THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ BORDEAUX 1

ÉCOLE DOCTORALE DE MATHEMATIQUE ET INFORMATIQUE

DOMAINE DE RECHERCHE : Statistique, Biostatistique

Présentée et soutenue par

MUHAMMAD RAMZAN TAHIR

Title: On Validation Of Parametric Models Applied In Survival Analysis And Reliability

Titre: Sur La Validation Des Modèles Paramétriques Appliqués En Analyse De Survie Et Fiabilité

Sous la direction de : **Mikhail Nikulin, Professeur des Universités**

le: 02/07/2012

Devant le jury composé de:

M. Mikhail NIKULIN	Professeur, Université de Bordeaux	Directeur
Mme. Catherine HUBER-CAROL	Professeur, Université René Descartes-Paris 5	Présidente
M. Fabrice GUERIN	Professeur, Université d'Angers	Rapporteur
M. Henning LAUTER	Professeur, University of Potsdam, Germany	Rapporteur
M. Nikolaos LIMNIOS	Professeur, Université de Technologie de Compiègne	Examinateur
M. Léo GERVILLE-REACHE	Maître de Conférences HDR, Université Bordeaux 2	Examinateur
M. Vincent COUALLIER	Maître de Conférences, Université Bordeaux 2	Examinateur

To the One I Love

Acknowledgements

I am thankful to Almighty Allah, the most beneficent Who gave me the thoughts, courage and strength to carry out this work.

I would like to express my gratitude to my advisor Prof. Dr. Mikhail Nikulin. His wonderful guidance, generous support, and patience enable me to complete this work. Also I am thankful to Madame Voila Nikulina.

I owe great gratitude to the president of the jury Mme. Catherine Huber for her kind suggestions, reporters of thesis M. Fabrice Guerin and M. Henning Lauter for their valuable reports, and examiners of the defence M. Vincent Couallier, M. Nikolaos Limnios, and M. Léo Gerville-Reache for their precious comments, which make this thesis presentable.

I would like to say thanks to my hosting lab for offering me a great environment for the research. I extend my gratitude to all the team members of the department and also to the secretaries for their administrative stuff.

I am grateful to Prof. Ahmadou Alioum for his useful advices and scientific guidance during my Masters and PhD. Also thanks to Dr. Daniel Commenges for his concern.

I am indebted always to those people that do their best to improve on my best. I would like to pay my special gratitude to Mr. Aijaz Asif and Ms. Li Qian for their silent continuous motivation to come up to this task.

I feel pleasure to say thanks to all Pakistani students in Bordeaux who make my stay pleasant.

I am thankful to Noureddine Saaidia for his collaboration and help during my PhD.

I pay my thanks to my dear friends Akhtar Ammar, Ch. Muhammad Saifullah, Fiaz Hussain, Dr. Ijaz Anjum, Dr. Junaid Bhatti, Muhammad Riaz, Rabnawaz Anjum, Tajammul Hussain and Yousaf khan for giving me the courage through their friendship.

My special thanks go to my parents, my brothers Ghulam Murtaza and Iftikhar Hussain, and all my family members for their support throughout my life.

I wish to express my earnest thanks to all my friends whose names are not mentioned but their friendship is valuable for me.

Thanks to the Higher Education Commission of Pakistan (HEC) for providing me the opportunity for higher studies.

Abstract

This is an increasing importance in survival analysis and reliability to select a suitable basic model for further inquiries of the data. Little deviation in basic model can cause serious problems in final results. The presence of censoring and accelerated stresses make this task more difficult. Chi-square type goodness of fit tests are most commonly used for model selection. Many modifications in chi-square tests have been proposed by various researcher. The first aim of the thesis is to present a goodness of fit test for wide rage of parametric models (shape-scale families) commonly used in survival analysis, social sciences, engineering, public health and demography, in presence of right censoring. We give the explicit forms of the quadratic form of the test statistic (RRN test) for various models and apply the test on real data. We develop a computer program in R-language for all models. A separate section is dedicated for the test in demography. We focus on the Birnbaum-Saunders (BS) distribution for goodness of fit test for parametric AFT-model and analysis of redundant system.

The other purpose of the thesis is to give the analysis of redundant system. To ensure high reliability of the main components of the systems, standby units are used. The main component is replaced by the standby unit automatically, if it fails. The standby unit can be in warm, hot, or cold state. We give the procedure of one main and $n-1$ standby units placed in hot state, and give the detailed analysis of one main and one standby unit using BS parametric family. We use Sedyakin's physical principal and the approach of accelerated failure time model for the analysis of redundant system. This approach is different from the traditional ones in the literature but difficulties in calculations arise. We calculate the reliability of the system in terms of distribution function (unreliability function) and asymptotic confidence interval.

Keywords : Accelerated failure time model, Birnbaum-Saunders distribution, Chi-Squared type goodness of fit tests, Demography, Redundant system, RRN test, Survival analysis.

Résumé de Thèse

Les méthodes de sélection de modèles deviennent de plus en plus importantes dans l'analyse de survie et de fiabilité. De petites variations dans le choix du modèle de base peuvent entraîner de grandes différences dans les résultats finaux. La présence de censure et de covariables rendent encore plus difficile les calculs liés à ces choix. Les tests d'adéquation de type chi-deux sont les plus couramment utilisés pour la sélection du modèle. Le premier objectif de la thèse est de présenter un test d'ajustement pour les modèles paramétriques couramment utilisés en analyse de survie, en sciences sociales, ingénierie, santé publique ou démographie, en présence de censure à droite. De nombreuses modifications dans les tests du chi-deux ont été proposées par divers chercheurs. La première modification pour les données censurées d'un test du chi-deux a été proposée par Habib et Thomas (1986) sur la base des différences entre estimateur de Kaplan-Meier $\hat{F}_n(t)$ et estimateur du maximum de vraisemblance paramétrique de fonction de survie $F(t, \hat{\theta}_n)$. Akritas (1988) a proposé une statistique du chi-deux basée sur l'idée de comparer le nombre observé et attendu de défaillances dans chaque classe. Hjort (1990) a développé une statistique de type chi-deux pour la validité d'un modèle paramétrique pour les données de durées de vie basé sur le processus de risque cumulatif. Plusieurs autres modifications furent proposées par les chercheurs (par example, Peña, 1992, Kim,1993, Nikulin et Solev, 1999, Bagdonavicius et al., 2010a). Au Chapitre 2 nous étudirons ce test et nous donnerons les formes explicites de la forme quadratique de la statistique de test (test de RRN) pour différents modèles avec des applications du test sur des données réelles. Nous développons également un logiciel en langue R pour ces modèles paramètriques. Dans ce chapitre, une section séparée est consacrée aux tests dans le domaine de la démographie. Nous nous concentrons ensuite sur le modèle de Birnbaum-Saunders (BS) pour le test d'ajustement pour les modèles AFT paramétriques (chapitre 3) et en analyse de systèmes redondants en Chapitre 4.

L'autre contribution porte sur l'analyse de systèmes redondants composés d'un composant en état actif hot et d'un autre composant en réserve dites tiède (ie en état warm). Les unités réservées sont utilisées pour augmenter la fiabilité du système. En cas de panne du composant principal, celui-ci est automatiquement remplacé par l'unité en veille. L'unité de réserve peut être dans un état dit hot, warm, ou cold. Nous donnons une procédure générale avec une unité principale et $n-1$ unités fonctionnant dans un warm, puis donnons l'analyse détaillée d'une unité en réserve en utilisant la famille paramétrique de Birnbaum-Saunders. Nous utilisons le principe de Sedyakin (1966) et l'approche du modèle de vie accéléré (AFT) pour l'analyse de systèmes redondants. Cette approche est différente des approches traditionnelles dans la littérature et elle pose des difficultés dans les calculs. Nous calculons la fiabilité du système en termes de fonction de répartition et nous donnons l'intervalle de confiance asymptotique au Chapitre 4.

Mots clés : Analyse de survie, Démographie, Modèle de Birnbaum-Saunders, Modèles de

vie accélérés, Système Redondant, Test d'ajustement de type de chi-deux, Test de RRN.

Table of Contents

List of Figures

General Introduction

1 Introduction

The thesis consists of three parts. All models, methods and tests are equally applied in survival analysis and in reliability. Some frequently used parametric models are explained in chapter 1. In the second chapter chi-squared type goodness-of-fit tests for several parametric models which are being used in survival analysis and reliability are given for right censored data. For all models elements of the quadratic form of the test are presented in a way that are easy to use for the practitioners. Third chapter is based on the analysis of accelerated failure time (AFT) model when the base line survival function is Birnbaum-Saunders (BS) distribution which include the estimation of parameters, survival function under normal stress and asymptotic confidence interval for survival function. Also a goodness-of-fit test is given for BS-AFT model. In fourth chapter statistical analysis of redundant systems with BS distribution is given with one operating unit and one standby unit operating in warm condition.

Here we give some introductory note on the contents of all chapters. Detailed methods and application with examples are presented in the relevant chapters.

2 Censored Data

Censoring means that the observations are partially known i.e. at the time of evaluation or at the end of the experiment the outcome or the event of interest does not occur on some individuals (time-to-event data) or some subject leave the study in-between with out any information (lost to follow-up). One can observe only the failure time T if it does not exceed the censoring time C. Censoring is very common in the experiment of health sciences and industry and it created some technical problems like what assumptions are made about censoring mechanism and whether censoring is independent of residual lifetime or not.

There are three types of *censoring mechanisms* depending on the relationship of the failure time and the censoring time(s), including left-censoring, right-censoring, and interval-censoring. An observation is *left censored* if the event of interest has already occurred when observation of time begins. *Right censoring* occurs when a subject has not experienced the failure at the end or has been lost to follow-up during the study. *Interval-censoring* occurs when the failure time happens between two successive observation times.

In this document we consider right censored data. Right censoring mechanism can be of various types. Let T be the failure time, C be a censoring time, and τ be some specified time for the experiment. $X = \min(T, C)$ is a censored observation of T and $\delta = 1_{\{T \leq C\}}$ is an indicator variable which indicates whether the event of interest is observed ($\delta = 1$) or not ($\delta = 0$). For each subject i, we actually observe the pair (X_i, δ_i).

- If n subjects are observed at fixed study time τ then censoring is called *Type I censoring.* For each subject we observe $X_i = \min(T_i, \tau)$, $\delta = 1_{\{T_i \leq \tau\}}$.
- In *Type II censoring* the study is terminated until a specified number of failure, say $r < n$, occurs. This type is often used in testing equipment life. For each subject we observe $X_i = \min(T_i, T_{(r)})$ and $\delta_i = 1_{\{X_i = T_i\}}$ where $T_{(r)}$ is the ordered failure time of *rth* subject.
- Let τ be the specified time of the study and units are entered in the study at different time points $(t_1, \cdots, t_n)$. We observe $X_i = \min(T_i, \tau - t_i)$ and $\delta_i = 1_{\{T_i \leq \tau - t_i\}}$. Here every subject has different fixed-censoring time which is non-random. This is called *progressive right censoring.*
- Censoring is called *independent right censoring* if the failure times $T_1, \cdots, T_n$ and the censoring times $C_1, \cdots, C_n$ are mutually independent random variables.

Right censored data can also be described by the stochastic process as

$$(N_1(t), Y_1(t), t > 0), \cdots, (N_n(t), Y_n(t), t > 0), \tag{1}$$

where $N_i(t) = \mathbf{1}_{\{X_i \leq t, \delta_i = 1\}}$ is the *process of failure*, $Y_i(t) = \mathbf{1}_{\{X_i \geq t\}}$ is the *process at risk*, and

$$N(t) = \sum_{i=1}^{n} N_i(t) \quad \text{and} \quad Y(t) = \sum_{i=1}^{n} Y_i(t).$$

3 Goodness-of-fit Test

Goodness-of-fit tests indicate whether or not it is reasonable to assume that a random sample comes from a specific distribution. Statistical techniques often rely on observations having come from a population that has a distribution of a specific form (e.g., normal, lognormal, Poisson, etc.). Standard control charts for continuous measurements, for instance, require that the data come from a normal distribution. Accurate lifetime modeling requires specifying the correct distributional model. There may be historical or theoretical reasons to assume that a sample comes from a particular population, as well. Past data may have consistently fit a known distribution, for example, or theory may predict that the underlying population should be of a specific form. The test requires that the data first be grouped. The actual number of observations in each group is compared to the expected number of observations and the test statistic is calculated as a function of this difference. The number of groups and how group membership is defined will affect the power of the test (i.e., how sensitive it is to detecting departures from the null hypothesis). Power will not only be affected by the number of groups and how they are defined, but by the sample size and shape of the null and underlying (true) distributions.

Selection of a suitable model in all types of statistical analysis is of great importance. For this purpose a lot of goodness-of-fit tests are proposed by the researchers. Chi-square type goodness-of-fit tests are mostly used where the random sample $\mathbf{X} = (X_1, \cdots, X_n)^T$ of size n is partitioned into k-subintervals and $\nu = (\nu_1, \cdots, \nu_k)^T$ is the vector of frequencies, where ν_j is frequency of j^{th} group and $\sum_{j=1}^{k} \nu_j = n$. The tests are based on the following Pearson's statistic

$$X_n^2(\boldsymbol{\theta}) = X_n^T(\boldsymbol{\theta})X_n(\boldsymbol{\theta}) = \sum_{j=1}^{k} \frac{(\nu_j - np_j(\boldsymbol{\theta}))^2}{np_j(\boldsymbol{\theta})}, \quad \boldsymbol{\theta} = (\theta_1, \cdots, \theta_s)^T \in \Theta \subset R^s, \tag{2}$$

where

$$X_n(\boldsymbol{\theta}) = \left(\frac{\nu_1 - np_1(\boldsymbol{\theta})}{\sqrt{np_1(\boldsymbol{\theta})}}, \cdots, \frac{\nu_k - np_k(\boldsymbol{\theta})}{\sqrt{np_k(\boldsymbol{\theta})}} \right)^T,$$

$\boldsymbol{p}(\boldsymbol{\theta}) = (p_1(\boldsymbol{\theta}), \cdots, p_k(\boldsymbol{\theta}))$ is the vector of probabilities and $\boldsymbol{\theta}$ is a vector of parameters which can be known (simple hypothesis) or unknown (composite hypothesis). In classical Pearson chi-square statistic $\boldsymbol{\theta}$ is supposed to be known but later the theory of the test developed and many researchers put their contribution to develop the test by using different estimation methods and interval selection procedures. Here we discuss this development in brief and in next chapter we give some details on estimation methods and their properties.

The standard form of Pearson's statistic (2) possesses a chi-squared distribution in limit under the simple hypothesis H_0 (when θ is known) with $k-1$ degrees of freedom. If parameter θ is unknown and replaced by $\sqrt{n}$-consistent estimate $\tilde{\theta}$ based on grouped data that is by the minimum chi-squared or grouped maximum likelihood estimates, then according to Fisher (1928) the statistic $X_n^2(\tilde{\boldsymbol{\theta}})$ still follow the chi-squared distribution but with $k-s-1$ degrees of freedom, where s is the number of estimated parameters. Further the problem arises when θ is replaced by ungrouped ML estimator which is the commonly used method for parameter estimation. In 1954, Chernoff and Lehmann (1954) showed that when MLE $\hat{\theta}_n$ is used the limit distribution of the Pearson's statistic does not follow a chi-squared distribution and in general it depends on the properties estimator θ^*. After that a lot of investigations have been done to recommend the application of chi-square testing in different fields.

In 1973, Nikulin (1973a, 1973b) proposed a modification in the standard chi-squared Pearson's test for continuous distributions also with shift and scale parameters. In 1974, Rao and Robson (1974) obtained the same result for exponential family, and later this statistic was well adapted by the researchers with the name as Rao-Robson-Nikulin (RRN) test, (see for example, Drost (1988), Van der Vaart (1998), Voinov et al. (2009), Zhang (1999), Bagdonavicius et al. (2010b)). This statistic can be written as

$$\begin{aligned} Y_n^2(\hat{\theta}_n) &= X_n^T(\hat{\theta})\Sigma^-(\hat{\theta})X_n(\hat{\theta}) \\ &= X_n^2(\hat{\theta}_n) + X_n^T(\hat{\theta}_n)B(\hat{\theta}_n)[I(\hat{\theta}_n) - J(\hat{\theta}_n)]^{-1}B^T(\hat{\theta}_n)X_n(\hat{\theta}_n), \end{aligned} \tag{3}$$

where $\hat{\theta}_n$ is the ML estimator of θ and the elements of the matrix $B(\theta)$ are

$$b_{js}(\theta) = \frac{1}{\sqrt{p_j(\theta)}}\frac{\partial p_j(\theta)}{\partial \theta_k}, \quad j = 1, 2, \cdots, k > s,$$

and

$$nJ(\theta) = nB^T(\theta)B(\theta)$$

is the *Fisher's information matrix* of the vector of frequencies ν and nI is the Fisher's information matrix of $\mathbf{X}$. The statistic $Y_n^2(\hat{\theta}_n)$ asymptotically follows a chi-squared distribution with $k-1$ degrees of freedom.

Another modification by Dzhaparidze and Nikulin (1974) valid for any $\sqrt{n}$-consistent estimator $\tilde{\theta}_n$ of θ (based on ungrouped data) showed that the statistic

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

in limit as $n \to \infty$ follows a χ^2_{k-s-1}, which coincides the Pearson-Fisher's test for based on grouped data (see Dzhaparidze and Nikulin (1992)). Voinov et al. (2009) showed that this test is not powerful for equiprobable intervals but is rather powerful with alternative hypothesis and with Neyman-Pearson classes (Greenwood and Nikulin, 1996).

From the literature one can see that a lot of research has been done on the modification of Pearson's chi-squared test. RRN-statistic is commonly used modified chi-squared test and many articles has been published based on RRN-statistics (see for example Dzhaparidze and

Nikulin (1982), Greenwood and Nikulin (1996), Mirvaliev (2001), Nikulin and Voinov (2006), Voinov et al. (2012)). In addition, the statistic $Y_n^2(\hat{\theta}_n)$ has a particularly convenient form when we construct a chi-square test with random cell boundaries for continuous distributions and for censored data, which we commonly have in reliability and survival analysis.

In Chapter-2 the RRN-test is applied for right censored data used in reliability and survival analysis for different parametric family and also in presence of covariates. For randomly censored samples, first modification in chi-squared test was proposed by Habib and Thomas (1986) based on the differences of Kaplan-Meier estimate $\hat{F}_n(t)$ and parametric ML estimators of survival functions $F(t, \hat{\theta}_n)$. Akritas (1988) proposed a chi-squared statistic based on the idea of comparing the observed and expected number of failures in each class. Hjort (1990) developed a chi-squared type statistic to test the validity of the parametric model for life history data based on the cumulative hazard process. Kim (1993) also proposed the chi-squared goodness-of-fit test based on the product limit estimator. Also other researchers like Hollander and Peña (1992), and Nikulin and Solev (1999) proposed modified chi-squared type tests for censored data. Bagdonavicius et al. (2010a) suggested modified chi-squared tests for randomly censored data. This tested is explained and applied on Arm-A head and neck cancer data in the next chapter 2.

4 Survival Analysis

Survival analysis also named as time-to-event analysis is a statistical method for data analysis where the outcome variable of interest is the time to the occurrence of an event. This method is applied in a number of applied fields, such as medicine, public health, epidemiology, engineering, and actuarial science. For example, time to event can be time to death, or time until the recurrence of some disease in medical science and in social sciences, it can be marriage, divorce etc. The survival analysis also contributed in the development of engineering systems by finding the lifetime of some machine to increase the reliability of the system.

Most of the development in the analysis of survival data has been made in second half 20^{th} century. In 1950, Berkson and Gage (1950) proposed a non-parametric method to compute the life table for analyzing survival data. Kaplan and Meier (1958) is another non-parametric method for survival curve. But when we have covariates along with the survival times, non-parametric methods are no more useful and we need semi-parametric and parametric regression methods. The standard regression is not adaptable to the survival data due to the presence of censoring and due the lack of normality for the survival time. Normally we have right censored and left truncated data and we know how to compute nonparametric estimator for them. But there is different method of estimation in interval censored data that is Turnbull estimator (Turnbull, 1976). Also Huber, Solov and Vonta (2006) studied the interval censored and truncated data.

Cox (1972) extended the methods of the non-parametric estimates to regression type arguments. He proposed a simple model and made no assumptions about the baseline hazard of individuals and only assumed that the hazard functions of different individuals remained proportional and constant over time. That's why this is also called proportional hazard model. Modifications in Cox PH model are made with time such as stratified Cox model (Kleinbaum and Klein, 2005) or Cox model with time-dependent variables (Tsiatis, 2006).

Accelerated failure time model (AFT) is a good alternative to the PH model (Tsiatis, 2006) for the analysis of survival time data or reliability data. In PH model we measure the direct effect of the explanatory variables on the hazard function while in the AFT models we measure that on the survival time. This provides us an easier interpretation of the results. AFT model is widely used in engineering systems to improve the reliability and increase the quality of the systems. Time-to-failure data under normal conditions of the systems or products is very time consuming due to their normal life. Due to this reason accelerated life testing (ALT) has been used to find the failure times in a short time by increasing the stress (temperature, pressure, dose of medicine etc.) on particular product or system.

A comprehensive work on AFT models is done by Bagdonavicius and Nikulin (1994, 1995, 2002) where they explained the construction of AFT model, failure time regression analysis, accelerated degradation models and gave the comparison of AFT model with various proportional hazard models. Also ALT is described very well in the literature such as Meeker and Escobar (1998), Lawless (2003), and Nelson (1990).

The Accelerated life testing of technical or bio-technical systems is an important practical statistical method of estimation of the reliability and the quality improvement of new systems without having to wait the operating life of an item. The ALT has been recognized as a necessary activity to ensure the reliability of electronic products used in military, aerospace, automotive and laptop computers. The accelerated testing of electronic products offers great potential to improve the quality in quick time. It is evident that the extrapolating reliability or quality from the ALT always carries the risk that the accelerated stresses do not properly excite the failure mechanism which dominates at operating (normal) stresses.

In ALT the choice of a good regression model is more important than in survival analysis. For example, in ALT units are tested under accelerated stresses which shorten the life. Using such experiments the life under the usual stress is estimated using some regression model. The values of the usual stress are often not in the range of the values of accelerated stresses, since the wide separation between experimental and usual stresses is possible. So if the model is misspecified, the estimators of survival under the usual stress may be very bad. In ALT we use the word stress for the covariates.

In chapter-3 the parametric AFT model based on Birnbaum-Saunders distribution is presented along with different type of stresses used in ALT. The survival function under normal stress and its confidence interval from BS-AFT model is estimated. A goodness-of-fit is given for this model.

5 Redundant Systems

The number of extra or reserve components with the same function in a parallel structure is called a redundancy. The use of extra components can enable a system to operate properly even in the case of failure of some components. The system composed of redundancy is called a redundant system. So in this way one can increase the reliability of the whole system, where reliability is a statistical probability i.e. $R(t) = 1 - F(t)$. Redundancy exists in living organs also, for example, in living organisms, vital organs and tissues (such as the liver, kidney, or pancreas) consist of many cells performing one and the same specialized function (Gavrilov and Gavrilova, 2006). There are different systems' structures for redundancy. Series and parallel structures are the basis for building more complicated structures (Figure-1). As the failure accumulated the redundancy

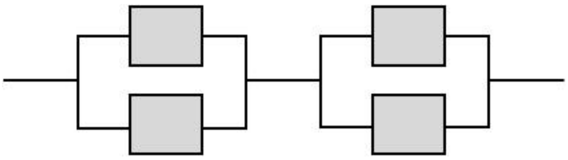

Figure 1 – A Series-Parallel System Structure

in the number of elements disappear. The number of extra components are determined by the cost and reliability measures during design. The positive effect of a system's redundancy is damage tolerance, which decreases the risk of failure and increases life span. Also redundancy can improve the monitoring system because the standby units are also monitoring the applications. It is supposed that the system's reliability increases with the number of redundant components (see Figure 2).

Other system's structure is standby or passive redundancy where a redundant unit is activated only when main unit fails and the redundant unit keep the system working. Here it is very important to determine the operating state of the standby unit i.e. hot, cold or warm. The main unit is working in hot conditions. If standby unit is functioning in hot condition (same as the main unit) then there is an equal probability of failure for both units. If it is placed at cold stage then it will take time to come into a hot state (switching time) which may cause an

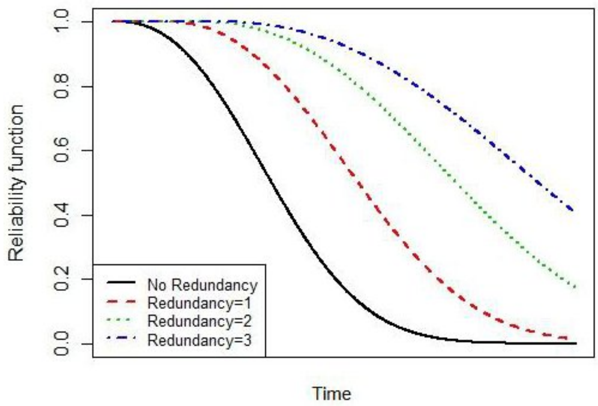

Figure 2 – Effect of redundancy on Reliability function (simulated).

interruption in functioning of the system. One need to consider the reliability of the switching mechanism that activates the standby units. It also requires switching time but less than that of cold standby (see Figure 4). Reijns and Gemund (2007) showed that cold standby redundancy provides a better mean time reliability than hot redundancy. In warm state, standby units operate under the less stress than the main unit. In Chapter-4 the switching mechanism is defined by the Sedyakin'a principal (Sedyakin, 1966) and AFT model. Let $x(\cdot), y(\cdot), z(\cdot)$ be the three levels of stresses for standby unit under hot, warm and cold states respectively. The functioning states and effect of switching on reliability from lower stress to higher of the standby unit is explained in Figure 3 and Figure 4 respectively. Different kinds of stress are used to calculate the reliability of engineering system. Two commonly used stresses are shown in the Figure 5.

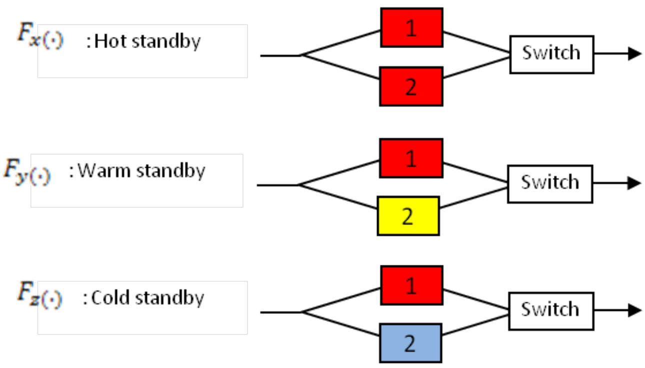

Figure 3 – The different states of standby unit ($F_{x(\cdot)} \geq F_{y(\cdot)} \geq F_{z(\cdot)}$)

Exponential family is commonly used by engineers to measure reliability characteristics where one have a constant failure rate. Gnedenko, Belyaev and Solov'ev (1968), and Kozlov and Ushakov (1970) give a good mathematical base to study the redundant systems and estimation of reliability characteristics. Recent research by Bagdonavicius, Masiulaityle, and Nikulin (2008a,

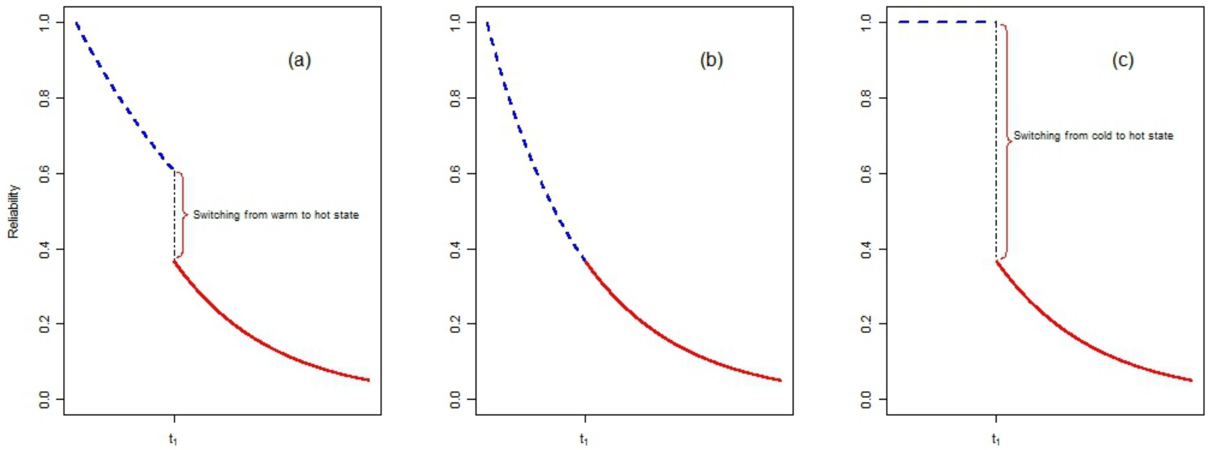

Figure 4 – The switching of three possible states of standby units to the hot state, t_1 is switching moment where standby unit replaces the main one. (a) From warm to hot state, (b) from hot to hot state (no effect before and after as it is already working under the same stress as that of hot), (c) from cold (off mode) to hot state (big difference of stresses can cause interruption during switching).

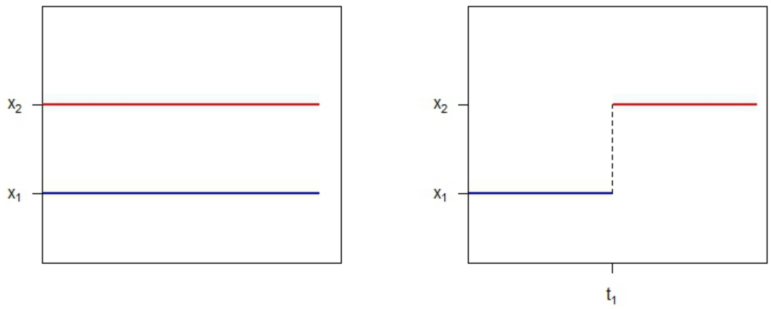

Figure 5 – Two parallel stresses : $x_2 > x_1$ (left), and increasing step stress at two levels (right)

2008b, 2009, 2010) is done for the statistical analysis of redundant systems $S(1, m-1)$ with warm standby units for exponential, Weibull and loglogistic family of distributions. They succeeded to apply the techniques of accelerated trials for the analysis of redundant system. Saaidia, Nikulin, and Tahir (2011) applied the same techniques with generalized Weibull model for the analysis of redundant system. Also Nikulin, Saaidia, and Tahir (2011c) give some simulated results for several models with unimodal hazard rate function. Chapter-4 contains the procedure of parameters' estimation, distribution function and confidence interval of the redundant system with one main unit and one standby unit operating in the warm conditions with Birnbaum-Saunders distributions. Notice that aging is another factor affecting the reliability which is a natural process with the age but this is not studied in the chapter of redundant systems.

Chapter 1

Parametric Failure Time Models

1 Introduction

In this chapter, various parametric models that are commonly applied to survival analysis and reliability are presented. *Survival function* for absolutely continuous random variable T of failure time can be defined as

$$S(t) = P\{T \geq t\}, t \geq 0;$$

i.e. the probability of survival up to fixed time t.

The *cumulative distribution function* $F(t) = 1 - S(t)$ is the probability of failure before time t and the probability density function is

$$f(t) = \frac{dF(t)}{dt} = -S'(t).$$

Mostly failure time data is plotted in terms of *hazard rate function* which can be defined as

$$\lambda(t) = \lim_{h \to 0} \frac{P(t \leq T < t + h | T \geq t)}{h} = \frac{f(t)}{S(t)}, \tag{1.1}$$

which means the risk of failure of the remaining subjects and the *cumulative hazard function* is

$$\Lambda(t) = \int_0^t \lambda(u) du = -\ln S(t).$$

2 Exponential Distribution

If the survival time T has a constant hazard rate then it is *exponentially distributed* with the following survival and hazard function

$$S(t, \theta) = e^{-t/\theta}, \qquad \lambda(t, \theta) = \frac{1}{\theta}, \quad \theta > 0, \, t \geq 0,$$

respectively. The hazard function is constant and is reciprocal of the mean. Also the exponential distribution has a memoryless property because the instantaneous failure rate is independent of time t. The hazard function, density function, and the survival function of the one parameter exponential model is shown in Figure 1.1.

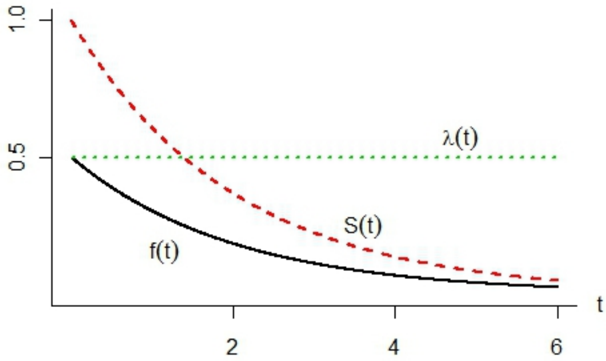

Figure 1.1 – The density, survival and hazard function of exponential model.

3 Gamma Distribution

Gamma distribution like Weibull distribution is also the generalization of the exponential distribution and belongs to shape-scale distribution family with the survival function

$$S(t) = 1 - \frac{1}{\Gamma(\nu)} \int_0^{t/\theta} u^{\nu-1} e^{-u} du, \quad t \geq 0; \quad \theta, \nu > 0,$$

and the hazard function

$$\lambda(t) = \frac{\frac{1}{\theta^\nu} t^{\nu-1} e^{-t/\theta}}{\Gamma(\nu) - \int_0^{t/\theta} u^{\nu-1} e^{-u} du}, \quad t > 0 \tag{1.2}$$

This distribution has a lot of applications in many field field (see Johnson et al., 1995). The hazard function is monotone increasing from zero to $1/\theta$ if $\nu > 1$, and monotone decreasing from ∞ to $1/\theta$ if $\nu < 1$ as t becomes large. If $\nu = 1$, the gamma distribution reduces to exponential distribution with constant hazard rate. Figure 1.2 shows the behavior of hazard function of gamma distribution. This distribution is difficult to apply with censored data due to the complexity of hazard function.

4 Weibull Distribution

Two parameter *Weibull distribution* is probably the most widely used for lifetimes especially to model fatigue failures, ball bearing failures. The survival function and the hazard function

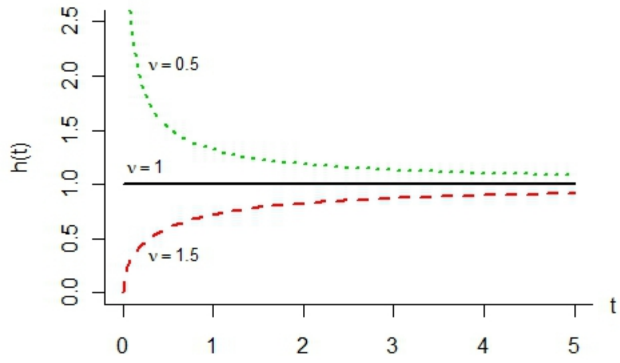

Figure 1.2 – Hazard functions of gamma distribution ($\theta = 1$).

respectively is written as

$$S(t;\theta,\nu) = \exp\{-(\frac{t}{\theta})^{\nu}\}, \quad \lambda(t,\theta,\nu) = \frac{\nu}{\theta^{\nu}}t^{\nu-1}, \quad (\theta,\nu > 0);\, t \geq 0.$$

ν is called the shape parameter and θ is the scale parameter. For $\nu = 1$ this is equal to the exponential distribution. Weibull distribution is capable of modeling decreasing failure rate (DFR) ($\nu < 1$), constant failure rate (CFR) ($\nu = 1$) and increasing failure rate (IFR) ($\nu > 1$) behavior, shown in Figure 1.3. That why this is equally applied in demographic studies (Gavrilov and Gavrilova, 2006).

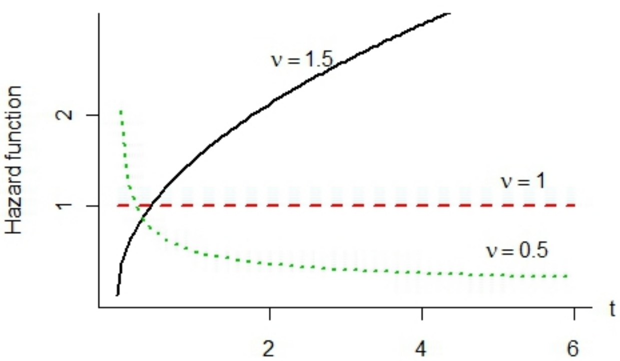

Figure 1.3 – Hazard functions of Weibull distribution for $\theta = 1$.

5 Lognormal Distribution

A random variable T is *lognormally distributed* if $Y = \log T$ is normally distributed and the survival function and the hazard function can be written as

$$S(t) = 1 - \Phi(\ln(t/\theta)^{\nu}), \quad \lambda(t) = \nu t^{-1}\frac{\phi(\ln(t/\theta)^{\nu})}{1-\Phi(\ln(t/\theta)^{\nu})}, \; (\theta, \nu > 0); \, t \geq 0,$$

where ϕ and Φ are the pdf and cdf of the standard normal distribution. The hazard function has $\bigcap$-shape (Figure 1.4). The computations for lognormal distribution is relatively complex particularly with censoring. In this case one can use the loglogistic distribution which gives a good approximation to the lognormal distribution.

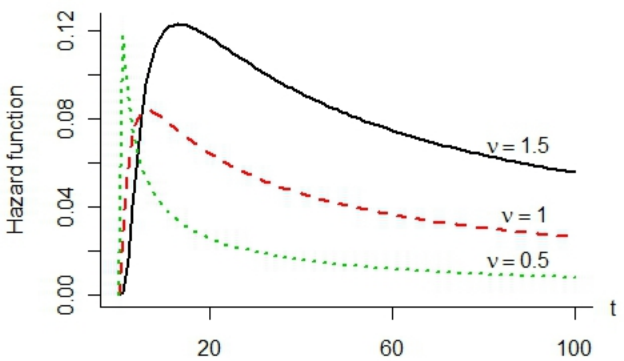

Figure 1.4 – Hazard functions of lognormal distribution for $\theta = 10$.

6 Loglogistic Distribution

The survivor and hazard functions are, respectively,

$$S(t) = \frac{1}{1+(\frac{t}{\theta})^{\nu}}, \quad \text{and} \quad \lambda(t) = \frac{\nu}{\theta^{\nu}}t^{\nu-1}\frac{1}{1+(\frac{t}{\theta})^{\nu}}, \quad (\theta, \nu > 0); \quad t \geq 0.$$

The *loglogistic distribution* has the advantage over the lognormal distribution due to its simple explicit form of hazard function. This distribution also has a $\bigcap$-shape hazard function. It is monotone decreasing from ∞ if $\nu < 1$. If $\nu > 1$, the hazard function is similar to that of lognormal i.e. it increases from zero to a maximum $t = \theta(\nu - 1)^{1/\nu}$ and then decreases to zero. This is show in Figure 1.5.

7 Exponentiated Weibull Distribution

Another interesting parametric family of distribution is the family of *exponentiated Weibull distributions*, induce by Efron (1988) and studied by Mudholkar, Srivastava and Kollia (1996).

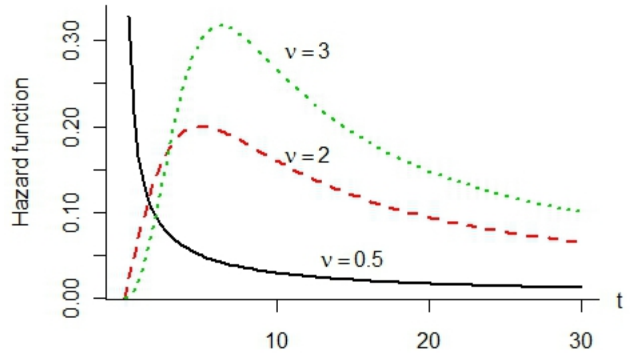

Figure 1.5 – Hazard functions of loglogistic distribution for $\theta = 5$.

The survival function from this family of distributions is given by formula

$$S(t,\theta,\nu,\gamma) = 1 - \left\{1 - \exp[-(\frac{t}{\theta})^{\nu}]\right\}^{1/\gamma}, \quad t \geq 0, \quad \theta,\nu,\gamma > 0,$$

and the hazard function is

$$\lambda(t,\theta,\nu,\gamma) = \frac{\nu \left\{1 - \exp[-(\frac{t}{\theta})^{\nu}]\right\}^{1/\gamma - 1} \exp[-(\frac{t}{\theta})^{\nu}](\frac{t}{\theta})^{\nu-1}}{\gamma\theta \left\{1 - (1 - \exp[-(\frac{t}{\theta})^{\nu}])^{1/\gamma}\right\}} \tag{1.3}$$

The hazard function of exponentiated Weibull distributions has also nice properties as that of generalized Weibull distributions (see Figure 1.6).

With $\nu > 1, \nu \geq \gamma$ the hazard rate is increasing from 0 to ∞.
With $\nu = 1, \gamma \leq 1$ the hazard rate is increasing from 0 to θ^{-1}.
With $0 < \nu < 1, \nu < \gamma$ the hazard rate is decreasing from ∞ to 0.
With $0 < \nu < 1, \nu = \gamma$ the hazard rate is decreasing from θ^{-1} to 0.
With $\gamma > \nu > 1$ the hazard rate is decreasing from ∞ to its minimal value $c > 0$ and then increases to ∞, it is $\bigcup -$*shaped*. With $\gamma < \nu < 1$ the hazard rate is increasing from 0 to its maximal value $c > 0$ and then decreases to 0, it is $\bigcap -$*shaped*.

8 Generalized Weibull Distribution

Three parameter *generalized Weibull distribution* was proposed recently as an alternative to exponentiated Weibull distribution by Bagdonavicious and Nikulin (2002) and its hazard function can be monotone, $\bigcap$-shaped and $\bigcup$-shaped according to the values of its parameters. All the moments of this distribution are finite. The survival function and hazard function,

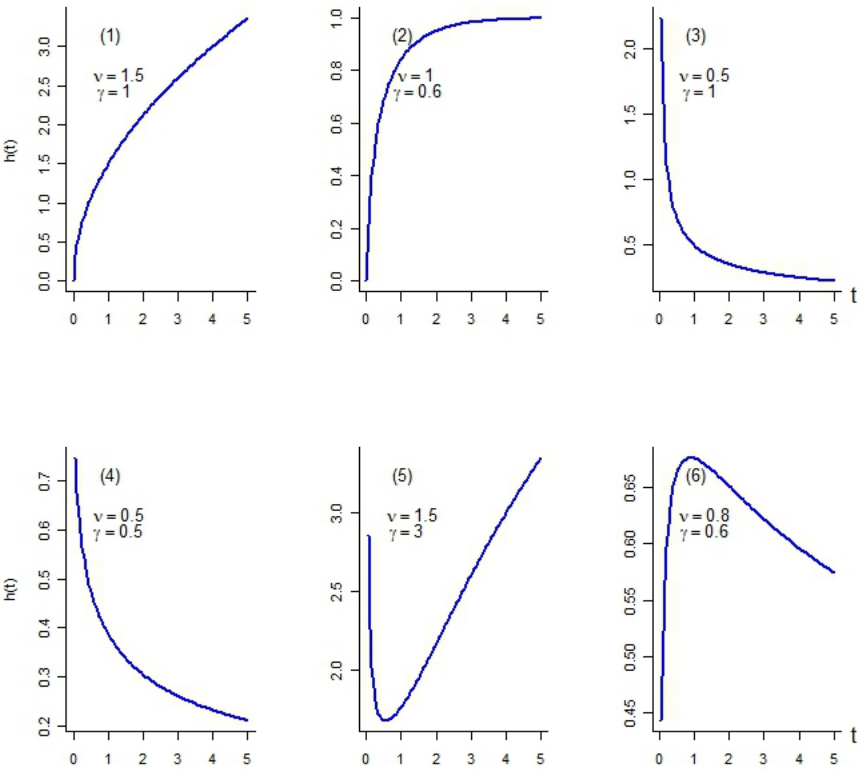

Figure 1.6 – Hazard functions of exponentiated Weibull distribution ($\theta = 1$).

respectively can be written as

$$S(t, \theta, \nu, \gamma) = \exp\left\{1 - \left(1 + (\frac{t}{\theta})^{\nu}\right)^{1/\gamma}\right\}, \quad t \geq 0;\ \theta, \nu, \gamma > 0$$

and

$$\lambda(t, \theta, \nu, \gamma) = \frac{\nu}{\gamma\theta^{\nu}} t^{\nu-1} \left(1 + (\frac{t}{\theta})^{\nu}\right)^{1/\gamma - 1}. \tag{1.4}$$

This distribution is easy to use with censored data due to simple form of hazard rate function. The hazard Rate has following properties and all are shown in the Figure 1.7 respectively :

If $\nu > 1, \nu > \gamma$, then the hazard rate increases from 0 to ∞.
If $\nu = 1, \gamma < 1$, then the hazard rate increases from $(\gamma\theta)^{-1}$ to ∞.
If $0 < \nu < 1, \nu < \gamma$, then the hazard rate decreases from ∞ to 0.
If $0 < \nu < 1, \nu = \gamma$, then the hazard rate decreases from ∞ to $1/\theta$.
If $\gamma > \nu > 1$, then the hazard rate increases from 0 to its maximum value and then decreases to 0 i.e. $\bigcap$ shape.
If $0 < \gamma < \nu < 1$, then the hazard rate decreases from ∞ to its minimum value and then

increases to ∞ i.e. $\bigcup$ shape.

