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.

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.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

University of Alberta

Nonparametric Sequential Change Detection in the Distribution of Randomly Censored Data

by

William Kai Midodzi 

A thesis submitted to the Faculty of Graduate Studies and Research in Partial fulfilment of the requirements for the Degree of Master of Science

in

Biostatistics

Department of Mathematical Sciences

Edmonton, Alberta

[Spring 2001]

University of Alberta

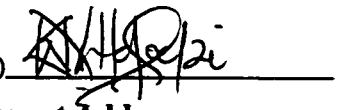
Library Release Form

Name of Author: William Kai Midodzi

Title of Thesis: Nonparametric Sequential Change Detection in the Distribution of Randomly Censored Data

Degree: Master of Science

Year this Degree Granted: 2001

(Signed) 

Permanent Address
Dept of Mathematical Sciences
University of Alberta
CAB 632
Edmonton, Alberta
Canada T6G 2G1

Date: December 19, 2000

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 Sequential Change Detection in the Distribution of Randomly Censored Data** submitted by **William Kai Midodzi** in partial fulfillment of the requirements for the degree of **Master of Science** in Biostatistics.

Dr. Edit Gombay (Supervisor)

Dr. Rohana Karunamuni (Chair)

Dr. A. Senthilselvan (Public Health Sciences)

Date: December 19, 2000

Abstract

Nonparametric statistical procedures are particularly useful in making inference in situations where serious doubt exists about the assumptions of the underlying stochastic process. In most cases, we will never know whether these assumptions hold in practical situations, but will often be reasonably certain that departure from these assumptions will be small enough so that the properties of the statistical procedure will be undisturbed.

In this thesis, we use sequential nonparametric techniques to validate the assumptions of identical distribution of time process which is common in most practice. We validate this assumption through a change detection procedure with particular attention to censored data.

For our proposed sequential test, *U*-statistics using anti-symmetric kernels are considered both under no-change and change hypothesis. We investigate the power of our test using simulation study, and demonstrate its application on two real and well-known data sets: the Stanford Heart Transplant data and the Radiation Therapy Oncology Group data.

Acknowledgements

I am very grateful to my supervisor, Prof. Edit Gombay, for leading me to the research field of this thesis, and for her immense support, suggestions and comments, encouragement and kindness throughout the preparation of this thesis. Without her guidance this thesis would have been impossible.

I also express my gratitude to my thesis committee members, Dr. R. Karunamuni and Dr. A. Senthilselvan, for their time spent in reading my thesis.

Many thanks also go to all my friends, in particular, Alex De Leon and my teaching mentor, Dr. Giseon Heo, for their time spent in reading my thesis and their extensive suggestions.

Special thanks go to the Department of Mathematical Sciences for providing me the opportunity to study here and for the financial support; and to all faculty members who taught me courses.

I also acknowledge the assistance of the department secretaries, especially Marion Benedict for her kindness and patience.

My final thanks go to my wife, Christine, for her moral support and encouragement throughout the preparation of this thesis.

Contents

List of Tables

List of Figures

Dedicated to my Parents
Emmanual Quaye Midodzi and Jane Catherine Appiah

Chapter 1

Introduction

In clinical trials, we are often interested in comparing the distribution of time to failure among several different treatment groups. Two common nonparametric statistical tests used in practice are the logrank statistic (Mantel, 1966; Cox, 1972; Peto and Peto, 1972) and modifications of the Wilcoxon statistic (Gilbert, 1962; Gehan, 1965a; Mantel, 1967; Breslow, 1970; Prentice, 1978). Clinical trials may be prophylactic or therapeutic. Prophylactic trials are conducted in preventive medicine, and the purpose is to asses the effectiveness of preventive treatment. A therapeutic trial is designed to compare a new treatment with the best of the current treatments. In a typical therapeutic trial, patients with similar characteristic are randomly allocated to two groups; one group is given the new treatment and the other, usually called the control is given the current treatment (or placebo). The interest is to see which group does better over a given time period. In a prophylactic trial, one group is given the prophylactic and the other is not. The primary interest in prophylactic trial is to investigate if the protected group has a lower incidence of specific disease than the unprotected. Thus the basic principles for prophylactic and therapeutic trials are similar. We discuss both trials in a single statistical context in this thesis.

