibull	Weibull-Pareto
35-40	5	11.3121	6.58432	3.12327	5.12015
40-45	5	18.5436	10.8356	8.54702	12.2578
45-50	14	27.8444	16.5835	18.3373	22.734
50-55	33	38.2767	23.726	31.7221	35.0143
55-60	40	47.8314	31.7438	45.2372	46.2852
60-65	49	53.6348	39.5699	54.1866	53.3184
65-70	44	52.9734	45.6122	55.4484	53.8479
70-75	52	44.9804	48.0683	49.2368	47.7321
75-80	44	31.8667	45.5942	38.4942	37.0826
80-85	28	18.1731	38.1322	26.8443	25.1793
85-90	29	7.99975	27.3789	16.8824	14.8929
90-95	13	2.5899	16.3023	9.65709	7.64642
95-100	9	0.583621	7.69315	5.054	3.39628
100-105	1	0.0860152	2.70823	2.42802	1.30088
105-110	1	0.00772955	0.654471	1.07209	0.42846
110-115	1	0.000391392	0.0964364	0.434938	0.121039
Total	368	356.704	361.284	366.706	366.358
Parameter Estimates		v=5.316	$\alpha = .1928$	$\alpha = 2.795$	$c = 6.694$
		$\lambda = 1.93 * 10^{-10}$	$\lambda = 0.395$	$\sigma = 52.74$	$\beta = .798$
			$\phi = 141$	$\theta = 4.502$	$\theta = 20.06$

Table 4.2: Observed and expected frequencies of Confusum 24°

x	Observed	Weibull	Generalized Weibull	Exponentiated Weibull	Weibull-Pareto
20-30	0	4.08727	3.81726	0.0706614	0
30-40	0	9.95255	9.05459	0.871671	0
40-50	3	19.2467	17.1896	4.93156	6.16691
50-60	9	32.1385	28.3509	16.6664	21.8221
60-70	39	48.2017	42.2456	38.9595	45.9356
70-80	53	66.232	58.0376	69.0166	73.1826
80-90	77	84.1816	74.2854	98.5644	97.3012
90-100	105	99.3362	89.0104	118.655	113.191
100-110	135	108.814	99.9526	124.529	118.255
110-120	114	110.347	105.026	116.965	112.746
120-130	113	103.112	102.891	100.38	99.159
130-140	92	88.2405	93.4732	80.0228	81.0757
140-150	59	68.6524	78.187	60.0404	61.9983
150-160	54	48.1565	59.6824	42.8357	44.5553
160-170	38	30.1769	41.1235	29.2925	30.2142
170-180	22	16.7261	25.2413	19.3165	19.4014
180-190	17	8.11328	13.5805	12.3391	11.8332
190-200	6	3.40556	6.27936	7.66069	6.87419
200-210	10	1.22242	2.4348	4.63371	3.81306
210-220	3	0.370604	0.767591	2.7354	2.02419
220-230	2	0.0936773	0.189087	1.57792	1.03055
230-240	0	0.0194784	0.0345429	0.890258	0.50416
240-250	1	0.00328556	0.00436008	0.491594	0.237433
250-260	0	0.000443128	0.000344343	0.265816	0.107824
260-270	0	0.0000470792	0.000014711	0.140805	0.0472914
Total	952	931.583	950.858	951.852	952.476
Parameter Estimates		v=3.711	$\alpha = 0.2757$	$\alpha = 1.75$	c=5.324
		$\lambda = 1.827 * 10^{-8}$	$\lambda = 0.03658$	$\sigma = 70.2$	$\beta = 0.6845$
			$\phi = 318$	$\theta = 5.935$	$\theta = 28.12$

