SCORE STATISTICS FOR THE GENERALIZED GAMMA DISTRIBUTION

A Thesis

Submitted to the Faculty of Graduate Studies and Research

In Partial Fulfillment of the Requirements

for the Degree of

Master of Science

In

Statistics

University of Regina

By

Olusesan Kolawole Ogunsanya

Regina, Saskatchewan

September 2018

© Copyright 2018: Olusesan Kolawole Ogunsanya
Olusesan Kolawole Ogunsanya, candidate for the degree of Master of Science in Statistics, has presented a thesis titled, *Score Statistics for the Generalized Gamma Distribution*, in an oral examination held on August 28, 2018. The following committee members have found the thesis acceptable in form and content, and that the candidate demonstrated satisfactory knowledge of the subject material.

External Examiner:         Dr. Abdul Bais, Electronic Systems Engineering

Supervisor:                Dr. Andrei Volodin, Department of Mathematics & Statistics

Committee Member:          Dr. DianLiang Deng, Department of Mathematics & Statistics

Chair of Defense:          Dr. Gale Russell, Faculty of Education
Abstract

Checking the lifetime distribution of a production for the purpose of locating and removing the defects in future and examining the reliability of a product is very important. The main goal of my research is to construct statistical inferential procedures for discriminating between three leading reliability models in survival analysis. These models are model of fatigue due to aging, model of a weak particle, and memoryless model of a defect at birth. We will be constructing some tests for pairwise distinctions between above-mentioned models and assess the results by Monte-Carlo methods of statistical simulations.
Acknowledgements

My first appreciation goes to God Almighty for giving me the knowledge and privilege to undertake this program and finish it satisfactorily. This achievement wouldn’t have been possible without him.

I would like to extend my outermost gratitude and appreciation to Dr. Andrei Volodin for his guidance, generous help, expert advice, encouragement and endless patience in making this thesis a success.

I would also like to express my enormous thanks to my beautiful wife Mrs. Patience E. Ogunsanya the best woman in the world and my gorgeous daughter Miss. Gabriella O. Ogunsanya for coming into my world during this research study. I also appreciate the financial supports of my brothers Mr. Akindele Ogunsanya and Mr. Oluyemi Ogunsanya during this program.

I am appreciative to my colleague Mr. Haruna Ismail for his continuous assistance and encouragement during this thesis.
The following notations are used throughout the thesis:

GG(\(\theta, \lambda, \mu\)) - Generalized Gamma distribution with parameters \(\theta, \lambda, \mu\);

G(\(\lambda\)) - Gamma distribution with parameter \(\lambda\);

W(\(\tau\)) - Weibull distribution with parameter \(\tau\);

GH(\(\mu\)) - Generalized Half - Gaussian distribution with parameter \(\mu\);

\(P\) - Probability measure;

\(F\) - Distribution function

\(f\) - Density of a continuous random variable

\(\bar{X}\) - The sample mean of the sample \(X\)

\(X\) - Stochastic variable;

\(X\) - Sample of observed random variables \(\{X_1, X_2, \ldots, X_n\}\)
# Contents

Abstract

Acknowledgements

Notations

Table of Contents

List of Figures

1 Introduction

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Importance of the Research</td>
<td>8</td>
</tr>
<tr>
<td>1.2 Lifetime Distribution</td>
<td>9</td>
</tr>
<tr>
<td>1.2.1 Mutually Independent Model</td>
<td>9</td>
</tr>
<tr>
<td>1.2.2 Memoryless Model</td>
<td>10</td>
</tr>
<tr>
<td>1.2.3 Aging and Deterioration Model</td>
<td>10</td>
</tr>
<tr>
<td>1.3 Research objective</td>
<td>10</td>
</tr>
</tbody>
</table>
1.4 Test Statistic .......................................................... 10

1.5 Testing statistical hypothesis ........................................ 11
  1.5.1 General framework ............................................... 11
  1.5.2 The likelihood ratio test. ........................................ 15
  1.5.3 Invariant Test .................................................... 18
  1.5.4 Locally Most Powerful Test ....................................... 18
  1.5.5 Historical remarks on Hypothesis Testing. ................... 20

1.6 Distributions we consider ............................................ 21
  1.6.1 Gamma Distribution .............................................. 21
  1.6.2 Weibull Distribution ............................................. 22
  1.6.3 Generalized Gaussian Distribution ............................. 23
  1.6.4 Generalized Gamma Distribution ................................. 23

1.7 Powerful Test Statistic to Discriminate Distributions Connected with
  Generalized Gamma Distribution ....................................... 25

2 Literature Review ......................................................... 27

3 Score Statistics ............................................................ 30
  3.1 Invariant statistical inference: The notion of maximal scale invariant
      statistic .......................................................... 30
  3.2 Most powerful invariant tests ...................................... 34
3.3 Score statistics for the Generalized Gamma distribution .......................... 42

4 Simulation for power of the tests based on Score statistics ........................ 49

4.1 Graphical illustrations of powers of Score statistics ................................. 51

5 Concluding remarks and Future Research ................................................. 66

References ........................................................................................................ 67
List of Figures

4.1 Power of statistic $T_1 + T_2$ for Gamma and Weibull distributions, sample size $n = 10$ .................................................. 52

4.2 Power of statistic $T$ for Gamma and Weibull distributions, sample size $n = 10$ .................................................. 53

4.3 Power of statistic $T_1 + T_2$ for Gamma and Weibull distributions, sample size $n = 50$ .................................................. 53

4.4 Power of statistic $T_1$ for Gamma and Weibull distributions, sample size $n = 50$ .................................................. 54

4.5 Power of statistic $T_1 + T_2$ for Gamma and Weibull distributions, sample size $n = 100$ .................................................. 54

4.6 Power of statistic $T_1$ for Gamma and Weibull distributions, sample size $n = 100$ .................................................. 55

4.7 Power of statistic $T_1 + T_2$ for Gamma and Weibull distributions, sample size $n = 200$ .................................................. 55
4.8 Power of statistic $T_1$ for Gamma and Weibull distributions, sample size $n = 200$  

4.9 Plot of the values of statistic $T_1 + T_2$ for Gamma and Weibull distributions, sample size $n = 10$, shape parameter $0$  

4.10 Plot of the values of statistic $T_1 + T_2$ for Gamma and Weibull distributions, sample size $n = 10$, shape parameter $1$  

4.11 Plot of the values of statistic $T_1 + T_2$ for Gamma and Weibull distributions, sample size $n = 10$, shape parameter $100$  

4.12 Plot of the values of statistic $T_1$ for Gamma and Weibull distributions, sample size $n = 10$, shape parameter $0$  

4.13 Plot of the values of statistic $T_1$ for Gamma and Weibull distributions, sample size $n = 10$, shape parameter $1$  

4.14 Plot of the values of statistic $T_1$ for Gamma and Weibull distributions, sample size $n = 10$, shape parameter $100$  

4.15 Plot of the values of statistic $T_1 + T_2$ for Gamma and Weibull distributions, sample size $n = 50$, shape parameter $0$  

4.16 Plot of the values of statistic $T_1 + T_2$ for Gamma and Weibull distributions, sample size $n = 50$, shape parameter $1$  

4.17 Plot of the values of statistic $T_1 + T_2$ for Gamma and Weibull distributions, sample size $n = 50$, shape parameter $100$
4.18  Plot of the values of statistic $T_1$ for Gamma and Weibull distributions, sample size $n = 50$, shape parameter 0 .......................... 60

4.19  Plot of the values of statistic $T_1$ for Gamma and Weibull distributions, sample size $n = 50$, shape parameter 1 .......................... 61

4.20  Plot of the values of statistic $T_1$ for Gamma and Weibull distributions, sample size $n = 50$, shape parameter 100 .......................... 61

4.21  Plot of the values of statistic $T_1 + T_2$ for Gamma and Weibull distributions, sample size $n = 100$, shape parameter 0 .......................... 62

4.22  Plot of the values of statistic $T_1 + T_2$ for Gamma and Weibull distributions, sample size $n = 100$, shape parameter 1 .......................... 62

4.23  Plot of the values of statistic $T_1 + T_2$ for Gamma and Weibull distributions, sample size $n = 100$, shape parameter 100 .......................... 62

4.24  Plot of the values of statistic $T_1$ for Gamma and Weibull distributions, sample size $n = 100$, shape parameter 0 .......................... 63

4.25  Plot of the values of statistic $T_1$ for Gamma and Weibull distributions, sample size $n = 100$, shape parameter 1 .......................... 63

4.26  Plot of the values of statistic $T_1$ for Gamma and Weibull distributions, sample size $n = 100$, shape parameter 100 .......................... 63

4.27  Plot of the values of statistic $T_1 + T_2$ for Gamma and Weibull distributions, sample size $n = 200$, shape parameter 0 .......................... 64
4.28 Plot of the values of statistic $\overline{T}_1 + \overline{T}_2$ for Gamma and Weibull distributions, sample size $n = 200$, shape parameter 1 . . . . . . . . . . . . 64

4.29 Plot of the values of statistic $\overline{T}_1 + \overline{T}_2$ for Gamma and Weibull distributions, sample size $n = 200$, shape parameter 100 . . . . . . . . . . . . 64

4.30 Plot of the values of statistic $\overline{T}_1$ for Gamma and Weibull distributions, sample size $n = 200$, shape parameter 0 . . . . . . . . . . . . . . . . . 65

4.31 Plot of the values of statistic $\overline{T}_1$ for Gamma and Weibull distributions, sample size $n = 200$, shape parameter 1 . . . . . . . . . . . . . . . . . 65

4.32 Plot of the values of statistic $\overline{T}_1$ for Gamma and Weibull distributions, sample size $n = 200$, shape parameter 100 . . . . . . . . . . . . . . . . . 65
Chapter 1

Introduction

1.1 Importance of the Research

Pressure is on manufacturers worldwide to produce products of high quality. Pressure due to extreme global competition among manufacturers, advancement in technology, improving and meeting customer expectations, development of highly sophisticated products, etc. Therefore, irrespective of what you produce it is very important to pay more attention on the reliability of the product. In few cases failure of technology (aircraft, hospital equipment, etc.) can be dangerous and life-threatening. In other cases, failure of some product (automobiles, kitchen appliances, mobile or home phones, etc.) can be inconvenient but not dangerous. Although failures are unavoidable, hence maximizing the reliability of our products and minimizing the effects of failures must be the tasks.