In most clinical trials, patients enter the study serially, and are assigned to treatment according to some random mechanism. They are followed until they either fail or the study is terminated. Almost all these studies are designed so that after sufficient amount of data is collected on patient's survival time a single terminal analysis will be made to test whether the failure time distribution is the same among the different treatment groups. However in practice, as well as for ethical considerations, the data should be monitored periodically and if sufficient differences are found between the treatment groups, decision has to be made to stop the study early. It is therefore very important to study the sequential properties of the tests used in analysis of follow-up studies in order that correct and efficient methods are employed in monitoring the data.

In 1969, Breslow provided sequential methods for comparing survival distributions. Recent works examine the more commonly used nonparametric statistics. Jones and Whitehead (1979) have studied the sequential logrank and modified Wilcoxon test. Tsiatis (1981) has derived the group sequential distribution of the logrank score and Slud and Wei (1992) that of the modified Wilcoxon score.

In this thesis, a nonparametric sequential method applied to clinical trials is considered through change detection procedures. Although the literature on change point detection problems is quite extensive, the case of sequential change point detection using randomly censored data has not received much consideration. Liu (1998) has studied change-point detection in distribution for censored data with fixed sample size data. In this thesis, we use different methods to set up sequential statistics by using information contained in a censored data. We discuss the test statistic and its properties under the null hypothesis.

Our discussion will be based on the generalization of Wilcoxon rank statistics which introduces a special anti-symmetric score function. We then generalize it to $\mathcal{U}$-statistics using non-degenerate anti-symmetric kernels under the null hypothesis of no change in survival time distribution.

1.1 The Change-point Problem

Let us assume that we have a sequence of independent continuous random variables $T_1, T_2, ...$ Our interest is to investigate any possible changes in the underlying stochastic mechanism. We wish to test the null hypothesis $H_o : T_i,\ i = 1, 2, ...n,$ are identically distributed with distribution function $F(t)$, against the alternative that after some k observations the distribution function $F(t)$ changes to a different distribution function $F^*(t)$. When we do not know $F(t)$, $F^*(t)$ or the change-point k, this problem is completely nonparametric. The problem is to determine whether a change in the distribution F occurred. This problem is called the change-point problem and has important applications in medical studies where situations often arise entailing investigation of the distribution of patient survival time T (i.e. time after treatment). It is possible that after some time, due to an improved medical method, or changed admission criteria, or some other reason, there could be a change in the distribution of the patient survival time.

Occasionally, the survival time cannot be observed because of the termination of the study or for some other reason. Natural causes of failure of follow up, drop-out or some other factors may happen and the survival time T of the patients may not be completely observed, but censored by a sequence of random variable C with an unknown distribution.

In situations like this, the random variable T is said to be censored on the right if we can only observed $X = min(T, C)$, where X is the observed survival time of the patient and C a random censoring variable or a constant. In most cases, patients will enter the study at any time and receive treatment with one of the several therapies available. Instead of completely observing the sequence of their lifetimes we can observe $(X_1, \delta_1), ..., (X_n, \delta_n)$, where $X_i = \min(T_i, C_i)$ and δ_i equals 0 or 1 according as $T_i > C_i$ or $T_i \leq C_i$. The index i of a patient corresponding to the observation (X_i, δ_i) is the chronological order in which the patient enters the study and receives the treatment.

In most of the published work to date, the mathematical theorems are developed for only completely observed failure time, which, of course, is rare in reality. In this thesis, we give consideration to typical cases of medical experiments, where the data of interest are measurements on time elapsing between the occurrence of two events, i.e., time a patient has spent in a follow up study and the index of the measurement corresponding to the chronological order in which the patient enters the study.

Chapter 2

Review of the Literature

This chapter reviews various Wilcoxon type test statistics that have been developed in the past to address nonparametric problems. We review the case of the two sample problem with complete data or using right censored data. We then look at the nonparametric change-point problem as a modification of the nonparametric two sample problem with particular attention to right censored data.

