12-31-2021

Model checks for two-sample location-scale

Atefeh Javidialsaadi
New Jersey Institute of Technology

ABSTRACT

MODEL CHECKS FOR TWO-SAMPLE LOCATION-SCALE

by
Atefeh Javidialsaadi

Two-sample location-scale refers to a model that permits a pair of standardized random variables to have a common distribution. This means that if X_1 and X_2 are two random variables with means μ_1 and μ_2 and standard deviations σ_1 and σ_2, then $(X_1 - \mu_1)/\sigma_1$ and $(X_2 - \mu_2)/\sigma_2$ have some common unspecified *standard* or *base* distribution F_0. Function-based hypothesis testing for these models refers to formal tests that would help determine whether or not two samples may have come from *some* location-scale family of distributions, without specifying the standard distribution F_0. For uncensored data, Hall et al. (2013) proposed a test based on empirical characteristic functions (ECFs), but it can not be directly applied for censored data. Empirical likelihood with minimum distance (MD) plug-ins provides an alternative to the approach based on ECFs (Subramanian, 2020). However, when working with standardized data, it appeared feasible to set up plug-in empirical likelihood (PEL) with estimated means and standard deviations as plug-ins, which avoids MD estimation of location and scale parameters and (hence) quantile estimation. This project addresses two issues: (i) Set up a PEL founded testing procedure that uses sample means and standard deviations as the plug-ins for uncensored case, and Kaplan–Meier integral based estimators as plug-ins for censored case, (ii) Extend the ECF test to accommodate censoring. Large sample null distributions of the proposed test statistics are derived. Numerical studies are carried out to investigate the performance of the proposed methods. Real examples are also presented for both the uncensored and censored cases.

MODEL CHECKS FOR TWO-SAMPLE LOCATION-SCALE

by
Atefeh Javidialsaadi

A Dissertation
Submitted to the Faculty of
New Jersey Institute of Technology
and Rutgers, The State University of New Jersey – Newark
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Mathematical Sciences

Department of Mathematical Sciences
Department of Mathematics and Computer Science, Rutgers-Newark

December 2021

APPROVAL PAGE

MODEL CHECKS FOR TWO-SAMPLE LOCATION-SCALE

Atefeh Javidialsaadi

Dr. Sundarraman Subramanian, Dissertation Advisor | Date
Associate Professor of Mathematical Sciences, NJIT

Dr. Sunil Dhar, Committee Member | Date
Professor of Mathematical Sciences, NJIT

Dr. Satrajit Roychoudhury, Committee Member | Date
Senior Director, Statistical Research and Data Science Center,
Pfizer, New York, NY

Dr. Zuofeng Shang, Committee Member | Date
Associate Professor of Mathematical Sciences, NJIT

Dr. Antai Wang, Committee Member | Date
Associate Professor of Mathematical Sciences, NJIT

BIOGRAPHICAL SKETCH

Author: Atefeh Javidialsaadi

Degree: Doctor of Philosophy

Date: December 2021

Undergraduate and Graduate Education:

- Doctor of Philosophy in Mathematical Sciences,
  New Jersey Institute of Technology, Newark, NJ, 2021
- Master of Science in Statistics,
  Tarbiat Modares University, Tehran, Iran, 2014
- Bachelor of Science in Statistics,
  Shiraz University, Shiraz, Iran, 2009

Major: Mathematical Sciences

Presentations and Publications:

A. Javidialsaadi, S. Mondal and S. Subramanian. Model checks for two-sample location-scale, 2021, Joint Statistical Meeting (JSM'2021), Virtual conference, August 8-12, 2021.

A. Javidialsaadi, S. Mondal and S. Subramanian. Model checks for two-sample location-scale. (Submitted)

To my love, Ehsan, who inspired me and left everything behind for this which he would not read it, EVER!

ACKNOWLEDGMENT

First and foremost, I would like to express my deep and sincere gratitude to my dissertation advisor, Dr. Sundarraman Subramanin, who has been actively helping me toward the completion of this work. It has been an amazing experience working with him in the past four years. His dynamism, vision, sincerity and motivation have deeply inspired me. He has taught me the methodology to carry out the research and to present the research works as clearly as possible. I thank Dr. Subramanian for his immense patience and unconditional support. It was a great privilege and honor to work and study under his guidance. I'm so grateful having such a tremendous mentor during my PhD study.

Besides, I would like to thank all my dissertation committee members Dr. Sunil Dhar, Dr. Satrajit Roychoudhury, Dr. Zuofeng Shang, and Dr. Antai Wang for their participation as the defense committee and also their helpful comments.

I would like to thank Department of Mathematical Sciences for supporting me financially through parts of my doctoral program, and thank Ms. Gonzalez-Lenahan and Dr. Ziavras who helped me to revise this work.

I would also like to thank my friends who made these four years easy and fun years, considering being away from our families, and made my experience at NJIT exciting. Thank you Beibei Li, Dr. Gan Luan, Axel Turnquist and Ruqi Pei.

Last but not the least, I would like to thank my family for their love and encouragement. To my Mom, thank you for supporting me in all my pursuits. I am so grateful for the way you raised me and for your love, caring and sacrifices that you made to educate me and prepare me for what was ahead of me in my life. To my sisters, thanks for your endless support. To my caring, loving, and supportive husband, Ehsan: my deepest gratitude, your encouragement when the times got rough

are much appreciated and duly noted. It was, is and always will be a great comfort and relief to know that you are there for me.

TABLE OF CONTENTS

Chapter **Page**

TABLE OF CONTENTS
(Continued)

Chapter **Page**

LIST OF TABLES

Table **Page**

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Two-sample *location-scale* arises in areas such as climate dynamics (Adamson and Nash, 2013), biomedical (Gerhard and Hothorn, 2010), medicine (Neuhauser et al. 2010), finance (Lunde and Timmermann, 2004) and the mining industry (Hall et al., 2013). In biomedical studies for examples, when two samples come from a location-scale family of distributions, incorporating this structure would improve efficacy comparisons between treatments. For efficient treatment comparisons within a location-scale framework to bear fruition, however, it is mandatory that a diagnostic test for location-scale membership be provided. This issue has been addressed, to a degree, see Hall et al. (2013), who developed a test founded on empirical characteristic functions (ECFs), but only for uncensored data, and Subramanian (2020), who developed a test founded on plug-in empirical likelihood (PEL), both for uncensored and censored data. In this dissertation, an alternative approach based on the PEL, will be developed. Furthermore, a censored ECF test, which is a nontrivial extension of its uncensored Hall et al. (2013) counterpart, will also be developed.

The knowledge that two distributions differ only in location and scale is reported to yield operational and economic advantages, enabling protocols we have for one type of data to be applied directly on another (Hall et al., 2013). For example, researchers may be interested in the effectiveness of a new treatment on a specific disease. Suppose that X and Y are the times taken to obtain relief from the malady when administered the new and standard treatments respectively. If compelling evidence can be provided that the distributions of X and Y belong to some location-scale family, then all the protocol and information available for the standard treatment like, say, the estimated relief proportion at time t, can be obtained for the new treatment directly by exploiting the model structure. Zhang and Yu (2002) used box plots of censored data to draw conclusions about probable distributional differences in scale and location. Conclusions based on box plots of censored responses rather than real lifetimes, on the other hand, may offer an incorrect picture and should be handled with caution. It is mandatory, therefore, that every such study using location-scale models include function-based hypothesis testing (Hall et al., 2013, Subramanian, 2020).

Consider two random variables X_1 and X_2 with distribution functions F_1 and F_2 respectively, both unspecified. Let the respective means be μ_1 and μ_2, and standard deviations be σ_1 and σ_2. Let F_{10} and F_{20} be the distribution functions of $U_1 = (X_1 - \mu_1)/\sigma_1$ and $U_2 = (X_2 - \mu_2)/\sigma_2$. Then $F_{10}(t) = F_1(\mu_1 + \sigma_1 t), t \in \mathbb{R}$ and $F_{20}(t) = F_2(\mu_2 + \sigma_2 t), t \in \mathbb{R}$. Letting $\varphi_i(t) = \mu_i + \sigma_i t, i = 1, 2$, then $F_{10}(t) = F_1(\varphi_1(t))$ and $F_{20}(t) = F_2(\varphi_2(t))$.

Suppose that $F_{10} = F_{20} := F_0$ (say). Note that F_0, designated as the *standard* or *base* distribution, has mean 0 and variance 1. In turn, the distributions F_1 and F_2 can be expressed in terms of F_0 by $F_i(t) = F_0((t - \mu_i)/\sigma_i), i = 1, 2$. Such considerations indicate that X_1 and X_2 may be represented by the equation $X_i = \mu_i + \sigma_i \epsilon, i = 1, 2$, where ϵ has distribution F_0.

Suppose that two independent samples $\{X_{ij}, j = 1, \ldots, n_i, i = 1, 2\}$ of sizes n_1 and n_2 are available from the X_1 and X_2 distributions respectively. Would these samples support the hypothesis that X_1 and X_2 belong to *some* location-scale family of distributions with standard member F_0, where F_0 is unspecified? Thus, the objective, stated formally, would be to devise a test for the hypothesis $H_0 : F_{10} = F_{20}$ or, equivalently, $F_1 \circ \varphi_1 = F_2 \circ \varphi_2$. Since F_0, the standard member, is unspecified, this kind of hypothesis testing may be regarded as function-based (Hall et al., 2013). Note that model adequacy tests for the case that F_0 is unspecified are inherently more difficult than tests that operate within the confines of specific location-scale, where F_0 is specified; e.g., standard normal or standard exponential.

It has been well documented that there is a general lack of function-based hypothesis testing techniques for two-sample location-scale families (Hall et al. 2013). To test equality of distributions of standardized variables, it would seem obvious to employ empirical distribution functions (EDFs) or empirical quantile functions (EQFs). According to Hall et al. (2013), however, both these strategies are impractical for the current situation. The weak limits of test statistics are not distribution free, hence bootstrap resampling from the data would be required under the null hypothesis. The bootstrap, on the other hand, does a poor job of approximating distributions in this case, and the bootstrap estimate of the null distribution converges particularly slowly. These difficulties prompted Hall et al. (2013) to devise function-based hypothesis testing in location-scale founded on ECFs. Such an approach is reported to work effectively and avoids the issues that plague EDF and EQF-based solutions.

Unfortunately, however, the cited advantages are nullified when there is censoring. Specifically, when the two censoring distributions are not the same, the bootstrap approach of Hall et al. (2013) would not be applicable. Therefore, any advantage that ECFs may have over other approaches is dissipated when there is censoring. Furthermore, when there is censoring to be accommodated, nontrivial adjustments not yet devised are needed for the ECF approach to work. Thus, a search for an alternative to the ECF is not without merit. It is this alternative, the PEL, that will be the first focus of this project. It will be seen that the proposed PEL approach applies readily to both uncensored and censored cases. The plug-ins for the censored case will be based on Kaplan–Meier (KM) integrals (Stute, 1995). Subsequently, a censored ECF test will also be developed that employs Stute's (1995)

KM integrals, with appropriate plug-ins for the means and standard deviations based on it thereof.

Csörgő (1981) and Marcus (1981) studied the ECF's convergence and limiting behavior. Several one-sample issues have been investigated by applying the ECF. Feuerverger and Mureika (1977) suggested an ECF-based test for symmetry, while Feuerverger and McDunnough (1981) explored efficient parameter estimation in one-sample problems when the kind of distribution is given. Tests for univariate normality based on the ECF were investigated by Epps and Pulley (1983) and Hall and Welsh (1983). ECF goodness-of-fit tests were studied by Koutrouvelis and Kellermeier (1981) and Koutrouvelis (1985). Matsui and Takemura (2005) dealt with goodness-of-fit ECF testing for Cauchy distributions, whereas Meintanis and Swanepoel (2007) and Huskovâ and Meintanis (2008) dealt with generalized goodness-of-fit ECF tests.

Prior to Hall et al. (2013), Epps and Singleton (1986) appear to be the only work on two-sample testing based on ECFs. Otherwise, the challenge of determining whether two samples are from some location-scale family has received little attention. One test proposed by Doksum and Sievers (1976) is to see if a nonparametric confidence band for the quantile comparison function comprises a straight line. However, because nonparametric confidence bands are generally quite wide, this test has poor power. Potgieter and Lombard (2008) developed a permutation test based on a confidence band for a linear quantile comparison function in the setting of matched pair data. See Hall et al. (2013) and Subramanian (2020) for related references.

For the case when censoring is completely absent, Hall et al. (2013) proposed their test that relies on integrated weighted squared modulus difference of two ECFs. Note that their ECF test requires the sample means and standard deviations as plug-ins. When there is censoring, these straightforward estimates are no more available, however. Hall et al. (2013) derived the large-sample null distribution of their test statistic. Since the weak limit is not distribution free, they proposed a bootstrap test. While their proposed test correctly applies when there is no censoring, we can not apply it when the observations are susceptible to random censoring. However, the plug-ins for the means and standard deviations given in this dissertation, along with the Kaplan–Meier (KM) estimator, will be utilized to propose a censored ECF test.

Note that the proposed PEL essentially owes its development to that proposed by Li (2003) to test the adequacy of a parametric distribution. See also Subramanian (2020), whose PEL test employed minimum distance (MD) estimators of the location and scale parameters, which require estimation of quantiles. In this dissertation, it was discovered that, working with standardized variables, it is possible to set up PEL with estimated means and standard deviations as plug-ins. This offers the advantage of comparing the ECF and PEL approaches on an even footing. Thus, we propose a PEL test that avoids MD estimation of location and scale parameters and (hence) quantile estimation, permitting direct comparisons with the ECF test. Necessarily, therefore, the likelihood ratio will be indexed by $\boldsymbol{\theta}$, where $\boldsymbol{\theta}$ represents the means and standard deviations, unlike in Subramanian (2020), where the likelihood ratio was indexed by $\boldsymbol{\theta}$ comprising of the location and scale parameters. Furthermore, the proposed PEL test is founded on estimated standardized variables and computed on the standardized scale. The PEL using MD estimators, however, was computed on the original scale. Thus, the apparent resemblance of the two PEL tests camouflages the fact that the two approaches lead to different sums for $R_{\boldsymbol{\theta}}(t)$, the likelihood ratio, see Equation (3.8). Also, as reported already, the plug-ins for the proposed PEL are founded on Stute's (1995) KM integrals and are quite different from the estimated MD plug-ins used in Subramanian (2020). In addition, because variables are standardized, we confront certain nuanced differences, requiring some altered technical treatments for the large-sample analysis. Finally the plug-ins obtained via Stute's (1995) KM integrals also prove decisively important for the censored ECF test that is proposed.

Just to make matters explicit, the proposed PEL test applies for both censored and uncensored data. More specifically, the proposed PEL test has the same form whether the data are censored or uncensored. As already reported, the only difference is that, for censored data, the plug-ins are based on Stute's (1995) KM integrals. For uncensored data, however, the sample means and sample standard deviations suffice for the plug-ins. Thus, unlike Equation (2.1), the extension of which to censoring would require non-trivial adjustments, the PEL ratio test statistic would apply for censored data also.

Because of plug-ins, the proposed PEL test is not asymptotically distribution free, a phenomenon already well noted (Hjort et al., 2009). Therefore, it is a challenge to obtain critical values needed for carrying out the test. The full bootstrap, the so-called "n bootstrap", did not work well for the PEL. Following Subramanian (2020), where failure of the full bootstrap was noted as well, use of smaller bootstrap sample sizes is warranted. The data-driven procedure given in Subramanian (2020), that adapts Bickel and Sakov's (2008) one-sample proposal, yielded a reasonably accurate choice for the bootstrap sample sizes. The bootstrap sample size was chosen after a search that involved a diagonal "trajectory". Subsequently, a relatively more refined search algorithm, involving two-steps, was investigated. Both approaches identified bootstrap sample sizes that returned realized levels closer to the nominal 5% level.

The rest of this dissertation is outlined as follows. The ECF test for uncensored data is introduced in Chapter 2. The plug-in PEL is studied in Chapter 3. In Section 3.1 for uncensored case, we introduce preliminaries and drive the log likelihood ratio (LR). The null sampling distribution of the uncensored PEL test is described. In Section 3.2 the extension to right censoring is considered. The algorithms for offering bootstrap sample size are described in Chapter 3.3. The ECF test for censored data is introduced in Chapter 4. Simulation results are reported in Chapter 5. Four real examples are presented in Chapter 6. A brief concluding discussion is given in Chapter 7. Technical details are coveredl, in the Appendix.

CHAPTER 2

ECF TEST FOR UNCENSORED DATA

We consider two random variables X_1 and X_2 with distribution functions F_1 and F_2 respectively, both unspecified. Let the respective means be μ_1 and μ_2, and standard deviations be σ_1 and σ_2. Let F_{10} and F_{20} be the distribution functions of $U_1 = (X_1 - \mu_1)/\sigma_1$ and $U_2 = (X_2 - \mu_2)/\sigma_2$. The null hypothesis is $H_0 : F_{10} = F_{20}$ or, equivalently, $F_1 \circ \varphi_1 = F_2 \circ \varphi_2$. Suppose that two independent samples $\{X_{kj}, j = 1, \ldots, n_k, k = 1, 2\}$ of sizes n_1 and n_2 are available from the X_1 and X_2 distributions respectively. Let $\chi_{U_k}(t) = E[exp(itU_{kj})]$ be the characteristic function of U_{kj}. It is estimated by

$$\begin{aligned} \hat{\chi}_{U_k}(t) &= \frac{1}{n_k} \sum_{j=1}^{n_i} \exp\left\{ it \left(\frac{X_{kj} - \hat{\mu}_k}{\hat{\sigma}_k} \right) \right\} \\ &= \frac{1}{n_k} \sum_{j=1}^{n_i} \cos\left\{ it \left(\frac{X_{kj} - \hat{\mu}_k}{\hat{\sigma}_k} \right) \right\} \\ &+ i.\frac{1}{n_k} \sum_{j=1}^{n_i} \sin\left\{ it \left(\frac{X_{kj} - \hat{\mu}_k}{\hat{\sigma}_k} \right) \right\} \end{aligned} \tag{2.1}$$

Hall et al.(2013) proposed the test statistic

$$\hat{L} = \frac{n_1 n_2}{n_1 + n_2} \int |\hat{\chi}_{U_1}(t) - \hat{\chi}_{U_2}(t)|^2 \, w(t) dt, \tag{2.2}$$

where $w(t)$ is a specified weight function. Let B denotes a Brownian bridge. Hall et al. (2013) proved that the process $n^{1/2}\{\hat{\chi}_{U_k}(t) - \chi_{U_k}(t)\}$ converges uniformly to a limiting process $\mathcal{G}_{F_{k0}}(t)$ in $t \in [-A, A]$ for any finite positive A, where

$$\mathcal{G}_{F_{k0}}(t) = \int \left\{ \exp(itx) - itx\chi_{U_k}(t) - \frac{t}{2}(x^2 - 1)\chi'_{U_k}(t) \right\} \times dB\{F_{k0}(x)\}, \tag{2.3}$$

and $\chi'_{U_k}(t)$ is the first derivative of $\chi_{U_k}(t)$. When the two sample are independent, Hall et al. (2013) proved that

$$\begin{aligned}\left(\frac{n_1 n_2}{n_1+n_2}\right)^{1/2}\{\hat{\chi}_{U_1}(t)-\hat{\chi}_{U_2}(t)\} \ &\xrightarrow{D}\ \int\left\{\exp(itx)-itx\chi_{U_1}(t)-\frac{t}{2}(x^2-1)\chi'_{U_1}(t)\right\}d\\ &\times\ [\sqrt{1-\lambda}B_1\{F_{10}(x)\}+\sqrt{\lambda}B_2\{F_{20}(x)\}]\\ &\stackrel{D}{=}\ \int\left\{\exp(itx)-itx\chi_{U_1}(t)-\frac{t}{2}(x^2-1)\chi'_{U_1}(t)\right\}\\ &\times\ dB\{F_{10}(x)\} \qquad (2.4)\end{aligned}$$

where B_1 and B_2 are independent copies of B, and $\lambda=\lim_{n_1,n_2\to\infty}\frac{n_1 n_2}{n_1+n_2}$ is considered to be within the range $(0,1)$. Hall et al. (2013) showed that $\hat{L}$ converges in distribution to the random variable

$$L=\int|\mathcal{G}_{F_{k0}}(t)|^2 w(t)dt. \qquad (2.5)$$

The random variable L is distributed as $\sum_{k\geq 1}\gamma_k(t)Z_k^2$, where $Z_1, Z_2, ...$ are independent and identically distributed standard normal random variables and $\gamma_1(t)>\gamma_2(t)>...$ are the eigenvalues of the covariance function of the process $w(t)\mathcal{G}_{F_{k0}}(t)$. The test statistic $\hat{L}$ is not asymptotically distribution free because the eigenvalues γ_j depend on the type of underlying distribution. Therefore, for practical applications, one will have to rely on the bootstrap to perform the test.