Reliability of a product is synonymous to a scale that measured the quality of a
product. Reliability is simply quality over time and it’s mostly use by manufacturers to predict product reliability in design stage so as to prevent failure of a product and also to determine warranty service of a product. Reliability data helps in checking manufacturers advertising claim and can also be used to compare products from different manufacturers.

1.2 Lifetime Distribution

We often assumed that the lifetime data of a product is coming from a particular lifetime distribution and that inference and statistical analysis of such lifetime data are done based on this lifetime distribution. We have different types of lifetime distributions since every product gives different information about its lifetime. Hence, we must be cautious when choosing a suitable lifetime distribution for a given lifetime data.

1.2.1 Mutually Independent Model

When the failure of any part of a system is considered as the total failure of the whole system, then the lifetime distribution of that system will follow a Weibull distribution.
1.2.2 Memoryless Model

Memoryless property means that the future and the past are independent. When we replace any element of a system with a new one after it breaks down, then the lifetime distribution of that system follow Generalized Gaussian distribution.

1.2.3 Aging and Deterioration Model

When the reliability of a given system is time dependent, then the lifetime distribution of the system follow a Gamma distribution.

1.3 Research objective

The aim of the thesis is to construct a test statistic for distinguishing between lifetime distributions that are connected with Generalized Gamma distribution proposed by Igor Volodin (1981) [15]. And the test is based on the separation of the score statistic by using Monte-Carlo method for our computer simulation. We derived the mean values, and also the variances and of course the covariances of the test statistic.

1.4 Test Statistic

The three main reliability models (Weibull, Gamma, and Generalized Gaussian) are connected to Generalized Gamma distribution proposed by Igor Volodin (1981)
[15]. Volodin (1981) [15] derived locally most powerful test statistic in this regard. In this thesis we considered a test statistic of the form

\[
T = \frac{\frac{1}{n} \sum_{i=1}^{n} \ln X_i - \ln \left( \frac{1}{n} \sum_{i=1}^{n} X_i \right)}{\ln \left( \frac{1}{n} \sum_{i=1}^{n} X_i \right) - \frac{1}{n} \sum_{i=1}^{n} \ln X_i},
\]

where \( X_i \sim iid.GGD(\theta, \lambda, \mu) \) \( i = 1, 2, \ldots, n \).

### 1.5 Testing statistical hypothesis

#### 1.5.1 General framework

The following material is taken from lecture notes on Statistical inference of my supervisor Dr. Volodin with his permission. This part describes the general framework of the notable procedure of statistical testing and is present in nearly all theses of Dr. Volodin’s students.

"Random samples of size \( n \) are represented by independent random variables \( X = \{X_1, \ldots, X_n\} \) having a common probability mass function or density function \( f(x|\theta) \), \( \theta \in \Theta \). Given competing hypotheses about possible values of \( \theta \), how do we decide between them?

The formal procedure, due to J. Neyman and E.S. Pearson (1920’s) and extending earlier ideas of R.A. Fisher, is to identify a

**Null hypothesis** \( H_0 : \theta \in \Theta_0 \),
which is contrasted with an

Alternative hypothesis \( H_A : \theta \in \Theta_1 \),

where \( \Theta_0 \) and \( \Theta_1 \) are disjoint subsets of the parameter space \( \Theta \). Call a hypothesis simple if it has the form \( \theta = \theta' \), a known constant, and composite otherwise.

Deciding between the null and alternative hypotheses involves a test statistic \( T = h(X) \) taking values in a space which is partitioned into disjoint subsets \( A \) and \( R \), called acceptance and rejection regions, and corresponding to \( \Theta_0 \) and \( \Theta_1 \), respectively. If an observed value of \( T \), \( t = h(X) \in R \), then \( H_0 \) is rejected in favour of \( H_A \), and if \( t \in A \) then \( H_0 \) is accepted. The latter term is usually taken to mean there is too little data evidence to opt decisively for \( H_0 \).

The law of \( T \) depends on the unknown value of \( \theta \). A crucial role is played by the law of \( T \) given that \( H_0 \) is true. This is well-defined only if \( H_0 \) is simple. If general, \( A \) and \( R \) are chosen so that if \( H_0 \) is true, the event \( \{ T \in R \} \) occurs with a small probability. Specifically, a small number \( \alpha \) is chosen by the statistician, and then \( R \) such that

\[
P(T \in R|\theta) \leq \alpha \text{ for all } \theta \in \Theta_0,
\]

trying to get as close to \( \alpha \) as possible.

If \( t \in R \) then we say that \( H_0 \) is rejected at the \( 100\alpha\% \) level of significance. Call \( \alpha \) the size of the test. With these choices, we expect that \( P(T \in R|\theta) > \alpha \text{ if } \theta \in \Theta_1 \), that is, the probability of rejection exceeds the chosen level of significance if \( H_0 \) is
false. In fact, this property cannot be inferred from the above test structure. A test which has this property is said to be *unbiased*.

We have the following rationale applicable to unbiased tests for making accept/reject decisions: If $t \in \mathcal{R}$, then:

(a) *Either* $H_0$ *is true and an event of small* ($\leq \alpha$) *probability has occurred: or*

(b) $H_0$ *is false, and an event has been observed whose probability exceeds* $\alpha$.

Option (b) is the better explanation of the observed outcome; it is consistent with the intuition supporting the maximum likelihood concept. This procedure gives rise to two possible errors:

*Type I error:* Reject $H_0$ when it is true, and

*Type II error:* Accept $H_0$ when it is false.

Type I error is held to be more serious, explaining why the test is designed to control its probability of occurrence:

$$P(\text{Type I error}) = P(T \in \mathcal{R}|H_0) \leq \alpha.$$ 

Computing the probability of a Type II error usually is possible only if $H_A$ is simple. In general, we define the power function $\beta_T(\theta) = P(T \in \mathcal{R}|\theta)$, all $\theta \in \Theta$. Thus the test is designed so that $\beta_T(\theta) \leq \alpha$ if $\theta \in \Theta_0$. Typically the power function is close to $\alpha$ if $\theta \in \Theta_1$ but close to its boundary, and increasing as $\theta$ moves away from the boundary. The sensitivity of a test can be judged in terms of how quickly $\beta_T(\theta)$ increases above $\alpha$ as $\theta \in T_1$ moves away from the boundary.
Remarks. 1. This Neyman-Pearson testing procedure is a frequentist concept: The operation meaning of the assertion ‘$H_0$ is rejected at the $100\alpha\%$ level if significance’ means that if this random experiment is independently replicated many times using the same population, then a Type I error occurs in a proportion $\leq \alpha$ of such replications.

2. In ‘scientific’ contexts $H_0$ represents accepted wisdom or a status quo, and experimental data has the express purpose of refuting rather than confirming $H_0$. Refutation should be compelling, beyond a reasonable doubt, thus explaining the special status accorded to Type I errors, and why $\alpha$ is chosen to be small. It follows that $H_0$ and $H_A$ are not inter-changeable.

On the other hand, $H_0$ could represent model assumptions, such as ‘errors are normally distributed’. In quality control situations $H_0$ could be ‘the process is in control’, i.e. the probability $p$ that a manufactured item is faulty is less than some very small number. For these cases, finding $t \in A$ gives weight to accepting $H_0$ as a viable working assumption, a desirable outcome.

3. An important question is how to choose a test statistic? Often the choice is made on a ‘common sense’ basis. But there are general results which can give guidance, so-called Likelihood Ratio tests that we are using in this thesis. Obviously,
we want a test statistic that is unbiased, and to have the property that $\beta_T(\theta)$ is as large as possible for all $\theta \in \Theta_1$, i.e. maximum power under $H_A$.

Say that a test statistic is *uniformly most powerful* if, for any other test statistic $T'$, we have $\beta_T(\theta) \geq \beta_{T'}(\theta)$ for all $\theta \in \Theta_1$. If $H_0$ and $H_A$ are simple hypotheses, then a fundamental result named the Neyman-Pearson lemma implies the existence of a uniformly most powerful test. If $H_0$ is simple and $H_A$ is composite, then $H_A$ can be considered as a union of simple hypotheses, and it may be that a uniformly most powerful test can be forged from the uniformly most powerful tests for each pair $(H_0, H_A(\theta'))$, where the second component is, for each $\theta' \in \Theta_1$, the assertion that $\theta = \theta'$.

If $H_0$ also is composite, it’s not obvious how to proceed. Instead there is another route based on the ‘common sense’ approach which is applicable to most problems arising in practice, so-called *likelihood ratio tests*.”

1.5.2 The likelihood ratio test.

How do we test between

$$H_0 : \theta \in \Theta_0 \quad \& \quad H_A : \theta \in \Theta_1,$$

where $\Theta_0 \cap \Theta_1 = \emptyset$? Often $\Theta_0 \cup \Theta_1$ is a proper subset of $\Theta$, but we will redefine $\Theta$ to equal this union. This redefinition is the first step in constructing a likelihood ratio test (LRT).
Let $L(\theta|X)$ denote the likelihood function, with $\theta$ in our cut-down parameter space $\Theta$, and let $\hat{\theta}$ be a Maximum Likelihood estimate computed using $\Theta$, and assumed to be the unique *global* maximizer. It can be that $\theta$ is a boundary point of $\Theta$. Next, let $\hat{\theta}_0$ be the Maximum Likelihood estimate restricted by $H_0$, i.e.,

$$
\hat{\theta}_0 = \text{Argmax}_{\theta \in \Theta_0} L(\theta|X).
$$

**Definition 1.** The likelihood ratio is

$$
\lambda_r = \lambda_r(X) = \frac{L(\hat{\theta}_0|X)}{L(\hat{\theta}|X)},
$$

and $L_r = \lambda_r(X)$ is the corresponding random variable, the LR test statistic.

