UNIVERSITY OF LATVIA

Faculty of Physics, Mathematics and Optometry

Department of Mathematics

Leonora Pahirko

EMPIRICAL LIKELIHOOD METHOD FOR TWO-SAMPLE PROBLEMS

A thesis presented for the Scientific doctoral degree
Doctor of Science (Ph.D.) in Natural Sciences
(A field of Mathematics and subfield of Probability Theory and Mathematical Statistics)

Supervisor: Professor, *Dr. math.* Jānis Valeinis

RIGA, 2024

Abstract

In this thesis, two aspects of the empirical likelihood method for two-sample problems are investigated. First, in the context of survival analysis, the two-sample empirical likelihood method for right-censored data is established. It is demonstrated that, under certain regularity conditions, the limiting distribution of the scaled test statistic is a chi-squared random variable with one degree of freedom. Additionally, an estimator of the scaling constant is proposed.

Second, for two-sample location-scale models, the plug-in empirical likelihood method is analyzed to develop a procedure for the goodness-of-fit test for validation of the location-scale model. Furthermore, a novel graphical empirical likelihood-based method is introduced for the construction of confidence regions for location and scale parameters, accompanied by an algorithm for the practical implementation.

All established methods and theoretical findings are supported by a simulation study. Real-world data examples are presented to illustrate the application of the proposed methods. The analysis of diabetes patient data is demonstrated using empirical likelihood approach for two-sample location-scale models.

Keywords: empirical likelihood, two-sample problems, plug-in estimators, survival analysis, right censored data, location-scale model, confidence intervals and regions

Contents

List of notations and abbreviations

EL – empirical likelihood

i.i.d. – independent and identically distributed

P-P plot – probability-probability plot

Q-Q plot – quantile-quantile plot

ROC curve – receiver operating characteristic curve

$I_{\{X<x\}}$ – indicator function of the event $\{X < x\}$

F – distribution function

f – probability density function

F^{-1} – quantile function

F_n – empirical distribution function

$\widehat{F}$ – estimator of distribution function F

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

$Z_{1-\alpha}$ – $(1 - \alpha)$ quantile of standard normal distribution

$N(\mu, \sigma^2)$ – normal distribution with mean μ and variance σ^2

χ^2_k – chi-squared distribution with k degrees of freedom

$\chi^2_{k,1-\alpha}$ – $(1 - \alpha)$ quantile of chi-squared distribution with k degrees of freedom

$exp(\lambda)$ – exponential distribution with the rate parameter λ

$U[a, b]$ – uniform distribution over the interval $[a, b]$

d_i – number of events observed at time t_i

r_i – number of subjects at risk at time t_i

RMST – restricted mean survival time

Introduction

Empirical likelihood method

The empirical likelihood (EL) method has gained widespread popularity among researchers since it was introduced by Owen in 1988 [30]. Owen demonstrated that it can be used to construct confidence intervals for the sample mean, a class of M-estimates including quantiles, and statistical functionals. Furthermore, EL method was extended to make inference about the mean of a vector of random variables [31] (see also [29]). In 1994, Qin and Lawless [36] also studied the EL method to link it with general estimating equations.

The EL method offers numerous advantages over conventionally used statistical methods (see, for example, [11] for review). It is nonparametric and Bartlett-correctable method, provides confidence regions that are not of predetermined shape, but automatically reflect the emphasis in the observed data. Additionally, it is range-preserving with respect to the parameter of interest.

In 2000, Qin and Zhao [37] extended the EL method for various differences of univariate parameters in the two-sample case. They introduced conditions ensuring a chi-squared limiting distribution of the EL statistic. This setup did not include applications with non-smooth estimating equations. However, Molanes Lopes *et al.* [26] later demonstrated that the EL method is applicable even in cases with non-smooth estimating equations. Alternatively, Chen and Hall [3] showed that by appropriately smoothing the estimating equations, coverage accuracy can be improved from order $n^{-1/2}$ to n^{-1}. Claeskens *et al.* [4] also studied the smoothed EL method for comparing two populations using probability-probability (P-P) plots and receiver operating characteristic (ROC) curves.

In 2009, Hjort *et al.* extended the EL method in the one-sample case to allow for nuisance parameters in the estimating equations, estimated by plug-in estimators. While they were not the first to study this approach, they aimed to establish a very general procedure covering a wide range of applications. It was demonstrated that, in the simplest cases, under several generic assumptions the EL statistic converges to a scaled chi-squared random variable, and the scaling

constant needs to be estimated from the data. In more complex cases where the scaling constant cannot be estimated, Hjort *et al.* [16] proposed the use of bootstrap resampling methods.

Various interesting statistical EL-based procedures have been developed in last decades (see [23] for review). However, there are only few implementations of those methods available in statistical software packages. The R package `EL` [50] was developed for several statistical two-sample inference problems such as the mean or median comparison, the difference of two quantile or distribution functions, P-P and quantile-quantile (Q-Q) plots, ROC curves, among others. Valeinis also studied the two-sample EL method for P-P and Q-Q plots [51], for structural relationship models [49], as well as for inference in robust statistics (see [52] and [7]).

Empirical likelihood method in survival analysis

Survival analysis is a branch of statistics that deals with non-negative time-to-event data. Right-censored data is common in survival analysis when some subjects do not experience the event during the study or exit before it occurs for unrelated reasons. Thus, for some patients actual survival times are not fully observed. The survival function, defined as a complement of the distribution function, is the central object in survival analysis along with is nonparametric Kaplan-Meier product limit estimator.

In two-sample case, standard methods include log-rank tests for comparing survival distributions and Cox regression models for estimating hazard ratios. However, these methods may perform poorly in specific cases or do not apply due to violation of the assumptions. In such cases, alternative measures like comparing survival probabilities at a fixed time, median survival times, mean residual times, or mean restricted times can be employed.

Thomas and Grunkemeier [48] were the first, who used the EL ratio to construct the confidence intervals for survival probabilities with right censored data. However, a significant contribution to the research involving censored data has been made by Mai Zhou, who summarized the standard survival analysis topics using the EL method, along with the accompanying R codes for the implementation, in his book [56] (see also R package `emplik`).

In 2001, Wang and Jing [53] developed an EL-based approach for a class of functionals of the survival distribution, determined by a non-negative measurable function ξ. They proposed a data modification that enabled the direct application of the EL method for randomly right censored data. However, in that case, the statistic no longer remains distribution-free and converges to a scaled chi-squared distribution. This arises from the use of plug-in estimator for the unknown censoring distribution as a part of the modification procedure.

In this thesis the result of Wang and Jing [53] is extended for the two-sample case. Similarly as in one-sample case, under certain regularity conditions, a scaled chi-squared limiting distribution of the EL test statistic is obtained. The scaling constant needs to be estimated from the data, and for this purpose the same technique as proposed in [53] is employed. Specifically, the asymptotic variance of the Kaplan-Meier integral is estimated using the jackknife procedure [45]. For the inference regarding the mean residual life, an alternative estimator was proposed in [35], however it is not investigated in more detail within this thesis.

The main advantage of the established two-sample approach is its generality, implying a wide set of applications. These include, but are not limited to the inference for the difference of k-th moments of survival distributions, including mean survival times; the difference of mean residual, as well as mean restricted survival times at a fixed point; the difference of survival probabilities and cumulative hazards at a fixed point.

The established two-sample method for right-censored data is implemented using the existing R library `EL` [50]. However, for censored data, additional programming is necessary to perform hypothesis testing and construct confidence intervals.

The EL-based inference for location-scale models

In medical research, comparisons between two groups are usually made by testing the equality of two measures of central tendency, such as the mean or the median. However, more refined statistical analysis of the differences between two distributions is advantageous (see [39]). The location model $F(x) = G(x - \mu)$, for $x \in \mathbb{R}$, demonstrates a uniform shift in all quantiles of the two distributions F and G, i.e., $F^{-1}(t) = G^{-1}(t) + \mu$ for $t \in (0, 1)$, where $\mu \in \mathbb{R}$ is a constant shift denoted as a location parameter. In terms of medical research this would lead to the conclusion that the treatment is beneficial for all members of population. Under the assumption of two-sample location or location-scale model many tests beyond the classical t-test and Wilcoxon test have been developed to verify the statistical significance of the treatment, for example Lepage, Podgor-Gastwirth, Cucconi and Neuhäuser test among others (see [25] for review). However, in order to apply those tests, a location-scale model needs to be validated between two distributions.

The assumption of semiparametric location-scale model generally might be true for randomized controlled trials, however not necessarily within other study designs. The goodness-of-fit test of the location-scale model itself has not been widely studied in the literature. Doksum and Sievers [9] introduced non-parametric simultaneous confidence bands for the general shift

function based on the Kolmogorov-Smirnov statistic. Freitag and Munk [10] described the general methodology to obtain a goodness-of-fit tests for the assessment of the validity of structural relationship models, including the location-scale models as well. Hall *et al.* [12] developed a test based on the empirical characteristic functions to verify if the distributions belong to a location-scale family. Recently, the validation of a location-scale model was also studied by Subramanian [47] for censored data case using the plug-in EL method.

Whithin this thesis an alternative procedure that can be used for goodness-of-fit testing of the two-sample location-scale model is developed. It is proposed first to estimate the location and scale parameters and afterwards to transform one of the samples by the location-scale transformation. The two-sample EL method is used to test the equality between the transformed and the other sample. Two EL-based statistics for comparison of two distribution functions are considered for this purpose: the difference between the distribution functions and the probability-probability plots at a fixed point. Due to the plug-in parameter estimators, the considered test statistics are not distribution-free and their limiting distribution converges to a scaled χ^2_1 random variable [49].

The second goal concerning location-scale models is to develop a new graphical tool for constructing confidence regions for both the location and scale parameters. To do this, a distribution-free EL test statistic for generalized probability-probability plot function is used. Other EL-based test statistics can be used for this purpose as well. A distribution free confidence region contains those values of location and scale parameter for which the null hypothesis of location-scale model is not rejected. Although the idea in principle comes from Doksum and Sievers [9], regarding the EL method this is a novel approach and could be used for any two sample problem.

Within this thesis the algorithm for the practical implementation of the proposed method is given for the location and the scale models only. In this case, simply the confidence interval instead of region at a fixed point is obtained for either a location or a scale parameter. The results obtained can be extended to the wider class of the semiparametric relationships between two distributions. Such models were first introduced in a general form by Freitag and Munk [10] and the most common examples include the two-sample location-scale model and the Lehmann alternative model, where the latter holds between two distributions in case of a proportional hazards model.

Statistical analysis of diabetes patients data

The author of this thesis has been actively engaged in a collaboration with the Personalized Laboratory of Medicine at the Faculty of Medicine, University of Latvia, since 2017. The primary focus of this laboratory is the investigation of risk factors associated with diabetes complications, leading to numerous research projects and an extensive collection of data from diabetes patients. This collaboration has resulted in six scientific papers (see List of ancillary publications), in which the thesis author has undertaken the statistical analysis of the datasets.

However, during this collaboration, it became apparent that traditional statistical methods were insufficient to provide highly detailed statistical analysis. Consequently, there arose a necessity to implement more advanced statistical methods, particularly those based on EL, in the analysis of data from diabetes patients. In this thesis, several applications of the proposed methods were demonstrated using data from diabetes studies, with the aim of promoting their use in medical research within Latvia.

Aim of the research

The aim of this research is twofold: to establish a two-sample EL method for right-censored data and to develop inferential methods for location-scale models based on the EL method. To achieve this aim, the following tasks and objectives have been outlined.

(1) To establish the two-sample EL method for right-censored data.

(2) To analyze the practical implementation and applications of the two-sample plug-in empirical likelihood for location-scale models.

(3) To develop a new method for constructing confidence regions for location and scale parameters based on the EL method for P-P plots.

(4) To analyze data from diabetes patients using advanced statistical methods, including EL method in two-sample case.

(5) To carry out the practical implementation of the analyzed and established methods in the R software.

(6) To perform a comprehensive simulation study that complements theoretical findings, and to illustrate the proposed methods with real-world datasets.

Structure of the thesis

This thesis is divided into two main parts: one part is devoted to the right censored data setup, and the other part is dedicated to the two-sample location-scale models.

The first chapter of the thesis gives the description of the empirical likelihood method in one and two-sample case based on [36] and [37]. Moreover, the plug-in EL method established by Hjort *et al.* [15] is presented for one-sample problems. Smoothed EL method is described as well following the setup by Claeskens *et al.* [4] for comparison distribution of two populations.

The second chapter is devoted to the extension of the EL method for the two-sample case for right censored data. The preliminaries of the survival analysis along with some other two-sample inferential methods is also discussed.

The third chapter focuses on two applications of the empirical likelihood-based inference for the location-scale models in two-sample setup. First, the validation of the semiparametric location-scale relationship between two distributions using the empirical likelihood method-based statistics is described. Second, the algorithm for the construction of the confidence intervals (regions) for the location and scale parameter between two samples is established.

Both second and third chapters are accompanied separately with the simulation studies and the applications to some real data analysis.

Approbation of the results and author's contribution

The results obtained in this thesis have been published in three peer-reviewed scientific papers, and one preprint has been prepared. All of them are listed below. The list of ancillary publications is found in the Appendix.

(1) L. Pahirko, J. Valeinis, J. Gredzens, M. Krumina. Validation of Two-sample Location-scale Model Using Empirical Likelihood Based Statistics. *In 5th International Conference on Statistics: Theory and Applications, ICSTA 2023, August 3–5, 2023, London, United Kingdom.* Paper ID 168, Avestia Publishing, 10 pages. `https://doi.org/10.11159/icsta23.168` (to be indexed in Scopus).

(2) L. Pahirko, J. Valeinis. Empirical Likelihood-Based Confidence Regions for the Structural Relationship Parameter. *In 6th International Conference on Mathematics and Statistics (ICoMS 2023), July 14–16, 2023, Leipzig, Germany.* ACM, New York, NY, USA, p. 38-47. `https://dl.acm.org/doi/abs/10.1145/3613347.3613354`.

(3) L. Pahirko, J. Valeinis. Two-sample empirical likelihood method for right censored data. Submitted for *International Journal of Biostatistics*, ID: DGIJB.2023.0132, 2023.

(4) A. Fedulovs, L. Pahirko, K. Jekabsons, L. Kunrade, J. Valeinis, U. Riekstina, V. Pirags, J. Sokolovska. Association of endotoxaemia with low grade inflammation, metabolic syndrome and distinct response to lipopolysaccharide in type 1 diabetes. *Biomedicines*, 11(**12**):3269, 2023. `https://www.mdpi.com/2227-9059/11/12/3269`.

In papers (1) - (3), the author wrote and edited the text, established and proved theoretical results, conducted simulation studies, and performed data analysis. In paper (4), the author conducted data analysis, described statistical methods and results sections, and edited the text.

The author of the thesis has participated in 13 scientific conferences since 2014 (see List of conferences in the Appendix): twelve international conferences (C1, C2, C4-C13) and one national conference in Latvia (C3).

Chapter 1

Empirical likelihood method

This chapter is dedicated to introducing the empirical likelihood method, which was established by Owen in 1988 [30]. A very basic definition of the method is stated briefly here, as a more rigorous description of the method can be found in a recent doctoral thesis by Delesa-Vēliņa [6]. Therefore, in this chapter the primarily focus is on the specific extensions of the method relevant to the thesis objectives, while also providing a fundamental definition of the method.

In Section 1.1, the empirical likelihood method is defined for a one-sample case in a general setup, following Qin and Lawless [36]. They examined the general estimating equations for a d-dimensional random vector with a p-dimensional parameter of interest, θ. The extension of their work was demonstrated by Hjort et al. [16], who introduced the concept of plug-in nuisance parameters allowed to be incorporated into the estimating equations. Their approach is described in Section 1.2. While their approach is quite general, in simpler cases, the limiting distribution converges to a scaled chi-squared random variable due to the use of plug-in estimators.

The method for the two-sample case was extended by Qin and Zhao [37], who investigated the empirical likelihood ratio for various differences in two populations denoted by Δ. They also introduced a set of smoothness conditions that must hold for the estimating equations (see Section 1.3), which excludes applications with non-smooth estimating functions. However, in Section 1.4, smoothed empirical likelihood is introduced, following the approach by Claeskens et al. [4], who studied the method for probability-probability (P-P) plots and receiver operating characteristic (ROC) curves using kernel density smoothing on the estimating equations. While the smoothed method allows for application in a wider range of scenarios, a new challenge arises concerning the selection of the smoothing bandwidth.

1.1 Empirical likelihood method

Let us follow the very general setup of the empirical likelihood method described in [36]. Suppose that X_1,...,X_n is a sequence of independent random vectors with unknown distribution function F. Let $\theta \in \mathbb{R}^p$ be a parameter of interest. The unknown true parameter θ_0 is uniquely determined by the system of $r \geq p$ unbiased estimating equations

$$E\{w(X, \theta_0)\} = 0, \tag{1.1}$$

where w is a vector of r functions $w_j(X, \theta)$ for $j = 1, \ldots, r$.

Example 1.1.1. The inference for mean value. Suppose the parameter of interest $\theta = \mu$. To obtain confidence region for μ, the estimating equation is $w(X, \mu) = X - \mu$.

Definition 1.1.1. For any $\theta \in \mathbb{R}^p$ the empirical likelihood ratio function is defined as

$$L(\theta) = n^n \sup \left\{ \prod_{i=1}^{n} p_i : p_i > 0, \sum_{i=1}^{n} p_i = 1, \sum_{i=1}^{n} p_i w_j(X_i, \theta) = 0,\ j = 1, \ldots, r \right\}. \tag{1.2}$$

where $p_i = P(X = x_i)$ with x_i, $i = 1, \ldots, n$, being the observations.

The supremum in 1.2 exists and is uniquely determined if 0 belongs to the interior of the convex hull of $\{w(X_1, \theta), \ldots, w(X_n, \theta)\}$. The standard Lagrange multiplier method is used to find the optimal p_i's. Thus

$$p_i = \frac{1}{n} \left\{ 1 + \sum_{j=1}^{r} \lambda_j(\theta) w_j(X_i, \theta) \right\}^{-1} \quad \text{for} \quad i = 1, \ldots, n. \tag{1.3}$$

Next, it is more convenient to use the minus log-likelihood ratio, which is expressed as

$$l_E(\theta) = -\log L(\theta) = \sum_{i=1}^{n} \log \left\{ 1 + \sum_{j=1}^{r} \lambda_j(\theta) w_j(X_i, \theta) \right\}, \tag{1.4}$$

where $\lambda_j(\theta)$ satisfy the following equations

$$\sum_{i=1}^{n} \frac{w_j(X_i, \theta)}{1 + \sum_{k=1}^{r} \lambda_k(\theta) w_k(X_i, \theta)} = 0 \quad \text{for} \quad j = 1, \ldots, r. \tag{1.5}$$

Qin and Lawless [36] introduce the following Lemma and Theorem for asymptotic behaviour of the empirical likelihood test statistic. Let $|| \cdot ||$ denote Euclidean norm.

Lemma 1.1.1. [36] *Assume that $E\{w(X,\theta_0)w^T(X,\theta_0)\}$ is positive definite, $\partial g(X,\theta)/\partial\theta$ is continuous in a neighbourhood of the true value θ_0, $||\partial w(X,\theta)/\partial\theta||$ and $||w(X,\theta)||^3$ are bounded by some integrable function $G(x)$ in the neighbourhood of θ_0, and the rank of $E\{\partial w(X,\theta)/\partial\theta\}$ is p. Then, as $n\to\infty$, $l(\theta)$ attains its minimum value at some point $\widehat{\theta}$ in the interior of the ball $||\theta-\theta_0||\leq n^{-1/3}$ with probability 1.*

Theorem 1.1.1. [36] *Assume the conditions of Lemma 1.1.1 are satisfied and $\partial^2 w(X,\theta)/(\partial\theta\partial\theta^T)$ is continuous in θ in a neighbourhood of the true value θ_0 and $||\partial^2 w(X,\theta)/(\partial\theta\partial\theta^T)||$ can be bounded by some integrable function $G(x)$ in the neighbourhood of θ_0. Then the following holds for the empirical likelihood ratio statistic for testing $H_0:\theta=\theta_0$ as $n\to\infty$,*

$$W_E(\theta_0)=2l_E(\theta_0)-2l_E(\widehat{\theta})\xrightarrow{d}\chi^2_p. \tag{1.6}$$

A level α test for the null hypothesis $H_0:\theta=\theta_0$ will reject H_0, if $W_E(\theta_0)>\chi^2_{p,1-\alpha}$, where $\chi^2_{p,1-\alpha}$ denotes $(1-\alpha)$ quantile of χ^2_p distribution. An approximate $100(1-\alpha)\%$ empirical likelihood confidence region for θ_0 is obtained by the following subset of $\mathbb{R}^p$

$$\{\theta: W_E(\theta)\leq\chi^2_{p,1-\alpha}\}. \tag{1.7}$$

Remark 1.1.1. *Molanes Lopez et al. [26] proved similar result for more general form of estimating equations $w(X,\theta)$ including a non-smooth functions with the respect to θ, such as the indicator functions. Their result was proved using different techique as compared to the Qin and Lawless [36]. However, within this thesis non-smooth estimating equations will be subject to smoothing (see Section 1.4).*

1.2 Empirical likelihood method with the plug-in parameter estimators

The use of plug-in for nuisance parameters in empirical likelihood have been previously studied in several survival analysis contexts (see, for example, [34], [53]) and in survey sampling with imputation for missing response [38]. The more widely applicable plug-in approach was established by Hjort *et al.* [16] to allow for nuisance parameters in the estimating equations and slower than $\sqrt{n}$-rates of convergence. They derived a very general result providing several generic assumptions that can be checked in a large number of applications. Moreover, they provide a bootstrap approximation that can be used in situations when the plug-in asymptotics

become too complex.

Let us again consider the setup as described in Section 1.1. However, in this setup the unbiased estimating equation is allowed to contain some nuisance parameter denoted by h, thus $E\{w(X, \theta, h)\} = 0$. The plug-in approach suggests that h might be estimated by some estimator, which will be denoted by $\widehat{h}$.

To establish the limiting distribution for $l(\theta, \widehat{h})$ in a very general case, Hjort *et al.* [16] provided four assumptions that need to be verified in specific applications. These assumptions ensure a non-degenerate limiting distribution, but do not require *i.i.d.* observations or consistency of $\widehat{h}$, although this would be helpful for checking the conditions in the applications.

The following notations for the vectors v are used: $||v||$ is the Euclidean norm and $v^{\otimes 2} = vv^T$. For matrices $V = (v_{ij})$, let $|V| = \max_{ij} |v_{ij}|$. Let U be a non-degenerate p-dimensional random vector, usually $U \sim N_p(0, V_1)$, where the covariance matrix V_1 is positive definite. Let V_2 denote a positive definite $p \times p$ covariance matrix. The assumptions are following.

(A0) $P(L(\theta_0, \widehat{h}) = 0) \to 0$, when $n \to \infty$.

(A1) $n^{-1/2} \sum_{i=1}^{n} w(X_i, \theta_0, \widehat{h}) \xrightarrow{d} U$.

(A2) $n^{-1} \sum_{i=1}^{n} w^{\otimes 2}(X_i, \theta_0, \widehat{h}) \xrightarrow{p} V_2$.

(A3) $\max_{1 \le i \le n} ||n^{-1/2} w(X_i, \theta_0, \widehat{h})|| \xrightarrow{p} 0$.

First assumption (A0) is the basic existence condition; (A1) and (A2) are necessary to establish the asymptotics; and (A3) is required to obtain the negligibility of the remainder term.

When the unknown parameter h_0 is substituted by its estimator $\widehat{h}$, the setup of empirical likelihood as in Section 1.1 can be extended with the following theorem.

Theorem 1.2.1. *If assumptions (A0) - (A3) hold, then*

$$l(\theta_0, \widehat{h}) = -2 \log L(\theta_0, \widehat{h}) \xrightarrow{d} U^T V_2^{-1} U. \tag{1.8}$$

The result in Theorem 1.2.1 provides a way of calibrating $\{\theta : l(\theta, \widehat{h}) \le c\}$ as a confidence region for θ_0.

Remark 1.2.1. *When $U \sim N_p(0, V_1)$ with V_1 positive definite, the limiting distribution in Theorem 1.2.1 may be expressed as $r_1\chi^2_{1,1} + \ldots + r_p\chi^2_{1,p}$, where the $\chi^2_{1,j}$'s are independent chi-squared random variables with one degree of freedom and the weights $r_1, \ldots, r_p$ are the eigenvalues of $V_2^{-1}V_2$. If, in addition, V_1 and V_2 coincide, the limiting distribution is the standard χ^2_p. If V_1*

and V_2 do not coincide, the weights $r_1, \ldots, r_p$ may need to be estimated via some consistent estimators $\widehat{V}_1$ and $\widehat{V}_2$ and computing the eigenvalues of $\widehat{V}_2^{-1}\widehat{V}_1$. When V_1 is difficult to estimate by other means, Hjort et.al. [16] suggest to use bootstrap approach.

Example 1.2.1. Functionals of survival distribution. In this example the inference for a class of functionals of survival distributions developed by Wang and Jing [53] is considered. Let F and G be a distribution functions of the survival and the censoring distributions, respectively. The parameter of interest is a linear function of F of the form $\theta = \theta(F) = \int_0^\infty \xi(t)dF(t)$, where $\xi(t)$ is some known non-negative measurable function and $\theta(F)$ is assumed finite. The estimating equation is

$$w(Z, \delta, \theta, G) = \frac{\xi(Z)\delta}{1 - G(Z)} - \theta,$$

where $X \sim F$ and $Y \sim G$, and both are assumed to be independent, $Z = \min(X, Y)$ and $\delta = I_{\{X<Y\}}$. The Kaplan-Meier estimator $\widehat{G}$ plays a role of the plug-in estimator. The conditions (A0) - (A3) are verified in [16]. In this example, V_1 represents the asymptotic variance of $\widehat{\theta} = \theta(\widehat{F})$ with $U \sim N(0, V_1)$. Furthermore, $V_2 = E\{w^2(Z, \delta, \theta_0, G)\}$. Consequently, the limiting distribution of the empirical likelihood statistic converges to a scaled χ_1^2 random variable as V_1 and V_2 do not coincide. For more details, refer to [53]. This example will be revisited in Chapter 2 of this thesis when discussing the extension of empirical likelihood for two populations.

1.3 Empirical likelihood inference for the differences of two populations

The previously described approach can be extended to the two-sample situation as well following the setup of Qin and Zhao [37]. Suppose $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$ are two independent random samples with unknown distribution functions F and G, respectively. The parameter of interest is the difference in the form $\Delta = \theta_1 - \theta_0$, where θ_0 and θ_1 are univariate parameters associated with F and G, respectively. Suppose that the parameters θ_0 and θ_1 are uniquely determined by the unbiased estimating equations

$$E\{w_1(X, \Delta_0, \theta_0)\} = 0 \quad \text{and} \quad E\{w_2(Y, \Delta_0, \theta_0)\} = 0. \tag{1.9}$$

Example 1.3.1. [37] Let $x \in \mathbb{R}$, $\theta_0 = F(x)$ and $\theta_1 = G(x)$, then for the difference of the

distribution functions $\Delta = G(x) - F(x)$, unbiased estimating equations are

$$w_1(X, \theta_0, \Delta_0) = I_{\{X \le x\}} - \theta_0, \tag{1.10}$$

$$w_2(Y, \theta_0, \Delta_0) = I_{\{Y \le x\}} - \theta_0 - \Delta_0. \tag{1.11}$$

Definition 1.3.1. For any $(\Delta, \theta) \in \mathbb{R}^2$ the empirical likelihood ratio function is defined as

$$\begin{aligned} L(\Delta, \theta) = n^n m^m \sup \Bigg\{ & \prod_{i=1}^{n} p_i \prod_{j=1}^{m} q_j : \\ & p_i > 0, \sum_{i=1}^{n} p_i = 1, \sum_{i=1}^{n} p_i w_1(X_i, \theta, \Delta) = 0, \\ & q_j > 0, \sum_{j=1}^{m} q_j = 1, \sum_{i=1}^{m} q_i w_2(Y_i, \theta, \Delta) = 0 \Bigg\}. \end{aligned} \tag{1.12}$$

The empirical log-likelihood statistic for Δ now is defined as

$$\begin{aligned} l(\Delta, \theta) &= -2 \log L(\Delta, \theta) \\ &= 2 \sum_{i=1}^{n} \log \left(1 + \lambda_1(\Delta, \theta) w_1(X_i, \theta, \Delta)\right) + 2 \sum_{i=1}^{m} \log \left(1 + \lambda_2(\Delta, \theta) w_2(Y_i, \theta, \Delta)\right), \end{aligned} \tag{1.13}$$

where the Lagrange multipliers $\lambda_1(\Delta, \theta)$ and $\lambda_2(\Delta, \theta)$ are defined analogous to the equation (1.5).

Qin and Zhao [37] introduced the following assumptions to establish the limiting distribution of the empirical likelihood ratio statistic.

(i) $\theta_0 \in \Omega$ and Ω is an open interval.

(ii) $E\{w_1^2(X, \Delta, \theta)\} > 0$ and $E\{w_2^2(Y, \Delta, \theta)\} > 0$;
$\partial w_1(X, \Delta, \theta)/\partial\theta$, $\partial w_2(Y, \Delta, \theta)/\partial\theta$ exist and are continuous in a neighbourhood of θ_0;
$\partial w_1(X, \Delta, \theta)/\partial\theta$ and $w_1^3(X, \Delta, \theta)$ are bounded by some integrable function $H_1(x)$ in the neighbourhood of θ_0;
$\partial w_2(Y, \Delta, \theta)/\partial\theta$ and $w_2^3(Y, \Delta, \theta)$ are bounded by some integrable function $H_2(x)$ in the neighbourhood of θ_0;
$E\{\partial w_1(X, \Delta, \theta)/\partial\theta\}$ and $E\{\partial w_2(Y, \Delta, \theta)/\partial\theta\}$ are non-zero.

(iii) $m/n \to k$ as $n, m \to \infty$ and $0 < k < \infty$.

Theorem 1.3.1. [37] *If the assumptions (i) - (iii) are satisfied, then*

$$l(\Delta_0, \widehat{\theta}(\Delta_0)) \xrightarrow{d} \chi^2_1,$$

where $\widehat{\theta}(\Delta) = \arg\min_\theta l(\Delta, \theta)$.

Based on Theorem 1.3.1, the $(1-\alpha)100\%$ confidence interval for Δ_0 is

$$\{\Delta : l(\Delta, \widehat{\theta}(\Delta)) \leq \chi^2_{1,1-\alpha}\},$$

where $\chi^2_{1,1-\alpha}$ denotes $(1-\alpha)$ quantile of χ^2_1 distribution.