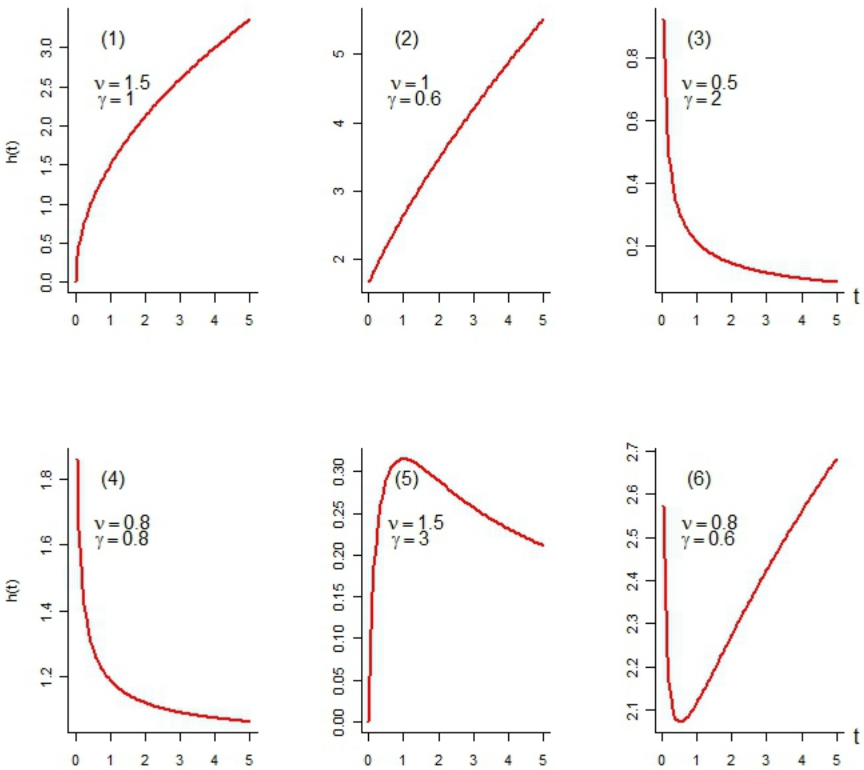

Figure 1.7 – Hazard functions of generalized Weibull distribution ($\theta = 1$).

9 Birnbaum-Saunders Distribution

The survival function of two-parameter fatigue life distribution known as *Birnbaum-Saunders distribution* (Birnbaum and Saunders, 1969) can be written as

$$S(t;\alpha,\beta) = 1 - \Phi\left[\frac{1}{\alpha}\left\{\left(\frac{t}{\beta}\right)^{\frac{1}{2}} - \left(\frac{\beta}{t}\right)^{\frac{1}{2}}\right\}\right], \quad 0 < t < \infty, \quad \alpha,\beta > 0,$$

where α is the shape parameter, β is the scale parameter and $\Phi(x)$ is the standard normal distribution function. The case where $\beta = 1$ is called the standard fatigue life distribution. The hazard rate function can be written as

$$\lambda(t;\alpha,\beta) = \frac{\frac{1}{2\sqrt{2\pi}\ \alpha\beta}\left\{\left(\frac{\beta}{t}\right)^{\frac{1}{2}} + \left(\frac{\beta}{t}\right)^{\frac{3}{2}}\right\}\exp\left[-\frac{1}{2\alpha^2}\left(\frac{t}{\beta} + \frac{\beta}{t} - 2\right)\right]}{1 - \Phi\left[\frac{1}{\alpha}\left\{\left(\frac{t}{\beta}\right)^{\frac{1}{2}} - \left(\frac{\beta}{t}\right)^{\frac{1}{2}}\right\}\right]}. \tag{1.5}$$

In recent year a lot of work is done the Birnbaum-Saunders(BS) distribution (see for example, Balakrishnan et al. (2007, 2009), Kundu et al. (2008), Lemonte et al. (2007)). Desmond (1986) worked on the relationship between Birnbaum-Saunders distribution and the family of inverse Gaussian distributions. Also Volodin and Dzhungurova (2000) introduced a family of so called crack distributions with the normal distribution, the inverse Gaussian distribution and the BS distribution. BS distribution describes the total time that passes until some type of cumulative damage produced by the development and growth of a dominant crack, surpasses a threshold, and causes a failure. Fatigue failure is due to repeated applications of a common cyclic stress pattern. Under the influence of this cyclic stress a dominant crack in the material grows until it reaches a critical size w, at that point fatigue failure occurs. The crack extension in each cycle are random variables and are statistically independent. Also the total extension of the crack is approximately normally distributed. This is not only applied in reliability but in a variety of fields. This distribution has unimodal density function and hazard rate function i.e. $\bigcap$-shaped hazard function. It is positively skewed distribution. The mean, variance, coefficient of variation (CV), skewness and kurtosis of the BS distribution are, respectively, given as

$$\mu = \frac{\beta}{2}(\alpha^2+2), \quad \sigma^2 = \frac{\beta^2}{4}(5\alpha^4+4\alpha^2), \quad CV = \frac{\sqrt{5\alpha^4+4\alpha^2}}{(\alpha^2+2)},$$
$$\mu_3 = \frac{16\alpha^2(11\alpha^2+6)}{(5\alpha^2+4)^3}, \quad \mu_4 = 3+\frac{6\alpha^2(93\alpha^2+43)}{(5\alpha^2+4)^2}.$$

Note that the CV, skewness, and kurtosis are independent of the scale parameter, β. If $T \sim BS(\alpha, \beta)$, then $T^{-1} \sim BS(\alpha, \beta^{-1})$.

Not a lot of work has been published on the analysis of censored data for the BS distribution which is common in reliability and survival analysis. Wang et al. (2006) give a comparison between the BS and the Weibull models in a real data application with censoring. Also they discussed the ML estimation of the parameters. Nikulin et al. (2011b) give a modifies chi-squared goodness-of-fit test for BS distribution. More details can be found from Johnson et al. (1995). The hazard function is shown in Figure 1.8.

10 Inverse Gaussian Distribution

The survival function of the unimodel *inverse Gaussian distribution* is

$$S(t) = \Phi\left(-\sqrt{\frac{\nu}{t}}\left(\frac{t}{\theta}-1\right)\right) - \exp(\frac{2\nu}{\theta})\Phi\left(-\sqrt{\frac{\nu}{t}}\left(\frac{t}{\theta}+1\right)\right), \; \theta, \nu > 0; \; t > 0$$

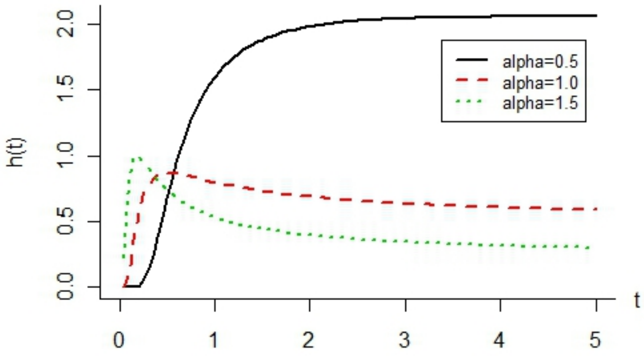

Figure 1.8 – Hazard functions of BS distribution ($\beta = 1$).

and the hazard rate function is

$$\lambda(t)) = \frac{\left(\frac{\nu}{2\pi t^3}\right)^{\frac{1}{2}} \exp\{-\frac{\nu(t-\theta)^2}{2\theta^2 t}\}}{\Phi\left(-\sqrt{\frac{\nu}{t}}\left(\frac{t}{\theta}-1\right)\right) - \exp(\frac{2\nu}{\theta})\Phi\left(-\sqrt{\frac{\nu}{t}}\left(\frac{t}{\theta}+1\right)\right)}, \qquad t > 0.$$

For all unimodel distribution like LN, LL, PGW, EW, the hazard rate increases from 0 to its maximum value and then decreases to zero, but for IG distribution the hazard rate increases from 0 to its maximum value and then decreases to $\nu/2\theta^2$ (Figure 1.9). In, Voinov and Nikulin (1993) one can find in table A6, the unbiased estimators of functions of parameters ν and θ of the inverse Gaussian distribution.

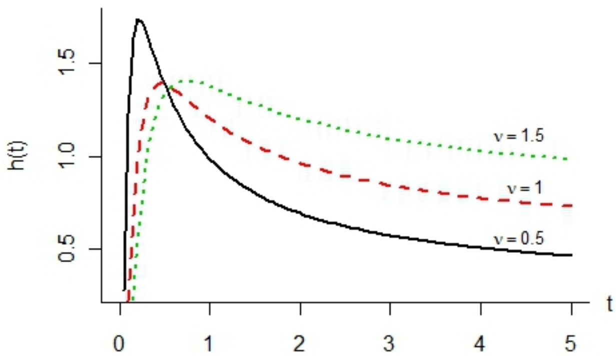

Figure 1.9 – Hazard functions of inverse Gaussian distribution ($\theta = 1$).

11 Gompertz-Makeham Distribution

Gompertz model (Gompertz, 1825) of aging is widely used in demography and other scientific disciplines e.g. medical sciences, survival analysis, actuarial sciences and reliability. In Gompertz Distribution mortality rate or hazard rate increases exponentially with age and can be formulated as

$$\lambda(t) = \theta e^{\nu t}, \quad \theta, \nu > 0; \quad t > 0, \tag{1.6}$$

where θ is known as the baseline mortality and ν is the age specific growth rate of the force of mortality. The Gompertz law has been the main demographic model since its discovering to fit the human mortality (see e.g. Gavrilov and Gavrilova (2001), Ricklefs and Scheuerlein (2002)). William Makeham (1860) modified the Gompertz model considering some other causes of failure independent of age by proposing the so called *Gompertz-Makeham* model as

$$\lambda(t) = \gamma + \theta e^{\nu t}, \quad \gamma, \theta, \nu > 0; \quad t > 0. \tag{1.7}$$

Here γ is a constant and non-aging component of failure rate and the second term $\theta e^{\nu t}$ is the Gompertz function depending on age (aging factor). Failure rate is shown in Figure 1.10. In chapter 2 one section is dedicated to the statistical inference of Gompertz-Makeham model with the example from demography (life table) and also from reliability with censored data.

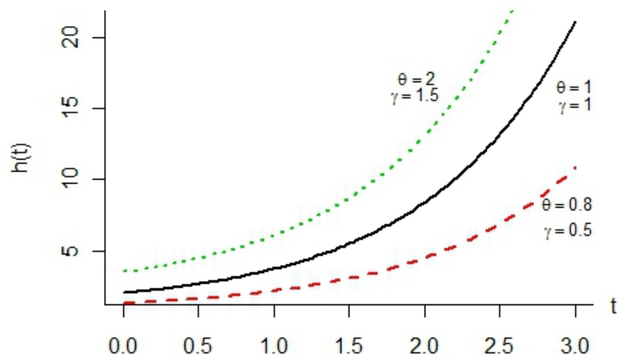

Figure 1.10 – Hazard functions of Gompertz-Makeham distribution ($\nu = 1$).

12 Some Other Distribution

There are many other distributions that can be used to model the survival or reliability data but here we attempted some commonly used distributions.

Extreme value distribution : The two parameter extreme value distribution refers to the distribution of the logarithm of a Weibull random variable, and belongs to the location scale family of distributions. The survival function and hazard function can be written as :

$$S(t) = e^{-e^{\frac{t-\mu}{\sigma}}}, \qquad \lambda(t) = \frac{1}{\sigma} e^{\frac{t-\mu}{\sigma}}, \quad -\infty < t < +\infty, \ \sigma > 0.$$

With $\mu = 0$ and $\sigma = 1$, it is called a smallest or standard extreme value distribution.

Rayleigh distribution : This is another special case of the Weibull distribution when the shape parameter is fixed at 2. The density function and hazard function is

$$f(t) = \frac{2t}{\theta^2} \exp\{-(\frac{t}{\theta})^2\}, \quad \lambda(t) = \frac{2t}{\theta^2}, \quad t \geq 0,\ \sigma > 0.$$

Rayleigh distribution can be used to model the magnitude of radial error (wind speed).

Pareto distribution : The Pareto distribution also known as the power law and is mostly applied in actuarial science and life testing. Its survival function and density function is given by

$$S(t) = \left(\frac{\theta}{t}\right)^{\nu}, \quad f(t) = \frac{\nu\theta^{\nu}}{t^{\nu+1}}, \quad \theta, \nu > 0,\ t \geq \theta.$$

The hazard function $\lambda(t) = \frac{\nu}{t}$, $t \geq \theta$ is a decreasing function.

Chapter 2

Chi-squared Type Goodness-of-Fit Tests

1 Pearson Statistic and Modifications

1.1 Introduction

Chi-square goodness-of-fit test can be the first step in further data analysis. It is a useful technique to see if the observed data are representative of a particular model or distribution. In hypothesis testing formulation of null and alternative hypotheses is required but GoF test can be proceeded only with the null hypothesis which tells that whether preselected model fits the data or not. The famous chi-square Gof test has been used in almost all areas of research since its discovery by Karl Pearson (1900). The limit distribution of Pearson statistic is chi-square only when we have the simple hypothesis. But it would not be chi-squared if the unknown parameters are to be estimated by a sample (Fisher, 1928) because the limit distribution depends on the method of estimation of the parameters. However, if the parameters are estimated through the minimum chi-square method or grouped maximum likelihood estimation method for grouped data, the limit distribution will remain chi-squared but the degree of freedom is reduced by the number of estimated parameters (Cramer, 1946). Moreover, for fixed grouping intervals Chernoff and Lehmann (1954) and for random grouping intervals Roy (1956) showed that if the unknown parameters are estimated by the maximum likelihood method the limit distribution is totally changed (see also LeCam et al. (1983)). To overcome the problem of parameter estimation for ungrouped data and the limit distribution of the statistic many modifications have been made in classical Pearson chi-square test and a variety of procedures have been implemented in many software applications. Nikulin (1973a, 1973b, 1973c) and Rao and Robson (1974) independently proposed a modification in Pearson statistic. Their statistic is now commonly used as Rao-Robson-Nikulin (RRN) statistic (Drost, 1988 ; Van der Vaart, 1998 ; Voinov et al., 2008) and the test in the limit follow a chi-squared distribution. Nikulin used a consistent estimate obtained from ungrouped data i.e. MLE. But in 1976 it was shown by Hsuan and Robson (1976) that in case of moment-type estimates the resulting modified statistic would be different. Dzhaparidze and Nikulin (1974) proposed a modification of Pearson's statistic (DN test) which follow a chi-square distribution for any $\sqrt{n}$-consistent estimator of unknown parameter from ungrouped data. Other important modification in chi-square goodness-of-fit test was proposed by Mirvaliev (2001). In this chapter RRN-test is focussed on censored data following Habib and Thomas (1986, Hjort (1990), Van der Vaart (1998), Bagdonavicius and Nikulin (2011) and Bagdonavicius et al. (2010a).

GoF test is commonly used in survival analysis and reliability and it is more critical in accelerated failure time (AFT) models where a small deviation in selecting the adequate model can totally change the final decision. Bagdonavicius et al. (2010a) extended the idea of RRN-test to the censored data. The details of the test for censored data are given in section 2. Although Gof test can be applied for both discrete and continuous distributions but here only continuous distributions are considered.

1.2 Pearson Statistic

Let $\mathbf{X} = (X_1, \cdots, X_n)^T$ be a random sample and suppose we want to test H_0 that the i.i.d. random variables $X_1, \cdots, X_n$ follow the same distribution, i.e.

$$\mathbf{P}\{X_i \leq x\} = F(x, \boldsymbol{\theta}), \quad \boldsymbol{\theta} = (\theta_1, \cdots, \theta_s)^T \in \Theta \subset R^s, \tag{2.1}$$

where $\boldsymbol{\theta}$ is the vector of parameters of dimension s of some distribution function F.

Let $-\infty = a_0 < a_1 < \cdots < a_{k-1} < a_k = \infty$ (k>s+1) be the boundary points of the groping intervals $I_1, I_2, \cdots, I_k$ in $(-\infty, \infty)$, and $\boldsymbol{\nu} = (\nu_1, \cdots, \nu_k)^T$ is the vector of frequencies

$$\nu_j = \sum_{i=1}^{n} 1_{\{X_i \in I_j\}}, \quad j = 1, \cdots, k$$

which lies in each of the successive intervals

$$(a_0, a_1), [a_0, a_1), \cdots, [a_{k-1}, a_k),$$

and the probability with respect to $F(x, \boldsymbol{\theta})$ is

$$p_j(\boldsymbol{\theta}) = \mathbf{P}\{X_i \in I_j | H_0\} = \int_{I_j} dF(x; \boldsymbol{\theta}) = \int_{I_j} f(x; \boldsymbol{\theta}) dx = \mathbf{P}_{\boldsymbol{\theta}}(a_{j-1} \leq X_i \leq a_j).$$

It is clear that $\boldsymbol{\nu}^T \mathbf{1}_k = n$, $(k > s+1)$ and

$$\mathbf{p} = \mathbf{p}(\boldsymbol{\theta}) = (p_1(\boldsymbol{\theta}), \cdots, p_k(\boldsymbol{\theta}))^T.$$

Under H_0 the statistic $\boldsymbol{\nu}$ follows the multinomial distribution $M_k(n, \mathbf{p}(\boldsymbol{\theta}))$, i.e.

$$\mathbf{P}\{\nu_1 = m_1, \cdots, \nu_k = m_k\} = \frac{n!}{m_1!, \cdots, m_k!}(p_1(\boldsymbol{\theta}))^{m_1}, \cdots, (p_k(\boldsymbol{\theta}))^{m_k},$$

$0 < p_j < 1$, $0 \leq m_i \leq n$, and $m_1 + \cdots + m_k = n$. We can write $\mathbf{E}\nu = n\boldsymbol{p}$ and covariance matrix as

$$\Sigma = \mathbf{E}(\nu - n\boldsymbol{p})(\nu - n\boldsymbol{p})^T = n(\mathbf{P} - \boldsymbol{p}\boldsymbol{p}^T),$$

where $\mathbf{P}$ is the diagonal matrix with the diagonal elements $p_1, \cdots, p_k$. The rank of Σ is $k-1$.

For each $\boldsymbol{\theta} \in \Theta$ we have the random vector

$$X_n(\boldsymbol{\theta}) = \left(\frac{\nu_1 - np_1(\boldsymbol{\theta})}{\sqrt{np_1(\boldsymbol{\theta})}}, \cdots, \frac{\nu_1 - np_k(\boldsymbol{\theta})}{\sqrt{np_k(\boldsymbol{\theta})}}\right)^T,$$

and the Pearson random variable

$$X_n^2(\boldsymbol{\theta}) = X_n^T(\boldsymbol{\theta})X_n(\boldsymbol{\theta}) = \sum_{j=1}^{k} \frac{(\nu_j - np_j(\boldsymbol{\theta}))^2}{np_j(\boldsymbol{\theta})}.$$

If $\boldsymbol{\theta}_0$ is the true value of the population parameter $\boldsymbol{\theta}$ under H_0, then we reject H_0 if

$$\lim_{n\to\infty} \mathbf{P}\{X_n^2(\boldsymbol{\theta}_0) \leq x | H_0\} = \mathbf{P}(\chi^2_{k-1} \leq x\}.$$

As in practical cases $\boldsymbol{\theta}$ is unknown, so we need to estimate it using the data. In this situation the limit distribution of the test statistic $X_n^2(\boldsymbol{\theta}^*)$ depends on the asymptotic properties of the estimator $\boldsymbol{\theta}^*$.

Two commonly used methods to construct a sequence $\{\boldsymbol{\theta}_n^*\}$ of estimators of $\boldsymbol{\theta}$ from grouped data are the *minimum chi-squared method* - $\{\widetilde{\boldsymbol{\theta}}_n\}$ and the *multinomial maximum likelihood method* - $\{\widetilde{\widetilde{\boldsymbol{\theta}}}_n\}$, and for ungrouped data the method of moments - $\{\bar{\boldsymbol{\theta}}_n\}$ and *maximum likelihood estimation method* - $\{\hat{\boldsymbol{\theta}}_n\}$ can be used. The ML method is considered as a particular case of method of moment (Greenwood and Nikulin, 1996). Under general conditions

$$\{\boldsymbol{\theta}_n^*\} \xrightarrow{p} \boldsymbol{\theta}, \quad \text{as } n \to \infty,$$

and the random vectors

$$\sqrt{n}(\boldsymbol{\theta}_n^* - \boldsymbol{\theta}) \approx AN(\mathbf{0_s}, \mathrm{Var}_\theta).$$

1.2.1 Minimum Chi-Squared Method

We supposed that $\widetilde{\boldsymbol{\theta}}_n$ is an estimate of $\boldsymbol{\theta}$ which gives the minimum value of the random variable $X_n^2(\boldsymbol{\theta})$, i.e.

$$X_n^2(\widetilde{\boldsymbol{\theta}}_n) = \min_{\theta\in\Theta} X_n^2(\theta), \Longleftrightarrow \widetilde{\boldsymbol{\theta}}_n = argmin X_n^2(\theta).$$

This is called *the minimum chi-squared method* of estimating $\boldsymbol{\theta}$ and $\widetilde{\boldsymbol{\theta}}_n$ is the minimum chi-squared estimator.

We suppose the following conditions of Cramer (1946),

1. $0 < p_i(\boldsymbol{\theta}) < 1, \quad \sum_{i=1}^{k} p_i(\boldsymbol{\theta}) = 1, \quad \forall \quad i = 1, \cdots, k$ and $\boldsymbol{\theta} \in \Theta, \quad (k > s+1).$
2. $\frac{\partial^2 p_j(\theta)}{\partial\theta_l \partial\theta_{l'}}$ are continuous functions on Θ.
3. The rank of the Fisher information matrix

$$\mathbf{J}(\boldsymbol{\theta}) = B(\boldsymbol{\theta})^T B(\boldsymbol{\theta})$$

is s, where

$$B(\boldsymbol{\theta}) = \left[\frac{1}{\sqrt{p_i(\boldsymbol{\theta})}} \frac{\partial p_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_j}\right]_{k\times s}.$$

and $n\mathbf{J}(\boldsymbol{\theta}) = nB(\boldsymbol{\theta})^T B(\boldsymbol{\theta})$ is the Fisher information matrix of the stochastic vector ν.

Theorem 1.1 *Fisher (1928) showed that if the regularity conditions of Cramer hold then*

$$\lim_{n\to\infty} \mathbf{P}\{X_n^2(\tilde{\boldsymbol{\theta}}_n) \leq x | H_0\} = \mathbf{P}\{\chi^2_{k-s-1} \leq x\}.$$

1.2.2 Multinomial Maximum Likelihood Estimation

Cramer (1946) showed that the result of Fisher remains valid if instead of $\tilde{\boldsymbol{\theta}}_n$ we choose multinomial maximum likelihood estimator $\widetilde{\widetilde{\boldsymbol{\theta}}}_n$ which return the maximum value of the likelihood function of the multinomial distribution $M_k(n, \mathbf{p}(\widetilde{\widetilde{\boldsymbol{\theta}}}_n))$, i.e.

$$l(\widetilde{\widetilde{\boldsymbol{\theta}}}_n) = \sup_{\boldsymbol{\theta}\in\Theta} l(\boldsymbol{\theta}), \quad \widetilde{\widetilde{\boldsymbol{\theta}}}_n = \arg\max l(\boldsymbol{\theta}),$$

where

$$l(\boldsymbol{\theta}) = \frac{n!}{\nu_1! \cdots \nu_k!} (p_1(\boldsymbol{\theta}))^{\nu_1} \cdots (p_k(\boldsymbol{\theta}))^{\nu_k},$$

is the likelihood function of the statistic $\boldsymbol{\nu} = (\nu_1, \cdots, \nu_k)^T$.
Cramer has shown that the estimator of Fisher $\widetilde{\boldsymbol{\theta}}_n$ and the multinomial maximum likelihood estimator $\widetilde{\widetilde{\boldsymbol{\theta}}}_n$ are asymptotically equivalent and hence $\forall x > 0$

$$\lim_{n\to\infty} \mathbf{P}\{X_n^2(\widetilde{\widetilde{\boldsymbol{\theta}}}_n) \leq x | H_0\} = \mathbf{P}(\chi^2_{k-s-1} \leq x).$$

We see that the estimator $\widetilde{\widetilde{\boldsymbol{\theta}}}_n$ is obtained from the grouped data, and if the distribution $F(x, \boldsymbol{\theta})$ is continuous then the statistic $\nu = (\nu_1, \cdots, \nu_k)^T$ is not the sufficient statistic and hence the estimator $\widetilde{\widetilde{\boldsymbol{\theta}}}_n$ may not be the best estimator.

1.2.3 Method of Moments

A common way of fitting a parametric family to data is to use estimates of the parameters that yield moments of the fitted density that match the sample moments. The second and higher moments used in this method of estimation are usually taken as the central moments. In some distributions, the method of moment estimates are the same as the maximum likelihood estimates. Hsuan and Robson (1976) provided the test statistic explicitly for the exponential family of distributions when method of moment estimates coincide with MLEs which confirms the Nikulin's result Voinov and Pya (2004). We supposed the followiing regularity conditions :

1. The MMEs are $\sqrt{n}$-consistent ;
2. The $s \times s$ matrix

$$K_{ij}(\theta) = \int x^i \frac{\partial f(x, \boldsymbol{\theta})}{\partial \theta_j} dx, \quad i, j = 1, \cdots, s, \tag{2.2}$$

   is singular ;
3. The population moments $m_s(\theta) = \mathbf{E}_\theta X_1^j, j = 1, \cdots, s$, exist.

Under the above regularity conditions, Hsuan and Robson (1976) showed that if we replace θ by $\sqrt{n}$-consistent estimator $\bar{\theta}$ in Pearson statistic, then

$$\lim_{n\to\infty} \mathbf{P}\{X_n^2(\bar{\boldsymbol{\theta}}_n) \leq x | H_0\} = \mathbf{P}\left\{\sum_{i=1}^{k-1} \lambda_i(\theta)\chi_i^2 \leq x\right\},$$

where χ_i^2 are the independent random variables with follow a chi-square distribution with one degree of freedom, and $\lambda_i(\theta)$ are non zero eigen vectors of the matrix

$$\Sigma(\theta) = I_k - qq^T + BK^{-1}V[K^{-1}]^T B^T - C[K^{-1}]^T B^T - BK^{-1}C^T, \qquad rank\Sigma(\theta) = k - 1,$$

where K is as in equation (2.2), and

$$C = C(\theta) = \left[\frac{1}{\sqrt{p_i(\theta)}}\left\{\int_{I_i} x^i f(x, \theta) dx - p_i(\theta) m_j(\theta)\right\}\right]_{r\times s},$$

$$V = V(\theta) = [m_{ij} - m_i m_j]_{r\times s}.$$

1.2.4 Maximum Likelihood Estimator

Let $\boldsymbol{X} = (X_1, \cdots, X_n)^T$ be a random sample and suppose we want to test the hypothesis

$$H_0 : P\{X_i \leq x\} = F(x, \boldsymbol{\theta}), \quad x \in R^1, \quad \boldsymbol{\theta} = (\theta_1, \cdots, \theta_s)^T \in \Theta \subset R^s,$$

where $\boldsymbol{\theta}$ is unknown which need to be estimated and F is the given distribution function. The likelihood function of the sample $\boldsymbol{X} = (X_1, \cdots, X_n)^T$ is

$$L_n(\boldsymbol{\theta}) = \prod_{i=1}^{n} f(X_i, \boldsymbol{\theta}), \quad \boldsymbol{\theta} \in \Theta.$$

For the family $\{f(x, \theta)\}$ we assume that for $n \to \infty$ the LeCam (1956) conditions of the local asymptotic normality (L.A.N.) and of the asymptotic differentiability of the likelihood function $L_n(\boldsymbol{\theta})$ in point $\boldsymbol{\theta}$ are satisfied (Dzhaparidze and Nikulin, 1995) :

1. $L_n(\boldsymbol{\theta} + \frac{1}{\sqrt{n}}h) - L_n(\boldsymbol{\theta}) = \frac{1}{\sqrt{n}}h^T\Lambda_n(\boldsymbol{\theta}) - \frac{1}{2}h^T I^{-1}(\boldsymbol{\theta})h + o_p(1), h \in R^s$,
2. $\mathcal{L}(\frac{1}{\sqrt{n}}\Lambda_n(\boldsymbol{\theta})) \to N_s(0_s, I^{-1}(\boldsymbol{\theta}))$,
3. for any $\sqrt{n}$-consistent sequence of estimators $\{\boldsymbol{\theta}_n^*\}$ of the parameter $\boldsymbol{\theta}$

$$\frac{1}{\sqrt{n}}(\Lambda_n(\boldsymbol{\theta}_n^*) - \Lambda_n(\boldsymbol{\theta})) = \sqrt{n}I(\boldsymbol{\theta})(\boldsymbol{\theta}_n^* - \boldsymbol{\theta}) + o(1_s),$$

where $\Lambda_n(\boldsymbol{\theta})) = grad \ln L_n(\boldsymbol{\theta})$ is the vector-informant of the sample $\boldsymbol{X}$, $1_s = (1, \cdots, 1)^T$ is the unit vector in R^s, 0_s is the zero vector in R^s, and

$$I(\boldsymbol{\theta}) = \frac{1}{n}\mathbf{E}_{\boldsymbol{\theta}}\Lambda_n(\boldsymbol{\theta})\Lambda_n^T(\boldsymbol{\theta})$$

is the Fisher information matrix, corresponding to the observation X_i.

ML estimator $\hat{\boldsymbol{\theta}}_n = \hat{\boldsymbol{\theta}}_n(X_1, \cdots, X_n)$ is based on the *individual* data that maximize the likelihood function

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} f(X_i, \boldsymbol{\theta}) : \quad L(\hat{\boldsymbol{\theta}}_n) = \sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} L(\boldsymbol{\theta}).$$

Under the assumed regularity conditions on $\{F(X, \boldsymbol{\theta})\}$, we know the asymptotic behavior of the sequence $\{\hat{\boldsymbol{\theta}}_n\}$, and hence we can write

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \approx AN(\mathbf{0}_s, \mathbf{I}^{-1}(\boldsymbol{\theta})), \quad \text{as} \quad n \to \infty.$$

Theorem 1.2 *By using these properties of ML estimator Chernoff and Lehmann (1954) showed that*

$$\lim_{n\to\infty} \mathbf{P}\{X_n^2(\hat{\theta}_n) \leq x | H_0\} = \mathbf{P}\{\chi^2_{k-s-1} + \lambda_1(\theta)\xi_1^2 + \cdots + \lambda_s(\theta)\xi_s^2 \leq x\}, \tag{2.3}$$

where $\xi_1, \cdots, \xi_s, \chi^2_{k-s-1}$ are the independent random variables, $\xi_i \sim N(0, 1)$, and $0 < \lambda_i(\theta) < 1$ are the roots of the equation,

$$|(1 - \lambda)I(\theta) - J(\theta)| = 0,$$

where $I(\theta)$ is the Fisher information matrix of the observation X_i and $nJ(\theta)$ is the Fisher *information matrix of the the vector of frequencies* ν.

It is clear from (2.3) that the limit distribution of $X_n^2(\hat{\theta}_n)$ is stochastically larger than χ^2_{k-s-1}.

1.3 A Modified Pearson's Statistic (Y_n^2)- RRN

RRN statistic Y_n^2 is a modified chi-square test for complete data based on the differences between two estimators of the probabilities in each interval. One estimator is based on the empirical distribution function and the other one is on the ML estimators of unknown parameters of the tested model from ungrouped data (See Nikulin, 1973b ; Rao and Robson, 1974 ; Drost,1988,1989 ; LeCam et al., 1983 ; Van der Vaart, 1998 ; Voinov et al., 2007).

From the theorem 1.2 of Chernoff and Lehmann we find that in general it is impossible to use the Pearson's statistic to test the composite hypothesis, when we use ML estimators $\hat{\theta}_n$ or their equivalents. But we can still construct a chi-squared test for the composite hypothesis. We denote by $\boldsymbol{G}(\theta)$ the covariance matrix of the vector $X_n(\boldsymbol{\theta}) = \frac{1}{\sqrt{n}}(\boldsymbol{\nu} - n\mathbf{p}(\boldsymbol{\theta}_\mathbf{n}))$ and one can show that (Dzhaparidge and Nikulin, 1974,1995 ; Greenwood and Nikulin, 1996)

$$\begin{pmatrix} X_n(\hat{\boldsymbol{\theta}}_n) \\ \sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \end{pmatrix} \sim AN\left(\mathbf{0}_{k+s}, \left\| \begin{array}{c|c} \boldsymbol{G}(\boldsymbol{\theta}) & 0_{k\times s} \\ \hline 0_{s\times k} & I^{-1}(\boldsymbol{\theta}) \end{array} \right\| \right), \tag{2.4}$$

and under H_0

$$X_n(\hat{\boldsymbol{\theta}}_n) \sim AN(\mathbf{0}_k, \boldsymbol{G}(\boldsymbol{\theta})),$$

where

$$\boldsymbol{G}(\boldsymbol{\theta}) = \mathbf{E}_k - \mathbf{q}\mathbf{q}^T - \mathbf{B}I^{-1}\mathbf{B}^\mathbf{T},$$

$$\mathbf{q} = \mathbf{q}(\boldsymbol{\theta}) = (\sqrt{p_1(\boldsymbol{\theta})}, \cdots, \sqrt{p_1(\boldsymbol{\theta})})^T, \quad \mathbf{B} = \mathbf{B}(\boldsymbol{\theta}) = [\mathbf{b}_{ij}]_{r\times s},$$

$$b_{ij} = \frac{1}{\sqrt{p_i(\boldsymbol{\theta})}} \frac{\partial p_i(\boldsymbol{\theta})}{\partial \theta_j}.$$

$\mathbf{E}_k$ is the unit matrix of rank k. $\boldsymbol{G}(\boldsymbol{\theta})$ is a singular and one can show that its rank is $k-1$ (Nikulin, 1973c ; Moore and Spruill, 1975).

Nikulin (1973a, 1973b, 1973c, 1974) proposed the following test statistic to test the composite hypothesis

$$Y_n^2 = Y_n^2(\hat{\boldsymbol{\theta}}_n) = X_n^T(\hat{\boldsymbol{\theta}}_n)\boldsymbol{G}^-(\hat{\boldsymbol{\theta}}_n)X_n(\hat{\boldsymbol{\theta}}_n), \tag{2.5}$$

where $\boldsymbol{G}^-(\boldsymbol{\theta})$ is generalized inverse matrix of $\boldsymbol{G}(\boldsymbol{\theta})$. Also he showed that this statistic follow the chi-square distribution with $k-1$ degrees of freedom. So one can write for any fixed $x > 0$

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

The same work is shown by Rao and Robson (1974) and now this statistic is known as RRN (Rao-Robson-Nikulin) statistic.

1.4 Some Other Modifications

Another modification by Dzhaparidze and Nikulin (1974) valid for any $\sqrt{n}$-consistent estimator $\tilde{\theta}_n$ of θ (based on ungrouped data) showed that the statistic

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

in limit follows a χ^2_{k-s-1}, which coincides the Pearson-Fisher's test for grouped data (see Dzhaparidze and Nikulin (1992)). Voinov et al. (2009) showed that this test is not powerful for equiprobable intervals but is rather powerful with alternative hypothesis and with Neyman-Pearson classes (Greenwood and Nikulin, 1996).

Hsuan and Robson (1976) provided the test statistic explicitly for the exponential family of distributions when MMEs coincide with MLEs which confirms the Nikulin's result Voinov and Pya (2004). Mirvaliev (2001) generalized the test based on method of moments estimate (MME) ($\bar{\theta}$) and Voinov et al. (2012) named this general test as Hsuan-Robson-Mirvaliev (HRM) statistic which can be written as

$$Y_n^2(\bar{\theta}_n) = X_n^2(\bar{\theta}_n) + R_n^2(\bar{\theta}_n) - Q_n^2(\bar{\theta}_n).$$

The above statistic in limit has χ^2_{k-1} distribution under some regularity conditions. McCulloch (1985) showed that the statistic $Y_n^2(\hat{\theta}_n) - U_n^2(\hat{\theta}_n)$ is asymptotically independent of the DN test $U_n^2(\hat{\theta}_n)$ that is

$$\lim_{n\to\infty} \mathbf{P}\left(Y_n^2(\hat{\theta}_n) - U_n^2(\hat{\theta}_n) \leq x|H_0\right) = \mathbf{P}(\chi^2_2 \leq x),$$

and the power of $U_n^2(\hat{\theta}_n)$ is negligible compared to that of $Y_n^2(\hat{\theta}_n) - U_n^2(\hat{\theta}_n)$. For details of the modifications and results see, Voinov et al. (2012). Here we show the same results for the power of the tests with equiprobable intervals by the Monte-Carlo simulations (Tahir and Saaidia (2012)).

1.4.1 Power Estimation : Simulation

We investigate the power of $Y_n^2(\hat{\theta}_n)$, $U_n^2(\hat{\theta}_n)$ and $Y_n^2(\hat{\theta}_n) - U_n^2(\hat{\theta}_n)$ tests for the BS distribution as null hypothesis against famous alternative lognormal, loglogistic, exponentiated Weibull and generalized Weibull distributions. These distributions are generally used in reliability when the hazard rate function is unimodal (i.e. $\cap$-shaped). The test is repeated 2000 times with equiprobable intervals k by taking a sample size $n = 200$ with significance level $\alpha = 0.05$. The results are shown in Figures 2.1-2.4.

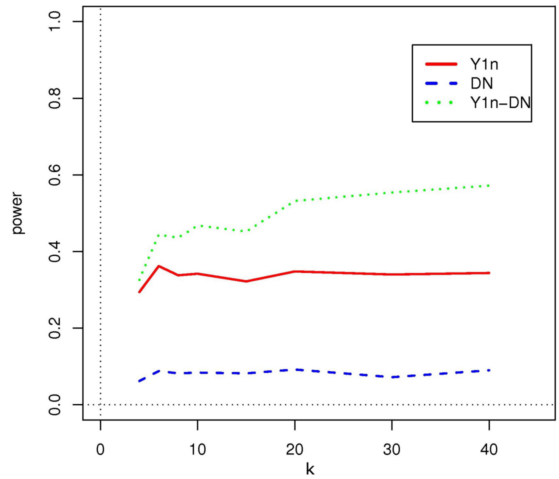

Figure 2.1 – Estimated powers of 3 tests against lognormal distribution as alternative.

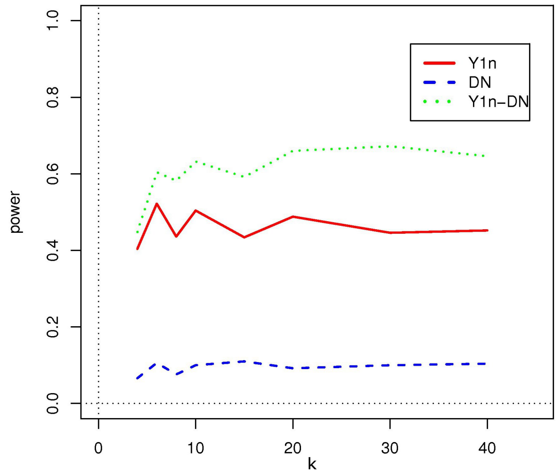

Figure 2.2 – Estimated powers of 3 tests against loglogistic distribution as alternative.

It is clear that the DN $U_n^2(\hat{\theta}_n)$ test for equiprobable intervals possesses low power for all alternative distributions. In contrast the $Y_n^2(\hat{\theta}_n)$ and $Y_n^2(\hat{\theta}_n) - U_n^2(\hat{\theta}_n)$ tests are the most powerful for all alternatives considered and for varying number of intervals k. Note that the case $r > 40$ needs further investigation because the expected intervals probabilities become small and the limit distribution of the above tests will not follows the chi-squared distribution. With equiprobable intervals it is recommended to take $k > 2$ as for $k = 2$ it will be more interesting to do the same study with the class of Neyman-Pearson and evaluate the power of these proposed tests.

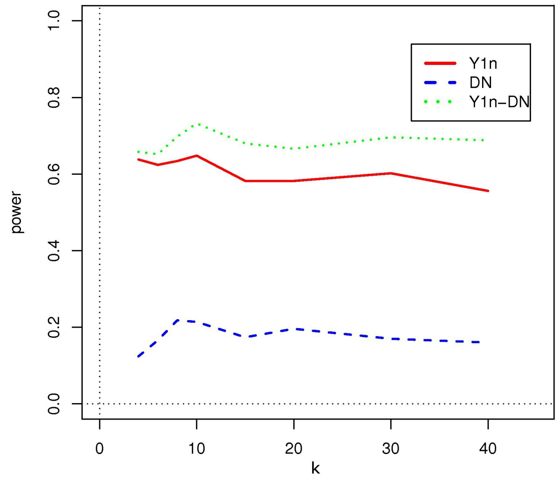

Figure 2.3 – Estimated powers of 3 tests against exponentiated Weibull distribution as alternative.

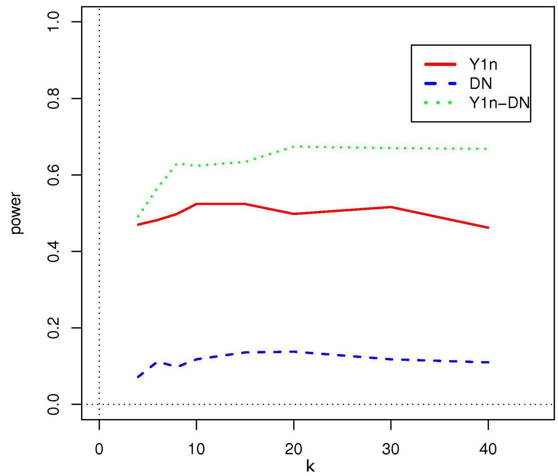

Figure 2.4 – Estimated powers of 3 tests against generalized Weibull distribution as alternative.

2 Goodness-of-Fit Test For Right Censored Data

In this section, following Bagdonavicius and Nikulin (2011), Bagdonavicius et al. (2010a), Bagdonavicius et al. (2010c), construction of a chi-squared test for testing composite parametric hypotheses for right censored data is explained. The modified chi-squared test for composite hypothesis for complete samples was first considered by Nikulin (1973a, 1973b), Rao and Robson (1974), and Dzhaparidze and Nikulin (1974). Several goodness-of-fit tests have been suggested by the statisticians for censored data. Habib and Thomas (1986), and Hollander and Pena (1992) developed a Pearson-type chi-squared statistic based on the differences of Kaplan-Meier estimate $\hat{F}_n(t)$ and parametric ML estimators of survival functions $F(t, \hat{\theta}_n)$. Akritas (1988) proposed a chi-squared statistic based on the idea of comparing the observed and expected number of failures in each class. Hjort (1990) developed a chi-squared type statistic to test the validity of the parametric model for life history data based on the cumulative hazard process. Kim (1993) also proposed the chi-squred goodness-of-fit test based on the product limit estimator.

Bagdonavicius et al. (2010a) extended the same idea of RRN statistic for complete data to censored data. They estimated end points of intervals a_j as the random data functions based on the idea to divide the interval $[0, t]$ into k subintervals with equal expected numbers of failures e_j (not necessarily integers). The test is based on the vector $Z = (Z_1, \cdots, Z_k)^T$, where $Z_j = \frac{1}{\sqrt{n}}(U_j - e_j)$, $j = 1, \cdots, k$, i.e. the differences between the number of observed and expected failures in the intervals $I_1, \cdots, I_k$. Zhang (1999) proposed a chi-squared type statistic for the logistic regression model by adapting the goodness-of fit test of Nikulin-Rao-Robson-Moore.

The maximum likelihood estimation method is almost the best method in the case of censored samples, whose asymptotic properties are well known. For testing a composite parametric hypothesis RRN statistic is well adapted for censored data and for application of M L estimators. We apply the test for the validity of exponential, Weibull, generalized Weibull, lognormal, loglogistic, and Birnbaum-Saunders distributions for different sets of data.

2.1 Composite Parametric Hypothesis

Let us consider the composite hypothesis

$$H_0 : F(x) \in \mathcal{F}_0 = \{F_0(x, \boldsymbol{\theta}), \quad x \in R^1, \boldsymbol{\theta} = (\theta_1, \cdots, \theta_s)^T \in \Theta \subset R^s\} \subset \mathcal{F}, \tag{2.6}$$

which means that the failure times T has the cdf F that belongs to the parametric family $\mathcal{F}_0$ and $\boldsymbol{\theta}$ is an unknown m-dimensional parameter and F_0 is a differentiable completely specified cdf with the support $(0, \infty)$. The class $\mathcal{F}$ contains all absolutely continuous cumulative distribution functions with the support $(0, \infty)$. We suppose also that τ is the time of studies.

2.2 ML Estimator And Its Properties

Suppose that

$$(X_1, \delta_1), \ldots, (X_n, \delta_n), \quad X_i = T_i \wedge C_i, \quad \delta_i = \mathbf{1}_{\{T_i \leq C_i\}}, \tag{2.7}$$

is a right censored sample. $T_1, \cdots, T_n$ are the failure times which are absolutely continuous i.i.d. random variables and C_i are the censoring times. We supposed that C_i are independent of T_i. The probability density function of the failure time T_1 belongs to a parametric family

$\{f(\cdot,\boldsymbol{\theta}),\, \boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \mathbf{R}^m\}$. Denote by

$$S(t,\boldsymbol{\theta}) = \mathbf{P}_{\boldsymbol{\theta}}\{T_1 > t\}, \quad \text{and} \quad \Lambda(t,\boldsymbol{\theta}) = -lnS(t,\boldsymbol{\theta}) = \int_0^t \lambda(u,\boldsymbol{\theta})du, \quad \boldsymbol{\theta} \in \boldsymbol{\Theta},$$

the survival function and the cumulative hazard function, respectively.

