1-1-2018

# Goodness-Of-Fit Test for Hazard Rate

Ralphntoine Vital

Goodness-of-fit test for hazard rate

By

Ralph-Antoine Vital

A Thesis
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Doctorate of Philosophy
in Statistics
in the Department of Mathematics and Statistics

Mississippi State, Mississippi

December 2018

Goodness-of-fit test for hazard rate

By

Ralph-Antoine Vital

Approved:

Prakash N. Patil
(Major Professor)

Jonathan R. Woody
(Committee Member)

Janice DuBien
(Committee Member)

Tung-Lung Wu
(Committee Member)

Qian Zhou
(Committee Member)

Mohammad Sepehrifar
(Graduate Coordinator)

Rick Travis
Dean
College of Arts & Sciences

Name: Ralph-Antoine Vital

Date of Degree: December 14, 2018

Institution: Mississippi State University

Major Field: Statistics

Major Professor: Dr. Prakash Patil

Title of Study: Goodness-of-fit test for hazard rate

Pages of Study: 100

Candidate for Degree of Doctorate of Philosophy

In certain areas such as Pharmacokinetic(PK) and Pharmacodynamic(PD), the hazard rate function, denoted by $\lambda$, plays a central role in modeling the instantaneous risk of failure time data. In the context of assessing the appropriateness of a given parametric hazard rate model, Huh and Hutmacher [22] showed that their hazard-based visual predictive check is as good as a visual predictive check based on the survival function. Even though Huh and Hutmacher's visual method is simple to implement and interpret, the final decision reached there depends on the personal experience of the user. In this thesis, our primary aim is to develop nonparametric goodness-of-fit tests for hazard rate functions to help bring objectivity in hazard rate model selections or to augment subjective procedures like Huh and Hutmacher's visual predictive check. Toward that aim two nonparametric goodness-of-fit (g-o-f) test statistics are proposed and they are referred to as chi-square g-o-f test and kernel-based nonparametric goodness-of-fit test for hazard rate functions, respectively. On one hand, the asymptotic distribution of the chi-square goodness-of-fit test for hazard rate functions is derived under the null hypothesis $H_0 : \lambda(x) = \lambda_0(x)\ \forall x \in \mathbb{R}^+$ as well as under the fixed alternative hypothesis $H_1 : \lambda(x) = \lambda_1(x)\ \forall x \in \mathbb{R}^+$. The results as expected are asymptotically similar to those of the usual Pearson chi-square test. That is, under the null hypothesis the proposed test converges to a chi-square distribution and under the fixed alternative hypothesis it converges to a non-central chi-square distribution. On the other hand, we showed that the power properties of the kernel-based nonparametric goodness-of-fit test for hazard rate functions are equivalent to those of the Bickel and Rosenblatt test, meaning the proposed kernel-based nonparametric goodness-of-fit test can detect alternatives converging to the null at the rate of $N^{\beta}, \beta < 1/2$, where $N$ is the sample size. Unlike the latter, the convergence rate of the kernel-base nonparametric g-o-f test is much greater; that is, one does not need a very large sample size for able to use the asymptotic distribution of the test in practice.

Key words: Goodness-of-fit test, Hazard rate function, Pearson chi-square test, Bickel-Rosenblatt test, Kernel-based nonparametric test, Smoothing parameter, Power function of the test, Pitman alternative.

DEDICATION

To my family.

## ACKNOWLEDGEMENTS

I would first like to express my deepest thanks to Dr. Patil, my advisor, who from the beginning gave me unprecedented support, council and recommendations for a better realization of the present research. More importantly, without his patience and valuable comments this dissertation would have been impossible to end.

In addition, I would like to extend my thanks to Dr. Woody, Dr. DuBien, Dr. Wu and Dr. Zhou for serving on my committee and for their support. Their feedback was incredibly helpful to me.

Special thanks to Dr. Miller for his invaluable help during the preparation of this dissertation.

Finally acknowledgement is also made to the colleagues and friends who helped and supported me during this process.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

## LIST OF SYMBOLS, ABBREVIATIONS, AND NOMENCLATURE

**PKPD** Pharmacokinetic and Pharmacodynamic

**ML** Maximum likelihood

**PDF** Probability density function

**CDF** Cumulative distribution function

**g-o-f** Goodness-of-fit

**HRF** Hazard rate function

**i.i.d** Independent and identically distributed

**B-R** Bickel and Rosemblatt

**BKN** Bagdonavičius, Kruopis and Nikulin

**ISE** Integrated square error

**MC** Monte Carlo

**df** Degree of freedom

$I(\cdot)$ Indicator function

$X, Y, Z, \cdots$ Random variables

$\mathbf{X}, \mathbf{Y}, \mathbf{Z} \cdots$ Random vectors

$A^{-1}$ Inverse of the matrix A

$A^{-}$ Generalized inverse of the matrix A

$EX$ The mean of the random variable $X$

$E(\mathbf{X})$ The mean of the random vector $\mathbf{X}$

$VarX$ The variance of the random variable $X$

$Var(\mathbf{X})$ The covariance matrix of the random vector $\mathbf{X}$

$E_0T$ The expected value of $T$ under the null hypothesis

$E_1 T$ The expected value of $T$ under the alternative hypothesis

$\mathbf{Z} = (Z_i)_{1 \le i \le k}$ A vector of length $k$

$\mathbf{Z}^t$ The transposed vector (i.e a vector-line)

$\lambda^0$ The hazard rate function associated with the probability density function $f^0$

$\lambda_0^0$ The hypothesized hazard rate function

$\mathbb{R}^+$ The set of positive real numbers

$Q^0$ The chi-square goodness-of-fit test for hazard rate functions

$N$ Sample size

$h$ Bandwidth parameter

$S_N(h)$ The Bickel and Rosenblatt test for hazard rate functions

$T_N(h)$ The kernel-based nonparametric goodness-of-fit test statistic for hazard rate functions

$N(0,1)$ The standard normal distribution

$N_n(0,I)$ The standard multi-normal distribution

$\chi_k^2$ Chi-square distribution with $k$ degree of freedom

$X_n \xrightarrow{d} N(0,1)$ Convergence in distribution ($n \to \infty$).

$X_n \xrightarrow{P} x$ Convergence in probability ($n \to \infty$.)

$X_n \sim N(0,1)$ $X_n$ asymptotically normally distributed with mean zero and standard deviation one

$X_n \sim \chi_k^2$ $X_n$ follows asymptotically a chi-square with $k$ degree of freedom.

$a_n \to x$ $a_n$ Converges to $x$ as $n \to \infty$.

**diag**$((e)_{1 \le i \le k})$ The diagonal matrix with diagonal terms equal to the elements of $(e)_{1 \le i \le k}$

$g'$ The first derivative of $g$

$\hat{\theta}$ An estimator of $\theta$

$\hat{\theta}_{ML}$ The maximum likelihood estimator of $\theta$

CHAPTER I

# INTRODUCTION

## 1.1 Background and Motivation

In survival analysis as well as in reliability theory, one of the other important functions, besides the survival function and the density function, is the hazard rate function. This is because the hazard rate function measures the instantaneous risk of failure of a subject, and, therefore, provides crucial information, which one finds very useful in survival studies. Moreover, in survival studies, sometimes researchers are interested in checking whether or not a specified hazard rate function can explain the rate of failure of a given data set over time or in testing if two or many subgroups in a population have the same hazard rate function.

The hazard rate function is defined as the probability that an item will fail immediately after time $t$ given that item has survived up to time $t$. To put this in another way, let $X^0$ be a positive random variable, which represents the lifespan of a component or a device or a person and let $f^0$ and $F^0$ denote, respectively, the probability density function (PDF) and the cumulative distribution function (CDF) of $X^0$. Then, the hazard rate function $\lambda^0(t)$ associated with the random variable $X^0$ is the probability that a component will die in the interval $[t, t + dt)$ given it was working before where $dt$ is very small. That is, the hazard rate function can be written as

$$\begin{aligned}\lambda^0(t) &= \lim_{dt\to 0}\frac{P[t \le X^0 < t+dt | X^0 \ge t]}{dt}\\ &= \lim_{dt\to 0}\frac{P[t \le X^0 < t+dt]}{dtP[X^0 \ge t]}\\ &= \frac{f^0(t)}{1-F^0(t)} \text{ for } F^0(t) < 1.\end{aligned}$$

In real life experiments it is not unusual for a participant (a component or a device or a person) of an experimental study to get relocated or removed, and hence lost to the study. In this case, let $X^0$ be a positive random variable that is now censored on the right by the i.i.d random variable $U$, which we assume to be independent of $X^0$. The hazard rate function is then given by

$$\begin{aligned}\lambda^0(t) &= \lim_{dt\to 0}\frac{P[t \le X^0 < t+dt | X^0 \ge t, U \ge t]}{dt}\\ &= \lim_{dt\to 0}\frac{P[t \le X^0 < t+dt]}{dtP[\{X^0 \ge t\} \cap \{U \ge t\}]}\\ &= \lim_{dt\to 0}\frac{P[t \le X^0 < t+dt]}{dtP[X^0 \ge t]}\\ &= \frac{f^0(t)}{1-F^0(t)} \text{ for } F^0(t) < 1.\end{aligned}$$

Observe that the final functional form of the hazard rate function in the censored case is equal to the one found in the uncensored case. This is a consequence of the random variable $X^0$ representing failure time and $U$ representing censoring time being independent of each other. Failure time data sets in which times of failure are independent of censoring times are referred to as independent right-censored data. It may be useful to note that $\Lambda^0(t) = \int_0^t \lambda^0(u)du$ is the cumulative hazard rate function, and thus the CDF of $X^0$ can be written as $F^0(t) = 1 - e^{\Lambda^0(t)}, t > 0$.

There are many areas where hazard rate functions are used to model the time at which an event of interest occurs. For instance, hazard rate functions are used in Pharmacokinetic and Pharmacodynamic (PKPD) to model the underlying relationship between time-to-event and drug exposure [22]. In the context of checking whether or not a specified hazard rate model is appropriate to explain the risk of failure of given time-to-event data, Huh and Hutmacher [22] showed through a graphical procedure that their selecting method based on nonparametric estimator for hazard rate functions is as efficient as methods based on nonparametric estimator for survival functions. In addition, their hazard-based visual method turns out to be a convenient and an appealing procedure for hazard rate model assessment [22]. Although their visual method is simple to implement and interpret, the final decision depends on the personal experience of the user. Therefore, the primary interest or aim in this thesis is to develop nonparametric goodness-of-fit tests for hazard rate functions to help bring objectivity into model selections or to augment subjective procedures like Huh and Hutmacher's visual predictive check. We achieve this aim by way of proposing a chi-square goodness-of-fit (g-o-f) test and a kernel-based nonparametric goodness-of-fit test for hazard rate functions.

In the next section, we give a brief introduction to parametric as well as to nonparametric methods used for hazard rate estimation. A brief survey of the parametric and nonparametric goodness-of-fit tests available in the literature is given in Section 1.3. The chi-square goodness-of-fit test for hazard rate functions is presented in Section 1.4. The construction of the kernel-based nonparametric goodness-of-fit test for hazard rate functions is exposed in Section 1.5 and finally an outline of the thesis is provided in Section 1.6.

### 1.2 Estimation Procedure for Hazard Rate Functions

In statistical inference, there exist two main approaches to estimate an unknown hazard rate function, and those are the parametric and the nonparametric approaches. The parametric approach is based on the assumption that the functional form of the hazard rate model is known up to the parameter; therefore, one only needs to estimate its parameter(s) to help draw inference on the unknown hazard rate function, whereas in the nonparametric approach the functional form of the hazard rate function is completely unknown.

#### 1.2.1 Parametric Hazard Rate Function Estimation

Maximum likelihood (ML) estimation is one of the most studied and used estimation procedures in statistics; therefore, other parametric hazard rate function estimation procedures will not be discussed in this thesis. Let $X_1^0, \cdots, X_n^0$ be independent and identically distributed (i.i.d.) failure times random variables that are censored on the right by the i.i.d. random variables $U_1, \cdots, U_n$, which are independent of the $X_i^0$'s. Then, observed failure time data sets are drawn from the pairs $(X_i, \delta_i), i = 1, \cdots, n$ where $X_i = \min\{X_i^0, U_i\}$ represents the time an event occurs and $\delta_i = I(X_i^0 \leq U_i)$ indicates if at that time the event is censored or not ($I(\cdot)$ is the indicator function). Assume that $x_1, \cdots, x_n$ are the realizations of the random sample $X_1, \cdots, X_n$. To convey as much information about the unknown parameter the likelihood function of $\theta$ is defined as

$$L(\theta) = \prod_{i=0}^{n} f^0(x_i;\theta)^{\delta_i} \bar{F}^0(x_i;\theta)^{1-\delta_i} \tag{1.1}$$

where $f^0(x_i;\theta)$ and $\bar{F}^0(x_i;\theta)$ are, respectively, the density and survival functions of the $X_i^0$'s with $\bar{F}^0(x_i;\theta) = 1 - F^0(x_i;\theta)$. Accordingly, when $\delta_i = 1$, the time of failure at

$x_i$ is observed with probability equivalent to $f^0(x_i;\theta)$, and when $\delta_i = 0$ then the time of failure at $x_i$ is unobserved with probability $\bar{F}^0(x_i;\theta) = 1 - F^0_\theta(x_i;\theta)$. Since $\bar{F}^0(x_i;\theta)$ can be written as $\exp[-\int_0^{x_i} \lambda^0(u,\theta)du]$, equation 1.1 could also be expressed as

$$L(\theta) = \prod_{i=0}^{n} \lambda^0(t_i;\theta)^{\delta_i} \exp[-\int_0^T \lambda^0(u;\theta)du], \tag{1.2}$$

where $\max\{x_i\}_{1\leq i\leq n} < T$. Therefore, the maximum likelihood estimator of $\theta$, denoted by $\hat{\theta}_{ML}$, is that value of $\theta$ which maximizes $L(\theta)$. Hence, parametric estimation of the hazard rate function is then $\lambda(t,\hat{\theta}_{ML})$ for $t > 0$. To find the maximum one can solve the equation $\dfrac{dL(\theta)}{d\theta} = 0$, which is also equivalent to solving $S(\theta) = 0$ where $S(\theta)$ is the first derivative of the log likelihood function and is referred to as the score function. Under certain regularity conditions, one can analytically derive the estimator of $\theta$ by solving the equation $S(\theta) = 0$, but when dealing with multivariate likelihood functions or incomplete data, it is natural to use a computational numerical method such as the Expectation-Maximization algorithm to solve $S(\theta;t) = 0$, see Patawan [30]. More detailed discussions on parametric estimation procedures for hazard rate functions can be found in Kalbfleish and Prentice [24] and Cox and Oakes [10].

### 1.2.2 Nonparametric Hazard Rate Function Estimation

When the distribution function $F^0(\cdot;\theta)$ is completely unknown, a nonparametric procedure is the only option that one has to estimate the hazard rate function associated with given failure time data. In the past few decades, many authors have proposed different types of nonparametric estimators for hazard rate functions. However, in this section, we discuss only the histogram hazard rate estimator.

Consider the pairs $(X_i, \delta_i), i = 1, \cdots, n$ given in the last subsection and a sequence of equally spaced points $x_{1:n} < x_{2:n} < \cdots < x_{k:n}$ over $\mathbb{R}^+$ with $\Delta = x_{j+1:n} - x_{j:n}$ for every $j \in \{1, \cdots, k-1\}$. For simplicity in the following paragraph we shall write $x_i$ instead of $x_{i:n}, i = 1, \cdots, k$. First, we define the statistics

$$f_1^0 = \#\{i : X_i \leq x_2 - \Delta/2 \text{ and } \delta_i = 1\}, \quad f_k^0 = \#\{i : X_i\delta_i > x_{k-1} + \Delta/2\},$$

$$f_j^0 = \#\{i : X_i\delta_i \in (x_j - \Delta/2, x_j + \Delta/2]\},\ 2 \leq j \leq k-1,$$

$$f_1 = \#\{i : X_i \leq x_2 - \Delta/2\}, \quad f_k = \#\{i : X_i > x_{k-1} + \Delta/2\},$$

$$f_j = \#\{i : X_i \in (x_j - \Delta/2, x_j + \Delta/2]\},\ 2 \leq j \leq k-1.$$

Then,

$$Y_j = \frac{1}{\Delta}\frac{f_j^0}{n - \sum_{l=1}^{j} f_l + 1},\ j = 1, 2, \cdots, k \tag{1.3}$$

is the histogram hazard rate estimator of $\lambda(x)$ for $x \in I_j = [x_j - \Delta, x_j + \Delta)$.

Note that $f_j^0$ is the number of items that have failed in the interval $I_j$ while $n - \sum_{l=1}^{j} f_l$ represents the number of items that have survived up to time $x_j$. Therefore, ignoring the width $\Delta$ one should consider the ratio $f_j^0$ to $n - \sum_{l=1}^{i} f_l$ as an estimator of the hazard rate function in the interval $I_j$. However, $n - \sum_{l=1}^{j} f_l$ is replaced by $n - \sum_{l=1}^{j} f_l + 1$ to avoid a potential division by zero. It can be shown that for sufficiently large $n$, the statistic $Y_j$ can be modeled as a normal random variable with mean $\lambda(x_j)$ and variance $\frac{1}{n\Delta}\frac{\lambda(x_j)}{G(x_j)}$ where $G(x_j) = (1 - F(x_j))(1 - H(x_j))$ and $H$ is the distribution function of $U_i$'s, see Patil and Bagkavos [29]. It is worth noting that the histogram hazard rate estimator presented in this subsection is a generalization of the hazard rate estimator studied by Watson and Leadbetter [39, 40].

There exist other nonparametric techniques that can be used to estimate the hazard rate function. To name a few, Liu and Van Ryzin [27] used a "nearest neighborhood" approach to propose a variant of the histogram hazard rate estimator, whereas kernel-based hazard rate estimators have been studied by numerous authors including Blum and Susarla [7], Tanner and Wong [34], Yandell [41], Burke [8], and Diehl and Stute [12] among others.

### 1.3 Goodness-of-fit Test for Hazard Rate Functions

In the literature, various goodness-of-fit test statistics have been proposed in order to check the adequacy of hazard rate functions. In this section, we present three of the main approaches used in statistical inference. The first approach is based on the ML test statistic, the second is the partial likelihood approach and the third is a Pearson-type goodness-of-fit test for hazard rate functions.

#### 1.3.1 Parametric Goodness-of-fit Test for Hazard Rate Functions

Recall from the last section that $L(\theta)$ and $S(\theta)$ denote, respectively, the likelihood function and the score function associated with the unknown parameter $\theta$. Now let $\mathcal{I}(\theta) = E(-\frac{\partial^2}{\partial\theta^2}\log L(\theta))$ denote the *expected Fisher information* and let $\hat{\theta}_{ML}$ be the maximum likelihood estimator of $\theta$. When the parameter space is one-dimensional, it is easy to establish under the null hypothesis $H_0 : \theta = \theta_0$ and certain regularity conditions that the statistics $(\hat{\theta}_{ML} - \theta)/\sqrt{\mathcal{I}(\theta)}$ and $(\hat{\theta}_{ML} - \theta)^2/\mathcal{I}(\theta)$ converge, respectively, to the standard normal distribution and the chi-square distribution with one degree of freedom.

To illustrate the implementation of the ML test procedure we consider the null hypothesis

$$H_0 : \theta = \theta_0$$

where $\theta = (\alpha, \beta)$ is the parameter of the Weibull hazard rate function, which functional expression is written as $\lambda(t, \theta) = \alpha \exp(\beta t)$ for $t > 0$. Given observed sample data $x_1, \cdots, x_n$ and using equation 1.2 the likelihood function of $\theta$ can be written as

$$L(\theta) = \exp[-\frac{\alpha}{\beta}(e^{\beta T} - 1)] \prod_{i=1}^{n} (\alpha e^{\beta x_i})$$

and the score function is given by

$$S(\theta) = -\frac{\alpha}{\beta}(e^{\beta} - 1) + n \log(\alpha) + \beta \sum x_i.$$

Without going into technicalities concerning the existence of the minimum, the solution $\hat{\theta}_{ML}$ such that $S(\hat{\theta}) = 0$ exists (see Patawan [30]), and we denote by $\hat{\alpha}_{ML}$ and $\hat{\beta}_{ML}$ the respective ML estimates of the parameters $\alpha$ and $\beta$. Hence, the maximum likelihood test statistic is given by

$$M = n(\hat{\theta}_{ML} - \theta_0)^t \mathcal{I}(\theta_0)(\hat{\theta}_{ML} - \theta_0).$$

Under $H_0$ the test statistic, $M$, converges in distribution to a $\chi^2_2$.

In survival analysis, it is also common to encounter hazard rate models in the form of $\lambda_p(t; x) = \lambda_0(t) \exp(Z^t \beta)$ where $t$ is the time, $x$ is the covariate, $Z$ is the derived covariate, $\beta$ is the parameter of the model, and $\lambda_0$ is a baseline hazard rate. The hazard rate model $\lambda_p$ is referred to as the proportional hazard rate model, and in order to test the null hypothesis

$$H_0 : \beta = \beta_0$$

Cox [9] in 1972 introduced the partial likelihood approach where the likelihood function of $\beta$ is defined as

$$L(\beta) = \prod_{j=1}^{n} \frac{\exp[Z_j(t_j)^t\beta]}{\sum_{l\in R(t_j)} \exp[Z_l(t_j)^t\beta]}.$$

Here, $t_j, j = 1, \cdots, n$ is a failure time data set and $R(t)$ is the set of items at risk of failure at time $t^-$, just prior time $t$. The partial likelihood function is based on the fact that the quotient $\frac{\exp[Z_j(t_j)^t\beta]}{\sum_{l\in R(t_j)} \exp[Z_l(t_j)^t\beta]}$ can be interpreted as the probability that an individual $j$ at time $t_j$ fails conditionally on the risk set $R(t_j)$ where the risk set is all the people that have survived up to $t_j$. Then, with score function $S(\beta) = \partial \log L/\partial\beta$ and *expect Fisher information matrix* $\mathcal{I}(\beta) = E(-\frac{\partial^2 \log L}{\partial\beta\partial\beta'})$, Cox showed that under certain regularity conditions and under $H_0 : \beta = \beta_0$

$$S(\beta_0)^t\mathcal{I}(\beta_0)^{-1}S(\beta_0) \sim \chi^2_p$$

where $p$ is the dimension of the parameter space.

As one would expect, test statistics based on the maximum likelihood method provide consistent and powerful tests against local alternatives compared to nonparametric goodness-of-fit tests. However, they are less efficient against global alternative hypotheses. The reason is that if the parametric model that is assumed for the data is true, the derived ML test statistic has greater performance than nonparametric g-o-f tests, but when the model is false, the ML test statistic leads to meaningless results.

### 1.3.2 Nonparametric Goodness-of-fit Test for Hazard Rate Functions

Let

$$H_0 : \lambda(x;\theta) = \lambda_0(x;\theta) \text{ for every } x \in \mathbb{R}^+ \text{ and } \theta \in \Theta \subset \mathbb{R}^m.$$

To test the null hypothesis $H_0$, one can consider $(X_i, \delta_i), i = 1, \cdots, n$ the right-censored failure time data defined in Section 1.2. Letting

$$N_i(x) = I(X_i \leq x, \delta_i = 1), Y_i(x) = I(X_i \geq x),$$

$$N(x) = \sum_{i=1}^{n} N_i(x) \text{ and } Y(x) = \sum_{i=1}^{n} Y_i(x),$$

and also considering the interval $I = [0, \tau]$ where $\tau$ is the time limit of the process. Then, one can divide the interval $I$ into $k$ subintervals $I_j = (a_j, a_{j+1}], j = 1, \cdots, k$ where $a_1 = 0$ and $a_{k+1} = \tau$. Moreover, with $U_j = N(a_{j+1}) - N(a_j)$ being the number of observed failures in the $j$-th class interval, $j = 1 \cdots, k$, under $H_0$ and certain regularity conditions it can be shown that $EN(x) = E\int_0^x \lambda(u, \theta)Y(u)du$. Meaning one can "expect" that the number of observed failures in the $I_j$ interval to be $\int_{I_j} \lambda(u, \hat{\theta}_{ML})Y(u)du$ where $\hat{\theta}_{ML}$ is the ML estimate of $\theta$. Similar to the usual Pearson chi-square g-o-f test, one can construct a chi-square goodness-of-fit test based on the difference between the number of observed and expected failures in the intervals $I_j, j = 1, \cdots, k$; that is, by setting

$$\mathbf{Z} = (\frac{1}{\sqrt{n}}(U_j - e_j)_{1 \leq j \leq k},$$

see Bagdonavičius, Kruopis and Nikulin [4].

Under certain regularity conditions, Bagdonavičius and Nikulin [3] showed that the stochastic process

$$H_n(t) = \frac{1}{\sqrt{n}}(N(t) - \int_0^t \lambda(u, \hat{\theta})Y(u)du) \tag{1.4}$$

converges in distribution to $H$ on $D[0, \tau]$ where $H$ is a zero mean Gaussian martingale such that for all $0 \leq s \leq t$ cov$(H(s), H(t)) = A(s) - C(s)^t I^{-1}(\theta_0)C(t)$, and $D[0, \tau]$ is the space of cadlag functions with Skorokhod metric [36]. More details on the form of the matrices $A$, $C$ and $I$ can be found in Bagdonavičius and Nikulin [3].

Letting $V = A - C^t I^{-1}(\theta_0)C$; Bagdonavičius and Nikulin showed that the statistic

$$Y^2 = Z^t \hat{V}^- Z \tag{1.5}$$

converges in distribution to a chi-square with degree of freedom $r = \text{rank}(V^-)$ where $V^-$ is the generalized inverse matrix of $V$. Observe that since the null hypothesis $H_0 : \lambda(x;\theta) = \lambda_0(x;\theta)$ for every $x \in \mathbb{R}^+$ is composite, the covariance-matrix $V$ is replaced by its ML estimate $\hat{V}$.

Akritas [1] was the first to propose such a procedure for right-censored failure time data in order to test the null hypothesis $H_0 : f^0(x) = f_0^0(x), \forall x \in \mathbb{R}$. Later, Hjort [21] augmented Akritas' goodness-of-fit test procedure to check the appropriateness of the proportional hazard rate model to given failure time data. Following Hjort's work, Bagdonavičius and Nikulin [2] extended the chi-square goodness-of-fit test procedure to test the adequacy of a myriad of parametric hazard rate models, including the accelerated failure time, the proportional hazards, the generalized proportional hazards, the frailty models, the transformation models, and models with cross effects of survival functions [3], to given failure times data . Lastly, in 2010, Bagdonavičius, Levuliene and Nikulin developed a variant of their Pearson-type g-o-f test for hazard rate functions in the context of accelerated failure time models with covariates [2].

