INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

University of Alberta

NONPARAMETRIC TESTS FOR CHANGE-POINT PROBLEMS WITH RANDOM CENSORSHIP

By

SHUANGQUAN LIU ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of **Doctor of Philosophy**

in

Statistics

Department of Mathematical Sciences

Edmonton, Alberta

Fall 1998

National Library of Canada
Bibliothèque nationale du Canada

Acquisitions and Bibliographic Services
Acquisitions et services bibliographiques

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-34803-2

Canada

University of Alberta

Library Release Form

Name of Author: Shuangquan Liu

Title of Thesis: Nonparametric Tests for Change-point Problems with Random Censorship

Degree: Doctor of Philosophy

Year this Degree Granted: 1998

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly, or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as hereinbefore provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

(Signed)________________

Permanent Address:

Dept of Computer Science

York University

4700 Keele Street

Toronto, Ontario

Canada M3J 1P3

Date: July 17, 1998

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **Nonparametric Tests for Change-point Problems with Random Censorship** submitted by **Shuangquan Liu** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy** in **Statistics**.

Edit Gombay
E. Gombay (Supervisor)

D. P. Wiens

Y. Wu

N. G. N. Prasad

F. Yeh (Dept. of Renewable Resources)

for L. Horváth (External Examiner, Univ. of Utah)

Date: May 12, 1998

Abstract

There are several papers and monographs on change-point analysis and on random censorship. However, there are only very few results when these two different topics are connected. In this thesis we deal with the change-point problem when a sequence of interested variables are not completely observed but rather, are randomly right-censored. The change-point problem considered here is indeed very important, especially in medical and reliability (quality control) settings.

We propose a new procedure of testing whether a change-point occured in randomly censored data. The procedure is based on an extension of Wilcoxon's rank statistics. We investigate the asymptotic distribution of the test statistic under the null hypothesis of no change and also under the alternative hypothesis of one change. One appealing property of the resulting test is that the critical values are easily obtained through the use of a computer.

Our discussion from Chapter 1 through Chapter 4 is based on the generalization of the Wilcoxon's rank statistics which introduces a special anti-symmetric score function. In Chapter 5 we generalize the U-statistics based test to the general symmetric or anti-symmetric kernel case to investigate the asymptotics under the alternative hypothesis of one change, and the distribution of the maximally selected U-statistic is found to be asymptotically normal.

In order to investigate and apply our test procedure, we did a series of simulations on the power of the test as well as the performance of proposed change-point

estimators. Also, we demonstrate the usefulness of the procedure on two well-known data sets, the Stanford heart transplant data and the Radiation Therapy Oncology Group data. The conclusions from these applications are consistent with some analysis using the Cox proportional hazards model.

To Yan and Mina

Acknowledgements

I am very grateful to my supervisor, Dr. E. Gombay, for her immense support, extensive suggestions and comments, and consistent encouragement, patience, generousness and kindness throughout the preparation of this thesis. Without her invaluable guidance, the successful completion of this thesis would have been impossible.

The deserving of special thanks go to my external examiner, Dr. L. Horváth, and Dr. Y. Wu for their thoughtful and helpful remarks, suggestions, and advice.

Many thanks also go to Dr. D. P. Wiens, Dr. N. G. N. Prasad, and Dr. F. Yeh for their valuable comments and support.

My final acknowledgement is to the department secretaries for their helpfulness and friendship.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Review and Problems

Historically, change-point problems originated from quality control. There, one typically observes the output of a production process, the quality of which may be measured in terms of a certain characteristic. We are interested in possible changes of the underlying stochastic mechanism. Such a situation can usually be modeled by saying that we have a random process that generates independent observations indexed by time and we wish to detect whether a change could have occurred in the distribution that governs this random process as time goes by.

There is already a rich literature dealing with change-point problems from both parametric and nonparametric points of view. For a review of historical perspectives of the classical change-point problems, we refer to Bhattacharya (1994), Csörgő and Horváth (1988), Wolfe and Schechtman (1984), and also see Csörgő and Horváth (1997) for an extensive reference list.

We note that, very little is known on the change-point problems under the case of random censoring which is certainly very important and is worth investigating. For example, Müller and Wang (1994) presented a review of parametric and nonparametric models and corresponding estimation procedures for change-points in hazard functions where the data are possibly subject to random censoring. In most of the published work to date the mathematical theories were developed for a failure time variable which is observable. This is of course rarely the case in reality. For example, in the case given in Matthews and Farewell (1982), which was subsequently analyzed by Worsley (1988) and Achcar (1989), data on 33 out of the 84 acute nonlymphoblastic leukemia patients were censored. Of those, 24 were censored at 182 days, when the patients were randomized to an experimental protocol. In another example in Matthews, Farewell and Pyke (1985), 11 out of the 31 advanced non-Hodgkin's lymphoma patients were still alive at the last time of follow up and were thus censored. Matthews and Farewell (1982) claimed that dropping the 24 censored observations at 182 days did not affect significantly the outcome of the likelihood ratio test. Subsequently, most work develops theory for the case of observable time variables and in applications, the censored observations are either discarded or the likelihood function is modified for censored data. Since the results under censoring may differ from those of the uncensored case such an approach is questionable, Loader (1991) presented a discussion of the effect of random censorship.

We know that change-point problems are the generalization of problems of comparing two or more samples. An interesting fact is that the two-sample (or more) problems with censored data have been studied by many authors. For a review see, for example, Gehan (1965a, 1965b), Mantel (1967), Efron (1967), Nelson (1969), Breslow (1970), R. Peto and J. Peto (1972), Prentice (1978), Prentice and Marek (1979), Fleming et al. (1980), Andersen et al. (1982), Harrington and Fleming (1982), Leurgans (1983, 1984), Breslow et al. (1984), Schumacher(1984), Wei (1984), Gill and Schumacher (1987).

In this thesis, we are about to combine these two types of problems getting change-point analysis with randomly censored data.

1.2 Randomly Censored Data and Test

Let $T_1, \cdots, T_n$ be a sequence of independent continuous random variables. We want to test the no-change null hypothesis

$H_0 : T_1, \cdots, T_n$ have the same distribution function (d.f.) F

against the one-change alternative hypothesis

H_1: there exists some $\lambda \in (0,1)$ such that

$$P\{T_1 \le t\} = \cdots = P\{T_{[n\lambda]} \le t\} = F^{(1)}(t),$$

$$P\{T_{[n\lambda]+1} \le t\} = \cdots = P\{T_n \le t\} = F^{(2)}(t) \text{ for all } t, \text{ and}$$

$$P\{T_{[n\lambda]} \le t_0\} \neq P\{T_{[n\lambda]+1} \le t_0\} \text{ for some } t_0.$$

We call this test the at-most-one-change (AMOC) change-point problem.

In survival analysis the data of interest typically are measurements on time elapsing between the occurrence of two events. For example, T may be the time until death a patient has spent in a follow-up study. The index of T corresponds to the chronological order in which one has entered the group. In such a situation, it is of interest to know if after some time, possibly due to an improved medical care, there has been a change in the survival distribution.

However, due to other causes of failure it may happen that the variables of interest T_i are not completely observable but right censored by random variables $C_i, 1 \leq i \leq n$. In other words, instead of $T_1, \cdots, T_n$ one can observe $(X_1, \delta_1), \cdots, (X_n, \delta_n)$ only, where

$$X_i = \min(T_i, C_i) = \begin{cases} T_i, & \text{if } T_i \leq C_i \text{, we say, } X_i \text{ is uncensored,} \\ C_i, & \text{if } T_i > C_i \text{, we say, } X_i \text{ is censored.} \end{cases} \tag{1.1}$$

and

$$\delta_i = I(T_i \leq C_i) = \begin{cases} 1 & \text{if } X_i \text{ is uncensored,} \\ 0 & \text{if } X_i \text{ is censored.} \end{cases} \tag{1.2}$$

The variable δ indicates whether T has been observed or not.

Random censoring arises in medical applications with animal studies or clinical trials. In a clinical trial, patients may enter the study at different times; then each is treated with one of several possible therapies. We want to observe their lifetimes, but censoring occurs in one of the following forms:

1. Loss to follow-up. The patient may decide to move elsewhere; we never see him again.

2. Drop out. The therapy may have such bad side effects that it is necessary to discontinue the treatment or the patient may still be in contact (he hasn't moved), but has refused to continue the treatment.

3. Termination of the study.

With random censoring we will make the following assumption:

Assumption 1.1. $C_1, \cdots, C_n$ is a sequence of independent and identically distributed (i.i.d.) and continuous random variables with censoring distribution function G. Furthermore, T_i and C_i are independent $(1 \leq i \leq n)$.

Please note that $\lambda, F, F^{(1)}, F^{(2)}$, and G are unknown.

By the independence of T_i and C_i, $X_1, \cdots, X_n$ have the same distribution function H under H_0, where

$$\begin{aligned} H(x) &= P_{H_0}\{X_1 \leq x\} \\ &= 1 - (1 - F(x))(1 - G(x)). \end{aligned} \tag{1.3}$$

Under H_1, $X_1, \cdots, X_{[n\lambda]}$ have the same distribution function $H^{(1)}$ and $X_{[n\lambda]+1}, \cdots$, X_n have the same distribution function $H^{(2)}$, where

$$\begin{aligned} H^{(1)}(x) &= P_{H_1}\{X_1 \leq x\} \\ &= 1 - (1 - F^{(1)}(x))(1 - G(x)) \end{aligned} \tag{1.4}$$

and

$$\begin{aligned} H^{(2)}(x) &= P_{H_1}\{X_n \leq x\} \\ &= 1 - (1 - F^{(2)}(x))(1 - G(x)). \end{aligned} \tag{1.5}$$

Since any change in the distribution function F results in a change in the distribution function H, one might argue that a method designed for detecting a change from a sequence of completely observable data could be applied to the X's as well to detect a change in the F's. This is true in principle. On the other hand, such a procedure would necessarily not incorporate the information contained in the δ's and therefore lead to an inefficient procedure.

Although the literature of the AMOC change-point problems is extensive, the important case when $\{T_1, \cdots, T_n\}$ are randomly censored has not received much attention. We are aware of the work of Stute (1996) only, in which the author investigated an estimator for the change-point in censored data. In this thesis we will use different methods to set up some test statistics by utilizing the information contained in both the X's and the δ's. We will discuss the asymptotics and the weighted asymptotics of the test statistics, and also the powers of the tests under the null hypothesis in Chapter 2 and Chapter 3. We will discuss the change-point estimators and the asymptotic distributions of the test statistics under the alternative hypothesis in Chapter 4. Our discussion from Chapter 1 through Chapter 4 is based on a generalization of the Wilcoxon's rank statistics which introduces a special anti-symmetric score function. In Chapter 5 we will generalize these special U-statistics to general symmetric or anti-symmetric kernel based U-statistics and discuss their asymptotics under the alternative hypothesis of one change.

Chapter 2

Test Statistics and Asymptotics under H_0

2.1 Wilcoxon-type Statistics

A generalization of Wilcoxon's test for comparing two samples has been proposed by Gehan (1965a) for use when the observations are subject to random right censorship. Mantel (1967), as well as Gehan (1965b), has considered a further generalization to the case of random two-sided censorship and simplified the calculations. Both Gehan (1965a) and Mantel (1967) proposed their non-parametric tests for comparing two samples under the additional assumption of identical censoring distributions.

Let

$$Z_i = (X_i, \delta_i), \ i = 1, \cdots, n, \tag{2.1}$$

where X_i and δ_i are defined in (1.1) and (1.2), respectively. Note that $Z_1, \cdots, Z_n$ is a sequence of two-dimensional independent random vectors.

We define the score function for comparing two observations Z_i and Z_j by

$$h(Z_i, Z_j) = \begin{cases} 1, & \text{if } (X_i > X_j, \delta_j = 1) \text{ or } (X_i = X_j, \delta_i = 0, \delta_j = 1), \\ -1, & \text{if } (X_i < X_j, \delta_i = 1) \text{ or } (X_i = X_j, \delta_i = 1, \delta_j = 0), \\ 0, & \text{otherwise.} \end{cases} \tag{2.2}$$

This score function was proposed by Gehan (1965a). His statistic for comparing two samples $\{Z_1, \cdots, Z_{[nt]}\}$ and $\{Z_{[nt]+1}, \cdots, Z_n\}$ is

$$U^*_{[nt]} = \sum_{i=1}^{[nt]} \sum_{j=[nt]+1}^{n} h(Z_i, Z_j), \quad 0 \le t \le 1. \tag{2.3}$$

Following Mantel (1967) we can write

$$U_i = \sum_{j=1}^{n} h(Z_i, Z_j),\ i = 1, \cdots, n, \tag{2.4}$$

and we have

$$U^*_{[nt]} = \sum_{i=1}^{[nt]} U_i\,, \quad 0 \le t \le 1 \tag{2.5}$$

as

$$h(Z_i, Z_j) = -h(Z_j, Z_i), \tag{2.6}$$

i.e. h is anti-symmetric and so

$$\sum_{i=1}^{[nt]} \sum_{j=1}^{n} h(Z_i, Z_j) = \sum_{i=1}^{[nt]} \sum_{j=[nt]+1}^{n} h(Z_i, Z_j), \quad 0 \le t \le 1,$$

or

$$\sum_{i=1}^{k} \sum_{j=1}^{n} h(Z_i, Z_j) = \sum_{i=1}^{k} \sum_{j=k+1}^{n} h(Z_i, Z_j), \quad 1 \le k \le n, \tag{2.7}$$

where $h(Z_i, Z_j)$'s cancel each other out in the sum of left hand side for $1 \le i, j \le k$.

Mantel's approach simplified both the method of computation and the determination of the permutation distribution of $U^*_{[nt]}$ and its variance, as well as the proof of asymptotic normality.

Note that (1.1) implies the convention in random right censorship that if an uncensored observation X_i and a censored observation X_j are tied, we consider the uncensored X_i to occur just before the censored X_j, i.e. we break the tie by considering $X_i < X_j$. On the other hand, since $X_1, \cdots, X_n$ are independent continuous random variables, the probability of ties is equal to zero, so for simplicity and convenience in calculations and exposition, we assume no ties without loss of generality. Therefore the score function can be simplified as

$$h(Z_i, Z_j) = I(X_i > X_j, \delta_j = 1) - I(X_i < X_j, \delta_i = 1). \tag{2.8}$$

We can think of the AMOC change-point problems as a series of two-sample problems under censored data, the following discussion is based on statistics (2.4), (2.5), and score function (2.8).

2.2 A Test Based on Exchangeable Variables

We say random variables $\xi_1, \cdots, \xi_n$ are exchangeable if each permutation of the set has the same joint distribution. We have

Theorem 2.1. *Under H_0,*

$$\frac{\sum_{i=1}^{[nt]} U_i}{\sqrt{\sum_{i=1}^{n} U_i^2}} \xrightarrow{\mathcal{D}} B(t),\ 0 \le t \le 1, \tag{2.9}$$

and

$$\sup_{0<t<1} \frac{|\sum_{i=1}^{[(n+1)t]} U_i|}{\sqrt{\sum_{i=1}^{n} U_i^2}} \xrightarrow{\mathcal{D}} \sup_{0<t<1} |B(t)|, \tag{2.10}$$

where U_i is defined in (2.4) *and $\{B(t), 0 \le t \le 1\}$ is a Brownian bridge.*

Proof. Note that the following proof is based on the Theorem 24.2 in Billingsley (1968), which says, if $\xi_1, \cdots, \xi_n$ are exchangeable and satisfy

$$\sum_{i=1}^{n} \xi_i \xrightarrow{P} 0, \quad \sum_{i=1}^{n} \xi_i^2 \xrightarrow{P} 1, \quad \max_{1 \le i \le n} |\xi_i| \xrightarrow{P} 0,$$

then $\sum_{i=1}^{[nt]} \xi_i \xrightarrow{\mathcal{D}} B(t)$. Define

$$Y_i = \frac{U_i}{\sqrt{\sum_{i=1}^n U_i^2}}, \quad 1 \le i \le n.$$

By (2.4) and Assumption 1.1, $Y_1, \cdots, Y_n$ are exchangeable random variables under H_0. As $h(Z_i, Z_j)$ given by (2.8) is anti-symmetric and

$$\sum_{i=1}^n U_i = \sum_{i=1}^n \sum_{j=1}^n h(Z_i, Z_j) = 0,$$

we have

$$\sum_{i=1}^n Y_i = 0.$$

Also,

$$\sum_{i=1}^n Y_i^2 = \frac{1}{\sum_{i=1}^n U_i^2} \sum_{i=1}^n U_i^2 = 1.$$

On the other hand, define the empirical distribution function of $X_1, \cdots, X_n$ by

$$H_n(t) = \frac{1}{n} \sum_{i=1}^n I(X_i \le t), \tag{2.11}$$

and the empirical subdistribution function of $X_1, \cdots, X_n$ by

$$\widetilde{H}_n(t) = \frac{1}{n} \sum_{i=1}^n I(X_i \le t, \delta_i = 1). \tag{2.12}$$

Then by (2.4) and (2.8)

$$\begin{aligned} U_i &= \sum_{j=1}^n I(X_i > X_j, \delta_j = 1) - \sum_{j=1}^n I(X_i < X_j, \delta_i = 1) \\ &= \sum_{j=1}^n I(X_j < X_i, \delta_j = 1) - \sum_{j=1}^n [1 - I(X_j \le X_i)]\delta_i \\ &= n\widetilde{H}_n(X_i-) - n[1 - H_n(X_i)]\delta_i \\ &= n\left\{\widetilde{H}_n(X_i-) - [1 - H_n(X_i)]\delta_i\right\}. \end{aligned} \tag{2.13}$$

Thus,

$$
\begin{aligned}
&\frac{1}{n^3}\sum_{i=1}^{n} U_i^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\left\{\widetilde{H}_n(X_i-) - [1-H_n(X_i)]\delta_i\right\}^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\widetilde{H}_n^2(X_i-) - \frac{2}{n}\sum_{i=1}^{n}\widetilde{H}_n(X_i-)[1-H_n(X_i)]\delta_i + \frac{1}{n}\sum_{i=1}^{n}[1-H_n(X_i)]^2\delta_i \\
&= \int \widetilde{H}_n^2(t-)dH_n(t) - 2\int \widetilde{H}_n(t-)(1-H_n(t))d\widetilde{H}_n(t) + \int (1-H_n(t))^2 d\widetilde{H}_n(t).
\end{aligned} \tag{2.14}
$$

Since

$$
H_n(t) \xrightarrow{a.s.} H(t) \quad \text{and} \quad \widetilde{H}_n(t) \xrightarrow{a.s.} \widetilde{H}(t)
$$

uniformly in t as $n \to \infty$ by the Glivenko-Cantelli theorem , where $H(t)$ is defined in (1.3) and $\widetilde{H}(t)$ is the subdistribution function of X's, i.e.

$$
\widetilde{H}(t) = P_{H_0}\{X_1 \le t, \delta_1 = 1\} = \int_{-\infty}^{t}(1-G(u))dF(u), \tag{2.15}
$$

it follows that as $n \to \infty$

$$
\frac{1}{n^3}\sum_{i=1}^{n} U_i^2 \xrightarrow{a.s.} \int \widetilde{H}^2(t)dH(t) - 2\int \widetilde{H}(t)(1-H(t))d\widetilde{H}(t) + \int (1-H(t))^2 d\widetilde{H}(t).
$$

Integration by parts gives

$$
\begin{aligned}
2\int \widetilde{H}(t)(1-H(t))d\widetilde{H}(t) &= \widetilde{H}^2(t)(1-H(t))\Big|_{-\infty}^{\infty} + \int \widetilde{H}^2(t)dH(t) \\
&= \int \widetilde{H}^2(t)dH(t),
\end{aligned}
$$

so the first two terms in the above limiting expression cancel. Thus

$$
\frac{1}{n^3}\sum_{i=1}^{n} U_i^2 \xrightarrow{a.s.} \int (1-H(t))^2 d\widetilde{H}(t) > 0. \tag{2.16}
$$

By (2.4) and (2.8)

$$
\begin{aligned}
\max_{1\le i\le n}|U_i| &= \max_{1\le i\le n}\left|\sum_{j=1}^{n} h(Z_i, Z_j)\right| \\
&\le \max_{1\le i\le n}\sum_{j=1}^{n}|h(Z_i, Z_j)| \\
&\le n.
\end{aligned}
$$

Thus, using (2.16), we get

$$\max_{1\le i\le n}|Y_i| = \frac{1}{\sqrt{\sum_{i=1}^n U_i^2}}\max_{1\le i\le n}|U_i|$$
$$\le \frac{n}{\sqrt{\sum_{i=1}^n U_i^2}}$$
$$= \frac{\frac{1}{\sqrt{n}}}{\sqrt{\frac{1}{n^3}\sum_{i=1}^n U_i^2}} \to 0 \text{ (as } n\to\infty).$$

Therefore the variables $\{Y_i\}$ satisfy the conditions of Theorem 24.2 in Billingsley (1968) under H_0, so

$$\frac{\sum_{i=1}^{[nt]} U_i}{\sqrt{\sum_{i=1}^n U_i^2}} = \sum_{i=1}^{[nt]} Y_i \xrightarrow{\mathcal{D}} B(t),\ 0\le t\le 1.$$

Relation (2.10) then follows from (2.9) and the theorem in Donsker (1952) (also, cf. Theorem 4.2.1 in Csörgő and Révész (1981)). □

We see that Theorem 2.1 gives the asymptotic distribution of the test statistic under the null hypothesis of no change in distribution. The distribution of $\sup_{0<t<1}|B(t)|$ is the well-known Kolmogorov-Smirnov-distribution (cf. Kolmogorov (1933)), that is,

$$P\{\sup_{0<t<1}|B(t)|>b\} = 2\sum_{i=1}^{\infty}(-1)^{i-1}\exp(-2i^2b^2), b\ge 0. \tag{2.17}$$

Thus we obtain a test statistic

$$\sup_{0<t<1}\frac{|\sum_{i=1}^{[(n+1)t]}U_i|}{\sqrt{\sum_{i=1}^n U_i^2}} = \max_{1\le k\le n}\frac{|\sum_{i=1}^k U_i|}{\sqrt{\sum_{i=1}^n U_i^2}}, \tag{2.18}$$

the approximated critical values of the test can be given by (2.17). Let c_α denote the $(1-\alpha)$-quantile of $\sup_{0<t<1}|B(t)|$; the test is defined by

$$I\left(\max_{1\le k\le n}\frac{|\sum_{i=1}^k U_i|}{\sqrt{\sum_{i=1}^n U_i^2}} > c_\alpha\right). \tag{2.19}$$

We reject H_0 vs H_1 if the value of the test statistic exceeds the critical value c_α at significance level α.

By use of (2.17) one can compute c_α and the P-values via computer. We list the critical values of c_α for some selected values of α in the following Table 2.1, and we'll use Table 2.1 in the following simulation study. We do a simulation study on the comparison of the approximated and simulated critical values, and the power of the test in the next section. See Section 4.4 for further applications.

Table 2.1. *Some selected critical values c_α*

α	0.175	0.15	0.125
c_α	1.104	1.138	1.177
α	0.1	0.075	0.05
c_α	1.224	1.281	1.358
α	0.025	0.01	0.0075
c_α	1.480	1.628	1.671
α	0.005	0.0025	0.001
c_α	1.731	1.828	1.949

Remark 2.1. If one is looking at the epidemic alternative hypothesis

H_1': there exist $\tau_1, \tau_2, 1 \le \tau_1 < \tau_2 < n$, such that

$$\begin{aligned} &P\{T_1 \le t\} = \cdots = P\{T_{\tau_1} \le t\} = P\{T_{\tau_2+1} \le t\} = \cdots = P\{T_n \le t\} = F^{(1)}(t), \\ &P\{T_{\tau_1+1} \le t\} = \cdots = P\{T_{\tau_2} \le t\} = F^{(2)}(t) \text{ for all } t, \text{ and} \\ &F^{(1)}(t_0) \ne F^{(2)}(t_0) \text{ for some } t_0, \end{aligned} \tag{2.20}$$

we should use statistic

$$\sup_{0<t_1<t_2<1} \frac{\left| \sum_{i=[(n+1)t_1]+1}^{[(n+1)t_2]} U_i \right|}{\sqrt{\sum_{i=1}^{n} U_i^2}}.$$

By Theorem 2.1, we have under H_0

$$\sup_{0<t_1<t_2<1} \frac{\left|\sum_{i=[(n+1)t_1]+1}^{[(n+1)t_2]} U_i\right|}{\sqrt{\sum_{i=1}^n U_i^2}} = \sup_{0<t_1<t_2<1} \left| \frac{\sum_{i=1}^{[(n+1)t_2]} U_i}{\sqrt{\sum_{i=1}^n U_i^2}} - \frac{\sum_{i=1}^{[(n+1)t_1]} U_i}{\sqrt{\sum_{i=1}^n U_i^2}} \right| \xrightarrow{\mathcal{D}} \sup_{0<t_1<t_2<1} |B(t_2) - B(t_1)|. \tag{2.21}$$

Thus one can use

$$\max_{1\leq k<l<n} \frac{\left|\sum_{i=k}^{l} U_i\right|}{\sqrt{\sum_{i=1}^n U_i^2}} \tag{2.22}$$

as a test statistic. The approximated critical values can be given by

$$P\Big\{\sup_{0<t_1<t_2<1} |B(t_2) - B(t_1)| > b\Big\} = 2\sum_{i=1}^{\infty}(4i^2b^2 - 1)\exp(-2i^2b^2),\ b \geq 0 \tag{2.23}$$

which is known as the limiting distribution of the Kuiper statistic (cf. Shorack and Wellner (1986)). Let c'_α denote the $(1-\alpha)$-quantile of $\sup_{0<t_1<t_2<1} |B(t_2) - B(t_1)|$; the test is defined by

$$I\left(\max_{1\leq k<l<n} \frac{\left|\sum_{i=k}^{l} U_i\right|}{\sqrt{\sum_{i=1}^n U_i^2}} > c'_\alpha\right). \tag{2.24}$$

We reject H_0 vs H'_1 if the value of test statistic exceeds the critical value c'_α at significance level α. We have $c'_\alpha = 2.00, 1.75$ and 1.62 for $\alpha = 0.01, 0.05$ and 0.10, respectively.

2.3 Simulation Study

To illustrate the proposed test, we would like to check the precision of the approximation and the power of the test through the Monte Carlo simulation study. We performed $N = 5000$ simulations for each case. In each case to be considered, we assumed that

$$F = \exp(\mu), \quad F^{(1)} = \exp(\mu_1), \quad F^{(2)} = \exp(\mu_2), \quad G = \exp(\mu_c).$$

That is, the simulated variables of interest T_i were exponentially distributed with mean μ under the null hypothesis of no change in distribution and exponentially distributed with mean μ_1 before change point λ and mean μ_2 after change under the alternative hypothesis of one change in distribution; the simulated censoring variables C_i were exponentially distributed with mean μ_c. Note that we obtained the simulated data $Z_i = (X_i, \delta_i)$ by (1.1) and (1.2), and the score function values $h(Z_i, Z_j)$ by (2.2) in order to handle the tied values in case. We took $\mu_1 = 1.0$ and $\mu_c = 3.0$ for all cases.

First, we did simulation to compare the simulated critical values with the approximated critical values for the test statistic (2.18). We considered sample sizes $n =$ 50, 100, 200, 500 and means $\mu =$ 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, significance levels $\alpha =$ 0.01, 0.05, 0.10. The results are reported in Table 2.2. It turned out that the critical value obtained from the limiting distribution overestimates the true one somehow, but is roughly close, thus we can use the approximated critical values to the proposed test and to estimate the power of the test reasonably.

Next, we did simulation to investigate the power of test (2.19). We performed simulation in two ways. One way was dealing with different sample sizes but the same location of change-point, i.e. $\lambda = 0.5$ so that pre- and after-change samples are balanced in size. We considered $n =$ 50, 100, 200, 500. The results are reported in Table 2.3. Another way simulation was performed was dealing

Table 2.2. *Comparison of simulated and approximated c.v. for statistic* (2.18)

n	μ	($\alpha = 0.01$) SimCV	AppCV	($\alpha = 0.05$) SimCV	AppCV	($\alpha = 0.10$) SimCV	AppCV
50	1.0	1.5570	1.6276	1.2963	1.3581	1.1399	1.2238
	1.5	1.5398	1.6276	1.2801	1.3581	1.1476	1.2238
	2.0	1.5159	1.6276	1.3001	1.3581	1.1506	1.2238
	2.5	1.5044	1.6276	1.2826	1.3581	1.1403	1.2238
	3.0	1.5112	1.6276	1.2892	1.3581	1.1359	1.2238
	3.5	1.5306	1.6276	1.2724	1.3581	1.1480	1.2238
100	1.0	1.5746	1.6276	1.3151	1.3581	1.1648	1.2238
	1.5	1.5656	1.6276	1.3095	1.3581	1.1641	1.2238
	2.0	1.5789	1.6276	1.2949	1.3581	1.1696	1.2238
	2.5	1.5650	1.6276	1.2878	1.3581	1.1742	1.2238
	3.0	1.5624	1.6276	1.3136	1.3581	1.1652	1.2238
	3.5	1.5471	1.6276	1.2839	1.3581	1.1835	1.2238
200	1.0	1.5675	1.6276	1.3208	1.3581	1.1781	1.2238
	1.5	1.5928	1.6276	1.3045	1.3581	1.1927	1.2238
	2.0	1.6067	1.6276	1.3168	1.3581	1.1810	1.2238
	2.5	1.5456	1.6276	1.3185	1.3581	1.2031	1.2238
	3.0	1.5893	1.6276	1.3193	1.3581	1.1713	1.2238
	3.5	1.5768	1.6276	1.3183	1.3581	1.1783	1.2238
500	1.0	1.5764	1.6276	1.3459	1.3581	1.1907	1.2238
	1.5	1.5999	1.6276	1.3368	1.3581	1.1925	1.2238
	2.0	1.6286	1.6276	1.3420	1.3581	1.2011	1.2238
	2.5	1.6072	1.6276	1.3542	1.3581	1.1964	1.2238
	3.0	1.6623	1.6276	1.3317	1.3581	1.2068	1.2238
	3.5	1.5978	1.6276	1.3449	1.3581	1.1963	1.2238

with different locations of change-points but with the same sample size. We assumed $n = 100$, and considered $\lambda = 0.1$, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, the results are reported in Table 2.4. We considered means $\mu_2 = 1.5$, 2.0, 2.5, 3.0, 3.5 and $\alpha = 0.01$, 0.05, 0.1 for each way.

From Table 2.3, we can see that the power increases as the distance of mean $|\mu_2 - \mu_1|$ increases. When $n \geq 100$ and $|\mu_2 - \mu_1| \geq 1.5$, the power is pretty large.