**Definition 2.** For a given $\alpha \in (0, 1)$, let $\lambda^*$ satisfy

$$
P(\Lambda_r \leq \lambda^*|H_0) \leq \alpha,
$$

with equality if possible. The likelihood ratio test rejects $H_0$ at the $100\alpha\%$ level if $\lambda_r \leq \lambda^*$.

**Rationale.** It should be clear that $\max_{\theta \in \Theta_0} L(\theta|X) \leq \max_{\theta \in \Theta} L(\theta|X)$, which implies that $0 < \lambda_r(X) \leq 1$. In particular, $\lambda_r(X) = 1$ if $\hat{\theta} \in \Theta_0$. Now argue as follows.

(i) If $H_0$ is true, then we expect with large probability that $\hat{\theta}_0 \approx \hat{\theta}$ and hence that $L(\hat{\theta}_0|X) \approx L(\hat{\theta}|X)$. This says that the data are explained almost as well by some
\( \theta \in \Theta_0 \) as by the best unrestricted choice from \( \Theta \), i.e. by \( \hat{\theta} \). So we expect with large probability that \( \Lambda_r \approx 1 \).

(ii) If \( H_0 \) is false, then the data are poorly explained by any \( \theta \in \Theta_0 \), and that with large probability the Maximum Likelihood estimator \( \hat{\theta} \) will take values in \( \Theta_1 \). So we expect to see values of \( L(\hat{\theta}_0|X) \) rather smaller than \( L(\hat{\theta}|X) \).

The implication of both of these is that data giving \( \lambda_r(X) \approx 1 \) provide strong evidential support for \( H_0 \), and conversely, data giving small values of \( \lambda_r(X) \) give strong evidence against \( H_0 \).

**Remarks.**

1. In most cases we use a monotonic transformation of \( \Lambda_r \) to give a test statistic whose law under \( H_0 \) is a familiar one.

2. Most standard tests can be derived as LRT's. In particular, there is a collection of tests for means and variances of \( N(\mu, \sigma^2) \) populations. These comprise one-sample tests: for \( \mu \) with \( \sigma^2 \) known (Z-test), for \( \mu \) with unknown \( \sigma^2 \) (t-test), and for \( \sigma \) with unknown \( \mu \) (chi-square test); and two-sample tests: a t-test for comparing two normal means, assuming each population has the same unknown variance \( \sigma^2 \), and the F-test for comparing two normal variances with unknown means. In this thesis we also consider LRT tests.

3. If \( H_0 \) and \( H_A \) are simple then the famous Neyman-Pearson lemma states that the LRT is the uniformly most powerful test with size \( \alpha \) among all tests of size \( \alpha \). If \( H_0 \) is simple and \( H_A \) is one-sided then a technical condition ensures that the LRT
is uniformly most powerful. This condition is satisfied in many important cases, and certainly in those where the minimal sufficient statistic has the same dimension as \( \Theta \). If \( H_A \) is two-sided then the LRT is \textit{not} uniformly most powerful; in fact no such test exists. The LRT still has desirable properties.

1.5.3 Invariant Test

Invariant property is when a test statistic remained unchanged even after series of class transformation.

According to Lehmann and Romano (2005), ‘If \( G \) is the group of transformations

\[
gx = (cx_1, cx_2, ..., cx_n), \quad c \neq 0,
\]

a special role is played by any zero coordinates. However, in statistical applications the set of points for which none of the coordinates is zero typically has probability 1.

1.5.4 Locally Most Powerful Test

According to Omelka (2005) [9], consider a random sample \( X = (X_1, X_2, ..., X_n) \) from a distribution with a density function \( f(x, \theta) \) which depends on the unknown parameter \( \theta \). Let \( H_0 \) be a null hypothesis about this parameter and \( \Phi \) be a test function defined on the sample space which give the probability of rejecting \( H_0 \) when the sample \( X = x \) is observed. Denote \( \beta_{\Phi}(\theta) = E_{\theta}\Phi(X) \) the power function of this test.
Definition 1.1 Let $d$ be a measure of the distance of an alternative $\theta \in H_1$ from a given hypothesis $H_0$. A level $\alpha$ test $\Phi_0$ is said to be locally most powerful (LMP) if, given any other level $\alpha$ test $\Phi$, there exists $\Delta > 0$ such that $\beta_{\Phi_0}(\theta) \geq \beta_{\Phi}(\theta)$ for all $\theta \in H_1$ with $0 < d(\theta) < \Delta$.

We shall restrict ourselves to the real $\theta$ and the null hypothesis $H_0 : \theta = \theta_0$; then it is natural to take $d(\theta) = \theta - \theta_0$ as a measure of the distance for one-sided alternatives $H_1 : \theta > \theta_0$ and $d(\theta) = |\theta - \theta_0|$ for two-sided alternatives.

In typical cases the Locally Most Powerful Test can be found as a test maximizing the first derivative of the power function at the point of the null hypothesis $\theta_0$. Computing the derivative of the power function $\beta_{\Phi}(\theta) = \int \Phi(x_1, x_2, ..., x_n) \prod_{i=1}^{n} f(x_i, \theta) dx_1 ... dx_n$ of an arbitrary test $\Phi$, we are often allowed to differentiate under the integral sign. Let $f'(x, \theta)$ denote the derivative of $f(x, \theta)$ with respect to $\theta$. Then

$$\frac{\partial \beta_{\Phi}(\theta)}{\partial \theta} = \int \Phi(x_1, ..., x_n) l'(x, \theta) \prod_{i=1}^{n} f(x_i, \theta) dx_1, ..., dx_n,$$

where $l'(x, \theta) = \sum_{i=1}^{n} \frac{f'(x_i, \theta)}{f(x_i, \theta)}$ is the well-known Fisher score function (calculated as the logarithmic derivative of the likelihood $L(x, \theta) = \prod_{i=1}^{n} f(x_i, \theta)$). From the Neyman-Pearson lemma we get that the Locally Most Powerful test has the critical region $r(x, \theta_0) \geq C_\alpha$ where $C_\alpha$ is appropriately chosen constant to reach the prescribed level $\alpha$.

Remark: Notice that if the Locally Most Powerful test is not simultaneously the Uniformly Most Powerful, then typically (with an exception of the finite sample space)
there does not exist a universal neighborhood over which the Locally Most Powerful test maximizes the power uniformly. To see it, it suffices to compare the power of Locally Most Powerful test with the power of the Neyman-Pearson test for an arbitrarily close simple alternative $\theta_1$.

1.5.5 Historical remarks on Hypothesis Testing.

Karl Pearson (father of Egon Pearson, the Pearson in Neyman-Pearson theories of testing hypotheses) was educated in England and Germany as an applied mathematician (and much else), and in 1884 (aged 27) he was appointed to a chair in that discipline. Under the influence of Francis Galton (Charles Darwin’s cousin) and Frank Weldon, Pearson’s interests radically changed in 1892. Weldon asked Pearson to analyze some biological data, causing him to rapidly learn what was known about mathematical statistics, then called ‘the theory of errors’. He quickly began to develop new methodology, and found a research school in biometrics and eugenics. Francis Galton and Frank Weldon are two founders of the so-called biometric school of evolution, which believed that genetic traits vary continuously in populations. By contrast, saltationists believed in discontinuous changes, both before the rediscovery of Mendel’s work. After this rediscovery, saltationists became known as Mendelians. Their differences developed into a heated dispute, but both are correct, in a sense. Genetic inheritance is discrete, but traits controlled by large gene complexes can
appear to vary continuously in large populations.

Among other contributions, Pearson invented the method of moment estimation (around 1895) together with a methodology for using sample moments to make a plausible choice of population law. This latter led to the Pearson system of frequency curves, i.e. families of density functions which solve the differential equation \( f'/f = (x - k)/(ax^2 + bx + c) \), where \( a, b, c \) and \( k \) are constants. His system contains all the most commonly used continuous laws, and others not so common. We also mention the chi-square in connection with goodness-of-fit tests, and tests for association using contingency tables. Both were introduced by Pearson, starting near 1900 and further developed thereafter.

1.6 Distributions we consider

1.6.1 Gamma Distribution

The Gamma distribution is a continuous probability distribution which has two parameters \( \lambda \) and \( \theta \), where \( \lambda \) is the shape parameter and \( \theta \) is the scale parameter. The scale parameter stretch out and compress the range of the distribution. Gamma distribution represents a family of shapes. When the shape parameter is one, the distribution reduces to one parameter exponential distribution with parameter \( \lambda \). When the shape parameter is less than one, the distribution reduces to exponentially shaped and asymptotic to both vertical and horizontal axis. And the distribution
assumed a mounded shape but skewed when the shape parameter is more than one. The skewness reduces as the value of the shape parameter increases. In this study we will focus our attention on the Gamma distribution used for waiting times. The form of the probability density function of the Gamma distribution $G(\lambda, \theta)$ considered here is:

$$f(x; \lambda, \theta) = \frac{1}{\Gamma(\lambda + 1)\theta^\lambda} x^\lambda \exp\left(-\frac{x}{\theta}\right) ; \ 0 < x < \infty, \ \lambda > -1, \theta > 0$$ \hspace{1cm} (1.1)

### 1.6.2 Weibull Distribution

The Weibull distribution was named after a Swedish mathematician Waloddi Weibul (1887-1979), who described it in detail in 1951. Although it was first identified by a French mathematician Rene Frechet in 1927. Due to its versatility and simplicity during lifetime data analysis, Weibull is mostly used in lifetime distribution analysis.

The general form of the density function of the Weibull distribution $W(\tau, h)$ is given as:

$$f(x; \tau, h) = \frac{\tau}{h} \left(\frac{x}{h}\right)^{\tau - 1} \exp\left\{-\left(\frac{x}{h}\right)^\tau\right\} ; \ 0 < x < \infty, \ h, \tau > 0.$$ \hspace{1cm} (1.2)

The Weibull distribution is a continuous probability distribution which has two parameters where $\tau$ is the shape parameter and $h$ is the scale parameter and the hazard rate is given as:

$$h(x|\tau, h) = \frac{\tau}{h} \left(\frac{x}{h}\right)^{\tau - 1}$$ \hspace{1cm} (1.3)
The application of hazard function of a Weibull distribution is so interesting due to the flexibility of hazard function when the shape parameter $\tau$ is adjusted. When $\tau$ is greater than 1 the hazard rate is an increasing function of the observations (Convex), if it is less than one the hazard rate is a decreasing function (concave) and a constant when it $\tau$ is 1. In this study, consideration is limited to the one parameter form of the first two situations.