To calculate the statistics stated in Equation (2.2), a suitable weight function must be chosen.

According to Hall et al. (2013) a popular choice of weight function is $w(t)=exp(-ht^2)$. The other boxcar weight function, $w(t)=1, \quad |t|\leq\pi/2h$, was another alternative. A third choice was $w(t)=(1-|th|)^2, \quad |t|\leq 1/h$; see Hall et al.(2013) for more insight on selecting $w(t)$.

To estimate the realized level of the test statistic, Hall et al. (2013) proposed using bootstrap as follows

- Generate samples X_1 and X_2 of size $n_k, k = 1, 2$, from a population that satisfies the null hypothesis.
- Compute the test statistic.
- Generate B bootstraps by resampling X_1^* and X_2^*, the bootstrap sample, from only the X_1 sample to estimate the bootstrap p-value.

The bootstrap p-value is based on M replications.

CHAPTER 3

PLUG-IN EMPIRICAL LIKELIHOOD RATIO

We first derive the test statistic when there is no censoring. The form of the derived likelihood ratio will indicate how, with minor tweaks, it will extend readily to the censored case.

3.1 The Uncensored Case

For $i = 1, 2$, let X_i have distribution function F_i with mean μ_i and standard deviation σ_i. Recall that $\varphi_i(t) = \mu_i + \sigma_i t$. The null hypothesis is

$$F_1(\varphi_1(t)) \quad = \quad F_2(\varphi_2(t)), \qquad \text{for all } t \in \mathbb{R}. \tag{3.1}$$

The observed data are $X_{ij}, j = 1, \ldots, n_i$, which are independent and identically distributed (iid) with distribution F_i. We now introduce some relevant results needed for our analysis.

3.1.1 Uncensored data preliminaries

Let $\Lambda_i(t)$ be the group-specific cumulative hazard with associated hazard function $\lambda_i(t)$. However, we will be dealing exclusively with the standardized variables $U_{ij} = (X_{ij} - \mu_i)/\sigma_i$. Accordingly, we let $\Lambda_{i0}(t)$ be the group-specific cumulative hazard of U_{ij} with associated hazard function $\lambda_{i0}(t)$. For each $t \in \mathbb{R}$, write $Y_{i0}(t) = n_i^{-1} \sum_{j=1}^{n_i} I(U_{ij} \geq t)$. Let $\hat{F}_{i0}$ be the empirical distribution function (EDF) of $U_{i1}, ..., U_{in_i}$. For ease of transition to the censored setting, we shall consider the estimator of $\Lambda_{i0}(t)$ given by

$$\hat{\Lambda}_{i0}(t) \quad = \quad \int_{-\infty}^{t} \frac{d\,\hat{F}_{i0}(s)}{Y_{i0}(s)}, \qquad i = 1, 2. \tag{3.2}$$

Let $y_{i0}(t)$ be the limit of $Y_{i0}(t)$. It can be shown that (cf. Major and Rejto, 1988),

$$\begin{aligned}
\hat{\Lambda}_{i0}(t)-\Lambda_{i0}(t) &= \int_{-\infty}^{t}\frac{d\left(\hat{F}_{i0}(u)-F_{i0}(u)\right)}{y_{i0}(u)}-\int_{-\infty}^{t}\frac{Y_{i0}(u)-y_{i0}(u)}{y_{i0}^{2}(u)}dF_{i0}(u)\\
&+ o_{\mathbb{P}}(n^{-1/2})\\
&= \frac{1}{n_i}\sum_{j=1}^{n_i}\frac{(I(U_{ij}\le t))}{y_{i0}(U_{ij})}-\frac{1}{n_i}\sum_{j=1}^{n_i}\int_{-\infty}^{t}\frac{I(U_{ij}\ge u)}{y_{i0}(u)}d\Lambda_{i0}(u)\\
&+ o_{\mathbb{P}}(n^{-1/2}). \qquad (3.3)
\end{aligned}$$

Let $\tilde{U}_{i1} < \cdots < \tilde{U}_{im_i} < \infty$ denote the $m_i \le n_i$ distinct standardized observations. Let r_{ij} and d_{ij} be the number at risk just before $\tilde{U}_{ij}$ and the number of failures at $\tilde{U}_{ij}$ respectively. Let $\eta_i(t)=\sum_{j=1}^{m_i} I(\tilde{U}_{ij}\le t)$. A uniformly consistent estimator of the asymptotic variance function of $n_i^{1/2}(\hat{\Lambda}_{i0}(\cdot)-\Lambda_{i0}(\cdot))$, is (cf. Kalbfleisch and Prentice, 2002)

$$\vartheta_{i0}^{2}(t) = n_i\sum_{j:\tilde{U}_{ij}\le t}\frac{d_{ij}}{r_{ij}(r_{ij}-d_{ij})} = n_i\sum_{j=1}^{\eta_i(t)}\frac{d_{ij}}{r_{ij}(r_{ij}-d_{ij})},\quad i=1,2. \qquad (3.4)$$

3.1.2 Proposed plug-in likelihood ratio

Fix $t>0$. For known $\boldsymbol{\theta}\equiv\left(\boldsymbol{\mu}^{\top},\boldsymbol{\sigma}^{\top}\right)^{\top}$, where $\boldsymbol{\mu}=(\mu_1,\mu_2)^{\top}$ and $\boldsymbol{\sigma}=(\sigma_1,\sigma_2)^{\top}$, we first derive an expression for $R_{\boldsymbol{\theta}}(t)$, the difference between the log of the constrained and unconstrained maximum likelihoods. For $i=1,2$, let Δ_i denote the space of all survival functions (with distributions having mean zero and variance 1) with support on the points $\tilde{U}_{i1},...,\tilde{U}_{im_i}$. For $K_i\in\Delta_i$, it is well known that the nonparametric likelihood for (K_1,K_2) is

$$L(K_1,K_2) = \prod_{i=1}^{2}\prod_{j=1}^{m_i}\left(K_i(\tilde{U}_{ij}-)-K_i(\tilde{U}_{ij})\right)^{d_{ij}}. \qquad (3.5)$$

The mechanism for maximizing $L(K_1,K_2)$ is via its parameterization through the discrete hazards, $\lambda_{ij}, j=1,\ldots,m_i, i=1,2$ (Kalbfleisch and Prentice, 2002).

Specifically, writing $K_i(\tilde{U}_{ij}) = \prod_{l=1}^{j}(1-\lambda_{il})$, the likelihood, expressed in terms of λ_{ij}, r_{ij} and d_{ij}, is

$$L(K_1, K_2) = \prod_{i=1}^{2}\prod_{j=1}^{m_i} \lambda_{ij}^{d_{ij}}(1-\lambda_{ij})^{r_{ij}-d_{ij}}. \tag{3.6}$$

Maximizing $L(K_1, K_2)$ is equivalent to maximizing each quantity in the double product, giving

$$\hat{\lambda}_{ij} = d_{ij}/r_{ij}, j = 1, \ldots, m_i, \quad i = 1, 2. \tag{3.7}$$

Note that $\prod_{l=1}^{j}(1-\hat{\lambda}_{il}) = \hat{S}_{i0}(\tilde{U}_{ij})$ gives the empirical survival function, the estimator of $S_{i0}(t) = 1 - F_{i0}(t)$, where F_{i0} is the distribution function of U_i. Plugging the estimates in (3.7) into Equation (3.6), the unconstrained maximum of the left hand side (LHS) of Equation (3.5) is

$$L(\hat{S}_{10}, \hat{S}_{20}) = \prod_{i=1}^{2}\prod_{j=1}^{m_i} \left(\frac{d_{ij}}{r_{ij}}\right)^{d_{ij}} \left(1 - \frac{d_{ij}}{r_{ij}}\right)^{r_{ij}-d_{ij}}.$$

Recall that $\eta_i(t) = \sum_{j=1}^{m_i} I(\tilde{U}_{ij} \leq t)$. In Appendix A.1, we maximize Equation (3.6) subject to $K_1(\cdot) = K_2(\cdot)$. The constrained estimates $\tilde{\lambda}_{ij}$ are given by Equation (A.1). The constrained survival function estimates are $\tilde{S}_{i0}(\tilde{U}_{ij}) = \prod_{l=1}^{j}(1-\tilde{\lambda}_{il}), j = 1, \ldots, m_i, i = 1, 2$. The Lagrange multiplier $\hat{\gamma}_{\boldsymbol{\theta}}(t)$ satisfies Equation (A.2). For notational simplicity, we shall suppress its dependence on t. Plugging the estimates into Equation (3.6), we obtain

$$
\begin{aligned}
L(\tilde{S}_{10}, \tilde{S}_{20}) &= \prod_{j=1}^{\eta_1(t)} \left(\frac{d_{1j}}{r_{1j} - \hat{\gamma}_{\boldsymbol{\theta}}}\right)^{d_{1j}} \left(1 - \frac{d_{1j}}{r_{1j} - \hat{\gamma}_{\boldsymbol{\theta}}}\right)^{r_{1j}-d_{1j}} \\
&\times \prod_{j=\eta_1(t)+1}^{m_1} \left(\frac{d_{1j}}{r_{1j}}\right)^{d_{1j}} \left(1 - \frac{d_{1j}}{r_{1j}}\right)^{r_{1j}-d_{1j}} \\
&\times \prod_{j=1}^{\eta_2(t)} \left(\frac{d_{2j}}{r_{2j} + \hat{\gamma}_{\boldsymbol{\theta}}}\right)^{d_{2j}} \left(1 - \frac{d_{2j}}{r_{2j} + \hat{\gamma}_{\boldsymbol{\theta}}}\right)^{r_{2j}-d_{2j}} \\
&\times \prod_{j=\eta_2(t)+1}^{m_2} \left(\frac{d_{2j}}{r_{2j}}\right)^{d_{2j}} \left(1 - \frac{d_{2j}}{r_{2j}}\right)^{r_{2j}-d_{2j}} .
\end{aligned}
$$

For fixed t and known $\boldsymbol{\theta}$ the log LR, denoted by $R_{\boldsymbol{\theta}}(t)$, is derived in Appendix A.1.1 and is

$$
\begin{aligned}
R_{\boldsymbol{\theta}}(t) &= \sum_{j=1}^{\eta_1(t)} \left[(r_{1j} - d_{1j}) \log\left(1 - \frac{\hat{\gamma}_{\boldsymbol{\theta}}}{r_{1j} - d_{1j}}\right) - r_{1j} \log\left(1 - \frac{\hat{\gamma}_{\boldsymbol{\theta}}}{r_{1j}}\right)\right] \\
&+ \sum_{j=1}^{\eta_2(t))} \left[(r_{2j} - d_{2j}) \log\left(1 + \frac{\hat{\gamma}_{\boldsymbol{\theta}}}{r_{2j} - d_{2j}}\right) - r_{2j} \log\left(1 + \frac{\hat{\gamma}_{\boldsymbol{\theta}}}{r_{2j}}\right)\right]. \qquad (3.8)
\end{aligned}
$$

To define our test statistic, for small enough $\eta > 0$, let $\nu_i = \inf\{t : F_{i0}(t) > \eta\}$. Let $\alpha_1 = \max(\nu_1, \nu_2)$. Let $\tau_i = \sup\{t : S_{i0}(t) > \eta\}$. Define $\alpha_2 = \min(\tau_1, \tau_2)$. Let $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\sigma}})^\top$ be a consistent estimator of $\boldsymbol{\theta}$. Our proposed PEL test statistic is $\|R_{\hat{\boldsymbol{\theta}}}\|_{\alpha_1}^{\alpha_2} = \sup_{t\in[\alpha_1,\alpha_2]} |R_{\hat{\boldsymbol{\theta}}}(t)|$.

Write $\hat{\eta}_i(t) = \eta_i(t|\hat{\mu}_i, \hat{\sigma}_i)$. Note that $\hat{\gamma} \equiv \hat{\gamma}_{\hat{\boldsymbol{\theta}}}(t)$ solves Equation (A.2) with $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ and satisfies $\hat{D}_2 < \hat{\gamma} < -\hat{D}_1$, where

$$
\hat{D}_1 = \max_{1\le j\le \hat{\eta}_1(t)} \{d_{1j} - r_{1j}\}, \qquad \hat{D}_2 = \max_{1\le j\le \hat{\eta}_2(t)} \{d_{2j} - r_{2j}\}. \qquad (3.9)
$$

After determining the bounds $\hat{D}_2$ and $-\hat{D}_1$, Brent's method can be used to compute the root. More specifically, the estimator $\hat{\gamma}$ is computed at $\tilde{U}_{1k}, k \in \Upsilon$, where $\Upsilon = \{k : \alpha_1 \le \tilde{U}_{1k} \le \alpha_2\}$, then $\hat{\gamma}$ and $\hat{\eta}_i(t)$ are plugged into the right hand side (RHS) of Equation (3.8) to obtain $R_{\hat{\boldsymbol{\theta}}}(\tilde{U}_{1k}), k \in \Upsilon$. The computed value of the PEL test statistic is then taken as $\max_{k\in\Upsilon} R_{\hat{\boldsymbol{\theta}}}(\tilde{U}_{1k})$.

3.1.3 Large-sample null distribution

Define $\hat{\zeta}_i(t) = -\log \hat{S}_i(t)$. The plug-ins are $\hat{\boldsymbol{\mu}} = (\hat{\mu}_1, \hat{\mu}_2)^\top$, where $\hat{\mu}_i = n_i^{-1}\sum_{j=1}^{n_i} X_{ij}$, and $\hat{\boldsymbol{\sigma}} = (\hat{\sigma}_1, \hat{\sigma}_2)^\top$ with $\hat{\sigma}_i$ being the i-th sample standard deviation, the plug-ins computed from the unstandardized X_{ij}. Recall that $\varphi_i(t) = \mu_i + t\sigma_i$. Its estimate is $\hat{\varphi}_i(t) = \hat{\mu}_i + t\hat{\sigma}_i$. Let $\lambda_{i0}(t)$ be the hazard function associated with $F_{i0}(t)$, the distribution function of U_i. Let $n = n_1 + n_2$ and $n_1/n \to \rho$ as $n_i \to \infty, i = 1, 2$. Note that the limit of n_2/n is $1-\rho$. Replace $\eta_i(t)$ in Equation (3.4) with $\hat{\eta}_i(t)$ to obtain $\hat{\vartheta}^2_{i0}(t)$, the estimate of $\vartheta^2_{i0}(t)$. Let $\vartheta^2_{\mathrm{c}}(t)$ be the limit of $\hat{\vartheta}^2_{\mathrm{c}}(t)$, where

$$\hat{\vartheta}^2_{\mathrm{c}}(t) = \frac{\hat{\vartheta}^2_{10}(t)}{\rho} + \frac{\hat{\vartheta}^2_{20}(t)}{1-\rho}. \tag{3.10}$$

In Appendix B, it is shown that $-2R_{\hat{\boldsymbol{\theta}}}(t)$ is asymptotically equivalent to $(n^{1/2}\hat{\mathbb{V}}(t))^2/\vartheta^2_{\mathrm{c}}(t)$, where $\hat{\mathbb{V}}(t) \equiv \hat{\zeta}_1(\hat{\varphi}_1(t)) - \hat{\zeta}_2(\hat{\varphi}_2(t))$. In Lemma 3 in Appendix A.1.2, it is shown that

$$\hat{\zeta}_1(\hat{\varphi}_1(t)) - \hat{\zeta}_2(\hat{\varphi}_2(t)) \equiv \Big(-\log \hat{S}_1(\hat{\varphi}_1(t))\Big) - \Big(-\log \hat{S}_2(\hat{\varphi}_2(t))\Big), t \in [\alpha_1, \alpha_2], \tag{3.11}$$

has the asymptotic representation $\hat{\mathbb{V}}(t) = A_1(t) + B_1(t) - (A_2(t) + B_2(t)) + o_{\mathbb{P}}(n^{-1/2})$, where

$$A_i(t) = \frac{\lambda_{i0}(t)}{\sigma_i}\left((\hat{\mu}_i - \mu_i) + \frac{t}{2\sigma_i}(\hat{\sigma}_i^2 - \sigma_i^2)\right) + o_{\mathbb{P}}(n^{-1/2}) \tag{3.12}$$

$$B_i(t) = \Lambda_{i0}(t) + \Big(\hat{\Lambda}_{i0}(t) - \Lambda_{i0}(t)\Big) + o_{\mathbb{P}}(n^{-1/2}). \tag{3.13}$$

Then, with $\bar{U}_i = (n_i)^{-1}\sum_{j=1}^{n_i} U_{ij}$, after elementary calculations it can be shown that

$$A_i(t) = \lambda_{i0}(t) \times \frac{1}{n_i}\sum_{j=1}^{n_i}\left[U_{ij} + \frac{t}{2}\Big((U_{ij} - \bar{U}_i)^2 - 1\Big)\right] + o_{\mathbb{P}}(n^{-1/2}). \tag{3.14}$$

Plug the asymptotic representation of Equation (3.3) into the RHS of Equation (3.13) to obtain

$$\begin{aligned} B_i(t) &= \Lambda_{i0}(t) + \frac{1}{n_i}\sum_{j=1}^{n_i}\left\{\frac{(I(U_{ij} \le t))}{y_{i0}(U_{ij})} - \int_{-\infty}^{\infty}\frac{I(u \le U_{ij} \wedge t)}{y_{i0}(u)}d\Lambda_{i0}(u)\right\} \\ &+ o_{\mathbb{P}}(n^{-1/2}). \end{aligned} \tag{3.15}$$

Because $\hat{\mathbb{V}}(t)$ is a difference, Λ_{10} and Λ_{20} cancel each other out under the null hypothesis. Thus

$$n^{1/2}\hat{\mathbb{V}}(t) = \frac{1}{\sqrt{\rho}}\left\{n_1^{-1/2}\sum_{j=1}^{n_1} J_{1j}(t)\right\} + \frac{1}{\sqrt{1-\rho}}\left\{n_2^{-1/2}\sum_{j=1}^{n_2} J_{2j}(t)\right\} + o_{\mathbb{P}}(1). \tag{3.16}$$

where

$$\begin{aligned} J_{ij}(t) &= \lambda_{i0}(t)\left[U_{ij} + \frac{t}{2}\left((U_{ij} - \bar{U}_i)^2 - 1\right)\right] \\ &+ \left[\frac{(I(U_{ij} \le t))}{y_{i0}(U_{ij})} - \int_{-\infty}^{U_{ij}\wedge t}\frac{d\Lambda_{i0}(u)}{y_{i0}(u)}\right]. \end{aligned} \tag{3.17}$$

Let $D[\alpha_1, \alpha_2]$ be the class of càdlàg functions defined on $[\alpha_1, \alpha_2]$ equipped with the supremum norm. The large-sample null distribution of $-2R_{\hat{\boldsymbol{\theta}}}(\cdot)$ is stated in Theorem 1. The proof is given in Appendix B.