Table 2.3. *Simulated power of test* (2.19) *for* $\lambda = 0.5$

n	μ_1	μ_2	μ_c	($\alpha = 0.01$) Simulated Power	($\alpha = 0.05$) Simulated Power	($\alpha = 0.10$) Simulated Power
50	1.0	1.5	3.0	0.0304	0.0982	0.1818
	1.0	2.0	3.0	0.1052	0.2598	0.3768
	1.0	2.5	3.0	0.2144	0.4384	0.5624
	1.0	3.0	3.0	0.3188	0.5748	0.6934
	1.0	3.5	3.0	0.4242	0.6800	0.7774
100	1.0	1.5	3.0	0.0854	0.2198	0.3170
	1.0	2.0	3.0	0.3032	0.5372	0.6628
	1.0	2.5	3.0	0.5592	0.7778	0.8638
	1.0	3.0	3.0	0.7582	0.9050	0.9482
	1.0	3.5	3.0	0.8714	0.9530	0.9758
200	1.0	1.5	3.0	0.2120	0.4296	0.5402
	1.0	2.0	3.0	0.6996	0.8664	0.9236
	1.0	2.5	3.0	0.9368	0.9784	0.9912
	1.0	3.0	3.0	0.9870	0.9966	0.9990
	1.0	3.5	3.0	0.9982	0.9996	1.0000
500	1.0	1.5	3.0	0.6502	0.8384	0.9076
	1.0	2.0	3.0	0.9924	0.9990	0.9998
	1.0	2.5	3.0	1.0000	1.0000	1.0000
	1.0	3.0	3.0	1.0000	1.0000	1.0000
	1.0	3.5	3.0	1.0000	1.0000	1.0000

From Table 2.4, we can see that the power almost always reaches the largest value when the change occurs in the middle ($\lambda = 0.5$), the power decreases when the change goes toward the two tails ($\lambda = 0.1$ or $\lambda = 0.9$), and power decreases very quickly and is small when the change occurs close to the two tails, so the test is not sensitive on the tails. Of course, this is not too surprising since a change occuring near the middle of a sequence of observations should be much easier to detect than one occuring at the beginning or the end of the sequence. In Chapter 3, we introduce weight functions that may remedy this situation somewhat on the tails.

Table 2.4. *Simulated power of test* (2.19) *for* $n = 100$

μ_1	μ_2	μ_c	λ	($\alpha = 0.01$) Simulated Power	($\alpha = 0.05$) Simulated Power	($\alpha = 0.10$) Simulated Power
1.0	1.5	3.0	0.1	0.0106	0.0544	0.1018
			0.2	0.0264	0.0948	0.1632
			0.3	0.0540	0.1584	0.2396
			0.4	0.0722	0.2014	0.3084
			0.5	0.0692	0.2124	0.3220
			0.6	0.0686	0.1902	0.2890
			0.7	0.0412	0.1476	0.2296
			0.8	0.0180	0.0840	0.1594
			0.9	0.0102	0.0448	0.0986
1.0	2.0	3.0	0.1	0.0170	0.0798	0.1438
			0.2	0.0780	0.2366	0.3588
			0.3	0.1908	0.4158	0.5402
			0.4	0.2710	0.5276	0.6370
			0.5	0.3100	0.5326	0.6674
			0.6	0.2678	0.4850	0.6112
			0.7	0.1674	0.3780	0.5008
			0.8	0.0620	0.1966	0.3228
			0.9	0.0136	0.0704	0.1306
1.0	2.5	3.0	0.1	0.0260	0.1224	0.2114
			0.2	0.1806	0.4154	0.5518
			0.3	0.4044	0.6490	0.7616
			0.4	0.5314	0.7584	0.8434
			0.5	0.5592	0.7764	0.8618
			0.6	0.4892	0.7334	0.8308
			0.7	0.3430	0.5966	0.7284
			0.8	0.1172	0.3340	0.4830
			0.9	0.0194	0.0810	0.1550
1.0	3.0	3.0	0.1	0.0408	0.1540	0.2796
			0.2	0.3066	0.5706	0.7042
			0.3	0.5958	0.8108	0.8932
			0.4	0.7332	0.8866	0.9338
			0.5	0.7476	0.9074	0.9452
			0.6	0.6896	0.8778	0.8778
			0.7	0.5056	0.7498	0.8462
			0.8	0.1910	0.4496	0.6128
			0.9	0.0306	0.1048	0.1820

Table 2.4. *Simulated power of test* (2.19) *for* $n = 100$ (*Continued*)

μ_1	μ_2	μ_c	λ	($\alpha = 0.01$) Simulated Power	($\alpha = 0.05$) Simulated Power	($\alpha = 0.10$) Simulated Power
1.0	3.5	3.0	0.1	0.0588	0.2060	0.3448
			0.2	0.4356	0.7014	0.8154
			0.3	0.7372	0.9000	0.9514
			0.4	0.8474	0.9502	0.9784
			0.5	0.8684	0.9576	0.9800
			0.6	0.8130	0.9390	0.9726
			0.7	0.6418	0.8584	0.9138
			0.8	0.2566	0.5616	0.7134
			0.9	0.0296	0.1206	0.2190

Chapter 3

Weighted Asymptotics of Test Statistics under H_0

3.1 Tests Based on U-Statistics

For a survey on U-statistics, we refer to Serfling (1980) and Lee (1990).

By (2.3) and (2.5), we have

$$U_k^* = \sum_{i=1}^{k} U_i = \sum_{i=1}^{k} \sum_{j=k+1}^{n} h(Z_i, Z_j), \ 1 \leq k < n. \tag{3.1}$$

That is, U_k^* has the form of U-statistic with an anti-symmetric kernel h.

Our next purpose is to consider different weighted versions of our test statistics by investigating the asymptotic properties of the U-statistics based processes (3.1). Note that our U-statistics with an anti-symmetric kernel $h(Z_i, Z_j)$ have two-dimensional arguments instead of one-dimensional arguments. However, we'll be able to see that we can reduce the problem to one-dimension.

Since the kernel function $h(Z_1, Z_2)$ is anti-symmetric and $|h(Z_1, Z_2)| \leq 1$, it follows that

$$Eh(Z_1, Z_2) = 0$$

and

$$Eh^2(Z_1, Z_2) \leq 1.$$

Define

$$\widetilde{h}(\underline{t}) = Eh(Z_1, \underline{t}) \tag{3.2}$$

with $\underline{t} = (t_1, t_2)$, and

$$\sigma^2 = E\widetilde{h}^2(Z_2). \tag{3.3}$$

The function defined in (3.2) is the projection of U-statistics.

Lemma 3.1. *Under* H_0,

$$\sigma^2 = \int (1 - H(t))^2 d\widetilde{H}(t) > 0, \tag{3.4}$$

and

$$\frac{1}{n^3} \sum_{i=1}^{n} U_i^2 \xrightarrow{a.s} \sigma^2, \tag{3.5}$$

where $H(t), \widetilde{H}(t)$ *and* U_i *are defined in* (1.3), (2.15) *and* (2.4), *respectively.*

Proof. Under H_0, by (2.8) we have

$$\begin{aligned}
\widetilde{h}(\underline{t}) &= Eh(Z_1, \underline{t}) \\
&= E\{h(Z_1, Z_2) | Z_2 = \underline{t}\} \\
&= E\{I(X_1 > X_2, \delta_2 = 1) | X_2 = t_1, \delta_2 = t_2\} \\
&\qquad - E\{I(X_1 < X_2, \delta_1 = 1) | X_2 = t_1, \delta_2 = t_2\} \\
&= E\{I(X_1 > t_1, \delta_2 = 1) | \delta_2 = t_2\} - EI(X_1 < t_1, \delta_1 = 1) \\
&= P\{X_1 > t_1, \delta_2 = 1 | \delta_2 = t_2\} - P\{X_1 < t_1, \delta_1 = 1\} \\
&= P\{X_1 > t_1\} I(t_2 = 1) - \widetilde{H}(t_1) \\
&= (1 - H(t_1)) I(t_2 = 1) - \widetilde{H}(t_1),
\end{aligned}$$

and

$$\begin{aligned}\tilde{h}(Z_i) &= \tilde{h}(X_i, \delta_i) \\ &= (1 - H(X_i))I(\delta_i = 1) - \widetilde{H}(X_i) \qquad (3.6) \\ &= (1 - H(X_i))\delta_i - \widetilde{H}(X_i), \ i = 1, \cdots, n.\end{aligned}$$

Thus,

$$\begin{aligned}\sigma^2 &= E\tilde{h}^2(Z_2) \\ &= E\big\{(1 - H(X_2))^2\delta_2 - 2\widetilde{H}(X_2)(1 - H(X_2))\delta_2 + \widetilde{H}^2(X_2)\big\} \qquad (3.7) \\ &= \int(1 - H(t))^2 d\widetilde{H}(t) - 2\int \widetilde{H}(t)(1 - H(t))d\widetilde{H}(t) + \int \widetilde{H}^2(t)dH(t).\end{aligned}$$

Integration by parts gives

$$2\int \widetilde{H}(t)(1 - H(t))d\widetilde{H}(t) = \int \widetilde{H}^2(t)dH(t).$$

Therefore

$$\sigma^2 = \int(1 - H(t))^2 d\widetilde{H}(t) > 0,$$

which gives (3.4). Thus (3.5) follows from (2.16) and (3.4). □

Lemma 3.1 provides a consistent estimator of σ^2.

We first strengthen the convergence in distribution result of Theorem 2.1 to convergence in probability in sup-norm by making use of U-statistic U_k^*.

Theorem 3.1. *Under H_0, we can define a sequence of Brownian bridges $\{B_n(t), 0 \le t \le 1\}$ such that*

$$\sup_{0<t<1}\left|\frac{\sum_{i=1}^{[(n+1)t]} U_i}{\sqrt{\sum_{i=1}^{n} U_i^2}} - B_n(t)\right| = o_P(1), \qquad (3.8)$$

where U_i is defined in (2.4).

Proof. By (3.1) we can write U_k^* as the sum of three U-statistics. Namely, we have

$$U_k^* = U_n^{(3)} - (U_k^{(1)} + U_k^{(2)}), \tag{3.9}$$

where

$$U_k^{(1)} = \sum_{1 \le i < j \le k} h(Z_i, Z_j),$$
$$U_k^{(2)} = \sum_{k+1 \le i < j \le n} h(Z_i, Z_j),$$

and

$$U_n^{(3)} = \sum_{1 \le i < j \le n} h(Z_i, Z_j).$$

By Theorem 2.1 of Janson and Wichura (1983), we have

$$\max_{1 \le k \le n} \left| U_k^{(1)} - \sum_{i=1}^{k} (k - 2i + 1)\tilde{h}(Z_i) \right| = O_P(n), \tag{3.10}$$

$$\max_{1 \le k \le n} \left| U_k^{(2)} - \sum_{i=k+1}^{n} (n + k - 2i + 1)\tilde{h}(Z_i) \right| = O_P(n), \tag{3.11}$$

and

$$\left| U_k^{(3)} - \sum_{i=1}^{n} (n - 2i + 1)\tilde{h}(Z_i) \right| = O_P(n). \tag{3.12}$$

By (3.9)-(3.12), we have

$$\begin{aligned}
&\max_{1 \le k \le n} \left| U_k^* - \left\{ n \sum_{i=1}^{k} \tilde{h}(Z_i) - k \sum_{i=1}^{n} \tilde{h}(Z_i) \right\} \right| \\
&= \max_{1 \le k \le n} \left| U_k^* - \left\{ (n-k) \sum_{i=1}^{k} \tilde{h}(Z_i) - k \Big(\sum_{i=1}^{n} \tilde{h}(Z_i) - \sum_{i=1}^{k} \tilde{h}(Z_i) \Big) \right\} \right| \\
&= O_P(n).
\end{aligned}$$

Thus,

$$\max_{1 \le k \le n} \left| \frac{1}{\sqrt{n^3}} U_k^* - \frac{1}{\sqrt{n}} \left\{ \sum_{i=1}^{k} \tilde{h}(Z_i) - \frac{k}{n} \sum_{i=1}^{n} \tilde{h}(Z_i) \right\} \right| = O_P\left(\frac{1}{\sqrt{n}}\right). \tag{3.13}$$

Since $\sigma^2 > 0$ by Lemma 3.1, (3.13) can be written as

$$\sup_{0<t<1}\left|\frac{1}{\sigma\sqrt{n^3}}U^*_{[(n+1)t]} - \frac{1}{\sigma\sqrt{n}}\left\{\sum_{i=1}^{[(n+1)t]}\tilde{h}(Z_i) - \frac{[(n+1)t]}{n}\sum_{i=1}^{n}\tilde{h}(Z_i)\right\}\right| = O_P(\frac{1}{\sqrt{n}}). \tag{3.14}$$

Note that $\tilde{h}(Z_i), i = 1, \cdots, n$ are i.i.d. random variables and $\sigma^2 > 0$, and hence, by the Skorohod-Dudley version of Donsker's theorem (the invariance principle for partial sums) (cf. Theorem 2.1.2 of Csörgő and Révész (1981)), we can define a sequence of Brownian bridges $\{B_n(t),\ 0 \le t \le 1\}$ such that

$$\sup_{0<t<1}\left|\frac{1}{\sigma\sqrt{n}}\left\{\sum_{i=1}^{[(n+1)t]}\tilde{h}(Z_i) - \frac{[(n+1)t]}{n}\sum_{i=1}^{n}\tilde{h}(Z_i)\right\} - B_n(t)\right| = o_P(1). \tag{3.15}$$

Thus, by (3.14) and (3.15),

$$\sup_{0<t<1}\left|\frac{1}{\sigma\sqrt{n^3}}U^*_{[(n+1)t]} - B_n(t)\right| = o_P(1). \tag{3.16}$$

On the other hand, by Lemma 3.1, we have

$$\frac{\sigma}{\sqrt{\frac{1}{n^3}\sum_{i=1}^n U_i^2}} \xrightarrow{a.s.} 1. \tag{3.17}$$

Thus, by (2.5), (2.17), (3.16), and (3.17),

$$\begin{aligned}
&\sup_{0<t<1}\left|\frac{\sum_{i=1}^{[(n+1)t]}U_i}{\sqrt{\sum_{i=1}^n U_i^2}} - \frac{1}{\sigma\sqrt{n^3}}U^*_{[(n+1)t]}\right| \\
&= \sup_{0<t<1}\left|\frac{\sigma}{\sqrt{\frac{1}{n^3}\sum_{i=1}^n U_i^2}} - 1\right|\left|\frac{1}{\sigma\sqrt{n^3}}U^*_{[(n+1)t]}\right| \\
&\stackrel{\mathcal{D}}{=} \sup_{0<t<1}\left|\frac{\sigma}{\sqrt{\frac{1}{n^3}\sum_{i=1}^n U_i^2}} - 1\right|\Big(|B(t)| + o_P(1)\Big) \\
&= o_P(1)O_P(1) \\
&= o_P(1).
\end{aligned} \tag{3.18}$$

Therefore (3.8) follows from (3.16) and (3.18). □

Theorem 3.1 obtains a stronger result than the one given by Theorem 2.1.

3.2 Weighted Approximations

Next, we use weight functions to construct tests and emphasize the possibility of having a change in distribution on the tails.

We say that $q(t)$ is a positive function on (0,1) if

$$\inf_{\varepsilon \le t \le 1-\varepsilon} q(t) > 0 \tag{3.19}$$

for all $0 < \varepsilon < \frac{1}{2}$. Define

$$I_{0,1}(q,c) = \int_0^1 \frac{1}{t(1-t)} \exp\left(-\frac{cq^2(t)}{t(1-t)}\right) dt, \ c > 0. \tag{3.20}$$

We have

Theorem 3.2. *Under H_0, we assume that $q(t)$ is a positive function on $(0,1)$, it increases in a neighbourhood of zero and decreases in a neighbourhood of one. Then*

(i) we can define a sequence of Brownian bridges $\{B_n(t), 0 \le t \le 1\}$ such that

$$\sup_{0<t<1} \left| \frac{\sum_{i=1}^{[(n+1)t]} U_i}{\sqrt{\sum_{i=1}^n U_i^2}} - B_n(t) \right| / q(t) = o_P(1), \tag{3.21}$$

if and only if $I_{0,1}(q,c) < \infty$ for all $c > 0$, where U_i is defined in (2.4) *and $I_{0,1}(q,c)$ is defined in* (3.20).

(ii)

$$\sup_{0<t<1} \frac{|\sum_{i=1}^{[(n+1)t]} U_i|}{\sqrt{\sum_{i=1}^n U_i^2}} / q(t) \xrightarrow{\mathcal{D}} \sup_{0<t<1} |B(t)|/q(t), \tag{3.22}$$

if and only if $I_{0,1}(q,c) < \infty$ *for some* $c > 0$*, where* U_i *is defined in* (2.4), $\{B(t), 0 \le t \le 1\}$ *is a Brownian bridge, and* $I_{0,1}(q,c)$ *is defined in* (3.20).

Proof. For given $q(t)$, by Theorem A.5.1 of Csörgő and Horváth (1997), we have

$$\limsup_{t\to 0} |B(t)|/q(t) = 0 \quad a.s. \tag{3.23}$$

and

$$\limsup_{t\to 1} |B(t)|/q(t) = 0 \quad a.s. \tag{3.24}$$

if and only if $I_{0,1}(q,c) < \infty$ for all $c > 0$. We also have

$$\limsup_{t\to 0} |B(t)|/q(t) < \infty \quad a.s. \tag{3.25}$$

and

$$\limsup_{t\to 1} |B(t)|/q(t) < \infty \quad a.s. \tag{3.26}$$

if and only if $I_{0,1}(q,c) < \infty$ for some $c > 0$.

First we prove part (i). Suppose that $I_{0,1}(q,c) < \infty$ for all $c > 0$. Then by Lemma A.5.1 of Csörgő and Horváth (1997), for given $q(t)$ we have

$$\lim_{t\to 0} q(t)/\sqrt{t} = \infty \quad \text{and} \quad \lim_{t\to 1} q(t)/\sqrt{1-t} = \infty. \tag{3.27}$$

By Theorem 3.1 and (3.19), we have

$$\begin{aligned} &\sup_{\varepsilon\le t\le 1-\varepsilon} \left| \frac{\sum_{i=1}^{[(n+1)t]} U_i}{\sqrt{\sum_{i=1}^{n} U_i^2}} - B_n(t) \right| / q(t) \\ &\le \frac{1}{\inf_{\varepsilon\le t\le 1-\varepsilon} q(t)} \sup_{\varepsilon\le t\le 1-\varepsilon} \left| \frac{\sum_{i=1}^{[(n+1)t]} U_i}{\sqrt{\sum_{i=1}^{n} U_i^2}} - B_n(t) \right| \\ &= o_P(1) \end{aligned} \tag{3.28}$$

for all $0 < \varepsilon < \frac{1}{2}$.

On the other hand, by Theorem 2.1.1 of Csörgő and Horváth (19

Theorem 3.1 with (3.23), (3.24) and (3.27), we get

$$\sup_{o<t<\varepsilon} |B_n(t)|/q(t) = o_P(1),$$

$$\sup_{0<t<\varepsilon} \frac{|\sum_{i=1}^{[(n+1)t]} U_i|}{\sqrt{\sum_{i=1}^{n} U_i^2}}/q(t) = o_P(1),$$

$$\sup_{1-\varepsilon<t<1} |B_n(t)|/q(t) = o_P(1),$$

and

$$\sup_{1-\varepsilon<t<1} \frac{|\sum_{i=1}^{[(n+1)t]} U_i|}{\sqrt{\sum_{i=1}^{n} U_i^2}}/q(t) = o_P(1).$$

Thus (3.21) follows from (3.28)-(3.32).

We now assume that (3.21) holds. We have to show that $I_{0,1}(q, c$

all $c > 0$.

Since the distribution of $B_n(t)$ does not depend on n, we get, by (

$$\sup_{0<t<1} \left| \frac{\sum_{i=1}^{[(n+1)t]} U_i}{\sqrt{\sum_{i=1}^{n} U_i^2}} - B(t) \right| /q(t) = o_P(1).$$

Thus we have

$$\sup_{0<t<1/(n+1)} |B(t)|/q(t) = o_P(1)$$

since

$$\sup_{0<t<1/(n+1)} \frac{\sum_{i=1}^{[(n+1)t]} U_i}{\sqrt{\sum_{i=1}^{n} U_i^2}} = 0,$$

and

$$\sup_{n/(n+1)\leq t<1} |B(t)|/q(t) = o_P(1)$$

which is, in turn, because

$$\begin{aligned}
&\sup_{n/(n+1)\le t<1} \frac{\sum_{i=1}^{[(n+1)t]} U_i}{\sqrt{\sum_{i=1}^n U_i^2}} \\
&= \frac{1}{\sqrt{\sum_{i=1}^n U_i^2}} \sum_{i=1}^n U_i \\
&= \frac{1}{\sqrt{\sum_{i=1}^n U_i^2}} \sum_{i=1}^n \sum_{j=1}^n h(Z_i, Z_j) \\
&= 0.
\end{aligned} \tag{3.37}$$

By (3.34) and (3.36), we get (3.23) and (3.24). Thus, $I_{0,1}(q,c) < \infty$ for all $c > 0$. Therefore part (i) is proven.

Next, we prove part (ii). Suppose that $I_{0,1}(q,c) < \infty$ for some $c > 0$. Then (3.27) is still true by Lemma A.5.1 of Csörgő and Horváth (1997). Thus, by (3.28), we have

$$\sup_{1/(n+1)\le t\le n/(n+1)} \left| \frac{\sum_{i=1}^{[(n+1)t]} U_i}{\sqrt{\sum_{i=1}^n U_i^2}} - B_n(t) \right| / q(t) = o_P(1). \tag{3.38}$$

Hence, to prove that (3.22) holds, by (3.38), combined with (3.35) and (3.37), we only need show that

$$\sup_{1/(n+1)\le t\le n/(n+1)} |B(t)|/q(t) \xrightarrow{\mathcal{D}} \sup_{0<t<1} |B(t)|/q(t),$$

which, fortunately, follows from (3.25) and (3.26).

Now we show that $I_{0,1}(q,c) < \infty$ for some $c > 0$ if (3.22) holds. Suppose that (3.22) holds. Then

$$\sup_{0<t<1} |B(t)|/q(t) < \infty \quad a.s.$$

which is equivalent to (3.25) and (3.26). Thus we have $I_{0,1}(q,c) < \infty$ for some $c > 0$. Therefore part (ii) is proven. □

3.3 Asymptotic Distributions of the Weighted Test Statistics

First we need to strengthen our earlier results to give a rate of convergence in the following lemma.

Lemma 3.2. *Under H_0,*

$$\left|\frac{1}{n^3}\sum_{i=1}^{n} U_i^2 - \sigma^2\right| = O_P\left(\frac{1}{\sqrt{n}}\right), \tag{3.39}$$

where U_i and σ^2 are defined in (2.4) *and* (3.3), *respectively.*

Proof. Under H_0, by (3.7) and Assumption 1.1, we have

$$\begin{aligned}\sigma^2 &= \int \widetilde{H}^2(t)dH(t) - 2\int \widetilde{H}(t)(1-H(t))d\widetilde{H}(t) + \int (1-H(t))^2 d\widetilde{H}(t)\\ &= \int \widetilde{H}^2(t-)dH(t) - 2\int \widetilde{H}(t-)(1-H(t))d\widetilde{H}(t) + \int (1-H(t))^2 d\widetilde{H}(t),\end{aligned}$$

where $H(t)$ and $\widetilde{H}(t)$ are defined in (1.3) and (2.15), respectively. Thus, by (2.14), we have

$$\begin{aligned}&\left|\frac{1}{n^3}\sum_{i=1}^{n} U_i^2 - \sigma^2\right|\\ &\quad= \Big|\Big[\int \widetilde{H}_n^2(t-)dH_n(t) - \int \widetilde{H}^2(t-)dH(t)\Big]\\ &\qquad - 2\Big[\int \widetilde{H}_n(t-)(1-H_n(t))d\widetilde{H}_n(t) - \int \widetilde{H}(t-)(1-H(t))d\widetilde{H}(t)\Big]\\ &\qquad + \Big[\int (1-H_n(t))^2 d\widetilde{H}_n(t) - \int (1-H(t))^2 d\widetilde{H}(t)\Big]\Big|.\end{aligned} \tag{3.40}$$

By the weak convergence of the empirical distribution function to a Brownian bridge (cf. Billingsley (1968)), we have

$$\sup_t |H_n(t) - H(t)| = O_P\left(\frac{1}{\sqrt{n}}\right)$$

and

$$\sup_t |\widetilde{H}_n(t) - \widetilde{H}(t)| = O_P(\frac{1}{\sqrt{n}}).$$

Thus under H_0,

$$\begin{aligned}
&\left|\int \widetilde{H}_n^2(t-)dH_n(t) - \int \widetilde{H}^2(t-)dH(t)\right| \\
&= \left|\int \widetilde{H}_n^2(t-)d\big[H_n(t) - H(t)\big]\right. \\
&\qquad \left. + \int (\widetilde{H}_n(t-) + \widetilde{H}(t-))\big[\widetilde{H}_n(t-) - \widetilde{H}(t-)\big]dH(t)\right| \\
&= O_P(\frac{1}{\sqrt{n}}),
\end{aligned} \tag{3.41}$$

$$\begin{aligned}
&\left|\int \widetilde{H}_n(t-)(1 - H_n(t))d\widetilde{H}_n(t) - \int \widetilde{H}(t-)(1 - H(t))d\widetilde{H}(t)\right| \\
&= \left|\int \big[\widetilde{H}_n(t-) - \widetilde{H}(t-)\big](1 - H_n(t))d\widetilde{H}_n(t)\right. \\
&\qquad + \int \widetilde{H}(t-)(1 - H_n(t))d\big[\widetilde{H}_n(t) - \widetilde{H}(t)\big] \\
&\qquad \left. - \int \widetilde{H}(t-)\big[H_n(t) - H(t)\big]d\widetilde{H}(t)\right| \\
&= O_P(\frac{1}{\sqrt{n}}),
\end{aligned} \tag{3.42}$$

and

$$\begin{aligned}
&\left|\int (1 - H_n(t))^2 d\widetilde{H}_n(t) - \int (1 - H(t))^2 d\widetilde{H}(t)\right| \\
&= \left|\int (1 - H_n(t))^2 d\big[\widetilde{H}_n(t) - \widetilde{H}(t)\big]\right. \\
&\qquad \left. - \int (2 - H_n(t) - H(t))\big[H_n(t) - H(t)\big]d\widetilde{H}(t)\right| \\
&= O_P(\frac{1}{\sqrt{n}}).
\end{aligned} \tag{3.43}$$

Therefore (3.39) follows by (3.40)-(3.43). □

The desirability of having weight function q in Theorem 3.2 is to make our statistical test more sensitive on the tails. Note that since the variance of $B(t)$

is $t(1-t)$, a typical choice of q is $\sqrt{t(1-t)}$, the asymptotic standard deviation, as a weight function. However, $I_{0,1}(\sqrt{t(1-t)}, c) = \infty$ for all $c > 0$, so we cannot apply Theorem 3.2 in the case of the natural weight function $\sqrt{t(1-t)}$. This case is studied in the next thorem.

Theorem 3.3. *Under H_0,*

$$\lim_{n\to\infty} P\left\{ A(\log n) \max_{1\le k<n} \frac{\sum_{i=1}^k U_i \Big/ \sqrt{\sum_{i=1}^n U_i^2}}{\sqrt{\frac{k}{n}\left(1-\frac{k-1}{n}\right)}} \le t + D(\log n) \right\} = \exp(-e^{-t}) \quad (3.44)$$

and

$$\lim_{n\to\infty} P\left\{ A(\log n) \max_{1\le k<n} \frac{\left|\sum_{i=1}^k U_i\right| \Big/ \sqrt{\sum_{i=1}^n U_i^2}}{\sqrt{\frac{k}{n}\left(1-\frac{k-1}{n}\right)}} \le t + D(\log n) \right\} = \exp(-2e^{-t}) \quad (3.45)$$

for all t, where U_i is defined in (2.4),

$$A(x) = \sqrt{2\log x}, \quad (3.46)$$

and

$$D(x) = 2\log x + \frac{1}{2}\log\log x - \frac{1}{2}\log \pi. \quad (3.47)$$

Proof. Under H_0 we know that, by (3.6),

$$\begin{aligned} E\widetilde{h}(Z_i) &= E\Big((1-H(X_i))\delta_i - \widetilde{H}(X_i)\Big) \\ &= \int (1-H(t))d\widetilde{H}(t) - \int \widetilde{H}(t)dH(t) \\ &= (1-H(t))\widetilde{H}(t)\Big|_{-\infty}^{+\infty} - \int \widetilde{H}(t)d(1-H(t)) - \int \widetilde{H}(t)dH(t) \quad (3.48) \\ &= \int \widetilde{H}(t)dH(t) - \int \widetilde{H}(t)dH(t) \\ &= 0, \end{aligned}$$

by (3.3),

$$E\Big(\frac{1}{\sigma}\tilde{h}(Z_i)\Big)^2 = 1,\ i = 1, \cdots, n,$$

and $\tilde{h}(Z_1), \cdots, \tilde{h}(Z_n)$ are i.i.d. random variables.