Denote by $\bar{G}_i$ the survival function of the censoring time C_i. For any $t > 0$ the value $\bar{G}_i(t)$ is probability for the i-th object not be censored to time t. Let us consider the distribution of the random vector (X_i, δ_i) in the case of random censoring with absolutely continuous censoring times C_i with the probability density function $g_i(t)$. We suppose that we observe the so-called *non-informative* censoring mechanism, it means that the survival function $\bar{G}_i$ and the density function $g_i(t)$ do not depend on the parameter $\boldsymbol{\theta}$. So in this case we obtain the following expressions for the likelihood function $L(\boldsymbol{\theta}), \{\boldsymbol{\theta} \in \boldsymbol{\Theta}\}$:

$$L(\boldsymbol{\theta}) = \prod_{i=1}^n f^{\delta_i}(X_i,\boldsymbol{\theta}) S^{1-\delta_i}(X_i,\boldsymbol{\theta})\, \bar{G}^{\delta_i}(X_i)\, g_i^{1-\delta_i}(X_i).$$

Since the problem is to estimate the parameter $\boldsymbol{\theta}$, we can skip the multipliers which do not depend on this parameter. So under non-informative censoring the likelihood function can be presented as :

$$L(\boldsymbol{\theta}) = \prod_{i=1}^n f^{\delta_i}(X_i,\boldsymbol{\theta}) S^{1-\delta_i}(X_i,\boldsymbol{\theta}), \{\boldsymbol{\theta} \in \boldsymbol{\Theta}\}.$$

Using the relation $f(t,\boldsymbol{\theta}) = \lambda(t,\boldsymbol{\theta})S(t,\boldsymbol{\theta})$ the likelihood function can be written as

$$L(\boldsymbol{\theta}) = \prod_{i=1}^n \lambda^{\delta_i}(X_i,\boldsymbol{\theta}) S(X_i,\boldsymbol{\theta}), \{\boldsymbol{\theta} \in \boldsymbol{\Theta}\}. \tag{2.8}$$

The estimator $\hat{\boldsymbol{\theta}}$, maximizing the likelihood function $L(\boldsymbol{\theta})$, is called *maximum likelihood estimator.*

The loglikelihood function

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^n \delta_i \ln \lambda(X_i,\boldsymbol{\theta}) + \sum_{i=1}^n \ln S(X_i,\boldsymbol{\theta}), \quad \{\boldsymbol{\theta} \in \boldsymbol{\Theta}\}, \tag{2.9}$$

is maximized at the same point as the likelihood function. If $\lambda(u,\boldsymbol{\theta})$ is sufficiently smooth function of the parameter $\boldsymbol{\theta}$ then the ML estimator satisfies the equation :

$$\dot{\ell}(\hat{\boldsymbol{\theta}}) = \mathbf{0}; \tag{2.10}$$

where $\dot{\ell}$ is the *score vector* and the Fisher's information matrix is

$$\boldsymbol{I}(\boldsymbol{\theta}) = -\mathbf{E}_{\boldsymbol{\theta}}\ddot{\ell}(\boldsymbol{\theta}). \tag{2.11}$$

The censored sample (2.7) may be written in the form of counting process

$$(N_1(t), Y_1(t), t \geq 0), \cdots, (N_n(t), Y_n(t), t \geq 0),$$

where

$$N_i(t) = \mathbf{1}_{\{X_i \leq t, \delta_i = 1\}}, \quad Y_i(t) = \mathbf{1}_{\{X_i \geq t\}}$$

$$N(t) = \sum_{i=1}^{n} N_i(t) \quad \text{and} \quad Y(t) = \sum_{i=1}^{n} Y_i(t). \tag{2.12}$$

Using these processes we obtain two useful relations :

$$\int_0^\infty \ln \lambda(u, \boldsymbol{\theta}) dN_i(u) = \begin{cases} \ln \lambda(X_i, \boldsymbol{\theta}), & \delta_i = 1, \\ 0, & \delta_i = 0. \end{cases} = \delta_i \ln \lambda(X_i, \boldsymbol{\theta})$$

and

$$\int_0^\infty Y_i(u)\lambda(u)du = \int_0^{X_i} \lambda(u)du = -\ln S(X_i, \theta).$$

From these relations we can write the loglikelihood function $\ell(\boldsymbol{\theta})$ and the score vector $\dot{\ell}(\boldsymbol{\theta})$ of the sample (2.7) in terms of stochastic processes N_i and Y_i under non-informative random censoring.

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^{n} \int_0^\infty \{\ln \lambda(u, \boldsymbol{\theta}) dN_i(u) - Y_i(u)\lambda(u, \boldsymbol{\theta})du\}, \tag{2.13}$$

from where it follows that

$$\dot{\ell}(\boldsymbol{\theta}) = \sum_{i=1}^{n} \int_0^\infty \frac{\partial}{\partial \boldsymbol{\theta}} \ln \lambda(u, \boldsymbol{\theta}) \{dN_i(u) - Y_i(u)\lambda(u, \boldsymbol{\theta})du\}, \tag{2.14}$$

and the Fisher's information matrix is

$$\boldsymbol{I}(\boldsymbol{\theta}) = -\mathbf{E}_{\boldsymbol{\theta}}\ddot{\ell}(\boldsymbol{\theta}) = \mathbf{E}_{\boldsymbol{\theta}} \sum_{i=1}^{n} \int_0^\infty \frac{\partial}{\partial \boldsymbol{\theta}} \ln \lambda(u, \boldsymbol{\theta}) \left(\frac{\partial}{\partial \boldsymbol{\theta}} \ln \lambda(u, \boldsymbol{\theta})\right)^T \lambda(u, \boldsymbol{\theta}) Y_i(u) du. \tag{2.15}$$

By tradition, accepted in survival analysis and reliability, we suppose that the processes N_i and Y_i are observed at finite time $\tau > 0$. It means that at time τ observation of all objects is censored, so in the place of censoring time C_i, censoring time $C_i \wedge \tau$ are used. We denote them once more by C_i. The process $N(t)$ shows for any $t > 0$ the number of observed failures in the interval $[0, \tau]$ and the process $Y(t)$ shows the number of objects which are *at risk* (not failed,

not truncated and not censored) just prior the time $t, t < \tau$, where τ is the maximum time of the study.
Consistency and asymptotic normality of the ML estimators $\hat{\boldsymbol{\theta}}$ hold under the following sufficient conditions (Hjort, 1990).

Conditions A :

1. There exists a neighborhood Θ_0 of $\boldsymbol{\theta}_0$ such that for all n and $\boldsymbol{\theta} \in \Theta_0$, and almost all $t \in [0, \tau]$, the partial derivatives of $\lambda(t, \boldsymbol{\theta})$ of the first, second, and the third order with respect to $\boldsymbol{\theta}$ exist and are continuous in $\boldsymbol{\theta}$ for $\boldsymbol{\theta} \in \Theta_0$. Moreover, they are bounded in $[0, \tau] \times \Theta_0$ and the log-likelihood function (2.9) may be differentiated three times with respect to $\boldsymbol{\theta} \in \Theta_0$ by interchanging the order of integration and differentiation.
2. $\lambda(t, \boldsymbol{\theta})$ is bounded away from zero in $[0, \tau] \times \Theta_0$.
3. There exists a positive deterministic function $y(t)$ such that

$$\underset{t\in[0,\tau]}{sup} |Y(t)/n - y(t)| \xrightarrow{P} 0.$$

4. The matrix $\boldsymbol{i}(\boldsymbol{\theta}_0) = \lim_{n\to\infty} \boldsymbol{I}(\boldsymbol{\theta}_0)/n$ (the limit exists under the conditions 1-3) is positive definite.

Let $\boldsymbol{\theta}_0$ be the true value of $\boldsymbol{\theta}$. The asymptotic properties of ML estimators and the score vector under conditions A) are :

1. $\hat{\boldsymbol{\theta}} \xrightarrow{d} \boldsymbol{\theta}_0$,
2. $\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) = \boldsymbol{i}^{-1}(\boldsymbol{\theta}_0)\frac{1}{\sqrt{n}}\dot{\ell}(\boldsymbol{\theta}_0) + O_p(1)$,
3. $-\frac{1}{n}\ddot{\ell}(\hat{\boldsymbol{\theta}}) \xrightarrow{P} \boldsymbol{i}(\boldsymbol{\theta_0})$,
4. $\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \xrightarrow{P} N_s(\mathbf{0}, \boldsymbol{i}^{-1}(\boldsymbol{\theta}_0))$,
5. $\frac{1}{\sqrt{n}}\dot{\ell}(\boldsymbol{\theta}) \xrightarrow{d} N_s(\mathbf{0}, \boldsymbol{i}(\boldsymbol{\theta}_0))$,

2.3 Construction Of The Test Statistic For Right Censored Data

Here a chi-squared test for testing composite parametric hypotheses (2.6) is explained. Suppose that the processes N_i, Y_i are observed at finite time τ. Then divide the interval $[0, \tau]$ into $k > s$ smaller intervals

$$I_j = (a_{j-1}, a_j], \quad a_0 = 0, \quad a_k = \tau,$$

and let denote by

$$U_j = N(a_j) - N(a_{j-1}) = \sum_{i:X_i\in I_j} \delta_i$$

the number of observed failures in the j-th interval, $j = 1, 2, ..., k$.

We need to estimate the "expected" number of failures in the interval I_j under the hypothesis H_0. Taking into account the equality

$$\mathbf{E}N(t) = \mathbf{E}\int_0^t \lambda(u, \boldsymbol{\theta}_0)Y(u)du,$$

we can "expect" to observe

$$e_j = \int_{a_{j-1}}^{a_j} \lambda(u, \hat{\boldsymbol{\theta}})Y(u)du \tag{2.16}$$

failures ; here $\hat{\boldsymbol{\theta}}$ is the MLE of the parameter $\boldsymbol{\theta}$. Following Bagdonavicius et al. (2010a) we shall construct the RRN type statistic based on the vector

$$\boldsymbol{Z} = (Z_1, ..., Z_k)^T, \quad Z_j = \frac{1}{\sqrt{n}}(U_j - e_j), \quad j = 1, ..., k. \tag{2.17}$$

To investigate the properties of the statistic $\boldsymbol{Z}$ one can use the properties of the stochastic process

$$H_n(t) = \frac{1}{\sqrt{n}}\left(N(t) - \int_0^t \lambda(u, \hat{\boldsymbol{\theta}})Y(u)du\right),$$

in the interval $[0, \tau]$, given in the next lemma.

Lemma 1 *Under conditions A) the following convergence holds :*

$$H_n \xrightarrow{d} V \quad on\ D[0, \tau];$$

where V is zero mean Gaussian martingale such that for all $0 \leq s \leq t$

$$\mathbf{Cov}\,(V(s), V(t)) = A(s) - \boldsymbol{C}^T(s)\boldsymbol{i}^{-1}(\boldsymbol{\theta}_0)\boldsymbol{C}(t);$$

where

$$A(t) = \int_0^t \lambda(u, \boldsymbol{\theta}_0)y(u)du, \quad \boldsymbol{C}(t) = \int_0^t \frac{\partial}{\partial \boldsymbol{\theta}} \ln \lambda(u, \boldsymbol{\theta}_0)\lambda(u, \boldsymbol{\theta}_0)y(u)du,$$

and $\xrightarrow{d}$ means week convergence in the space $D[0, \tau]$ of cadlag functions with Skorokhod metric.

Set for $i = 1, ..., s\,;\ j, j' = 1, ..., k$

$$V_j = V(a_j) - V(a_{j-1}), \quad v_{jj'} = \mathbf{Cov}\,(V_j, V_{j'}),$$

$$A_j = A(a_j) - A(a_{j-1}), \quad \boldsymbol{C}_j = (C_{1j}, ..., C_{sj})^T = \boldsymbol{C}(a_j) - \boldsymbol{C}(a_{j-1}),$$

$$\boldsymbol{V} = [v_{jj'}]_{k \times k}, \quad \boldsymbol{C} = [C_{ij}]_{s \times k},$$

and denote by $\boldsymbol{A}$ the $k \times k$ diagonal matrix with diagonal elements $A_1, ..., A_k$.

It is easy to verify, see Bagdonavicius and Nikulin (2011), that under conditions A)

$$\boldsymbol{Z} \xrightarrow{d} \boldsymbol{Y} \sim N_k(\boldsymbol{0}, \boldsymbol{V}) \quad \text{as } n \to \infty,$$

where

$$\boldsymbol{V} = \boldsymbol{A} - \boldsymbol{C}^T \boldsymbol{i}^{-1}(\boldsymbol{\theta}_0)\boldsymbol{C}.$$

Remark : Set the matrix

$$\boldsymbol{G} = \boldsymbol{i} - \boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{C}^T.$$

If $\boldsymbol{G}$ is non-degenerate then a generalized inverse of the matrix $\boldsymbol{V}$ is

$$\boldsymbol{V}^- = \boldsymbol{A}^{-1} + \boldsymbol{A}^{-1}\boldsymbol{C}^T\boldsymbol{G}^{-1}\boldsymbol{C}\boldsymbol{A}^{-1}.$$

We need to inverse only diagonal $k \times k$ matrix $\boldsymbol{A}$ and $m \times m$ matrix $\boldsymbol{G}$, (usually $m = 1$, $m = 2$ or $m = 3$).

Theorem 2.1 *Under conditions A) the following estimators of A_j, $\boldsymbol{C}_j$, $\boldsymbol{i}(\boldsymbol{\theta}_0)$ and $\boldsymbol{V}$ are consistent (see Bagdonavicius et al., 2010a) :*

$$\hat{A}_j = U_j/n, \quad \hat{\boldsymbol{C}}_j = \frac{1}{n}\int_{a_{j-1}}^{a_j} \frac{\partial}{\partial \boldsymbol{\theta}} \ln \lambda(u, \hat{\boldsymbol{\theta}}) dN(t),$$

and

$$\hat{\boldsymbol{i}} = \frac{1}{n}\int_0^\tau \left(\frac{\partial}{\partial \boldsymbol{\theta}} \ln \lambda(u; \hat{\boldsymbol{\theta}})\right)\left(\frac{\partial}{\partial \boldsymbol{\theta}} \ln \lambda(u; \hat{\boldsymbol{\theta}})\right)^T dN(u), \qquad \hat{\boldsymbol{V}} = \hat{\boldsymbol{A}} - \hat{\boldsymbol{C}}^T \hat{\boldsymbol{i}}^{-1} \hat{\boldsymbol{C}}.$$

The statistic $\hat{\boldsymbol{i}} = -\frac{1}{n}\ddot{\ell}(\hat{\boldsymbol{\theta}})$ is also a consistent estimator of $\boldsymbol{i}(\boldsymbol{\theta}_0)$ but it is recommended to use the above estimator to ensure that both components of the following RRN test statistic are non-negative for any n

$$Y_n^2 = \boldsymbol{Z}^T \hat{\boldsymbol{V}}^- \boldsymbol{Z},$$

where $\hat{\boldsymbol{V}}^-$ is the special general inverse of the matrix $\hat{\boldsymbol{V}}$ as

$$\hat{\boldsymbol{V}}^- = \hat{\boldsymbol{A}}^{-1} + \hat{\boldsymbol{A}}^{-1}\hat{\boldsymbol{C}}^T\hat{\boldsymbol{G}}^-\hat{\boldsymbol{C}}\hat{\boldsymbol{A}}^{-1}, \quad \hat{\boldsymbol{G}} = \hat{\boldsymbol{i}} - \hat{\boldsymbol{C}}\hat{\boldsymbol{A}}^{-1}\hat{\boldsymbol{C}}^T.$$

So the test statistic can be written in the form

$$Y_n^2 = \boldsymbol{Z}^T\hat{\boldsymbol{A}}^{-1}\boldsymbol{Z} + \boldsymbol{Z}^T\hat{\boldsymbol{A}}^{-1}\hat{\boldsymbol{C}}^T\hat{\boldsymbol{G}}^-\hat{\boldsymbol{C}}\hat{\boldsymbol{A}}^{-1}\boldsymbol{Z} = \sum_{j=1}^k \frac{(U_j - e_j)^2}{U_j} + Q,$$

where

$$Q = \boldsymbol{W}^T \hat{\boldsymbol{G}}^{-} \boldsymbol{W}, \quad \boldsymbol{W} = \hat{\boldsymbol{C}} \hat{\boldsymbol{A}}^{-1} \boldsymbol{Z} = (W_1, ..., W_s)^T,$$

$$\hat{\boldsymbol{G}} = [\hat{g}_{ll'}]_{s \times s}, \quad \hat{g}_{ll'} = \hat{i}_{ll'} - \sum_{j=1}^{k} \hat{C}_{lj} \hat{C}_{l'j} \hat{A}_j^{-1}, \quad W_l = \sum_{j=1}^{k} \hat{C}_{lj} \hat{A}_j^{-1} Z_j,$$

$$\hat{\boldsymbol{i}}_{ll'} = \frac{1}{n} \sum_{i=1}^{n} \delta_i \frac{\partial \ln \lambda(X_i; \hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}_l} \frac{\partial \ln \lambda(X_i; \hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}_{l'}}, \quad \hat{\boldsymbol{C}}_{lj} = \frac{1}{n} \sum_{i:X_i \in I_j} \delta_i \frac{\partial}{\partial \boldsymbol{\theta}} \ln \lambda(X_i, \hat{\boldsymbol{\theta}}),$$

$$\hat{A}_j = U_j/n, \quad U_j = \sum_{i:X_i \in I_j} \delta_i, \quad Z_j = \frac{1}{\sqrt{n}}(U_j - e_j),$$

$i = 1, \cdots, n, \quad j = 1, \cdots, k, \quad l, l' = 1, \cdots, s.$
Denote by $\hat{g}^{ll'}$ the elements of $\hat{\boldsymbol{G}}^{-}$. The quadratic form Q can be written as follows :

$$Q = \sum_{l=1}^{m} \sum_{l'=1}^{m} W_l g^{ll'} W_{l'}.$$

Under H_0 the limit distribution of the statistic Y_n^2 is chi-square with $r = rank(\boldsymbol{V}^{-}) = Tr(\boldsymbol{V}^{-}\boldsymbol{V})$ degrees of freedom. If G is non-degenerate then $r = k$.

Statistical inference for the hypothesis H_0 : The hypothesis is rejected with approximate significance level α if $Y_n^2 > \chi^2_\alpha(r)$.

2.3.1 Choice of random grouping intervals $\hat{a}_j$

Set

$$b_i = (n-i)\Lambda(X_{(i)}, \hat{\boldsymbol{\theta}}) + \sum_{l=1}^{i} \Lambda(X_{(l)}, \hat{\boldsymbol{\theta}}),$$

where $X_{(i)}$ is the *ith* element in the ordered statistics $(X_{(1)}, \cdots, X_{(n)})$.
If i is the smallest natural number verifying $E_j \in [b_{i-1}, b_i], j = 1, \cdots, k-1$ then

$$(n-i+1)\Lambda(a_j, \hat{\boldsymbol{\theta}}) + \sum_{l=1}^{i-1} \Lambda(X_{(l)}, \hat{\boldsymbol{\theta}}) = E_j$$

where a_j are the end points of the intervals and can be estimated as

$$\hat{a}_j = \Lambda^{-1}\left([E_j - \sum_{l=1}^{i-1} \Lambda(X_{(l)}, \hat{\boldsymbol{\theta}})]/(n-i+1), \hat{\boldsymbol{\theta}}\right), \quad \hat{a}_k = \max(X_{(n)}, \tau) \tag{2.18}$$

where Λ^{-1} is the inverse of cumulative hazard function Λ. We have $0 < \hat{a}_1 < \hat{a}_2, \cdots, \hat{a}_k = \tau$. With this choice of intervals

$$ej = E_k/k$$

for any j, where $E_k = \sum\limits_{i=1}^{n} \Lambda(X_i, \hat{\boldsymbol{\theta}})$. Usually in real application we fix k. Bagdonavicius et al. (2010a) and Greenwood and Nikulin (1996) give some recommendations for the choice of intervals. If there is no alternative hypothesis, the number of intervals k can be taken such as $n/k > 5$.
By random change of time theorem (Billingsley, 1968)

$$(H_n(\hat{a}_1), \cdots, H_n(\hat{a}_1))^T \xrightarrow{d} (V(a_1), \cdots, V(a_1))^T,$$

which means that replacing a_j by $\hat{a}_j$ in the expression of the statistic Y_n^2, the limit distribution of the statistic remain chi-square with r degrees of freedom as in the case of fixed a_j.

In classical way of selecting *equiprobable intervals* we fix k and take $0 < P_1 < ... < P_k < 1$ in such a way that $P_j = j/(k+1)$, $j = 1, ..., k$. For example, taking $k = 9$ we have $P_1 = 0.1, P_2 = 0.2, ..., P_9 = 0.9$ and we make the intervals from the following cut-points

$$a_j = F^{-1}(P_j, \hat{\boldsymbol{\theta}}) = \inf\{t : F_\tau(t, \hat{\boldsymbol{\theta}}) \geq P\}.$$

2.3.2 Choice of $\hat{a}_j$ in shape and scale distribution families

Set

$$b_i = (n-i)\Lambda_0(Y_{(i)}) + \sum_{l=1}^{i} \Lambda_0(Y_{(l)}), \qquad Y_{(i)} = (\frac{X_{(i)}}{\hat{\theta}})^{\hat{\nu}};$$

where Λ_0 is the cumulative hazard function. If i is the smallest natural number verifying $E_j \in [b_{i-1}, b_i], j = 1, \cdots, k-1$ then

$$\hat{a}_j = \hat{\theta}\left\{\Lambda_0^{-1}\left([E_j - \sum_{l=1}^{i-1} \Lambda_0(Y_{(l)})]/(n-i+1), \boldsymbol{\theta}\right)\right\}, \quad \hat{a}_k = \max(X_{(n)}, \tau)$$

where Λ_0^{-1} is the inverse of cumulative hazard function Λ_0.
For such choices of intervals we have $ej = E_k/k$ for any j, where $E_k = \sum_{i=1}^{n} \Lambda_0(Y_i)$.

2.4 Application Of The RRN Test

Here we apply the RRN test and give the elements of the quadratic form of different parametric models such as exponential, Weibull, generalized Weibull, exponentiated Weibull, loglogistic, lognormal, and Birnbaum-Saunders. Also we apply the test on the original data of arm A head and neck cancer. R-statistical programming language is used to apply the goodness of fit tests for all models.

Arm-A head and neck cancer data

The data for arm A of head and beck cancer study was conducted by northern California oncology group. The survival times in days for the patients ($n = 51$) were as below ($\delta = \sum_{i}^{n} \delta_i = 42$).

7, 34, 42, 63, 64, 74*, 83, 84, 91, 108, 112, 129, 133, 133, 139, 140,
140, 146, 149, 154, 157, 160, 160, 165, 173, 176, 185*, 218, 225, 241,
248, 273, 277, 279*, 297, 319*, 405, 417, 420, 440, 523*, 523, 583, 594,
1101, 1116*, 1146, 1226*, 1349*, 1412*, 1417

*censoring

First this data was used by Efron (1988) for logistic distribution. Mudholkar et al. (1996) and Nikulin and Haghighi (2006) reanalysed the same data and give the acceptable fit (chi-square type test) to the exponentiated Weibull and generalized Weibull distribution families respectively. We use the data after transforming the *survival times in months* (1 month=30.438 days). The hazard functions using this data for the shape-scale families are shown in Figure 2.5.

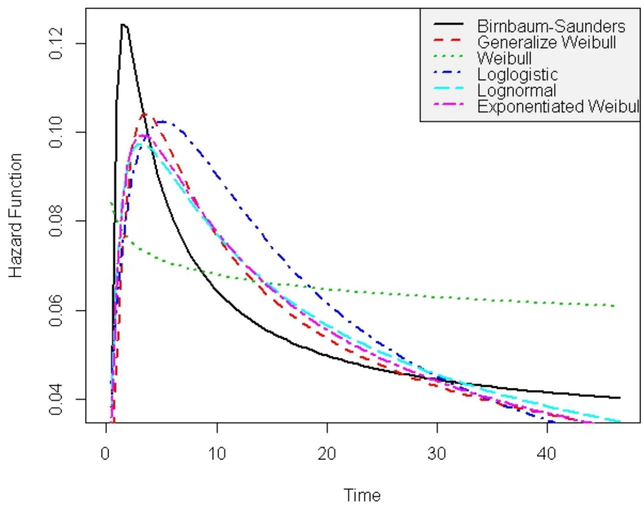

Figure 2.5 – Hazard plots for shape-scale families.

Now following the theory in previous section-2, we consider several examples widely used in reliability and survival analysis.

2.4.1 Exponential Distribution

The exponential distribution is primarily used in reliability applications to model the data with a constant failure rate. Suppose the hypothesis that the distribution of failure times in exponential i.e.

$$H_0 : F(t) = 1 - e^{\lambda t}, t \geq 0, \lambda > 0.$$

Let

$$S(t, \theta) = e^{\lambda t}, \qquad \lambda(t, \theta) = \lambda,$$

be the survival and the hazard function respectively.
Set $S_n = \sum_{i=1}^{n} X_i$, we have

$$\hat{\lambda} = \delta / S_n, \qquad U_j = \sum_{i:X_i \in I_j} \delta_i, \quad \hat{C}_j = \frac{U_j}{n\hat{\lambda}},$$

$$\hat{i} = \frac{\delta}{n\hat{\lambda}^2}, \qquad \hat{G} = \hat{g}_{11} = \frac{\delta}{n\hat{\lambda}^2} - \sum_{j=1}^{k} \frac{U_j^2}{n^2\hat{\lambda}^2} \frac{n}{U_j} = 0.$$

It means G is degenerated and we can not find the general inverse of $\hat{G}$, so we find the general inverse of $\hat{V}$ directly.

Under the exponential distribution the elements of the matrix $\hat{V}$ are

$$\hat{v}_{jj} = \hat{A}_j - \hat{C}_j^2 \, \hat{i}^{-1} = \frac{U_j}{n} - \frac{U_j^2}{n\delta},$$

and for $j \neq j'$

$$\hat{v}_{jj'} = -\hat{C}_j \, \hat{i}^{-1} \hat{C}_{j'} = -\frac{U_j U_{j'}}{n\delta}.$$

Set

$$\hat{\pi}_j = \frac{U_j}{\delta}, \qquad \sum_{j=1}^{k} \hat{\pi}_j = 1, \qquad \hat{\boldsymbol{\pi}} = (\hat{\pi}_1, \cdots, \hat{\pi}_k)^T.$$

Denote by $\boldsymbol{D}$ the diagonal matrix with the diagonal elements $\hat{\pi}$. The matrix $\hat{V}$ and its generalized inverse $\hat{\boldsymbol{V}}^-$ have the form

$$\hat{V} = \frac{\delta}{n}(\hat{\boldsymbol{D}} - \hat{\boldsymbol{\pi}}\hat{\boldsymbol{\pi}}^T), \qquad \hat{\boldsymbol{V}}^- = \frac{n}{\delta}(\hat{\boldsymbol{D}}^{-1} + \mathbf{1}\mathbf{1}^{\mathbf{T}}),$$

by using the equalities

$$\mathbf{1}^T \hat{\boldsymbol{D}} = \hat{\boldsymbol{\pi}}^T, \quad \mathbf{1}^T \hat{\boldsymbol{\pi}} = \hat{\boldsymbol{\pi}}^T \mathbf{1} = 1, \quad \hat{\boldsymbol{D}}\mathbf{1} = \hat{\boldsymbol{\pi}}, \quad \hat{\boldsymbol{\pi}}^T \hat{\boldsymbol{D}}^{-1} = \mathbf{1}^T,$$

we obtain

$$\hat{\boldsymbol{V}} \hat{\boldsymbol{V}}^- \hat{\boldsymbol{V}} = \hat{\boldsymbol{V}}.$$

This we can show as

$$\begin{aligned}
\hat{\boldsymbol{V}}^{-}\hat{\boldsymbol{V}} &= \frac{1}{\hat{A}}(\hat{\boldsymbol{D}}^{-1} + \mathbf{1}\mathbf{1}^T)\hat{A}(\hat{\boldsymbol{D}} - \hat{\boldsymbol{\pi}}\hat{\boldsymbol{\pi}}^T) = \mathbf{E} - \mathbf{1}\hat{\boldsymbol{\pi}}^T, \\
\hat{\boldsymbol{V}}\hat{\boldsymbol{V}}^{-}\hat{\boldsymbol{V}} &= \hat{A}(\hat{\boldsymbol{D}} - \hat{\boldsymbol{\pi}}\hat{\boldsymbol{\pi}}^T)(\mathbf{E} - \mathbf{1}\hat{\boldsymbol{\pi}}^T) \\
&= \hat{A}(\hat{\boldsymbol{D}} - \hat{\boldsymbol{D}}\mathbf{1}\hat{\boldsymbol{\pi}}^T - \hat{\boldsymbol{\pi}}\hat{\boldsymbol{\pi}}^T - \hat{\boldsymbol{\pi}}\hat{\boldsymbol{\pi}}^T\mathbf{1}\hat{\boldsymbol{\pi}}^T) \\
&= \hat{A}(\hat{\boldsymbol{D}} - \hat{\boldsymbol{\pi}}\hat{\boldsymbol{\pi}}^T) = \hat{\boldsymbol{V}}.
\end{aligned}$$

So the quadratic form can be written as

$$\begin{aligned}
Y_n^2 &= \boldsymbol{Z}^T\hat{\boldsymbol{V}}^{-}\boldsymbol{Z} = \frac{n}{\delta}\boldsymbol{Z}^T\hat{\boldsymbol{D}}^{-1}\boldsymbol{Z} + \frac{n}{\delta}(\boldsymbol{Z}^T\mathbf{1})^2 \\
&= \sum_{j=1}^{k}\frac{(U_j - e_j)^2}{U_j} + \frac{1}{\delta}[\sum_{j=1}^{k}(U_j - e_j)]^2.
\end{aligned}$$

The limit distribution of the statistic Y_n^2 is chi square with $Tr(\boldsymbol{V}^{-}\boldsymbol{V}) = k-1$ degrees of freedom because

$$\hat{A}_j \xrightarrow{P} A_j > 0, \quad \delta/n \xrightarrow{P} A = \sum_{j=1}^{k} Aj \in (0,1), \quad \hat{\pi}_j \xrightarrow{P} A_j/A = \pi_j, \quad \hat{\boldsymbol{D}} \xrightarrow{P} \boldsymbol{D},$$

so

$$\boldsymbol{V}^{-}\boldsymbol{V} = \frac{1}{A}(\boldsymbol{D}^{-1} + \mathbf{1}\mathbf{1}^T)A(\boldsymbol{D} - \boldsymbol{\pi}\boldsymbol{\pi}^T) = \mathbf{E} - \mathbf{1}\boldsymbol{\pi}^T,$$

$$Tr(\mathbf{E} - \mathbf{1}\boldsymbol{\pi}^T) = k - \sum_{j=1}^{k}\pi_j = k - 1.$$

Note that

$$\sum_{j=1}^{k} e_j = \hat{\lambda}\int_0^{\tau} Y(u)d(u) = \hat{\lambda}\sum_{i=1}^{n} X_i = \hat{\lambda}S_n = \delta = \sum_{j=1}^{k} U_j.$$

So

$$\frac{1}{\delta}[\sum_{j=1}^{k}(U_j - e_j)]^2 = 0.$$

Choice of $\hat{a}_j$: Set

$$S_0 = 0, \quad S_i = (n-i)X_{(i)} + \sum_{l=1}^{i} X_{(l)}, \quad i = 1, \cdots, n.$$

The formula (2.18) implies that the limits of the intervals I_j are chosen as : if i is the smallest natural number verifying the inequalities $S_{i-1} \leq \frac{j}{k} S_n \leq S_i$, then

$$\hat{a}_j = \left(\frac{j}{k} S_n - \sum_{l=1}^{i-1} X_{(l)} \right) / (n - i + 1), \qquad j = 1, \cdots, k-1, \quad \hat{a}_k = X_{(n)}.$$

Under this choice of intervals we have $e_j = \delta / k$ for any j.

Chi-squared test for the exponential distribution : the null hypothesis is rejected with approximate significance level α if

$$Y_n^2 = \sum_{j=1}^{k} \frac{(U_j - e_j)^2}{U_j} > \chi^2_\alpha(k-1).$$

2.4.2 Weibull Distribution

Weibull distribution is most commonly used distribution in reliability to model the failure times due its flexible parameters. The failure rate can be decreasing, constant or increasing depending upon the values of parameters. Modeling the monotone hazard rates, Weibull distribution can be the initial choice due to its positively and negatively skewed density shapes.

Suppose that under H_0 the failure times follow the weibull distribution with

$$S(t; \theta, \nu) = \exp\{-(\frac{t_i}{\theta})^\nu\}, \quad \lambda(t, \theta, \nu) = \frac{\nu}{\theta^\nu} t_i^{\nu-1}, \quad \Lambda(t, \theta, \nu) = (\frac{t_i}{\theta})^\nu,$$

$(\theta, \nu > 0); t \geq 0,$
the survival function, the hazard function and the cumulative hazard function respectively.
The Log likelihood function is

$$\ell(\theta, \nu) = \sum_{i=1}^{n} \{\delta_i [(\nu - 1) \ln X_i - \nu \ln \theta + \ln \nu] - (\frac{X_i}{\theta})^\nu\}.$$

Denote by $\hat{\theta}$ and $\hat{\nu}$ the maximum likelihood estimator of the parameters $\boldsymbol{\theta}$ and ν.
The matrix $\boldsymbol{G}$ is degenerate and has rank 1. So we need only $\hat{g}_{22}$ to find $\hat{G}^-$. We have

$$\hat{i}_{22} = \frac{1}{n\hat{\nu}^2} \sum_{i=1}^{n} \delta_i \left(1 + \ln Y_i\right)^2.$$

Choice of $\hat{a}_j$: Set

$$Y_i = (\frac{X_i}{\hat{\theta}})^{\hat{\nu}}, \quad b_i = (n - i) Y_{(i)} + \sum_{l=1}^{i} Y_{(l)}, \quad i = 1, \cdots, n.$$

If i is the smallest natural number verifying the inequalities

$$b_{i-1} \leq \frac{j}{k} b_n \leq b_i$$

then

$$\hat{a}_j = \hat{\theta} \left\{ [\frac{j}{k} b_n - \sum_{l=1}^{i-1} Y_{(l)}]/(n-i+1) \right\}^{1/\hat{\nu}},$$

$$j = 1, \cdots, k-1, \quad \hat{a}_k = \max(X_{(n)}, \tau).$$

For such choice of intervals we have $e_j = \delta/k$ for any j.
The test statistic is

$$Y_n^2 = \sum_{j=1}^{k} \frac{(U_j - e_j)^2}{U_j} + Q.$$

Since matrix $\boldsymbol{G}$ is degenerate so we have

$$Q = \frac{W_2}{\hat{g}_{22}}, \quad \hat{g}_{22} = \hat{\imath}_{22} - \sum_{j=1}^{k} \hat{C}_{2j}^2 \hat{A}_j^{-1}, \quad W_2 = \sum_{j=1}^{k} \hat{C}_{2j} \hat{A}_j^{-1} Z_j,$$

$$\hat{C}_{2j} = \frac{1}{n\hat{\nu}} \sum_{i:X_i \in I_j} \delta_i \{1 + \ln Y_i\},$$

$$\hat{A}_j = U_j/n, \quad U_j = \sum_{i:X_i \in I_j} \delta_i, \quad Z_j = \frac{1}{\sqrt{n}}(U_j - e_j),$$

The zero hypothesis is rejected with an approximate significance level α if $Y_n^2 > \chi_\alpha^2(k-1)$.

The maximum likelihood estimators of Weibull distribution by taking the survival times in months are; $\hat{\theta} = 14.0242$, $\hat{\nu} = 0.9297$. We take 5 intervals i.e. k=5. Further results to calculate the Y^2 are shown below :

j	1	2	3	4	5
$\hat{a}_j$	2.0873	4.8731	10.3402	21.9896	46.5537
U_j	4	13	15	7	3
e_j	8.4	8.4	8.4	8.4	8.4

$$\hat{\boldsymbol{\imath}}_{22} = 0.807105, \quad \hat{g}_{22} = 0.091314, \quad W_2 = -0.703186.$$

The value of test statistic is

$$Y^2 = X^2 + Q = 19.3717 + 5.4150 = 24.7867,$$

and

$$pv = P\{\chi^2_4 > 24.7867\} = 0.000056.$$

So from the result we reject the hypothesis and conclude that the head and neck cancer data doesn't follow the weibull distribution.

Remark : If we take the classical way of selecting equiprobable intervals ($\hat{a}_j$) then we have :

j	1	2	3	4	5
$\hat{a}_j$	2.7937	6.8090	12.7655	23.3984	46.5537
U_j	7	18	7	7	3
e_j	10.8031	9.6099	7.2415	6.5775	7.7680

$$\hat{\boldsymbol{\imath}}_{22} = 1.67699, \quad \hat{g}_{22} = 0.97414, \quad W_2 = -0.75799.$$

The value of test statistic is

$$Y^2 = X^2 + Q = 13.5888 + 0.5898 = 14.1786,$$

and

$$pv = P\{\chi^2_4 > 14.1786\} = 0.0067.$$

So from the result we reject the hypothesis and conclude the same inference as above but the new method strongly reject the hypothesis as one can see from the p-value.

2.4.3 Generalized Weibull Distribution

This distribution is the extension of Weibull distribution and contains four shapes of the hazard function and is used in the reliability and survival analysis.

Suppose under H_0 the distribution of failure times is Generalized Weibull. Then

$$S(t,\theta,\nu,\gamma) = \exp\left\{1 - \left(1 + (\frac{t}{\theta})^\nu\right)^{1/\gamma}\right\}, \quad t \geq 0, (\theta,\nu,\gamma > 0),$$

$$\lambda(t,\theta,\nu,\gamma) = \frac{\nu}{\gamma\theta^\nu} t^{\nu-1}\left(1 + (\frac{t}{\theta})^\nu\right)^{1/\gamma - 1},$$

be the survival function and the hazard function respectively.

The log-likelihood function is

$$\ell = \sum_{i=1}^n \delta_i \left\{\ln\nu - \ln\gamma - \nu\ln\theta + (\nu-1)\ln X_i + (\frac{1}{\gamma} - 1)\ln(1+Y_i)\right\} + n - \sum_{i=1}^n (1+Y_i)^{1/\gamma}.$$

where $Y_i = (\frac{X_i}{\theta})^{\nu}$.

The RRN statistic is

$$Y_n^2 = \sum_{j=1}^{k} \frac{(U_j - e_j)^2}{U_j} + Q,$$

where

$$U_j = \sum_{i:X_i \in I_j} \delta_i, \quad Q = \sum_{l=1}^{3} \sum_{l'=1}^{3} W_l g_{ll'}^{-1} W_{l'}, \qquad l, l' = 1, 2, 3.$$

$$W_l = \sum_{j=1}^{k} \hat{C}_{ij} \hat{A}_j^{-1} Z_j, \qquad \hat{g}_{ll'} = \hat{i}_{ll'} - \sum_{j=1}^{k} \hat{C}_{ij} \hat{C}_{i'j} \hat{A}_j^{-1}, \qquad [\hat{g}^{ll'}]_{3\times 3} = [\hat{g}_{ll'}]_{3\times 3}^{-1},$$

$$\hat{A}_j = U_j/n, \quad \hat{C}_{lj} = \frac{1}{n} \sum_{i:X_i \in I_j} \delta_i \frac{\partial}{\partial \theta} \ln \lambda(X_i; \hat{\theta}),$$

$$\begin{aligned}
\hat{C}_{1j} &= -\tfrac{\hat{\nu}}{n\hat{\theta}} \textstyle\sum_{i:X_i \in I_j} \delta_i \left\{ 1 + (\tfrac{1}{\hat{\gamma}} - 1) \tfrac{Y_i}{1+Y_i} \right\}, \\
\hat{C}_{2j} &= \tfrac{1}{n\hat{\nu}} \textstyle\sum_{i:X_i \in I_j} \delta_i \left\{ 1 + \ln Y_i + (\tfrac{1}{\hat{\gamma}} - 1) \tfrac{Y_i \ln Y_i}{1+Y_i} \right\}, \\
\hat{C}_{3j} &= -\tfrac{1}{n\hat{\gamma}} \textstyle\sum_{i:X_i \in I_j} \delta_i \left\{ 1 + \tfrac{1}{\hat{\gamma}} \ln(1+Y_i) \right\}.
\end{aligned}$$

The elements of the symmetric matrix $\hat{\boldsymbol{i}} = [\hat{\boldsymbol{i}}_{ll'}]_{3\times 3}$ are

$$\hat{i}_{11} = \frac{\hat{\nu}^2}{n\hat{\theta}^2} \sum_{i=1}^{n} \delta_i \left\{ 1 + (\frac{1}{\hat{\gamma}} - 1) \frac{Y_i}{1+Y_i} \right\}^2,$$

$$\hat{i}_{22} = \frac{1}{n\hat{\nu}^2} \sum_{i=1}^{n} \delta_i \left\{ 1 + \ln Y_i + (\frac{1}{\hat{\gamma}} - 1) \frac{Y_i \ln Y_i}{1+Y_i} \right\}^2,$$

$$\hat{i}_{33} = \frac{1}{n\hat{\gamma}^2} \sum_{i=1}^{n} \delta_i \left\{ 1 + \frac{1}{\hat{\gamma}} \ln(1+Y_i) \right\}^2,$$

$$\hat{i}_{12} = -\frac{1}{n\hat{\theta}} \sum_{i=1}^{n} \delta_i \left\{ 1 + (\frac{1}{\hat{\gamma}} - 1) \frac{Y_i}{1+Y_i} \right\} \left\{ 1 + \ln Y_i + (\frac{1}{\hat{\gamma}} - 1) \frac{Y_i \ln Y_i}{1+Y_i} \right\},$$

$$\hat{i}_{13} = \frac{\hat{\nu}}{n\hat{\theta}\hat{\gamma}} \sum_{i=1}^{n} \delta_i \left\{ 1 + (\frac{1}{\hat{\gamma}} - 1) \frac{Y_i}{1+Y_i} \right\} \left\{ 1 + \frac{1}{\hat{\gamma}} \ln(1+Y_i) \right\},$$

$$\hat{i}_{23} = -\frac{1}{n\hat{\gamma}\hat{\nu}} \sum_{i=1}^{n} \delta_i \left\{ 1 + \ln Y_i + (\frac{1}{\hat{\gamma}} - 1) \frac{Y_i \ln Y_i}{1+Y_i} \right\} \left\{ 1 + \frac{1}{\hat{\gamma}} \ln(1+Y_i) \right\}.$$

Choice of $\hat{a}_j$: Set

$$b_i = (n-i)\Lambda_0(Y_{(i)}) + \sum_{l=1}^{i} \Lambda_0(Y_{(l)}),$$

where

$$\Lambda_0(t) = (1+t)^{1/\gamma} - 1.$$

If i is the smallest natural number verifying $E_j \in [b_{i-1}, b_i], j = 1, \cdots, k-1$, then

$$\hat{a}_j = \hat{\theta}\left\{\Lambda_0^{-1}\left([E_j - \sum_{l=1}^{i-1}\Lambda_0(Y_{(l)})]/(n-i+1)\right)\right\}^{1/\hat{\nu}}, \quad \hat{a}_k = \max(X_{(n)}, \tau),$$

where

$$\Lambda_0^{-1}(t) = (1+t)^{\gamma} - 1.$$

For such choices of intervals we have $ej = E_k/k$ for any j, where

$$E_k = \sum_{i=1}^{n} \Lambda_0(Y_i).$$

By putting $\gamma = 1$ we can deduce the elements of the estimator $\hat{\boldsymbol{i}} = [\hat{i}_{ll'}]_{2\times 2}$ and $\hat{C}_j$ for the Weibull distribution.

Chi-squared test for the hypothesis H_0 : The hypothesis is rejected with approximate significance level α, if $Y_n^2 > \chi^2_\alpha(k)$, where k is the number of classes.

The maximum likelihood estimators of generalized Weibull distribution by taking into account the survival times in months are :

$$\hat{\theta} = 2.5458, \quad \hat{\nu} = 2.1887, \quad \hat{\gamma} = 4.9946.$$

Here $\hat{\gamma} = 4.9946$ which is much greater than 1 and it justifies the rejection of Weibull distributoin because for Weibull distribution $\gamma = 1$ in generalized Weibull distribution.

We take 5 intervals i.e. k=5. Further results to calculate the Y^2 are shown below

j	1	2	3	4	5
$\hat{a}_j$	2.7900	4.7846	8.3689	16.3947	46.5536
U_j	7	9	13	7	6
e_j	8.2184	8.2184	8.2184	8.2184	8.2184

$$\hat{g}_{3\times 3} = \begin{pmatrix} 0.0127094 & -0.0041379 & -0.0011513 \\ -0.0041379 & 0.0405083 & -0.0077945 \\ -0.0011513 & -0.0077945 & 0.0031218 \end{pmatrix}.$$

$$W_l = (-0.022203 \quad 0.134012 \quad 0.005736)^T$$

Here the matrix $\boldsymbol{G}$ is not degenerated. The value of test statistic is

$$Y^2 = X^2 + Q = 3.0710 + 1.1424 = 4.2134,$$

and

$$pv = P\{\chi_5^2 > 4.2134\} = 0.5191.$$

So from the result we have no reason to reject the null hypothesis and conclude that the data follow the generalized Weibull distribution.