Figure 4.2: Overlay of Graphs

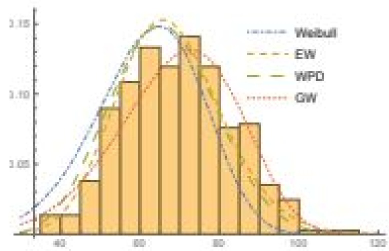


Figure 4.3: Histogram and Curves

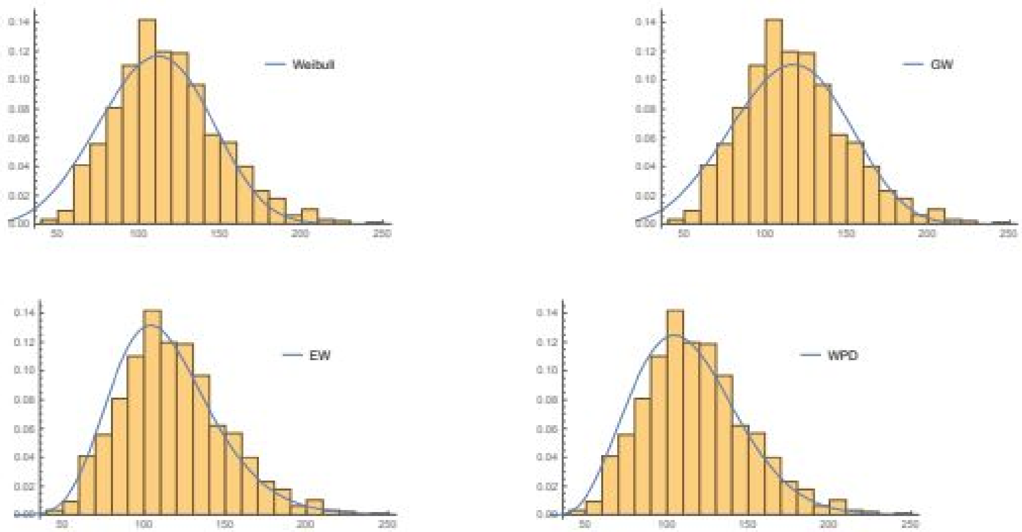
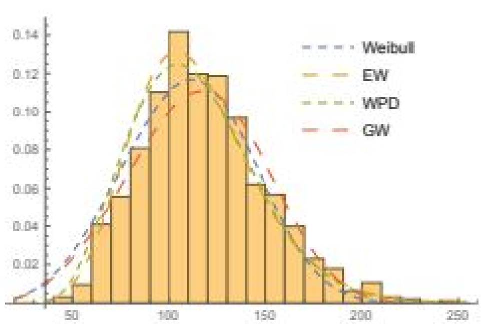


Figure 4.4: Overlay of Graphs



4.2 Censoring with Melanoma data

This component of the report analyzes the data as part of a study on cutaneous melanoma in order to evaluate a specific medication that was given after surgery. The research was carried out during the years 1991 and 1995. After then, in the year 1998, there was a subsequent investigation. Park selected a random sample of 417 patients to study; the event of interest was the amount of time that passed before the patient passed away, also known as the censoring time. The patient's age is also taken into account. The Weibull-Pareto model is utilized in both the construction of the survival model as well as its subsequent utilization in the application of the appropriate censoring data. The survival function can be summarized as follows:

$$S(x) = 1 - F(x) = exp(-(\beta log(\frac{x}{\theta}))^c)$$

In this case, t denotes the checking time indicator, and with the help of the data presented in Section 3.4, we are able to formulate the likelihood as follows:

$$Likelihood = [f(x)]^t * [S(x)]^{(1-t)}$$

We would like to model one parameter of the Weibull-Pareto utilizing the age covariate that is at our disposal. This model will serve as an example. The first option that comes to mind is the employment of the *log(scale)*, which involves the application of a linear function; to put it another way,

$\frac{1}{\beta} = e^{\psi_{age} * x_{age}}$ (ψ has been employed as regression coefficient to communicate efficiently).

Because θ and c would really be difficult to read, the choice to represent the cavariate was made to employ β instead. In spite of the fact that this is the case, the prior on θ within the Bayesian model is capped at $x_{(1)}$. When θ is modeled through the covariate, it is not necessarily a straightforward exercise to ensure that this restriction is met. If option c were picked instead, the model's price would be higher than it is currently.

The previous specification is the phase that comes after this one, and it covers the parameters ψage, θ, andc. Flat gamma priors are used for both θ and c as their respective priors. We make use of normal priors for the ψ_{age} calculation.

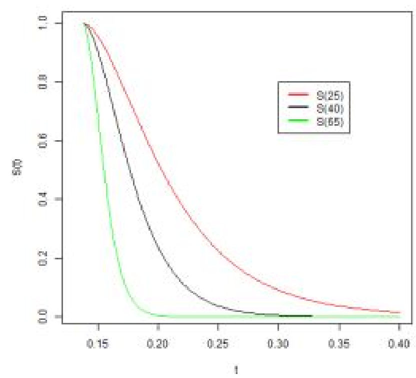
Priors:

$c \sim \Gamma(0, 0.1)$

$\theta \sim \Gamma(1, 0.1) I_{(0, x_{(1)})}(\theta)$

$\psi_{age} \sim N(0, 0.1)$

Figure 4.5: Graph of Survival Function



- $\hat{c} = 1.778$
- $\hat{\psi}_{age} = 0.0299$
- $\hat{\theta} = 0.1381$

In the very last stage of the procedure, which is the estimation of the survival function, the role that the covariate performs is the primary focus of our attention. The fact that the positive gauge of the regression parameters demonstrates that the rate parameter of the Weibull-Pareto distribution, (β), decreases as age increases demonstrates that the tail of the distribution will drop at a much more moderate rate, which will result in a level bend in the distribution. This will lead to a noticeable acceleration in the deterioration of the survival function as a consequence. Taking a look at Figure 4.5, which illustrates the survival curves for our population at the ages of 25, 40, and 65, is a good way to illustrate this point. It should come as no surprise that a person who is 25 years old has a significantly higher probability of surviving the challenges of life than an individual who is 65 years old does. For instance, the survival curve for 65 implies that there is no possibility of survival when t equals 0.25, whereas the survival curve for 25 indicates that there is approximately a 20 percent probability of survival. Both curves are shown below.