Note that $Z_1, \cdots, Z_n$ are two-dimensional i.i.d. random vectors. Using (3.10)-(3.12), and Theorem A.4.1 and Theorem 2.4.12 of Csörgő and Horváth (1997), we get

$$\lim_{n\to\infty} P\Big\{A(\log n)\max_{1\le k<n}\frac{\sum_{i=1}^k U_i}{\sigma\sqrt{k(n-k+1)n}} \le t + D(\log n)\Big\} = \exp(-e^{-t}) \tag{3.49}$$

and

$$\lim_{n\to\infty} P\Big\{A(\log n)\max_{1\le k<n}\frac{|\sum_{i=1}^k U_i|}{\sigma\sqrt{k(n-k+1)n}} \le t + D(\log n)\Big\} = \exp(-2e^{-t}) \tag{3.50}$$

for all t. Thus,

$$A(\log n)\max_{1\le k<n}\frac{|\sum_{i=1}^k U_i|}{\sigma\sqrt{k(n-k+1)n}} = O_P(\log\log n), \tag{3.51}$$

since

$$A(\log n) = \sqrt{2\log\log n}$$

and

$$D(\log n) = 2\log\log n + \frac{1}{2}\log\log\log n - \frac{1}{2}\log\pi.$$

On the other hand, by Lemma 3.2, we have

$$\begin{aligned}
&\left|\frac{\sigma}{\sqrt{\frac{1}{n^3}\sum_{i=1}^n U_i^2}} - 1\right| \\
&= \frac{\left|\sigma - \sqrt{\frac{1}{n^3}\sum_{i=1}^n U_i^2}\right|}{\sqrt{\frac{1}{n^3}\sum_{i=1}^n U_i^2}} \\
&= \frac{\left|\frac{1}{n^3}\sum_{i=1}^n U_i^2 - \sigma^2\right|}{\sqrt{\frac{1}{n^3}\sum_{i=1}^n U_i^2}\Big(\sqrt{\frac{1}{n^3}\sum_{i=1}^n U_i^2} + \sigma\Big)} \\
&= O_P\Big(\frac{1}{\sqrt{n}}\Big).
\end{aligned} \tag{3.52}$$

Hence, by (3.51) and (3.52), we have

$$\begin{aligned}
&A(\log n)\max_{1\le k<n}\left|\frac{\sum_{i=1}^{k}U_i\Big/\sqrt{\sum_{i=1}^{n}U_i^2}}{\sqrt{\frac{k}{n}\Big(1-\frac{k-1}{n}\Big)}}-\frac{\sum_{i=1}^{k}U_i}{\sigma\sqrt{k(n-k+1)n}}\right| \\
&\quad= A(\log n)\max_{1\le k<n}\frac{|\sum_{i=1}^{k}U_i|}{\sigma\sqrt{k(n-k+1)n}}\left|\frac{\sigma}{\sqrt{\frac{1}{n^3}\sum_{i=1}^{n}U_i^2}}-1\right| \qquad (3.53)\\
&\quad= O_P(\log\log n)O_P\Big(\frac{1}{\sqrt{n}}\Big)\\
&\quad= o_P(1).
\end{aligned}$$

Therefore (3.44) and (3.45) follow from (3.49), (3.50), and (3.53). □

By Theorem 3.3, we obtain a weighted test statistic

$$\max_{1\le k<n}\frac{\left|\sum_{i=1}^{k}U_i\right|\Big/\sqrt{\sum_{i=1}^{n}U_i^2}}{\sqrt{\frac{k}{n}\Big(1-\frac{k-1}{n}\Big)}}. \qquad (3.54)$$

The approximated critical values can be given by (3.45). Let $c_\alpha(n)$ denote the critical value of the test for sample size n at significance level α; the test is defined by

$$I\left(\max_{1\le k<n}\frac{\left|\sum_{i=1}^{k}U_i\right|\Big/\sqrt{\sum_{i=1}^{n}U_i^2}}{\sqrt{\frac{k}{n}\Big(1-\frac{k-1}{n}\Big)}}>c_\alpha(n)\right). \qquad (3.55)$$

Using (3.45) from Theorem 3.3, we get

$$c_\alpha(n)=\frac{-\log(-\log\sqrt{1-\alpha})+D(\log n)}{A(\log n)}. \qquad (3.56)$$

We list several critical values of $c_\alpha(n)$ for some selected values of sample size n and α in Table 3.1. We'll use Table 3.1 in the following simulation study.

We do a simulation study to investigate the power on the tails of the weighted statistic test (3.55) in the next section.

Table 3.1. *Some selected critical values $c_\alpha(n)$*

n \ α	0.01	0.05	0.10
50	4.6039	3.6171	3.1813
100	4.5701	3.6374	3.2256
200	4.5513	3.6588	3.2646
500	4.5389	3.6862	3.3096

3.4 Simulation Study

To see if any improvement can be made on detecting change-points on the tails for the proposed weighted test, we would like to compare the power of test (2.19) and power of weighted test (3.55) through the Monte Carlo simulation study.

We performed $N = 5000$ simulations with sample size $n = 200$ for each case. In each case to be considered, we assumed that

$$F^{(1)} = \exp(\mu_1), \quad F^{(2)} = \exp(\mu_2), \quad G = \exp(\mu_c),$$

where $\mu_1 = 1.0$ and $\mu_c = 3.0$. We considered $\mu_2 = 2.5, 3.0, 3.5$ and change-points varied at $\lambda = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9$ and significance levels $\alpha = 0.05, 0.10$. The results are reported in Table 3.2. From Table 3.2, we can see that the powers of weighted test (3.55) are larger than the corresponding powers of test (2.19) on the two tails ($\lambda = 0.1$ or $\lambda = 0.9$). On the other hand, they are smaller than the corresponding powers of test (2.19) in any other case. So, the weighted statistical test (3.55) indeed improves only the situation when the change-point occurs on the tails.

Table 3.2. *Comparison of powers of test* (2.19) *and weighted test* (3.55)

μ_1	μ_2	λ	($\alpha = 0.05$) power	power(weighted)	($\alpha = 0.10$) power	power(weighted)
1.0	2.5	0.1	0.2636	0.3282	0.3996	0.4994
		0.2	0.7926	0.6494	0.8698	0.7866
		0.3	0.9414	0.7908	0.9736	0.8988
		0.4	0.9768	0.8354	0.9908	0.9240
		0.5	0.9832	0.8562	0.9940	0.9414
		0.6	0.9682	0.8136	0.9882	0.9138
		0.7	0.9208	0.7228	0.9592	0.8486
		0.8	0.6976	0.4932	0.8194	0.6842
		0.9	0.1846	0.1550	0.3052	0.3156
1.0	3.0	0.1	0.3860	0.5172	0.5492	0.6820
		0.2	0.9236	0.8468	0.9612	0.9320
		0.3	0.9920	0.9398	0.9966	0.9756
		0.4	0.9960	0.9630	1.0000	0.9872
		0.5	0.9974	0.9640	0.9986	0.9900
		0.6	0.9946	0.9470	0.9984	0.9828
		0.7	0.9800	0.8920	0.9902	0.9532
		0.8	0.8476	0.7084	0.9224	0.8528
		0.9	0.2476	0.2616	0.3994	0.4562
1.0	3.5	0.1	0.5080	0.6752	0.6762	0.8126
		0.2	0.9712	0.9408	0.9868	0.9728
		0.3	0.9976	0.9804	0.9992	0.9936
		0.4	0.9996	0.9920	1.0000	0.9982
		0.5	1.0000	0.9942	0.9998	0.9992
		0.6	0.9988	0.9840	1.0000	0.9968
		0.7	0.9932	0.9616	0.9990	0.9894
		0.8	0.9266	0.8414	0.9690	0.9288
		0.9	0.3070	0.3594	0.4814	0.5898

Chapter 4

Asymptotics under H_1

4.1 Introduction

The asymptotic distribution of test statistics under the alternative hypothesis has not been considered so far. In this chapter, we consider the change-point estimation and the asymptotic distribution of the test statistic under H_1.

First we give some lemmas. Denote the true change-point under H_1 by

$$\tau = [n\lambda], \ 0 < \lambda < 1. \tag{4.1}$$

We give a series of definitions. Define the subdistribution function of $X_1, \cdots, X_\tau$ by

$$\widetilde{H}^{(1)}(t) = P\{X_1 \le t, \delta_1 = 1\}, \tag{4.2}$$

the empirical distribution function of $X_1, \cdots, X_\tau$ by

$$H_\tau^{(1)}(t) = \frac{1}{\tau}\sum_{i=1}^{\tau} I(X_i \le t), \tag{4.3}$$

the empirical subdistribution function of $X_1, \cdots, X_\tau$ by

$$\widetilde{H}_\tau^{(1)}(t) = \frac{1}{\tau}\sum_{i=1}^{\tau} I(X_i \le t, \delta_i = 1), \tag{4.4}$$

the subdistribution function of $X_{\tau+1}, \cdots, X_n$ by

$$\widetilde{H}^{(2)}(t) = P\{X_n \le t, \delta_n = 1\}, \tag{4.5}$$

the empirical distribution function of $X_{\tau+1}, \cdots, X_n$ by

$$H_{n-\tau}^{(2)}(t) = \frac{1}{n-\tau} \sum_{i=\tau+1}^{n} I(X_i \le t), \tag{4.6}$$

and the empirical subdistribution function of $X_{\tau+1}, \cdots, X_n$ by

$$\widetilde{H}_{n-\tau}^{(2)}(t) = \frac{1}{n-\tau} \sum_{i=\tau+1}^{n} I(X_i \le t, \delta_i = 1). \tag{4.7}$$

Let $\mathcal{F}_{ni}$ denote the sigma–field generated by $\{Z_1, \cdots, Z_i, Z_{\tau+1}, \cdots, Z_n\}, i = 1, \cdots, \tau$ and $\mathcal{F}_{n0}$ denote the sigma-field generated by $\{Z_{\tau+1}, \cdots, Z_n\}$. Define

$$\mu_{ni} = E\Big\{ \sum_{j=\tau+1}^{n} h(Z_i, Z_j) | \mathcal{F}_{n,i-1} \Big\}, \ i = 1, \cdots, \tau, \tag{4.8}$$

and

$$\begin{aligned} \sigma_{ni}^2 &= E\Big\{\Big[\sum_{j=\tau+1}^{n} h(Z_i, Z_j) - \mu_{ni} \Big]^2 | \mathcal{F}_{n,i-1} \Big\} \\ &= E\Big\{\Big[\sum_{j=\tau+1}^{n} h(Z_i, Z_j) \Big]^2 | \mathcal{F}_{n,i-1} \Big\} - \mu_{ni}^2, \ i = 1, \cdots, \tau, \end{aligned} \tag{4.9}$$

where h is given by (2.8).

We have

Lemma 4.1. *Under H_1,*

$$\begin{aligned} \mu_{n1} &= \cdots = \mu_{n\tau} \\ &= \sum_{j=\tau+1}^{n} \big\{ [1 - H^{(1)}(X_j)]\delta_j - \widetilde{H}^{(1)}(X_j) \big\} \\ &= (n-\tau)\Big\{ \int \widetilde{H}_{n-\tau}^{(2)}(t-) dH^{(1)}(t) - \int [1 - H_{n-\tau}^{(2)}(t)] d\widetilde{H}^{(1)}(t) \Big\}, \end{aligned} \tag{4.10}$$

and

$$\begin{aligned} \sigma_{n1}^2 &= \cdots = \sigma_{n\tau}^2 \\ &= (n-\tau)^2 [A_1(\tau) + A_2(\tau) - 2A_3(\tau)] - \mu_{n\tau}^2, \end{aligned} \tag{4.11}$$

where

$$A_1(\tau) = \int [\widetilde{H}_{n-\tau}^{(2)}(t-)]^2 dH^{(1)}(t),$$

$$A_2(\tau) = \int [1 - H_{n-\tau}^{(2)}(t)]^2 d\widetilde{H}^{(1)}(t),$$

$$A_3(\tau) = \int \widetilde{H}_{n-\tau}^{(2)}(t-)[1 - H_{n-\tau}^{(2)}(t)] d\widetilde{H}^{(1)}(t),$$

and $H^{(1)}(t)$ *is defined in* (1.4), *and* τ *is given by* (4.1).

Proof. Under H_1, the variables $X_1, \cdots, X_\tau$ are i.i.d. each with d.f. $H^{(1)}$ and $X_{\tau+1}, \cdots, X_n$ are i.i.d. each with d.f. $H^{(2)}$. For $i \le \tau$, by the definitions of $h(Z_i, Z_j), H^{(1)}, \widetilde{H}^{(1)}$, and $\mathcal{F}_{ni}$, we have

$$\begin{aligned}
\mu_{ni} &= \sum_{j=\tau+1}^{n} \Big\{E\big\{I(X_i > X_j, \delta_j = 1)|\mathcal{F}_{n,i-1}\big\} - E\big\{I(X_i < X_j, \delta_i = 1)|\mathcal{F}_{n,i-1}\big\}\Big\} \\
&= \sum_{j=\tau+1}^{n} \Big\{P\big\{X_i > X_j, \delta_j = 1|\mathcal{F}_{n,i-1}\big\} - P\big\{X_i < X_j, \delta_i = 1|\mathcal{F}_{n,i-1}\big\}\Big\} \\
&= \sum_{j=\tau+1}^{n} \Big\{P\big\{X_i > X_j, \delta_j = 1|(X_j, \delta_j)\big\} - P\big\{X_i < X_j, \delta_i = 1|(X_j, \delta_j)\big\}\Big\} \\
&= \sum_{j=\tau+1}^{n} \big\{[1 - H^{(1)}(X_j)]I(\delta_j = 1) - \widetilde{H}^{(1)}(X_j)\big\} \\
&= \sum_{j=\tau+1}^{n} \big\{[1 - H^{(1)}(X_j)]\delta_j - \widetilde{H}^{(1)}(X_j)\big\},
\end{aligned}$$

which gives the first expression of μ_{ni} in (4.10). On the other hand, by the definitions of $H^{(2)}_{n-\tau}$ and $\widetilde{H}^{(2)}_{n-\tau}$, we have

$$\begin{aligned}
&\sum_{j=\tau+1}^{n} h(Z_i, Z_j) \\
&\qquad = \sum_{j=\tau+1}^{n} I(X_i > X_j, \delta_j = 1) - \sum_{j=\tau+1}^{n} I(X_i < X_j, \delta_i = 1) \\
&\qquad = (n-\tau)\widetilde{H}^{(2)}_{n-\tau}(X_i-) - (n-\tau)[1 - H^{(2)}_{n-\tau}(X_i)]\delta_i \\
&\qquad = (n-\tau)\big\{\widetilde{H}^{(2)}_{n-\tau}(X_i-) - [1 - H^{(2)}_{n-\tau}(X_i)]\delta_i\big\}.
\end{aligned} \tag{4.12}$$

Thus,

$$\begin{aligned}
\mu_{ni} &= E\big\{\sum_{j=\tau+1}^{n} h(Z_i, Z_j)|\mathcal{F}_{n,i-1}\big\} \\
&= (n-\tau)\Big\{E\big\{\widetilde{H}^{(2)}_{n-\tau}(X_i-)|\mathcal{F}_{n,i-1}\big\} - E\big\{[1 - H^{(2)}_{n-\tau}(X_i)]\delta_i|\mathcal{F}_{n,i-1}\big\}\Big\} \\
&= (n-\tau)\Big\{E\big\{\widetilde{H}^{(2)}_{n-\tau}(X_i-)|\mathcal{F}_{n0}\big\} - E\big\{[1 - H^{(2)}_{n-\tau}(X_i)]\delta_i|\mathcal{F}_{n0}\big\}\Big\} \\
&= (n-\tau)\big\{\int \widetilde{H}^{(2)}_{n-\tau}(t-)dH^{(1)}(t) - \int [1 - H^{(2)}_{n-\tau}(t)]d\widetilde{H}^{(1)}(t)\big\},
\end{aligned}$$

which gives the second expression of μ_{ni} in (4.10).

Now, by (4.12), we have

$$
\begin{aligned}
&[\sum_{j=\tau+1}^{n} h(Z_i, Z_j)]^2 \\
&= (n-\tau)^2 \Big\{ \widetilde{H}^{(2)}_{n-\tau}(X_i-) - [1 - H^{(2)}_{n-\tau}(X_i)]\delta_i \Big\}^2 \\
&= (n-\tau)^2 \Big\{ [\widetilde{H}^{(2)}_{n-\tau}(X_i-)]^2 + [1 - H^{(2)}_{n-\tau}(X_i)]^2\delta_i - 2\widetilde{H}^{(2)}_{n-\tau}(X_i-)[1 - H^{(2)}_{n-\tau}(X_i)]\delta_i \Big\}.
\end{aligned} \tag{4.13}
$$

Thus,

$$
\begin{aligned}
&E\Big\{[\sum_{j=\tau+1}^{n} h(Z_i, Z_j)]^2 | \mathcal{F}_{n,i-1}\Big\} \\
&= (n-\tau)^2 \Big\{ E\Big\{[\widetilde{H}^{(2)}_{n-\tau}(X_i-)]^2 | \mathcal{F}_{n,i-1}\Big\} + E\Big\{[1 - H^{(2)}_{n-\tau}(X_i)]^2\delta_i | \mathcal{F}_{n,i-1}\Big\} \\
&\qquad - 2E\Big\{\widetilde{H}^{(2)}_{n-\tau}(X_i-)[1 - H^{(2)}_{n-\tau}(X_i)]\delta_i | \mathcal{F}_{n,i-1}\Big\}\Big\} \\
&= (n-\tau)^2 \Big\{ E\Big\{[\widetilde{H}^{(2)}_{n-\tau}(X_i-)]^2 | \mathcal{F}_{n0}\Big\} + E\Big\{[1 - H^{(2)}_{n-\tau}(X_i)]^2\delta_i | \mathcal{F}_{n0}\Big\} \\
&\qquad - 2E\Big\{\widetilde{H}^{(2)}_{n-\tau}(X_i-)[1 - H^{(2)}_{n-\tau}(X_i)]\delta_i | \mathcal{F}_{n0}\Big\}\Big\} \\
&= (n-\tau)^2 \Big\{ \int [\widetilde{H}^{(2)}_{n-\tau}(t-)]^2 dH^{(1)}(t) + \int [1 - H^{(2)}_{n-\tau}(t)]^2 d\widetilde{H}^{(1)}(t) \\
&\qquad - 2\int \widetilde{H}^{(2)}_{n-\tau}(t-)[1 - H^{(2)}_{n-\tau}(t)] d\widetilde{H}^{(1)}(t) \Big\} \\
&= (n-\tau)^2 [A_1(\tau) + A_2(\tau) - 2A_3(\tau)],
\end{aligned} \tag{4.14}
$$

and therefore (4.11) follows from (4.9), (4.10), and (4.14). □

So, by Lemma 4.1, (4.8) and (4.9) can be written as

$$
\mu_{ni} = E\Big\{ \sum_{j=\tau+1}^{n} h(Z_1, Z_j) | \mathcal{F}_{n0} \Big\},\ i = 1, \cdots, \tau, \tag{4.15}
$$

and

$$
\begin{aligned}
\sigma^2_{ni} &= E\Big\{\Big[\sum_{j=\tau+1}^{n} h(Z_1, Z_j) - \mu_{n1}\Big]^2 | \mathcal{F}_{n0}\Big\} \\
&= E\Big\{\Big[\sum_{j=\tau+1}^{n} h(Z_1, Z_j)\Big]^2 | \mathcal{F}_{n0}\Big\} - \mu^2_{n1},\ i = 1, \cdots, \tau,
\end{aligned} \tag{4.16}
$$

respectively, which are independent of $i, 1 \leq i \leq \tau$.

4.2 Change-point Estimators

Define

$$\begin{aligned} \theta_{1,2} &= Eh(Z_\tau, Z_{\tau+1}) \\ &= P\{X_\tau > X_{\tau+1}, \delta_{\tau+1} = 1\} - P\{X_\tau < X_{\tau+1}, \delta_\tau = 1\}. \end{aligned} \tag{4.17}$$

We have

Lemma 4.2. *Under H_1, for $1 \leq i \leq \tau$,*

$$\begin{aligned} EU_i &= E\Big\{ \sum_{j=\tau+1}^{n} h(Z_i, Z_j)\Big\} \\ &= E\mu_{ni} = (n-\tau)\theta_{1,2}, \end{aligned} \tag{4.18}$$

and for $\tau + 1 \leq i \leq n$,

$$\begin{aligned} EU_i &= E\Big\{ \sum_{j=1}^{\tau} h(Z_i, Z_j)\Big\} \\ &= -\tau\theta_{1,2}, \end{aligned} \tag{4.19}$$

where U_i is defined in (2.4)*, h, τ, and μ_{ni} are given by* (2.8), (4.1)*, and* (4.10)*, respectively.*

Proof. Under H_1, the variables $X_1, \cdots, X_\tau$ are i.i.d. and $X_{\tau+1}, \cdots, X_n$ are i.i.d., so we have

$$E\Big\{ \sum_{j=1}^{\tau} h(Z_i, Z_j)\Big\} = 0 \quad \text{if } 1 \leq i \leq \tau, \tag{4.20}$$

and

$$E\Big\{ \sum_{j=\tau+1}^{n} h(Z_i, Z_j)\Big\} = 0 \quad \text{if } \tau + 1 \leq i \leq n \tag{4.21}$$

since h is anti-symmetric. Hence, by (2.4), (4.17), and (4.20), we have

$$
\begin{aligned}
EU_i &= E\Big\{\sum_{j=1}^{n} h(Z_i, Z_j)\Big\} \\
&= E\Big\{\sum_{j=\tau+1}^{n} h(Z_i, Z_j)\Big\} \\
&= \sum_{j=\tau+1}^{n} Eh(Z_\tau, Z_{\tau+1}) \\
&= (n-\tau)\theta_{1,2}, \; i = 1, \cdots, \tau,
\end{aligned} \tag{4.22}
$$

and then by (4.8),

$$
\begin{aligned}
E\mu_{ni} &= E\Big\{E\Big\{\sum_{j=\tau+1}^{n} h(Z_i, Z_j)|\mathcal{F}_{n,i-1}\Big\}\Big\} \\
&= E\Big\{\sum_{j=\tau+1}^{n} h(Z_i, Z_j)\Big\} = EU_i.
\end{aligned} \tag{4.23}
$$

(4.22) and (4.23) give (4.18).

For $\tau + 1 \leq i \leq n$, by (2.4), (4.17), and (4.21), we have

$$
\begin{aligned}
EU_i &= E\Big\{\sum_{j=1}^{n} h(Z_i, Z_j)\Big\} \\
&= E\Big\{\sum_{j=1}^{\tau} h(Z_i, Z_j)\Big\} \\
&= -E\Big\{\sum_{j=1}^{\tau} h(Z_j, Z_i)\Big\} \\
&= -\sum_{j=1}^{\tau} Eh(Z_\tau, Z_{\tau+1}) \\
&= -\tau\theta_{1,2},
\end{aligned}
$$

which gives (4.19). □

Remark 4.1. By Lemma 4.2, we note that, under H_1, if $\theta_{1,2} > 0$, then

$$EU_i > 0 \quad \text{for } 1 \leq i \leq \tau, \text{ and } EU_i < 0 \quad \text{for } \tau + 1 \leq i \leq n.$$

If $\theta_{1,2} < 0$, then

$$EU_i < 0 \quad \text{for } 1 \leq i \leq \tau, \text{and } EU_i > 0 \quad \text{for } \tau + 1 \leq i \leq n.$$

Note that we have

$$E(\sum_{i=1}^{k} U_i) = kE(U_1) \text{ for } 1 \le k \le \tau,$$

and

$$\begin{aligned} E(\sum_{i=1}^{k} U_i) &= E(\sum_{i=1}^{\tau} U_i) + E(\sum_{i=\tau+1}^{k} U_i) \\ &= \tau E(U_1) + (k-\tau)E(U_n) \text{ for } \tau+1 \le k \le n. \end{aligned}$$

Thus, by Lemma 4.2 we get

Corollary 4.1. *Under* H_1,

$$\theta_{1,2} = \int \widetilde{H}^{(2)}(t)dH^{(1)}(t) - \int [1 - H^{(2)}(t)]d\widetilde{H}^{(1)}(t), \tag{4.24}$$

$$E\Big(\sum_{i=1}^{k} U_i\Big) = k(n-\tau)\theta_{1,2},\ k = 1, \cdots, \tau, \tag{4.25}$$

and

$$E\Big(\sum_{i=1}^{k} U_i\Big) = (n-k)\tau\theta_{1,2},\ k = \tau+1, \cdots, n, \tag{4.26}$$

where $H^{(1)}(t)$, $H^{(2)}(t)$, $\widetilde{H}^{(1)}(t)$, *and* $\widetilde{H}^{(2)}(t)$ *are defined in* (1.4), (1.5), (4.2), *and* (4.5), *respectively, and* U_i, τ, *and* $\theta_{1,2}$ *are defined in* (2.4), (4.1), *and* (4.17), *respectively.*

Remark 4.2. By Corollary 4.1, we can see that, if $\theta_{1,2} > 0$, then

$$E\Big(\sum_{i=1}^{k} U_i\Big) > 0,\ 1 \le k < n,$$

and $E(\sum_{i=1}^{k} U_i)$ increases before change point τ $(1 \le k \le \tau)$ and decreases after change point $(\tau+1 \le k \le n)$. If $\theta_{1,2} < 0$, then

$$E\Big(\sum_{i=1}^{k} U_i\Big) < 0,\ 1 \le k < n,$$

and $E(\sum_{i=1}^{k} U_i)$ decreases before change point τ $(1 \le k \le \tau)$ and increases after change point $(\tau + 1 \le k \le n)$.

Thus we can get

Corollary 4.2. *Under* H_1,

$$\Big|E\Big(\sum_{i=1}^{\tau} U_i\Big)\Big| = \max_{1\le k\le n} \Big|E\Big(\sum_{i=1}^{k} U_i\Big)\Big| = \tau(n-\tau)|\theta_{1,2}|. \tag{4.27}$$

Remark 4.3. By Remark 4.2 and Corollary 4.2 we can see that, if $\theta_{1,2} > 0$, then $E(\sum_{i=1}^{k} U_i)$ is a hat-type function and change-point τ is the maximizer of $E(\sum_{i=1}^{k} U_i)$ $(1 \le k \le n)$. If $\theta_{1,2} < 0$, then $E(\sum_{i=1}^{k} U_i)$ is a U-type function and change-point τ is the minimizer of $E(\sum_{i=1}^{k} U_i)$ $(1 \le k \le n)$. These properties can be used to help detect the change point. See Example 4.1 in Section 4.3 for details.

We define an estimator of the change-point τ by

$$\hat{\tau} = \operatorname*{argmax}_{1\le k<n} \Big|\sum_{i=1}^{k} U_i\Big|, \tag{4.28}$$

where U_i is defined in (2.4), and

$$\begin{aligned}\operatorname*{argmax}_{1\le k<n} \Big|\sum_{i=1}^{k} U_i\Big| &= \min\Big\{k : \Big|\sum_{i=1}^{k} U_i\Big| = \max_{1\le m<n} \Big|\sum_{i=1}^{m} U_i\Big|\Big\}\\ &= \min\left\{k : \frac{|\sum_{i=1}^{k} U_i|}{\sqrt{\sum_{i=1}^{n} U_i^2}} = \max_{1\le m<n} \frac{|\sum_{i=1}^{m} U_i|}{\sqrt{\sum_{i=1}^{n} U_i^2}}\right\}.\end{aligned}$$

We define the weighted analogue of $\hat{\tau}$ by

$$\begin{aligned}\tilde{\tau} &= \operatorname*{argmax}_{1\le k<n} \frac{\big|\sum_{i=1}^{k} U_i\big|/\sqrt{\sum_{i=1}^{n} U_i^2}}{\sqrt{k(n-k+1)}}\\ &= \min\left\{k : \frac{\big|\sum_{i=1}^{k} U_i\big|}{\sqrt{k(n-k+1)}} = \max_{1\le m<n} \frac{\big|\sum_{i=1}^{m} U_i\big|}{\sqrt{m(n-m+1)}}\right\}.\end{aligned} \tag{4.29}$$

We have

Theorem 4.1. *Under H_0,*

(i) $$\lim_{n\to\infty} P\{\frac{\hat{\tau}}{n} \le x\} = x, \quad 0 \le x \le 1. \tag{4.30}$$

(ii) $$\frac{\tilde{\tau}}{n} \xrightarrow{\mathcal{D}} \eta, \tag{4.31}$$

where η is a random variable such that

$$P\{\eta = 0\} = P\{\eta = 1\} = \frac{1}{2}.$$

Proof. Both (4.30) and (4.31) follow directly from Theorem 2.4.14 of Csörgő and Horváth (1997). □

From Theorem 4.1(i), we see that under the null hypothesis of no change, the estimator $\hat{\tau}$ is asymptotically uniformly distributed on the range $\{1, \cdots, n\}$, and views any of the integers between 1 and n as a possible time of change with the same probability irrespective of the value of the test statistic. This corresponds well to the interpretation of $\max|\sum_{i=1}^{k} U_i|$ as a uniform prior distribution of the place of change and is considered a desirable property (cf. Henderson (1990)). On the other hand, Theorem 4.1(ii) says that, the change-point estimator $\tilde{\tau}$ will be close to 0 or n for large n. This is a good property in the sense that H_0 can be interpreted as the case when the time of change is at the first or after the last observation.

We also have

Theorem 4.2. *Under H_1, if $\theta_{1,2} \neq 0$, then*

$$|\hat{\tau} - \tau| = O_P(1), \tag{4.32}$$

where $\theta_{1,2}$ *is defined in* (4.17), τ *and* $\hat{\tau}$ *are defined in* (4.1) *and* (4.28), *respectively.*

Proof. For the sake of simplicity of exposition we do the proof for statistic

$$\max_{1\le k<n} \frac{\sum_{i=1}^{k} U_i}{\sqrt{\sum_{i=1}^{n} U_i^2}}$$

and the corresponding change-point estimator

$$\hat{\tau} = \operatorname*{argmax}_{1\le k<n} \frac{\sum_{i=1}^{k} U_i}{\sqrt{\sum_{i=1}^{n} U_i^2}},$$

which are to be used for the one-sided alternative of $\theta_{1,2} > 0$. The proof of the more general statement for the two-sided alternative $\theta_{1,2} \neq 0$ follows immediately, here it will be omitted.

By the definition of $\hat{\tau}$, note that relation (4.32) is equivalent to

$$\begin{aligned} &\lim_{k\to\infty} \limsup_{n\to\infty} P\Big\{ \max_{1\le m\le \tau-k} \sum_{i=1}^{m} U_i \ge \max_{\tau-k<m<\tau+k} \sum_{i=1}^{m} U_i \Big\} \\ &\quad + \lim_{k\to\infty} \limsup_{n\to\infty} P\Big\{ \max_{\tau+k\le m<n} \sum_{i=1}^{m} U_i \ge \max_{\tau-k<m<\tau+k} \sum_{i=1}^{m} U_i \Big\} = 0, \end{aligned} \tag{4.33}$$

so we only need prove (4.33).