Remark :

Equiprobable intervals

j	1	2	3	4	5
$\hat{a}_j$	3.2743	6.1343	11.0292	22.6327	46.5536
U_j	8	17	7	7	3
e_j	10.3304	9.8179	8.0722	7.4611	5.4106

$$\hat{g}_{3\times 3} = \begin{pmatrix} 0.0131205 & -0.0023642 & -0.0011198 \\ -0.0023642 & 0.0465760 & -0.0076510 \\ -0.0011198 & -0.0076510 & 0.0029554 \end{pmatrix},$$

$$W_l = (-0.011408 \quad 0.175988 \quad 0.024349)^T.$$

The value of test statistic is

$$Y^2 = X^2 + Q = 5.8447 + 2.5072 = 8.3519,$$

and

$$pv = P\{\chi_5^2 > 8.3519\} = 0.1379.$$

So from the result we have no reason to reject the hypothesis that the head and neck cancer data follow the generalized Weibull distribution. It means that we have same inference as in the case of random grouping intervals but from $p - value$ we can see that the new method is more likely to accept the hypothesis.

Remark on testing $\gamma = 1$:

Weibull distribution is a special case of generalized Weibull distribution when $\gamma = 1$, but for Weibull distribution the matrix $\mathbf{G}$ is degenerated and consequently the limit distribution of the statistic Y_n^2 is chi-square distribution with $k-1$ degrees of freedom. We can get the elements of the quadratic form of Weibull distribution directly from the formulas of generalized Weibull distribution by putting $\gamma = 1$.

$$Q = \frac{W_2^2}{\hat{g}_{22}}, \qquad W_2^2 = \sum_{j=1}^{k} \hat{C}_{2j}\hat{A}_j^{-1}Z_j, \qquad \hat{g}_{22} = \hat{i}_{22} - \sum_{j=1}^{k} \hat{C}_{2j}^2 \hat{A}_j^{-1},$$

$$\hat{i}_{22} = \frac{1}{n\hat{\nu}^2}\sum_{i=1}^{n} \delta_i \left\{1 + \ln Y_i\right\}^2, \qquad \hat{C}_{2j} = \frac{1}{n\hat{\nu}} \sum_{i:X_i \in I_j} \delta_i \left\{1 + \ln Y_i\right\}.$$

In general we don't know that some data follow Weibull or generalized Weibull distribution, if we want to fit one from these two. One way is to get the idea about the model by plotting the hazard function. The other way is that we can use the likelihood ratio test for testing $\gamma = 1$ in case of selecting the model between these two distributions. If we do not reject the null hypothesis $\gamma = 1$, it means the data may follow the Weibull distribution, otherwise it is better to go for the generalized Weibull distribution.

For example, we apply the likelihood ratio test on the head and neck cancer data and the value of test statistic is

$$LR = -2[LL_w - LL_{gw}] = -2[-153.46 + 148.82] = 9.28.$$

As the critical significance level for 1 degree of freedom is $\chi^2(1) = 3.84$, so we can reject the null hypothesis that the data follow the Weibull distribution. In the same way we can use the likelihood ratio test for exponential and Weibull distributions, as exponential is also a case of Weibull distribution when $\nu = 1$.

2.4.4 Exponentiated Weibull Distribution

This distribution is also an extension of the Weibull distribution. We suppose that the failure times follow the exponentiated Weibull distribution under H_0 with the survival function

$$S(t, \theta, \nu, \gamma) = 1 - \left\{1 - exp[-(\frac{t}{\theta})^{\nu}]\right\}^{1/\gamma}, \quad t \geq 0, \quad \theta, \nu, \gamma > 0,$$

and the hazard function

$$\lambda(t, \theta, \nu, \gamma) = \frac{\nu \left\{1 - \exp[-(\frac{t}{\theta})^{\nu}]\right\}^{1/\gamma - 1} \exp[-(\frac{t}{\theta})^{\nu}](\frac{t}{\theta})^{\nu-1}}{\gamma\theta \left\{1 - \left(1 - \exp[-(\frac{t}{\theta})^{\nu}]\right)^{1/\gamma}\right\}}.$$

The parameters can be estimated by maximizing the following log-likelihood function

$$\ell = \ln L \quad = \quad \sum_{i=1}^{n} \delta_i \ln f(X_i, \theta, \nu, \gamma) + \sum_{i=1}^{n} (1 - \delta_i) \ln S(X, \theta, \nu, \gamma)$$

$$= \sum_{i=1}^{n}\delta_i\left\{\ln\nu - \ln\gamma - \nu\ln\theta + (\nu-1)\ln X_i - Y_i + (\frac{1}{\gamma}-1)\ln(1-e^{-Y_i})\right\} + \sum_{i=1}^{n}(1-\delta_i)\ln[1-(1-e^{-Y_i})^{\frac{1}{\gamma}}],$$

where $Y_i = (\frac{X_i}{\theta})^{\nu}$.
The elements of the estimators for $\hat{C}_j$ and for the Fisher information matrix $\hat{i}_{ls}$ can be calculated as

$$\hat{C}_j = \frac{1}{n}\sum_{i:X_i\in I_j}\delta_i\frac{\partial}{\partial\boldsymbol{\theta}}\ln\lambda(X_i;\hat{\boldsymbol{\theta}}), \qquad \hat{C}_j = (\hat{C}_{1j},\hat{C}_{2j},\hat{C}_{3j})^T,$$

$$\hat{C}_{1j} = -\frac{\hat{\nu}}{n\hat{\theta}}\sum_{i:X_i\in I_j}\delta_i\left\{1-\hat{Y}_i+(\frac{1}{\hat{\gamma}}-1)\frac{\hat{Y}_ie^{-\hat{Y}_i}}{1-e^{-\hat{Y}_i}}+\frac{1}{\hat{\gamma}}\frac{\hat{Y}_ie^{-\hat{Y}_i}(1-e^{-\hat{Y}_i})^{1/\hat{\gamma}-1}}{1-(1-e^{-\hat{Y}_i})^{1/\hat{\gamma}}}\right\},$$

$$\hat{C}_{2j} = \frac{1}{n\hat{\nu}}\sum_{i:X_i\in I_j}\delta_i\left\{1+\ln\hat{Y}_i-\hat{Y}_i\ln\hat{Y}_i+(\frac{1}{\hat{\gamma}}-1)\frac{\hat{Y}_ie^{-\hat{Y}_i}\ln\hat{Y}_i}{1-e^{-\hat{Y}_i}}+\frac{\hat{Y}_ie^{-\hat{Y}_i}\ln\hat{Y}_i(1-e^{-\hat{Y}_i})^{1/\hat{\gamma}-1}}{\hat{\gamma}[1-(1-e^{-\hat{Y}_i})^{1/\hat{\gamma}}]}\right\},$$

$$\hat{C}_{3j} = -\frac{1}{n\hat{\gamma}}\sum_{i:X_i\in I_j}\delta_i\left\{1+\frac{1}{\hat{\gamma}}\ln(1-e^{-\hat{Y}_i})[1+\frac{(1-e^{-\hat{Y}_i})^{1/\hat{\gamma}}}{1-(1-e^{-\hat{Y}_i})^{1/\hat{\gamma}}}]\right\},$$

where $\hat{Y}_i = \left(\frac{X_i}{\hat{\theta}}\right)^{\hat{\nu}}$. The information matrix is

$$\hat{\boldsymbol{i}}_{ls} = \frac{1}{n}\sum_{i=1}^{n}\delta_i\frac{\partial}{\partial\theta_l}\ln\lambda(X,\hat{\boldsymbol{\theta}})\frac{\partial}{\partial\theta_s}\ln\lambda(X,\hat{\boldsymbol{\theta}}), \quad l,s=1,2,3,$$

where

$$\hat{i}_{11} = \frac{\hat{\nu}^2}{n\hat{\theta}^2}\sum_{i=1}^{n}\delta_i\left(1-\hat{Y}_i+(\frac{1}{\hat{\gamma}}-1)\frac{\hat{Y}_ie^{-\hat{Y}_i}}{1-e^{-\hat{Y}_i}}+\frac{\hat{Y}_ie^{-\hat{Y}_i}(1-e^{-\hat{Y}_i})^{\frac{1}{\hat{\gamma}}-1}}{\hat{\gamma}\{1-(1-e^{-\hat{Y}_i})^{\frac{1}{\hat{\gamma}}}\}}\right)^2,$$

$$\hat{i}_{22} = \frac{1}{n\hat{\nu}^2}\sum_{i=1}^{n}\delta_i\left(1+\ln\hat{Y}_i-\hat{Y}_i\ln\hat{Y}_i+(\frac{1}{\hat{\gamma}}-1)\frac{\hat{Y}_ie^{-\hat{Y}_i}\ln\hat{Y}_i}{1-e^{-\hat{Y}_i}}+\frac{\hat{Y}_ie^{-\hat{Y}_i}\ln\hat{Y}_i(1-e^{-\hat{Y}_i})^{\frac{1}{\hat{\gamma}}-1}}{\hat{\gamma}\{1-(1-e^{-\hat{Y}_i})^{\frac{1}{\hat{\gamma}}}\}}\right)^2,$$

$$\hat{i}_{33} = \frac{1}{n\hat{\gamma}^2}\sum_{i=1}^{n}\delta_i\left(1+\frac{1}{\hat{\gamma}}\ln(1-e^{-\hat{Y}_i})+\frac{(1-e^{-\hat{Y}_i})^{\frac{1}{\hat{\gamma}}}\ln(1-e^{-\hat{Y}_i})}{\hat{\gamma}\{1-(1-e^{-\hat{Y}_i})^{\frac{1}{\hat{\gamma}}}\}}\right)^2,$$

$$\hat{i}_{12} = -\frac{1}{n\hat{\theta}}\sum_{i=1}^{n}\delta_i\left(1-\hat{Y}_i+(\frac{1}{\hat{\gamma}}-1)\frac{\hat{Y}_ie^{-\hat{Y}_i}}{1-e^{-\hat{Y}_i}}+\frac{\hat{Y}_ie^{-\hat{Y}_i}(1-e^{-\hat{Y}_i})^{\frac{1}{\hat{\gamma}}-1}}{\hat{\gamma}\{1-(1-e^{-\hat{Y}_i})^{\frac{1}{\hat{\gamma}}}\}}\right)\times \left(1+\ln\hat{Y}_i-\hat{Y}_i\ln\hat{Y}_i+(\frac{1}{\hat{\gamma}}-1)\frac{\hat{Y}_ie^{-\hat{Y}_i}\ln\hat{Y}_i}{1-e^{-\hat{Y}_i}}+\frac{\hat{Y}_ie^{-\hat{Y}_i}\ln\hat{Y}_i(1-e^{-\hat{Y}_i})^{\frac{1}{\hat{\gamma}}-1}}{\hat{\gamma}\{1-(1-e^{-\hat{Y}_i})^{\frac{1}{\hat{\gamma}}}\}}\right),$$

$$\hat{i}_{13} = \frac{\hat{\nu}}{n\hat{\theta}\hat{\gamma}}\sum_{i=1}^{n}\delta_i\left(1-\hat{Y}_i+(\frac{1}{\hat{\gamma}}-1)\frac{\hat{Y}_i e^{-\hat{Y}_i}}{1-e^{-\hat{Y}_i}}+\frac{\hat{Y}_i e^{-\hat{Y}_i}(1-e^{-\hat{Y}_i})^{\frac{1}{\hat{\gamma}}-1}}{\hat{\gamma}\{1-(1-e^{-\hat{Y}_i})^{\frac{1}{\hat{\gamma}}}\}}\right)\times$$
$$\left(1+\frac{1}{\hat{\gamma}}\ln(1-e^{-\hat{Y}_i})+\frac{(1-e^{-\hat{Y}_i})^{\frac{1}{\hat{\gamma}}}\ln(1-e^{-\hat{Y}_i})}{\hat{\gamma}\{1-(1-e^{-\hat{Y}_i})^{\frac{1}{\hat{\gamma}}}\}}\right),$$

$$\hat{i}_{23} = -\frac{1}{n\hat{\nu}\hat{\gamma}}\sum_{i=1}^{n}\delta_i\left(1+\frac{1}{\hat{\gamma}}\ln(1-e^{-\hat{Y}_i})+\frac{(1-e^{-\hat{Y}_i})^{\frac{1}{\hat{\gamma}}}\ln(1-e^{-\hat{Y}_i})}{\hat{\gamma}\{1-(1-e^{-\hat{Y}_i})^{\frac{1}{\hat{\gamma}}}\}}\right)\times$$
$$\left(1+\ln\hat{Y}_i-\hat{Y}_i\ln\hat{Y}_i+(\frac{1}{\hat{\gamma}}-1)\frac{\hat{Y}_i e^{-\hat{Y}_i}\ln\hat{Y}_i}{1-e^{-\hat{Y}_i}}+\frac{\hat{Y}_i e^{-\hat{Y}_i}\ln\hat{Y}_i(1-e^{-\hat{Y}_i})^{\frac{1}{\hat{\gamma}}-1}}{\hat{\gamma}\{1-(1-e^{-\hat{Y}_i})^{\frac{1}{\hat{\gamma}}}\}}\right),$$

Choice of $\hat{a}_j$:

Set

$$b_i = (n-i)\Lambda_0(Y_{(i)}) + \sum_{l=1}^{i}\Lambda_0(Y_{(l)}),$$

where

$$\Lambda_0(t) = -\ln[1-(1-e^{-t})^{\frac{1}{\gamma}}].$$

If i is the smallest natural number verifying $E_j \in [b_{i-1}, b_i], j = 1, \cdots, k-1$ then

$$\hat{a}_j = \hat{\theta}\left\{\Lambda_0^{-1}\left([E_j - \sum_{l=1}^{i-1}\Lambda_0(Y_{(l)})]/(n-i+1)\right)\right\}^{1/\hat{\nu}}, \quad \hat{a}_k = \max(X_{(n)}, \tau),$$

where

$$\Lambda_0^{-1}(t) = -\ln[1-(1-e^{-t})^{\gamma}].$$

For such choices of intervals we have $ej = E_k/k$ for any j, where

$$E_k = \sum_{i=1}^{n}\Lambda_0(Y_i).$$

From the above calculations we can find the elements for RRN test of Weibull distribution by putting $\gamma = 1$.

We return to our example with :

$$\hat{\theta} = 0.14405, \quad \hat{\nu} = 0.29435, \quad \hat{\gamma} = 0.05545.$$

Here also $\hat{\gamma} = 0.05545$ is much smaller than 1, which justifies the rejection of Weibull distribution. We take 5 intervals i.e. k=5. Further results to calculate the Y^2 are shown below

j	1	2	3	4	5
$\hat{a}_j$	2.5818	4.6728	8.3907	16.4711	46.5536
U_j	5	11	13	7	6
e_j	8.3936	8.3936	8.3936	8.3936	8.3936

Fisher information matrix is

$$\hat{\boldsymbol{i}}_{3\times 3} = \begin{pmatrix} 33.2772 & -60.3813 & -101.2478 \\ -60.3813 & 133.8186 & 140.5539 \\ -101.2478 & 140.5539 & 387.1453 \end{pmatrix}.$$

$$W_l = (-0.082478 \quad 0.0341998 \quad 0.479222)^T.$$

The value of test statistic is

$$Y^2 = X^2 + Q = 5.7854 + 0.0000 = 5.7854,$$

and

$$pv = P\{\chi^2_5 > 4.2134\} = 0.3277.$$

So from the result we have no reason to reject the hypothesis that the data follow exponentiated Weibull distribution. We can test $\gamma = 1$ (Weibull distribution) in the same way as in the case of generalized Weibull distribution.

2.4.5 Loglogistic Distribution

Loglogistic distribution is one the models having unimodal hazard rate function. It is convenient to apply due to its simple algebraic expressions than the lognormal distribution.

Suppose that under H_0 the failure times follow the loglogistic distribution with

$$S(t;\theta,\nu) = \frac{1}{1+(\frac{t}{\theta})^\nu}, \quad \lambda(t,\theta,\nu) = \frac{\nu}{\theta^\nu} t^{\nu-1} \frac{1}{1+(\frac{t}{\theta})^\nu},$$

$$\Lambda(t,\theta,\nu) = \ln[1+(\frac{t}{\theta})^\nu], \quad (\theta,\nu > 0); \quad t \geq 0,$$

the survival function, the hazard function and the cumulative hazard function respectively. The Log likelihood function is

$$\ell = \sum_{i=1}^n \delta_i \{\ln\nu - \nu\ln\theta + (\nu-1)\ln X_i - \ln(1+(\frac{X_i}{\theta})^\nu)\} - \sum_{i=1}^n \ln(1+(\frac{X_i}{\theta})^\nu).$$

Denote by $\hat{\theta}$ and $\hat{\nu}$ the maximum likelihood estimator of the parameters θ and ν.

The elements of the estimator $\hat{\boldsymbol{i}} = [\hat{i}_{ll'}]_{2\times 2}$ are

$$\hat{i}_{11} = \frac{\hat{\nu}^2}{n\hat{\theta}^2}\sum_{i=1}^{n}\frac{\delta_i}{(1+Y_i)^2}, \quad \hat{i}_{12} = -\frac{1}{n\hat{\theta}}\sum_{i=1}^{n}\delta_i\frac{1+Y_i+\ln Y_i}{(1+Y_i)^2},$$

$$\hat{i}_{22} = \frac{1}{n\hat{\nu}^2}\sum_{i=1}^{n}\delta_i\left(1+\frac{\ln Y_i}{1+Y_i}\right)^2.$$

Choice of $\hat{a}_j$: Set

$$Y_i = (\frac{t_i}{\hat{\theta}})^{\hat{\nu}}, \quad b_i = (n-i)\ln(1+Y_{(i)}) + \sum_{l=1}^{i}\ln(1+Y_{(l)}), \quad E_j = \frac{j}{k}\,b_n.$$

If i is the smallest natural number verifying the inequalities

$$b_{i-1} \le E_j \le b_i,$$

then

$$\hat{a}_j = \hat{\theta}\left[\exp\left\{[E_j - \sum_{l=1}^{i-1}\ln(1+Y_{(l)})]/(n-i+1)\right\} - 1\right]^{1/\hat{\nu}},$$

$$j = 1,\cdots,k-1, \quad \hat{a}_k = \max(X_{(n)},\tau),$$

and with this interval we have $e_j = b_k/n$ for any j.
The other elements of the test statistic are

$$\hat{C}_{1j} = -\frac{\hat{\nu}}{n\hat{\theta}}\sum_{i:X_i\in I_j}(\frac{\delta_i}{1+Y_i}), \quad \hat{C}_{2j} = \frac{1}{n\hat{\nu}}\sum_{i:X_i\in I_j}\delta_i\{1+\frac{\ln Y_i}{1+Y_i}\},$$

$$\hat{A}_j = U_j/n, \quad U_j = \sum_{i:X_i\in I_j}\delta_i, \quad Z_j = \frac{1}{\sqrt{n}}(U_j - e_j).$$

The zero hypothesis is rejected with an approximate significance level α if $Y_n^2 > \chi^2_\alpha(k)$.

The maximum likelihood estimators of loglogistic distribution by taking the survival times in months are ;

$$\hat{\theta} = 7.8303, \quad \hat{\nu} = 1.5265.$$

Taking k=5, further results to calculate the Y_n^2 are shown below :

j	1	2	3	4	5
$\hat{a}_j$	2.7627	5.0426	8.8026	16.3783	46.5537
U_j	7	11	11	7	6
e_j	8.4	8.4	8.4	8.4	8.4

$$\hat{g}_{2\times 2} = \begin{pmatrix} 0.019187 & 0.014895 \\ 0.014895 & 0.167373 \end{pmatrix},$$

and

$$W_l = (-0.035557 \quad 0.010180)^T.$$

The matrix G is non-degenerate. The value of test statistic is

$$Y^2 = X^2 + Q = 2.7491 + 0.0751 = 2.8242,$$

and

$$pv = P\{\chi^2_5 > 2.8242\} = 0.7271.$$

So from the result we have no reason to reject the hypothesis that the head and neck cancer data follow the loglogistic distribution.

2.4.6 Lognormal Distribution

Lognormal distribution is commonly used to model the failure times. Many properties of this distribution follow directly from the properties of normal distribution.

Suppose the distribution of the failure times is lognormal and under H_0

$$S(t;\theta,\nu) = 1 - \Phi(\ln(t/\theta)^\nu), \quad \lambda(t,\theta,\nu) = \nu t^{-1}\frac{\phi(\ln(t/\theta)^\nu)}{1-\Phi(\ln(t/\theta)^\nu)},$$

be the survival function and the hazard function respectively, where ϕ and Φ are the pdf and cdf of the standard normal distribution respectively, θ and ν are unknown scalar parameters. The log likelihood function is

$$\ell(\theta,\nu) = \sum_{i=1}^n \delta_i\{\ln\nu - \ln X_i + \ln\frac{\phi(V_i)}{1-\Phi(V_i)}\} + \ln[1-\Phi(V_i)],$$

where $V_i = \ln(X_i/\theta)^\nu$, ϕ is the pdf of the standard normal distribution. The estimator $\hat{\boldsymbol{i}} = [\hat{i}_{ls}]_{2\times 2}$ has the form :

$$\hat{i}_{11} = \frac{\hat{\nu}^2}{n\hat{\theta}^2}\sum_{i=1}^n \delta_i g_1^2(Y_i), \quad \hat{i}_{12} = -\frac{1}{n\hat{\theta}}\sum_{i=1}^n \delta_i g_1(Y_i)[1 + g_1(Y_i)\ln Y_i],$$

$$\hat{i}_{22} = \frac{1}{n\hat{\nu}^2}\sum_{i=1}^n \delta_i[1 + g_1(Y_i)\ln Y_i]^2,$$

where

$$Y_i = (X_i/\theta)^\nu, \quad g_1(T) = \frac{\phi(\ln t)}{1-\Phi(\ln t)} - \ln t.$$

Choice of $\hat{a}_j$: Set

$$b_i = (n-i)\Lambda_0(Y_i) + \sum_{l=1}^{i} \Lambda_0(Y_{(l)}),$$

where

$$\Lambda_0(t) = -\ln[1-\Phi(\ln t)].$$

If i is the smallest natural number verifying the inequalities $b_{i-1} \le E_j \le b_i$, then for $j = 1, \cdots, k-1$

$$\hat{a}_j = \hat{\theta}\left\{\Lambda_0^{-1}\left([E_j - \sum_{l=1}^{i-1}\Lambda_0(Y_{(l)})]/(n-i+1)\right)\right\}^{1/\hat{\nu}}, \quad \hat{a}_k = \max(X_{(n)}, \tau),$$

where

$$\Lambda_0^{-1}(t) = \exp\{\Phi^{-1}(1-e^{-t})\}.$$

For such choices of intervals we have $e_j = E_k/k$ for any j, where

$$E_k = \sum_{i=1}^{n} \Lambda_0(Y_i).$$

The other elements for the test statistic are

$$\hat{C}_{1j} = -\frac{\hat{\nu}}{n\hat{\theta}} \sum_{i:X_i \in I_j} \delta_i g_1(Y_i), \quad \hat{C}_{2j} = \frac{1}{n\hat{\nu}} \sum_{i:X_i \in I_j} \delta_i\{1 + g_1(Y_i)\ln Y_i\},$$

$$\hat{A}_j = U_j/n, \quad U_j = \sum_{i:X_i \in I_j} \delta_i, \quad Z_j = \frac{1}{\sqrt{n}}(U_j - e_j).$$

The zero hypothesis is rejected with an approximate significance level α if $Y_n^2 > \chi^2_\alpha(k)$.

The maximum likelihood estimators of lognormal distribution by taking the survival times in months are :

$$\hat{\theta} = 8.2135, \qquad \hat{\nu} = 0.8495.$$

The results to calculate the Y^2 with k=5 are shown below :

j	1	2	3	4	5
$\hat{a}_j$	2.5521	4.6866	8.5134	16.8011	46.5537
U_j	5	11	13	7	6
e_j	8.4	8.4	8.4	8.4	8.4

$$\hat{\boldsymbol{G}} = \hat{g}_{2\times 2} = \begin{pmatrix} 0.000434 & 0.022043 \\ 0.022043 & 1.148915 \end{pmatrix},$$

$$W_l = (-0.022228 \quad 1.252508)^T.$$

The matrix G is non-degenerate. The value of test statistic is

$$Y^2 = X^2 + Q = 5.7989 + 1.6605 = 7.4594,$$

and

$$pv = P\{\chi^2_5 > 7.4594\} = 0.1887.$$

So from the result we have no reason to reject the hypothesis that the head and neck cancer data follow the lognormal distribution.

2.4.7 Birnbaum-Saunders Distribution

An argument of *fatigue or cumulative damage* justifies the use of the Birnbaum-Saunders (BS) distribution. The Birnbaum-Saunders distribution (fatigue life distribution) is used commonly in reliability applications to model failure times.

Suppose that under H_0 the failure times T_i follow the two-parameter Birnbaum-Saunders distribution with cumulative distribution function

$$F(t;\alpha,\beta) = \Phi\left[\frac{1}{\alpha}\left\{\left(\frac{t}{\beta}\right)^{\frac{1}{2}} - \left(\frac{\beta}{t}\right)^{\frac{1}{2}}\right\}\right], \quad 0 < t < \infty, \quad \alpha,\beta > 0,$$

where α is the shape parameter, β is the scale parameter and $\Phi(x)$ is the standard normal distribution function. The probability density function can be written as

$$\begin{aligned} f(t;\alpha,\beta) = \quad & \frac{1}{2\sqrt{2\pi}\ \alpha\beta}\left\{\left(\frac{\beta}{t}\right)^{\frac{1}{2}} + \left(\frac{\beta}{t}\right)^{\frac{3}{2}}\right\}\exp\left[-\frac{1}{2\alpha^2}\left(\frac{t}{\beta} + \frac{\beta}{t} - 2\right)\right], \\ & 0 < t < \infty, \quad \alpha,\beta > 0. \end{aligned}$$

The loglikelihood function is

$$\begin{aligned} \ell \quad = \quad & \sum_{i=1}^{n}\delta_i\left[\ln\left(\frac{1}{2\sqrt{2\pi}}\right) - \ln\alpha - \ln\beta + \ln\left\{\left(\frac{\beta}{X_i}\right)^{\frac{1}{2}} + \left(\frac{\beta}{X_i}\right)^{\frac{3}{2}}\right\} - \frac{1}{2\alpha^2}\left(\frac{X_i}{\beta} + \frac{\beta}{X_i} - 2\right)\right] \\ & + \sum_{i=1}^{n}(1-\delta_i)\ln\left(1 - \Phi\left[\frac{1}{\alpha}\left\{\left(\frac{X_i}{\beta}\right)^{\frac{1}{2}} - \left(\frac{\beta}{X_i}\right)^{\frac{1}{2}}\right\}\right]\right). \end{aligned}$$

We can estimate the elements of Fisher information matrix as

$$\hat{\boldsymbol{i}}_{ll'} = \frac{1}{n}\sum_{i=1}^{n}\delta_i\frac{\partial\ln\lambda(X_i;\hat{\boldsymbol{\theta}})}{\partial\boldsymbol{\theta}_l}\frac{\partial\ln\lambda(X_i;\hat{\boldsymbol{\theta}})}{\partial\boldsymbol{\theta}_{l'}},$$

where

$$\hat{\boldsymbol{i}}_{11} = \frac{1}{n\hat{\alpha}^2}\sum_{i=1}^{n}\delta_i\left[-1+(A(X_i))^2-\frac{A(X_i)\varphi(A(X_i))}{1-\Phi(A(X_i))}\right]^2,$$

$$\hat{\boldsymbol{i}}_{22} = \frac{1}{n\hat{\beta}^2}\sum_{i=1}^{n}\delta_i\left[-1+\frac{1}{2}\left(\frac{1+3\frac{\hat{\beta}}{X_i}}{1+\frac{\hat{\beta}}{X_i}}\right)+\frac{1}{2}A(X_i)B(X_i)-\frac{1}{2}\frac{B(X_i)\varphi(A(X_i))}{1-\Phi(A(X_i))}\right]^2,$$

$$\hat{\boldsymbol{i}}_{12} = \frac{1}{n\hat{\alpha}\hat{\beta}}\sum_{i=1}^{n}\delta_i\left[-1+(A(X_i))^2-\frac{A(X_i)\varphi(A(X_i))}{1-\Phi(A(X_i))}\right]\times \left[-1+\frac{1}{2}\left(\frac{1+3\frac{\hat{\beta}}{X_i}}{1+\frac{\hat{\beta}}{X_i}}\right)+\frac{1}{2}A(X_i)B(X_i)-\frac{1}{2}\frac{B(X_i)\varphi(A(X_i))}{1-\Phi(A(X_i))}\right].$$

where

$$A(X_i)=\frac{1}{\alpha}\left\{\left(\frac{X_i}{\beta}\right)^{\frac{1}{2}}-\left(\frac{\beta}{X_i}\right)^{\frac{1}{2}}\right\},\quad B(X_i)=\frac{1}{\alpha}\left\{\left(\frac{X_i}{\beta}\right)^{\frac{1}{2}}+\left(\frac{\beta}{X_i}\right)^{\frac{1}{2}}\right\},$$

φ and Φ are the density function and cumulative distribution function of the standard normal distribution. The elements of matrix $\hat{\boldsymbol{C}}$ are

$$\hat{\boldsymbol{C}}_{1j} = \frac{1}{n\hat{\alpha}}\sum_{i:X_i\in I_j}\delta_i\left[-1+(A(X_i))^2-\frac{A(X_i)\varphi(A(X_i))}{1-\Phi(A(X_i))}\right],$$

$$\hat{\boldsymbol{C}}_{2j} = \frac{1}{n\hat{\beta}}\sum_{i:X_i\in I_j}\delta_i\left[-1+\frac{1}{2}\left(\frac{1+3\frac{\hat{\beta}}{X_i}}{1+\frac{\hat{\beta}}{X_i}}\right)+\frac{1}{2}A(X_i)B(X_i)-\frac{1}{2}\frac{B(X_i)\varphi(A(X_i))}{1-\Phi(A(X_i))}\right].$$

Remark : Bagdonavicius et al. (2010a) give the explicit formula to estimate a_j for the shape-scale families of distributions in the form of inverse hazard function. As there is no explicit form of the inverse hazard function of Birnbaum-Saunders distribution, so we estimate intervals by iterative method.

The maximum likelihood estimators of Birnbaum-Saunders distribution by taking into account the survival times in months are ; $\hat{\alpha} = 1.4390$, $\hat{\beta} = 7.6851$. We take 5 intervals i.e. k=5. Further results to calculate the Y^2 are shown below :

j	1	2	3	4	5
$\hat{a}_j$	2.0371	3.8279	7.1816	15.8970	46.5537
U_j	3	7	16	10	6
e_j	8.55876	8.55876	8.55876	8.55876	8.55876

$$\hat{G} = \begin{pmatrix} 1.21220 & -0.06967 \\ -0.06967 & 0.01340 \end{pmatrix}.$$

$$W_l = (-3.68787 \quad 0.30150)^T$$

The value of test statistic is

$$Y^2 = X^2 + Q = 15.4067 + 12.0731 = 27.4798,$$

and

$$pv = P\{\chi_5^2 > 27.4798\} = 0.000046.$$

So from the result we can say that Birnbaum-Saunders distribution does not fits the head and neck cancer data.

In Figure 2.6 the empirical survival function i.e. Kaplan Meier curve for head and neck cancer data is compared with fitted survival functions based on ML estimators for various shape-scale models.

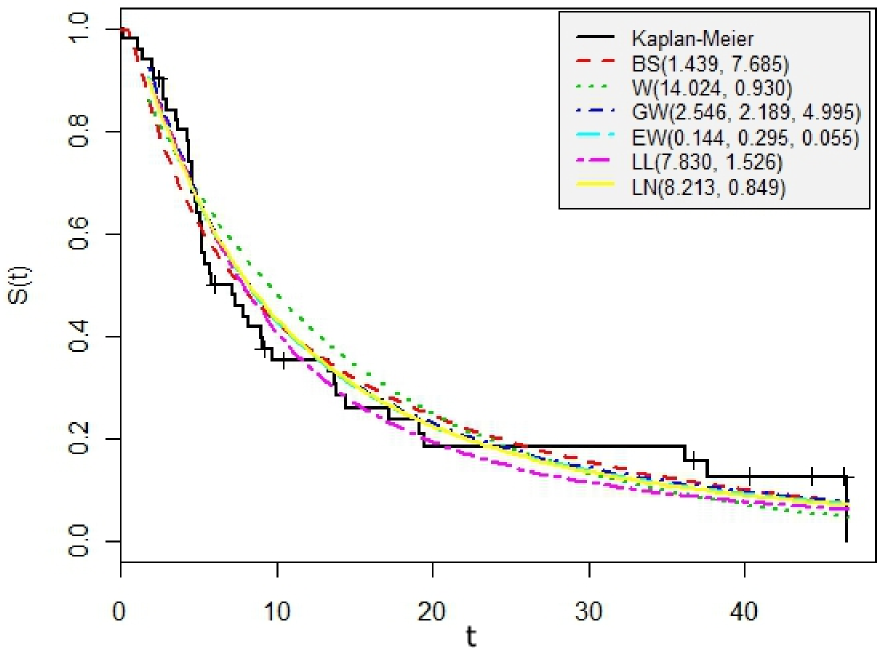

Figure 2.6 – The empirical survival function (Kaplan-Meier) and the fitted survival functions (MLE).

2.4.8 Power Of The Test

Here power of the RRN test is calculated for different percentage of censoring. We take the same values of parameters for simulations as estimated from the head and neck cancer data for different models. In table 2.1 the power of test is shown for loglogistic distribution taking generalized Weibull as alternative, and in table 2.2 for Birnbaum-Saunders distribution against generalized Weibull distribution. From table 2.2 one can see that Birnbaum-Saunders is found to be more powerful which is expected as BS distribution does not fits the data and also power increases with the increase in sample size and decreases with the increase in censoring percentage.

	n					
Censoring (%)	50	100	150	200	300	500
10	16.52	21.95	24.67	30.76	33.97	36.05
20	18.70	22.20	26.10	25.88	28.23	23.30
30	21.40	28.21	36.36	30.69	23.10	23.50

Table 2.1 – Power of RRN (Loglogistic Vs GW)

	Censoring	n					
Model	(%)	50	100	150	200	300	500
GW	10	32.20	44.80	57.50	64.80	74.60	89.90
	20	26.18	40.40	46.30	53.00	62.40	71.40
	30	26.30	35.50	42.00	48.40	54.60	57.63

Table 2.2 – Power of RRN (BS Vs GW)

3 Goodness-of-Fit Tests In Demography and Assurance

In demography, Gompertz and Makeham models have significant role in modeling and in analysis of mortality and ageing. Till the end of 20th century, researchers have used the tables of mortalities (also called life tables) for demographic analysis but in the end of 20th century due to the development in statistical methods of survival analysis and reliability one can treat the individuals data even with the information of censoring. The Gompertz, Makeham, and Weibull models are compared with respect to the goodness-of-fit to the table of mortality and to the individuals data in the presence of censoring. For data from the table of mortality, the test statistic considered by Gerville-Reache and Nikulin (2000) is used. For censored individual data the chi-squared type test proposed by Bagdonavicius et al. (2010a) is used. The choice of random grouping intervals is made to overcome the problem of very small expected number of events for some interval. This can happen in demography because the number of deaths at early age is very small.

In reliability and demography model selection for some specific data is vital for further analysis and decision making. Testing the two-parameter Gompertz distribution (Gompertz, 1825) to model the rate of mortality has been used for a long time, where the rate of mortality increases with the age. Gompertz-Makeham model (William Makeham, 1860) with one additional parameter covers the mortality independent of age. The researchers have used the life and mortality tables to find the force of mortality. Gerville-Reache and Nikulin (2000) gave a chi-square type goodness-of-fit test for Makeham model using the table of mortality (grouped data). In section 3 we briefly discuss their proposed statistic and also we compare Makeham model with Gompertz and Weibull models for different age groups. But now with the advanced data collection techniques, one can have the individual's information (ungrouped data) also with censoring mechanism. Gompertz and Makeham models are frequently used in demography but in reliability Weibull model is considered the alternative for Gompertz model (Juckett and Rosenberg, 1993).

Most researchers compare Gompertz model with the Weibull model due to its flexible parameters (Gavrilov and Gavrilova, 2001). Logistic distribution can be another alternative for Gompertz (Wilson, 1994). The Gompertz function is a better choice for all causes of mortality and combined disease categories while the Weibull model has been shown to be a better choice over Gompertz model for a specific cause of mortality (Juckett and Rosenberg, 1993). Nikulin et al. (2011a) presented several models in demography but here we consider the Gompertz-Makeham and Weibull models for censored data and the following results are published by Gerville-Rache, Nikulin and Tahir 2012.

3.1 Gompertz-Makeham and Weibull Models in Demograpgy

Gompertz model of aging is widely used in demography and other scientific disciplines e.g. medical sciences, survival analysis, actuarial sciences and reliability. Gompertz (1825) gave the first mathematical model to explain the exponential increase in mortality rate with age. He explained that the law of geometric progression pervades in mortality after a certain age. Gompertz mortality rate can be presented as

$$\mu_x = \theta e^{\nu x}, \quad (\theta, \nu) > 0, \quad x > 0, \tag{2.19}$$

where θ is known as the baseline mortality and ν is the age specific growth rate of the force of mortality.

Mortality rate μ_x in demographic notation is equivalent to the failure rate $\mu(x)$ in reliability or hazard rate $\lambda(x)$ in survival analysis. The Gompertz law has been the main demographic model since its discovering to fit the human mortality (see for example Gavrilov and Gavrilova, 2001 ; Ricklefs and Scheuerlein, 2002).

Since Gompertz model gives the rate of mortality only related to age and does not take into account the other factors independent of age, other researchers tried to modify this model to fulfill the requirement of real data. William Makeham (1860) modified the Gompertz model considering some other causes of death independent of age by proposing the so called ***Gompertz-***

Makeham law of mortality as

$$\mu_x = \gamma + \theta e^{\nu x}, \quad \text{where} \quad (\gamma, \theta, \nu) > 0 \quad x > 0. \tag{2.20}$$

Here the first term γ (Makeham parameter) is a constant and non-aging component of failure rate (e.g. accidents, independent of age) and the second term $\theta e^{\nu t}$ is the Gompertz function depending on age (aging factor).

The ***Weibull*** distribution is one of the most widely used distributions in survival analysis and reliability due to the characteristics of its shape parameter ν. The mortality rate or hazard function is

$$\mu_x = \frac{\nu}{\theta^\nu} x^{\nu-1}, \quad \text{for} \quad x \geq 0 \quad (\theta, \nu) > 0. \tag{2.21}$$

The hazard function of the Weibull distribution can be decreasing, constant or increasing according to the value of its shape parameter i.e. three Weibull models can make a bathtub shape, but now there are some models like the generalized Weibull model which can have bathtub shape (Bagdonavicius and Nikulin, 2002). The Weibull law is more commonly applicable for technical devices while the Gompertz law is more common for biological systems (Gavrilov & Gavrilova, 1991). When the Gompertz law fails to follow some biological failure mechanism, the best alternative is Weibull law due to its basis on reliability theory. If the probability of failure at the start of the system is almost zero, the failure rate increases with the power function with age i.e. Weibull law and if the system has defects at the beginning, the failure rate increases exponentially with age i.e. Gompertz law. So, to apply the Weibull law in demography, the biological population should be independent of initial deaths. Logistic distribution is considered as the other alternative for Gompertz distribution (Vanfleteren et al., 1998).

3.2 Test Statistic For The Table Of Mortality

Consider $x = 0$ as the origin of time for an individual of age x, and T_x is a random variable for its residual life from this origin. The probability of death is

$${}_tq_x = \mathbf{P}\{0 < T_x \leq t\}, \quad t > 0, x > 0.$$

So the annual rate of mortality for the people having age x can be defined as

$$q_x = \mathbf{P}\{0 < T_x \leq 1\}, \quad x > 0.$$

A relation between the rate of mortality and the instantaneous rate of mortality μ_x is

$$q_x = 1 - \exp\left(-\int_x^{x+1} \mu_y dy\right), \quad x > 0.$$

The theoretical annual rate of mortality in the case of *Gompertz* model can be written as

$$q_x = 1 - \exp\left(-\frac{\theta}{\nu}e^{\nu x}(e^{\nu} - 1)\right), \quad \theta, \nu > 0. \tag{2.22}$$

In the same way we can find the theoretical annual rate of mortality for Makeham, Weibull and other parametric models.

We observe the n persons independent of mortality and we regroup them in the same age, say ω groups, where ω is the maximum age in years. The group G_x contains ℓ_x persons of age x $(x = 0, \cdots, \omega - 1)$ and q_x is the probability of death of each individual in the year. Let denote by D_x the number of deaths in the group G_x.
Using the data D_x and ℓ_x from the table of mortality, we can obtain the empirical annual rate of mortality observed at age x, such that

$$Q_x = \frac{D_x}{\ell_x},$$

which follows the binomial law with parameters ℓ_x and q_x. According to the central limit theorem if $\min_x(\ell_x) \to \infty$ when $n \to \infty$, then $Q = (Q_0, \cdots, Q_{\omega-1})' \sim^{as} N_\omega(q, P)$, where $q = (q_0, \cdots, q_{\omega-1})'$ and P is the diagonal matrix of the elements $\frac{q_x(1-q_x)}{\ell_x}$ for $x = 0, 1, \cdots, \omega - 1$.
So we can write that

$$\frac{(D_x - \ell_x q_x)^2}{\ell_x q_x(1 - q_x)} \sim^{as} \chi^2_1.$$

As it is shown in Gerville-Reache and Nikulin (2000),

$$X^2_\omega = \sum_{x=0}^{\omega-1} \frac{(D_x - \ell_x q_x)^2}{\ell_x q_x(1 - q_x)} \sim^{as} \chi^2_\omega.$$

One can use this statistic for testing simple hypotheses, as one uses the classical Pearson statistic for testing simple hypotheses (see Greenwood and Nikulin, 1996).

3.2.1 Estimation Of Parameters In Composite Hypothesis

Let consider the composite hypothesis

$$H_0 : q_x = q_x(\boldsymbol{\theta}), \qquad \boldsymbol{\theta} = (\theta_1, \cdots, \theta_s)' \in \Theta \subseteq R^s, s < \omega.$$

We estimate the parameters by the maximum likelihood method using the data from the table of mortality. We have the random variable D_x which follows the binomial law with parameters

ℓ_x and q_x. The likelihood function is

$$L(\boldsymbol{\theta}) = \prod_{x=0}^{\omega-1} \binom{\ell_x}{D_x} [q_x(\boldsymbol{\theta})]^{D_x} [1 - q_x(\boldsymbol{\theta})]^{\ell_x - D_x}.$$

We take the estimator $\hat{\boldsymbol{\theta}}$ that maximizes the likelihood function, i.e. $\hat{\boldsymbol{\theta}} = \operatorname{argmax} L(\boldsymbol{\theta})$.

One can find the maximum likelihood estimator $\hat{\boldsymbol{\theta}}$ for $\boldsymbol{\theta}$ by solving the following score vector

$$\frac{\partial \ln L}{\partial \theta_i} = 0, \forall\, i = 1, \cdots, s.$$

Let consider the statistic

$$X^2_\omega(\hat{\boldsymbol{\theta}}) = \sum_{x=0}^{\omega-1} \frac{(D_x - \ell_x q_x(\hat{\boldsymbol{\theta}}))^2}{\ell_x q_x(\hat{\boldsymbol{\theta}})(1 - q_x(\hat{\boldsymbol{\theta}}))} \sim^{as} \chi^2_{\omega-s}.$$

Gerville-Reache and Nikulin (2000) showed that under the hypothesis H_0, $X^2_\omega(\hat{\boldsymbol{\theta}})$ asymptotically follows a chi-square statistic with $\omega - s$ degrees of freedom, where s is the number of parameters to be estimated, from where it follows that we may use this statistic for testing H_0. One can see that the statistic $X^2_\omega(\hat{\boldsymbol{\theta}})$ is different from the classical Pearson statistic.

3.2.2 Example : Data Analysis From The Table Of Mortality

The data in Table 2.3 is from INSEE Aquitaine-France and give the number of deaths D_x in 1990 for each 5-year age group, where ℓ_x is the number of habitants for each age group on January 1^{st} 1990. This data is used for the validity of three models i.e. *Gompertz*, *Makeham*,

age	ℓ_x	D_x	*age*	ℓ_x	D_x
5-9	75498	14	45-49	64575	195
10-14	77284	16	50-54	57974	247
15-19	90337	45	55-59	61871	384
20-24	102544	91	60-64	62473	622
25-29	91339	92	65-69	61122	958
30-34	90769	128	70-74	36425	944
35-39	93324	156	75-79	37124	1341
40-44	96692	226	80-84	29541	2020

Table 2.3 – Table of Mortality (INSEE, Gironde 1990)

and *Weibull*. The rate of mortality for these three models is adjusted with maximum likelihood estimators and then the value of chi-square is calculated. In case of the adjustment between 5 and 84 years of age, the annual rate of mortality follows neither the Gompertz and Makeham nor the Weibull model. But when the adjustment is made for the age groups between 30 and 74 years, the Makeham model is accepted. The Gompertz model also becomes valid with Makeham when the annual rate of mortality is adjusted for the age between 50 and 79 years. It means that Gompertz model is validated in the older age and it coincides with the theory regarding Gompertz model as discussed in the previous section. The Weibull model gets close but still it does not fit the data significantly. The calculated values of the test statistic with corresponding p-values are shown in table 2.3 and the fitted models are presented in figures 2.7-2.9.