Theorem 1 *Let $V_i(s,t) = \mathbb{E}\left(J_{i1}(s)J_{i1}(t)\right)$. Under the null hypothesis, $-2R_{\hat{\boldsymbol{\theta}}}(\cdot)$ converges weakly in $D[\alpha_1, \alpha_2]$ to $\mathbb{V}^2(\cdot)/\vartheta_{\mathrm{c}}^2(\cdot)$, where $\mathbb{V}$ is a centered Gaussian process with covariance function*

$$\mathrm{Cov}(\mathbb{V}(s), \mathbb{V}(t)) = \frac{V_1(s,t)}{\rho} + \frac{V_2(s,t)}{1-\rho}. \tag{3.18}$$

The continuous mapping theorem guarantees the weak convergence of the PEL test statistic.

3.2 Extension to Censored Data

The censoring variables C_i are assumed to be independent with distribution functions $G_i, i = 1, 2$. As we know, $X_{ij}, j = 1, \ldots, n_i$, are iid failure times having common distributions as $X_i, i = 1, 2$. Let $C_{ij}, j = 1, \ldots, n_i$, be iid censoring times having the same distribution as C_i. The observed data are random samples $\{(Z_{ij}, \delta_{ij}), j = 1, \ldots, n_i, i = 1, 2\}$, where $Z_{ij} = \min(X_{ij}, C_{ij})$ and $\delta_{ij} = I(X_{ij} \le C_{ij})$. As in the uncensored case, U_{ij} are the standardized failure times not all of which will be observed due to censoring. They have distribution function F_{i0} and cumulative hazard Λ_{i0}. Let $\tilde{Z}_{ij} = (Z_{ij} - \mu_i)/\sigma_i$. Those $\tilde{Z}_{ij}$ with $\delta_{ij} = 1$ coincide with the U_{ij} that are uncensored. The Nelson–Aalen estimator, $\hat{\Lambda}_{i0}(t)$, of $\Lambda_{i0}(t)$ is given by Equation (3.2), but will have $\hat{H}_{i0}^{(1)}(t) \equiv n_i^{-1} \sum_{j=1}^{n_i} I(\tilde{Z}_{ij} \le t, \delta_{ij} = 1)$ as the integrating measure instead of $\hat{F}_{i0}(t)$. Let $\tilde{y}_i(t)$ be the limit of $\tilde{Y}_i(t) = n_i^{-1} \sum_{j=1}^{n_i} I(\tilde{Z}_{ij} \ge t)$. It can be shown that [cf. Equation (3.3)], modulo a remainder of $o_{\mathbb{P}}(n^{-1/2})$ (e.g., Major and Rejto, 1988),

$$\begin{aligned} \hat{\Lambda}_{i0}(t) - \Lambda_{i0}(t) &= \frac{1}{n_i} \sum_{j=1}^{n_i} \frac{\Big(I(\tilde{Z}_{ij} \le t, \delta_{ij} = 1\Big)}{\tilde{y}_i(\tilde{Z}_{ij})} \\ &- \frac{1}{n_i} \sum_{j=1}^{n_i} \int_{-\infty}^{t} \frac{I(\tilde{Z}_{ij} \ge u)}{\tilde{y}_i(u)} d\Lambda_{i0}(u). \end{aligned} \tag{3.19}$$

Let $\tilde{U}_{i1} < \cdots < \tilde{U}_{im_i} < \infty$ be the distinct $\tilde{Z}_{ij}$ with $\delta_{ij} = 1$. Let d_{ij} and r_{ij} be the number of observed failures at $\tilde{U}_{ij}$ and the number at risk just before $\tilde{U}_{ij}$ respectively. The r_{ij} are obtained by summing all $d_{ir}, r = j, \ldots, m_i$, and the censored outcomes greater than $\tilde{U}_{ij}$ (assuming no ties between failed and censored items). Then, Equation (3.8), the log LR, holds. Continuing with the same notation as for the uncensored case, let $\eta_i(t) = \sum_{j=1}^{m_i} I(\tilde{U}_{ij} \le t)$. Subsequently, Equation (3.4) defines $\vartheta_{i0}^2(\cdot)$, a uniformly consistent estimator of the asymptotic variance function of $n_i^{1/2}(\hat{\Lambda}_{i0}(\cdot) - \Lambda_{i0}(\cdot))$.

Let $\kappa_i = \inf\{u : H_i(u) = 1\}$, where H_i is the common distribution function of $Z_{ij}, j = 1, \ldots, n_i$. Also, let $T_{i1} < T_{i2} < \ldots < T_{im_i}$ be the distinct uncensored Z_{ij}'s. To obtain consistent plug-ins for μ_i and σ_i, we assume that $\kappa_i = \infty$ and that κ_i is not an atom of H_i (Stute and Wang, 1993; Stute, 1995). Let ψ_i be the second moment X_i. Then $\Delta\hat{S}_i(T_{ij}) \equiv \hat{S}_i(T_{ij}) - \hat{S}_i(T_{i(j-1)}) = -\hat{S}_i(T_{i(j-1)})d_{ij}/r_{ij}$, where $\hat{S}_i(T_{i0}) \equiv 1$. The estimates of μ_i and ψ_i are

$$\hat{\mu}_i \;=\; -\sum_{j=1}^{m_i} T_{ij}\Delta\hat{S}_i(T_{ij}); \qquad \hat{\psi}_i \;=\; -\sum_{j=1}^{m_i} T_{ij}^2\Delta\hat{S}_i(T_{ij}). \tag{3.20}$$

Remark 1 Suppose that $h : \mathbb{R}^1 \to \mathbb{R}^1$ satisfies $\int |h| dF < \infty$. If either $\kappa_i < \infty$ or κ_i is an atom of H_i, then the KM integral $\int h\, d\hat{F}_i$, converges with probability 1 to $\int h\, d\tilde{F}_i$, where $\tilde{F}_i(x) = F_i(x)I(x < \kappa_i) + F_i(x-)I(x \geq \kappa_i)$ (Stute, 1995). When $\kappa_i = \infty$ and κ_i is not an atom of H_i, then $\int h\, d\hat{F}_i$ converges with probability 1 to $\int h\, dF_i$. Thus, when these conditions hold, the choices $h(x) = x$ and $h(x) = x^2$ yield consistent moment estimators in Equations (3.20).

Remark 2 If either $\kappa_i < \infty$ or κ_i is an atom of H_i, Equations (3.20) yield inconsistent estimators of μ_i and ψ_i. Consequently, when used as plug-ins, the process $\hat{\mathbb{V}}$ [see Equation (3.11)] would be asymptotically biased, the non-vanishing bias being the difference of the cumulative hazards of $(X_1 - \tilde{\mu}_1)/\tilde{\sigma}_1$ and $(X_2 - \tilde{\mu}_2)/\tilde{\sigma}_2$ where $\tilde{\mu}_i = \int x d\tilde{F}_i$ and $\tilde{\sigma}_i^2 = \tilde{\psi}_i^2 - (\tilde{\mu}_i)^2$ and $\tilde{\psi}_i = \int x^2 d\tilde{F}_i$.

To define the test statistic, as for the uncensored case, for small enough $\eta > 0$, let $\nu_i = \inf\{t : H_{i0}(t) > \eta\}$. Let $\alpha_1 = \max(\nu_1, \nu_2)$. Let $\tau_i = \sup\{t : 1 - H_{i0}(t) > \eta\}$. Define $\alpha_2 = \min(\tau_1, \tau_2)$. With $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\sigma}})^\top$, the censoring adjusted PEL test statistic is

$$\|R_{\hat{\boldsymbol{\theta}}}\|_{\alpha_1}^{\alpha_2} = \sup_{t \in [\alpha_1, \alpha_2]} |R_{\hat{\boldsymbol{\theta}}}(t)|.$$

Note that the Lagrange multiplier function $\hat{\gamma}_{\hat{\boldsymbol{\theta}}}(t)$, for fixed t, lies between $\hat{D}_2$ and $-\hat{D}_1$, see Equation (3.9). The censoring adjusted PEL is essentially computed as for the uncensored case.

As for the uncensored case $\hat{\mathbb{V}}(t)$, defined by Equation (3.11), has the asymptotic representation $\hat{\mathbb{V}}(t) = A_1(t) + B_1(t) - (A_2(t) + B_2(t)) + o_{\mathbb{P}}(n^{-1/2})$, where $A_i(t)$ and $B_i(t)$ are given by Equation (3.12) and Equation (3.13) respectively. From here on, there is considerable departure from the uncensored representation. Let $\Lambda_{G_i}(t)$ be the cumulative hazard associated with G_i. We have

$$\begin{aligned} A_i(t) &= \frac{\lambda_{i0}(t)}{\sigma_i}\left((\hat{\mu}_i - \mu_i) + \frac{t}{2\sigma_i}(\hat{\sigma}_i^2 - \sigma_i^2)\right) + o_{\mathbb{P}}(n^{-1/2}) \\ &= \frac{\lambda_{i0}(t)}{\sigma_i}\left((1 - t\mu_i/\sigma_i)(\hat{\mu}_i - \mu_i) + \frac{t}{2\sigma_i}(\hat{\psi}_i - \psi_i)\right) + o_{\mathbb{P}}(n^{-1/2}). \quad (3.21) \end{aligned}$$

Let H_i be the distribution function of Z_{ij} and let $H_i^{(0)}$ and $H_i^{(1)}$ be the subdistribution functions of Z_{ij} associated with $\delta_{ij} = 0$ and $\delta_{ij} = 1$ respectively. For $i = 1, 2$, let (Stute, 1995)

$$\begin{aligned} \beta_{i0}(x) &= \exp\left(\int_{-\infty}^{x-} \frac{H_i^{(0)}(dz)}{(1 - H_i(z))}\right) = \exp(-\log(1 - G_i(x-))) = \frac{1}{1 - G_i(x-)}; \\ \beta_{i1}(x) &= \frac{1}{1 - H_i(x)} \int 1_{\{x<w\}} h(w)\gamma_{i0}(w) H_i^{(1)}(dw) = \frac{1}{1 - H_i(x)} \int_x^{\infty} h(w) dF_i(w); \\ \beta_{i2}(x) &= \int\int 1_{\{v<x, v<w\}} h(w)\beta_{i0}(w) H_i^{(0)}(dv) H_i^{(1)}(dw) \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{x\wedge w} \frac{\Lambda_{G_i}(dv)}{1 - H_i(v)} h(w) F_i(dw). \end{aligned}$$

To derive the influence function of $\int h\, d(\hat{F}_i - F_i)$, Stute (1995) introduced the conditions

$$\int \left(\frac{h(x)}{1 - G_i(x-)}\right)^2 H_i^{(1)}(dx) < \infty; \quad (3.22)$$

$$\int |h(x)| \left(\int_{-\infty}^{x-} \frac{\Lambda_{G_i}(dy)}{1 - H_i(y)}\right)^{1/2} F_i(dx) < \infty. \quad (3.23)$$

Assuming (3.22) and (3.23), Stute (1995) derived the influence function of $\int h d\hat{F}_i$, given by

$$h(Z_{ij})\beta_{i0}(Z_{ij})\delta_{ij} + \beta_{i1}(Z_{ij})(1-\delta_{ij}) - \beta_{i2}(Z_{ij}). \tag{3.24}$$

Plugging $\beta_{ik}(x), k = 0, 1, 2,$ into (3.24) and applying $h(x) = x$ or $h(x) = x^2$, the influence functions of $\hat{\mu}_i - \mu_i = \int x\, d\,(\hat{F}_i(x) - F_i(x))$ and $\hat{\psi}_i - \psi_i = \int x^2\, d\,(\hat{F}_i(x) - F_i(x))$ are

$$\begin{aligned} I_{ij}^{(1)} &= \left(\frac{Z_{ij}\delta_{ij}}{1-G_i(Z_{ij}-)} - \mu_i\right) + \frac{1-\delta_{ij}}{1-H_i(Z_{ij})}\int_{Z_{ij}}^{\infty} u\, dF_i(u) \\ &- \int_{-\infty}^{\infty}\int_{-\infty}^{Z_{ij}\wedge u} \frac{d\Lambda_{G_i}(v)}{1-G_i(v)} u d\Lambda_i(u); \\ I_{ij}^{(2)} &= \left(\frac{Z_{ij}^2\delta_{ij}}{1-G_i(Z_{ij}-)} - \psi_i\right) + \frac{1-\delta_{ij}}{1-H_i(Z_{ij})}\int_{Z_{ij}}^{\infty} u^2\, dF_i(u) \\ &- \int_{-\infty}^{\infty}\int_{-\infty}^{Z_{ij}\wedge u} \frac{d\Lambda_{G_i}(v)}{1-G_i(v)} u^2 d\Lambda_i(u). \end{aligned}$$

Note that $I_{ij}^{(1)}$ and $I_{ij}^{(2)}$ are centered. In particular, the remark on page 426 of Stute (1995) implies that the second and third terms of $I_{ij}^{(1)}$ and $I_{ij}^{(2)}$ have identical expectations. Thus

$$\hat{\mu}_i - \mu_i = \frac{1}{n_i}\sum_{j=1}^{n_i} I_{ij}^{(1)}; \qquad \hat{\psi}_i - \psi_i = \frac{1}{n_i}\sum_{j=1}^{n_i} I_{ij}^{(2)}. \tag{3.25}$$

Plugging Equation (3.25) into the RHS of Equation (3.21), we obtain the iid representation

$$A_i(t) = \frac{\lambda_{i0}(t)}{\sigma_i}\left[(1 - t\mu_i/\sigma_i)\left(\frac{1}{n_i}\sum_{j=1}^{n_i} I_{ij}^{(1)}\right) + \frac{t}{2\sigma_i}\left(\frac{1}{n_i}\sum_{j=1}^{n_i} I_{ij}^{(2)}\right)\right] + o_{\mathbb{P}}(n^{-1/2}). \tag{3.26}$$

Furthermore, plugging the RHS of Equation (3.19) into the RHS of Equation (3.13), we obtain

$$B_i(t) = \Lambda_{i0}(t) + \frac{1}{n_i}\sum_{j=1}^{n_i}\left(\frac{\left(I(\tilde{Z}_{ij} \le t, \delta_{ij} = 1)\right)}{\tilde{y}_i(\tilde{Z}_{ij})} - \int_{-\infty}^{\infty} \frac{I(u \le \tilde{Z}_{ij} \wedge t)}{\tilde{y}_i(u)} d\Lambda_{i0}(u)\right) \tag{3.27}$$

Thus, mirroring Equation (3.16), we have the iid representation

$$n^{1/2}\hat{\mathbb{V}}(t) = \frac{1}{\sqrt{\rho}}\left\{n_1^{-1/2}\sum_{j=1}^{n_1} J_{1j}^{(c)}(t)\right\} + \frac{1}{\sqrt{1-\rho}}\left\{n_2^{-1/2}\sum_{j=1}^{n_2} J_{2j}^{(c)}(t)\right\} + o_{\mathbb{P}}(1). \quad (3.28)$$

where, for $j = 1, \ldots, n_i$ and $i = 1, 2$,

$$\begin{aligned} J_{ij}^{(c)}(t) &= \frac{\lambda_{i0}(t)}{\sigma_i}\left\{\left(1 - t\frac{\mu_i}{\sigma_i}\right) I_{ij}^{(1)} + \frac{t}{2\sigma_i} I_{ij}^{(2)}\right\} \\ &+ \left\{\frac{\delta_{ij} I(\tilde{Z}_{ij} \le t)}{\tilde{y}_i(\tilde{Z}_{ij})} - \int_{-\infty}^{\tilde{Z}_{ij}\wedge t} \frac{d\Lambda_{i0}(u)}{\tilde{y}_i(u)}\right\}. \end{aligned} \quad (3.29)$$

Let $V_i^{(c)}(s,t) = \mathbb{E}\left(J_{i1}^{(c)}(s) J_{i1}^{(c)}(t)\right)$. The estimate $\hat{\vartheta}_c^2(t)$ continues to be as in Equation (3.10) with the necessary plug-in adjustments. The large-sample null distribution of $-2R_{\hat{\boldsymbol{\theta}}}(\cdot)$, the censored data extension of the PEL statistic, is stated in Theorem 2. As in the uncensored case, $-2R_{\hat{\boldsymbol{\theta}}}(t)$ is asymptotically equivalent to the square of $n^{1/2}\hat{\mathbb{V}}(t)$ scaled by the reciprocal of $\hat{\vartheta}_c^2(t)$. The proof of Theorem 2 follows exactly like the proof of Theorem 1 and is omitted.

Theorem 2 *Assume conditions (3.22) and (3.23) and that κ_i, the endpoint of the support of H_i, is $+\infty$ and κ_i is not an atom of H_i, where $i = 1, 2$. Suppose that $F_{10}(\cdot) = F_{20}(\cdot)$. Then the process $-2R_{\hat{\boldsymbol{\theta}}}(\cdot)$ converges weakly in $D[\alpha_1, \alpha_2]$ to $\left(\mathbb{V}^{(c)}\right)^2(\cdot)/\vartheta_c^2(\cdot)$, where $\mathbb{V}^{(c)}$ is centered Gaussian with covariance function*

$$\mathrm{Cov}(\mathbb{V}^{(c)}(s), \mathbb{V}^{(c)}(t)) = \frac{V_1^{(c)}(s,t)}{\rho} + \frac{V_2^{(c)}(s,t)}{1-\rho}. \quad (3.30)$$

The continuous mapping theorem guarantees the weak convergence of the PEL test statistic.