First, we show that the first term on the left side of (4.33) is zero. We have

$$\begin{aligned} &P\Big\{ \max_{1\le m\le \tau-k} \sum_{i=1}^{m} U_i \ge \max_{\tau-k<m<\tau+k} \sum_{i=1}^{m} U_i \Big\} \\ &\quad \le P\Big\{ \max_{1\le m\le \tau-k} \sum_{i=1}^{m} U_i \ge \max_{\tau-k<m\le\tau} \sum_{i=1}^{m} U_i \Big\} \\ &\quad = P\Big\{ \max_{1\le m\le \tau-k} \sum_{i=1}^{m} U_i \ge \sum_{i=1}^{\tau-k} U_i + \max_{\tau-k<m\le\tau} \sum_{i=\tau-k+1}^{m} U_i \Big\} \\ &\quad = P\Big\{ \exists j, 1\le j\le \tau-k : \sum_{i=1}^{j} U_i \ge \sum_{i=1}^{\tau-k} U_i + \max_{\tau-k<m\le\tau} \sum_{i=\tau-k+1}^{m} U_i \Big\} \\ &\quad = P\Big\{ \exists j, 1\le j\le \tau-k : \sum_{i=j+1}^{\tau-k} U_i + \max_{\tau-k<m\le\tau} \sum_{i=\tau-k+1}^{m} U_i \le 0 \Big\} \end{aligned}$$

$$= 1 - P\Big\{ \min_{1\le m\le \tau-k} \sum_{i=m+1}^{\tau-k} U_i + \max_{\tau-k<m\le\tau} \sum_{i=\tau-k+1}^{m} U_i > 0 \Big\}$$

$$= P\Big\{ \min_{1\le m\le \tau-k} \sum_{i=m+1}^{\tau-k} U_i + \max_{\tau-k<m\le\tau} \sum_{i=\tau-k+1}^{m} U_i \le 0 \Big\} \tag{4.34}$$

$$= P\Big\{ \frac{1}{\sqrt{k}} \min_{1\le m\le \tau-k} \sum_{i=m+1}^{\tau-k} U_i + \frac{1}{\sqrt{k}} \max_{\tau-k<m\le\tau} \sum_{i=\tau-k+1}^{m} U_i \le 0 \Big\}.$$

We note that $\{U_i = \sum_{j=1}^{n} h(Z_i, Z_j),\ i = 1, \cdots, \tau\}$ are identically distributed random variables under H_1, and hence

$$\min_{1\le m\le \tau-k} \sum_{i=m+1}^{\tau-k} U_i \stackrel{\mathcal{D}}{=} \min_{1\le m<\tau-k} \sum_{i=1}^{m} U_i. \tag{4.35}$$

Let

$$H_{12}(t) = \lambda H^{(1)}(t) + (1-\lambda) H^{(2)}(t),$$

$$\widetilde{H}_{12}(t) = \lambda \widetilde{H}^{(1)}(t) + (1-\lambda) \widetilde{H}^{(2)}(t),$$

where $H^{(1)}$, $H^{(2)}$, $\widetilde{H}^{(1)}$ and $\widetilde{H}^{(2)}$ are given by (1.4), (1.5), (4.2), and (4.5), respectively.

We can write

$$\begin{aligned} U_i &= \sum_{j=1}^{n} h(Z_i, Z_j) \\ &= \sum_{j=1}^{n} I(X_i > X_j, \delta_j = 1) - \sum_{j=1}^{n} I(X_i < X_j, \delta_i = 1) \\ &= n\Big\{ \widetilde{H}_n(X_i-) - [1 - H_n(X_i)]\delta_i \Big\} \\ &= n\Big\{ \widetilde{H}_{12}(X_i) - [1 - H_{12}(X_i)]\delta_i + O_P(\frac{1}{\sqrt{n}}) \Big\} \\ &= Q_i + nO_P(\frac{1}{\sqrt{n}}),\ i = 1, \cdots, \tau, \end{aligned}$$

where H_n and $\widetilde{H}_n$ are given by (2.11) and (2.12), respectively, and

$$Q_i = n\Big\{ \widetilde{H}_{12}(X_i) - [1 - H_{12}(X_i)]\delta_i \Big\}.$$

Note that $Q_1, \cdots, Q_\tau$ are independent and identically distributed with $EQ_i > 0$ by Remark 4.1, and

$$\begin{aligned}
D_q^2 &= Var(Q_i) = E(Q_i^2) - (EQ_i)^2 \\
&= n^2\Big\{E\widetilde{H}_{12}^2(X_i) + E[(1 - H_{12}(X_i))^2\delta_i] - 2E[\widetilde{H}_{12}(X_i)(1 - H_{12}(X_i))\delta_i]\Big\} \\
&\quad - n^2\Big\{\int \widetilde{H}_{12}(t)dH^{(1)}(t) - \int[1 - H_{12}(t)]d\widetilde{H}^{(1)}(t)\Big\}^2 \\
&= n^2\Big\{\int \widetilde{H}_{12}^2(t)dH^{(1)}(t) + \int[1 - H_{12}(t)]^2 d\widetilde{H}^{(1)}(t) \\
&\qquad - 2\int \widetilde{H}_{12}(t)[1 - H_{12}(t)]d\widetilde{H}^{(1)}(t) \\
&\qquad - \Big[\int \widetilde{H}_{12}(t)dH^{(1)}(t) - \int(1 - H_{12}(t))d\widetilde{H}^{(1)}(t)\Big]^2\Big\}.
\end{aligned}$$

We have

$$\sum_{i=1}^{m}\frac{U_i}{D_q} = \sum_{i=1}^{m}\frac{Q_i}{D_q} + mO_P(\frac{1}{\sqrt{n}}). \tag{4.36}$$

Let $\{g(k)\}$ be a sequence such that

$$g(k) < \tau - k,\ g(k) \to \infty, \frac{g(k)}{k} \to 0 \quad \text{as} \quad k \to \infty. \tag{4.37}$$

Denote

$$[x]^- = -\min(x, 0),$$

so if $x \geq 0$, then $[x]^- = 0$. Thus, by (4.36) and (4.37), we have

$$\begin{aligned}
&-\Big[\frac{1}{\sqrt{k}}\min_{1\leq m\leq g(k)}\sum_{i=1}^{m}\frac{U_i}{D_q}\Big]^- \\
&= -\Big[\min_{1\leq m\leq g(k)}\Big\{\sqrt{\frac{g(k)}{k}}\frac{1}{\sqrt{g(k)}}\sum_{i=1}^{m}\frac{Q_i - EQ_i}{D_q} + \frac{m}{\sqrt{k}}\Big(E\Big(\frac{Q_i}{D_q}\Big) + O_P(\frac{1}{\sqrt{n}})\Big)\Big\}\Big]^-.
\end{aligned} \tag{4.38}$$

As $EQ_i > 0$ and the partial sum process

$$\Big\{\frac{1}{\sqrt{g(k)}}\sum_{i=1}^{m}\frac{Q_i - EQ_i}{D_q}, m = 1, \cdots, g(k)\Big\} \stackrel{\mathcal{D}}{=} \Big\{W\Big(\frac{m}{g(k)}\Big) + o_P(1), m = 1, \cdots, g(k)\Big\},$$

where $W(\cdot)$ is a Wiener process, we get

$$-\Big[\frac{1}{\sqrt{k}}\min_{1\le m\le g(k)}\sum_{i=1}^{m}\frac{U_i}{D_q}\Big]^- = o_P(1) \tag{4.39}$$

when k is large enough by (4.38) and the continuity of the Wiener process at 0.

In addition, by the law of iterated logarithm and $EQ_i > 0$, we have

$$\begin{aligned}
&-\Big[\frac{1}{\sqrt{k}}\min_{g(k)<m<\tau-k}\sum_{i=1}^{m}\frac{U_i}{D_q}\Big]^- \\
&= -\Big[\min_{g(k)<m<\tau-k}\Big\{\sqrt{\frac{m}{k}}\frac{1}{\sqrt{m}}\sum_{i=1}^{m}\frac{Q_i-EQ_i}{D_q}+\frac{m}{\sqrt{k}}E\Big(\frac{Q_i}{D_q}\Big)+\frac{m}{\sqrt{k}}O_P\Big(\frac{1}{\sqrt{n}}\Big)\Big\}\Big]^- \\
&= -\Big[\min_{g(k)<m<\tau-k}\Big\{\frac{m}{\sqrt{k}}\Big(O_P\Big(\sqrt{\frac{\log\log m}{m}}\Big)+E\Big(\frac{Q_i}{D_q}\Big)+O_P\Big(\frac{1}{\sqrt{n}}\Big)\Big)\Big\}\Big]^- \\
&= o_P(1)
\end{aligned} \tag{4.40}$$

when k is large enough. Thus, by (4.35), (4.39) and (4.40), we have

$$\begin{aligned}
&-\Big[\frac{1}{\sqrt{k}}\min_{1\le m\le \tau-k}\sum_{i=m+1}^{\tau-k}\frac{U_i}{D_q}\Big]^- \\
&\overset{\mathcal{D}}{=} -\Big[\frac{1}{\sqrt{k}}\min_{1\le m<\tau-k}\sum_{i=1}^{m}\frac{U_i}{D_q}\Big]^- \\
&= -\Big[\frac{1}{\sqrt{k}}\min\Big\{\min_{1\le m\le g(k)}\sum_{i=1}^{m}\frac{U_i}{D_q},\min_{g(k)<m<\tau-k}\sum_{i=1}^{m}\frac{U_i}{D_q}\Big\}\Big]^- \\
&= \min\Big\{-\Big[\frac{1}{\sqrt{k}}\min_{1\le m\le g(k)}\sum_{i=1}^{m}\frac{U_i}{D_q}\Big]^-,-\Big[\frac{1}{\sqrt{k}}\min_{g(k)<m<\tau-k}\sum_{i=1}^{m}\frac{U_i}{D_q}\Big]^-\Big\} \\
&= o_P(1).
\end{aligned} \tag{4.41}$$

On the other hand,

$$\begin{aligned}
&\frac{1}{\sqrt{k}}\max_{\tau-k<m\le\tau}\sum_{i=\tau-k+1}^{m}\frac{U_i}{D_q} \\
&\ge \frac{1}{\sqrt{k}}\sum_{i=\tau-k+1}^{\tau}\frac{U_i}{D_q} \\
&\overset{\mathcal{D}}{=} \frac{1}{\sqrt{k}}\sum_{i=1}^{k}\frac{U_i}{D_q} \\
&= \sqrt{k}\Big(\frac{1}{k}\sum_{i=1}^{k}\frac{Q_i}{D_q}+O_P\Big(\frac{1}{\sqrt{n}}\Big)\Big).
\end{aligned} \tag{4.42}$$

By the strong law of large numbers, as $k \to \infty$

$$\frac{1}{k}\sum_{i=1}^{k}\frac{Q_i}{D_q} \xrightarrow{a.s} E\Big(\frac{Q_i}{D_q}\Big) > 0.$$

Thus,

$$\begin{aligned}&\frac{1}{\sqrt{k}}\max_{\tau-k<m\le\tau}\sum_{i=\tau-k+1}^{m}\frac{U_i}{D_q}\\ &\ge \sqrt{k}\Big(\frac{1}{k}\sum_{i=1}^{k}\frac{Q_i}{D_q}+O_P\big(\frac{1}{\sqrt{n}}\big)\Big) \to +\infty \quad \text{as} \quad k\to\infty.\end{aligned} \tag{4.43}$$

Putting (4.41) and (4.43) together, we get

$$\lim_{k\to\infty}\limsup_{n\to\infty}P\Big\{\frac{1}{\sqrt{k}}\min_{1\le m\le\tau-k}\sum_{i=m+1}^{\tau-k}\frac{U_i}{D_q}+\frac{1}{\sqrt{k}}\max_{\tau-k<m\le\tau}\sum_{i=\tau-k+1}^{m}\frac{U_i}{D_q}\le 0\Big\}=0. \tag{4.44}$$

Therefore, by (4.34) and (4.44), we can get the first term on the left side of (4.33)

$$\lim_{k\to\infty}\limsup_{n\to\infty}P\Big\{\max_{1\le m\le\tau-k}\sum_{i=1}^{m}U_i \ge \max_{\tau-k<m<\tau+k}\sum_{i=1}^{m}U_i\Big\}=0. \tag{4.45}$$

To show that the second term on the left side of (4.33) is zero, we note that

$$\sum_{i=1}^{m}U_i = \sum_{i=m+1}^{n}(-U_i)$$

since $\sum_{i=1}^{n} U_i = 0$. If we consider the sequence from the opposite direction, let

$$Z_i' = Z_{n+1-i},\ i = 1,\cdots,n$$

and

$$U_i' = -U_{n+1-i},\ i = 1,\cdots,n.$$

Thus,

$$(Z_1',\cdots,Z_n') = (Z_n,\cdots,Z_1),$$

$$(U_1',\cdots,U_n') = (-U_n,\cdots,-U_1),$$

$$\sum_{i=1}^{m} U_i = \sum_{i=1}^{n-m} U_i',$$

and

$$\tau' = n + 1 - (\tau + 1) = n - \tau$$

is the change-point of the opposite direction sequence.

Let $m' = n - m$, so

$$\lim_{k\to\infty} \limsup_{n\to\infty} P\Big\{ \max_{\tau+k\le m<n} \sum_{i=1}^{m} U_i \ge \max_{\tau-k<m<\tau+k} \sum_{i=1}^{m} U_i \Big\}$$

$$= \lim_{k\to\infty} \limsup_{n\to\infty} P\Big\{ \max_{1\le n-m\le \tau'-k} \sum_{i=1}^{n-m} U_i' \ge \max_{\tau'-k<n-m<\tau'+k} \sum_{i=1}^{n-m} U_i' \Big\} \qquad (4.46)$$

$$= \lim_{k\to\infty} \limsup_{n\to\infty} P\Big\{ \max_{1\le m'\le \tau'-k} \sum_{i=1}^{m'} U_i' \ge \max_{\tau'-k<m'<\tau'+k} \sum_{i=1}^{m'} U_i' \Big\}.$$

Thus, we can use the same method as above to prove that the second term on the left side of (4.33) is zero; the proof is omitted. □

Note that the relation (4.32) can be restated in terms of the change-point parameter λ and its estimator $\hat{\lambda} = \hat{\tau}/n$ as

$$|\hat{\lambda} - \lambda| = O_P(\frac{1}{n}),$$

which means the optimal rate of convergence for a point estimator of λ is $\frac{1}{n}$.

We do some simulation study to investigate the precision of the change-point estimators to the true value in Section 4.4.

4.3 Asymptotic Distribution of the Test Statistic

Next we discuss the asymptotics of test statistic (2.18) based on $\sum_{i=1}^{k} U_i$. First of all, let's quote a useful result due to Dvoretzky (1972) in the following Theorem (cf. Theorem 2.4.1 of Sen (1981)).

Theorem (Dvoretzky, 1972). *For a sequence $\{(X_{n,k},\ 0 \le k \le k_n);\ n \ge 1\}$ of random variables (not necessarily independent), we set $X_{n,0} = 0$,*

$$S_{n,k} = X_{n,0} + \cdots + X_{n,k} \ \text{ for } k = 1, \cdots, k_n,$$

and let $\mathcal{G}_{n,k}$ be the sigma-field generated by $S_{n,k}$ for $k \ge 1$ and $n \ge 1$ (where $\mathcal{G}_{n,0}$ is the trivial sigma-field for every $n \ge 1$). Assume that $EX^2_{n,i} < \infty$, $\forall i \ge 1, n \ge 1$. Let then

$$\mu_{n,k} = E(X_{n,k}|\mathcal{G}_{n,k-1}) \ \text{ and } \ \sigma^2_{n,k} = E(X^2_{n,k}|\mathcal{G}_{n,k-1}) - \mu^2_{n,k}$$

for $1 \le k \le k_n$, $n \ge 1$ and assume that $k_n \to \infty$ as $n \to \infty$. Suppose that as $n \to \infty$,

$$\sum_{k=1}^{k_n} \mu_{n,k} \xrightarrow{P} 0, \quad \sum_{k=1}^{k_n} \sigma^2_{n,k} \xrightarrow{P} 1,$$

and the conditional Lindeberg condition holds, that is, for every $\varepsilon > 0$,

$$\sum_{k=1}^{k_n} E(X^2_{n,k} I(|X_{n,k}| > \varepsilon)|\mathcal{G}_{n,k-1}) \xrightarrow{P} 0 \quad \text{as} \quad n \to \infty.$$

Then

$$S_{n,k_n} \xrightarrow{\mathcal{D}} N(0,1).$$

We'll use this Theorem for the proof of the following Lemma 4.3.

Lemma 4.3. *Under H_1,*

$$\frac{1}{\sqrt{\tau}\sigma_{n\tau}}\Big(\sum_{i=1}^{\tau} U_i - \tau\mu_{n\tau}\Big) \xrightarrow{\mathcal{D}} N(0,1), \tag{4.47}$$

where U_i is defined in (2.4), τ, $\mu_{n\tau}$, and $\sigma^2_{n\tau}$ are given by (4.1), (4.10), and (4.11), respectively.

Proof. Let

$$\xi_{ni} = \frac{\sum_{j=\tau+1}^{n} h(Z_i, Z_j) - \mu_{n\tau}}{\sqrt{\tau}\sigma_{n\tau}},\ i = 1, \cdots, \tau, \tag{4.48}$$

$$a_{ni} = E(\xi_{ni}|\mathcal{F}_{n,i-1}),\ i = 1, \cdots, \tau,$$

and

$$\begin{aligned} b_{ni}^2 &= Var(\xi_{ni}|\mathcal{F}_{n,i-1}) \\ &= E\Big\{(\xi_{ni} - a_{ni})^2|\mathcal{F}_{n,i-1}\Big\},\ i = 1, \cdots, \tau. \end{aligned}$$

Now, under H_1, by (4.8), and by (4.10) of Lemma 4.1, we have

$$\begin{aligned} a_{ni} &= E\left\{\frac{\sum_{j=\tau+1}^{n} h(Z_i, Z_j) - \mu_{n\tau}}{\sqrt{\tau}\sigma_{n\tau}}\Big|\mathcal{F}_{n,i-1}\right\} \\ &= \frac{1}{\sqrt{\tau}\sigma_{n\tau}}\Big(E\Big\{\sum_{j=\tau+1}^{n} h(Z_i, Z_j)|\mathcal{F}_{n,i-1}\Big\} - \mu_{n\tau}\Big) \\ &= \frac{1}{\sqrt{\tau}\sigma_{n\tau}}(\mu_{ni} - \mu_{n\tau}) \\ &= 0,\ i = 1, \cdots, \tau, \end{aligned}$$

and hence

$$\sum_{i=1}^{\tau} a_{ni} = 0. \tag{4.49}$$

Also, by (4.9) and Lemma 4.1,

$$\begin{aligned} b_{ni}^2 &= E\left\{\left[\frac{\sum_{j=\tau+1}^{n} h(Z_i, Z_j) - \mu_{n\tau}}{\sqrt{\tau}\sigma_{n\tau}}\right]^2\Big|\mathcal{F}_{n,i-1}\right\} \\ &= \frac{1}{\tau\sigma_{n\tau}^2}E\Big\{\Big[\sum_{j=\tau+1}^{n} h(Z_i, Z_j) - \mu_{n\tau}\Big]^2\Big|\mathcal{F}_{n,i-1}\Big\} \\ &= \frac{1}{\tau\sigma_{n\tau}^2}\sigma_{n\tau}^2 \\ &= \frac{1}{\tau},\ i = 1, \cdots, \tau, \end{aligned}$$

and so

$$\sum_{i=1}^{\tau} b_{ni}^2 = 1. \tag{4.50}$$

In addition, we note that $\xi_{n1}, \cdots, \xi_{n\tau}$ are conditionally i.i.d. given $Z_{\tau+1}, \cdots, Z_n$. So for every $\varepsilon > 0$, by Lemma 4.1 and (4.16),

$$\begin{aligned}
&\sum_{i=1}^{\tau} E\Big\{\xi_{ni}^2 I(|\xi_{ni}| > \varepsilon)|\mathcal{F}_{n,i-1}\Big\} \\
&= \sum_{i=1}^{\tau} E\Big\{\xi_{ni}^2 I(|\xi_{ni}| > \varepsilon)|\mathcal{F}_{n0}\Big\} \\
&= \sum_{i=1}^{\tau} E\left\{\left[\frac{\sum_{j=\tau+1}^{n} h(Z_i, Z_j) - \mu_{n\tau}}{\sqrt{\tau}\sigma_{n\tau}}\right]^2 I\Big(\Big|\sum_{j=\tau+1}^{n} h(Z_i, Z_j) - \mu_{n\tau}\Big| > \varepsilon\sqrt{\tau}\sigma_{n\tau}\Big)|\mathcal{F}_{n0}\right\} \\
&= \frac{\tau}{\tau\sigma_{n\tau}^2} E\Big\{\Big[\sum_{j=\tau+1}^{n} h(Z_\tau, Z_j) - \mu_{n\tau}\Big]^2 I\Big(\Big|\sum_{j=\tau+1}^{n} h(Z_\tau, Z_j) - \mu_{n\tau}\Big| > \varepsilon\sqrt{\tau}\sigma_{n\tau}\Big)|\mathcal{F}_{n0}\Big\} \\
&= \frac{1}{\sigma_{n\tau}^2} E\Big\{\Big[\sum_{j=\tau+1}^{n} h(Z_\tau, Z_j) - \mu_{n\tau}\Big]^2 I\Big(\Big|\sum_{j=\tau+1}^{n} h(Z_\tau, Z_j) - \mu_{n\tau}\Big| > \varepsilon\sqrt{\tau}\sigma_{n\tau}\Big)|\mathcal{F}_{n0}\Big\} \\
&\to 0, \text{ as } n \to \infty.
\end{aligned} \tag{4.51}$$

Thus the conditions of Theorem (Dvoretzky, 1972) are satisfied by (4.49)-(4.51), and the central limit theorem holds for $\sum_{i=1}^{\tau} \xi_{ni}$, that is,

$$\sum_{i=1}^{\tau} \xi_{ni} \xrightarrow{\mathcal{D}} N(0,1). \tag{4.52}$$

Note that by (2.3), (2.5), and (4.48),

$$\begin{aligned}
\sum_{i=1}^{\tau} \xi_{ni} &= \sum_{i=1}^{\tau} \left(\frac{\sum_{j=\tau+1}^{n} h(Z_i, Z_j) - \mu_{n\tau}}{\sqrt{\tau}\sigma_{n\tau}}\right) \\
&= \frac{1}{\sqrt{\tau}\sigma_{n\tau}} \sum_{i=1}^{\tau} \Big[\sum_{i=\tau+1}^{n} h(Z_i, Z_j) - \mu_{n\tau}\Big] \\
&= \frac{1}{\sqrt{\tau}\sigma_{n\tau}} \Big[\sum_{i=1}^{\tau} \sum_{i=\tau+1}^{n} h(Z_i, Z_j) - \tau\mu_{n\tau}\Big] \\
&= \frac{1}{\sqrt{\tau}\sigma_{n\tau}} \Big(\sum_{i=1}^{\tau} U_i - \tau\mu_{n\tau}\Big).
\end{aligned} \tag{4.53}$$

Thus (4.47) follows from (4.52) and (4.53). □

Let

$$A_1^0 = \int [\widetilde{H}^{(2)}(t)]^2 dH^{(1)}(t),$$
$$A_2^0 = \int [1 - H^{(2)}(t)]^2 d\widetilde{H}^{(1)}(t),$$
$$A_3^0 = \int \widetilde{H}^{(2)}(t)[1 - H^{(2)}(t)] d\widetilde{H}^{(1)}(t),$$

and

$$\sigma_\tau^2 = (n-\tau)^2 (A_1^0 + A_2^0 - 2A_3^0 - \theta_{1,2}^2), \tag{4.54}$$

where $H^{(1)}(t), H^{(2)}(t), \widetilde{H}^{(1)}(t)$, and $\widetilde{H}^{(2)}(t)$ are defined in (1.4), (1.5), (4.2), and (4.5), respectively, and $\tau, \theta_{1,2}$ are defined in (4.1) and (4.17), respectively. We have

Corollary 4.3. *Under H_1,*

$$\frac{1}{\sqrt{\tau}\sigma_\tau}\Big(\sum_{i=1}^{\tau} U_i - \tau\mu_{n\tau}\Big) \xrightarrow{\mathcal{D}} N(0,1), \tag{4.55}$$

where U_i is defined in (2.4), $\tau, \mu_{n\tau}$, *and* σ_τ^2 *are given by* (4.1), (4.10), *and* (4.54), *respectively.*

Proof. By the weak convergence of the empirical distribution function to a Brownian bridge, we know that

$$\sup_t \left| H_{n-\tau}^{(2)}(t) - H^{(2)}(t) \right| = O_P\left(\frac{1}{\sqrt{n}}\right)$$

and

$$\sup_t \left| \widetilde{H}_{n-\tau}^{(2)}(t) - \widetilde{H}^{(2)}(t) \right| = O_P\left(\frac{1}{\sqrt{n}}\right).$$

Thus,

$$\int [\widetilde{H}_{n-\tau}^{(2)}(t-)]^2 dH^{(1)}(t) \xrightarrow{P} \int [\widetilde{H}^{(2)}(t)]^2 dH^{(1)}(t),$$

i.e.

$$A_1(\tau) \xrightarrow{P} A_1^0, \tag{4.56}$$

$$\int [1 - H^{(2)}_{n-\tau}(t)]^2 d\widetilde{H}^{(1)}(t) \xrightarrow{P} \int [1 - H^{(2)}(t)]^2 d\widetilde{H}^{(1)}(t),$$

that is,

$$A_2(\tau) \xrightarrow{P} A_2^0, \tag{4.57}$$

$$\int \widetilde{H}^{(2)}_{n-\tau}(t-)[1 - H^{(2)}_{n-\tau}(t)] d\widetilde{H}^{(1)}(t) \xrightarrow{P} \int \widetilde{H}^{(2)}(t)[1 - H^{(2)}(t)] d\widetilde{H}^{(1)}(t),$$

i.e.

$$A_3(\tau) \xrightarrow{P} A_3^0, \tag{4.58}$$

and

$$\int \widetilde{H}^{(2)}_{n-\tau}(t-) dH^{(1)}(t) - \int [1 - H^{(2)}_{n-\tau}(t)] d\widetilde{H}^{(1)}(t)$$
$$\xrightarrow{P} \int \widetilde{H}^{(2)}(t) dH^{(1)}(t) - \int [1 - H^{(2)}(t)] d\widetilde{H}^{(1)}(t),$$

that is,

$$\frac{1}{n-\tau}\mu_{n\tau} \xrightarrow{P} \theta_{1,2} \tag{4.59}$$

by (4.10) of Lemma 4.1 and (4.24) of Corollary 4.1.

Thus, by (4.11), (4.54), (4.56)-(4.59), we have

$$\frac{\sigma^2_{n\tau}}{\sigma^2_\tau} \xrightarrow{P} 1. \tag{4.60}$$

Therefore (4.55) follows from Lemma 4.3 and (4.60). □

Theorem 4.3. *Under* H_1, *if* $\theta_{1,2} \neq 0$, *then*

$$\frac{1}{\sqrt{n\lambda}\sigma_\tau}\Big(\sum_{i=1}^{\hat{\tau}} U_i - \tau\mu_{n\tau}\Big) \xrightarrow{\mathcal{D}} N(0,1), \tag{4.61}$$

where $U_i, \tau, \theta_{1,2}$, *and* $\hat{\tau}$ *are defined in* (2.4), (4.1), (4.17), *and* (4.28), *respectively, and* $\mu_{n\tau}$ *and* σ^2_τ *are given in* (4.10) *and* (4.54), *respectively.*

Proof. If $\hat{\tau} < \tau$, then, by (4.1),

$$\begin{aligned} &\frac{1}{\sqrt{n\lambda}\sigma_\tau}\Big(\sum_{i=1}^{\hat{\tau}} U_i - \tau\mu_{n\tau}\Big) \\ &= \frac{1}{\sqrt{n\lambda}\sigma_\tau}\Big(\sum_{i=1}^{\tau} U_i - \tau\mu_{n\tau}\Big) - \frac{1}{\sqrt{n\lambda}\sigma_\tau}\sum_{i=\hat{\tau}+1}^{\tau} U_i \\ &= \sqrt{\frac{[n\lambda]}{n\lambda}} \cdot \frac{1}{\sqrt{\tau}\sigma_\tau}\Big(\sum_{i=1}^{\tau} U_i - \tau\mu_{n\tau}\Big) - \frac{1}{\sqrt{n\lambda}\sigma_\tau}\sum_{i=\hat{\tau}+1}^{\tau} U_i. \end{aligned} \tag{4.62}$$

Now, under H_1, by (2.4), (4.1), (4.54), (4.10) of Lemma 4.1, and Theorem 4.2,

$$\begin{aligned} &\left| -\frac{1}{\sqrt{n\lambda}\sigma_\tau}\sum_{i=\hat{\tau}+1}^{\tau} U_i \right| \\ &\leq \frac{1}{\sqrt{n\lambda}\sigma_\tau}\sum_{i=\hat{\tau}+1}^{\tau} |U_i| \\ &\leq \frac{1}{\sqrt{n\lambda}\sigma_\tau} n(\tau - \hat{\tau}) \\ &= \frac{n(\tau - \hat{\tau})}{\sqrt{n\lambda}(n-\lambda)\sqrt{A_1^0 + A_2^0 - 2A_3^0 - \theta_{1,2}^2}} \\ &= o_P(1). \end{aligned} \tag{4.63}$$

In addition,

$$\lim_{n\to\infty} \frac{[n\lambda]}{n\lambda} = 1. \tag{4.64}$$

Thus (4.61) follows from (4.62)-(4.64) and Corollary 4.3.

If $\hat{\tau} > \tau$, we have

$$\begin{aligned} &\frac{1}{\sqrt{n\lambda}\sigma_\tau}\Big(\sum_{i=1}^{\hat{\tau}} U_i - \tau\mu_{n\tau}\Big) \\ &= \sqrt{\frac{[n\lambda]}{n\lambda}} \cdot \frac{1}{\sqrt{\tau}\sigma_\tau}\Big(\sum_{i=1}^{\tau} U_i - \tau\mu_{n\tau}\Big) + \frac{1}{\sqrt{n\lambda}\sigma_\tau}\sum_{i=\tau+1}^{\hat{\tau}} U_i. \end{aligned} \tag{4.65}$$

Similarly, we can get

$$\left| \frac{1}{\sqrt{n\lambda}\sigma_\tau}\sum_{i=\tau+1}^{\hat{\tau}} U_i \right| = o_P(1). \tag{4.66}$$

Thus (4.61) follows immediately from (4.64)-(4.66) and Corollary 4.3. □

By Theorem 4.3, we can also get

Corollary 4.4. *Under* H_1, *if* $\theta_{1,2} > 0$, *then*

$$\frac{1}{\sqrt{n\lambda}\sigma_\tau}\Big(\max_{1\le k<n}\sum_{i=1}^{k} U_i - \tau\mu_{n\tau}\Big) \xrightarrow{\mathcal{D}} N(0,1), \tag{4.67}$$

and if $\theta_{1,2} < 0$, *then*

$$\frac{1}{\sqrt{n\lambda}\sigma_\tau}\Big(\min_{1\le k<n}\sum_{i=1}^{k} U_i - \tau\mu_{n\tau}\Big) \xrightarrow{\mathcal{D}} N(0,1). \tag{4.68}$$

Proof. By the definition of $\hat{\tau}$ in (4.28), we note that

$$\sum_{i=1}^{\hat{\tau}} U_i = \begin{cases} \max\limits_{1\le k<n} \sum\limits_{i=1}^{k} U_i & \text{if } \theta_{1,2} > 0, \\ \min\limits_{1\le k<n} \sum\limits_{i=1}^{k} U_i & \text{if } \theta_{1,2} < 0. \end{cases} \tag{4.69}$$

Thus (4.67) and (4.68) follow immediately from (4.69) and Theorem 4.3. □

The consistency of the test based on (2.19) easily follows from Theorem 4.3. It can also be used to approximate the power of the test for large sample size and its proof shows that the two-sample and the change-point test statistics have identical asymptotic distributions. A similar statement is true for tests of change-point based on maximum likelihood, rank or sign statistics (cf. Gombay (1994), Gombay and Horváth (1996), Gombay and Jin (1996)). This is important for the asymptotic efficiency considerations, provided the measure of efficiency is not based on contiguous alternatives.