Table 2.4 – Results from the table of mortality

	Gompertz		Makeham		Weibull	
Age Groups	$X^2_\omega(\hat{\theta})$	p-value	$X^2_\omega(\hat{\theta})$	p-value	$X^2_\omega(\hat{\theta})$	p-value
5-84	214.19	≈ 0	99.98	≈ 0	2363.98	≈ 0
30-74	45.62	≈ 0	3.70	0.72	158.93	≈ 0
50-79	9.01	0.11	8.48	0.08	25.44735	0.0001

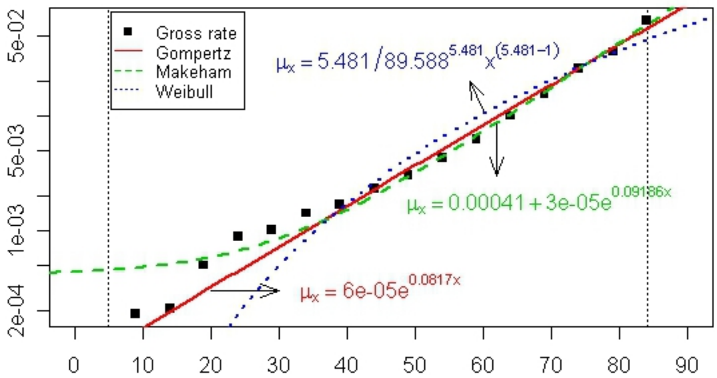

Figure 2.7 – Model fitted for age between 5 and 84 years (log scale)

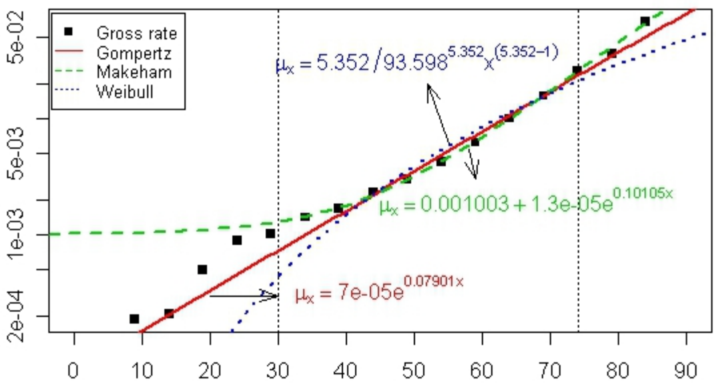

Figure 2.8 – Model fitted for age between 30 and 74 years (log scale)

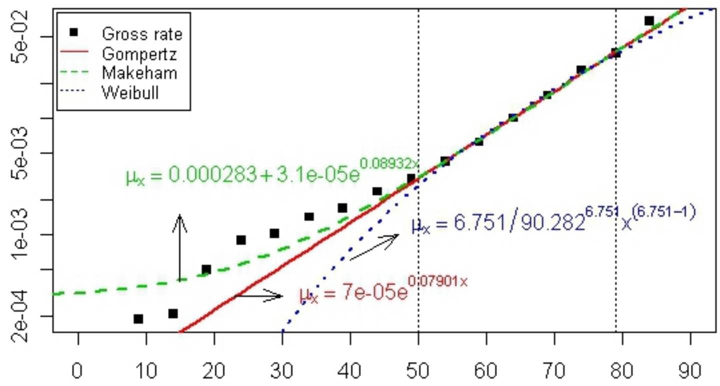

Figure 2.9 – Model fitted for age between 50 and 79 years (log scale)

3.3 Comparison of Gompertz, Weibull and Makeham Models

Here we use the same test as proposed in section 2 for individual right censored data in demography. Let consider the hypothesis that under H_0 the distribution of the failure times is *Gompertz* with hazard function and cumulative hazard function given by ;

$$\mu_x = \theta e^{\nu x}, \quad \Lambda_x = \frac{\theta}{\nu}(e^{\nu x} - 1) \quad x > 0, \quad (\theta, \nu) > 0.$$

The loglikelihood function is

$$\ell(\theta, \nu) = \sum_{i=1}^{n} \left\{ \delta_i [\ln \theta + \nu X_i] - \frac{\theta}{\nu}(e^{\nu X_i} - 1) \right\}.$$

Let denote by $\hat{\theta}$ and $\hat{\nu}$ the ML estimators of θ and ν.

Since the matrix G is found to be degenerated, the quadratic form can be written as :

$$Q = \frac{W_2^2}{\hat{g}_{22}},$$

where

$$\hat{g}_{22} = \hat{i}_{22} - \sum_{j=1}^{k} \hat{C}_{2j}^2 \hat{A}_j^{-1}, \quad \hat{i}_{22} = \frac{1}{n} \sum_{i=1}^{n} \delta_i X_i^2, \quad \hat{C}_{2j} = \frac{1}{n} \sum_{i:X_i \in I_j} \delta_i X_i,$$

$$\hat{A}_j = \frac{U_j}{n}, \quad W_2 = \sum_{j=1}^{k} \hat{C}_{2j} \hat{A}_j^{-1} Z_j, \quad Z_j = \frac{1}{\sqrt{n}}(U_j - e_j).$$

Choice of $\hat{a}_j$: Set

$$b_i = (n-i)\frac{\hat{\theta}}{\hat{\nu}}(e^{\hat{\nu} X_{(i)}} - 1) + \frac{\hat{\theta}}{\hat{\nu}} \sum_{l=1}^{i} (e^{\hat{\nu} X_{(l)}} - 1), \quad i = 1, \cdots, n.$$

If i is the smallest natural number satisfying the inequalities

$$b_{i-1} \leq E_j \leq b_i, \qquad E_j = \frac{j}{k} b_n,$$

then for $j = 1, \cdots, k-1$

$$\hat{a}_j = \frac{1}{\hat{\nu}} \ln \left\{ 1 + \frac{\hat{\nu}}{\hat{\theta}} \left(\frac{j}{k} b_n - \frac{\hat{\theta}}{\hat{\nu}} \sum_{l=1}^{i-1} (e^{\hat{\nu} X_{(l)}} - 1) \right) / (n - i + 1) \right\}, \quad \hat{a}_k = \max(X_{(n)}, \tau).$$

For such choices of intervals we have $e_j = E_k/k$ for any j.

Example : This data is taken from the book of Bagdonavicius et al. (2010a). n = 120 electronic devices were observed for time $\tau = 5.54$ (years). The number of failures is δ =113 :

1.7440	1.9172	2.1461	2.3079	2.3753	2.3858	2.4147	2.5404	2.6205	2.6471
2.8370	2.8373	2.8766	2.9888	3.0720	3.1586	3.1730	3.2132	3.2323	3.3492
3.3507	3.3514	3.3625	3.3802	3.3855	3.4012	3.4382	3.4438	3.4684	3.5019
3.5110	3.5297	3.5363	3.5587	3.5846	3.5992	3.654	3.6574	3.6674	3.7062
3.7157	3.7288	3.7502	3.7823	3.8848	3.8902	3.9113	3.9468	3.9551	3.9728
3.9787	3.9903	4.0078	4.0646	4.1301	4.1427	4.2300	4.2312	4.2525	4.2581
4.2885	4.2919	4.2970	4.3666	4.3918	4.4365	4.4919	4.4932	4.5388	4.5826
4.5992	4.6001	4.6324	4.6400	4.7164	4.7300	4.7881	4.7969	4.8009	4.8351
4.8406	4.8532	4.8619	4.8635	4.8679	4.8858	4.8928	4.9466	4.9846	5.0008
5.0144	5.0517	5.0898	5.0929	5.0951	5.1023	5.1219	5.1223	5.1710	5.1766
5.1816	5.2441	5.2546	5.3353	5.4291	5.4360	5.4633	5.4842	5.4860	5.4903
5.5199	5.5232	5.5335.							

Suppose the failure times have a *Gompertz* distribution. The maximum likelihood estimators of *Gompertz model* are ; $\hat{\theta} = 0.0051$, $\hat{\nu} = 1.1586$. We take 10 intervals i.e. k=10. Further results to calculate Y_n^2 are shown below :

j	1	2	3	4	5	6	7	8	9	10
$\hat{a}_j$	2.70	3.33	3.74	4.07	4.34	4.57	4.78	5.00	5.25	5.54
U_j	10	9	23	12	9	6	7	13	13	11
e_j	11.3	11.3	11.3	11.3	11.3	11.3	11.3	11.3	11.3	11.3

$$\hat{i}_{22} = 16.7779, \quad \hat{g}_{22} = 0.0141, \quad W_2 = -0.3737.$$

The matrix G is degenerate, so $r = k - 1 = 9$. The value of test statistic is

$$Y_n^2 = X^2 + Q = 15.1130 + 9.8867 = 24.9997,$$

and

$$pv = P\{\chi_9^2 > 24.9997\} = 0.0053.$$

So from the result we can say that failure times don't follow Gompertz distribution.

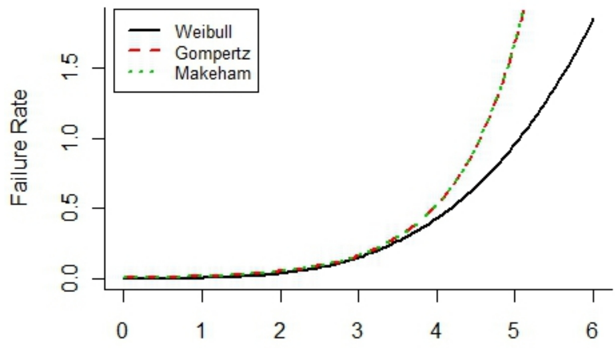

Figure 2.10 – The failure rate of electronic devices

Suppose that the failure times follow a *Weibull model*. The maximum likelihood estimators of Weibull model are ; $\hat{\theta} = 4.6078$, $\hat{\nu} = 4.9554$. We take 10 intervals i.e. $k = 10$. Further results to calculate Y_n^2 are shown below :

j	1	2	3	4	5	6	7	8	9	10
$\hat{a}_j$	2.89	3.36	3.70	3.98	4.24	4.47	4.68	4.90	5.16	5.54
U_j	13	9	17	12	7	8	8	13	11	15
e_j	11.3	11.3	11.3	11.3	11.3	11.3	11.3	11.3	11.3	11.3

$$\hat{i}_{22} = 0.0618, \quad \hat{g}_{22} = 0.0027, \quad W_2 = -0.0545.$$

The matrix G is degenerate, so $r = k - 1 = 9$. The value of test statistic is

$$Y_n^2 = X^2 + Q = 9.2692 + 1.0845 = 10.3536,$$

and

$$pv = P\{\chi_9^2 > 10.3536\} = 0.3226.$$

So from the result we have no reason to reject that the failure times follow the Weibull distribution. In the same way we can apply the test for Makeham model.

Here Weibull model gives the better fit which is expected since the data refer to technical devices and according to Gavrilov and Gavrilova (2001) technical deceives fail according to the Weibull law. Also from Figure 2.10 one can observe the behavior of Gompertz model, according to which in later times the failure rate increases very fast. The hazard plot for Makeham model coincides with the Gompertz because the estimated value of Makeham parameter (non-aging component) γ is negligibly small for this data.

The hazard function and the cumulative hazard function of Makeham model is given as

$$\mu_x = \gamma + \theta e^{\nu x}, \qquad \Lambda_x = \gamma x + \frac{\theta}{\nu}(e^{\nu x} - 1).$$

The loglikelihood function is

$$\ell(\theta, \nu) = \sum_{i=1}^{n} \left\{ \delta_i [\ln(\gamma + \theta e^{\nu x})] - \gamma x - \frac{\theta}{\nu}(e^{\nu x} - 1) \right\}.$$

The information matrix $\hat{i}$ and the vector $\hat{C}_j$ are

$$\hat{i}_{11} = \frac{1}{n}\sum_{i=1}^{n} \delta_i \left(\frac{1}{\gamma + \theta e^{\nu x}}\right)^2, \quad \hat{i}_{12} = \frac{1}{n}\sum_{i=1}^{n} \delta_i \frac{e^{\nu x}}{(\gamma + \theta e^{\nu x})^2},$$

$$\hat{i}_{13} = \frac{1}{n}\sum_{i=1}^{n} \delta_i \frac{x\theta e^{\nu x}}{(\gamma + \theta e^{\nu x})^2}, \quad \hat{i}_{23} = \frac{1}{n}\sum_{i=1}^{n} \delta_i \frac{x\theta e^{2\nu x}}{(\gamma + \theta e^{\nu x})^2},$$

$$\hat{i}_{22} = \frac{1}{n}\sum_{i=1}^{n} \delta_i \left(\frac{e^{\nu x}}{\gamma + \theta e^{\nu x}}\right)^2, \quad \hat{i}_{33} = \frac{1}{n}\sum_{i=1}^{n} \delta_i \left(\frac{x\theta e^{\nu x}}{\gamma + \theta e^{\nu x}}\right)^2,$$

$$\hat{C}_{1j} = \frac{1}{n}\sum_{i:X_i \in I_j} \delta_i \left(\frac{1}{\gamma + \theta e^{\nu x}}\right), \quad \hat{C}_{2j} = \frac{1}{n}\sum_{i:X_i \in I_j} \delta_i \left(\frac{e^{\nu x}}{\gamma + \theta e^{\nu x}}\right),$$

$$\hat{C}_{3j} = \frac{1}{n}\sum_{i:X_i \in I_j} \delta_i \left(\frac{x\theta e^{\nu x}}{\gamma + \theta e^{\nu x}}\right).$$

The MLE for Makeham distribution are :

$$\gamma = 6.2022 \times 10^{-08}, \quad \theta = 0.0051, \quad \nu = 1.1586.$$

The end points of the intervals with observed and expected frequencies are

j	1	2	3	4	5	6	7	8	9	10
$\hat{a}_j$	2.70	3.33	3.74	4.07	4.34	4.57	4.78	4.99	5.24	5.54
U_j	10	9	23	12	9	6	7	13	13	11
e_j	11.3	11.3	11.3	11.3	11.3	11.3	11.3	11.3	11.3	11.3

$$\hat{i}_{3\times 3} = \begin{pmatrix} 24.7020 & 557.3764 & 8.7870 \\ 557.3764 & 35891.5809 & 757.3056 \\ 8.7870 & 757.3056 & 16.7779 \end{pmatrix},$$

$$W_l = (-0.2191 \quad -0.0075 \quad -0.3739)^T.$$

The value of test statistic is

$$Y_n^2 = X^2 + Q = 15.1131 + 29.9280 = 45.0411,$$

and

$$pv = P\{\chi_9^2 > 45.0411\} = 2.1379 \times 10^{-6}.$$

So we can say that the Makeham model does not fits the data. One can see that due to the quadratic form Q the hypothesis is rejected otherwise it is accepted.

Chapter 3

AFT Regression Analysis With BS Distribution

1 Introduction

In ALT, Accelerated Failure Time model is commonly used (Bagdonavicius and Nikulin, 2002). The objective of this theory is to know the influence of the stresses (covariates) on the life duration of the items. Our purpose is to estimate the reliability under specified values of interest of these variables. The AFT model allows us to control the degradation process and to search the optimal condition for the products or systems. This model allows to foresee the product or system reliability under normal conditions or standard stress. Accelerated experiments occurred under higher stress than usual stress or standard stress.

Types of stress :

- Constant stress : the stress remains constant during the time, and each item is tested at a constant stress level,
- Step-stress : these are the mostly used time varying stresses in ALT i.e. the units are tested under lower initial stress and if they do not fail in specified time, the stress is increased. If they do not fall in the next specified time, the stress is increased again and so on.
- Progressive stress : this type of stress increases continuously with time.
- Cyclic stress : some products repeatedly undergo a cyclic stress loading with amplitude.
- Random stress : the products are tested under a variable continuous in time stress.

AFT model is called parametric when baseline survival function belongs to the parametric families of distributions such as Weibull, log-normal or generalized Weibull, inverse Gaussian, and Birnbaum-Saunders. AFT model is semi-parametric when the survival function is unknown and is non-parametric when the survival function and life-stress relationship are unknown. Here we consider the parametric AFT when the base line survival function belongs to Birnbaum-Saunders distribution. We estimate the parameters, give all the elements of information matrix, and calculate the confidence interval for the survival function of the AFT model. One can find sufficient information on ALT method from literature (see for example, Singpurwalla (1971), Meeker and Escobar (1998), Bagdonavicius and Nikulin (2002), Lawless (2003), Nelson (2004), Huber-Carol et al. (2008)).

Let consider E is a set of all possible admissible m-dimensional time dependent stresses or covariates

$$E = \{x(\cdot) = (x_0(\cdot), x_1(\cdot), \cdots, x_m(\cdot))^T \quad x : [0\,, \infty[\in \boldsymbol{R}^m\}$$

We write x instead of $x(\cdot)$ if the stress is constant and we denote by $E_1 \subset E$ the set of all constant stresses and $x_0 \in E_0 \subset E$ be the usual stress (standard or normal). We suppose that the failure time $T_{x(\cdot)}$ under stress $x(\cdot)$ is a positive random variable with survival function

$$S_{x(\cdot)}(t) = P\{T_{x(\cdot)} > t\}\,, \, t > 0, \quad x(\cdot) \in E$$

Let denote by

$$\lambda_{x(\cdot)}(t) = \frac{f_{x(\cdot)}(t)}{S_{x(\cdot)}(t)} = -\frac{S'_{x(\cdot)}(t)}{S_{x(\cdot)}(t)}, \quad \text{and} \quad \Lambda_{x(\cdot)}(t) = -\ln S_{x(\cdot)}(t)$$

the hazard function and cumulative hazard function respectively.

Definition 1 *A stress $x_2(\cdot)$ is accelerated with respect $x_1(\cdot)$, that is $x_2(\cdot) > x_1(\cdot)$ if $S_{x_1(\cdot)}(t) \geq S_{x_2(\cdot)}(t)$ (see Figure 3.1).*

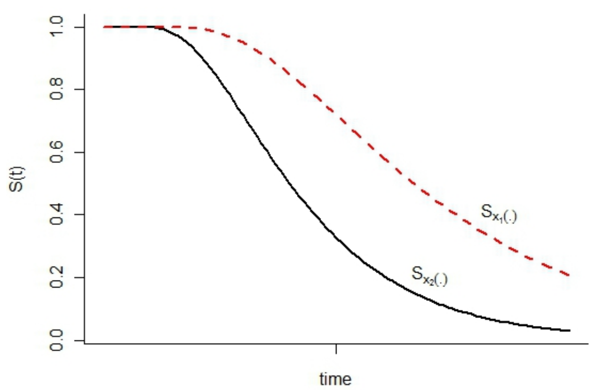

Figure 3.1 – Survival curves with two different stresses.

Let $f_{x(\cdot)}(t)$ be a used S_0-resource under stress $x(\cdot)$ until the moment t (failure time) and is expressed as

$$f_{x(\cdot)}(t) = S_0^{-1}(S_{x(\cdot)}(t)), \quad \text{and} \quad S_0^{-1}(t) = \inf\{s\, : S_0(s) \geq p\},$$

where S_0 is the survival function at normal stress. So we can write for all $x(\cdot) \in E$

$$S_0(f_{x(\cdot)}(t)) = S_{x(\cdot)}(t)\,, t > 0,$$

which means that the survival probability at t under stress $x(\cdot)$ is the same as that of at the moment $f_{x(\cdot)}(t)$ under the normal stress. This transfer functional is explained in Figure 3.2.

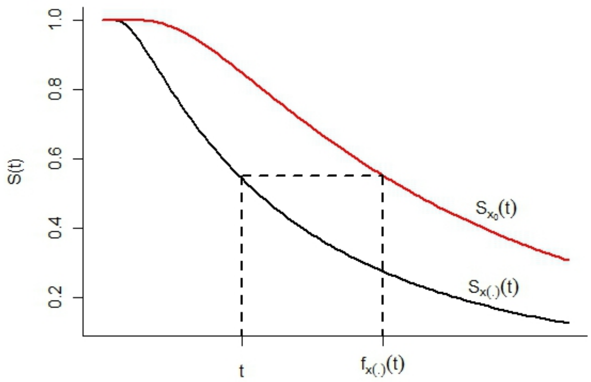

Figure 3.2 – Transfer functional of the survival curves.

Definition 2 *The AFT model is defined on E, if the survival function under the stress $x(\cdot) \in E$ is :*

$$S_{x(\cdot)}(t) = S_0\left(\int_0^t r(x(u))du\right), \quad x(\cdot) \in E. \tag{3.1}$$

If $x(\tau) = x$ is constant, then the model (3.1) becomes :

$$S_{x(\cdot)}(t) = S_0(r(x)t), \quad x \in E_1. \tag{3.2}$$

The function r can be chosen from the certain class of functions. Often the AFT model is parameterized as

$$r(x) = e^{-\beta^T z},$$

where $\beta = (\beta_o, \beta_1, \cdots, \beta_m) \in \boldsymbol{R}^{m+1}$ is the vector of unknown parameters, and

$$z = (z_o, z_1, \cdots, z_m) = (\psi_0(x), \cdots, \psi_m(x))^T$$

is a vector of known stress functions ψ_i, where $\psi_0(t) = 1$.
So the parametric AFT model on E is

$$S_{x(.)}(t) = S_0\left(\int_0^t \exp\{-\beta^T x(\tau)\}d\tau\right), \quad x(\cdot) \in E, \tag{3.3}$$

and on the set of stresses E_1 i.e. constant in time

$$S_{x(.)}(t) = S_0(e^{-\beta^T x} t), \quad x \in E_1, \tag{3.4}$$

and the logarithm of the failure time T_x under x maybe written as

$$\ln(T_x) = \beta^T x + \varepsilon,$$

where the survival function of ε does not depend on x and is $S(t) = S_0(\ln t)$. Often S_0 belongs to a specified shape-scale class of survival functions :

$$S_0(t) = G_0\left((\frac{t}{\theta})^\nu \right),$$

where

$$G_0(t) = e^{-t}, \qquad (1+t)^{-1}, \qquad 1 - \Phi(\ln t),$$

belongs to the Weibull, loglogistic, lognormal distributions, respectively. Φ is the distribution function of the standard normal distribution. If the model (3.2) holds on E_0, then for all $x_1, x_2 \in E_0$:

$$S_{x_2}(t) = S_{x_1}(\rho(x_1, x_2)t)$$

where $\rho(x_1, x_2) = r(x_2)/r(x_1)$ shows the degree of scale variation. It is evident that $\rho(x, x) = 1$. If we suppose that this stress $x \in E_0$ is one-dimensional, the rate of scale variation is thus expressed (Viertl-1988) by :

$$\delta(x) = \lim_{\Delta x \to 0} \frac{\rho(x, x + \Delta x) - \rho(x, x)}{\Delta x} = [\log r(x)]'. \tag{3.5}$$

So for all $x \in E_0$

$$r(x) = r(x_0) \exp\left\{ \int_{x_0}^{x} \delta(v) dv \right\},$$

where x_0 is a fixed stress. One can remark that the random variable

$$R = \int_0^{T_x(\cdot)} e^{-\beta^T x(s)} ds,$$

is a parameter free with the survival function $S_0(t)$.

2 AFT Model Parametrization

Suppose that $\delta(x)$ is proportional to specified function $u(x)$

$$\delta(x) = \alpha u(x)$$

In one dimensional stress case

$$r(x) = \exp\{-\beta_0 - \beta_1 \psi_1(x)\}$$

where β_0, β_1 are unknown parameters and ψ is a given function of x which can be parameterized in many ways such as :

- $r(x) = e^{-\beta_0 - \beta_1 x}, \quad \psi(x) = x$, this is the log-linear model.
- $r(x) = e^{-\beta_0 - \beta_1 \ln x}, \quad \psi(x) = \ln(x)$, this is the power-rule model.
- $r(x) = e^{-\beta_0 - \beta_1 / x}, \quad \psi(x) = 1/x$, this is the Arrhenius model.

The above three models are the particular cases of the

$$\delta(x) = \alpha x^{\gamma}$$

where γ is unknown, and we can write in terms of $r(x)$ as

$$r(x) = \begin{cases} e^{-\beta_0 - \beta_1 (x^{\epsilon} - 1)/\epsilon}, & \text{if } \epsilon \neq 0; \\ e^{-\beta_0 - \beta_1 \log(x)}, & \text{if } \epsilon = 0. \end{cases}$$

If $\delta(x) = 1/x + \alpha/x^2$ then

$$r(x) = e^{-\beta_0 - \beta_1 \log(x) - \beta_2 / x} = \alpha_1 x e^{-\beta_2 / x}$$

where $\beta_1 = -1$ and it is the Eyring model, applied when the explanatory variable x is the temperature.

If the explanatory variable or stress $x = (x_1, ..., x_m)$ is constant and m-dimensional and if there is no interaction between them, then the model can be generalized as :

$$r(x) = \exp\{-\beta_0 - \sum_{i=1}^{m} \sum_{j=1}^{n_i} \beta_{ij} z_{ij}(x_i)\}, \tag{3.6}$$

where $z_{ij}(x_i)$ are known functions and β_{ij} are unknown parameters.

3 Plans of Experiments

As mentioned in the previous section that the purpose of accelerated life testing is to find the reliability of the components under usual stress from the data obtained by using the higher stresses that is the accelerated stresses. So for using higher stresses several plans of experiments can be found in the literature (see for example, Bagdonavicius et al. (2002) and Nelson (2004)).

3.1 First plan of Experiments

Let denote by $x^{(0)} = (x_{00}, x_{01}, \cdots, x_{0m})$, where $x_{00} = 1$ is the usual stress. Generally one-dimensional stress ($m = 1$), or two-dimensional stress ($m = 2$) are used for ALT experiments. Let $x_1, \cdots, x_k$ be constant over time accelerated stresses :

$$x_0 < x_1 < \cdots < x_k,$$

here $x_i = (x_{i0}, x_{i1}, \cdots, x_{im})^T$, $x_{i0} = 1$. The usual stress x_0 is not used during experiments. According to this plan of experiments k groups of units are tested. The i^{th} group of n_i units , $\sum_{i=1}^{k} n_i = n$, is tested under the stress x_i. The data can be complete or independently right censored.

3.2 Second plan of experiments

In second plan of experiments the step-stress are used where n units are tested in such a way that at first they have to undergo at low stress and if they do not fail in predetermined time then the stress is increased. If they do not fail at the increased stress in the given time the stress is increase again and so on. The step-stress can be formulated as

$$x(\omega) = \begin{cases} x_1, & 0 \le \omega < t_1, \\ x_2, & t_1 \le \omega < t_2, \\ \vdots & \\ x_k, & t_{k-1} \le \omega < t_k = +\infty, \end{cases} \tag{3.7}$$

where $x_j = (x_{j0}, x_{j1}, \cdots, x_{jm})^T \in E_m$, $x_{j0} = 1$. The function $r(x)$ should be parameterized, and the survival function under any stress $x(\cdot)$ of the plan (3.7) can be written as for $i = 1, \cdots, k$

$$S_{x(\cdot)}(t) = S_0 \left\{ 1_{\{i>1\}} \sum_{j=1}^{i-1} e^{-\beta^T x_j}(t_j - t_{i-1}) + e^{-\beta^T x_i}(t - t_{i-1}) \right\}, \quad t \in [t_{i-1}, t_i[.$$

3.3 Third plan of experiments

In this plan of experiment two groups of units are tested. The first group of n_1 units is tested under a constant accelerated stress x_1 and the second group of n_2 units are tested under a step stress x_2 ; stress x_1 until the moment t_1 then under the usual stress x_0 until the moment t_2 i.e.

$$x_2(\omega) = \begin{cases} x_1, & 0 \le \omega \le t_1, \\ x_0, & t_1 < \omega \le t_2. \end{cases} \tag{3.8}$$

Units use much of their resources until the moment t_1 under the accelerated stress x_1, so after the switch-up failures occur in the interval $[t_1, t_2]$ even the usual or normal stress. The survival function in AFT model is

$$S_{x_1}(u) = S_{x_0}(ru),$$

where $r = r(x_1)/r(x_0)$, and

$$S_{x_2(\cdot)}(u) = \begin{cases} S_{x_0}(ru), & 0 \le u \le t_1 \\ S_{x_0}(ru)(rt_1 + u - t_1), & t_1 < u \le t_2, \end{cases} \tag{3.9}$$

or

$$S_{x_2(\cdot)}(t) = S_{x_0}(r(u \wedge t_1) + (u - t_1) \vee 0), \tag{3.10}$$

where $a \wedge b = \min(a, b)$ and $a \vee b = \max(a, b)$.

3.4 Fourth plan of experiments

This plan of experiment is used when the failure-time distribution is exponential. For this plan k groups of units are observed. The $i - th$ group of n_i units is tested under one-dimensional constant stress $x^{(i)}$ until the $r_i - th$ failure $(r_i \le n_i)$ (type II censoring). The failure moment of $i - th$ group are $T_{i1} \le \cdots \le T_{ir_i}$, where $i = 1, \cdots, k$.

Note : In this thesis, only the first two plans of experiments are considered for parameter estimation and confidence interval calculation. In first plan the constant stress x_j is used all time of the experiment for the jth group of units while in the second plan the constant stress x_j is used in the interval $[t_{j-1}, t_j)$ for all units.

4 Failure Time Regression Analysis

Suppose that n units are observed. The ith unit is tested under the stress

$$x^{(i)}(\cdot) = (x_0^{(i)}(\cdot), x_1^{(i)}(\cdot), \cdots, x_m^{(i)}(\cdot))^T,$$

and consider that

$$(X_1, \delta_1, x_1), \ldots, (X_n, \delta_n, x_n), \quad X_i = T_i \wedge C_i, \quad \delta_i = \mathbf{1}_{\{T_i \leq C_i\}}, \tag{3.11}$$

is a right censored failure time regression sample. $T_1, \cdots, T_n$ are the failure times which are absolutely continuous i.i.d. random variables, C_i are the censoring times, and $x_i = (1, x_{i1}, \cdots, x_{im})$ is a vector of covariates.

The purpose of the section is to give the analysis (i.e. estimation of parameters, confidence interval for survival or reliability function, and goodness-of-fit test) of accelerated failure time regression data when the base line survival function belongs to the Birnbaum-Saunders distributions family. First some general calculations are given for the shape-Scale distributions (Weibull, loglogistic and lognormal).

4.1 Shape-Scale families of distributions

Consider the AFT model

$$S_{x(\cdot)}(t) = S_0\left(\int_0^t \exp\{-\beta^T x(u)\} du\right), \tag{3.12}$$

where S_0 belongs to the a specified shape-scale parametric class of distributions, i.e.

$$S_0(t) = G_0\{(\frac{t}{\eta})^\nu\}, \quad \eta, \nu > 0.$$

For example, if

$$G_0(t) = e^{-t}, \qquad G_0(t) = (1+t)^{-1}, \qquad G_0(t) = 1 - \Phi(\ln t),$$

then we get the class of distributions of Weibull, loglogistic, lognormal distributions, respectively. Φ is the distribution function of the standard normal distribution. The parameter η maybe included in the coefficient β_0, we suppose that :

$$S_0(t, \sigma) = G_0(t^{\frac{1}{\sigma}}), \quad \sigma = \frac{1}{\nu}.$$

The model (3.12) can be written in the form :

$$S_i(t,\, \beta,\, \sigma) = G_0\left\{\left(\int_0^t e^{-\beta^T x^{(i)}(u)} du\right)^{\frac{1}{\sigma}}\right\}.$$

If $x^{(i)}$ is constant then

$$S_i(t) = G\left(\frac{\ln t - \beta^T x^{(i)}}{\sigma} du\right),$$

where

$$G(u) = G_0(e^u), \quad u \in \boldsymbol{R}.$$

Set

$$g(u) = -G'(u), \quad h(u) = \frac{g(u)}{G(u)}$$

For Weibull distribution

$$G(u) = e^{-e^u}, \quad g(u) = e^u e^{-e^u}, \quad h(u) = e^u, \quad (\ln h(u))' = 1.$$

In the same way one can write the formulas for loglogistic and lognormal distributions. So the likelihood function for shape-scale families for constant $x^{(i)}$ can be written as

$$L(\beta, \sigma) = \prod_{i=1}^{n} \left\{ \frac{1}{\sigma X_i} h\left(\frac{\ln X_i - \beta^T x^{(i)}}{\sigma}\right)\right\}^{\delta_i} G\left(\frac{\ln X_i - \beta^T x^{(i)}}{\sigma}\right).$$

Bagdonavicius and Nikulin (2002) gave the parameter estimation and all the reliability characteristics of these models. Also they calculated the reliability characteristics along with asymptotic confidence intervals for generalized Weibull distribution which can have a $\bigcap$-shape and $\bigcup$-shape (for certain parameters) form of the hazard function, which is very common for reliability data. Saaidia et al. (2010) calculated the reliability characteristics of the AFT model using inverse Gaussian distribution with the asymptotic confidence intervals of the survival or reliability function.

In the next section parameter estimation and the asymptotic confidence interval of the AFT model are given when the base line survival function belongs to the Birnbaum-Saunders distribution.

5 Birnbaum-Saunders AFT model

Birnbaum-Saunders is a life distribution model based on cycles of stress causing degradation or crack growth. This distribution is reasonable alternative to Weibull, lognormal, inverse Gaussian, etc. Due to its theoretical argument of fatigue life distribution it is commonly used by the engineers in the reliability studies of the fatigue process. Here we use this model for the failure time regression analysis in accelerated life testing. Details are given in section 9 of chapter 1.

Suppose that the baseline survival function of the AFT belongs to the BS family as

$$S_0(t) = 1 - \Phi\left[\frac{1}{a}\left\{\left(\frac{t}{b}\right)^{\frac{1}{2}} - \left(\frac{b}{t}\right)^{\frac{1}{2}}\right\}\right], \quad 0 < t < \infty, \quad a, b > 0,$$

here Φ is the standard normal distribution function.
Under the AFT model the survival function $S_{x^{(i)}(\cdot)}$ can be written as

$$S_i(t;\beta,a,b) = 1 - \Phi\left[\frac{1}{a}\left\{\left(\frac{\int_0^t e^{-\beta^T x^{(i)}(u)}du}{b}\right)^{\frac{1}{2}} - \left(\frac{b}{\int_0^t e^{-\beta^T x^{(i)}(u)}du}\right)^{\frac{1}{2}}\right\}\right],$$

and for constant stresses $x^{(i)}$

$$S_i(t;\beta,a,b) = 1 - \Phi\left[\frac{1}{a}\left\{\left(\frac{te^{-\beta^T x^{(i)}}}{b}\right)^{\frac{1}{2}} - \left(\frac{b}{te^{-\beta^T x^{(i)}}}\right)^{\frac{1}{2}}\right\}\right].$$

5.1 Estimation Of Parameters

For the parameter estimation in the BS-AFT model we consider the first plan of experiments and suppose that t_i be the maximum time of experiment for ith group of size n_i under the accelerated stress x_i, $(i = 1, \cdots, k)$. $\beta = (\beta_0, \beta_1, \cdots, \beta_m)^T$ is the vector of regression parameters. The life time of the jth component or unit from the ith group is T_{ij}, such that $X_{ij} = T_{ij} \wedge t_i$ and $\delta_{ij} = 1_{\{T_{ij} < ti\}}$ is the indicator variable for censoring.

The likelihood function for the AFT model under the constant stress $x_i \in E_0$ can be written as

$$L = \prod_{i=1}^{k}\prod_{j=1}^{n_i}(e^{-\beta^T x^{(i)}})^{\delta_{ij}}\left(f_0(e^{-\beta^T x^{(i)}}X_{ij})\right)^{\delta_{ij}}\left(S_0(e^{-\beta^T x^{(i)}}X_{ij})\right)^{1-\delta_{ij}}.$$

For BS-AFT model the log likelihood function is

$$\begin{aligned}\ell(\beta,a,b) \quad = \quad & \sum_{i=1}^{k}\sum_{j=1}^{n_i}\delta_{ij}\beta^T x^{(i)} + \sum_{i=1}^{k}\sum_{j=1}^{n_i}\delta_{ij}\left[-\ln a - \ln b + \ln\left\{\left(\frac{b}{X_i e^{-\beta^T x^{(i)}}}\right)^{\frac{1}{2}} + \left(\frac{b}{X_i e^{-\beta^T x^{(i)}}}\right)^{\frac{3}{2}}\right\}\right. \\ & \left. -\frac{1}{2a^2}\left(\frac{X_i e^{-\beta^T x^{(i)}}}{b} + \frac{b}{X_i e^{-\beta^T x^{(i)}}} - 2\right)\right] + \\ & \sum_{i=1}^{k}\sum_{j=1}^{n_i}(1-\delta_{ij})\ln\left(1 - \Phi\left[\frac{1}{a}\left\{\left(\frac{X_i e^{-\beta^T x^{(i)}}}{b}\right)^{\frac{1}{2}} - \left(\frac{b}{X_i e^{-\beta^T x^{(i)}}}\right)^{\frac{1}{2}}\right\}\right]\right).\end{aligned}$$

We put $K_{ij} = X_{ij}e^{-\beta^T x^{(i)}}$, the loglikelihood function becomes

$$\begin{aligned}\ell(\beta,a,b) &= \sum_{i=1}^{k}\sum_{j=1}^{n_i}\delta_{ij}\beta^T x^{(i)} + \sum_{i=1}^{k}\sum_{j=1}^{n_i}\delta_{ij}\left[-\ln a - \ln b + \ln\left\{\left(\frac{b}{K_{ij}}\right)^{\frac{1}{2}} + \left(\frac{b}{K_{ij}}\right)^{\frac{3}{2}}\right\}\right.\\ &\left. -\frac{1}{2a^2}\left(\frac{K_{ij}}{b} + \frac{b}{K_{ij}} - 2\right)\right] + \\ &\sum_{i=1}^{k}\sum_{j=1}^{n_i}(1-\delta_{ij})\ln\left(1 - \Phi\left[\frac{1}{a}\left\{\left(\frac{K_{ij}}{b}\right)^{\frac{1}{2}} - \left(\frac{b}{K_{ij}}\right)^{\frac{1}{2}}\right\}\right]\right). \qquad (3.13)\end{aligned}$$

Parameters are estimated by putting the partial derivatives with respect to $\beta_l, (l = 0, 1, \cdots, m)$, a and b equal to zero. The score vector $U_l(\beta, a, b), (l = 0, 1, \cdots, m+2)$ is written as

$$\begin{aligned}U_l(\beta,a,b) = \frac{\partial\ell(\beta,a,b)}{\partial\beta_l} &= \sum_{i=1}^{k}\sum_{j=1}^{n_i}\delta_{ij}x_{il} + \sum_{i=1}^{k}\sum_{j=1}^{n_i}\delta_{ij}\frac{x_{il}}{2}\left[\frac{1+3\left(\frac{b}{K_{ij}}\right)}{1+\left(\frac{b}{K_{ij}}\right)} + A_{ij}B_{ij}\right] + \\ &\frac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{n_i}(1-\delta_{ij})\left[\frac{x_{il}B_{ij}\varphi(A_{ij})}{1-\Phi(A_{ij})}\right], \quad l = 0, 1, \cdots, m,\end{aligned}$$

$$U_{m+1}(\beta,a,b) = \frac{\partial\ell(\beta,a,b)}{\partial a} = \frac{1}{a}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\delta_{ij}\left[A_{ij}^2 - 1\right] + \frac{1}{a}\sum_{i=1}^{k}\sum_{j=1}^{n_i}(1-\delta_{ij})\left[\frac{A_{ij}\varphi(A_{ij})}{1-\Phi(A_{ij})}\right],$$

$$\begin{aligned}U_{m+2}(\beta,a,b) = \frac{\partial\ell(\beta,a,b)}{\partial b} &= \frac{1}{b}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\delta_{ij}\left[-1 + \frac{1}{2}\left(\frac{1+3\frac{b}{K_{ij}}}{1+\frac{b}{K_{ij}}}\right) + \frac{1}{2}A_{ij}B_{ij}\right] + \\ &\frac{1}{2b}\sum_{i=1}^{k}\sum_{j=1}^{n_i}(1-\delta_{ij})\left[\frac{B_{ij}\varphi(A_{ij})}{1-\Phi(A_{ij})}\right],\end{aligned}$$

where

$$A_{ij} = \frac{1}{a}\left\{\left(\frac{K_{ij}}{b}\right)^{\frac{1}{2}} - \left(\frac{b}{K_{ij}}\right)^{\frac{1}{2}}\right\}, \quad B_{ij} = \frac{1}{a}\left\{\left(\frac{K_{ij}}{b}\right)^{\frac{1}{2}} + \left(\frac{b}{K_{ij}}\right)^{\frac{1}{2}}\right\}.$$

The elements of the Fisher information matrix $I(\beta, a, b) = I_{ls}(\beta, a, b)$ are given as

$$\begin{aligned}I_{ls} &= -\frac{\partial^2\ell(\beta,a,b)}{\partial\beta_l\beta_s} = -\frac{1}{a^2}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\delta_{ij}x_{il}x_{is}\left[\frac{1}{B_{ij}^2} - \frac{a^2A_{ij}^2}{2} - 1\right] \\ &-\frac{1}{4}\sum_{i=1}^{k}\sum_{j=1}^{n_i}(1-\delta_{ij})x_{il}x_{is}\varphi(A_{ij})\left[\frac{(1-\Phi(A_{ij}))A_{ij}\left[B_{ij}^2 - 1\right] - B_{ij}^2\varphi(A_{ij})}{(1-\Phi(A_{ij}))^2}\right],\end{aligned}$$

$$\begin{aligned} I_{l,m+1} &= -\frac{\partial^2 \ell(\beta,a,b)}{\partial\beta_l \partial a} = \frac{1}{a}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\delta_{ij}x_{il}A_{ij}B_{ij} - \\ &\frac{1}{2a}\sum_{i=1}^{k}\sum_{j=1}^{n_i}(1-\delta_{ij})x_{il}B_{ij}\varphi(A_{ij})\left[\frac{(1-\Phi(A_{ij}))\left[A_{ij}^2-1\right]-A_{ij}\varphi(A_{ij})}{(1-\Phi(A_{ij}))^2}\right], \end{aligned}$$

$$\begin{aligned} I_{l,m+2} &= -\frac{\partial^2 \ell(\beta,a,b)}{\partial\beta_l \partial b} = -\frac{1}{a^2 b}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\delta_{ij}x_{il}\left[\frac{1}{B_{ij}^2}-\frac{a^2A_{ij}^2}{2}-1\right] \\ &-\frac{1}{4b}\sum_{i=1}^{k}\sum_{j=1}^{n_i}(1-\delta_{ij})x_{il}\varphi(A_{ij})\left[\frac{(1-\Phi(A_{ij}))A_{ij}\left[B_{ij}^2-1\right]-B_{ij}^2\varphi(A_{ij})}{(1-\Phi(A_{ij}))^2}\right], \end{aligned}$$

$$\begin{aligned} I_{m+1,m+1} &= -\frac{\partial^2 \ell(\beta,a,b)}{\partial a^2} = -\frac{1}{a^2}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\delta_{ij}\left[1-3A_{ij}^2\right] - \\ &\frac{1}{a^2}\sum_{i=1}^{k}\sum_{j=1}^{n_i}(1-\delta_{ij})A_{ij}\varphi(A_{ij})\left[\frac{(1-\Phi(A_{ij}))\left[A_{ij}^2-2\right]-A_{ij}\varphi(A_{ij})}{(1-\Phi(A_{ij}))^2}\right], \end{aligned}$$

$$\begin{aligned} I_{m+1,m+2} &= -\frac{\partial^2 \ell(\beta,a,b)}{\partial a \partial b} = \frac{1}{ab}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\delta_{ij}A_{ij}B_{ij} - \\ &\frac{1}{2ab}\sum_{i=1}^{k}\sum_{j=1}^{n_i}(1-\delta_{ij})B_{ij}\varphi(A_{ij})\left[\frac{(1-\Phi(A_{ij}))\left[A_{ij}^2-1\right]-A_{ij}\varphi(A_{ij})}{(1-\Phi(A_{ij}))^2}\right], \end{aligned}$$

$$\begin{aligned} I_{m+2,m+2} &= -\frac{\partial^2 \ell(\beta,a,b)}{\partial b^2} = -\frac{1}{b^2}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\delta_{ij}\left[1-\frac{\frac{K_{ij}}{b}+2+\frac{3b}{K_{ij}}}{2a^2B_{ij}^2}-\frac{K_{ij}}{a^2 b}\right] - \\ &\frac{1}{4b^2}\sum_{i=1}^{k}\sum_{j=1}^{n_i}(1-\delta_{ij})\varphi(A_{ij})\left[\frac{(1-\Phi(A_{ij}))\left[-2B_{ij}-A_{ij}+A_{ij}B_{ij}^2\right]-B_{ij}^2\varphi(A_{ij})}{(1-\Phi(A_{ij}))^2}\right], \end{aligned}$$

where $\varphi(t)$, $\Phi(t)$ are respectively the pdf and the cdf of the standard normal distribution and the following expression are used for simplification

$$K_{ij} = X_{ij}e^{-\beta^T x^{(i)}},$$

$$A_{ij} = \frac{1}{a}\left\{\left(\frac{K_{ij}}{b}\right)^{\frac{1}{2}} - \left(\frac{b}{K_{ij}}\right)^{\frac{1}{2}}\right\}, \quad B_{ij} = \frac{1}{a}\left\{\left(\frac{K_{ij}}{b}\right)^{\frac{1}{2}} + \left(\frac{b}{K_{ij}}\right)^{\frac{1}{2}}\right\}.$$

The asymptotic distribution of $(\hat{\beta}, \hat{a}, \hat{b})^T$ when n_i are large is approximately normally distributed that is

$$(\hat{\beta}, \hat{a}, \hat{b})^T \approx N((\beta, a, b)^T, \Sigma(\beta, a, b)), \tag{3.14}$$

and the covariance matrix $\Sigma(\beta, a, b)$ can be estimated by $I^{-1}(\hat{\beta}, \hat{a}, \hat{b}) = I^{ls}(\hat{\beta}, \hat{a}, \hat{b})_{(m+3)\times(m+3)}$ (Greenwood and Nikulin, 1996).

