Exact Comparison of Hazard Rate Functions of Log-Logistic Survival Distributions

Except where reference is made to the work of others, the work described in this thesis is my own or was done in collaboration with my advisory committee. This thesis does not include proprietary or classified information.

Asha Dixit

Certificate of Approval:

Nedret Billor Associate Professor Mathematics and Statistics	Asheber Abebe, Chair Associate Professor Mathematics and Statistics
Hyejin Shin Assistant Professor Mathematics and Statistics	Joe F. Pittman Interim Dean Graduate School

Exact Comparison of Hazard Rate Functions of Log-Logistic Survival Distributions

Asha Dixit

A Thesis

Submitted to

the Graduate Faculty of

Auburn University

in Partial Fulfillment of the

Requirements for the

Degree of

Master of Science

Auburn, Alabama
Aug 09, 2008

Exact Comparison of Hazard Rate Functions of Log-Logistic Survival Distributions

Asha Dixit

_____

Signature of Author

_____

Date of Graduation

Vita

Asha Dixit was born on July 17, 1980 in Sringeri, India. She graduated with Bachelor of Science from Kuvempu University, Shimoga, Karnataka, India in 2000. She also completed her Bachelor of Education from Kuvempu University in 2001 where she got training for teaching physics and mathematics. In 2002, she joined Bangalore University for MS program in Applied Mathematics. She was awarded Master of Science in 2003. Following her graduation, she taught mathematics to undergraduate students. She joined Auburn University in fall 2006 in the department of mathematics and statistics.

Thesis Abstract

Exact Comparison of Hazard Rate Functions of Log-Logistic Survival Distributions

Asha Dixit

Master of Science, Aug 09, 2008
(M.S., Bangalore University, 2003)
(B.Ed., Kuvempu University, 2001)
(B.S., Kuvempu University, 2000)

69 Typed Pages

Directed by Asheber Abebe

A comparison of hazard rates of multiple treatments are compared under the assumption that survival times follow the log-logistic distribution. Exact test procedures are developed for ordered comparisons of the worst case hazard rates of several log-logistic survival functions. In particular, critical constants are computed for testing the null hypothesis that all dose levels give the same maximum hazard rates versus the alternative that the maximum hazard rates are decreasing with increasing dose level. In addition, critical constants are given for comparing equal maximum hazard rates against the alternative of valley ordered hazard rates. A procedure for building simultaneous confidence intervals for certain contrasts is provided. The procedure proposed in this thesis is then compared to two nonparametric simultaneous inference procedures compare it to two nonparametric procedures: the Jonckheere-Terpstra test and the Mack-Wolfe test.

Acknowledgments

This dissertation could not have been written without Dr. Abebe who not only served as my supervisor but also encouraged and challenged me throughout my academic program. I thank Dr.Abebe and my committee members, Dr. Billor and Dr.Shin, for their guidance.

Thanks to all the faculty members and GTA's in Department of Mathematics and Statistics. I thank my Father Markandeya Dixit, my mother Nagalaxmi Dixit, my brother Vishwanath Dixit for their love and continuous support.

I dedicate this work to my belloved husband Madhu Kirugulige, for his love and support.

Style manual or journal used Journal of Approximation Theory (together with the style known as "aums"). Bibliograpy follows van Leunen's *A Handbook for Scholars.*

Computer software used The document preparation package TEX (specifically LATEX) together with the departmental style-file `aums.sty`.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

In survival analysis, the survival and hazard functions are very functions used to characterize the important characteristics of survival up to a certain specified time and instantaneous death or break-down at a specific time. Although the Weibull distribution is frequently used in modeling survival data, its use is restricted since its hazard function is either monotonically increasing or decreasing. The log-logistic distribution is similar in appearance to the log-normal distribution but its hazard and survival functions can be computed efficiently. Thus it is a good choice when the hazard rate function is desired to have increasing and then decreasing shapes in addition to monotone increasing/decreasing shapes. Moreover, the log-logistic distribution can be easily employed in the presence of censored data which is very common in survival or reliability analysis. It is very difficult to use log-normal distribution in such cases. Consequently, the log-logistic distribution has seen increasing use recently. For example, Diekmann [12] used the log-logistic distribution as a model for event history analysis, Bennett [8] used it to model survival data, and Singh, Lee and George [41] used it to model censored survival data.

The log-logistic distribution is a derivative of the very popular logistic distribution. The logistic distribution was initially developed to model population growth by Verhulst [47, 48]. Verhulst [47] noticed that exponential distribution was used

in studies involving growth of biological populations such as cancer cells and bacteria. But when there is a limitation of food and space for a large population then its growth will not follow the exponential curve but rather the logistic curve. The use of the logistic distribution for economic and demographic purposes was very popular in the nineteenth century. The logistic distribution is also known by names such as growth function, autocatalytic curve and so on depending on its application. The name 'logistic 'was coined by Reed and Berkson [32]. Berkson [10] noted that under some circumstances if the dosage of a drug is expressed in proportion to its logarithm, the effect, as a percentage, follows the form of a more or less symmetric sigmoidal curve, the integral of a normal curve has been employed for the estimation of the potency of a drug. The logistic distribution has been used on human population by Pearl and Read [31], on fish by Jensen [21], on animals by Miller and Botkin [11], on bacteria and cells by Tan [43] and on tumor cells by Eisen [15], and on breast tumor by Moolgavkar [30].

In many practical situations theories and previous evidences or conditions suggest an expected ordering among the treatment effects. Example of such situations include severity of disease, drug dosage level etc. In our study, we consider a dose-response relationship where increasing dose levels lead to certain order relationships among the hazard rates. In particular we consider the situation where increasing dose leads to decreasing hazard rate and the situation where increasing dose leads to decreasing hazard rate up to a certain level and any more increase in the dose level results in an increase in hazard rate.

Such type of ordered alternatives have been considered in the past. Robertson [33] considers umbrella ordering to fit multinomial distributions to cell counts. Bartholomew [5, 6, 7] proposed a likelihood ratio test (LRT) for umbrella alternatives. Simpson and Margolin [40] considered umbrella ordering in dose-response relationships. Hayter and Liu [18] considered umbrella alternative for the normal distribution and Singh *et al.*[42] for the logistic distribution. More recently, the test of a null of no difference against that of a u-shaped alternative was developed for the exponential distribution location parameters by Abebe and Singh [1]. A nonparametric test for umbrella alternatives was given by Mack and Wolfe [28].

When it comes to simple (increasing or decreasing) ordering of parameters, Hayter and Liu [17] developed tests for the normal distribution location parameters while Tebbs and Bilder [44] developed such tests for comparing proportions. A nonparametric test for the simple order was given by Jonckheere [22] and Terpstra [45].

In Chapter 2 the log logistic distribution and its properties are discussed along with its comparison with the log-normal distribution. In Chapter 3, we develop an exact simultaneous testing procedure to compare the maximum hazard rates of k treatments (or doses) under a simple ordering restriction. From the union-intersection test statistic, the required critical constants are computed using a recursive algorithm and tables of critical constants are provided. A Monte Carlo simulation is performed to verify the results and compare the new test procedure with the nonparametric test due to Jonckheere and Terpstra. The fourth chapter deals with the valley ordering

restriction of the maximum hazard rates. A test statistic is introduced for testing for a valley ordering and a recursive algorithm is given for computing the critical constants. The power of this test is then compared to that of the Mack-Wolfe procedure for umbrella alternative. Chapter 5 presents a discussion of simultaneous confidence intervals for a certain set of contrasts.

Chapter 2

The Log-Logistic Distribution and its Properties

The log-logistic distribution is used in survival and reliability analysis as a model for survival times and is similar in shape to the log-normal distribution (see for example Kalbfleish and Prentice [24]). Its use is appealing because, like the lognormal distribution, its hazard rate function takes several different shapes depending on a value of a shape parameter. Recently it has seen increased use in hydrology to model stream flow and precipitation and also in economics as a simple model of the distribution of wealth or income due to its relationship to the generalized Pareto distribution. For more on the use of the log-logistic distribution in hydrology, the reader is referred to Shoukri *et al.*[39], Robson and Reed [35], or Ahmad *et al.*[3].

2.1 Density and Distribution Functions

A random variable T is said to follow the log-logistic distribution with scale parameter γ and shape parameter β, henceforth denoted by $T \sim L_L(\gamma, \beta)$, if its probability density function (pdf) is given by

$$f(t;\gamma,\beta) = \frac{(\beta/\gamma)(t/\gamma)^{\beta-1}}{[1+(t/\gamma)^{\beta}]^2}\ , \quad t>0\ , \tag{2.1}$$

where $\gamma > 0$ and $\beta > 0$. The corresponding cumulative function (cdf) is given by

$$F(t;\gamma,\beta) = \frac{t^{\beta}}{\gamma^{\beta} + t^{\beta}} \ . \tag{2.2}$$

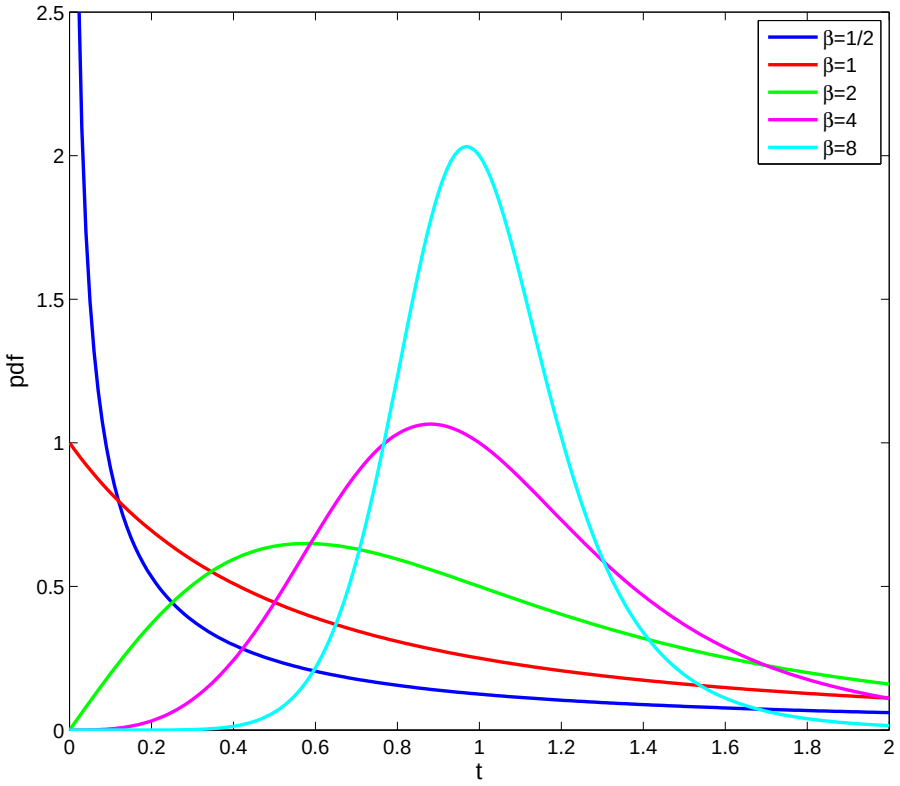

Figure 2.1: The pdf of $L_L(1,\beta)$ for different values of β

In Figure 2.1, the pdf of log-logistic distribution is shown for various values of β. This is a non negative distribution of random variables which takes various shapes. For $\beta \leq 1$ the pdf is a decreasing function, where as for $\beta > 1$, the pdf is an increasing decreasing function and there by has a peak. As value of β increases, the peak of the pdf shifts towards one and becomes more symmetric.