### 1.6.3 Generalized Gaussian Distribution

For a random variable ($X$), the two parameters Generalized Gaussian distribution has the probability distribution function as:

$$f(x; \sigma, \mu) = \frac{1 + \mu}{2\sigma \Gamma(\frac{1}{1+\mu})} \exp\{-\left(\frac{|x|^{1+\mu}}{\sigma}\right)\} \quad -\infty < x < \infty, \sigma > 0, \mu > -1 \quad (1.4)$$

The transformation $Z = X^{1+\mu}$, the random variable $Z$ have the Gamma distribution with $\theta = 1$ and $\lambda = \frac{1}{1+\mu}$. In this study we focus our attention on the one parameter Generalized Half Gaussian distribution, the case where $x > 0$.

$$f(x; \mu) = \frac{1 + \mu}{\Gamma(\frac{1}{1+\mu})} \exp\left(-x^{1+\mu}\right) \quad 0 < x < \infty, \mu > -1 \quad (1.5)$$

### 1.6.4 Generalized Gamma Distribution

The Generalized Gamma distribution was made known by an American mathematician Stacy (1962) [12]. It is a continuous probability distribution with three
parameters. It is a generalization of two parameters Gamma distribution. The dis-
tribution helps us to determine which parametric model is appropriate for a given set of
lifetime data. The probability density function of Generalized Gamma distribution
(GGD1) is

\[ f(x; a, d, p) = \frac{(a/d)^d x^{d-1} \exp\left\{-\left(\frac{x}{a}\right)^p\right\}}{\Gamma\left(\frac{d}{p}\right)} ; x > 0, a, p, d > 0. \]  \hspace{1cm} (1.6)

This distribution is popular due to its flexibility. Many well-known distributions
such as Exponential, Weibull, Gamma, and Log Normal distribution are special case
of the Generalized Gamma distribution.

Stacy and Mihran (1965) [13] introduced a further generalization of the General-
ized Gamma distribution and discuss parameters estimation using modified method
of moment technique. The density function of a further Generalized Gamma distri-
bution (GGD2) for a random variable \( X \) is

\[ f(x; \alpha, \gamma, \delta) = \frac{\gamma}{\delta \Gamma(\alpha)} \left(\frac{x}{\delta}\right)^{\alpha \gamma - 1} \exp\left\{-\left(\frac{x}{\delta}\right)^\gamma\right\} ; x > 0, \alpha, \gamma, \delta > 0. \] \hspace{1cm} (1.7)

Igor Volodin(1981) [15] used this form of definition for Generalized Gamma distribu-
tion (GGD3):

\[ \frac{1}{\theta} f\left(\frac{x}{\theta}; \lambda, \beta\right) = \frac{1 + \beta}{\Gamma\left(\frac{1 + \lambda}{1 + \beta}\right)} \theta^{\left(\frac{1 + \lambda}{1 + \beta}\right)} x^\lambda \exp\left\{-\frac{x^{1+\beta}}{\theta}\right\}; x > 0, \theta > 0, \lambda, \beta > -1 \] \hspace{1cm} (1.8)

This was used for constructing hypothesis test in order to discriminate between differ-
ent types of distributions that are connected to the Generalized Gamma distribution.
proposed by Volodin (1981) [15] and so many other papers have used this form in reliability model.

In order to combine the three main reliability models discussed above which their properties are the problem of this research, we considered parametrization of the [12] definition of Generalized Gamma distribution as given bellow

\[
f(x; \theta, \lambda, \mu) = \frac{1 + \mu}{\Gamma\left(\frac{(1+\lambda)(1+\mu)}{(1+\lambda)}\right)} x^\lambda \exp\left\{-\frac{x^{1+\mu}}{\theta}\right\}; \quad x > 0, \theta > 0, \lambda, \mu > -1 \quad (1.9)
\]

The subfamilies of Generalized Gamma distribution considered in the literature are Gamma for \( \mu = 0 \), Weibull for \( \lambda = \beta = \tau \), \( \theta = h^{\mu+1} \) and Generalized Half Gaussian for \( \lambda = 0, \theta = 1 \).

### 1.7 Powerful Test Statistic to Discriminate Distributions Connected with Generalized Gamma Distribution

According to Volodin (1981) [15], there is a locally most powerful test statistic to discriminate two types of distributions connected with the Generalized Gamma distribution. Let \( X_1, X_2, X_3, \ldots \) to be a succession of independent and identically distributed random variables that follow a Generalized Gamma distribution. The following statistics has been suggested by him: \( \bar{\xi} = c_1 \xi_1 + c_2 \xi_2 \), where

\[
\xi_1 = \frac{1}{n} \sum_{i=1}^{n} \ln X_i - \ln \left(\frac{1}{n} \sum_{i=1}^{n} X_i\right), \quad \xi_2 = \ln \left(\frac{1}{n} \sum_{i=1}^{n} X_i\right) - \frac{\frac{1}{n} \sum_{i=1}^{n} X_i \ln X_i}{\frac{1}{n} \sum_{i=1}^{n} X_i}.
\]

25
We know that the statistic $\mathcal{T}$ is a linear combination of statistics $\mathcal{T}_1$ and $\mathcal{T}_2$. The rejection region of the powerful test statistic introduced above will have the usual form $\mathcal{T} > c$. We can use statistic $\mathcal{T}$ to test one of the combinations of models related to Generalized Gamma distribution.
Chapter 2

Literature Review

The problem of assigning a probability density function to a random variable \( X \) from two separate families of distributions was emphasized by Cox (1961) [1] to create more awareness of its existence. In his effort to address this problem, he developed the theory of hypothesis testing for separate families of distributions which are overlapping Cox (1961, 1962) [1], [2]. In 1961, he gave some examples of this problem and developed a general large-sample procedure based on the modification of the Neyman-Pearson maximum likelihood ratio and then he gave some general comments on the formulation of the problem in 1962. He also suggests combining the two hypotheses in a general model which would both be special cases.

Dumonceaux, Antle and Haas (1973) [5] considered the problem of selecting a model from two models with unknown location and scale parameters where the distribution of the ratio of maximum likelihoods does not depend upon the values of the...
nuisance location and scale parameters. Consequently, this ratio provides a convenient
test for discriminating between two location and scale parameter models when the
parameters are unknown. For any given location and scale parameter distribution one
can construct tables of critical values for the Komogorov-Smirnov test. In the same
year, Dumonceaux and Antle [4] proposed a likelihood ratio test for discriminating
between the Lognormal and Weibull distributions.

the distribution is independent of the nuisance parameter. Hence it is possible to
reduce the calculation necessary for hypothesis testing. This publication is especially
important for us before we follow the techniques developed there.

Pace and Salvan (1990) [10] studied the possibility of constructing best similar
tests. They characterized sufficient statistics of the union of the two families.

between the Generalized Exponential and Gamma distributions and used them to
determine the necessary sample size for a user specified probability of correct selection
and a tolerance limit. His work was observed by a monte carlo simulation study which
worked quite well even when the sample size is not too large.

Kundu and Manglick (2004) [8] considered the problem of discriminating between
the Weibull and Lognormal distributions. They used the ratio of the maximized like-
lihood in discriminating between the two distribution functions which is independent
Day and Kundu (2009) [3] considered the model discrimination among the Lognormal, Weibull and Generalized Exponential distributions. All these three distributions have been used quite effectively to analyze lifetime data in the reliability analysis.

Orawan Supapueng (2016) [11] worked on the same problem of discriminating between Generalized Exponential and Gamma distribution. She used the asymptotic method for discriminating these two distributions. This method is independent of the scale and location nuisance parameters and it was observed to work for different sample sizes. This thesis considered details literature view of [11] with an extension in test cases.
Chapter 3

Score Statistics

3.1 Invariant statistical inference: The notion of maximal scale invariant statistic

There exists a wide class of statistical inference problems, where a specific transformation of sample data should not cause to a decision making which contradicts to the decision made on the bases of the original data. For example, if the sample data consist of measurements of the amount of harmful impurities in a diesel fuel, then the decision of its quality should not depend on the units in which the measurements have been made (in grams, kilograms, or in percentage of the total fuel weight). Especially for our topic, if we are making inference on lifetime measurements, then our decisions should also be free of the way we measure the time (in seconds, or in years, or in percentage of a standard lifetime). Hence, the statistical inference should obey
invariance properties to scale transformations of the sample data.

If the statistical problem is consists in the estimation of the distribution function (typical problem in the Reliability Theory or Survival Analysis or when a decision should be made on the guarantee lifetime of a product), then the estimate should not depend on the order in which we obtain the sample values. This means that the statistical inference should be invariant under permutations of the sample components.

There are many other examples of types of invariance requirements for statistical inference, for example with respect to a choice of the coordinate system for observations (temperature can be measured in Fahrenheit and Celsius), invariance with respect to orthogonal transformations and so on. For us the scale invariance is the most important, the invariance under permutations will follow from the formula.

The general mathematical expression of such requirement of “equanimity” of the statistical inference is formulated in terms of its invariance with respect of an appropriate group of transformations of the sample space. We will not go into details of the Group Theory because we are interested only in scale invariance.