5.2 Estimation Of Survival Function

If $\hat{\beta}, \hat{a}, \hat{b}$ are the ML estimator then the estimator of survival function under constant accelerated stress $x \in E_0$ is

$$\hat{S}_x(t) = 1 - \Phi\left[\frac{1}{\hat{a}}\left\{\left(\frac{te^{-\hat{\beta}^T x}}{\hat{b}}\right)^{\frac{1}{2}} - \left(\frac{\hat{b}}{te^{-\hat{\beta}^T x}}\right)^{\frac{1}{2}}\right\}\right].$$

And the estimated survival function under the normal or usual stress when $x = x^{(0)}$ can be written as

$$\hat{S}_x^{(0)}(t) = 1 - \Phi\left[\frac{1}{\hat{a}}\left\{\left(\frac{te^{-\hat{\beta}^T x^{(0)}}}{\hat{b}}\right)^{\frac{1}{2}} - \left(\frac{\hat{b}}{te^{-\hat{\beta}^T x^{(0)}}}\right)^{\frac{1}{2}}\right\}\right]$$

5.3 Asymptotic Confidence Interval For Survival Function

By using the properties of ML estimators (3.14), under the usual stress an approximate $(1-\alpha)$-percent confidence interval for the survival function $S_x^{(0)}(t)$ is

$$\left(1 + \frac{1 - \hat{S}_{x^{(0)}}(t)}{\hat{S}_{x^{(0)}}(t)} e^{\pm\hat{\sigma}_{Q_{x^{(0)}}}\omega_{1-\frac{\alpha}{2}}}\right)^{-1},$$

where ω_α is the α-quantile of the normal distribution and

$$\hat{\sigma}^2_{Q_{x^{(0)}}} = \frac{J_g^T(\hat{\theta}) I^{-1}(\hat{\theta}) J_g(\hat{\theta})}{(\hat{S}_{x^{(0)}}(t))^2 (1 - \hat{S}_{x^{(0)}}(t))^2},$$

where

$$
\begin{aligned}
J_g^T(\hat{\theta})I^{-1}(\hat{\theta})J_g(\hat{\theta}) &= \left(\frac{\partial S}{\partial \beta_0}, \cdots, \frac{\partial S}{\partial \beta_m}, \frac{\partial S}{\partial a}, \frac{\partial S}{\partial b}\right) \begin{pmatrix} I^{00} & \cdots & I^{0(m+2)} \\ \vdots & & \\ I^{i0} & \cdots & I^{i(m+2)} \\ \vdots & & \\ I^{(m+2)0} & \cdots & I^{(m+2)(m+2)} \end{pmatrix} \begin{pmatrix} \frac{\partial S}{\partial \beta_0} \\ \vdots \\ \frac{\partial S}{\partial \beta_m} \\ \frac{\partial S}{\partial a} \\ \frac{\partial S}{\partial b} \end{pmatrix} \\
&= \left(\frac{\partial S}{\partial \beta_0}, \cdots, \frac{\partial S}{\partial \beta_m}, \frac{\partial S}{\partial a}, \frac{\partial S}{\partial b}\right) \begin{pmatrix} \sum_{j=0}^{m+2} I^{0j} \frac{\partial S}{\partial \beta_j} \\ \sum_{j=0}^{m+2} I^{1j} \frac{\partial S}{\partial \beta_j} \\ \vdots \\ \sum_{j=0}^{m+2} I^{(m+2)j} \frac{\partial S}{\partial \beta_j} \end{pmatrix} \\
&= \sum_{k=0}^{m+2} \sum_{j=0}^{m+2} \frac{\partial S}{\partial \beta_k} I^{kj} \frac{\partial S}{\partial \beta_j}.
\end{aligned}
$$

Or it can be written in simple form as

$$
\hat{\sigma}^2_{Q_{x(0)}} = \frac{1}{\left(\hat{S}_{x^{(0)}}(t)\right)^2 \left(1 - \hat{S}_{x^{(0)}}(t)\right)^2} \sum_{k=0}^{m+2} \sum_{j=0}^{m+2} a_k(t; \hat{\beta}, \hat{a}, \hat{b}) I^{kj}(\hat{\beta}, \hat{a}, \hat{b}) a_j(t; \hat{\beta}, \hat{a}, \hat{b}), \tag{3.15}
$$

where $a_k(\hat{\beta}, \hat{a}, \hat{b})$ and $a_j(\hat{\beta}, \hat{a}, \hat{b})$ are the partial derivatives of the survival function with respect to the parameters and for BS-AFT model such as

$$
a_k(t; \hat{\beta}, \hat{a}, \hat{b}) = \frac{1}{2} x_k B_t \varphi(A_t), quadk = 0, 1, \cdots, m,
$$

$$
a_{m+1}(t; \hat{\beta}, \hat{a}, \hat{b}) = \frac{1}{a} A_t \varphi(A_t), \qquad a_{m+2}(t; \hat{\beta}, \hat{a}, \hat{b}) = \frac{1}{2b} B_t \varphi(A_t),
$$

with

$$
A_t = \frac{1}{a} \left\{ \left(\frac{te^{-\beta^T x^{(0)}}}{b} \right)^{\frac{1}{2}} - \left(\frac{b}{te^{-\beta^T x^{(0)}}} \right)^{\frac{1}{2}} \right\},
$$

$$
B_t = \frac{1}{a} \left\{ \left(\frac{te^{-\beta^T x^{(0)}}}{b} \right)^{\frac{1}{2}} + \left(\frac{b}{te^{-\beta^T x^{(0)}}} \right)^{\frac{1}{2}} \right\},
$$

and $I^{kj}(\hat{\beta}, \hat{a}, \hat{b})$ are the elements of the matrix $I^{-1}(\hat{\beta}, \hat{a}, \hat{b})$. The Fisher information matrix $I^{kj}(\beta, a, b)$ is estimated by

$I(\hat{\beta}, \hat{a}, \hat{b}) = -\frac{\partial^2 \ell(\hat{\beta}, \hat{a}, \hat{b})}{\partial \theta_i \partial \theta_j}, \qquad \theta = (\beta_0, \beta_1, \cdots, \beta_m, a, b)^T.$

6 Goodness-of-fit Test for Parametric AFT Models

Due to the complexity of estimation in parametric models, for a long time people used non-parametric and semiparametric approch to calculate the reliability measures. But now with a lot of development in computer softwares, it is more convenient to use the parametric models in reliabilty studies. Several parametric models have successfully served as population models for failure times arising from a wide range of failure mechanisms. Exponential, Weibull, and lognormal are mostly used due to their presence in many softwares. The estimation of parameters and other reliability measure from some complicated model is a difficult task and especially with covariates in the presence of censoring.

A modified chi-squared type test for parametric AFT model (RRN) based on the Pearson statistic is given in this section proposed by Bagdonavicius and Nikulin (2011). They proposed the goodness of fit test for different parametric accelerate failure time models (AFT). Here goodness-of fit test is given for concrete Birnbaum-Saunders AFT model. Elements of the quadratic form of the test are calculated. ML estimation method is used for the parameter estimation due to its flexible properties. Random grouping intervals is considered. We estimated all the elements of the test and develop a program on R-software.

6.1 Hypothesis, Data and Test construction

Let consider the hypothesis of parametric AFT model on E as in 4.1 :

$$H_0 : S(t|z) = S_0(\int_0^t e^{\beta z(u)} du; \gamma),$$

where $z(t) = (1, z_1(t), ..., z_m(t))^T$ is a vector of possibly time dependent covariates, $\beta = (\beta_0, ..., \beta_m)^T$ is a vector of unknown regression parameters, the function S_0 does not depend on z_i and belongs to a specified class of survival functions :

$$S_0(t, \gamma), \quad \gamma = (\gamma_1, ..., \gamma_q)^T \in G \subset R^q, \quad \theta = (\beta^T, \gamma^T)^T.$$

If explanatory variables are constant over time then the parametric AFT model has the form

$$S(t|z) = S_0(e^{\beta z} t; \gamma).$$

The logarithm of the failure times T under z may be written as

$$\ln T = \beta^T z + \varepsilon, \quad z \in E_1,$$

where the survival function of ε does not depend on z and is $S(t) = S_0(\ln(t))$. If ε is normally distributed then the AFT model is the standard multiple linear regression model. For AFT model suppose we have following censored sample

$$(X_1, \delta_1, z_1), \ldots, (X_n, \delta_n, z_n),$$

where

$$X_i = T_i \wedge C_i, \quad \delta_i = \mathbf{1}_{\{T_i \leq C_i\}}, \quad z_i(t) = (1, z_{i1}(t), ..., z_{im}(t))^T,$$

where T_i is the failure time, C_i the censoring time, and z_i is the vector of covariates.
Set

$$N_i(t) = \mathbf{1}_{\{X_i \leq t, \delta_i = 1\}},\ Y_i(t) = \mathbf{1}_{\{X_i \geq t\}}, \quad N(t) = \sum_{i=1}^{n} N_i(t),\ Y(t) = \sum_{i=1}^{n} Y_i(t).$$

Suppose that the processes N_i, Y_i, z_i are observed at finite time τ and censoring is non-informative. The compensators of the counting processes N_i with respect to the history of the observed processes are $\int_0^t Y_i \lambda_i du$.
Let divide the interval into $k > m+q+1 = s$ classes such that $I_j = (a_{j-1}, a_j], \quad a_0 = 0, \quad a_k = \tau$, and let denote by

$$U_j = N(a_j) - N(a_{j-1}) = \sum_{i: X_i \in I_j} \delta_i,$$

the number of observed failures in the jth interval, $j = 1, 2, ..., k$. To estimate the "expected" number of failures in the interval I_j under the hypothesis H_0, consider the equality

$$\mathbf{E} N_i(t) = \mathbf{E} \int_0^t \lambda_i(u, \theta) Y_i(u) du,$$

we can "expect" to observe

$$e_j = \sum_{i=1}^{n} \int_{I_j} \lambda_i(u, \hat{\theta}) Y_i(u) du \tag{3.16}$$

failures ; here $\theta = (\beta^T, \gamma^T)^T$ and $\lambda_i(t, \theta)$ is the hazard function of T_i under z_i. and $\hat{\theta}$ is the ML estimator of the parameter θ under H_0.
So the test statistic is based on the vector

$$\boldsymbol{Z} = (Z_1, \cdots, Z_k)^T, \qquad Z_j = \frac{1}{\sqrt{n}}(U_j - e_j),\ j = 1, \cdots, k. \tag{3.17}$$

Following we consider the asymptotic properties of $\boldsymbol{Z}$.

6.2 Asymptotic Distribution of Z

Properties of the statistic Z can be investigated through the stochastic process

$$H_n(t) = \frac{1}{\sqrt{n}}\left(N(t) - \sum_{i=1}^{n}\int_0^t \lambda_i(u,\hat{\theta})Y_i(u)du\right).$$

The properties of ML estimators are summed up in conditions 1 as :

Conditions 1

$$\hat{\theta} \xrightarrow{P} \theta_0; \quad \frac{1}{\sqrt{n}}\dot{\ell}(\theta_0) \xrightarrow{d} N_m(0,\, i(\theta_0)); \quad \frac{-1}{n}\ddot{\ell}(\theta_0) \xrightarrow{P} i(\theta_0);$$

and

$$\sqrt{n}(\hat{\theta} - \theta_0) = i^{-1}(\theta_0)\frac{1}{\sqrt{n}}\dot{\ell}(\theta_0) + O_P(1),$$

where

$$\dot{\ell}(\theta) = \sum_{i=1}^{n}\int_0^{+\infty}\frac{\partial}{\partial\theta}\ln\left(\lambda_{x^{(i)}(\cdot)}(u,\theta)\right)dM_i(u)$$

and $i(\theta_0) = \lim_{n\to\infty} I(\theta_0)/n$. *Notice that here* $\theta = (\beta^T, \gamma^T)^T$.

Set

$$S^{(0)}(t,\theta) = \sum_{i=1}^{n} Y_i(t)\lambda_i(t,\theta), \quad S^{(1)}(t,\theta) = \sum_{i=1}^{n} Y_i(t)\frac{\partial \ln\lambda_i(t,\theta)}{\partial\theta}\lambda_i(t,\theta),$$

$$S^{(2)}(t,\theta) = \sum_{i=1}^{n} Y_i(t)\frac{\partial^2 \ln\lambda_i(t,\theta)}{\partial\theta^2}\lambda_i(t,\theta).$$

Conditions 2 *There exist a neighborhood* Θ *of* θ_0 *and continuous bounded on* $\Theta\times[0,\tau]$ *functions*

$$s^{(0)}(t,\theta), \quad s^{(1)}(t,\theta) = \frac{\partial s^{(0)}(t,\theta)}{\partial\theta}, \quad s^{(2)}(t,\theta) = \frac{\partial^2 s^{(0)}(t,\theta)}{\partial\theta^2},$$

such that for j=0,1,2

$$\sup_{t\in[0,\tau],\theta\in\Theta} \|\frac{1}{n}S^{(j)}(t,\theta) - s^{(j)}(t,\theta)\| \xrightarrow{P} 0 \text{ as } n \to \infty.$$

The conditions 2 imply that uniformly for $t \in [0,\, \tau]$

$$\frac{1}{n}\sum_{i=1}^{n}\int_0^t Y_i(u)\lambda_i(u,\,\theta_o)du \xrightarrow{P} A(t)\,; \quad \frac{1}{n}\sum_{i=1}^{n}\int_0^t Y_i(u)\frac{\partial\lambda_i(u,\,\theta_o)}{\partial\theta}du \xrightarrow{P} C(t),$$

where A and C are finite functions.

Lemma 2 *Under conditions 1 and 2 the following convergence holds :*

$$H_n \xrightarrow{d} \boldsymbol{V} \quad on \quad D[0,\, \tau]$$

where $D[0,\, \tau]$ *is space of cadlag functions with Skorokhod metric,* $\boldsymbol{V}$ *is zero mean Gaussian martingale such that* $\forall\ 0 \le u \le v \le T$

$$cov(V(u), V(v)) = A(u) - C^T(u) i^{-1}(\theta_o) C(v).$$

For $i = 0, \cdots, m+q\,;\ j, j' = 1, \cdots, k$, set :

$$V_j = V(a_j) - V(a_{j-1}); \quad \sigma_{jj'} = cov(V_j, V_{j'}),$$

$$A_j = A(a_j) - A(a_{j-1});\ C_{ij} = C_i(a_j) - C_i(a_{j-1}); \quad C_j = (C_{0j},\, C_{1j},\, \cdots,\, C_{m+q,\,j})^T,$$

$$\Sigma = [\sigma_{jj'}]_{k\times k}; \quad C = [C_{ij}]_{s\times k},$$

and A is a $k \times k$ diagonal matrix with diagonal elements $A_1, A_2, \cdots, A_k$.

Theorem 6.1 *Under conditions 1 and 2*

$$Z \xrightarrow{d} Y \sim N_k(0, \Sigma) \quad as\ n \to \infty,$$

where

$$\Sigma = A - C^T i^{-1}(\theta_0) C.$$

Remark : Set the matrix

$$G = i - CA^{-1}C^T.$$

If G is non-degenerate then the generalized inverse of the matrix V is

$$\Sigma^- = A^{-1} + A^{-1}C^T G^{-1} C A^{-1}.$$

We need to inverse only diagonal $k \times k$ matrix A and $s \times s$ matrix G.

Theorem 6.2 *Under the conditions 1 and 2 the following estimators are consistent* A_j*,* C_j*,* $i(\theta_o)$ *and* Σ *:*

$$\hat{A}_j = \frac{U_j}{n}; \quad \hat{C}_j = \frac{1}{n}\sum_{i=1}^{n}\int_{I_j} \frac{\partial}{\partial\theta} \ln \lambda_i(u,\, \hat{\theta}) dN_i(u),$$

$$\hat{i} = \frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau} \frac{\partial \ln \lambda_i(u,\, \hat{\theta})}{\partial\theta} \left(\frac{\partial \ln \lambda_i(u,\, \hat{\theta})}{\partial\theta}\right)^T dN_i(u),$$

$$\hat{\Sigma} = \hat{A} - \hat{C}^T i^{-1} \hat{C}.$$

Note : The proofs of the theorems and lemma are given recently in Bagdonavicius and Nikulin (2011). Using these results we shall express the test statistic of chi-squared type for testing H_0.

6.3 Test statistic

The test statistic is according to the idea of chi-squared test. From theorems 6.1 and 6.2 the test for hypothesis H_0 can be based on the statistic

$$Y_n^2 = Z^T \hat{\Sigma}^- Z$$

where

$$\hat{\Sigma}^- = \hat{A}^{-1} + \hat{A}^{-1}\hat{C}^T\hat{G}^-\hat{C}\hat{A}^{-1}, \quad \hat{G} = \hat{i} - \hat{C}\hat{A}^{-1}\hat{C}^T$$

The test statistic can be written in the following form

$$Y_n^2 = \sum_{j=1}^{k} \frac{(U_j - e_j)^2}{U_j} + Q,$$

where

$$Q = W^T\hat{G}^-W, \quad W = \hat{C}\hat{A}^{-1}Z = (W_0, W_1, ..., W_{m+q})^T, \quad \hat{G} = [\hat{g}_{ll'}]_{s\times s},$$

$$\hat{g}_{ll'} = \hat{i}_{ll'} - \sum_{j=1}^{k} \hat{C}_{lj}\hat{C}_{l'j}\hat{A}_j^{-1}, \quad W_l = \sum_{j=1}^{k} \hat{C}_{lj}\hat{A}_j^{-1}Z_j,$$

$$\hat{C}_{lj} = \frac{1}{n}\sum_{i:X_i\in I_j} \delta_i \frac{\partial}{\partial\theta} \ln\lambda(X_i, \hat{\theta}), \quad \hat{A}_j = U_j/n, \quad Z_j = \frac{1}{\sqrt{n}}(U_j - e_j),$$

$$\hat{i}_{ll'} = \frac{1}{n}\sum_{i=1}^{n} \delta_i \frac{\partial \ln\lambda(X_i;\hat{\theta})}{\partial\theta_l}\frac{\partial \ln\lambda(X_i;\hat{\theta})}{\partial\theta_{l'}}, \qquad U_j = \sum_{i:X_i\in I_j} \delta_i, \tag{3.18}$$

$i = 1, \cdots, n, \quad j = 1, \cdots, k, \quad l, l' = 0, 1, \cdots, m + q.$

The limit distribution of the statistic Y_n^2 is chi-square with

$$r = rank(\Sigma^-) = Tr(\Sigma^-\Sigma)$$

degrees of freedom. If G is non-degenerate then $r = k$.

Statistical inference for the hypothesis H_0 : The hypothesis is rejected with approximate significance level α if $Y_n^2 > \chi^2_\alpha(r)$.

This is a good success in the application of chi-square type tests to solve the complicated problems. But the problem to develop a software is still there because for testing esch model one has to make a computer program for all the elements of the test.

6.4 Choice of random grouping intervals

It is recommended to take intervals as random data functions. The idea is to divide the interval $[0, t]$ into k intervals with equal expected numbers of failures (which are not necessary integers) to avoid small number or no failures in several first and last intervals.

Define

$$E_k = \sum_{i=1}^{n} \int_0^{\tau} \lambda_i(u, \hat{\theta}) Y_i(u) du = \sum_{i=1}^{n} \Lambda_i(X_i, \hat{\theta}), \quad E_j = \frac{j}{k} E_k, \quad j = 1, \cdots, k.$$

$\hat{a}_j$ verify the following equalities to have equal number of expected failure in all intervals

$$g(\hat{a}_j) = E_j, \quad g(a) = \sum_{i=1}^{n} \int_0^{a} \lambda_i(u, \hat{\theta}) Y_i(u) du.$$

Denote by the $X_{(1)} \leq \cdots \leq X_{(n)}$ the ordered sample from $X_i, \cdots, X_n$. The function

$$\begin{aligned} g(a) &= \sum_{i=1}^{n} \Lambda_i(X_i \wedge a, \hat{\theta}) \\ &= \sum_{i=1}^{n} \left[\sum_{l=i}^{n} \Lambda_{(l)}(a, \hat{\theta}) + \sum_{l=1}^{i-1} \Lambda_{(l)}(X_{(l)}, \hat{\theta}) \right] 1_{[X_{(i-1)}, X_{(i)}]}(a) \end{aligned}$$

is continuous and increasing on $[0, \tau]$; here $X_{(0)} = 0$, $\sum_{l=1}^{0} c_l = 0$. Set

$$b_i = \sum_{l=i+1}^{n} \Lambda_{(l)}(X_{(i)}, \hat{\theta}) + \sum_{l=1}^{i} \Lambda_{(l)}(X_{(l)}, \hat{\theta}).$$

If $E_j \in [b_{i-1}, b_i]$ then $\hat{a}_j$ is the unique solution of the equation

$$\sum_{l=i}^{n} \Lambda_{(l)}(\hat{a}_j, \hat{\theta}) + \sum_{l=1}^{i-1} \Lambda_{(l)}(X_{(l)}, \hat{\theta}) = E_j \tag{3.19}$$

We have $0 < \hat{a}_1 < \hat{a}_2, \cdots, \hat{a}_k = \tau$. With this choice of intervals

$$ej = E_k / k$$

for any j.
Now we consider an example.

6.5 Application of the test for BS Distribution

The parameter $\theta = (\beta^T, a, b)^T$ can be estimated by maximizing the loglikelihood function (3.13). The hazard function for AFT models with constant stress can be written as

$$\lambda_i(t,\theta) = e^{-\beta^T x^{(i)}} \lambda_0 \left(e^{-\beta^T x^{(i)}} t, \theta \right).$$

We can write the hazard function for AFT model with BS distribution as

$$\lambda_i(t,\theta) = e^{-\beta^T x^{(i)}} \left[\frac{\frac{1}{2\sqrt{2\pi}ab} \left\{ \left(\frac{b}{K_i}\right)^{1/2} + \left(\frac{b}{K_i}\right)^{3/2} \right\} \exp\left\{ -\frac{1}{2a^2} \left(\frac{K_i}{b} + \frac{b}{K_i} - 2 \right) \right\}}{1 - \Phi\left\{ \frac{1}{a} \left[\left(\frac{K_i}{b}\right)^{1/2} - \left(\frac{b}{K_i}\right)^{1/2} \right] \right\}} \right],$$

here $\theta = (\beta^T, a, b)^T$ and $K_i = X_i e^{-\beta^T x^{(i)}}$.
The parameters can be estimated by maximizing the the following loglikelihood function

$$\begin{aligned} \ell(\beta, a, b) \quad = \quad & \sum_{i=1}^{n} \delta_i \beta^T x_{(i)} + \sum_{i=1}^{n} \delta_i \left[-\ln a - \ln b + \ln \left\{ \left(\frac{b}{K_i} \right)^{\frac{1}{2}} + \left(\frac{b}{K_i} \right)^{\frac{3}{2}} \right\} \right. \\ & \left. - \frac{1}{2a^2} \left(\frac{K_i}{b} + \frac{b}{K_i} - 2 \right) \right] + \\ & \sum_{i=1}^{n} (1-\delta_i) \ln \left(1 - \Phi \left[\frac{1}{a} \left\{ \left(\frac{K_i}{b} \right)^{\frac{1}{2}} - \left(\frac{b}{K_i} \right)^{\frac{1}{2}} \right\} \right] \right). \end{aligned} \tag{3.20}$$

The expression for the elements of the matrix $\hat{i} = [\hat{i}_{ls}]_{(m+3)\times(m+3)}$ and $\hat{C}_{lj}$ elements for the test are calculated by using the formula (3.18) and given as

$$\begin{aligned} \hat{i}_{ls} \quad = \quad & \frac{1}{n} \sum_{i=1}^{n} \delta_i \left[x_{il} + \frac{x_{il}}{2} \left\{ \frac{1 + 3\left(\frac{b}{K_i}\right)}{1 + \left(\frac{b}{K_i}\right)} + A_i B_i \right\} - \frac{1}{2} \left(\frac{x_{il} B_i \varphi(A_i)}{1 - \Phi(A_i)} \right) \right] \times \\ & \left[x_{is} + \frac{x_{is}}{2} \left\{ \frac{1 + 3\left(\frac{b}{K_i}\right)}{1 + \left(\frac{b}{K_i}\right)} + A_i B_i \right\} - \frac{1}{2} \left(\frac{x_{is} B_i \varphi(A_i)}{1 - \Phi(A_i)} \right) \right], \quad l, s = 0, 1, \cdots, m, \end{aligned}$$

where $A_i = \frac{1}{a} \left\{ \left(\frac{K_i}{b}\right)^{\frac{1}{2}} - \left(\frac{b}{K_i}\right)^{\frac{1}{2}} \right\}$, $\quad B_i = \frac{1}{a} \left\{ \left(\frac{K_i}{b}\right)^{\frac{1}{2}} + \left(\frac{b}{K_i}\right)^{\frac{1}{2}} \right\}$,

$$\begin{aligned} \hat{i}_{l,m+1} \quad = \quad & \frac{1}{n} \sum_{i=1}^{n} \delta_i \left[x_{il} + \frac{x_{il}}{2} \left\{ \frac{1 + 3\left(\frac{b}{K_i}\right)}{1 + \left(\frac{b}{K_i}\right)} + A_i B_i \right\} - \frac{1}{2} \left(\frac{x_{il} B_i \varphi(A_i)}{1 - \Phi(A_i)} \right) \right] \times \\ & \left[\frac{1}{a} \left(A_i^2 - 1 \right) - \frac{1}{a} \left(\frac{A_i \varphi(A_i)}{1 - \Phi(A_i)} \right) \right], \end{aligned}$$

$$\hat{i}_{l,m+2} = \frac{1}{n}\sum_{i=1}^{n}\delta_i\left[x_{il}+\frac{x_{il}}{2}\left\{\frac{1+3\left(\frac{b}{K_i}\right)}{1+\left(\frac{b}{K_i}\right)}+A_iB_i\right\}-\frac{1}{2}\left(\frac{x_{il}B_i\varphi(A_i)}{1-\Phi(A_i)}\right)\right]\times \left[\frac{1}{b}\left(-1+\frac{1}{2}\left(\frac{1+3\frac{b}{K_i}}{1+\frac{b}{K_i}}\right)+\frac{1}{2}A_iB_i\right)-\frac{1}{2b}\left(\frac{B_i\varphi(A_i)}{1-\Phi(A_i)}\right)\right],$$

$$\hat{i}_{m+1,m+1} = \frac{1}{n}\sum_{i=1}^{n}\delta_i\left[\frac{1}{a}\left(A_i^2-1\right)-\frac{1}{a}\left(\frac{A_i\varphi(A_i)}{1-\Phi(A_i)}\right)\right]^2,$$

$$\hat{i}_{m+1,m+2} = \frac{1}{n}\sum_{i=1}^{n}\delta_i\left[\frac{1}{a}\left(A_i^2-1\right)-\frac{1}{a}\left(\frac{A_i\varphi(A_i)}{1-\Phi(A_i)}\right)\right]\times \left[\frac{1}{b}\left(-1+\frac{1}{2}\left(\frac{1+3\frac{b}{K_i}}{1+\frac{b}{K_i}}\right)+\frac{1}{2}A_iB_i\right)-\frac{1}{2b}\left(\frac{B_i\varphi(A_i)}{1-\Phi(A_i)}\right)\right],$$

$$\hat{i}_{m+2,m+2} = \frac{1}{n}\sum_{i=1}^{n}\delta_i\left[\frac{1}{b}\left(-1+\frac{1}{2}\left(\frac{1+3\frac{b}{K_i}}{1+\frac{b}{K_i}}\right)+\frac{1}{2}A_iB_i\right)-\frac{1}{2b}\left(\frac{B_i\varphi(A_i)}{1-\Phi(A_i)}\right)\right]^2,$$

and

$$\hat{C}_{lj} = \frac{1}{n}\sum_{i:X_i\in I_j}\delta_i\left[x_{il}+\frac{x_{il}}{2}\left\{\frac{1+3\left(\frac{b}{K_i}\right)}{1+\left(\frac{b}{K_i}\right)}+A_iB_i\right\}-\frac{1}{2}\left(\frac{x_{il}B_i\varphi(A_i)}{1-\Phi(A_i)}\right)\right],$$

$$\hat{C}_{m+1,j} = \frac{1}{n}\sum_{i:X_i\in I_j}\delta_i\left[\frac{1}{a}\left(A_i^2-1\right)-\frac{1}{a}\left(\frac{A_i\varphi(A_i)}{1-\Phi(A_i)}\right)\right],$$

$$\hat{C}_{m+2,j} = \frac{1}{n}\sum_{i:X_i\in I_j}\delta_i\left[\frac{1}{b}\left(-1+\frac{1}{2}\left(\frac{1+3\frac{b}{K_i}}{1+\frac{b}{K_i}}\right)+\frac{1}{2}A_iB_i\right)-\frac{1}{2b}\left(\frac{B_i\varphi(A_i)}{1-\Phi(A_i)}\right)\right].$$

Since the inverse of the cumulative hazard function of the BS distribution can not be written in an explicit form, so we can estimate the end points of the random intervals by numerical method using the equation (3.19).

6.5.1 Example 1

The BPD Data set from chapter 2 of Hosmer and Lemeshow (2008) is used for the goodness of fit test. The data contains $n = 78$ subjects and with four variables (id, surfact,ondays,censor). Surfact is a covariate in the the application of the test statistic. The data is given in Appendix 3.1. The results of the test are given below :

The values of ML estimators are :

$$\hat{\beta}_0 = 4.9969, \quad \hat{\beta}_1 = -0.2001, \quad \hat{a}_{bs} = 1.2661, \quad \hat{b}_{bs} = 0.8108.$$

With $k = 8$, the random intervals and other estimates are given in the following table.

j	1	2	3	4	5	6	7	8
$\hat{a}_j$	27.16	43.98	64.92	97.80	149.18	220.85	346.79	733.00
U_j	6	5	15	14	8	6	8	11
e_j	9.30	9.30	9.30	9.30	9.30	9.30	9.30	9.30
C_{1j}	-0.04833	0.0689	0.2508	0.2430	0.1400	0.1018	0.1305	0.1696
C_{2j}	0.0125	0.0438	0.1521	0.0873	0.0873	0.0504	0.0816	0.0155
C_{3j}	0.2354	0.0124	-0.0700	-0.1046	-0.0851	-0.0787	-0.1177	-0.1790
C_{4j}	-0.2493	-0.0732	-0.1650	-0.1430	-0.0803	-0.0642	-0.0921	-0.1386

The Fisher's information matrix is

$$\hat{\boldsymbol{i}}_{4\times 4} = \begin{pmatrix} 1.7690446 & 0.6868031 & -1.4546406 & -0.4236595 \\ 0.6868031 & 0.6868031 & -0.3479777 & -0.4618151 \\ -1.4546406 & -0.3479777 & 2.1312279 & -0.8386943 \\ -0.4236595 & -0.4618151 & -0.8386943 & 1.9582996 \end{pmatrix}.$$

We calculate

$$W = (W_1, W_2, W_3, W_4)^T = (0.6170, -0.0933, -1.3528, 1.1656)^T$$

Notice that in the case of BS-distribution the matrix $\hat{G}$ is found to be degenerated. So the rank of the matrix Σ is $k-1$, and we have

$$X_n^2 = 11.7744, \quad Q = 3.2343, \quad Y_n^2 = 15.0087, \quad pv = 0.0359.$$

So from the above result we can say that the hypothesis is rejected.

6.5.2 Example 2

This data is taken from Lawless (2003) which gives failure times of 76 electric insulating fluid tested at voltage ranging from 26 to 38 kilovolts (kV). This experiment was run long enough to observe the failures of all items. The data is given in Appendix 3.2. The main purpose of the study was to investigate the distribution of time to breakdown for the insulating fluid and to relate this to the voltage level.

Lawless (2003) suggest Weibull AFT-power rule model. Lawless used parametric methods based on loglikelihood ratios to verify separately Weibull, AFT and power rule assumptions of the

model. All tests did not contradict the Weibull AFT-power rule model. Bagdonavicius et al. (2012) used this data for exponential-AFT and Weibull-AFT models. The data rejected for the exponential AFT but did not reject for the Weibull AFT model. Let us use this data for the BS-AFT model.
The values of ML estimators are :

$$\hat{\beta}_0 = 53.7138, \quad \hat{\beta}_1 = -13.8958, \quad \hat{a}_{bs} = 2.0726, \quad \hat{b}_{bs} = 0.0377.$$

We choose $k = 8$, the elements of the test statistic are calculated as below.

j	1	2	3	4	5	6	7	8
$\hat{a}_j$	0.532	1.450	3.417	8.103	18.289	40.967	116.415	2323.700
U_j	7	12	10	12	7	9	10	9
e_j	10.51	10.51	10.51	10.51	10.51	10.51	10.51	10.51

The Fisher information matrix is

$$\hat{\boldsymbol{i}}_{4\times 4} = \begin{pmatrix} 2.0879419 & 7.316385 & -0.7999923 & -15.123080 \\ 7.3163849 & 25.656078 & -2.7819900 & -53.174564 \\ -0.7999923 & -2.781990 & 0.6746561 & -3.649846 \\ -15.1230799 & -53.174564 & -3.6498462 & 544.176774 \end{pmatrix}.$$

Notice that for this data the matrix $\hat{G}$ is also found to be degenerated. So the rank of the matrix Σ is $k-1$, and we have

$$X_n^2 = 4.4633, \quad Q = 37.7796, \quad Y_n^2 = 42.2429, \quad pv = 4.6688^{-7}.$$

The hypothesis of the BS-AFT model for this data is rejected. One can see that without quadratic form the hypothesis can not be rejected.

Chapter 4

Analysis of Redundant Systems

1 A Historical Review

Redundancy is one of the common ways to increase the systems' realisability where extra units are attached in the system with the main functioning unit. The extra units or components are called standby units or redundant units. Standby redundancy can be reparable or unreparable. In our work we study the system with unreparable redundancy. Before going to further analysis of the redundant systems it is important to specify the operating state or type of standby units. In literature three states of the standby units are mentioned which are :

– Hot or active redundancy : the standby units are subject to the the same operating rule as the main unit, they can fail before the main one that is equal failing probabilities,
– Cold or inactive redundancy : the standby units are placed in off mode (idol), they cannot fail until take the function of main one,
– Warm or partially operating redundancy : the standby units are placed in partially operating state, can fail before replacing the main unit but with smaller probability than the main one.

Suppose that a system compose of one main unit and $n-1$ standby units and let denote by

$$T_1, T_2, \cdots, T_n$$

the random functioning time of the main and standby units. Then for hot standby system the total operation time can be written as

$$\tau_1 = \max(T_1, T_2, \cdots, T_n),$$

and for the cold standby system

$$\tau_2 = T_1 + T_2 + \cdots + T_n.$$

Its is obvious that

$$\tau_1 \leq \tau_2.$$

On the base of this result we can conclude that a cold redundancy is more suitable than the hot redundancy subject to the condition that the switching does not cause any damage. Moreover people in the field of reliability used simple methods for this purpose. For example Gnedenko, Belyaev and Solovyev (1968), and Kozlov and Ushakov (1970) give the simple formula to calculate the reliability function $R(t)$ or unreliability function (distribution function) $F(t) = 1 - R(t)$. For hot standby system when the failure are all independent of each other, the distribution function can be written as

$$F_n^{(1)}(t) = q_1(t)q_n(t)\cdots q_n(t) = \prod_{i=1}^{n}(1 - p_i(t)),$$

where $p_i(t), i = 1, \cdots, n$ is the reliability or failure-free operation of the ith unit and $q_i(t)$ is its unreliability, and for cold standby system the distribution function is

$$F_n^{(2)}(t) \approx q_1(t)q_n(t)\cdots q_n(t)/n!.$$

From the ratio

$$\frac{F_n^{(1)}(t)}{F_n^{(2)}(t)} \approx n!,$$

it is clear that when we shift to a cold standby redundancy, the distribution function decreases to $1/n!$ its preceding value.

In practice it is not convenient to apply hot standby because it does not improve the reliability upto desired level due to high failure probability of standby units. Also it is not suitable to use the cold standby due to switching time which can cause the interruption in functioning of the system. In such situation it is better to use the warm or partially operating standby units.

Suppose that we have one main unit and $m - 1$ standby units. Let denote by $p_k^{(n)}(t)$ the reliability of the mth unit in warm stat and denote by $p_k^{(0)}(t^*, t)$ the conditional reliability of the mth unit operating in hot state in the interval (t^*, t) under the assumption that it has not failed in the interval $(0, t^*)$ when it was in nonoperating state. t^* is the moment when the unit is absorbed into the system. We denote $t_1^*, t_2^*, \cdots, t_k^*$ the random operating times of the basic and standby units. Let $F_1(t)$ and $F_m(t)$ are the distribution functions of the basic unit and the whole standby system with the main and $(m-1)$st standby unit. And $F_n(t)$ is the distribution function or of the standby group. The life length of the standby unit depends on the preceding times say t_i^*. We can write the distribution or unreliability function for two consecutive functions $F_k(t)$ and $F_{k+1}(t)$.

$$
\begin{aligned}
F_k(t) = \mathbf{P}\{T_k < t\} &= \mathbf{P}\{T_{k-1} < t, T_k < t\} = \int_0^t \mathbf{P}\{t^* < T_{k-1} < t^* + dt^*, t_k^* < t\} \\
&= \int_0^t \mathbf{P}\{t_k^* < t | T_{k-1} = t^*\} dF_{k-1}t^* = \int_0^t [1 - \mathbf{P}\{t_k^* > t | T_{k-1} = t^*\}] dF_{k-1}t^* \\
&= \int_0^t [1 - p_k^{(n)}(t^*) p_k^{(0)}(t^*, t)] dF_{k-1}t^* \qquad (4.1)
\end{aligned}
$$

This recursion formula presented by Gnedenko, Belyaev and Solovyev (1968) to find the unreliability function of the standby system. The function $p_k^{(n)}$ is unknown and can be parameterized by some parametric family of distributions. The same approach was used by some other author like Kozlov and Ushakov (1970). It is very common to see application of exponential and Weibull distribution in reliability studies due to their simple forms.

The researcher used the different mathematical and statistical techniques to improve the systems reliability but without using accelerated stress. Due to the excessive use of accelerated life testing in industry and the use of resources by the warm standby units till the failure of main unit which requires a transfer functional, opens a new horizon of research for reliability analysts. Also the censoring mechanism make the things more complicated. Recently Bagdonavicius, Masiulaityle, and Nikulin (2008a, 2008b, 2009, 2010) study all these phenomenons in a series of paper on the statistical analysis of redundant systems. They calculated the reliability characteristics in terms of distribution function or unreliability functions. They generalized their approach to m standby units and also in case of censoring. In this work we extend their work with some modifications and apply the techniques for Birnbaum-Saunders family of distributions which is well studied model in fatigue failure data in industry.

2 Introduction

Recently the series of papers on the redundant system was published by Bagdonavicius, Masiulaityle, and Nikulin (2008a, 2008b, 2009, 2010) where they studies the reliability of the system with exponential, Weibull and loglogistic models. Here we give the reliability of redundant system based on parametric Birnbaum-Saunders family of distributions and we develop a software for the redundant system. Redundancy which means the duplication of critical components of a system is a common and useful approach to increase the reliability of the system. The redundant systems contain more than one subcomponents and all must fail before the system fails. The functioning component or unit is called the operating component or main unit and the other subcomponents are called stand-by units. If the main unit fails then the first stand-by unit is commuted automatically and if this unit fails the second stand-by unit is commuted and so on. So the redundant system fails only if all units fails. A system with two, three or many replications of each element is respectively termed as dual modular redundant, triple modular redundant, and multi-modular redundant. Redundant system with one main unit and $m-1$ stand-by units is denoted by $S(1, m-1)$, $m \geq 2$. Redundant system is different from the backup system in a way that with redundant system data is continuously passed from the primary to secondary component while backup systems may loos data and may take many hours to become operational. Redundancy increases the cost and complexity of the system design and sometimes it is the unique way to provide high reliability to modern electrical, biotechnical and mechanical systems. Many low-risk industries do not need redundancy in order to be successful. However, in high-risk industries such as aerospace and nuclear, where the cost of failure is high enough, the redundancy becomes essential to have the high reliability.

Here we consider a system $S(1, 1)$ with one main unit operating in hot conditions and one stand-by unit operating in warm conditions. In the terminology of accelerated life testing main unit is working under accelerated stress with respect to stand by unit. We suppose that the switching on from the warm to hot state does not do any damage to the unit and the switching time is stochastic. The model with fluent switching, important for practice, is considered. On the base of this supposition Bagdonavicius et al. (2008b) proposed the test for *general fluent switching hypothesis* H_0 and for *particular fluent switching hypothesis* H_0^*. The hypothesis H_0 is formulated exactly using *Sedyakin's reliability principle* (Sedyakin, 1966) and the hypothesis H_0^* is formulated using AFT model for the reliability of redundant systems. Sedyakin's model explains that after the failure of main unit the stand-by unit follow the same reliability curve as that of the main unit. Following precedent results of Bagdonavicius et al. (2008b) we study the asymptotic properties of the test statistics. After testing, we construct parametric estimators of the cumulative distribution function $K_2(t)$ of redundant system, using censored reliability data of components under different stresses. Then confidence intervals for the commutative distribution function of redundant system $S(1, 1)$ are constructed under the assumption that the distribution of failure times of both units follow the Birnbaum-Saunders family.

3 Redundant System With Warm Stand-By Unit

Redundancy is mostly used technique to increase the reliability of the technical systems. We consider the redundant system with one main and one stand-by unit in parallel and we denote the system by $S(1, 1)$. For configuring the redundant system the functioning state of the stand-by unit is very important according to the criticality of the system. In the last century several people study the system with hot and cold stand-by units but recently there are many papers where the intermediate warm conditions are used for stand-by units. If the stand-by unit is working in hot conditions as the main unit, the failing probability of stand-by unit is equal to the main unit i.e. 0.5, because both are functioning under the same stress. And if the stand-by unit is in cold state, it requires commuting time to come in hot conditions after the failure of main unit which can interrupt the system's continuity and also can cause the increase in failure rate due to the sudden change in stress that is due to the burn-in period (failure in early life). So the warm reserving can be the better choice where the stand-by unit is functioning under lower stress than the main one. In this way the probability of failure of the stand-by unit is smaller than that of the main unit and commuting is fluent but we don't discuss here the problem of the choice of optimality of the warm state which depends on the practical problem. Recently a lot of work has been done for statistical analysis of redundant system by Bagdonavicius, Masiulaityle, and Nikulin (2008a, 2008b, 2009, 2010) and with warm stand-by units. They used many parametric models such as exponential, Weibull, lognormal, loglogistic, generalized Weibull, and inverse Gaussian for determining the reliability of redundant system (see also Nikulin et al. (2011c)). Here we follow their results and use the famous Birnbaum-Saunders (BS) family of distributions as the distribution of failure times (see Nikulin and Tahir, 2011).

4 Accelerated Life Testing (ALT) In Reliability

Failure times data of high reliable units may take many years for reliability analysis. One way of obtaining quick reliability information is to use accelerated life testing (ALT) method, where higher level of experimental factors, stresses or covariates (temperature, voltage or speed) are applied on the units to increase the number of failures in less time (Bagdonavicius and Nikulin (2002), Meeker and Escobar (1998)).

In ALT, it is supposed that a stress (or explanatory variable) is a deterministic time function, may be multidimensional :

$$x(\cdot) = (x_1(\cdot), \dots, x_m(\cdot))^T : [0, \infty[\to R^m, \tag{4.2}$$

which is a vector of covariates itself or a realization of a stochastic process $X(\cdot)$. This process is also called the covariate process, $X(\cdot) = (X_1(\cdot), ..., X_m(\cdot))^T$. We denote E a set of all possible (admissible) stresses, and by E_1 a set of all constant over time covariates, $E_1 \subset E$. By tradition if $x(\cdot)$ is constant in time, we write x instead of $x(\cdot)$.

There are many types of stresses (e.g. step stresses, continuous cyclic stress, cyclic stress of type switch-on-switch-off, degradation stress) but step stresses are commonly used in ALT where the units are placed initially at low stress and if they do not fail till certain moment then stress is increased continuously in steps. Step-stress can be increasing or decreasing. For our problem it is interesting to introduce the next class of step-stresses. We denote $E_2 \subset E$ a set of step-stresses of the form

$$x(t) = x_1 1_{\{0 \le t < t_1\}} + x_2 1_{\{t_1 \le t\}}, \quad x_1, x_2 \in E_1. \tag{4.3}$$

Let $T_{x(\cdot)}$ be the failure time under the stress $x(\cdot)$ and

$$S_{x(\cdot)}(t) = \mathbf{P}\{T_{x(\cdot)} > t\}, \quad F_{x(\cdot)}(t) = \mathbf{P}\{T_{x(\cdot)} \le t\}, \quad f_{x(\cdot)}(t) = -S'_{x(\cdot)}(t),$$

$$\lambda_{x(\cdot)}(t) = \lim_{h \to 0} \frac{1}{h} \mathbf{P}\{T_{x(\cdot)} \in [t, t+h) | T_{x(\cdot)} \ge t\} = -\frac{S'_{x(\cdot)}(t)}{S_{x(\cdot)}(t)},$$

be the survival function, cumulative distribution function, probability density function, and hazard function respectively.

Let $x(\cdot), y(\cdot) \in E$ be the two stresses. A stress $y(\cdot)$ is accelerated with respect to the stress $x(\cdot)$ if $S_{y(\cdot)}(t) \le S_{y(\cdot)}(t)$, or equivalently we can write (also shown in Figure 4.1)

$$F_{y(\cdot)}(t) \ge F_{x(\cdot)}(t), \quad t \ge 0.$$

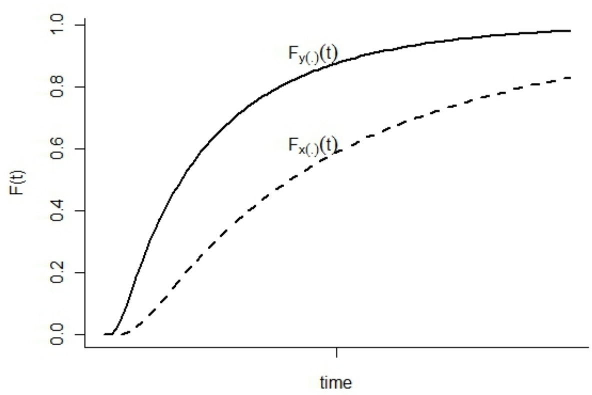

Figure 4.1 – Distribution functions for two levels of stresses where $y(\cdot) > x(\cdot)$.