1.4 Smoothed empirical likelihood method for P-P plots and ROC curves

Let us consider the two-sample framework as in previous section. Claeskens *et al.* [4] studied the empirical likelihood method to compare two populations using the distribution functions F and G. First approach is to study the comparison distribution defined as $D(t) = F\{G^{-1}(t)\}$, for $0 \leq t \leq 1$. The test for the equality of the distribution functions F and G is the same as to test for the uniformity of D. The graph of D is referred as the probability-probability (P-P) plot. The equality of both distributions is not rejected if the P-P plot lies on the 45-degree diagonal. However, a plot strictly above or below the diagonal indicates that one of the distributions is stochastically larger than other. The second approach is to use the receiver operating characteristic (ROC) curve defined as the graph of $R(t) = 1 - F\{G^{-1}(1-t)\}$, for $0 \leq t \leq 1$. A ROC curve is an important tool used to summarize the performance of a medical diagnostic test.

The main reason to use the empirical likelihood method for above mentioned problems was the construction of confidence regions, which are determined by the sample [4]. For P-P plots, define the parameter of interest Δ as

$$\Delta = F\{G^{-1}(t)\},\ t \in (0,1), \tag{1.14}$$

for which the unbiased estimating equations are

$$w_1(X, \Delta_0, \theta_0) = I_{\{X \leq \theta_0\}} - \Delta_0, \tag{1.15}$$

$$w_2(Y, \Delta_0, \theta_0) = I_{\{Y \leq \theta_0\}} - t, \tag{1.16}$$

where $\theta_0 = G^{-1}(t)$ denotes the true nuisance parameter. For ROC curves, define the parameter of interest Δ as

$$\Delta = 1 - F\{G^{-1}(1-t)\},\ t \in (0,1), \tag{1.17}$$

then the unbiased estimating equations are

$$w_1(X, \Delta_0, \theta_0) = I_{\{X \le \theta_0\}} + \Delta_0 - 1, \tag{1.18}$$

$$w_2(Y, \Delta_0, \theta_0) = I_{\{Y \le \theta_0\}} + t - 1, \tag{1.19}$$

where $\theta_0 = G^{-1}(1-t)$ denotes the true nuisance parameter.

The functions w_1 and w_2 in both cases are not smooth with the respect to θ, in which case Claeskens et.al. [4] suggested to use the smoothed empirical likelihood. Chen and Hall [3] studied the empirical likelihood method in the context of quantile estimation and argued that non-smooth empirical likelihood method cannot outperform other existing methods in such cases. Moreover, they demonstrated that by appropriately smoothing of the estimating equations, coverage accuracy may be improved from order $n^{-1/2}$ to n^{-1} and using Bartlett correction for scale of smoothed empirical likelihood can reduce the size of coverage error from order n^{-1} to n^{-2} (see [3]).

In order to make inference on Δ, let us introduce the smoothed empirical likelihood method considering the following setup by Claeskens *et al.* [4]. Define $H_j(t) = \int_{u \le t} K_j(u)du$ for $j = 1, 2$, where K_j is a compactly supported r-th order kernel with $r \ge 2$. Next, define $H_{b_j}(t) = H_j(t/b_j)$, where $b_1 = b_1(n)$ and $b_2 = b_2(m)$ are bandwidth sequences converging to zero as $n, m \to \infty$. Define the estimators

$$\widehat{F}_{b_1,p}(x) = \sum_{i=1}^{n} p_i H_{b_1}(x - X_i) \quad \text{and} \quad \widehat{F}_{b_2,q}(y) = \sum_{j=1}^{m} q_j H_{b_2}(y - Y_j). \tag{1.20}$$

Now, to carry out the smoothed version of the two-sample empirical likelihood method for Δ, the function $L(\Delta, \theta)$ as defined in (1.12) is subject to the constraints

$$\widehat{F}_{b_1,p}(\theta_0) = \Delta_0 \quad \text{and} \quad \widehat{F}_{b_2,q}(\theta_0) = t \tag{1.21}$$

for P-P plots, and

$$\widehat{F}_{b_1,p}(\theta_0) = 1 - \Delta_0 \quad \text{and} \quad \widehat{F}_{b_2,q}(\theta_0) = 1 - t \tag{1.22}$$

for ROC curves. The smoothed unbiased estimating equations are in the form

$$\tilde{w}_1(X_i, \Delta_0, \theta_0) = H_{b_1}(\theta_0 - X_i) - \Delta_0, \tag{1.23}$$

$$\tilde{w}_2(Y_j, \Delta_0, \theta_0) = H_{b_2}(\theta_0 - Y_j) - t \tag{1.24}$$

for P-P plots, and

$$\tilde{w}_1(X_i, \Delta_0, \theta_0) = H_{b_1}(\theta_0 - X_i) + \Delta_0 - 1, \tag{1.25}$$

$$\tilde{w}_2(Y_j, \Delta_0, \theta_0) = H_{b_2}(\theta_0 - Y_j) + t - 1 \tag{1.26}$$

Claeskens *et al.* [4] introduce the following conditions for samples $j = 1, 2$.

(C1) For density functions $f(x) = F'(x)$ and $g(x) = G'(x)$, there exists an integer $r \geq 2$ such that derivatives $f^{(r-1)}$ and $g^{(r-1)}$ exist in a neighbourhood of θ and is continuous at θ. Furthermore, $f(\theta)g(\theta) > 0$.

(C2) As $\min(n, m) \to \infty$, $\frac{n}{n+m} \to \gamma_1$ and $\frac{m}{n+m} \to \gamma_2$, where $0 < \gamma_1, \gamma_2 < 1$.

(C3) Let $c \neq 0$ and K_j is an r-th order ($r \geq 2$) kernel satisfying

$$\int u^k K_j(u) du = \begin{cases} 1, & \text{if } k = 0 \\ 0, & \text{if } 1 \geq k \geq r-1 \\ c, & \text{if } k = r \end{cases}.$$

(C4) $nb_1^{4r} \to 0$ and $nb_1^{2r} / \log(n) \to \infty$ as $n \to \infty$;
$mb_2^{4r} \to 0$ and $mb_2^{2r} / \log(m) \to \infty$ as $m \to \infty$.

The latter condition (C4) dictates the rates at which the bandwidths b_1 and b_2 tends to zero and ensures that this convergence is fast enough [4]. The next result established by Claeskens *et al.* [4] is the nonparametric version of Wilks theorem for the P-P plots and ROC curves.

Theorem 1.4.1. [4] *Under conditions (C1) - (C4), the following holds*

$$\tilde{l}(\Delta_0, \widehat{\theta}(\Delta_0)) \xrightarrow{d} \chi_1^2,$$

where $\tilde{l}(\Delta, \theta)$ *denotes the empirical likelihood statistic* (1.13) *under estimating equations* $\tilde{w}_1(X_i, \Delta, \theta)$ *and* $\tilde{w}_2(Y_i, \Delta, \theta)$.

Based on Theorem 1.4.1, the $(1-\alpha)100\%$ confidence interval for Δ_0 is

$$\{\Delta : \tilde{l}(\Delta, \widehat{\theta}(\Delta)) \leq \chi^2_{1,1-\alpha}\}, \tag{1.27}$$

where $\chi^2_{1,1-\alpha}$ denotes $(1-\alpha)$ quantile of χ^2_1 distribution.

Remark 1.4.1. *The practical implementation of the two-sample empirical likelihood method for various parameters of interest, including the difference of distribution functions, P-P plots and ROC curves, can be carried out using the R library* `EL`*. This package implements the smoothing in case the estimating equations are not smooth with the respect to* θ*.*

Chapter 2

Two-sample empirical likelihood for right censored data

Survival analysis is a branch of statistics which deals with time-to-event data. Data can be expressed in days, months, years and other time measures. These type of data frequently arise in medicine, biology, public health, epidemiology, engineering, economics, demography, reliability theory etc. Historically, terminology in survival analysis was developed based on analysing mortality tables, thus the event of interest usually was associated with the death of the patient. However, in different applications one can be interested in other events such as diagnosis of some disease, equipment lifetime, duration of unemployment etc.

It is often impossible to observe the exact survival time of all subjects. For example, if the subject never experiences the event of interest by the end of the study, or dies for a reason other than the disease under study, or changes state of residence during the study and can no longer continue to participate. In these cases survival time is considered to be censored. Such incomplete data are a feature of survival analysis datasets that requires special statistical methods for data analysis.

In this chapter the preliminaries of the survival analysis is described in Section 2.1 covering the definitions of basic quantities and summary measures. In Section 2.2 methods for the two-sample inference are described. The main result of this chapter is introduced in Section 2.3 along with the proof of the main theorem in Section 2.4. The simulation study results are described in Section 2.5, and an application to two real datasets analysis is illustrated in Section 2.6 of this chapter.

2.1 Preliminaries of survival analysis

In this section, the basic concepts of survival analysis are defined along with theoretical aspects behind some standard procedures that are used for the inference based on book [21].

2.1.1 Basic quantities of survival analysis

There are four basic quantities that describe the distribution in survival analysis: a distribution function, a survival function, a hazard function and a cumulative hazard function. If any of these functions is known, the rest can be expressed immediately.

Let X be some non-negative random variable denoting time to the event of interest with distribution function F. The main object in the survival analysis is survival function, which describes the probability of a subject that has survived until time t, to survive at a time t.

Definition 2.1.1. The survival function of a random variable X is defined as

$$S(t) = P(X > t) = 1 - F(t), \quad t \in [0, \infty).$$

If X is a continuous random variable, a survival function $S(t)$ has following common characteristics: it is a decreasing function; $S(0) = 1$ as all subjects are alive at time $t = 0$; at time $t = \infty$ all subjects are dead, thus $S(\infty) = 0$. Moreover, if the probability density function $f(t)$ exists, the survival function is expressed as

$$S(t) = P(X > t) = \int_t^\infty f(x)dx. \tag{2.1}$$

From (2.1) it follows that $f(t) = -dS(t)/dt$.

However, if X is a discrete random variable, $S(t)$ is a decreasing step function. Suppose, X takes values x_i for $i = 1, 2, \ldots$, such that $x_1 < x_2 < \ldots$, then a survival function $S(t)$ can be expressed as

$$S(t) = P(X > t) = \sum_{x_i > t} P(X = x_i), \quad t \in [0, \infty).$$

Definition 2.1.2. The hazard function $h(t)$ of a random variable X is defined as

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \le X < t + \Delta t | X \ge t)}{\Delta t}, \quad t \in [0, \infty). \tag{2.2}$$

The hazard function describes the immediate risk for a subject who survived by time t to experience an event of interest in the next instant of time. The hazard function $h(t)$ does not

have any common characteristics of its shape, except that $h(t) \geq 0$. If X is a continuous random variable, the hazard function $h(t)$ can be expressed as

$$h(t) = \frac{f(t)}{S(t)} = -\frac{d(\log S(t))}{dt}. \tag{2.3}$$

Definition 2.1.3. The cumulative hazard function $H(t)$ of a random variable X is defined as

$$H(t) = \int_0^t h(x)dx = -\log S(t), \quad t \in [0, \infty).$$

A survival function $S(t)$ is expressed by a cumulative hazard function as

$$S(t) = e^{-H(t)} = \exp\left(-\int_0^t h(x)dx\right).$$

Summary measures

Definition 2.1.4. The mean residual life of a random variable X at a time $t_0 \in [0, \infty)$ is defined as

$$mrl(t_0) = E(X - t_0 | X > t_0). \tag{2.4}$$

Mean survival time (mean life) is defined as $\mu = mlr(0) = EX$.

The mean life μ coincides with the area under the survival curve $S(t)$, but $mlr(t_0)$ coincides with the area under the survival curve to the right from t_0 divided by $S(t)$. For the continuous random variables

$$mrl(t_0) = \frac{\int_{t_0}^{\infty}(t - t_0)f(t)dt}{S(t_0)} = \frac{\int_{t_0}^{\infty} S(t)dt}{S(t_0)} \tag{2.5}$$

and

$$\mu = \int_0^{\infty} tf(t)dt = \int_0^{\infty} S(t)dt. \tag{2.6}$$

Moreover, the variance of X is expressed by the survival function in the form

$$D(X) = 2\int_0^{\infty} tS(t)dt - \left[\int_0^{\infty} S(t)\right]^2. \tag{2.7}$$

Definition 2.1.5. A p-th quantile of the distribution F is the smallest t_p such that

$$S(t_p) \leq 1 - p, \text{ i.e., } t_p = \inf\{t : S(t) \leq 1 - p\}.$$

The median of X is defined as 0.5-th quantile of the distribution F.

For the continuous random variables the equality $S(t_{0.5}) = 0.5$ holds.

A restricted mean survival time is a summary measure that has gained much attention in the literature, recently (see, e.g., [41], [28]). It calculates the area under the survival curve $S(t)$ up to time t_0. The area above the survival curve $S(t)$ up to time t_0 is denoted as the restricted mean time lost.

Definition 2.1.6. [40] The restricted mean survival time (RMST) of a random variable X at a time $t_0 \in (0, \infty)$ is defined as

$$\mu(t_0) = E\{\min(X, t_0)\}. \tag{2.8}$$

The restricted mean time lost (RMTL) of a random variable X up to time t_0 is defined as $rmtl(t_0) = t_0 - \mu(t_0)$.

If X is a continuous random variable

$$\mu(t_0) = E\{\min(X, t_0)\} = \int_{t_0}^{\infty} S(t)dt/S(t_0). \tag{2.9}$$

Let $T := \min(X, t_0)$ a random variable, then $\mu(t_0) = ET$ and the variance is

$$Var(T) = ET^2 - (ET)^2 = 2\int_0^{t_0} tS(t)dt - \left(\int_0^{t_0} S(t)dt\right)^2. \tag{2.10}$$

2.1.2 Common parametric distributions in survival analysis

Although, this thesis is devoted to study a nonparametric statistical method, the analysis of survival data often encounters parametric models that are based on several widely used distributions. Moreover, these distributions have not only been chosen because of their popularity among survival data analysts, but also because they provide a better understanding of the quantities and functions discussed in the previous chapter, in particular the risk function $h(t)$.

A summary of hazard, survival, probability density functions and mean survival times for five well-known distributions is given in Table 2.1.

Exponential distribution

The exponential distribution is not only one of the most historically significant but also one of the mathematically simplest distributions in survival analysis. It is characterised by a constant risk function $h(t) = \lambda$. Moreover, the exponential distribution has the lack of memory property

$$P(X \geq t + z | X \geq t) = P(X \geq z). \tag{2.11}$$

Due to the lack of memory, mean residual life is also a constant function

$$E(X - x | X > x) = E(X) = \frac{1}{\lambda}, \tag{2.12}$$

as time-to-event is independent from past events. The latter can also be observed from the hazard function; as time goes by, risk of the event does not change. However, this restricts the applications of exponential distribution in some of the areas such as public health or engineering.

Weibull distribution

Although Weibull was not first who introduced the Weibull distribution, but he studied this distribution in [54] to describe the lifetime of materials and later it was named after him. The Weibull distribution is also one of the most commonly used probability distributions in survival analysis context, especially in the reliability theory. Also various extended forms of the Weibull distribution have been developed, e.g., exponentiated Weibull, inverse Weibull etc.

The Weibull distribution is defined by the survival function $S(t) = e^{-\lambda t^{\alpha}}$, $t > 0$. The exponential distribution is a special case of the Weibull distribution when $\alpha = 1$. The Weibull distribution can be fitted to an increasing, decreasing or constant hazard function by choosing $\alpha > 1$, $\alpha < 1$ or $\alpha = 1$, respectively. Therefore, α is referred to as the shape parameter. An example of the increasing Weibull hazard function would be leukaemia patients, whose survival chances worsen with time. However, decreasing hazard function would fit the surgery patients, whose survival chances are worse immediately after the surgery, but increases over time goes by.

For a detailed description of other common distributions used to model time-to-event data, see, for example, the book [21].

2.1.3 Observation censoring

In practice, it is often impossible to observe the exact survival time for all subjects. For example, if the subject never experiences the event of interest by the end of the study, or dies for a reason other than the disease under study, or changes state of residence during the study and can no longer continue to participate. In these cases survival time is considered to be censored. Such incomplete data are a feature of survival analysis datasets.

Table 2.1: Hazard, survival and probability density functions, and expected life for most frequently used distributions in survival analysis.

Distribution	$h(t)$	$S(t)$	$f(t)$	EX
Exponential $\lambda > 0$ $t \geq 0$	λ	$e^{-\lambda t}$	$\lambda e^{-\lambda t}$	$\frac{1}{\lambda}$
Weibull $\alpha, \lambda > 0$ $t \geq 0$	$\alpha\lambda t^{\alpha-1}$	$e^{-\lambda t^{\alpha}}$	$\alpha\lambda t^{\alpha-1}e^{-\lambda t^{\alpha}}$	$\frac{\Gamma(1+1/\alpha)}{\lambda^{1/\alpha}}$
Gamma $\beta, \lambda > 0$ $t \geq 0$	$\frac{f(t)}{S(t)}$	$1 - \frac{\int_0^{\lambda t} u^{\beta-1}e^{-u}du}{\Gamma(\beta)}$	$\frac{\lambda^{\beta}t^{\beta-1}e^{-\lambda t}}{\Gamma(\beta)}$	$\frac{\beta}{\lambda}$
Pareto $\theta, \lambda > 0$ $t \geq \lambda$	$\frac{\theta}{t}$	$\frac{\lambda^{\theta}}{t^{\theta}}$	$\frac{\theta\lambda^{\theta}}{t^{\theta+1}}$	$\frac{\theta\lambda}{\theta-1}$, if $\theta > 1$

There are several types of censoring, however, within this thesis only right randomly censored survival times are considered. Random censoring appears when the aim is to assess the marginal distribution of an event, but some subjects face circumstances that prevent them from continuing to participate in the study. These conditions are called competing risks.

Let X_1,...,X_n denote the survival times of n subjects as *i.i.d.* random variables with some unknown distribution function F. Let Y_1,...,Y_n denote the censoring times for each subject. In right randomly censored data case, the exact survival time of the i-th subject is known only if X_i is less than or equal to Y_i, otherwise the patient is called a "survivor" and his censoring time Y_i is observed.

Definition 2.1.7. Right randomly censored survival data are represented by pairs of random variables (Z_i, δ_i) for $i = 1, \ldots, n$, where $Z_i = \min(X_i, Y_i)$ and $\delta_i = I_{\{X_i \leq Y_i\}}$ is censoring indicator.

Other censoring types are described in more detail in [21].

2.1.4 Nonparametric estimators

In this section, it is assumed that right randomly censored data (Z_i, δ_i) for $i = 1, \ldots, n$ as in Definition 2.1.7 are given. The aim is to make inference on random variable X, which denotes real survival times. Suppose that the survival and censoring times are independent, and there are k distinct observations such that $t_1 < t_2 < \ldots < t_k$.

Definition 2.1.8. The Kaplan-Meier (1958) estimator of the survival function $S(t)$ is defined as

$$\widehat{S}(t) = \begin{cases} 1, & t < t_1, \\ \prod_{i:t_i \leq t} \left(1 - \frac{d_i}{r_i}\right) & t \geq t_1, \end{cases} \tag{2.13}$$

where $0 \leq t \leq t_k$, d_i is the number of events observed at time t_i, and r_i denotes the number of subjects at risk at time t_i.

For values of t beyond t_k, the Kaplan-Meier estimator is not well-defined (see Practical Notes 1 and 2 in [21] for further details). As can be seen from Definition 2.1.8, the Kaplan-Meier estimator is a step function with jumps in the observed event times. The size of the jumps depends not only on the number of observed events at time t_i, but also on the number of censored observations before t_i. In the absence of censored observations, the Kaplan-Meier estimator simply reduces to the empirical survival function.

Corollary 2.1.1. *The variance of the Kaplan-Meier estimator is obtained by the Greenwood formula in the form*

$$\widehat{V}(\widehat{S}(t)) = \widehat{S}^2(t) \sum_{i:t_i \leq t} \frac{d_i}{r_i(r_i - d_i)}. \tag{2.14}$$

The cumulative hazard function $H(t)$ is also estimated through the Kaplan-Meier estimator of the survival function in the form $\widehat{H}(t) = -\ln(\widehat{S}(t))$. Alternatively, cumulative hazard function is estimated using the Nelson-Allen (1972, 1978) estimator, which is better suited for small sample sizes.

Definition 2.1.9. Nelson-Allen estimator of the cumulative hazard function $H(t)$ is defined as

$$\tilde{H}(t) = \sum_{i:t_i \leq t} \frac{d_i}{r_i}, \quad 0 \leq t \leq t_k. \tag{2.15}$$

Corollary 2.1.2. *The variance of the Nelson-Allen estimator is*

$$\widehat{V}(\tilde{H}(t)) = \sum_{i:t_i \leq t} \frac{d_i}{r_i^2}. \tag{2.16}$$

The survival function $S(t)$ can be estimated using the Nelson-Allen estimator as well, yielding $\tilde{S}(t) = \exp\{-\tilde{H}(t)\}$.

Pointwise confidence intervals for survival function $S(t)$

Let us denote $\sigma_S^2 = \hat{V}(\hat{S}(t))/\hat{S}^2(t)$. The commonly used $100(1-\alpha)\%$ confidence intervals for the survival function at fixed time t_0 are called linear confidence intervals and are defined as

$$\left[\hat{S}(t_0) - Z_{1-\alpha/2}\sigma_S(t_0)\hat{S}(t_0),\ \hat{S}(t_0) + Z_{1-\alpha/2}\sigma_S(t_0)\hat{S}(t_0)\right], \tag{2.17}$$

where $Z_{1-\alpha/2}$ is $1-\alpha/2$ quantile of a standard normal distribution.

Alternatively, the transformation of $S(t_0)$ can be utilized. For example, the logarithmic transformation gives $100(1-\alpha)\%$ confidence intervals for the survival function at time t_0 in the form

$$\left[\hat{S}(t_0)^{1/\theta},\ \hat{S}(t_0)^{\theta}\right], \quad \text{where} \quad \theta = \exp\left\{\frac{Z_{1-\alpha/2}\sigma_S(t_0)}{\ln(\hat{S}(t_0))}\right\}. \tag{2.18}$$

Note that the logarithmic transformation gives intervals that are not symmetric around the survival function estimate $\hat{S}(t_0)$.

Simultaneous confidence bands for survival function $S(t)$

Pointwise confidence intervals are useful if one are only interested in a fixed time t_0, but often in practical applications researchers are interested in finding upper and lower bounds on the true survival function with some level of significance at all points in time t. In particular, the aim is to find two functions $L(t)$ and $U(t)$ such that $1-\alpha = P\{L(t) \le S(t) \le U(t),\ \text{for all } t_L \le t \le t_U\}$. Let $[L(t), U(t)]$ be called the $100(1-\alpha)\%$ simultaneous confidence band for the survival function $S(t)$.

First, choose $t_L < t_U$ such that t_L is greater than or equal to t_1 and t_U is less than or equal to t_k. Suppose n is the sample size for which the confidence bands are constructed, then define

$$a_L = \frac{n\sigma_S^2(t_L)}{1 + n\sigma_S^2(t_L)}$$

and

$$a_U = \frac{n\sigma_S^2(t_U)}{1 + n\sigma_S^2(t_U)},$$

where further required that $0 < a_L < a_U < 1$.

To construct a $100(1-\alpha)\%$ confidence band for the survival function $S(t)$ in the region $[t_L, t_U]$, find the confidence coefficient $c_\alpha(a_L, a_U)$ in Table C.3 of the Appendix C in [21]. Next,

as in the case of pointwise confidence intervals, the linear confidence bands are defined as

$$[\hat{S}(t) - c_\alpha(a_L, a_U)\sigma_S(t)\hat{S}(t),\ \hat{S}(t) + c_\alpha(a_L, a_U)\sigma_S(t)\hat{S}(t)]$$

and the logarithmic transformation bands are defined in the form

$$[\hat{S}(t)^{1/\theta},\ \hat{S}(t)^{\theta}], \quad \text{where} \quad \theta = \exp\left\{\frac{c_\alpha(a_L, a_U)\sigma_S(t)}{\ln(\hat{S}(t))}\right\}.$$

Estimators of mean and median survival times

The mean survival time and median survival time can be expressed as functions of the survival function $S(t)$. The nonparametric estimators of these quantities can be obtained simply by replacing the unknown survival function by its Kaplan-Meier estimator in the corresponding formula.

Recall that the average survival time is defined as $\mu = \int_0^\infty S(t)dt$, and the estimator is obtained by replacing $S(t)$ with $\hat{S}(t)$. Note that this estimator is only appropriate when the maximum observed survival time is uncensored, otherwise it can be artificially transformed into a survival time or an interval $[0, \tau]$ can be chosen where τ is the maximum possible survival time chosen by the investigator. Then the estimator of men survival up to time τ is

$$\widehat{\mu}_\tau = \int_0^\tau \hat{S}(t)dt,$$

where τ is maximal observed survival time or maximal bound defined by the researcher. The variance of this estimator is

$$\widehat{V}(\widehat{\mu}_\tau) = \sum_{i=1}^{k}\left[\int_{t_i}^{\tau}\hat{S}(t)dt\right]^2 \frac{d_i}{r_i(r_i - d_i)}.$$

$100(1-\alpha)\%$ confidence interval for mean survival time μ is obtained in the form

$$\left[\widehat{\mu}_\tau - Z_{1-\alpha/2}\sqrt{\widehat{V}(\widehat{\mu}_\tau)},\ \widehat{\mu}_\tau + Z_{1-\alpha/2}\sqrt{\widehat{V}(\widehat{\mu}_\tau)}\right].$$

Remark 2.1.1. *In the absence of censoring, the estimator of mean survival reduces to the sample mean.*

The Kaplan-Meier estimator can also be used to obtain estimates of the distribution quantiles x_p. Recall that $x_p = \inf\{t : S(t) \leq 1 - p\}$. When $p = 1/2$, x_p is the median survival time, whose estimator is $\widehat{x}_p = \inf\{t : \hat{S}(t) \leq 1 - p\}$. The $100(1-\alpha)\%$ confidence interval for the

quantile x_p, based on the linear confidence interval of the survival function, is the set of all time points t satisfying the condition

$$-Z_{1-\alpha/2} \leq \frac{\widehat{S}(t) - (1-p)}{\widehat{V}^{1/2}(\widehat{S}(t))} \leq Z_{1-\alpha/2}.$$

Moreover, the $100(1-\alpha)\%$ confidence interval for the quantile x_p based on the logarithmic transformation is the set of all points t satisfying the condition

$$-Z_{1-\alpha/2} \leq \frac{\Big(\ln\{-\ln[\widehat{S}(t)]\} - \ln\{-\ln[1-p]\}\Big)\Big(\widehat{S}(t)\ln[\widehat{S}(t)]\Big)}{\widehat{V}^{1/2}(\widehat{S}(t))} \leq Z_{1-\alpha/2}.$$

2.2 Two-sample inference

Let us now consider the right censored data framework for the two-sample case. Suppose that $X_{11}, \ldots, X_{1n_1}$ and $X_{21}, \ldots, X_{2n_2}$ are two independent random samples denoting survival times with the unknown distribution functions F_1 and F_2, respectively. Let us denote the corresponding censoring times with $Y_{j1}, \ldots, Y_{jn_j}$ with the distribution functions G_j for $j = 1, 2$. Survival times and censoring times are assumed to be independent. Thus, a randomly right censored observations are (Z_{ji}, δ_{ji}), where $Z_{ji} = \min(X_{ji}, Y_{ji})$ and $\delta_{ji} = I(X_{ji} \leq Y_{ji})$ for $i = 1, \ldots, n_j$ and $j = 1, 2$.

The respective survival functions are defined as $S_j(t) = 1 - F_j(t)$, $t \in [0, \infty)$ for $j = 1, 2$ and can be estimated via classical Kaplan-Meier product limit estimator

$$\widehat{S}_j(t) = \prod_{i:t_{ji} \leq t} \left(1 - \frac{d_{ji}}{r_{ji}}\right), \quad \text{for} \quad j = 1, 2, \tag{2.19}$$

where t_{ji}, are distinct event times, d_{ji} denotes number of events and r_{ji} denotes number at risk at time t_{ji} in the jth group. The variances of the Kaplan-Meier estimators obtained by Greenwood formula in two-sample setup are denoted as

$$\widehat{V}\left(\widehat{S}_j(t)\right) = \widehat{S}_j^2(t)\widehat{\sigma}_j^2, \quad \text{for} \quad j = 1, 2, \quad \text{where} \quad \widehat{\sigma}_j^2(t) = \sum_{i:t_{ji} \leq t} \frac{d_{ji}}{r_{ji} - d_{ji}}. \tag{2.20}$$

The simplest two-sample inference, sometimes called naive or crude approach, is to use the normal approximation for construction of the confidence intervals. This approach will be used in simulation study for two parameters of interest Δ defined in the examples below.

Example 2.2.1. The difference between two survival curves at a fixed point in time [20].

Let $S_j(t_0)$ denote the survival probability at time t_0 in jth group. Then, the parameter of interest is $\Delta = S_2(t_0) - S_1(t_0)$.

Example 2.2.2. The difference between two restricted mean survival times (RMST) [40]. Let $\mu_j(t_0)$ denote the t_0 restricted mean survival time in jth group. Then, the parameter of interest is $\Delta = \mu_2(t_0) - \mu_1(t_0)$.

Let θ_j denote either RMST $\mu_j(t_0)$ or survival probability $S_j(t_0)$ for $j = 1, 2$. Clearly, the estimator $\widehat{\theta}_j$ is obtained by employing the Kaplan-Meier estimator of survival function $\widehat{S}_j(t)$ for $j = 1, 2$.

Then, the estimator of Δ is $\widehat{\Delta} = \widehat{\theta}_2 - \widehat{\theta}_1$, and it is assumed that the variance is obtained by $\widehat{V}(\widehat{\Delta}) = \widehat{V}(\widehat{\theta}_1) + \widehat{V}(\widehat{\theta}_2)$, hence the standard error is $SE(\widehat{\Delta}) = \sqrt{\widehat{V}(\widehat{\Delta})}$. Then by the central limit theorem [44], under $H_0 : \Delta = 0$, the following statistic

$$Z = \frac{\widehat{\Delta}}{SE(\widehat{\Delta})} \tag{2.21}$$

is asymptotically normally distributed random variable with mean zero and variance 1 (see [40] and [20]). From equation (2.21), the $(1-\alpha)100\%$ confidence interval is obtained as

$$I_N = [\widehat{\Delta} \pm Z_{1-\alpha/2} SE(\hat{\Delta})], \tag{2.22}$$

where $Z_{1-\alpha/2}$ denotes $(1-\alpha/2)$ quantile of a standard normal distribution.

Log-rank type tests

To test the following null hypothesis $H_0 : S_1(t) = S_2(t)$ for all $0 < t \leq \tau$, where $\tau = \inf\{t : S_1(t) = 0 \,\&\, S_2(t) = 0\}$, i.e, τ denotes the largest time at which both samples have at least one subject at risk.

Definition 2.2.1. [19] The generalized weighted log-rank test statistic is defined as

$$L = \left(\sum_i \omega_i \left[d_{ji} - d_j \frac{r_{ji}}{r_j} \right] \right)^2 \Big/ \sum_{i=1}^{k} \omega_i^2 \frac{r_{1i} r_{2i} d_i (r_i - d_i)}{r_i^2 (r_i - 1)}, \tag{2.23}$$

where $j = 1, 2$ and d_{ji} is the number of events in group j at time i, d_i is the total number of events in both groups at time i, r_{ji} is the number of subjects at risk in group j at time i, r_i is the total number at risk in both groups at time i, ω_i are some weights.