3.3 Location-Scale Appropriate Resampling

As was noted in the Introduction section, the bootstrap sample sizes need to be smaller than n_1 and n_2. For each case, uncensored and censored, we describe the bootstrap and employ an adaptive rule for selecting the bootstrap sample sizes. Since two bootstrap sample sizes need to be determined, the adaptive rule will be a two-dimensional extension of that proposed by Bickel and Sakov (2008). Starting with $0 < q < 1$, sample sizes n_1 and n_2 are scaled down successively by employing scaling factors of higher powers of q, inducing the search to traverse along a "diagonal" path. Note that Bickel and Sakov's (2008) search trajectory was horizontal and can only be applied to the one-dimensional setting; see Subramanian (2020).

3.3.1 Uncensored case

Let m_{B_1} and m_{B_2} be the bootstrap sample sizes that are to be determined, see below. Given m_{B_1} and m_{B_2}, obtain $\{X^*_{ij}, j = 1, \ldots, m_{B_i}, i = 1, 2\}$ by resampling from the EDF of $\{X_{11}, ..., X_{1,n_1}\}$. Then compute the RHS of Equation (3.8) to obtain $R^*_{\hat{\theta}^*}$, the bootstrap log LR, where $\hat{\boldsymbol{\theta}}^*$ is the bootstrap estimate of $\boldsymbol{\theta}$.

To obtain m_{B_1} and m_{B_2}, write $\boldsymbol{m}_B = (m_{B_1}, m_{B_2})^\top$. Let $L_{\boldsymbol{n}} = \|R_{\hat{\boldsymbol{\theta}}}\|^{\alpha_2}_{\alpha_1}$ be the test statistic, and $L_{\boldsymbol{m}_B,\boldsymbol{n}} = \|R^*_{\hat{\theta}^*}\|^{\alpha_2}_{\alpha_1}$ be its bootstrap version, where $\boldsymbol{n} = (n_1, n_2)^\top$. Fix $0 < q < 1$ and let $[\alpha]$ be the smallest integer greater than or equal to α. The steps are:

1. Consider sequence $\boldsymbol{m}_{B_k} = [m_{B_{1k}}, m_{B_{2k}}]^\top$, where $m_{B_{ik}} = [q^k n_i], i = 1, 2, k = 0, 1, 2, ...$
2. For each $\boldsymbol{m}_{B_k}$, obtain B bootstrap replicates of $L_{\boldsymbol{m}_{B_k},\boldsymbol{n}}$. Let $F_{L_{B_k,\boldsymbol{n}}}$ be the EDF computed from the B values of $L_{\boldsymbol{m}_{B_k},\boldsymbol{n}}$.
3. Compute $\hat{k} = \text{argmin}_k(\sup_x |F_{L_{B_k,\boldsymbol{n}}}(x) - F_{L_{B_{k+1},\boldsymbol{n}}}(x)|)$, the index minimizing the Kolmogorov–Smirnov (KS) distance. If the distance is minimized for multiple k, pick the largest one.
4. The bootstrap sample size vector is then $\hat{\boldsymbol{m}}_B \equiv \boldsymbol{m}_{B_{\hat{k}}} = [m_{B_{1\hat{k}}}, m_{B_{2\hat{k}}}]^\top \equiv [\hat{m}_{B_1}, \hat{m}_{B_2}]^\top$.

The bootstrap PEL is computed 1,000 times and the critical value is the 950-th ordered value.

3.3.2 Censored case

We generate $\{(Z^*_{1j}, \delta^*_{1j}), j = 1, ..., m_{B_1}\}$ from the EDF of $\{(Z_{11}, \delta_{11}), ..., (Z_{1,n_1}, \delta_{1,n_1})\}$, but not generate $\{(Z^*_{2j}, \delta^*_{2j}), j = 1, ..., m_{B_2}\}$ from the EDF of $\{(Z_{21}, \delta_{21}), ..., (Z_{2,n_2}, \delta_{2,n_2})\}$, since it may not satisfy the null hypothesis. Also, we don't generate $\{(Z^*_{2j}, \delta^*_{2j}), j = 1, ..., m_{B_2}\}$ from the EDF of the first sample since the censoring distributions could be different. Therefore,

1. Generate $X^*_{1j}, j = 1, ..., m_{B2}$, from $\hat{F}_1$, the KM estimator of F_1. Realizations greater than the highest uncensored value were assigned an arbitrarily large value. The rationale is that such an action would not affect LR calculations over $[\alpha_1, \alpha_2]$.

2. Compute $X^*_{2j} = \hat{\mu}_2 + (X^*_{1j} - \hat{\mu}_1)\hat{\sigma}_2/\hat{\sigma}_1, j = 1, \ldots, m_{B_2}$, where $\hat{\mu}_i$ and $\hat{\sigma}_i$ are the estimators of μ_i and σ_i respectively [see Equation (3.20)], and are computed from the original data.

3. Generate $C^*_{2j}, j = 1, \ldots, m_{B_2}$, from $\hat{G}_2$, the KM estimator of G_2. Realizations greater than the highest censored value were assigned an arbitrarily large value. This action would not affect LR calculations over $[\alpha_1, \alpha_2]$.

4. Obtain $\{(Z^*_{2j}, \delta^*_{2j}), j = 1, ..., m_{B_2}\}$, where $Z^*_{2j} = \min(X^*_{2j}, C^*_{2j})$ and $\delta^*_{2j} = I(X^*_{2j} \leq C^*_{2j})$.

Adaptive selection of bootstrap sample sizes was carried out as for the uncensored case.

CHAPTER 4

CENSORED ECF TEST

In this chapter we extend Hall et al's (2013) approach for censored data. The extension is non-trivial since one would need plug-ins for the means and standard deviations. In the proposed extension detailed in this chapter, the plug-ins will be based on KM integrals (Stute, 1995), described in Chapter 3.

Recall that, for $k = 1, 2$, $\hat{F}_k(x) = 1 - \hat{S}_k(x)$, and $\hat{S}_k(x)$ is the KM estimator of $S_k(x)$ based on the original (non-standardized) data. The estimated characteristic function of the standardized variable $U_k = (X_k - \mu_k)/\sigma_k$ is

$$\begin{aligned} \hat{\chi}_k(t) &= \int_{\mathbb{R}} \exp\left\{ it \left(\frac{x - \hat{\mu}_k}{\hat{\sigma}_k} \right) \right\} d\hat{F}_k(x) \\ &= \int \cos\left(t \left(\frac{x - \hat{\mu}_k}{\hat{\sigma}_k} \right) \right) d\hat{F}_k(x) + i \int_{\mathbb{R}} \sin\left(t \left(\frac{x - \hat{\mu}_k}{\hat{\sigma}_k} \right) \right) d\hat{F}_k(x). \end{aligned} \tag{4.1}$$

The important insight is to express the estimated characteristic function as an integral, with the KM estimator as the integrating measure. Additionally, as already stated, the plug-ins are based on Stute's (1995) KM integrals.

With these changes, the proposed test statistic will be the same as in Hall et al. (2013) and is

$$L_{n_1,n_2} := \frac{n_1 n_2}{n_1 + n_2} \int |\hat{\chi}_1(t) - \hat{\chi}_2(t)|^2 w(t) dt, \tag{4.2}$$

where $w(t)$ is a specified weight function that must be chosen to calculate the test statistic. According to Hall et al. (2013) a popular choice of weight function is $w(t) = exp(-ht^2)$. The other boxcar weight function, $w(t) = 1, \quad |t| \leq \pi/2h$, was another alternative. A third choice was $w(t) = (1 - |th|)^2, \quad |t| \leq 1/h$; see Hall et al. (2013) for more insight on selecting $w(t)$.

4.1 ECF Censored Case

We obtain an asymptotic iid representation for L_{n_1,n_2}. We use Taylor expansions of the sine and cosine functions about (μ_k, σ_k) and disregard second order terms. From the sequel it will be clear that those second order terms contribute remainder terms which are $o_{\mathbb{P}}(n^{-1/2})$. Therefore, considering the dominant (linear) terms, we have

$$\begin{aligned}
\cos\left(t\left(\frac{X_k-\hat{\mu}_k}{\hat{\sigma}_k}\right)\right) &= \cos\left(t\left(\frac{X_k-\mu_k}{\sigma_k}\right)\right) + (\hat{\mu}_k-\mu_k)\frac{t}{\sigma_k}\sin\left(t\left(\frac{X_k-\mu_k}{\sigma_k}\right)\right) \\
&\quad + (\hat{\sigma}_k-\sigma_k)\ \frac{t}{\sigma_k}\frac{X_k-\mu_k}{\sigma_k}\sin\left(t\left(\frac{X_k-\mu_k}{\sigma_k}\right)\right) \\
&= \cos\left(t\left(\frac{X_k-\mu_k}{\sigma_k}\right)\right) - i^2(\hat{\mu}_k-\mu_k)\frac{t}{\sigma_k}\sin\left(t\left(\frac{X_k-\mu_k}{\sigma_k}\right)\right) \\
&\quad - (\hat{\sigma}_k-\sigma_k)\frac{t}{\sigma_k}\frac{\partial}{\partial t}\left\{\cos\left(t\left(\frac{X_k-\mu_k}{\sigma_k}\right)\right)\right\}. \qquad (4.3)
\end{aligned}$$

Likewise, we also have

$$\begin{aligned}
\sin\left(t\left(\frac{X_k-\hat{\mu}_k}{\hat{\sigma}_k}\right)\right) &= \sin\left(t\left(\frac{X_k-\mu_k}{\sigma_k}\right)\right) - (\hat{\mu}_k-\mu_k)\frac{t}{\sigma_k}\cos\left(t\left(\frac{X_k-\mu_k}{\sigma_k}\right)\right) \\
&\quad - (\hat{\sigma}_k-\sigma_k)\ \frac{t}{\sigma_k}\frac{X_k-\mu_k}{\sigma_k}\cos\left(t\left(\frac{X_k-\mu_k}{\sigma_k}\right)\right) \\
&= \sin\left(t\left(\frac{X_k-\mu_k}{\sigma_k}\right)\right) - (\hat{\mu}_k-\mu_k)\frac{t}{\sigma_k}\cos\left(t\left(\frac{X_k-\mu_k}{\sigma_k}\right)\right) \\
&\quad - (\hat{\sigma}_k-\sigma_k)\frac{t}{\sigma_k}\frac{\partial}{\partial t}\left\{\sin\left(t\left(\frac{X_k-\mu_k}{\sigma_k}\right)\right)\right\}. \qquad (4.4)
\end{aligned}$$

It follows that, modulus remainder terms which are $o_{\mathbb{P}}(n^{-1/2})$,

$$\begin{aligned}
\exp\left\{it\left(\frac{X_k-\hat{\mu}_k}{\hat{\sigma}_k}\right)\right\} &= \cos\left(t\left(\frac{X_k-\hat{\mu}_k}{\hat{\sigma}_k}\right)\right) + i\sin\left(t\left(\frac{X_k-\hat{\mu}_k}{\hat{\sigma}_k}\right)\right) \\
&= \cos\left(t\left(\frac{X_k-\hat{\mu}_k}{\hat{\sigma}_k}\right)\right) + i\sin\left(t\left(\frac{X_k-\hat{\mu}_k}{\hat{\sigma}_k}\right)\right) \\
&\quad - i(\hat{\mu}_k-\mu_k)\frac{t}{\sigma_k}\left\{\cos\left(t\left(\frac{X_k-\mu_k}{\sigma_k}\right)\right) + i\sin\left(t\left(\frac{X_k-\mu_k}{\sigma_k}\right)\right)\right\} \\
&\quad - (\hat{\sigma}_k-\sigma_k)\frac{t}{\sigma_k}\frac{\partial}{\partial t}\left[\cos\left(t\left(\frac{X_k-\mu_k}{\sigma_k}\right)\right) + i\sin\left(t\left(\frac{X_k-\mu_k}{\sigma_k}\right)\right)\right] \\
&= \exp\left(it\left(\frac{X_k-\hat{\mu}_k}{\hat{\sigma}_k}\right)\right) - i(\hat{\mu}_k-\mu_k)\frac{t}{\sigma_k}\exp\left(it\left(\frac{X_k-\mu_k}{\sigma_k}\right)\right) \\
&\quad - (\hat{\sigma}_k-\sigma_k)\frac{t}{\sigma_k}\frac{\partial}{\partial t}\left[\exp\left(it\left(\frac{X_k-\mu_k}{\sigma_k}\right)\right)\right]. \qquad (4.5)
\end{aligned}$$

The estimated characteristic function given by Equation (4.1) has the asymptotic representation

$$\begin{aligned}\hat{\chi}_k(t) &= \int_{\mathbb{R}} \exp\left\{it\left(\frac{x-\mu_k}{\sigma_k}\right)\right\} d\hat{F}_k(x) \\ &\quad - i\left(\hat{\mu}_k-\mu_k\right)\frac{t}{\sigma_k}\int_{\mathbb{R}} \exp\left\{it\left(\frac{x-\mu_k}{\sigma_k}\right)\right\} d\hat{F}_k(x) \\ &\quad - \left(\hat{\sigma}_k^2-\sigma_k^2\right)\frac{t}{2\sigma_k^2}\int_{\mathbb{R}} \frac{\partial}{\partial t}\left[\exp\left(it\left(\frac{x-\mu_k}{\sigma_k}\right)\right)\right] d\hat{F}_k(x) + o_{\mathbb{P}}(n^{-1/2}). \quad (4.6)\end{aligned}$$

The integrals in the second and third expressions on the right hand side (RHS) of Equation (4.6) converge in probability to $\chi_k(t)$ and $\chi_k'(t)$ respectively (Csörgő, 1981; Stute, 1995). Applying the delta method, it follows that the third expression of Equation (4.6) is

$$\left[(\hat{\psi}_k-\psi_k) - 2\mu_k(\hat{\mu}_k-\mu_k)\right]\frac{t}{2\sigma_k^2}\chi_k'(t) + o_{\mathbb{P}}(n^{-1/2}). \quad (4.7)$$

It follows from Equation (4.6) and expression (4.7) that

$$\begin{aligned}\hat{\chi}_k(t) &= \chi_k(t) + \int_{\mathbb{R}} \exp\left\{it\left(\frac{x-\mu_k}{\sigma_k}\right)\right\} d\left\{\hat{F}_k(x)-F_k(x)\right\} \\ &- (\hat{\mu}_k-\mu_k)\frac{t}{\sigma_k}\left(i\chi_k(t)+\frac{\mu_k}{\sigma_k}\chi_k'(t)\right) \\ &\quad - \left(\hat{\psi}_k-\psi_k\right)\frac{t}{2\sigma_k^2}\chi_k'(t) + o_{\mathbb{P}}(n^{-1/2}) \\ &\equiv \chi_k(t) + I_{k1}(t) + I_{k2}(t) + I_{k3}(t), \quad (4.8)\end{aligned}$$

where $I_{k2}(t)$ and $I_{k3}(t)$ are the third and fourth expressions on the RHS of Equation (4.8) and

$$\begin{aligned}I_{k1}(t) &= \int_{\mathbb{R}} \exp\left\{it\left(\frac{x-\mu_k}{\sigma_k}\right)\right\} d\left\{\hat{F}_k(x)-F_k(x)\right\} \\ &= \int_{\mathbb{R}} \cos\left\{t\left(\frac{x-\mu_k}{\sigma_k}\right)\right\} d(\hat{F}_k(x)-F_k(x)) \\ &+ i\int_{\mathbb{R}} \sin\left\{t\left(\frac{x-\mu_k}{\sigma_k}\right)\right\} d(\hat{F}_k(x)-F_k(x)).\end{aligned}$$

The influence function of the KM integral process $\int h\, d\,(\hat{F}_k - F_k)$ is given by Equation (1.7) and Equation (1.8) of Stute (1995). When $h(x) = \cos(x)$ and $h(x) = \sin(x)$, the influence functions connected with the asymptotic representations of the two expressions on the RHS above can be obtained. Thus $I_{k1}(t) = \frac{1}{n_k}\sum_{j=1}^{n_k} I_{k11}^{(j)}(t) + \frac{i}{n_k}\sum_{j=1}^{n_k} I_{k12}^{(j)}(t) + o_{\mathbb{P}}(n^{-1/2})$, where

$$\begin{aligned}
I_{k11}^{(j)}(t) &= \left(\frac{\cos\left\{t\left(\frac{Z_{kj}-\mu_k}{\sigma_k}\right)\right\}\delta_{kj}}{1-G_k(Z_{kj}-)} - E\left(\cos\left\{t\left(\frac{Z_{kj}-\mu_k}{\sigma_k}\right)\right\}\right)\right) \\
&+ \frac{1-\delta_{kj}}{1-H_k(Z_{kj})}\int_{Z_{kj}}^{\infty}\cos\left\{t\left(\frac{u-\mu_k}{\sigma_k}\right)\right\}dF_k(u) \\
&- \int_{-\infty}^{\infty}\int_{-\infty}^{Z_{kj}\wedge u}\frac{d\Lambda_{G_k}(v)}{1-G_k(v)}\cos\left\{t\left(\frac{u-\mu_k}{\sigma_k}\right)\right\}d\Lambda_k(u); \\
I_{k12}^{(j)}(t) &= \left(\frac{\sin\left\{t\left(\frac{Z_{kj}-\mu_k}{\sigma_k}\right)\right\}\delta_{kj}}{1-G_k(Z_{kj}-)} - E\left(\sin\left\{t\left(\frac{Z_{kj}-\mu_k}{\sigma_k}\right)\right\}\right)\right) \\
&+ \frac{1-\delta_{kj}}{1-H_k(Z_{kj})}\int_{Z_{kj}}^{\infty}\sin\left\{t\left(\frac{u-\mu_k}{\sigma_k}\right)\right\}dF_k(u) \\
&- \int_{-\infty}^{\infty}\int_{-\infty}^{Z_{kj}\wedge u}\frac{d\Lambda_{G_k}(v)}{1-G_k(v)}\sin\left\{t\left(\frac{u-\mu_k}{\sigma_k}\right)\right\}d\Lambda_k(u).
\end{aligned}$$