## 1.4 A proposed Chi-square Goodness-of-fit Test for Hazard Rate Functions

Hereinafter, the hazard rate functions considered are denoted by $\lambda^0$ where

$$\lambda^0(x) = \frac{f^0(x)}{1 - F^0(x)}, \text{ for } F^0(x) < 1.$$

The null hypothesis of interest is then written as

$$H_0 : \lambda^0(x) = \lambda_0^0(x) \text{ for every } x \in \mathbb{R}^+.$$

To test $H_0$, one can use the Pearson-type goodness-of-fit test described in equation (1.5); however, one also needs to compute the covariance matrix $V$, which can be cumbersome for non-trivial hazard rate models

On the other hand, many authors have studied the asymptotic behavior of the histogram hazard rate estimator for censored as well as for uncensored data [27, 39, 40]. Until now there is not a goodness-of-fit test for hazard rate functions based on the histogram hazard rate estimator. Therefore, our first objective is to propose and to study a chi-square goodness-of-fit test for hazard rate functions based on the histogram hazard rate estimator.

First, the domain of the hazard rate function is discretized into $k$ disjoint sets of intervals $I_j$ for $j \in \{1, \cdots, k\}$. Then, the hazard rate estimate of $\lambda^0(x)$ for $x \in I_i$ is computed using equation 1.3. Finally, we propose the statistic defined as

$$Q^0 \quad = \mathbf{Z^0}^t W^- \mathbf{Z^0} \tag{1.6}$$

where $W$ represents the covariance matrix of $\mathbf{Z^0}^t = (Y_1 - EY_1, \cdots, Y_k - EY_k)$ to test $H_0$. Recall that $Y_i$ is defined in Section 1.2. One can note that the statistic $Q^0$ measures the discrepancy between an observed failure time data and the hypothesized hazard rate function. A complete analysis of the performance and the power properties of $Q^0$ is given in Chapter 2.

## 1.5 A proposed Kernel-based Nonparametric Goodness-of-fit Test for Hazard Rate Functions

Often, nonparametric goodness-of-fit test statistics are based on the integrated square error (I.S.E.) functional of the hypothesized function of interest. In 1973, Bickel and Rosenblatt proposed one of the first nonparametric goodness-of-fit tests based on the I.S.E. approach to check whether or not a density model is appropriate to describe the variation within a given data set. In 1975, Rosenblatt [31], extended the B-R test to the two-dimensional case. Later, using a Central Limit Theorem for degenerate U-statistics and under mild conditions, Hall [18] established the normality convergence of the test statistic $I_n(h) = \int (\hat{f} - f)^2 dx$ in the setting of multivariate random variables where $\hat{f}$ is a second order kernel estimator of $f$.

In this section, our aim is to develop a kernel-based nonparametric goodness-of-fit test for hazard rate functions based on the I.S.E. approach. But before that we note that properties of $L_2$ error based test statistics have also been discussed in the settings of regression models by Hall [19] and Hardle [20]. Following Hall's work, Bagkavos $et\,al.$ [5] proposed a nonparametric goodness-of-fit test statistic capable of checking the adequacy of the survival function to given failure times data.

With all the considerations above, a natural nonparametric goodness-of-fit test statistic to test the null hypothesis $H_0 : \lambda^0(x) = \lambda_0^0(x)$ for every $x \in \mathbb{R}^+$ can be defined as

$$S_n(h) = \int [\hat{\lambda}_n(x) - E\lambda_n(x)]^2 dx \tag{1.7}$$

where $\hat{\lambda}_n(x) = \sum_{i=1}^n \frac{1}{nb_n} K((x - X_i)/b_n)/(1 - \hat{F}(X_i))$ with $\hat{F}(u) = \frac{1}{n+1} \sum_{i=1}^n I[X_i \leq u]$. Under similar conditions found in Bickel and Rosenblatt [6] and for $\delta \in (0, 1/4)$, one can show that

$$Z_n(\lambda_0) = n^{\delta/2} \sigma_0^{-1} (S_n(h) - \mu_0) \xrightarrow{d} N(0,1) \text{ as } N \to \infty \tag{1.8}$$

where $\mu_0 = \int K^2(x)dx \times \int \frac{\lambda_0(x)}{1-F(x)} a(x)dx$ and $\sigma_0^2 = 2\int [\int K(x+y)K(x)dx]^2 dy \times \int \frac{\lambda_0^2(x)}{(1-F(x))^2} a^2(x)dx$. However, although the test statistic $S_n(h)$ defined above shows good power properties against the Pitman alternatives considered in Ghosh and Huang [16], $S_n(h)$ also converges in distribution to a normal variable at a very slow rate, meaning one needs a very large "finite" sample for the asymptotic distribution to play a role in practice. Therefore, in the following section, a second kernel-based nonparametric goodness-of-fit test for hazard rate functions is proposed based on a different rationale.

Recall that the density and distribution functions of the failure time random variables $X_i^0 > 0,\ i = 1, \cdots, N$ are denoted by $f^0$ and $F^0$, respectively; in this section, the sample size is denoted by $N$. Moreover, we denote the distribution function of the $U_i$'s by $H$ where $H$ is continuous. For a given sample size $N$, we let $x_{1:N} < x_{2:N} < \cdots < x_{n:N}$ be a sequence of equally spaced points such that $\Delta = x_{i+1:N} - x_{i:N}$. For simplicity, we write $x_i$ instead $x_{i:N}$ for $i = 1, \cdots, n$. We define as in Section 1.2 the statistics

$$f_1^0 = \#\{i : X_i \leq x_2 - \Delta/2 \text{ and } \delta_i = 1\}, \quad f_n^0 = \#\{i : X_i\delta_i > x_{n-1} + \Delta/2\},$$

$$f_j^0 = \#\{i : X_i\delta_i \in (x_j - \Delta/2, x_j + \Delta/2]\},\ 2 \leq j \leq n-1,$$

$$f_1 = \#\{i : X_i \leq x_2 - \Delta/2\}, \quad f_n = \#\{i : X_i > x_{n-1} + \Delta/2\},$$

$$f_j = \#\{i : X_i \in (x_j - \Delta/2, x_j + \Delta/2]\},\ 2 \leq j \leq n-1.$$

Hence the histogram estimator of the hazard rate function is given by

$$Y_i = \frac{1}{\Delta} \frac{f_i^0}{N - \sum_{j=1}^{i} f_j + 1},\ i = 1, 2, \cdots, n. \tag{1.9}$$

The construction of the histogram hazard rate estimator in this section is similar to the one found in Section 1.2; in addition, we assume that $\Delta = cN^{-\alpha}$ for some $c$ positive and $1/2 < \alpha < 1$. This rassumption is one of the fundamental differences between the two nonparametric g-o-f test statistics proposed in this thesis.

To construct our kernel-based nonparametric g-o-f test statistic for hazard rate functions, first, we consider an $h$-neighborhood of $x_i$, for some $h > 0$. Then, to test the null hypothesis $H_0 : \lambda^0(x) = \lambda_0^0(x)$ for every $x \in \mathbb{R}^+$ one can compute the statistic

$$T_{x_i} = \sum_{j=1, j\neq i}^{n} I(|x_j - x_i| < h)(Y_j - \lambda_0^0(x_j)), i = 1, 2, \cdots, n.$$

Observe that if one let both $\Delta$ and $h$ converge to zero at appropriate rates, which are defined later in Chapter 3; then under $H_0$ it follows that $E(T_{x_i}) \to 0$ as $N \to \infty$. That is, if $H_0$

is not true, $E[T_{x_i}]$ will be away from zero. Thus, in order to collect evidence against the null hypothesis across the domain of the hazard rate function, we consider a weighted and appropriately scaled sum of $T_{x_i}$ :

$$\tilde{T}_N(h) = \frac{1}{n(n-1)h} \sum_{i=1}^{n} u(x_i) T_{x_i}$$

where $u(x_i) = (Y_i - \lambda_0^0(x_i))$; clearly, when the null is true, this sum is expected to be close to zero and away from zero if not.

One can observe that the construction of the test first consists of testing for mean equality at different special locations (i.e. the centers), and second it also involves a summation of all those tests in order to collect evidence across the domain of the hazard rate function. Similar test procedures are used in nonparametric g-o-f test for regression functions, see Guerre and Lavergne [17].

## 1.6 Outline of the Thesis

In Chapter 2, we introduce the chi-square goodness-of-fit test for hazard rate functions presented in Section 1.4 by outlining some of the motivations behind this particular type of test. In addition, we derive its limiting distribution under the null hypothesis $H_0 : \lambda^0(x) = \lambda_0^0(x)$ for every $x \in \mathbb{R}^+$ where $\lambda_0^0$ is completely known and also under the null hypothesis $H_0 : \lambda^0(x) = \lambda_0^0(x;\theta)$ for every $x \in \mathbb{R}^+$ and $\theta \in \Theta$ where $\Theta$ is the parameter space of the unknown parameter $\theta$. It turns out that the limiting distribution of the test statistic under both types of null hypotheses is a chi-square distribution just as in the usual Pearson chi-square goodness-of-fit test, whereas against fixed alternative hypotheses it converges in distribution to a non-central chi-square distribution. Furthermore, simulation studies are carried out at the end of Chapter 2 in order to investigate the performance and power properties of the chi-square goodness-of-fit test for hazard rate functions.

In Chapter 3, the asymptotic distribution of the kernel-based nonparametric goodness-of-fit test for hazard hazard rate functions, constructed in Section 1.5, is established under the null hypothesis $H_0 : \lambda^0(x) = \lambda_0^0(x)$ for every $x \in \mathbb{R}^+$ using the Central Limit Theorem for degenerate U-Statistics introduced by Hall [18]. For analytical purposes we also study the asymptotic distribution of the kernel-based nonparametric g-o-f test statistic under the alternative hypothesis $H_1 : \lambda^0(x) = \lambda_1^0(x)$ for every $x \in \mathbb{R}^+$ where $\lambda_0^0 \neq \lambda_1^0$. Lastly, the power functions of the test statistic are given against the two Pitman alternatives found in Ghosh and Huang [16].

In Chapter 4, a series of simulation studies are conducted in order to evaluate the behavior of the kernel-based nonparametric goodness-of-fit test for hazard rate functions in finite-samples. First, we compare plots of the kernel density estimate of the density function of the test at different sample sizes to the plot of the standard normal density function. At a qualitative level, it is harder for the proposed test to distinguish local alternative hypotheses that are converging to the null hypothesis; therefore, we investigate the power properties of the test first when the null and alternative functions are derived from the same family of distribution functions and second when the alternative is a Pitman alternative. Furthermore, in Chapter 4 an application of the test to real life data is included in our series of simulation studies and we also compare our kernel-based nonparametric g-o-f test with the Pearson-type g-o-f test presented in section 1.3.

In nonparametric kernel testing theory, the choice of the smoothing parameter is paramount not only because it plays a central part in the asymptotic convergence of the test under the null but ill-chosen parameters can produce critical values that may lessen the performance of the test. To study this issue, in Chapter 4, we carry out a series of simulations with different values of the smoothing parameter and study the influence of the smoothing parameter on the performance and power of the test. Moreover, a Monte Carlo procedure is implemented in order to select objectively the smoothing parameter $h,$ and we also examine its influence on the observed critical values of the kernel-based nonparametric goodness-of-fit test for hazard rate functions.

CHAPTER II

# A CHI-SQUARE GOODNESS-OF-FIT TEST FOR HAZARD RATE FUNCTIONS

## 2.1 Introduction

In this chapter, a chi-square goodness-of-fit (g-o-f) test is proposed to check the adequacy of an hypothesized hazard rate function. It involves replacing the observed and expected frequencies in the usual Pearson chi-square g-o-f test by the observed histogram hazard rate defined in equation 1.3 and its expected value, respectively. The chi-square g-o-f test for hazard rate functions is motivated by three observations. The first one arises from the fact that in survival analysis as well as in reliability theory, more often than not, the time of occurrence of certain events is discretized for simplification or because of the circumstances. For example, it is at times more informative to have hazard rate models in terms of the number of tasks completed before failure occurs instead of the age of the item. To be more precise, in certain economic studies it is not uncommon to have information available only in terms of days or months or years (e.g. the number of days unemployed before returning to the work or the length of a strike in days); see Karlis & Patilea and Tutz & Pritscher [25, 35] for more on discrete hazard rate models. In certain survival studies, failure time data are collected by way of aggregation since it can be very difficult to monitor certain experiments continuously; one example is that a person time of failure, usually, is given in term of number of years, see Wang, Muller and Capra [38]. Secondly, it is easier and simpler to implement a chi-square g-o-f test for hazard rate models similar to the usual Pearson chi-square g-o-f test instead of using test statistics such as the kernel-based nonparametric g-o-f test defined in Section 1.5. Lastly, Huh and Hutmacher [22] showed that an hazard-based visual predictive check for hazard rate model selection is as good as one based on a survival estimator. Although their approach is easy to implement in the context of hazard rate selections, the conclusion reached depends on the experience of the user. Thus, by proposing a chi-square g-o-f type for hazard rate functions our aim is to add objectivity to the visual procedure of Huh and Hutmacher [22] and to propose an easy-to-implement g-o-f test for hazard rate functions.

With aforementioned considerations, in Section 2.2 we first define the chi-square g-o-f test statistic for hazard rate functions, and then study its limiting distribution in the settings of right-censored data. The limiting distribution of the proposed test statistic is derived in two different cases. First, when the hypothesized hazard rate function is completely known and second, when the hypothesized function is known up to the parameter; that is, testing $H_0 : \lambda^0(x) = \lambda^0(x;\theta)$ for $x \in \mathbb{R}^+$ where $\theta$ is unknown. In all the cases considered above, the derivation of the distribution of the chi-square g-o-f test for hazard rate functions is established through an application of the Central Limit Theorem for multivariate random variables. The power properties of the test are investigated in Section 2.3. Lastly, a simulation study is carried out in Section 2.4, followed by the proofs in Section 2.5.

## 2.2 The Chi-square Test Statistic and its Asymptotic Distribution

Recall from Chapter 1 that $X_1^0, \cdots, X_n^0$ are independent and identically distributed failure times random variables that are censored on the right by the i.i.d random variables $U_1, \cdots, U_n$, which are independent from the $T_i$'s. The distribution function of the $X_i^0$'s and the $U_i$'s are denoted by $F^0$ and $H$, respectively. In addition, we consider the pairs $(X_i^0, \delta_i), i = 1, \cdots, n$ where $X_i = \min\{X_i^0, U_i\}$ and $\delta_i = (X_i^0 \leq U_i). i = 1, \cdots, n$, and we denote by $F$ the cumulative distribution function of the $X_i$'s.

### 2.2.1 Hypothesized Function Completely Known

Since hazard rate estimates at a given point $x$ are expected to be highly unstable whenever $x$ is in the extreme right tail, we restrict our chi-square g-o-f type test on the interval $[0, T]$ where $T = \sup\{t; F(t) < 1 - \epsilon\}$ for some small $\epsilon > 0$, in order to test the null hypothesis

$$H_0 : \lambda^0(x) = \lambda_0^0(x) \text{ for every } x \in \mathbb{R}^+. \tag{2.1}$$

Then, similar to the usual Pearson g-o-f test, the interval $[0, T]$ is divided into $k$ subintervals $I_i = [a_i, a_{i+1})$, for $i = 1, \cdots, k$ with $a_1 = 0$, $a_{k+1} = T$. In addition, the center of the subintervals $I_i$ is denoted by $x_i, i = 1, \cdots, k$. Given the subintervals, $I_i, i = 1, \cdots, k$, the histogram hazard rate estimator of $\lambda(x)$ at $x = x_i$, as defined in equation 1.3, is written as $q^0(x_i) = q_{ni}^0 = \frac{f_i^0}{\Delta_i(n-m_i+1)}$ for $1 \leq i \leq k$ where $m_i = \sum_{j=1}^n f_j$. For simplicity, in this chapter, $\Delta_i = \Delta$ for every $i, i = 1, \cdots, k$.

Let $f_0^0$ be the density function associated with the hypothesized hazard rate function $\lambda_0^0$; then, we set $\pi_i = \int_{I_i} f_0^0(x)dx$, $\Pi_i = \sum_{j=1}^{i} \pi_i$ and $\Gamma_i = P(U_i \leq x_i)$ for $i = 1, \cdots, k$. It can be shown that under the null hypothesis (see Patil and Bagkavos [29])

$$Eq_{ni}^0 = \frac{\pi_i}{\Delta(1-\Pi_i)} + o(1/n),$$

$$\text{var}q_{ni}^0 = \frac{\pi_i}{\Delta^2(1-\Pi_i)(1-\Gamma_i)}(1 - \frac{\pi_i}{1-\Pi_i}) + o(1/n),$$

and

$$\text{cov}(q_{ni}^0, q_{nj}^0) = o(1/n).$$

Thus, the chi-square g-o-f test for hazard rate function is given by

$$Q^0 = \mathbf{Y^0}^t W^- \mathbf{Y^0} \tag{2.2}$$

where $\mathbf{Y^0} = \mathbf{q^0}_n - E\mathbf{q^0}_n$ with $\mathbf{q^0}_n = (q_{n1}^0, \cdots, q_{n1}^0)$, and $W^-$ is the generalized inverse matrix of $W$ the covariance matrix of $\mathbf{Y^0}$. Then, observe that when $H_0$ is true, $E\{\mathbf{Y^0}\} \to 0$ as $n \to \infty$ and that when $H_0$ is false, $E\{\mathbf{Y^0}\}$ will be bound away from 0. The limiting distribution of the test statistic $Q^0$ is stated in the next theorem.

**Theorem 1**

*Under the null hypothesis*

$$Q^0 = \sum_{i=1}^{k} n\Bigg(\frac{\frac{n_i^0}{n-m_i+1} - \frac{\pi_i}{(1-\Pi_i)}}{\sqrt{\frac{\pi_i(1-\Pi_{i-1})}{(1-\Pi_i)^3(1-\Gamma_i)}}}\Bigg)^2 \sim \chi_k^2 \tag{2.3}$$

*where $k$ is the number of classes, $\Pi_0 = 0$, $\Gamma_i = P(U_i \leq x_i)$ and $\Pi_i = \sum_{j=1}^{i} \pi_j$ for $i = 1, \cdots, k$.*

Proof: The proof is a direct application of the Central Limit Theorem for multivariate random variables. A detailed proof is given in Section 2.5. ■

The usual Pearson chi-square g-o-f test converges to a chi-square distribution with degrees of freedom equal to the number of class intervals minus 1. Our chi-square goodness-of-fit test for hazard rate functions also converges to a chi-square, but the degree of freedom is equal to the number of class intervals. In the Pearson case, one can show that the covariance matrix of the vector $(O_i - E_i)_{1\leq i\leq k}$ where $O_i$ is the frequency and $E_i$ the expected frequency, is of rank $k-1$; hence, the $k-1$ degrees of freedom. As we shall see in the proof of Theorem 1, the covariance matrix of the random vector $\mathbf{q^0}_n$ is of rank $k$.

### 2.2.2 Hypothesized Function Known up to the Parameter

In this subsection, we consider the null hypothesis

$$H_0 : \lambda^0(x) = \lambda^0(x;\theta) \text{ for every } x \in \mathbb{R}^+ \tag{2.4}$$

where the parameter $\theta$ is unknown, but lies in a well specified parameter space. Consequently, the chi-square g-o-f test defined in Subsection 2.2 becomes

$$Q^0(\theta) = \sum_{i=1}^{k} n\Bigg(\frac{\frac{n_i^0}{n-m_i+1} - \frac{\pi_i(\theta)}{1-\Pi_i(\theta)}}{\sqrt{\frac{\pi_i(\theta)(1-\Pi_{i-1}(\theta))}{(1-\Pi_i(\theta))^3}}}\Bigg)^2$$

where $\pi_i(\theta) = \int_{I_i} f^0(x;\theta)dx$. Therefore, to test the null hypothesis given in equation 2.4, one also needs to estimate the unknown parameter $\theta$. In statistical inference, one of the most studied and used estimation procedures is the maximum likelihood; however, here, due to the structure of the test statistic, the parameter $\theta$ is estimated using the minimum chi-square procedure.

The minimum chi-square procedure was introduced by Pearson in order to test the null hypothesis $H_0 : f(x) = f(x;\theta)$ for $x \in \mathbb{R}$ where $f(x,\theta)$ is known up to the parameter

$\theta$. Generally speaking, the minimum chi-square method is suited for a dataset that has a multinomial structure [28]. Under regularity conditions, Neyman [28] showed that the statistic

$$\chi_p^2(\theta) = \sum_{i=1}^{k} \frac{(n_i - n\pi_i(\theta))^2}{n_i}$$

has a minimum with respect to $\theta$ and that minimum, denoted by $\hat{\theta}$, is a Best Asymptotic Normal estimator (B.A.N.) for $\theta$. Observe that $\chi_p^2(\theta)$ is a modified version of the Pearson chi-square test statistic, which is easier to work within practice compare to the usual $\chi^2(\theta)$ statistic. Lastly, Neyman proved that under regularity conditions the minimum chi-square method is asymptotically equivalent to the maximum likelihood estimator. In conclusion, estimators based on, both the statistics $\chi^2(\theta)$ and $\chi_p^2(\theta)$, produce B.A.N. estimators.

Before we state the theorem, which establishes the asymptotic distribution of the test statistic under the null defined in equation (2.4), we state a lemma that is used in the proof of this theorem. For that define the multivariate function,

$$\begin{aligned} \mathbf{g} : \quad (0,1)^k &\rightarrow (0,1)^k \\ x = (x_1, \cdots, x_k) &\rightarrow \mathbf{g}(x) = (g_1(x), g_2(x), \cdots, g_k(x)) \end{aligned} \tag{2.5}$$

where $g_i(x_1, \cdots, x_k) = \frac{x_i}{1-\sum_1^i x_j + 1/n}$ for $i \in 1, \cdots, k-1$ and $g_k(x_1, \cdots, x_k) = x_k$.

**Remark 1** *First, one can observe that the evaluation of* $\mathbf{g}$ *at* $x_n = (n_1/n, \cdots, n_k/n)$ *is equal to* $\Delta\mathbf{q}_n^0$. *As a consequence, the multivariate function* $\mathbf{g}$ *is a one-to-one multivariate function that transforms the density histogram into the hazard rate histogram and vice-versa.*

**Lemma 1**

*Let* $\mathbf{g}$ *be defined as in 2.5. Then, if* $\mathbf{g}$ *satisfies the following :*

1. g *is a one to one bi-continous mapping from a neighborhood of* $\pi(\theta)$ *to* $(0,1)^k$,

2. g *has continuous partial derivative of the second order,*

3. $\mathbf{g}'(\pi(\theta))$ *is non singular for each* $\theta \in (0,1)^k$,

4. *and* $\mathbf{g}'(\pi(\hat{\theta}))^t\Sigma(\hat{\theta})\mathbf{g}'(\pi(\hat{\theta})) \xrightarrow{P} \mathbf{g}'(\pi(\theta_0))^t\Sigma(\theta_0)\mathbf{g}(\pi(\theta_0))$

*the value* $\theta_g$ *which minimizes the transformed* $\chi^2$ *statistics exists, is a B.A.N estimator, and converges in distribution to* $\theta_0$.

Proof: First, the multivariate function g is continuous as a product of continuous function and also each of its coordinate is continuous. Since g is defined on the open interval $(0,1)^k$, it follows from Remark 1 that g is a one-to-one multivariate function. Thus, condition 1 is guaranteed.

Condition 2 is also true since the denominator of the first and second derivatives is always greater than zero; in addition, the first and second derivatives are products of continuous functions.

By construction of g, the matrix $\mathbf{g}'(P(\theta))$ is triangular; hence, it is non singular. In particular, g is a well defined and continuously differentiable on $(0,1)^k$.

The last conditions is a result of the composition of continuous functions. If we assume that $\hat{\theta}$ converges in probability to $\theta$ and since $\mathbf{g}'(\pi(\cdot))$ is a continuous function, condition 4 is verified.

Therefore, the proof of Lemma 1 follows from Ferguson [14], Theorem 1. ■

The limiting distribution of the vector $\mathbf{Z}^0 = (\frac{n_i^0}{n-m_i+1} - \frac{\pi_i(\theta)}{1-\Pi_i(\theta)})_{1\le i\le k}$ can be derived from the classical Central Limit Theorem for multivariate random variables. However, the next theorem provides an alternative way to derive the asymptotic distribution of $Z$.

**Theorem 2**

*Suppose that* $b_n^{-1}X_n = (X_{n1}, \cdots, X_{nk})$ *is asymptotically a* $N_n(\mu, \Sigma)$ *with* $\Sigma$ *a covariance matrix and* $b_n \to 0$. *Let* $\mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), \cdots, g_m(\mathbf{x}))$ *and* $\mathbf{x} = (x_1, \cdots, x_k)$ *be a vector-valued function for which each component function of* $\mathbf{g}(\mathbf{x})$ *is real valued and has nonzero differential* $g_l(\mu; t)$ , $t = (t_1, \cdots, t_k)$, *at* $x = \mu$. *Put*

$$D = \left[\frac{\partial g_i}{\partial x_j}\Big|_{x=\mu}\right]_{m\times k}. \tag{2.6}$$

*Then*

$$b_n^{-1} g(X_n) \xrightarrow{d} N(g(\mu), D\Sigma D^t). \tag{2.7}$$

Proof: See Serfling [32] for a proof of Theorem 2. ■

A version of Theorem 1 for $\theta$ unknown is stated below.

**Theorem 3**

*Let* g *be defined as in 2.5 and* $\hat{\theta}$ *the minimum chi-square estimate of* $\theta$. *Suppose further that* $\pi(\hat{\theta})$ *is a consistent estimator of* $\pi(\theta)$ *and that* $\mathbf{g}'(\pi(\hat{\theta}))^t\Sigma(\hat{\theta})\mathbf{g}'(\pi(\hat{\theta})) \xrightarrow{P} \mathbf{g}'(\pi(\theta))^t\Sigma(\theta)\mathbf{g}'(\pi(\theta))$.

*Then under* $H_0$,

$$Q^0(\hat{\theta}) = \sum_{i=1}^{k} n\left(\frac{\frac{n_i^0}{n-m_i+1} - \frac{\pi_i(\hat{\theta})}{1-\Pi_i(\hat{\theta})}}{\sqrt{\frac{\pi_i(\hat{\theta})(1-\Pi_{i-1}(\hat{\theta}))}{(1-\Pi_i(\hat{\theta}))^3}}}\right)^2 \sim \chi^2_{k-s} \tag{2.8}$$

*where* $s$ *is the dimension of the parameter space.*

Proof: The proof of Theorem 3 is given in section 2.5. ■

### 2.3 Power of the Test Statistic

In this section, we establish the limiting distribution of the test statistic $Q^0$ under fixed alternatives. For that consider the null hypothesis

$$H_0 : \lambda^0(x) = \lambda^0_0(x) \text{ for every } x \in \mathbb{R}^+$$

against the alternative hypothesis

$$H_1 : \lambda^0(x) = \lambda^0_1(x) \text{ for every } x \in \mathbb{R}^+.$$

In general, one can consider a sequence of alternatives

$$H_{1n} : \lambda^0(x) = \lambda^0_{1n}(x) \text{ for every } x \in \mathbb{R}^+$$

where $\lambda^0_{1n}(x)$ converges to $\lambda^0_1(x)$ for every $x \in \mathbb{R}^+$. The distribution of the test statistic $Q^0$ under the aforementioned fixed alternative hypotheses is stated in the next theorem.