Let \( \mathbf{X} = \{X_1, X_2, \ldots, X_n\} \) be a sample of size \( n \). A statistics \( T(\mathbf{X}) = T(X_1, X_2, \ldots, X_n) \) is called \textit{scale invariant} if for any appropriate \( \theta \) (in our case we should have that \( \theta > 0 \) because we are dealing with samples from a lifetime distributions) \( T\left(\frac{\mathbf{X}}{\theta}\right) = T(\mathbf{X}) \), where the sample

\[
\frac{\mathbf{X}}{\theta} = \left\{ \frac{X_1}{\theta}, \frac{X_2}{\theta}, \ldots, \frac{X_n}{\theta} \right\}.
\]
The simplest but the most important for further discussion example of such scale invariant statistics would be

\[ T(X) = \left( \frac{X_1}{X_n}, \frac{X_2}{X_n}, \ldots, \frac{X_{n-1}}{X_n} \right). \]

Note also that this choice of statistic \( T(x) \) provides us so called maximum scale invariant statistic, that is, if we have two samples \( X' \) and \( X'' \) such that \( T(X') = T(X'') \), then \( X'' = X'/\theta \) for some suitable number \( \theta \). This means that the statistic \( T(X) \) is constant for any scale transformation of the sample of the form \( X/\theta \) and it takes different values for samples that are not a scale transformation of each other.

Now we derive the distribution of this \((n - 1)\)-dimensional statistic \( T \) under the assumption that the sample is taken from the distribution with the support on the positive numbers with the density function \( f(x), x > 0 \).

**Proposition 1.** The density function of statistic \( T \) can be written as:

\[
f_T(y_1, \ldots, y_{n-1}) = \int_0^\infty x^{n-1} f(x) \prod_{i=1}^{n-1} f(y_i/x) dx.
\]

**Proof:** Consider case \( n = 3 \). Then random variables \( X_1, X_2, X_3 \) have marginal density functions \( f(x) \) and by independence their joint density function is

\[
f_{X_1,X_2,X_3}(x_1, x_2, x_3) = f(x_1) f(x_2) f(x_3).
\]

We are interested in the density function \( f_{Y_1,Y_2}(y_1, y_2) \), where \( Y_1 = \frac{X_1}{X_3} \) and \( Y_2 = \frac{X_2}{X_3} \). Introduce one “temporary” variable \( Y_3 = X_3 \) more, then we have that the support of
all three variables $Y_1, Y_2, Y_3$ is $y_1 > 0, y_2 > 0, y_3 > 0$ and

$$J = \begin{cases} 
  y_1 = x_1/x_3 \\
  y_2 = x_2/x_3 \\
  y_3 = x_3 
\end{cases} \quad \text{or} \quad \begin{cases} 
  x_1 = y_1x_3 \\
  x_2 = y_2x_3 \\
  x_3 = y_3 
\end{cases}$$

The Jacobian of this transformation is

$$\det \begin{pmatrix}
  \frac{\partial x_1}{\partial y_1} & \frac{\partial x_2}{\partial y_1} & \frac{\partial x_3}{\partial y_1} \\
  \frac{\partial x_1}{\partial y_2} & \frac{\partial x_2}{\partial y_2} & \frac{\partial x_3}{\partial y_2} \\
  \frac{\partial x_1}{\partial y_3} & \frac{\partial x_2}{\partial y_3} & \frac{\partial x_3}{\partial y_3}
\end{pmatrix} = \det \begin{pmatrix}
  y_3 & 0 & y_1 \\
  0 & y_3 & y_2 \\
  0 & 0 & 1
\end{pmatrix} = y_3^2$$

Hence, by the Transformation Theorem (see, for example Gut (2009) [7] Theorem 2.2.1.), the joint density function of the random variables $Y_1, Y_2, Y_3$ is

$$f_{Y_1, Y_2, Y_3}(y_1, y_2, y_3) = f(y_1y_3)f(y_2y_3)f(y_3)y_3^2.$$ 

Because we do not need the “temporary” variable $Y_3$, we integrate it out

$$f_{Y_1, Y_2}(y_1, y_2) = \int_0^\infty y_3^2 f(y_3)f(y_1y_3)f(y_2y_3)dy_3.$$ 

The last expression after the change of variables $x = y_3$ can be rewritten as

$$f_{Y_1, Y_2}(y_1, y_2) = \int_0^\infty x^2 f(x) \prod_{i=1}^{2} f(y_i x) dx,$$

which is exactly what is stated in the Proposition for the case $n = 3$.

The generalization on the case of arbitrary $n$ can be done in the same way by letting $Y_i = \frac{X_i}{X_n}, 1 \leq i \leq n-1$ and $Y_n = X_n$. Note that the Jacobian $J = y_n^{n-1}$. 

33
Obviously, there exist many other maximal shale invariant statistics, for example

\[ T(X) = (\log(X_2) - \log(X_1), \log(X_3) - \log(X_1), \ldots, \log(X_n) - \log(X_1)) \].

In the particular case \( n = 1 \) there is no nontrivial scale invariant statistics.

### 3.2 Most powerful invariant tests

Consider a construction of invariant test for hypotheses testing when there exists a nuisance scale parameter \( \theta \). This means that the distribution of the observed random element \( \xi \) depends on \( \theta \) in a “multiplicative way”. In our thesis we consider random variables with the distribution function of the form \( F_{\lambda}(\frac{x}{\theta}) \theta \geq 0 \). This means that the density function has the form \( f_{\lambda}(x; \theta) = \frac{1}{\theta} p_{\lambda}(\frac{x}{\theta}) \). As an example of such density function, consider Gamma distribution as follows:

\[ f_{\lambda}(x; \theta, \lambda) = \frac{1}{\theta^\lambda \Gamma(\lambda)} x^{\lambda-1} \exp \left\{ -\frac{x}{\theta} \right\}, x > 0; \theta > 0, \lambda > 0. \]

Notice that this function can be rewritten:

\[ f_{\lambda}(x; \theta, \lambda) = \frac{1}{\theta^{\lambda} \Gamma(\lambda)} (\frac{x}{\theta})^{\lambda-1} \exp \left\{ -\frac{x}{\theta} \right\}. \]

If we denote

\[ p_{\lambda}(t) = \frac{1}{\Gamma(\lambda)} t^{\lambda-1} \exp\{-t\}, \]

then obviously \( f_{\lambda}(x; \theta) = \frac{1}{\theta} p_{\lambda}(\frac{x}{\theta}) \).
Next we observe that the function \( p_\lambda(\cdot) \) depends on \( \frac{x}{\theta} \), not by \( \theta \) itself. This is the main advantage of the form \( \frac{1}{\theta} p_\lambda(\frac{x}{\theta}) \).

For the parameter \( \lambda \) we are testing some hypothesis with the nuisance scale parameter \( \theta \).

In order to construct the most powerful test with the help of Neyman-Pearson lemma we need to derive the distribution of the invariant statistic. As it has been mentioned above, we can consider the following maximum scale invariant statistic

\[
T(X) = \left( \frac{X_1}{X_n}, \frac{X_2}{X_n}, \ldots, \frac{X_{n-1}}{X_n} \right).
\]

The likelihood for \( T \) by the proposition can be written as

\[
L(T|X) = f_T \left( \frac{X_1}{X_n}, \ldots, \frac{X_{n-1}}{X_n} \right) = \int_0^\infty x^{n-1} f_0(x) \prod_{k=1}^{n-1} f \left( \frac{X_k}{X_n} x \right) \, dx.
\]

Change of variables \( x = tX_n \) leads to the following likelihood

\[
L(T|X) = X_n^n \int_0^\infty t^{n-1} \prod_{k=1}^{n-1} f \left( \frac{X_k}{X_n} t \right) \, dt.
\]

Lemma. If the density functions are of the form \( f(x) = \frac{1}{\theta} p \left( \frac{x}{\theta} \right) \), then

\[
L(T|X) = X_n^n \int_0^\infty t^{n-1} \prod_{k=1}^{n} p \left( X_k t \right) \, dt.
\]
Proof: Notice that

\[
L(T|X) = X_n^n \int_0^\infty t^{n-1} \prod_{k=1}^n f(X_k t) \, dt
\]

\[
= X_n^n \int_0^\infty t^{n-1} \prod_{k=1}^n \frac{1}{\theta^p} \left( X_k \frac{t}{\theta} \right) \, dt
\]

change of variables \(u = t/\theta\)

\[
= X_n^n \int_0^\infty u^{n-1} \prod_{k=1}^n p(X_k t) \, du.
\]

Remark. The main conclusion of the Lemma is that the scale parameter \(\theta\) disappears and we do not need to deal with it anymore.

Consider the problem of discriminating two simple hypotheses with nuisance scale parameter. Notice that this is a different procedure than we discussed in the section on the Likelihood Ratio test because we are not using Maximum Likelihood estimates.

In this case, we test a simple hypothesis \(H_0\) : the sample \(X = (X_1, X_2, \ldots, X_n)\) is taken from a continues distribution with density function \(f_0 = \frac{1}{\theta^p} p_0 \left( \frac{x}{\theta} \right), \ x > 0, \theta > 0\) with the alternative \(H_1\) which suggests that the distribution of each \(X_i, i = 1, \ldots, n\), has density \(f_1 = \frac{1}{\theta^p} p_1 \left( \frac{x}{\theta} \right)\) with the same support \(x > 0\) and the same possible values of the nuisance parameter \(\theta > 0\). Neyman-Pearson test for testing \(H_0\) with alternative \(H_1\) is based on the likelihood ratio. By the Lemma, if the density functions that correspond to hypotheses \(H_0\) and \(H_1\) are of the form \(f_0(x) = \frac{1}{\theta^p} p_0 \left( \frac{x}{\theta} \right)\) and \(f_1(x) = \frac{1}{\theta^p} p_1 \left( \frac{x}{\theta} \right)\)
\[ \frac{1}{\theta}p_1(\frac{x}{\theta}) \], respectively, then the likelihoods for \( H_0 \) and \( H_1 \) can be written as

\[
L_0(T|X) = X_n^n \int_0^\infty t^{n-1} \prod_{k=1}^n p_0(X_k t) \, dt
\]

\[
L_1(T|X) = X_n^n \int_0^\infty t^{n-1} \prod_{k=1}^n p_1(X_k t) \, dt.
\]

Because the distribution from which the sample is taken is considered to be taken is continuous, the most powerful test is not randomized and its the Critical Region in the form \( L_1(T)/L_0(T) > C \), where the constant \( C = C(\alpha) \) is defined according the given significance level \( \alpha \) by the equation

\[
P_0 \left( \frac{L_1(T|X)}{L_0(T|X)} > C \right) = \alpha,
\]

where the probability \( P_0 \) is defined through the density function \( f_0 \).