Note that $I_{k11}^{(j)}(t)$ and $I_{k12}^{(j)}(t)$ are centered. In particular, the remark on page 426 of Stute (1995) implies that the second and third terms of $I_{k11}^{(j)}(t)$ and $I_{k12}^{(j)}(t)$ have identical expectations. However, we can combine the two influence functions and obtain

$$\begin{aligned}
I_{k11}^{(j)}(t) + I_{k12}^{(j)}(t) &= \left(\frac{\exp\left\{it\left(\frac{Z_{kj}-\mu_k}{\sigma_k}\right)\right\}\delta_{kj}}{1-G_k(Z_{kj}-)} - E\left(\exp\left\{it\left(\frac{Z_{kj}-\mu_k}{\sigma_k}\right)\right\}\right)\right) \\
&+ \frac{1-\delta_{kj}}{1-H_k(Z_{kj})}\int_{Z_{kj}}^{\infty}\exp\left\{it\left(\frac{u-\mu_k}{\sigma_k}\right)\right\}dF_k(u) \\
&- \int_{-\infty}^{\infty}\int_{-\infty}^{Z_{kj}\wedge u}\frac{d\Lambda_{G_k}(v)}{1-G_k(v)}\exp\left\{it\left(\frac{u-\mu_k}{\sigma_k}\right)\right\}d\Lambda_k(u)) \;\equiv\; I_{k1}^{(j)}(t).
\end{aligned}$$

The influence functions of $\hat{\mu}_k - \mu_k = \int x\, d\left(\hat{F}_k(x) - F_k(x)\right)$ and $\hat{\psi}_k - \psi_k = \int x^2\, d\left(\hat{F}_k(x) - F_k(x)\right)$ were determined in Chapter 3. They are given by

$$\begin{aligned}
I_{k2}^{(j)} &= \left(\frac{Z_{kj}\delta_{kj}}{1 - G_k(Z_{kj}-)} - \mu_k\right) + \frac{1-\delta_{kj}}{1 - H_k(Z_{kj})}\int_{Z_{kj}}^{\infty} u\, dF_k(u) \\
&- \int_{-\infty}^{\infty}\int_{-\infty}^{Z_{kj}\wedge u} \frac{d\Lambda_{G_k}(v)}{1 - G_k(v)} u d\Lambda_k(u); \\
I_{k3}^{(j)} &= \left(\frac{Z_{kj}^2\delta_{kj}}{1 - G_k(Z_{kj}-)} - \psi_k\right) + \frac{1-\delta_{kj}}{1 - H_k(Z_{kj})}\int_{Z_{kj}}^{\infty} u^2 dF_k(u) \\
&- \int_{-\infty}^{\infty}\int_{-\infty}^{Z_{kj}\wedge u} \frac{d\Lambda_{G_k}(v)}{1 - G_k(v)} u^2 d\Lambda_k(u).
\end{aligned}$$

Thus

$$\hat{\mu}_k - \mu_k = \frac{1}{n_k}\sum_{j=1}^{n_k} I_{k2}^{(j)}; \qquad \hat{\psi}_k - \psi_k = \frac{1}{n_k}\sum_{j=1}^{n_k} I_{k3}^{(j)}. \tag{4.9}$$

Therefore we have

$$\begin{aligned}
I_{k2}(t) &= -\frac{t}{\sigma_k}\left(i\chi_k(t) + \frac{\mu_k}{\sigma_k}\chi_k'(t)\right)\left(\frac{1}{n_k}\sum_{j=1}^{n_k} I_{k2}^{(j)}\right) + o_{\mathbb{P}}(n^{-1/2}); \\
I_{k3}(t) &= -\frac{t}{\sigma_k^2}\chi_k'(t)\left(\frac{1}{n_k}\sum_{j=1}^{n_k} I_{k3}^{(j)}\right) + o_p(n^{-1/2}).
\end{aligned}$$

Plugging the above results into the RHS of Equation (4.6), we obtain the iid representation

$$\begin{aligned}
\hat{\chi}_k(t) - \chi_k(t) &= \frac{1}{n_k}\sum_{j=1}^{n_k}\left[I_{k1}^{(j)}(t) + -\frac{t}{\sigma_k}\left(i\chi_k(t) + \frac{\mu_k}{\sigma_k}\chi_k'(t)\right) I_{k2}^{(j)} - \frac{t}{\sigma_k^2}\chi_k'(t) I_{k3}^{(j)}\right] \\
&\quad + o_p(n^{-1/2}).
\end{aligned} \tag{4.10}$$

Recall that we work with independent samples. We now have

$$\begin{aligned}
\hat{\chi}_1(t) - \hat{\chi}_2(t) &= \hat{\chi}_1(t) - \chi_1(t) - \{\hat{\chi}_2(t) - \chi_2(t)\} + \chi_1(t) - \chi_2(t) \\
&= \frac{1}{n_1}\sum_{j=1}^{n_1}\left[I_{11}^{(j)}(t) - \frac{t}{\sigma_1}\left(i\chi_1(t) + \frac{\mu_1}{\sigma_1}\chi_1'(t)\right) I_{12}^{(j)} - \frac{t}{\sigma_1^2}\chi_1'(t) I_{13}^{(j)}\right] \\
&\quad - \frac{1}{n_2}\sum_{j=1}^{n_2}\left[I_{21}^{(j)}(t) - \frac{t}{\sigma_2}\left(i\chi_2(t) + \frac{\mu_2}{\sigma_2}\chi_2'(t)\right) I_{22}^{(j)} - \frac{t}{\sigma_2^2}\chi_2'(t) I_{23}^{(j)}\right] \\
&\quad + \chi_1(t) - \chi_2(t) + o_p(n^{-1/2}).
\end{aligned} \tag{4.11}$$

Under the null hypothesis $\chi_1(t) = \chi_2(t)$. Let ρ be the limit of $n_1/(n_1+n_2)$ as $n_1 \to \infty$ and $n_2 \to \infty$. Then

$$\begin{aligned}\left\{\frac{n_1 n_2}{n_1+n_2}\right\}^{1/2} \hat{\mathbb{V}}^{(\mathrm{cf})}(t) &:= \left\{\frac{n_1 n_2}{n_1+n_2}\right\}^{1/2} (\hat{\chi}_1(t) - \hat{\chi}_2(t)) \\ &= (\sqrt{1-\rho}) n_1^{-1/2} \sum_{j=1}^{n_1} J_{1j}^{(\mathrm{cf})}(t) - (\sqrt{\rho}) n_2^{-1/2} \sum_{j=1}^{n_2} J_{2j}^{(\mathrm{cf})}(t) + o_{\mathbb{P}}(1),\end{aligned}$$

where

$$J_{kj}^{(cf)}(t) \;=\; I_{k1}^{(j)}(t) - \frac{t}{\sigma_k}\left(i\chi_k(t) + \frac{\mu_k}{\sigma_k}\chi_k'(t)\right) I_{k2}^{(j)} - \frac{t}{\sigma_k^2}\chi_k'(t) I_{k3}^{(j)}. \qquad (4.12)$$

Let $V_k^{(\mathrm{cf})}(s,t) = \mathbb{E}\left(J_{k1}^{(cf)}(s) J_{k1}^{(cf)}(t)\right)$. Then $\{n_1 n_2/(n_1+n_2)\}^{1/2}\, \hat{\mathbb{V}}^{(\mathrm{cf})}(\cdot)$ converges in distribution to $\mathcal{G}(\cdot)$, a zero-mean Gaussian process with covariance function at (s,t) given by

$$\mathrm{Cov}(\mathcal{G}(s), \mathcal{G}(t)) = (1-\rho) V_1^{(\mathrm{cf})}(s,t) + \rho V_2^{(\mathrm{cf})}(s,t).$$

The proposed censored test statistic L_{n_1,n_2} converges in distribution to the random variable

$$L = \int |\mathcal{G}(t)|^2 w(t) dt.$$

Proof of this follows the steps given in Hall et al. (2013).

The censored ECF test statistic L_{n_1,n_2} is not asymptotically distribution free. Therefore, it is a challenge to obtain critical values needed for carrying out the test. The full bootstrap that Hall et al. (2013) proposed did not work well for the censored case. As is explained in Chapter 3, we are required to use smaller bootstrap sample sizes. The data-driven procedure given in Subramanian (2020) that adapts Bickel and Sakov's (2008) one-sample proposal yielded a reasonably accurate choice for the bootstrap sample size vector.

For clarity, we once again describe the algorithm to obtain the bootstrap sample size vector. To obtain m_{B1} and m_{B2}, the bootstrap sample sizes, write $\boldsymbol{m}_B = (m_{B1}, m_{B2})^\top$. Let $L_{\boldsymbol{n}} = \hat{L}_{n_1,n_2}$ be the test statistic, and $L_{\boldsymbol{m}_B,\boldsymbol{n}} = \hat{L}^*_{n_1,n_2}$ be its bootstrap version, where $\boldsymbol{n} = (n_1, n_2)^\top$. Fix $0 < q < 1$ and let $[\alpha]$ be the smallest integer greater than or equal to α. The steps are:

1. Consider sequence $\boldsymbol{m}_{Bk} = [m_{B1k}, m_{B2k}]^\top$, where $m_{Bik} = [q^k n_i], i = 1, 2, k = 0, 1, 2, ...$
2. For each $\boldsymbol{m}_{Bk}$, obtain B bootstrap replicates of $L_{\boldsymbol{m}_{Bk},\boldsymbol{n}}$. Let $F_{L_{k,\boldsymbol{n}}}$ be the EDF computed from the B values of $L_{\boldsymbol{m}_{Bk},\boldsymbol{n}}$.
3. Compute $\hat{k} = \text{argmin}_k \left(\sup_x |F_{L_{k,\boldsymbol{n}}}(x) - F_{L_{k+1,\boldsymbol{n}}}(x)|\right)$. If the supnorm within the parenthesis is minimized for more than one k, then pick the largest as $\hat{k}$.
4. The bootstrap sample size vector is then $\hat{\boldsymbol{m}}_B = [\hat{m}_{B1}, \hat{m}_{B2}]^\top$.

The bootstrap PEL is computed 1,000 times and the critical value is the 950-th ordered value. Using the identified bootstrap sample sizes, we found that the realized levels are closer to the nominal 5% level.

CHAPTER 5

NUMERICAL STUDIES

We first report a power comparison study of the uncensored ECF and PEL tests. We then report simulation results for the uncensored PEL test. The power comparison study of the censored PEL an d ECF tests will be reported next and at the end the simulation results for the censored PEL and ECF tests are reported.

5.1 Power Comparisons for the Uncensored Case

We carried out the first power study as in Hall et al. (2013). The X_1 and X_2 samples, each of size 250, were generated from the same location-scale family, namely the t distribution with 5 degrees of freedom. For each of 10,000 samples, the PEL and ECF statistics were computed and were arranged in increasing order. The 9,500-th ordered value was chosen as the critical value. Then, 1,000 X_1 samples were generated as they were when determining the critical value but the X_2 samples were drawn from a skew-t distribution with 5 degrees of freedom, and the skewness parameter was taken over a range of values between $0.0 \le \delta \le 0.99$. The PEL and ECF test statistics were calculated again for each of these 1,000 samples. The weight function $w(t) = e^{-t^2}$ was used to calculate the ECF test. The estimated power at the 5% level of significance for each test is the proportion of test statistics that exceeded the respective critical values. The estimated power was plotted against skewness value and reported in Fig. 5.1. The ECF performs slightly better when $\delta \in (0, 0.5)$. For $\delta \in (0.5, 1)$, the PEL performs slightly better. Thus there seems to be no clear winner.

The critical value for the first power comparison study is obtained for a specified distribution under the null hypothesis. Since, in practice, no such distribution will be recognized, one might wish to investigate test efficacy based on a more realistic power comparison, namely that which makes use of the proposed bootstrap procedure.

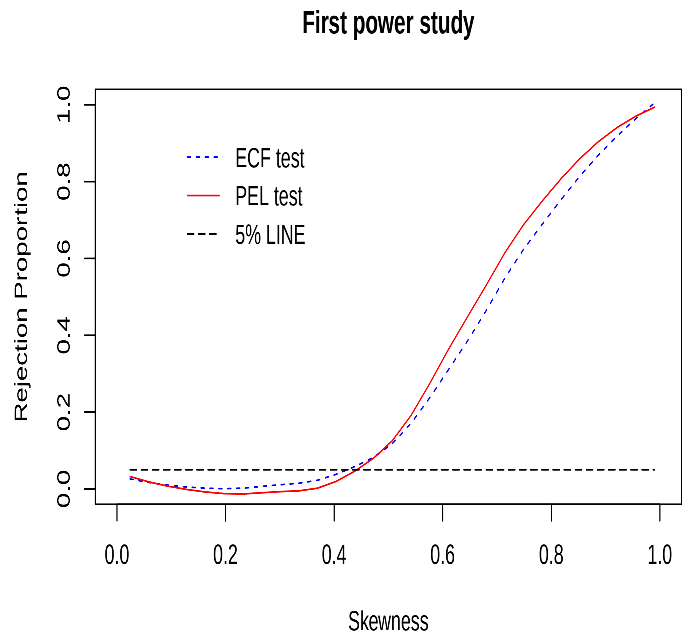

Figure 5.1 First power comparison study of the uncensored PEL and ECF tests.

Therefore, a second power comparison study was done as follows. The X_1 sample of size 250 was generated from the t distribution with 5 degrees of freedom. The X_2 sample of size 250 was generated from the skew-t distribution with 5 degrees of freedom, with the skewness parameter taken over a range of values between $0.0 \leq \delta \leq 0.99$. The PEL and ECF statistics were computed for each of such 1,000 simulated data sets. Approximate ECF and PEL critical values for each data set were also obtained — Hall et al.'s (2013) bootstrap for the ECF test and the proposed model-appropriate bootstrap for the PEL test. The ECF and PEL tests based rejection indicators were obtained for each data set. The rejection proportions give the respective empirical powers. The weight function $w(t) = e^{-t^2}$ was used for the ECF test, and two values of q (0.9 and 0.85), were investigated for the PEL test. The diagonal trajectory for selecting the bootstrap sample sizes (Subsection 3.3.1) was applied. The empirical power was plotted against skewness value and reported in Fig. 5.2. The PEL test appears to show improved performance over the ECF test.

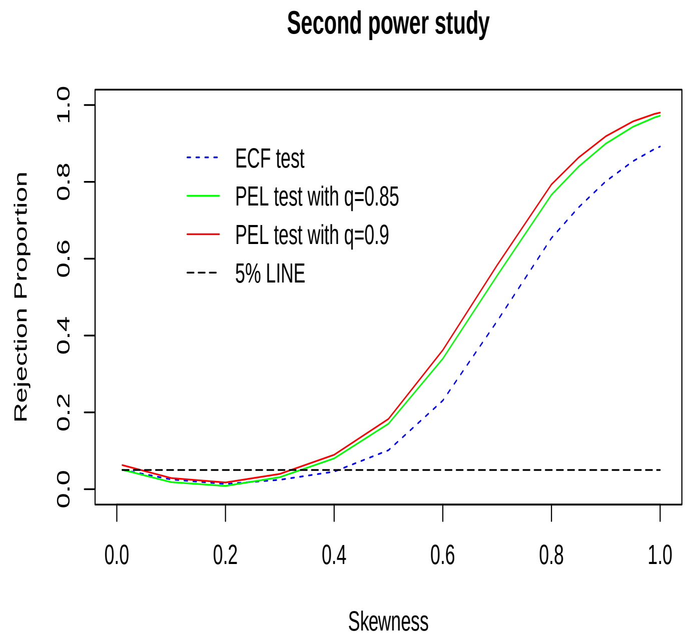

Figure 5.2 Second power comparison study of uncensored PEL and ECF (diagonal path).

For the second power study, we also applied an alternate search trajectory to zero-in on the bootstrap sample sizes m_{B_1} and m_{B_2}. Taking $q = 0.9$, we first navigated through a one-dimensional horizontal path $\boldsymbol{m}_{B_{kl}} = [m_{B_{1k}}, m_{B_{2l}}]^\top$, where $m_{B_{1k}} = [q^k n_1], m_{B_{2l}} = [q^l n_2]$, by fixing index k and running through index l $(k = 0, \ldots, 8, l = 0, \ldots, 8)$. The bootstrap sample size vector for each k that minimized the KS distance between successive EDFs (see Section 3.3.1) is $\boldsymbol{m}_{B_{.k}} = [m_{B_{1k}}, \hat{m}_{B_{2k}}]^\top, k = 0, \ldots 8$. In the next step we performed a "downward" search trajectory along the sequence $\boldsymbol{m}_{B_{.k}}, k = 0, \ldots, 8$ and zoomed-in on the bootstrap sample size $\hat{\boldsymbol{m}}_B = [\hat{m}_{B_1}, \hat{m}_{B_2}]^\top$ that minimized the KS distance between successive EDF's when traversing downward along the sequence. The PEL power curve was plotted and compared with that of ECF, see Fig. 5.3. The PEL test again appears to show superiority over the ECF test.

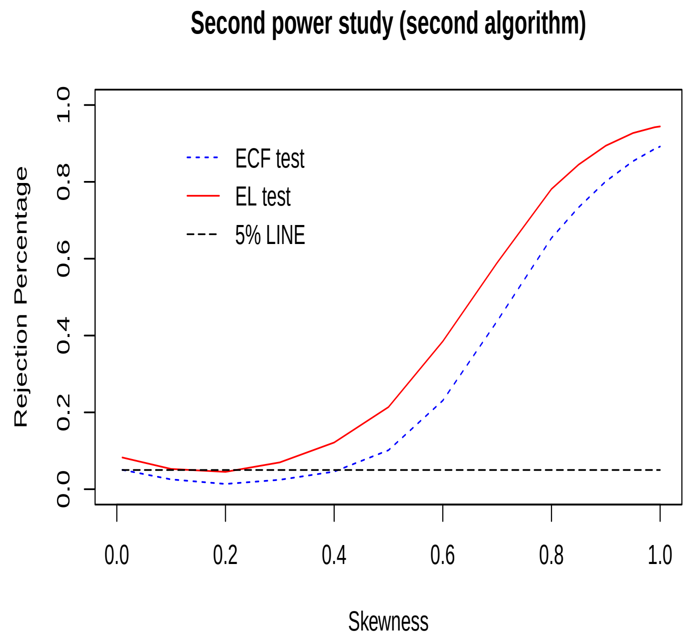

Figure 5.3 Second power comparison study (horizontal and downward trajectory).

5.2 Level Studies for PEL Uncensored Case

The steps for obtaining realized levels of the proposed test are: (i) Generate X_1 and X_2 samples from populations satisfying the null hypothesis. (ii) Compute $\|R_{\hat{\theta}}\|_{\alpha_1}^{\alpha_2}$, the PEL statistic. (iii) As explained in Section 5.1, apply adaptive rule to determine the bootstrap sample sizes m_{B1} and m_{B2}, and compute the bootstrap PEL from each of the 1,000 bootstrap resamples. The proportion of rejections in 1,000 replications gives the realized level.

For our simulations, we used three different populations, namely, normal, skew-normal with skew parameter δ being 0.99 and 0.95, and logistic. Sample sizes, indicated by n, were the same for both the samples. A number of values for the factor q were investigated. Realized levels, $\hat{p}$, as well as standard errors, $(\hat{p}(1-\hat{p})/1000)^{1/2}$, are reported.

For the normal location-scale study, $X_1 \sim N(2, 0.25)$ and $X_2 \sim N(4, 0.36)$, which satisfied the null hypothesis. Table 5.1 gives the estimated levels for different sample sizes. It can be seen that all the realized levels are reasonably close to the nominal level of 5%.