**Theorem 4**

*Under $H_1$, $(\mathbf{q}_n - \mathbf{h}_0)^t W^-(\mathbf{q}_n - \mathbf{h}_0)$ converges in distribution to $\chi^2_k(\delta^t W^{-1}\delta)$ where $\delta = \mathbf{h}_1\mathbf{h}_0$ with $\mathbf{h}_1 = E_1(\mathbf{q}_n)$.*

Proof: The proof of Theorem 4 is given in Section 2.5. ■

Moreover, in order to prove Theorem 4, the next lemma is required.

**Lemma 2**

*Suppose $\mathbf{Z} \sim N_k(\mu, P)$ where $P$ is a projection matrix of rank $r \leq k$ and $P\mu = \mu$. Then $\mathbf{Z}^t\mathbf{Z} \sim \chi^2_r(\mu^t\mu)$.*

Proof: $P$ being a covariance matrix implies that $P$ is symmetric. Therefore, there exists an orthogonal matrix $Q$ such that $QPQ^{-1} = \mathbf{diag}(e)$ and $e$ is the vector of eigenvalues of $P$. Because $P$ is an projection matrix, the eigenvalues of $P$ are 0 or 1, and the number of 1's is equal to the rank of $P$. Since $P$ has rank $k$, the random vector $QZ \sim N_n(Q\mu, \mathbf{diag}(e))$ and $Z^tZ = (QZ)^tQZ$ is distributed as $\chi^2_r(\varphi)+\phi$ where $\varphi = \sum_{i=1}^{r}(Q\mu)_i$ and $\phi = \sum_{i=r+1}^{k}(Q\mu)_i$. Since

$$Q\mu = QP\mu = QPQ^tQ\mu = \mathbf{diag}(e)Q\mu$$

we must have $\phi = 0$ and $\varphi = \mu^t\mu$. ■

### 2.4 Simulation

The aim of this section is to evaluate the performance of the proposed chi-square g-o-f test for hazard rate functions and to investigate its power properties in finite-samples. In addition, the density functions of the test statistic $Q^0$, at specified sample sizes, are plotted against the chi-square distribution function for comparison purposes.

In this investigation, we generate 1000 samples from the distribution function associated with the specified hazard rate function $\lambda_0^0$ (e.g. the Weibull distribution function with parameter (1.5,1)). Then, the histogram hazard rate estimator is computed using equation 1.3 followed by the computation of the observed test statistic $Q^0$ using equation 2.2. In particular, we take $T$ equal to the $80th$ percentile of the distribution function associated with $\lambda_0^0$. In this simulation study, the number of class intervals are set to fourteen.

To plot the estimated density function of $Q^0$, we use the function `density` in R. Figure 2.1 shows the plots of the estimated density function of $Q^0$ in different colors and line type for sample of size $n = 80, 100,$ and $200$, respectively. The solid line represents the plot of the chi-square with 15 degree of freedom. Observed that the estimated density functions of the test are closer to the $\chi^2_{14}$'s density function as $n$ increases.

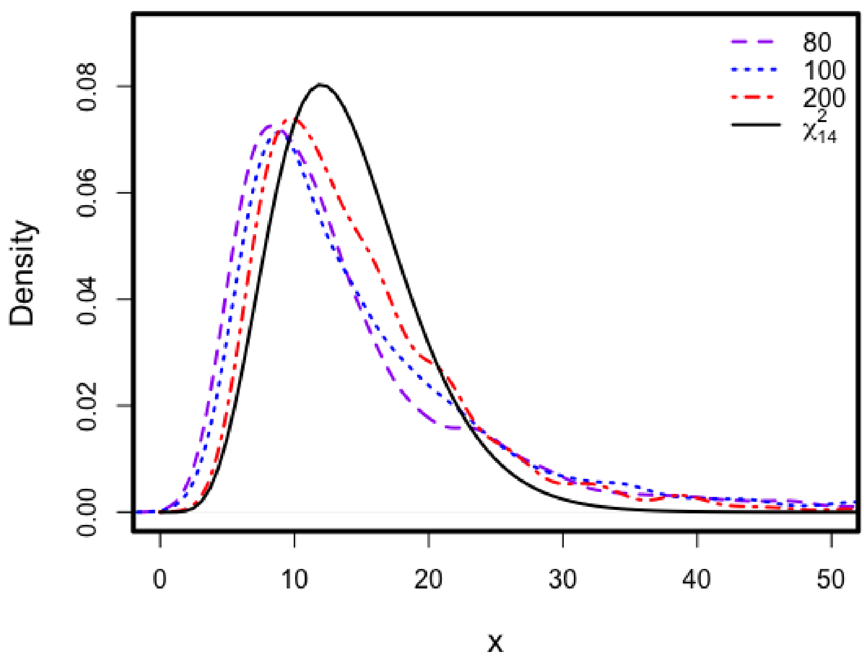


Figure 2.1: Plot of the asymptotic distribution of $Q^0$ versus the plot of the estimated density function of the test at sample sizes $n = 80, 100, 200$.

### 2.4.1 Observed Critical Values

In this subsection, the hazard rate functions considered in this simulation study are associated with the following distribution functions :

- the Weibull (W) distribution with density $f(x) = (a/b)(x/b)^{(a-1)} \exp(-(x/b)^a)$,
- the lognormal (LN) distribution with density $(x) = \frac{1}{s} exp((x-\mu)/\sigma)(1 + exp((x-\mu)/\sigma))^{-2}$,
- the folded normal (FN) distribution with density given by $\phi(y, \mu, \sigma)$ + $\phi(y, -\mu, \sigma)$ where $\phi$ is the density of the standard normal,
- the Birnbaum-Saunders (BS) distribution with cumulative distribution $F(y; a, b) = \Phi[xi(y/b)/a]$ where $\Phi$ is the cumulative distribution function of a standard normal, $xi(t) = t^{0.5} - t^{-0.5}, y > 0, a > 0, b > 0$,
- the Generalized Gamma (GG) distribution with density $f(x|\alpha, \beta, \gamma) = \frac{\gamma x^{\gamma\beta - 1} e^{(x/\alpha)^\gamma}}{\alpha^{\gamma\beta}\Gamma(\beta)}$.

In addition, we take arbitrarily $a = 1.5$ and $b = 1$ for the Weibull (W) model parameters, whereas for the Generalized Gamma model we let $\alpha = 1$, $\beta = 0.6$ and $\gamma = 4$, for the lognormal (LN) model we set $m = 0$ and $s = 1$, for the folded normal (FN) model we take $\mu = 0$ and $\sigma = 1$, and for the Birnbaum-Saunders (BS) model we take $a = 1.75$ and $b = 1$. Figure 2.2 displays the plots of the W, the LN, the FN, the GG, and the BS hazard rate functions, respectively.

The null hypothesis of interest is

$$H_0 : \lambda^0(x) = \lambda_0^0(x) \text{ for every } x \in \mathbb{R}^+$$

where $\lambda_0^0$ is associated with one of the distribution functions above.

To evaluate the performance of the test, 1000 data sets of the observed events denoted by $x_i^0, i = 1, \cdots, n$ are generated from the distribution function associated with the null hypothesis. We also generate 1000 data sets of unobserved events denoted by $u_i, i =$

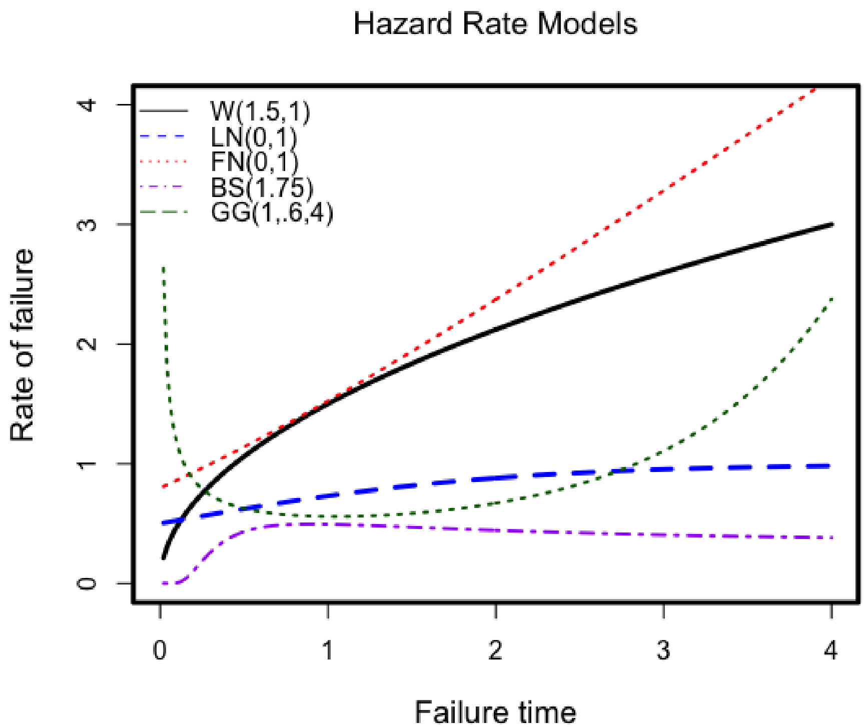


Figure 2.2: Example of hazard rate functions.

$1, \cdots, n$ using the exponential distribution. The censored data set is then given by the relation $x_i = \min\{x_i^0, t_i\}, i = 1, \cdots, n$ and the observed test statistic, $Q^0$, is computed using equation 2.2.

In this simulation study, the different levels of censoring considered are 0%,5%,15% and 28%, respectively. Each level of censoring is set by adjusting the value of the parameter of the exponential distribution. The observed critical value of the test is then computed at different sample sizes ($n$) using the function `quantile` in R with type=7. In particular, we let $n$ =80,100 and 200 respectively. Table 2.1 contains the observed critical value for $n$ =80,100,200 at $\alpha = 0.10$. One can observe that as $n$ increases the observed critical values converge to the nominal ones. The results for the other hazard rate functions are given in Appendix A.1.

Table 2.1: Observed critical values of the test $Q^0$versus chi-square critical values

| df | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|
| $C_{80,.10}$ | 14.74 | 16.04 | 17.22 | 17.37 | 20.01 | 21.60 | 22.81 |
| $C_{100,.10}$ | 13.63 | 16.30 | 17.04 | 18.16 | 19.35 | 21.25 | 21.78 |
| $C_{200,.10}$ | 13.44 | 14.85 | 16.41 | 18.36 | 19.06 | 21.39 | 22.42 |
| $c_{0.10}$ | 14.06 | 15.50 | 16.91 | 18.30 | 19.67 | 21.03 | 22.36 |

### 2.4.2 Observed Power

In this subsection, our goal is to evaluate the consistency of $Q^0$ by computing the observed power of the test against fixed alternatives. Toward that aim, we consider the null hypothesis $H_0 : \lambda^0(x) = \lambda_0^0(x)$ for every $x \in \mathbb{R}^+$ where $\lambda_0^0$ is associated with the

Weibull distribution with parameters (shape=1.5, scale=1) versus the alternative hypothesis $H_1 : \lambda^0(x) = \lambda_1^0(x)$ where $\lambda_1^0$ is derived from the folded normal (FN(0,1)), the lognormal (LN(0,1)), the Birnbaum-Saunders (BS(1.75,1)), and the generalized gamma GG(1,0.6,4) distributions respectively.

Although the data is generated using similar procedures to the process presented above, it is worth mentioning that the failure time data sets are simulated from the distribution function associated with the alternative hazard rate function. From there, at a given sample size, the observed test statistic, $Q^0$, is computed 1000 times; then, the observed power of test is computed as follow

$$\text{Power} = \frac{\#\{Q^0 > \hat{c}_{0.05}\}}{1000}$$

where $\hat{c}_{0.05}$ is the observed critical values of the test at the nominal level $\alpha = 0.05$.

Based on Figure 2.2, one expects that a consistent test statistic may have relatively smaller power against the folded normal hazard rate function comparatively to the other hazard rate functions, but as $n$ increases its observed power should converge to one. Table 2.2 displays the observed power of $Q^0$ against the four different hazard rate functions considered in this simulation study. As expected one can note that the observed power of the test increases to one as $n \to \infty$ .

Table 2.2: Observed power of the test $Q^0$ with $H_0 : \lambda^0 = \lambda^0_W$ versus $H_1 : \lambda^0 = \lambda^0_A$ where $W$ is the Weibull hazard rate model and $A$ stands for the Log-normal, the Folded normal, the Birnbaum-Saunders and the generalized gamma hazard rate models, respectively.

| $n$ | FN(0,1) | BS(1.75) | LN(0,1) | GG(1,.6,4) |
|---|---|---|---|---|
| Observed Power | | | | |
| 100 | 0.21 | 0.99 | 0.01 | 0.98 |
| 150 | 0.35 | 1.00 | 0.19 | 1.00 |
| 200 | 0.49 | 1.00 | 0.75 | 1.00 |
| 300 | 0.67 | 1.00 | 1.00 | 1.00 |

## 2.5 Proofs

Proof of Theorem 1: Let $\Delta \mathbf{q}^0_n = (\frac{f^0_i}{n - \sum_{j=1}^{i} f_j + 1})_{1 \le i \le k}$ be the relative histogram hazard rate estimator defined is Section 2.2 vector where $f^0_i$ and $f_i$ are computed using equation 1.3. Thus, the covariance matrix $W$ of $\Delta \mathbf{q}^0_n$ is given by

$$\begin{pmatrix} \frac{\pi_1(1-\Pi_1-\pi_1)}{(1-\Pi_1)^2(1-\Pi_1)} + o(\frac{1}{n}) & o(1/n) & \cdots & o(\frac{1}{n}) \\ o(\frac{1}{n}) & \frac{\pi_2(1-\Pi_2-\pi_2)}{(1-\Pi_2)^2(1-\Pi_2)} + o(\frac{1}{n}) & \cdots & o(\frac{1}{n}) \\ \vdots & \cdots & \ddots & \vdots \\ o(\frac{1}{n}) & \cdots & \cdots & \frac{\pi_k(1-\Pi_k-\pi_k)}{(1-\Pi_k)^2(1-\Pi_k)} + o(\frac{1}{n}) \end{pmatrix} \tag{2.9}$$

By the Central Limit Theorem for multivariate random variables, $\Delta \mathbf{q}^0_n$ converges to a multivariate normal distribution with mean $\mathbf{0}$ and diagonal covariance matrix with diagonal terms given by $\frac{\pi_i(1-\Pi_i-\pi^0_i)}{(1-\Pi_i)^2(1-\Pi_i)}$ for $1 \le i \le k$.

Let $\mathbf{Y^0} = \mathbf{q^0}_n - E\mathbf{q^0}_n$. Then, again by a second application of the Central Limit Theorem

$$n\mathbf{Y^0}' W^- \mathbf{Y^0} \sim \chi^2_k$$

since $W$ is asymptotically a diagonal matrix of dimension $k \times k$ and clearly it is a projection matrix. ■

Proof of Theorem 3: First, let $k_i(\theta) = g_i(\pi(\theta)) = \frac{\pi_i(\theta)}{(1-\Pi_i(\theta))}$ where $\theta$ is a vector with $s$ components; then, the first derivative of $k$ with respect to $\theta_l$, for $l = 1, \cdots, s$ is denoted as $\frac{\partial}{\partial \theta_l} k_i(\theta)$ for $i = 1, \cdots, k$.

By Lemma 1 the minimizer of $Q_p^0(\theta)$ exists, it is a B.A.N. estimate of $\theta$, and $\theta_n \xrightarrow{P} \theta_0$.

Now expanding $k(\theta_n)$ about $\theta_0$ and under the differentiability assumption on $k_i$, one can show that

$$\frac{n_i}{n - m_i + 1} - \frac{\pi_i(\theta_n)}{1 - \Pi_i(\theta_n)} = \frac{\pi_i(\theta_0)}{1 - \Pi_i(\theta_0)} - \sum_{l=1}^{r} [\frac{\partial}{\partial \theta_l} k_i(\theta_0) + o_p(1)](\theta_{nl} - \theta_{0l})$$

or

$$(\frac{n_i}{n - m_i + 1})c_i^{-1} - (\frac{\pi_i(\theta_n)}{1 - \Pi_i(\theta_n)})c_i^{-1} = (\frac{\pi_i(\theta_0)}{1 - \Pi_i(\theta_0)})c_i^{-1} - \sum_{l=1}^{r} [\frac{\partial}{\partial \theta_l} k_i(\theta_0) c_i^{-1} + o_p(1)](\theta_{nl} - \theta_{0l}) \tag{2.10}$$

where $c_i$ is an appropriate estimator of $\sqrt{\frac{\pi_i(\theta)(1-\Pi_{i-1}(\theta))}{(1-\Pi_i(\theta))^3}}$. Note that

$$\sum_{i=1}^{k} [\frac{n_i}{n - m_i + 1} - \frac{\pi_i(\theta_n)}{1 - \Pi_i(\theta_n)}] \frac{\partial}{\partial \theta_l} k_i(\theta_n) = 0$$

since $\theta_n$ minimizes $Q_p(\theta)$and that $\frac{\partial}{\partial \theta_l} k_i(\theta_n) = \frac{\partial}{\partial \theta_l} k_i(\theta_0) + o_p(1) = b_{il} + o_p(1)$ (recall that $k_i$ is differentiable with respect to $\theta$). Let $C = b_{il} c_i^{-1}$ for $i = 1, \cdots, k$ and $l = 1, \cdots, s$; then, equation 2.10 turns into $W(\theta_n) = W(\theta_0) - [C + o_p(1)](\theta_n - \theta_0)$ where $W(\theta_0) = ((\frac{n_i}{n-m_i} - \frac{\pi_i(\theta_0)}{1-\Pi_i(\theta_0)})c_i^{-1})_{1 \le i \le k}$. Since $C^t W(\theta_n) = 0$, it implies that $(C^t C)^{-1} C^t W(\theta_0) = (\theta_n - \theta_0)$ asymptotically. Moreover, it follows that $W(\theta_n) = [I_k - C(C^t C)^{-1} C^t] W(\theta_0)$. By Theorem 2, the distribution of $W(\theta_0)$ is a multi-normal distribution with mean equal to 0 and matrix covariance given by the relation $D^t\Sigma D$. Since $C(C^tC)^{-1}C^t$ is a $k \times k$ matrix with rank equal to $s$, the rest of the theorem follows from standard procedure found in Cramer [11]. ■

Proof of Theorem 4: Define $\Gamma = \mathbf{diag}((\frac{\pi_i(1-\Pi_{i-1})}{(1-\Pi_i)^3})_{1\leq i\leq n})$, and let $W$ be the covariance matrix of $\mathbf{q}_n$ under the null hypothesis, i.e., $\sqrt{n}(\mathbf{q}_n - \mathbf{h}_0) \to N_k(0, W)$ if $\mathbf{q}_n \sim N_{k-1}(\mathbf{h}_0, W)$ and $\mathbf{h}^t = (\frac{\pi_1}{1-\Pi_1}, \cdots, \frac{\pi_k}{1-\Pi_k})$ and $Eq_{n_i} = \frac{\pi_i}{1-\Pi_i} + o(n^{-1})$ for $i = 1, \cdots, k$.

Since $E\mathbf{q}_{n_i} \to \frac{\pi_{i(n)}}{1-\Pi_{i(n)}}$ as $n \to \infty$, it can be shown under the earlier assumption (i.e. $(\mathbf{h}_n - \mathbf{h}_0) \to \delta$) that :

$$\sqrt{n}(\mathbf{q}_n - \mathbf{h}_n) \xrightarrow{d} N_k(0, W). \tag{2.11}$$

Now we claim that the limit of 2.11 implies that the statistic $n(\mathbf{q}_n - \mathbf{h}_0)^t\Gamma^{-1}(\mathbf{q}_n - \mathbf{h}_0)$ converges in distribution to $\chi^2_k(\delta^t\Gamma^{-1}\delta)$. Let

$$V^{(n)} = \sqrt{n}(\mathbf{q}_n - \mathbf{h}_0).$$

Then

$$V^{(n)} = \sqrt{n}(\mathbf{q} - \mathbf{h}_n) + \sqrt{n}(\mathbf{h}_n - \mathbf{h}_0).$$

The first term on the right hand side converges in distribution to $N_k(0, W)$ and the second converges to $\delta$. By the Slutsky's theorem, we have $V^{(n)} \xrightarrow{d} N_k(\delta, W)$. Hence

$$\Gamma^{-1/2}V^{(n)} \xrightarrow{d} N_k(\Gamma^{-1/2}\delta, \Gamma^{-1/2}W\Gamma^{-1/2}).$$

If we can prove that $\Gamma^{-1/2}W\Gamma^{-1/2}\Gamma^{-1/2}\delta = \Gamma^{-1/2}\delta$; then by Lemma 4 the proof is over. It is clear that $\Gamma^{-1/2}W\Gamma^{-1/2} = \mathbb{I}_k$ where $\mathbb{I}_k$ is the identity matrix; hence $\Gamma^{-1/2}W\Gamma^{-1/2}$ is a projection matrix. By Lemma 2, we conclude that

$$V^{(n)}\Gamma^{-1}V^{(n)} \xrightarrow{d} \chi_k^2(\delta^t\Gamma^{-1}\delta)$$

under the sequence of alternatives $H_{1n}$. ■

# CHAPTER III

# A KERNEL-BASED NONPARAMETRIC GOODNESS-OF-FIT TEST FOR HAZARD RATE FUNCTIONS

## 3.1 Introduction

In statistical inference, there exists a wide body of literature on nonparametric goodness-of-fit test procedures. In the last few decades, numerous authors have proposed nonparametric g-o-f procedures based on the integrated square error functional. One of the first tests and important test of this type is the one by Bickel and Rosenblatt, henceforth referred to as the B-R test. Following Bickel and Rosenblatt [6] it is straightforward to use the test statistic defined in equation 1.7 to test the null hypothesis

$$H_0 : \lambda^0(x) = \lambda_0^0(x)$$

where $\lambda_0^0(x)$ is an hypothesized functional form of the hazard rate function. Accordingly, the performance of the test statistic defined in 1.7 will be very much the same as the B-R test statistic. In fact, under appropriate regularity conditions and for $\delta \in (0, 1/4)$, one can show that the performance and power properties of the test statistic $S_n(h)$ are similar to those of the B-R test, but as pointed out in Section 1.5 one needs a very large finite sample for asymptotics to play a role when implementing such a test.

In this chapter, our main objective is to show that the kernel-based nonparametric goodness-of-fit test statistic, $\tilde{T}_N(h)$, proposed in Section 1.5 avoids the drawback mentioned in the preceding paragraph. Towards that end, in section 3.2 we state the main results, which establish the asymptotic properties of the test statistic together with the assumptions that we make for theses results to be true. The proofs of the main theorems are given in Section 3.3.

## 3.2 The Nonparametric Test Statistic and its Asymptotic Distribution

Recall that $X_i^0 > 0,\ i = 1, \cdots, N$ are independent and identically distributed (i.i.d.) failure times that are censored on the right by the i.i.d. random variables $U_i,\ i = 1, \cdots, N$, which are independent of the $X_i^0$'s. We also denote by $F^0$ and $H$ the distribution functions of the $X_i^0$'s and $U_i$'s, respectively; in addition, we assume that $F^0$ is absolutely continuous with its density function denoted by $f^0$ and $H$ is a continuous distribution function.

To study the asymptotic properties of the test statistic defined in Section 1.4, in Subsection 3.2.1 we state the assumptions that we make. The main result establishing the asymptotic distribution of the test statistic under the null hypothesis is given in Subsection 3.2.2. In Subsection 3.2.3, we establish the asymptotic distribution of the test statistic under the fixed alternative and its power properties against Pitman alternatives.

### 3.2.1 Assumptions

Recall that the test statistic defined in Section 1.5 is written as

$$\tilde{T}_N(h) = \sum\sum_{1 \leq i < j \leq n} \frac{I(|x_j - x_i| < h)}{n(n-1)h} (Y_i - \lambda_0(x_i))(Y_j - \lambda_0(x_j)).$$

To study its asymptotic properties, we make the following assumptions.

$A_1$ Let $\omega(x) = I(0 < x < T)$ be a weighted function on $\mathbb{R}$ where $T < \min(T_H, T_{F^0})$ with $T_H = \sup\{t; H(t) < 1 - \epsilon\}$ and $T_{F^0} = \sup\{t; F^0(t) < 1 - \epsilon\}$ for some small $\epsilon > 0$.

$A_2$ Assume $\Delta = N^{-\alpha}$ where $\frac{1}{2} < \alpha < 1$; as $N \to \infty, n = n(N)$ is such that $n\Delta \to \infty$ and $x_1 \to 0, x_n \to \infty$; and $h = h(N)$ is such that $h \to 0$ and $h/\Delta \to \infty$ as $N \to \infty$.

$A_3$ The hazard rate function ($\lambda^0$) , the density function ($f^0$), and the first derivative of the density function ($f^{0\prime}$) are bounded and uniformly continuous on compact sets of $\mathbb{R}$.

Since the hazard rate estimate $Y_i$ is expected to be highly unstable whenever $x_i$ is in the extreme right tail, as in Chapter 2 we confine ourselves on the interval $[0, T]$ where $T$ is determined as in $A_1$ to test the hypothesis

$$H_0 : \lambda^0(x) = \lambda^0_0(x) \text{ for every } x \in \mathbb{R}^+.$$

Accordingly, we also consider the following modification to $\tilde{T}_N$,

$$T_N(h) = \sum\sum_{1 \le i < j \le n} \frac{I(|x_j - x_i| < h)}{n(n-1)h} \omega(x_i)\omega(x_j)(Y_i - \lambda_0(x_i))(Y_j - \lambda_0(x_j)). \qquad (3.1)$$

Assumption $A_2$ states at what rate our smoothing parameters are allowed to converge to zero as $N$ becomes large. It is important to note that though both $h$ and $\Delta$ tend to zero, the ration $h/\Delta$ tends to infinity as $N \to \infty$. The last assumption $A_3$ implies that the integral over $[0, T]$ of the hazard rate function, the density function, and its first derivative are finite.

### 3.2.2 Distribution of the Test Statistic under $H_0$

Consider the null hypothesis

$$H_0 : \lambda^0(x) = \lambda^0_0(x) \text{ for every } x \in \mathbb{R}^+$$

where $\lambda^0_0$ is entirely known. The main result of this chapter is stated in the next theorem.

**Theorem 5**

*Under assumptions* $A_1, A_2, A_3$ *and* $H_0 : \lambda^0(x) = \lambda_0^0(x)$

$$N\sqrt{h}T_N(h) \quad \xrightarrow{d} N(0, \sigma_0^2) \;\; \text{as } N \;\; \to \infty$$

*where* $\sigma_0^2 = 4\int (\frac{\lambda_0^0(x)}{(1-F(x))T^2})^2 \omega(x)dx$ *and* $1 - F = (1 - F_0^0)(1 - H)$.