If the likelihood ratio \( L_1(T|X)/L_0(T|X) \) is a monotone function of a statistic \( U(T) \) and the hypotheses \( H_0 \) and \( H_1 \) are statements about the parameter \( \lambda \) of the distribution with the density function \( \frac{1}{\theta}f_\lambda(\frac{x}{\theta}) \) (with some additional regularity conditions that we do not present here), then we can construct the uniformly most powerful invariant tests for such hypothesis.

Consider the following preliminary example on the uniformly most powerful invariant test construction, which is of special interest for testing of probability models for the lifetime testing (Reliability Theory). Our main goal is to generalize this test for a more general model, which is done in the next section.
Example. Uniformly most powerful invariant test for testing the probability model of “manufacture defect (defect at birth or fragility)” sometimes also called the model of “absence of aftereffects” for the lifetime distributions.

Let sample $X = (X_1, X_2, \ldots, X_n)$ be taken from Gamma distribution with density function

$$f(x; \theta, \lambda) = \frac{1}{\theta^\lambda \Gamma(\lambda)} x^{\lambda-1} \exp\left\{-\frac{x}{\theta}\right\}, \quad x > 0; \theta > 0, \lambda > 0.$$ 

If we denote

$$p_\lambda(t) = \frac{1}{\Gamma(\lambda)} t^{\lambda-1} \exp\{-t\},$$

then the density function can be written in the desired form $f(x; \theta, \lambda) = \frac{1}{\theta} p_\lambda \left(\frac{x}{\theta}\right)$.

The hypotheses to test are $H_0 : \lambda = 1$ with the alternative $H_1 : \lambda > 1$. In the framework of the lifetime testing, when the sample data are the lifetimes of $n$ experimental objects, then the hypothesis $H_0$ corresponds to the probability model of absence of aftereffects (exponential lifetime distribution), while the alternative $H_1$ corresponds to the model of aging and deterioration of the object under the influence of stress loads.

Hypothesis $H_0$ is scale invariant, hence its testing should be done based on the maximal invariant statistic

$$T(X) = \left(\frac{X_1}{X_n}, \ldots, \frac{X_{n-1}}{X_n}\right).$$
By Lemma, the corresponding likelihood function is

\[ L(\lambda|X) = X^n \int_0^\infty t^{n-1} \prod_{k=1}^n p_\lambda(X_k t) \, dt = \]

\[ \frac{X^n}{\Gamma^n(\lambda)} \prod_{k=1}^n X_k^{\lambda-1} \int_0^\infty t^{n\lambda-1} \exp \left\{ -t \sum_{k=1}^n X_k \right\} \, dt = X^n \frac{\Gamma(n\lambda)}{\Gamma^n(\lambda)} \frac{\prod_{k=1}^n X_k^{\lambda-1}}{\left( \sum_{k=1}^n X_k \right)^n \lambda}. \]

Notice that the log likelihood ratio \( \ln L(\lambda'') - \ln L(\lambda') \) for \( \lambda'' > \lambda' \) is the monotonically increasing function of the statistic

\[ U(T) = \sum_{k=1}^n \ln X_k - n \ln \left( \sum_{k=1}^n X_k \right), \]

hence the uniformly most powerful test is not randomized and rejects the model of absence of aftereffects in favour of the model of aging and deterioration if \( U(T) > C \).

The same result can be obtained by the consideration of likelihood ratio of conditional distributions of the sample vector with respect to the sufficient statistic (for the nuisance parameter) \( \sum_{k=1}^n X_k \).

Anyway, now we left with the problem of finding the critical constant \( C \) by the given significance level \( \alpha \). Unfortunately, the exact distribution of the statistic \( U(T) \) has the form of a double improper integral and it is impossible to work with it. For large sample sizes \( n \) it is preferable to apply the asymptotic distribution of this statistic. In order to simplify the derivation of the asymptotic distribution of \( U(T) \) we introduce the statistic \( V(T) = U(T)/n \).
Proposition 2. Statistic \( V = V(T) \) is asymptotically \( (n \to \infty) \) normal with parameters

\[
\mu(\lambda) = E_\lambda V = \psi(\lambda) - \psi(n\lambda), \quad \sigma^2(\lambda) = \text{Var}_\lambda V = \psi'(\lambda)/n - \psi'(n\lambda),
\]

(3.1)

where \( \psi(x) = d\ln \Gamma(x)/dx \) is the \( \psi \)-Euler function and \( \psi'(x) \) is the derivative of this function.

The power function of the test \( V > C \) obtains the following asymptotic representation

\[
m(\lambda) = 1 - \Phi\left( \frac{C - \mu(\lambda)}{\sigma(\lambda)} \right) + O\left( \frac{1}{\sqrt{n}} \right),
\]

where the critical constant \( C = C(\alpha) = \mu(1) + \sigma(1)\Phi^{-1}(1 - \alpha) \).

**Proof.** Asymptotic normality of the statistic

\[
V = \frac{1}{n} \sum_{k=1}^{n} \ln X_k - \ln \left( \sum_{k=1}^{n} X_k \right)
\]

can be established by the classical Delta-method expanding the function of sample mean \( \bar{X} \) by Taylor series in a neighborhood of \( E\bar{X} \) and keeping only linear terms. In our case

\[
V = \frac{1}{n} \sum_{k=1}^{n} \ln X_k - \ln n - \ln \bar{X}.
\]

Knowing the moments of Gamma-distribution \( E_\lambda \bar{X} = \lambda, E_\lambda \ln X = \psi(\lambda) \), we have

\[
\sqrt{n}V = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} [\ln X_k - \psi(\lambda)] + \sqrt{n} \psi(\lambda) - \sqrt{n} \ln n - \sqrt{n} \ln \bar{X}
\]

\[
= \frac{1}{\sqrt{n}} \sum_{k=1}^{n} [\ln X_k - \psi(\lambda)] + \sqrt{n} \psi(\lambda) -
\]

40
\[
\sqrt{n} \ln n - \frac{\sqrt{n}(\bar{X} - \lambda)}{\lambda} + \frac{\sqrt{n}(\bar{X} - \lambda)^2}{2[\lambda + \gamma(\bar{X} - \lambda)]^2},
\] (3.2)

where \(0 < \gamma < 1\) and the remaining term (last expression in (3.2)) converges in probability to zero with rate \(O(1/\sqrt{n})\). The Linear terms of the expansion are the sums of independent identically distributed random variables with the finite second moments. Hence they are asymptotically normal by the Central Limit Theorem. This is the statement of the Delta-method.

Of course, the expansion (3.2) allows us to calculate the parameters of the asymptotic normality, but in this case it is possible to find the exact value of the mathematical expectation and variance of \(U\) using its Moment Generating Function. It is interesting that in this case we can derive the Moment Generating Function in closed form:

\[
\phi(t) = E\exp(tU) = \frac{1}{\Gamma^n(\lambda)} \int_0^\infty \cdots \int_0^\infty \prod_{k=1}^n x_k^{\lambda-1} \times \exp \left\{ t \left[ \sum_{k=1}^n \ln x_k - n \ln \left( \sum_{k=1}^n x_k \right) \right] - \sum_{k=1}^n x_k \right\} dx_1 \cdots dx_n =
\]

\[
\frac{1}{\Gamma^n(\lambda)} \int_0^\infty \cdots \int_0^\infty \left( \sum_{k=1}^n x_k \right)^{-nt} \prod_{k=1}^n x_k^{\lambda+t-1} \exp \left\{ - \sum_{k=1}^n x_k \right\} dx_1 \cdots dx_n.
\]

The last \(n\)-fold integral up to a constant is the mathematical expectation of \(\sum_{k=1}^n X_k\), when \(X_1, \ldots, X_n\) are independent and identically Gamma distributed with the form parameter \(\lambda + t\). This in some sense artificial interpretation simplifies the
calculation of the integral. Therefore
\[
\phi(t) = \frac{\Gamma^n(\lambda + t)}{\Gamma^n(\lambda)} \frac{1}{\Gamma(n(\lambda + t))} \int_0^\infty x^{-nt}x^{n(\lambda + t) - 1}e^{-x}dx = \frac{\Gamma(n\lambda)\Gamma^n(\lambda + t)}{\Gamma^n(\lambda)\Gamma(n(\lambda + t))}.
\]

Expanding \( \ln \phi(t) \) into Taylor series by the powers of \( t \), we obtain the mean value of \( U \) (the coefficient by \( t \)) and variance of \( U \) (the coefficient by \( (t)^2/2 \)). Formulae (3.1) for the mean and variance of the statistic \( V \) are obtained by the division \( E_\lambda U \) by \( n \) and division \( \text{Var}_\lambda U \) by \( n^2 \).

### 3.3 Score statistics for the Generalized Gamma distribution

For the derivation of the score statistics we consider a sample

\[
X = \{X_1, X_2, \ldots, X_n\}
\]

of size \( n \) from the Generalised Gamma distribution with density presented in Volodin (1974) [14]

\[
f(x; \theta, \lambda, \mu) = \frac{\mu + 1}{\theta^{\lambda+1}\Gamma\left(\frac{\lambda+1}{\mu+1}\right)} x^\lambda \exp\left\{ -\left(\frac{x}{\theta}\right)^{\mu+1} \right\}, x > 0, \theta > 0, \lambda, \mu > -1. \tag{3.3}
\]

Note that this expression can be rewritten:

\[
f(x; \theta, \lambda, \mu) = \frac{1}{\theta} \frac{\mu + 1}{\Gamma\left(\frac{\lambda+1}{\mu+1}\right)} \left(\frac{x}{\theta}\right)^\lambda \exp\left\{ -\left(\frac{\theta}{x}\right)^{\mu+1} \right\},
\]

42
that is, if we denote

\[ p(t; \lambda, \mu) = \frac{\mu + 1}{\Gamma\left(\frac{\lambda + 1}{\mu + 1}\right)} t^{\lambda} \exp \left\{ -t^{\mu+1} \right\}, \]

then the density function can be represented in the desired form

\[ f(x; \theta, \lambda, \mu) = \frac{1}{\theta} p\left(\frac{x}{\theta}; \lambda, \mu\right) \]

and the technique of the scale invariant tests developed above can be applied.