Table 5.1 Normal Study for Uncensored PEL with $q = 0.93$

n	Realized level	Standard error
100	0.051	0.0069
200	0.053	0.0070
250	0.055	0.0072

For logistic location-scale study, we generated the X_1 and X_2 samples from a logistic distribution with location and scale parameters 0 and 1 respectively, denoted by LG$(0, 1)$. Table 5.2 reports the realized levels. Here too, they were close to the nominal 5%.

Table 5.2 Logistic Study for Uncensored PEL with $q = 0.91$

n	Realized level	Standard error
100	0.051	0.0069
200	0.053	0.0070
250	0.047	0.0080

For our last simulation study, we carried out three different simulations with the skew-normal (SN) distribution. For the first, the X_1 and X_2 samples were both SN with location and scale parameters 0 and 1 respectively, and shape parameter 27.85 yielding skewness $\delta = 0.99$. For the second and third cases, the distribution for the X_1 sample remained unchanged but the X_2 sample was from SN with location and scale parameters 2 and 4 respectively. The skewness for the second study was $\delta = 0.99$ and for the third study it was $\delta = 0.95$ (shape parameter 9.34). The results are reported in Table 5.3. In all cases, the realized levels were close to the nominal 5%.

Table 5.3 Skew-normal Study for Uncensored PEL

Study no.	n	q	Realized level	Standard error
I	100	0.92	0.055	0.0072
	200	0.92	0.053	0.0071
	250	0.92	0.051	0.0069
II	100	0.91	0.043	0.0064
	200	0.91	0.048	0.0069
	250	0.91	0.049	0.0068
III	100	0.91	0.049	0.0068
	200	0.91	0.046	0.0066
	250	0.91	0.051	0.0069

5.3 Level Studies for PEL Censored Case

We used normal, logistic, and extreme value families for our level studies. We carried out two simulations for the normal case, I and II. For normal study I, we took $X_1 \sim N(0,1)$ and $C_1 \sim N(0.75, 0.81)$, giving a censoring rate (CR) of about 29%. We also took $X_2 \sim N(2,9)$ and $C_2 \sim N(4,25)$ giving about 37% CR. For normal study II, we took $X_1 \sim N(0,1)$ and $C_1 \sim N(0.9,1)$, giving about 25% CR. We also took $X_2 \sim N(2,9)$ and $C_2 \sim N(10,9)$ giving about 25% CR. The realized levels, presented in Table 5.4, are close to the nominal 5% level.

For logistic location-scale, we carried out two different simulations. For the logistic study I, $X_1 \sim \mathrm{LG}(0,1)$ and $C_1 \sim \mathrm{LG}(0.5, 0.25)$, giving CR of about 39%. Also, $X_2 \sim \mathrm{LG}(1,4)$ and $C_2 \sim \mathrm{LG}(3.5, 0.56)$, giving CR of about 24%.

Table 5.4 Normal Study for Censored PEL

Study no.	n	q	Realized level	Standard error
I	100	0.84	0.047	0.0066
	200	0.84	0.047	0.0066
	250	0.84	0.049	0.0068
II	100	0.87	0.048	0.0068
	200	0.87	0.049	0.0068
	250	0.87	0.045	0.0065

For logistic study II, $X_1 \sim \mathrm{LG}(0,1)$ and $C_1 \sim \mathrm{LG}(1.5,1)$, giving CR of about 25%. Also, $X_2 \sim \mathrm{LG}(0,1)$ and $C_2 \sim \mathrm{LG}(1.5,1)$ giving CR of about 25%.

In both cases, the realized levels, presented in Table 5.5, are close to the nominal 5% level.

Table 5.5 Logistic Study for Censored PEL

Study no.	n	q	Realized level	Standard error
I	100	0.88	0.053	0.007
	200	0.88	0.054	0.0071
	250	0.88	0.049	0.0068
II	100	0.86	0.048	0.0068
	200	0.86	0.056	0.0072
	250	0.86	0.048	0.0068

We write $\mathrm{EVD}(a,b)$ to denote the extreme value distribution, where a and b are the location and scale parameters respectively. We sampled $X_1 \sim \mathrm{EVD}(0,1)$ and $X_2 \sim \mathrm{EVD}(0,1)$ Also, $C_1 \sim N(0,1)$ giving CR of about 38%, $C_2 \sim N(0,2.25)$ giving

CR of about 40%. The realized levels presented in Table 5.6 are close to the nominal 5%.

Table 5.6 Extreme Value Study for Censored PEL

n	q	Realized level	Standard error
100	0.86	0.049	0.0069
200	0.86	0.049	0.0069
250	0.86	0.048	0.0068

In conclusion, the PEL method competes favorably with ECF for the uncensored case while providing a readily feasible extension for the censored case.

5.4 Level Studies for ECF Censored Case

We used normal, logistic, and extreme value families for our level studies. We carried out two simulations for the normal case, I and II. For normal study I, we took $X_1 \sim N(0,1)$ and $C_1 \sim N(0.75, 0.81)$, giving a censoring rate (CR) of about 29%. We also took $X_2 \sim N(2,9)$ and $C_2 \sim N(4,25)$ giving about 37% CR. For normal study II, we took $X_1 \sim N(0,1)$ and $C_1 \sim N(0.9,1)$, giving about 25% CR. We also took $X_2 \sim N(2,9)$ and $C_2 \sim N(10,9)$ giving about 25% CR. The realized levels, presented in Table 5.7, are close to the nominal 5% level.

For logistic location-scale, we carried out two different simulations. For the logistic study I, $X_1 \sim \text{LG}(0,1)$ and $C_1 \sim \text{LG}(0.5, 0.25)$, giving CR of about 39%. Also, $X_2 \sim \text{LG}(1,4)$ and $C_2 \sim \text{LG}(3.5, 0.56)$, giving CR of about 24%.

For logistic study II, $X_1 \sim \text{LG}(0,1)$ and $C_1 \sim \text{LG}(1.5,1)$, giving CR of about 25%. Also, $X_2 \sim \text{LG}(0,1)$ and $C_2 \sim \text{LG}(1.5,1)$ giving CR of about 25%.

In both cases, the realized levels, presented in Table 5.8, are close to the nominal 5% level.

Table 5.7 Normal Studies for Censored ECF

Study no.	n	q	Realized level	Standard error
I	100	0.85	0.053	0.0070
	200	0.85	0.047	0.0080
	250	0.85	0.049	0.0068
II	100	0.87	0.048	0.0068
	200	0.87	0.049	0.0068
	250	0.87	0.046	0.0068

Table 5.8 Logistic Studies for Censored ECF

Study no.	n	q	Realized level	Standard error
I	100	0.89	0.052	0.0069
	200	0.89	0.053	0.007
	250	0.89	0.052	0.0069
II	100	0.87	0.049	0.0068
	200	0.87	0.051	0.0069
	250	0.87	0.049	0.0068

We write EVD(a, b) to denote the extreme value distribution, where a and b are the location and scale parameters respectively. We sampled $X_1 \sim$ EVD$(0, 1)$ and $X_2 \sim$ EVD$(0, 1)$ Also, $C_1 \sim N(0, 1)$ giving CR of about 38%, $C_2 \sim N(0, 2.25)$ giving CR of about 40%. The realized levels presented in Table 5.9 are close to the nominal 5%.

Table 5.9 Extreme Value Study for Censored ECF

n	q	Realized level	Standard error
100	0.86	0.049	0.0068
200	0.86	0.053	0.0072
250	0.86	0.052	0.007

In conclusion, the PEL method competes favorably with ECF for the uncensored case while providing a readily feasible extension for the censored case.

CHAPTER 6

REAL DATA ANALYSIS

In this chapter, we provide several real examples illustrating the PEL and censored ECF methods.

6.1 Uncensored Case Illustrations

The first example is based on a biomedical experiment concerned with smoking and its effect on cholesterol. We consider a data set from Kössler and Mukherjee (2019) related to an experiment that investigated the health risks of smoking, measured by cholesterol levels of randomly selected persons of two groups of similar ages. One group had a history of smoking for at least 25 years (smokers) and the other group had smoked for no more than 5 years and then stopped (ex-smokers). Information was available on 43 smokers and 33 ex-smokers. Our goal was to test the null hypothesis of membership of the distributions of cholesterol levels for smokers and ex-smokers in some location-scale family. The calculated value of the ECF test statistic was 0.08486. The critical value obtained by the simple bootstrap of Hall et al. (2013) was 0.2786621. It was based on 2,000 bootstrap samples. Furthermore, the p-value was found to be 0.409, providing little evidence against the null hypothesis. The computed value of the PEL test statistic was 1.758. The proposed location-scale appropriate resampling procedure was employed to obtain the critical value. Table 6.1 shows the critical values obtained for different values of q. In all the cases the null hypothesis cannot be rejected, suggesting that the samples are likely to be from distributions belonging to some location-scale family.

We also applied the alternate algorithm described in the power study (see Section 5.1). The bootstrap sample size vector was determined as $(39, 18)^\top$. The

Table 6.1 Location-scale Model Checks for Smokers Data

q	Critical value	p-value	$\hat{\boldsymbol{m}}_B$
0.98	4.965	0.504	(39,30)
0.97	4.888	0.481	(42,33)
0.96	4.995	0.498	(40,31)
0.95	4.574	0.512	(36,27)
0.94	4.993	0.493	(36,28)
0.93	5.003	0.366	(40,31)
0.92	5.036	0.521	(40,31)
0.91	4.758	0.492	(40,31)
0.90	4.662	0.488	(39,30)

critical value was 4.19 and the p-value was 0.445. The findings appear consistent with Table 6.1 figures.

The second illustration compares the distributions of the length of stay (LOS) at hospital of two groups of patients in a clinical trial conducted at the Institute of Living (Hartford Hospital, Hartford CT). The CYP-guides trial data, on which this illustration is based, were collected by Tortora et al. (2020). Out of a total of 1459 patients, 477 were randomly assigned to standard therapy (S) and 982 to genetically-guided therapy (G). Both distributions being strictly positive, we performed a location-scale analysis on the logarithmic scale.

The ECF test statistic was 0.149 and the critical value was 0.404. The bootstrap p-value, based on 2,000 bootstrap samples, was 0.335, and does not provide evidence against the null hypothesis that the S and G groups belong to the same location-scale family.

The PEL test statistic was 3.41. Table 6.2 shows critical values for different values of q and the data-driven value of $\boldsymbol{m}_B$. In all cases the null hypothesis is not rejected. The result of the PEL study is consistent with findings from the ECF method. The alternate search algorithm described in the power study returned $(884, 282)^\top$ as the bootstrap sample size vector. Here also $q = 0.9$. The critical value was 7.89 and the p-value was 0.415. Again, the findings appear to be consistent with Table 6.2 figures.

Table 6.2 Location-scale Model Checks for CYP-guides Trial Data

q	Critical value	p-value	$\hat{\boldsymbol{m}}_B$
0.95	8.51	0.45	(842,409)
0.94	7.57	0.49	(721,351)
0.93	7.89	0.52	(790,384)
0.92	7.66	0.51	(765,372)
0.91	6.95	0.43	(580,282)

6.2 Censored Case Illustrations

Our censored case illustration is concerned with a study that was designed to evaluate a new body-cleansing method using 4% chlorhexidine gluconate with a standard method (initial surface decontamination with 10% povidone-iodine followed with regular bathing with Dial soap) on burn subjects, see Ichida et al. (1993). The study period was 18 months. The time until staphylococcus infection (in days) was recorded, along with a binary variable indicating presence or absence of infection. The group which received the new bathing solution had 84 subjects, with 24 right censored. The control group had 70 subjects, with 31 right-censored. The computed value of the PEL test statistic was 2.1485. Table 6.3 shows critical values and $\hat{\boldsymbol{m}}_B$

arrived at from different values of q. In all cases the null hypothesis is not rejected, suggesting that the samples may have arisen from some location-scale family.

Table 6.3 Location-scale Model Checks for the Burn Data (PEL)

q	Critical value	p-value	$\hat{\boldsymbol{m}}_B$
0.9	7.88	0.7	(57,69)
0.89	6.37	0.53	(35,42)
0.88	6.37	0.58	(37,45)
0.87	6.46	0.54	(35,42)
0.86	7.14	0.61	(45,54)
0.85	6.47	0.59	(37,44)

The alternate search algorithm returned $(57, 69)$ as the estimated bootstrap sample size vector. Here again $q = 0.9$. The critical value was 8.43 and the p-value was 0.72. As was the case for the two illustrations before, the findings appear to be consistent with Table 6.3 figures.

The computed value of the ECF test statistic was 0.2802. Table 6.4 shows critical values and $\hat{\boldsymbol{m}}_B$ arrived at from different values of q. In all cases the null hypothesis is not rejected, suggesting that the samples may have arisen from some location-scale family. The result of the ECF study is consistent with findings from the PEL method.

Table 6.4 Location-scale Model Checks for the Burn Data (ECF)

q	Critical value	p-value	$\hat{\boldsymbol{m}}_B$
0.9	2.81	0.84	(57,69)
0.85	2.312	0.77	(60,72)

CHAPTER 7

CONCLUSION

The first part of this dissertation introduced a plug-in empirical likelihood approach for testing the hypothesis that two samples are from some location-scale family. Tests were developed for both the uncensored and censored cases. Using these tests one can check whether two samples are being drawn from some location-scale family. Knowing that distributions differ only in location and scale can facilitate comparisons and potentially save time and cost. The method would be useful in biomedical studies where interest may center on comparing the effect of a new treatment effects over a standard treatment. More generally, in some situations location-scale may provide greater leverage over fully parametric or fully nonparametric inference, offering a compromise between the two extremes. In such instances it is imperative to have the location-scale assumption checked in advance of data analysis. The PEL test would be a highly useful preliminary two-sample testing device.

Unlike Hall et al.'s (2013) uncensored ECF test, the proposed PEL test applies equally well to censored data also. This is due to the fact that the PEL ratio assumes the same form whether the data are censored or uncensored, the plug-ins being the only difference. For uncensored data, the sample mean and sample standard deviation are plugged in. For censored data estimators based on Stute's (1995) KM integrals are employed. This approach is different from that of Subramanian (2020), who employed minimum distance plug-ins of location and scale parameters. Since the proposed approach works with standardized variables, a direct comparison with Hall et al.'s (2013) uncensored ECF test is rendered feasible. In fact, the plug-ins based on Stute's (1995) KM integrals paved the way for setting up a censored ECF test. This facility is not available with the approach that uses different plug-ins (Subramanian, 2020).

The power studies indicate that the proposed PEL has the potential to outperform the ECF test. Several Illustrations with real examples are given.

The failure of the full bootstrap, noted by Bickel and Sakov (2008), recurs with PEL as well. Resampling with smaller bootstrap sample sizes appears to be effective. As in Subramanian (2020), Bickel et al.'s (2008) adaptive rule for the one-dimensional setting is extended to the two-dimensional setting to obtain the estimated bootstrap sample size vector. Unlike in Bickel and Sakov (2008), however, the two-sample setting requires minimization over a two-dimensional grid, creating a relatively more complicated data-driven procedure for determining the bootstrap sample sizes. The search trajectory proceeded diagonally downward. An alternate algorithm, used to illustrate the proposed PEL test, employed a two stage approach, first searching horizontally and then moving downward.

In the second part of this dissertation, the ECF test is extended to accommodate right censoring. The plug-ins are again based on Stute's (1995) KM integrals. Critical values are again based on a model appropriate resampling procedure, which produced realized levels close to the nominal level. The censored ECF test is illustrated with a real examples.

There are some potential future research directions. The first would be to construct an algorithm for computing bootstrap sample sizes that are independent of the choice of q. Theoretical justification that the algorithm produces bootstrap sample sizes that, when combined with the location-scale model-appropriate sampling, yields correct realized levels would be another important direction. Another important study would be a comprehensive power evaluation of the three methods, namely censored ECF and PEL (Two methods).

APPENDIX A

PEL RATIO TEST

This appendix show the lemmas and details we used to derive asymptotic distribution of PEL test statistics.

A.1 Derivation of the constrained estimates

For known $\boldsymbol{\theta}$, let $\gamma_{\boldsymbol{\theta}}(t)$ be a Lagrange multiplier. We find $\tilde{\lambda}_{ij}$, the λ_{ij} that maximize

$$D = \log[L(K_1, K_2)] + \gamma_{\boldsymbol{\theta}}(t)\left(\sum_{j=1}^{\eta_2(t)} \log(1-\lambda_{2j}) - \sum_{j=1}^{\eta_1(t)} \log(1-\lambda_{1j})\right).$$

Applying Equation (3.6), it can be readily checked that (cf. Thomas and Grunkemeier, 1975)

$$\begin{aligned} D &= \sum_{i=1}^{2}\Bigg[\sum_{j=1}^{\eta_i(t)}\left[d_{ij}\log(\lambda_{ij}) + \big(r_{ij} - d_{ij} + (-1)^i\gamma_{\boldsymbol{\theta}}(t)\big)\log(1-\lambda_{ij})\right] \\ &\quad + \sum_{j=\eta_i(t)+1}^{m_i}\left[d_{ij}\log(\lambda_{ij}) + (r_{ij}-d_{ij})\log(1-\lambda_{ij})\right]\Bigg]. \end{aligned}$$

Maximizing D, it can also be checked that the constrained estimates $\tilde{\lambda}_{ij}, i = 1, 2$, are

$$\tilde{\lambda}_{ij} = \begin{cases} I(i=2) + (-1)^{i-1} d_{ij}/(r_{ij} + (-1)^i\hat{\gamma}_{\boldsymbol{\theta}}(t)), & j = 1, \ldots, \eta_i(t) \\ I(i=2) + (-1)^{i-1} d_{ij}/r_{ij}, & j = \eta_i(t)+1, \ldots, m_i; \end{cases} \tag{A.1}$$

Furthermore, $\hat{\gamma}_{\boldsymbol{\theta}}(t)$ satisfies the equation

$$\prod_{j=1}^{\eta_1(t)}\left(1 - \frac{d_{1j}}{r_{1j} - \gamma_{\boldsymbol{\theta}}(t)}\right) - \prod_{j=1}^{\eta_2(t)}\left(1 - \frac{d_{2j}}{r_{2j} + \gamma_{\boldsymbol{\theta}}(t)}\right) = 0. \tag{A.2}$$

Note that $\hat{\gamma}_{\boldsymbol{\theta}}$ is hemmed-in between D_2 and $-D_1$, where

$$D_1 = \max_{1\le j\le \eta_1(t)}\{d_{1j} - r_{1j}\}, \qquad D_2 = \max_{1\le j\le \eta_2(t)}\{d_{2j} - r_{2j}\}. \tag{A.3}$$

Then $\prod_{l=1}^{j}(1-\tilde{\lambda}_{il}) = \tilde{S}_{i0}(\tilde{U}_{ij})$, are the constrained survival function estimates.