The test statistic L follows a χ^2_1 distribution [21]. The classical log-rank test is obtained by choosing weights $\omega_i = 1$. Such choice assigns equal weights throughout the whole time period.

Pseudo-observations approach

Here the pseudo-observations approach is briefly described based on the papers [2], [1] and [20]. To introduce the method, suppose the pooled sample is defined by $X_1, \ldots, X_N$ *i.i.d.* random variables, where $N = n_1 + n_2$. Suppose that the parameter of interest is the expectation θ of some function ξ of X, it is

$$\theta = E(\xi(X_i)) = \int_0^\infty \xi(t) dF(t). \tag{2.24}$$

Let us assume that there is an (at least approximately) unbiased estimator $\widehat{\theta}$. Let $\boldsymbol{T_i}$, $i = 1, \ldots, N$ be *i.i.d.* explanatory variables. The conditional expectation is $\theta_i = E\left(\xi(X_i)|\boldsymbol{T}_i\right)$.

Definition 2.2.2. The i-th pseudo-observation is defined as

$$\widehat{\theta}_i = N\widehat{\theta} - (N-1)\widehat{\theta}^{-i}, \tag{2.25}$$

where $\widehat{\theta}^{-i}$ is the "leave-one-out" estimator for θ based on X_k, $k = 1, \ldots, i-1, i+1, \ldots, n$.

If all X_i are observed, i.e., there is no presence of censoring, then θ is estimated by the average of the $\xi(X_i)$. However, further the censored data model only is assumed, i.e., the observations are presented as pairs (Z_i, δ_i) for $i = 1, \ldots, N$. Next, the generalized linear model is used to model the association of how θ_i depends on T_i. This takes the following form

$$g(\theta_i) = \beta^T \boldsymbol{T}_i, \tag{2.26}$$

where β is a vector of the regression coefficients and a column $T_{i0} = 1$ is added to the matrix $\boldsymbol{T}$ corresponding to an intercept β_0, and $g(\cdot)$ denotes the link function. Next, it is proposed to estimate the regression coefficients β using generalized estimating equations in the form

$$U(\beta) = \sum_{i=1}^{N} U_i(\beta)) = \sum_{i=1}^{N} \left(\frac{\partial}{\partial \beta} g^{-1}\left(\beta^T \boldsymbol{T}_i\right) \boldsymbol{R}_i^{-1}(\beta) \left(\widehat{\theta}_i - g^{-1}\left(\beta^T \boldsymbol{T}_i\right) \right) \right) = 0, \tag{2.27}$$

where $\boldsymbol{R}_i(\beta)$ is a working covariance matrix in general setup. The variances of β was obtained

in [2] using standard sandwich estimator

$$\widehat{\boldsymbol{V}} = I\left(\widehat{\beta}\right)^{-1} \widehat{Var}\left(U(\beta)\right) I\left(\widehat{\beta}\right)^{-1}, \tag{2.28}$$

where

$$I\left(\beta\right) = \sum_{i=1}^{N} \left(\frac{\partial g^{-1}\left(\beta^T \boldsymbol{T}_i\right)}{\partial \beta}\right)^T \boldsymbol{R}_i^{-1}(\beta) \left(\frac{\partial g^{-1}\left(\beta^T \boldsymbol{T}_i\right)}{\partial \beta}\right) \tag{2.29}$$

$$\widehat{Var}\left(U(\beta)\right) = \sum_{i=1}^{N} U_i(\beta) U_i(\beta)^T. \tag{2.30}$$

Once the pseudo-observations are calculated, coefficient estimates and their standard errors are obtained using standard statistical software for generalized estimation equation (for example, R library `geepack`).

Within the simple two-sample setup without any other adjustment variables $\boldsymbol{T}_i = (1, T_{i1})$ for $i = 1, \ldots, N$, where T_{i1} is a univariate random variable indicating the group of a subject. A test based on the regression coefficient and their variance estimates of the hypothesis that the corresponding coefficient $\beta_1 = 0$, is a test of equality of parameter of interest θ in two samples. More precisely, hypothesis equals $H_0 : \Delta = \theta_2 - \theta_1 = 0$, where θ_1 is the parameter of interest in the first and θ_2 in the second sample.

While any reasonable link function can be used, some standard choices are given in [20]. For the censored data case, reasonable choice for working covariance matrix is the identity matrix.

Example 2.2.3. The difference between two survival curves at a fixed point in time [20]. Here the pooled sample is considered. The Kaplan-Meier estimator based on all $n_1 + n_2 = N$ observations is denoted as $\widehat{S}_p(t)$. The Kaplan-Meier estimator based on the sample of size $N - 1$ with the ith observation removed is denoted as $\widehat{S}_p^{-i}(t)$ for $i = 1, \ldots, N$. The ith pseudo-observation at some pre-specified time point t_0 is defined as

$$\widehat{\theta}_i = N\widehat{S}_p(t_0) - (N-1)\widehat{S}_p^{-i}(t_0), \quad \text{for} \quad i = 1, \ldots, N. \tag{2.31}$$

Example 2.2.4. The difference between two restricted mean survival times (RMST) [1]. The definition of RMST is given by the equation (2.8). The pooled sample of $n_1 + n_2 = N$ observations is considered. The estimator of RMST is obtained by simply plugging in the Kaplan Meier estimator. Thus, the ith pseudo-observation at some pre-specified time point t_0 is defined

as

$$\widehat{\theta}_i = N \int_0^{t_0} \widehat{S}_p(t)dt - (N-1) \int_0^{t_0} \widehat{S}^{-i}(t)dt. \tag{2.32}$$

Proportional hazards models

The proportional hazards regression model, also known as Cox's regression model in honor of its founder, Sir David R. Cox [5], stands as one of the most widely used models in the field of survival analysis. Here, it is supposed that survival times are only continuous random variables. Let $\boldsymbol{T}_i$ denote covariate vector for ith individual, where $i = 1, \ldots, N$ (based on pooled sample as in previous paragraph).

Definition 2.2.3. [5] The Cox regression model is defined as

$$h(t, T) = e^{T\beta} h_0(t), \tag{2.33}$$

where β is a vector of regression coefficients and $h_0(t)$ is a baseline hazard function.

For two-sample problem, suppose that there is only one covariate T taking values 0 and 1, indicating one of two samples. Then, according to the equation (2.33), the hazard for the first sample ($T = 0$) is equal to baseline hazard $h_0(t)$, and for the second sample ($T = 1$) it is $h_1(t) = e^{\beta} h_0(t)$.

From the latter, it follows that the ratio of both hazard functions is constant, i.e.,

$$\frac{h_1(t)}{h_0(t)} = \frac{e^{\beta} h_0(t)}{h_0(t)} = e^{\beta}. \tag{2.34}$$

Hence, the name proportional hazards, i.e., the regression model is semiparametric as the proportionality of hazard functions needs to be assumed.

2.3 Main results

In this section, the empirical likelihood method in the two-sample setup for right censored data is described. The method was established following the approach by Wang and Jing [53] in one-sample case. The main theoretical result of this chapter is stated in Theorem 2.3.1 and its proof is given in Section 2.4.

2.3.1 Empirical likelihood for censored data in two-sample setup

Consider the previously described two-sample framework with right censored data. The classical Kaplan-Meier product limit estimator of the survival function (2.13) is rewritten in the for as follows

$$\widehat{S}_j(t) = 1 - \widehat{F}_j(t) = \prod_{i=1}^{n_j} \left[\frac{n_j - i}{n_j - i + 1} \right]^{I_{\{Z_{(ji)} \leq t, \delta_{(ji)}=1\}}} \quad \text{for all} \quad t \leq Z_{(jn_j)}, \tag{2.35}$$

where $Z_{(j1)} \leq \ldots \leq Z_{(jn_j)}$ are the ordered statistics, and $\delta_{(ji)}$ is the corresponding censoring indicator. Moreover, $\widehat{S}_j(t) = 0$, when $t > Z_{(jn_j)}$.

The parameter of interest is a functional of distribution functions F_1 and F_2 in the form

$$\Delta := \Delta(F_1, F_2) = \int_0^\infty \xi_2(t) dF_2(t) - \int_0^\infty \xi_1(t) dF_1(t), \tag{2.36}$$

where $\xi_j(t)$ for $j = 1, 2$ is some (non-negative) measurable function. A list of common choices for ξ_j is given below, derived from examples discussed in the literature for one or two-sample problems.

Example 2.3.1. Δ is the difference of k-th moments of survival times, including the difference of mean survival times, if $\xi_j(t) = t^k$ for $j = 1, 2$ (see [46], [53] and [14]).

Example 2.3.2. Δ is the difference of survival probabilities at a fixed time $t_0 \in (0, \infty)$, i.e., $\Delta = S_2(t_0) - S_1(t_0)$, if $\xi_j(t) = I_{\{t \geq t_0\}}$ for $j = 1, 2$ (see [46], [53] and [14]).

Example 2.3.3. Δ is the difference of RMST's at a fixed point $t_0 \in (0, \infty)$, if $\xi_j(t) = \min(t, t_0)$ for $j = 1, 2$ (see [57]).

Example 2.3.4. Δ is the difference of cumulative hazard functions at a fixed time $t_0 \in (0, \infty)$, if $\xi_j(t) = I_{\{t \leq t_0\}}/(1 - F_j(t))$ for $j = 1, 2$ (see [53]).

Example 2.3.5. Δ is the difference of mean residual times at $t_0 \in (0, \infty)$, i.e., $\Delta = mrl_2(t_0) - mrl_1(t_0)$, if $\xi_j(t) = (t - t_0) I_{\{t \geq t_0\}}/(1 - F_j(t_0))$ for $j = 1, 2$ (see [46],[35], [58] and [14]).

The empirical likelihood function in the two-sample case is defined as

$$L(F_1, F_2) = \prod_{i=1}^{n_1} p_i \prod_{k=1}^{n_2} q_k, \tag{2.37}$$

where $p = (p_1, \ldots, p_{n_1})$ and $q = (q_1, \ldots, q_{n_2})$ denote two probability vectors (see [37] for the reference). It is known that $L(F_1, F_2)$ has maximum value $n_1^{-n_1} n_2^{-n_2}$. This leads to the empirical

likelihood ratio function in the form

$$R(F_1, F_2) = \prod_{i=1}^{n_1}(n_1 p_i) \prod_{k=1}^{n_2}(n_2 q_k). \tag{2.38}$$

Now, we extend the two-sample empirical likelihood method for right-censored data following the setup of Wang and Jing [53] for the one sample case (see also Example 1.2.1). First, we will utilize the fact that

$$E\{\xi_j(X_{ji})\} = E\left\{ \frac{\xi_j(Z_{ji})\delta_{ji}}{1 - G_j(Z_{ji})} \right\}, \quad j = 1, 2, \quad \text{and} \quad i = 1, \ldots, n_j. \tag{2.39}$$

This gives us additional information on Δ, which is employed to construct the unbiased estimating equations

$$w_1(Z_{1i}, \Delta, \theta, G_1) = \frac{\xi_1(Z_{1i})\delta_{1i}}{1 - G_1(Z_{1i})} - \theta, \quad i = 1, \ldots, n_1, \tag{2.40}$$

$$w_2(Z_{2k}, \Delta, \theta, G_2) = \frac{\xi_2(Z_{2k})\delta_{2k}}{1 - G_2(Z_{2k})} - \theta - \Delta, \quad k = 1, \ldots, n_2, \tag{2.41}$$

where $\theta := \int_0^\infty \xi_1(t) dF_1(t)$ is a univariate nuisance parameter with the true value $\theta_0 \in \mathbb{R}$. The parameter θ is profiled out by the empirical likelihood method. Censoring functions G_1 and G_2 are nuisance parameters as well, and will be estimated by the Kaplan-Meier product limit estimator in the form (2.35), where $\delta_{(ji)} = 1$ are replaced with $\delta_{(ji)} = 0$ in the power of multipliers. The profile empirical likelihood ratio function is defined as

$$R(\Delta, \theta, G_1, G_2) = \sup_{\theta, p, q} \prod_{i=1}^{n_1}(n_1 p_i) \prod_{k=1}^{n_2}(n_2 q_k), \tag{2.42}$$

where the probability vectors p and q are subject to constraints

$$\begin{aligned} &p_i > 0, \quad \sum_{i=1}^{n_1} p_i = 1, \quad \sum_{i=1}^{n_1} p_i w_1(Z_{1i}, \Delta, \theta, G_1) = 0, \\ &q_k > 0, \quad \sum_{k=1}^{n_2} q_k = 1, \quad \sum_{k=1}^{n_2} q_k w_2(Z_{2k}, \Delta, \theta, G_2) = 0. \end{aligned}$$

For a given θ, a unique maximum of (2.42) exists, provided that 0 is inside the convex hull of the points $w_1(Z_{1i}, \theta, \Delta, G_1)$, $i = 1, \ldots, n_1$, and $w_2(Z_{2k}, \theta, \Delta, G_2)$, $k = 1, \ldots, n_2$. (see [34]). The maximum may be found via Lagrange multiplier method, which gives the probability vectors p

and q in the explicit form

$$p_i = \frac{1}{n_1(1+\lambda_1(\theta)w_1(Z_{1i},\Delta,\theta,G_1))}, \quad i=1,\ldots,n_1,$$
$$q_k = \frac{1}{n_2(1+\lambda_2(\theta)w_2(Z_{2k},\Delta,\theta,G_2))}, \quad k=1,\ldots,n_2,$$

where λ_1 and λ_2 are the Lagrange multipliers. Thus, the minus two log-likelihood ratio statistic is expressed as

$$\begin{aligned}-2\log R(\Delta,\theta,G_1,G_2) = 2\sum_{i=1}^{n_1}\log(1+\lambda_1(\theta)w_1(Z_{1i},\Delta,\theta,G_1)) \\ + 2\sum_{k=1}^{n_2}\log(1+\lambda_2(\theta)w_2(Z_{2k},\Delta,\theta,G_2)).\end{aligned} \tag{2.43}$$

To profile out the unknown parameters θ, $\lambda_1(\theta)$, and $\lambda_2(\theta)$, the following derivative is set $\partial\{-2\log R(\Delta,\theta,\widehat{G}_1,\widehat{G}_2)\}/\partial\theta = 0$, where $\widehat{G}_1$ and $\widehat{G}_2$ are plug-in estimators, and the three-equation system is obtained

$$\frac{\lambda_1(\theta)}{n_1}\sum_{i=1}^{n_1}\{1+\lambda_1(\theta)w_1(Z_{1i},\Delta,\theta,\widehat{G}_1)\}^{-1} + \frac{\lambda_2(\theta)}{n_1}\sum_{k=1}^{n_2}\{1+\lambda_2(\theta)w_2(Z_{2k},\Delta,\theta,\widehat{G}_2)\}^{-1} = 0, \tag{2.44}$$

$$\frac{1}{n_1}\sum_{i=1}^{n_1}\frac{w_1(Z_{1i},\Delta,\theta,\widehat{G}_1)}{1+\lambda_1(\theta)w_1(Z_{1i},\Delta,\theta,\widehat{G}_1)} = 0, \tag{2.45}$$
$$\frac{1}{n_2}\sum_{k=1}^{n_2}\frac{w_2(Z_{2j},\Delta,\theta,\widehat{G}_2)}{1+\lambda_2(\theta)w_2(Z_{2j},\Delta,\theta,\widehat{G}_2)} = 0. \tag{2.46}$$

2.3.2 The conditions and main theorem

To state the following lemma and the main theorem, the following conditions from [53] are defined for both samples. Define for $j=1,2$

$$H_j(s) = P(Z_{j1}\le s), \quad \bar{H}_j(s) = P(Z_{j1} > s),$$
$$\tilde{H}_{0j}(s) = P(Z_{j1} > s, \delta_{j1}=0), \quad \tilde{H}_{1j}(s) = P(Z_{j1} > s, \delta_{j1}=1),$$
$$\gamma_{0j}(x) = \exp\left\{\int_0^{x-}\frac{d\tilde{H}_{0j}(s)}{\bar{H}_j(s)}\right\},$$
$$C_j(x) = \int_0^{x-}\frac{dG_j(s)}{(1-H_j(s))(1-G_j(s))}, \quad \tau_{H_j} = \inf\{t: H_j(t)=1\},$$

where "$x-$" denotes that x is not included. And let

$$\gamma_{1j}(x) = \frac{1}{\bar{H}_j(x)} \int I_{\{x<s\}} \xi_j(s) \gamma_{0j}(s) d\tilde{H}_{1j}(s),$$
$$\gamma_{2j}(x) = \iint \frac{I_{\{s<x, s<t\}} \xi_j(t) \gamma_{0j}(t)}{\bar{H}_j^2(s)} d\tilde{H}_{0j}(s) d\tilde{H}_{1j}(t).$$

The conditions are following

(P1) $\int_0^{\tau_{H_j}} \xi_j^2(x) \gamma_{0j}^2(x) d\tilde{H}_{1j}(x) < \infty$,

(P2) $\int_0^{\tau_{H_j}} \xi_j(x) C_j^{1/2}(x) dF_j(x) < \infty$,

(P3) $\int_0^{\tau_{H_j}} \frac{\xi_j^2(x) dF_j(x)}{1-G_j(x)} < \infty$,

(P4) $\tau_{F_j} = \tau_{H_j}$ and $F_j(\tau_{F_j}) = F_j(\tau_{F_j}-)$.

To derive the limiting distribution of the empirical likelihood statistic (2.43), the following lemma is necessary.

Lemma 2.3.1. *Under the conditions (P1) – (P4) for both samples, the following holds for* $j = 1, 2$*,*

$$\frac{1}{\sqrt{n_j}} \sum_{i=1}^{n_j} w_j(Z_{ji}, \Delta_0, \theta_0, \widehat{G}_j) \xrightarrow{d} N(0, \sigma_j^2), \tag{2.47}$$

$$\frac{1}{n_j} \sum_{i=1}^{n_j} w_j^2(Z_{ji}, \Delta_0, \theta_0, \widehat{G}_j) \xrightarrow{p} V_j. \tag{2.48}$$

The variance σ_j^2, used in Lemma 2.3.1, defines the asymptotic variance of Kaplan-Meier integral (see [45] and [53]) and is expressed in a complex form as follows

$$\sigma_j^2 = \text{Var}(\xi_j(Z_{j1}) \gamma_{0j}(Z_{j1}) \delta_{j1} + \gamma_{1j}(Z_{j1})(1 - \delta_{j1}) - \gamma_{2j}(Z_{j1})) \quad \text{for} \quad j = 1, 2. \tag{2.49}$$

The proof of (2.47) in Lemma 2.3.1 is given in Section 4 of [53]. The proof of (2.48) in Lemma 2.3.1 can be found in Section 3.3 of [16].

Theorem 2.3.1. *Assume* $n_2/n_1 \to m$ *(as* $n_1, n_2 \to \infty$*), where* $0 < m < \infty$*,* $\theta_0 \in \Omega$*, where* Ω *is an open interval, and conditions (P1) – (P4) from [53] are satisfied for both samples. Then*

$$-2r \log R(\Delta_0, \widehat{\theta}, \widehat{G}_1, \widehat{G}_2) \xrightarrow{d} \chi_1^2,$$

as $n_1, n_2 \to \infty$*, where* $r = (V_2 + mV_1)/(\sigma_2^2 + m\sigma_1^2)$ *and* $\widehat{\theta}$ *is a consistent estimator of* θ *obtained by solving the system of equations* (2.44)*,* (2.45) *and* (2.46)*.*

The proof of the Theorem 2.3.1 is given in Section 2.4. Using the result from Theorem 2.3.1, the empirical likelihood-based confidence intervals for the parameter of interest Δ are in the form

$$\{\Delta : -2\widehat{r} \log R(\Delta, \widehat{\theta}, \widehat{G}_1, \widehat{G}_2) \leq \chi^2_{1,1-\alpha}\}, \tag{2.50}$$

where $\chi^2_{1,1-\alpha}$ denotes $1-\alpha$ quantile of the chi-squared distribution with one degree of freedom and $\widehat{r}$ is an estimator for the scaling constant r.

Remark 2.3.1. *There are two particular cases, when the scaling constant in Theorem 2.3.1 is equal to 1, resulting in a limiting distribution of the statistic* (2.43) *as a simple* χ^2_1 *random variable. First, when there is no perturbation due to the censoring, i.e., all survival times have been observed. Second, when the true censoring functions* G_1 *and* G_2 *are known and used rather than their plug-in estimators.*

2.3.3 The estimator of the scaling constant

To find the estimator of the scaling constant r, the variances defined in Lemma 2.3.1 need to be estimated. Similar estimators as proposed by Wang and Jing [53] in one-sample case are used for this purpose.

Regarding V_j, for $j = 1, 2$, a consistent estimator follows from the equation (4.14) of [53] in the form

$$\widehat{V}_j := n_j^{-1} \sum_{i=1}^{n_j} \left(V_{n_j i} - \bar{V}_{n_j}\right)^2, \tag{2.51}$$

where

$$V_{n_j i} = \frac{\xi_j(Z_{ji})\delta_{ji}}{1 - \widehat{G}_j(Z_{ji})}, \quad i = 1, \ldots, n_j, \quad \text{and} \quad \bar{V}_{n_j} = \frac{1}{n_j} \sum_{i=1}^{n_j} V_{n_j i} \quad \text{for} \quad j = 1, 2.$$

Wang and Jing [53] also proposed to use a consistent jackknife estimator for σ_j^2 for $j = 1, 2$, which was established in [45] and is defined below. Another consistent estimator of σ_j^2 was considered by Qin and Zhao [35] specifically for application of the mean residual life in one-sample setup, however, it was not studied within the scope of this thesis.

For simplicity, assume that $j = 1, 2$ throughout this section. Let us denote

$$\theta_j := \theta(F_j) = \int_0^\infty \xi_j(t) dF_j(t).$$

The aim is to estimate the asymptotic variance of a Kaplan-Meier integral defined as

$$\widehat{\theta}_j := \theta(\widehat{F}_j) = \int_0^{Z_{(jn_j)}} \xi_j(t) d\widehat{F}_j(t), \tag{2.52}$$

where $\widehat{F}_j$ is obtained by (2.35). Denote

$$T_{ji} = \frac{\delta_{(ji)}}{n_j - i + 1} \prod_{k=1}^{i-1} \left[\frac{n_j - k}{n_j - k + 1} \right]^{\delta_{(jk)}},$$

then $\widehat{\theta}_j$ can be rewritten as empirical integral

$$S_{n_j} = \sum_{i=1}^{n_j} T_{ji} \xi_j(Z_{(ji)}).$$

Stute [44] proved that central limit theorem holds for censored data in the following form

$$\sqrt{n_j} \left(S_{n_j} - S_j \right) \xrightarrow{d} N(0, \sigma_j^2),$$

where $S_j := \lim_{n_j \to \infty} S_{n_j}$.

Let $\widehat{F}_j^{-k}$ denote the Kaplan-Meier estimator for the sample from which the observation $(Z_{(jk)}, \delta_{(jk)})$ is removed. Then $S_{n_j}^{-k} := \theta(\widehat{F}_j^{-k})$ for $k = 1, \ldots, n_j$ denote the pseudo-values of a jackknife (also referred to as leave-one-out) method. The mean value of the pseudo-values is expressed as

$$\bar{S}_{n_j} = S_{n_j} - \xi_j(Z_{(jn_j)}) \frac{\delta_{(jn_j)}(1 - \delta_{(j(n_j-1))})}{n_j} \prod_{i=1}^{n_j-2} \left[\frac{n_j - i - 1}{n_j - i} \right]^{\delta_{(ji)}}. \tag{2.53}$$

Finally, the jackknife variance estimator is defined as

$$\widehat{\sigma}_j^2 = (n_j - 1) \sum_{k=1}^{n_j} \left(S_{n_j}^{-k} - \bar{S}_{n_j} \right)^2. \tag{2.54}$$

Theorem 2.3.2. [45] *Suppose condition (P1) holds, then $\widehat{\sigma}_j^2 \to \sigma_j^2$ almost sure.*

Corollary 2.3.1. *Under the conditions of Theorem 2.3.1, the estimator of the scaling constant r is obtained by*

$$\widehat{r} = \frac{\widehat{V}_2 + m\widehat{V}_1}{\widehat{\sigma}_2^2 + m\widehat{\sigma}_1^2}. \tag{2.55}$$

2.4 Proof of Theorem 2.3.1

In this section, first two additional lemmas are stated, which are necessary to prove the result in Theorem 2.3.1 from Section 2.3.2. Afterwards, the proof of the Theorem 2.3.1 is given.

2.4.1 Technical lemmas

Lemma 2.4.1. *Suppose that* $1/3 < \eta < 1/2$ *and the conditions of Theorem 2.3.1 are satisfied, then*

$$\lambda_1(\theta) = O_p(n_1^{-\eta}) \quad \textit{and} \quad \lambda_2(\theta) = O_p(n_1^{-\eta}) \tag{2.56}$$

uniformly around $\theta \in \{\theta : |\theta - \theta_0| \leq cn_1^{-\eta}\}$*, where* c *is some positive constant. Moreover, with probability tending to 1, there exists a root* $\widehat{\theta}$ *of* (2.44) *such that*

$$|\widehat{\theta} - \theta_0| = O_p(n_1^{-\eta}), \tag{2.57}$$

and $R(\Delta_0, \theta, \widehat{G}_1, \widehat{G}_2)$ *attains its local maximum value at* $\widehat{\theta}$*.*

The proof of Lemma 2.4.1 is given in [37] with the second assumption (ii) from Section 1.3 satisfied in our setup. See also Section 3.3 in [16] for details.

Lemma 2.4.2. *Assume that the conditions of Theorem 2.3.1 are satisfied, then with root* $\widehat{\theta}$ *of* (2.44)*, as given in Lemma 2.4.1, we have*

$$\sqrt{n_1}(\widehat{\theta} - \theta_0) \xrightarrow{d} N\left(0, \frac{\sigma_1^2 V_2 + m\sigma_2^2 V_1}{(V_2 + mV_1)^2}\right), \tag{2.58}$$

$$\lambda_1(\widehat{\theta}) = -m\lambda_2(\widehat{\theta}) + o_p(n_1^{-1/2}), \tag{2.59}$$

$$\sqrt{n_1}\lambda_2(\widehat{\theta}) \xrightarrow{d} N\left(0, \frac{\sigma_2^2 + m\sigma_1^2}{m(V_2 + mV_1)^2}\right). \tag{2.60}$$

Proof. Denote $\lambda_1 := \lambda_1(\theta)$, $\widehat{\lambda}_1 := \lambda_1(\widehat{\theta})$, $\lambda_2 := \lambda_2(\theta)$, $\widehat{\lambda}_2 := \lambda_2(\widehat{\theta})$ for simplicity. Denote equations (2.45), (2.46) and (2.44) by

$$Q_1(\theta, \lambda_1, \lambda_2) = \frac{1}{n_1}\sum_{i=1}^{n_1} \frac{w_1(Z_{1i}, \Delta, \theta, \widehat{G}_1)}{1 + \lambda_1 w_1(Z_{1i}, \theta, \Delta, \widehat{G}_1)},$$

$$Q_2(\theta, \lambda_1, \lambda_2) = \frac{1}{n_2}\sum_{k=1}^{n_2} \frac{w_2(Z_{2k}, \Delta, \theta, \widehat{G}_2)}{1 + \lambda_2 w_2(Z_{2k}, \theta, \Delta, \widehat{G}_2)},$$

$$Q_3(\theta, \lambda_1, \lambda_2) = \lambda_1 \frac{1}{n_1}\sum_{i=1}^{n_1} \{1 + \lambda_1 w_1(Z_{1i}, \Delta, \theta, \widehat{G}_1)\}^{-1} + \lambda_2 \frac{1}{n_1}\sum_{k=1}^{n_2} \{1 + \lambda_2 w_2(Z_{2k}, \Delta, \theta, \widehat{G}_2)\}^{-1},$$

respectively. From Lemma 2.4.1 we have $Q_i(\widehat{\theta}, \widehat{\lambda}_1, \widehat{\lambda}_2) = 0$ for $i = 1, 2, 3$. By Lemma 2.4.1 and using Taylor expansion, the following holds

$$\begin{aligned}
0 =& Q_i(\widehat{\theta}, \widehat{\lambda}_1, \widehat{\lambda}_2) \\
=& Q_i(\theta_0, 0, 0) + \frac{\partial Q_i(\theta_0, 0, 0)}{\partial \theta}(\widehat{\theta} - \theta_0) \\
&+ \frac{\partial Q_i(\theta_0, 0, 0)}{\partial \lambda_1}\widehat{\lambda}_1 + \frac{\partial Q_i(\theta_0, 0, 0)}{\partial \lambda_2}\widehat{\lambda}_2 + O_p(n_1^{-2\eta}), \quad i = 1, 2, 3.
\end{aligned}$$

The partial derivatives of Q_i, for $i = 1, 2, 3$, give

$$\begin{array}{lll}
\frac{\partial Q_1(\theta_0, 0, 0)}{\partial \theta} \to -1 \text{ a.s.}, & \frac{\partial Q_1(\theta_0, 0, 0)}{\partial \lambda_1} \to -V_1 \text{ a.s.}, & \frac{\partial Q_1(\theta_0, 0, 0)}{\partial \lambda_2} = 0, \\
\frac{\partial Q_2(\theta_0, 0, 0)}{\partial \theta} \to -1 \text{ a.s.}, & \frac{\partial Q_2(\theta_0, 0, 0)}{\partial \lambda_1} = 0, & \frac{\partial Q_2(\theta_0, 0, 0)}{\partial \lambda_2} \to -V_2 \text{ a.s.}, \\
\frac{\partial Q_3(\theta_0, 0, 0)}{\partial \theta} = 0, & \frac{\partial Q_3(\theta_0, 0, 0)}{\partial \lambda_1} \to -1 \text{ a.s.}, & \frac{\partial Q_3(\theta_0, 0, 0)}{\partial \lambda_2} \to -m \text{ a.s.}
\end{array}$$

Therefore

$$\begin{pmatrix} \widehat{\theta} - \theta_0 \\ \widehat{\lambda}_1 \\ \widehat{\lambda}_2 \end{pmatrix} = S^{-1} \begin{pmatrix} Q_1(\theta_0, 0, 0) \\ Q_2(\theta_0, 0, 0) \\ 0 \end{pmatrix} + o_p(n_1^{-1/2}),$$

where

$$S = \begin{pmatrix} -1 & -V_1 & 0 \\ -1 & 0 & -V_2 \\ 0 & -1 & -m \end{pmatrix} \quad \text{and} \quad S^{-1} = \frac{1}{V_2 + mV_1} \begin{pmatrix} -V_2 & -mV_1 & V_1V_2 \\ -m & m & -V_2 \\ 1 & -1 & -V_1 \end{pmatrix}.$$

Finally, the following equations are obtained

$$\widehat{\theta} - \theta_0 = \frac{1}{V_2 + mV_1}\left(-V_2 Q_1(\theta_0, 0, 0) - mV_1 Q_2(\theta_0, 0, 0)\right) + o_p(n_1^{-1/2}), \tag{2.61}$$

$$\lambda_1 = m\left(\frac{1}{V_2 + mV_1}\right)\left(-Q_1(\theta_0, 0, 0) + Q_2(\theta_0, 0, 0)\right) + o_p(n_1^{-1/2}), \tag{2.62}$$

$$\lambda_2 = -\left(\frac{1}{V_2 + mV_1}\right)\left(-Q_1(\theta_0, 0, 0) + Q_2(\theta_0, 0, 0)\right) + o_p(n_1^{-1/2}). \tag{2.63}$$

From Lemma 2.3.1, it holds that

$$\sqrt{n_1}\begin{pmatrix} Q_1(\theta_0, 0, 0) \\ Q_2(\theta_0, 0, 0) \end{pmatrix} \xrightarrow{d} N\left(0, \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & m^{-1}\sigma_2^2 \end{bmatrix}\right),$$

which together with (2.61), (2.62) and (2.63) completes the proof of Lemma 2.4.2.