2.1 U-Statistics

Definition

Let F be a set of distribution functions on $\mathcal{R}$, the set of real measurements. Consider a sequence $X_1, ..., X_n$ of independent random variables, identically distributed with unknown distribution function $F \in F$. Let θ be a functional defined on F. Suppose on the basis of a sample $X_1, ..., X_n$, we wish to estimate $\theta(F)$, and that for a sufficiently large sample size n, there is a function $\psi_n(X_1, ..., X_n)$ such that

$$\theta(F) = E\{\psi_n(X_1, ..., X_n)\} \tag{2.1}$$

for all F in F. Then θ is said to admit an unbiased estimator if and only if there is a function h of k variables such that

$$\theta(F) = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} h(x_1, ..., x_k) dF(x_1) ... F(x_k), \tag{2.2}$$

for all $F \in F$. The unbiased estimator $\psi_n(X_1, ..., X_n) = h(x_1, ..., x_k)$ satisfying (2.1) for all $n \geq k$ is called a regular statistical functional of degree k and h is called the kernel of the functional.

For the sub-sample $x_1, ..., x_k$, define

$$U_k = \frac{1}{k!} \sum h(x_{i1}, ..., x_{ik}) \tag{2.3}$$

where the sum is taken over permutations $(i_1, ...i_k)$ of all sub-samples of size k. Write

$$U_n = \binom{n}{k}^{-1} \sum_{(n,k)} U_k, \tag{2.4}$$

where the sum is over all subsets $1 \leq i_1 < ... < i_k \leq n$ of $\{1, ...n\}$.

The statistic (2.4) is an unbiased estimator of $\theta(F)$ and is called a $\mathcal{U}$-statistic (Hoeffding, 1948). Such statistic has desirable properties as an estimator of the regular functional. The variance of $\mathcal{U}$-statistics based on $i.i.d.$ sample can be usefully expressed in terms of certain conditional expectations. For a survey on the theory of $\mathcal{U}$-statistics, refer to Serfling (1980) and Lee (1990).

2.2 Wilcoxon Type Nonparametric Methods in Survival Analysis

In biomedical research, a clinician may be interested in comparing the ability of two or more treatments to prolong life or maintain health. A laboratory researcher may want to compare the tumor-free time of two or more groups of rats exposed to carcinogens. Almost invariably, the survival time of the different groups vary. These differences can be illustrated by drawing graphs of the estimated survivorship function, but that only gives a rough idea of the difference between the distributions. It does not reveal whether differences are significant or merely chance variations. In clinical trials, only a statistical test will be able to establish any real differences between the different treatment groups.

There are two types of methods in experimental clinical trials: fixed-sample method and sequential method. In fixed-sample methods, the number of patients allocated to the two treatment is fixed before the study begins. In sequential methods, the decision whether to continue taking new patients or new observations is determined by the result accumulated up to that time. In the following subsections, we discuss briefly some methods of fixed-

sample and sequential analysis already widely used based on the modified and generalized Wilcoxon statistic.

2.2.1 Wilcoxon (Mann-Whitney) Statistics

Suppose there are two groups of patients who receive treatment 1 and treatment 2. Let $X_1, ..., X_{n_1}$ and $Y_1, ..., Y_{n_2}$ be the independent samples from the two populations with continuous but unknown distribution functions F_1 and F_2, respectively. The two-sample nonparametric problem is to test

$$H_0 : F_1 = F_2 \; versus \; H_1 : F_1 \neq F_2. \tag{2.5}$$

Let S_1 and S_2 denote the set of subscripts corresponding to the samples from populations 1 and 2, respectively. Let R_k denote the rank of X_k in the combined sample $X_1, ..., X_{n_1}, Y_1, ..., Y_{n_2}$, of size n $(= n_1 + n_2)$, where $k = 1, ..., n$. When there are no ties or censored observations, the Wilcoxon rank sum statistic is

$$W = \sum_{k=1}^{n} R_k I(k \in S_1), \tag{2.6}$$

where I is the indicator function defined by

$$I(k \in S_1) = \begin{cases} 1 & k \in S_1 \\ 0 & k \notin S_1 \end{cases}.$$

The statistic in (2.6) has an asymptotic normal distribution under H_0. The null hypothesis H_0 is rejected if W is too large or too small.