4.4 Simulation Study and Applications

In this section, we first do some simulation study on the change-point estimators and then we apply our test based on (2.19) to some examples from clinical trials and also discuss the change-point estimator $\hat{\tau}$ based on (4.28) to illustrate the calculations.

4.4.1 Simulation study

To illustrate the proposed change-point estimators, we would like to check the precision of the estimated mean of the change-point estimator to the true

Table 4.1. *Comparison of true values and estimating values of change-point* ($\lambda = 0.5$)

n	μ_1	μ_2	μ_c	$\tau = [n\lambda]$	Simulated Mean of $\hat{\tau}$
50	1.0	1.5	3.0	25	24.78
	1.0	2.0	3.0	25	24.77
	1.0	2.5	3.0	25	24.57
	1.0	3.0	3.0	25	24.74
	1.0	3.5	3.0	25	24.55
100	1.0	1.5	3.0	50	49.80
	1.0	2.0	3.0	50	49.78
	1.0	2.5	3.0	50	49.53
	1.0	3.0	3.0	50	49.50
	1.0	3.5	3.0	50	49.55
200	1.0	1.5	3.0	100	100.32
	1.0	2.0	3.0	100	99.37
	1.0	2.5	3.0	100	99.57
	1.0	3.0	3.0	100	99.54
	1.0	3.5	3.0	100	99.39
500	1.0	1.5	3.0	250	248.99
	1.0	2.0	3.0	250	249.80
	1.0	2.5	3.0	250	249.33
	1.0	3.0	3.0	250	249.04
	1.0	3.5	3.0	250	249.45

Table 4.2. *Comparison of simulated values $\hat{\tau}$ and $\tilde{\tau}$ with true values* $(n = 100)$

μ_1	μ_2	μ_c	λ	$\tau = [n\lambda]$	Simulated Mean of $\hat{\tau}$	Simulated Mean of $\tilde{\tau}$
1.0	1.5	3.0	0.3	30	41.59	39.87
			0.4	40	45.04	43.27
			0.5	50	49.43	46.84
			0.6	60	54.46	50.19
			0.7	70	57.35	52.05
1.0	2.0	3.0	0.3	30	37.08	34.84
			0.4	40	43.27	41.44
			0.5	50	49.52	47.82
			0.6	60	55.88	53.75
			0.7	70	61.26	60.12
1.0	2.5	3.0	0.3	30	35.54	32.21
			0.4	40	42.27	40.48
			0.5	50	49.57	48.39
			0.6	60	57.01	56.31
			0.7	70	63.56	63.46
1.0	3.0	3.0	0.3	30	33.97	30.79
			0.4	40	41.56	39.77
			0.5	50	49.59	48.54
			0.6	60	57.24	57.33
			0.7	70	64.42	65.45
1.0	3.5	3.0	0.3	30	33.42	30.55
			0.4	40	41.13	39.51
			0.5	50	49.40	48.53
			0.6	60	57.62	58.06
			0.7	70	64.94	66.71

value of change-point $\tau = [n\lambda]$ through the Monte Carlo simulation study. We performed $N = 5000$ simulations for each case. In each case to be considered, we assumed that

$$F^{(1)} = \exp(\mu_1), \quad F^{(2)} = \exp(\mu_2), \quad G = \exp(\mu_c).$$

We considered means $\mu_2 = 1.5, 2.0, 2.5, 3.0, 3.5$. The particular feature of the underlying sampling situation is that in each case we have $\lambda = 0.5$ so that preand after-change samples are balanced in size. In this case, we considered sample sizes $n = 50$, 100, 200, 500, the results are reported in Table 4.1. From Table 4.1, we can see that, the simulated mean of $\hat{\tau}$ fits the true value of τ quite well.

The precision of the estimated mean of the change-point estimator $\hat{\tau}$ and the estimated mean of the weighted change-point estimator $\tilde{\tau}$ was compared for sample size $n = 100$ and varying λ's, $\lambda = 0.3, 0.4, 0.5, 0.6, 0.7$. The results are reported in Table 4.2. From Table 4.2, we can see that, the estimation precision increases as the distance of mean $|\mu_2 - \mu_1|$ increases or the true value of change-point is close to middle, which is quite reasonable. Overall, the performance of $\hat{\tau}$ gets somewhat worse for unbalanced sample sizes. In such a situation slight improvements have been obtained by considering the weighted version $\tilde{\tau}$.

4.4.2 Applications

Example 4.1. Crowley and Hu (1977) gave survival times of 103 potential heart transplant recipients from their date of acceptance into the Stanford heart transplant program. The survival time is said to be uncensored or censored depending on whether the date last seen is the date of death or the closing date of the present study, April 1, 1974. (The survival time for two patients who were deselected was censored by dates prior to the closing date on which they were lost to follow-up.). The patients entered the study randomly between 1967 and 1974. See Crowley and Hu (1977), and Kalblfeisch and Prentice (1980) for detailed descriptions.

We reproduce the data provided by Crowley and Hu (1977), and Kalblfeisch

Table 4.3. *Stanford heart transplant data*

i	1	2	3	4	5	6	7	8	9
X_i	6	50	6	16	39	18	3	675	40
Tr_i	0	0	0	1	1	0	0	1	0
i	10	11	12	13	14	15	16	17	18
X_i	85	58	153	8	81	1386	1	308	36
Tr_i	0	1	1	0	1	1	0	1	0
i	19	20	21	22	23	24	25	26	27
X_i	43	37	28	1032	51	733	219	1799+	1400+
Tr_i	1	0	1	1	1	1	1	1	0
i	28	29	30	31	32	33	34	35	36
X_i	263	72	35	852	16	77	1586+	1571+	12
Tr_i	0	1	0	1	0	1	1	1	0
i	37	38	39	40	41	42	43	44	45
X_i	100	66	5	53	1407+	1321+	3	2	40
Tr_i	1	1	1	1	1	1	0	0	0
i	46	47	48	49	50	51	52	53	54
X_i	45	995	72	9	1141+	979	285	102	188
Tr_i	1	1	1	0	1	1	1	0	1
i	55	56	57	58	59	60	61	62	63
X_i	3	61	941+	149	342	915+	427+	68	2
Tr_i	0	1	1	0	1	1	0	1	0
i	64	65	66	67	68	69	70	71	72
X_i	69	841+	583	78	32	285	68	670+	30
Tr_i	1	1	1	1	1	1	1	1	1
i	73	74	75	76	77	78	79	80	81
X_i	620+	596+	90	17	2	545+	21	515+	96
Tr_i	1	1	0	1	0	1	0	1	1
i	82	83	84	85	86	87	88	89	90
X_i	482+	445+	80	334	5	397+	110	370+	207
Tr_i	1	1	1	1	1	1	1	1	1
i	91	92	93	94	95	96	97	98	99
X_i	186	340	340+	265+	165	16	180+	131+	109+
Tr_i	1	0	1	1	1	1	1	1	1
i	100	101	102	103					
X_i	21	39+	31+	11+					
Tr_i	0	1	0	0					

Sources: Crowley and Hu (1977), and Kalbfleisch and Prentice (1980).

and Prentice (1980), and list sequentially in order of date of acceptance in Table 4.3, where i denotes the sequential order based on the date of patient acceptance. X_i denotes the ith patient's survival time which was recorded in days starting with the date of acceptance, the plus sign "+" indicates a censored observation. Tr_i denotes the ith patient's transplant status, 1 = transplanted, 0 = not transplanted.

Let's first analyse these data of size $n = 103$. Using the original definition of score function h, which is given by (2.2), we calculate the values of U_i ($i = 1, \cdots, 103$) by (2.4) and obtain the value of the test statistic

$$\max_{1 \le k \le 103} \frac{|\sum_{i=1}^{k} U_i|}{\sqrt{\sum_{i=1}^{103} U_i^2}} = 1.398.$$

Our test based on (2.19) gives P-value 0.040, it indicates a change in the survival time distribution over the study period 1967-1974. We get the estimating value of the change-point $\hat{\tau}$=49. This concurs with Kalbfleisch and Prentice's use of time dependent covariates in the Cox proportional hazards model. To analyse these data in a more detailed fashion, we made a scatterplot of $\{\sum_{i=1}^{k} U_i, k = 1, \cdots, 103\}$ in Figure 4.1. It shows that all values of $\sum_{i=1}^{k} U_i$ $(1 \le k < 103)$ are negative, the values of $\sum_{i=1}^{k} U_i$ have a decrease trend before $\hat{\tau}$=49 and have an increase trend after $\hat{\tau}$=49, which support the claims described in Corollary 4.2, Remark 4.2, and Remark 4.3.

Next we only consider the data which were from those patients who received a transplant. The data set size is 69. The value of the test statistic is

$$\max_{1 \le k \le 69} \frac{|\sum_{i=1}^{k} U_i|}{\sqrt{\sum_{i=1}^{69} U_i^2}} = 1.028.$$

Our test based on (2.19) gives P-value 0.241, so the null hypothesis of no change in survival time distribution is not rejected. This indicates the increase in survival time of the patients is not due to the change in transplant technique.

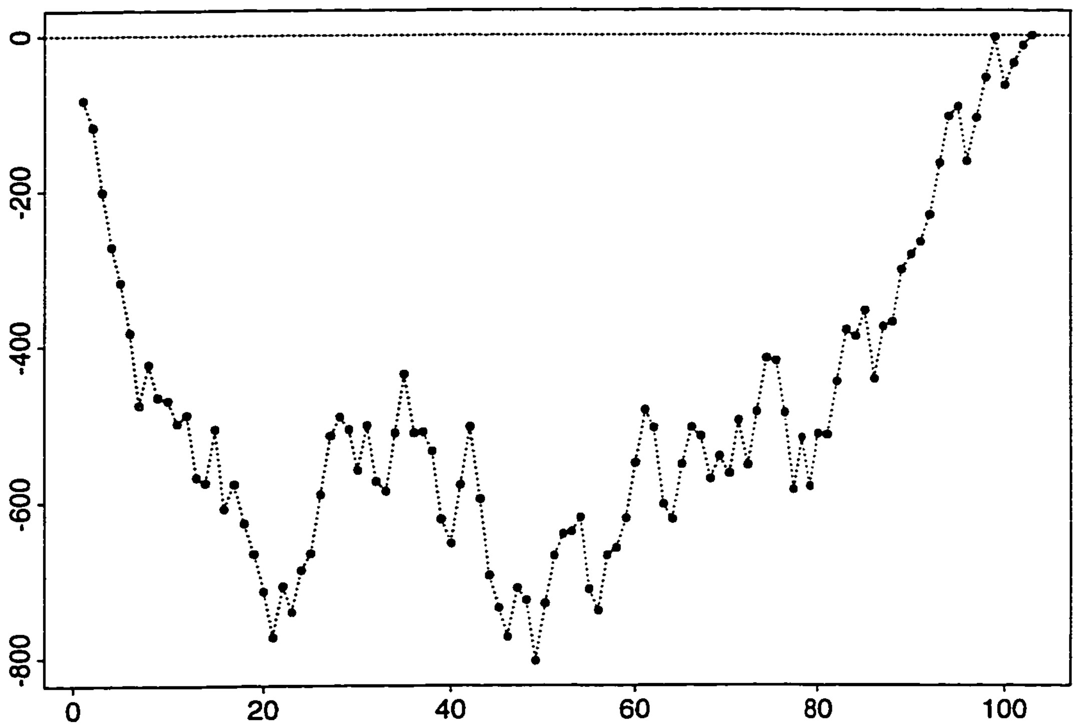

Figure 4.1. *Plot of* $\sum_{i=1}^{k} U_i, k = 1, \cdots, 103$ *for the Stanford heart transplant data.*

Example 4.2. The Data Set II of Kalbfleisch and Prentice (1980) gave the data for a part of a large clinical trial carried out by the Radiation Therapy Oncology Group in the United States. 195 patients with squamous carcinoma entered the study randomly between 1968 and 1972. See Kalbfleisch and Prentice (1980) for a description. These data are reproduced and listed sequentially in order of the patient entry date in Table 4.4.

Table 4.4. *Radiation Therapy Oncology Group data*

i	1	2	3	4	5	6	7	8
X_i	631	270	327	243	916	1823+	637	235
i	9	10	11	12	13	14	15	16
X_i	255	184	1064	414	216	324	480	245
i	17	18	19	20	21	22	23	24
X_i	1565+	560	376	911	279	144	1092	94
i	25	26	27	28	29	30	31	32
X_i	177	1472+	526	173	575	222	167	1565
i	33	34	35	36	37	38	39	40
X_i	134	256	404	1495+	162	262	307	782
i	41	42	43	44	45	46	47	48
X_i	661	1609+	546	1766+	374	1489+	1446+	74
i	49	50	51	52	53	54	55	56
X_i	301	328	459	446	1644+	494	279	915
i	57	58	59	60	61	62	63	64
X_i	228	127	1574	561	370	805	192	273
i	65	66	67	68	69	70	71	72
X_i	1377+	407	929	548	1317+	1317+	517	1307+
i	73	74	75	76	77	78	79	80
X_i	230	763	172	1455+	1234+	544	800	1460+
i	81	82	83	84	85	86	87	88
X_i	785	714	338	432	1312+	351	205	1219+
i	89	90	91	92	93	94	95	96
X_i	11	666	147	1060+	477	1058+	1312+	696
i	97	98	99	100	101	102	103	104
X_i	112	308	15	130	296	293	545	1086+
i	105	106	107	108	109	110	111	112
X_i	1250+	147	726	310	599	998+	1089+	382
i	113	114	115	116	117	118	119	120
X_i	932+	264	11	911+	89	525	532+	637
i	121	122	123	124	125	126	127	128
X_i	112	1095+	170	943+	191	928+	918+	825+
i	129	130	131	132	133	134	135	136
X_i	99	99	933+	461	347	372	731+	363
i	137	138	139	140	141	142	143	144
X_i	238	593+	219	465	446	553	274	723+

Table 4.4. *Radiation Therapy Oncology Group data (Continued)*

i	145	146	147	148	149	150	151	152
X_i	532	154	369	541	107	854+	822+	775
i	153	154	155	156	157	158	159	160
X_i	336	513	914+	757	794+	105	733+	600+
i	161	162	163	164	165	166	167	168
X_i	266	317	407	346	518	395	81	608
i	169	170	171	172	173	174	175	176
X_i	760+	343	324	254	751+	334	275	546+
i	177	178	179	180	181	182	183	184
X_i	112	182+	209	208	174	651+	672+	291
i	185	186	187	188	189	190	191	192
X_i	498	276+	90+	213	38	128	445+	159
i	193	194	195					
X_i	219	173	413+					

Source: Kalbfleisch and Prentice (1980).
NOTE: i denotes the sequential order based on the patient entry date. X_i denotes the ith patient's survival time which was recorded in days from day of diagnosis, the plus sign "+" indicates a censored onservation.

The value of the test statistic is

$$\max_{1\le k\le 195} \frac{|\sum_{i=1}^{k} U_i|}{\sqrt{\sum_{i=1}^{195} U_i^2}} = 0.779.$$

Our test based on (2.19) gives P-value 0.578 which gives support to H_0.

Furthermore, we made a scatterplot of $\{\sum_{i=1}^{k} U_i, k = 1, \cdots, 195\}$ in Figure 4.2. The graph suggests a test based on (2.24) for the epidemic alternative H_1' which is described in (2.20), but the result is not significant as the P-value is 0.23. Hence, Kalbfleisch and Prentice's use of the proportional hazards model with time independent covariates is justified.

We should point out that, both the Stanford heart transplant data and the Radiation Therapy Oncology Group data have been used many times in the literature; the conclusions from above applications are consistent with some analysis using the Cox proportional hazards model.

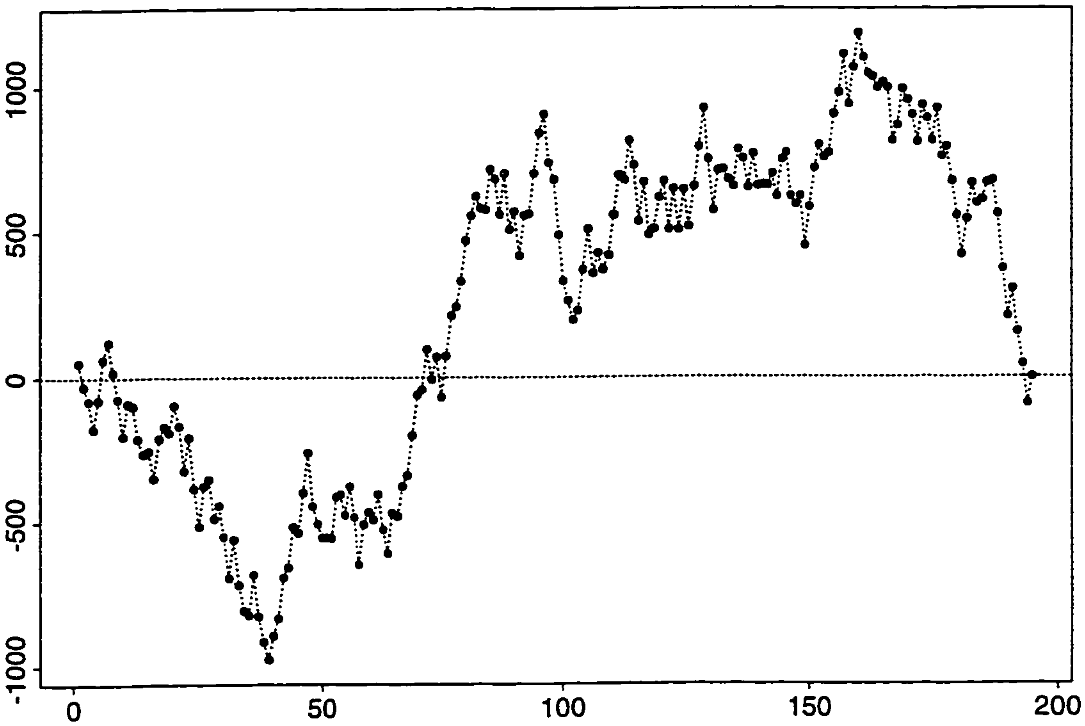

Figure 4.2. *Plot of* $\sum_{i=1}^{k} U_i, k = 1, \cdots, 195$ *for the Radiation Therapy Oncology Group data.*

Chapter 5

Further Discussion of U-Statistics Based Processes

5.1 Introduction

We notice that all the previous discussion is based on the Wilcoxon-Gehan-Mantel score function h given in (2.8). We also discussed the tests for change based on U-statistics with h as a special anti-symmetric kernel both under the null hypothesis of no change and under the alternative hypothesis of one change.

U-statistics are often used to test for change in the distribution of a sequence of random variables. In this chapter, we generalize the U-statistic based test to the general symmetric kernel case and the anti-symmetric kernel case separately. We call h symmetric if

$$h(x, y) = h(y, x), \tag{5.1}$$

and anti-symmetric if

$$h(x, y) = -h(y, x). \tag{5.2}$$

Now, let $X_1, \cdots, X_n$ be a sequence of independent real-valued observations. In this chapter, we would like to test the no-change null hypothesis

$$H_0 : X_1, \cdots, X_n \text{ i.i.d.} F(\cdot)$$

against the one-change alternative hypothesis

H_1: there exists some $\tau \in \{1, 2, \cdots, n-1\}$ such that

$$X_1, \cdots, X_\tau \text{ i.i.d. } F^{(1)}(\cdot),$$

$$X_{\tau+1}, \cdots, X_n \text{ i.i.d. } F^{(2)}(\cdot), \text{ and}$$

$$F^{(1)}(x) \neq F^{(2)}(x) \text{ for some } x,$$

where distribution functions $F, F^{(1)}, F^{(2)}$ and change-point τ are unknown. We assume that

$$\tau = [n\lambda] \text{ for some } \lambda,\ 0 < \lambda < 1. \tag{5.3}$$

The U-statistic with kernel $h(x, y)$ is

$$Z_k = \sum_{i=1}^{k} \sum_{j=k+1}^{n} h(X_i, X_j),\ 1 \leq k < n. \tag{5.4}$$

We assume, that kernel $h(x, y)$ is of bounded variation as a function of x or of y with the other variable fixed at any value.

Also, we assume that

$$Eh^2(X_1, X_2) < \infty,\ Eh^2(X_{n-1}, X_n) < \infty, \quad \text{and} \quad Eh^2(X_1, X_n) < \infty. \tag{5.5}$$

Let

$$\theta_1 = Eh(X_1, X_2), \tag{5.6}$$

$$\theta_2 = Eh(X_{n-1}, X_n), \tag{5.7}$$

and

$$\theta_{12} = Eh(X_1, X_n). \tag{5.8}$$

Define

$$\tilde{h}_1(t) = Eh(X_1, t) - \theta_1 \tag{5.9}$$

and

$$\tilde{h}_2(t) = Eh(X_n, t) - \theta_2. \tag{5.10}$$

Condition (5.5) implies that $E\tilde{h}_1^2(X_1) < \infty$ and $E\tilde{h}_2^2(X_n) < \infty$. In addition, we assume that the kernel is non-degenerate, i.e.

$$E\tilde{h}_1^2(X_1) > 0 \quad \text{and} \quad E\tilde{h}_2^2(X_n) > 0, \tag{5.11}$$

and consider the processes

$$U_k = Z_k - k(n-k)\theta_1, \ 1 \le k < n. \tag{5.12}$$

Our test will be based on (5.12). Let an estimator of the change-point τ be

$$\hat{\tau} = \min\{k : U_k = \max_{1 \le m < n} U_m\}. \tag{5.13}$$

The limiting distribution based on $\max_{1 \le k < n} U_k$ under H_0 is known. We refer to Csörgő and Horváth (1997) for the results and for a detailed list of references, however, the asymptotic distribution of the test statistic under the alternative hypothesis has seldom been considered so far. We discuss the limiting distribution of test statistic $\max_{1 \le k < n} U_k$ under the alternative hypothesis H_1 based on symmetric and anti-symmetric kernels separately.

Let $\mathcal{F}_{n,i}$ be the sigma-field generated by $\{X_1, \cdots, X_i, X_{\tau+1}, \cdots, X_n\}, i = 1, \cdots, \tau$, and $\mathcal{F}_{n,0}$ be the sigma-field generated by $\{X_{\tau+1}, \cdots, X_n\}$. Define

$$\mu_n = E\Big\{\sum_{j=\tau+1}^{n} h(X_1, X_j)|\mathcal{F}_{n,0}\Big\} \tag{5.14}$$

and

$$\sigma_n^2 = Var\Big\{\sum_{j=\tau+1}^{n} h(X_1, X_j)|\mathcal{F}_{n,0}\Big\}. \tag{5.15}$$

Denote the empirical distribution function of $X_{\tau+1}, \cdots, X_n$ by

$$F_{n-\tau}^{(2)}(t) = \frac{1}{n-\tau} \sum_{j=\tau+1}^{n} I(X_j \le t). \tag{5.16}$$

We have

$$\begin{aligned}
\mu_n &= E\Big\{ \sum_{j=\tau+1}^{n} h(X_1, X_j) | \mathcal{F}_{n,0} \Big\} \\
&= \sum_{j=\tau+1}^{n} E\{h(X_1, X_j) | \mathcal{F}_{n,0}\} \\
&= \sum_{j=\tau+1}^{n} E\{h(X_1, X_j) | X_j\} \\
&= \sum_{j=\tau+1}^{n} \tilde{h}_1(X_j) + (n-\tau)\theta_1
\end{aligned} \tag{5.17}$$

and

$$\sum_{j=\tau+1}^{n} \tilde{h}_1(X_j) = (n-\tau) \int \tilde{h}_1(t) dF_{n-\tau}^{(2)}(t). \tag{5.18}$$

Thus,

$$\mu_n = (n-\tau)\Big(\int \tilde{h}_1(t) dF_{n-\tau}^{(2)}(t) + \theta_1 \Big). \tag{5.19}$$

Further, we have

$$\begin{aligned}
\sigma_n^2 &= Var\Big\{ \sum_{j=\tau+1}^{n} h(X_1, X_j) | \mathcal{F}_{n,0} \Big\} \\
&= Var\Big\{ (n-\tau) \int h(X_1, t) dF_{n-\tau}^{(2)}(t) | \mathcal{F}_{n,0} \Big\} \\
&= (n-\tau)^2 Var\Big\{ \int h(X_1, t) dF_{n-\tau}^{(2)}(t) | \mathcal{F}_{n,0} \Big\}.
\end{aligned} \tag{5.20}$$

We get our first lemma:

Lemma 5.1. *We assume that h is symmetric or anti-symmetric and that, under H_1,* (5.5) *and* (5.11) *hold. Then*

$$\frac{1}{\sqrt{\tau}\sigma}(Z_\tau - \tau\mu_n) \xrightarrow{\mathcal{D}} N(0,1), \tag{5.21}$$

where τ, Z_τ and μ_n are defined in (5.3), (5.4) *and* (5.14), *respectively, and*

$$\sigma^2 = \sigma^2(n) = (n-\tau)^2\Big\{\int\Big[\int h(s,t)dF^{(2)}(t)\Big]^2 dF^{(1)}(s) - \theta_{12}^2\Big\}. \qquad (5.22)$$

Proof. We are going to use Theorem 2.4.1 of Sen (1981) for the proof. Let

$$\xi_{ni} = \frac{1}{\sqrt{\tau}\sigma_n}\Big\{\sum_{j=\tau+1}^{n} h(X_i, X_j) - \mu_n\Big\},\ i = 1, \cdots, \tau,$$

where σ_n^2 is defined in (5.15). By (5.17), we note that under H_1, μ_n is a partial sum of i.i.d. random variables. We have

$$\begin{aligned} E(\xi_{ni}|\mathcal{F}_{n,i-1}) &= \frac{1}{\sqrt{\tau}\sigma_n}\Big(E\Big\{\sum_{j=\tau+1}^{n} h(X_i,X_j)|\mathcal{F}_{n,i-1}\Big\} - \mu_n\Big) \\ &= \frac{1}{\sqrt{\tau}\sigma_n}(\mu_n - \mu_n) \\ &= 0,\ i = 1, \cdots, \tau. \end{aligned}$$

Thus,

$$\sum_{i=1}^{\tau} E(\xi_{ni}|\mathcal{F}_{n,i-1}) = 0. \qquad (5.23)$$

Further, we have

$$\begin{aligned} Var(\xi_{ni}|\mathcal{F}_{n,i-1}) &= \frac{1}{\tau\sigma_n^2} Var\Big\{\Big[\sum_{j=\tau+1}^{n} h(X_i,X_j) - \mu_n\Big]|\mathcal{F}_{n,i-1}\Big\} \\ &= \frac{1}{\tau\sigma_n^2} Var\Big\{\sum_{j=\tau+1}^{n} h(X_i,X_j)|\mathcal{F}_{n,i-1}\Big\} \\ &= \frac{1}{\tau\sigma_n^2}\cdot\sigma_n^2 \\ &= \frac{1}{\tau} > 0,\ i = 1, \cdots, \tau, \end{aligned}$$

and so

$$\sum_{i=1}^{\tau} Var(\xi_{ni}|\mathcal{F}_{n,i-1}) = 1. \qquad (5.24)$$

On the other hand, since $\xi_{n1}, \cdots, \xi_{n\tau}$ are conditionally independent and identically distributed with mean $= 0$ and variance > 0 given $\mathcal{F}_{n,0}$, the conditional Lindeberg condition holds, that is, for every $\varepsilon > 0$,

$$\sum_{i=1}^{\tau} E\Big\{\xi_{ni}^2 I(|\xi_{ni}| > \varepsilon)|\mathcal{F}_{n,i-1}\Big\} \xrightarrow{P} 0 \tag{5.25}$$

as $n \to \infty$ (hence $\tau = [n\lambda] \to \infty$). Hence, by (5.23)-(5.25), we can apply Theorem 2.4.1 of Sen (1981) and get that the central limit theorem holds for $\sum_{i=1}^{\tau} \xi_{ni}$, that is

$$\sum_{i=1}^{\tau} \xi_{ni} \xrightarrow{\mathcal{D}} N(0,1),$$

where

$$\begin{aligned}\sum_{i=1}^{\tau} \xi_{ni} &= \sum_{i=1}^{\tau} \Big\{\frac{1}{\sqrt{\tau}\sigma_n}\Big(\sum_{j=\tau+1}^{n} h(X_i, X_j) - \mu_n\Big)\Big\} \\ &= \frac{1}{\sqrt{\tau}\sigma_n}\Big\{\sum_{i=1}^{\tau}\sum_{j=\tau+1}^{n} h(X_i, X_j) - \tau\mu_n\Big\} \\ &= \frac{1}{\sqrt{\tau}\sigma_n}(Z_\tau - \tau\mu_n).\end{aligned}$$

Thus we get

$$\frac{1}{\sqrt{\tau}\sigma_n}(Z_\tau - \tau\mu_n) \xrightarrow{\mathcal{D}} N(0,1). \tag{5.26}$$

On the other hand, by (5.22), since

$$\begin{aligned}\sigma^2 &= (n-\tau)^2 Var\Big\{E[h(X_1, X_n)|X_1]\Big\} \\ &= (n-\tau)^2 Var\Big\{\int h(X_1, t)dF^{(2)}(t)\Big\}\end{aligned} \tag{5.27}$$

and $Eh^2(X_1, X_n) < \infty$, comparing (5.20) with (5.27) we get

$$\frac{\sigma^2}{\sigma_n^2} \xrightarrow{P} 1. \tag{5.28}$$

Thus (5.26) and (5.28) give (5.21). □

By Lemma 5.1 we have

Theorem 5.1. *We assume that h is symmetric or anti-symmetric and that, under* H_1, (5.5) *and* (5.11) *hold. Then*

$$\frac{Z_\tau - \tau(n-\tau)\theta_{12}}{n^{3/2}\sqrt{\lambda(1-\lambda)^2 D_1^2 + \lambda^2(1-\lambda)D_2^2}} \xrightarrow{\mathcal{D}} N(0,1), \tag{5.29}$$

where τ, Z_τ *and* θ_{12} *are defined in* (5.3), (5.4) *and* (5.8), *respectively, and*

$$D_1^2 = \int \Big[\int h(s,t)dF^{(2)}(t)\Big]^2 dF^{(1)}(s) - \theta_{12}^2, \tag{5.30}$$

$$D_2^2 = \int \Big[\int h(s,t)dF^{(1)}(s)\Big]^2 dF^{(2)}(t) - \theta_{12}^2. \tag{5.31}$$