2.4.2 Proof of Theorem 2.3.1

Here the proof of the main theorem is given. First, following the argument for equation (4.8) in [53] and multiplying both sides by $\lambda_1(\widehat{\theta})$ and $\lambda_2(\widehat{\theta})$, respectively, the following equations are obtained

$$\lambda_1(\widehat{\theta})\sum_{i=1}^{n_1} w_1(Z_{1i},\Delta_0,\widehat{\theta},\widehat{G}_1) = \lambda_1^2(\widehat{\theta})\sum_{i=1}^{n_1} w_1^2(Z_{1i},\Delta_0,\widehat{\theta},\widehat{G}_1) + o_p(1)\,, \tag{2.64}$$

$$\lambda_2(\widehat{\theta})\sum_{k=1}^{n_2} w_2(Z_{2k},\Delta_0,\widehat{\theta},\widehat{G}_2) = \lambda_2^2(\widehat{\theta})\sum_{k=1}^{n_2} w_2^2(Z_{2k},\Delta_0,\widehat{\theta},\widehat{G}_2) + o_p(1)\,. \tag{2.65}$$

Next, from Taylor expansion

$$\begin{aligned}
-2\log R(\Delta_0,\widehat{\theta},\widehat{G}_1,\widehat{G_2}) =& 2\lambda_1(\widehat{\theta})\sum_{i=1}^{n_1} w_1(Z_{1i},\Delta_0,\widehat{\theta},\widehat{G}_1) - \lambda_1^2(\widehat{\theta})\sum_{i=1}^{n_1} w_1^2(Z_{1i},\Delta_0,\widehat{\theta},\widehat{G}_1)\\
&+2\lambda_2(\widehat{\theta})\sum_{k=1}^{n_2} w_2(Z_{2k},\Delta_0,\widehat{\theta},\widehat{G}_2) - \lambda_2^2(\widehat{\theta})\sum_{k=1}^{n_2} w_2^2(Z_{2k},\Delta_0,\widehat{\theta},\widehat{G}_2) + o_p(1)\\
=&\lambda_1^2(\widehat{\theta})\sum_{i=1}^{n_1} w_1^2(Z_{1i},\Delta_0,\widehat{\theta},\widehat{G}_1) + \lambda_2^2(\widehat{\theta})\sum_{k=1}^{n_2} w_2^2(Z_{2k},\Delta_0,\widehat{\theta},\widehat{G}_2) + o_p(1)\\
=&m^2\lambda_2^2(\widehat{\theta})\sum_{i=1}^{n_1} w_1^2(Z_{1i},\Delta_0,\widehat{\theta},\widehat{G}_1) + \lambda_2^2(\widehat{\theta})\sum_{k=1}^{n_2} w_2^2(Z_{2k},\Delta_0,\widehat{\theta},\widehat{G}_2) + o_p(1)\\
=&n_1m^2\lambda_2^2(\widehat{\theta})V_1 + n_2\lambda_2^2(\widehat{\theta})V_2 + o_p(1)\\
=&m(V_2+mV_1)\left(\sqrt{n_1}\lambda_2(\widehat{\theta}\right)^2 + o_p(1)
\end{aligned}$$

And finally, employing the equation (2.60) from Lemma 2.4.2, it is concluded that

$$-2r\log R(\Delta_0,\widehat{\theta},\widehat{G}_1,\widehat{G}_2) \xrightarrow{d} \chi_1^2 \quad \text{with} \quad r=\frac{V_2+mV_1}{\sigma_2^2+m\sigma_1^2},$$

which completes the proof of Theorem 2.3.1.

2.5 Simulation study

In this section the results of a simulation study are presented. The coverage accuracy of confidence intervals was calculated for two scenarios listed below.

I: the difference of survival probabilities at fixed time t_0 (see Example 2.3.2 from Section

2.3);

II: the difference of restricted mean survival times (RMST's) at time t_0 (see Example 2.3.3 from Section 2.3).

In both cases, two other methods were used for comparison. The first method is based on normal approximation (see Section 2.2). The second method is based on pseudo-observations (see Section 2.2). In the case of RMSTs, the empirical likelihood-based approach by Zhou [57] has also been considered for comparison. However, it was computationally very intensive and the solution to the optimization problem could not be found in some cases. Consequently, it was not included in this simulation study.

To simulate censored data, first survival times and censoring times from the exponential distribution with parameters 1 and c, respectively, i.e., $F_j(t) = 1 - \exp(-t)$ and $G_j(t) = 1 - \exp(-ct)$ for $t \geq 0$ were generated. Then the right censored data is obtained as pairs (Z_{ji}, δ_{ji}), $j = 1, 2$. The constant c was chosen as 0.111, 0.250 and 0.429 such that on average 10%, 20% or 30% of the observations were censored, respectively. Samples were generated $N = 10\,000$ times with equal sample sizes $n_1 = n_2$. For each case the 95% confidence interval for Δ was constructed at the specified value of t_0, and the coverage accuracy was calculated for $H_0 : \Delta = 0$.

The implementation of the two-sample empirical likelihood method for right censored data established in Section 2.3.2 of this thesis, is based on an existing R library `EL` [50], which covers many different two-sample problems for *i.i.d.* data, including the inference for the difference of two mean values. However for right censored data, only the statistic value calculated by the function `EL.means` was used. Additional steps of implementation was needed to obtain the estimate of the scaling constant $\widehat{r}$ to construct the confidence intervals for Δ. The normal approximation was implemented using packages `survival` and `survRM2` for survival probabilities and RMST's, respectively. The pseudo-observations approach was carried out using packages `pseudo` and `geepack`.

The results of the simulations for the 95% confidence interval coverage accuracy are summarized in Tables 2.2 and 2.3 for the difference of the survival probabilities (Scenario I) and for the difference of the RMST's (Scenario II), respectively. When sample sizes were $n_1 = n_2 = 15$, coverage accuracy could not be calculated for certain cases (e.g., when there were samples without events after the selected time t_0), denoted as "NA". Results of the simulations lead to the following conclusions.

(1) The empirical likelihood method gives better confidence interval coverage accuracy com-

Table 2.2: Confidence interval coverage accuracy (mean interval length) for the difference of survival probabilities at the point t_0 obtained by normal approximation, pseudo-observations (PO) and empirical likelihood (EL). Survival times were simulated from exponential distribution with rate parameter 1, and censoring times were simulated from exponential distribution with rate parameter c, chosen such that the pre-specified censoring proportion was obtained. Number of replications $N = 10\,000$, $\alpha = 0.05$. NA – result not available.

Censoring proportion	(n_1, n_2)	$t_0 = 0.5$			$t_0 = 0.7$		
		Normal approx.	PO	EL	Normal approx.	PO	EL
0.1	(15, 15)	0.930 (0.683)	0.932 (0.685)	0.950 (0.696)	NA	0.936 (0.707)	0.949 (0.720)
	(20, 20)	0.933 (0.597)	0.933 (0.599)	0.949 (0.606)	0.936 (0.616)	0.936 (0.618)	0.951 (0.627)
	(30, 30)	0.936 (0.493)	0.939 (0.493)	0.949 (0.498)	0.941 (0.508)	0.941 (0.509)	0.949 (0.514)
	(50, 50)	0.944 (0.384)	0.945 (0.385)	0.950 (0.387)	0.946 (0.396)	0.947 (0.397)	0.951 (0.399)
	(70, 70)	0.949 (0.326)	0.950 (0.326)	0.953 (0.327)	0.945 (0.336)	0.946 (0.336)	0.951 (0.338)
	(100, 100)	0.951 (0.273)	0.951 (0.274)	0.954 (0.274)	0.949 (0.282)	0.950 (0.282)	0.952 (0.283)
	(150, 150)	0.951 (0.224)	0.952 (0.224)	0.953 (0.224)	0.951 (0.230)	0.950 (0.231)	0.952 (0.231)
	(200, 200)	0.948 (0.194)	0.948 (0.194)	0.949 (0.194)	0.947 (0.200)	0.947 (0.200)	0.949 (0.200)
0.2	(15, 15)	0.928 (0.695)	0.933 (0.700)	0.946 (0.718)	NA	0.933 (0.729)	0.942 (0.754)
	(20, 20)	0.934 (0.609)	0.937 (0.611)	0.948 (0.624)	0.934 (0.633)	0.936 (0.637)	0.946 (0.654)
	(30, 30)	0.936 (0.502)	0.939 (0.503)	0.946 (0.511)	0.938 (0.522)	0.939 (0.524)	0.947 (0.534)
	(50, 50)	0.944 (0.392)	0.945 (0.392)	0.949 (0.396)	0.945 (0.407)	0.946 (0.408)	0.950 (0.413)
	(70, 70)	0.949 (0.332)	0.949 (0.333)	0.951 (0.335)	0.945 (0.345)	0.946 (0.346)	0.949 (0.349)
	(100, 100)	0.952 (0.279)	0.953 (0.279)	0.955 (0.280)	0.947 (0.290)	0.947 (0.290)	0.949 (0.292)
	(150, 150)	0.952 (0.228)	0.952 (0.228)	0.954 (0.229)	0.952 (0.237)	0.952 (0.237)	0.953 (0.238)
	(200, 200)	0.948 (0.198)	0.950 (0.198)	0.951 (0.198)	0.948 (0.205)	0.949 (0.205)	0.949 (0.206)
0.3	(15, 15)	0.927 (0.711)	0.934 (0.719)	0.943 (0.745)	NA	0.932 (0.759)	0.939 (0.800)
	(20, 20)	0.931 (0.623)	0.936 (0.627)	0.945 (0.646)	0.929 (0.655)	0.937 (0.662)	0.941 (0.688)
	(30, 30)	0.935 (0.514)	0.940 (0.517)	0.944 (0.528)	0.938 (0.541)	0.942 (0.544)	0.945 (0.560)
	(50, 50)	0.943 (0.401)	0.946 (0.403)	0.948 (0.409)	0.942 (0.422)	0.944 (0.424)	0.947 (0.432)
	(70, 70)	0.949 (0.341)	0.950 (0.341)	0.952 (0.345)	0.944 (0.358)	0.946 (0.359)	0.948 (0.364)
	(100, 100)	0.949 (0.286)	0.951 (0.286)	0.953 (0.289)	0.947 (0.300)	0.948 (0.301)	0.949 (0.304)
	(150, 150)	0.952 (0.234)	0.952 (0.234)	0.954 (0.236)	0.951 (0.246)	0.952 (0.246)	0.952 (0.248)
	(200, 200)	0.951 (0.203)	0.951 (0.203)	0.952 (0.204)	0.948 (0.213)	0.948 (0.213)	0.949 (0.215)

pared to other studied methods for small to moderate sample sizes (up to $n_1 = n_2 = 100$). This is true for almost all studied cases.

(2) For small to moderate sample sizes, the coverage accuracy of our proposed method tends to be closer to the nominal level as the proportion of censoring decreases. Meanwhile, it appears that an increase in censoring proportion does not significantly affect the performance of other methods.

(3) For larger sample sizes, the empirical likelihood-based confidence intervals has slightly inflated coverage accuracy in some cases, especially for the difference of survival probabilities at $t_0 = 0.5$ (see Table 2.2).

(4) The confidence intervals obtained by the empirical likelihood method tend to be wider on average compared to other methods, especially for small to moderate sample sizes. This holds true for larger sample sizes as well, either when the censoring proportion is increased or t_0 is chosen further from the origin.

Table 2.3: Confidence interval coverage accuracy (mean interval length) for the difference of restricted mean survival times at the point t_0 obtained by normal approximation, pseudo-observations (PO) and empirical likelihood (EL) method. Survival times were simulated from exponential distribution with rate parameter 1, and censoring times were simulated from exponential distribution with rate parameter c, chosen such that the pre-specified censoring proportion was obtained. Number of replications $N = 10\,000$, $\alpha = 0.05$. NA – result not available.

Censoring proportion	(n_1, n_2)	$t_0 = 0.5$			$t_0 = 1$		
		Normal approx.	PO	EL	Normal approx.	PO	EL
0.1	(15, 15)	0.933 (0.221)	0.933 (0.221)	0.948 (0.228)	0.931 (0.502)	0.932 (0.502)	0.945 (0.516)
	(20, 20)	0.937 (0.193)	0.937 (0.193)	0.949 (0.198)	0.934 (0.439)	0.934 (0.439)	0.946 (0.448)
	(30, 30)	0.940 (0.160)	0.940 (0.160)	0.948 (0.162)	0.940 (0.362)	0.940 (0.362)	0.949 (0.367)
	(50, 50)	0.941 (0.125)	0.941 (0.125)	0.947 (0.126)	0.943 (0.282)	0.944 (0.282)	0.947 (0.285)
	(70, 70)	0.947 (0.106)	0.948 (0.106)	0.951 (0.106)	0.947 (0.239)	0.948 (0.239)	0.951 (0.241)
	(100, 100)	0.950 (0.089)	0.949 (0.089)	0.952 (0.089)	0.951 (0.201)	0.951 (0.201)	0.953 (0.202)
	(150, 150)	0.948 (0.073)	0.948 (0.073)	0.950 (0.073)	0.949 (0.164)	0.949 (0.164)	0.951 (0.165)
	(200, 200)	0.948 (0.063)	0.949 (0.063)	0.950 (0.063)	0.947 (0.142)	0.947 (0.142)	0.948 (0.143)
0.2	(15, 15)	0.932 (0.223)	0.932 (0.223)	0.944 (0.230)	0.929 (0.510)	0.930 (0.512)	0.940 (0.530)
	(20, 20)	0.937 (0.195)	0.938 (0.195)	0.946 (0.200)	0.934 (0.447)	0.935 (0.448)	0.944 (0.460)
	(30, 30)	0.939 (0.161)	0.940 (0.161)	0.945 (0.164)	0.938 (0.368)	0.938 (0.369)	0.945 (0.376)
	(50, 50)	0.941 (0.126)	0.942 (0.126)	0.945 (0.127)	0.945 (0.287)	0.946 (0.288)	0.949 (0.291)
	(70, 70)	0.947 (0.107)	0.947 (0.107)	0.949 (0.108)	0.946 (0.243)	0.946 (0.244)	0.948 (0.246)
	(100, 100)	0.949 (0.090)	0.950 (0.090)	0.951 (0.090)	0.950 (0.204)	0.950 (0.204)	0.951 (0.206)
	(150, 150)	0.949 (0.073)	0.949 (0.073)	0.950 (0.074)	0.950 (0.167)	0.950 (0.167)	0.951 (0.168)
	(200, 200)	0.949 (0.063)	0.949 (0.064)	0.950 (0.064)	0.949 (0.145)	0.949 (0.145)	0.950 (0.145)
0.3	(15, 15)	0.932 (0.225)	0.933 (0.226)	0.940 (0.234)	NA	0.929 (0.525)	0.935 (0.547)
	(20, 20)	0.936 (0.197)	0.936 (0.198)	0.943 (0.203)	0.932 (0.457)	0.935 (0.459)	0.938 (0.474)
	(30, 30)	0.941 (0.163)	0.941 (0.163)	0.946 (0.166)	0.939 (0.377)	0.939 (0.378)	0.942 (0.387)
	(50, 50)	0.942 (0.127)	0.942 (0.127)	0.945 (0.129)	0.942 (0.294)	0.943 (0.294)	0.946 (0.299)
	(70, 70)	0.947 (0.108)	0.947 (0.108)	0.949 (0.109)	0.947 (0.249)	0.947 (0.249)	0.949 (0.253)
	(100, 100)	0.950 (0.091)	0.950 (0.091)	0.951 (0.091)	0.946 (0.209)	0.947 (0.209)	0.947 (0.211)
	(150, 150)	0.948 (0.074)	0.948 (0.074)	0.949 (0.074)	0.950 (0.171)	0.950 (0.171)	0.951 (0.172)
	(200, 200)	0.948 (0.064)	0.949 (0.064)	0.950 (0.064)	0.950 (0.148)	0.951 (0.148)	0.951 (0.149)

2.6 Applications to real data sets

In this section two real right censored data examples are studied for the illustration of the two-sample empirical likelihood method.

2.6.1 Lung cancer survival time analysis

A comparison of lung cancer survival times between males and females are presented. The data is sourced from the `cancer` dataset included in the R library `survival`. The dataset comprises 138 survival times for males and 90 for females, with censoring rates of approximately 19% and 41%, respectively. The Kaplan-Meier survival curves for both male and female groups are demonstrated in panel A of Figure 2.1.

As suggested by the estimated survival curves, overall survival is better in females, which is also confirmed by the log-rank test ($p = 0.0013$). The difference of the survival probabilities

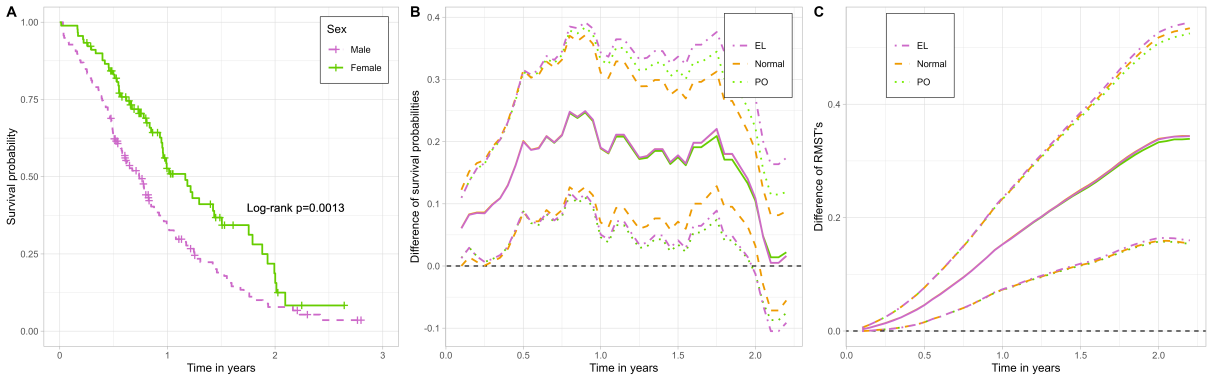


Figure 2.1: **A** – Kaplan-Meier estimates of survival curves in groups of males and females for lung cancer survival data, p-value of log-rank test for survival curves is represented; **B** – The estimated difference function of lung cancer survival probabilities between females and males with 95% pointwise confidence intervals obtained using the empirical likelihood method (EL), normal approximation (Normal), and the pseudo-observations approach (PO); **C** – The estimated difference function of lung cancer restricted mean survival times (RMST's) between females and males with 95% pointwise confidence intervals obtained using the empirical likelihood method (EL), normal approximation (Normal), and the pseudo-observations approach (PO)

and RMST's at $t_0 = 1$ and 2 years was analyzed. The obtained results, along with the results for two other methods (similar to those used in Section 3.5.2) for comparison, are summarized in Table 2.4.

Table 2.4: Results of lung cancer survival time analysis stratified by groups of sex. Parameters of interest analyzed were the difference of survival probabilities and the difference of RMST's at time $t_0 = 1$ year and 2 years. Estimates, along with their corresponding 95% confidence intervals and interval lengths, were obtained using three methods: normal approximation, pseudo-observations approach, and two-sample empirical likelihood method for right-censored data.

Lung cancer: female vs. male		**Difference of survival probabilities**			**Difference of RMST's**		
t_0	Method	Estimate	Confidence interval	Length	Estimate	Confidence interval	Length
1 year	Normal approx.	0.190	[0.070, 0.311]	0.241	0.153	[0.073, 0.234]	0.161
	Pseudo-observations	0.189	[0.046, 0.332]	0.285	0.153	[0.073, 0.233]	0.160
	Empirical likelihood	0.190	[0.051, 0.343]	0.293	0.153	[0.074, 0.235]	0.162
2 years	Normal approx.	0.109	[0.032, 0.186]	0.153	0.339	[0.159, 0.518]	0.359
	Pseudo-observations	0.105	[-0.012, 0.222]	0.234	0.333	[0.158, 0.508]	0.350
	Empirical likelihood	0.109	[-0.013, 0.272]	0.285	0.338	[0.164, 0.527]	0.363

The estimates obtained using the empirical likelihood method coincide with those derived in the normal approximation approach, but exhibit slight differences from the estimates provided by the pseudo-observations approach. On average, empirical likelihood-based confidence intervals are wider compared to the other methods, with this disparity being particularly evident when comparing survival probabilities. At the two-year timepoint ($t_0 = 2$ years), only the normal approximation confidence interval indicates higher survival probability in females com-

pared to males. Both the empirical likelihood method and the pseudo-observations approach do not reject the equality of survival probabilities at $t_0 = 2$ years.

If no particular pre-specified time point of interest t_0 is known, one can construct the difference function for all $t_0 \in (0, \tau)$, where $\tau = \inf\{t : S_1(t) = 0 \,\&\, S_2(t) = 0\}$, along with pointwise confidence intervals. Panels B and C in Figure 2.1 demonstrate the difference function, along with 95% pointwise confidence intervals, for both the survival probabilities and the restricted mean survival times (RMSTs) in the cancer survival data. It is evident that the RMST's significantly differ for all values of t_0. However, the difference function for survival probabilities indicates equality between males and females in survival probabilities after approximately $t_0 = 2$ years.

2.6.2 Time to successful treatment response in patients with tuberculosis

The patients with pulmonary and extrapulmonary drug-susceptible tuberculosis at Riga East University Hospital, Centre of Tuberculosis and Lung Diseases, were included in the study. Isoniazid (INH) is a key drug of the currently recommended treatment regimen. N-acetyltransferase 2 (NAT2) is an enzyme responsible for metabolizing various drugs and chemicals in the human body, including INH. Glutathione S-transferase M1 (GSTM1) is an enzyme involved in the detoxification of various endogenous and exogenous compounds, including drugs, carcinogens, and environmental toxins. Both NAT2 acetylator phenotypes and GSTM1 genotypes were investigated to assess their impact on patients' treatment response time. The time to successful treatment response was measured as the days until sputum culture conversion to culture-negative for tuberculosis bacteria. Individuals who did not achieve a successful treatment response were considered as censored.

A total of 40 patients were included in the study cohort. Among them, 21 patients had the null GSTM1 genotype, and 19 patients had the plus GSTM1 genotype. Within the plus genotype group, 6 patients were considered censored, and in the null genotype group, 4 patients were censored, resulting in a censoring rate of 32% and 19%, respectively. Furthermore, 14 patients exhibited the intermediate NAT2 acetylator phenotype (IA), while 26 patients had the slow NAT2 acetylator phenotype (SA). The censoring rates were 21% (3 patients) for IA and 27% (7 patients) for SA when stratifying by NAT2 acetylator phenotype. The estimated cumulative probabilities of successful treatment response by groups obtained using complement of Kaplan-Meier curves are illustrated in panel A of Figure 2.2 and 2.3 for GSTM1 and NAT2, respectively.

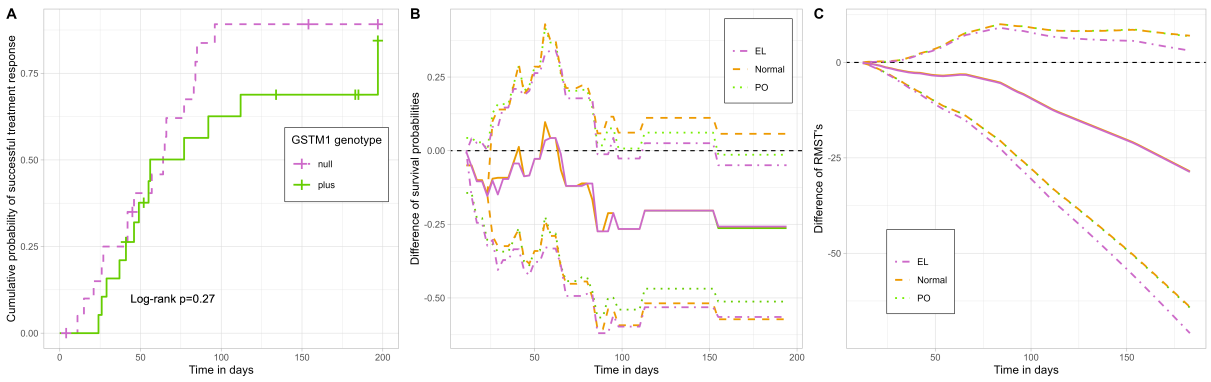


Figure 2.2: **A** – Estimated probabilities of successful treatment response obtained using complement of Kaplan-Meier curves by groups of GSTM1 genotypes in tuberculosis patient data, p-value of log-rank test for survival curves are represented; **B** – The estimated difference function in tuberculosis patients data between survival probabilities of null and plus GSTM1 genotypes with 95% pointwise confidence intervals obtained using the empirical likelihood method (EL), normal approximation (Normal), and the pseudo-observations approach (PO); **C** – The estimated difference function in tuberculosis patients data between restricted mean survival times (RMST's) of null and plus GSTM1 genotypes with 95% pointwise confidence intervals obtained using the empirical likelihood method (EL), normal approximation (Normal), and the pseudo-observations approach (PO).

The empirical likelihood method for two-sample right censored data introduced in this paper was employed to conduct separate comparisons of time to successful treatment response for GSTM1 groups and NAT2 groups. The obtained results, along with the results for two other methods for comparison, are summarized in Table 2.5 for the differences of survival probabilities and RMST's at $t_0 = 60, 90, 120, 180$ days. The chosen value of $t_0 = 60$ days corresponds to the standard two-month intensive phase of the tuberculosis treatment regimen, a time point of significant interest among healthcare professionals.

Interpreting the results requires caution, particularly when considering healing time as the event of interest, as in this case. However, here we only intend to draw conclusions about the methods used for analysis, and the following observations were made. The estimates obtained by the empirical likelihood method coincides with estimates obtained by normal approximation and are close to pseudo-observations estimates in case of survival probabilities. In case of RMST's, estimates slightly differs from both other methods. As already concluded in Section 3.5.2, confidence intervals obtained using our method are wider for the difference of RMST's for both enzymes and for all t_0. However, for the difference of survival probabilities in some cases normal approximation gives wider intervals compared to our method.

The results in Table 2.5 show discrepancies among the employed methods in only a few cases, all regarding to the difference in survival probabilities. For instance, discrepancies are

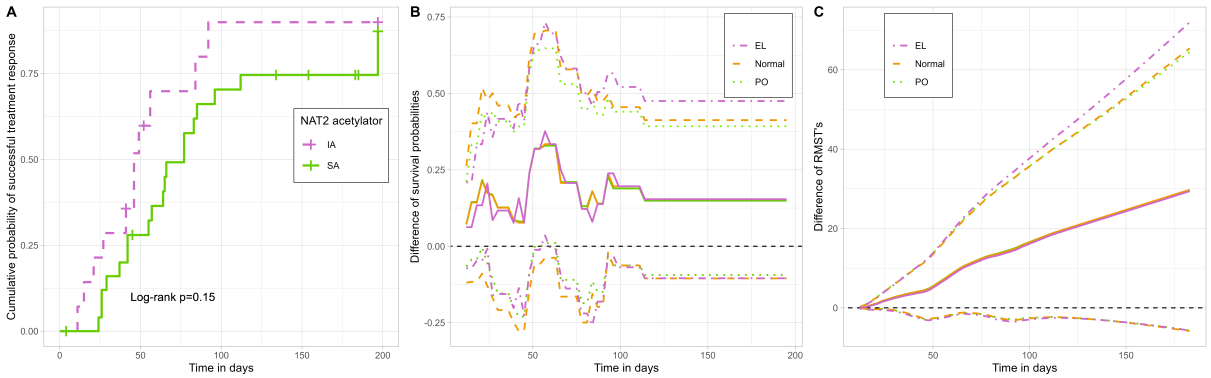


Figure 2.3: **A** – Estimated probabilities of successful treatment response obtained using complement of Kaplan-Meier curves by groups of NAT2 acetylator phenotypes in tuberculosis patient data, p-value of log-rank test for survival curves are represented; **B** – The estimated difference function in tuberculosis patients data between survival probabilities of SA and IA NAT2 acetylator phenotypes with 95% pointwise confidence intervals obtained using the empirical likelihood method (EL), normal approximation (Normal), and the pseudo-observations approach (PO); **C** – The estimated difference function in tuberculosis patients data between restricted mean survival times (RMST's) of SA and IA NAT2 acetylator phenotypes with 95% pointwise confidence intervals obtained using the empirical likelihood method (EL), normal approximation (Normal), and the pseudo-observations approach (PO).

observed in groups of GSTM1 genotypes at $t_0 = 90$ and 180 days, and in the NAT2 acetylator phenotype at $t_0 = 60$ days.

The plots of the estimated difference functions with 95% pointwise confidence intervals were also constructed (see panels B and C of Figures 2.2 and 2.3), to better observe the overall trends of differences over all values of t_0.

Table 2.5: Results of time to successful treatment response analysis in patients with tuberculosis stratified by groups of the GSTM1 genotype and NAT2 acetylator phenotype. Parameters of interest analyzed were the difference of survival probabilities and the difference of RMST's at time $t_0 = 60, 90, 120, 180$ days. Estimates, along with their corresponding 95% confidence intervals and interval lengths, were obtained using three methods: normal approximation, pseudo-observations approach, and two-sample empirical likelihood method for right-censored data.