If the data are censored then we have to consider the influence of $x(\cdot)$ on the distribution of censoring time C, i.e. we write $C = C_{x(\cdot)}$, and we observe

$$X_{x(\cdot)} = \min(T_{x(\cdot)}, C_{x(\cdot)}).$$

5 Sedyakin'S Physical Principle In Reliability

Accelerated stresses are used to reduce the time on test. So a transfer functional is needed to interpolate the accelerated failure times to the failure times under usual stress (Bagdonavicius

(1978), Bagdonavicius and Nikulin (1997)). The physical principle in reliability proposed by N. Sedyakin (1966) states that for two identical populations of units functioning under different stresses $x_1 \neq x_2$, two moments t_1 and t_1^* are equivalent if the probabilities of survival are equal until these moments, i.e.

$$S_{x_1}(t_1^*) = S_{x_2}(t_1), \quad t_1^* = g(t_1), \quad \text{where} \quad g(t) = S_{x_1}^{-1}(S_{x_2}(t)). \tag{4.4}$$

This principle gives an interesting way to prolong any class of survival functions $\{S_x(\cdot), x \in E_1\}$ indexed by constant in time stresses to a class of survival functions indexed by step-stresses, for example from E_2 given by (4.2). Figure 4.2 shows the increasing step-stress. x_1 and x_2 are the

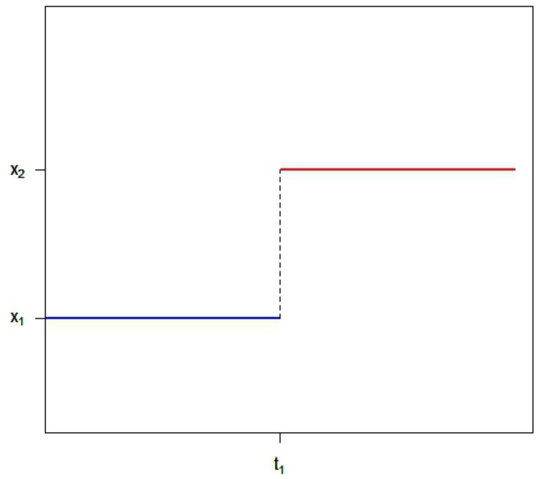

Figure 4.2 – Increasing step-stress for the warm stand-by unit.

stresses corresponding to warm and hot conditions respectively. We suppose that stress x_2 is accelerated with respect to stress x_1. The moment t_1 is random in our case.

According to Sedyakin we may consider the model on E_2 for all $s \geq 0$

$$\lambda_{x(\cdot)}(t_1 + s) = \lambda_{x_2}(t_1^* + s). \tag{4.5}$$

In terms of the survival function $S_{x(\cdot)}(t), x(\cdot) \in E_2$ that satisfies the same rule of time-shift

$$S_{x(\cdot)} = \begin{cases} S_{x_1}(t), & 0 \leq t < t_1, \\ S_{x_2}(t - t_1 + t_1^*), & t \geq t_1, \end{cases} \tag{4.6}$$

where t_1^* is determined by the equation (4.4). The model given by (4.5) and (4.6) is called the Sedyakin model on E_2.

The generalized Sedyakin (GS) model on a set of stresses E can be written by supposing that the hazard rate $\lambda_{x(\cdot)}(t)$ at any moment t is a function of the value of the stress at this moment and of the probability of survival until this moment.

$$\lambda_{x(\cdot)}(t) = h(x(t), S_{x(\cdot)}(t)), \quad x(\cdot) \in E,$$

where h is a positive function. Note that the AFT model verifies this rule.

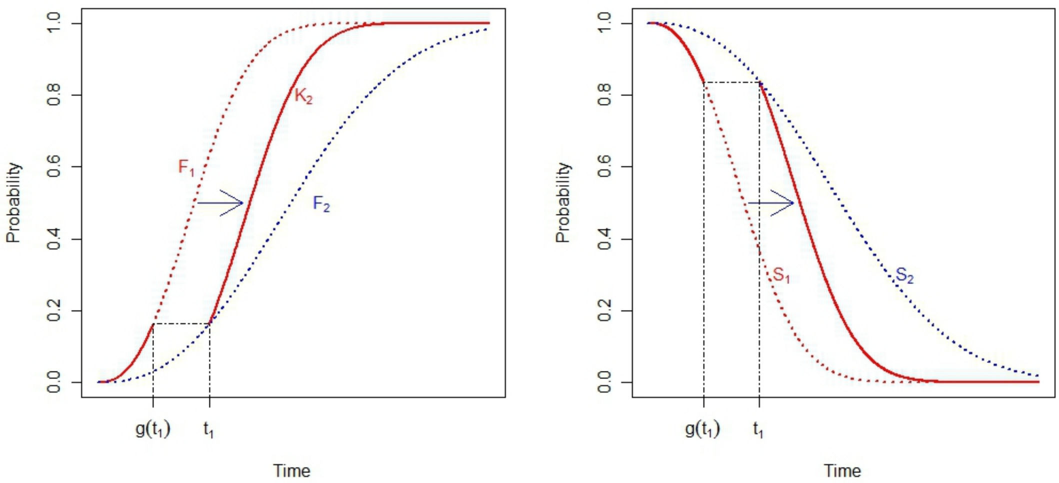

Figure 4.3 – Cumulative distribution function (left) and survival function (right) of the system under Sedyakin's principal.

6 Sedyakin Model And Its Application In Redundant System

Let denote by T_1, F_1 and f_1 the failure time, the cumulative distribution function and the probability density function of the main unit. Suppose the failure time of the stand-by unit be T_2. If it is working in hot condition its distribution function is also F_1. In warm conditions the distribution function of T_2 is F_2 and the p.d.f. is f_2. After the failure of main unit the stand-by unit is switched to hot conditions and its distribution function is different from F_1 and F_2. The system fails if both units fail i.e. the failure time of the system is $T = \max(T_1, T_2)$. Let the conditional density function of T_2 given that the main unit fails at moment y is denoted by

$$f_2^{(y)}(x) = f_{T_2|T_1=y}(x)$$

and let denote by K_2 and k_2 the distribution function and the density function of the system's failure time T. The cumulative distribution function K_2 can be written as

$$\begin{aligned} K_2(t) &= \mathbf{P}(T \le t) = \mathbf{P}(T_1 \le t, T_2 \le t) = \int_0^t \mathbf{P}(T_2 \le t \mid T_1 = y) dF_1(y) \\ &= \int_0^t \left\{ \int_0^y f_2(x)dx + \int_y^t f_2^{(y)}(x)dx \right\} f_1(y)dy. \end{aligned} \tag{4.7}$$

When the stand-by unit is in **cold** state then

$$f_2(x) = 0 \quad \text{if} \quad x \le y, \quad \text{and} \quad f_2^{(y)}(x) = f_1(x-y) \quad \text{if} \quad x > y,$$

so from equation (4.7), it follows that

$$K_2(t) = \int_0^t \left\{ \int_y^t f_1(x-y)dx \right\} f_1(y)dy = \int_0^t F_1(t-y)dF_1(y).$$

When the stand-by unit is in **hot** state then

$$f_2^{(y)}(x) = f_2(x) = f_1(x),$$

so using equation (4.7) we can write

$$\begin{aligned} K_2(t) &= \int_0^t \left\{ \int_0^y f_2(x)dx + \int_y^t f_2(x)dx \right\} f_1(y)dy \\ &= \int_0^t \left\{ \int_0^t f_2(x)dx \right\} dF_1(y) = \int_0^t \left\{ \int_0^t f_1(x)dx \right\} dF_1(y) = [F_1(t)]^2. \end{aligned}$$

When the stand-by unit is in **warm** conditions the following hypothesis is assumed :

$$H_0 : f_2^{(y)}(x) = f_1(x + g(y) - y), \quad \forall \quad x \ge y \ge 0, \tag{4.8}$$

where $g(y)$ is the moment which in hot conditions corresponds to the moment y in warm conditions in the sense that

$$F_1(g(y)) = \mathbf{P}(T_1 \le g(y)) = \mathbf{P}(T_2 \le y) = F_2(y),$$

so

$$g(y) = F_1^{-1}(F_2(y)).$$

Conditionally (given $T_1 = y$) the hypothesis (4.8) corresponds to the Sedyakin's model. In the situation considered here the switch off moments are random. The equation (4.7) implies that under hypothesis H_0 the distribution function of of the system $S(1,1)$ is

$$\begin{aligned} K_2(t) &= \int_0^t \{F_2(y) + \int_y^t f_1(x+g(y)-y)dx\} f_1(y)dy \\ &= \int_0^t \{F_2(y) + F_1(t+g(y)-y)dx - F_1(g(y))\} f_1(y)dy \\ &= \int_0^t F_1(t+g(y)-y)dF_1(y). \end{aligned} \tag{4.9}$$

In particular, if we suppose that the distribution of the units functioning in warm and hot conditions differ only in scale, i.e.

$$F_2(t) = F_1(rt),$$

for some unknown $r > 0$, then $g(y) = ry$. This make the sense of AFT model. In such case the following hypothesis is to be verified :

$$H_0^* : f_2^{(y)}(x) = f_1(x + ry - y), \quad \forall \quad x \geq y \geq 0, \tag{4.10}$$

and we can write the cumulative distribution function of the system as

$$K_2(t) = \int_0^t F_1(t + ry - y)dF_1(y). \tag{4.11}$$

The hypothesis H_0^* can also be considered as the generalization of the accelerated failure time (AFT) model to the case of stress with random switch-on (Bagdonavicius, 1978). In this text we suppose that the cumulative distribution function of failure times for both units belongs to the Birnbaum-Saunders family. So we need to estimate the parameters of the models and other reliability characteristics, for example, distribution function of the system from the data which can be noncensored or censored. But before going to the estimation problem, it is necessary to test the model given by the hypotheses H_0^* in equation (4.10).

7 Generalization Of The Redundant System $S(1, m-1)$

Following we consider redundant systems with one main unit and $m-1$, $m \geq 2$, standby units operating in *warm* conditions, i.e. under lower stress than the main one. We denote $S(1, m-1)$ for such systems.

Let denote by T_1, F_1 and f_1 the failure time, the c.d.f. and the probability density function of the main unit. The failure times of the standby units denote by $T_2, \ldots, T_m$, the c.d.f. of T_i is F_2 and the p.d.f. is f_2, $i = 2, \ldots, m$. If a stand-by unit is switched to *hot* conditions, its c.d.f. is different from F_1 and F_2.

The failure time of the system $S(1, m-1)$ is

$$T^{(m)} = T_1 \vee T_2 \vee \cdots \vee T_m.$$

Denote by K_j and k_j the c.d.f. and the p.d.f. of $T^{(j)}$, respectively, $(j = 2, \ldots, m)$, $K_1 = F_1$, $k_1 = f_1$. The c.d.f K_j can be written in terms of the c.d.f. K_{j-1} and F_1 :

$$K_j(t) = \mathbf{P}(T^{(j)} \leq t) = \mathbf{P}(T^{(j-1)} \leq t, T_j \leq t) = \int_0^t \mathbf{P}(T_j \leq t | T^{(j-1)} = y) dK_{j-1}(y).$$

The *fluent switch on* hypothesis H_0, formulated by Bagdonavičius et al. (2008), states that

$$f_{T_j|T^{(j-1)}=y}(t) = \begin{cases} f_2(t) & \text{if} \quad t \leq y, \\ f_1(t + g(y) - y) & \text{if} \quad t > y, \end{cases}$$

where

$$g(y) = F_1^{-1}(F_2(y))$$

is the so-called *transfer functional*, (Bagdonavicius and Nikulin (2002)).

This model implies that

$$K_j(t) = \int_0^t F_1(t + g(y) - y) dK_{j-1}(y), \quad K_1(t) = F_1(t), \tag{4.12}$$

from where it follows that the distribution function K_m of the system with $m-1$ stand-by units is defined recurrently using the formula (4.12). It is often in practice that the c.d.f. of units functioning in *hot* and *warm* conditions belong to the same parametric families of distribution. Here we do the accent on the family of BS-distribution.

8 Birnbaum-Saunders (BS) Family Of Life Distributions

The family of Birnbaum-Saunders (BS) distributions is widely used for failure time data especially when the failures are due to crack. This family was proposed by Birnbaum and Saunders (1969) with two parameters, named as shape and scale parameters. Fatigue failure is due to repeated applications of a common cyclic stress pattern. The PDF and hazard function of this distribution is unimodal and is very popular in modeling fatigue failures in industry as an alternative to other unimodal distributions such as the lognormal and inverse Gaussian. Desmond (1986) worked on the relationship between Birnbaum-Saunders distribution and the family of inverse Gaussian distributions. The hazard functions for both of these distributions are very similar (for details see chapter 1 section 9).

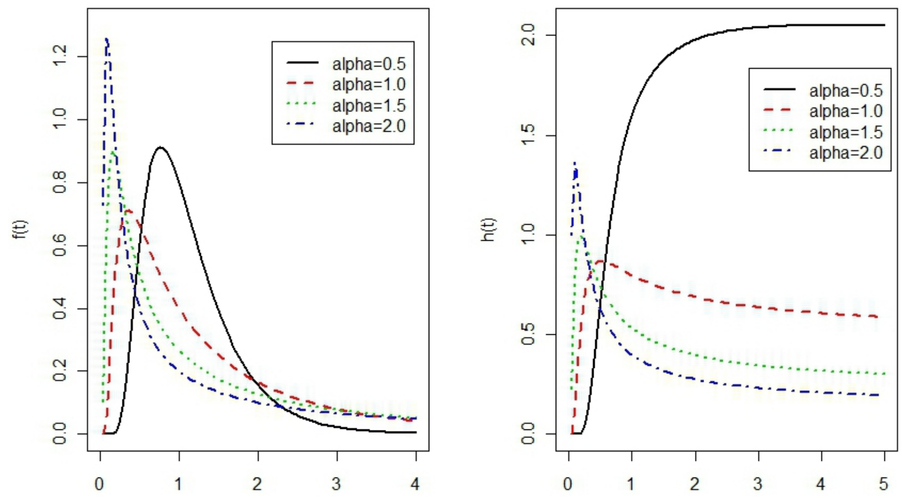

Figure 4.4 – Pdf (left) and hazard function (right) for BS distribution for $\beta = 1$.

The cumulative distribution function of two-parameter Birnbaum-Saunders distribution is

$$F(t;\alpha,\beta) = \Phi\left[\frac{1}{\alpha}\left\{\left(\frac{t}{\beta}\right)^{\frac{1}{2}} - \left(\frac{\beta}{t}\right)^{\frac{1}{2}}\right\}\right], \quad 0 < t < \infty, \quad \alpha, \beta > 0, \tag{4.13}$$

where α is the shape parameter, β is the scale parameter and $\Phi(x)$ is the standard normal distribution function. The probability density function can be written as

$$f(t;\alpha,\beta) = \frac{1}{2\sqrt{2\pi}\ \alpha\beta}\left\{\left(\frac{\beta}{t}\right)^{\frac{1}{2}} + \left(\frac{\beta}{t}\right)^{\frac{3}{2}}\right\}\exp\left[-\frac{1}{2\alpha^2}\left(\frac{t}{\beta} + \frac{\beta}{t} - 2\right)\right], \quad t > 0, \quad \alpha, \beta > 0.$$

A lot of work has been done on the Birnbaum-Saunders distribution and its application in the failure time data but not a lot of work is done on its uses in the redundant system for to enhance the reliability (Balakrishnan et al. (2007, 2009), Johnson et al. (1995), Kundu et al. (2008), Lemonte et al. (2007), Volodin & Dzhungurova, (2000)).

9 Goodness-of-fit Test For Hypotheses H_0^*

For testing the hypotheses we suppose the following plan of experiment for data :

a) the failure times $T_{11}, \cdots, T_{1n_1}$ of n_1 units tested in hot conditions,

b) the failure times $T_{21}, \cdots, T_{2n_2}$ of n_2 units tested in warm conditions,

c) the failure times $T_1, \cdots, T_n$ of n redundant systems (with warm standby units).

The test is based on the difference of two estimators of the cumulative distribution function $F(\cdot)$ of the system failure time T. The first estimator is the nonparametric estimator based on the empirical distribution function obtained from the data $T_1, \cdots, T_n$

$$\hat{F}^{(1)}(t) = \frac{1}{n} \sum_{i=1}^{n} 1_{\{T_i \leq t\}}, \quad t \geq 0,$$

using the data from the third sample. The second estimator, by using the data $T_{11}, \cdots, T_{1n_1}$ and $T_{21}, \cdots, T_{2n_2}$ and is based on the

$$\hat{F}^{(2)}(t) = \int_0^t \hat{F}_1(t + \hat{g}(y) - y) d\hat{F}_1(y),$$

where (if we *test the hypothesis* H_0)

$$\hat{g}(y) = \hat{F}_1^{-1}(\hat{F}_2(y)), \quad \hat{F}_j(t) = \frac{1}{n_j} \sum_{i=1}^{n_j} 1_{\{T_{ji} \leq t\}}, \quad \hat{F}_1^{-1}(y) = \inf\{s : \hat{F}_1(s) \geq t\},$$

and (if we *test the hypothesis* H_0^*), then

$$\hat{F}^{(2)}(t) = \int_0^t \hat{F}_1(t + \hat{r}y - y) d\hat{F}_1(y),$$

where

$$\hat{r} = \frac{\hat{\mu}_1}{\hat{\mu}_2}, \quad \hat{\mu}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} T_{ji}, \quad j = 1, 2,$$

where $\hat{\mu}_1$ and $\hat{\mu}_2$ are the means of two empirical distributions based on the first and second samples respectively.
The test is based on the statistic

$$X = \sqrt{n} \int_0^\infty [\hat{F}^{(1)}(t) - \hat{F}^{(2)}(t)] dt \tag{4.14}$$

It is natural generalization of Student's type t-test for comparing the means of two populations. Indeed, the mean failure time of the system with distribution function F is

$$\mu = \int_0^\infty [1 - F(s)] ds,$$

so the statistic (4.14) is the normed difference of two estimators (the second being not the empirical mean) of the mean μ. Student's type t-test is based on the difference of empirical means of two populations. To find the asymptotic distribution of the statistic (4.14), we consider the following theorem.

Theorem 9.1 *Bagdonavičius, Masiulaityle, and Nikulin (2008b) : Suppose that* $n_i/n \to l_i \in (0,1), n \to \infty$ *and the densities* $f_i(x), i = 1, 2$ *are continuous and positive on* $(0, \infty)$*. Then under*

H_0^* the statistic X converges in distribution to the normal law $N(0,\sigma^2)$, where

$$\sigma^2 = Var(T_i) + \frac{1}{l_1} Var(H(T_{1i})) + \frac{c^2 r^2}{l_2^2} Var(T_{2i}),$$

where

$$H(x) = x[c + r - 1 - F_1(x/r) - rF_2(x)] + r\mathbf{E}(1_{\{T_{1i} \le x/r\}} T_{1i}) + r\mathbf{E}(1_{\{T_{2i} \le x\}} T_{2i}),$$

$$c = \frac{1}{\mu^2} \int_0^\infty y[1 - F_2(y)] dF_1(y).$$

The test statistic is

$$Y_n = \frac{X}{\hat{\sigma}}, \tag{4.15}$$

where $\hat{\sigma}$ is a consistent estimator of σ and is estimated as

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (T_i - \hat{\mu})^2 + \frac{n}{n_1^2} \sum_{i=1}^{n_1} [\hat{H}(T_{1i}) - \hat{\bar{H}}]^2 + \frac{\hat{c}^2 \hat{r}^2 n}{n_2^2} \sum_{i=1}^{n_2} (T_{2i} - \hat{\mu}_2)^2,$$

where

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} T_i, \quad \hat{c} = \frac{1}{\hat{\mu}_2} \int_0^\infty y[1 - \hat{F}_2(y)] d\hat{F}_1(y) = \frac{1}{\hat{\mu}_2} \sum_{i=1}^{n_1} T_{1i}[1 - \hat{F}_2(T_{1i})],$$

$$\hat{H}(x) = x[\hat{c} + \hat{r} - 1 - \hat{F}_1(x/\hat{r}) - \hat{r}\hat{F}_2(x)] + \frac{\hat{r}}{n_1} \sum_{i=1}^{n_1} 1_{\{T_{1i} \le x/\hat{r}\}} T_{1i} + \frac{\hat{r}}{n_2} \sum_{i=1}^{n_2} 1_{\{T_{2i} \le x\}} T_{2i},$$

$$\hat{\bar{H}} = \frac{1}{n} \sum_{i=1}^{n_1} \hat{H}(T_{1i}).$$

The distribution of the statistic Y_n is approximated by the standard normal distribution and the hypothesis H_0^* is rejected with approximative significance value α if $|Y_n| > z_{\alpha/2}$, where $z_{\alpha/2}$ is the upper-$(\alpha/2)$ critical value of the standard normal distribution.

Remark : If switching from warm to hot conditions does not damage units in the system $S(1,1)$ then it is natural that this is true for the system $S(1, m-1)$, $m > 2$. So it is sufficient to use goodness-of-fit tests for the hypotheses H_0 and H_0^* when only one stand-by unit is used. Now we check the power of the test with the alternative. Bagdonavičius et al. (2008b) proved the similar results for nonparametric case.

9.1 Power of the tests

Now we want to investigate the power of the above goodness-of-fit tests when the distribution of the units in *warm* and *hot* conditions is Birnbaum-Saunders. Let us consider the following alternative hypothesis $\tilde{H}_0^*$:

$$f_2^{(y)}(x) = f_1[x + F_1^{-1}(F_2(y) + p(1 - F_2(y)) - y], 0 < p < 1.$$

It means that at the switching time y the c.d.f. of the standby unit has a jump of size $p(1-F_2(y))$. Set

$$g_p(y) = F_1^{-1}(F_2(y) + p[1 - F_2(y)])).$$

Under the alternative hypothesis the c.d.f. of the standby system is

$$\begin{aligned} F(t) &= \int_0^t F_1(t + g_p(y) - y) dF_1(y) \\ &= F_1(t) - \int_0^t S_1(t + g_p(y) - y) dF_1(y). \end{aligned} \tag{4.16}$$

Now we apply this approach to the case when the elements of the system belongs to the BS distribution, that is

$$T_{1i} \sim BS(\alpha_1, \beta_1), \qquad T_{2j} \sim BS(\alpha_2, \beta_2).$$

Under the hypothesis H_0 the function g can be written as

$$g(t) = \beta_1 \left\{ \alpha_1 \Phi^{-1} \left(F_2(t) \right) /2 + \sqrt{(\alpha_1 \Phi^{-1}(F_2(t))/2)^2 + 1} \right\}^2$$

We consider $\alpha_1 = \alpha_2$, then the hypothesis H_0 coincides with the hypothesis H_0^*. In this case $g(t) = rt$ and we get $r = \beta_1/\beta_2$.

Under the alternative hypothesis $\tilde{H}_0^*$ with BS distribution we can write

$$g_p(y) = \beta_1 \left\{ \alpha_1 \Phi^{-1} \left(F_2(y) + p[1 - F_2(y)] \right) /2 + \sqrt{(\alpha_1 \Phi^{-1}(F_2(y) + p[1 - F_2(y)])/2)^2 + 1} \right\}^2 \tag{4.17}$$

where $F_2(y)$ is the cumulative distribution function of BS distribution with parameters α_2 and β_2. The distribution function of the redundant system under alternative hypothesis can be written as

$$\begin{aligned} F(t) &= \frac{1}{2\sqrt{2\pi}\alpha\beta} \int_0^t \Phi \left\{ \frac{1}{\alpha} \left[\left(\frac{t + g_p(y) - y}{\beta} \right)^{1/2} - \left(\frac{\beta}{t + g_p(y) - y} \right)^{1/2} \right] \right\} \times \\ & \left\{ \left(\frac{\beta}{y} \right)^{1/2} + \left(\frac{\beta}{y} \right)^{3/2} \right\} \exp \left\{ -\frac{1}{2\alpha^2} \left(\frac{y}{\beta} + \frac{\beta}{y} - 2 \right) \right\} dy, \end{aligned} \tag{4.18}$$

where $g_p(y)$ is given as in equation (4.17).

10 Estimation Of Distribution Function Of Redundant System $S(1,1)$

Bagdonavičius, Masiulaityle, and Nikulin (2008b, 2009, 2010) studied asymptotic properties of point and interval estimators of reliability characteristics of redundant system for uncensored and censored data based on various parametric models like exponential, Weibull, and loglogistic. They also give the nonparametric estimation of the cumulative distribution function but here we study only the parametric estimation. We use their results to apply on the BS family of life distributions. Separate plans of experiments for data are considered for complete and censored data.

From equation (4.11) using the Birnbaum-Saunders family of distributions, the cumulative distribution function $K_2(t)$ of the system system $S(1,1)$ is estimated by

$$\begin{aligned}\hat{K}_2(t) &= \frac{1}{2\sqrt{2\pi}\hat{\alpha}\hat{\beta}}\int_0^t \Phi\left\{\frac{1}{\hat{\alpha}}\left[\left(\frac{t+\hat{r}y-y}{\hat{\beta}}\right)^{1/2} - \left(\frac{\hat{\beta}}{t+\hat{r}y-y}\right)^{1/2}\right]\right\}\times \\ &\left\{\left(\frac{\hat{\beta}}{y}\right)^{1/2}+\left(\frac{\hat{\beta}}{y}\right)^{3/2}\right\}\exp\left\{-\frac{1}{2\hat{\alpha}^2}\left(\frac{y}{\hat{\beta}}+\frac{\hat{\beta}}{y}-2\right)\right\}dy, \qquad (4.19)\end{aligned}$$

where $\hat{\alpha}, \hat{\beta}, \hat{r}$ are the maximum likelihood estimators.

10.1 Asymptotic Confidence Interval For $K_2(t)$

Denote by $I_n(\gamma) = -\mathbf{E}\ddot{\ell}(\gamma)$ the Fisher information matrix and suppose that $\frac{1}{n}I_n(\gamma) \to i(\gamma)$. Under classical assumptions on the family of distributions $f_1(t,\theta)$ the maximum likelihood estimator γ^* is asymptotically normal :

$$\sqrt{n}(\gamma^*-\gamma) \xrightarrow{d} Y = (Y_1, Y_2^T)^T \sim N_{k+1}(0, i^{-1}(\gamma)).$$

Y_1 is one dimensional, Y_2 is k-dimensional. Using the delta method we get :

$$\sqrt{n}(\hat{K}_2(t) - K_2(t)) \xrightarrow{\mathcal{D}} W_2(t) = Y^T C_2(t;\gamma),$$

where

$$C_2(t,\gamma) = (C_{21}(t,\gamma), C_{22}(t,\gamma), C_{23}(t,\gamma))^T,$$

$$C_{21}(t,\gamma) = \int_0^t \frac{\partial F_1}{\partial r}(t+ry-y;\alpha,\beta)dF_1(y;\alpha,\beta),$$

$$C_{22}(t,\gamma) = \int_0^t \frac{\partial F_1}{\partial \alpha}(t+ry-y;\alpha,\beta)dF_1(y,\alpha,\beta) + F_1(t+ry-y;\alpha,\beta)d(\frac{\partial F_1}{\partial \alpha}(y;\alpha,\beta)),$$

$$C_{23}(t,\gamma) = \int_0^t \frac{\partial F_1}{\partial \beta}(t+ry-y;\alpha,\beta)dF_1(y,\alpha,\beta) + F_1(t+ry-y;\alpha,\beta)d(\frac{\partial F_1}{\partial \beta}(y;\alpha,\beta)).$$

The random variable $W_2(t)$ is linear function of Y. If $j \geq 2$ then

$$\sqrt{n}(\hat{K}_j(t) - K_j(t)) \xrightarrow{\mathcal{D}} W_j(t).$$

So the variance

$$Var(W_j(t)) = Var(C_j(t,\gamma)^T Y) = C_j^T(t,\gamma) i^{-1}(\gamma) C_j(t,\gamma)$$

is estimated by $nC_j^T(t,\hat{\gamma})I^{-1}(\hat{\gamma})C_j(t,\hat{\gamma})$, and variance $\sigma^2_{\hat{K}_j(t)}$ of the estimator $\hat{K}_j(t)$ is estimated by

$$\hat{\sigma}^2_{\hat{K}_2(t)} = C_2^T(t,\hat{\gamma})I^{-1}(\hat{\gamma}_1)C_2(t,\hat{\gamma}),$$

and the matrix $I(\hat{\gamma})$ is replaced by $-\ddot{\ell}(\hat{\gamma})$. So the asymptotic $(1-\alpha)$ confidence interval for $K_j(t)$ is

$$\hat{K}_j(t) \pm \hat{\sigma}_{\hat{K}_j(t)} z_{1-\alpha/2}. \tag{4.20}$$

And one can write the asymptotic confidence interval $(\underline{K}_2(t), \overline{K}_2(t))$ for $K_2(t)$, where

$$\underline{K}_2(t) = \left(1 + \frac{1-\hat{K}_2(t)}{\hat{K}_2(t)} exp\left\{ \frac{\hat{\sigma}_{\hat{K}_2} z_{1-\alpha/2}}{\sqrt{\hat{K}_2(t)(1-\hat{K}_2(t))}} \right\}\right)^{-1},$$

$$\overline{K}_2(t) = \left(1 + \frac{1-\hat{K}_2(t)}{\hat{K}_2(t)} exp\left\{ -\frac{\hat{\sigma}_{\hat{K}_2} z_{1-\alpha/2}}{\sqrt{\hat{K}_2(t)(1-\hat{K}_2(t))}} \right\}\right)^{-1},$$

10.2 Noncensored (complete) data

Suppose that the hypothesis H_0^* is true and we have following data,

a) the failure times $T_{11}, \cdots, T_{1n_1}$ of n_1 units tested in hot conditions,

b) the failure times $T_{21}, \cdots, T_{2n_2}$ obtained by testing of n_2 units in warm conditions,

and in hot conditions the cumulative distribution function $F_1(t;\boldsymbol{\theta})$ is absolutely continuous and belongs to the parametric Birnbaum-Saunders family, here $\boldsymbol{\theta}^T = (\alpha,\beta)$. Set $\gamma = (r,\boldsymbol{\theta}^T)^T$. The maximum likelihood estimator $\hat{\gamma} = (\hat{r},\hat{\boldsymbol{\theta}}^T)^T$ of the parameter γ maximizes following loglikelihood function

$$\ell(\gamma) = -n\ln(\alpha\beta) + \frac{n}{\alpha^2} - \frac{1}{2\alpha^2\beta}\sum_{i=1}^{n_1} T_{1i} - \frac{\beta}{2\alpha^2}\sum_{i=1}^{n_1} T_{1i}^{-1} - \frac{r}{2\alpha^2\beta}\sum_{j=1}^{n_2} T_{2j} - \frac{\beta}{2r\alpha^2}\sum_{j=1}^{n_2} T_{2j}^{-1}$$
$$+n_2\ln r + \sum_{i=1}^{n_1}\ln\left\{\left(\frac{\beta}{T_{1i}}\right)^{1/2} + \left(\frac{\beta}{T_{1i}}\right)^{3/2}\right\} + \sum_{j=1}^{n_2}\ln\left\{\left(\frac{\beta}{rT_{2j}}\right)^{1/2} + \left(\frac{\beta}{rT_{2j}}\right)^{3/2}\right\},$$

where $n = n_1 + n_2$. Parameters can be estimated by solving $\ell(\gamma) = 0$. The plot of the trajectories of $\hat{F}_1$ and $\hat{K}_2$ is shown in Figure 4.5.

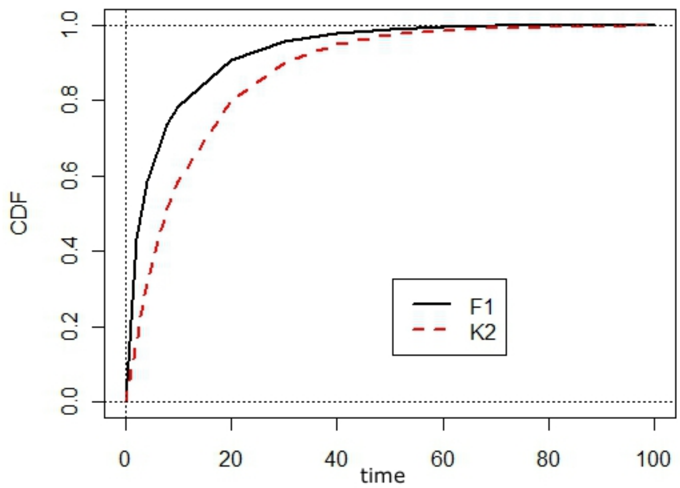

Figure 4.5 – Trajectories of the parametric estimators $\hat{F}_1$ and $\hat{K}_2$ (BS distribution).

Fisher information matrix can be estimated from the second derivatives of the loglikelihood function i.e.

$$I(\gamma) = -E\ddot{\ell}(\gamma),$$

and it may be replaced by $-\ddot{\ell}(\hat{\gamma})$. The second partial derivatives of loglikelihood function are

$$\frac{\partial^2\ell}{\partial\beta^2} = \frac{n}{\beta^2} - \frac{1}{2\beta^2\alpha^2}\sum_{i=1}^{n_1}\frac{\frac{T_{1i}}{\beta} + 2 + 3\left(\frac{\beta}{T_{1i}}\right)}{B_i^2} - \frac{1}{2\beta^2\alpha^2}\sum_{j=1}^{n_2}\frac{\frac{rT_{2j}}{\beta} + 2 + 3\left(\frac{\beta}{rT_{2j}}\right)}{A_j^2} - \frac{1}{\alpha^2\beta^3}\left[\sum_{i=1}^{n_1} T_{1i} + r\sum_{j=1}^{n_2} T_{2j}\right],$$

$$\frac{\partial^2\ell}{\partial r^2} = -\frac{n_2}{r^2} + \frac{1}{2r^2\alpha^2}\sum_{j=1}^{n_2}\frac{\frac{rT_{2j}}{\beta} + 6 + 3\frac{\beta}{rT_{2j}}}{A_j^2} - \frac{1}{\alpha^2 r^2}\sum_{j=1}^{n_2}\frac{\beta}{rT_{2j}},$$

$$\frac{\partial^2 \ell}{\partial \alpha^2} = \frac{n}{\alpha^2} - \frac{3}{\alpha^2}\left[\sum_{i=1}^{n_1} A_i^2 + \sum_{j=1}^{n_2} B_j^2\right], \qquad \frac{\partial^2 \ell}{\partial \alpha \partial \beta} = -\frac{1}{\alpha\beta}\left[\sum_{i=1}^{n_1} A_i B_i + \sum_{j=1}^{n_2} A_j B_j\right],$$

$$\frac{\partial^2 \ell}{\partial \beta \partial r} = -\frac{1}{\beta\alpha^2 r}\sum_{j=1}^{n_2}\frac{1}{A_j^2} + \frac{1}{2\alpha^2\beta r}\sum_{j=1}^{n_2}\left(\alpha^2 A_j^2 - 2\right), \qquad \frac{\partial^2 \ell}{\partial \alpha \partial r} = \frac{1}{r\alpha}\sum_{j=1}^{n_2} A_j B_j,$$

where

$$A_i = \frac{1}{\alpha}\left\{\left(\frac{T_{1i}}{\beta}\right)^{\frac{1}{2}} - \left(\frac{\beta}{T_{1i}}\right)^{\frac{1}{2}}\right\}, \quad B_i = \frac{1}{\alpha}\left\{\left(\frac{T_{1i}}{\beta}\right)^{\frac{1}{2}} + \left(\frac{\beta}{T_{1i}}\right)^{\frac{1}{2}}\right\},$$

$$B_j = \frac{1}{\alpha}\left\{\left(\frac{rT_{2j}}{\beta}\right)^{\frac{1}{2}} - \left(\frac{\beta}{rT_{2j}}\right)^{\frac{1}{2}}\right\}, \quad A_j = \frac{1}{\alpha}\left\{\left(\frac{rT_{2j}}{\beta}\right)^{\frac{1}{2}} + \left(\frac{\beta}{rT_{2j}}\right)^{\frac{1}{2}}\right\}.$$

And the elements of the vector $C_2(t,\gamma)$ for BS distribution are

$$C_{21}(t,\gamma) = \frac{1}{2\alpha\beta}\int_0^t y\left\{\left(\frac{\beta}{t+ry-y}\right)^{\frac{1}{2}} + \left(\frac{\beta}{t+ry-y}\right)^{\frac{3}{2}}\right\}\varphi\left(A(t+ry-y)\right) f_1(y;\alpha,\beta)dy,$$

$$C_{22}(t,\gamma) = \int_0^t \left(\frac{\partial}{\partial \alpha}F_1(t+ry-y;\alpha,\beta)f_1(y,\alpha,\beta) + F_1(t+ry-y;\alpha,\beta)\left|\frac{\partial}{\partial \alpha}F_1(y;\alpha,\beta)\right|_y^{\prime}\right)dy,$$

$$C_{23}(t,\gamma) = \int_0^t \left(\frac{\partial}{\partial \beta}F_1(t+ry-y;\alpha,\beta)f_1(y,\alpha,\beta) + F_1(t+ry-y;\alpha,\beta)\left|\frac{\partial}{\partial \beta}F_1(y;\alpha,\beta)\right|_y^{\prime}\right)dy,$$

where

$$\frac{\partial}{\partial \alpha}F_1(t) = -\frac{1}{\alpha}A(t)\varphi(A(t)), \qquad \frac{\partial}{\partial \beta}F_1(t) = -\frac{1}{2\beta}B(t)\varphi(A(t)),$$

$$A(t) = \frac{1}{\alpha}\left\{\left(\frac{t}{\beta}\right)^{\frac{1}{2}} - \left(\frac{\beta}{t}\right)^{\frac{1}{2}}\right\}, \qquad B(t) = \frac{1}{\alpha}\left\{\left(\frac{t}{\beta}\right)^{\frac{1}{2}} + \left(\frac{\beta}{t}\right)^{\frac{1}{2}}\right\},$$

$$\left|\frac{\partial}{\partial \alpha}F_1(y;\alpha,\beta)\right|_y^{\prime} = -\frac{1}{2\alpha^2\beta}\left\{\left(\frac{\beta}{y}\right)^{\frac{1}{2}} + \left(\frac{\beta}{y}\right)^{\frac{3}{2}}\right\}\varphi(A(y))[1-A^2(y)],$$

$$\left|\frac{\partial}{\partial \beta}F_1(y;\alpha,\beta)\right|_y^{\prime} = -\frac{1}{4\alpha\beta^2}\varphi(A(y))\left(\left\{\left(\frac{\beta}{y}\right)^{\frac{1}{2}} - \left(\frac{\beta}{y}\right)^{\frac{3}{2}}\right\} - B(y)A(y)\left\{\left(\frac{\beta}{y}\right)^{\frac{1}{2}} + \left(\frac{\beta}{y}\right)^{\frac{3}{2}}\right\}\right),$$

and $f_1(y)$ is the pdf of the BS distribution.

10.2.1 Simulation Study

In this section some simulated results are given for confidence intervals with Birnbaum-Saunders distribution and Lognormal distribution. The mathematical formulas for Lognormal distribution are given in Appendix 2.

10.2.2 Birnbaum Saunders Distribution

Let us consider the case of complete sample of size $n_1 = n_2 = 100$. The experiment is repeated 2000 times. We find by simulation the confidence levels of intervals using formulas with $1 - \alpha = 0.9$. We simulated failure times T_{1i} and T_{2j} from Birnbaum-Saunders distribution with the parameters :

$$T_{1i} \sim BS(\alpha_1, \beta_1), \quad T_{2j} \sim BS(\alpha_2, \beta_2)$$

$$\alpha_1 = \alpha_2 = 2, \quad \beta_1 = 2, \quad \beta_2 = 4.$$

For various values of t the proportions of confidence interval (C.L.) realizations covering the true value of the distribution function $K_2(t)$ are given in table 4.1.

t	10	20	50	70	100
$K_2(t)$	0.6428	0.8374	0.9813	0.9953	0.9994
Confidence level (%)	80.35	75.23	69.33	57.20	45.89

Table 4.1 – Confidence level for the c.d.f. of the redundant system (BS distribution).

10.2.3 Lognormal Distribution

Let us consider the case of complete sample of size $n_1 = n_2 = 100$. Each sample is repeated 5000 times. We find by simulation the confidence levels of intervals using formulas with $1 - \alpha = 0.9$. We simulated failure times T_{1i} and T_{2j} from Log-normal distribution with the parameters :

$$T_{1i} \sim LN(m_1, \sigma_1), \quad T_{2j} \sim LN(m_2, \sigma_2)$$

$$m_1 = 1, \quad m_2 = -0.3862944, \quad \sigma_1 = \sigma_2 = 1.5.$$

For various values of t the proportions of confidence interval (C.L.) realizations covering the true value of the distribution function $K_2(t)$ are given in table 4.2.

The simulation studies for inverse Gaussian and generalized Weibull distributions are published in Nikulin et al. (2011c).

t	10	50	100	200	300
$K_2(t)$	0.7119	0.9556	0.9855	0.9961	0.9983
Confidence level (%)	82.25	88.87	83.55	86.60	89.04

Table 4.2 – Confidence level for the c.d.f. of the redundant system (Log-normal distribution).

10.3 Censored data

The cumulative distribution function of the redundant system for censored data can be obtain by using the same formula (4.19), although the estimation of parameters and information matrix can be calculated from the following way.

Suppose that we have following censored data :

a) right censored sample

$$(X_{11}, \delta_{11}), \ldots, (X_{1n_1}, \delta_{1n_1}),$$

of size n_1, where

$$X_{1i} = T_{1i} \wedge C_{1i}, \quad \delta_{1i} = \mathbf{1}_{\{T_{1i} \leq C_{1i}\}},$$

T_{1i} are the failure time of units tested in hot conditions, C_{1i} are the censored times ;

b) right censored sample

$$(X_{21}, \delta_{21}), \ldots, (X_{2n_1}, \delta_{2n_1}),$$

of size n_2, where

$$X_{2i} = T_{2j} \wedge C_{2j}, \quad \delta_{2j} = \mathbf{1}_{\{T_{2j} \leq C_{2j}\}},$$

T_{2j} are the failure time of units tested in warm conditions, C_{2j} are the censored times.

Let denote

$$m_1 = \sum_{i=1}^{n_1} \delta_{1i}, \quad m_2 = \sum_{j=1}^{n_2} \delta_{2j}, \quad \text{and} \quad m = m_1 + m_2.$$

By using the data from above plan of experiment, the maximum likelihood estimator $\hat{\gamma} = (\hat{r}, \hat{\boldsymbol{\theta}}^T)^T$ of the parameter $\gamma = (r, \boldsymbol{\theta}^T)^T = (r, \alpha, \beta)^T$ can be estimated from the following loglikelihood function

$$\begin{aligned} \ell(\gamma) &= \sum_{i=1}^{n_1} \delta_{1i} \ln f_1(X_{1i}; \boldsymbol{\theta}) + \sum_{i=1}^{n_1} (1 - \delta_{1i}) \ln S_1(X_{1i}; \boldsymbol{\theta}) + m_2 \ln r \\ &+ \sum_{j=1}^{n_2} \delta_{2j} \ln f_1(rX_{2j}; \boldsymbol{\theta}) + \sum_{j=1}^{n_2} (1 - \delta_{2j}) \ln S_1(rX_{2j}; \boldsymbol{\theta}). \end{aligned} \tag{4.21}$$

The loglikelihood function in terms of Birnbaum-Saunders distribution can be written as

$$\begin{aligned}\ell(\gamma) &= -m(\ln\alpha + \ln\beta) + m_2 \ln r \\ &+ \sum_{i=1}^{n_1} \delta_{1i} \ln\left\{\left(\frac{\beta}{X_{1i}}\right)^{\frac{1}{2}} + \left(\frac{\beta}{X_{1i}}\right)^{\frac{3}{2}}\right\} + \sum_{j=1}^{n_2} \delta_{2j} \ln\left\{\left(\frac{\beta}{rX_{2j}}\right)^{\frac{1}{2}} + \left(\frac{\beta}{rX_{2j}}\right)^{\frac{3}{2}}\right\} \\ &- \frac{1}{2\alpha^2}\left[\sum_{i=1}^{n_1} \delta_{1i}\left(\frac{X_{1i}}{\beta} + \frac{\beta}{X_{1i}} - 2\right) + \sum_{j=1}^{n_2} \delta_{2j}\left(\frac{rX_{2j}}{\beta} + \frac{\beta}{rX_{2j}} - 2\right)\right] \\ &+ \sum_{i=1}^{n_1}(1-\delta_{1i}) \ln S_1(X_{1i};\alpha,\beta) + \sum_{j=1}^{n_2}(1-\delta_{2j}) \ln S_1(rX_{2j};\alpha,\beta).\end{aligned}$$

Maximum likelihood estimator $\hat{\gamma}$ can be find by equating the score vector to zero, i.e. $\dot{\ell}(\gamma) = 0$. The estimated values of $\hat{F}_1$ and $\hat{K}_2$ with censored data are plotted in Figure 4.6.