Proof. Let

$$A_n = Z_\tau - \tau(n-\tau)\theta_{12}. \tag{5.32}$$

We have

$$\begin{aligned} A_n &= \sum_{i=1}^{\tau}\sum_{j=\tau+1}^{n}[h(X_i,X_j)-\theta_{12}] \\ &= \Big\{\sum_{i=1}^{\tau}\sum_{j=\tau+1}^{n}[h(X_i,X_j)-\theta_{12}] - \tau\sum_{j=\tau+1}^{n}[\tilde{h}_1(X_j)+(\theta_1-\theta_{12})]\Big\} \\ &\quad + \tau\sum_{j=\tau+1}^{n}[\tilde{h}_1(X_j)+(\theta_1-\theta_{12})] \\ &= (A_n - B_n) + B_n, \end{aligned}$$

where

$$B_n = \tau\sum_{j=\tau+1}^{n}[\tilde{h}_1(X_j)+(\theta_1-\theta_{12})]. \tag{5.33}$$

Note that by (5.32) and (5.33)

$$\begin{aligned} E(A_n|\mathcal{F}_{n,0}) &= \sum_{i=1}^{\tau}\sum_{j=\tau+1}^{n}\Big\{E[h(X_i,X_j)|\mathcal{F}_{n,0}] - \theta_{12}\Big\} \\ &= \sum_{i=1}^{\tau}\sum_{j=\tau+1}^{n}\Big\{[\tilde{h}_1(X_j)+\theta_1] - \theta_{12}\Big\} \\ &= \tau\sum_{j=\tau+1}^{n}[\tilde{h}_1(X_j)+(\theta_1-\theta_{12})] \\ &= B_n, \end{aligned}$$

that is

$$B_n = E(A_n|\mathcal{F}_{n,0}). \tag{5.34}$$

Now, by (5.17), (5.32) and (5.33),

$$\begin{aligned} A_n - B_n &= Z_\tau - \tau(n-\tau)\theta_{12} - \tau \sum_{j=\tau+1}^{n} [\tilde{h}_1(X_j) + (\theta_1 - \theta_{12})] \\ &= Z_\tau - \tau \sum_{j=\tau+1}^{n} [\tilde{h}_1(X_j) + \theta_1] \\ &= Z_\tau - \tau\mu_n. \end{aligned} \tag{5.35}$$

Thus, by Lemma 5.1, we have

$$\frac{A_n - B_n}{\sqrt{\tau}\sigma} \xrightarrow{\mathcal{D}} N(0,1). \tag{5.36}$$

By (5.22) and (5.30), since

$$\sigma^2 = (n-\tau)^2 D_1^2, \tag{5.37}$$

we get, by (5.36) and (5.37),

$$\frac{A_n - B_n}{\sqrt{\tau}(n-\tau)D_1} \xrightarrow{\mathcal{D}} N(0,1). \tag{5.38}$$

In addition, by (5.33), as B_n is a partial sum of i.i.d. random variables, by the central limit theorem, we have

$$\frac{B_n - E(B_n)}{\sqrt{Var(B_n)}} \xrightarrow{\mathcal{D}} N(0,1). \tag{5.39}$$

Now, by (5.8), (5.32) and (5.34),

$$\begin{aligned} E(B_n) &= E[E(A_n|\mathcal{F}_{n,0})] \\ &= E(A_n) \\ &= \sum_{i=1}^{\tau} \sum_{j=\tau+1}^{n} [Eh(X_i, X_j) - \theta_{12}] \\ &= \sum_{i=1}^{\tau} \sum_{j=\tau+1}^{n} (\theta_{12} - \theta_{12}) \\ &= 0 \end{aligned} \tag{5.40}$$

and, by (5.8), (5.9), (5.31) and (5.33),

$$\begin{aligned} Var(B_n) &= Var\Big\{\tau \sum_{i=\tau+1}^{n} [\tilde{h}_1(X_j) + (\theta_1 - \theta_{12})]\Big\} \\ &= \tau^2 \sum_{i=\tau+1}^{n} Var[\tilde{h}_1(X_j)] \\ &= \tau^2(n-\tau)Var[\tilde{h}_1(X_n)] \\ &= \tau^2(n-\tau)Var[E(h(X_1,X_n)|X_n)] \\ &= \tau^2(n-\tau)\Big\{E[E(h(X_1,X_n)|X_n)]^2 - \big[E[E(h(X_1,X_n)|X_n)]\big]^2\Big\} \\ &= \tau^2(n-\tau)\Big\{\int\Big[\int h(s,t)dF^{(1)}(s)\Big]^2 dF^{(2)}(t) - \theta_{12}^2\Big\} \\ &= \tau^2(n-\tau)D_2^2. \end{aligned} \tag{5.41}$$

Thus, by (5.39)-(5.41), we get

$$\frac{B_n}{\tau\sqrt{n-\tau}D_2} \xrightarrow{\mathcal{D}} N(0,1). \tag{5.42}$$

On the other hand, by (5.34),

$$\begin{aligned} &E[(A_n - B_n)B_n|\mathcal{F}_{n,0}] \\ &= E(A_nB_n|\mathcal{F}_{n,0}) - E(B_n^2|\mathcal{F}_{n,0}) \\ &= B_nE(A_n|\mathcal{F}_{n,0}) - B_n^2 \\ &= 0, \end{aligned}$$

and so

$$E[(A_n - B_n)B_n] = 0, \tag{5.43}$$

i.e. $A_n - B_n$ and B_n are uncorrelated. Thus, using (5.38), (5.42) and (5.43), combined with the Skorokhod representation theorem, the property of multivariate normality and the Cramér-Wold device (cf. Billingsley (1968) and Serfling (1980)), we can get that the joint distribution of $A_n - B_n$ and B_n is asymptotically bivariate normal with

$$\left(\frac{A_n - B_n}{\sqrt{\tau}(n-\tau)D_1}, \frac{B_n}{\tau\sqrt{n-\tau}D_2}\right) \xrightarrow{\mathcal{D}} N(0,\Sigma), \tag{5.44}$$

where

$$\Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

is the convariance matrix. Thus,

$$\frac{(A_n - B_n) + B_n}{\sqrt{\tau(n-\tau)^2 D_1^2 + \tau^2(n-\tau)D_2^2}} \xrightarrow{\mathcal{D}} N(0,1),$$

i.e.

$$\frac{A_n}{\sqrt{\tau(n-\tau)^2 D_1^2 + \tau^2(n-\tau)D_2^2}} \xrightarrow{\mathcal{D}} N(0,1). \tag{5.45}$$

Therefore, (5.29) follows by (5.32) and (5.45), combined with $\tau = [n\lambda]$ and

$$\lim_{n\to\infty} \frac{[n\lambda]}{n} = \lambda.$$

□

In the following sections we'll discuss the asymptotics of the test statistic $\max_{1\leq k<n} U_k$ under the alternative hypothesis of one change H_1 based on symmetric and anti-symmetric kernels separately.

5.2 Asymptotics for Symmetric Kernel

We have

Lemma 5.2. *We assume that h is symmetric and that, under H_1,* (5.5) *and* (5.11) *hold, $\theta_{12} > \theta_1$, and for some $\varepsilon > 0$,*

$$\lambda(\theta_{12} - \theta_1) - t(\theta_2 - \theta_1) > \varepsilon,\ 0 < t < 1 - \lambda. \tag{5.46}$$

Then

$$|\hat{\tau} - \tau| = O_P(1), \tag{5.47}$$

where θ_1, θ_2, θ_{12} *and* $\hat{\tau}$ *are defined in* (5.6), (5.7), (5.8) *and* (5.13), *respectively.*

Proof. We only need show that

$$\lim_{k\to\infty} \limsup_{n\to\infty} P\Big\{ \max_{1\le m\le \tau-k} U_m \ge \max_{\tau-k<m<\tau+k} U_m \Big\} = 0 \tag{5.48}$$

and

$$\lim_{k\to\infty} \limsup_{n\to\infty} P\Big\{ \max_{\tau+k\le m<n} U_m \ge \max_{\tau-k<m<\tau+k} U_m \Big\} = 0. \tag{5.49}$$

For notational simplicity, let

$$h_1(x,y) = h(x,y) - \theta_1. \tag{5.50}$$

Thus,

$$\begin{aligned} U_k &= \sum_{i=1}^{k} \sum_{j=k+1}^{n} h(X_i, X_j) - k(n-k)\theta_1 \\ &= \sum_{i=1}^{k} \sum_{j=k+1}^{n} h_1(X_i, X_j) \end{aligned} \tag{5.51}$$

and

$$Eh_1(X_1, X_n) = \theta_{12} - \theta_1 > 0. \tag{5.52}$$

We prove (5.48) first. We have

$$\begin{aligned} &P\Big\{ \max_{1\le m\le \tau-k} U_m \ge \max_{\tau-k<m<\tau+k} U_m \Big\} \\ &\le P\Big\{ \max_{1\le m\le \tau-k} U_m \ge U_\tau \Big\} \\ &= P\{\exists l, 1 \le l \le \tau - k : U_l \ge U_\tau\} \\ &= P\Big\{\exists l, 1 \le l \le \tau - k : \sum_{i=1}^{l} \sum_{j=l+1}^{n} h_1(X_i, X_j) \ge \sum_{i=1}^{\tau} \sum_{j=\tau+1}^{n} h_1(X_i, X_j) \Big\} \end{aligned}$$

$$= P\Big\{\exists l, 1 \le l \le \tau - k : \sum_{i=1}^{l}[\sum_{j=l+1}^{\tau} h_1(X_i, X_j) + \sum_{j=\tau+1}^{n} h_1(X_i, X_j)] \ge$$

$$\sum_{i=1}^{l}\sum_{j=\tau+1}^{n} h_1(X_i, X_j) + \sum_{i=l+1}^{\tau}\sum_{j=\tau+1}^{n} h_1(X_i, X_j)\Big\}$$

$$= P\Big\{\exists l, 1 \le l \le \tau - k : \sum_{i=1}^{l}\sum_{j=l+1}^{\tau} h_1(X_i, X_j) \ge \sum_{i=l+1}^{\tau}\sum_{j=\tau+1}^{n} h_1(X_i, X_j)\Big\}$$

$$= 1 - P\Big\{\min_{1\le m\le \tau-k}\Big[-\sum_{i=1}^{m}\sum_{j=m+1}^{\tau} h_1(X_i, X_j) + \sum_{i=m+1}^{\tau}\sum_{j=\tau+1}^{n} h_1(X_i, X_j)\Big] > 0\Big\}$$

$$= P\Big\{\min_{1\le m\le \tau-k}\Big[-\sum_{i=1}^{m}\sum_{j=m+1}^{\tau} h_1(X_i, X_j) + \sum_{i=m+1}^{\tau}\sum_{j=\tau+1}^{n} h_1(X_i, X_j)\Big] \le 0\Big\}$$

$$= P\Big\{\min_{1\le m\le \tau-k}\Big[-\frac{1}{n}\sum_{i=m+1}^{\tau}\sum_{j=1}^{m} h_1(X_i, X_j) + \frac{1}{n}\sum_{i=m+1}^{\tau}\sum_{j=\tau+1}^{n} h_1(X_i, X_j)\Big] \le 0\Big\}, \tag{5.53}$$

where the last line was obtained by the symmetry of $h_1(x, y)$.

Denote the empirical distribution function of $X_1, \cdots, X_m$ by

$$F_m^{(1)}(t) = \frac{1}{m}\sum_{j=1}^{m} I(X_j \le t),\ m \le \tau. \tag{5.54}$$

Thus,

$$\begin{aligned} &-\frac{1}{n}\sum_{i=m+1}^{\tau}\sum_{j=1}^{m} h_1(X_i, X_j) \\ &= -\frac{m}{n}\sum_{i=m+1}^{\tau}\int h_1(X_i, t)dF_m^{(1)}(t) \\ &= -\frac{m}{n}\sum_{i=m+1}^{\tau}\int h_1(X_i, t)dF^{(1)}(t) \\ &\quad -\frac{\sqrt{m}}{n}\sum_{i=m+1}^{\tau}\int h_1(X_i, t)d\Big[\sqrt{m}\Big(F_m^{(1)}(t) - F^{(1)}(t)\Big)\Big], \end{aligned} \tag{5.55}$$

and

$$\frac{1}{n}\sum_{i=m+1}^{\tau}\sum_{j=\tau+1}^{n} h_1(X_i, X_j) = \frac{1}{n}\sum_{i=m+1}^{\tau}(n-\tau)\int h_1(X_i, t)dF_{n-\tau}^{(2)}(t). \tag{5.56}$$

Using the uniform approximation of an empirical process by a Brownian bridge (Skorohod-Dudley version of Donsker's theorem, cf. Theorem 4.2.1 of Csörgő and Révész (1981)), we have

$$\sqrt{m}\Big(F_m^{(1)}(t) - F^{(1)}(t)\Big) \stackrel{\mathcal{D}}{=} B(F^{(1)}(t)) + o_P(1),$$

where $\{B(u), 0 \le u \le 1\}$ is a Brownian bridge. So by (5.55),

$$\begin{aligned}
&-\frac{1}{n}\sum_{i=m+1}^{\tau}\sum_{j=1}^{m} h_1(X_i, X_j) \\
&\stackrel{\mathcal{D}}{=} -\frac{m}{n}\sum_{i=m+1}^{\tau}\int h_1(X_i,t)dF^{(1)}(t) - \frac{\sqrt{m}}{n}\sum_{i=m+1}^{\tau}\int h_1(X_i,t)d\Big[B(F^{(1)}(t)) + o_P(1)\Big].
\end{aligned}$$

Note that

$$\begin{aligned}
E\int h_1(X_i,t)dF^{(1)}(t) &= E\{E[h_1(X_1,X_2)|X_1]\} \\
&= Eh_1(X_1,X_2) \\
&= Eh(X_1,X_2) - \theta_1 \\
&= 0, \ i \le \tau
\end{aligned}$$

and $Eh^2(X_1,X_2) < \infty$. Thus by the law of iterated logarithm

$$\begin{aligned}
&-\frac{1}{n}\sum_{i=m+1}^{\tau}\sum_{j=1}^{m} h_1(X_i, X_j) \\
&= O_P\Big(\sqrt{(\tau-m)\log\log(\tau-m)}\Big) + O_P\left(\sqrt{\frac{(\tau-m)\log\log(\tau-m)}{n}}\right) \qquad (5.57) \\
&= O_P\Big(\sqrt{(\tau-m)\log\log(\tau-m)}\Big).
\end{aligned}$$

On the other hand, by (5.56),

$$\begin{aligned}
&\frac{1}{n}\sum_{i=m+1}^{\tau}\sum_{j=\tau+1}^{n} h_1(X_i,X_j) \\
&\quad = \frac{n-\tau}{n}\sum_{i=m+1}^{\tau}\int h_1(X_i,t)dF_{n-\tau}^{(2)}(t)
\end{aligned}$$

$$
\begin{aligned}
&= \frac{n-\tau}{n} \sum_{i=m+1}^{\tau} \int h_1(X_i,t)dF^{(2)}(t) \\
&\quad + \frac{\sqrt{n-\tau}}{n} \sum_{i=m+1}^{\tau} \int h_1(X_i,t)d\Big[\sqrt{n-\tau}\Big(F^{(2)}_{n-\tau}(t) - F^{(2)}(t)\Big)\Big] \\
&= \frac{n-\tau}{n} \sum_{i=m+1}^{\tau} \Big\{\Big(\int h(X_i,t)dF^{(2)}(t) - \theta_{12}\Big) + (\theta_{12} - \theta_1)\Big\} \\
&\quad + \frac{\sqrt{n-\tau}}{n} \sum_{i=m+1}^{\tau} \int h_1(X_i,t)d\Big[\sqrt{n-\tau}\Big(F^{(2)}_{n-\tau}(t) - F^{(2)}(t)\Big)\Big] \\
&= \frac{n-\tau}{n} \sum_{i=m+1}^{\tau} \Big(\int h(X_i,t)dF^{(2)}(t) - \theta_{12}\Big) + \frac{n-\tau}{n}(\tau - m)(\theta_{12} - \theta_1) \\
&\quad + \frac{\sqrt{n-\tau}}{n} \sum_{i=m+1}^{\tau} \int h_1(X_i,t)d\Big[\sqrt{n-\tau}\Big(F^{(2)}_{n-\tau}(t) - F^{(2)}(t)\Big)\Big].
\end{aligned}
\tag{5.58}
$$

Again, using the uniform approximation of an empirical process by a Brownian bridge, we have

$$
\sqrt{n-\tau}\Big(F^{(2)}_{n-\tau}(t) - F^{(2)}(t)\Big) \stackrel{\mathcal{D}}{=} B(F^{(2)}(t)) + o_P(1),
$$

where $\{B(u), 0 \le u \le 1\}$ is a Brownian bridge. Thus, by (5.58),

$$
\begin{aligned}
&\frac{1}{n} \sum_{i=m+1}^{\tau} \sum_{j=\tau+1}^{n} h_1(X_i, X_j) \\
&\quad \stackrel{\mathcal{D}}{=} \Big(1 - \frac{[n\lambda]}{n}\Big) \sum_{i=m+1}^{\tau} \Big(\int h(X_i,t)dF^{(2)}(t) - \theta_{12}\Big) + (1 - \frac{[n\lambda]}{n})(\theta_{12} - \theta_1)(\tau - m) \\
&\qquad + \frac{\sqrt{n-\tau}}{n} \sum_{i=m+1}^{\tau} \int h_1(X_i,t)d\Big(B(F^{(2)}(t)) + o_P(1)\Big).
\end{aligned}
$$

Note that

$$
\begin{aligned}
&E\Big\{\int h(X_i,t)dF^{(2)}(t) - \theta_{12}\Big\} \\
&= E\{E[h(X_1, X_n)|X_1]\} - \theta_{12} \\
&= Eh(X_1, X_n) - \theta_{12} \\
&= 0, \ i \le \tau.
\end{aligned}
$$

Thus by the law of iterated logarithm

$$\begin{aligned}
&\frac{1}{n}\sum_{i=m+1}^{\tau}\sum_{j=\tau+1}^{n} h_1(X_i,X_j)\\
&= O_P\Big(\sqrt{(\tau-m)\log\log(\tau-m)}\Big) + O_P\left(\sqrt{\frac{(\tau-m)\log\log(\tau-m)}{n}}\right) \qquad (5.59)\\
&\quad + (1-\lambda)(\theta_{12}-\theta_1)(\tau-m)\\
&= O_P\Big(\sqrt{(\tau-m)\log\log(\tau-m)}\Big) + (1-\lambda)(\theta_{12}-\theta_1)(\tau-m).
\end{aligned}$$

Thus, by (5.57) and (5.59),

$$\begin{aligned}
&-\frac{1}{n}\sum_{i=m+1}^{\tau}\sum_{j=1}^{m} h_1(X_i,X_j) + \frac{1}{n}\sum_{i=m+1}^{\tau}\sum_{j=\tau+1}^{n} h_1(X_i,X_j)\\
&= O_P\Big(\sqrt{(\tau-m)\log\log(\tau-m)}\Big) + (1-\lambda)(\theta_{12}-\theta_1)(\tau-m) \qquad (5.60)\\
&= \sqrt{(\tau-m)\log\log(\tau-m)}\left(O_P(1) + \frac{(1-\lambda)(\theta_{12}-\theta_1)(\tau-m)}{\sqrt{(\tau-m)\log\log(\tau-m)}}\right).
\end{aligned}$$

Since $1-\lambda>0$, $\theta_{12}-\theta_1>0$, and $m\le\tau-k$, i.e. $\tau-m\ge k$, we can say

$$(1-\lambda)(\theta_{12}-\theta_1)(\tau-m)\to+\infty \quad \text{as} \quad k\to+\infty,$$

and, furthermore,

$$\frac{(1-\lambda)(\theta_{12}-\theta_1)(\tau-m)}{\sqrt{(\tau-m)\log\log(\tau-m)}}\to+\infty \quad \text{as} \quad k\to+\infty.$$

Thus, by (5.60),

$$-\frac{1}{n}\sum_{i=m+1}^{\tau}\sum_{j=1}^{m} h_1(X_i,X_j) + \frac{1}{n}\sum_{i=m+1}^{\tau}\sum_{j=\tau+1}^{n} h_1(X_i,X_j)\to+\infty \text{ as } k\to+\infty. \quad (5.61)$$

Therefore (5.48) follows by (5.53) and (5.61).

Next, we prove (5.49). Note that the relation (5.49) is not completely symmetrical to (5.48). We consider the opposite direction sequence of observations $X_n,\cdots,X_1$. Let

$$X_i' = X_{n+1-i},\ i=1,\cdots,n,$$

$$m' = n - m,$$

and

$$\tau' = (n+1) - (\tau+1) = n - \tau.$$

τ' is the change-point of the opposite direction sequence. Thus, the alternative hypothesis H_1 is equivalent to

H_1': there exists some $\tau' \in \{1, 2, \cdots, n-1\}$ such that

$X_1', \cdots, X_{\tau}'$ i.i.d. $F^{(2)}(\cdot)$,

$X_{\tau+1}', \cdots, X_n'$ i.i.d. $F^{(1)}(\cdot)$, and

$F^{(2)}(x) \neq F^{(1)}(x)$ for some x.

Also,

$$U_m = \sum_{i=1}^{m} \sum_{j=m+1}^{n} \Big(h(X_i, X_j) - \theta_1\Big), \tau + k \leq m < n$$

is the same as

$$U'_{m'} = \sum_{i=1}^{m'} \sum_{j=m'+1}^{n} \Big(h(X_i', X_j') - \theta_1\Big), 1 \leq m' \leq \tau' - k,$$

and

$$U_m = \sum_{i=1}^{m} \sum_{j=m+1}^{n} \Big(h(X_i, X_j) - \theta_1\Big), \tau - k < m < \tau + k$$

is the same as

$$U'_{m'} = \sum_{i=1}^{m'} \sum_{j=m'+1}^{n} \Big(h(X_i', X_j') - \theta_1\Big), \tau' - k < m' < \tau' + k$$

due to the symmetry of h. Note that the only difference between (5.48) and (5.49) is that

$$Eh_1(X_1', X_2') = Eh(X_1', X_2') - \theta_1 = Eh(X_n, X_{n-1}) - \theta_1 = \theta_2 - \theta_1 \qquad (5.62)$$

instead of zero. Thus, we have

$$
\begin{aligned}
&P\Big\{\max_{\tau+k\le m<n} U_m \ge \max_{\tau-k<m<\tau+k} U_m\Big\} \\
&\quad = P\Big\{\max_{1\le m'\le \tau'-k} U'_{m'} \ge \max_{\tau'-k<m'<\tau'+k} U'_{m'}\Big\} \\
&\quad \le P\Big\{\min_{1\le m'\le \tau'-k}\Big[-\frac{1}{n}\sum_{i=m'+1}^{\tau'}\sum_{j=1}^{m'} h_1(X'_i, X'_j) + \frac{1}{n}\sum_{i=m'+1}^{\tau'}\sum_{j=\tau'+1}^{n} h_1(X'_i, X'_j)\Big] \le 0\Big\}.
\end{aligned}
\tag{5.63}
$$

Denote the empirical distribution function of $X'_1, \cdots, X'_{m'}$ by

$$
F^{(2)}_{m'}(t) = \frac{1}{m'}\sum_{j=1}^{m'} I(X'_j \le t),\ m' \le \tau', \tag{5.64}
$$

and the empirical distribution function of $X'_{\tau'+1}, \cdots, X'_n$ by

$$
F^{(1)}_{n-\tau'}(t) = \frac{1}{n-\tau'}\sum_{j=\tau'+1}^{n} I(X'_j \le t). \tag{5.65}
$$

Using the law of iterated logarithm and the uniform approximation of an empirical process by a Brownian bridge, we get

$$
\begin{aligned}
&-\frac{1}{n}\sum_{i=m'+1}^{\tau'}\sum_{j=1}^{m'} h_1(X'_i, X'_j) \\
&= -\frac{m'}{n}\sum_{i=m'+1}^{\tau'}\int h_1(X'_i, t)\,dF^{(2)}_{m'}(t) \\
&= -\frac{m'}{n}\sum_{i=m'+1}^{\tau'}\int h_1(X'_i, t)\,dF^{(2)}(t) \\
&\quad - \frac{\sqrt{m'}}{n}\sum_{i=m'+1}^{\tau'}\int h_1(X'_i, t)\,d\Big[\sqrt{m'}\Big(F^{(2)}_{m'}(t) - F^{(2)}(t)\Big)\Big] \\
&= -\frac{m'}{n}\sum_{i=m'+1}^{\tau'}\Big\{\Big(\int h_1(X'_i, t)\,dF^{(2)}(t) - (\theta_2-\theta_1)\Big) + (\theta_2-\theta_1)\Big\} \\
&\quad - \frac{\sqrt{m'}}{n}\sum_{i=m'+1}^{\tau'}\int h_1(X'_i, t)\,d\Big[\sqrt{m'}\Big(F^{(2)}_{m'}(t) - F^{(2)}(t)\Big)\Big]
\end{aligned}
$$

$$
\begin{aligned}
&= -\frac{m'}{n} \sum_{i=m'+1}^{\tau'} \Big(\int h_1(X_i', t) dF^{(2)}(t) - (\theta_2 - \theta_1) \Big) \\
&\quad - \frac{\sqrt{m'}}{n} \sum_{i=m'+1}^{\tau'} \int h_1(X_i', t) d\Big[\sqrt{m'}\Big(F_{m'}^{(2)}(t) - F^{(2)}(t)\Big)\Big] - \frac{m'}{n}(\theta_2 - \theta_1)(\tau' - m') \\
&= O_P\Big(\sqrt{(\tau' - m')\log\log(\tau' - m')}\Big) + O_P\Bigg(\sqrt{\frac{(\tau' - m')\log\log(\tau' - m')}{n}}\Bigg) \\
&\quad - \frac{m'}{n}(\theta_2 - \theta_1)(\tau' - m') \\
&= O_P\Big(\sqrt{(\tau' - m')\log\log(\tau' - m')}\Big) - \frac{m'}{n}(\theta_2 - \theta_1)(\tau' - m')
\end{aligned} \tag{5.66}
$$

since, by (5.62),

$$
\begin{aligned}
&E\Big(\int h_1(X_i', t) dF^{(2)}(t) - (\theta_2 - \theta_1) \Big) \\
&\quad = E\{E[h_1(X_1', X_2')|X_1']\} - (\theta_2 - \theta_1) \\
&\quad = Eh_1(X_1', X_2') - (\theta_2 - \theta_1) \\
&\quad = 0, \; i \le \tau'.
\end{aligned}
$$

Also,

$$
\begin{aligned}
&\frac{1}{n} \sum_{i=m'+1}^{\tau'} \sum_{j=\tau'+1}^{n} h_1(X_i', X_j') \\
&\quad = \frac{n - \tau'}{n} \sum_{i=m'+1}^{\tau'} \int h_1(X_i', t) dF_{n-\tau'}^{(1)}(t) \\
&\quad = \frac{\tau}{n} \sum_{i=m'+1}^{\tau'} \Big\{ \Big(\int h_1(X_i', t) dF^{(1)}(t) - (\theta_{12} - \theta_1) \Big) + (\theta_{12} - \theta_1) \Big\} \\
&\qquad + \frac{\sqrt{\tau}}{n} \sum_{i=m'+1}^{\tau'} \int h_1(X_i', t) d\Big[\sqrt{n - \tau'}\Big(F_{n-\tau'}^{(1)}(t) - F^{(1)}(t)\Big)\Big] \\
&\quad = \frac{\tau}{n} \sum_{i=m'+1}^{\tau'} \Big(\int h_1(X_i', t) dF^{(1)}(t) - (\theta_{12} - \theta_1) \Big) \\
&\qquad + \frac{\sqrt{\tau}}{n} \sum_{i=m'+1}^{\tau'} \int h_1(X_i', t) d\Big[\sqrt{n - \tau'}\Big(F_{n-\tau'}^{(1)}(t) - F^{(1)}(t)\Big)\Big] + \frac{\tau}{n}(\theta_{12} - \theta_1)(\tau' - m')
\end{aligned}
$$

$$= O_P\Big(\sqrt{(\tau' - m')\log\log(\tau' - m')}\Big) + O_P\Bigg(\sqrt{\frac{(\tau' - m')\log\log(\tau' - m')}{n}}\Bigg)$$
$$+ \lambda(\theta_{12} - \theta_1)(\tau' - m')$$
$$= O_P\Big(\sqrt{(\tau' - m')\log\log(\tau' - m')}\Big) + \lambda(\theta_{12} - \theta_1)(\tau' - m') \tag{5.67}$$

since

$$E\Big(\int h_1(X_i', t)dF^{(1)}(t) - (\theta_{12} - \theta_1)\Big) = 0,\ i \le \tau'.$$

Thus, by (5.66) and (5.67), we have

$$-\frac{1}{n}\sum_{i=m'+1}^{\tau'}\sum_{j=1}^{m'} h_1(X_i', X_j') + \frac{1}{n}\sum_{i=m'+1}^{\tau'}\sum_{j=\tau'+1}^{n} h_1(X_i', X_j')$$
$$= O_P\Big(\sqrt{(\tau' - m')\log\log(\tau' - m')}\Big) + \Big[\lambda(\theta_{12} - \theta_1) - \frac{m'}{n}(\theta_2 - \theta_1)\Big](\tau' - m')$$
$$= \sqrt{(\tau' - m')\log\log(\tau' - m')}\left(O_P(1) + \frac{\Big[\lambda(\theta_{12} - \theta_1) - \frac{m'}{n}(\theta_2 - \theta_1)\Big](\tau' - m')}{\sqrt{(\tau' - m')\log\log(\tau' - m')}}\right). \tag{5.68}$$

Now, let

$$t' = \frac{m'}{n}.$$

Since $1 \le m' \le \tau' - k = n - (\tau + k) = n - ([n\lambda] + k)$, we get

$$0 < \frac{1}{n} \le t' = \frac{m'}{n} \le 1 - \frac{[n\lambda] + k}{n} < 1 - \lambda, \quad \text{and} \quad \tau' - m' \ge k.$$

Thus, by (5.46), we get

$$\lambda(\theta_{12} - \theta_1) - \frac{m'}{n}(\theta_2 - \theta_1) = \lambda(\theta_{12} - \theta_1) - t'(\theta_2 - \theta_1) > \varepsilon > 0.$$

Hence,

$$\Big[\lambda(\theta_{12} - \theta_1) - \frac{m'}{n}(\theta_2 - \theta_1)\Big](\tau' - m') \to +\infty \quad \text{as} \quad k \to +\infty,$$

and also

$$\frac{\left[\lambda(\theta_{12}-\theta_1)-\frac{m'}{n}(\theta_2-\theta_1)\right](\tau'-m')}{\sqrt{(\tau'-m')\log\log(\tau'-m')}} \to +\infty \quad \text{as} \quad k \to +\infty.$$

Thus, by (5.68),

$$-\frac{1}{n}\sum_{i=m'+1}^{\tau'}\sum_{j=1}^{m'} h_1(X_i', X_j') + \frac{1}{n}\sum_{i=m'+1}^{\tau'}\sum_{j=\tau'+1}^{n} h_1(X_i', X_j') \to +\infty \quad \text{as} \quad k \to +\infty. \tag{5.69}$$

Therefore (5.49) follows by (5.63) and (5.69). □

Theorem 5.2. *We assume that h is symmetric and that, under H_1,* (5.5), (5.11) *and* (5.46) *hold, and also $\theta_{12} > \theta_1$. Then*

$$\frac{\max_{1\le k<n} U_k - \tau(n-\tau)(\theta_{12}-\theta_1)}{n^{3/2}\sqrt{\lambda(1-\lambda)^2 D_1^2 + \lambda^2(1-\lambda)D_2^2}} \xrightarrow{\mathcal{D}} N(0,1), \tag{5.70}$$

where θ_1, θ_{12}, U_k, D_1^2 and D_2^2 are defined in (5.6), (5.8), (5.12), (5.30) *and* (5.31), *respectively.*

Proof. We note that, by the definition of $\hat{\tau}$ given in (5.13) and the definition of U_k given in (5.12)

$$\max_{1\le k<n} U_k = U_{\hat{\tau}} = Z_{\hat{\tau}} - \hat{\tau}(n-\hat{\tau})\theta_1. \tag{5.71}$$

Thus,

$$\begin{aligned}
&\max_{1\le k<n} U_k - \tau(n-\tau)(\theta_{12}-\theta_1) \\
&\quad = Z_{\hat{\tau}} - \hat{\tau}(n-\hat{\tau})\theta_1 - \tau(n-\tau)(\theta_{12}-\theta_1) \qquad (5.72)\\
&\quad = [Z_\tau - \tau(n-\tau)\theta_{12}] + [\tau(n-\tau) - \hat{\tau}(n-\hat{\tau})]\theta_1 + (Z_{\hat{\tau}} - Z_\tau).
\end{aligned}$$

Let

$$D^2 = n^3[\lambda(1-\lambda)^2 D_1^2 + \lambda^2(1-\lambda)D_2^2],\ D > 0. \tag{5.73}$$

Thus,

$$D = O(n^{3/2}). \tag{5.74}$$

We have

$$\begin{aligned}
&\frac{\max_{1\le k<n} U_k - \tau(n-\tau)(\theta_{12}-\theta_1)}{n^{3/2}\sqrt{\lambda(1-\lambda)^2 D_1^2 + \lambda^2(1-\lambda)D_2^2}} \\
&\quad = \frac{Z_\tau - \tau(n-\tau)\theta_{12}}{D} + \frac{[\tau(n-\tau) - \hat{\tau}(n-\hat{\tau})]\theta_1}{D} + \frac{Z_{\hat{\tau}} - Z_\tau}{D}.
\end{aligned} \tag{5.75}$$

Since

$$\frac{Z_\tau - \tau(n-\tau)\theta_{12}}{D} \xrightarrow{\mathcal{D}} N(0,1) \tag{5.76}$$

by Theorem 5.1, in order to prove (5.70) we only need show that

$$\frac{[\tau(n-\tau) - \hat{\tau}(n-\hat{\tau})]\theta_1}{D} = o_P(1), \tag{5.77}$$

and

$$\frac{Z_{\hat{\tau}} - Z_\tau}{D} = o_P(1). \tag{5.78}$$

Now,

$$[\tau(n-\tau) - \hat{\tau}(n-\hat{\tau})]\theta_1 = (\tau-\hat{\tau})(n-\tau-\hat{\tau})\theta_1 = O_P(n) \tag{5.79}$$

by Lemma 5.2, and so (5.77) follows by (5.74) and (5.79).