GSTM1 genotype: null vs. plus		**Difference of survival probabilities**			**Difference of RMST's**		
t_0	Method	Estimate	Confidence interval	Length	Estimate	Confidence interval	Length
60 days	Normal approxim.	0.043	[-0.289, 0.375]	0.67	-3.22	[-12.60, 6.17]	18.77
	Pseudo-observations	0.043	[-0.281, 0.367]	0.65	-3.22	[-12.57, 6.13]	18.71
	Empirical likelihood	0.043	[-0.332, 0.340]	0.67	-3.38	[-13.66, 5.71]	19.37
90 days	Normal approx.	-0.274	[-0.607, 0.059]	0.665	-6.86	[-23.38, 9.66]	33.03
	Pseudo-observations	-0.274	[-0.567, 0.018]	0.584	-6.87	[-23.39, 9.66]	33.05
	Empirical likelihood	-0.274	[-0.620, -0.012]	0.607	-7.03	[-25.66, 8.56]	34.23
120 days	Normal approx.	-0.203	[-0.518, 0.111]	0.63	-14.14	[-36.51, 8.24]	44.74
	Pseudo-observations	-0.204	[-0.469, 0.061]	0.53	-14.15	[-36.58, 8.28]	44.86
	Empirical likelihood	-0.203	[-0.532, 0.025]	0.56	-14.30	[-40.11, 6.36]	46.47
180 days	Normal approx.	-0.258	[-0.572, 0.057]	0.63	-27.75	[-62.65, 7.16]	69.81
	Pseudo-observations	-0.263	[-0.512, -0.014]	0.50	-27.92	[-62.85, 7.02]	69.87
	Empirical likelihood	-0.258	[-0.565, -0.049]	0.52	-27.91	[-69.37, 3.35]	72.72
NAT2 acetylator: SA vs. IA		**Difference of survival probabilities**			**Difference of RMST's**		
t_0	Method	Estimate	Confidence interval	Length	Estimate	Confidence interval	Length
60 days	Normal approx.	0.334	[-0.038, 0.706]	0.74	8.65	[-1.65, 18.95]	20.60
	Pseudo-observations	0.329	[0.010, 0.647]	0.64	8.71	[-1.77, 19.20]	20.97
	Empirical likelihood	0.334	[-0.010, 0.696]	0.71	8.29	[-1.99, 19.46]	21.44
90 days	Normal approx.	0.138	[-0.198, 0.474]	0.67	14.49	[-3.04, 32.01]	35.05
	Pseudo-observations	0.140	[-0.149, 0.430]	0.58	14.63	[-2.86, 32.12]	34.98
	Empirical likelihood	0.138	[-0.181, 0.505]	0.69	14.13	[-3.45, 33.62]	37.07
120 days	Normal approx.	0.154	[-0.105, 0.413]	0.52	20.08	[-2.46, 42.63]	45.09
	Pseudo-observations	0.149	[-0.094, 0.392]	0.49	20.04	[-2.44, 42.52]	44.97
	Empirical likelihood	0.154	[-0.105, 0.475]	0.58	19.72	[-2.57, 45.57]	48.14
180 days	Normal approx.	0.154	[-0.105, 0.413]	0.52	29.30	[-5.66, 64.27]	69.93
	Pseudo-observations	0.149	[-0.094, 0.392]	0.49	28.99	[-5.44, 63.43]	68.88
	Empirical likelihood	0.154	[-0.105, 0.475]	0.58	28.94	[-5.51, 70.66]	76.17

2.7 Conclusions

In this chapter, basic concepts of survival analysis were introduced, classical and also novel methods for the two-sample inference were described. The main result of this chapter is given in Section 2.3.2 along with the proof in Section 2.4. The generalized approach of empirical likelihood method for two-sample right censored data problems was established. A frequently advocated advantage of the empirical likelihood method is its range-preserving property, which aligns well for the survival data analysis. While the empirical likelihood method for right censored data in a two-sample case has been studied before (see, e.g., [58], [57]), the approach used enables a more general application. The parameter of interest under consideration is the difference between two functionals of survival distributions, determined by some non-negative

measurable functions. While several distinct examples were provided, the choice of the parameter of interest is not limited to only those.

It was proved that the two-sample empirical likelihood statistic for right censored data follows a scaled chi-squared distribution, and the form of the theoretical scaling constant was derived. Following the setup in one-sample case established in [53], the consistent estimator for the scaling constant was proposed.

Two particular parameters of interest in more detail using simulations and real data applications were studied: the difference of survival probabilities at a fixed point and the difference of restricted mean survival times. Both parameters have recently been recognized as highly significant within the context of survival data analysis (see [20], [1], and [40]).

The method was implemented using the existing R library `EL`, though additional computations were necessary to estimate the scaling constant for constructing confidence intervals.

In the simulation study, the comparison for the coverage accuracy of confidence intervals obtained using the empirical likelihood method with the normal approximation and pseudo-observations approach was carried out. Empirical likelihood-based confidence intervals have greater widths compared to other methods across nearly all studied sample sizes and censoring scenarios. Based on the results, we conclude that the proposed empirical likelihood approach is generally comparable to the other methods under consideration, while it particularly performs substantially better for sample sizes of approximately up to 100 observations each. The empirical likelihood method's performance is affected by the increase of the censoring proportion, but it still retains superior performance for smaller sample sizes.

Two real datasets to illustrate the application of the proposed method for right censored data were used. The `cancer` dataset available in the R library `survival`, contains lung cancer survival times compared across sex-based groups. The other dataset was acquired from a study conducted at the Latvian Biomedical Research and Study Centre. In the latter case, we analyzed time to successful treatment response in tuberculosis patients stratified by subgroups of enzymes responsible for drug metabolism. For all studied examples the graphs for the difference over all time points were constructed. Such illustrations can be utilized if no pre-specified time point of interest is known.

Several ideas for future research regarding the proposed method are following:

(1) to investigate the scaling constant estimator in greater detail and explore alternative estimation techniques;

(2) to implement influence functions in the estimating equations to avoid the need for variance

estimation, as established for the one-sample case in [14];

(3) to extend the adjusted empirical likelihood method for the two-sample setup with right censored data, following the approach described in [55];

(4) to incorporate covariate adjustment within the proposed empirical likelihood method for rigth censored data.

Chapter 3

Inference for location-scale models using empirical likelihood method

In this chapter, the study of two-sample location-scale models, which represent a specialized subset within the broader class of structural relationship models, is presented. The empirical likelihood method for structural relationship models was previously investigated in [49], employing probability-probability and quantile-quantile plots at a fixed point. In this work, the ideas from [49] are revisited and extended regarding the location-scale models.

In Section 3.1, the location-scale model is defined, the hypothesis test for the two-sample location-scale model is described, and two methods focused on assessing the goodness-of-fit of the location-scale model are outlined. The issue of plug-in parameter estimators are addressed, as well as two methods for parameter estimation are discussed.

Section 3.2 introduces the two-sample empirical likelihood method for location-scale models. The main results of this chapter are presented in Section 3.3. It includes a description of the appropriate bootstrap resampling techniques used to generate critical values for practical goodness-of-fit hypothesis testing of the two-sample location-scale model. Additionally, it outlines a bootstrap procedure for constructing simultaneous confidence bands for the parameter of interest over a specified interval. The algorithm for constructing confidence intervals for either the location or scale parameter using the empirical likelihood statistic for a probability-probability plot at a fixed point is described. This approach was inspired from the idea of confidence bands for the general shift function, as developed by Doksum and Sievers [9]. Regarding the empirical likelihood statistic, such approach is novel to our knowledge.

Section 3.4 introduces the semiparametric structural relationship models and demonstrates that the theory encompass the entire class of these models. Furthermore, Lehmann alternative

models are considered, which serve as another illustrative example of structural relationship models.

Finally, Sections 3.5 and 3.6 present the results of a simulation study for the proposed methods and illustrate a real data analysis using proposed methods, respectively.

3.1 Preliminaries

Let X and Y denote two independent random variables with distribution functions F and G, respectively.

Definition 3.1.1. The location-scale model between two distribution functions F and G is defined by the equation

$$F(x) = G\left(\frac{x-\mu}{\sigma}\right), \quad x \in \mathbb{R}, \tag{3.1}$$

where $\mu \in \mathbb{R}$ is the location and $\sigma > 0$ is the scale parameter.

The location-scale model can be also expressed in terms of quantile functions as

$$F^{-1}(t) = \sigma G^{-1}(t) + \mu, \quad t \in (0,1),\ \mu \in \mathbb{R},\ \sigma > 0, \tag{3.2}$$

where F^{-1} and G^{-1} are the inverse (quantile) functions of F and G.

Let us denote the means and standard deviations of X and Y with μ_X, μ_Y, σ_X and σ_Y, respectively. Then the location-scale model parameters are expressed as

$$\sigma = \frac{\sigma_X}{\sigma_Y} \quad \text{and} \quad \mu = \mu_X - \sigma\mu_Y. \tag{3.3}$$

If $\sigma = 1$, the model (3.1) simplifies to the location model $F(x) = G(x-\mu)$; if $\mu = 0$, the model (3.1) simplifies to the scale model $F(x) = G(x/\sigma)$.

Let us introduce the quantile-quantile (Q-Q) and the probability-probability (P-P) plots that are valuable tools for assessing equality of distributions and differences in locations and/or scales.

Definition 3.1.2. A quantile-quantile plot is defined as a graph of a function

$$\{F^{-1}(G(x)) : x \in \mathbb{R}\}. \tag{3.4}$$

Definition 3.1.3. A probability-probability plot is defined as a graph of a function

$$\{F(G^{-1}(t)) : t \in (0,1)\}. \tag{3.5}$$

It can be verified that in case of location-scale model (3.1) Q-Q plot will be constructed as a straight line. If the distributions come from the same family (Q-Q is a straight line), P-P plots can be constructed. The P-P plot curves above or below the 45° diagonal indicating of shift in locations. If P-P plot crosses the diagonal, it indicates the change in scales. The estimates of Q-Q and P-P plots are obtained by simply replacing the theoretical distribution functions in (3.4) and (3.5) by the empirical distribution functions.

3.1.1 Hypothesis testing

Now, suppose two independent random samples, denoted as $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$, are given. For simplicity, let one of the groups represent the treatment group, and the other group be the control group. In practice, the usual approach is to apply Student's t-test in case of normal data or nonparametric Wilcoxon test to compare the locations in case of non-normality. However, according to Doksum and Sievers [9], there are additional questions that need to be considered:

(1) Is the treatment beneficial for all the members of the population?

(2) If the answer to the question (1) is "No", then for which part of the population is the treatment beneficial?

(3) Does a location (shift) model hold, i.e., $\mu \neq 0$?

(4) If the answer to the question (3) is "No", does a location-scale model hold, i.e., $\mu \neq 0$ and $\sigma \neq 1$?

Therefore, instead of testing the general two-sample problem $H_0 : F(x) = G(x)$ for all $x \in \mathbb{R}$, the aim is to test, whether the two-sample location-scale model holds. The problem of testing whether two samples belong to the same location-scale family has not been widely studied in the literature. Doksum and Sievers [9] proposed a test equivalent to checking whether a nonparametric confidence band for the quantile comparison function contains a straight line. However, this test has low power, since nonparametric confidence bands are typically very wide. Another test based on empirical characteristic functions was proposed by Hall *et al.* [12].

In order to verify, whether the location-scale model (3.1) holds between two distributions, the aim is to test the hypothesis

$$H_0: \text{ There exist constants } \mu \in \mathbb{R} \text{ and } \sigma > 0 \text{ such that } X \stackrel{d}{=} \sigma Y + \mu, \tag{3.6}$$

where d denotes the equality of random variables in distribution.

Nonparametric confidence bands for general shift function

Here we describe the construction of nonparametric simultaneous confidence bands for some response function $\Delta(x)$ following the Doksum and Sievers approach described in [9] (see also [8]).

Definition 3.1.4. Assume F and G are continuous. Let $G^{-1}(u) = \inf\{x : G(x) \geq u\}$ be the left inverse of G. Then the general shift function is defined as

$$\Delta(x) = G^{-1}\{F(x)\} - x, \quad x \in \mathbb{R}. \tag{3.7}$$

If the location (shift) model holds, then $\Delta(x) = \mu$. A natural estimator of $\Delta(x)$ is obtained by

$$\widehat{\Delta}(x) = G_m^{-1}\{F_n(x)\} - x, \quad x \in \mathbb{R}, \tag{3.8}$$

where F_n and G_m denote the empirical distribution functions based on the X and Y samples.Let $G_{\Delta,m}(y) = G_m(\Delta(y) + y)$. If $\phi(F_n, G_m)$ is a distribution-free level α test function for H_0 : $F = G$, then

$$\{\Delta(\cdot) : \phi(F_n, G_{\Delta,m}) = 0\} \tag{3.9}$$

is a distribution-free level $(1 - \alpha)$ confidence region for the response function $\Delta(\cdot)$. These regions reduce to a simple bands if the distribution-free test statistic $T(F_n, G_m)$ is considered with the property that the inequality $T(F_n, G_m) \leq K$ is equivalent to

$$h_*\{F_n(x)\} \leq G_m(x) \leq h^*\{F_n(x)\} \tag{3.10}$$

for all x and some typically non-decreasing functions h_* and h^*. Then the confidence bands are derived as follows. Let $G_m^{-1}(u) = \inf\{x : G_m(x) \geq u\}$ and $G_m^{-I} = \sup\{x : G_m(x) \leq u\}$ be the left and right inverses of G_m, and suppose K is chosen so that

$$P(T(F_n, G_m) \leq K) = 1 - \alpha \tag{3.11}$$

under the hypothesis $H_0 : F = G$. Then, simultaneously for all x, the following holds

$$\begin{aligned}
1-\alpha &= P(T(F_n, G_m) \leq K) \\
&= P(T(F_n, G_{\Delta,m}) \leq K) \\
&= P\left(h_* F_n(x)\} \leq G_{\Delta,m}(x) \leq h^*\{F_n(x)\}\right) \\
&= P\left(h_*\{F_n(x)\} \leq G_m(\Delta(x)+x) \leq h^*\{F_n(x)\}\right) \\
&= P\left(G_m^{-1}\{h_*(F_n(x))\} - x \leq \Delta(x) \leq G_m^{-I}\{h^*(F_n(x))\} - x\right).
\end{aligned}$$

Theorem 3.1.1. [9] *If* (3.10) *and* (3.11) *hold, then*

$$\left[G_m^{-1}\{h_*(F_n(x))\} - x,\ G_m^{-I}\{h^*(F_n(x))\} - x\right] \tag{3.12}$$

as $-\infty < x < \infty$ *gives a level* $(1-\alpha)$ *simultaneous, distribution-free confidence band for the shift function* $\Delta(x)$.

Suppose that the test statistic $T(F_n, G_m)$ is in the form of Kolmogorov-Smirnov statistic

$$T(F_n, G_m) = \sqrt{\frac{mn}{m+n}}\sup\left|F_n(x) - G_m(x)\right|, \tag{3.13}$$

and $K_{S,\alpha}$ is chosen from the Kolmogorov-Smirnov tables so that $P(T(F_n, G_m) \leq K_{S,\alpha}) = 1-\alpha$. Then level $(1-\alpha)$ simultaneous, distribution-free confidence bands for $\Delta(x)$ based on Kolmogorov-Smirnov statistic as $-\infty < x < \infty$ is given by

$$\left[G_m^{-1}\left(F_n(x) - K_{S,\alpha}/\sqrt{mn/(m+n)}\right) - x,\ G_m^{-I}\left(F_n(x) + K_{S,\alpha}/\sqrt{mn/(m+n)}\right) - x\right]. \tag{3.14}$$

The bands given by (3.14) is often referred to as S bands in the literature [9].

If the location model holds between two distributions, horizontal line would fit within the simultaneous bands for the general shift function. In turn, if the scale or location-scale model holds, some linear increasing or decreasing line would fit within the bands. Therefore, to test the hypothesis (3.6), one might construct the bands and check if any straight line fits within the bands.

Effect of parameter estimation on conventional function-based two-sample statistics

To test the hypothesis (3.6) the approach suggested in [12] is based on standardization of both random variables. It is, to utilize standardized random variables

$$U := \frac{X - \mu_X}{\sigma_X} \sim F_0 \quad \text{and} \quad V := \frac{Y - \mu_Y}{\sigma_Y} \sim G_0, \tag{3.15}$$

and rewrite the test hypothesis (3.6) as $H_0 : F_0 = G_0$. A natural choice of test statistic for the null hypothesis of this form is of Cramér-von Mises type statistic, which is defined as

$$S_{mn} = \frac{mn}{m+n} \int_{\mathbb{R}} \left(\widehat{F}_0(x) - \widehat{G}_0(x) \right)^2 \omega(x) dx, \tag{3.16}$$

where $\widehat{F}_0$ and $\widehat{G}_0$ denote estimators of F_0 and G_0, and $\omega(x)$ is some weight function.

Let us now discuss the case for random variable X. Let $\tilde{X}_i := (X_i - \bar{X})/S_X$ for $i = 1, \ldots, n$, where $\bar{X}$ denotes the sample mean and S_X denotes the sample standard deviation. Let $\widehat{F}_0$ denote the empirical distribution function (ECDF) based on the observations $\tilde{X}_i$. Then

$$\widehat{F}_0(x) = F_n(\bar{X} + S_X \cdot x) = F(\bar{X} + S_x \cdot x) + n^{-1/2} \Delta_F(\bar{X} + S_X \cdot x),$$

where

$$F_n(x) = n^{-1} \sum_{i=1}^{n} I_{\{X_i \leq x\}}, \quad \text{and} \quad \Delta_F(x) = n^{1/2} \left(F_n(x) - F(x) \right).$$

Assuming without loss of generality that $(\mu_X, \sigma_X) = (0, 1)$, then by Taylor expansion it follows that

$$F(\bar{X} + S_X \cdot x) = F_0(x) + \{\bar{X} + (S_X - 1)x\} f_0(x) + o_p(n^{-1/2})$$

and $\Delta_F(\bar{X} + S_X \cdot x) = \Delta_{F_0}(x) + o_p(1)$. Therefore,

$$\widehat{F}_0(x) = F_0(x) + \{\bar{X} + (S_X - 1)x\} f_0(x) + n^{-1/2} \Delta_{F_0}(x) + o_p(n^{-1/2}),$$

where the unknown density function f_0 appears in this expansion. The same can be derived for random varible Y as well. It is apparent that the distribution of S_{mn} will depend on the unknown density functions f_0 and g_0. One possibility is to employ the bootstrap resampling method. However, any test statistic based on $\widehat{F}_0(x)$ and $\widehat{G}_0(x)$ will encounter problems since the bootstrap cannot estimate the density functions $n^{1/2}-$consistently [12]. Other possibility would be to apply smoothed bootstrap (see e.g. [18]), but in that case the next issue would be

choice of the appropriate smoothing bandwidth.

Example 3.1.1. In this example we consider the two-sample Kolmogorov-Smirnov test. In the one-sample case, effect of the parameter estimation on the limiting distribution can be easily shown using simulations. For testing the normality in a composite hypothesis, Lilliefor's correction is provided to carry out the hypothesis test correctly. Similar situation arises in the asymptotic behaviour of the two-sample Kolmogorov-Smirnov statistic as well, when location and scale parameters are estimated. The two-sample test statistic is

$$\sqrt{\frac{nm}{n+m}}\sup|F_n(x)-G_m(x)| \xrightarrow{d} \sup_{0<t<1}|B(t)|, \tag{3.17}$$

where $B(t)$ denotes the standard Brownian bridge. However, to test for location-scale model by using $H_0 : F_0(x) = G_0(x)$, one would use the statistic

$$\sqrt{\frac{nm}{n+m}}\sup\left|\widehat{F}_0(x)-\widehat{G}_0(x)\right|. \tag{3.18}$$

In panel A of Figure 3.1, it is demonstrated through simulations that the statistic given by equation (3.18) exhibits a left-skewed limiting distribution (in green) when compared to the distribution of the statistic (3.17) based on the Brownian bridge (in blue). The second issue arises from the fact that, with fixed sample sizes n and m, only specific values of the statistic can be observed. As a result, the obtained limiting distribution displays gaps between these values, as illustrated by histograms in panel B of Figure 3.1.

Goodness-of-fit test for a two-sample location-scale model based on empirical characteristic functions

To overcome the issues with conventional test statistics based on empirical distribution or empirical quantile functions, Hall *et al.* [12] propose a new test based on the characteristic functions. Their approach is based on the standardized variables U and V as defined in the equation (3.15).

Definition 3.1.5. [12] The characteristic function of a random variable X is defined as

$$\psi_X(t) = E\{e^{itX}\}, \quad t \in \mathbb{R}, \tag{3.19}$$

with the empirical counterpart for a random sample $X_1, \ldots, X_n$ in the form

$$\widehat{\psi}_X(t) = \frac{1}{n}\sum_{i=1}^{n} e^{itX_i}, \quad t \in \mathbb{R}. \tag{3.20}$$

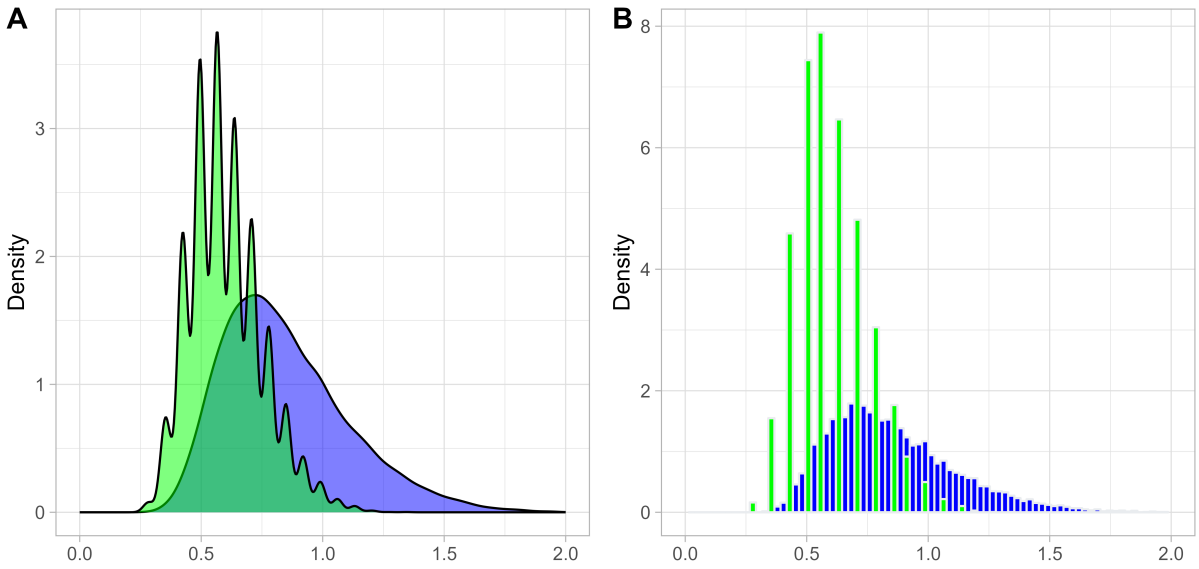


Figure 3.1: The kernel density estimates (panel **A**) and histograms (panel **B**) representing the simulated limiting distribution of the two-sample Kolmogorov-Smirnov statistic under $H_0 : F(x) = G(x)$ with X and Y distributed as $N(2, 1)$ (in blue) and under $H_0 : F_0(x) = G_0(x)$ with $X \sim N(2, 1)$ and $Y \sim N(0, 1)$ (in green). The number of replications is $N = 10000$.

Let the estimators of unknown μ_X, μ_Y, σ_X and σ_Y are obtained by their corresponding sample estimators $\bar{X}$, $\bar{Y}$, S_X and S_Y, respectively. Also let

$$\begin{aligned}\widehat{\psi}_U(t) &= \frac{1}{n}\sum_{i=1}^{n} \exp\left(it\frac{X_i - \bar{X}}{S_X}\right) \\ &= \frac{1}{n}\sum_{i=1}^{n} \cos\left(t\frac{X_i - \bar{X}}{S_X}\right) + \frac{i}{n}\sum_{i=1}^{n} \sin\left(t\frac{X_i - \bar{X}}{S_X}\right).\end{aligned}$$

Similar definition is obtained also for $\widehat{\psi}_V$ using $\bar{Y}$ and S_Y. The test statistic for the hypothesis (3.6) proposed in [12] is

$$L_{mn} = \frac{mn}{m+n}\int_{\mathbb{R}} \left|\widehat{\psi}_U(t) - \widehat{\psi}_V(t)\right|^2 \omega(t)dt, \tag{3.21}$$

where $\omega(t)$ is a specified weight function.

Theorem 3.1.2. [12] *For two independent random samples $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$ under the null hypothesis* (3.6) *the following holds*

$$\sqrt{\frac{mn}{m+n}}\left(\widehat{\psi}_U(t) - \widehat{\psi}_V(t)\right) \xrightarrow{d} \int_{\mathbb{R}} \left(e^{itx} - itx\,\psi_U(t) - \frac{t}{2}(x^2-1)\psi'_U(t)\right) dB(F_0(x)), \tag{3.22}$$

where $B(x)$ is a standard Brownian bridge.

Remark 3.1.1. *It is possible to relax the assumption of two finite moments for a random sample by assuming that one-sample location and scale estimators $\widehat{\mu}_X$ and $\widehat{\sigma}_X$ have the properties*

$$\widehat{\mu}_X - \mu_X = \frac{1}{n}\sum_{i=1}^{n}\omega_1(X_i) + o_p\left(n^{-1/2}\right),$$
$$\widehat{\sigma}_X^2 - \sigma_X^2 = \frac{1}{n}\sum_{i=1}^{n}\omega_2(X_i) + o_p\left(n^{-1/2}\right),$$

where $E\{\omega_1(X)\} = E\{\omega_2(X)\} = 0$ and $\max\{Var(\omega_1(X)), Var(\omega_2(X))\} < \infty$. *For example, robust choices for one-sample location and scale parameters could be utilized such as the median and interquartile range or M-estimators of location and scale.*

In [12], it is argued that the test statistic L_{mn} defined by the equation (3.21) is not asymptotically distribution free. Therefore in practical applications, bootstrap resampling procedure needs to be employed. The standard choice for weight function is considered as $\omega(t) = e^{-bt^2}$, where b denotes the bandwidth in corresponding kernel density estimation. Other possible choices for $\omega(t)$ are given in [12].

When it can be assumed beforehand that the samples come from a symmetric distribution, an alternative statistic is derived. Since $\psi_U(t)$ and $\psi_V(t)$ are real for distributions symmetric about 0, the following estimators for $\psi_U(t)$ and $\psi_V(t)$ can be employed

$$\widehat{c}_U(t) = \frac{1}{n}\sum_{i=1}^{n}\cos\left(t\frac{X_i - \bar{X}}{S_X}\right),$$
$$\widehat{c}_V(t) = \frac{1}{m}\sum_{k=1}^{m}\cos\left(t\frac{Y_k - \bar{Y}}{S_Y}\right).$$

Then the statistic analogous to (3.21) is defined as

$$Q_{mn} = \frac{mn}{m+n}\int_{\mathbb{R}}\left(\widehat{c}_U(t) - \widehat{c}_V(t)\right)^2\omega(t)dt. \tag{3.23}$$

Theorem 3.1.3. [12] *For two independent random samples $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$ with symmetric distributions under the null hypothesis* (3.6) *the following holds*

$$\sqrt{\frac{mn}{m+n}}\left(\widehat{c}_U(t) - \widehat{c}_V(t)\right) \xrightarrow{d} \int_{\mathbb{R}}\left(\cos(tx) - \frac{t}{2}(x^2-1)\psi_U'(t)\right)dB(F_0(x)), \tag{3.24}$$

where $B(x)$ denotes a standard Brownian bridge.

The proofs for Theorem 3.1.2 and 3.1.3 are given in the Appendix of [12].

3.1.2 Methods for location and scale parameter estimation

Previously described approach to test the hypothesis (3.6) was based on the standardized random variables $U \sim F_0$ and $V \sim G_0$. In this thesis, another slightly different approach is proposed. Let $Y^* = \sigma Y + \mu$ with some distribution function G^*. Under the assumption of location-scale model, Y^* and X are equally distributed, i.e.,

$$\begin{aligned} G^*(x) &= P(\sigma Y + \mu \leq x) = P\left(Y \leq \frac{x-\mu}{\sigma}\right) \\ &= G\left(\frac{x-\mu}{\sigma}\right) = F(x), \quad x \in \mathbb{R}. \end{aligned}$$

Therefore, the test hypothesis (3.6) for validation of the two-sample location-scale model is expressed as $H_0 : F(x) = G^*(x)$. In practice, to carry out such hypothesis test, the estimators for μ and σ have to be considered.

In this thesis, two methods for the location and scale parameter estimation are considered and studied in more detail: method of moments (without trimming) and Mallows distance. Potgieter and Lombard also studied the estimation of two-sample location and scale parameters and proposed a nonparametric asymptotic likelihood estimator in [32]. Moreover, both authors considered two other methods as well in [33]. The first approach minimizes a weighted L^2 distance between the empirical characteristic functions of the X and Y data (*WI* method). The second approach is based on a construction of a quadractic form based on the diference of the empirical characteristic functions of samples of X and Y at a pre-selected number of points (*k-L* method).

Method of moments (MOM)

For the estimation of μ and σ a natural choice is to directly replace the theoretical values in (3.3) with their sample estimators leading to the estimators

$$\hat{\sigma} = S_X / S_Y \quad \text{and} \quad \hat{\mu} = \bar{X} - \hat{\sigma}\bar{Y}, \tag{3.25}$$

where $\bar{X}$ and $\bar{Y}$ denote the corresponding sample means and S_X and S_Y denote the corresponding sample standard deviations. This method previously has been studied in [32]. Moreover, the modification of the MOM is obtained using the trimmed versions of sample estimators for X leading to

$$\bar{X}^{\gamma} = \frac{1}{n - 2\lfloor n\gamma \rfloor} \sum_{i=\lfloor n\gamma \rfloor + 1}^{n - \lfloor n\gamma \rfloor} X_i, \quad \text{and} \quad S_X^{2\gamma} = \frac{1}{n - 2\lfloor n\gamma \rfloor} \sum_{i=\lfloor n\gamma \rfloor + 1}^{n - \lfloor n\gamma \rfloor} (X_i - \bar{X}^{\gamma})^2, \tag{3.26}$$

where γ denotes the trimming proportion, and $\lfloor x \rfloor$ denotes the integer part of x. S_X^γ is obtained by taking the square-root from $S_X^{2\gamma}$. Analogous definitions for $\bar{Y}^\gamma$ and S_Y^γ are obtained.

Mallows distance

The Mallows distance was first used for two distribution functions in [24]. Later, Munk and Czado [27] used trimmed Mallows distance to test the equivalence between two distributions. Freitag and Munk [10] studied properties of Mallows distance for structural relationship models, which include the location-scale models as well. The definition of the Mallows distance is based on the quantile function form of a two-sample location-scale model as given in equation (3.2).

Definition 3.1.6. [10] The Mallows distance between two distribution functions F and G is defined as

$$M(F, G) := \int_0^1 \left(F^{-1}(u) - \sigma G^{-1}(u) - \mu\right)^2 \mathrm{d}u. \tag{3.27}$$

Thus, the estimators $\hat{\mu}$ and $\hat{\sigma}$ can be obtained by minimizing the Mallows distance (3.27), in which F^{-1} and G^{-1} are replaced by the corresponding empirical quantile functions F_n^{-1} and G_m^{-1}, respectively.