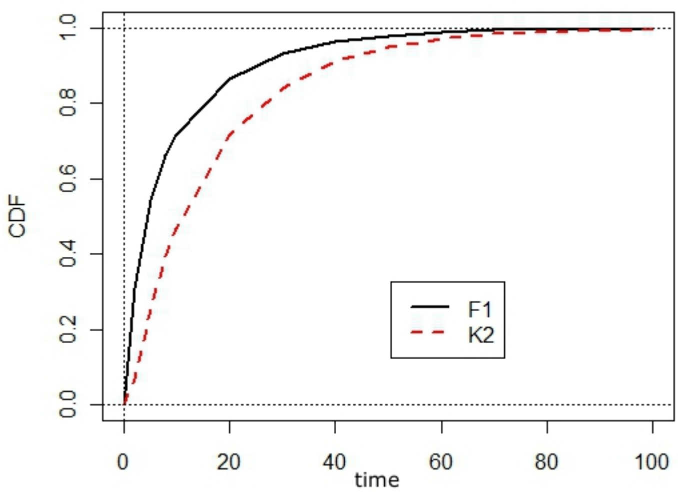

Figure 4.6 – Trajectories of the parametric estimators $\hat{F}_1$ and $\hat{K}_2$ (censored data).

The Fisher's information matrix is

$$I(\gamma) = -E\ddot{\ell}(\gamma),$$

and it may be estimated by $-\ddot{\ell}(\hat{\gamma})$. The second partial derivatives are

$$\begin{aligned}\frac{\partial^2 \ell}{\partial r^2} &= -\frac{m_2}{r^2} + \frac{1}{2r^2\alpha^2}\sum_{j=1}^{n_2} \delta_{2j} \frac{\frac{rX_{2j}}{\beta} + 6 + 3(\frac{\beta}{rX_{2j}})}{A_j^2} - \frac{\beta}{\alpha^2 r^2}\sum_{j=1}^{n_2} \delta_{2j}\frac{1}{rX_{2j}} + \\ &\frac{1}{4r^2}\sum_{j=1}^{n_2}(1-\delta_{2j})\varphi(B_j)\frac{[1-\Phi(B_j)]\left[\frac{1}{\alpha}\left\{\left(\frac{rX_{2j}}{\beta}\right)^{\frac{1}{2}} + 3\left(\frac{\beta}{rX_{2j}}\right)^{\frac{1}{2}}\right\} + A_j^2 B_j\right] - A_j^2\varphi(B_j)}{(1-\Phi(B_j))^2},\end{aligned}$$

$$
\begin{aligned}
\frac{\partial^2 \ell}{\partial \alpha^2} &= \frac{m}{\alpha^2} - \frac{3}{\alpha^2}\left[\sum_{i=1}^{n_1}\delta_{1i}A_i^2 + \sum_{j=1}^{n_2}\delta_{2j}B_j^2\right] + \\
&\quad \frac{1}{\alpha^2}\sum_{i=1}^{n_1}(1-\delta_{1i})\frac{A_i\varphi(A_i)\left[(1-\Phi(A_i))\left[A_i^2-2\right]-A_i\varphi(A_i)\right]}{[1-\Phi(A_i)]^2} \\
&\quad + \frac{1}{\alpha^2}\sum_{j=1}^{n_2}(1-\delta_{2j})\frac{B_j\varphi(B_j)\left[(1-\Phi(B_j))(B_j^2-2)-B_j\varphi(B_j)\right]}{[1-\Phi(B_j)]^2},
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 \ell}{\partial \beta^2} &= \frac{m}{\beta^2} - \frac{1}{2\beta^2\alpha^2}\sum_{i=1}^{n_1}\delta_{1i}\frac{\frac{X_{1i}}{\beta}+2+3(\frac{\beta}{X_{1i}})}{B_i^2} - \frac{1}{2\beta^2\alpha^2}\sum_{j=1}^{n_2}\delta_{2j}\frac{\frac{rX_{2j}}{\beta}+2+3(\frac{\beta}{rX_{2j}})}{A_j^2} \\
&\quad - \frac{1}{\alpha^2\beta^3}\left[\sum_{i=1}^{n_1}\delta_{1i}X_{1i} + r\sum_{j=1}^{n_2}\delta_{2j}X_{2j}\right] - \\
&\quad \frac{1}{4\beta^2}\sum_{i=1}^{n_1}(1-\delta_{1i})\varphi(A_i)\frac{(1-\Phi(A_i))\left[\frac{1}{\alpha}\left\{3\left(\frac{X_{1i}}{\beta}\right)^{\frac{1}{2}}+\left(\frac{\beta}{X_{1i}}\right)^{\frac{1}{2}}\right\}-B_i^2A_i\right]+B_i^2\varphi(A_i)}{(1-\Phi(A_i))^2} - \\
&\quad \frac{1}{4\beta^2}\sum_{j=1}^{n_2}(1-\delta_{2j})\varphi(B_j)\frac{(1-\Phi(B_j))\left[\frac{1}{\alpha}\left\{3\left(\frac{rX_{2j}}{\beta}\right)^{\frac{1}{2}}+\left(\frac{\beta}{rX_{2j}}\right)^{\frac{1}{2}}\right\}-A_j^2B_j\right]+A_j^2\varphi(B_j)}{(1-\Phi(B_j))^2},
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 \ell}{\partial \beta \partial r} &= -\frac{1}{r\alpha^2\beta}\sum_{j=1}^{n_2}\delta_{2j}\frac{1}{A_j^2} + \frac{1}{2r\alpha^2\beta}\sum_{j=1}^{n_2}\delta_{2j}\left(\alpha^2A_j^2-2\right) + \\
&\quad \frac{1}{4\beta r}\sum_{j=1}^{n_2}(1-\delta_{2j})\varphi(B_j)\frac{(1-\Phi(B_j))B_j(9+2\alpha^2B_j^2)+A_j^2\varphi(B_j)}{(1-\Phi(B_j))^2},
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 \ell}{\partial \alpha \partial \beta} &= -\frac{1}{\alpha\beta}\left[\sum_{i=1}^{n_1}\delta_{1i}A_iB_i + \sum_{j=1}^{n_2}\delta_{2j}A_jB_j\right] + \\
&\quad \frac{1}{2\alpha\beta}\sum_{j=1}^{n_2}(1-\delta_{1i})B_i\varphi(A_i)\frac{(1-\Phi(A_i))(A_i^2-1)-\varphi(A_i)A_i}{(1-\Phi(A_i))^2} + \\
&\quad \frac{1}{2\alpha\beta}\sum_{j=1}^{n_2}(1-\delta_{2j})A_j\varphi(B_j)\frac{(1-\Phi(B_j))(B_j^2-1)-\varphi(B_j)B_j}{[1-\Phi(B_j)]^2},
\end{aligned}
$$

$$
\frac{\partial^2 \ell}{\partial \alpha \partial r} = \frac{1}{r\alpha}\sum_{j=1}^{n_2}\delta_{2j}A_jB_j + \frac{1}{2r\alpha}\sum_{j=1}^{n_2}(1-\delta_{2j})\varphi(B_j)A_j\frac{(1-\Phi(B_j))(1-B_j^2)+B_j\varphi(B_j)}{(1-\Phi(B_j))^2},
$$

where

$$A_i = \frac{1}{\alpha}\left\{\left(\frac{X_{1i}}{\beta}\right)^{\frac{1}{2}} - \left(\frac{\beta}{X_{1i}}\right)^{\frac{1}{2}}\right\}, \quad B_i = \frac{1}{\alpha}\left\{\left(\frac{X_{1i}}{\beta}\right)^{\frac{1}{2}} + \left(\frac{\beta}{X_{1i}}\right)^{\frac{1}{2}}\right\},$$

$$B_j = \frac{1}{\alpha}\left\{\left(\frac{rX_{2j}}{\beta}\right)^{\frac{1}{2}} - \left(\frac{\beta}{rX_{2j}}\right)^{\frac{1}{2}}\right\}, \quad A_j = \frac{1}{\alpha}\left\{\left(\frac{rX_{2j}}{\beta}\right)^{\frac{1}{2}} + \left(\frac{\beta}{rX_{2j}}\right)^{\frac{1}{2}}\right\}.$$

The estimator of the variance $\hat{\sigma}^2_{\hat{K}_2(t)}$ of the estimator $\hat{K}_2(t)$ is

$$\hat{\sigma}^2_{\hat{K}_2(t)} = C_2^T(t;\hat{r},\hat{\alpha},\hat{\beta})(-\ddot{\ell}(\hat{r},\hat{\alpha},\hat{\beta}))^{-1}C_2(t;\hat{r},\hat{\alpha},\hat{\beta}), \tag{4.22}$$

where $C_2(t;\gamma) = (C_{21}, C_{22}, C_{23})^T$ is given in previous section. The asymptotic $(1-\alpha)$ confidence interval for $K_2(t)$ is calculated from 4.20.

Remark : we can generalize this with multiple standby units based on BS family of life distribution to increase the reliability. But increase in the number of units increase the cost of system and complications in the calculations. So, the number of stand-by units depend on the optimality of the particular under-study system from the industry. The study of the optimality is the perspective of our work.

10.3.1 Simulation

Let us consider the right censored sample of size $n_1 = n_2 = 100$, with 15 percent random censoring. The experiment is repeated 2000 times. We find by simulation the confidence levels of intervals using formulas with $1-\alpha = 0.9$. We simulated censoring times from uniform distribution and failure times T_{1i} and T_{2j} from BS distribution with the parameters :

$$T_{1i} \sim BS(\alpha_1, \beta_1), \quad T_{2j} \sim BS(\alpha_2, \beta_2)$$

$$\alpha_1 = \alpha_2 = 2, \quad \beta_1 = 3, \quad \beta_2 = 6.$$

For various values of t the proportions of confidence interval (C.L.) realizations covering the true value of the distribution function $K_2(t)$ are given as

t	10	20	50	70	100
$K_2(t)$	0.6428	0.8374	0.9813	0.9953	0.9994
Confidence level (%)	50.44	44.00	41.03	36.72	29.55

Table 4.3 – Confidence level for the c.d.f. of the redundant system (censored data).

Conclusion and Perspectives

Chi-square type goodness-of-fit test using RRN statistic is applied to various parametric models and explicit forms of the elements of quadratic forms are presented. The calculation is still a problem. As the future work a R-package can be developed for RRN test for the selection of suitable parametric model. Presence of covariates can also be treated i.e. test for the parametric AFT model for various parametric families. In this thesis we explained the test for Birnbaum-Saunders AFT model.

Tests for general *fluent switching hypothesis* formulated using *Sedyakin's reliability principle* and for particular fluent switching hypothesis formulated using accelerated failure time model are explained. Parametric estimators of the cumulative distribution function $\hat{K}_2$ of redundant system $S(1,1)$ using reliability data of components tested under different stresses are calculated. Estimation of $\hat{K}_3, \cdots, \hat{K}_m$ for redundant system and simulation work can be interesting to do in further research.

Asymptotic confidence intervals for cumulative distribution function of redundant system are constructed. Parametric estimators of the system are investigated by simulation using finite samples. We studied the redundant system for most commonly used models in reliability analysis like exponential, Weibull, loglogistic, lognormal, inverse Gaussian, generalized Weibull, and Birnbaum-Saunders distributions. A R-package can be developed for reliability analysis of redundant system for all these models.

Bibliography

[1] Akritas, M. (1988). Pearson-type goodness of fit tests : the univariate case. *J. Amer.Statist.Assoc.*, 83 :222–230. 5, 32

[2] Bagdonavicius, V., Kruopis, J., and Nikulin, M. (2010a). *Nonparametric tests for Censored Data*. ISTE and J. Wiley. 5, 24, 32, 33, 37, 38, 40, 58, 60, 67

[3] Bagdonavicius, V., Kruopis, J., and Nikulin, M. (2010b). *Nonparametric tests for complete Data*. ISTE and J. Wiley. 4

[4] Bagdonavicius, V. and Nikulin, M. (1994). Stochastic models of accelerated life. In Gutierrez, J. and Valderrama, M., editors, *Advanced Topics in Stochastic Modelling*. World Scient. 6

[5] Bagdonavicius, V. and Nikulin, M. (1995). Semiparametric models in accelerated life testing. Queen's Papers in Pure and Applied Mathematics, Queen's University, Kingston, Ontario, Canada. 6

[6] Bagdonavicius, V. and Nikulin, M. (2002). *Accelerated Life Models*. Chapman and Hall-CRC, Boca Raton. 6, 15, 62, 71, 72, 76, 79, 99

[7] Bagdonavicius, V. and Nikulin, M. (2011). Chi-square goodness-of-fit test for right censored data. *Int.J.Appl.Math.Stat.*, 24(SI-11A) :30–50. 24, 32, 38, 85, 89

[8] Bagdonavicius, V., Nikulin, M., and Tahir, R. (2010c). Chi-square goodness-of-fit test for right censored data. In *3rd International conference on accelerated life testing, reliability-based analysis and design*, pages 223–230, Clermont-Ferrand, France. 32

[9] Bagdonavičius, V. (1978). Testing the hypothesis of the additive accumulation of damages. *Probability Theory and its Applications*, 23(2) :403–408. 101, 104

[10] Bagdonavičius, V., Levuliene, R., and Nikulin, M. (2012). Chi-squared goodness-of-fit tests for parametric accelerated failure time models. *Communications in Statistics - Theory and Methods*. to appear. 94

[11] Bagdonavičius, V., Masiulaityle, I., and Nikulin, M. (2008a). *Statistical analysis of redundant system with one stand-by unit*, chapter In : Mathematical methods in Survival analysis, reliability and quality of life, pages 189–202. Eds. Huber C., Limnios N., Balakrishnan N., Messbah M., and Nikulin M. ISTE & Wiley, London. 8, 97, 99

[12] Bagdonavičius, V., Masiulaityle, I., and Nikulin, M. (2008b). Statistical analysis of redundant system with "warm" stand-by units. *Stochastics : An International Journal of Probability and Stochastical Processes*, 80(2-3) :115–128. 9, 97, 98, 99, 107, 108, 110

[13] Bagdonavičius, V., Masiulaityle, I., and Nikulin, M. (2009). *Asymptotic Properties of Redundant Systems Reliability Estimators*, chapter In : Advances in Degradation Modeling : Applications to Reliability, Survival Analysis, and Finance, pages 293–310. Eds. Nikulin M., Limnios N., Balakrishnan N., Kahle W., and Huber C. Birkhäuser, Boston, 1st edition edition. 9, 97, 99, 110

[14] Bagdonavičius, V., Masiulaityle, I., and Nikulin, M. (2010). *Parametric estimation of redundant system reliability from censored data*, chapter In : Mathematical and Statistical Methods in Reliability, pages 51–64. Eds. Balakrishnan, N. and Nikulin, M. and Rykov, V. Springer Verlag, Boston. 9, 97, 99, 110

[15] Bagdonavičius, V. and Nikulin, M. (1997). Transfer functional and semiparametric regression models. *Biometrika*, 84 :365–378. 101

[16] Balakrishnan, N., Leiva, V., and Léopez, J. (2007). Acceptance sampling plans from truncated life tests based on the generalized Birnbaum-Saunders distribution. *Communications in Statistics : Simulation and Computation*, 36 :643–656. 18, 106

[17] Balakrishnan, N., Leiva, V., Sanhueza, A., and Vilca, F. (2009). Estimation in the Birnbaum-Saunders distribution based on scale-mixture of normals and the em-algorithm. *SORT*, 33(2) :171–192. 18, 106

[18] Berkson, J. and Gage, R. (1950). Calculation of survival rates for cancer. In *Proceedings of the staff meetings of the Mayo clinic*, volume 25, pages 270–286. 5

[19] Billingsley, P. (1968). *Convergence of Probability Measures*. J.Wiley, New York. 40

[20] Birnbaum, Z. W. and Saunders, S. C. (1969). A new family of life distributions. *Journal of Applied Probability*, 6 :637–652. 17, 105

[21] Chernoff, H. and Lehmann, E. (1954). The use of maximum likelihood estimates in tests for goodness of fit. *Ann. Math. Statist.*, 25 :579–586. 4, 23, 28

[22] Cox, D. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society*, 34 :187–220. 6

[23] Cramer, U. (1946). *Mathematical Methods of statistics.* Princeton university press, New York. 23, 26

[24] Desmond, A. (1986). On the relationship between two fatigue-lide models. *Transactions on Reliability*, 35 :167–169. 18, 105

[25] Drost, F. (1988). Asymptotics for generalized chi-squared goodness-of-fit tests. *CWI Tracs, Amsterdam*, 48. Centre for Mathematics and Computer Sciences. 4, 23, 29

[26] Drost, F. (1989). Generalized goodness-of-fit tests for location scale models when the number of classes tends to infinity. *The Annals of Statistics*, 17(3) :1285–1300. 29

[27] Dzhaparidze, K. and Nikulin, M. (1982). Probability distributions of the kolmogorov and omega-square statistics for continuous distributions with shift and scale parameters. *Journal of Soviet Mathemetics*, 20 :2147–2164. 5

[28] Dzhaparidze, K. and Nikulin, M. (1992). On evaluation of statistics of chi-square type tests. *Problems of the Theory of Probability Distributions*, 12 :59–90. 4, 30

[29] Dzhaparidze, K. O. and Nikulin, M. S. (1974). On a modification of the standard statistic of Pearson. *Theory of probability and its applications*, 19 :851–852. 4, 23, 29, 30, 32

[30] Dzhaparidze, K. O. and Nikulin, M. S. (1995). On the computation of chi-square-type statistics. *Journal of Mathematical Sciences*, 75(5) :1910–1921. 28, 29

[31] Efron, B. (1988). Logistic regression, survival analysis, and the Kaplan-Meier curve. *J. Amer. Statist. Assoc.*, 83 :414–425. 14, 41

[32] Fisher, R. (1928). On a property connecting the chi-square measure of discrepancy with the method of maximum likelihood. In *Atti de Congresso Internazionale di Mathematici*, volume 6, pages 95–100, Bologna. 3, 23, 26

[33] Gavrilov, L. and Gavrilova, N. (1991). *The Biology of Life Span : A Quantitative Approach.* New York : Harwood Academic Publisher. 62

[34] Gavrilov, L. and Gavrilova, N. (2001). The reliability theory of aging and longevity. *Journal of Theoretical Biology*, 213(4) :527–545. 20, 61, 68

[35] Gavrilov, L. and Gavrilova, N. (2006). *Reliability Theory of Aging and Longevity*, chapter In : Handbook of the Biology of Aging, pages 3–42. Eds : Masoro, E.J. and Austad, S.N. Elsevier Academic Press, San Diego, CA, USA, 6th edition. 7, 13

[36] Gerville-Rache, L. and Nikulin, M. (2000). Analyse statistique du modèle de Makeham. *Revue Roumaine Math. Pure et Appl.*, 45(6) :947–957. 60, 61, 63, 64

[37] Gerville-Rache, L., Nikulin, M., and Tahir, R. (2012). On reliability approach and statistical inference in demography. *Journal of Reliability and Statistical Studies*, 5 :37–49. Issue Special. 61

[38] Gnedenko, B. V., Belyaev, Y. K., and Solov'ev, A. D. (1968). *Mathematical methodes in reliability theory*. New York : Academic Press, Inc. 8, 96, 97

[39] Gompertz, B. (1825). On the nature of the function expressive of the law of human mortality. *Philosophical Transactions of the Royal Society of London, Ser. A*, 115(6) :513–585. reprint : Haberman S and Sibbett, TA (1995), Vol. II, 119-191. 19, 60, 61

[40] Greenwood, P. and Nikulin, M. (1996). *A guide to Chi-squared Testing*. John Wiley and Sons, New York. 4, 5, 25, 29, 30, 40, 63, 83

[41] Habib, M. and Thomas, D. (1986). Chi-squared goodness-of-fit tests for randomly censored data. *The Annals of Statistics*, 14 :759–765. 5, 24, 32

[42] Hjort, N. (1990). Goodness of fit tests in models for life history data based on cumulative hazard rates. *The Annals of Statistics*, 18(3) :1221–1258. 5, 24, 33, 36

[43] Hollander, M. and Pena, E. (1992). Chi-square goodness-of-fit test for randomly censored data. *Journal of American Statistical Association*, 87(417) :458–463. 5, 32

[44] Hosmer, D., Lemeshow, S., and May, S. (2008). *Applied Survival Analysis : Regression Modeling of Time to Event Data*. John Wiley and Sons Inc., New York, NY, second edition. xii, 92, 136

[45] Hsuan, T. and Robson, D. (1976). The goodness-of-fit tests with moment type estimators. *Communications in Statistics -Theory and Methods*, A5 :1509–1519. 23, 27, 30

[46] Huber, C., Solev, V., and Vonta, I. (2006). Interval censored and truncated data : rate of convergence of npmle of the density. *Journal of Statistical Planning and Inference*. 5

[47] Huber-Carol, C., Limnios, N., Mesbah, M., and Nikulin, M., editors (2008). *Mathematical Methods in Survival Analysis, Reliability and Quality of Life*. Hermes(ISTE)/Wiley. 72

[48] Johnson, N., Kotz, S., and Balakrishnan, N. (1995). *Continuous Univariate Distributions*, volume 2. John Wiley, New York, 2nd edition. 12, 18, 106

[49] Jockett, D. and Rosenberg, B. (1993). Comparison of the Gompertz and Weibull functions as descriptors for human mortality distributions and their intersections. *Mechanisms of Ageing and Development*, 69(131). 61

[50] Kaplan, E. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of American Statistical Association*, 53 :457–481. 5

[51] Kim, J. (1993). Chi-square goodness-of-fit tests for randomly censored data. *The Annals of Statistics*, 21 :1621–39. 5, 33

[52] Kleinbaum, D. and Klein, M. (2005). *Survival Analysis : A Self learning text*. Springer : New York, 2nd edition. 6

[53] Kozlov, B. A. and Ushakov, I. A. (1970). *Reliabilty handbook*. Internation series in decision process. Holt, Rinehart and Winston, Inc. 8, 96, 97

[54] Kundu, D., Kannan, N., and Balakrishnan, N. (2008). On the hazard function of Birnbaum-Saunders distribution and associated inference. *Computational Statistics and Data Analysis*, 52 :2692–2702. 18, 106

[55] Lawless, J. (2003). *Statistical Models and Methods for Lifetime Data*. Wiley : New York. 6, 72, 93

[56] LeCam, L. (1956). Locally asymptotically normal families of distributions. *University of California publication in statistics*, 3 :37–98. 28

[57] LeCam, L., Mahan, C., and Singh, A. (1983). *An extension of a Theorem of H. Chernoff and E.L. Lehmann*, chapter In : Recent advances in statistics, pages 303–332. Academic Press, Orlando. 29

[58] Lemonte, A., Cribari-Neto, F., and Vasconcellos, K. (2007). Improved statistical inference for the two-parameter Birnbaum-Saunders distribution. *Computational Statistics and Data Analysis*, 51 :4656–4681. 18, 106

[59] Makeham, W. (1860). On the law of mortality and the construction of annuity tables. *Journal of the Institute of Actuaries*, 8 :301–310. 20, 60, 61

[60] McCulloch, C. E. (1985). Relationships among some chi-squared goodness of fit statistics. *Communications in Statistics-Theory and Methods*, 14 :593–603. 30

[61] Meeker, W. and Escobar, L. (1998). *Statistical Methods for Reliability Data*. John Wiley and Sons. 6, 72, 99

[62] Mirvaliev, M. (2001). *An investigation of generalized chi-squared type statistics*. PhD thesis, Academy of Sciences of the Republic of Uzbekistan, Tashkent. Doctoral dissertation. 5, 24, 30

[63] Moore, D. S. and Spruill, M. C. (1975). Unified large-sample theory of general chi-squared statistics for tests of fit. *The Annals of Statistics*, 3 :599–616. 29

[64] Mudholkar, G., Srivastava, D., and Kollia, G. D. (1996). A generalization of the Weibull distribution with application. *J. Amer. Statist. Assoc.*, 91 :1575–1583. 14, 41

[65] Nelson, W. (1990). *Accelerated Testing : Statistical Models, Test Plans, and Data Analyses*. John Wiley and Sons, New York. xii, 6, 137

[66] Nelson, W. (2004). *Accelerated Life Testing : Statistical Models, Test Plans, and Data Analysis*. Wiley-Interscience : New York, 2nd edition. 72, 76

[67] Nikulin, M. (1973a). Chi-square test for continuous distribution with shift and scale parameters. *Theory of Probability and its Applications*, 19(3) :559–568. 4, 23, 29, 32

[68] Nikulin, M. (1973b). Chi-square test for normality. In *International Vilnius Conference on Probability Theory and Mathematical Statistics*, volume 2, pages 119–122. 4, 23, 29, 32

[69] Nikulin, M. (1973c). On a chi-square test for continuous distribution. *Theory of Probability and its Applications*, 19(3) :638–639. 23, 29

[70] Nikulin, M. (1974). On a quantile criterion. *Theory of Probability and its Applications*, 19 :431–434. 29

[71] Nikulin, M., Anisimov, V., and Nikulin, A. (2011a). Statistical models of longevity, aging and degradation in demography, gerontology and oncology. *Advances in Gerontology*, 24(3) :366–379. 61

[72] Nikulin, M., Balakrishnan, N., Tahir, R., and Saaidia, N. (2011b). Modified chi-squared goodness-of-fit test for Birnbaum-Saunders distribution. In Lemeshko, B., Nikulin, M., and Balakrishnan, N., editors, *Applied methods of statistical analysis : Simulation and statistical inference*, pages 87–99, Novosibirsk, Russia. Publishing house of NSTU. 18

[73] Nikulin, M., Saaidia, N., and Tahir, R. (2011c). Reliability analysis of redundant systems by simulation for data with unimodal hazard rate functions. *Mathematics in Engineering, Science and Aerospace*, 2(3) :105–114. 9, 99, 114

[74] Nikulin, M. and Solev, V. (1999). *Chi-squared goodness of fit test for doubly censored data, applied in Survival Analysis and Reliability*, chapter In : Probabilistic and Statistical Models in Reliability (Eds. N. Limnios and D. Ionescu), pages 101–112. Birkhauser, Boston. 5

[75] Nikulin, M. and Tahir, R. (2011). Application of sedyakin's model and bs family for statistical analysis of redundant systems with one warm standby unit. *Journal of Mathematical sciences*, 396 :155–171. 99

[76] Nikulin, M. and Voinov, V. (2006). Chi-squared testing. In *Encyclopedia of Statistical Sciences*, volume 1 of *Eds. S. Kotz, N. Balakrishnan, C.B. Read, B. Vidakovich*, pages 912–921. John Wiley and Sons, Hoboken, New Jersey, 2nd edition. 5

[77] Nikulin, N. and Haghighi, F. (2006). A chi-squared test for the generalized power weibull family for the head-and-neck cancer censored data. *Journal of Mathematical Sciences*, 133(3) :1333–1341. 41

[78] Pearson, K. (1900). On the criterion that a given system of deviation from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Phil. Mag.*, 50 :157–175. 23

[79] Rao, k. and Robson, D. (1974). A chi-square statistic for goodness-of-fit tests within the exponential family. *Communication in Statistics*, 3 :1139–1153. 4, 23, 29, 30, 32

[80] Reijns, G. and Gemund, A. v. (2007). Reliability analysis of hierarchical systems using statistical moments. *IEEE Transactions On Reliability*, 56(3) :525–533. 8

[81] Ricklefs, R. and Scheuerlein, A. (2002). Biological implications of the Weibull and Gompertz model of aging. *Journals of Gerontology : Biological Sciences*, 57A(2) :B69–B76. 20, 61

[82] Roy, A. R. (1956). On χ^2-statistics with variable intervals. Technical report N1, Stanford University, Statistics Department, Palo Alto, CA. 23

[83] Saaidia, N., Nikulin, M., and Tahir, R. (2011). An application of the aft model based on the generalized Weibull distribution for reliability analysis of redundant systems. In *43rd Journée de Statistique*, Tunizia. Societe Franacaise de statistique. 9

[84] Saaidia, N., Tahir, R., and Seddik-Ameur, N. (2010). The inverse Gaussian distribution and its applications in accelerated life testing when the hazard rate function has $\bigcap$-shape. In *3rd International conference on accelerated life testing, reliability-based analysis and design*, pages 71–75, Clermont-Ferrand, France. 79

[85] Sedyakin, N. (1966). On one physical principle in reliability theory (in russian). *Techn. Cybernetics*, 3 :80–87. 8, 98, 101

[86] Singpurwalla, N. (1971). Inference from accelerated life tests when observations are obtained from censored data. *Technometrics*, 13 :161–170. 72

[87] Tahir, R. and Saaidia, N. (2012). A modified chi-squared goodness-of-fit test for the Birnbaum-Saunders distribution. In *44th Journée de Statistique*, Bruxelles. Societe Francaise de statistique. 30

[88] Tsiatis, A. (2006). *Semiparametric theory and missing data.* Springer Series in Statistics. New York : Springer. 6

[89] Turnbull, B. (1976). The empirical distribution function with arbitrary grouped, censored and truncated data. *Journal of the Royal Statistical Society*, 38 :290–295. 5

[90] Van der Vaart, A. (1998). *Asymptotic Statistics.* Cambridge, UK. 4, 23, 24, 29

[91] Vanfleteren, J., DeVreese, A., and Braeckman, B. P. (1998). Two-parameter logistic and Weibull equations provide better fits to survival data from isogenic populations of caenorhabditis elegans in axenic culture than does the Gompertz model. *Journal of Gerontology*, 53 :393–403. 62

[92] Voinov, N. and Nikulin, M. (1993). *Unbiased estimators and their applications*, volume 1. Kluwer Academic Publishers, Dordrecht. 19

[93] Voinov, V., Alloyarova, R., and Pya, N. (2008). *Recent achievements in modified chi-squared goodness of-fit testing*, chapter In : Statistical Models and Methods for Biomedical and Technical Systems, pages 245–262. (Eds. Vonta, F., Nikulin, M., Limnios, N., and Huber, C.). Birkhuser, Boston. 23

[94] Voinov, V., Nikulin, M., and Pya, N. (2007). *Independent χ_1^2 distributed in the limit components of some chi-squared tests*, chapter In : Recent Advance in Stochastic Modelling and Data Analysis, pages 243–250. Skiadis C.H., ed. World Scientific. 29

[95] Voinov, V., Nikuline, M., and Balakrishnan, N. (2012). *Chi-Squared goodness-of-fit tests with applications.* Wiley. to appear. 5, 30

[96] Voinov, V. and Pya, N. (2004). On the power of modified chi-squared goodness-of-fit tests for the family of logistic and normal distributions. In *7th Iranian Statististic Conference, Allameh-Tabatabaie University*, pages 385–403. 27, 30

[97] Voinov, V., Pya, N., and Alloyarova, R. (2009). A comparative study of some modified chi-squared tests. *Communications in Statistics : Simulation and Computation*, 38 :355–367. 4, 30

[98] Volodin, I. and Dzhungurova, O. (2000). On limit distribution emerging in the generalized Birnbaum-Saunders model. *Journal of Mathematical Science*, 99 :1348–1366. 18, 106

[99] Wang, Z., Desmond, A., and Lu, X. (2006). Modified censored moment estimation for the two-parameter Birnbaum-Saunders distribution. *Computational Statistics ans Data Analysis*, 50 :1033–1051. 18

[100] Wilson, D. (1994). The analysis of survival mortality data : fitting Gompertz, Weibull, and logistic functions. *Mechanisms of Ageing and Development*, 74 :15–31. 61

[101] Zhang, B. (1999). A chi-squared goodness-of-fit test for the logistic regression models based on case-control data. *Biometrika*, 86 :531–539. 4, 33

Appendix

1 Publications During PhD

1. Gerville-Réache L, Nikulin, M., Tahir, R. (2012). On reliability approach and statistical inference in demography. Journal of Reliability and Statistical Studies, Vol. 5, Issue Special, 37-49.
2. Tahir, R. and Saaidia, N. (2012). A modified chi-squared goodness-of-fit test for the Birnbaum-Saunders distribution. 44th Journée de Statistique, Bruxelles. 21-25 May.
3. Nikulin, M. and Tahir, R. (2011). Application of Sedyakin's model and BS family for statistical analysis of redundant systems with one warm standby unit. Journal of Mathematical Sciences, V. 396, 155-171.
4. Nikulin, M., Saaidia, N., and Tahir, R. (2011). Reliability analysis of redundant systems by simulation for data with unimodal hazard rate functions. Mathematics in Engineering, Science and Aerospace. Vol. 2, No. 3, 105-114.
5. Nikulin, M., Balakrishnan, N., Tahir, R., and Saaidia, N. (2011). Modified chi-squared goodness-of-fit test for Birnbaum-Saunders distribution. In : Applied methods of statistical analysis : Simulation and statistical inference, (editors : Lemeshko, B., Nikulin, M., and Balakrishnan, N.). Novosibirsk, Russia, 87-99.
6. Nikulin, M., Saaidia, N., and Tahir, R. (2011). An application of the aft model based on the log-normal distribution for reliability analysis of redundant systems. 9ème Congrès International QUALITA2011. Angers, France.
7. Saaidia, N., Nikulin, M., and Tahir, R. (2011). An application of the aft model based on the generalized weibull distribution for reliability analysis of redundant systems. 43rd Journée de Statistique, Tunizia.
8. Tahir, R. and Nikulin, M., (2010). Goodness-of-fit test for right censored samples : application on head and neck cancer data. 3rd International Conference on Statistical Sciences. Lahore Pakistan. (Oral Presentation).

9. Bagdonavicius, V., Nikulin, M., and Tahir, R. (2010). Chi-squared goodness-of-fit test for generalized Weibull distribution when data are right censored. 3rd International Conference on Accelerated Life Testing, Reliability-based Analysis and Design. Clermont-Ferrand, France.

2 Redundant System for Lognormal Distribution

Suppose that the distribution of failure times in hot and warm conditions is log-normal, i.e.

$$S_1(t) = 1 - F_1(t) = 1 - \Phi\left(\frac{\ln t - m}{\sigma}\right).$$

In the figure 2, we represent the trajectories of the parametric estimators $\hat{F}_1$, $\hat{K}_2$, $\hat{K}_3$ and $\hat{K}_4$ for the log-normal distribution.

The loglikelihood function

$$\ell(r, m, \sigma) = -n\left(\ln\sigma + \frac{m^2}{2\sigma^2}\right) + \frac{n_2 \ln r}{\sigma^2}\left(m - \frac{\ln r}{2}\right) + \left(\frac{m}{\sigma^2} - 1\right)\sum_{i=1}^{n_1} \ln(T_{1i}) +$$

$$\left(\frac{m}{\sigma^2} - \frac{\ln r}{\sigma^2} - 1\right)\sum_{j=1}^{n_2} \ln(T_{2j}) - \frac{1}{2\sigma^2}\sum_{i=1}^{n_1} \ln^2(T_{1i}) - \frac{1}{2\sigma^2}\sum_{j=1}^{n_2} \ln^2(T_{2j}),$$

where $n = n_1 + n_2$.

The score functions are

$$\frac{\partial \ell}{\partial r} = \frac{n_2}{r\sigma^2}(m - \ln r) - \frac{1}{r\sigma^2}\sum_{j=1}^{n_2} \ln(T_{2j}),$$

$$\frac{\partial \ell}{\partial m} = -\frac{nm}{\sigma^2} + \frac{n_2 \ln r}{\sigma^2} + \frac{1}{\sigma^2}\sum_{i=1}^{n_1} \ln(T_{1i}) + \frac{1}{\sigma^2}\sum_{j=1}^{n_2} \ln(T_{2j}),$$

$$\frac{\partial \ell}{\partial \sigma} = -n\left(\frac{1}{\sigma} - \frac{m^2}{\sigma^3}\right) - \frac{n_2 \ln r}{\sigma^3}(2m - \ln r) - \frac{2m}{\sigma^3}\sum_{i=1}^{n_1} \ln(T_{1i}) +$$

$$\frac{2}{\sigma^3}(-m + \ln r)\sum_{j=1}^{n_2} \ln(T_{2j}) + \frac{1}{\sigma^3}\sum_{i=1}^{n_1} \ln^2(T_{1i}) + \frac{1}{\sigma^3}\sum_{j=1}^{n_2} \ln^2(T_{2j}).$$

To find the estimator $\hat{\gamma}$ one can solve the system formed by equalizing the score functions to zero.

Second partial derivatives of the loglikelihood function are

$$\frac{\partial^2 \ell}{\partial r^2} = -\frac{n_2}{r^2\sigma^2}(m + \ln r + 1) + \frac{1}{r^2\sigma^2}\sum_{j=1}^{n_2} \ln(T_{2j}); \quad \frac{\partial^2 \ell}{\partial r \partial m} = \frac{n_2}{r\sigma^2};$$

$$\frac{\partial^2 \ell}{\partial r \partial \sigma} = \frac{2n_2}{r\sigma^3}(-m + \ln r) + \frac{2}{r\sigma^3}\sum_{j=1}^{n_2} \ln(T_{2j}); \quad \frac{\partial^2 \ell}{\partial m^2} = -\frac{n}{\sigma^2};$$

$$\frac{\partial^2 \ell}{\partial m \partial \sigma} = \frac{2nm}{\sigma^3} - 2n_2 \frac{\ln r}{\sigma^3} - \frac{2}{\sigma^3}\sum_{i=1}^{n_1}\ln(T_{1i}) - \frac{2}{\sigma^3}\sum_{j=1}^{n_2}\ln(T_{2j});$$

$$\frac{\partial^2 \ell}{\partial \sigma^2} = -n\left(-\frac{1}{\sigma^2} + \frac{3m^2}{\sigma^4}\right) + \frac{3n_2}{\sigma^4}\left(2m\ln r - \ln^2 r\right) + \frac{6m}{\sigma^4}\sum_{i=1}^{n_1}\ln(T_{1i}) +$$

$$\frac{6}{\sigma^4}(m - \ln r)\sum_{j=1}^{n_2}\ln(T_{2j}) - \frac{3}{\sigma^4}\sum_{i=1}^{n_1}\ln^2(T_{1i}) - \frac{3}{\sigma^4}\sum_{j=1}^{n_2}\ln^2(T_{2j}).$$

So the Fisher's information matrix is

$$I(\gamma) = \begin{pmatrix} \frac{n_2}{r^2\sigma^2} & -\frac{n_2}{r\sigma^2} & 0 \\ -\frac{n_2}{r\sigma^2} & \frac{n}{\sigma^2} & 0 \\ 0 & 0 & \frac{2n}{\sigma^2} \end{pmatrix}.$$

The inverse of the Fisher's information matrix is

$$I^{-1}(\gamma) = \begin{pmatrix} \frac{nr^2\sigma^2}{n_1 n_2} & \frac{r\sigma^2}{n_1} & 0 \\ \frac{r\sigma^2}{n_1} & \frac{\sigma^2}{n_1} & 0 \\ 0 & 0 & \frac{\sigma^2}{2n} \end{pmatrix}.$$

Taking $j = 2$, the c.d.f. $K_2(t)$ is estimated by

$$\hat{K}_2(t) = \frac{1}{\sqrt{2\pi}\hat{\sigma}}\int_0^t \frac{1}{y}\Phi\left(\frac{\ln(t + \hat{r}y - y) - \hat{m}}{\hat{\sigma}}\right) exp\left\{-\frac{\ln y - \hat{m}}{2\hat{\sigma}^2}\right\} dy.$$

$$\mathbf{C}_2(t,\gamma) = (\mathbf{C}_{21}(t,\gamma), \mathbf{C}_{22}(t,\gamma), \mathbf{C}_{23}(t,\gamma))^T,$$

$$\mathbf{C}_{21}(t,\gamma) = \int_0^t \frac{\partial F_1}{\partial r}(t + ry - y, m, \sigma) dF_1(y, m, \sigma),$$

$$\mathbf{C}_{22}(t,\gamma) = \int_0^t \frac{\partial F_1}{\partial m}(t + ry - y, m, \sigma) dF_1(y, m, \sigma) + F_1(t + ry - y, m, \sigma) d(\frac{\partial F_1}{\partial m}(y, m, \sigma)),$$

$$\mathbf{C}_{23}(t,\gamma) = \int_0^t \frac{\partial F_1}{\partial \sigma}(t + ry - y, m, \sigma) dF_1(y, \mu, \lambda) + F_1(t + ry - y, m, \sigma) d(\frac{\partial F_1}{\partial \sigma}(y, m, \sigma)).$$

The partial derivatives are

$$\frac{\partial F_1}{\partial r}(t + ry - y, m, \sigma) = \frac{y}{\sigma(t + ry - y)}\varphi\left(\frac{\ln(t + ry - y) - m}{\sigma}\right),$$

$$\frac{\partial F_1}{\partial m}(t+ry-y;m,\sigma) = -\frac{1}{\sigma}\varphi\left(\frac{\ln(t+ry-y)-m}{\sigma}\right),$$

$$\frac{\partial F_1}{\partial \sigma}(t+ry-y;m,\sigma) = -\frac{\ln(t+ry-y)-m}{\sigma^2}\varphi\left(\frac{\ln(t+ry-y)-m}{\sigma}\right),$$

we can write also

$$\mathbf{C}_{21}(t,\gamma) = \frac{1}{\sqrt{2\pi}}\int_0^t \frac{1}{\sigma^2(t+ry-y)} exp\left\{-\frac{(\ln y-m)^2}{2\sigma^2}\right\}\varphi\left(\frac{\ln(t+ry-y)-m}{\sigma}\right)dy,$$

$$\mathbf{C}_{22}(t,\gamma) = -\frac{1}{\sigma^2\sqrt{2\pi}}\int_0^t \frac{1}{y} exp\left\{-\frac{(\ln y-m)^2}{2\sigma^2}\right\}\varphi\left(\frac{\ln(t+ry-y)-m}{\sigma}\right)dy+$$

$$\frac{1}{\sigma^3}\int_0^t \Phi\left(\frac{\ln(t+ry-y)-m}{\sigma}\right)\left(\frac{\ln y-m}{y}\right)\varphi\left(\frac{\ln y-m}{\sigma}\right)dy,$$

$$\mathbf{C}_{23}(t,\gamma) = -\frac{1}{\sigma^3\sqrt{2\pi}}\int_0^t \frac{\ln(t+ry-y)-m}{\sigma^2} exp\left\{-\frac{(\ln y-m)^2}{2\sigma^2}\right\}\varphi\left(\frac{\ln(t+ry-y)-m}{\sigma}\right)dy+$$

$$\frac{1}{\sigma^2}\int_0^t \frac{1}{y}\Phi\left(\frac{\ln(t+ry-y)-m}{\sigma}\right)\left(\frac{(\ln y-m)^2}{\sigma^2}-1\right)\varphi\left(\frac{\ln y-m}{\sigma}\right)dy,$$

where $\varphi(.)$ is the density function of the standard normal distribution.

3 Data Sets

3.1 Bronchopulmonary dysplasia (BPD)-data, Hosmer and Lemeshow (2008)

id	surfact	ondays	censor	id	surfact	ondays	censor	id	surfact	ondays	censor
1	0	59	1	27	0	553	1	53	1	30	1
2	0	514	1	28	0	76	1	54	1	45	1
3	0	313	1	29	0	134	1	55	1	23	1
4	0	631	1	30	0	116	1	56	1	54	1
5	0	107	1	31	0	83	1	57	1	63	1
6	0	71	1	32	0	33	1	58	1	14	1
7	0	583	1	33	0	317	1	59	1	96	1
8	0	91	1	34	0	600	1	60	1	103	1
9	0	66	1	35	0	362	1	61	1	71	1
10	0	95	1	36	0	333	1	62	1	71	1
11	0	13	1	37	0	68	1	63	1	64	1
12	0	5	1	38	0	217	1	64	1	253	1
13	0	85	1	39	0	733	0	65	1	54	1
14	0	619	0	40	0	546	1	66	1	236	1
15	0	580	1	41	0	546	1	67	1	51	1
16	0	196	1	42	0	56	1	68	1	134	1
17	0	475	1	43	0	48	1	69	1	31	1
18	0	32	1	44	1	43	1	70	1	274	0
19	0	161	1	45	1	250	1	71	1	204	1
20	0	193	0	46	1	110	1	72	1	118	1
21	0	59	1	47	1	249	1	73	1	424	1
22	0	62	1	48	1	181	1	74	1	56	1
23	0	95	1	49	1	70	1	75	1	310	0
24	0	63	1	50	1	197	1	76	1	108	1
25	0	26	1	51	1	306	1	77	1	51	1
26	0	16	1	52	1	53	1	78	1	70	1

Variables in the data set are :

Variable	Description	Codes/Values
id	Study ID	1 - 78
surfact	Surfactant use	0 = No, 1 = Yes
ondays	Days in Oxygen	Days
censor	Censoring Indicator	1 = Off Oxygen, 0 = Still on Oxygen at Study End

3.2 Time to breakdown at each of voltage levels (Nelson((1990)))

Voltage Level (ν_i)	n_i	Breakdown Time T_i
26	3	5.79, 1579.52, 2323.7
28	5	68.85, 426.07, 110.29, 108.29, 1067.6
30	11	17.05, 22.66, 21.01, 175.88, 139.07, 144.12, 20.46, 43.40, 194.90, 47.30, 7.74
32	15	0.40, 82.85, 9.88, 89.29, 215.10, 2.75, 0.79, 15.93, 3.91, 0.27, 0.69, 100.58, 27.80, 13.95, 53.24
34	19	0.96, 4.15, 0.19, 0.78, 8.01, 31.75, 7.35, 6.50, 8.27, 33.91, 32.52, 3.16, 4.85, 2.78, 4.67, 1.31, 12.06, 36.71, 72.89
36	15	1.97, 0.59, 2.58, 1.69, 2.71, 25.50, 0.35, 0.99, 3.99, 3.67, 2.07, 0.96, 5.35, 2.90, 13.77
38	8	0.47, 0.73, 1.40, 0.74, 0.39, 1.13, 0.09, 2.38