Denote the conditional mean and variance of W under H_0 given the group sample sizes n_1 and n_2 by $E(W|\ H_0, n_1, n_2)$ and $var(W|H_0, n_1, n_2)$ respectively. Then

$$E(W|H_0, n_1, n_2) = \frac{n_1(n_1 + n_2 + 1)}{2} \tag{2.7}$$

and

$$var(W|H_0, n_1, n_2) = \frac{n_1 n_2(n_1 + n_2 + 1)}{12}. \tag{2.8}$$

To test H_0 against H_1, a large sample approximation (say n_1, n_2 about 25 or more) uses the standardized score

$$Z_W = \frac{W - E(W|H_0, n_1, n_2)}{\sqrt{var(W|H_0, n_1, n_2)}} \xrightarrow{D} N(0, 1). \tag{2.9}$$

For a specified significant level α, H_0 is rejected if $|Z_W| \geq Z_{\alpha/2}$, where $Z_{\alpha/2}$ is the upper $100\alpha/2$ percentile of the standard normal distribution.

A simple modification was made in the Wilcoxon statistic to generalize the test further to "tied" observations. For $i = 1, ...n_1$ and $j = 1, ..., n_2$, define the score function

$$U_{ij} = U(X_i, Y_j) = \begin{cases} 1 & if\ X_i > Y_j \\ 0 & if\ X_i = Y_j \\ -1 & if\ X_i < Y_j \end{cases} \tag{2.10}$$

and

$$U = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} U_{ij}. \tag{2.11}$$

The score function (2.10) is the Mann-Whitney sign score function and the statistic U is the Mann-Whitney $\mathcal{U}$-$\mathcal{SUM}$ statistic, which has an asymptotic normal distribution under H_0. The Mann-Whitney test rejects H_0 if U or $|U|$ is too large. The test is conditional on the fact that there are no censored observations in the given sample from each groups. The

conditional mean and variance of U are given by

$$E(U|H_0, n_1, n_2) = 0$$

and

$$var(U|H_0, n_1, n_2) = \frac{n_1 n_2 (n_1 + n_2 + 1)}{3}.$$

Similarly, large sample approximation uses

$$Z_M = \frac{U - E(U|H_0, n_1, n_2)}{\sqrt{var(U|H_0, n_1, n_2)}} \xrightarrow{D} N(0, 1) \tag{2.12}$$

under H_0 and for a specified significant level α, H_0 is rejected if $|Z_M| \geq Z_{\alpha/2}$.

2.2.2 Gehan Test

Gehan test is a generalization of the Wilcoxon-Mann-Whitney test. It assumes that we have n_1, n_2 independent observations with distribution functions $F_1(x)$ and $F_2(y)$. respectively. The conditions on sampling are such that there is a probability that each patient in both groups will be right censored so that each of the samples consist of censored and uncensored observations. It assumes that the conditions leading to the censored observations are the same in the two groups. Ties are permitted among both censored and uncensored observations.

Let $T_1, ..., T_{n_1}$ be the $i.i.d.$ sample from group 1 with $d.f.$ F_1 and $C_{11}, ...C_{1n_1}$ be $i.i.d.$ with $d.f.$ G_1, where C_{1i} is the censoring time associated with T_i. We observe $(X_1, \delta_1), ..., (X_{n_1}, \delta_{n_1})$, where

$$X_i = \min(T_i, C_{1i}) \text{ and } \delta_i = I(T_i < C_{1i}),$$

and I is the indicator variable.

Similarly, for the second group, let $V_1, ..., V_{n_1}$ be the $i.i.d$ sample each with $d.f.$ F_2 and $C_{21}, ...C_{2n_2}$ be $i.i.d$ with $d.f.$ G_2, where C_{2j} is the censoring time associated with V_j. We observe $(Y_1, \varepsilon_1), ..., (Y_{n_2}, \varepsilon_{n_2})$, where

$$Y_j = \min(V_j, C_{2j}) \text{ and } \varepsilon_j = I(V_j < C_{2j}).$$

Gehan (1965) considered an extension of the Wilcoxon-Mann-Whitney test under the more restrictive null hypothesis

$$H_0^* : F_1 = F_2 \; and \; G_1 = G_2. \tag{2.13}$$

For right censored data, Gehan (1965b) proposed the sign score function

$$U_{ij}^* = U^*(X_i, Y_j) = \begin{cases} +1 & if\ (X_i > Y_j, \varepsilon_j = 1) \; or \; (X_i = Y_j, \delta_i = 0, \varepsilon_j = 1) \\ 0 & otherwise \\ -1 & if\ (X_i < Y_j, \varepsilon_i = 1) \; or \; (X_i = Y_j, \delta_i = 1, \varepsilon_j = 0) \end{cases} \tag{2.14}$$

and defined

$$U^* = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} U_{ij}^*. \tag{2.15}$$