3.2 Two-sample EL method for location-scale model

In this section the two-sample empirical likelihood method is briefly described in the context of location-scale models. Two EL-based statistics are considered: the difference of distribution functions and the probability-probability plots. Let $X_1, \ldots, X_n$ and $Y_1, \ldots Y_m$ be two independent random samples with distributions F and G, respectively. Let us denote the univariate parameter of interest with Δ and the vector of a location and scale parameters with ν for simplicity, i.e., $\nu = (\mu, \sigma)$. The true parameters are determined by the unbiased estimating equations

$$E\{w_1(X, \Delta_0, \theta_0, \nu_0)\} = 0,$$
$$E\{w_2(Y, \Delta_0, \theta_0, \nu_0)\} = 0,$$

where Δ_0 denotes the true value of Δ, and θ_0 is the true value of the nuisance parameter $\theta \in \mathbb{R}$ associated with one of the distribution functions, and $\nu_0 = (\mu_0, \sigma_0)$ denotes the vector of the true location and scale parameter values.

The two-sample empirical likelihood function is defined as

$$L(\Delta,\theta,\nu) = n^n m^m \sup_{\theta,p,q}\left\{ \prod_{i=1}^{n} p_i \prod_{j=1}^{m} q_j \; : \right.$$
$$p_i > 0,\, i=1,\ldots,n,\, \sum_{i=1}^{n} p_i = 1,\, \sum_{i=1}^{n} p_i w_1(X_i,\Delta,\theta,\nu) = 0,$$
$$\left. q_j > 0,\, j=1,\ldots,m,\, \sum_{j=1}^{m} q_j = 1,\, \sum_{j=1}^{m} q_j w_2(Y_j,\Delta,\theta,\nu) = 0 \right\}, \tag{3.28}$$

where $p = (p_1,\ldots,p_n)$ and $q = (q_1,\ldots,q_m)$ are probability vectors. For the parameters Δ, θ and ν fixed, the supremum in $L(\Delta,\theta,\nu)$ exists and is unique provided that 0 belongs to the interior of the convex hull of $w_1(X_1,\Delta,\theta,\nu)$, …, $w_1(X_n,\Delta,\theta,\nu)$ and $w_2(Y_1,\Delta,\theta,\nu)$, …, $w_2(Y_m,\Delta,\theta,\nu)$. Optimal p and q can be found by the Lagrange multiplier method yielding the minus twice log-empirical likelihood statistic in the form of

$$\begin{aligned} l(\Delta,\theta,\nu) &:= -2\log L(\Delta,\theta,\nu) \\ &= 2\sum_{i=1}^{n} \log(1+\lambda_1(\Delta,\theta) w_1(X_i,\Delta,\theta,\nu)) \\ &+ 2\sum_{j=1}^{m} \log(1+\lambda_2(\Delta,\theta) w_2(Y_j,\Delta,\theta,\nu)), \end{aligned} \tag{3.29}$$

where $\lambda_1(\Delta,\theta)$ and $\lambda_2(\Delta,\theta)$ are Lagrange multipliers determined by the equations

$$\sum_{i=1}^{n} \frac{w_1(X_i,\Delta,\theta,\nu)}{1+\lambda_1(\Delta,\theta) w_1(X_i,\Delta,\theta,\nu)} = 0 \tag{3.30}$$

and

$$\sum_{j=1}^{m} \frac{w_2(Y_j,\Delta,\theta,\nu)}{1+\lambda_2(\Delta,\theta) w_2(Y_j,\Delta,\theta,\nu)} = 0. \tag{3.31}$$

Additionally, the nuisance parameter θ is estimated by minimizing the likelihood statistic $l(\Delta,\theta,\nu)$ for the parameters Δ and ν fixed, i.e.,

$$\hat{\theta}(\Delta,\nu) = \arg\min_{\theta} l(\Delta,\theta,\nu).$$

The optimization with the respect to θ can be done also by taking the derivative of $l(\Delta,\theta,\nu)$ and introducing the third equation besides the equations for both Lagrange multipliers (3.30) and (3.31) (for details, see Qin and Zhao [37]). Note that R package `EL` finds λ_1, λ_2 and θ values simultaneously minimizing the $l(\Delta,\theta,\nu)$ function for fixed Δ value.

Example 3.2.1. Location-scale goodness-of-fit test at a fixed point based on the difference of distribution functions. For the location-scale model let the parameter of interest be $\Delta = G\left(\frac{x-\mu}{\sigma}\right) - F(x)$, $x \in \mathbb{R}$, and denote the true nuisance parameter as $\theta_0 = F(x)$. In this case the estimating equations are

$$w_1(X, \Delta_0, \theta_0, \nu_0) = I_{\{X \leq x\}} - \theta_0, \tag{3.32}$$

$$w_2(Y, \Delta_0, \theta_0, \nu_0) = I_{\{\sigma_0 Y + \mu_0 \leq x\}} - \theta_0 - \Delta_0. \tag{3.33}$$

If $\mu_0 = 0$ and $\sigma_0 = 1$, then $\Delta_0 = G(x) - F(x)$, which is the case considered by Qin and Zhao [37]. Additionally, from Theorem 1.3.1 it holds, that EL statistic $l(\Delta_0, \widehat{\theta}(\Delta_0, \nu_0), \nu_0)$ converges to a χ^2_1 distributed random variable under the estimating equations (3.32) and (3.33).

Example 3.2.2. Location-scale goodness-of-fit test at a fixed point based on P-P plots. Define the parameter of interest Δ as

$$\Delta = F\{\sigma G^{-1}(t) + \mu\},\ t \in (0, 1). \tag{3.34}$$

The estimating equations in this case are

$$w_1(X, \Delta_0, \theta_0, \nu_0) = I_{\{X \leq \theta_0\}} - \Delta_0, \tag{3.35}$$

$$w_2(Y, \Delta_0, \theta_0, \nu_0) = I_{\{\sigma_0 Y + \mu_0 \leq \theta_0\}} - t, \tag{3.36}$$

where $\theta_0 = \sigma_0 G^{-1}(t) + \mu_0$ denotes the true nuisance parameter. If $\mu_0 = 0$ and $\sigma_0 = 1$, then $\Delta_0 = F\{G^{-1}(t)\}$ is the classical probability-probability plot as studied in [4] (see also Section 1.4). The functions (3.35) and (3.36) are not smooth with respect to θ, therefore similarly as in Claeskens *et al.* [4] the smoothed unbiased estimating equations are used

$$\tilde{w}_1(X, \Delta_0, \theta_0, \nu_0) = H_{b_1}(\theta_0 - X) - \Delta_0, \tag{3.37}$$

$$\tilde{w}_2(Y, \Delta_0, \theta_0, \nu_0) = H_{b_2}(\theta_0 - (\sigma_0 Y + \mu_0)) - t, \tag{3.38}$$

where $H_{b_j}(t) = H_j(t/b_j)$ denotes the kernel density estimator, and $b_1 = b_1(n)$, $b_2 = b_2(m)$ are bandwidth sequences converging to zero as $n, m \to \infty$ (see [4] for more detail). If the conditions of Theorem 1.4.1 holds, it follows that the EL statistic $\tilde{l}(\Delta_0, \widehat{\theta}(\Delta_0, \nu_0), \nu_0)$ converges to a χ^2_1 random variable under smoothed estimating equations (3.37) and (3.38).

Hjort *et al.* [16] introduced the plug-in empirical likelihood method demonstrating the ef-

fect of plug-in parameter estimators on the limiting distribution under the null hypothesis (see Section 1.2). In the one-sample case, the plug-in empirical likelihood statistic typically follows a scaled χ^2_1 distribution. For various two-sample problems the plug-in empirical likelihood method was extended in [49]. The theorem presented below follows from [49] for the parameter of interest Δ as described in Examples 3.2.1 and 3.2.2 using the estimator for ν.

Theorem 3.2.1. *Under some regularity conditions (see* [49]*) the scaled two-sample plug-in empirical likelihood statistic converges to a χ^2_1 distribution, i.e.,*

$$rl(\Delta_0, \hat{\theta}(\Delta_0, \hat{\nu}), \hat{\nu}) \xrightarrow{d} \chi^2_1, \tag{3.39}$$

where $\hat{\nu} = (\hat{\mu}, \hat{\sigma})$ and r is some scaling constant.

The proof of this theorem is based on [37] and [16], and is presented in [49]. Using Theorem 3.2.1 $(1-\alpha)100\%$ confidence interval for Δ is obtained as

$$I_{EL} = \{\Delta : -2 \log L(\Delta, \hat{\theta}(\Delta, \hat{\nu}), \hat{\nu}) < c\}, \tag{3.40}$$

where c is $(1-\alpha)$ quantile of the χ^2_1 distribution.

The asymptotic behaviour of the limiting distribution was analysed using simulation study (see Section 3.5.2), figure 3.4 shows the comparison of simulated limiting distributions of the test statistics with theoretical χ^2_1 distribution for Δ as described both in Example 3.2.1 and 3.2.2. The scaling constant r depends on the underlying distribution of samples X and Y, and is complicated to estimate. Instead it is proposed to use the bootstrap resampling method for practical applications.

3.3 Main results

In this section two inferential statistical methods are derived to test the location-scale models based on the previously described El method. The first procedure is derived to test the hypothesis of the location-scale model between two-samples using the parameter estimates. The second approach is a graphical test by constructing the confidence intervals for the location and/or scale parameter.

3.3.1 Procedure for the goodness-of-fit test of location-scale model

To test the hypothesis (3.6) in practice, the following procedure is derived, which allows the use of R built-in functions from the package `EL` to calculate the EL statistic for the difference of the distribution functions (will be denoted by *nsfdiff*) and for the P-P plot (will be denoted by *pp*).

(1) Obtain the estimates of location and scale parameters $\widehat{\mu}$ and $\widehat{\sigma}$;

(2) Test the hypothesis $H_0 : F(x) = G^*(x)$ at a fixed point $x \in \mathbb{R}$ using the parameter estimates and one of the proposed EL statistics: *nsfdiff* or *pp*. The asymptotic behaviour of the statistic is described in Theorem 3.2.1.

(3) Repeat the hypothesis test for all $x \in \mathbb{R}$ and construct the (pointwise) confidence intervals.

(4) Compare the obtained intervals with a reference line indicating the location-scale model between the two samples (in the case of *nsfdiff*, the reference line is $y = 0$, and in the case of *pp*, the reference line is $y = x$).

(5) Draw the conclusion on the (pointwise) presence of the location-scale model.

To carry out the hypothesis test in the step (2) of the derived procedure, the appropriate location-scale bootstrap resampling has to be utilized to obtain the critical value. Slightly different bootstrap procedure was described in [12] to obtain the critical values for the test statistic based on empirical characteristic functions. However in censored data case, the appropriate location-scale bootstrap resampling was described by Subramanian [47]. The procedure for location-scale appropriate bootstrap resampling is given below.

(i) Obtain the estimates of location and scale parameters $\widehat{\mu}$ and $\widehat{\sigma}$;

(ii) Fix one of the two samples, say X. Use the inverse location-scale transformation $h^{-1}(x) = (x - \widehat{\mu})/\widehat{\sigma}$ to transform X, i.e.,

$$X^* = h^{-1}(X) = \frac{X - \widehat{\mu}}{\widehat{\sigma}}. \tag{3.41}$$

Such transformation using estimates $\hat{\mu}$ and $\hat{\sigma}$ ensure that the null hypothesis of location-scale model holds true for X and X^*;

(iii) Draw nonparametric bootstrap samples from X and X^* 10 000 times and calculate the statistic for each pair of samples to obtain the limiting distribution of the statistic.

The location-scale hypothesis test is performed by calculating the empirical likelihood statistic between X and Y and comparing to the bootstrapped critical value $q_{1-\alpha}$, which is the empirical $(1-\alpha)$ quantile of the distribution obtained in the last step (iii) of the bootstrap procedure.

Construction of the simultaneous confidence bands

The procedure for constructing a simultaneous confidence bands over some interval (a, b) using bootstrap resampling method first was introduced in [13], but was used in [51] as well. The procedure without plug-in parameter estimation is implemented in R package `EL` [50]. To carry out the procedure, define the maximum empirical likelihood estimator as

$$\hat{\Delta} = \arg\max_{\Delta} L(\Delta, \hat{\theta}, \hat{\nu}), \tag{3.42}$$

where $\hat{\theta} = \hat{\theta}(\Delta, \hat{\nu})$, and calculate $\hat{\Delta}$ for all $x \in (a, b)$. The bootstrap critical value c^* is chosen such that

$$P\left(-2\log L(\hat{\Delta}, \hat{\theta}, \hat{\nu}) \leq c^* \quad \text{for} \quad a \leq x \leq b\right) = 1-\alpha, \tag{3.43}$$

where L is the empirical likelihood function for bootstrapped samples from X and Y. Next, find maximum statistic over the interval (a, b) and repeat the bootstrapping to find the distribution of maxima, which gives us the $(1-\alpha)$ quantile for construction of the simultaneous confidence bands.

By constructing simultaneous confidence bands, one can carry out the goodness-of-fit of the location-scale model over some interval (a, b). For the difference of the distribution functions (*nsfdiff*) location-scale model holds if the horizontal line $y = 0$ falls within the confidence bands, for the probability-probability plots (*pp*) one needs to check if the diagonal $y = x$ falls within the bands. The graphical approach of the goodness-of-fit testing demonstrates, how exactly both distributions differ and where the location-scale model does not hold in case the hypothesis is rejected.

3.3.2 Confidence regions for location and scale parameters

In this subsection, only the EL statistic for the P-P plots (see example 3.2.2) is considered, but other EL-based statistics can be considered as well. The aim is not to estimate the nuisance parameter ν to make inference on Δ, but to construct a confidence region for ν itself. Similar idea was used by Doksum and Sievers [9] to construct the simultaneous confidence bands for the general shift function using the Kolmogorov-Smirnov statistic for two-sample hypothesis test

(see Section 3.1.1). To author's best knowledge, this approach is new regarding the EL method.

Let us consider the EL statistic for P-P plot at a fixed point t, to test the following hypothesis

$$H_0 : \Delta = \Delta_0 \text{ against } H_1 : \Delta \neq \Delta_0, \tag{3.44}$$

and to construct the $(1-\alpha)100\%$ pointwise or simultaneous confidence interval I_{EL} for Δ. Thus, a distribution-free $(1-\alpha)100\%$ confidence region for the parameter ν is

$$\{\nu : \Delta_0 \in I_{EL}\} \tag{3.45}$$

based on the distribution functions F and G^*. For a fixed value ν_0 for ν, the EL statistic follows a χ^2_1 random variable in case of P-P plots.

Corollary 3.3.1. *Let* $\nu_* = \inf\{\nu : \Delta_0 \in I_{EL}\}$ *and* $\nu^* = \sup\{\nu : \Delta_0 \in I_{EL}\}$ *denoting the lower and upper confidence limits, respectively. The empirical likelihood-based* $(1-\alpha)100\%$ *confidence region for* ν *at a fixed point* t *is* $\{\nu : \Delta_0 \in I_{EL}\} = (\nu_*, \nu^*)$*, where* I_{EL} *is found by* (3.40) *and* Δ_0 *corresponds to the null hypothesis* (3.44)*.*

Figure 3.2 demonstrates the illustration of the method for a simple location model with $\mu_0 = 1$ (and $\sigma_0 = 1$). Panel A demonstrates the EL statistic function for two arbitrary chosen fixed values of μ with the corresponding 95% confidence intervals I_{EL} for Δ (obtained as the intersection with the 95th quantile of the χ^2_1 distribution). Panel B demonstrates the confidence intervals for Δ for all μ values over a partition of some interval around the true parameter value. The 95% confidence interval for μ at $t = 0.5$ is obtained by taking those values of μ for which $0.5 \in I_{EL}$, as $\Delta_0 = 0.5$ for P-P plot.

The procedure is summarized in the Algorithm 1 separately for the location and for the scale parameter. However, it is easily to extend the algorithm for both parameters simultaneously leading to a two-dimensional confidence region. Moreover, the algorithm can be applied for multiple values of $t \in (0, 1)$, for example, taking deciles $t = \{0.1, \ldots, 0.9\}$. Thus, obtained confidence intervals allow to test the location or scale changes between two distributions graphically.

Remark 3.3.1. *The algorithm can be used to construct also simultaneous confidence intervals for* μ *or* σ*. The construction of simultaneous bootstrap bands for* Δ *is similar as considered in Section 3.3.1. However, in this case (when fixed value for the parameter* ν *is used) the R library* `EL` *gives already correct confidence bands.*

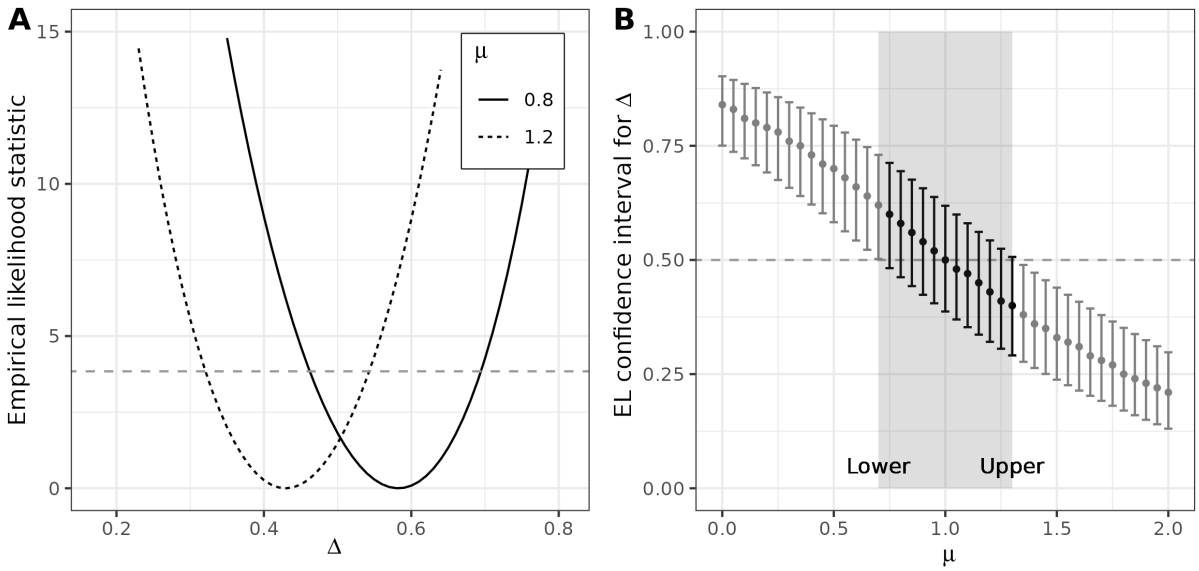
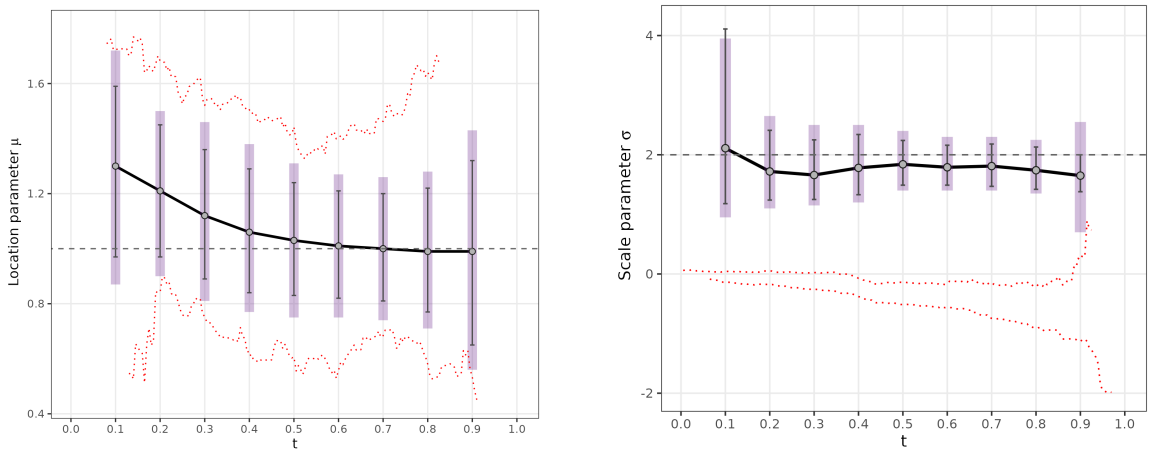


Figure 3.2: A simulated data example from $X \sim N(1,1)$ and $Y \sim N(0,1)$ illustrating the procedure of the confidence interval construction using statistic for probability-probability plot at $t = 0.5$ in case of the location model with the true location parameter $\mu_0 = 1$. Sample size $n = m = 100$, bandwidths calculated using bw.nrd0. EL statistic functions with the 95% confidence intervals of Δ for two arbitrary chosen μ values $\{0.8, 1.2\}$ are demonstrated in the panel **A**; 95% EL confidence intervals of Δ for all $\mu \in [0, 2]$ are demonstrated in the panel **B**, the confidence interval for μ is obtained by the intersection with the line $y = 0.5$.

By using the Algorithm 1 for multiple t values, one can obtain the illustration of the confidence intervals as demonstrated in Figure 3.3. They are based on pointwise intervals for Δ (vertical error bars), as well as on simultaneous bands for Δ (shaded bars). In case (a), the confidence intervals are constructed for the location parameter μ with the true value $\mu_0 = 1$. It can be seen, that the line $y = 1$ lies within the intervals. Dotted lines represent the Doksum and Sievers [9] confidence bands for the general shift function, which show really similar pattern. In case (b), the confidence intervals for the parameter σ are constructed using the exponential distribution. It can be seen that line $y = 2$, representing the true value σ_0, lies within the confidence intervals. In this case Doksum and Sievers' bands perform differently and do not allow to draw conclusions on the scale changes. This occurs because the general shift function is designed to detect a possible (uniform) shift between two distributions, which is not present between two distributions with only the scale differences.

Algorithm 1 The confidence interval for the parameter ν in the location or the scale model

Require: choose α, fix $t \in (0,1)$ and $\Delta_0 \leftarrow t$
  if location model **then**
    $\nu \leftarrow \mu$ and $\sigma \equiv 1$
  else if scale model **then**
    $\nu \leftarrow \sigma$ and $\mu \equiv 0$
  end if
  Fix interval $[a, b]$ around some estimator of $\hat{\nu}$
  for all $\nu \in [a, b]$ **do**
    Obtain transformed sample as $Y^* = \sigma Y + \mu$
    Find $(1-\alpha)100\%$ I_{EL} for Δ on X and Y^*
    Check whether $H_0 : \Delta = \Delta_0$ is not rejected
  end for
  Obtain $(1-\alpha)100\%$ confidence interval for ν as $\{\nu : \Delta_0 \in I_{EL}\} = (\nu_*, \nu^*)$
  if $\nu_* = a$ or $\nu^* = b$ **then**
    Adjust the bounds of $[a, b]$ and repeat the procedure
  end if



(a) $X \sim N(1,1)$ and $Y \sim N(0,1)$, bandwidths selected by *bw.nrd0*.

(b) $X \sim exp(1)$ and $Y \sim exp(2)$, bandwidths selected by *bw.SJ*.

Figure 3.3: Empirical likelihood-based confidence intervals for simulated data, $n = m = 200$. Shaded bars represent the EL-based confidence intervals based on the simultaneous confidence intervals for Δ at the points $t = \{0.1, \ldots, 0.9\}$. Dotted lines illustrate the Doksum's confidence bands for the general shift function.

3.4 Extension to the class of structural relationship models

The EL approach is not limited to location-scale models only; it extends to more general semi-parametric relationship models, referred to as structural relationship models. The concept of structural relationship models was introduced by Freitag and Munk in [10].

Definition 3.4.1. [10] F and G are related by a structural relationship if

$$F^{-1}(t) = \phi_1\{G^{-1}\left(\phi_2(t,h)\right), h\}, \quad t \in [0,1], \tag{3.46}$$

where $\mathcal{H} \subseteq \mathbb{R}^l$, $\phi_1 : \mathbb{R} \times \mathcal{H} \to \mathbb{R}$ and $\phi_2 : [0,1] \times \mathcal{H} \to [0,1]$ are continuous functions with respect to both arguments, and $h \in \mathcal{H}$ is some l-dimensional structural parameter.

Alternatively, the equation (3.46) can be rewritten as

$$F(x) = \phi_2^{-1}\{G\left(\phi_1^{-1}(x,h)\right), h\}, \quad x \in \mathbb{R}, \tag{3.47}$$

where the inverse functions of ϕ_1 and ϕ_2 exist with the respect to their first argument.

Example 3.4.1. The classical two-sample location-scale model is a structural relationship model, if

$$\phi_1(x,h) = \mu + \sigma x \text{ and } \phi_2(t,h) \equiv t,$$

where $h = \nu = (\mu, \sigma)$ denotes a two-dimensional structural parameter. Specifically, for the location model $h = \mu$ and $\phi_1(x,h) = x + h$, whereas for the scale model $h = \sigma$ and $\phi_1(x,h) = hx$.

In the case of structural relationship model (3.46), the parameter of interest Δ for P-P plot is defined as

$$\Delta = F\{\phi_1\left(G^{-1}(\phi_2(t,h)), h\right)\}, \; t \in (0,1), \tag{3.48}$$

which can be seen as a generalization of the P-P plot. The true parameters are determined by the estimating equations

$$E\{w_1(X, \theta_0, \Delta_0, h_0)\} = 0,$$
$$E\{w_2(Y, \theta_0, \Delta_0, h_0)\} = 0$$

with

$$w_1(X, \theta_0, \Delta_0, h_0) = I_{\{X \le \theta_0\}} - \Delta_0, \tag{3.49}$$

$$w_2(Y, \theta_0, \Delta_0, h_0) = I_{\left\{Y \le \phi_1^{-1}(\theta_0, h_0)\right\}} - \phi_2(t, h_0), \tag{3.50}$$

where $\theta_0 = \phi_1\{G^{-1}(\phi_2(t, h_0)), h_0\}$ denotes the true univariate nuisance parameter and h_0 is the true structural relationship parameter. The first function $w_1(X, \theta, \Delta, h)$ is fixed, but $w_2(X, \theta, \Delta, h)$ takes different forms for any particular problem of interest (see Section 3.2 for estimating equations of location-scale models).

Lehmann alternative models

Let us consider the Lehmann alternative model that was introduced in [22] and is also an example of the structural relationship models.

Definition 3.4.2. The Lehmann alternative model between two distribution functions F and G is defined as

$$F(x) = 1 - \{1 - G(x)\}^{\frac{1}{h}}, \quad x \in \mathbb{R},\ h > 0. \tag{3.51}$$

The model (3.51) can be expressed in the terms of quantile functions as

$$F^{-1}(t) = G^{-1}\{1 - (1 - t)^h\}, \quad t \in (0, 1),\ h > 0. \tag{3.52}$$

Lehmann alternative models are commonly employed to study the power of rank tests [43]. They also hold significant relevance in survival analysis, since Lehmann alternative models coincide with proportional hazards models (see Section 2.2 in Chapter 2). To show the latter statement, suppose that model (3.51) is given. Let us first derive both sides with the respect to x and obtain

$$f(x) = \frac{1}{h}(1 - G(x))^{\frac{1}{h} - 1} g(x). \tag{3.53}$$

Denote the corresponding hazard functions by h_1 and h_2, and apply simple calculations to obtain

$$\begin{aligned} h_1(x) &= \frac{f(x)}{1-F(x)} \\ &= \frac{(1-G(x))^{\frac{1}{h}-1}g(x)}{h(1-G(x))^{\frac{1}{h}}} \\ &= \frac{g(x)}{h(1-G(x))} \\ &= \frac{1}{h}h_2(x). \end{aligned}$$

Thus, hazard functions $h_1(x)$ and $h_2(x)$ follows the proportionality assumption under Lehmann alternative model. Moreover, the proportionality constant is $h_2(x)/h_1(x) = h$.

Lehmann alternative model follows as a special case of structural relationships with structural parameter h, if

$$\phi_1(x,h) \equiv x \quad \text{and} \quad \phi_2(t,h) = 1-(1-t)^h.$$

For Lehmann alternative model (3.51), the second estimating equation for the empirical likelihood statistic for Δ defined by the equation (3.48) is

$$w_2(Y,\theta_0,\Delta_0,h_0) = I_{\{Y\le\theta_0\}} - 1 + (1-t)^{h_0}. \tag{3.54}$$

Example 3.4.2. It is known that the proportional hazards model holds between two exponential distributions. Let $X \sim exp(\lambda_1)$ and $Y \sim exp(\lambda_2)$, then with $h = \lambda_2/\lambda_1$ Lehmann alternative model holds

$$\begin{aligned} (1-F(x))^h &= \left(e^{-\lambda_1 x}\right)^h = e^{-\lambda_1 h x} \\ &= e^{-\lambda_2 x} = 1-G(x). \end{aligned}$$

Note that for exponential distribution h coincides also with the scale parametr σ from a scale model.

Remark 3.4.1. *For the Lehmann alternative models there are no such transformation as for the location-scale models that allows to implement the method directly using R library* `EL`*. However, for now a simple scale transformation can be used to do the inference for the structural parameter h in the Lehmann model implicitly. For two exponential distributions the scale parameter σ coincides with the structural parameter h from the Lehmann alternative model. Consider the*

random variables X and Y from two Weibull distributions with the scale parameters σ_1 and σ_2 and a common shape parameter a (see Section 2.1.2 for the definition). In this case $\sigma = \sigma_1/\sigma_2$, but the structural parameter is $h = (\sigma_1/\sigma_2)^a$. As it can be seen, h is a constant, which is a function of the scale parameter σ. By the proposed algorithm we do not get confidence intervals directly for the structural parameter h of the Lehmann alternative model. However, if any constant fits into the confidence intervals for the scale parameter σ, we can not reject the constant scale model and the constant structural parameter h as well.

3.5 Simulation study results

In this section the results for the simulation study approving the theoretical results are presented.

3.5.1 Comparison of the methods for parameter estimation

Here a small simulation study is demonstrated to compare two methods described in Section 3.1.2 for estimation of location parameter μ and scale parameter σ: the method of moments giving the estimators (3.3) and the minimizing of Mallows distance defined by the equation (3.27). Two samples of size $n = m = 100, 250$ from $N(0, 1)$, Student's distributions with 3 and 5 degrees of freedom (t_3 and t_5, respectively) were drawn 10 000 times. The exact setup from [33] was used in order to compare results with two methods (*k-L* and *WI*) based on ECF summarized in Table 4 - Table 6 in [33]. The true location and scale parameters in all cases are chosen as $\mu_0 = 0$ and $\sigma_0 = 1$. The empirical standard errors and biases of the estimators for all scenarios are summarized in Table 3.1, the rows corresponding to the methods k-L and WI were taken from [33] according to the parameters with the best performance. The Mallows distance estimates were obtained using R function `optim` for two-parameter optimization with quasi-Newton method `"BFGS"`.

As can be seen from the simulation results in Table 3.1, both studied methods perform similarly for the normal distribution. For t_5 distribution the Mallows distance gives slightly better estimates of μ and substantially better estimates of σ. In case of t_3 distribution, the Mallows distance performs noticeably better. In comparison with the results from [33], only in some cases their methods give better results than method of moments and the Mallows distance.