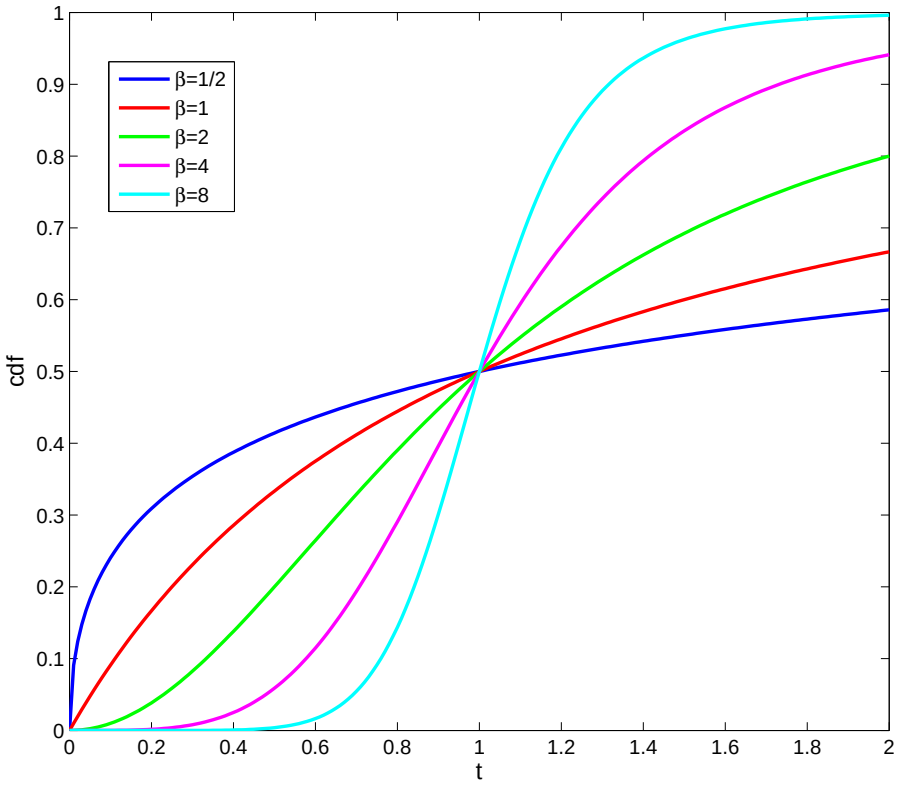

Figure 2.2: The cdf of $L_L(1, \beta)$ for different values of β

In Figure 2.2 the cdf of the log-logistic distribution is shown for various values of β.

2.2 Relationship with the Logistic Distribution

A random variable X is said to follow the logistic distribution with location μ and scale σ, written $X \sim L(\mu, \sigma)$, if its probability density function (pdf) is given by

$$g(x; \mu, \sigma) = \frac{\exp\left(\frac{-(x-\mu)}{\sigma}\right)}{\sigma\left[1 + \exp\left(\frac{-(x-\mu)}{\sigma}\right)\right]^2} , \quad |x| < \infty ,$$

where $-\infty < \mu < \infty$ and $\sigma > 0$. The cdf is given by

$$G(x; \mu, \sigma) = \left[1 + e^{\frac{-(x-\mu)}{\sigma}}\right]^{-1} .$$

The relationship between the logistic and the log-logistic distributions is analogous to that between the normal and the log-normal distributions. In particular, using a simple change-of-variable technique, one can show that $X \sim L(\mu, \sigma)$ if and only if $T \equiv \exp(X) \sim L_L(\exp(\mu), 1/\sigma)$ distribution.

In later sections, we will exploit this relationship between the two random variables when constructing simultaneous tests and confidence intervals.

2.3 Moments, Mode and Median

Let $T \sim L_L(\gamma, \beta)$. The kth moment of T for $k < \beta$ can be shown to be

$$ET^k = \gamma^k B(1 - k/\beta \ , \ 1 + k/\beta) = \gamma^k (k\pi/\beta) \csc(k\pi/\beta) ,$$

where $B(\cdot\ ,\ \cdot)$ is the beta function given by $B(a,b) = \int_0^1 s^{a-1}(1-s)^{b-1}ds$ and $\csc(\cdot)$ is the cosecant function. The kth moment is undefined if $k \geq \beta$. In particular, if $\beta > 1$, we can show that the mean and variance of T are (see Tadikamalla and Johnson [2])

$$ET = \gamma(\pi/\beta)\csc(\pi/\beta)$$

and

$$ET^2 - E^2T = \gamma^2[(2\pi/\beta)\csc(2\pi/\beta) - (\pi/\beta)^2\csc^2(\pi/\beta)]\ .$$

The $L_L(\gamma,\beta)$ distribution is unimodal with

$$\gamma\left(\frac{\beta-1}{\beta+1}\right)^{1/\beta}\ .$$

for $\beta > 1$. The mode is zero for $\beta \leq 1$. The median m of $L_L(\gamma,\beta)$ is found by solving $F(m;\gamma,\beta) = 0.5$ which gives $m = \gamma$.

2.4 Survival and Hazard Functions

The survival function, also known as the reliability function in engineering, is the characteristic of an explanatory variable that maps a set of events, usually associated with mortality or failure of some system onto time. It is the probability that the system will survive beyond a specified time. The term reliability function is common in engineering while the term survival function is used commonly in many fields, including human mortality. Lately, the log-logistic survival model is being used

increasingly. Conkin [14] used the log-logistic survival model as a model for hypobaric decompression sickness as a consequence of flying high in the atmosphere. According to Jones [23], even when the underlying survival function was best described by a negative power curve, a log-logistic model fits the data well and provides more versatility for fitting individual populations. In summarizing survival data, there are two important functions, namely survival function and hazard function. The actual survival time of an individual, represented by t is regarded as the realized value of the survival time T, which is a random variable that can take any non negative value.

The survival function of $T \sim L_L(\gamma, \beta)$ is

$$S(t) = P(T > t) = \left[1 + (t/\gamma)^{\beta}\right]^{-1}$$

.

This survival function is plotted in Figure 2.3. These decreasing survival functions cross each other at $t = 1$ as $S(t)$ becomes independent of β for this particular value of t.

The corresponding hazard function is the probability that an individual dies or an equipment fails at instantaneously at time t, provided that the individual survived up to time t. So the hazard function represents the death rate at a given point of time t. The hazard function at t is given by

$$\lambda(t) = \frac{f(t)}{S(t)} = (\beta/\gamma)(t/\gamma)^{\beta-1}\left[1 + (t/\gamma)^{\beta}\right]^{-1} .$$

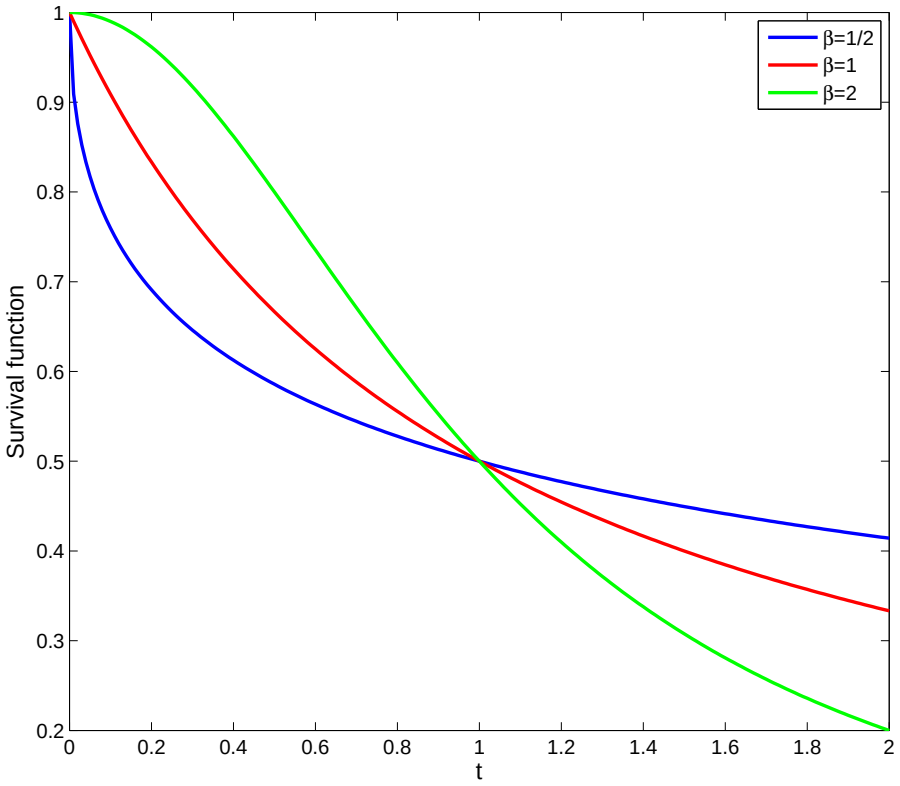

Figure 2.3: The pdf of $L_L(1, \beta)$ for different values of β

Figure 2.4 explains the hazard rate function for different values of β. Later on we assume $\beta > 1$ so that hazard rate function increases and decreases. We compare the worst case scenarios.

Thus besides the similarity of the log-logistic distribution to the log-normal distribution, one using the log-logistic distribution to model survival data has the advantage of knowing explicit forms of the hazard and survival functions. This is not the case with the log-normal distribution.

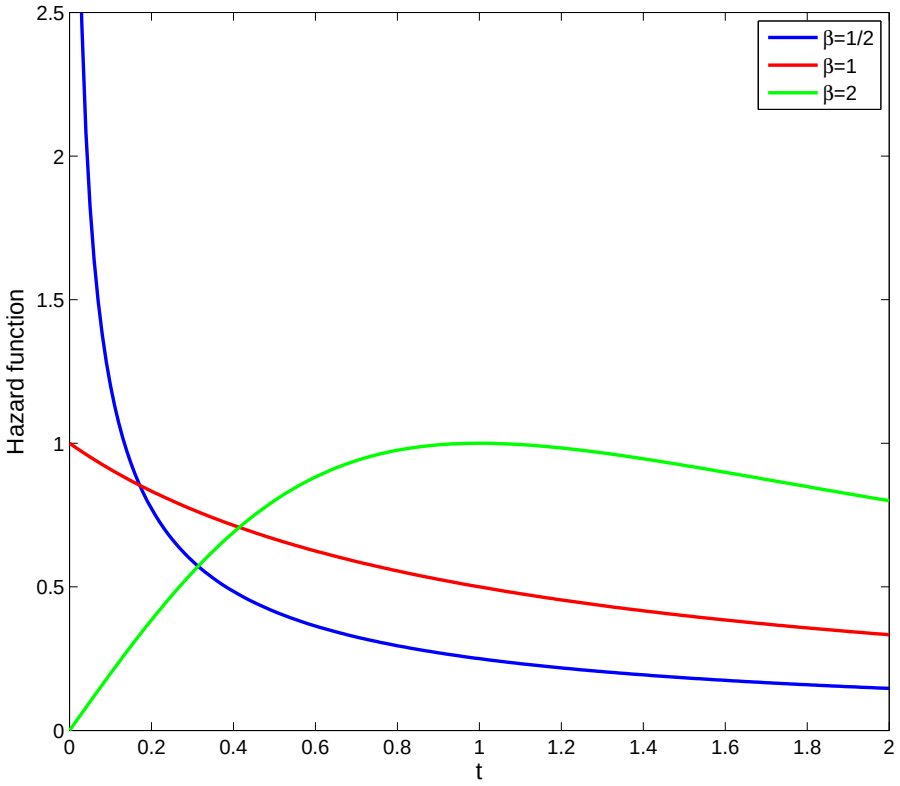

Figure 2.4: The cdf of $L_L(1, \beta)$ for different values of β

2.5 Comparison of Log-Normal and Log-Logistic Distributions

In this section we will compare several characteristics of the log-logistic and log-normal distributions.

Considering the shape of the distributions, the log-normal distribution, like the Weibull distribution, is a very flexible model that can empirically fit many types of failure time data. The log-normal distribution has two parameters, shape parameter

and a scale parameter. The log-logistic distribution is also defined using a scale and a shape parameter and can also take a variety of shapes like the log-normal distribution.

Since scale parameter will not affect the tails of either the log-normal or the log-logistic distributions, we can take the scale parameter to be unity without any loss of generality. The log-logistic distribution has heavier tail than log-normal distribution when the shape parameter of the log-logistic distribution is less than or equal to $\frac{4}{\sqrt{2\pi}}$ times the shape parameter of the log-normal distribution Yanagimoto [50].

According to Bennett [8], the log-logistic distribution is very similar in shape of the log-normal distribution but the log-logistic distribution is more suitable for survival analysis than the log-normal distribution when the data contain censored observations. Censored observations are quite common in survival analysis and the log-normal distribution cannot be used directly in the presence of censored data. With the Weibull distribution being monotonic increasing or decreasing, the log-logistic distribution is a popular choice.

The differences between the probability density functions of the log-normal and log-logistic distributions is illustrated in Fig. 2.5 for the same grid values. Similarly, the cdf of the log-normal and log-logistic distributions are shown in Fig. 2.6.

Some properties of the log-logistic and log-normal distributions are listed in Table 2.1.