A.1.1 Formula for log likelihood ratio statistic

The log LR statistic is

$$R_{\boldsymbol{\theta}}(t) \quad := \quad \log(L(\tilde{S}_{10}, \tilde{S}_{20})) - \log(L(\hat{S}_{10}, \hat{S}_{20})). \tag{A.4}$$

where the expressions for $L(\hat{S}_{10}, \hat{S}_{20})$ and $L(\tilde{S}_{10}, \tilde{S}_{20})$ on the RHS of Equation (A.4) were derived in Section 3.2. Using these derived expressions, we compute the right hand side of Equation (A.4). Suppressing the dependence on t and $\boldsymbol{\theta}$, we write $R = \tilde{R}_1 + \tilde{R}_2 + \tilde{R}_3 + \tilde{R}_4$, where

$$\tilde{R}_1 = \sum_{j=1}^{\eta_1(t)} d_{1j} \log\left(\frac{d_{1j}}{r_{1j} - \hat{\gamma}_{\boldsymbol{\theta}}}\right) - \sum_{j=1}^{\eta_1(t)} d_{1j} \log\left(\frac{d_{1j}}{r_{1j}}\right),$$

$$\tilde{R}_2 = \sum_{j=1}^{\eta_1(t)} (r_{1j} - d_{1j}) \log\left(1 - \frac{d_{1j}}{r_{1j} - \hat{\gamma}_{\boldsymbol{\theta}}}\right) - \sum_{j=1}^{\eta_1(t)} (r_{1j} - d_{1j}) \log\left(1 - \frac{d_{1j}}{r_{1j}}\right),$$

$$\tilde{R}_3 = \sum_{j=1}^{\eta_2(t))} d_{2j} \log\left(\frac{d_{2j}}{r_{2j} + \hat{\gamma}_{\boldsymbol{\theta}}}\right) - \sum_{j=1}^{\eta_2(t))} d_{2j} \log\left(\frac{d_{2j}}{r_{2j}}\right),$$

$$\tilde{R}_4 = \sum_{j=1}^{\eta_2(t))} (r_{2j} - d_{2j}) \log\left(1 - \frac{d_{2j}}{r_{2j} + \hat{\gamma}_{\boldsymbol{\theta}}}\right) - \sum_{j=1}^{\eta_2(t))} (r_{2j} - d_{2j}) \log\left(1 - \frac{d_{2j}}{r_{2j}}\right).$$

Writing $\tilde{R}_1 = R_1 - \breve{R}_1$, where

$$\breve{R}_1 = \sum_{j=1}^{\eta_1(t)} (r_{1j} - d_{1j}) \log\left(\frac{d_{1j}}{r_{1j} - \hat{\gamma}_{\boldsymbol{\theta}}}\right) - \sum_{j=1}^{\eta_1(t)} (r_{1j} - d_{1j}) \log\left(\frac{d_{1j}}{r_{1j}}\right),$$

we see that

$$R_1 \equiv \sum_{j=1}^{\eta_1(t)} r_{1j} \log\left(\frac{d_{1j}}{r_{1j} - \hat{\gamma}_{\boldsymbol{\theta}}}\right) - \sum_{j=1}^{\eta_1(t)} r_{1j} \log\left(\frac{d_{1j}}{r_{1j}}\right) = -\sum_{j=1}^{\eta_1(t)} r_{1j} \log\left(1 - \frac{\hat{\gamma}_{\boldsymbol{\theta}}}{r_{1j}}\right).$$

Next we write $\tilde{R}_1 + \tilde{R}_2 \equiv R_1 + R_2$, where $R_2 = \tilde{R}_2 - \breve{R}_1$. Then, after some algebra, we obtain

$$R_2 = \sum_{j=1}^{\eta_1(t)} (r_{1j} - d_{1j}) \log\left(1 - \frac{\hat{\gamma}_{\boldsymbol{\theta}}}{r_{1j} - d_{1j}}\right).$$

Applying exactly the same technique, we have that $\tilde{R}_3 + \tilde{R}_4 \equiv R_3 + R_4$, where

$$R_3 = -\sum_{j=1}^{\eta_2(t))} r_{2j} \log\left(1 + \frac{\hat{\gamma}_{\boldsymbol{\theta}}}{r_{2j}}\right),$$
$$R_4 = \sum_{j=1}^{\eta_2(t))} (r_{2j} - d_{2j}) \log\left(1 + \frac{\hat{\gamma}_{\boldsymbol{\theta}}}{r_{2j} - d_{2j}}\right).$$

We have shown that $R = R_1 + R_2 + R_3 + R_4$. Equation (3.8) follows from these calculations. □

A.1.2 Some lemmas with proofs

Recall that $\eta_i(t) = \sum_{j=1}^{m_i} I(\tilde{U}_{ij} \le t)$. Also note that $\hat{\zeta}_{i0}(t) = -\log \hat{S}_{i0}(t)$, where $\hat{S}_{i0}(t) = 1 - \hat{F}_{i0}(t)$ and $F_{i0}(t)$ is the distribution function of U_i. In the notation incorporating ties,

$$\hat{\Lambda}_{i0}(t) = \sum_{j=1}^{\eta_i(t)} \left(\frac{d_{ij}}{r_{ij}}\right), \qquad \hat{\zeta}_{i0}(t) = -\sum_{j=1}^{\eta_i(t)} \log\left(1 - \frac{d_{ij}}{r_{ij}}\right), \qquad i = 1, 2, \quad \text{(A.5)}$$

It can be shown that (e.g., Subramanian 2020)

$$\|\hat{\Lambda}_{i0} - \hat{\zeta}_{i0}\|_{\alpha_1}^{\alpha_2} := \sup_{t \in [\alpha_1, \alpha_2]} |\hat{\Lambda}_{i0}(t) - \hat{\zeta}_{i0}(t)| = o_{\mathbb{P}}(n^{-1/2}). \qquad \text{(A.6)}$$

Remark A.1 To prove that certain remainder terms are negligible we require an important result. Note that $S_{i0}(t) = S_i(\varphi_i(t))$ and $\hat{S}_{i0}(t) = \hat{S}_i(\varphi_i(t))$, where $S_i(t)$ is the survival function of X_i, $\hat{S}_i(t)$ is the empirical distribution of $X_{i,1}, \ldots, X_{i,n_i}$, and $\varphi_i(t) = \mu_i + \sigma_i t$. Write $\hat{\varphi}_i(t) = s'$, where $\hat{\varphi}_i(t) = \hat{\mu}_i + \hat{\sigma}_i t$, and $\varphi_i(t) = t'$. Then $|s' - t'| = |\hat{\mu}_i - \mu_i| + |t||\hat{\sigma}_i - \sigma_i| = o_{\mathbb{P}}(n^{-1/2+\epsilon})$, uniformly on $[\alpha_1, \alpha_2]$. From Lemma 1 of Ying et al. (1995),

$$\sup_{t \in [\alpha_1, \alpha_2]} \left|\hat{\zeta}_i(\hat{\varphi}_i(t)) - \zeta_i(\hat{\varphi}_i(t)) - \hat{\zeta}_i(\varphi_i(t)) + \zeta_i(\varphi_i(t))\right| = o_{\mathbb{P}}(n^{-1/2}), \quad i = 1, 2. \text{(A.7)}$$

Before deploying Taylor expansions of functions of $\hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}(\cdot)$ about 0, we need to obtain a proper uniform rate for $\hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}(\cdot)$ (cf. Li, 1995). Upper bounds are derived in Lemma 1.

Lemma 1 *The Lagrange multiplier, $\hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}(t)$, solving Eq. (A.2), satisfies*

$$|\hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}(t)| \;\leq\; n_2\left(\hat{\zeta}_1(\varphi_1(t)) - \hat{\Lambda}_2(\varphi_2(t))\right)/\hat{\Lambda}_2(\varphi_2(t)), \qquad \text{when } \hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}(t) < 0. \quad \text{(A.8)}$$

$$|\hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}(t)| \;\leq\; n_1\left(\hat{\zeta}_2(\varphi_2(t)) - \hat{\Lambda}_1(\varphi_1(t))\right)/\hat{\Lambda}_1(\varphi_1(t)), \qquad \text{when } \hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}(t) > 0. \quad \text{(A.9)}$$

Proof From Equation (A.2), $\hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}(t)$ satisfies the equation $I_1(t) - I_2(t) = 0$, where

$$I_1(t) = \sum_{j=1}^{\eta_1(t)} \log\left(1 - \frac{d_{1j}}{r_{1j} - \hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}(t)}\right); \qquad I_2(t) = \sum_{j=1}^{\eta_2(t)} \log\left(1 - \frac{d_{2j}}{r_{2j} + \hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}(t)}\right).$$

We consider the case $\hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}(t) < 0$. Then, $\hat{\boldsymbol{\gamma}}_{\theta}(t) = -|\hat{\boldsymbol{\gamma}}_{\theta}(t)|$. Following Li's (1995) approach that leads to his equation (2.12), we likewise use $-\log(1-x) \geq x$, for $0 \leq x < 1$, to obtain

$$\begin{aligned}
-I_2(t) &= \sum_{j=1}^{\eta_2(t)} -\log\left(1 - \frac{d_{2j}}{r_{2j} + \hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}(t)}\right) \\
&\geq \sum_{j=1}^{\eta_2(t)} \left(\frac{d_{2j}}{r_{2j} + \hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}(t)}\right) = \sum_{j=1}^{\eta_2(t)} \frac{d_{2j}}{r_{2j}} \left(\frac{r_{2j}}{r_{2j} - |\hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}(t)|}\right) \\
&\geq \sum_{j=1}^{\eta_2(t)} \frac{d_{2j}}{r_{2j}} \left(\frac{n_2}{n_2 - |\hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}(t)|}\right) = \hat{\Lambda}_2(\varphi_2(t)) \left(\frac{1}{1 - |\hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}(t)|/n_2}\right).
\end{aligned}$$

Note that, for large enough n_2, $0 < |\hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}(t)|/n_2 < 1$ almost surely, see below. Then, as in McKeague and Zhao (2005), we use the fact that $1/(1-x) \geq 1 + x$ for $0 < x < 1$, to obtain

$$-I_2(t) \;\geq\; \hat{\Lambda}_2(\varphi_2(t)) + \hat{\Lambda}_2(\varphi_2(t))|\hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}(t)|/n_2. \qquad \text{(A.10)}$$

Turning to $I_1(t)$, because $-\hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}(t) > 0$, we can write

$$I_1(t) = \sum_{j=1}^{\eta_1(t)} \log\left(1 - \frac{d_{1j}}{r_{1j} + |\hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}(t)|}\right).$$

We again follow Li (1995) by exploiting that $\log(1-x)+x$ is decreasing on $(0,1)$, to obtain

$$\begin{aligned} I_1(t) &\geq \sum_{j=1}^{\eta_1(t)} \left[-\left(\frac{d_{1j}}{r_{1j}+|\hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}(t)|} \right) + \log\left(1-\frac{d_{1j}}{r_{1j}}\right) + \frac{d_{1j}}{r_{1j}} \right] \\ &= -\sum_{j=1}^{\eta_1(t)} \frac{d_{1j}}{r_{1j}} \left(\frac{r_{1j}}{r_{1j}+|\hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}(t)|} \right) + \sum_{j=1}^{\eta_1(t)} \left[\log\left(1-\frac{d_{1j}}{r_{1j}}\right) + \frac{d_{1j}}{r_{1j}} \right] \\ &\geq -\sum_{j=1}^{\eta_1(t)} \frac{d_{1j}}{r_{1j}} \left(\frac{n_1}{n_1+|\hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}(t)|} \right) + \sum_{j=1}^{\eta_1(t)} \left[\log\left(1-\frac{d_{1j}}{r_{1j}}\right) + \frac{d_{1j}}{r_{1j}} \right] \\ &= -\hat{\Lambda}_2(\varphi_2(t)) \left(\frac{n_1}{n_1+|\hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}(t)|} \right) - \hat{\zeta}_1(\varphi_1(t)) + \hat{\Lambda}_1(\varphi_1(t)) \\ &= \hat{\Lambda}_1(\varphi_1(t)) \left(\frac{|\hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}(t)|}{n_1+|\hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}(t)|} \right) - \hat{\zeta}_1(\varphi_1(t)) \;\geq\; -\hat{\zeta}_1(\varphi_1(t)). \end{aligned} \tag{A.11}$$

Combine inequalities (A.10) and (A.11) to obtain inequality (A.8) from

$$0 \;=\; I_1(t) - I_2(t) \;\geq\; -\hat{\zeta}_1(\varphi_1(t)) + \hat{\Lambda}_2(\varphi_2(t)) + \hat{\Lambda}_2(\varphi_2(t))|\hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}(t)|/n_2.$$

It remains to show that, for large enough n_2, $0 < |\hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}(t)|/n_2 < 1$ a.s. Indeed, assuming no ties, we note from Equation (3.9) that $D_2 < \hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}} < 0$, where $D_2 = \max_{1\leq j\leq \eta_2(t)}\{d_{2j} - r_{2j}\}$. Therefore,

$$\max_{1\leq j\leq \eta_2(t)}\{d_{2j}-r_{2j}\} < \hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}} \equiv -|\hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}| < 0 \iff 0 < |\hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}| < \min_{1\leq j\leq \eta_2(t)}\{r_{2j}-d_{2j}\}.$$

Clearly $0 < \min_{1\leq j\leq \eta_2(t)}\{r_{2j}-d_{2j}\} < n_2$ when $t \in [\alpha_1, \alpha_2]$. Hence $0 < |\hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}(t)|/n_2 < 1$ a.s.

When $\hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}(t) > 0$, then $\hat{\boldsymbol{\gamma}}_{\theta}(t) = |\hat{\boldsymbol{\gamma}}_{\theta}(t)|$, the above approach can be repeated by applying Li's (1995) technique in reverse order. First obtain the inequality

$$I_2(t) \;\geq\; -\hat{\Lambda}_2(\varphi_2(t)) \left(\frac{n_2}{n_2+|\hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}(t)|} \right) - \hat{\zeta}_2(\varphi_2(t)) + \hat{\Lambda}_2(\varphi_2(t)) \;\geq\; -\hat{\zeta}_2(\varphi_2(t)). \tag{A.12}$$

Then obtain the inequality

$$-I_1(t) \;\geq\; \hat{\Lambda}_1(\varphi_1(t)) \left(\frac{n_1}{n_1-|\hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}(t)|} \right) \;\geq\; \hat{\Lambda}_1(\varphi_1(t))\left(1+|\hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}(t)|/n_1\right). \tag{A.13}$$

Combine inequalities (A.12) and (A.13) to obtain inequality (A.9) from

$$0 \;=\; I_2(t) - I_1(t) \;\geq\; -\hat{\zeta}_2(\varphi_2(t)) + \hat{\Lambda}_1(\varphi_1(t)) + \hat{\Lambda}_1(\varphi_1(t))\hat{\boldsymbol{\gamma}}_{\boldsymbol{\theta}}(t)|/n_1. \qquad \square$$

Lemma 2 *The estimate $\hat{\boldsymbol{\gamma}}_{\hat{\boldsymbol{\theta}}}(t)$, solving Equation (A.2) with $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$, satisfies $\|\hat{\boldsymbol{\gamma}}_{\hat{\boldsymbol{\theta}}}\|_{\alpha_1}^{\alpha_2} = O_{\mathbb{P}}(n^{1/2})$.*

Proof Note from Equation (A.2) that $\hat{\boldsymbol{\gamma}}_{\hat{\boldsymbol{\theta}}}(t)$ solves the equation $\hat{I}_1(\gamma_{\boldsymbol{\theta}}(t)) - \hat{I}_2(\gamma_{\boldsymbol{\theta}}(t)) = 0$, where

$$\hat{I}_1(\gamma_{\boldsymbol{\theta}}(t)) = \sum_{j=1}^{\hat{\eta}_1(t)} \log\left(1 - \frac{d_{1j}}{r_{1j} - \gamma_{\boldsymbol{\theta}}(t)}\right); \qquad \hat{I}_2(\gamma_{\boldsymbol{\theta}}(t)) = \sum_{j=1}^{\hat{\eta}_2(t)} \log\left(1 - \frac{d_{2j}}{r_{2j} + \gamma_{\boldsymbol{\theta}}(t)}\right).$$

Following closely the proof of Lemma 1, it follows that the Lagrange multiplier, $\hat{\boldsymbol{\gamma}}_{\hat{\boldsymbol{\theta}}}(t)$, satisfies

$$|\hat{\boldsymbol{\gamma}}_{\hat{\boldsymbol{\theta}}}(t)| \leq n_2\left(\hat{\zeta}_1(\hat{\varphi}_1(t)) - \hat{\Lambda}_2(\hat{\varphi}_2(t))\right)/\hat{\Lambda}_2(\hat{\varphi}_2(t)), \qquad \text{when } \hat{\boldsymbol{\gamma}}_{\hat{\boldsymbol{\theta}}}(t) < 0. \quad \text{(A.14)}$$

$$|\hat{\boldsymbol{\gamma}}_{\hat{\boldsymbol{\theta}}}(t)| \leq n_1\left(\hat{\zeta}_2(\hat{\varphi}_2(t)) - \hat{\Lambda}_1(\hat{\varphi}_1(t))\right)/\hat{\Lambda}_1(\hat{\varphi}_1(t)), \qquad \text{when } \hat{\boldsymbol{\gamma}}_{\hat{\boldsymbol{\theta}}}(t) > 0. \quad \text{(A.15)}$$

Because of consistency of $\hat{\mu}_i$ and $\hat{\sigma}_i$, and strong consistency of $\hat{\Lambda}_i$ on $\mathbb{R}^1$, the denominators on the RHS of inequalities (A.14) and (A.15) are each $O_{\mathbb{P}}(1)$, and are bounded away from 0 uniformly for $t \in [\alpha_1, \alpha_2]$. It remains to show that the numerators on the RHS of inequalities (A.14) and (A.15) are each $O_{\mathbb{P}}(n^{1/2})$ uniformly for $t \in [\alpha_1, \alpha_2]$. Consider the numerator on the RHS of inequality (A.14). We obtain the decomposition

$$\begin{aligned}\hat{\zeta}_1(\hat{\varphi}_1(t)) &= \hat{\zeta}_1(\hat{\varphi}_1(t)) - \zeta_1(\hat{\varphi}_1(t)) - \hat{\zeta}_1(\varphi_1(t)) + \zeta_1(\varphi_1(t)) \\ &+ \zeta_1(\hat{\varphi}_1(t)) + \hat{\zeta}_1(\varphi_1(t)) - \zeta_1(\varphi_1(t)),\end{aligned}$$

with a similar representation for $\hat{\Lambda}_2(\hat{\varphi}_2(t))$ [replace $(\hat{\zeta}_1, \zeta_1)$ above with $(\hat{\Lambda}_2, \Lambda_2)$ respectively and $(\hat{\varphi}_1, \varphi_1)$ above with $(\hat{\varphi}_2, \varphi_2)$ respectively]. After applying **Remark A.1** (and its variant) to the first four terms of each representation, giving $o_{\mathbb{P}}(n^{-1/2})$ for each, we are left with

$$\begin{aligned}\zeta_1(\hat{\varphi}_1(t)) - \zeta_1(\varphi_1(t)) + \hat{\zeta}_1(\varphi_1(t)) &= \Lambda_1(\hat{\varphi}_1(t)) - \Lambda_1(\varphi_1(t)) + \hat{\zeta}_1(\varphi_1(t)) \\ &= \Lambda_1(\hat{\varphi}_1(t)) + O_{\mathbb{P}}(n^{-1/2}),\end{aligned}$$

the last step after applying Equation (A.6) on the original (non-standardized) data. Likewise we obtain $\Lambda_2(\hat{\varphi}_2(t)) + O_{\mathbb{P}}(n^{-1/2})$ for the three terms remaining in the second representation. Noting that $\Lambda_1(\varphi_1(t)) = \Lambda_2(\varphi_2(t))$ under the null hypothesis, the numerator of Equation (A.14) is

$$\begin{aligned} n(\Lambda_1(\hat{\varphi}_1(t)) - \Lambda_2(\hat{\varphi}_2(t))) + O_{\mathbb{P}}(n^{1/2}) &= n(\Lambda_1(\varphi_1(t)) - \Lambda_2(\varphi_2(t))) + O_{\mathbb{P}}(n^{1/2}) \\ &= O_{\mathbb{P}}(n^{1/2}). \end{aligned}$$