**Remark.** Note that, asymptotically, the test will reject $H_0$ at level $\alpha$ when

$$T_N(h) > z_\alpha \times \frac{\sigma_0}{N\sqrt{h}} = C(\alpha, h, \lambda_0^0(x), F(x)) \tag{3.2}$$

where $z_\alpha$ is such that $P(Z > z_\alpha) = \alpha$ and $Z$ is the standard normal variable. Clearly, the cut-off point depends on unknown $H$ through $F$ and in almost all real life situations the knowledge on the precise form of $H$ seems quite unrealistic. Of course in the uncensored case, since the whole mass of the censoring distribution is at infinity it is not a problem, but it is so for the censored case. To overcome this we use the most flexible and commonly used Weibull distribution in place of $H$. In this thesis, we do not investigate a possible alternative approach of using a nonparametric estimator of $\sigma_0^2$.

### 3.2.3 Distribution of the Test Statistic under $H_1$

Let

$$H_1 : \lambda^0(x) = \lambda_1^0(x), \text{ for every } x \in \mathbb{R}^+$$

where $\lambda_1^0(x) = \frac{f_1^0(x)}{(1-F_1^0(x))}$ for $F_1^0(x) < 1$ is a fixed alternative hazard rate function. The asymptotic distribution of the test statistic under $H_1$ is given in the next corollary.

**Corollary 1**

*Under $A_1$, $A_2$, $A_3$ and $H_1$*

$$N\sqrt{h}T_N(h) \xrightarrow{d} N\Big(\frac{2}{T^2}\int(\lambda_1^0(x)-\lambda_0^0(x))^2\omega(x)dx,\ 4\int(\frac{\lambda_1^0(x)}{Q_1(x)T^2})^2\omega(x)dx\Big)$$

*as $N\to\infty$ where $Q_1(x)=(1-F_1^0)(1-H)$.*

Let $h\sim N^{-\delta}$ such that $0<\delta<\alpha<1$. Then, the Pitman alternatives we consider to evaluate the performance of our test statistic are,

$H_{1N}:\lambda_{1N}^0(x)=\lambda_0^0(x)+N^{-\beta}\eta(x)+o(N^{-\beta}),\forall x\in\mathbb{R}^+$ and $\beta>0,$

$H_{2N}:\lambda_{2N}^0(x)=\lambda_0^0(x)+N^{-\epsilon}\sum_{j=1}^k\eta_j(\frac{x-c_j}{N^{-\gamma}})+o(N^{-\epsilon-\gamma}),\forall x\in\mathbb{R}^+$ and $0<\epsilon,\ 0<\gamma<\delta.$

where $\eta$ is a continuous function on $\mathbb{R}$ and the $\eta_j$'s are twice continuously differentiable on $\mathbb{R}$ and centered about the $c_j$'s.

**Corollary 2**

*Under assumptions $A_1$,$A_2$, $A_3$ and $H_{1N}$*

$$P(T_N(h)>C(\alpha,h,\lambda_0^0(x))|\lambda^0=\lambda_{1N}^0)=\begin{cases}\alpha, & \text{if } \beta>(2-\delta)/4,\\ \Phi(l), & \text{if } \beta=(2-\delta)/4,\\ 1, & \text{if } 0<\beta<(2-\delta)/4,\end{cases}$$

*as $N\to\infty$ where $l=\sigma_0^{-1}\int(\eta(x))^2\omega(x)dx-z_\alpha$. and $\Phi$ is the CDF of the standardized normal variable.*

*And under $A_1$, $A_2$, $A_3$ and $H_{2N}$*

$$P(T_N(h)>C(\alpha,h,\lambda_0^0(x))|\lambda^0=\lambda_{2N}^0)=\begin{cases}\alpha, & \text{if } 1-2\epsilon-\gamma<\delta/2,\\ \Phi(l'), & \text{if } 1-2\epsilon-\gamma=\delta/2,\\ 1, & \text{if } 1-2\epsilon-\gamma>\delta/2>0,\end{cases}$$

as $N \to \infty$ where $l' = \sigma_0^{-1} \sum_{j=1}^k \Big\{ \int (\eta_j(x))^2 \omega(x) dx \Big\} - z_\alpha$.

The proof of corollaries 1 and 2 are given in Section 3.3.

## 3.3 Proofs

Proof of Theorem 5: Let

$$T_N^*(h) = \sum\sum_{1 \le i \ne j \le n} \frac{I(|x_i - x_j| < h)}{n(n-1)h} \omega(x_i)\omega(x_j)(Y_j^* - \lambda_0^0(x_j))(Y_i^* - \lambda_0^0(x_i)) \tag{3.3}$$

where

$$Y_i^* = \frac{1}{\Delta} \frac{f_i^0}{N(1 - F(x_i))}, \ i = 1, 2, \cdots, n. \tag{3.4}$$

Then, from Lemma 3, $T_N(h) = T_N^*(h) + o_p(1)$. As a consequence, Theorem 5 can be proved by deriving the distribution of $T_N^*(h)$. Noting the fact that

$$Y_i^* = \sum_{k=1}^N \frac{k(x_i, X_k)}{N \Delta Q(x_i)}$$

where $k(x_i, X_k) = K(\frac{x_i - X_k}{\Delta/2}) C_k$ and $Q(x_i) = 1 - F(x_i)$, $T_n^*(h)$ can be expressed as

$$T_N^*(h) = J_{N2} + J_{N1} + J_{N3} + J_{N4} + J_{N5}$$

where

$$\begin{aligned}
J_{N1} &= \sum_{k=1}^N \Big( \sum\sum_{1 \le i \ne j \le n} \frac{I(|x_i - x_j| < h)}{n(n-1)h} \omega(x_i)\omega(x_j) s(x_i, X_k) s(x_j, X_k) \Big), \\
J_{N2} &= \sum\sum_{1 \le k \ne l \le N} \Big( \sum\sum_{1 \le i \ne j \le n} \frac{I(|x_i - x_j| < h)}{n(n-1)h} \omega(x_i)\omega(x_j) s(x_i, X_k) s(x_j, X_l) \Big), \\
J_{N3} &= \sum_{k=1}^N \Big( \sum\sum_{1 \le i \ne j \le n} \frac{I(|x_i - x_j| < h)}{n(n-1)h} \omega(x_i)\omega(x_j) s(x_j, X_k)(EY_i^* - \lambda_0^0(x_i)) \Big), \\
J_{N4} &= \sum_{k=1}^N \Big( \sum\sum_{1 \le i \ne j \le n} \frac{I(|x_i - x_j| < h)}{n(n-1)h} \omega(x_i)\omega(x_j) s(x_i, X_k)(EY_j^* - \lambda_0^0(x_j)) \Big), \\
J_{N5} &= \sum\sum_{1 \le i \ne j \le n} \frac{I(|x_i - x_j| < h)}{n(n-1)h} \omega(x_i)\omega(x_j)(EY_j^* - \lambda_0^0(x_j))(EY_i^* - \lambda_0^0(x_i))
\end{aligned}$$

with $s(x_i, X_k) = (\frac{k(x_i,X_k)}{N\Delta Q(x_i)} - E\frac{k(x_i,X_k)}{N\Delta Q(x_i)})$.

Now to prove Theorem 5, we analyze each of the terms $J_{N1}$,$J_{N2}$,$J_{N3}$, $J_{N4}$ and $J_{N5}$. Hereinafter, we set $I(|x_i - x_j| < h) = I_{i,j}(h)$ and $\omega_{i,j} = \omega(x_i)\omega(x_j)$. First, let $J_{N2} = 2(N\Delta)^{-2}U_N$ where $U_N = \sum\sum_{1\leq k<l\leq N} H_N(X_k, X_l)$ with

$$\begin{aligned} H_N(X_k, X_l) &= \sum\sum_{1\leq i\neq j\leq n} \frac{I_{i,j}(h)\omega_{i,j}(x_i)}{n(n-1)hQ(x_i)Q(x_j)}(k(x_i, X_k) - Ek(x_i, X_k)) \\ &\times(k(x_j, X_l) - Ek(x_j, X_l)). \end{aligned} \tag{3.5}$$

Observe that $H_N(x, y)$ is a symmetric function and by Lemma 8 $H_N(X_k, X_l)$ is a degenerate U-statistic. Let us define

$$\begin{aligned} G_N(x, y) &= \tfrac{1}{(n(n-1))^2h^2} \sum_{i_1\neq j_1}\sum_{i_2\neq j_2} I_{i_1,j_1}(h)I_{i_2,j_2}(h)\omega_{i_1,j_1,i_2,j_2} \\ &\times E\{(\tfrac{k(x_{i_1}-X_1)}{Q(x_{i_1})} - E\tfrac{k(x_{i_1}-X_1)}{Q(x_{i_1})})(\tfrac{k(x_{i_2}-X_1)}{Q(x_{i_2})} - E\tfrac{k(x_{i_2}-X_1)}{Q(x_{i_2})})\} \\ &\times(\tfrac{k(x_{j_1}-x)}{Q(x_{j_1})} - E\tfrac{k(x_{j_1}-x)}{Q(x_{j_1})})(\tfrac{k(x_{j_2}-y)}{Q(x_{j_2})} - E\tfrac{k(x_{j_2}-y)}{Q(x_{j_2})}), \end{aligned} \tag{3.6}$$

and observe that as $N \to \infty$ it follows from Lemma 8 that

$$[EG_N^2(X_1, X_2) + N^{-1}EH_N^4(X_1, X_2)]/[EH_N^2(X_1, X_2)]^2 \to 0. \tag{3.7}$$

Therefore, using Theorem 1 of Hall [18] , $U_N$ is asymptotically normally distributed with mean zero and variance equal to $\frac{N^2}{2}EH_N^2(X_1, X_2)$; hence, one can conclude that $J_{N2}$ is asymptotically normal with mean zero and variance $\sigma^2_{N2}$ where $\sigma^2_{N2} = 4N^{-2}h^{-1}T^{-4}\sigma^2_2$ with $\sigma^2_2 = \int(\frac{\lambda^0_0(x)}{Q(x)})^2\omega(x)dx$ from Lemma 8. Since analysis and conclusions about $J_{N3}$ are applicable exactly to $J_{N4}$, we give analysis only for $J_{N3} = \sum_{k=1}^N(N\Delta)^{-1}J_{N3k}$ where

$$J_{N3k} = \frac{1}{n(n-1)h}\sum\sum_{1\leq i\neq j\leq n} \frac{I_{i,j}(h)\omega_{i,j}}{Q(x_j)}(k(x_j, X_k) - Ek(x_j, X_k))(EY_i - \lambda^0_0(x_i)). \tag{3.8}$$

If we set $s_N^2 = \sum_{k=1}^N EJ_{N3k}^2$, it follows from Lemma 9 that

$$s_N^{-2} \sum_{i=1}^N E(J_{N3i}^2 I(|J_{N3i}| > \varepsilon s_N)) \leq \varepsilon^{-2} s_N^{-4} \sum_{i=1}^N EJ_{N3i}^4 \to 0 \text{ as } N \to \infty.$$

As a consequence, $J_{N3}$ is asymptotically normally distributed with mean zero and variance $\sigma_{N3}^2 = N^{-1}\Delta^4 T^{-4} k^2 \sigma_3^2$ where $k = \int t^2 K(t)dt$ and $\sigma_3$ is given in lemma 9. Thus $J_{N3} = O_p(N^{-1/2}\Delta^2)$.

Now let us consider $J_{N1}$; it follows from Lemma 7 that

$$EJ_{N1} = -\frac{1}{N}\frac{1}{T^2}\int \lambda_0^0(x)^2\omega(x)dx + o(\Delta),$$

$$\text{Var}J_{N1} = O((N^3h)^{-1}) \text{ and } J_{N1} = O_p(\frac{1}{N\sqrt{Nh}}).$$

Hence $J_{N1} = o_p(J_{N2})$. The last term, $J_{N5}$, is deterministic. Therefore, using the Riemann integration approximation techniques given in Lemma 5, it can be shown that

$$EJ_{N5} = 2\Delta^4 k^2 \int \kappa(x)\omega(x)dx + o(\Delta^4)$$

where $\kappa = f_0^{0\prime\prime}/(1-F)$. Consequently,

$$ET_N^*(h) = -\frac{1}{N}\frac{1}{T^2}\int \lambda_0^0(x)^2\omega(x)dx + o(\Delta)$$

which goes to zero as $N \to \infty$. With $J_{N3} = O_p(N^{-1/2-2\alpha})$, $J_{N2} = O_p(N^{-1+\delta/2})$, and $0 < \delta < \alpha < 1$ the asymptotic distribution of $T_N^*(h)$ is determined by $J_{N2}$. That completes the proof. ∎

**Lemma 3**

*Assuming* $A_1, A_2, A_3$ *and under* $H_0$

$$N\sqrt{h}(T_N(h) - T_N^*(h)) = o_p(1).$$

Proof of Lema 3: Using the expansion of the following product:

$$(Y_i - Y_i^* + Y_i^* - \lambda_0(x_i))(Y_j - Y_j^* + Y_j^* - \lambda_0(x_j)),$$

it follows that

$$T_N(h) - T_N^*(h) = D_1 + D_2 + D_3$$

where

$$\begin{aligned} D_1 &= \tfrac{1}{n(n-1)h} \sum\sum_{1\le i\ne j\le n} I_{i,j}(h)\omega_{i,j}(Y_i - Y_i^*)(Y_j - Y_j^*) \\ D_2 &= \tfrac{1}{n(n-1)h} \sum\sum_{1\le i\ne j\le n} I_{i,j}(h)\omega_{i,j}(Y_i^* - \lambda_0(x_i))(Y_j - Y_j^*) \\ D_3 &= \tfrac{1}{n(n-1)h} \sum\sum_{1\le i\ne j\le n} I_{i,j}(h)\omega_{i,j}(Y_i - Y_i^*)(Y_j^* - \lambda_0(x_j)). \end{aligned}$$

Thus,

$$|T_N(h) - T_N^*(h)| \quad \le |D_1| + |D_2| + |D_3|.$$

Observe that the first term, $D_1$, is the sum of the following product terms :

$$\frac{f_i^0}{N\Delta}\left(\frac{1}{1-\hat{F}_n(x_i)+1/N} - \frac{1}{1-F(x_i)}\right)\frac{f_j^0}{N\Delta}\left(\frac{1}{1-\hat{F}_n(x_j)+1/N} - \frac{1}{1-F(x_j)}\right)$$

where $\hat{F}_n(u) = \frac{1}{N}\sum_{i=1}^N I[X_i \le u]$.

By the Central Limit Theorem, for each $x_i$ we have $\hat{F}_n(x_i) \xrightarrow{d} F(x_i)$ at the standard rate of $\sqrt{n}$ (see Serfling [32]) and $1/N \to 0$ as $N \to \infty$. Since $x \to \frac{1}{1-x}$ is a continuous function on $[0,1)$, by the mapping theorem $\frac{1}{1-\hat{F}_n(x_i)+1/N} - \frac{1}{1-F(x_i)} = O_p(\sqrt{n})$. Using Lemma 4 and the convergence rate of $\hat{F}_n(x_i)$,

$$|D_1| \le M_3^2 O_p(\frac{1}{N}) + o_p(1)$$

where $M_3 = \max_i\{f_0^0(x_i)(1 - H(x_i))\}$. The terms, $D_2$ and $D_3$, can be treated similarly; therefore, we only present the analysis of $D_2$. It is easy to check that

$$|D_2| \leq \Big|\frac{1}{n(n-1)h}\sum\sum_{1\leq i\neq j\leq n} I_{i,j}(h)\omega_{i,j}(Y_i^* - \lambda_0^0(x_i))\Big|\Big|(Y_j - Y_j^*)\Big|,$$

which by Lemma 6 implies that

$$|D_2| \leq M_3 O_p(\frac{1}{\sqrt{N}})O_p(\frac{1}{\sqrt{N}}) = O_p(\frac{1}{N}).$$

Therefore,

$$T_N(h) - T_N^*(h) = (O_p(\frac{1}{N}) + o_p(\frac{1}{N}))$$

and

$$N\sqrt{h}[T_N(h) - T_N^*(h)] = N\sqrt{h}(O_p(\frac{1}{N}) + o_p(\frac{1}{N})) = O_p(\sqrt{h}) + o_p(\sqrt{h}) = o_p(1).$$

■

We now state and prove the lemmas used to prove Theorem 5 and Lemma 3.

**Lemma 4**

*Under assumption* $A_2$, $\sum\sum_{1\leq i\neq j\leq n} I_{i,j}(h)\omega_{i,j} = 2\frac{nh}{\Delta} + O(n).$

Proof: Let $R = \sum\sum_{1\leq i\neq j\leq n} I_{i,j}(h)\omega_{i,j}\Delta^2 - \int_0^T\int_0^T I_{u,v}(h)\omega(u)\omega(v)dudv,$ $s_1 = 0 < \cdots < s_{n+1} = T$ such that $s_i = x_i - \Delta/2$ for $i = 1, \cdots, n$. One can note that $s_i < x_i < s_{i+1}$ for $1 \leq i \leq n$ and show that

$$|R| \;\; = \Big|\sum\sum_{1\leq i\neq j\leq n}\{I_{i,j}(h)\Delta^2 - \int_{s_j}^{s_{j+1}}\int_{s_i}^{s_{i+1}} I_{u,v}(h)dudv\}\Big|.$$

By definition of $I(\cdot)$ and the Mean Value Theorem, there exist $\alpha_i$, $\beta_j$ such that

$$\int_{s_j}^{s_{j+1}}\int_{s_i}^{s_{i+1}} I_{u,v}(h)dudv = I_{\alpha_i,\beta_j}(h)(s_{i+1} - s_i)(s_{j+1} - s_j)$$

and $s_i < \alpha_i < s_{i+1}$, $s_j < \beta_j < s_{j+1}$. Moreover, since $s_{i+1} - s_i = \Delta = \frac{T}{n}$ for $1 \le i \le n$, it follows that

$$|R| \; = \Delta^2 \Big| \sum\sum_{1\le i\ne j\le n} \{I_{i,j}(h) - I_{\alpha_i,\beta_j}(h)\} \Big|.$$

By a counting argument, we observe that $I_{i,j}(h) - I_{\alpha_i,\beta_j}(h)$ can be different from zero at most on four intervals; therefore,

$$|R| \; \le \Delta^2 \sum\sum_{1\le i\ne j\le n} |\{I_{i,j}(h) - I_{\alpha_i,\beta_j}(h)\} \Big| < 4\Delta^2 n = O(n^{-1}).$$

Finally, $\int_0^T \int_0^T I_{u,v}(h)dudv = 2hT$ implies $\sum\sum_{1\le i\ne j\le n} I_{i,j}(h) = \dfrac{2nh}{\Delta} + O(n)$. ■

**Lemma 5**

*Under assumption $A_2$,*

$$\sum\sum_{1\le i\ne j\le n} \frac{I(|x_i - x_j| < h)}{n(n-1)h} \omega(x_i)\omega(x_j)g(x_i)g(x_j) = \int g(u)^2\omega(u)du + o(1)$$

*where $g$ is a continuous function.*

Proof: First we know that

$$\sum\sum_{1\le i\ne j\le n} \frac{I(|x_i - x_j| < h)}{n(n-1)h} \omega(x_i)\omega(x_j)g(x_i)g(x_j) =$$
$$\sum\sum_{1\le i\le n, 1\le j\le n} \frac{I(|x_i - x_j| < h)}{n(n-1)h} \omega(x_i)\omega(x_j)g(x_i)g(x_j) \quad + \sum_{1\le i\le n} \frac{I(0)}{n(n-1)h}\omega(x_i)g(x_i)^2;$$

hence,

$$\sum\sum_{1\le i\ne j\le n} \frac{I(|x_i - x_j| < h)}{n(n-1)h} \omega(x_i)\omega(x_j)g(x_i)g(x_j) =$$
$$\sum\sum_{1\le i\le n, 1\le j\le n} \frac{I(|x_i - x_j| < h)}{n(n-1)h} \omega(x_i)\omega(x_j)g(x_i)g(x_j) + O(n^{-1})$$

Let $s_1 = 0 < \cdots < s_{n+1} = T$ such that $s_i = x_i - \Delta/2$ for $i = 1, \cdots, n$ with $s_i < x_i < s_{i+1}$ for $1 \le i \le n$, and set

$$R = | \sum\sum_{1\le i\le n, 1\le j\le n} \frac{I(|x_i - x_j| < h)m(x_i, x_j)}{n(n-1)h} - \int\int \frac{I(|u - v| < h)m(u, v)}{h} dudv|$$

where $m(x_i, x_j) = \omega(x_i)\omega(x_j)g(x_i)g(x_j)$. By application of the Mean Value Theorem one can show that

$$R = |\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{n(x_i, x_j)}{n(n-1)h} - \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{n(\epsilon_i, \epsilon_j)}{h}(s_{i+1} - s_i)(s_{j+1} - s_j)|$$

where $n(x_i, x_j) = I(|x_i - x_j| < h)m(x_i, x_j)$ and $\epsilon_i, \epsilon_j$ is such that $s_i < \epsilon_i < s_{i+1}$, $s_j < \epsilon_j < s_{j+1}$. Recall that $\Delta = T/n$ and that $g$ is a continuous function therefore

$$R < \frac{M\Delta^2}{h}|\sum\sum_{1\le i\le n, 1\le j\le n} I(|x_i - x_j| < h) - I(|\epsilon_i - \epsilon_j| < h)|$$

where $M = \max\{m(x_i, x_j)g(x_i)g(x_j)\}_{1\le i,j\le n}$. Using the counting argument given in the proof of Lemma 4, one can prove that $R < \frac{Cn\Delta^2}{h} = o(1)$ for some constant $C > 0$. To complete the proof of the lemma, we need to establish $\int_0^T\int_0^T \frac{I(|u-v|<h)}{h}g(u)g(v)dudv = \int_0^T g(u)^2du$. First note that

$$\int_0^T\int_0^T \frac{I(|u-v| < h)}{h}g(u)g(v)dudv = \frac{1}{h}\int_0^T\int_{-h+v}^{h+v} g(u)g(v)dudv$$

Now since $\int_{-h+v}^{h+v} g(u)du = G(-h+v) - G(h+v)$ where $G$ is the antiderivative of $g$ that implies $\int_{-h+v}^{h+v} g(u)du/h \to g(v)$ as $h \to 0$; hence, the result. ■

**Lemma 6**

*Let*

$$A_N^*(h) = \frac{1}{n(n-1)h}\sum\sum_{1\le i\ne j\le n} I_{i,j}(h)\omega(x_i)\omega(x_j)(Y_j^* - \lambda_0^0(x_j))$$

*then assuming* $A_1, A_2, A_3$ *and* $H_0$

$$A_N^*(h) = O_p(N^{-\frac{1}{2}}).$$

Proof: First, observe that

$$\begin{aligned} E\{A_N^*(h)\} &= E\{\tfrac{1}{n(n-1)h}\sum\sum_{1\le i\ne j\le n} I_{i,j}(h)\omega(x_i)\omega(x_j)(Y_j^* - \lambda_0^0(x_j))\} \\ &= \tfrac{1}{n(n-1)h}\sum\sum_{1\le i\ne j\le n} I_{i,j}(h)\frac{\omega(x_i)\omega(x_j)}{N\Delta Q_0^0(x_j)}(Np_j - Nf_0^0(x_j)\Delta). \end{aligned}$$

where $Q_0^0 = 1 - F_0^0$. Since $p_j = \int_{x_j-\Delta/2}^{x_j+\Delta/2} f_0^0(x)dx = \Delta f_0^0(x_j) + \Delta^2 f_0^{0'}(x_j) + o(\Delta^2)$, by Lemma 4 it turns out that $|E\{A_N^*(h)\}| \le \frac{2M\Delta}{TN} + o(\Delta^2)$ with $M = \max_j\{\frac{f_0^{0'}(x_j)}{1-F_0^0(x_j)}\}$. Therefore,

$$E\{A_N^*(h)\} = o(\frac{\Delta}{N^{1+\alpha}}).$$

For the variance analysis, let $D = \frac{1}{n(n-1)h}\frac{1}{N\Delta}$; then, by definition of $Y_i^*$ for $1 \le i \le n$ we write $A_N^*(h)$ as

$$A_N^*(h) = \frac{1}{D}\sum_{k=1}^{N} A_{Nk}^*(h)$$

where

$$A_{Nk}^*(h) = \sum\sum_{1\le i\ne j\le n} \frac{I_{i,j}(h)\omega_{i,j}}{Q(x_j)}(k(x_j, X_k) - f_0^0(x_j)(1 - H(x_j))\Delta).$$

Note that $A_N^*(h)$ is written as a sum of independent random variable. Now, let $\gamma_j = f_0^0(x_j)(1 - H(x_j))\Delta$; then,

$$\begin{aligned} &\tfrac{1}{D^2}A_{Nk}^{*2}(h) \\ = \quad & \tfrac{1}{D^2}\sum\sum\sum\sum_{1\le i\ne j\ne k\ne l\le n} \frac{I_{i,j}(h)I_{k,l}(h)\omega_{i,j,k,l}}{Q(x_l)Q(x_j)}(k(x_j, X_k) - \gamma_j)(k(x_l, X_k) - \gamma_l) \\ & + \tfrac{1}{D^2}\sum\sum\sum_{1\le i\ne j\le n, 1\le k\ne j\le n} \frac{I_{i,j}(h)I_{k,j}(h)\omega_{i,j,k}}{Q(x_j)^2}(k(x_j, X_k) - \gamma_j)^2 \\ & + \tfrac{1}{D^2}\sum\sum_{1\le i\ne j\le n} \frac{I_{i,j}(h)\omega_{i,j}}{Q(x_j)^2}(k(x_j, X_k) - \gamma_j)^2. \end{aligned}$$

Since $k$ is the uniform kernel, it implies that $E[k(x_j, X_k)k(x_l, X_k)] = 0$ for $j \neq l$. In addition, using the fact that $E((k(x_j, X_k) - \gamma_j)^2) = \Delta(f_0^0(x_j)(1 - H(x_j)) + o(\Delta))$, one can show that

$$E[\tfrac{1}{D^2} A_{Nk}^{*2}(h)] = \sum\sum\sum\sum_{1 \leq i \neq j \neq k \neq l \leq n} \frac{I_{i,j}(h) I_{k,l}(h) \omega_{i,j,k,l}}{D^2} v(x_i, x_j)$$
$$+ \sum\sum\sum_{1 \leq i \neq j \leq n, 1 \leq k \neq j \leq n} \frac{I_{i,j}(h) I_{k,j}(h) \omega_{i,j,k}}{D^2} \frac{\lambda_0^0(x_j)}{Q(x_j)} + \sum\sum_{1 \leq i \neq j \leq n} \frac{I_{i,j}(h) \omega_{i,j}}{D^2} \frac{\lambda_0^0(x_j)}{Q(x_j)}$$

and by Lemma 4

$$|E[\frac{1}{D^2} A_{Nk}^{*2}(h)]| \leq \frac{2M_1}{T^2 n N^2 h} + \frac{4M_1}{T^2 N^2} + \frac{8M_2}{T N^2} + o(\frac{8}{N^2 \Delta^2})$$

where $v(x_i, x_j) = -\frac{\lambda_0^0(x_i) f_0^0(x_j)}{Q(x_j)} - \frac{\lambda_0^0(x_j) f_0^0(x_i)}{Q(x_i)} + \frac{f_0^0(x_j)}{Q(x_j)}$ and $M_1 = \max_j \{\frac{\lambda_0^0(x_j)}{Q(x_j)}\}$ and $M_2 = \max_{j,l}\{\lambda_0^0(x_j)\lambda_0^0(x_l)\}$. Consequently,

$$|E[A_N^{*2}(h)]| \leq \frac{2M_1}{T^2 n N h} + \frac{4M_1}{T^2 N} + \frac{8M_2}{T N} + o(\frac{8}{N \Delta^2}).$$

■

**Lemma 7**

*Let* $J_{N1} = (N\Delta)^{-2} \sum_{k=1}^{N} J_{N1k}$ *where*

$$J_{N1k} = \frac{1}{n(n-1)h} \sum\sum_{1 \leq i \neq j \leq n} \frac{I_{i,j}(h) \omega_{i,j}}{Q(x_i) Q(x_j)} (k(x_i, X_k) - E_{i,k})(k(x_j, X_k) - E_{j,k})$$

*where* $E_{i,k} = Ek(x_i, X_k)$. *Then under assumptions* $A_1, A_2, A_3$ *and* $H_0$

$$EJ_{N1k} = \Delta^2 \int \lambda_0^0(x)^2 \omega(x) dx + o(\Delta^2) \tag{3.9}$$

*and*

$$E(^2_{N1k} = O(\frac{\Delta^4}{h}) \tag{3.10}$$

*as* $\Delta \to 0$.