In this case \( \theta \) is scale nuisance parameter that allows us to use the maximum invariant statistic

\[ T(X) = \left( \frac{X_1}{X_n}, \frac{X_2}{X_n}, \ldots, \frac{X_{n-1}}{X_n} \right) \]

when we make inferences on parameters \( \lambda \) and \( \mu \).

By Lemma, the likelihood function of statistic \( T \) when sample is taken from the Generalized Gamma distribution can be written as:

\[
L(T|X) = X_n \int_0^\infty t^{n-1} \prod_{k=1}^{n} \frac{\mu + 1}{\Gamma\left(\frac{\lambda + 1}{\mu + 1}\right)} \left( X_k t \right)^{\lambda} \exp \left\{ -(X_k t)^{\mu+1} \right\} dt
\]

\[
= \frac{X_n^{n}(\mu + 1)^{n} \prod_{k=1}^{n} X_k^{-\lambda}}{\Gamma^n \left(\frac{\lambda + 1}{\mu + 1}\right)} \int_0^\infty t^{n(\lambda+1)-1} \exp \left\{ -t^{\mu+1} \sum_{k=1}^{n} X_k^{\mu+1} \right\} dt.
\]

Denote

\[ a = \frac{X_n^{n}(\mu + 1)^{n} \prod_{k=1}^{n} X_k^{\lambda}}{\Gamma^n \left(\frac{\lambda + 1}{\mu + 1}\right)}, \quad b = \left( \sum_{k=1}^{n} X_k^{\mu+1} \right)^{-\frac{1}{\mu+1}}, \quad c = n(\lambda + 1) - 1. \]

With these notations the likelihood function can be rewritten:

\[
L(T|X) = a \int_0^\infty t^c \exp \left\{ -\left( \frac{t}{b} \right)^{\mu+1} \right\} dt.
\]
Multiplying and dividing by $\frac{\Gamma\left(\frac{1+c}{\mu+1}\right)b^{c+1}}{\mu+1}$, the likelihood function can be rewritten:

$$L(T|X) = a\Gamma\left(\frac{1+c}{\mu+1}\right)b^{c+1} \int_0^\infty \frac{\mu+1}{\Gamma\left(\frac{1+c}{\mu+1}\right)b^{c+1}} t^\mu \exp\left\{-\left(\frac{t}{b}\right)^{\mu+1}\right\} dt.$$  

Notice that under the integral in the last expression we the density function of the Generalized Gamma distribution with parameters $\theta = b, \lambda = c, \mu = \mu$. Therefore this integral is equal to one as an integral of a density function. Hence

$$L(T|X) = \frac{a\Gamma\left(\frac{1+c}{\mu+1}\right)b^{c+1}}{\mu+1}.$$  

Substituting back the values of $a, b$ and $c$, we obtain:

$$L(T|X) = \frac{(\mu+1)^{n-1} \Gamma\left(\frac{n(\lambda+1)}{\mu+1}\right)}{\Gamma_n\left(\frac{\lambda+1}{\mu+1}\right)} \frac{X_n^n \prod_{k=1}^n X_k^\lambda}{\left(\sum_{k=1}^n X_k^{\mu+1}\right)^{\frac{n(\lambda+1)}{\mu+1}}}.$$  

According to the last formula, the likelihood functions for particular distributions are as follows.

i. Sample from the Gamma distribution $X \sim G(\lambda)$. Take $\mu = 0$ to obtain:

$$L^G(T|X) = \frac{\Gamma(n(\lambda+1))}{\Gamma(n+1)} \frac{X_n^n \prod_{k=1}^n X_k^\lambda}{\left(\sum_{k=1}^n X_k\right)^{n(\lambda+1)}}.$$  

ii. Sample from the Weibull distribution $X \sim W(\tau)$. Take $\mu = \lambda = \tau$ to obtain:

$$L^W(T|X) = (\tau + 1)^{n-1} \Gamma(n) \frac{X_n^n \prod_{k=1}^n X_k^\tau}{\left(\sum_{k=1}^n X_k^{\tau+1}\right)^n}.$$
iii. Sample from the Generalized Half-Gaussian distribution $X \sim GH(\mu)$. Take $\lambda = 0$ to obtain:

$$L^{GH}(T|X) = \frac{(\mu + 1)^{n-1} \Gamma \left( \frac{n}{\mu + 1} \right)}{\Gamma \left( \frac{1}{\mu + 1} \right)} \frac{X_n \prod_{k=1}^{n} X_k}{\left( \sum_{k=1}^{n} X_k^{\mu + 1} \right)^{\frac{n}{\mu + 1}}}.$$ 

The log-likelihood function of statistic $T$ when sample is taken from the Generalized Gamma distribution can be written as:

$$\ln L(T|X) = (n - 1) \ln(\mu + 1) + \ln \left( \Gamma \left( \frac{\lambda + 1}{\mu + 1} \right) \right) - n \ln \left( \Gamma \left( \frac{\lambda + 1}{\mu + 1} \right) \right)$$

$$+ n \ln(X_n) + \lambda \sum_{k=1}^{n} \ln(X_k) - n \left( \frac{\lambda + 1}{\mu + 1} \right) \ln \left( \sum_{k=1}^{n} X_k^{\mu + 1} \right).$$

According to the last formula, the log-likelihood functions for particular distributions are as follows.

i. Sample from the Gamma distribution $X \sim G(\lambda)$. Take $\mu = 0$ to obtain:

$$\ln L^G(T|X) = \ln(\Gamma(n(\lambda + 1))) - n \ln(\Gamma(\lambda + 1)) + n \ln(X_n)$$

$$+ \lambda \sum_{k=1}^{n} \ln(X_k) - n(\lambda + 1) \ln \left( \sum_{k=1}^{n} X_k^{\mu + 1} \right).$$

ii. Sample from the Weibull distribution $X \sim W(\tau)$. Take $\mu = \lambda = \tau$ to obtain:

$$\ln L^W(T|X) = (n - 1) \ln(\tau + 1) + \ln \Gamma(n) + n \ln(X_n)$$

$$+ \tau \sum_{k=1}^{n} \ln(X_k) - n \ln \left( \sum_{k=1}^{n} X_k^{\tau + 1} \right).$$
iii. Sample from the Generalized Half-Gaussian distribution $X \sim GH(\mu)$. Take $\lambda = 0$ to obtain:

$$
\ln L^{GH}(T|X) = (n - 1) \ln(\mu + 1) + \ln \left( \Gamma \left( n \frac{1}{\mu + 1} \right) \right) - n \ln \left( \Gamma \left( \frac{1}{\mu + 1} \right) \right) \\
+ n \ln(X_n) - n \frac{1}{\mu + 1} \ln \left( \sum_{k=1}^{n} X_{k}^{\mu+1} \right).
$$

By the definition, the score statistic is the derivative of the log-likelihood function by the parameter when the value of this parameter equals zero.

Therefore, the score statistics for particular distributions can be derived as follows. Notice that we used the definition of the $\psi$-Euler function $\psi(x) = d\ln \Gamma(x)/dx$ in all these derivations.

i. Sample from the Gamma distribution $X \sim G(\lambda)$. Take derivative by $\lambda$ to obtain:

$$
\frac{d \ln L^{G}(T|X)}{d\lambda} = n \psi(n(\lambda + 1)) - n \psi(\lambda + 1) + \sum_{k=1}^{n} \ln(X_k) - n \ln \left( \sum_{k=1}^{n} X_k \right).
$$

Let $\lambda = 0$ to obtain the score statistic for Gamma distribution $X \sim G(\lambda)$:

$$
T^{G}(X) = n \psi(n) - n \psi(1) + \sum_{k=1}^{n} \ln(X_k) - n \ln \left( \sum_{k=1}^{n} X_k \right),
$$

or, ignoring constants:

$$
T^{G}(X) = \sum_{k=1}^{n} \ln(X_k) - n \ln \left( \sum_{k=1}^{n} X_k \right).
$$
ii. Sample from the Weibull distribution $X \sim W(\tau)$. Take derivative by $\tau$ to obtain:

$$\frac{d \ln L^W(T|X)}{d\tau} = \frac{n - 1}{\tau + 1} + \sum_{k=1}^{n} \ln(X_k) - \frac{n}{\sum_{k=1}^{n} X_k^{\tau+1}} \sum_{k=1}^{n} \ln X_k X_k^{\tau+1}.$$ 

Let $\tau = 0$ to obtain the score statistic for Weibull distribution $X \sim W(\tau)$:

$$T^W(X) = n - 1 + \sum_{k=1}^{n} \ln(X_k) - \frac{\sum_{k=1}^{n} X_k \ln X_k}{\sum_{k=1}^{n} X_k},$$

or, ignoring constants:

$$T^W(X) = \sum_{k=1}^{n} \ln(X_k) - \frac{\sum_{k=1}^{n} X_k \ln X_k}{\sum_{k=1}^{n} X_k}.$$  

iii. Sample from the Generalized Half-Gaussian distribution $X \sim GH(\mu)$. Take derivative by $\mu$ to obtain

$$\frac{d \ln L^{GH}(T|X)}{d\mu} = \frac{n - 1}{\mu + 1} - \frac{n}{(\mu + 1)^2} \psi\left(\frac{n}{\mu + 1}\right) + \frac{n}{(\mu + 1)^2} \psi\left(\frac{1}{\mu + 1}\right)$$

$$+ \frac{n}{(\mu + 1)^2} \ln\left(\sum_{k=1}^{n} X_k^{\mu+1}\right) - \frac{n}{\mu + 1} \frac{1}{\sum_{k=1}^{n} X_k^{\mu+1}} \ln(X_k).$$

Let $\mu = 0$ to obtain the score statistic for Generalized Half-Gaussian distribution $X \sim GH(\mu)$:

$$T^{GH}(X) = n - 1 - \psi(n) + n\psi(1) + n \ln\left(\sum_{k=1}^{n} X_k\right) - \frac{n}{\sum_{k=1}^{n} X_k \ln(X_k)},$$

47
or, ignoring constants:

\[ T^{GH}(\mathbf{X}) = n \ln \left( \sum_{k=1}^{n} X_k \right) - \frac{n}{\sum_{k=1}^{n} X_k \ln(X_k)}. \]

We ignore constants for the score statistics because they do not influence the corresponding tests.
Chapter 4

Simulation for power of the tests based on Score statistics

According to the articles [15] and [14] by I. Volodin, uniformly most powerful test for discriminating Generalized Half-Gaussian distribution against Gamma distribution has the formula:

\[ T_1 = \sum_{k=1}^{n} \ln X_k - n \ln \left( \sum_{k=1}^{n} X_k \right). \]

Locally most powerful test for discriminating Generalized Half-Gaussian distribution against Weibull distribution has the formula:

\[ T_1 + T_2 = n \left( \sum_{k=1}^{n} \ln X_k - \frac{\sum_{k=1}^{n} X_k \ln X_k}{\sum_{k=1}^{n} X_k} \right). \]

The test is consistent and performs very well when we discriminate Generalized Half-Gaussian distribution against Gamma or Weibull distribution.