Gehan test uses the $\mathcal{U}$-$\mathcal{SUM}$ statistic U^* and rejects H_0^* if U^* or $|U^*|$ is too large. The statistic U^* is asymptotically normally distributed by the theory of the two sample $\mathcal{U}$-statistics under H_0^*. To calculate the critical values, we need to know the moments under the permutation model of U^*. We may consider two cases as follows.

CASE 1: Complete data

Denote the permutation mean and variance of U^* by $E_p(U^*|H_0^*, n_1, n_2)$ and $var_p(U^*|H_0^*, n_1, n_2)$. Then

$$E_p(U^*|H_0^*, n_1, n_2) = 0 = E(U|H_0, n_1, n_2)$$

and

$$var_p(U^*|H_0^*, n_1, n_2) = \frac{n_1 n_2(n_1 + n_2 + 1)}{3} = var(U|H_0, n_1, n_2).$$

Under the null hypothesis H_0^*,

$$Z_G = \frac{U^*}{\sqrt{var_p(U^*|H_0^*, n_1, n_2)}} \xrightarrow{D} N(0,1), \tag{2.16}$$

for large samples. H_0^* is rejected for given significant value α if $|Z_G| \geq Z_{\alpha/2}$. Note that with no censored observations, the Gehan statistic reduces exactly to the Mann-Whitney statistic.

CASE 2: With censoring observation

With censoring observations, Gehan uses permutation under the restrictive null hypothesis. He considered the combined sample $(Z_1, \xi_1), ..., (Z_n, \xi_n)$ of patients, where $n = n_1 + n_2$ is the combined sample size, and considered sampling n_1 patients from n patients labelled with identity $(Z_1, \xi_1), ..., (Z_n, \xi_n)$. Suppose the labels on the n_1 patients are $(X_1, \delta_1), ...(X_{n_1}, \delta_{n_1})$ and the labels on the n_2 patients are $(Y_1, \varepsilon_1), ..., (Y_1, \varepsilon_{n_2})$. Then we have

$$E_p(U^*|H_0^*, n_1, n_2) = 0$$

and $var_P(U^*|H_0^*, n_1, n_2)$ is given on page 206 of Gehan (1965a). Gehan's formula for $var_P(U^*|H_0^*, n_1, n_2)$ is too complicated and introduces computational inconvenience. Man-

tel (1967) considered a computational form for $var_P(U^* \mid H_0^*, n_1, n_2)$ which is much easier to work with, and we discuss this as follows.

Define

$$U_{kl}^* = U^*((Z_k, \xi_k), (Z_l, \xi_l)) \tag{2.17}$$

$$= \begin{cases} +1 & if\ (Z_k > Z_l, \xi_l = 1) \text{ or } (Z_k = Z_l, \xi_k = 0, \xi_l = 1), \\ 0 & otherwise, \\ -1 & if\ (Z_k < Z_l, \xi_k = 1) \text{ or } (Z_k = Z_l, \xi_k = 1, \xi_l = 0), \end{cases}$$

and

$$U_k^* = \sum_{l \neq k:\, l=1}^{n} U_{kl}, \tag{2.18}$$

and

$$U^* = \sum_{k=1}^{n} U_k^* I(k \in S_1), \tag{2.19}$$

where S_1 is the set of integers corresponding to the subscripts of observations from population 1. U^* is the Mantel statistic which is equal to the Gehan statistic (because $U_{k'l'}^* = -U_{l'k'}^*$ so if $k', l' \in S_1$, they cancel each other out in the sum). Mantel considered calculating the permutation distribution of U^* as follows. Suppose we are given $U_1^*, \ldots U_n^*$, where $n = n_1 + n_2$. Under H_0^*, sample n_1 of U_k^*'s without replacement and calculate U^* using (2.19), the sum of these n_1 observations. Using results from sampling from finite populations, Mantel showed that

$$E_p(U^* | H_0^*, n_1, n_2) = 0$$

and

$$var_P(U^*|H_0^*, n_1, n_2) = \frac{n_1 n_2}{(n_1 + n_2 - 1)(n_1 + n_2)} \sum_{k=1}^{n} (U_k^*)^2. \tag{2.20}$$

For large sample approximation, we use

$$Z_G = \frac{U^*}{\sqrt{var_p(U^*|H_0, n_1, n_2)}} \xrightarrow{D} N(0,1), \tag{2.21}$$

under H_0^* and H_0^* is rejected for a given significance level value α if $|Z_G| \geq Z_{\alpha/2}$.