Proof: Noting that

$$
\begin{aligned}
J_{N1k} &= \tfrac{1}{n(n-1)h} \sum\sum_{1\le i\ne j\le n} \frac{I_{i,j}(h)\omega_{i,j}}{Q(x_i)Q(x_j)} \Big( k(x_i,X_k)k(x_j,X_k) \\
&- k(x_i,X_k)E(k(x_j,X_1)) - k(x_j,X_k)E(k(x_j,X_1)) \\
&+ E(k(x_i,X_1))E(k(x_j,X_1)) \Big),
\end{aligned}
$$

it follows that

$$
EJ_{N1k} = -\tfrac{1}{n(n-1)h} \sum\sum_{1\le i\ne j\le n} \frac{I_{i,j}(h)\omega_{i,j}}{Q(x_i)Q(x_j)} E(k(x_i,X_k))E(k(x_j,X_k))
$$

since $E[k(x_i,X_k)k(x_j,X_k)] = 0$ (i.e. $X_k$ cannot be in two different intervals at the same time). Now $E[k(x_j,X_1)] = \Delta f_0^0(x_j)\bar{H}(x_j) + o(\Delta)$ implies that

$$
EJ_{N1k} = -\sum\sum_{1\le i\ne j\le n} \frac{I_{i,j}(h)\omega_{i,j}\Delta^2 f_0^0(x_j)\bar{H}(x_j)f_0^0(x_i)\bar{H}(x_i)}{n(n-1)hQ(x_i)Q(x_j)} + o(\Delta^2).
$$

Finally, using the Riemann integral approximation we deduce that

$$
EJ_{N1k} = -\frac{\Delta^2}{T^2} \int \lambda_0^0(x)^2\omega(x)dx + o(\Delta^2).
$$

Since using similar argument, one can shown that $EJ_{N1k}^2 = O(\frac{\Delta^4}{h})$, we omit the details. ■

**Lemma 8**

*Under conditions $A_1, A_2, A_3$ and $H_0$*

$$
EH_N^2(X_1,X_2) = \frac{2\Delta^4}{T^4 h} \int (\frac{\lambda_0^0(x_i)}{Q(x)})^2\omega(x)dx + O(\Delta^4), \tag{3.11}
$$

$$
EH_N^4(X_1,X_2) = O(\frac{\Delta^8}{h^3}) \tag{3.12}
$$

*and*

$$
EG_N^2(X_1,X_2) = O(\Delta^8) \tag{3.13}
$$

as $\Delta \to 0$ and $h \to 0$.

Proof: $H_N(X_1, X_2)$ is a degenerate U-statistic since

$$\begin{aligned} EH_N(X_k, X_l)|X_l &= \tfrac{1}{n(n-1)h} \sum\sum_{1\leq i\neq j\leq n} \frac{I_{i,j}(h)\omega_{i,j}}{Q(x_i)Q(x_j)} E\{(k(x_i, X_k) - E_{x_i,k}) \\ &\quad \times (k(x_j, X_l) - E_{x_j,l})\} \\ &= 0. \end{aligned}$$

Now considering the fact that

$$\begin{aligned} H_N(X_k, X_l)^2 &= \tfrac{1}{(n(n-1))^2h^2} \sum_{\substack{i_1=1 \\ i_1\neq j_1}}^{n} \sum_{j_1=1}^{n} \sum_{\substack{i_2=1 \\ i_2\neq j_2}}^{n} \sum_{j_2=1}^{n} I_{i_1,j_1}(h) I_{i_2,j_2}(h) \omega_{i_1,j_1,i_2,j_2} \\ &\quad \times (\tfrac{k(x_{i_1},X_k)}{Q(x_{i_1})} - E\tfrac{k(x_{i_1},X_k)}{Q(x_{i_1})})(\tfrac{k(x_{j_1},X_l)}{Q(x_{j_1})} - E\tfrac{k(x_{j_1},X_l)}{Q(x_{j_1})}) \\ &\quad \times (\tfrac{k(x_{i_2},X_k)}{Q(x_{i_2})} - E\tfrac{k(x_{i_2},X_k)}{Q(x_{i_2})})(\tfrac{k(x_{j_2},X_l)}{Q(x_{j_2})} - E\tfrac{k(x_{j_2},X_l)}{Q(x_{j_2})}). \end{aligned}$$

and that $k$ is the uniform kernel,

$$EH_N(X_k, X_l)^2 = \Big( \sum\sum_{1\leq i_1\neq j_1\leq n} \frac{I_{i_1,j_1}(h)\omega_{i_1,j_1}}{(n(n-1))^2h^2} h(x_{i_1}, X_k) h(x_{j_1}, X_l) + L \Big)$$

where $h(x_{i_1}, X_k) = E\{(\frac{k(x_{i_1},X_k)}{Q(x_{i_1})} - E\frac{k(x_{i_1},X_k)}{Q(x_{i_1})})^2\}$ and $L$ contains terms of the form :

1. $\sum\sum\sum \frac{I_{i_1,j_1}(h)I_{i_1,j_2}(h)}{(n(n-1))^2h^2} Eh(\cdot,\cdot)Eh(\cdot,\cdot)Eh^2(\cdot,\cdot)$
2. and $\sum\sum\sum\sum \frac{I_{i_1,j_1}(h)I_{i_2,j_2}(h)}{(n(n-1))^2h^2} Eh(\cdot,\cdot)Eh(\cdot,\cdot)Eh(\cdot,\cdot)Eh(\cdot,\cdot)$.

In particular, one can shown that $L = O(\Delta^4)$ using similar arguments presented in Lemma 6. Therefore, by the Riemann integral approximation and with $h(x_{i_1}, X_k) = \Delta \frac{\lambda_0^0(x_{i_1})}{Q(x_{i_1})} + O(\Delta^2)$,

$$EH_N(X_k, X_l)^2 = \tfrac{2\Delta^4}{T^4h} \int (\tfrac{\lambda_0^0(x)}{Q(x)})^2 \omega(x)dx + O(\Delta^4).$$

This proves (3.11).

To prove (3.12), first, let us consider the expansion

$$
\begin{aligned}
H_N(X_k,X_l)^4 &= \tfrac{1}{(n(n-1))^4h^4}\sum_{i_1\neq j_1}\sum_{i_2\neq j_2}\sum_{i_3\neq j_3}\sum_{i_4\neq j_4} I_{i_1,j_1}(h)I_{i_2,j_2}(h)\\
&\times I_{i_3,j_3}(h)I_{i_4,j_4}(h)\omega_{i_1,j_1,i_2,j_2,i_3,j_3,i_4,j_4}\\
&\times(\tfrac{k(x_{i_1},X_k)}{Q(x_{i_1})}-E\tfrac{k(x_{i_1},X_k)}{Q(x_{i_1})})(\tfrac{k(x_{j_1},X_l)}{Q(x_{j_1})}-E\tfrac{k(x_{j_1},X_l)}{Q(x_{j_1})})\\
&\times(\tfrac{k(x_{i_2},X_k)}{Q(x_{i_2})}-E\tfrac{k(x_{i_2},X_k)}{Q(x_{i_2})})(\tfrac{k(x_{j_2},X_l)}{Q(x_{j_2})}-E\tfrac{k(x_{j_2},X_l)}{Q(x_{j_2})})\\
&\times(\tfrac{k(x_{i_3},X_k)}{Q(x_{i_3})}-E\tfrac{k(x_{i_3},X_k)}{Q(x_{i_3})})(\tfrac{k(x_{j_3},X_l)}{Q(x_{j_3})}-E\tfrac{k(x_{j_3},X_l)}{Q(x_{j_3})})\\
&\times(\tfrac{k(x_{i_4},X_k)}{Q(x_{i_4})}-E\tfrac{k(x_{i_4},X_k)}{Q(x_{i_4})})(\tfrac{k(x_{j_4},X_l)}{Q(x_{j_4})}-E\tfrac{k(x_{j_4},X_l)}{Q(x_{j_4})}).
\end{aligned}
$$

Then, taking term-by-term expectation we have

$$
\begin{aligned}
EH_N(X_k,X_l)^4 &= \tfrac{1}{(n(n-1))^4h^4}\sum_{i_1\neq j_1}\sum_{i_2\neq j_2}\sum_{i_3\neq j_3}\sum_{i_4\neq j_4} I_{i_1,j_1}(h)I_{i_2,j_2}(h)\\
&\times I_{i_3,j_3}(h)I_{i_4,j_4}(h)\omega_{i_1,j_1,i_2,j_2,i_3,j_3,i_4,j_4}\\
&\times E\{(\tfrac{k(x_{i_1},X_k)}{Q(x_{i_1})}-E\tfrac{k(x_{i_1},X_k)}{Q(x_{i_1})})(\tfrac{k(x_{i_2},X_k)}{Q(x_{i_2})}-E\tfrac{k(x_{i_2},X_k)}{Q(x_{i_2})})\\
&\times(\tfrac{k(x_{i_3},X_k)}{Q(x_{i_3})}-E\tfrac{k(x_{i_3},X_k)}{Q(x_{i_3})})(\tfrac{k(x_{i_4},X_k)}{Q(x_{i_4})}-E\tfrac{k(x_{i_4},X_k)}{Q(x_{i_4})})\}\\
&\times E\{(\tfrac{k(x_{j_1},X_l)}{Q(x_{j_1})}-E\tfrac{k(x_{j_1},X_l)}{Q(x_{j_1})})(\tfrac{k(x_{j_2},X_l)}{Q(x_{j_2})}-E\tfrac{k(x_{j_2},X_l)}{Q(x_{j_2})})\\
&\times(\tfrac{k(x_{j_3},X_l)}{Q(x_{j_3})}-E\tfrac{k(x_{j_3},X_l)}{Q(x_{j_3})})(\tfrac{k(x_{j_4},X_l)}{Q(x_{j_4})}-E\tfrac{k(x_{j_4},X_l)}{Q(x_{j_4})})\}.
\end{aligned}
$$

Arguing as before and noting that $k$ is the uniform kernel density, one can rearrange the previous expectation as

$$
EH_N(X_k,X_l)^4 = \sum\sum_{1\le i_1\neq j_1\le n}\frac{I_{i_1,j_1}(h)\omega_{i_1,j_1}}{(n(n-1))^4h^4}(m(x_{i_1},X_k)m(x_{j_1},X_l))+L_1
$$

where

$$
m(x_{i_1},X_k) = \Delta\tfrac{\lambda_0^0(x_{i_1})}{Q^3(x_{i_1})}+O(\Delta^2)\text{and } L_1=O(\Delta^8).
$$

Therefore, by the Riemann integral approximation the preceding expectation becomes

$$EH_N(X_k, X_l)^4 = \tfrac{2\Delta^8}{T^8h^3}\int(\tfrac{\lambda_0^0(x)}{Q(x)^3})^2\omega(x)dx + O(\Delta^8) = O(\tfrac{\Delta^8}{h^3}).$$

To prove 3.13, observe that

$$\begin{aligned} G_N(x,y)^2 &= \tfrac{1}{(n(n-1))^4h^4}\sum_{i_1\neq j_1}\sum_{i_2\neq j_2}\sum_{i_3\neq j_3}\sum_{i_4\neq j_4} I_{i_1,j_1}(h)I_{i_2,j_2}(h) \\ &\times I_{i_3,j_3}(h)I_{i_4,j_4}(h)\omega_{i_1,j_1,i_2,j_2,i_3,j_3,i_4,j_4} \\ &\times E\{(\tfrac{k(x_{i_1},X_1)}{Q(x_{i_1})} - E\tfrac{k(x_{i_1},X_1)}{Q(x_{i_1})})(\tfrac{k(x_{i_2},X_1)}{Q(x_{i_2})} - E\tfrac{k(x_{i_2},X_1)}{Q(x_{i_2})})\} \\ &\times E\{(\tfrac{k(x_{i_3},X_1)}{Q(x_{i_3})} - E\tfrac{k(x_{i_3},X_1)}{Q(x_{i_3})})(\tfrac{k(x_{i_4},X_1)}{Q(x_{i_4})} - E\tfrac{k(x_{i_4},X_1)}{Q(x_{i_4})})\} \\ &\times(\tfrac{k(x_{j_1},x)}{Q(x_{j_1})} - E\tfrac{k(x_{j_1},x)}{Q(x_{j_1})})(\tfrac{k(x_{j_2},y)}{Q(x_{j_2})} - E\tfrac{k(x_{j_2},y)}{Q(x_{j_2})}) \\ &\times(\tfrac{k(x_{j_3},x)}{Q(x_{j_3})} - E\tfrac{k(x_{j_3},x)}{Q(x_{j_3})})(\tfrac{k(x_{j_4},y)}{Q(x_{j_4})} - E\tfrac{k(x_{j_4},y)}{Q(x_{j_4})}). \end{aligned}$$

By similar arguments given in Lemma 6 and the Riemann integration approximation, one can prove that

$$EG_N(X_1, X_2)^2 = O(\Delta^8).$$

■

**Lemma 9**

*Let* $J_{N3} = (N\Delta)^{-1}\sum_{k=1}^{N} J_{N3k}$ *where*

$$J_{N3k} = \frac{1}{n(n-1)h}\sum\sum_{1\leq i\neq j\leq n}\frac{I_{i,j}(h)\omega_{i,j}}{Q(x_j)}(k(x_j, X_k) - Ek(x_j, X_k))(EY_i^* - \lambda_0^0(x_i)).$$

*Then, under conditions* $A_1, A_2, A_3$ *and* $H_0$

$$EJ_{N3k} = 0, \tag{3.14}$$

$$EJ_{N3k}^2 \sim 4\frac{\Delta^6}{T^4}k^4\Big[\int\{\kappa(x)\}^2\frac{\lambda_0^0(x)}{(Q(x))}\omega(x)dx - \{\int(\kappa(x))\lambda_0^0(x)\omega(x)dx\}^2\Big] \tag{3.15}$$

*and*

$$EJ_{N3k}^4 = O(\Delta^{12}) \tag{3.16}$$

*as* $\Delta \to 0$ *where* $\kappa = \frac{f_0^{0\prime\prime}}{1-F}$.

Proof: Let $g(x_i) = (EY_i^* - \lambda_0(x_i))$ for $i \in \{1, \cdots, n\}$; then,

$$J_{N3k} = \frac{1}{n(n-1)h} \sum\sum_{1 \leq i \neq j \leq n} I_{i,j}(h)\omega_{i,j}(\frac{k(x_j, X_k)}{Q(x_j)} - E\frac{k(x_j - X_k)}{Q(x_j)})g(x_i).$$

First, note that $E(J_{N3k}) = 0$, which implies that the variance of $J_{N3k}$ is equal to $E(J_{N1k}^2)$.

Moreover,

$$J_{N3k}^2 = \sum\sum_{1 \leq i \neq j \leq n} \sum\sum_{1 \leq m \neq l \leq n} \frac{I_{i,j}(h)I_{m,l}(h)\omega_{i,j,m,l}}{(n(n-1))^2h^2} z(x_j, X_k, x_l, X_k)g(x_i)g(x_m)$$

where $z(x_j, X_k, x_l, X_k) = (\frac{k(x_j,X_k)}{Q(x_j)} - E\frac{k(x_j,X_k)}{Q(x_j)}) \times (\frac{k(x_l,X_k)}{Q(x_l)} - E\frac{k(x_l,X_k)}{Q(x_l)})$. As a consequence,

$$EJ_{N3k}^2 = \sum\sum_{1 \leq i \neq j \leq n} \sum\sum_{1 \leq m \neq l \leq n} \frac{I_{i,j}(h)I_{m,l}(h)\omega_{i,j,m,l}}{(n(n-1))^2h^2} Ez(x_j, X_k, x_l, X_k)g(x_m)g(x_i).$$

Since $k$ is the uniform kernel, the preceding equation becomes

$$\begin{aligned} EJ_{N3k}^2 = & \sum\sum\sum\sum_{1 \leq i \neq j \neq m \neq l \leq n} \frac{I_{i,j}(h)I_{m,l}(h)\omega_{i,j,m,l}}{(n(n-1))^2h^2} Ez(x_j, X_k, x_l, X_k)g(x_m)g(x_i) \\ & + \sum\sum\sum_{1 \leq i \neq j \neq m \leq n} \frac{I_{i,j}(h)I_{m,l}(h)}{(n(n-1))^2h^2} Ez(x_j, X_k, x_j, X_k)g(x_m)g(x_i) \\ & + \sum\sum_{1 \leq i \neq j \leq n} \frac{I_{i,j}(h)}{(n(n-1))^2h^2} Ez(x_j, X_k, x_j, X_k)g(x_i)^2. \end{aligned}$$

Now replacing the expectations by their analytical expression, using the Riemann integral approximation, and having $g(x_i) = \Delta^2 k\kappa(x_i) + o(\Delta^2)$ imply that

$$\begin{aligned} EJ_{N3k}^2 = & 4\frac{\Delta^6}{T^4}k^2 \int (\kappa(x))^2 \frac{\lambda_0^0(x)}{Q(x)} \omega(x)dx \\ & -4k^2 \tfrac{\Delta^6}{T^4} \{\textstyle\int (\kappa(x))\lambda_0^0(x)\omega(x)dx\}^2 + o(\Delta^6), \end{aligned}$$

which proves (3.15.) Finally, using similar arguments, one can shown that

$$E(J_{N3k}^4) = O(\Delta^{12}).$$

■

Finally, the proofs of corollaries 1 and 2.

Proof of Corollary 1: Under the fixed alternative hypothesis,

$$T_N(h) = \sum\sum_{1\le i<j\le n} \frac{I(|x_j - x_i| < h)}{n(n-1)h} \omega(x_i)\omega(x_j)(Y_i - \lambda_1(x_i))(Y_j - \lambda_1(x_j)),$$

$$+ \sum\sum_{1\le i<j\le n} \frac{I(|x_j - x_i| < h)}{n(n-1)h} \omega(x_i)\omega(x_j)(\lambda_1(x_i) - \lambda_0(x_i))(Y_j - \lambda_0(x_j)),$$

$$+ \sum\sum_{1\le i<j\le n} \frac{I(|x_j - x_i| < h)}{n(n-1)h} \omega(x_i)\omega(x_j)(Y_i - \lambda_0(x_i))(\lambda_1(x_j) - \lambda_0(x_j)),$$

$$+ \sum\sum_{1\le i<j\le n} \frac{I(|x_j - x_i| < h)}{n(n-1)h} \omega(x_i)\omega(x_j)(\lambda_1(x_i) - \lambda_0(x_i))(\lambda_1(x_j) - \lambda_0(x_j)).$$

The last term on the right side is entirely deterministic and converges to

$$\mu_1 = \frac{2}{T^2} \int (\lambda_1(x) - \lambda_0(x))^2 dx,$$

the two terms on the right side in the middle are converging to zero as $N$ increases, and the first term on the right side has an asymptotic normal distribution with mean 0 and variance,

$$\sigma_1^2 = \frac{4}{N^2 h T^4} \int (\frac{\lambda_1}{1 - F_1(x)})^2 dx.$$

Therefore,

$$N\sqrt{h} T_{N0}(h) = N\sqrt{h} T_{N1}(h) + N\sqrt{h}(\int (\lambda_1(x) - \lambda_0(x))^2 dx + o_p(1)),$$

which implies the result. ■

Proof of Corollary 2: To prove Corollary 2, one can use the relation

$$Z^* = N\sqrt{h}T_{N0}(h) = N\sqrt{h}T_{N1}(h) + N\sqrt{h}(\int(\lambda_1(x) - \lambda_0(x))^2 dx + o_p(1)).$$

Under $H_{1N}$,

$$Z^* = N\sqrt{h}T_{N0}(h) = N\sqrt{h}T_{N1}(h) + N\sqrt{h}N^{-2\beta}(\int(\eta(x))^2 dx + o_p(1)).$$

Now, observe that $N\sqrt{h}N^{-2\beta} \to \infty$ for $\beta < (2-\delta)/4$ while when $\beta > (2-\delta)/4$, it converges to zero. This proves the first part of the corollary. The second part of the corollary can be proved in a similar fashion. That is, under $H_{2N}$,

$$Z^* = N\sqrt{h}T_{N0}(h) = N\sqrt{h}T_{N1}(h) + N\sqrt{h}N^{-2\epsilon}N^{\gamma}(\int(\eta(x))^2 dx + o_p(1)).$$

In this case $N\sqrt{h}N^{-2\epsilon}N^{\gamma} \to \infty$ for $1 - 2\epsilon - \gamma > \delta/2 > 0$, while for $1 - 2\epsilon - \gamma < \delta/2$, it converges to zero. ∎

# CHAPTER IV

# A SIMULATION STUDY

## 4.1 Introduction

The asymptotic distribution of the test statistic $T_N(h)$ was established in chapter 3; in this chapter, to complement the theoretical findings of Chapter 3, we shall conduct a series of simulation studies. First, in Section 4.2, a verification of the distribution of the kernel-based nonparametric g-o-f test under the null hypothesis is carried out by comparing the plot of the distribution of the test statistic, $T_N(h)$, at a fixed sample size to the plot of the standard normal distribution. Second, in Section 4.3, the performance and power properties of $T_N(h)$ are examined while in Section 4.4 we present an application of the test to real life data. Furthermore, one of the objectives of these stimulation studies is to evaluate the influence of the parameter $h$ on the performance of $T_N(h)$ and to provide some suggestions on how to choose objectively $h$. Therefore, in the last section, Section 4.5, a Monte Carlo procedure is proposed in order to select $h$ and we also investigate the influence of the selected smoothing parameter on the observed power of the test.

To study the performance and power properties of the test in finite-samples, in this chapter, we restrict ourselves to certain well-known hazard models, and for that we consider the hazard rate models given in Chapter 2. That is, (1) the Weibull (W), (2) the lognormal (LN), (3) the folded normal (FN), (4) the Birnbaum-Saunders (BS), and (5) the generalized gamma (GG) distributions. In addition, in order to evaluate the power properties of $T_N(h)$, we present a detailed simulation study case for the Weibull hazard rate model since it is a classical hazard rate function, which is widely used to model different life testing experiments, and, also because complex hazard rate models such as the bathtub hazard rate model can be simulated by mixing Weibull hazard rate functions.

## 4.2 Verification of the Asymptotic Distribution of $T_N(h)$ under $H_0$

In this section, in order to illustrate the functional shape of the distribution of the test $T_N(h)$ in finite-samples, we plot the distribution function of $\sigma_0^{-1}T_N(h)$ where $\sigma_0$ is the standard deviation of $T_N(h)$ at different values of $h$ and $N = 100$.

In this simulation study, the censored dataset is generated using the simulation procedure presented in Chapter 2, Section 2.4. That is, we generate two independent data sets, one for the observed events ($x_i^0, i = 1, \cdots, n$ ) and one for the unobserved events ($u_i, i = 1, \cdots, n$); then, we determine the censored data set using the relation $x_i = \min\{x^0, u_i\}$. For each censored data set, we compute the observed value of the standardized test statistic

$$Z_N(h) = \sigma_0^{-1} \sum\sum_{1 \leq i \neq j \leq n} \frac{I(|x_i - x_j| < h)}{n(n-1)h} \omega(x_i)\omega(x_j)(Y_j - \lambda^{(k)}(x_j))(Y_i - \lambda^{(k)}(x_i)) \quad (4.1)$$

at some chosen values of the pair of parameters $(h, \Delta)$ where $T_N(h)$ is computed using equation (3.1). In particular, we take $\Delta = c_1 N^{-\alpha}$ where $\alpha = 3.5/5, c_1 = 1$, $h = 0.04$, $0.05$, $0.06$, $0.07$, respectively and $\omega(x) = I(0 \leq x \leq T_0)$ with $T_0$ equals to the 80th percentile of the distribution function associated with the hypothesized hazard rate function. The foregoing process is replicated 1000 times for each $h$. Then, we compute a kernel density estimator, $\sum_{j=1}^{n_0} \frac{1}{nb} K(\frac{x - x_j}{b})$, for the observed values of the test statistic $Z_N(h)$

where $K$ is the Gaussian Kernel and the normal reference bandwidth $b = 1.06\hat{s}_{n_0} n_0^{-1/5}$ with $n_0 = 1000$, see Fan and Gijbels [23], Wand and Jones [37] and Silverman [33].

The function `bkde` in R is used to compute the kernel density estimate of $Z_N(h)$ and the function `plot` to plot it. The plots are displayed in Figure 4.1 at $N = 100$ and $h = 0.04, 0.05, 0.06, 0.07$. The solid line represents the kernel density estimate of the distribution function of the test statistic $Z_N(h)$, whereas the dashed line is the plot of the standard normal distribution $N(0, 1)$. Noticed that, visually, the difference between the kernel density estimate of the density function of $Z_N(h)$ and the standard normal density functiion is at their minimum for $h = 0.07$. In this simulation study, the data set is censored up to 15%.

### 4.3 Structure of the Simulation Study

The null hypothesis we want to test is

$$H_0 : \lambda^0(x) = \lambda_0^0(x) \text{ for every } x \in \mathbb{R}^+ \tag{4.2}$$

where $\lambda_0^0$ is the hazard rate associated with a given distribution function, say the Weibull distribution for example. The alternative hypothesis is given as

$$H_1 : \lambda^0(x) = \lambda_1^0(x) \text{ for every } x \in \mathbb{R}^+ \tag{4.3}$$

where $\lambda_1^0$ corresponds to a hazard rate function that differs from $\lambda_0^0$.