On the other hand,

$$Z_{\hat{\tau}} - Z_\tau = \sum_{i=1}^{\hat{\tau}} \sum_{j=\hat{\tau}+1}^{n} h(X_i, X_j) - \sum_{i=1}^{\tau} \sum_{j=\tau+1}^{n} h(X_i, X_j). \tag{5.80}$$

(i) If $\hat{\tau} < \tau$, we have

$$\begin{aligned}
&\sum_{i=1}^{\hat{\tau}} \sum_{j=\hat{\tau}+1}^{n} h(X_i, X_j) - \sum_{i=1}^{\tau} \sum_{j=\tau+1}^{n} h(X_i, X_j) \\
&\quad = \sum_{i=1}^{\hat{\tau}} \Big(\sum_{j=\hat{\tau}+1}^{\tau} + \sum_{j=\tau+1}^{n} \Big) h(X_i, X_j) - \Big(\sum_{i=1}^{\hat{\tau}} + \sum_{i=\hat{\tau}+1}^{\tau} \Big) \sum_{j=\tau+1}^{n} h(X_i, X_j)
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{\hat{\tau}} \sum_{j=\hat{\tau}+1}^{\tau} h(X_i, X_j) - \sum_{i=\hat{\tau}+1}^{\tau} \sum_{j=\tau+1}^{n} h(X_i, X_j) \\
&= \sum_{i=\hat{\tau}+1}^{\tau} \sum_{j=1}^{\hat{\tau}} h(X_i, X_j) - \sum_{i=\hat{\tau}+1}^{\tau} \sum_{j=\tau+1}^{n} h(X_i, X_j)
\end{aligned} \tag{5.81}
$$

since h is symmetric. By (5.60) and Lemma 5.2, we get

$$
\begin{aligned}
&\sum_{i=\hat{\tau}+1}^{\tau} \sum_{j=1}^{\hat{\tau}} h(X_i, X_j) - \sum_{i=\hat{\tau}+1}^{\tau} \sum_{j=\tau+1}^{n} h(X_i, X_j) \\
&\quad = n\Big\{O_P\Big(\sqrt{(\tau-\hat{\tau})\log\log(\tau-\hat{\tau})}\Big) - (1-\lambda)(\theta_{12}-\theta_1)(\tau-\hat{\tau})\Big\} \\
&\quad = O_P(n).
\end{aligned} \tag{5.82}
$$

Thus (5.78) follows by (5.74), (5.80)-(5.82).

(ii) If $\hat{\tau} > \tau$, we consider the opposite direction sequence of the observations which was used in the proof of Lemma 5.2. Let

$$X'_i = X_{n+1-\tau}, \; i = 1, \cdots, n,$$

$$\tau' = n - \tau \quad \text{and} \quad \hat{\tau}' = n - \hat{\tau}.$$

τ' and $\hat{\tau}'$ are the change-point and the change-point estimator of the opposite direction sequence, respectively. Thus

$$\hat{\tau}' < \tau' \quad \text{if } \hat{\tau} > \tau$$

and

$$
\begin{aligned}
&\sum_{i=1}^{\hat{\tau}} \sum_{j=\hat{\tau}+1}^{n} h(X_i, X_j) - \sum_{i=1}^{\tau} \sum_{j=\tau+1}^{n} h(X_i, X_j) \\
&\quad = \sum_{j=\hat{\tau}'+1}^{n} \sum_{i=1}^{\hat{\tau}'} h(X'_j, X'_i) - \sum_{j=\tau'+1}^{n} \sum_{i=1}^{\tau'} h(X'_j, X'_i) \\
&\quad = \sum_{i=1}^{\hat{\tau}'} \sum_{j=\hat{\tau}'+1}^{n} h(X'_i, X'_j) - \sum_{i=1}^{\tau'} \sum_{j=\tau'+1}^{n} h(X'_i, X'_j)
\end{aligned} \tag{5.83}
$$

since h is symmetric. Similar to the proof in part (i), we can prove

$$\sum_{i=1}^{\hat{\tau}'}\sum_{j=\hat{\tau}'+1}^{n} h(X_i', X_j') - \sum_{i=1}^{\tau'}\sum_{j=\tau'+1}^{n} h(X_i', X_j') = O_P(n). \tag{5.84}$$

Thus (5.78) follows by (5.74), (5.80), (5.83) and (5.84). Therefore, (5.70) follows by (5.75)-(5.78). □

Theorem 5.2 can give an approximation to the power of the test based on $\max_{1\le k<n} U_k$ for large n.

5.3 Asymptotics for Anti-symmetric Kernel

If h is anti-symmetric, then we get

$$\theta_1 = Eh(X_1, X_2) = 0, \tag{5.85}$$

$$\theta_2 = Eh(X_{n-1}, X_n) = 0, \tag{5.86}$$

and

$$\begin{aligned} U_k = Z_k &= \sum_{i=1}^{k}\sum_{j=k+1}^{n} h(X_i, X_j) \\ &= \sum_{i=1}^{k}\sum_{j=1}^{n} h(X_i, X_j),\ 1 \le k < n. \end{aligned} \tag{5.87}$$

Let

$$V_i = \sum_{j=1}^{n} h(X_i, X_j),\ i = 1, \cdots, \tau. \tag{5.88}$$

Thus

$$U_k = Z_k = \sum_{i=1}^{k} V_i \tag{5.89}$$

and

$$\hat{\tau} = \min\left\{k : Z_k = \max_{1\leq m<n} Z_m\right\}$$
$$= \min\left\{k : \sum_{i=1}^{k} V_i = \max_{1\leq m<n} \sum_{i=1}^{m} V_i\right\}. \tag{5.90}$$

We have

Lemma 5.3. *We assume that h is anti-symmetric and that, under H_1,* (5.5), (5.11) *hold and $\theta_{12} > 0$. Then*

$$|\hat{\tau} - \tau| = O_P(1), \tag{5.91}$$

where θ_{12} and $\hat{\tau}$ are defined in (5.8) *and* (5.13), *respectively.*

Proof. As h is anti-symmetric, we only need show that

$$\lim_{k\to\infty} \limsup_{n\to\infty} P\Big\{\max_{1\leq m\leq \tau-k} \sum_{i=1}^{m} V_i \geq \max_{\tau-k<m<\tau+k} \sum_{i=1}^{m} V_i\Big\} = 0 \tag{5.92}$$

and

$$\lim_{k\to\infty} \limsup_{n\to\infty} P\Big\{\max_{\tau+k\leq m<n} \sum_{i=1}^{m} V_i \geq \max_{\tau-k<m<\tau+k} \sum_{i=1}^{m} V_i\Big\} = 0. \tag{5.93}$$

We prove (5.92) first. Denote the empirical distribution function of $X_1, \cdots,$ X_n by

$$H_n(t) = \frac{1}{n}\sum_{i=1}^{n} I(X_i \leq t) \tag{5.94}$$

and

$$H(t) = \lambda F^{(1)}(t) + (1-\lambda)F^{(2)}(t). \tag{5.95}$$

Note that $V_1, \cdots, V_\tau$ are identically distributed although not independent, using the uniform approximation of an empirical process by a Brownian bridge again, we may write

$$
\begin{aligned}
V_i &= \sum_{j=1}^{n} h(X_i, X_j) \\
&= n \int h(X_i, t) dH_n(t) \\
&= n \int h(X_i, t) dH(t) + \sqrt{n} \int h(X_i, t) d\Big[\sqrt{n}\Big(H_n(t) - H(t)\Big)\Big] \\
&\overset{\mathcal{D}}{=} nQ_i + \sqrt{n} \int h(X_i, t) d[B(H(t)) + o_P(1)], \; i \le \tau,
\end{aligned}
\tag{5.96}
$$

where $\{B(u), 0 \le u \le 1\}$ is a Brownian bridge, and

$$
\begin{aligned}
Q_i &= \int h(X_i, t) dH(t) \\
&= \lambda \int h(X_i, t) dF^{(1)}(t) + (1 - \lambda) \int h(X_i, t) dF^{(2)}(t) \\
&= \lambda E[h(X_i, X_1)|X_i] + (1 - \lambda) E[h(X_i, X_n)|X_i] \\
&= -\lambda \tilde{h}_1(X_i) - (1 - \lambda) \tilde{h}_2(X_i)
\end{aligned}
\tag{5.97}
$$

since h is anti-symmetric.

Note that $Q_1, \cdots, Q_\tau$ are independent and identically distributed, by (5.97),

$$
\begin{aligned}
EQ_i &= \lambda Eh(X_2, X_1) + (1 - \lambda) Eh(X_1, X_n) \\
&= (1 - \lambda)\theta_{12} > 0, \; i \le \tau,
\end{aligned}
\tag{5.98}
$$

and

$$
\begin{aligned}
Var(Q_i) &= \lambda^2 Var[\tilde{h}_1(X_i)] + (1 - \lambda)^2 Var[\tilde{h}_2(X_i)] + 2\lambda(1 - \lambda) Cov[\tilde{h}_1(X_i), \tilde{h}_2(X_i)] \\
&= \lambda^2 E\tilde{h}_1^2(X_1) + (1 - \lambda)^2 [E\tilde{h}_2^2(X_1) - \theta_{12}^2] \\
&\quad + 2\lambda(1 - \lambda) E\tilde{h}_1(X_1)\tilde{h}_2(X_1), \; i \le \tau.
\end{aligned}
\tag{5.99}
$$

We have

$$
\begin{aligned}
&P\Big\{ \max_{1 \le m \le \tau - k} \sum_{i=1}^{m} V_i \ge \max_{\tau - k < m < \tau + k} \sum_{i=1}^{m} V_i \Big\} \\
&\le P\Big\{ \max_{1 \le m \le \tau - k} \sum_{i=1}^{m} V_i \ge \max_{\tau - k < m \le \tau} \sum_{i=1}^{m} V_i \Big\}
\end{aligned}
$$

$$= P\Big\{\exists j, 1 \leq j \leq \tau - k : \sum_{i=1}^{j} V_i \geq \sum_{i=1}^{\tau-k} V_i + \max_{\tau-k<m\leq\tau} \sum_{i=\tau-k+1}^{m} V_i\Big\}$$
$$= P\Big\{\frac{1}{n\sqrt{k}} \min_{1\leq m\leq\tau-k} \sum_{i=m+1}^{\tau-k} V_i + \frac{1}{n\sqrt{k}} \max_{\tau-k<m\leq\tau} \sum_{i=\tau-k+1}^{m} V_i \leq 0\Big\}. \quad (5.100)$$

Now,

$$\frac{1}{n\sqrt{k}} \min_{1\leq m\leq\tau-k} \sum_{i=m+1}^{\tau-k} V_i \overset{\mathcal{D}}{=} \frac{1}{n\sqrt{k}} \min_{1\leq m<\tau-k} \sum_{i=1}^{m} V_i. \quad (5.101)$$

Let $\{g(k)\}$ be a sequence such that

$$g(k) < \tau - k,\ g(k) \to \infty,\ \frac{g(k)}{k} \to 0 \quad \text{as} \quad k \to \infty. \quad (5.102)$$

Also, denote

$$[x]^- = -\min(x, 0). \quad (5.103)$$

Thus, by (5.96), (5.98) and $\theta_{12} > 0$, we get

$$-\Big[\frac{1}{n\sqrt{k}} \min_{1\leq m\leq g(k)} \sum_{i=1}^{m} V_i\Big]^-$$
$$\overset{\mathcal{D}}{=} -\Big[\frac{1}{\sqrt{k}} \min_{1\leq m\leq g(k)} \Big\{\sum_{i=1}^{m} Q_i + \frac{1}{\sqrt{n}} \sum_{i=1}^{m} \int h(X_i, t) d(B(H(t)) + o_P(1))\Big\}\Big]^-$$
$$= -\Big[\frac{1}{\sqrt{k}} \min_{1\leq m\leq g(k)} \Big\{\sum_{i=1}^{m} (Q_i - EQ_i) + m(EQ_i + \frac{1}{\sqrt{n}} O_P(1))\Big\}\Big]^-$$
$$= -\Big[\min_{1\leq m\leq g(k)} \Big\{\sqrt{\frac{g(k)}{k}} \frac{1}{\sqrt{g(k)}} \sum_{i=1}^{m} (Q_i - EQ_i) + \frac{m}{\sqrt{k}}((1-\lambda)\theta_{12} + \frac{1}{\sqrt{n}} O_P(1))\Big\}\Big]^-$$
$$= o_P(1), \quad (5.104)$$

which can be seen by the Wiener approximation to the partial sum process

$$\Big\{\sum_{i=1}^{m} \Big(\frac{Q_i - EQ_i}{\sqrt{Var(Q_i)}}\Big),\ m = 1, \cdots, g(k)\Big\}$$

and the continuity of the Wiener process at 0.

Furthermore, by the law of iterated logarithm and $\theta_{12} > 0$,

$$
\begin{aligned}
&-\Big[\frac{1}{n\sqrt{k}} \min_{g(k)<m<\tau-k} \sum_{i=1}^{m} V_i\Big]^- \\
&\stackrel{\mathcal{D}}{=} -\Big[\min_{g(k)<m<\tau-k} \Big\{\frac{1}{\sqrt{k}} \sum_{i=1}^{m}(Q_i - EQ_i) + \frac{m}{\sqrt{k}}\Big((1-\lambda)\theta_{12} + \frac{1}{\sqrt{n}} O_P(1)\Big)\Big\}\Big]^- \\
&= -\Big[\min_{g(k)<m<\tau-k} \Big\{\frac{m}{\sqrt{k}}\Big(O_P\Big(\sqrt{\frac{\log\log m}{m}}\Big) + (1-\lambda)\theta_{12} + \frac{1}{\sqrt{n}} O_P(1)\Big)\Big\}\Big]^- \\
&= o_P(1)
\end{aligned}
\tag{5.105}
$$

when k, and hence $g(k)$, is large enough.

Thus, by (5.101), (5.104) and (5.105), we have

$$
\begin{aligned}
&-\Big[\frac{1}{n\sqrt{k}} \min_{1\le m\le\tau-k} \sum_{i=m+1}^{\tau-k} V_i\Big]^- \\
&\stackrel{\mathcal{D}}{=} -\Big[\frac{1}{n\sqrt{k}} \min_{1\le m<\tau-k} \sum_{i=1}^{m} V_i\Big]^- \\
&= \min\Big\{ -\Big[\frac{1}{n\sqrt{k}} \min_{1\le m\le g(k)} \sum_{i=1}^{m} V_i\Big]^-, -\Big[\frac{1}{n\sqrt{k}} \min_{g(k)<m<\tau-k} \sum_{i=1}^{m} V_i\Big]^- \Big\} \\
&= o_P(1).
\end{aligned}
\tag{5.106}
$$

On the other hand, by the law of large numbers, the law of iterated logarithm and $\theta_{12} > 0$, we have

$$
\begin{aligned}
&\frac{1}{n\sqrt{k}} \max_{\tau-k<m\le\tau} \sum_{i=\tau-k+1}^{m} V_i \\
&\quad\ge \frac{1}{n\sqrt{k}} \sum_{i=\tau-k+1}^{\tau} V_i \\
&\quad\stackrel{\mathcal{D}}{=} \frac{1}{n\sqrt{k}} \sum_{i=1}^{k} V_i \\
&\quad\stackrel{\mathcal{D}}{=} \frac{1}{\sqrt{k}} \sum_{i=1}^{k} Q_i + \frac{1}{\sqrt{nk}} \sum_{i=1}^{k} \int h(X_i,t)d[B(H(t)) + o_P(1)] \\
&\quad= \sqrt{k}\Big(\frac{1}{k} \sum_{i=1}^{k} Q_i + \frac{1}{\sqrt{nk}\sqrt{k}} \sum_{i=1}^{k} \int h(X_i,t)d[B(H(t)) + o_P(1)]\Big)
\end{aligned}
$$

$$= \sqrt{k}\Big[(1-\lambda)\theta_{12} + \frac{1}{\sqrt{n}} O_P\Big(\sqrt{\frac{\log\log k}{k}}\Big)\Big] + o_P(1) \tag{5.107}$$

$$\to +\infty \quad \text{as} \quad k \to +\infty$$

since $\theta_{12} > 0$.

Therefore, putting together (5.100), (5.106) and (5.107), we can get (5.92).

In order to prove (5.93), we note that

$$\sum_{i=1}^{n} V_i = 0$$

and

$$\sum_{i=1}^{m} V_i = \sum_{i=m+1}^{n} (-V_i).$$

If we consider the opposite direction sequence of observations and use the method in the proof of Lemma 5.2, we'll find relation (5.93) is a symmetrical version of (5.92), so its proof is the same, hence it will be omitted. □

Theorem 5.3. *We assume that h is anti-symmetric and that, under H_1,* (5.5), (5.11) *hold and* $\theta_{12} > 0$. *Then*

$$\frac{\max_{1 \le k < n} U_k - \tau(n-\tau)\theta_{12}}{n^{3/2}\sqrt{\lambda(1-\lambda)^2 D_1^2 + \lambda^2(1-\lambda) D_2^2}} \xrightarrow{\mathcal{D}} N(0,1), \tag{5.108}$$

where θ_{12}, U_k, D_1^2 and D_2^2 are defined in (5.8), (5.12), (5.30) *and* (5.31), *respectively.*

Proof. We note that, by (5.71) and (5.85),

$$\max_{1 \le k < n} U_k = Z_{\hat{\tau}}. \tag{5.109}$$

Thus,

$$\max_{1 \le k < n} U_k - \tau(n-\tau)\theta_{12} = [Z_\tau - \tau(n-\tau)\theta_{12}] + (Z_{\hat{\tau}} - Z_\tau). \tag{5.110}$$

Similar to the proof of Theorem 5.2, in order to prove (5.108), we only need show that

$$\frac{Z_{\hat{\tau}} - Z_\tau}{D} = o_P(1), \tag{5.111}$$

where D^2 is given by (5.73). Now, by (5.89),

$$Z_{\hat{\tau}} - Z_\tau = \sum_{i=1}^{\hat{\tau}} V_i - \sum_{i=1}^{\tau} V_i. \tag{5.112}$$

If $\hat{\tau} > \tau$, by (5.96) and Lemma 5.3, we have

$$\begin{aligned}
&\sum_{i=1}^{\hat{\tau}} V_i - \sum_{i=1}^{\tau} V_i \\
&\quad = \sum_{i=\tau+1}^{\hat{\tau}} V_i \\
&\quad \stackrel{\mathcal{D}}{=} \sum_{i=1}^{\hat{\tau}-\tau} V_i \\
&\quad \stackrel{\mathcal{D}}{=} \sum_{i=1}^{\hat{\tau}-\tau} \left\{ nQ_i + \sqrt{n} \int h(X_i, t) d[B(H(t)) + o_P(1)] \right\} \\
&\quad = n \sum_{i=1}^{\hat{\tau}-\tau} \left\{ Q_i + \frac{1}{\sqrt{n}} \int h(X_i, t) d[B(H(t)) + o_P(1)] \right\} \\
&\quad = O_P(n).
\end{aligned} \tag{5.113}$$

Thus (5.111) follows by (5.74), (5.112), (5.113).

Similarly, we can prove (5.111) if $\hat{\tau} < \tau$. □

Theorem 5.3 can give an approximation to the power of the test based on $\max_{\leq k < n} U_k$ for large n.

In Section 5.4, we do some simulation study to investigate the precision of the change-point estimator to the true value, and also to investigate the precision of the power approximation of the test.

5.4 Simulation Study

To illustrate the proposed change-point estimator and the power approximation of the test, we would like to check the precision of the estimated mean of the change-point estimator to the true value of change-point $\tau = [n\lambda]$, and also to compare the simulated power with the approximated power of the test to check the precision through the Monte Carlo simulation study.

For computational simplicity, we only consider the anti-symmetric kernel. In this case, if $Eh^2(X_1, X_2) < \infty$ and $E\tilde{h}_1^2(X_1) > 0$, by Theorem 2.4.10 of Csörgő and Horváth (1997), we get, under the null hypothesis H_0 of no change,

$$\frac{1}{n^{3/2}D_3} \max_{1\le k<n} U_k \xrightarrow{\mathcal{D}} \sup_{0<t<1} B(t), \tag{5.114}$$

where $\{B(t), 0 \le t \le 1\}$ is a Brownian bridge, and

$$D_3^2 = E\tilde{h}_1^2(X_1) = \int \Big[\int h(s,t)dF^{(1)}(s)\Big]^2 dF^{(1)}(t). \tag{5.115}$$

The test statistic is

$$\frac{1}{n^{3/2}D_3} \max_{1\le k<n} U_k. \tag{5.116}$$

For simplicity, consider the one-sided test which is defined by

$$I\Big(\frac{1}{n^{3/2}D_3} \max_{1\le k<n} U_k > c_{1\alpha}\Big), \tag{5.117}$$

where $c_{1\alpha}$ is the $(1-\alpha)$-quantile of $\sup_{0<t<1} B(t)$ which can be given by (cf. Shorack and Wellner (1986))

$$P\{\sup_{0<t<1} B(t) > x\} = \exp(-2x^2),\ x \ge 0. \tag{5.118}$$

So from (5.118) we get

$$c_{1\alpha} = \sqrt{-\ln\sqrt{\alpha}}. \tag{5.119}$$

We list several critical values of $c_{1\alpha}$ for some selected values of α in Table 5.1. We'll use Table 5.1 in the following simulation study.

Table 5.1. *Some selected critical values* $c_{1\alpha}$

α	0.175	0.15	0.125
$c_{1\alpha}$	0.9335	0.9739	1.0197
α	0.1	0.075	0.05
$c_{1\alpha}$	1.0730	1.1380	1.2239
α	0.025	0.01	0.0075
$c_{1\alpha}$	1.3581	1.5174	1.5641
α	0.005	0.0025	0.001
$c_{1\alpha}$	1.6276	1.7308	1.8585

Now, by Theorem 5.3, under H_1 we have

$$
\begin{aligned}
&P\Big\{\frac{1}{n^{3/2}D_3}\max_{1\le k<n} U_k > c_{1\alpha}\Big\} \\
&= P\Big\{\frac{1}{n^{3/2}}\max_{1\le k<n} U_k > c_{1\alpha}D_3\Big\} \\
&= P\Bigg\{\frac{\max_{1\le k<n} U_k - \tau(n-\tau)\theta_{12}}{n^{3/2}\sqrt{\lambda(1-\lambda)^2D_1^2+\lambda^2(1-\lambda)D_2^2}} > \frac{n^{3/2}c_{1\alpha}D_3 - \tau(n-\tau)\theta_{12}}{n^{3/2}\sqrt{\lambda(1-\lambda)^2D_1^2+\lambda^2(1-\lambda)D_2^2}}\Bigg\} \\
&\approx 1-\Phi(R_n).
\end{aligned}
\tag{5.120}
$$

(5.120) gives the power approximation of the one-sided test, where

$$
R_n = \frac{n^{3/2}c_{1\alpha}D_3 - \tau(n-\tau)\theta_{12}}{n^{3/2}\sqrt{\lambda(1-\lambda)^2D_1^2+\lambda^2(1-\lambda)D_2^2}}
$$

and Φ is the standard normal distribution function.

In our simulation study we took the anti-symmetric kernel

$$
h(x,y) = \text{sign}(y-x). \tag{5.121}
$$

Thus, by (5.115), we have

$$
\begin{aligned}
D_3^2 &= \int \Big[\int h(s,t)dF^{(1)}(s)\Big]^2 dF^{(1)}(t) \\
&= \int \Big[\int_{-\infty}^{t} \operatorname{sign}(t-s)dF^{(1)}(s) + \int_{t}^{\infty} \operatorname{sign}(t-s)dF^{(1)}(s)\Big]^2 dF^{(1)}(t) \\
&= \int \Big[\int_{-\infty}^{t} dF^{(1)}(s) - \int_{t}^{\infty} dF^{(1)}(s)\Big]^2 dF^{(1)}(t) \\
&= \int [2F^{(1)}(t) - 1]^2 dF^{(1)}(t) \\
&= \frac{1}{3},
\end{aligned} \tag{5.122}
$$

which is true for any distribution functions $F^{(1)}$ and $F^{(2)}$.

We performed the simulation study in two ways with $N = 5000$ repetitions. One way is for the particular sampling situation, and in each case we have $\lambda = 0.5$ so that pre- and after-change samples are balanced in size. In this case we considered the different sample sizes $n = 50, 100, 200, 500$. Another way is for the unbalanced sample sizes, and in each case we took the same sample size $n = 100$ with varying λ's.

First, we did simulation study on the precision of the change-point estimator which is given by (5.13). For the generated data, we considered three cases:

Case 1. The generated data had uniform distribution with

$$
F^{(1)} \sim U(0,1) \quad \text{and} \quad F^{(2)} \sim U(d, 1+d),\ 0 < d < 1. \tag{5.123}
$$

The results are reported in Table 5.2 and Table 5.3.

Case 2. The generated data had exponential distribution with

$$
F^{(1)} = \exp(1) \quad \text{and} \quad F^{(2)} = \exp(m),\ m > 1. \tag{5.124}
$$

The results are reported in Table 5.4 and Table 5.5.

Case 3. The generated data had normal distribution with

$$F^{(1)} \sim N(0,1) \quad \text{and} \quad F^{(2)} \sim N(\mu,1), \quad \mu > 0. \tag{5.125}$$

The results are reported in Table 5.6 and Table 5.7.

From Tables 5.2-5.7 we can see that, the simulated mean of $\hat{\tau}$ pretty well fits the true values of τ for the balanced sample sizes. On the other hand, the performance of $\hat{\tau}$ is still good when the change-point is close to the middle and gets somewhat worse for the unbalanced sample sizes especially when the change-point is close to the two tails. In such a situation, slight improvements may be obtained by considering the weighted version as pointed out in Chapter 4.

Next, we did simulation study on the power approximation of the one-sided test which is given by (5.120). For the generated data, we considered two cases: Case 1. $F^{(1)}$ and $F^{(2)}$ were uniform distribution functions which are given by (5.123). Thus, by (5.8), we have

$$\begin{aligned}
\theta_{12} &= \int\int h(s,t)dF^{(1)}(s)dF^{(2)}(t) \\
&= \int \Big[\int_{-\infty}^{t} \operatorname{sign}(t-s)dF^{(1)}(s) + \int_{t}^{+\infty} \operatorname{sign}(t-s)dF^{(1)}(s)\Big]dF^{(2)}(t) \\
&= \int \Big[\int_{-\infty}^{t} dF^{(1)}(s) - \int_{t}^{+\infty} dF^{(1)}(s)\Big]dF^{(2)}(t) \\
&= \int [2F^{(1)}(t) - 1]dF^{(2)}(t) \\
&= \int_{d}^{1+d} [2F^{(1)}(t) - 1]dt \\
&= 2d - d^2 > 0,
\end{aligned}$$

by (5.30), we have

$$D_1^2 = \int \Big[\int h(s,t)dF^{(2)}(t)\Big]^2 dF^{(1)}(s) - \theta_{12}^2$$
$$= \int [1 - 2F^{(2)}(s)]^2 dF^{(1)}(s) - \theta_{12}^2$$
$$= \int_0^1 [1 - 2F^{(2)}(s)]^2 ds - \theta_{12}^2$$
$$= (\frac{1}{3} + 2d^2 - \frac{4}{3}d^3) - (2d - d^2)^2$$
$$= \frac{1}{3} - 2d^2 + \frac{8}{3}d^3 - d^4,$$

and by (5.31), we have

$$D_2^2 = \int \Big[\int h(s,t)dF^{(1)}(s)\Big]^2 dF^{(2)}(t) - \theta_{12}^2$$
$$= \int [2F^{(1)}(t) - 1]^2 dF^{(2)}(t) - \theta_{12}^2$$
$$= \int_d^{1+d} [2F^{(1)}(t) - 1]^2 dt - \theta_{12}^2$$
$$= (\frac{1}{3} + 2d^2 - \frac{4}{3}d^3) - (2d - d^2)^2$$
$$= \frac{1}{3} - 2d^2 + \frac{8}{3}d^3 - d^4.$$

We took $d = 0.15, 0.20, 0.25, 0.30$, getting $\theta_{12} = 0.2775, 0.3600, 0.4375, 0.5100$, respectively for this simulation study. The results are reported in Table 5.8 and Table 5.9.