Table 3.1: Empirical standard errors (SE) and biases of 10 000 times calculated estimates of location parameter μ and scale parameter σ by the method of moments and the Mallows distance for two samples simulated from given distribution of size $n = m = 100, 250$. The true parameter values are $\mu_0 = 0$ and $\sigma_0 = 1$. For methods k-L and WI results taken from Potgieter and Lombard [33] for comparison.

		$n = m = 100$				$n = m = 250$			
Distribution	Method	$SE(\hat{\mu})$	$bias(\hat{\mu})$	$SE(\hat{\sigma})$	$bias(\hat{\sigma})$	$SE(\hat{\mu})$	$bias(\hat{\mu})$	$SE(\hat{\sigma})$	$bias(\hat{\sigma})$
$N(0,1)$	Method of moments	0.142	-0.003	0.101	0.004	0.089	-0.001	0.064	0.001
	Mallows distance	0.141	0.004	0.108	-0.003	0.089	0.000	0.066	-0.002
	Potgieter & Lombard (*k-L*)	0.183	0.005	0.161	0.136	0.105	0.001	0.097	0.074
	Potgieter & Lombard (*WI*)	0.143	0.003	0.105	0.006	0.090	0.000	0.065	0.002
t_3	Method of moments	0.246	-0.002	0.365	0.049	0.153	-0.001	0.257	0.026
	Mallows distance	0.198	-0.001	0.155	-0.005	0.127	-0.001	0.092	0.005
	Potgieter & Lombard (*k-L*)	0.246	-0.003	0.339	0.200	0.118	-0.001	0.128	0.061
	Potgieter & Lombard (*WI*)	0.175	-0.001	0.148	0.017	0.111	-0.001	0.089	0.008
t_5	Method of moments	0.180	-0.001	0.173	0.015	0.116	0.000	0.112	0.007
	Mallows distance	0.170	0.002	0.131	-0.004	0.108	-0.001	0.081	0.001
	Potgieter & Lombard (*k-L*)	0.242	-0.001	0.312	0.233	0.122	0.000	0.167	0.098
	Potgieter & Lombard (*WI*)	0.166	0.002	0.131	0.011	0.104	0.000	0.082	0.004

3.5.2 Asymptotic behaviour of EL statistic with estimated parameters

The limiting distribution of the scaled plug-in empirical likelihood statistic is a χ^2_1 distribution (see Theorem 3.2.1). To demonstrate that a small simulation study is conducted to examine the 95% critical values corresponding to the empirical 0.95-th quantiles for both statistics given in Example 3.2.1 and 3.2.2. Two samples of size $n = m =$10 000 are generated under the location-scale model from

I: the normal distribution family with $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(0,1)$,

II: the uniform distribution family with $X \sim U[a,b]$ and $Y \sim U[0,1]$,

where different combinations of parameters μ_X, σ_Y, a and b were chosen. Next, 10 000 times the statistic using samples X and Y^* was calculated and the empirical 0.95-th quantile of the obtained limiting distribution was found for $t = 0.3, 0.5, 0.7$. For the difference of the distribution functions statistic was calculated at the point $x = q_t$, where q_t denotes t-th quantile of the theoretical distribution, and for the probability-probability plot at the point t (see Table 3.4 and Figure 3.4 for the results).

The estimates of μ and σ were obtained using only the method of moments due to high calculation time for the optimisation of the Mallows distance in practice. The statistic values were calculated by the function *EL.statistic* from R library EL [50], using the methods *"nsfdiff"* and *"pp"* for the difference of the distribution functions and probability-probability plot at a fixed point, respectively. For probability-probability plot three different smoothing bandwidths

were selected using built-in R functions *bw.nrd0*, *bw.nrd* and *bw.SJ*.

Results in Table 3.4 demonstrates that empirical 0.95-th quantiles are quite similar within the same distribution family with chosen different parameter values. Therefore, the limiting distribution of the statistic with estimated parameters might not be affected by the values of the true location and scale parameters. However, for probability-probability plots we can observe that quantiles differ along with the choice of bandwidth selection method. Empirical quantiles also depend on the point at which the statistic is calculated.

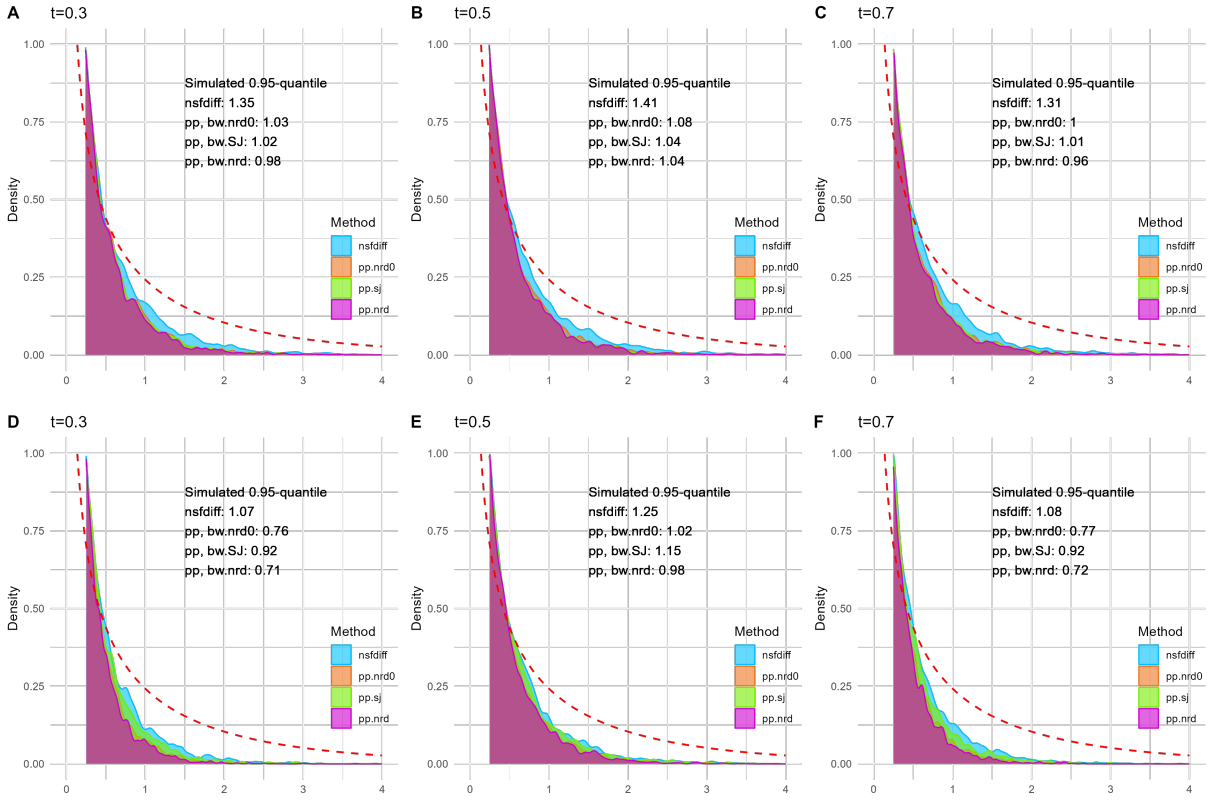


Figure 3.4: Comparison of the χ^2_1 distribution (dashed line) with the simulated limiting distributions for the empirical likelihood statistic for the difference of distribution functions (*nsfdiff*) at a fixed point $x = q_t$ and for the probability-probability plots (*pp*) at a fixed point t, where $t = 0.3, 0.5, 0.7$. Samples of size $n = m =$10 000 were drawn 10 000 times from normal distribution family in the upper panels (A, B, C), and from the uniform distribution family in the lower panels (D, E, F) with the true parameters $\mu_0 = 0$ and $\sigma_0 = 1$. Three methods for the smoothing bandwidth selection were used in case of *pp*: *bw.nrd0*, *bw.SJ*, and *bw.nrd*.

3.5.3 Asymptotic distribution of EL statistic under location-scale transformation

Here a small simulation study is conducted to verify the asymptotic χ^2_1 limiting distribution of the test statistic for the location or scale models after the transformation of the Y sample. The

Table 3.2: Empirical 0.95-th quantiles of the distribution of 10 000 times simulated empirical likelihood statistic for the difference of the distribution functions (*nsfdiff*) calculated at point $x = q_t$ and the probability-probability plot (*pp*) calculated at point t. For normal distributions $Y \sim N(0,1)$, for uniform distributions $Y \sim U[0,1]$, sample size $n = m =$10 000. Three methods were used for bandwidth selection in case of *pp*.

X	$t = 0.3$				$t = 0.5$				$t = 0.7$			
	nsfdiff	*pp*			*nsfdiff*	*pp*			*nsfdiff*	*pp*		
		bw.nrd0	*bw.nrd*	*bw.SJ*		*bw.nrd0*	*bw.nrd*	*bw.SJ*		*bw.nrd0*	*bw.nrd*	*bw.SJ*
$N(0,1)$	1.30	1.01	0.95	1.00	1.45	1.13	1.07	1.08	1.30	1.00	0.96	1.00
$N(0.5,1)$	1.35	0.99	0.94	0.98	1.38	1.09	1.04	1.04	1.35	1.00	0.96	0.98
$N(1,1)$	1.30	1.01	0.96	0.99	1.35	1.09	1.03	1.03	1.34	1.02	0.98	1.00
$N(2,1)$	1.36	1.01	0.96	0.99	1.41	1.12	1.09	1.09	1.25	1.00	0.96	0.99
$N(0,0.5^2)$	1.36	1.04	0.98	1.04	1.38	1.08	1.03	1.04	1.29	1.02	0.96	1.02
$N(0,3.14^2)$	1.32	1.00	0.95	0.99	1.38	1.08	1.04	1.05	1.34	0.97	0.93	0.97
$N(0,7^2)$	1.33	1.06	1.01	1.05	1.34	1.10	1.05	1.06	1.26	1.05	1.00	1.04
$N(1,2^2)$	1.30	1.01	0.96	1.01	1.38	1.09	1.04	1.05	1.33	1.01	0.97	1.00
$N(2,0.5^2)$	1.30	1.04	0.99	1.02	1.41	1.11	1.07	1.07	1.28	1.00	0.95	0.98
$N(5,5^2)$	1.33	1.01	0.96	0.99	1.41	1.06	1.01	1.01	1.27	1.00	0.96	0.98
$U[0,1]$	1.05	0.75	0.70	0.91	1.25	1.05	1.01	1.16	1.04	0.77	0.72	0.92
$U[0.5,1.5]$	1.05	0.75	0.70	0.91	1.25	1.02	0.99	1.14	1.04	0.77	0.72	0.90
$U[-2,-1]$	1.06	0.78	0.72	0.91	1.31	1.05	1.01	1.16	1.01	0.75	0.71	0.91
$U[7,8]$	1.04	0.78	0.74	0.93	1.28	1.02	0.99	1.13	1.05	0.77	0.72	0.92
$U[-1,2]$	1.06	0.76	0.71	0.92	1.31	1.03	1.00	1.13	1.04	0.79	0.74	0.92
$U[0.25,0.75]$	1.08	0.74	0.70	0.90	1.28	1.03	0.99	1.14	1.05	0.75	0.72	0.90
$U[0.4,0.6]$	1.06	0.76	0.71	0.92	1.28	1.05	1.00	1.17	1.04	0.77	0.72	0.89
$U[0,6]$	1.04	0.74	0.70	0.92	1.34	1.04	1.00	1.17	1.09	0.76	0.71	0.91
$U[-3,-1]$	1.05	0.76	0.71	0.92	1.28	1.07	1.03	1.18	1.03	0.78	0.73	0.94
$U[-3.14,3.14]$	1.04	0.76	0.72	0.92	1.25	1.05	1.00	1.18	1.07	0.79	0.74	0.93

procedure is described in Algorithm 1 and is carried out using the empirical likelihood method for P-P plot implemented in the library `EL`. The results are expected to be similar to the results in Table 1 and 2 from [4], where the empirical confidence interval coverage accuracy for Δ between two normal distributions was illustrated under $H_0 : F(x) = G(x)$, $x \in \mathbb{R}$. To demonstrate the scale model, simulations were carried out for two exponential distributions as well.

A number of 10 000 samples of sizes $n = m = 30, 50, 70, 100$ were generated and the coverage accuracy of 95% confidence intervals for the parameter Δ at fixed values $t = 0.2, 0.3, \ldots, 0.8$ was calculated for two following scenarios.

I: $X \sim N(1,1)$, $Y \sim N(0,1)$, where the true location and scale parameters are $\mu = 1$ and $\sigma = 1$, respectively.

II: $X \sim exp(1)$, $Y \sim exp(2)$, where the true location and scale parameters are $\mu = 0$ and $\sigma = 2$, respectively. This model yields proportional hazards with a proportionality parameter $h = 2$ as well.

Three different approaches for bandwidths are used: (1) the default smoothing method (`bw.nrd0`) for both samples (it implements a rule-of-thumb method which works well for the dis-

Table 3.3: Coverage accuracy for $\Delta = F\{\sigma G^{-1}(t) + \mu\}$ at a fixed point $t \in (0, 1)$ based on 10 000 replications for the location model $N(1, 1)$ vs. $N(0, 1)$ and for the scale model (proportional hazards model) $exp(1)$ vs. $exp(2)$. Three different choices of bandwidths were used: (1) *bw.nrd0*, (2) $b_1 = n^{-3/20}$, $b_2 = m^{-3/20}$, (3) *bw.SJ*.

F vs. G	bw	(n, m)	t						
			0.2	0.3	0.4	0.5	0.6	0.7	0.8
$N(1,1)$ vs. $N(0,1)$	(1)	(30, 30)	0.938	0.943	0.947	0.944	0.949	0.945	0.941
		(50, 50)	0.944	0.944	0.952	0.947	0.949	0.950	0.941
		(70, 70)	0.944	0.944	0.949	0.953	0.948	0.945	0.943
		(100, 100)	0.947	0.946	0.949	0.950	0.950	0.946	0.943
	(2)	(30, 30)	0.945	0.947	0.947	0.946	0.952	0.951	0.949
		(50, 50)	0.952	0.950	0.953	0.947	0.950	0.953	0.951
		(70, 70)	0.950	0.949	0.951	0.950	0.950	0.948	0.950
		(100, 100)	0.951	0.950	0.951	0.949	0.950	0.950	0.949
	(3)	(30, 30)	0.929	0.938	0.945	0.942	0.946	0.940	0.931
		(50, 50)	0.933	0.939	0.949	0.946	0.946	0.943	0.932
		(70, 70)	0.936	0.940	0.947	0.950	0.946	0.941	0.937
		(100, 100)	0.939	0.940	0.947	0.948	0.946	0.943	0.937
$exp(1)$ vs. $exp(2)$	(1)	(30, 30)	0.958	0.951	0.946	0.922	0.921	0.927	0.937
		(50, 50)	0.958	0.950	0.942	0.926	0.924	0.928	0.941
		(70, 70)	0.954	0.953	0.941	0.931	0.927	0.933	0.937
		(100, 100)	0.946	0.951	0.940	0.934	0.930	0.933	0.938
	(2)	(30, 30)	0.951	0.944	0.954	0.945	0.948	0.946	0.949
		(50, 50)	0.948	0.948	0.952	0.948	0.950	0.949	0.949
		(70, 70)	0.949	0.950	0.949	0.948	0.949	0.950	0.949
		(100, 100)	0.943	0.950	0.950	0.950	0.951	0.952	0.949
	(3)	(30, 30)	0.966	0.943	0.940	0.931	0.937	0.941	0.948
		(50, 50)	0.963	0.938	0.934	0.938	0.940	0.944	0.953
		(70, 70)	0.959	0.940	0.940	0.940	0.943	0.947	0.949
		(100, 100)	0.949	0.941	0.943	0.943	0.945	0.946	0.947

tributions close to the normal distribution); (2) fixed bandwidths $b_1 = n^{-3/20}$ and $b_2 = m^{-3/20}$ for X and Y, respectively, taken from Claeskens *et al.* paper [4]; (3) Sheather and Jones's method (`bw.SJ`), which selects the bandwidth using the pilot estimation of derivatives [42].

Coverage accuracy for 95% confidence intervals for Δ are summarized in Table 3.3 for both scenarios. For the location model between two normal distributions the respective coverage quickly converge to 95% as n, m increase. Sheather and Jones's method seems to be slightly outperformed by two other methods.

For the scale model between two exponential distributions, `bw.nrd0` does not perform very well, as it was expected. The choice of bandwidths from Claeskens et.al. [4] performs superior compared to both other methods for bandwidth selection.

3.5.4 Coverage accuracy of confidence intervals for location or scale parameter

A small simulation study is also conducted to examine the coverage accuracy of the pointwise confidence intervals constructed for the parameter $\nu = (\mu, \sigma)$ using probability-probability plots as described in the Algorithm 1 of the Section 3.3.2. For the sake of computation time, only three values for t are chosen here, $t = 0.3, 0.5, 0.7$. Samples of equal size $n = m = 30, 50, 70, 100$ are drawn regarding one of two scenarios considered here:

I: the location model from the normal distribution family with the location parameter $\mu = 0, 0.5, 1$ and the scale parameter $\sigma = 1$.

II: the scale model from the exponential distribution family, provided that $\sigma = 0.5, 1, 1.5$.

The same three methods for bandwidth selection as in the previous subsection are used. Due to computational considerations only 1000 replications were used in each case to calculate the coverage accuracy of the confidence intervals at the fixed point t for parameters μ and σ.

Table 3.4: The empirical coverage accuracy of 95% confidence intervals for the parameter μ of the location model ($\sigma = 1$). 1000 samples were drawn from two normal distributions of sizes n and m, respectively. Three different choices of bandwidths were used: (1) *bw.nrd0*, (2) $b_1 = n^{-3/20}$, $b_2 = m^{-3/20}$, (3) *bw.SJ*.

t	(n, m)	$N(0,1)$ vs. $N(0,1)$			$N(0,1)$ vs. $N(0.5,1)$			$N(0,1)$ vs. $N(1,1)$		
		(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
0.3	(30, 30)	0.944	0.955	0.945	0.953	0.954	0.937	0.933	0.949	0.924
	(50, 50)	0.937	0.953	0.937	0.939	0.941	0.939	0.939	0.957	0.941
	(70, 70)	0.936	0.944	0.944	0.943	0.938	0.942	0.943	0.955	0.940
	(100, 100)	0.927	0.943	0.934	0.939	0.936	0.946	0.933	0.941	0.945
0.5	(30, 30)	0.942	0.945	0.934	0.949	0.957	0.951	0.943	0.956	0.960
	(50, 50)	0.941	0.948	0.953	0.947	0.948	0.947	0.943	0.953	0.952
	(70, 70)	0.960	0.953	0.941	0.944	0.938	0.940	0.964	0.952	0.941
	(100, 100)	0.948	0.942	0.944	0.946	0.937	0.946	0.943	0.938	0.942
0.7	(30, 30)	0.942	0.933	0.918	0.939	0.947	0.943	0.944	0.952	0.943
	(50, 50)	0.947	0.943	0.944	0.942	0.951	0.946	0.952	0.945	0.944
	(70, 70)	0.950	0.938	0.955	0.944	0.954	0.934	0.933	0.951	0.945
	(100, 100)	0.944	0.945	0.960	0.942	0.933	0.941	0.945	0.947	0.951

The results for the first simulation scenario are summarized in Table 3.4 and for the second scenario in Table 3.5. As can be seen in the both tables, the coverage accuracy fluctuates slightly around 95% and is not so stable as in Table 3.3 where 10 000 replicates have been made.

Table 3.5: The empirical coverage accuracy of 95% confidence intervals for the parameter σ of the scale (proportional hazards) model ($\mu = 0$). 1000 samples were drawn from two exponential distributions of size n and m, respectively. Three different choices of bandwidths were used: (1) *bw.nrd0*, (2) $b_1 = n^{-3/20}$, $b_2 = m^{-3/20}$, (3) *bw.SJ*.

t	(n, m)	$exp(1)$ vs. $exp(0.7)$			$exp(1)$ vs. $exp(1)$			$exp(1)$ vs. $exp(2)$		
		(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
0.3	(30,30)	0.947	0.943	0.940	0.961	0.946	0.940	0.960	0.949	0.947
	(50,50)	0.955	0.948	0.950	0.942	0.943	0.928	0.950	0.943	0.941
	(70,70)	0.947	0.942	0.936	0.958	0.951	0.937	0.942	0.948	0.931
	(100,100)	0.941	0.943	0.946	0.954	0.942	0.929	0.949	0.937	0.944
0.5	(30,30)	0.935	0.929	0.939	0.913	0.955	0.932	0.919	0.946	0.929
	(50,50)	0.929	0.934	0.932	0.924	0.951	0.942	0.903	0.957	0.948
	(70,70)	0.926	0.942	0.943	0.928	0.952	0.936	0.923	0.941	0.929
	(100,100)	0.924	0.950	0.944	0.926	0.948	0.947	0.938	0.950	0.953
0.7	(30,30)	0.925	0.949	0.946	0.917	0.941	0.941	0.923	0.927	0.934
	(50,50)	0.924	0.942	0.932	0.923	0.950	0.936	0.933	0.959	0.958
	(70,70)	0.904	0.934	0.952	0.918	0.943	0.944	0.928	0.958	0.926
	(100,100)	0.911	0.951	0.948	0.932	0.954	0.953	0.942	0.946	0.946

3.6 Applications of the proposed methods

In this section, three examples illustrating the application of the proposed methods are presented. The first two examples use the inflammatory marker data of type 1 diabetes patients in Latvia. The first example demonstrates the application for validation of the location-scale model between two samples. In the second example, the construction of confidence intervals for the location parameter using an empirical likelihood (EL)-based statistic for the probability-probability (P-P) plot is demonstrated. Finally, the third example provides an illustration of constructing confidence intervals for the scale parameter using EL-based statistic for the P-P plot using mice survival data.

3.6.1 Normalization of data measured on separate plates

Potential markers of inflammation were studied in Latvian patients with type 1 diabetes mellitus within the scope of the project "Intestinal inflammation as a potentially modifiable risk factor for complications in type 1 diabete" (Project No. 1.1.1.2/VIAA/3/19/525). Lipopolysachharide binding protein (LBP) was measured using ELISA kit on three separate plates along with the four other potential markers of inflammation. Numbers of blood samples tested on plate 1, 2 and 3 were 36, 36 and 31, respectively. In this section the use of the location-scale goodness-of-fit testing to carry out the normalization of data measured on separate plates is illustrated.

Variability of LBP measurements across the plates was detected visually (see panel A of figure 3.5), and confirmed using Kruskal-Wallis test for locations ($p < 0.001$). Moreover, significant differences were observed specifically between plate 1 and 3 and between plate 2 and 3 (see panel B and C of figure 3.5).

In order to use the data for further analysis, it was decided to normalize data on plate 3 with respect to the combined data on plate 1 and 2. To perform the normalization we consider the following steps: combine observations from plates 1 and 2 in one sample as X and take plate 3 as Y; test, whether the two-sample location-scale model holds between X and Y; if the previous hypothesis is not rejected, transform Y by location-scale transformation using parameter estimates to normalize the data on plate 3 with the respect to combined data on plate 1 and 2.

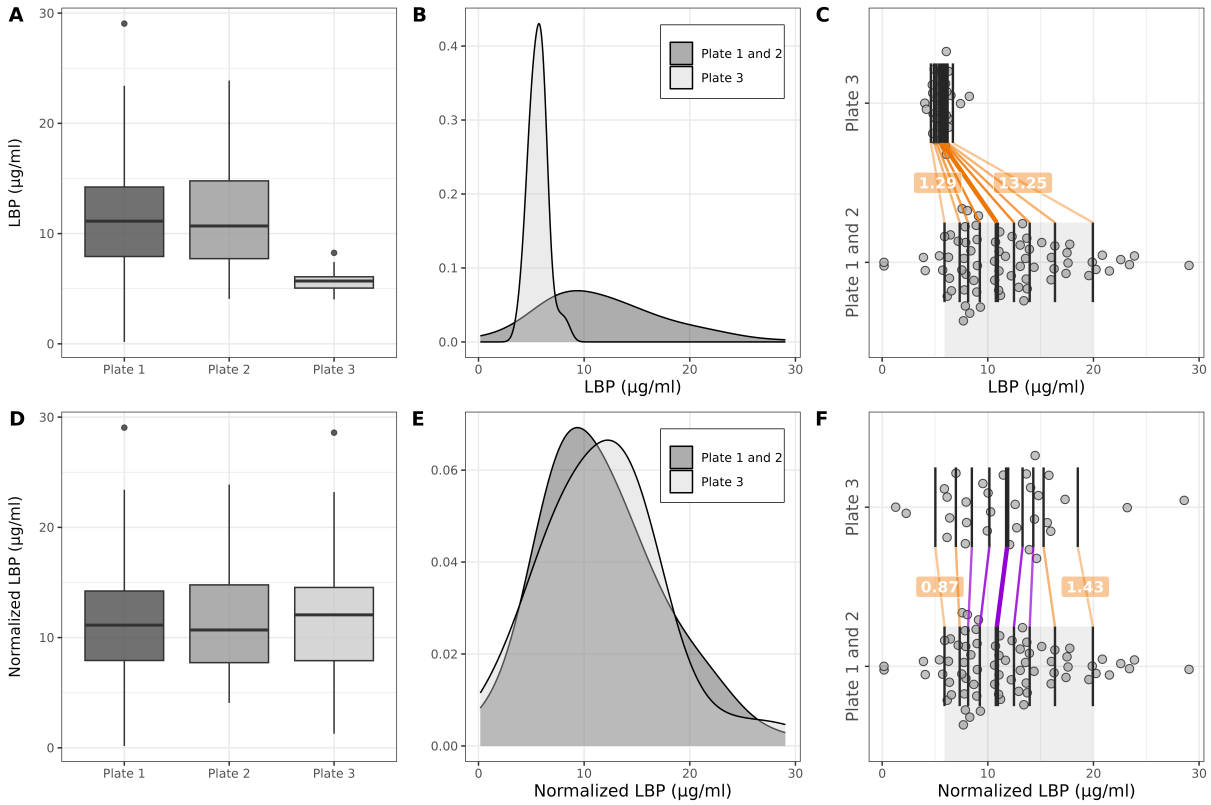


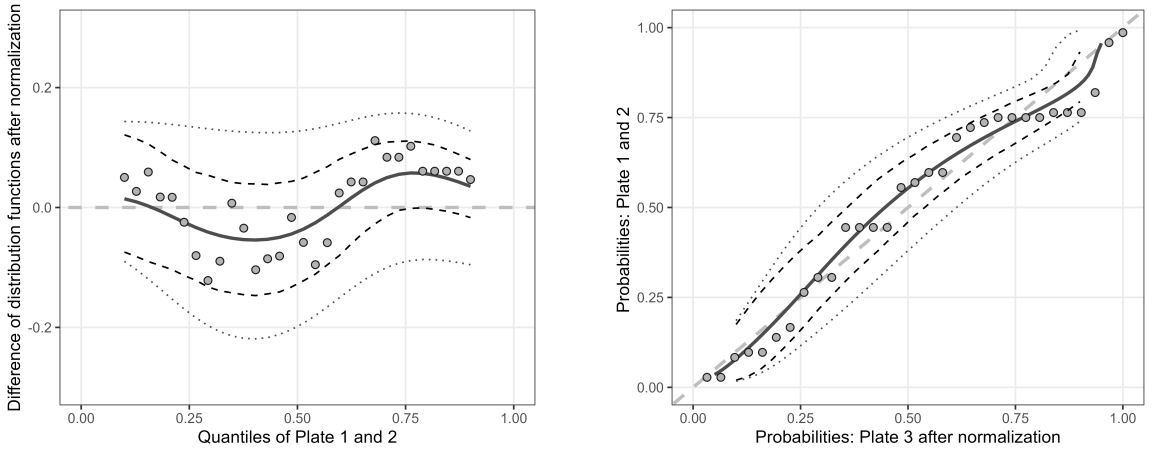
Figure 3.5: Boxplots by plates, density plots and one-dimensional scatterplots for LBP data before normalization (panels **A**, **B** and **C**) and after normalization (panels **D**, **E** and **F**) using the location-scale transformation with the parameter estimators obtained by the method of moments.

The location and scale parameter estimates with 95% bootstrap percentile confidence intervals using sample estimators were $\hat{\mu} = -24.88\,(-43.4, -15.3)$ and $\hat{\sigma} = 6.48\,(4.68, 9.97)$, and using Mallows distance were $\hat{\mu} = -27.25\,(-40.2, -20.2)$ and $\hat{\sigma} = 6.93\,(5.39, 9.48)$. According to the test based on empirical characteristic functions by Hall *et al.* [12] the location-scale model between distributions of both samples is not rejected (test statistic was 0.34, 500 times bootstrapped critical value was 2.77, and $p = 0.757$, see Section 3.1.1 for the description of the

test).

For further analysis, only the method of moments with the sample estimators is utilized to reduce the computation time for bootstrap resampling. The empirical likelihood statistic is calculated for the difference of distribution functions (smoothed version *nsfdiff* is used for better visualization) and for the probability-probability plot to construct pointwise confidence intervals and simultaneous confidence bands (see figure 3.6). The method `bw.SJ` was used to select the smoothing bandwidth. If the horizontal line $y = 0$ falls within the simultaneous confidence bands in case of *nsfdiff* or if the diagonal $y = x$ falls within the bands in case of *pp* the location-scale model between both distributions can not be rejected. As can be seen in Figure 3.6, the hypothesis of the location-scale model between both samples X and Y is not rejected. Additionally, the largest deviation from the null hypothesis is observed at the right tail of the distributions as suggested by plots. The pointwise interval bounds cross the reference lines approximately at $t = 0.9$.

Finally, as the hypothesis of location-scale model is not rejected, normalized data on plate 3 is obtained by applying the location-scale transformation with the estimates of the parameters. Observations across the plates are visually homogeneous after the normalization (see panel D, E and F in figure 3.5).



(a) Smoothed plot for the difference of the distribution functions (*nsfdiff*)

(b) Smoothed plot for probability-probability plot (*pp*)

Figure 3.6: Plots for goodness-of-fit testing of location-scale model in LBP data example. The smoothing bandwidth was selected by the method of Sheater and Jones (*bw.SJ*).

3.6.2 Confidence intervals for a location parameter in diabetes patient data

Here, once again the diabetes patient study of inflammatory markers is considered. The construction of the confidence intervals for location parameter μ is demonstrated. Previous research findings have indicated that the ratio of bacterial lipopolysaccharides (LPS, EU/ml) to high-density lipoproteins (HDL, mmol/l) may serve as a potential indicator of chronic inflammation in Type 1 Diabetes (T1D) patients. Therefore, the LPS/HDL ratio within patient groups, stratified based on the presence of metabolic syndrome (MS) is investigated.

The distributions of both groups are skewed to the right (see panel A in Figure 3.7), and also the normality assumption is violated (Shapiro-Wilk test $p < 0.001$). The medians are significantly different between both groups of patients according to the Wilcoxon test ($p = 0.01$). The P-P plot for the probabilities of patients with MS versus patients without MS is constructed (panel B in Figure 3.7) using the smoothed EL method (bandwidths selected by `bw.SJ`). The P-P plot lies strictly above the diagonal, which suggests the uniform shift in the locations between both study groups.