Chapter 5

CONCLUSION

Through the use of simulations, we investigated the effectiveness of Bayes estimates for the parameters of the Weibull-Pareto Distribution. The estimates generated by Bayes were evaluated next to the actual values. The results of the simulation showed that the model worked well for values of c that were lower than one, which is proof that the distribution has a positive positive skew. When both c and β are greater than1, as in the previous illustration, there is a significant amount of bias present in the parameters. The simulation that included the requisite right censoring demonstrated how correctly the model works for censored data when applied to both symmetrically and skewedly formed Weibull-Pareto distributions.

As the degree of censoring increased, so did the level of bias, and so did the number of times our credible interval successfully projected the value for WPD (2,2,2). Due to the fact that β was improperly calculated at a censoring rate of 35 percentage points for WPD(5,2,2), we discovered that the credible interval was unable to take into account β. As a consequence of this, we were unable to determine the true value of β. When the Tribolium Confusum was used in the research, it was discovered that the WPD Bayesian estimates were comparable to the EW and GW estimates. We note, as part of the process of carrying out a survival analysis, how the covariate age reveals that the likelihood of survival decreases with increasing age. This discovery is in line with what we would anticipate seeing to be the case. R, version 4.0.3, was used to do the analysis on the data.

The continuation of this research effort will require the examination of additional loss functions, such as the Linex loss, amongst others. It is necessary to conduct additional research into WPD using a number of different censoring methods, such as type I, type II, and progressive. The application of the approach to data sets that are both more relevant and have a longer tail is being considered. The efficiency of utilizing Bayesian WPD with a restricted amount of available data.

Bibliography

1. Alzaatreh, A., Lee, C., Famoye, F. (2013). Weibull-Pareto Distribution and It's Applications. Communications in Statistics-Theory and Methods, 42: 1673-1691.

2. Alzaatreh, A., Lee, C., Famoye, F. (2013). A new method for generating families of continuous distributions. Metron: International Journal of Statistics. 71:6379

3. Famoye, F., Lee, C., Olumolade, O. (2005).The beta-weibull distribution.J.Statist. Theor. Appl., 4(2):121136. Hosmer,D., Lemeshow, S., May, S., Applied Survival Analysis,2nd ed., John Wiley Sons, Inc., (New Jersey)(2008).

4. Khan, H., M., R., Saxena, A., Vera, V., Abdool-Ghany, F. Gabbidon, K., Perea, N., Stewart, T. S., Ramamoorthy, V. (2014). Black Hispanic and Black NonHispanic Breast Cancer Survival Data Analysis with Half-normal Model Application. Asian PAcific Journal of Cancer Prevention 15 9453-9458.

5. Lynch, Scott. Introduction to Applied Statistics and Estimation for Social Scientists, Springer, (New York) (2007).

6. Martin,J., Perez,C.J (2009),Bayesian analysis of a generalized lognormal distribution,Computational Statistics and Data Analysis, 53:1377-1387.

7. Mudholkar, G. S., Kollia, G. D. (1994). Generalized Weibull family: A structural analysis. Commun. Statist. Theor. Meth., 23(4):11491171.

8. Mudholkar, G. S., Srivastava, D. K., Freimer, M. (1995). The exponentiated Weibull family: a reanalysis of the bus-motor-failure data. Technometrics 37(4):436445.

9. Noor,F. and Aslam M. (2013), Bayesian inference of the inverse Weibull mixture distribution using type-I censoring,Journal of Applied Statistics, Vol.40, No.5,1076-1089.

10. Park, T. (1954). Experimental studies of interspecies competition II. Temperature, humidity, and competition in two species of Tribolium. Physiological Zoology 27:177238.

11. Park,T.,Leslie, P.H., Mertz., D. B. (1964). Genetic strains and competition in population of Tribolium.Physiological Zoology 37:97-162.

12. Rodrigues, J., Concho, V.G, Castro, M., and Balakrishnan,N. (2010). A Bayesian destructive weighted Poisson cure rate model and an application to a cutaneous melanoma data. Statistical Methods in Medical Research 21(6) 585-597.

13. Singh, S.K, Singh, U., and Kumar, D.(2013), Bayesian estimation of parameters of inverse Weibull distribution,Journal of Applied Statistics, Vol.40, No.7, 1597-1607.

14. Smith, L.R. (1985). Maximum likelihood in a class of nonregular cases.Biometrika, 72 (1) : 67-90