In each simulation study, the number of replicated data sets is set to 1000 datasets. Then, the observed critical values of the test statistic are computed using the 1000 values of the observed test statistic $T_N(h)$. To illustrate, for $\alpha = 0.05$, the observed critical value

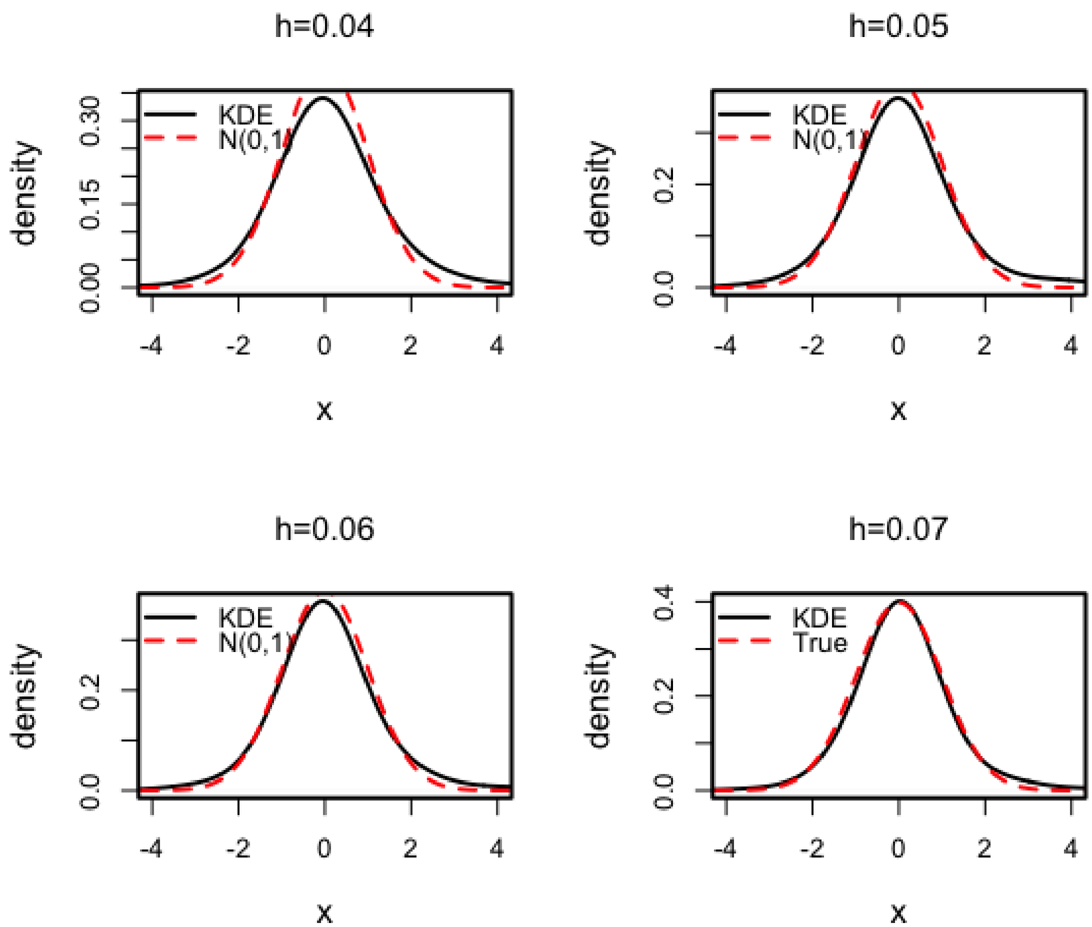


Figure 4.1: Kernel density estimation (KDE) for the observed values of $\sigma_0^{-1}T_N(h)$ under the null hypothesis for censored data (15%) at $h = 0.04, 0.05, 0.06, 0.07$ and $N = 100$.

denoted by $\hat{c}_{0.05}$ is the 950-$th$ value of the increasing sequence of the 1000 observed values of $T_N(h)$. As one can observe the rate at which $h \to 0$ depends on the parameter $\Delta$ since $h/\Delta \to \infty$ as $N \to \infty$. It is known that kernel-based nonparametric goodness-of-fit test are very sensitive to the choice of the smoothing parameter; hence, producing observed critical values that are much greater than the true critical value of the test see Fan [13].

### 4.3.1 Observed Critical Values

Based on the panel given in Figure 4.1, one may conclude that the critical values of the test might be sensitive to the choice of the smoothing parameter. In such a case, a bootstrapping procedure can be helpful in determining more stable cut-off points. In statistical inference, bootstrapping is well known to produce accurate approximation of the distribution of a test under the null hypothesis. Li and Wang [26] showed that bootstrapping under regularity conditions gives more stable critical values, which are robust to the choice of the smoothing parameter.

As to illustrate, we consider the null hypothesis $H_0 : \lambda^0(x) = \lambda_0^0(x),$ for every $x \in \mathbb{R}^+$ where $\lambda_0^0$ is a specified hazard rate function derived from the Weibull distribution with shape parameter equal to 1.5 and scale parameter equal to 1 with density function given by $f(x) = \frac{3}{2}x^{1/2}\exp(-x^{3/2})$ for $x > 0$. Then, we resample a sample of size $N$ generated from the Weibull distribution and calculate the statistic $T_N(h)$ using equation 3.1. The foregoing process is repeated $M$ times for $M$ very large. Table 4.1 shows the estimated cut-off points for $10000$ values of $Z_N(h)$ under the null hypothesis for $h = 0.3N^{-1/4}$ and $\Delta = N^{-3.5/5}$. The cut-off points represent the 95th percentile of the replicated data and are obtained using the function `quantile` in R with type=7. As $N \to \infty$ the estimated cut-off point also converges to the nominal critical value (1.65) of the limiting distribution of $Z_N(h)$ at $\alpha = 0.05$; however, one can observe that the convergence is not monotone. As mentioned in the last paragraph, our selected $h$ seems to produce unstable observed critical values. The tables display in Appendix A.2 contain the critical values of the test statistic $Z_N(h)$ $-$the standardized form of $T_N(h)-$ for the hazard rate models enumerated in Section 4.1. As one can note the results are similar to the those in Table 4.1.

Table 4.1: Observed critical values of the test $T_N(h)$ at $N =$ 50,80, $100, 150, 200, 300$.

| Sample size | 50 | 80 | 100 | 150 | 200 | 300 |
|---|---|---|---|---|---|---|
| Cut-off points $(\hat{l})$ | 1.40 | 1.79 | 1.84 | 1.65 | 1.69 | 1.73 |

### 4.3.2 Observed Power

Under the alternative hypothesis the data set is generated from the distribution function associated with the hazard rate function $\lambda_1^0$. This process is replicated 1000 times and the observed power of the test statistic $T_N(h)$ is computed as

$$P(T_N(h) > \hat{c}_{0.05}) = \frac{\#\{T_{N0}(h) > \hat{c}_{0.05}\}}{1000} \tag{4.4}$$

where $\hat{c}_{0.05}$ is the observed critical value obtained using the procedure described in the subsection above. In this section, the observed power of the test is computed using the observed critical values given in Table 4.1.

#### 4.3.2.1 Case 1: Global Alternatives

The null hypothesis is defined as in equation 4.2 where $\lambda_0^0$ is the Weibull hazard rate function with parameter (1.5,1). In this simulation study, the alternative hypothesis is similar to equation 4.3 where $\lambda_1^0$ is one the following hazard rate models:

1. the loggormal (LN(0,1)),
2. the folded normal (FN (0,1)),
3. the Birnbaum-Saunders (BS(1.75,1)),
4. the generalized gamma (GG(1,.6,4)).

Those are classical hazard rate functions that are often studied in survival analysis in the context of hazard rate model assessment [24]. Recall that their plots are given in Figure 2.2. Table 4.2 contains the observed power of the test against the folded normal, the Birnbaum-Saunders, the generalized gamma, and the lognormal distributions, respectively for $h = 0.3N^{-1/4}, \Delta = N^{-3.5/5}$. As $N \to \infty$ the estimated power of the test converges to one, and as in the simulation section in Chapter 2 the power of test seems to be lower against the folded normal(0,1) hazard rate function.

#### 4.3.2.2 Case 2: Local Alternatives

In the last subsection, we showed that the test statistic $T_N(h)$ is consistent against fixed alternatives. Here, we examine the observed power properties of the kernel-based nonparametric g-of test against local alternatives. Therefore, the null hypothesis and alternative hypotheses are expressed as in equations 4.2 and 4.3, respectively; taking $\lambda_0$ as the Weibull hazard rate function is (1.5,1) while $\lambda_1^0$ is a Weibull hazard rate function with shape parameter ($s$) and $s$ is equal to $1.75, 2, 2.25$ and $2.5,$ respectively. Figure 4.2 displays the

Table 4.2: Observed power of the test $T_N(h)$ against fixed alternatives when $\lambda_0^0$ is the Weibull hazard rate function with parameters (1.5,1).

| $N$ | FN(0,1) | BS(1.75) | LN(0,1) | GG(1,0.6,4) |
|---|---|---|---|---|
| | | Observed Power | | |
| 50 | 0.12 | 0.90 | 0.12 | 0.77 |
| 80 | 0.13 | 1.00 | 0.44 | 0.97 |
| 100 | 0.14 | 1.00 | 0.50 | 0.99 |
| 200 | 0.19 | 1.00 | 0.93 | 1.00 |
| 300 | 0.20 | 1.00 | 0.98 | 1.00 |

plots of the Weibull hazard rate models for $s$=1.5, 1.75, and 2, respectively. Observe that as the shape parameter $s$ is approaching 1.5, the $W(s,1)$ hazard rate functions also converges to the $W(1.5,1)$ hazard rate function. Table 4.3 contains the estimated power of the test statistic against the Weibull hazard rate functions where $s \geq 1.75$, $T = 80th$ percentile of the Weibull distribution with parameters (1.5,1), $h = 0.3N^{-1/4}$, and $\Delta = N^{-3.5/5}$. As $N$ increases, the observed power converges to one as expected.

#### 4.3.2.3 Case 3: Pitman Alternatives

The Pitman alternative is one of the difficult type of local alternatives to detect since as $N$ increases the alternative hypothesis converges to the null hypothesis. In this simulation study, the Pitman alternative is in the form of

$$H_{1N} : \lambda_1^0(x) = \lambda_0^0(x) + N^{-\beta}\eta(x) + o(N^{-\beta}) \text{ for every } x \in \mathbb{R}^+$$

where $0 < \beta$ and $\lambda_0^0$ is the Weibull. Observe that as $\beta$ increases it becomes harder for the test to distinguish the null hypothesis from the alternative hypothesis.

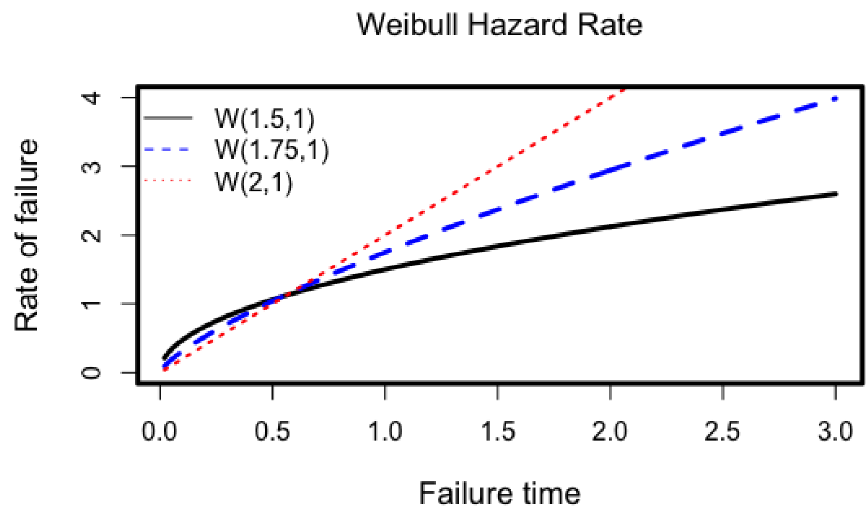


Figure 4.2: Example of Weibull hazard rate functions.

Table 4.3: Observed power of the test $T_N(h)$ when the null and alternative hypotheses are both Weibull hazard rate functions.

| $N$ | W(1.75,1) | W(2,1) | W(2.25,1) | W(2.5,1) |
|---|---|---|---|---|
| | | Observed power | | |
| 50 | 0.14 | 0.26 | 0.43 | 0.61 |
| 100 | 0.15 | 0.40 | 0.65 | 0.87 |
| 200 | 0.24 | 0.66 | 0.95 | 1.00 |
| 300 | 0.24 | 0.76 | 0.98 | 1.00 |

The estimated power is computed using the procedure described at the beginning of this section. Figure 4.3 compares the plots of Weibull hazard rate functions for different $\beta$'s. Model 0 is the plot of the W(1.5,1) hazard rate function (i.e. the null hypothesis), Model 1 is the plot of the W(1.5,1) hazard rate function augmented by $N^{-\beta}\eta(x)$ where $N = 100, \beta = 0.3$, and $\eta$ equal to the W(1.5,1) hazard rate function, and Model 2 is similar to Model 1 except that $\beta = 0.01$. The results in Table 4.4 show as expected that the power fo the test increases to one as $\beta$ decreases.

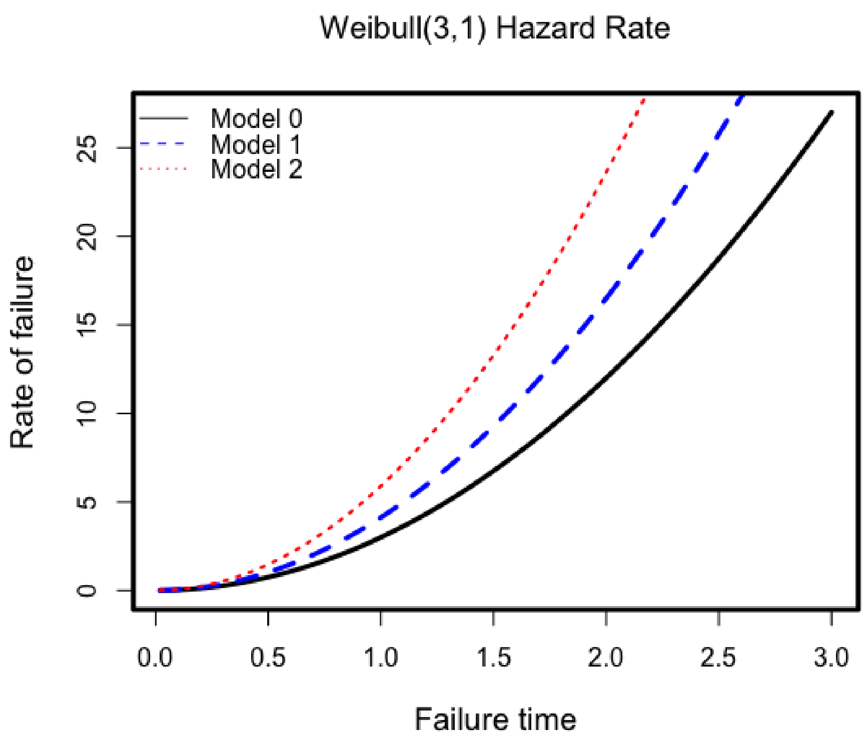


Figure 4.3: Example of Pitman hazard rate functions.

Table 4.4: Observed power of the test $T_N(h)$ against Pitman alternative when $\lambda_0^0$ is the Weibull hazard rate function.

| $\beta$ | 0.3 | 0.2 | 0.1 | 0.01 |
|---|---|---|---|---|
| | | W(1.5,1) | | |
| Power | 0.296 | 0.476 | 0.782 | 0.953 |

## 4.4 Application

In this section, we apply the kernel-based nonparametric g-o-f test statistic, $T_N(h)$, to two sets of data found in Bagdonavicius, Kruopis and Nikulin, Chapter 2 [4]. In their book, Bagdonavicius, Kruopis and Nikulin modeled the first and second dataset using an exponential and a Weibull hazard rate models, respectively. In addition, the first data set is a type II right-censored data while the second is an independent right-censored data set. A failure time data set of size $n$ is a type II right-censored data if the experiment ends after $k$ items have failed. Recall that the independent right-censored data is defined in Section 1.1. The authors did not provide any information concerning the collection of the data sets.

To be able to implement the nonparametric g-o-f test statistic proposed in Chapter 3, we use the function `survreg` in R to estimate the unknown parameters of each hypothesized distribution. The function `survreg` also supports different types of censoring mechanism including the ones we mention in this paragraph.

#### 4.4.0.1 Exponential Hazard Rate Model

Similar to Bagdonavicius, Kruopis and Nikulin for the first dataset we consider the null hypothesis

$$H_0 : \lambda(t) = \mu, \; t \geq 0 \tag{4.5}$$

which means that $\lambda$ is the hazard rate function of the exponential distribution with a parameter $\mu$ and a cumulative distribution function, $1 - \exp(-\mu t) \; t \geq 0$. The p-values of the test are given in Table 4.5 for different values of $h$.

Table 4.5: P-value under $H_0 : \lambda(t) = \mu, \; t \geq 0$ at $h = 40, 50, 60$, and $70$, respectively.

| $h$ | 40 | 50 | 60 | 70 |
|---|---|---|---|---|
| P-value | 1 | 1 | 1 | 1 |

Using the nonparametric test statistic for hazard rate functions described in Section 1.3 the authors found a p-value equal to 0.8945. Also note that the smoothing parameters $h$ are relatively large; this is a consequence of the range of the dataset.

#### 4.4.0.2 Weibull Hazard Rate Model

Following Bagdonavičius, Kruopis and Nikulin we also test the second data set under the null hypothesis

$$H_0 : \lambda(t) = \frac{\gamma}{\theta}(t/\theta)^{\gamma-1}, \; t \geq 0. \tag{4.6}$$

In this case, $\lambda$ is the hazard rate function associated with the Weibull distribution where the scale parameter is $\theta$ and shape parameter is $\gamma$. Its cumulative distribution function is written as $1 - \exp(t/\theta)^{\gamma}$. The p-values of our test are given in Table 4.6 with four different $h$ values.

Table 4.6: P-value under $H_0 : \lambda(t) = \frac{\gamma}{\theta}(t/\theta)^{\gamma-1},\ t \geq 0$ at $h = 40, 50, 60,$ and $70$, respectively.

| $h$ | 40 | 50 | 60 | 70 |
|---|---|---|---|---|
| P-value | 1.8332e-08 | 4.220134e-07 | 4.220134e-07 | 4.220134e-07 |

This result is similar to the one found in Bagdonavičius, Kruopis and Nikulin [4] even though their p-value ($1.22 \times 10^{-15}$) clearly provides more evidence in favor of the alternative. A complete analysis should also compares the power properties of both tests but the authors did not investigate the power properties of their test against the type of alternatives considered in this chapter. Therefore, we conduct a simulation study in order to compare the power properties of the two test statistics under fixed and Pitman alternatives. For that, we considered

$$H_0 : \lambda^0(x) = \lambda_0^0(x) \text{ for every } x \in \mathbb{R}^+$$

versus

$$H_1 : \lambda^0(x) = \lambda_1^0(x) \text{ for every } x \in \mathbb{R}^+$$

where $\lambda_0$ is associated with the W(1.5,1) and $\lambda_1$ to the W(2,1). Also we consider the case

$$H_0 : \lambda^0(x) = \lambda_0^0(x) \text{ against } H_1 : \lambda^0(x) = \lambda_0^0(x) + N^{-\beta}\lambda_0^0(x) \text{ for every } x \in \mathbb{R}^+$$

where $\lambda_0^0$ is the W(1.5,1) hazard rate function and $\beta = 0.01$. We refer to their test as the BKN test and it is denoted by $Y$ in the tables below. The results are provide in Table 4.7 and Table 4.8 for $h = 0.04$ and $0.07$ and $k = 6$ is the number of class intervals. As one can see, against fixed alternatives their test has greater observed powers than that of $T_N(h)$;

however, against Pitman alternatives $T_N(h)$ shows power greater than 50% at $N = 300$ while the BKN test is unable to distinguish the two hypotheses even at $N = 100$. The observed critical values used in this section are, respectively, 1.66,1.62, 1.64 at $N$ =100,200 and 300.

Table 4.7: Observed power of the test $T_N(h)$ versus observed power of the BKN test against $H_1 : \lambda = \lambda_{W(2,1)}$.

| | $T_N(h)$ | | $Y$ |
|---|---|---|---|
| $h$ | 0.04 | 0.07 | $k = 6$ |
| 100 | 0.348 | 0.376 | 0.965 |
| 200 | 0.610 | 0.619 | 1.000 |
| 300 | 0.556 | 0.575 | 1.000 |

### 4.5 Bandwidth Selection

In nonparametric kernel testing, finding the optimum bandwidth consists of selecting the smoothing parameter(s) such that the power function is maximized for a specified size level. For instance, Gao and Gijbels [15] have proposed a bandwidth selection procedure based on the Edgeworth expansion of the power and the size functions; however, their technique is more appropriate for large sample size since the power function and the size function are obtained by using expansion techniques. In failure time experiments it is common to encounter relatively small sample size, and using a procedure based on function's expansion tends to produce highly biased estimators. Therefore, bootstrapping constitutes

Table 4.8: Observed power of the test $T_N(h)$ versus observed power of the BKN test against Pitman alternatives.

| | $T_N(h)$ | | $Y$ |
|---|---|---|---|
| $h$ | 0.04 | 0.07 | $k=6$ |
| 100 | 0.894 | 0.890 | 0.060 |
| 200 | 0.558 | 0.559 | 0.065 |
| 300 | 0.500 | 0.508 | 0.040 |

an alternative procedure to select the bandwidth parameter, but ,in this section, we propose a different method based on the Monte Carlo (MC) approach.

First, we generate $M$ samples of size $N$ from the distribution function associated with the hazard rate function under the null hypothesis (W(1.5,1)); then, for each sample generated we compute the observed statistic $T_N(h)$. The selected bandwidth is therefore the value of $h$ that minimizes the percentile distance function defined as

$$D(h) = \sum_{i=1}^{99}(d_i - \hat{d}_i)^2 \tag{4.7}$$

where $d_i$ is the theoretical percentile of the test distribution and $\hat{d}_i$ the estimated one. The estimated percentile are chosen from the $M$ observed values of $N\sqrt{h}\sigma_0^{-1}T_N(h)$ using the `quantile` function in R with type=7. Moreover, by Theorem 5 $N\sqrt{h}\sigma_0^{-1}T_N(h)$ is asymptotically a standard normal distribution; hence, the theoretical percentiles are computed using the function `qnorm` in R.

**Remark 2** *First, one can observe that $D(h)$ is the numerical version of the well-known Q-Q plot procedure used in descriptive statistic when checking for normality assumption. Since the limiting distribution of $T_N(h)$ is normal, it is expected that the estimated per-*

*centile, obtained from the standardized observed values of $T_N(h)$, to converge to the standard normal percentiles as $N \to \infty$. Second, one can derive a similar procedure based on the integrated square error function $I(\hat{f}) = \int (f(x) - \hat{f}(x))^2 dx$ where $\hat{f}$ is a nonparametric estimate of $f$.*

Furthermore, we take a sequence of equally spaced bandwidths from $\mathcal{H}_m = \{h_1, \cdots, h_m\}$ where $h_1 = a\Delta$ and $h_m = 16\Delta$ with $a > 1$. Since the choice of $h_1$, $h_m$, and $m$ depends on the user of the test, we let $m = 20$. Finally, the selected bandwidth $h^*$ is such that

$$h^* = \arg \min_{h \in \mathcal{H}_m} D(h).$$

Table 4.9 contains the chosen bandwidth for sample size $N = 50, 80, 100, 150,$ and $200$. Note that $h^*$ converges to zero as $N$ increases to infinity with $Nh^*$ being a non decreasing sequence. The second row of Table 4.9 contains their corresponding observed critical values. It seems that although these values are less than the ones given in Table 4.1, the MC procedure produces more stable observed critical values. This result is the same for the other hazard rate functions considered in this simulations studies. The observed critical values and power of the test when the hypothesized function is the folded normal, the lognormal, the generalized gamma, and the Birnbaum-Saunders distributions, respectively, are given in Appendix A.2. In the next subsection, we investigate their influence on the power properties of $T_N(h)$.

Table 4.9: Selected bandwidth based on a percentile distance function for the the Weibull(1.5,1) hazard rate function.

| Sample size | 50 | 80 | 100 | 150 | 200 |
|---|---|---|---|---|---|
| $h^*$ | 0.12 | 0.09 | 0.07 | 0.05 | 0.04 |
| $\hat{c}_{0.05}$ | 1.83 | 1.52 | 1.71 | 1.59 | 1.62 |

### 4.5.1 Observed Power of the Test for the MC-driven Bandwidth

Lastly, we carry out a simulation study in order to evaluate the influence of $h^*$ on the test. For that, consider the null hypothesis

$$H_0 : \lambda^0(x) = \lambda_0^0(x) \text{ for every } x \in \mathbb{R}^+$$

where $\lambda_0^0$ is the Weibull hazard rate with parameters (shape=1.5, scale=1) against the alternative hypothesis

$$H_1 : \lambda^0(x) = \lambda_1^0(x) \text{ for every } x \in \mathbb{R}^+$$

where $\lambda_1$ is the Weibull hazard rate with shape parameters equal to 2.5 and 2, respectively and scale=1. The results are given in Table 4.10 for sample sizes equal to 50, 80,100,150 and 200, respectively. As expected the test is consistent that is as $N \to \infty$, the observed power increases to one.

Table 4.10: Observe power of the test $T_N(h)$ for the MC-driven smoothing parameters.

| $N$ | 50 | 80 | 100 | 150 | 200 |
|---|---|---|---|---|---|
| | | | $H_1 : \lambda = \lambda_{W(2.5,1)}$ | | |
| Power | 0.577 | 0.782 | 0.816 | 0.932 | 0.969 |
| | | | $H_1 : \lambda = \lambda_{W(2,1)}$ | | |
| Power | 0.229 | 0.267 | 0.339 | 0.410 | 0.472 |

CHAPTER V

# CONCLUSION

## 5.1 Conclusion

In this thesis, two nonparametric goodness-of-fit tests − a chi-square goodness-of-fit test and a kernel-based nonparametric g-o-f test − were proposed to check whether or not a hazard rate model assesses appropriately the rate of failure of given failure times data. The main objective was to study and establish their asymptotic distribution. The work began in Chapter 2 with a presentation of the motivations behind the chi-square goodness-of-fit test for hazard rate functions. Then, we showed via the Central Limit Theorem for multivariate variables that under regularity conditions and when the null hypothesis is completely known, the proposed test converges to a chi-square with degrees of freedom equal to the number of class intervals, whereas when the null hypothesis was known up to the parameter $\theta$, the test converges in distribution to a chi-square with degrees of freedom equal to the number of class intervals minus the dimension of the parameter space. Furthermore, in the simulation studies, as expected we found that the estimated critical values $\hat{c}_\alpha$ converged to the nominal critical values $c_\alpha$ for $\alpha = 0.05$ and $0.10$, respectively. We also proved that against fixed alternatives the estimated power of the chi-square goodness-of-fit test for hazard rate functions converges to one.