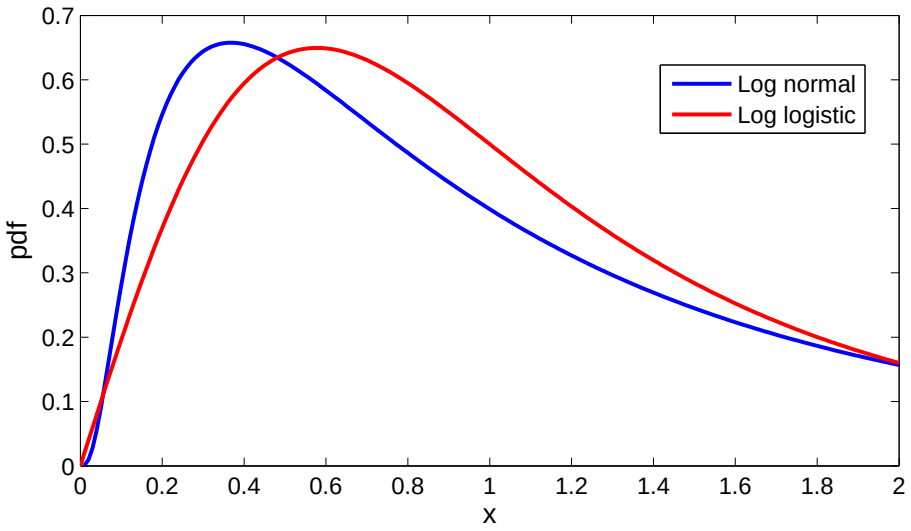

Figure 2.5: The pdf of log-logistic and log-normal distributions

2.6 Sampling Distribution of the Logistic Sample Median

We will be using the sampling distribution of the median of a random sample from the logistic distribution to construct tests and intervals relating to the scale parameter of the log-logistic distribution. The median is considered to be a good estimator of the location parameter of the logistic distribution since

1. the logistic distribution is symmetric,

2. the logistic distribution is long-tailed,

3. the median is easy to compute, and

4. the closed form expressions of the pdf and cdf of the sampling distribution of the logistic median are known.

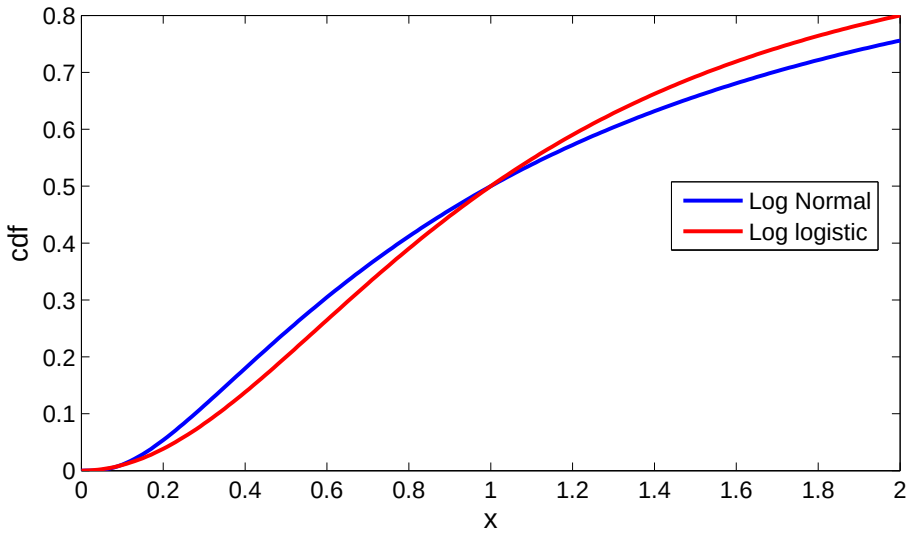

Figure 2.6: The cdf of log-logistic and log-normal distributions

Consider a random sample of size n from the logistic distribution with location parameter μ and scale parameter σ. Without loss of generality, let $m = (2n - 1)$. The probability density function (pdf) of logistic distribution $L(\mu, \sigma)$ in variable X is given by Eq. 2.3

$$f(x; \mu, \sigma) = \frac{a}{\sigma} \frac{e^{\frac{-a(x-\mu)}{\sigma}}}{1 + e^{\frac{-a(x-\mu)}{\sigma}}} \tag{2.3}$$

and corresponding cumulative function (cdf) is given by

$$F(x; \mu, \sigma) = \frac{1}{1 + e^{\frac{-a(x-\mu)}{\sigma}}} \tag{2.4}$$

where $a = \pi/\sqrt{3}$.

	Log-normal Distribution	Log-logistic Distribution
Support	$[0, \infty)$	$[0, \infty)$
Probability Density Function	$\frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(ln(x)-\mu)^2}{2\sigma^2}\right)$	$\frac{(\beta/\gamma)(t/\gamma)^{\beta-1}}{\left[1+(t/\gamma)^{\beta}\right]^2}$
Cumulative Density Function	$\frac{1}{2} + \frac{1}{2}\Phi\left(\frac{(ln(x)-\mu)}{\sigma\sqrt{2}}\right)$	$\frac{t^{\beta}}{\gamma^{\beta}+t^{\beta}}$
Mean	$e^{\mu+\sigma^2}$	$\frac{\alpha\pi/\beta}{sin(\pi/\beta)}$ if $\beta > 1$, else not defined
Median	e^{μ}	γ
Mode	$e^{(\mu-\sigma^2)}$	$\gamma\left(\frac{\beta-1}{\beta+1}\right)^{1/\beta}$ if $\beta > 1, 0$ otherwise

Table 2.1: comparison of log-logistic and log-normal distribution

Given a random sample $X_1, \ldots, X_n$ from the $L(0, 1)$ distribution, by ordering

$$X_{(1)} \leq X_{(2)} \leq \cdots X_{(n)}$$

we can get the median as $X_{(m)}$. Using the formula for the distribution of order statistics we can write the pdf of the sample median as

$$\psi_m(x) = \frac{\Gamma(2m)}{\Gamma^2(m)} F^{(m-1)}(x)[1 - F(x)]^{m-1} f(x)$$

that, after some simplification, can be written as

$$\psi_m(x) = \frac{\Gamma(2m)}{\Gamma^2(m)} a(e^{-ax})^m (1 + e^{-ax})^{-2m} \tag{2.5}$$

The cdf is given by

$$\Psi_m(x) = \frac{\Gamma(2m)}{\Gamma^2(m)} \sum_{j=0}^{m-1} \binom{m-1}{j} (2m-j-1)^{-1} (-1)^{m-1-j} (1+e^{-ax})^{j+1-2m} \,. \quad (2.6)$$

The probability density functions and cumulative distribution function of the logistic distribution with median as location parameter is shown in Fig. 2.7 and Fig. 2.8.

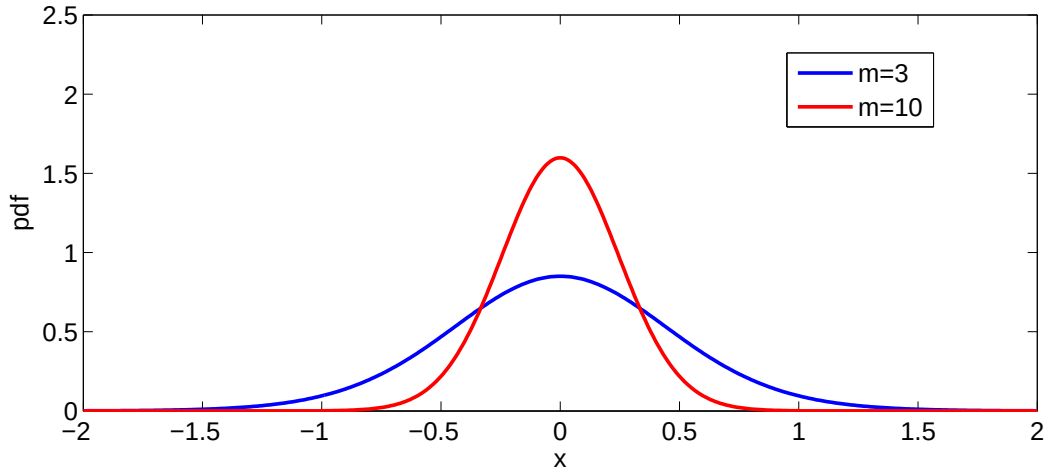

Figure 2.7: The pdf of logistic distributions with median as location parameter.

The asymptotic distribution of the standardized sample median is normal with mean zero and variance $[4nf^2(0)]^{-1}$ since $f^2(0) = a^2/16$ this variance reduces to $12/n\pi^2$ Ref. [27]

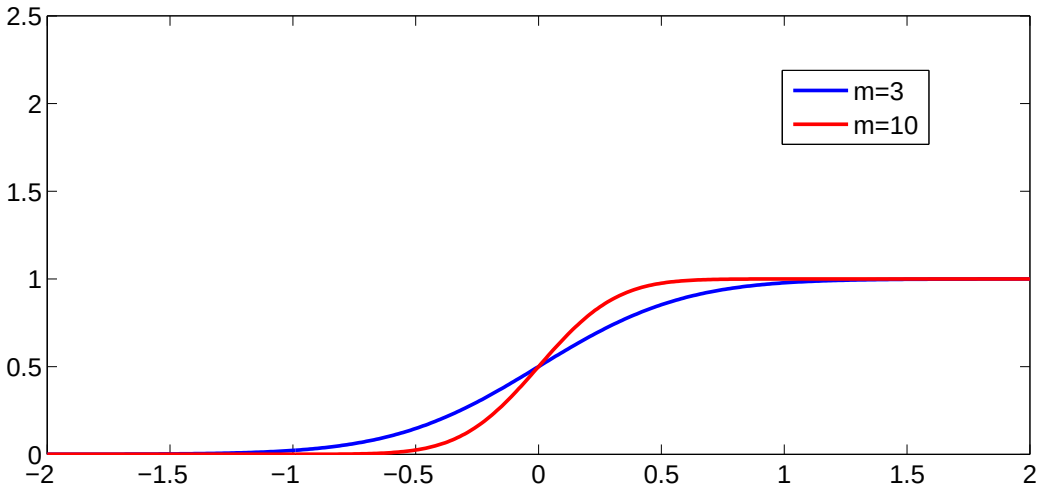

Figure 2.8: The cdf of logistic distributions with median as location parameter.

Chapter 3

Simple Ordering of Hazard Rates

Consider the situation where we have k dose levels of a certain treatment. Suppose that n subjects have been randomly allocated to each of the k dose levels. Let the survival times be given according to the log-linear model

$$\log T_{ij} = \mu_i + \epsilon_{ij} \ , \qquad 1 \leq j \leq n \ , \ 1 \leq i \leq k \ , \tag{3.1}$$

where ϵ_{ij} are independent $L(0, \sigma)$ random variables and μ_i is the unknown center of the ith treatment and the scale parameter σ is assumed to be known. As we have shown earlier, this condition is equivalent to assuming that $T_{ij} \sim L_L(\exp(\mu_i), 1/\sigma)$ for $1 \leq j \leq n \ , \ 1 \leq i \leq k$.

Let $\lambda_i(\cdot)$ be the hazard function of the ith dose level and let

$$\lambda_i^{max} = \sup_t \lambda_i(t) \ ,$$

$1 \leq i \leq k$. We are interested in finding out if increasing dose leads to a decrease of the worst case hazard rate of individuals. In particular, we would like to perform the simultaneous test of the null hypothesis

$$H_0 : \lambda_1^{max} = \cdots = \lambda_k^{max}$$

versus the simple ordering alternative

$$H_s : \lambda_1^{max} \geq \cdots \geq \lambda_k^{max}$$

with at least one strict inequality.

As the following theorems show, for the log-logistic distribution, performing this test is equivalent to making multiple comparisons of the scale parameters of the log-logistic distribution.

Theorem 3.1. *Let $\lambda(\cdot;\gamma,\beta^*)$ be the hazard function of $L_L(\gamma,\beta^*)$, for a given value of β^*. Then $\lambda(\cdot;\gamma,\beta^*)$ is a monotone decreasing function of γ.*

Proof. The proof follows since for any given t, we can write

$$\lambda(t;\gamma,\beta^*) = \frac{\beta^* t^{\beta^*-1}}{\gamma^{\beta^*} + t^{\beta^*}} .$$

□

Theorem 3.2. *Let $\lambda_{max}(\gamma,\beta^*)$ be worst case of the hazard function of $L_L(\gamma,\beta^*)$, for a given value of β^*. Then $\lambda_{max}(\gamma,\beta^*)$ is a monotone decreasing function of γ.*

Proof. For any given t, we can write log-logistic distribution as

$$\lambda(t;\gamma,\beta) = \frac{\left(\frac{\beta}{\gamma}\right)\left(\frac{t}{\gamma}\right)^{\beta-1}}{1+\left(\frac{t}{\gamma}\right)^{\beta}} .$$

If $\beta > 1$, we can find the value of t that maximizes $\lambda(t;\gamma,\beta)$ by solving

$$0 = \frac{\partial\lambda(t)}{\partial t} = \frac{\left(\frac{\beta}{\gamma}\right)(\beta-1)\left(\frac{t}{\gamma}\right)^{\beta-2}\left[1+\left(\frac{t}{\gamma}\right)^{\beta}\right] - \beta\left(\frac{t}{\alpha}\right)^{\beta-1}\left(\frac{\beta}{\gamma}\right)\left(\frac{t}{\gamma}\right)^{\beta-1}}{\left[1+\left(\frac{t}{\gamma}\right)^{\beta}\right]^2}.$$

This gives

$$0 = (\beta-1) + (\beta-1)\left(\frac{t}{\alpha}\right)^{\beta} - \beta\left(\frac{t}{\gamma}\right)^{\beta}$$

and solving for t gives

$$t_{max} = \gamma\left(\beta-1\right)^{\frac{1}{\beta}}$$

Substituting in $\lambda(t;\gamma,\beta)$ we get

$$\begin{aligned}
\lambda(t_{max};\gamma,\beta) &= \frac{\left(\frac{\beta}{\gamma}\right)\left(\frac{t_{max}}{\gamma}\right)^{\beta-1}}{1+\left(\frac{t_{max}}{\gamma}\right)^{\beta}} \\
&= \frac{\left(\frac{\beta}{\gamma}\right)\left[\frac{\gamma\left(\beta-1\right)^{\frac{1}{\beta}}}{\gamma}\right]^{\beta-1}}{1+\left[\frac{\gamma\left(\beta-1\right)^{\frac{1}{\beta}}}{\gamma}\right]^{\beta}} \\
&= \frac{\left(\beta-1\right)^{\beta-1/\beta}}{\gamma}
\end{aligned}$$

Thus for a fixed $\beta = \beta^*$

$$\lambda_{max} = \frac{k}{\gamma}$$

where k is a constant. □

One may consider a Cox proportional hazards model by imposing the restriction

$$\lambda_i(t) = \lambda_{0i}(t) \exp(\delta_i)$$

where δ_i is an unknown constant. This restriction produces a models that are in the so-called Lehmann class (Lehmann [25]). Although this is used often in the literature, it imposes a proportionality assumption that is not necessary in our case. Thus, this model will not be considered in this thesis.