We show that these tests are inconsistent for discriminating Gamma and Weibull
distributions. This is achieved by simulating powers of both test statistics mentioned above.

The hypotheses we are testing are:

**Case 1:**

- $H_0$: the sample is obtained from Weibull distribution.
- $H_1$: the sample is obtained from Gamma distribution

**Case 2:**

- $H_0$: the sample is obtained from Gamma distribution.
- $H_1$: the sample is obtained from Weibull distribution

Sample sizes $n = 10, 30, 80, 100, 300$ and $400$ with significance level $\alpha = 0.05$ are considered. Number of simulations, $N = 20,000$.

Now we describe steps for Case 2 and statistic $\bar{T}_1$. For Case 1 and statistic $\bar{T}_1 + \bar{T}_2$ all steps are similar.

**Step 1:** The procedure for finding the critical constant by the sample from exponential distribution defined by the 5% upper percentile for the statistic. Simulate $N$ times a sample of size $n$ from Exponential distribution with parameter 1. For each sample calculate the statistic $\bar{T}_1$. Take $(1 - \alpha)100$ percentile for these statistic values. That is, arrange the values of the statistic from smaller to bigger $(\bar{T}_1 \leq \bar{T}_2 \leq \cdots \leq \bar{T}_N)$ and find the value at the place $(1 - \alpha) \times N$, this will give us
the constant for the rejection region

\[ C_1 = \mathcal{I}_{(\alpha N)}. \]

**Step 2.** Generate sequences of random numbers from Gamma and Weibull distributions with different shape parameters for modelling simulating statistic \( \mathcal{I}_1 \), namely:

Fix \( \lambda \) and simulate \( N \) times a sample of size \( n \) from Gamma distribution with parameter \( \lambda \).

Fix \( \tau \) and simulate \( N \) times a sample of size \( n \) from Weibull distribution with parameter \( \tau \).

For each sample calculate the statistic of interest \( \mathcal{I}_1 \).

**Step 3.** Calculate the number \( A(\lambda) \) of the statistic \( \mathcal{I}_i, 1 \leq i \leq N \) which are greater than \( C_1 \). For each \( \lambda \), calculate type I error of the test as

\[ \alpha(\lambda) = A(\lambda)/N. \]

**Step 4.** Calculate the number \( M(\tau) \) of the statistic \( \mathfrak{T}_i, 1 \leq i \leq N \) which are greater than \( C_1 \). For each \( \tau \) calculate power of the test as

\[ m(\tau) = M(\tau)/N. \]

### 4.1 Graphical illustrations of powers of Score statistics

Below we presents plots for the behavior of power functions for tests \( \mathcal{I}_1 \), and \( \mathcal{T}_1 + \mathcal{T}_2 \) as functions of the shape parameters of Gamma and Weibull distributions for
different sample sizes \( n \) according to the procedure (Steps 1-4) described above. The dot-dashed line represents the power function of the Gamma distribution and solid line represents the Weibull distribution.

From plots it is possible to see that the test’s power function and type 1 error behave themselves nearly identical for discriminating Gamma and Weibull distributions. Therefore, Score statistics have practically the same power as for testing for the null hypothesis as for the alternative. This means that the test in inconsistent for discriminating Gamma and Weibull distributions.

Additional argument that the test is inconsistent for discriminating Gamma and Weibull distributions can be made based on simulating each of statistics separately and presenting their values as functions of the sample size \( n \) and different values of shape parameters \( \lambda \) or \( \tau \). For the following plots the dot-dashed line represents the
Figure 4.2: Power of statistic $\xi$ for Gamma and Weibull distributions, sample size $n = 10$

Figure 4.3: Power of statistic $\xi_1 + \xi_2$ for Gamma and Weibull distributions, sample size $n = 50$
Figure 4.4: Power of statistic $\Xi_1$ for Gamma and Weibull distributions, sample size $n = 50$

Figure 4.5: Power of statistic $\Xi_1 + \Xi_2$ for Gamma and Weibull distributions, sample size $n = 100$
Figure 4.6: Power of statistic $\bar{T}_1$ for Gamma and Weibull distributions, sample size $n = 100$

Figure 4.7: Power of statistic $\bar{T}_1 + \bar{T}_2$ for Gamma and Weibull distributions, sample size $n = 200$
Figure 4.8: Power of statistic $T_1$ for Gamma and Weibull distributions, sample size $n = 200$

value of statistics for the Gamma distribution and solid line represents the the value of statistics for Weibull distribution.
Figure 4.9: Plot of the values of statistic $\Xi_1 + \Xi_2$ for Gamma and Weibull distributions, sample size $n = 10$, shape parameter 0

Figure 4.10: Plot of the values of statistic $\Xi_1 + \Xi_2$ for Gamma and Weibull distributions, sample size $n = 10$, shape parameter 1

Figure 4.11: Plot of the values of statistic $\Xi_1 + \Xi_2$ for Gamma and Weibull distributions, sample size $n = 10$, shape parameter 100
Figure 4.12: Plot of the values of statistic $\mathcal{T}_1$ for Gamma and Weibull distributions, sample size $n = 10$, shape parameter 0

Figure 4.13: Plot of the values of statistic $\mathcal{T}_1$ for Gamma and Weibull distributions, sample size $n = 10$, shape parameter 1

Figure 4.14: Plot of the values of statistic $\mathcal{T}_1$ for Gamma and Weibull distributions, sample size $n = 10$, shape parameter 100
Figure 4.15: Plot of the values of statistic $\xi_1 + \xi_2$ for Gamma and Weibull distributions, sample size $n = 50$, shape parameter 0

Figure 4.16: Plot of the values of statistic $\xi_1 + \xi_2$ for Gamma and Weibull distributions, sample size $n = 50$, shape parameter 1
Figure 4.17: Plot of the values of statistic $\mathcal{T}_1 + \mathcal{T}_2$ for Gamma and Weibull distributions, sample size $n = 50$, shape parameter 100

Figure 4.18: Plot of the values of statistic $\mathcal{T}_1$ for Gamma and Weibull distributions, sample size $n = 50$, shape parameter 0
Figure 4.19: Plot of the values of statistic $T_1$ for Gamma and Weibull distributions, sample size $n = 50$, shape parameter 1

Figure 4.20: Plot of the values of statistic $T_1$ for Gamma and Weibull distributions, sample size $n = 50$, shape parameter 100
Figure 4.21: Plot of the values of statistic $\mathcal{I}_1 + \mathcal{I}_2$ for Gamma and Weibull distributions, sample size $n = 100$, shape parameter 0

Figure 4.22: Plot of the values of statistic $\mathcal{I}_1 + \mathcal{I}_2$ for Gamma and Weibull distributions, sample size $n = 100$, shape parameter 1

Figure 4.23: Plot of the values of statistic $\mathcal{I}_1 + \mathcal{I}_2$ for Gamma and Weibull distributions, sample size $n = 100$, shape parameter 100
Figure 4.24: Plot of the values of statistic $T_1$ for Gamma and Weibull distributions, sample size $n = 100$, shape parameter 0

Figure 4.25: Plot of the values of statistic $T_1$ for Gamma and Weibull distributions, sample size $n = 100$, shape parameter 1

Figure 4.26: Plot of the values of statistic $T_1$ for Gamma and Weibull distributions, sample size $n = 100$, shape parameter 100
Figure 4.27: Plot of the values of statistic $T_1 + T_2$ for Gamma and Weibull distributions, sample size $n = 200$, shape parameter 0

Figure 4.28: Plot of the values of statistic $T_1 + T_2$ for Gamma and Weibull distributions, sample size $n = 200$, shape parameter 1

Figure 4.29: Plot of the values of statistic $T_1 + T_2$ for Gamma and Weibull distributions, sample size $n = 200$, shape parameter 100
Figure 4.30: Plot of the values of statistic $\mathcal{I}_1$ for Gamma and Weibull distributions, sample size $n = 200$, shape parameter 0

Figure 4.31: Plot of the values of statistic $\mathcal{I}_1$ for Gamma and Weibull distributions, sample size $n = 200$, shape parameter 1

Figure 4.32: Plot of the values of statistic $\mathcal{I}_1$ for Gamma and Weibull distributions, sample size $n = 200$, shape parameter 100
Chapter 5

Concluding remarks and Future Research

In this thesis, some probability models of lifetime distributions have been investigated. The main purpose of the thesis is to investigate the consistency of Score test statistics for discriminating Gamma and Weibull distributions by the method of statistical simulations.

From the results obtained, it is possible to make the conclusion that Score test statistics are inconsistent for discriminating Gamma and Weibull distributions. We should not apply it if there is any doubt that the alternative hypothesis is true. If it is required to check that the alternative is Gamma and if in the reality the alternative is Weibull, then this alternative will be accepted with the same probability as the null hypothesis with Gamma distribution.
Bibliography