Besides the chi-square g-o-f test for hazard rate functions, a kernel-based nonparametric g-o-f test for hazard rate functions was constructed by way of expansion of the Bickel-Rosenblatt kernel-based goodness-of-fit test. The rationale behind the motivation and the construction of this test was explained in Chapters 1 and 3. Accordingly, under regularity conditions given in Chapter 3 and using the Central Limit Theorem of Hall [18], we established the limiting distribution of the test statistic $T_N(h)$ under the null hypothesis $H_0 : \lambda^0(x) = \lambda_0^0(x)$ for every $x \in \mathbb{R}^+$ as well as under the alternative hypothesis $H_1 : \lambda^0(x) = \lambda_1^0(x)$ for every $x \in \mathbb{R}^+$. It should be noted that against alternative hypotheses in the form of $H_{1N,}$ the kernel-based nonparametric g-o-f test statistic, $T_N(h)$, has non-trivial powers for $\beta = (2 - \delta)/2$; in addition, the power of the test converges to one for $0 < \beta < 1/2$. More importantly, unlike the B-R test, as $\delta$ decreases to 0 the power of the test approaches one while the convergence rate of the test increases to one. For instance, if we let $\delta = 1/8$, our kernel-based nonparametric g-o-f test converges to the normal distribution at the rate of $N^{15/16}$, which is greater than the convergence rate of the Bickel-Rosenblatt test (i.e. $N^{1/16}$). Against alternative hypotheses in the form of $H_{2N}$ the order of magnitude of the integral of $N^{-\epsilon} \sum_{j=1}^{k} \eta_j(\frac{x-c_j}{N^{-\gamma}})$ can be chosen to be equal to $n^{-d}$ where $d > 1/2$. For instance, one can let $\epsilon = 1/6$ and $\gamma = 5/12$; as a consequence, our proposed kernel-based nonparametric g-o-f test will have power bounded away from 0 and 1 when $\delta = 1/2$ while tests like Kolmogorov- Smirnov test shows a power, which converges to the nominal significance level.

A simulation study was conducted in Chapter 4 to evaluate the performance and the power properties of the kernel-based nonparametric g-o-f test statistic, $T_N(h)$, in finite-samples. The results showed as expected that against fixed alternatives the power of the test increases to one as the sample size tends to infinity. In particular, under certain conditions, the test shows power properties against Pitman alternative in the form of $H_{1N}$ that are similar to those of the B-R test. Moreover, a comparison was conducted in Chapter 4 with the Pearson-type g-o-f test for hazard rate functions proposed by Bagdonavičius, Kruopis and Nikulin (BKN) [4]. We found that their test showed greater power against fixed alternative than $T_N(h)$; however, against Pitman alternatives the BKN test was unable to distinguish the null from the alternative hypothesis, whereas our kernel-based nonparametric test statistic $T_N(h)$ turned out to have relatively good power properties as Corrolory 2 suggested. The simulations also revealed the importance of the choice of the smoothing parameter and its influence on the performance of the test. Henceforth, the smoothing parameter, $h$, shall not be chosen too large or to small, especially for relatively small sample size; accordingly, a Monte Carlo procedure was implemented for one to be able to select in an objective and practical fashion the smoothing parameter $h$. Simulation studies also showed that the test performed well under the null hypothesis and that $T_N(h)$ were consistent against fixed alternatives for the selected smoothing parameter.

### 5.2 Discussion and Future Works

The asymptotic properties of the kernel-based nonparametric g-o-f test for hazard rate functions, in this thesis, were established under the null hypothesis

$$H_0 : \lambda(x) = \lambda_{\theta_0}(x), \text{ for every } x \in \mathbb{R}^+$$

where the parameter $\theta_0$ is known, but it might be of interest to extend $T_N(h)$ to test the null hypothesis

$$H_0 : \lambda(x) = \lambda_\theta(x; \theta), \text{ for every } x \in \mathbb{R}^+$$

where $\theta$ is unknown and lies in a known parameter space. In addition, a possible extension of this work would be to explore the distribution of the test statistic $T_N(h)$ when the data set drawn from a strictly stationary random process satisfying certain regularity conditions.

Furthermore, a Monte Carlo procedure was used to select the bandwidth parameter $h.$ Although this procedure is very practical and convenient, it does not produce bandwidth parameters that are optimal in the sense Gao and Gijbels [15]. A derivation of such a bandwidth might be considered in future researches.

Last but not least, we argued that the test needs to be implemented on a compact support of $\mathbb{R}^+$ in the form of $[0, T]$ since hazard rate estimates are highly biased at the extreme right tail of their domain of definition. Accordingly, in our simulation studies we, arbitrarily, set $T$ to the $80$th percentile of the distribution function associated with the hypothesized hazard rate function. Although the performance of the test statistic is acceptable and consistent, in future work it is worth considering a detailed analysis of the influence of $T$ on the performance and power properties of the test while providing suggestions on how to select $T$.

# BIBLIOGRAPHY


[1] M. G. Akritas, "Pearson-Type Goodness-of-Fit Tests: The Univariate Case," *Journal of the American Statistical Association*, vol. 83, no. 401, 1988, pp. 222–230.

[2] V. Bagdonavicius and M. Nikulin, "Chi-squared tests for general composite hypotheses from censored samples," *Comptes Rendus Mathematique*, vol. 349, no. 3, 2011, pp. 219–223.

[3] V. Bagdonavičius and M. Nikulin, "Chi-squared tests for general composite hypotheses from censored samples," *Comptes Rendus Mathematique*, vol. 349, no. 3, 2011, pp. 219 – 223.

[4] V. B. Bagdonavičius, julius Kruopis, and M. S. Nikulin, *Non-parametric Tests for Censored Data*, ISTE, 2011.

[5] D. Bagkavos, D. Ioannides, and A. Kalamatianou, "A goodness of fit test for the survival function under random right censoring," *Electron. J. Statist.*, vol. 7, 2013, pp. 2550–2576.

[6] P. J. Bickel and M. Rosenblatt, "On Some Global Measures of the Deviations of Density Function Estimates," *Ann. Statist.*, vol. 1, no. 6, 11 1973, pp. 1071–1095.

[7] J. Blum and V. Susarla, "Maximal deviation theory of density and failure rate function estimates based on censored data," *In: Krishnaiah, P.R. (ed.) Mult. Analysis*, 1980, pp. 213–222.

[8] M. D. Burke, "Approximations of some hazard rate estimators in a competing risks model," *Stochastic Processes and their Applications*, vol. 14, no. 2, February 1983, pp. 157–174.

[9] D. R. Cox, "Regression Models and Life-Tables," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 34, no. 2, 1972, pp. 187–220.

[10] D. R. Cox and D. Oakes, *Analysis of Survival Data*, Chapman and Hall, London ? New York, 1994.

[11] H. Cramer, *Mathematical Methods of Statistics*, 1946.

[12] S. Diehl and W. Stute, "Kernel density and hazard function estimation in the presence of censoring," *Journal of Multivariate Analysis*, vol. 25, no. 2, 1988, pp. 299 – 310.

[13] Y. Fan, "Testing the Goodness of Fit of a Parametric Density Function by Kernel Method," *Econometric Theory*, vol. 10, no. 2, 1994, pp. 316–356.

[14] T. S. Ferguson, "A Method of Generating Best Asymptotically Normal Estimates with Application to the Estimation of Bacterial Densities," *The Annals of Mathematical Statistics*, vol. 29, no. 4, 1958, pp. 1046–1062.

[15] J. Gao and I. Gijbels, "Bandwidth selection in nonparametric kernel testing," *Journal of the American Statistical Association*, vol. 103, no. 484, 2008, pp. 1584–1594.

[16] B. K. Ghosh and W.-M. Huang, "The Power and Optimal Kernel of the Bickel-Rosenblatt Test for Goodness of Fit," *Ann. Statist.*, vol. 19, no. 2, 6 1991, pp. 999–1009.

[17] E. Guerre and P. Lavergne, "Data-Driven Rate-Optimal Specification Testing in Regression Models," *The Annals of Statistics*, vol. 33, no. 2, 2005, pp. 840–870.

[18] P. Hall, "Central limit theorem for integrated square error of multivariate nonparametric density estimators," *Journal of Multivariate Analysis*, vol. 14, no. 1, 1984, pp. 1 – 16.

[19] P. Hall, "Integrated Square Error Properties of Kernel Estimators of Regression Functions," *Ann. Statist.*, vol. 12, no. 1, 03 1984, pp. 241–260.

[20] W. Hardle and E. Mammen, "Comparing Nonparametric Versus Parametric Regression Fits," *The Annals of Statistics*, vol. 21, no. 4, 1993, pp. 1926–1947.

[21] N. L. Hjort, "Goodness of Fit Tests in Models for Life History Data Based on Cumulative Hazard Rates," *The Annals of Statistics*, vol. 18, no. 3, 1990, pp. 1221–1258.

[22] Y. Huh and M. Hutmacher, "Application of a hazard-based visual predictive check to evaluate parametric hazard models.," *J. Pharmacokin et. Pharmacodyn.*, vol. 43, 2016, pp. 57–71.

[23] J. J Fan and I. Gijbels, *Local Polynomial Modelling and Its Applications: Monographs on Statistics and Applied Probability*, Chapman and Hall, London, 1996.

[24] J. D. Kalbfleisch and R. L. Prentice, *The statistical Analysis of Failure Time Data*, second edition edition, Wiley Series in Probability and Statistics, 2002.

[25] D. Karlis and V. Patilea, "Confidence intervals of the hazard rate function for discrete distributions using mixtures," *Computational Statistics and Data Analysis*, vol. 51, no. 11, 2007, pp. 5388 – 5401, Advances in Mixture Models.

[26] Q. Li and S. Wang, "A simple consistent bootstrap test for a parametric regression function," *Journal of Econometrics*, vol. 87, no. 1, 1998, pp. 145 – 165.

[27] R. Y. C. Liu and J. V. Ryzin, "A Histogram Estimator of the Hazard Rate with Censored Data," *The Annals of Statistics*, vol. 13, no. 2, 1985, pp. 592–605.

[28] J. Neyman, "Contribution to the theory of the chi square test," *Berkley Symposium on Mathematical Statistics and Probability, University of California Press*, 1949.

[29] P. Patil and D. Bagkavos, "Histogram for hazard rate estimation," *Sankhy?: The Indian Journal of Statistics, Series B (2008-)*, vol. 74, no. 2, 2012, pp. 286–301.

[30] Y. Pawitan, *In All Likelihood*, Oxford, 2013.

[31] M. Rosenblatt, "A Quadratic Measure of Deviation of Two-Dimensional Density Estimates and A Test of Independence," *Ann. Statist.*, vol. 3, no. 1, 01 1975, pp. 1–14.

[32] R. J. Serfling, *Approximation Theorems of Mathematical Statistics*, Wiley Series in Probability and Statistics, 1980.

[33] B. Silverman, *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, 1986.

[34] M. A. Tanner and W. H. Wong, "The Estimation of the Hazard Function from Randomly Censored Data by the Kernel Method," *Ann. Statist.*, vol. 11, no. 3, 09 1983, pp. 989–993.

[35] G. Tutz and L. Pritscher, *Nonparametric estimation of discrete hazard functions*, Kluwer Academic Publishers, 1996.

[36] A. W. v. d. Vaart, *Asymptotic Statistics*, cambridge edition, Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1998.

[37] M. Wand and M. Jones, *Multivariate Density Estimation: Theory, Practice, and Visualization*, Wiley New York, 1992.

[38] J.-L. Wang, H.-G. Muller, and W. B. Capra, "Analysis of oldest-old mortality: lifetables revisited," *Ann. Statist.*, vol. 26, no. 1, 02 1998, pp. 126–163.

[39] G. S. Watson and M. R. Leadbetter, "Hazard Analysis. I," *Biometrika*, vol. 51, no. 1/2, 1964, pp. 175–184.

[40] G. S. Watson and M. R. Leadbetter, "Hazard Analysis II," *Sankhy: The Indian Journal of Statistics, Series A (1961-2002)*, vol. 26, no. 1, 1964, pp. 101–116.

[41] B. S. Yandell, "Nonparametric Inference for Rates with Censored Survival Data," *The Annals of Statistics*, vol. 11, no. 4, 1983, pp. 1119–1135.

APPENDIX A

TABLES

## A.1 Observed Critical Values of the Chi-square Goodness-of-fit Test for Hazard Rate Functions

This section contains tables of the observed critical values of the chi-square type g-o-f $Q^0$ at sample size equals to 80,100, and 200, respectively. The significance levels are equals to $\alpha = 0.05, 0.1$, respectively. We denote by $C_{N,\alpha}$ the observed critical value for $Q^0$. The theoretical critical value at level $\alpha$ is denoted by $c_\alpha$ and the abbreviation df stands for degree of freedom.

Table A.1: Observed critical values of the test statistic $Q^0$ under under the null hypothesis $H_0 : \lambda = \lambda_w$ where $w$ is the Weibull(3,1).

| df | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|---|---|---|---|---|---|---|---|---|---|---|
| $C_{80,.10}$ | 14.74 | 16.04 | 17.22 | 17.37 | 20.01 | 21.60 | 22.81 | 24.87 | 24.88 | 26.93 |
| $C_{100,.10}$ | 13.63 | 16.30 | 17.04 | 18.16 | 19.35 | 21.25 | 21.78 | 24.14 | 24.11 | 26.92 |
| $C_{200,.10}$ | 13.44 | 14.85 | 16.41 | 18.36 | 19.06 | 21.39 | 22.42 | 22.46 | 24.94 | 25.53 |
| $C_{80,.05}$ | 18.21 | 19.62 | 21.58 | 21.25 | 23.45 | 25.78 | 26.95 | 30.26 | 29.29 | 32.05 |
| $C_{100,0.05}$ | 16.79 | 20.144 | 20.46 | 21.61 | 23.37 | 24.91 | 26.91 | 27.81 | 28.45 | 31.82 |
| $C_{200,0.05}$ | 16.10 | 18.56 | 18.97 | 21.41 | 22.63 | 23.60 | 24.77 | 27.67 | 28.81 | 29.77 |

Table A.2: Observed critical values of the test statistic $Q^0$ under the null hypothesis $H_0$ : $\lambda = \lambda_l$ where $l$ is the lognormal(0,1).

| df | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|---|---|---|---|---|---|---|---|---|---|---|
| $C_{80,.10}$ | 14.34 | 16.32 | 17.76 | 19.21 | 19.70 | 21.59 | 23.25 | 24.37 | 26.21 | 27.01 |
| $C_{100,.10}$ | 14.23 | 15.86 | 17.56 | 19.30 | 20.01 | 21.32 | 22.99 | 24.57 | 25.67 | 26.59 |
| $C_{200,.10}$ | 13.55 | 15.04 | 16.15 | 17.89 | 19.24 | 20.65 | 21.81 | 23.38 | 25.47 | 25.68 |
| $C_{80,.05}$ | 18.48 | 20.16 | 23.60 | 23.81 | 25.26 | 24.88 | 28.03 | 29.15 | 32.28 | 32.59 |
| $C_{100,0.05}$ | 18.25 | 19.77 | 20.93 | 23.68 | 23.94 | 26.01 | 27.35 | 29.36 | 30.71 | 30.42 |
| $C_{200,0.05}$ | 15.78 | 18.28 | 19.45 | 21.16 | 21.44 | 24.53 | 24.89 | 27.08 | 29.65 | 29.20 |

Table A.3: Observed critical values of the test statistic $Q^0$ under the null hypothesis $H_0$ : $\lambda = \lambda_f$ where $f$ is the folded normal(0,1).

| df | 7 | 8 | 9 | 10 | 11 | 012 | 13 | 14 | 15 | 16 |
|---|---|---|---|---|---|---|---|---|---|---|
| $C_{80,.10}$ | 8.86 | 10.46 | 12.37 | 14.86 | 15.74 | 16.44 | 18.47 | 19.55 | 22.08 | 21.99 |
| $C_{100,.10}$ | 9.18 | 11.09 | 12.41 | 15.09 | 14.73 | 17.18 | 18.35 | 20.40 | 22.19 | 23.50 |
| $C_{200,.10}$ | 8.60 | 11.03 | 12.11 | 13.96 | 16.41 | 16.724 | 19.38 | 20.45 | 22.28 | 23.48 |
| $C_{80,.05}$ | 10.92 | 13.49 | 14.80 | 17.15 | 18.12 | 19.63 | 21.36 | 23.67 | 25.88 | 25.45 |
| $C_{100,0.05}$ | 12.91 | 14.19 | 15.56 | 17.59 | 18.63 | 20.18 | 21.45 | 24.48 | 25.83 | 27.54 |
| $C_{200,0.05}$ | 11.74 | 13.90 | 15.40 | 16.74 | 18.58 | 19.67 | 23.42 | 23.44 | 27.06 | 27.49 |

Table A.4: Observed critical values of the test statistic $Q^0$ under the null hypothesis $H_0$ : $\lambda = \lambda_b$ where $b$ is the Birnbaum-Saunders(1.75,1).

| df | 7 | 8 | 9 | 10 | 11 | 012 | 13 | 14 | 15 | 16 |
|---|---|---|---|---|---|---|---|---|---|---|
| $C_{80,.10}$ | 14.19 | 16.99 | 16.78 | 18.62 | 20.06 | 22.22 | 24.19 | 24.59 | 26.96 | 27.35 |
| $C_{100,.10}$ | 14.64 | 15.97 | 17.72 | 18.65 | 20.52 | 22.36 | 22.94 | 23.93 | 25.08 | 27.56 |
| $C_{200,.10}$ | 13.27 | 15.52 | 16.85 | 17.18 | 19.95 | 20.44 | 22.14 | 23.74 | 24.01 | 24.76 |
| $C_{80,.05}$ | 18.63 | 20.63 | 20.82 | 23.69 | 24.11 | 26.56 | 29.85 | 28.85 | 31.18 | 32.52 |
| $C_{100,0.05}$ | 17.74 | 21.022 | 21.08 | 21.72 | 24.65 | 26.47 | 28.33 | 27.99 | 29.23 | 32.72 |
| $C_{200,0.05}$ | 15.73 | 19.67 | 19.69 | 20.44 | 23.21 | 24.16 | 26.04 | 28.33 | 28.01 | 29.43 |

Table A.5: Observed critical values of the test statistic $Q^0$ under the null hypothesis $H_0$ : $\lambda = \lambda_g$ where $g$ is the generalized gamma(1,0.6,4).

| df | 7 | 8 | 9 | 10 | 11 | 012 | 13 | 14 | 15 | 16 |
|---|---|---|---|---|---|---|---|---|---|---|
| $C_{80,.10}$ | 14.12 | 15.21 | 17.27 | 17.96 | 21.09 | 22.23 | 23.53 | 25.18 | 28.60 | 27.83 |
| $C_{100,.10}$ | 14.27 | 15.48 | 17.67 | 20.21 | 21.47 | 21.73 | 23.72 | 24.46 | 24.91 | 26.89 |
| $C_{200,.10}$ | 13.34 | 15.09 | 16.08 | 18.45 | 19.57 | 20.93 | 22.74 | 23.87 | 24.98 | 26.16 |
| $C_{80,.05}$ | 18.05 | 18.88 | 20.67 | 22.75 | 25.09 | 26.99 | 30.77 | 30.30 | 34.99 | 34.89 |
| $C_{100,0.05}$ | 18.47 | 19.64 | 20.58 | 24.79 | 25.42 | 27.05 | 28.91 | 29.26 | 29.92 | 32.487 |
| $C_{200,0.05}$ | 15.82 | 17.98 | 19.63 | 21.64 | 22.40 | 24.77 | 26.58 | 26.68 | 28.92 | 30.46 |

Table A.6: Theoretical critical values of the chi-square distribution for $\alpha$=0.05 and 0.10.

| df | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|---|---|---|---|---|---|---|---|---|---|---|
| $c_{\alpha=0.05}$ | 14.06 | 15.50 | 16.91 | 18.30 | 19.67 | 21.026 | 22.36 | 23.68 | 24.99 | 26.29 |
| $c_{\alpha=0.10}$ | 12.01 | 13.36 | 14.68 | 15.98 | 17.27 | 18.549 | 19.81 | 21.06 | 22.30 | 23.54 |

## A.2 Observed Critical Values of the Kernel-Based Nonparametric Goodness-of-fit Test for Hazard Rate Functions

This section contains tables of the observed critical values of the standardized test $Z_N(h)$ at sample size equals to 50,80,100,150 and 200 and significance level $\alpha = 0.05$. The abbreviation NA stands for Not Available. In this section, the observed critical value for $Z_N(h)$ at a given sample size $N$ is deboted by $C_N$. The theoretical critical value at $\alpha = 0.05$ is equal to 1.645.

Table A.7: Observed critical values of the standardized test statistic $Z_N(h)$ under the null hypothesis $H_0 : \lambda = \lambda_w$ where $w$ is the Weibull(3,1).

| $h$ | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 | 0.10 | 0.11 | 0.12 | 0.13 | 0.14 | 0.15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $C_{50}$ | NA | NA | NA | 2.56 | 2.23 | 2.43 | 2.26 | 1.67 | 1.85 | 2.55 | 2.57 | 2.34 |
| $C_{80}$ | NA | 2.16 | 1.88 | 1.79 | 1.66 | 1.52 | 2.22 | 1.95 | 2.06 | 2.24 | 2.28 | 2.37 |
| $C_{100}$ | 2.16 | 1.71 | 1.75 | 1.71 | 1.92 | 1.89 | 2.18 | 1.71 | 2.22 | 2.71 | 1.98 | 2.10 |
| $C_{150}$ | 1.80 | 1.59 | 2.07 | 2.20 | 2.04 | 2.00 | 2.27 | 1.92 | 2.06 | 2.04 | 1.61 | 2.20 |
| $C_{200}$ | 1.62 | 2.25 | 1.95 | 1.93 | 1.91 | 1.95 | 2.15 | 2.17 | 1.70 | 1.82 | 1.71 | 2.23 |

Table A.8: Observed critical values of the standardized test statistic $Z_N(h)$ under the null hypothesis $H_0 : \lambda = \lambda_l$ where $l$ is the lognormal(0,1)

| $h$ | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 | 0.10 | 0.11 | 0.12 | 0.13 | 0.14 | 0.15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $C_{50}$ | NA | NA | NA | 1.88 | 1.62 | 1.52 | 1.45 | 1.54 | 1.38 | 2.23 | 1.98 | 1.75 |
| $C_{80}$ | NA | 1.79 | 1.62 | 1.53 | 1.59 | 1.34 | 1.73 | 1.89 | 1.55 | 1.52 | 1.89 | 1.82 |
| $C_{100}$ | 1.70 | 1.59 | 1.56 | 1.51 | 1.88 | 1.65 | 1.76 | 1.80 | 1.63 | 1.80 | 1.66 | 1.69 |
| $C_{150}$ | 1.56 | 1.45 | 1.82 | 1.66 | 1.58 | 1.73 | 1.85 | 1.61 | 1.74 | 1.89 | 1.62 | 1.86 |
| $C_{200}$ | 1.35 | 1.72 | 1.51 | 1.56 | 1.82 | 1.58 | 1.81 | 1.76 | 1.73 | 1.96 | 1.61 | 1.78 |

Table A.9: Observed critical values of the standardized test statistic $Z_N(h)$ under the null hypothesis $H_0 : \lambda = \lambda_f$ where $f$ is the folded normal(0,1).

| $h$ | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 | 0.10 | 0.11 | 0.12 | 0.13 | 0.14 | 0.15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $C_{50}$ | NA | NA | NA | 1.73 | 2.05 | 1.68 | 1.55 | 1.63 | 1.44 | 2.00 | 1.99 | 1.93 |
| $C_{80}$ | NA | 1.81 | 1.79 | 1.55 | 1.28 | 1.26 | 1.98 | 1.77 | 1.49 | 1.55 | 2.15 | 1.69 |
| $C_{100}$ | 1.85 | 1.54 | 1.63 | 1.51 | 1.81 | 1.82 | 1.83 | 1.63 | 2.10 | 1.60 | 1.73 | 1.90 |
| $C_{150}$ | 1.52 | 1.38 | 1.74 | 1.75 | 1.65 | 2.00 | 1.72 | 1.56 | 2.06 | 1.73 | 1.64 | 1.93 |
| $C_{200}$ | 1.56 | 1.89 | 1.51 | 1.39 | 1.75 | 1.66 | 1.91 | 1.65 | 1.74 | 1.57 | 1.60 | 1.85 |

Table A.10: Observed critical values of the standardized test statistic $Z_N(h)$ under the null hypothesis $H_0 : \lambda = \lambda_b$ where $b$ is the Birnbaum-Saunders(1.75,1)

| $h$ | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 | 0.10 | 0.11 | 0.12 | 0.13 | 0.14 | 0.15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $C_{50}$ | NA | NA | NA | 1.75 | 1.55 | 1.50 | 1.71 | 1.45 | 1.34 | 1.83 | 1.72 | 1.74 |
| $C_{80}$ | Na | 1.82 | 1.50 | 1.72 | 1.42 | 1.34 | 1.70 | 1.60 | 1.69 | 1.56 | 1.81 | 1.81 |
| $C_{100}$ | 1.71 | 1.73 | 1.53 | 1.34 | 1.90 | 1.68 | 1.51 | 1.41 | 1.88 | 1.74 | 1.73 | 1.54 |
| $C_{150}$ | 1.39 | 1.43 | 1.61 | 1.61 | 1.58 | 1.92 | 1.62 | 1.70 | 1.86 | 1.55 | 1.62 | 1.75 |
| $C_{200}$ | 1.19 | 1.68 | 1.58 | 1.61 | 1.64 | 1.49 | 1.74 | 1.74 | 1.68 | 1.67 | 1.64 | 1.74 |

Table A.11: Observed critical values of the standardized test statistic $Z_N(h)$ under the null hypothesis $H_0 : \lambda = \lambda_g$ where $g$ is the Generalized Gamma(1,0.6,4)).

| $h$ | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 | 0.10 | 0.11 | 0.12 | 0.13 | 0.14 | 0.15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $C_{50}$ | NA | NA | NA | 2.08 | 2.23 | 1.87 | 1.76 | 1.64 | 1.62 | 2.07 | 2.07 | 2.07 |
| $C_{80}$ | NA | 1.89 | 1.86 | 1.67 | 1.63 | 1.47 | 2.03 | 1.79 | 1.92 | 1.79 | 2.28 | 2.26 |
| $C_{100}$ | 1.90 | 1.88 | 1.69 | 1.51 | 2.30 | 1.90 | 1.74 | 1.58 | 2.23 | 1.95 | 1.92 | 1.84 |
| $C_{150}$ | 1.64 | 1.57 | 2.23 | 2.04 | 1.77 | 1.89 | 1.88 | 1.83 | 2.08 | 1.87 | 1.88 | 2.34 |
| $C_{200}$ | 1.70 | 2.01 | 1.85 | 1.66 | 1.95 | 1.92 | 2.01 | 2.03 | 2.05 | 2.23 | 1.91 | 2.25 |