In a similar manner to the proof of Theorem 3.1, one can show that the survival function of the $L_L(\gamma, \beta)$ distribution is a monotone increasing function of γ for a fixed value of β. So, a similar analysis may be performed for survival functions. In our case, it is true that decreasing $\lambda(\cdot)$ is equivalent to increasing $S(\cdot)$ and thus the inferences that we develop for $\lambda(\cdot)$ may be used directly for $S(\cdot)$.

Let us now consider the test of H_0 versus the simple ordering H_s. By Theorem 3.2, the hypotheses can be written down as

$$H_0 : \gamma_1 = \cdots = \gamma_k$$

versus

$$H_s : \gamma_1 \leq \cdots \leq \gamma_k$$

with at least one strict inequality. Here $\gamma_i = \exp(\mu_i)$ and μ_i is given in Equation 3.1 for $1 \leq i \leq k$. Since $\exp(\cdot)$ is a monotone increasing function, this is equivalent to testing

$$H_0 : \mu_1 = \cdots = \mu_k$$

versus

$$H_s : \mu_1 \leq \cdots \leq \mu_k$$

with at least one strict inequality. Recall that this corresponds to the case where increasing dose levels are expected to give decreasing worst case hazard rates.

Thus we have reframed the problem as a multiple comparison of logistic location parameters. This has been considered in Singh *et al.*[42]. Similar comparisons for the case where the underlying distribution is normal has been considered in the literature (see Robertson *et al.*[34]). Williams [49] gave a test based on the maximum likelihood estimators of the normal means $\mu_1, \ldots, \mu_k$ under the restriction given by $H_s : \mu_1 \leq \cdots \leq \mu_k$. Hayter [16] gave a studentized range test for the normal case while Hayter and Liu [17] gave a recursive computational procedure for computing the exact critical values of this studentized range test.

We reiterate that if we know any information specifically about the data from theory or from the previous experience that the hazard rates follow a certain order prior to collecting the data, then it is very important to incorporate this information

into our analysis. If one neglects these factors then the potential consequence will be improper interpretation of results.

3.1 Test Statistic

Consider the model given by Equation 3.1. Let $Y_i = \text{median}\{\log T_{i1}, \ldots, \log T_{in}\}$ for $1 \leq i \leq k$. For H_0 versus H_s given above, the statistic for the union-intersection test (Roy [36]; Sen [38]) is given by

$$W_{k,m} = \min_{1 \leq i < j \leq k} (Y_i - Y_j)$$

and the null is rejected in favor of the alternative for small values of $W_{k,m}$. In particular, H_0 is rejected at with a familywise error rate of α if and only if

$$W_{k,m} \leq -c_{k,m,\alpha}$$

where $c_{k,m,\alpha}$ is chosen such that

$$P_0(W_{k,m} \leq -c_{k,m,\alpha}) = \alpha \ .$$

Most of our remaining discussion of this chapter considers the computation of the critical constants $c_{k,m,\alpha}$.

3.2 Computing the Critical Constants

Since the variance of $L(\mu, \sigma)$ is $\pi^2\sigma^2/3$, we will re-parameterize it using $\log T/a$ in place of $\log T$ and then taking, without loss of generality, the known $\sigma = 1$, where $a = \pi/\sqrt{3}$. Under the null, $W_{k,m}$ is invariant to log-logistic scale transformations, we can take $\gamma = 1$ (or equivalently $\mu = 0$ for the logistic distribution). This means that, under the null, T_{ij} are iid $L_L(1, a)$ random variables, $1 \leq i \leq k,\ 1 \leq j \leq n$. Consequently, under the null, the medians $Y_1, \ldots, Y_k$ are iid random variables.

Let us take $n = 2m-1$. The even sample size case is analogous. From the theory of order statistics (David and Nagaraja [4] or see the discussion above), the pdf and cdf of $Y_1, \ldots, Y_k$ are

$$\psi_m(x) = \frac{\Gamma(2m)}{\Gamma^2(m)} a(e^{-ax})^m(1 + e^{-ax})^{-2m} \tag{3.2}$$

and

$$\Psi_m(x) = \frac{\Gamma(2m)}{\Gamma^2(m)} \sum_{j=0}^{m-1} \binom{m-1}{j} (2m - j - 1)^{-1}(-1)^{m-1-j}(1 + e^{-ax})^{j+1-2m} \tag{3.3}$$

respectively.

The following theorem gives an integral equation that can be used in computing the desired critical constants.

Theorem 3.3. *The critical constant $c = c_{k,m,\alpha}$ is the solution of*

$$1 - \alpha \;=\; \int_{-\infty}^{\infty} \psi_m(y) A_k(y, c) dy \;,$$

where $A_k(y, c) = P_0\left(\min_{1 \le i < j \le k}(Y_i - Y_j) \ge -c | Y_k = y\right)$.

Proof. From the definition of the critical constants we have that $P_0(W_{k,m} \ge -c_{k,m,\alpha}) = 1 - \alpha$. Now with $c = c_{k,m,\alpha}$, we have

$$1 - \alpha = P_0(W_{k.m} \ge -c) = P_0\left\{\min_{1 \le i < j \le k}(Y_i - Y_j) \;\ge -c\right\} \;.$$

Conditioning on Y_k gives

$$1 - \alpha = \int_{-\infty}^{\infty} P_0\left(\min_{1 \le i < j \le k}(Y_i - Y_j) \ge -c \Big| Y_k = y\right) \psi_m(y)\; dy$$

which is nothing but

$$1 - \alpha = \int_{-\infty}^{\infty} \psi_m(y) A_k(y, c) dy \;.$$

□

Note that the computation of the critical constants requires finding the numerical solution of the k-variate integral equation $1 - \alpha = \int_{-\infty}^{\infty} \psi_m(y) A_k(y, c) dy$. This is not feasible for any practical value of k. However, the advantage of the representation

given in the above theorem is that instead of solving a k-variate integral equation, we can now recursively compute A_k that can then be solved numerically to give the critical constants. The following theorem gives the recursive algorithm that is used to compute A_k.

Theorem 3.4. *Let*

$$\begin{aligned} A_0(y,c) &= 1 \\ A_t(y,c) &= P_0\left(\min_{1\le i<j\le t+1}(Y_i - Y_j) \ge -c | Y_{t+1} = y\right) \ , \quad t \ge 1 \ . \end{aligned}$$

Then, for $t \ge 2$, $A_t(y;c)$ can be recursively computed as

$$A_t(y,c) = \int_y^\infty A_{t-1}(x,c)\psi_m(x)dx \ + \ A_{t-1}(y,c)[\Psi_m(y) - \Psi_m(y-c)] \ .$$

Proof. For $t = 1$

$$\begin{aligned} A_1(y,c) &= P_0\left(\min_{1\le i<j\le 2}(Y_i - Y_j) \ge -c | Y_2 = y\right) \\ &= P_0\left(Y_1 - Y_2 \ge -c | Y_2 = y\right) \\ &= P_0(Y_1 \ge y - c) \\ &= 1 - \Psi_m(y) + \Psi_m(y) - \Psi_m(y-c) \\ &= \int_y^\infty A_0(x,c)\psi_m(x)dx + A_0(y,c)[\Phi_m(y) - \Phi_m(y-c)] \ . \end{aligned}$$

For $t \geq 2$

$$
\begin{aligned}
A_t(y,c) &= P_0\left(\min_{1\leq i<j\leq t}(Y_i - Y_j) \geq -c\ ,\ \min_{1\leq i\leq t}(Y_i - y) \geq -c\right) \\
&= \int_{y-c}^{\infty} P_0\left(\min_{1\leq i<j\leq t}(Y_i - Y_j) \geq -c\ ,\ \min_{1\leq i\leq t}(Y_i - y) \geq -c \mid Y_t = x\right)\psi_m(x)dx \\
&= \int_{y-c}^{\infty} P_0\left(\min_{1\leq i<j\leq t-1}(Y_i - Y_j) \geq -c\ ,\ \min_{1\leq i\leq t-1} Y_i \geq \max(y-c, x-c)\right)\psi_m(x)dx \\
&= \int_{y-c}^{y} P_0\left(\min_{1\leq i<j\leq t-1}(Y_i - Y_j) \geq -c\ ,\ \min_{1\leq i\leq t-1} Y_i \geq \max(y-c, x-c)\right)\psi_m(x)dx\ + \\
&\qquad \int_{y}^{\infty} P_0\left(\min_{1\leq i<j\leq t-1}(Y_i - Y_j) \geq -c\ ,\ \min_{1\leq i\leq t-1} Y_i \geq \max(y-c, x-c)\right)\psi_m(x)dx \\
&= \int_{y-c}^{y} P_0\left(\min_{1\leq i<j\leq t-1}(Y_i - Y_j) \geq -c\ ,\ \min_{1\leq i\leq t-1}(Y_i - y) \geq -c\right)\psi_m(x)dx\ + \\
&\qquad \int_{y}^{\infty} P_0\left(\min_{1\leq i<j\leq t-1}(Y_i - Y_j) \geq -c\ ,\ \min_{1\leq i\leq t-1}(Y_i - x) \geq -c\right)\psi_m(x)dx \\
&= \int_{y-c}^{y} P_0\left(\min_{1\leq i<j\leq t}(Y_i - Y_j) \geq -c | Y_t = y\right)\psi_m(x)dx\ + \\
&\qquad \int_{y}^{\infty} P_0\left(\min_{1\leq i<j\leq t}(Y_i - Y_j) \geq -c | Y_t = x)\right)\psi_m(x)dx \\
&= \int_{y}^{\infty} A_{t-1}(x,c)\psi_m(x)dx\ +\ A_{t-1}(y,c)[\Psi_m(y) - \Psi_m(y-c)]
\end{aligned}
$$

□

The theorem above shows that A_k can be represented as

$$A_k(y,c) = \int_{y}^{\infty} A_{k-1}(x,c)\psi_m(x)dx\ +\ A_{k-1}(y,c)[\Psi_m(y) - \Psi_m(y-c)]$$

setting up a recursion of univariate integrals. After computing A_k using this recursive algorithm, the integral equation

$$1 - \alpha \;=\; \int_{-\infty}^{\infty} \psi_m(y) A_k(y, c) dy,$$

can easily be solved using Gauss quadrature. The Matlab code that is used is found in the appendix. The computed critical constants for $k = 3(1)10$ and $m = 3(1)10$ for $\alpha = 0.01, 0.05$ are given in Table 3.1.

Values of the critical constants $c_{k,m,\alpha}$, decreases as the value of the location of the median increases and $c_{k,m,\alpha}$ increases as the number of groups increases. These changes ensure that the test maintains level α.