The confidence intervals for location parameter μ between both study groups at nine values of t using EL-based method proposed in Section 3.3.2 is demonstrated in panel C of Figure 3.7. The hypothesis of no shift between both distributions is rejected at every fixed value of t. Moreover, the shift in both distributions is approximately equal at the first six values of t, but increases linearly at the last three values of t. The latter observation indicates a significant increase in scale or variance at the right tail of the distribution of patients with MS compared to the patients without MS.

3.6.3 Confidence intervals for a scale parameter in mice survival data

The data on laboratory male mice survival after exposure of radiation was analysed in and acquired from [17]. Two groups of RFM strain male mice at the age 5-6 weeks received a dose of 300 rads of radiation. The group of 99 mice was kept in a conventional laboratory environment (control), and the other group of 82 mice was kept in germ-free environment. Survival time (in years) for all subjects were observed, there were no presence of the censoring.

The survival curves (panel A in Figure 3.8) are fairly similar up to about 1 year, but after that we see a higher probabilities of survival in mice kept in the germ-free environment. A similar trend can be observed from the P-P plot (panel B in Figure 3.8), where a diagonal lies within the point-wise confidence bands approximately in the interval $(0, 0.4)$. However, on the rest part of

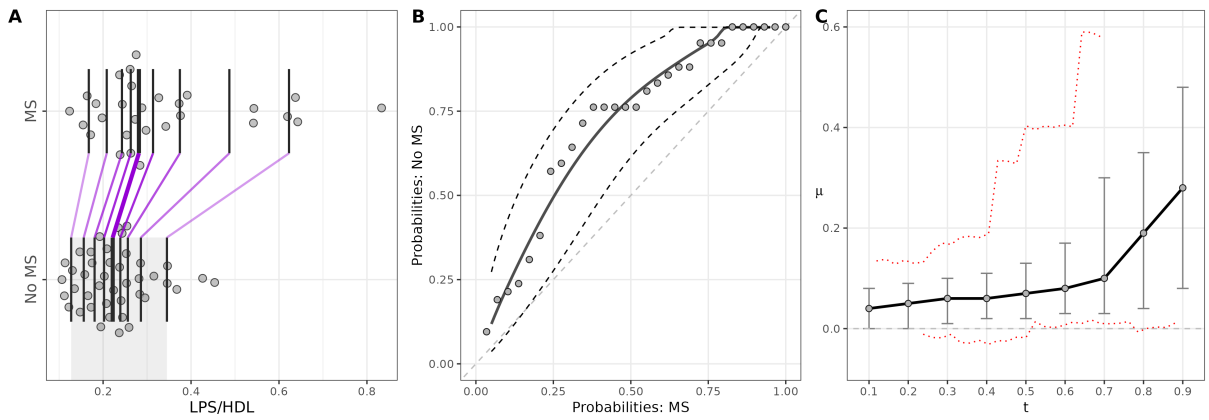


Figure 3.7: Illustrations for LPS and HDL ratio (LPS/HDL) between the groups of T1D patients stratified by the presence of metabolic syndrome (MS). **A**–one-dimensional scatter plots of both distributions with linked deciles; **B**–P-P plot of patients with MS versus patients without MS with 95% pointwise EL-based confidence intervals; **C**–95% EL-based confidence intervals for location parameter μ between groups of patients calculated at nine values of t, dotted lines representing the Doksum's confidence bands. Smoothing bandwiths in **B** and **C** were chosen by the method of Sheather and Jones (*bw.SJ*).

the interval $(0, 1)$ P-P plot lies strictly above the diagonal. Such trend in P-P plot, when it is not uniformly strictly above or below the diagonal, indicates of the presence of scale diferences.

In panel C of Figure 3.8 the EL-based confidence intervals for a scale parameter σ between both study groups of mice are demonstrated. The confidence intervals at $t = 0.1, \ldots, 0.4$ include the value 1, so we do not reject the hypothesis that survival times are the same at respective t values. However, at $t = 0.5, \ldots, 0.9$ the intervals become substantially narrower indicating the different scales of distributions. In addition, since the confidence intervals include a horizontal straight line at about $y = 0.8$, we draw a conclusion that the scale of survival times in a germ-free group is uniformly larger compared to survival times in a conventional group. Moreover, we do not reject a constant scale model, thus the Lehman alternative model might hold with some constant structural parameter h as well.

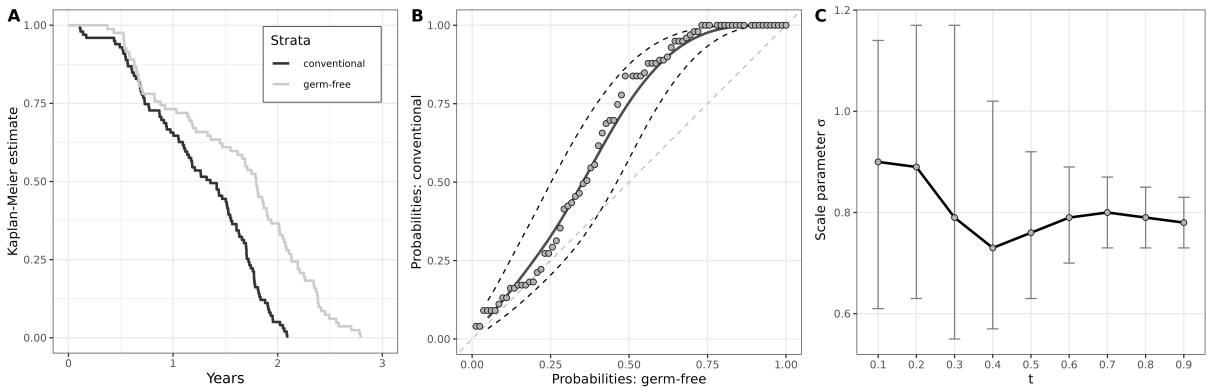


Figure 3.8: Illustrations for conventional versus germ-free environment mice survival. **A**–Kaplan-Meier estimates of survival curves for both groups of mice; **B**–P-P plot of germ-free versus conventional mice survival times with 95% pointwise EL-based confidence intervals; **C**–95% EL-based confidence intervals for scale parameter σ between both study groups calculated at nine values of t. Smoothing bandwidths in **B** and **C** were chosen by the method of Sheather and Jones (*bw.SJ*).

3.7 Conclusions

In this chapter the goodness-of-fit procedure for the two-sample location-scale models is established using the plug-in empirical likelihood method for the difference of the distribution functions and for the probability-probability plots at a fixed point.

Two methods for the plug-in estimators of the location and scale parameters were studied: method of moments and minimizing the Mallows distance. Recently two new methods based on empirical characteristic functions were introduced in [33]. Previously, the method for parameter estimation using asymptotic likelihood estimator was studied in [32]. The comparison of four methods are provided in Table 3.1. For the future work, two additional methods can be proposed by using more robust estimators in the method of moments and by replacing the empirical distribution functions in the Mallows distance with the smoothed versions. However, in the latter case the choice of the smoothing parameter needs to be considered.

The limiting distribution of the empirical likelihood statistic is affected by the parameter estimation procedure (see Table 3.4). Simulation study shows that scaling constant depends on the underlying distribution, but might be the same for the different true parameter values. The method for the difference of the distribution functions (*nsfdiff*) does not depend on the choice of the smoothing parameter, which is an advantage over the probability-probability plot method (*pp*).

A graphical illustration of the location-scale model testing is proposed using pointwise and simultaneous confidence bands. This is convenient for a visual assessment of the differences between two samples to examine at which points exactly and what kind of the discrepancies are

observed. Such a graphical tool can be used in addition to other testing procedures, for example, with the test based on empirical characteristic functions developed in [12].

The proposed method for location-scale goodness-of-fit testing also can be used as a pretest for the validation of the location-scale model between two distributions. If the location-scale model is not rejected, a test for simultaneous detection of the location and scale differences such as Lepage, Cucconi, Podgor-Gastwirth or Neuhauser (see [25] for the review) can be applied. The application of the procedure to test the location-scale model for the laboratory data observed on separate plates is demonstrated. If the variability between two plates is observed and the location-scale model is not rejected, normalization of the data can be carried out using the location-scale transformation. In practice, more than two separate plates are often used, therefore the extension for the normalization across multiple plates can be studied in the future.

A novel method based on the EL statistic was developed for the constructing of the confidence regions for the location and scale parameter in two-sample location-scale models. The extension to the more general class of structural relationship models including Lehmann alternative models was shown. The algorithm for the practical implementation of the method was established and is suitable for the constructing intervals for location and scale parameters separately. However, algorithm can be easily extended for the constructing confidence region for structural parameter in the two-sample location-scale model. For the moment, inference for the Lehmann alternative model is to be done through the scale model, which typically holds between two distributions with proportional hazards (it is true for exponential, Weibull and Pareto distributions among others).

In the case of the two-sample location model obtained confidence intervals are comparable to Doksum's bands for the general shift function. However, for the scale transformation the completely different results are obtained as compared to the Doksum's bands. This occurs due to the fact that Doksum's bands can only indicate of different scales but not model the scale parameter itself.

The results indicate only the first steps into this problem. First of all, it is possible to use various other test statistics, which do not necessarily include the smoothing parameter (EL test statistic based on the difference of two distribution functions or quantile-quantile plots, for example). Next, much more extensive simulation study needs to be considered to compare the proposed method with some other approaches and also to analyse the bandwidth's behavior in more detail.

In conclusion, the proposed method for constructing the confidence regions (intervals) has several important advantages. First, it is a graphical method, which demonstrates where and how

two distributions differ between each other. Second, it allows to make inference on the equality of two distributions and on the presence of structural relationship model simultaneously. Finally, the proposed method is applicable for different practical problems and is implemented using an existing R library `EL`.

Conclusions

Both main aspects of the aim of this thesis are achieved. It is, the empirical likelihood (EL) for the two-sample case with right censored data is established and the inferential methods based on EL method for two-sample location-scale model were developed and analyzed.

The established two-sample EL method for right-censored data is applicable to a class of differences between two functionals of survival distributions, thus retaining some extent of generality. The thesis provides several examples of applications, with two of them studied in greater detail through simulations and real data problems. The theorem demonstrating that the limiting distribution of the test statistic is a scaled chi-squared random variable has been proven. An estimator for the scaling constant has been proposed, following the ideas presented by Wang and Jing [53]. There are several possible extensions of the research concerning EL for right-censored data. For instance, exploring the adjusted EL method introduced by Zheng *et al.* [55] or incorporating influence functions into the estimating equations (as discussed in He *et al.* [14]) to address the variance estimation problem. For other ideas see Section 2.7.

In the context of inferential methods for location-scale models, two main objectives were studied. Regarding the first objective, a procedure for conducting the goodness-of-fit test for a two-sample location-scale model was developed, employing plug-in estimators for the location and scale parameters. This procedure recommends the use of an appropriate bootstrap resampling method for location-scale models. For the second objective, a new graphical method was established to construct confidence regions for the location and scale parameters. This method relies on the inversion of the EL statistic for P-P plots. Although using the statistic for P-P plots is the most convenient approach, other EL-based statistics can also be employed. However, one drawback of P-P plots is the need to select an appropriate smoothing bandwidth parameter.

In Section 3.6, two applications are demonstrated using diabetes patient data. Firstly, a goodness-of-fit test was conducted to validate the presence of a two-sample location-scale model in the data. Subsequently, a data normalization procedure for measurements on separate plates was performed using the location-scale transformation. The second application involved con-

structing confidence intervals for the location parameter between groups of diabetes patients with and without metabolic syndrome. These intervals enable more detailed and comprehensive conclusions regarding observed differences in the data.

All the theoretical results in this thesis are complemented by a simulation study encompassing several scenarios in each case. A significant advantage of the studied methods is that their practical implementation relies on an existing R library `EL`. Consequently, a future goal is to implement these methods in a dedicated R library for greater accessibility and ease of use among practitioners.

Author's publications

(1) L. Pahirko, J. Valeinis, J. Gredzens, M. Krumina. Validation of Two-sample Location-scale Model Using Empirical Likelihood Based Statistics. *In 5th International Conference on Statistics: Theory and Applications, ICSTA 2023, August 3–5, 2023, London, United Kingdom.* Paper ID 168, Avestia Publishing, 10 pages. `https://doi.org/10.11159/icsta23.168` (to be indexed in Scopus).

(2) L. Pahirko, J. Valeinis. Empirical Likelihood-Based Confidence Regions for the Structural Relationship Parameter. *In 6th International Conference on Mathematics and Statistics (ICoMS 2023), July 14–16, 2023, Leipzig, Germany.* ACM, New York, NY, USA, p. 38-47. `https://dl.acm.org/doi/abs/10.1145/3613347.3613354`.

(3) L. Pahirko, J. Valeinis. Two-sample empirical likelihood method for right censored data. Submitted for *International Journal of Biostatistics*, ID: DGIJB.2023.0132, 2023.

(4) A. Fedulovs, L. Pahirko, K. Jekabsons, L. Kunrade, J. Valeinis, U. Riekstina, V. Pirags, J. Sokolovska. Association of endotoxaemia with low grade inflammation, metabolic syndrome and distinct response to lipopolysaccharide in type 1 diabetes. *Biomedicines*, 11(**12**):3269, 2023. `https://www.mdpi.com/2227-9059/11/12/3269`.

List of ancillary publications

Listed below are the ancillary publications that were prepared during the doctoral studies. These publications were developed in collaboration with the Personalized Laboratory of Medicine (Faculty of Medicine, University of Latvia), whose primary focus is researching risk factors for diabetes complications. This collaboration allowed the author to immerse herself in the field of diabetes research and identify areas where statistical analysis of diabetes data could be improved, thus contributing to the development of new and necessary methods.

(5) Z. Svikle, **L. Pahirko**, L. Zariņa, K. Baumane, D. Kardonaite, L. Radzeviciene, L. Daugintyte-Petrusiene, V.J. Balciuniene, R. Verkauskiene, A. Tiščuka, V. Rovite, N. Sjakste, J. Sokolovska. Telomere lengths and serum proteasome concentrations in patients with Type 1 diabetes and different severities of diabetic retinopathy in Latvia and Lithuania. *Journal of Clinical Medicine*, 11(**10**), p.2768, 2022. `http://www.scopus.com/inward/record.url?eid=2-s2.0-85129827483&partnerID=MN8TOARS`.

(6) I. Salna, E. Salna, **L. Pahirko**, S. Skrebinska, R. Krikova, I. Folkmane, V. Pīrāgs, J. Sokolovska. Achievement of treatment targets predicts progression of vascular complications in Type 1 diabetes. *Journal of Diabetes and its Complications*, 35(**12**), p.108072, 2021. `http://www.scopus.com/inward/record.url?eid=2-s2.0-85116739936&partnerID=MN8TOARS`.

(7) E. Rostoka, I. Salna, A. Dekante, **L. Pahirko**, V. Borisovs, L. Celma, J. Valeinis, N. Sjakste, J. Sokolovska. DNA damage in leukocytes and serum nitrite concentration are negatively associated in Type 1 diabetes. *Mutagenesis*, 36(**3**), p.213-222, 2021. `http://www.scopus.com/inward/record.url?eid=2-s2.0-85110763761&partnerID=MN8TOARS`.

(8) J. Sokolovska, J. Stefanovics, G. Gersone, **L. Pahirko**, J. Valeinis, S. Kalva-Vaivode, V. Rovite, L. Blumfelds, V. Pirags, P. Tretjakovs. Angiopoietin 2 and Neuropeptide Y are associated with diabetic kidney disease in Type 1 diabetes mellitus. *Experimental*

and Clinical Endocrinology & Diabetes, 128(**10**), p.654-662, 2020. https://pubmed.ncbi.nlm.nih.gov/31958847/.

(9) J. Sokolovska, K. Ostrovska, **L. Pahirko**, G. Varblane, K. Krilatiha, A. Cirulnieks, I. Folkmane, V. Pirags, J. Valeinis, A. Klavina, L. Selavo. Impact of interval walking training managed through smart mobile devices on albuminuria and leptin/adiponectin ratio in patients with Type 2 diabetes. *Physiological reports*, 8(**13**), e14506, 2020. https://physoc.onlinelibrary.wiley.com/doi/full/10.14814/phy2.14506.

(10) J. Sokolovska, A. Dekante, L. Baumane, **L. Pahirko**, J. Valeinis, K. Dislere, V. Rovite, V. Pirags, N. Sjakste. Nitric oxide metabolism is impaired by Type 1 diabetes and diabetic nephropathy. *Biomedical Reports*, 12(**5**), p.251-258, 2020. https://www.spandidos-publications.com/10.3892/br.2020.1288#.

List of conferences

C1 **L. Pahirko**, J. Valeinis. Two-Sample Problems of Survival Data. *19th International Conference Mathematical Modelling and Analysis (MMA 2014)*. Druskininkai, Lithuania, May 28 - 29, 2014.

C2 **L. Pahirko**, J. Valeinis. Validation of the two-sample location-scale model using confidence bands. *25th European Meeting of Statisticians (EMS 2015)*. Amsterdam, The Netherlands, July 6 - 10, 2015.

C3 **L. Pahirko**. Asymptotic variance estimation for survival data. *12th Conference of Latvian Mathematical Society*. Ventspils, Latvia, April 13 - 14, 2018.

C4 **L. Pahirko**. Empirical likelihood method for two-sample censored data. *27th Nordic Conference in Mathematical Statistics (NORDSTAT 2018).* Tartu, Estonia, June 26 - 29, 2018.

C5 **L. Pahirko**. Lehmann alternative model. *79th International Scientific Conference of the University of Latvia.* Riga, Latvia (online), February 18, 2021.

C6 Z. Svikle, A. Tiščuka, L. Voitoviča, **L. Pahirko**, K. Baumane, D. Kardonaite, M. Kazokaite, L. Radzeviciene, L. Daugintyte-Petrusiene, J. Balciuniene, R. Verskauskiene, J. Sokolovska. Telomere length and serum proteasome concentration in patients with type 1 diabetes and different severity of diabetic retinopathy in Latvia and Lithuania. *International Scientific Conference on Medicine, 79th International Scientific Conference of the University of Latvia.* Riga, Latvia (online), April 23-24, 2021.

C7 **L. Pahirko**. Shift function and empirical process approach in location-scale hypothesis testing. *80th International Scientific Conference of the University of Latvia.* Riga, Latvia, February 17, 2022.

C8 **L. Pahirko**. Comparison of Non-Parametric Simultaneous Tests for Location and Scale

With Application to Latvian Diabetes Patient Data. *31st International Biometric Conference (IBC 2022).* Riga, Latvia, July 10 - 15, 2022.

C9 **L. Pahirko**. Restricted mean survival time for two samples. *81st International Scientific Conference of the University of Latvia.* Riga, Latvia, March 30, 2023.

C10 **L. Pahirko**, J. Valeinis. Validation of the Two-Sample Location-Scale Model. *26th International Conference Mathematical Modelling and Analysis (MMA 2023)* Jurmala, Latvia, May 30 – June 2, 2023.

C11 J. Valeinis, **L. Pahirko**, J. Sokolovska. Inference for Structural Relationship Models. *29th Nordic Conference in Mathematical Statistics (NORDSTAT 2023).* Gothenburg, Sweden, June 19 - 22, 2023.

C12 **L. Pahirko**, J. Valeinis. Empirical Likelihood-Based Confidence Regions for the Structural Relationship Parameter. *6th International Conference on Mathematics and Statistics (ICoMS 2023).* Leipzig, Germany, July 14 - 16, 2023.

C13 **L. Pahirko**, J. Valeinis, J. Gredzens, M. Krūmiņa. Validation of Two-sample Location-scale Model Using Empirical Likelihood Based Statistics. *5th International Conference on Statistics: Theory and Applications (ICSTA 2023).* London, United Kingdom (virtual participation), August 3 - 5, 2023.

Bibliography

[1] F. Ambrogi, S. Iacobelli, and P.K. Andersen. Analyzing differences between restricted mean survival time curves using pseudo-values. *BMC Medical Research Methodology*, 22(1):1–12, 2022.

[2] P.K Andersen, J.P. Klein, and S. Rosthøj. Generalised linear models for correlated pseudo-observations, with applications to multi-state models. *Biometrika*, 90(1):15–27, 2003.

[3] S.X. Chen and P. Hall. Smoothed empirical likelihood confidence intervals for quantiles. *The Annals of Statistics*, 21(3):1166–1181, 1993. ISSN: 0090-5364.

[4] G. Claeskens, B. Jing, L. Peng, and W. Zhou. Empirical likelihood confidence regions for comparison distributions and roc curves. *Canadian Journal of Statistics*, 31(2):173–190, 2003.

[5] D.R. Cox. Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 34(2):187–202, 1972.

[6] M. Delesa-Vēliņa. *Empirical Likelihood Method for a Location Parameter Using Some Robust Estimators*. PhD thesis, Latvijas Universitāte, 2022.

[7] M. Delesa-Vēliņa, J. Valeinis, and G. Luta. Comparing two independent populations using a test based on empirical likelihood and trimmed means. *Lithuanian Mathematical Journal*, 61:199–216, 2021.

[8] K. Doksum. Empirical probability plots and statistical inference for nonlinear models in the two-sample case. *The Annals of Statistics*, 2(2):267–277, 1974.

[9] K.A. Doksum and G.L. Sievers. Plotting with confidence: graphical comparisons of two populations. *Biometrika*, 63(3):421–434, 1976.

[10] G. Freitag and A. Munk. On hadamard differentiability in k-sample semiparametric models–with applications to the assessment of structural relationships. *Journal of Multivariate Analysis*, 94(1):123–158, 2005.

[11] P. Hall and B. La Scala. Methodology and algorithms of empirical likelihood. *International Statistical Review/Revue Internationale de Statistique*, 58(2):109–127, 1990.

[12] P. Hall, F. Lombard, and C.J. Potgieter. A new approach to function-based hypothesis testing in location-scale families. *Technometrics*, 55(2):215–223, 2013.

[13] P. Hall and A.B. Owen. Empirical likelihood confidence bands in density estimation. *Journal of Computational and Graphical Statistics*, 2(3):273–289, 1993.

[14] S. He, W. Liang, J. Shen, and G. Yang. Empirical likelihood for right censored lifetime data. *Journal of the American Statistical Association*, 111(514):646–655, 2016.

[15] N.L. Hjort. *Towards semiparametric bandwidth selectors for kernel density estimators*. Department of Mathematics, University of Oslo, 1999.

[16] N.L. Hjort, I.W. McKeague, and I. Van Keilegom. Extending the scope of empirical likelihood. *The Annals of Statistics*, 37(3):1079–1111, 2009.

[17] D.G. Hoel. A representation of mortality data by competing risks. *Biometrics*, 28(2):475–488, 1972.

[18] L. Horváth, Z. Horváth, and W. Zhou. Confidence bands for roc curves. *Journal of Statistical Planning and Inference*, 138(6):1894–1904, 2008.

[19] P.G. Karadeniz and I. Ercan. Examining tests for comparing survival curves with right censored data. *Statistics in Transition*, 18(2):311–28, 2017.

[20] J.P. Klein, B. Logan, M. Harhoff, and P.K. Andersen. Analyzing survival curves at a fixed point in time. *Statistics in Medicine*, 26(24):4505–4519, 2007.

[21] J.P. Klein and M.L. Moeschberger. *Survival analysis: techniques for censored and truncated data*. Springer, 2003.

[22] E.L. Lehmann. The power of rank tests. *The Annals of Mathematical Statistics*, 24(1):23–43, 1953.

[23] P. Liu and Y. Zhao. A review of recent advances in empirical likelihood. *Wiley Interdisciplinary Reviews: Computational Statistics*, e1599, 2022.

[24] C.L. Mallows. A note on asymptotic joint normality. *The Annals of Mathematical Statistics*:508–515, 1972.

[25] M. Marozzi. Nonparametric simultaneous tests for location and scale testing: a comparison of several methods. *Communications in Statistics-Simulation and Computation*, 42(6):1298–1317, 2013.

[26] E.M. Molanes Lopez, I. Van Keilegom, and N. Veraverbeke. Empirical likelihood for non-smooth criterion functions. *Scandinavian Journal of Statistics*, 36(3):413–432, 2009.

[27] A. Munk and C. Czado. Nonparametric validation of similar distributions and assessment of goodness of fit. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 60(1):223–241, 1998.

[28] S. Nemes, E. Bülow, and A. Gustavsson. A brief overview of restricted mean survival time estimators and associated variances. *Stats*, 3(2):107–119, 2020.

[29] A.B. Owen. *Empirical likelihood*. CRC press, 2001.

[30] A.B. Owen. Empirical likelihood ratio confidence intervals for a single functional. *Biometrika*, 75(2):237–249, 1988.

[31] A.B. Owen. Empirical likelihood ratio confidence regions. *The Annals of Statistics*, 18(1):90–120, 1990.

[32] C.J. Potgieter and F. Lombard. Nonparametric estimation of location and scale parameters. *Computational Statistics & Data Analysis*, 56(12):4327–4337, 2012.

[33] C.J. Potgieter and F. Lombard. Nonparametric two-sample estimation of location and scale parameters from empirical characteristic functions. *Journal of Statistical Computation and Simulation*, 86(16):3225–3242, 2016.

[34] G. Qin and B.Y. Jing. Empirical likelihood for censored linear regression. *Scandinavian Journal of Statistics*, 28(4):661–673, 2001.

[35] G. Qin and Y. Zhao. Empirical likelihood inference for the mean residual life under random censorship. *Statistics & Probability Letters*, 77(5):549–557, 2007.

[36] J. Qin and J. Lawless. Empirical likelihood and general estimating equations. *The Annals of Statistics*, 22(1):300–325, 1994.

[37] Y.S. Qin and L.C. Zhao. Empirical likelihood ratio confidence intervals for various differences of two populations. *Journal of Systems Science and Complexity*, 13(1):23–30, 2000.

[38] J.N.K. Rao and Q.H. Wang. Empirical likelihood-based inference under imputation for missing response data. *The Annals of Statistics*, 30(3):896–924, 2002.

[39] G.A. Rousselet, C.R. Pernet, and R.R. Wilcox. Beyond differences in means: robust graphical methods to compare two groups in neuroscience. *European Journal of Neuroscience*, 46(2):1738–1748, 2017.

[40] P. Royston and M.K.B. Parmar. Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. *BMC Medical Research Methodology*, 13(1):1–15, 2013.

[41] P. Royston and M.K.B. Parmar. The use of restricted mean survival time to estimate the treatment effect in randomized clinical trials when the proportional hazards assumption is in doubt. *Statistics in Medicine*, 30(19):2409–2421, 2011.

[42] S.J. Sheather and M.C. Jones. A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society: Series B (Methodological)*, 53(3):683–690, 1991.

[43] E. Stoimenova and N. Balakrishnan. A class of exceedance-type statistics for the two-sample problem. *Journal of Statistical Planning and Inference*, 141(9):3244–3255, 2011.

[44] W. Stute. The central limit theorem under random censorship. *The Annals of Statistics*:422–439, 1995.

[45] W. Stute. The jackknife estimate of variance of a kaplan-meier integral. *The Annals of Statistics*, 24(6):2679–2704, 1996.

[46] W. Stute and J.L. Wang. The strong law under random censorship. *The Annals of Statistics*:1591–1607, 1993.

[47] S. Subramanian. Function-based hypothesis testing in censored two-sample location-scale models. *Lifetime Data Analysis*, 26(1):183–213, 2020.

[48] D.R. Thomas and G.L. Grunkemeier. Confidence interval estimation of survival probabilities for censored data. *Journal of the American Statistical Association*:865–871, 1975. ISSN: 0162-1459.

[49] J. Valeinis. *Confidence bands for structural relationship models*. PhD thesis, Niedersachsische Staats-und Universitatsbibliothek Gottingen, 2007.

[50] J. Valeinis and E. Cers. *EL: Two-Sample Empirical Likelihood*. R package version 1.2. 2022. URL: `https://CRAN.R-project.org/package=EL`.

[51] J. Valeinis, E. Cers, and J. Cielens. Two-sample problems in statistical data modelling. *Mathematical modelling and analysis*, 15(1):137–151, 2010.

[52] M. Velina, J. Valeinis, and G. Luta. Empirical likelihood-based inference for the difference of two location parameters using smoothed m-estimators. *Journal of Statistical Theory and Practice*, 13:1–21, 2019.

[53] Q.H. Wang and B.Y. Jing. Empirical likelihood for a class of functionals of survival distribution with censored data. *Annals of the Institute of Statistical Mathematics*, 53(3):517–527, 2001.

[54] W. Weibull. A statistical theory of strength of materials. *IVB-Handl.*, 1939.

[55] J. Zheng, J. Shen, and S. He. Adjusted empirical likelihood for right censored lifetime data. *Statistical Papers*, 55(3):827–839, 2014.

[56] M. Zhou. *Empirical likelihood method in survival analysis*, volume 79. CRC Press, 2015.

[57] M. Zhou. Restricted mean survival time and confidence intervals by empirical likelihood ratio. *Journal of Biopharmaceutical Statistics*, 31(3):362–374, 2021.

[58] M. Zhou and J.-H. Jeong. Empirical likelihood ratio test for median and mean residual lifetime. *Statistics in Medicine*, 30(2):152–159, 2011.

Acknowledgements

I would like to express my gratitude to my family for their support both during my doctoral studies and throughout the writing of my doctoral thesis. Special thanks to my life partner, Normunds, and my children, Valdis, Jānis, and Rasa, for their understanding, patience, love, as well as to my parents and Normunds' mother for their support and assistance in our daily lives. I extend my thanks to my cousins, Antra and Dace, for their care and support, including technical assistance when needed.

A tremendous thank you to my colleague Māra, for her inspiration and emotional support, as well as for her practical assistance and valuable advice. I would also like to express my gratitude to the rest of my SPDAL colleagues, especially Reinis, Artis A., Artis L., and Jānis G. Thanks also to all the students who explored topics related to my research in their bachelor's theses and contributed to the completion of this work.

I would like to thank the family of the Department of Mathematics and administrative staff at the Faculty of Physics, Mathematics, and Optometry of the University of Latvia for their significant support during the writing of my doctoral thesis. Special thanks to my former high school mathematics teacher and current colleague, Baiba Āboltiņa, whose guidance led me to pursue studies in the field of statistics.

My sincere thanks to Dr.med. Jeļizaveta Sokolovska for the invaluable experience gained in scientific collaboration projects and for granting me the opportunity to use data from diabetes patients (Project No. 1.1.1.2/VIAA/3/19/525). I am also grateful to Viktorija Ulanova for engaging discussions and for allowing me to use data on tuberculosis patients (Project No. lzp-2020/1-0050) for method illustrations.

Last but certainly not least, I extend my deepest gratitude to my doctoral thesis advisor, Dr.math., Professor Jānis Valeinis, for his selfless and immeasurable support throughout the preparation of publications, the writing of the doctoral thesis, as well as during my doctoral studies and my career development in teaching. Thank you for your motivation and trust! I am delighted to be a part of Jānis's purpose-driven and enthusiastically led team.