Table A.12: Observed critical values of the standardized test $Z_N(h)$ under the null hypothesis $H_0 : \lambda = \lambda_w$ where $w$ is the Weibull(3,1) and the level of censoring is 5%.

| $h$ | 0.06 | 0.07 | 0.08 | 0.09 | 0.10 | 0.11 | 0.12 | 0.13 | 0.14 | 0.15 |
|---|---|---|---|---|---|---|---|---|---|---|
| $C_{50}$ | NA | NA | NA | 3.10 | 3.01 | 2.82 | 2.82 | 2.18 | 2.27 | 2.75 |
| $C_{80}$ | NA | NA | 2.85 | 2.60 | 3.09 | 2.46 | 2.08 | 2.38 | 2.07 | 2.60 |
| $C_{100}$ | NA | 2.74 | 2.59 | 2.32 | 2.28 | 2.07 | 2.18 | 2.63 | 2.38 | 2.71 |
| $C_{150}$ | 2.54 | 2.08 | 1.98 | 1.62 | 3.05 | 2.22 | 2.39 | 2.29 | 1.97 | 2.81 |
| $C_{200}$ | 1.89 | 1.68 | 1.73 | 2.42 | 2.45 | 2.11 | 1.80 | 2.20 | 2.75 | 2.42 |
| $C_{5000}$ | 1.75 | 1.68 | 2.07 | 1.73 | 2.24 | 1.87 | 2.06 | 2.23 | 1.76 | 2.33/ |

Table A.13: Observed critical values of the standardized test $Z_N(h)$ under the null hypothesis $H_0 : \lambda = \lambda_w$ where $w$ is the Weibull(3,1) and the level of censoring is 16%.

| $h$ | 0.06 | 0.07 | 0.08 | 0.09 | 0.10 | 0.11 | 0.12 | 0.13 | 0.14 | 0.15 |
|---|---|---|---|---|---|---|---|---|---|---|
| $C_{50}$ | NA | NA | NA | NA | 3.07 | 3.04 | 3.33 | 3.23 | 3.54 | 2.55 |
| $C_{80}$ | NA | NA | 2.72 | 2.68 | 2.44 | 2.79 | 2.37 | 2.43 | 1.98 | 3.16 |
| $C_{100}$ | NA | 3.12 | 2.58 | 2.62 | 2.58 | 2.55 | 2.12 | 3.50 | 2.84 | 2.55 |
| $C_{150}$ | 2.44 | 2.25 | 2.13 | 1.89 | 3.05 | 2.84 | 2.47 | 2.36 | 2.16 | 2.45 |
| $C_{200}$ | 2.12 | 2.08 | 1.67 | 2.47 | 2.29 | 2.28 | 2.50 | 2.79 | 2.64 | 2.36 |
| $C_{500}$ | 2.01 | 2.17 | 2.42 | 1.89 | 2.41 | 2.29 | 2.02 | 2.43 | 2.03 | 2.14 |

Table A.14: Observed critical values of the standardized test $Z_N(h)$ under the null hypothesis $H_0 : \lambda = \lambda_w$ where $w$ is the Weibull(3,1) and the level of censoring is 28%.

| $h$ | 0.06 | 0.07 | 0.08 | 0.09 | 0.10 | 0.11 | 0.12 | 0.13 | 0.14 | 0.15 |
|---|---|---|---|---|---|---|---|---|---|---|
| $C_{80}$ | NA | NA | 5.36 | 3.69 | 3.74 | 3.40 | 3.04 | 3.31 | 3.81 | 4.22 |
| $C_{100}$ | NA | 4.82 | 3.73 | 3.61 | 4.10 | 3.09 | 3.31 | 4.46 | 4.91 | 4.40 |
| $C_{150}$ | 2.98 | 3.12 | 3.05 | 2.44 | 3.86 | 3.88 | 3.61 | 3.63 | 3.52 | 4.15 |
| $C_{200}$ | 2.48 | 2.09 | 2.15 | 2.81 | 2.66 | 3.04 | 3.02 | 3.13 | 2.99 | 3.15 |
| $C_{500}$ | 2.33 | 2.43 | 2.37 | 2.57 | 3.05 | 2.67 | 2.28 | 2.64 | 2.40 | 2.63 |

## A.3 Observed Power of the Kernel-Based Nonparametric Goodness-of-fit Test for Hazard Rate Functions

Table A.15: Observed power of the test $T_N(h)$ at $\alpha = 0.05$ ($H_0 : \lambda = \lambda_{W(3,1)}$ vs $H_1 : \lambda = \lambda_{W(2.5,1)}$).

| $h$ | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 | 0.10 | 0.11 | 0.12 | 0.13 | 0.14 | 0.15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $C_{50}$ | NA | NA | NA | 0.02 | 0.02 | 0.02 | 0.01 | 0.03 | 0.03 | 0.02 | 0.03 | 0.03 |
| $C_{80}$ | NA | 0.03 | 0.04 | 0.03 | 0.04 | 0.04 | 0.05 | 0.06 | 0.04 | 0.03 | 0.05 | 0.05 |
| $C_{100}$ | 0.04 | 0.07 | 0.04 | 0.03 | 0.07 | 0.07 | 0.04 | 0.06 | 0.08 | 0.04 | 0.08 | 0.06 |
| $C_{150}$ | 0.05 | 0.06 | 0.09 | 0.06 | 0.05 | 0.10 | 0.08 | 0.10 | 0.15 | 0.13 | 0.16 | 0.10 |
| $C_{200}$ | 0.07 | 0.09 | 0.11 | 0.01 | 0.14 | 0.15 | 0.14 | 0.12 | 0.18 | 0.18 | 0.19 | 0.16 |

Table A.16: Observed power of the test $T_N(h)$ at $\alpha = 0.05$ ($H_0 : \lambda = \lambda_{W(3,1)}$ vs $H_1 : \lambda = \lambda_{W(2,1)}$).

| $h$ | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 | 0.10 | 0.11 | 0.12 | 0.13 | 0.14 | 0.15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $C_{50}$ | NA | NA | NA | 0.06 | 0.08 | 0.03 | 0.05 | 0.12 | 0.06 | 0.09 | 0.09 | 0.11 |
| $C_{80}$ | NA | 0.16 0. | 18 | 0.17 | 0.18 | 0.17 | 0.26 | 0.28 | 0.24 | 0.18 | 0.27 | 0.24 |
| $C_{100}$ | 0.23 | 0.26 | 0.22 | 0.20 | 0.38 | 0.33 | 0.27 | 0.37 | 0.37 | 0.25 | 0.40 | 0.36 |
| $C_{150}$ | 0.33 | 0.38 | 0.52 | 0.44 | 0.46 | 0.59 | 0.50 | 0.59 | 0.64 | 0.58 | 0.71 | 0.62 |
| $C_{200}$ | 0.49 | 0.60 | 0.62 | 0.60 | 0.74 | 0.71 | 0.75 | 0.72 | 0.80 | 0.823 | 0.83 | 0.78 |

Table A.17: Observed power of the test $T_N(h)$ at $\alpha = 0.05$ ($H_0 : \lambda = \lambda_{LN(0,1)}$ vs $H_1 : \lambda = \lambda_{LN(0.25,1)}$).

| $h$ | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 | 0.10 | 0.11 | 0.12 | 0.13 | 0.14 | 0.15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $C_{50}$ | NA | NA | NA | 0.02 | 0.01 | 0.01 | 0.03 | 0.01 | 0.01 | 0.01 | 0.01 | 0.02 |
| $C_{80}$ | NA | 0.02 | 0.01 | 0.02 | 0.01 | 0.01 | 0.04 | 0.03 | 0.04 | 0.04 | 0.04 | 0.05 |
| $C_{100}$ | 0.03 | 0.03 | 0.01 | 0.01 | 0.04 | 0.06 | 0.03 | 0.03 | 0.10 | 0.06 | 0.07 | 0.07 |
| $C_{150}$ | 0.03 | 0.02 | 0.07 | 0.07 | 0.07 | 0.12 | 0.09 | 0.11 | 0.18 | 0.10 | 0.17 | 0.16 |
| $C_{200}$ | 0.04 | 0.09 | 0.12 | 0.08 | 0.13 | 0.18 | 0.21 | 0.18 | 0.17 | 0.18 | 0.29 | 0.27 |

Table A.18: Observed power of the test $T_N(h)$ at $\alpha = 0.05$ ($H_0 : \lambda = \lambda_{LN(0,1)}$ vs $H_1 : \lambda = \lambda_{LN(0.5,1)}$).

| $h$ | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 | 0.10 | 0.11 | 0.12 | 0.13 | 0.14 | 0.15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $C_{50}$ | NA | NA | NA | 0.02 | 0.03 | 0.03 | 0.03 | 0.02 | 0.02 | 0.05 | 0.08 | 0.13 |
| $C_{80}$ | NA | 0.09 | 0.11 | 0.08 | 0.04 | 0.1 | 0.32 | 0.21 | 0.34 | 0.3 | 0.42 | 0.44 |
| $C_{100}$ | 0.21 | 0.18 | 0.12 | 0.10 | 0.41 | 0.45 | 0.37 | 0.27 | 0.66 | 0.6 | 0.60 | 0.54 |
| $C_{150}$ | 0.38 | 0.31 | 0.66 | 0.67 | 0.68 | 0.85 | 0.79 | 0.85 | 0.90 | 0.85 | 0.90 | 0.91 |
| $C_{200}$ | 0.61 | 0.86 | 0.89 | 0.82 | 0.89 | 0.94 | 0.96 | 0.96 | 0.94 | 0.97 | 0.98 | 0.99 |

Table A.19: Observed power of the test $T_N(h)$ at $\alpha = 0.05$ ($H_0 : \lambda = \lambda_{FN(0,1)}$ vs $H_1 : \lambda = \lambda_{FN(0.5,1)}$).

| $h$ | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 | 0.10 | 0.11 | 0.12 | 0.13 | 0.14 | 0.15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $C_{50}$ | NA | NA | NA | 0.03 | 0.02 | 0.02 | 0.02 | 0.01 | 0.02 | 0.02 | 0.03 | 0.02 |
| $C_{80}$ | NA | 0.01 | 0.02 | 0.02 | 0.02 | 0.03 | 0.03 | 0.02 | 0.04 | 0.03 | 0.03 | 0.04 |
| $C_{100}$ | 0.02 | 0.04 | 0.01 | 0.01 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.07 | 0.05 | 0.03 |
| $C_{150}$ | 0.03 | 0.03 | 0.06 | 0.05 | 0.05 | 0.04 | 0.08 | 0.08 | 0.06 | 0.09 | 0.09 | 0.09 |
| $C_{200}$ | 0.02 | 0.06 | 0.09 | 0.09 | 0.09 | 0.09 | 0.11 | 0.13 | 0.10 | 0.17 | 0.17 | 0.14 |

Table A.20: Observed power of the test $T_N(h)$ at $\alpha = 0.05$ ($H_0 : \lambda = \lambda_{FN(0,1)}$ vs $H_1 : \lambda = \lambda_{FN(1,1)}$).

| $h$ | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 | 0.10 | 0.11 | 0.12 | 0.13 | 0.14 | 0.15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $C_{50}$ | NA | NA | NA | 0.15 | 0.04 | 0.09 | 0.13 | 0.07 | 0.13 | 0.29 | 0.25 | 0.25 |
| $C_{80}$ | NA | 0.32 | 0.23 | 0.32 | 0.42 | 0.37 | 0.48 | 0.57 | 0.63 | 0.60 | 0.61 | 0.73 |
| $C_{100}$ | 0.47 | 0.55 | 0.39 | 0.41 | 0.75 | 0.70 | 0.66 | 0.70 | 0.78 | 0.85 | 0.83 | 0.77 |
| $C_{150}$ | 0.76 | 0.74 | 0.89 | 0.88 | 0.90 | 0.93 | 0.95 | 0.96 | 0.96 | 0.97 | 0.98 | 0.98 |
| $C_{200}$ | 0.83 | 0.96 | 0.98 | 0.98 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 1 | 1 | 1 |

Table A.21: Observed power of the test $T_N(h)$ at $\alpha = 0.05$ ($H_0 : \lambda = \lambda_{GG(1,0.6,4)}$ vs $H_1 : \lambda = \lambda_{GG(1.5,0.6,4)}$).

| $h$ | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 | 0.10 | 0.11 | 0.12 | 0.13 | 0.14 | 0.15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $C_{50}$ | NA | NA | NA | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.04 | 0.03 | 0.02 |
| $C_{80}$ | NA | 0.03 | 0.02 | 0.02 | 0.01 | 0.02 | 0.09 | 0.12 | 0.08 | 0.08 | 0.12 | 0.09 |
| $C_{100}$ | 0.06 | 0.04 | 0.02 | 0.02 | 0.29 | 0.34 | 0.09 | 0.13 | 0.33 | 0.31 | 0.13 | 0.11 |
| $C_{150}$ | 0.15 | 0.10 | 0.25 | 0.23 | 0.29 | 0.51 | 0.46 | 0.46 | 0.54 | 0.59 | 0.56 | 0.55 |
| $C_{200}$ | 0.14 | 0.49 | 0.46 | 0.50 | 0.63 | 0.58 | 0.73 | 0.68 | 0.67 | 0.73 | 0.80 | 0.77 |

Table A.22: Observed power of the test $T_N(h)$ at $\alpha = 0.05$ ($H_0 : \lambda = \lambda_{GG(1,,0.6,4)}$ vs $H_1 : \lambda = \lambda_{GG(2,0.6,4)}$).

| $h$ | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 | 0.10 | 0.11 | 0.12 | 0.13 | 0.14 | 0.15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $C_{50}$ | NA | NA | NA | 0.06 | 0.01 | 0.04 | 0.04 | 0.06 | 0.05 | 0.36 | 0.33 | 0.27 |
| $C_{80}$ | NA | 0.40 | 0.33 | 0.35 | 0.31 | 0.40 | 0.71 | 0.77 | 0.68 | 0.69 | 0.80 | 0.78 |
| $C_{100}$ | 0.62 | 0.52 | 0.41 | 0.46 | 0.94 | 0.94 | 0.79 | 0.84 | 0.95 | 0.95 | 0.86 | 0.85 |
| $C_{150}$ | 0.91 | 0.88 | 0.96 | 0.96 | 0.99 | 0.99 | 0.99 | 0.99 | 1 | 1 | 1 | 1 |
| $C_{200}$ | 0.96 | 1 | 0.99 | 0.1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

Table A.23: Observed power of the test $T_N(h)$ at $\alpha = 0.05$ ($H_0 : \lambda = \lambda_{BS(1.75,1)}$ vs $H_1 : \lambda = \lambda_{BS(2.5,1)}$).

| $h$ | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 | 0.10 | 0.11 | 0.12 | 0.13 | 0.14 | 0.15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $C_{50}$ | NA | NA | NA | 0.01 | 0.02 | 0.01 | 0.00 | 0.00 | 0.01 | 0.02 | 0.03 | 0.02 |
| $C_{80}$ | NA | 0.01 | 0.02 | 0.00 | 0.01 | 0.01 | 0.05 | 0.06 | 0.03 | 0.03 | 0.09 | 0.06 |
| $C_{100}$ | 0.02 | 0.01 | 0.01 | 0.02 | 0.05 | 0.06 | 0.08 | 0.1 | 0.10 | 0.13 | 0.11 | 0.14 |
| $C_{150}$ | 0.10 | 0.02 | 0.19 | 0.15 | 0.13 | 0.20 | 0.27 | 0.22 | 0.29 | 0.43 | 0.31 | 0.44 |
| $C_{200}$ | 0.14 | 0.25 | 0.24 | 0.15 | 0.42 | 0.42 | 0.49 | 0.43 | 0.42 | 0.60 | 0.58 | 0.63 |

Table A.24: Observed power of the test $T_N(h)$ at $\alpha = 0.05$ ($H_0 : \lambda = \lambda_{BS(1.75,1)}$ vs $H_1 : \lambda = \lambda_{BS(3,1)}$).

| $h$ | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 | 0.10 | 0.11 | 0.12 | 0.13 | 0.14 | 0.15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $C_{50}$ | NA | NA | NA | 0.01 | 0.02 | 0.01 | 0.01 | 0.01 | 0.01 | 0.08 | 0.09 | 0.08 |
| $C_{80}$ | NA | 0.04 | 0.07 | 0.01 | 0.04 | 0.047 | 0.27 | 0.26 | 0.17 | 0.22 | 0.42 | 0.38 |
| $C_{100}$ | 0.14 | 0.05 | 0.08 | 0.11 | 0.31 | 0.34 | 0.39 | 0.47 | 0.49 | 0.52 | 0.53 | 0.61 |
| $C_{150}$ | 0.40 | 0.24 | 0.71 | 0.68 | 0.66 | 0.79 | 0.85 | 0.77 | 0.86 | 0.93 | 0.89 | 0.93 |
| $C_{200}$ | 0.68 | 0.86 | 0.86 | 0.80 | 0.95 | 0.95 | 0.98 | 0.97 | 0.96 | 0.99 | 0.98 | 0.99 |

APPENDIX B

GRAPH

## B.1 Graph

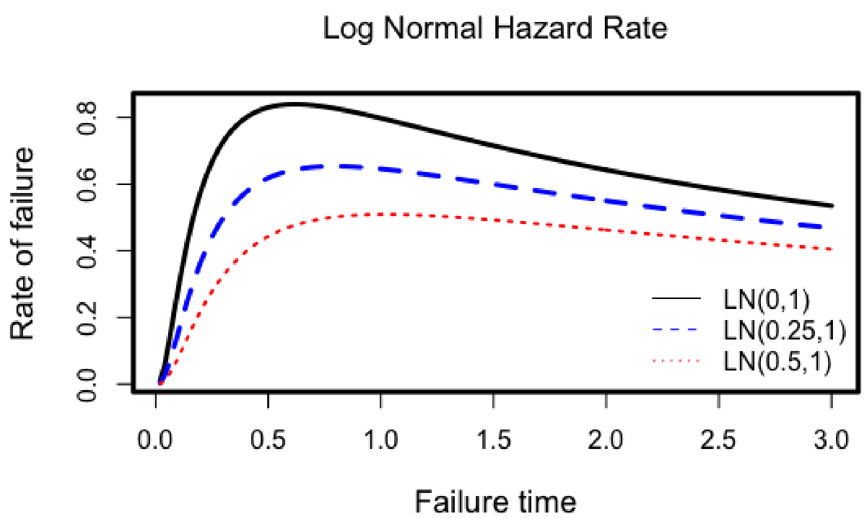


Figure B.1: Example of lognormal hazard rate functions.

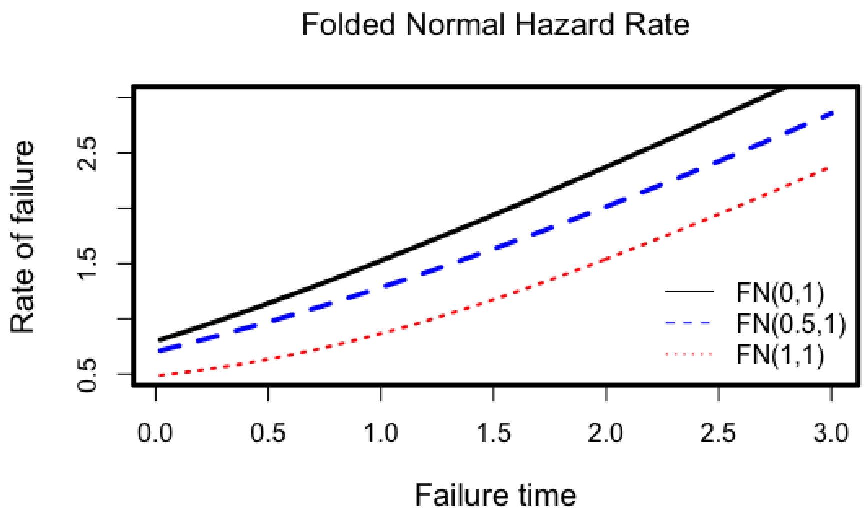


Figure B.2: Example of folded normal hazard rate functions.

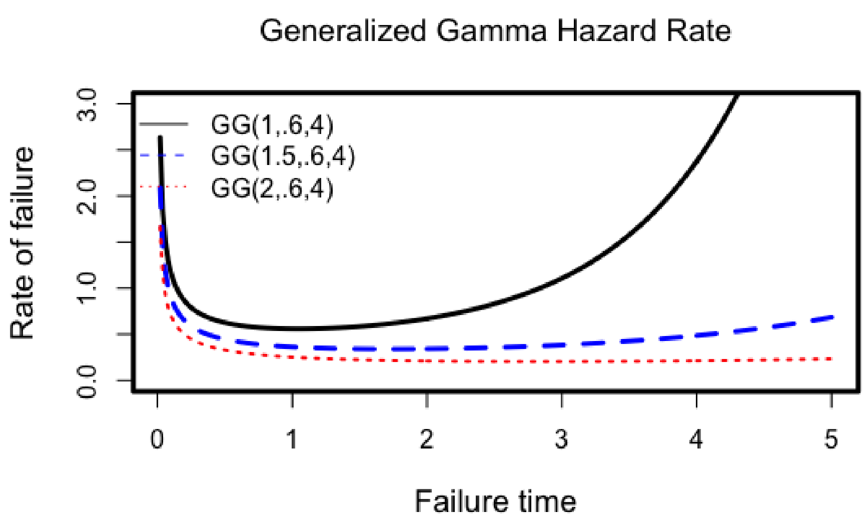


Figure B.3: Example of generalized gamma hazard rate functions.

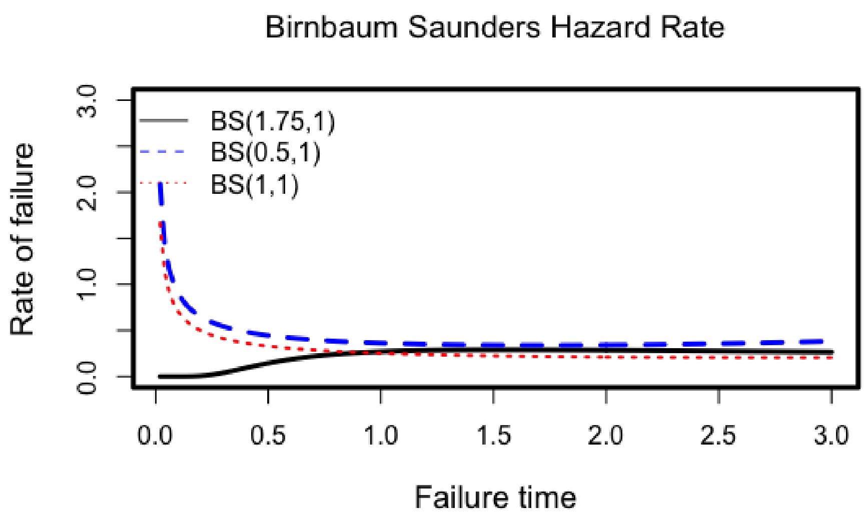


Figure B.4: Example of Birnbaum-Saunders hazard rate functions.

# APPENDIX C

# DATA

## C.1 Data

This section contains the data used in Section 4.4.

### C.1.1 Exponential Hazard Rate Model

The first set of data contains 248 units tested up to time $t = 2000$. The failure times of 125 units are given as follow

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1402 | 1921 | 408 | 1891 | 142 | 161 | 1222 | 307 | 718 | 1664 | 1801 | 36 | 396 |
| 192 | 1758 | 832 | 486 | 1454 | 640 | 1099 | 1691 | 3 | 734 | 1069 | 155 | 667 |
| 907 | 1688 | 138 | 674 | 1947 | 895 | 791 | 1203 | 282 | 1938 | 1737 | 1494 | 633 |
| 1892 | 424 | 799 | 654 | 880 | 1214 | 219 | 862 | 1290 | 1231 | 1263 | 810 | 1032, |
| 337 | 389 | 335 | 728 | 136 | 641 | 1587 | 471 | 591 | 293 | 1992 | 1925 | 1043 |
| 510 | 1194 | 859 | 1552 | 344 | 1256 | 481 | 578 | 15 | 474 | 759 | 1210 | 935 |
| 1212 | 823 | 383 | 1545 | 1446 | 1655 | 125 | 1154 | 453 | 381 | 1881 | 180 | 1458 |
| 525 | 1214 | 115 | 1452 | 1060 | 1000 | 1403 | 1289 | 1447 | 1460 | 1815 | 595 | 697 |
| 405 | 1143 | 368 | 760 | 16 | 401 | 537 | 363 | 1702 | 888 | 1022 | 550 | 218 |
| 20 | 157 | 1353 | 796 | 1699 | 1617 | 1838 | 649 | | | | | |

### C.1.2 Weibull Hazard Rate Model

The second set of data are divided into two groups. First the failure times are given as follow

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 278 | 317 | 327 | 342 | 354 | 361 | 370 | 380 | 395 | 401 | 431 | 438 | 482 |
| 484 | 507 | 513 | 521 | 549 | 553 | 568 | 575 | 588 | 596 | 599 | 627 | 629 |
| 633 | 633 | 636 | 641 | 642 | 645 | 659 | 680 | 685 | 692 | 700 | 704 | 741 |
| 743 | 757 | 767 | 772 | 784 | 788 | 790 | 790 | 793 | 798 | 823 | 825 | 830 |
| 838 | 846 | 852 | 853 | 860 | 863 | 869 | 871 | 889 | 901 | 902 | 911 | 913 |
| 921 | 935 | 944 | 947 | 965 | 994 | 999 | 1003 | 1012 | 1023 | 1045 | 1049 | 2115 |
| 1050 | 1051 | 1053 | 1058 | 1069 | 1078 | 1081 | 1087 | 1095 | 1103 | 1118 | 1887 | 1566 |
| 1118 | 1137 | 1140 | 1149 | 1186 | 1198 | 1223 | 1227 | 1271 | 1283 | 1339 | 1757 | 1622 |
| 1342 | 1357 | 1358 | 1372 | 1373 | 1377 | 1413 | 1436 | 1436 | 1444 | 1493 | 1574 | 1586 |
| 1494 | 1496 | 1511 | 1528 | 1538 | | | | | | | | |

The censoring times are :

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 470 | 504 | 626 | 717 | 781 | 813 | 860 | 886 | 906 | 947 | 973 | 982 | 1002 |
| 2427 | 1015 | 1023 | 1069 | 1122 | 1150 | 1182 | 1211 | 1313 | 1332 | 1409 | 1426 | 1476 |
| 1542 | 1577 | 1588 | 1606 | 1683 | 1738 | 1855 | 1911 | 1946 | 1960 | 1969 | 2075 | 2078 |
| 2092 | 2133 | 2241 | 2242 | 2342 | 2367 | 2385 | | | | | | |