3.3 Monte Carlo Simulation

3.3.1 Nominal Level Simulation

After finding the values of the critical constants for different combinations of group sizes and sample sizes, we performed a simulation study to evaluate the correctness, in terms of retaining the nominal α, of the computed critical constants. This is done by repeatedly generating $k \times (2m-1)$ random variates from the $L_L(1,1)$ distribution, splitting these into k equal parts o f $2(m-1)$ and computing the hazard rate function. We then compute the test statistic $W_{k,m}$ and compute the proportion of times H_0 is rejected. 10,000 iterations are performed and these values are given

$\alpha = 0.01$								
	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10
m=3	1.9203	2.0877	2.2075	2.3007	2.3770	2.4414	2.4972	2.5463
m=4	1.6163	1.7542	1.8525	1.9288	1.9910	2.0435	2.0888	2.1287
m=5	1.4211	1.5407	1.6258	1.6916	1.7453	1.7905	1.8295	1.8638
m=6	1.2826	1.3895	1.4655	1.5242	1.5720	1.6122	1.6468	1.6773
m=7	1.1778	1.2754	1.3446	1.3980	1.4414	1.4780	1.5095	1.5371
m=8	1.0951	1.1853	1.2493	1.2986	1.3387	1.3724	1.4014	1.4269
m=9	1.0276	1.1119	1.1716	1.2177	1.2551	1.2865	1.3136	1.3373
m=10	0.9712	1.0507	1.1069	1.1502	1.1854	1.2149	1.2404	1.2627
$\alpha = 0.05$								
	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10
m=3	1.4516	1.6352	1.7650	1.8653	1.9468	2.0154	2.0745	2.1265
m=4	1.2285	1.3818	1.4897	1.5728	1.6402	1.6968	1.7455	1.7882
m=5	1.0837	1.2178	1.3120	1.3843	1.4429	1.4921	1.5343	1.5713
m=6	0.9803	1.1009	1.1854	1.2503	1.3028	1.3467	1.3845	1.4176
m=7	0.9016	1.0121	1.0894	1.1487	1.1967	1.2368	1.2712	1.3014
m=8	0.8393	0.9418	1.0135	1.0684	1.1128	1.1499	1.1818	1.2097
m=9	0.7883	0.8843	0.9515	1.0029	1.0444	1.0791	1.1089	1.1349
m=10	0.7456	0.8362	0.8996	0.9480	0.9871	1.0199	1.0479	1.0724

Table 3.1: Table of critical values $c_{k,m,\alpha}$ for simple ordered alternatives

in Table 3.2. It can be seen that the computed constants maintain give Type I error rates that is very close to the nominal levels.

k=h, $\alpha = 0.01$								
	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10
m=3	0.0109	0.0114	0.0101	0.0076	0.0115	0.0113	0.0075	0.0095
m=4	0.0100	0.0100	0.0105	0.0102	0.0122	0.0092	0.0095	0.0086
m=5	0.0120	0.0097	0.0095	0.0100	0.0105	0.0104	0.0096	0.0083
m=6	0.0093	0.0117	0.0094	0.0090	0.0096	0.0103	0.0112	0.0097
m=7	0.0087	0.0113	0.0097	0.0106	0.0105	0.0096	0.0098	0.0095
m=8	0.0102	0.0110	0.0105	0.0098	0.0088	0.0090	0.0096	0.0099
m=9	0.0109	0.0112	0.0090	0.0100	0.0097	0.0099	0.0084	0.0105
m=10	0.0096	0.0097	0.0109	0.0104	0.0102	0.0114	0.0097	0.0101
k=h, $\alpha = 0.05$								
	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10
m=3	0.0468	0.0509	0.0506	0.0510	0.0483	0.0540	0.0466	0.0529
m=4	0.0483	0.0499	0.0482	0.0551	0.0508	0.0484	0.0523	0.0469
m=5	0.0500	0.0547	0.0539	0.0554	0.0487	0.0502	0.0464	0.0498
m=6	0.0506	0.0471	0.0488	0.0495	0.0518	0.0497	0.0497	0.0506
m=7	0.0510	0.0507	0.0495	0.0524	0.0485	0.0518	0.0468	0.0475
m=8	0.0485	0.0517	0.0489	0.0494	0.0493	0.0527	0.0449	0.0531
m=9	0.0481	0.0516	0.0499	0.0494	0.0471	0.0481	0.0497	0.0502
m=10	0.0512	0.0501	0.0492	0.0477	0.0520	0.0534	0.0478	0.0518

Table 3.2: Table of significance level α for simple ordered alternatives

3.3.2 Power Simulation versus the Jonckheere-Terpstra Test

The corresponding nonparametric test when the treatment can be labeled a priori in such a way that the experimenter expects any deviation from H_0 to be in the particular direction associated with H_s is the Jonckheere and Terpstra of Jonckheere [22] and Terpstra [45], that can be found in Hollander and Wolfe [19]. As above, we

label the treatments so that the ordered alternatives are appropriate. The labeling must be done in accordance with the nature of the experimental design and not the data observed.

After labeling the treatments so that they are in the expected order in the alternative Jonckheere-Terpstra statistic J is then the sum of the $k(k-1)/2$ Mann-Whitney counts U_{uv} given by

$$U_{uv} = \sum_{i=1}^{n_u} \sum_{j=1}^{n_v} \phi(\log T_{iu}, \log T_{jv}), \quad 1 \leq u < v \leq k$$

where $\phi(a, b) = 1$ if $a < b$, 0 otherwise. That is, U_{uv} is the number of sample u before sample v precedences. The Jonckheere-Terstra statistic J, is then sum of these $k(k-1)/2$ Mann-Whitney counts,

$$J = \sum_{u=1}^{v-1} \sum_{v=2}^{k} U_{uv}$$

For testing the null hypothesis

$$H_0 : \lambda_1^{max} = \cdots = \lambda_k^{max}$$

versus the simple ordering alternative

$$H_s : \lambda_1^{max} \geq \cdots \geq \lambda_k^{max}$$

with at least one strict inequality at α level of significance, the JT test rejects H_0 if $J \geq j_\alpha$. The value of j_α is taken from table A.13 of Hollander and Wolfe [19] chosen to make Type I error probability equal to α.

n	α	j_α	γ_1	γ_2	γ_3	c	$P(W < -c)$	$P(J \geq j_\alpha)$
3	0.0369	22	1	1	1.65	1.9869	0.0901	0.0657
5	0.0456	54	1	1	1.28	1.4808	0.0781	0.0779
7	0.0471	100	1	1	1.65	1.2443	0.9181	0.2046

Table 3.3: Power comparison against JT for $k = 3$

We performed a Monte Carlo simulation study comparing the performance of the test procedure developed in this thesis with that of the Jonckheere-Terpstra test. We performed 10000 repetitions where data are generated from the log-logistic distribution with the given parameters. The worst case hazard rates and the proportion of times that the null is rejected are computed for both tests. The results are given in Table 3.3.

It is evident from Table 3.3 that our test W rejects the null hypothesis more often than the JT test J. Thus for these particular alternative hypothesis configurations, our test proves to be more powerful than the Jonckheere-Terstra test in detecting decreasing worst case hazard rate patterns.

Chapter 4

Valley Ordering of Hazard Functions

4.1 Dose Response experiment

In toxicology and drug development studies, several increasing dose levels of a substance are usually compared with the zero-dose control to investigate the effect of the substance. According to Chen [13] for this purpose, a dose-response experiment is often conducted in a one way layout in which the doses of the substance under consideration are administered to separate groups of subjects. There are different concerns in these studies. In toxicological studies, the major concern is the safety of the toxin under consideration. Therefore, the goal is to estimate the highest dose that shows no significance difference from the zero-dose control, which is generally called the no statistical significance of trend (NOSTASOT; Tukey, Cimenra and Heyse, [46]) or no observed adverse event level (NOAEL; Ryan, [26]) dose. In drug developmental studies, however the primary interest is identifying the lowest dose level producing a desirable effect over that of the zero-dose control, which is commonly referred as the minimum effect dose(MED; Ruberg, [37]).

The approach in toxicological studies is to identify the NOSTASOT or NOAEL dose and apply appropriate safety factors to reach a safe dose level.

The regression based quantitative approach is not commonly used in drug development studies it is impractical to specify such an amount of increase in effect over

the zero-dose control so that the corresponding dose level causes a desirable effect. In dose-response experiments one cannot do extrapolation. So a test based approach to identifying the MED in drug development studies is very crucial. In such experiments it is expected that increasing dose level will produce stronger or at least equal treatment effects. However in many situations due to the toxic effects at high doses, an ordering in the treatment effects is anticipated, monotonically increasing up to a point then monotonically decrease. This up and down ordering of the treatment is called as umbrella alternate hypothesis (Mack and Wolfe, [28]) and the turning point is called the peak or valley of the umbrella.

To identify the MED for normal distribution several methods have been developed. But many a times the data do not fallow normal distribution. Then for such scenarios there are several nonparametric methods have been developed.

In this thesis we are interested in dose-response experiments where the hazard rate is generally decreasing with increasing dose level until MED is reached and then any more increase in dose starts to have an adverse effect increasing the hazard rate. In a sense, we are testing if this inverted umbrella pattern is satisfied by the worst cases of the hazard functions, that is, we wish to test

$$H_0 : \lambda_1^{max} = \cdots = \lambda_k^{max}$$

versus

$$H_v : \lambda_1^{max} \geq \cdots \geq \lambda_h^{max} \leq \cdots \leq \lambda_k^{max}$$

with at least one strict inequality, where $1 \leq h \leq k$ is a known valley point. In other words, the investigator has an idea of the ordering and what the most effective dose will be, but wants to test if it is statistically significant with a specified familywise error rate of α.

4.2 Test Statistic

Once again, we will assume the log-linear model given by Equation 3.1. Let $Y_i = \text{median}\{\log T_{i1}, \ldots, \log T_{in}\}$ for $1 \leq i \leq k$. By Theorem 3.1, for testing H_0 versus H_v given above, the statistic for the union-intersection test (Roy [36]; Sen [38]) is given by

$$V_{k,h,m} = \min \left\{ \min_{1 \leq i < j \leq h} (Y_i - Y_j) \ , \ \min_{h \leq j < i \leq k} (Y_i - Y_j) \right\}$$

where small values of $V_{k,h,m}$ indicate the alternative. Thus H_0 is rejected with a familywise error rate of α if and only if

$$V_{k,h,m} \leq -q_{k,h,m,\alpha}$$

where the critical constant $q_{k,h,m,\alpha}$ is the solution of

$$P_0(V_{k,h,m} \leq -q_{k,h,m,\alpha}) = \alpha \ .$$

Once again, solving for $q_{k,h,m,\alpha}$ requires finding the solution of a multiple integral equation. Since this is not practically feasible, like the simple ordering case, we will find a recursive method for obtaining the critical constants.

4.3 Computing the Critical Constants

As in the simple ordering case, assume that $n = 2m - 1$. Then under the null hypothesis, $Y_1, \ldots, Y_k$ are iid random variables with density ψ_m and distribution Ψ_m given in Equations 3.2 and 3.3, respectively.

The following theorem gives the integral equation that will be solved to give the critical constants $q_{k,h,m,\alpha}$.

Theorem 4.1. *The critical constant $q = q_{k,h,m,\alpha}$ is the solution of*

$$1 - \alpha \;=\; \int_{-\infty}^{\infty} \psi_m(y) A_{h-1}(y,q) A_{k-h}(y,q) dy \;,$$

where $A_t(y,q) = P_0\left(\min_{1\le i<j\le t+1}(Y_i - Y_j) \ge -q | Y_{t+1} = y\right)$ for $t \ge 1$ and $A_0(y,q) = 1$.