Case 2. $F^{(1)}$ and $F^{(2)}$ were exponential distribution functions which are given by (5.124). Thus, we have

$$\theta_{12} = \int [2F^{(1)}(t) - 1]dF^{(2)}(t)$$
$$= \int_0^{+\infty} [2(1 - e^{-t}) - 1]\frac{1}{m}e^{-t/m}dt$$
$$= \frac{m-1}{m+1} > 0,$$

$$D_1^2 = \int [1 - 2F^{(2)}(s)]^2 dF^{(1)}(s) - \theta_{12}^2$$
$$= \int_0^{+\infty} [1 - 2(1 - e^{-s/m})]^2 e^{-s} ds - \theta_{12}^2$$
$$= 1 + \frac{4m}{m+2} - \frac{4m}{m+1} - \Big(\frac{m-1}{m+1}\Big)^2,$$

and

$$\begin{aligned} D_2^2 &= \int [2F^{(1)}(t) - 1]^2 dF^{(2)}(t) - \theta_{12}^2 \\ &= \int_0^{+\infty} [2(1 - e^{-t}) - 1]^2 \frac{1}{m} e^{-t/m} dt - \theta_{12}^2 \\ &= 1 + \frac{4}{2m+1} - \frac{4}{m+1} - \left(\frac{m-1}{m+1}\right)^2. \end{aligned}$$

In order to get the same θ_{12} values as in Case 1, we took $m = 1.768166$, 2.125000, 2.555556, 3.081633 corresponding to $\theta_{12} = 0.2775$, 0.3600, 0.4375, 0.5100, respectively for this simulation study. The results are reported in Table 5.10 and Table 5.11.

From Tables 5.8-5.11 we can see that the approximations can be considered as good at sample size $n = 100$ as at $n = 500$. Our approximation is pretty close to the real power except for the situation that the change-point occurs on the tails although it somewhat underestimates the real power. The results indicate that the test statistic (5.116) is not sensitive for detecting change-point on the tails.

5.5 Conclusion and Future Research

We should point out that the topic of thesis is new and the results developed in this thesis confirm earlier findings and also give new insights. This thesis contains very useful, easily understood tests for changes in censored data. Such tests did not exist before and we have demonstrated their practical applications. Also, the proofs are based on one of the most recent results in mathematical statistics. Weak convergence of stochastic processes in weighted metrics as well as weighted approximations of partial sums and empirical processes are used.

Next, we list several problems which could be done in the future research.

The low power problems illustrated in the simulation study is also due to the fact that the censoring proportion of the simulated data is too big. It would also be helpful to see the effect of censoring on our procedures by comparing the simulated powers (and sizes) under various degrees of censoring, e.g. 025reported as well.

Note that the censoring data discussed are just right censoring data only, we could consider doubly censored data, that is either right censored or left censored. We point out that Gehan (1965b) generalized a two-sample Wilcoxon test further from right censored samples to doubly censored samples.

The variables of interests (the "lifetimes") and censoring distributions are assumed to be continuous in this thesis. We could extend our results to non-continuous lifetime and censorng distributions.

We considered at-most-one-change change-point problem under the alternative, we could consider at-most-two-change or more change cases under the alternative.

As we know that the proposed test for censored data are based on Gehan's generalization of the Wilcoxon rank test. It is known that the Gehan's approach does not have optimal efficiency properties in the two-sample problem as it assigns a zero value to the score function $h(.,.)$ defined in formula (2.2) when the smaller one of a pair of observations is censored. A more efficient change-point test, though technically more difficult to handle, could be derived by replacing the score function in (2.2) by Efron's (1967) version which assigns a positive score with values strictly between zero and one to such a pair of observations.

Table 5.2. *Comparison of true value τ and simulated $\hat{\tau}$ for uniform observations* ($\lambda = 0.5$)

n	d	$\tau = [n\lambda]$	Simulated Mean of $\hat{\tau}$
50	0.2	25	24.88
	0.3	25	24.99
	0.4	25	24.99
	0.5	25	24.95
	0.6	25	24.99
100	0.2	50	50.03
	0.3	50	49.88
	0.4	50	49.98
	0.5	50	50.01
	0.6	50	49.97
200	0.2	100	99.81
	0.3	100	100.14
	0.4	100	100.10
	0.5	100	99.98
	0.6	100	99.97
500	0.2	250	250.03
	0.3	250	249.98
	0.4	250	250.05
	0.5	250	250.00
	0.6	250	250.02

Table 5.3. *Comparison of true value τ and simulated $\hat{\tau}$ for uniform observations* ($n = 100$)

d	λ	$\tau = [n\lambda]$	Simulated Mean of $\hat{\tau}$
0.2	0.3	30	36.24
	0.4	40	42.58
	0.5	50	49.73
	0.6	60	57.65
	0.7	70	63.55
0.3	0.3	30	33.80
	0.4	40	41.55
	0.5	50	50.03
	0.6	60	58.41
	0.7	70	66.12
0.4	0.3	30	32.47
	0.4	40	40.96
	0.5	50	50.02
	0.6	60	58.94
	0.7	70	67.47
0.5	0.3	30	31.79
	0.4	40	40.66
	0.5	50	49.97
	0.6	60	59.27
	0.7	70	68.06
0.6	0.3	30	31.34
	0.4	40	40.49
	0.5	50	49.97
	0.6	60	59.52
	0.7	70	68.65

Table 5.4. *Comparison of true value τ and simulated $\hat{\tau}$ for exponential observations* ($\lambda = 0.5$)

n	m	$\tau = [n\lambda]$	Simulated Mean of $\hat{\tau}$
50	1.5	25	25.73
	2.0	25	25.34
	2.5	25	25.40
	3.0	25	25.46
	3.5	25	25.33
100	1.5	50	50.77
	2.0	50	50.93
	2.5	50	50.78
	3.0	50	50.58
	3.5	50	50.45
200	1.5	100	101.63
	2.0	100	100.91
	2.5	100	100.80
	3.0	100	100.61
	3.5	100	100.63
500	1.5	250	252.05
	2.0	250	251.23
	2.5	250	250.91
	3.0	250	250.69
	3.5	250	250.64

Table 5.5. *Comparison of true value τ and simulated $\hat{\tau}$ for exponential observations* ($n = 100$)

m	λ	$\tau = [n\lambda]$	Simulated Mean of $\hat{\tau}$
1.5	0.3	30	41.47
	0.4	40	45.88
	0.5	50	50.75
	0.6	60	55.64
	0.7	70	59.89
2.0	0.3	30	37.06
	0.4	40	43.67
	0.5	50	50.86
	0.6	60	57.82
	0.7	70	63.94
2.5	0.3	30	35.46
	0.4	40	42.65
	0.5	50	50.60
	0.6	60	58.58
	0.7	70	65.93
3.0	0.3	30	34.37
	0.4	40	42.15
	0.5	50	50.56
	0.6	60	59.05
	0.7	70	66.68
3.5	0.3	30	33.69
	0.4	40	41.83
	0.5	50	50.55
	0.6	60	59.14
	0.7	70	67.36

Table 5.6. *Comparison of true value τ and simulated $\hat{\tau}$ for normal observations* ($\lambda = 0.5$)

n	μ	$\tau = [n\lambda]$	Simulated Mean of $\hat{\tau}$
50	0.5	25	24.84
	1.0	25	25.00
	1.5	25	24.97
	2.0	25	24.99
	2.5	25	24.97
100	0.5	50	49.88
	1.0	50	50.00
	1.5	50	50.03
	2.0	50	50.01
	2.5	50	49.98
200	0.5	100	100.07
	1.0	100	100.02
	1.5	100	99.95
	2.0	100	99.99
	2.5	100	99.99
500	0.5	250	249.90
	1.0	250	249.99
	1.5	250	250.04
	2.0	250	249.97
	2.5	250	250.02

Table 5.7. *Comparison of true value τ and simulated $\hat{\tau}$ for normal observations* ($n = 100$)

μ	λ	$\tau = [n\lambda]$	Simulated Mean of $\hat{\tau}$
0.5	0.3	30	38.61
	0.4	40	43.69
	0.5	50	50.18
	0.6	60	56.28
	0.7	70	61.64
1.0	0.3	30	33.57
	0.4	40	41.49
	0.5	50	49.96
	0.6	60	58.37
	0.7	70	66.48
1.5	0.3	30	31.94
	0.4	40	40.73
	0.5	50	50.01
	0.6	60	59.28
	0.7	70	68.01
2.0	0.3	30	31.33
	0.4	40	40.48
	0.5	50	49.95
	0.6	60	59.52
	0.7	70	68.70
2.5	0.3	30	31.00
	0.4	40	40.35
	0.5	50	50.00
	0.6	60	59.63
	0.7	70	69.00

Table 5.8. *Comparison of simulated and approximated power for uniform observations* ($\lambda = 0.5$)

n	d	θ_{12}	α	Simulated Power	Approximated Power
50	0.15	0.2775	0.05	0.3728	0.2139
	0.15	0.2775	0.10	0.5134	0.3180
	0.20	0.3600	0.05	0.5806	0.3941
	0.20	0.3600	0.10	0.7110	0.5258
	0.25	0.4375	0.05	0.7558	0.6061
	0.25	0.4375	0.10	0.8510	0.7325
	0.30	0.5100	0.05	0.8794	0.7986
	0.30	0.5100	0.10	0.9366	0.8869
100	0.15	0.2775	0.05	0.6698	0.4812
	0.15	0.2775	0.10	0.7806	0.6074
	0.20	0.3600	0.05	0.8780	0.7704
	0.20	0.3600	0.10	0.9304	0.8585
	0.25	0.4375	0.05	0.9668	0.9407
	0.25	0.4375	0.10	0.9832	0.9721
	0.30	0.5100	0.05	0.9940	0.9926
	0.30	0.5100	0.10	0.9990	0.9975
200	0.15	0.2775	0.05	0.9234	0.8432
	0.15	0.2775	0.10	0.9592	0.9078
	0.20	0.3600	0.05	0.9918	0.9849
	0.20	0.3600	0.10	0.9968	0.9938
	0.25	0.4375	0.05	0.9998	0.9996
	0.25	0.4375	0.10	0.9998	0.9999
	0.30	0.5100	0.05	1.0000	1.0000
	0.30	0.5100	0.10	1.0000	1.0000
500	0.15	0.2775	0.05	0.9994	0.9990
	0.15	0.2775	0.10	0.9998	0.9997
	0.20	0.3600	0.05	1.0000	1.0000
	0.20	0.3600	0.10	1.0000	1.0000
	0.25	0.4375	0.05	1.0000	1.0000
	0.25	0.4375	0.10	1.0000	1.0000
	0.30	0.5100	0.05	1.0000	1.0000
	0.30	0.5100	0.10	1.0000	1.0000

Table 5.9. *Comparison of simulated and approximated power for uniform observations* ($n = 100$)

d	θ_{12}	λ	α	Simulated Power	Approximated Power
0.15	0.2775	0.1	0.05	0.1252	0.0026
0.15	0.2775	0.1	0.10	0.2200	0.0118
0.15	0.2775	0.2	0.05	0.3314	0.1141
0.15	0.2775	0.2	0.10	0.4868	0.2103
0.15	0.2775	0.3	0.05	0.5356	0.3099
0.15	0.2775	0.3	0.10	0.6706	0.4415
0.15	0.2775	0.4	0.05	0.6338	0.4395
0.15	0.2775	0.4	0.10	0.7602	0.5692
0.15	0.2775	0.5	0.05	0.6616	0.4812
0.15	0.2775	0.5	0.10	0.7754	0.6074
0.15	0.2775	0.6	0.05	0.6302	0.4395
0.15	0.2775	0.6	0.10	0.7486	0.5692
0.15	0.2775	0.7	0.05	0.5334	0.3099
0.15	0.2775	0.7	0.10	0.6794	0.4415
0.15	0.2775	0.8	0.05	0.3212	0.1141
0.15	0.2775	0.8	0.10	0.4862	0.2103
0.15	0.2775	0.9	0.05	0.1264	0.0026
0.15	0.2775	0.9	0.10	0.2222	0.0118
0.20	0.3600	0.1	0.05	0.1728	0.0073
0.20	0.3600	0.1	0.10	0.2984	0.0297
0.20	0.3600	0.2	0.05	0.5162	0.2660
0.20	0.3600	0.2	0.10	0.6810	0.4176
0.20	0.3600	0.3	0.05	0.7666	0.5817
0.20	0.3600	0.3	0.10	0.8574	0.7157
0.20	0.3600	0.4	0.05	0.8476	0.7307
0.20	0.3600	0.4	0.10	0.9170	0.8302
0.20	0.3600	0.5	0.05	0.8750	0.7704
0.20	0.3600	0.5	0.10	0.9280	0.8585
0.20	0.3600	0.6	0.05	0.8446	0.7307
0.20	0.3600	0.6	0.10	0.9136	0.8302
0.20	0.3600	0.7	0.05	0.7618	0.5817
0.20	0.3600	0.7	0.10	0.8624	0.7157
0.20	0.3600	0.8	0.05	0.5094	0.2660
0.20	0.3600	0.8	0.10	0.6722	0.4176
0.20	0.3600	0.9	0.05	0.1734	0.0073
0.20	0.3600	0.9	0.10	0.3046	0.0297

Table 5.9. *Comparison of simulated and approximated power for uniform observations* ($n = 100$) (*Continued*)

d	θ_{12}	λ	α	Simulated power	Approximated power
0.25	0.4375	0.1	0.05	0.2386	0.0178
0.25	0.4375	0.1	0.10	0.4002	0.0647
0.25	0.4375	0.2	0.05	0.7118	0.4867
0.25	0.4375	0.2	0.10	0.8504	0.6575
0.25	0.4375	0.3	0.05	0.9140	0.8246
0.25	0.4375	0.3	0.10	0.9596	0.9060
0.25	0.4375	0.4	0.05	0.9632	0.9212
0.25	0.4375	0.4	0.10	0.9816	0.9618
0.25	0.4375	0.5	0.05	0.9712	0.9407
0.25	0.4375	0.5	0.10	0.9848	0.9721
0.25	0.4375	0.6	0.05	0.9610	0.9212
0.25	0.4375	0.6	0.10	0.9834	0.9618
0.25	0.4375	0.7	0.05	0.9130	0.8246
0.25	0.4375	0.7	0.10	0.9606	0.9060
0.25	0.4375	0.8	0.05	0.7160	0.4867
0.25	0.4375	0.8	0.10	0.8310	0.6575
0.25	0.4375	0.9	0.05	0.2302	0.0178
0.25	0.4375	0.9	0.10	0.3968	0.0647
0.30	0.5100	0.1	0.05	0.3184	0.0383
0.30	0.5100	0.1	0.10	0.4914	0.1255
0.30	0.5100	0.2	0.05	0.8648	0.7213
0.30	0.5100	0.2	0.10	0.9436	0.8541
0.30	0.5100	0.3	0.05	0.9750	0.9560
0.30	0.5100	0.3	0.10	0.9902	0.9827
0.30	0.5100	0.4	0.05	0.9944	0.9883
0.30	0.5100	0.4	0.10	0.9982	0.9959
0.30	0.5100	0.5	0.05	0.9966	0.9926
0.30	0.5100	0.5	0.10	0.9992	0.9975
0.30	0.5100	0.6	0.05	0.9948	0.9883
0.30	0.5100	0.6	0.10	0.9978	0.9959
0.30	0.5100	0.7	0.05	0.9776	0.9560
0.30	0.5100	0.7	0.10	0.9920	0.9827
0.30	0.5100	0.8	0.05	0.8658	0.7213
0.30	0.5100	0.8	0.10	0.9426	0.8541
0.30	0.5100	0.9	0.05	0.3044	0.0383
0.30	0.5100	0.9	0.10	0.5010	0.1255

Table 5.10. *Comparison of simulated and approximated power for exponential observations* ($\lambda = 0.5$)

n	m	θ_{12}	α	Simulated Power	Approximated Power
50	1.768166	0.2775	0.05	0.3902	0.2160
	1.768166	0.2775	0.10	0.5300	0.3196
	2.125000	0.3600	0.05	0.5672	0.3957
	2.125000	0.3600	0.10	0.7070	0.5254
	2.555556	0.4375	0.05	0.7498	0.6038
	2.555556	0.4375	0.10	0.8378	0.7278
	3.081633	0.5100	0.05	0.8744	0.7909
	3.081633	0.5100	0.10	0.9360	0.8793
100	1.768166	0.2775	0.05	0.6636	0.4814
	1.768166	0.2775	0.10	0.7728	0.6065
	2.125000	0.3600	0.05	0.8754	0.7669
	2.125000	0.3600	0.10	0.9292	0.8548
	2.555556	0.4375	0.05	0.9702	0.9364
	2.555556	0.4375	0.10	0.9846	0.9691
	3.081633	0.5100	0.05	0.9932	0.9909
	3.081633	0.5100	0.10	0.9976	0.9968
200	1.768166	0.2775	0.05	0.9180	0.8410
	1.768166	0.2775	0.10	0.9624	0.9058
	2.125000	0.3600	0.05	0.9928	0.9836
	2.125000	0.3600	0.10	0.9968	0.9931
	2.555556	0.4375	0.05	0.9996	0.9995
	2.555556	0.4375	0.10	0.9998	0.9999
	3.081633	0.5100	0.05	1.0000	1.0000
	3.081633	0.5100	0.10	1.0000	1.0000
500	1.768166	0.2775	0.05	0.9994	0.9989
	1.768166	0.2775	0.10	0.9998	0.9996
	2.125000	0.3600	0.05	1.0000	1.0000
	2.125000	0.3600	0.10	1.0000	1.0000
	2.555556	0.4375	0.05	1.0000	1.0000
	2.555556	0.4375	0.10	1.0000	1.0000
	3.081633	0.5100	0.05	1.0000	1.0000
	3.081633	0.5100	0.10	1.0000	1.0000

Table 5.11. *Comparison of simulated and approximated power for exponential observations* ($n = 100$)

m	θ_{12}	λ	α	Simulated Power	Approximated Power
1.768166	0.2775	0.1	0.05	0.1148	0.0013
1.768166	0.2775	0.1	0.10	0.2138	0.0075
1.768166	0.2775	0.2	0.05	0.3274	0.1023
1.768166	0.2775	0.2	0.10	0.4874	0.1983
1.768166	0.2775	0.3	0.05	0.5394	0.3046
1.768166	0.2775	0.3	0.10	0.6636	0.4397
1.768166	0.2775	0.4	0.05	0.6288	0.4389
1.768166	0.2775	0.4	0.10	0.7454	0.5699
1.768166	0.2775	0.5	0.05	0.6730	0.4814
1.768166	0.2775	0.5	0.10	0.7818	0.6065
1.768166	0.2775	0.6	0.05	0.6210	0.4412
1.768166	0.2775	0.6	0.10	0.7526	0.5673
1.768166	0.2775	0.7	0.05	0.5132	0.3178
1.768166	0.2775	0.7	0.10	0.6462	0.4441
1.768166	0.2775	0.8	0.05	0.3470	0.1290
1.768166	0.2775	0.8	0.10	0.4806	0.2248
1.768166	0.2775	0.9	0.05	0.1228	0.0049
1.768166	0.2775	0.9	0.10	0.2334	0.0184
2.125000	0.3600	0.1	0.05	0.1674	0.0036
2.125000	0.3600	0.1	0.10	0.3046	0.0190
2.125000	0.3600	0.2	0.05	0.5296	0.2522
2.125000	0.3600	0.2	0.10	0.6886	0.4121
2.125000	0.3600	0.3	0.05	0.7638	0.5848
2.125000	0.3600	0.3	0.10	0.8634	0.7230
2.125000	0.3600	0.4	0.05	0.8542	0.7328
2.125000	0.3600	0.4	0.10	0.9216	0.8327
2.125000	0.3600	0.5	0.05	0.8768	0.7669
2.125000	0.3600	0.5	0.10	0.9312	0.8548
2.125000	0.3600	0.6	0.05	0.8406	0.7227
2.125000	0.3600	0.6	0.10	0.9112	0.8206
2.125000	0.3600	0.7	0.05	0.7528	0.5768
2.125000	0.3600	0.7	0.10	0.8416	0.7037
2.125000	0.3600	0.8	0.05	0.5102	0.2831
2.125000	0.3600	0.8	0.10	0.6752	0.4243
2.125000	0.3600	0.9	0.05	0.1778	0.0142
2.125000	0.3600	0.9	0.10	0.3034	0.0452

Table 5.11. *Comparison of simulated and approximated power for exponential observations* ($n = 100$) (*Continued*)

m	θ_{12}	λ	α	Simulated Power	Approximated Power
2.555556	0.4375	0.1	0.05	0.2266	0.0089
2.555556	0.4375	0.1	0.10	0.3792	0.0436
2.555556	0.4375	0.2	0.05	0.7340	0.4856
2.555556	0.4375	0.2	0.10	0.8478	0.6699
2.555556	0.4375	0.3	0.05	0.9256	0.8351
2.555556	0.4375	0.3	0.10	0.9604	0.9154
2.555556	0.4375	0.4	0.05	0.9600	0.9230
2.555556	0.4375	0.4	0.10	0.9790	0.9631
2.555556	0.4375	0.5	0.05	0.9670	0.9364
2.555556	0.4375	0.5	0.10	0.9834	0.9691
2.555556	0.4375	0.6	0.05	0.9564	0.9098
2.555556	0.4375	0.6	0.10	0.9784	0.9534
2.555556	0.4375	0.7	0.05	0.8944	0.8050
2.555556	0.4375	0.7	0.10	0.9476	0.8874
2.555556	0.4375	0.8	0.05	0.7068	0.4881
2.555556	0.4375	0.8	0.10	0.8354	0.6420
2.555556	0.4375	0.9	0.05	0.2430	0.0331
2.555556	0.4375	0.9	0.10	0.4046	0.0924
3.081633	0.5100	0.1	0.05	0.3062	0.0204
3.081633	0.5100	0.1	0.10	0.4824	0.0925
3.081633	0.5100	0.2	0.05	0.8758	0.7403
3.081633	0.5100	0.2	0.10	0.9548	0.8765
3.081633	0.5100	0.3	0.05	0.9792	0.9632
3.081633	0.5100	0.3	0.10	0.9938	0.9867
3.081633	0.5100	0.4	0.05	0.9936	0.9887
3.081633	0.5100	0.4	0.10	0.9978	0.9961
3.081633	0.5100	0.5	0.05	0.9970	0.9909
3.081633	0.5100	0.5	0.10	0.9992	0.9968
3.081633	0.5100	0.6	0.05	0.9902	0.9828
3.081633	0.5100	0.6	0.10	0.9966	0.9933
3.081633	0.5100	0.7	0.05	0.9652	0.9383
3.081633	0.5100	0.7	0.10	0.9864	0.9718
3.081633	0.5100	0.8	0.05	0.8420	0.6962
3.081633	0.5100	0.8	0.10	0.9266	0.8218
3.081633	0.5100	0.9	0.05	0.3262	0.0663
3.081633	0.5100	0.9	0.10	0.5100	0.1648

References

Achcar, J. A. (1989). Constant Hazard Against a Change-point Alternative: a Bayesian Approach With Censored Data. *Communications in Statistics-Theory and Methods* **18**, 3801-3819.

Andersen, P. K., Borgan, O., Gill, R. D. and Kieding, N. (1982). Linear Non-parametric Tests for Comparison of Counting Processes, With Applications to Censored Data. *International Statistical Review* **50**, 219-258. Correction (1984), **52**, 225.

Bhattacharya, P. K. (1994). Some Aspects of Change-point Analysis. In *Change-point Problems, IMS Lecture Notes-Monograph Series* **23**, 28-56.

Billingsley, P. (1968). *Convergence of Probability Measures.* New York: John Wiley.

Breslow, N. (1970). A Generalized Kruskal-Wallis Test for Comparing K Samples Subject to Unequal Patterns of Censorship. *Biometrika* **57**, 579-594.

Breslow, N., Edler, L. and Berger, J. (1984). A Two-Sample Censored Data Rank Test for Acceleration. *Biometrics* **40**, 1049-1062.

Crowley, J. and Hu, M. (1977). Covariance Analysis of Heart Transplant Survival Data. *Journal of the American Statistical Association* **72**, 27-36.

Csörgő, M. and Horváth, L. (1988). Nonparametric Methods for Change-point Problems. In *Handbook of Statistics* **7**, 403-425.

Csörgő, M. and Horváth, L. (1993). *Weighted Approximations in Probability and Statistics.* Chichester: John Wiley.

Csörgő, M. and Horváth, L. (1997). *Limit Theorems in Change-point Analysis.* New York: John Wiley.

Csörgő, M. and Révész, P. (1981) *Strong Approximations in Probability and Statistics.* New York: Academic Press.

Donsker, M. D. (1952). Justification and Extension of Doob's Heuristic Approach to the Kolmogorov-Smirnov Theorems. *Annals of Mathematical Statistics* **23**, 277-281.

Dvoretzky, A. (1972). Central Limit Theorems for Dependent Random Variables. In *Proceedings of the Sixth Berkeley Symposium on Mathematics, Statistics and Probability* **2**, 513-553. Los Angeles: University of California Press.

Efron, B. (1967). The Two Sample Problem With Censored Data. In *Proceedings of Fifth Berkeley Symposium* **4**, 831-853.

Fleming, T. R., O'Fallon, J. R., O'Brien, P. C. & Harrington, D. P. (1980). Modified Kolmogorov-Smirnov Test Procedures With Application to Arbitrarily Censored Data. *Biometrics* **36**, 607-625.

Gehan, E. (1965a). A Generalized Wilcoxon Test for Comparing Arbitrarily Single Censored Samples. *Biometrika* **52**, 203-223.

Gehan, E. (1965b). A Generalized Two-sample Wilcoxon Test for Doubly Censored Data. *Biometrika* **52**, 650-653.

Gill, R. and Schumacher, M. (1987). A Simple Test of the Proportional Hazards Assumption. *Biometrika* **74**, 289-300.

Gombay, E. (1994). A Limit Theorem for Rank Tests of the Change-point Problem. In *Proceedings of the Twelfth Prague Conference, Academy of Sciences of the Czech Republic & Charles University.* 81-83.

Gombay, E. and Horváth, L. (1996). Approximations for the Time of Change and the Power Function in Change-point Models. *Journal of Statistical Planning and Inference* **52**, 43-66.

Gombay, E. and Jin, X. (1996). Sign Tests for Change Under Alternatives. Manuscript.

Harrington, D. P. and Fleming, T. R. (1982). A Class of Rank Test Procedures for Censored Survival Data. *Biometrika* **69**, 553-566.

Henderson, R. (1990). A Problem With the Likelihood Ratio Test for a Change-point Hazard Rate Model. *Biometrika* **77**, 835-843.

Janson, S. and Wichura, M. J. (1983). Invariance Principles for Stochastic Area and Related Stochastic Integrals. *Stochastic Processes and their Applications* **16**, 71-84.

Kalbfleisch, J. D. and Prentice, R. L. (1980). *The Statistical Analysis of Failure Time Data.* New York: John Wiley.

Kolmogorov, A. N. (1933). Sulla Determinazione Empirica di Une Legge di Distribuzione. *Giorn. Inst. Ital. Attuari* **4**, 83-91.

Lee, A. J. (1990). *U-Statistics.* Marcel Dekker Inc. New York.

Leurgans, S. (1983). Three Classes of Censored Data Rank Tests: Strengths and Weakness Under Censoring. *Biometrika* **70**, 651-658.

Leurgans, S. (1984). Asymptotic Bahaviour of Two-sample Rank Tests in the Presence of Random Censoring. *Annals of Statistics* **12**, 572-589.

Loader, C. R. (1991). Inference for a Hazard Rate Change-point Model. *Biometrika* **78**, 749-757.

Mantel, N. (1967). Ranking Procedures for Arbitrarilly Restricted Observation.

Biometrics **23**, 65-78.

Matthews, D. E. and Farewell, V. T. (1982). On Testing for a Constant Hazard Against a Change-point Alternative. *Biometrics* **38**, 463-468.

Matthews, D. E., Farewell, V. T. and Pyke, R. (1985). Asymptotic Score-statistic Processes and Tests for Constant Hazard Against a Change-point Alternative. *Annals of Statistics* **13**, 583-591.

Müller, H. G. and Wang, J.-L. (1994). Change-point Models for Hazard Functions. In *Change-point Problems, IMS Lecture Notes-Monograph Series* **23**, 224-241.

Nelson, W. (1969). Hazard Plotting for Incomplete Failure Data. *J. Qual. Technol.* **1**, 27-52.

Peto, R. and Peto, J. (1972). Asymptotically Efficient Rank Invariant Test Procedures (With Discussion). *Journal of the Royal Statistical Society, Ser. A* **135**, 185-206.

Prentice, R. L. (1978). Linear Rank Tests With Right Censored Data. *Biometrika* **63**, 291-298.

Prentice, R. L. and Marek, P. (1979). A Qualitative Discrepency Between Censored Data Rank Tests. *Biometrics* **35**, 861-867.

Schumacher, M. (1984). Two-sample Tests of Cramér-von Mises- and Kolmogorov-Smirnov-type for Randomly Censored Data. *International Statistical Review* **52**, 263-281.

Sen, P. K. (1981). *Sequential Nonparametrics.* New York: John Wiley.

Serfling, R. J. (1980). *Approximation Theorem of Mathematical Statistics.* New York: John Wiley.

Shorack, G. R. and Wellner, J. A. (1986). *Empirical Processes With Application to Statistics*. New York: John Wiley.

Stute, W. (1996). Changepoint Problems Under Random Censorship. *Statistics* **27**, 255-266.

Wei, L. J. (1984). Testing Goodness-of-Fit for Proportional Hazards Model With Censored Observations. *Journal of the American Statistical Association.* **79**, 649-652.

Wolfe, D. A. and Schechtman, E. (1984). Nonparametric Statistical Procedures for the Changepoint Problem. *Journal of Statistical Planning and Inference* **9**, 389-396.

Worsley, K. J. (1988). Exact Percentage Points of the Likelihood-ratio Test for a Change-point Hazard-rate Model. *Biometrics* **44**, 259-263.

IMAGE EVALUATION TEST TARGET (QA-3)

150mm

6"

APPLIED IMAGE . Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989

© 1993, Applied Image, Inc., All Rights Reserved