Analogous treatment of the RHS of (A.15) yields $O_{\mathbb{P}}(n^{1/2})$, completing proof of the lemma. □

Lemma 3 *Under the null hypothesis, for each* $i = 1, 2$, $\hat{\zeta}_i(\hat{\varphi}_i(t)) = A_i(t) + B_i(t) + o_{\mathbb{P}}(n^{-1/2})$, *where* $A_i(t)$ *is defined by Equation (3.12) and* $B_i(t)$ *is defined by Equation (3.13).*

Proof The seven-term decomposition given for $\hat{\zeta}_i(\hat{\varphi}_i(t))$ in the proof of Lemma 2, and the consequent application of **Remark A.1** through which the first four terms jointly account for $o_{\mathbb{P}}(n^{-1/2})$ implies that $\hat{\zeta}_i(\hat{\varphi}_i(t)) = A_i(t) + B_i(t) + o_{\mathbb{P}}(n^{-1/2})$, where $A_i(t) = -\log S_i(\hat{\varphi}_i(t)) + \log S_i(\varphi_i(t))$ and $B_i(t) = -\log \hat{S}_i(\varphi_i(t))$. Applying Taylor's expansions, we have

$$A_i(t) \;=\; -\frac{S_i(\hat{\varphi}_i(t)) - S_i(\varphi_i(t))}{S_i(\varphi_i(t))} + o_{\mathbb{P}}(n^{-1/2}) \;=\; (\hat{\varphi}_i(t) - \varphi_i(t))\lambda_i(\varphi_i(t)) + o_{\mathbb{P}}(n^{-1/2}).$$

Since $\lambda_i(\varphi_i(t)) = \lambda_{i0}(t)/\sigma_i$, and $\hat{\varphi}_i(t) = \hat{\mu}_i + t\hat{\sigma}_i$, elementary manipulations yield Equation (3.12). An application of the delta method for $B_i(t) \equiv -\log \hat{S}_i(\varphi_i(t)) = -\log \hat{S}_{i0}(t)$ produces

$$B_i(t) \;=\; -\log S_{i0}(t) - \log \hat{S}_{i0}(t) + \log S_{i0}(t) \;=\; \Lambda_{i0}(t) - \frac{\hat{S}_{i0}(t) - S_{i0}(t)}{S_{i0}(t)} + o_{\mathbb{P}}(n^{-1/2}).$$

Applying the Duhamel equation (see Andersen et al., 1993) we obtain Equation (3.13).

□

APPENDIX B
PROOF OF THEOREM 1

The proof follows standard techniques. From Equation (A.2), $h_1\left(-\hat{\boldsymbol{\gamma}}_{\hat{\boldsymbol{\theta}}}(t)\right) - h_2\left(\hat{\boldsymbol{\gamma}}_{\hat{\boldsymbol{\theta}}}(t)\right) = 0$, where

$$h_i(\gamma) \;=\; \sum_{j=1}^{\hat{\eta}_i(t)} \log\left(1 - \frac{d_{ij}}{r_{ij}+\gamma}\right). \tag{B.1}$$

Note that $h_1(0) = -\hat{\zeta}_1(\hat{\varphi}_1(t))$ and $h_2(0) = -\hat{\zeta}_2(\hat{\varphi}_2(t))$, see Equation (A.5). Note also that

$$h_i^{'}(\gamma) \;=\; \sum_{j=1}^{\hat{\eta}_i(t)} \frac{d_{ij}}{(r_{ij}+\gamma)(r_{ij}+\gamma-d_{ij})}, \qquad h_i^{''}(\gamma) \;=\; \sum_{j=1}^{\hat{\eta}_i(t)} \frac{d_{ij}(2(r_{ij}+\gamma)-d_{ij})}{(r_{ij}+\gamma)^2(r_{ij}+\gamma-d_{ij})^2}.$$

Note that $n_i h_i^{'}(0) = \hat{\vartheta}^2_{i0}(t), i = 1,2$, where Equation (3.4) defines $\vartheta^2_{i0}(t)$. Let $|\hat{\xi}_i| \leq |\hat{\boldsymbol{\gamma}}_{\hat{\boldsymbol{\theta}}}(t)|, i = 1,2$. Taylor's expansion about 0 yields

$$h_1\left(-\hat{\boldsymbol{\gamma}}_{\hat{\boldsymbol{\theta}}}(t)\right) \;=\; -\hat{\zeta}_1(\hat{\varphi}_1(t)) - \frac{1}{n_1}\hat{\vartheta}^2_{10}(t)\hat{\boldsymbol{\gamma}}_{\hat{\boldsymbol{\theta}}}(t) + \frac{1}{2}h_1^{''}(\hat{\xi}_1)\hat{\boldsymbol{\gamma}}^2_{\hat{\boldsymbol{\theta}}}(t), \tag{B.2}$$

$$h_2\left(\hat{\boldsymbol{\gamma}}_{\hat{\boldsymbol{\theta}}}(t)\right) \;=\; -\hat{\zeta}_2(\hat{\varphi}_2(t)) + \frac{1}{n_2}\hat{\vartheta}^2_{20}(t)\hat{\boldsymbol{\gamma}}_{\hat{\boldsymbol{\theta}}}(t) + \frac{1}{2}h_1^{''}(\hat{\xi}_2)\hat{\boldsymbol{\gamma}}^2_{\hat{\boldsymbol{\theta}}}(t). \tag{B.3}$$

First we argue that $\|h_i''(\hat{\xi}_i)\|_{\alpha_1}^{\alpha_2} = O_{\mathbb{P}}(n_i^{-2})$. Consider the two denominator terms of $h_i''(\hat{\xi}_i)$. By the Glivenko–Cantelli lemma and Lemma 3, each term within parenthesis is dominated by $r_{ij} = O_{\mathbb{P}}(n_i)$. The denominator is thus $O_{\mathbb{P}}(n_i^4)$, uniformly over $j = 1,\ldots,\hat{\eta}_i(t)$. Assuming continuity (guaranteeing no ties), the numerator is dominated by $r_{ij} = O_{\mathbb{P}}(n_i)$ uniformly over $j = 1,\ldots,\hat{\eta}_i(t)$. The summation contributes to $O_{\mathbb{P}}(n_i)$. Thus $\|h_i''(\hat{\xi}_i)\|_{\alpha_1}^{\alpha_2} = O_{\mathbb{P}}(n_i^{-2})$. By Lemma 2, therefore

$$\left\|h_i^{''}(\hat{\xi}_i)\left(\hat{\boldsymbol{\gamma}}_{\hat{\boldsymbol{\theta}}}(t)\right)^2\right\|_{\alpha_1}^{\alpha_2} = O_{\mathbb{P}}(1/n_i) = O_{\mathbb{P}}(1/n).$$

Thus $h_1\left(-\hat{\boldsymbol{\gamma}}_{\hat{\boldsymbol{\theta}}}(t)\right) - h_2\left(\hat{\boldsymbol{\gamma}}_{\hat{\boldsymbol{\theta}}}(t)\right) = 0$, combined with Equation (3.10), Equation (B.2), and Equation (B.3) give

$$\begin{aligned} 0 \;&=\; -\hat{\zeta}_1(\hat{\varphi}_1(t)) + \hat{\zeta}_2(\hat{\varphi}_2(t)) - \hat{\boldsymbol{\gamma}}_{\hat{\boldsymbol{\theta}}}(t)\left(\vartheta^2_{10}(t)/n_1 + \vartheta^2_{20}(t)/n_2\right) + O_{\mathbb{P}}\left(1/n\right) \\ &=\; -\hat{\mathbb{V}}(t) - \hat{\boldsymbol{\gamma}}_{\hat{\boldsymbol{\theta}}}(t)\vartheta^2_{\mathrm{c}}(t)/n + O_{\mathbb{P}}\left(1/n\right). \end{aligned}$$

Solving for $\hat{\boldsymbol{\gamma}}_{\hat{\boldsymbol{\theta}}}(t)$, we obtain

$$\hat{\boldsymbol{\gamma}}_{\hat{\boldsymbol{\theta}}}(t) \quad = \quad \frac{n}{\vartheta_{\mathrm{c}}^{2}(t)}\left\{-\hat{\mathbb{V}}(t)+O_{\mathbb{P}}\left(1/n\right)\right\}. \tag{B.4}$$

To complete the proof of theorem 1, consider Equation (3.8). Using Taylor expansions of $\log(1+x)$ and $\log(1-x)$ about 0 and applying Equation (B.4), the leading term of $-2R_{\hat{\boldsymbol{\theta}}}(t)$ equals

$$(\hat{\boldsymbol{\gamma}}_{\hat{\boldsymbol{\theta}}}(t))^2\sum_{i=1}^{2}\sum_{j=1}^{\hat{\eta}_i(t)}\frac{d_{ij}}{r_{ij}(r_{ij}-d_{ij})} \quad = \quad (\hat{\boldsymbol{\gamma}}_{\hat{\boldsymbol{\theta}}}(t))^2\left(\hat{\vartheta}_{10}^{2}(t)/n_1+\hat{\vartheta}_{20}^{2}(t)/n_2\right). \tag{B.5}$$

Applying Equation (B.4) for $\hat{\boldsymbol{\gamma}}_{\hat{\boldsymbol{\theta}}}(t)$ appearing on the RHS of Equation (B.5), the leading term is

$$\begin{aligned}
&\left(n\left\{\hat{\mathbb{V}}(t)+O_{\mathbb{P}}(1/n)\right\}^2/\vartheta_{\mathrm{c}}^{2}(t)\right)\left(n\,\vartheta_{10}^{2}(t)/n_1+n\,\vartheta_{20}^{2}(t)/n_2\right)\\
&=\frac{1}{\left(\vartheta_{\mathrm{c}}^{2}(t)\right)^2}\left\{n^{1/2}\hat{\mathbb{V}}(t)+o_{\mathbb{P}}(1)\right\}^2\vartheta_{\mathrm{c}}^{2}(t)+o_{\mathbb{P}}(1) \;=\; \frac{1}{\vartheta_{\mathrm{c}}^{2}(t)}\left\{n^{1/2}\hat{\mathbb{V}}(t)\right\}^2+o_{\mathbb{P}}(1),
\end{aligned}$$

uniformly over $[\alpha_1,\alpha_2]$. The subsequent terms of $-2R_{\hat{\boldsymbol{\theta}}}(t)$ are proportional to

$$(\hat{\boldsymbol{\gamma}}_{\hat{\boldsymbol{\theta}}}(t))^l\left[\sum_{i=1}^{2}\sum_{j=1}^{\hat{\eta}_i(t)}\left(\frac{1}{(r_{ij}-d_{ij})^{l-1}}-\frac{1}{r_{ij}^{l-1}}\right)\right], \qquad l=3,4,\ldots,$$

each of which is $o_{\mathbb{P}}(1)$, uniformly for $t\in[\alpha_1,\alpha_2]$. For example, when $l=3$, we have

$$\begin{aligned}
&\frac{2}{3}\,(\hat{\boldsymbol{\gamma}}_{\hat{\boldsymbol{\theta}}}(t))^3\sum_{j=1}^{\hat{\eta}_i(t)}\left(\frac{1}{r_{ij}-d_{ij}}-\frac{1}{r_{ij}}\right)\left(\frac{1}{r_{ij}-d_{ij}}+\frac{1}{r_{ij}}\right)\\
&=O_{\mathbb{P}}(n^{3/2})\,O_{\mathbb{P}}\left(\max_{1\le j\le\hat{\eta}_i(t)}\left|\frac{1}{r_{ij}-d_{ij}}\right|+\max_{1\le j\le\hat{\eta}_i(t)}\left|\frac{1}{r_{ij}}\right|\right)\,n_i\hat{\vartheta}_{i0}^{2}(t)\\
&=O_{\mathbb{P}}(n^{1/2})O_{\mathbb{P}}(n_i^{-1})O_{\mathbb{P}}(n_i^{-1}) \;=\; o_{\mathbb{P}}(1). \qquad \square
\end{aligned}$$

REFERENCES

[1] G. C. D. Adamson and D. J. Nash. Long-term variability in the date of monsoon onset over western india. *Climate Dynamic*, 40:2589–2603, 2013.

[2] P. K. Andersen, O. Borgan, R. D. Gill, and N. Keiding. *Statistical Models Based on Counting Processes*. New York, NY: Springer, 1993.

[3] P. J. Bickel and A. Sakov. On the choice of m in the m out of n bootstrap and confidence bounds for extrema. *Statistica Sinica*, 18:967–985, 2008.

[4] S. Csörgô. Limit behaviour of the empirical characteristic function. *Annals of Probability*, 9:130–144, 1981.

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

[6] T. W. Epps and L. B. Pulley. A test for normality based on the empirical characteristic function. *Biometrika*, 70:723–726, 1983.

[7] T.W. Epps and K. J. Singleton. An omnibus test for the two-sample problem using the empirical characteristic function. *Journal of Statistical Computation and Simulation*, 26:177–203, 1986.

[8] A. Feuerverger and P. McDunnough. The empirical characteristic function and its applications. *The Annals of Statistics*, 5:88–97, 1981.

[9] A. Feuerverger and R. A Mureika. On the efficiency of empirical characteristic function procedures. *Procedures," Journal of the Royal Statistical Society*, 43:20–27, 1977.

[10] D. Gerhard and L. A. Hothorn. Rank transformation in haseman-elston regression using scores for location-scale alternatives. *Human Heredity*, 69:143–151, 2010.

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

[12] P. Hall and A. Welsh. A test for normality based on the empirical characteristic function. *Biometrika*, 70:485–489, 1983.

[13] N. L. Hjort, I. W. McKeague, and I. V. Keilegom. Extending the scope of empirical likelihood. *Annals of Statistics*, 37:1079–1111, 2009.

[14] Huskovâ and Meintanis. Tests for the multivariate k -sample problem based on the empirical characteristic function. *Journal of Nonparametric Statistics*, 20:263–277, 2008.

[15] J. M. Ichida, J. T. Wassell, M. D. Keller, and L. W. Ayers. Evaluation of protocol change in burn-care management using the cox proportional hazards model with time-dependent covariates. *Statistics in Medicine*, 12:301–310, 1993.

[16] J. D. Kalbfleisch and R. L. Prentice. *Statistical Analysis of Failure Time Data*. Hoboken, NJ: Wiley, 2002.

[17] I.A. Koutrouvelis. Distribution-free procedures for location and symmetry inference problems based on the empirical characteristics function. *Scandinavian Journal of Statistics*, 12:257–269, 1985.

[18] I.A. Koutrouvelis and L. Kellermeier. A goodness-of-fit test based on the empirical characteristic function when parameters must be estimated. *Journal of the Royal Statistical Society*, 43:173–176, 1981.

[19] R. Kyle, T. Therneau, V. Rajkumar, J. Offord, D. Larson, M. Plevak, and L. J Melton. A long-term study of prognosis in monoclonal gammopathy of undetermined significance. *New England Journal of Medicine*, 346:564–569, 2002.

[20] W. Kössler and A. Mukherjee. Distribution-free simultaneous tests for location–scale and lehmann alternative in two-sample problem. *Biometrical*, 62:99–123, 2019.

[21] G. Li. On nonparametric likelihood ratio estimation of survival probabilities for censored data. *Statistics and Probability Letters*, 25:95–104, 1995.

[22] G. Li. Nonparametric likelihood ratio goodness of fit tests for survival data. *Journal of Multivariate Analysis*, 86:166–182, 2003.

[23] A. Lunde and A. Timmermann. Duration dependence in stock prices: an analysis of bull and bear markets. *Journal of Business and Economic Statistics*, 22:253–273, 2004.

[24] P. Major and L. Rejto. Strong embedding of the estimator of the distribution function under random censorship. *Annals of Statistics*, 16:1113–1132, 1998.

[25] M. B. Marcus. Weak convergence of the empirical characteristic function. *Annals of Probability*, 9:194–201, 1981.

[26] M. Matsui and A. Takemura. Empirical characteristic function ap proach to goodness-of-fit tests for the cauchy distribution with parame ters estimated by mle or eise. *Test*, 57:183–199, 2005.

[27] I. W. McKeague and Y. Zhao. Comparing distribution functions via empirical likelihood. *Journal of Biostatistics*, 1, 2005.

[28] S. Meintanis and J. Swanepoel. Bootstrap goodness-of-fit tests with estimated parameters based on empirical transforms. *Statistics and Probability Letter*, 77:1004–1013, 2007.

[29] M. Neuhauser, J. Kotzmann, M. Walier, and R. Poulin. The comparison of mean crowding between two groups. *Journal of Parasitology*, 96:477–481, 2010.

[30] C. J. Potgieter and F. Lombard. Estimation of linear treatment effects from matched pair data. *South African Statistical*, 42:17–45, 2008.

[31] W. Stute. The central limit theorem under random censorship. *Annals of Statistics*, 23:422–439, 1995.

[32] S. Subramanian. Function-based hypothesis testing in censored two-sample location-scale models. *Lifetime Data Analysis*, 26:183–213, 2020.

[33] D. R. Thomas and G. L. Grunkemeier. Confidence interval estimation of survival probabilities for censored data. *Journal of the American Statistical Association*, 70:865–871, 1975.

[34] J. Tortora, S. Robinson, S. Baker, and G. Ruaño. Clinical dataset of the cyp-guides trial. *Mendeley. Data.*, 1:253–273, 2020.

[35] Z. Ying, S. H. Jung, and L. J. Wei. Survival analysis with median regression models. *Journal of the American Statistical Association*, 90:178–184, 1995.

[36] Z. Zhang and Q. Yu. A minimum distance estimation approach to the two-sample location-scale problem. *Lifetime Data Anal*, 8:289–305, 2002.