Proof. If $h = 1$ or $h = k$, then we have simple ordering. The result follows by Theorem 3.3 since $A_0(y,q) = 1$. If $1 < h < k$, then from the definition of the critical constants we have that $P_0(V_{k,h,m} \ge -q_{k,h,m,\alpha}) = 1 - \alpha$. Thus

$$1 - \alpha = P_0\left(\min\left\{\min_{1\le i<j\le h}(Y_i - Y_j)\;,\;\min_{h\le j<i\le k}(Y_i - Y_j)\right\} \ge -q\right)\;.$$

Conditioning on Y_h gives

$$
\begin{aligned}
1-\alpha &= \int_{-\infty}^{\infty} P_0\left(\min\left\{\min_{1\le i<j\le h}(Y_i-Y_j)\ ,\ \min_{h\le j<i\le k}(Y_i-Y_j)\right\}\ge -q\Big|Y_h=y\right)\ \psi_m(y)\ dy \\
&= \int_{-\infty}^{\infty} P_0\Bigg(\min\Bigg\{\min_{1\le i<j\le h-1}(Y_i-Y_j)\ ,\ \min_{1\le i\le h-1}(Y_i-y)\ , \\
&\qquad\qquad \min_{h+1\le j<i\le k}(Y_i-Y_j)\ ,\ \min_{h+1\le i\le k}(Y_i-y)\Bigg\}\ge -q\Bigg)\ \psi_m(y)\ dy \\
&= \int_{-\infty}^{\infty} P_0\left(\min\left\{\min_{1\le i<j\le h-1}(Y_i-Y_j)\ ,\ \min_{1\le i\le h-1}(Y_i-y)\right\}\ge -q\right)\times \\
&\qquad\qquad P_0\left(\min\left\{\min_{h+1\le j<i\le k}(Y_i-Y_j)\ ,\ \min_{h+1\le i\le k}(Y_i-y)\right\}\ge -q\right)\ \psi_m(y)\ dy \\
&= \int_{-\infty}^{\infty} P_0\left(\min_{1\le i<j\le h}(Y_i-Y_j)\ge -q\Big|Y_h=y\right)\times \\
&\qquad\qquad P_0\left(\min_{h\le j<i\le k}(Y_i-Y_j)\ge -q\Big|Y_h=y\right)\ \psi_m(y)\ dy\ .
\end{aligned}
$$

Since the probabilities are taken under the null, $Y_1,\ldots,Y_k$ are iid. Thus making a switch of variables in the second probability gives

$$
\begin{aligned}
1-\alpha &= \int_{-\infty}^{\infty} P_0\left(\min_{1\le i<j\le h}(Y_i-Y_j)\ge -q\Big|Y_h=y\right)\times \\
&\qquad\qquad P_0\left(\min_{1\le i<j\le k-h+1}(Y_i-Y_j)\ge -q\Big|Y_{k-h+1}=y\right)\ \psi_m(y)\ dy \\
&= \int_{-\infty}^{\infty} \psi_m(y)A_{h-1}(y,q)A_{k-h}(y,q)\ dy\ .
\end{aligned}
$$

□

Although, Theorem 4.1 provides us a representation that can be solved to find the critical constants $q_{k,h,m,\alpha}$, it is still not numerically feasible, especially for large

k, since it is a k-variate integral equation. A direct numerical solution would require us to generate grid points on $\Re^k$ and then perform numerical integration which then has to be solved for $q_{k,h,m,\alpha}$.

Fortunately, by Theorem 3.4, the values of A_t for $0 \leq t \leq k$ can be found recursively requiring only single dimensional numerical integration. Thus the eventual solution will require the solution of a single dimensional integral equation.

Using Matlab 7.5, we performed a trapezoidal numerical integration is used to calculate the critical constants on a grid from -8 to 8. As in Hayter and Liu [16], the grids are selected such that the grid points are denser near 0 and more sparse near the tails. The critical values $q_{k,h,m,\alpha}$ are given for $k = 3(1)10$, $m = 3(1)10$, $h = 2(1)5$ in Table 4.1 and Table 4.2, for $\alpha = 0.05$ and $\alpha = 0.01$, respectively.

Values of the critical constants $q_{k,h,m,\alpha}$ follows the same pattern as $c_{k,m,\alpha}$ for Simple ordering. $q_{k,h,m,\alpha}$ decreases as the value of the location of the median increases and increases as the number of groups increases. These changes ensure that the test maintains level α.

4.4 Monte Carlo Simulation Studies

4.4.1 Nominal FWER Simulation

Just as in the simple ordering case, after finding the values of the critical constants for different combinations of group sizes and sample sizes, we performed a simulation study to evaluate the correctness, in terms of retaining the nominal familywise error rate α, of the computed critical constants. This is done by repeatedly

h=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10
m=3	1.3314	1.5270	1.6749	1.7894	1.8818	1.9587	2.0244	2.0816
m=4	1.1280	1.2915	1.4147	1.5098	1.5863	1.6499	1.7041	1.7512
m=5	0.9957	1.1388	1.2465	1.3294	1.3961	1.4513	1.4984	1.5392
m=6	0.9010	1.0298	1.1265	1.2010	1.2607	1.3102	1.3523	1.3889
m=7	0.8290	0.9470	1.0356	1.1037	1.1583	1.2035	1.2419	1.2752
m=8	0.7718	0.8814	0.9635	1.0267	1.0772	1.1191	1.1546	1.1855
m=9	0.7250	0.8828	0.9047	0.9638	1.0111	1.0502	1.0835	1.1123
m=10	0.6858	0.7828	0.8554	0.9112	0.9558	0.9927	1.0240	1.0511
h=3		k=4	k=5	k=6	k=7	k=8	k=9	k=10
m=3		1.5270	1.6391	1.7447	1.8365	1.9155	1.9839	2.0440
m=4		1.2915	1.3846	1.4724	1.5485	1.6139	1.6705	1.7200
m=5		1.1388	1.2201	1.2967	1.3630	1.4199	1.4691	1.5121
m=6		1.0298	1.1028	1.1715	1.2310	1.2820	1.3261	1.3645
m=7		0.9470	1.0138	1.0767	1.3110	1.1776	1.2179	1.2530
m=8		0.8814	0.9433	1.0016	1.0520	1.0951	1.1324	1.1649
m=9		0.8274	0.8857	0.9403	0.9874	1.0278	1.0626	1.0930
m=10		0.7828	0.8375	0.8890	0.9335	0.9715	1.0043	1.0329
h=4			k=5	k=6	k=7	k=8	k=9	k=10
m=3			1.6749	1.7447	1.8201	1.8919	1.9575	2.0167
m=4			1.4147	1.4724	1.5348	1.5942	1.6484	1.6973
m=5			1.2465	1.2967	1.3510	1.4026	1.4498	1.4923
m=6			1.1265	1.1715	1.2202	1.2665	1.3087	1.3467
m=7			1.0336	1.0767	1.1211	1.1634	1.2020	1.2367
m=8			0.9635	1.0016	1.0428	1.0819	1.1176	1.1497
m=9			0.9047	0.9403	0.9788	1.0154	1.0488	1.0788
m=10			0.8554	0.8890	0.9253	0.9598	0.9913	1.0196
h=5				k=6	k=7	k=8	k=9	k=10
m=3				1.7894	1.8365	1.8919	1.9482	2.0023
m=4				1.5098	1.5485	1.5942	1.6407	1.6853
m=5				1.3294	1.3630	1.4026	1.4430	1.4818
m=6				1.2010	1.2310	1.2665	1.3026	1.3373
m=7				1.1037	1.1311	1.1634	1.1964	1.2280
m=8				1.0267	1.0520	1.0819	1.1125	1.1417
m=9				0.9638	0.9874	1.0154	1.0440	1.0713
m=10				0.9112	0.9335	0.9598	0.9867	1.0125

Table 4.1: Values of $q_{k,h,m,.05}$

h=2		k=4	k=5	k=6	k=7	k=8	k=9	k=10
m=3	1.8154	1.9895	2.1241	2.2299	2.3158	2.3878	2.4496	2.5036
m=4	1.5297	1.6733	1.7840	1.8707	1.9410	1.9998	2.0501	2.0940
m=5	1.3459	1.4706	1.5665	1.6415	1.7022	1.7528	1.7961	1.8339
m=6	1.2153	1.3268	1.4126	1.4794	1.5335	1.5786	1.6172	1.6507
m=7	1.1164	1.2182	1.2964	1.3573	1.4065	1.4475	1.4825	1.5130
m=8	1.0383	1.1325	1.2047	1.2610	1.3064	1.3443	1.3766	1.4046
m=9	0.9745	1.0625	1.1300	1.1826	1.2250	1.2603	1.2904	1.3166
m=10	0.9211	1.0041	1.0677	1.1171	1.1571	1.1903	1.2186	1.2432
h=3		k=4	k=5	k=6	k=7	k=8	k=9	k=10
m=3		1.9895	2.0898	2.1864	2.2716	2.3455	2.4100	2.4667
m=4		1.6733	1.7556	1.8348	1.9046	1.9650	2.0176	2.0638
m=5		1.4706	1.5418	1.6103	1.6706	1.7227	1.7680	1.8078
m=6		1.3268	1.3904	1.4516	1.5053	1.5518	1.5921	1.6275
m=7		1.2182	1.2762	1.3319	1.3808	1.4230	1.4597	1.4919
m=8		1.1325	1.1860	1.2375	1.2826	1.3216	1.3555	1.3851
m=9		1.0625	1.1125	1.1606	1.2028	1.2392	1.2707	1.2984
m=10		1.0041	1.0512	1.0964	1.1361	1.1704	1.2001	1.2261
h=4			k=5	k=6	k=7	k=8	k=9	k=10
m=3			2.1241	2.1864	2.2552	2.3217	2.3832	2.4392
m=4			1.7840	1.8348	1.8910	1.9453	1.9955	2.0411
m=5			1.5665	1.6103	1.6588	1.7057	1.7489	1.7882
m=6			1.4126	1.4516	1.4947	1.5365	1.5750	1.6100
m=7			1.2964	1.3319	1.3711	1.4091	1.4441	1.4759
m=8			1.2047	1.2375	1.2737	1.3088	1.3411	1.3704
m=9			1.1300	1.1606	1.1944	1.2271	1.2573	1.2846
m=10			1.0677	1.0964	1.1283	1.1591	1.1874	1.2131
h=5				k=6	k=7	k=8	k=9	k=10
m=3				2.2299	2.2716	2.3217	2.3737	2.4243
m=4				1.8707	1.9046	1.9453	1.9877	2.0289
m=5				1.6415	1.6706	1.7057	1.7421	1.7776
m=6				1.4794	1.5053	1.5365	1.5689	1.6005
m=7				1.3573	1.3808	1.4091	1.4386	1.4673
m=8				1.2610	1.2826	1.3088	1.3360	1.3624
m=9				1.1826	1.2028	1.2271	1.2525	1.2772
m=10				1.1171	1.1361	1.1591	1.1829	1.2061

Table 4.2: Values of $q_{k,h,m,.01}$

generating k groups, each with $(2m-1)$ iid random variates from the $L_L(1,1)$ distribution and computing the hazard rate function. We then compute the test statistic $V_{k,h,m}$ and compute the proportion of times H_0 is rejected in favor of H_v. These values are given in Table 4.3 and Table 4.4 for $\alpha = 0.05$ and $\alpha = 0.01$, respectively. It can be seen that the computed constants maintain the FEWR at the desired nominal levels.

4.4.2 Power Simulation versus the Mack-Wolfe Test

A nonparametric test of H_0 versus H_v for the known valley is the Mack-Wolfe test given by Mack and Wolfe [28]. We would like to compare the power of the test that we proposed in this thesis with that of the Mack-Wolfe test.

The Mack-Wolfe test statistic is essentially the sum of two Jonckheere-Terpstra test statistics, one for the decreasing trend at first and another one for the increasing trend after the valley. Thus we must first label the treatments so that they are in the prescribed ordered relationships to the known valley, h, corresponding to H_v as discussed earlier. By Theorem 3.1, this turns into an umbrella pattern of logistic location parameters with peak h. To calculate peak-known Mack-Wolfe statistic, A_h, we first compute the $h(h-1)/2$ Mann-Whitney counts U_{rs} for every pair of treatments with labels less than or equal to the hypothesized peak, that is for every $1 \leq r < s \leq h$. These Mann-Whitney counts are given by

$$U_{rs} = \sum_{i=1}^{n_r} \sum_{j=1}^{n_s} \phi(\log T_{ir}, \log T_{js}) \ , \qquad 1 \leq r < s \leq h \ ,$$

h=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10
m=3	0.0505	0.0482	0.0502	0.0472	0.0536	0.0509	0.0511	0.0526
m=4	0.0551	0.0473	0.0476	0.0526	0.0520	0.0514	0.0469	0.0479
m=5	0.0516	0.0492	0.0486	0.0499	0.0515	0.0483	0.0503	0.0492
m=6	0.0498	0.0539	0.0484	0.0524	0.0523	0.0461	0.0507	0.0530
m=7	0.0514	0.0471	0.0524	0.0515	0.0488	0.0470	0.0517	0.0475
m=8	0.0538	0.0474	0.0480	0.0511	0.0506	0.0497	0.0464	0.0482
m=9	0.0460	0.0357	0.0518	0.0459	0.0499	0.0529	0.0522	0.0516
m=10	0.0467	0.0483	0.0472	0.0475	0.0520	0.0465	0.0502	0.0486
h=3		k=4	k=5	k=6	k=7	k=8	k=9	k=10
m=3		0.0517	0.0494	0.053	0.0527	0.0518	0.0525	0.0513
m=4		0.0526	0.0483	0.0505	0.0486	0.0483	0.0513	0.0506
m=5		0.0574	0.0482	0.0517	0.051	0.0506	0.0494	0.0475
m=6		0.0513	0.0517	0.0482	0.0466	0.0494	0.0578	0.0469
m=7		0.0489	0.0478	0.0509	0.0178	0.0463	0.0523	0.0504
m=8		0.0471	0.0509	0.0467	0.0527	0.0531	0.0488	0.0491
m=9		0.0511	0.0467	0.0497	0.0502	0.0489	0.0494	0.0497
m=10		0.0491	0.0508	0.0518	0.0508	0.0509	0.0526	0.0465
h=4			k=5	k=6	k=7	k=8	k=9	k=10
m=3			0.0498	0.0509	0.0519	0.0512	0.0502	0.0457
m=4			0.0507	0.0505	0.0506	0.0547	0.0527	0.0498
m=5			0.0521	0.0520	0.0503	0.0506	0.0517	0.0479
m=6			0.0522	0.0478	0.0475	0.0509	0.0502	0.0488
m=7			0.0507	0.0511	0.0474	0.0513	0.0529	0.0532
m=8			0.0481	0.0507	0.0516	0.0472	0.0510	0.0463
m=9			0.0486	0.0504	0.0475	0.0510	0.0519	0.0507
m=10			0.0471	0.0498	0.0481	0.0450	0.0552	0.0487
h=5				k=6	k=7	k=8	k=9	k=10
m=3				0.0501	0.0494	0.0492	0.0513	0.0487
m=4				0.0513	0.0470	0.0522	0.0473	0.0531
m=5				0.0530	0.0518	0.0487	0.0469	0.0487
m=6				0.0502	0.0512	0.0523	0.0508	0.0525
m=7				0.0534	0.0489	0.0491	0.0548	0.0488
m=8				0.0502	0.0484	0.0523	0.0495	0.0501
m=9				0.0534	0.0510	0.0529	0.0476	0.0452
m=10				0.0473	0.0501	0.0511	0.0528	0.0544

Table 4.3: Values of α from simultaneous confidence intervals for $\alpha = 0.05$

h=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10
m=3	0.0103	0.0089	0.0088	0.0109	0.0116	0.0118	0.0096	0.0099
m=4	0.0105	0.0090	0.0082	0.0104	0.0109	0.0076	0.0097	0.0103
m=5	0.0131	0.0099	0.0101	0.0108	0.0092	0.0106	0.0087	0.0105
m=6	0.0109	0.0103	0.0108	0.0099	0.0099	0.0091	0.0100	0.0096
m=7	0.0101	0.0097	0.0105	0.0101	0.0112	0.0084	0.0082	0.0104
m=8	0.0086	0.0121	0.0097	0.0089	0.0095	0.0112	0.0102	0.0096
m=9	0.0108	0.0118	0.0099	0.0093	0.0086	0.0094	0.0114	0.0097
m=10	0.0096	0.0113	0.0096	0.0098	0.0102	0.0106	0.0091	0.0101
h=3		k=4	k=5	k=6	k=7	k=8	k=9	k=10
m=3		0.0092	0.0111	0.0104	0.0104	0.0102	0.0107	0.0103
m=4		0.0099	0.0106	0.0090	0.0113	0.0096	0.0093	0.0112
m=5		0.0095	0.0088	0.0118	0.0118	0.0095	0.0103	0.0090
m=6		0.0090	0.0091	0.0109	0.0117	0.0102	0.0085	0.0078
m=7		0.0093	0.0090	0.0095	0.0102	0.0084	0.0118	0.0100
m=8		0.0104	0.0089	0.0092	0.0086	0.0097	0.0085	0.0085
m=9		0.0091	0.0095	0.0095	0.0087	0.0095	0.0101	0.0102
m=10		0.0096	0.0103	0.0100	0.0102	0.0095	0.0103	0.0105
h=4			k=5	k=6	k=7	k=8	k=9	k=10
m=3			0.0095	0.0109	0.0114	0.0105	0.0109	0.0099
m=4			0.0094	0.0107	0.0093	0.0102	0.0100	0.0119
m=5			0.0103	0.0107	0.0109	0.0117	0.0094	0.0097
m=6			0.0110	0.0119	0.0100	0.0088	0.0095	0.0082
m=7			0.0098	0.0082	0.0091	0.0102	0.0100	0.0098
m=8			0.0110	0.0100	0.0095	0.0093	0.0112	0.0088
m=9			0.0096	0.0102	0.0100	0.0102	0.0106	0.0105
m=10			0.0100	0.0122	0.0085	0.0093	0.0097	0.0099
h=5				k=6	k=7	k=8	k=9	k=10
m=3				0.0106	0.0084	0.0085	0.0109	0.0104
m=4				0.0091	0.0106	0.0114	0.0106	0.0108
m=5				0.0108	0.0109	0.0109	0.0080	0.0098
m=6				0.0103	0.0122	0.0099	0.0119	0.0090
m=7				0.0109	0.0096	0.0099	0.0097	0.0091
m=8				0.0110	0.0106	0.0114	0.0099	0.0100
m=9				0.0108	0.0111	0.0091	0.0109	0.0084
m=10				0.0091	0.0097	0.0113	0.0099	0.0087

Table 4.4: Values of α from simultaneous confidence intervals for $\alpha = 0.01$

where $\phi(a, b) = 1$ if $a < b$, 0 otherwise. That is to say U_{rs} is the number of sample r before sample s precedences. The Jonckheere-Terpstra statistic, J_1, is then the sum of these $h(h-1)/2$ Mann-Whitney counts,

$$J_1 = \sum_{u=1}^{v-1} \sum_{v=2}^{h} U_{uv} \ .$$

Since we have an umbrella alternative, after the order is reversed after the peak and the Mann-Whitney counts are computed. That is $(k-h+1)(k-h)/2$ reverse Mann-Whitney counts U_{rs} for every pair of treatments with labels greater than or equal to the hypothesized peak for each pair of remaining treatments $(h \leq r < s \leq k)$. Thus U_{sr} is the number of sample s before sample r precedences. The Jonckheere-Terpstra statistic, J_2, is then the sum of these $(k-h+1)(k-h)/2$ reverse Mann-Whitney counts,

$$J_2 = \sum_{u=h}^{v-1} \sum_{v=h+1}^{k} U_{vu} \ .$$

Then Mack-Wolfe peak known statistic, A_h, is then the sum of the Mann-Whitney counts to the left of the peak and the reverse Mann-Whitney counts to the right of the peak in accordance with the umbrella alternative, that is

$$A_h = J_1 + J_2 \ .$$

To test the umbrella alternative against no difference at α level significance, we reject H_0 in favor of H_v if

$$A_h \geq a_{h,\alpha} \ ,$$

where the constant $a_{h,a}$ is chosen to make the Type I error probability equal to α. Values of $a_{h,\alpha}$ are given in Table A.14 Hollander and Wolfe [19] for each selected h and k combinations.

We considered the hazard function given by

$$\lambda(t;\gamma) = \frac{(\beta/\gamma)(t/\gamma)^{\beta-1}}{1+(t/\gamma)^{\beta}}$$

and fixed $\beta = 1$. This reduces the hazard function to $\lambda(t;\gamma) = (\gamma + t)^{-1}$. We then generated several hazard functions with different values of γ providing valley patterned hazard rate function configurations. The Type I error rates of under these different configurations are given in Table 4.5.

λ_1^{max}	λ_2^{max}	λ_3^{max}	λ_4^{max}	$P_0(V_{4,3,3} \leq -q_{4,3,3,.0497})$	$P_0(A_3 > a_{3,.0497})$
1.0	1.0	1.0	1.0	0.0495	0.0423
$(1.7)^{-1}$	$(1.7)^{-1}$	$(4.5)^{-1}$	$(1.0)^{-1}$	0.6932	0.5800
$(1.0)^{-1}$	$(2.7)^{-1}$	$(4.5)^{-1}$	$(2.7)^{-1}$	0.7327	0.5895
$(2.7)^{-1}$	$(2.7)^{-1}$	$(4.5)^{-1}$	$(2.7)^{-1}$	0.5715	0.1956
$(2.7)^{-1}$	$(2.7)^{-1}$	$(7.4)^{-1}$	$(1.0)^{-1}$	0.8912	0.6424
$(2.7)^{-1}$	$(7.4)^{-1}$	$(7.4)^{-1}$	$(1.0)^{-1}$	0.9096	0.6375

Table 4.5: Simulated Power of the Mach-Wolfe Test against our Test

Based on the simulated power analysis, the following observations are noted:

- our test maintains the level of the test $\alpha = 0.0497$ when all the hazard functions are the same level of accuracy
- our test is more powerful in detecting valley patterns in the hazard functions compared to the Mack-Wolfe test.

Chapter 5

Simultaneous Confidence Intervals

When the null hypothesis is rejected in favor of the alternative hypothesis, the investigator is usually interested in determining exactly which hazard functions are different from each other. If the null is not rejected, then no further action is needed.

Let us focus on the valley shaped alternative that was discussed in Chapter 4. This case is chosen since it is more general than the simple ordering alternative considered in Chapter 3. If one is interested in the simple ordering case, then one can get the simultaneous confidence intervals by simply replacing h by $1(k)$ for the decreasing (increasing) arrangement of hazard functions.

The simultaneous confidence intervals with simultaneous confidence coefficient of $1-\alpha$ for the pairwise differences $(\lambda_j^{max} - \lambda_i^{max})$, $1 \le i < j \le h$ and $h \le j < i \le k$, may be constructed by inverting the α-level test $P_{H_0}(V_{k,h,m} < -q_{k,h,m,\alpha}) = \alpha$. The $100(1-\alpha)\%$ simultaneous confidence interval may be obtained as

$$
\begin{aligned}
1-\alpha &= P_{H_0}(V_{k,h,m} \ge -q_{k,h,m,\alpha}) \\
&= P\left(\lambda_j^{max} - \lambda_i^{max} \ge (Y_j - Y_i) - q_{k,h,m,\alpha}\ , \quad 1 \le i < j \le h,\ h \le j < i \le k\right)
\end{aligned}
$$

The investigator may also be interested in more general contrasts of the hazard functions rather than just simple pairwise differences. Let $\mathbf{Y} = (Y_1, \ldots, Y_k)'$. Letting $\boldsymbol{\lambda}^{max} = (\lambda_1^{max}, \ldots, \lambda_k^{max})'$, one may want to construct a confidence interval for $\mathbf{c}'\boldsymbol{\lambda}^{max}$

for a given constant $\mathbf{c} \in \Re^k$ such that $\mathbf{c}'\mathbf{1} = 0$, where $\mathbf{1}$ is the vector of all ones in $\Re^k$. More generally, one may want simultaneous intervals for all contrasts of a certain type.

Following Berk and Marcus [9] (see also Hayter and Liu [17]), one may invert the α level test to produce a set of simultaneous confidence intervals for the restricted set of contrasts

$$\left\{ \mathbf{c}'\boldsymbol{\lambda}^{max} : \ \mathbf{c} \ \in \ \mathcal{C}_h \right\}, \tag{5.1}$$

where the class of contrasts is given by

$$\mathcal{C}_h = \left\{ \mathbf{c} \ : \ \sum_{i=1}^{j} c_i \geq 0 \text{ for } j = 1, \ldots, h{-}1, \ \sum_{i=k-t+1}^{k} c_i \geq 0 \text{ for } t = 1, \ldots, k{-}h, \ \text{ and } \mathbf{c}'\mathbf{1} = 0 \right\}.$$

The set simultaneous confidence intervals for the all the contrasts given in (5.1) is

$$1 - \alpha = P\left(\mathbf{c}'\boldsymbol{\lambda}^{max} \geq \mathbf{c}'\mathbf{Y} - q_{k,h,m,\alpha}\|\mathbf{c}\| \ , \quad \forall \mathbf{c} \in \mathcal{C}_h\right) \ . \tag{5.2}$$

Whenever the null hypothesis is rejected, we can find which of the means are different. We can find any class of contrast including pairwise differences which follows Eqs. (5.1) and (5.2). As an example in Table 5.1, we considered few combinations of contrasts and inverted to get the simultaneous confidence intervals. Table 5.1 gives SCI's of $(1 - \alpha)$.

	μ_i	c	$(1-\alpha)$
$h=2, k=5;\ r=1.2465$	(0.5 2 0.5 0.5 0)	(-1 4 -1 -1 -1)	0.9519
$h=3, k=5;\ r=1.2201$	(0.5 0.5 2 0.5 0)	(-1 -1 4 -1 -1)	0.9436
$h=3, k=6;\ r=1.2967$	(0.5 1 2 1 0.5 0)	(-1 -1 5 -1 -1 -1)	0.9580

Table 5.1: Simultaneous confidence interval for $m = 5$ and $\alpha = 0.05$

If the ordering given in the alternative hypothesis is certain, then this information can be used to improve the simultaneous confidence intervals. One instance is when a negative lower bound is obtained for the contrast $\mathbf{c}'\boldsymbol{\lambda}^{max}$, which is impossible under the alternative hypothesis. One then sharpens the interval by using zero as the lower bound. In general, if the ordering given in the alternative hypothesis is believed to be true, then the simultaneous confidence intervals for all sets of contrasts with coefficients in $\mathcal{C}_h$ are

$$\mathbf{c}'\boldsymbol{\lambda}^{max} \geq \max\{0\ ,\ \mathbf{c}'\mathbf{Y} - q_{k,h,m,\alpha}\|\mathbf{c}\|\}\ .$$

For more discussion on simultaneous confidence intervals under order restrictions, see Hayter and Liu [18], Marcus and Genizi [29], or Hwang and Peddada [20].

Chapter 6

Summary

In this thesis we have developed an exact testing strategy for comparing maximum hazard rates of multiple treatments or doses under the assumption that the survival times follow the log-logistic distribution. In particular, exact test procedures are formulated for testing the null hypothesis of no difference against a simple ordered alternative in the comparison of the worst case hazard rates of several log-logistic survival functions. A recursive algorithm is given for computing the required critical constants along with the computed constants and a Monte Carlo simulation study to evaluate the power of the given procedure versus a suitable nonparametric procedure.

The method of obtaining the critical constants for the simple ordering case was extended to include umbrella and valley ordering on the maximum hazard rates. Once again critical constants were computed using a recursive algorithm and they were evaluated for correctness using a simulation study. We also compared the power performance of the newly developed procedure against the nonparametric Mack-Wolfe test under several valley-shaped configurations of the maximum hazard rate functions. These comparisons show that the new test is much more powerful than the Mack-Wolfe test in detecting valley-shaped patterns of the maximum hazard rate functions.

The test procedures developed were inverted to give simultaneous confidence intervals for a class of contrasts including pairwise comparisons. This can be used to

identify the treatments or doses that are significantly different from each other in case the null hypothesis of no difference is rejected in favor of the valley- or simple-ordered alternative.

We are currently working on developing test procedures for the entire survival curve as well as other interesting characteristics of the survival function of log-logistic survival times.

Bibliography

[1] Abebe A, Singh P (2008), "An Exact Test for U-Shaped Alternatives of Location Parameters: the Exponential Distribution Case ", *Communications in Statistics Theory and Methods* ; **37(1) 2008**: 37-45.

[2] Ahmad M.I., Sinclair C.D., Werrity A.,(1982) "Systems of frequency curves generated by transformations of logistic variables "; *Biometrika* ; **69(2) 1982** : 461-465.

[3] Ahmad M.I., Sinclair C.D., Werrity A.,(1988) "Log-Logistic Flood Frequency Analysis "; *Journal of Hydrology* ; **98(3-4) 1988** : 205-224.

[4] Arnold B.C., David H.A., Kent J.T., Nagaraja H.N., Ferreira J.T.A.S., Steel M.F.J., Jones M.C. (2004), "Families of distributions arising from distributions of order statistics - Discussion"; *TEST* ; **13(1) 2004**: 23-37.

[5] Bartholomew DJ (1959), "A test of homogeneity for ordered alternatives ", *Biometrika*; **46(1-2) 1959**: 36-48.

[6] Bartholomew DJ (1959), "A test of homogeneity for ordered alternatives 2. ", *Biometrika*; **46(3-4) 1959**: 328-335.

[7] Bartholomew DJ (1961), "A test of homogeneity of means under restricted alternatives 2. ", *Journal of the Royal Statistical Society. Series B (Methodological)*; **23(2) 1961**: 239-281.

[8] Bennett S (1983) "Log-logistic regression models for survival data ", *Applied Statistics Journal of Royal statistical society seties C*; **32(2) 1983)**: 165-171.

[9] Berk R., Marcus R.(1996), "Dual Cones, Dual Norms and Simultaneous Inference for Partially Ordered Means. ", *Journal of the American Statistical Association*; **91**.

[10] Berkson J. (1944), "Application of the logistic function to bio-assay ", *Journal of the American Statistical Association*; **39(227) 1944**: 357-365.

[11] Botkin DB, Richard SM(1974), "Mortality rates and survival of birds ", *The American Naturalist*; **108(960) 1974**: 181-192 .

[12] Brüederl J, Diekmann A (1995) "The Log-logistic rate model- 2 genealizations with an application to demogaphic data ", *Socialogical Methods and Reasearch*; **24(2)**: 158-186.

[13] Chen Y.I. (1999), "Nonparametric Identification of the Minimum Effective Dose ", *Biometrics*; **55(4)**: 1236-1240.

[14] Conkin J.(2001), "A Log Logistic Survival Model Applied to Hypobaric Decompression Sickness ";

[15] Eisen, M. (1979), "Mathematical Models in Cell Biology and Cancer Chemotherapy ", *Lecture Notes in Biomathematics Springer-Verlag, Berlin***30 1979** :

[16] Hayter A. J. (1990), "A One-Sided Studentized Range Test for Testing Against a Simple Ordered Alternative"; *Journal of the American Statistical Association*; **85(411)**: 778-785.

[17] Hayter A. J. , Liu W. (1996), "Exact calculations for the one sided studentized range test for testing against a simple ordered alternative "; *Computational statistics and data analysis*; **22 1996**: 17-25.

[18] Hayter A. J. and Liu W. (1999), "A new test against an umbrella alternative and the associated simultaneous confidence intervals "; *Computational statistics and data analysis*; **30**: 393-401.

[19] Hollander M. and Wolfe D.A. , "Nonparametric statistical methods ", 2^{end} Ed. John Wiley and Sons, New York - 1999.

[20] Hwang J. T. Gene and Shyamal Das Peddada (1994), "Confidence Interval Estimation Subject to Order Restrictions"; *The Annals of Statistics*; **22(1)**: 67-93.

[21] Jenson (1975), "Comaprision of logistic equations for population growth ", *Biometrics*; **31(4) 1975**: 853-862 .

[22] Jonckheere(1954), "A distribution-free fc-sample test against ordered alternatives "; *Biometrika*; **41(1-2) 1954**: 133-145.

[23] Jones, R.H. and R.R. Sharitz. (1998), "Survival and growth of woody plant seedlings in the understorey of floodplain forests in South Carolina. ", *Journal of Ecology*; **86**: 574-587

[24] Kalbfleisch J. D., Prentice R.L. (1973), "Marginal likelihoods based on Cox's regression and life model "; *Biometrika* ; **60(2) 1973** : 267-278.

[25] Lehmann E.L.(1953), "The Power of Rank Tests ", *The Annals of Mathematical Statistics*; **24(1)**: 23-43.

[26] Leisenring W., Ryan L. (1992), "Statistical properties of NOAEL. ", *Biometrics*; **15(2) 1992**: 161-171.

[27] Lorenzen T.J., McDonald G.C. (1981), "Selecting Logistic populations using the sample medians ", *Journal of the american statistical association*; **39(227)**: 357-365.

[28] Mack G.A., Wolfe D.A. (1981), "K-Sample Rank Tests for Umbrella Alternatives ", *Journal of the American Statistical Association* ; **76(373) 1981**: 175-181.

[29] Marcus R., Genizi A. (1994), "Simultaneous confidence intervals for umbrella contrasts of normal means ", *Computational Statistics and Data Analysis*; **17(4)**: 393-407.

[30] Moolgavkar SH, Lustbader ED, Vennzon DJ (1985), "Assessing the adequacy of the logistic regression model for matched case control studies ", *The American Naturalist*; **4(4) 1985**: 425-435.

[31] Pearl R, Reed LJ (1920), "On the rate of growth of the population of the united states since 1790 and its mathematical representation ", *Proceedings of the National Academy of Sciences of the United States of America*; **6(6) 1920**: 275-288.

[32] Reed LJ, Berkson J (1929), "The application of the logistic function to experimental data ", *Journal of Physical Chemistry*; **33 1929)**: 760-779.

[33] Robertson T (1978), "Testing for and against an order restriction on multinomial parameters. ", *Journal of American Statistical Association*; **73(361) 1978**: 197-202.

[34] Robertson T., Wright F.T., Dykstra R.L. , "Order Restricted Statistical Inference ", John Wiley and Sons, New York - 1988.

[35] Robson A. J., Reed D.W., "Statistical Procedures for Flood Frequency Estimation "; *The Canadian Journal of Statistics* ; **16(3) 1999** : 223-236.

[36] Roy S. N. (1953), "On a heuristic method of test construction and its use in multivariate analysis"; *The Annals of Mathematical Statistics*; **24(2)**: 220-238.

[37] Ruberg S. J.(1989), "Contrasts for identifying the minimum effective dose ", *Journal of the American Statistical Association*; **84(407) 1989**: 816-822.

[38] Sen K.P (2007), Union intersection principle and constrained statistical inference *Journal of Statistical Planning and Inference*; **137(11)**: 3741-3752.

[39] Shoukri M. M., Mian I.U.M., Tracy D.S. (1988), "Sampling Properties of Estimators of the Log-Logistic Distribution with Application to Canadian Precipitation Data "; *The Canadian Journal of Statistics* ; **16(3) 1988** : 223-236.

[40] Simpson DG, Margolin BH (1986), "Recursive nonparametric testing for dose-response relaflonships subject to downturns at high doses ", *Biometrika*; **73(3) 1986**: 589-596.

[41] Singh KP, Lee CMS, George EO (1998), "On generalized log-logistic model for censored survival data ", *Biometrical Journal* ; **30(7) 1988)**: 843-850.

[42] Singh P., Carpenter M., Mishra S.N. (2006), "Multiple comparison procedures for location parameters: logistic population case ", *American journal of Mathematical and Management Sciences* ; **26(3-4) 2006**: 371-385.

[43] Tan W.Y. (1983), "On the distribution of number of mutants at the hypoxanthine-quanine phorsphoribosal transferase locus in Chinese hamster ovary cells ", *Math. Biosciences***18(1) 1983** : 67-192.

[44] Tebbs J.M., Bilder C.R. (2006), "Hypothesis tests for and against a simple order among proportions estimated by pooled testing "; *Biometrical Journal*; **48(5) 2006**: 792-804.

[45] Terpstra T.J.(1952), "The asymptotic normality and consistency of Kendall's test against trend, when ties are present in one ranking "; *Indagationes Mathematicae*; **14 1952**: 327-333.

[46] Tukey J.W., Cimenera J.L., Heyse J.F.(1985), "Testing the statistical certainty of a response to increasing doses of a drug. ", *Biometrics*; **41**: 295-301.

[47] Verhulst, P. F. (1838), "Notice sur la loi que la population pursuit dans son accroissement ", *Correspondance math m atique et physique***10 1838** : 113-121.

[48] Verhulst, P. F. (1845), "Recherches Mathematiques sur La Loi D'Accroissement de la Population ", *Nouveaux Memoires de l'Academie Royale des Sciences et Belles-Lettres de Bruxelles***18(1) 1845**: 1-45.

[49] Williams D.A.(1977), "Some inference procedures for monotonically ordered normal means ", *Biometrika*; **64(1)**: 9-14.

[50] Yanagimoto, T., Sibuya, M. (1980), "Comparision of tails of distributions in models for estimating safe doses ", *Annals of the Institute of Statistical Mathematics*; **32(1) 1980)**: 325-340.

APPENDIX

MATLAB PROGRAM

```
MAINPROGAM.m
clear all;
close all;
c100=8;
lambda=6*c100/(100*(100+1)*(2*100+1));
i= -101:0.001:100;
    c=lambda*i.*(i+1).*(2*i+1)./6;
h=3; k=7; m=5; alpha=0.95;
f=fzero(@(r) myfunlogistic(r,c,h,k,alpha,m), 0)
 % Solve for 'r'

MYFUN.m

function F=myfunlogistic(r,y,h,k,alpha,m)
a=pi/3^0.5;
term1=(gamma(2*m)/(gamma(m))^2);
phi=term1*a*exp(-(a*y.*m)).*((1+exp(-a*y)).^(-2*m));
%pdf
for j=0:m-1,
term2(j+1,:)=nchoosek((m-1),j)*(2*m-j-1)^(-1)*(-1)^(m-
1-j)*(1+exp(-a*(y+r))).^(j+1-2*m);
end

for j=0:m-1,
term3(j+1,:)=nchoosek((m-1),j)*(2*m-j-1)^(-1)*(-1)^(m-
1-j)*(1+exp(-a*y)).^(j+1-2*m);
end

f1=term1*sum(term2);   % cdf
f12=term1*sum(term3);
dif1=k-h;
dif2=h-1;
if (dif2 == 1)
    p1=f1;
elseif (dif2 == 0)
    p1=1;
```

```
else
 p1 = myfunlogistic_rec(r,y,dif2,phi,f1,f12);
end

if (dif1 == 1)
    p2=f1;
else

    p2=myfunlogistic_rec(r,y,dif1,phi,f1,f12);
end
    FB=(phi.*p1.*p2);

  F=trapz(y,FB)-alpha; %y1 is phi(y)---pdf
```

MYFUNREC.m

```
function p=  myfunlogistic_rec(r,y,dif,phi,f1,f12)
if (dif == 2)
p=cumtrapz(y,phi.*f1)+f1.*( f1-f12);
else
p=cumtrapz(y,phi.*myfunlogistic_rec(r,y,dif-
1,phi,f1,f12))+myfunlogistic_rec(r,y,dif-
1,phi,f1,f12).*( f1-f12);
end
```