2.3 The AMOC Change-point Problem

The two sample problem (2.5) is to test

$$H_0 : F_1 = F_2 \; versus \; H_1 : F_1 \neq F_2.$$

In a fashion similar to the Gehan or Mantel procedure, we could consider the combined sample of lifetimes $T_1, ..., T_n$, and test the hypothesis

$$H_o : T_1, ..., T_n \text{ i.i.d } F(t) \tag{2.22}$$

against the alternative

$$H_a : \text{ there exists some } k \in \{1, 2, ..., n-1\} \text{ such that}$$

$$T_1, ..., T_k \text{ i.i.d } F^{(1)}(t) \text{ and } T_{k+1}, ..., T_n \text{ i.i.d } F^{(2)}(t), \tag{2.23}$$

and $F^{(1)}(t) \neq F^{(2)}(t)$ for some t, where the distribution functions F, $F^{(1)}$, $F^{(2)}$ and k are unknown and n $(= n_1 + n_2)$ is the size of combined sample. The aim is to find a detecting procedure in order to raise an alarm that a change occurred in the sequence. If we are interested in only one change in the sequence, then the problem is called At Most-One-Change (AMOC) change-point problem. Therefore we can think of the AMOC change-

point problem as a series of tests of two sample problems. The change-point problem has many applications in a variety of areas such as the surveillance of a system, monitoring the quality of production process, etc.

In medical applications, since we cannot completely observe the survival times $T_1, ..., T_n$, we instead observe $(X_1, \delta_1), ..., (X_n, \delta_n)$, where

$$X_i = \min(T_i, C_i) \text{ and } \delta_i = I(T_i < C_i),\ i = 1, ..., n. \tag{2.24}$$

The variables T_i and C_i are independent and δ indicates whether T has been censored or not. In the AMOC change-point problem, under the assumption that $C_1, ..., C_n$ are identically distributed as G, we assume that the random variables $X_1, ..., X_n$ have the same distribution function $\mathcal{H}$ under H_o, where

$$\begin{aligned} \mathcal{H}(x) &= P_{H_o}\{X_1 \leq x\} \\ &= 1 - (1 - F(x))(1 - G(x)). \end{aligned} \tag{2.25}$$

If H_o is not true, then a change in the distribution function F results in a change in the distribution function $\mathcal{H}$. In view of this, we might think that methods designed for detecting a change from a sequence of completely observable data could be applied to X's as well to detect changes in $\mathcal{H}$. However, this procedure do not incorporate the information contained in the δ's and therefore leads to inefficient results.

Under H_a, let $X_1, .., X_k$ have the same distribution function $\mathcal{H}^{(1)}$ and $X_{k+1}, ..., X_n$ have the same distribution function $\mathcal{H}^{(2)}$, where

$$\begin{aligned} \mathcal{H}^{(1)}(x) &= P_{H_a}\{X_1 \leq x\} \\ &= 1 - (1 - F^{(1)}(x))(1 - G(x)) \end{aligned} \tag{2.26}$$

and

$$\begin{aligned} \mathcal{H}^{(2)}(x) &= P_{H_a}\{X_n \leq x\} \\ &= 1-(1-F^{(2)}(x))(1-G(x)). \end{aligned} \tag{2.27}$$

If H_o is not true, then $\mathcal{H}^{(1)}(x) \neq \mathcal{H}^{(2)}(x)$ for some $x = \min(t, c)$.

Let

$$Z_i = (X_i, \delta_i),\ i = 1, ..., n. \tag{2.28}$$

Note that $Z_1, ..., Z_n$ are a sequence of two-dimensional independent random vectors. We define the score function for comparing two observations Z_i and Z_j in the sample as

$$h(Z_i, Z_j) = \begin{cases} +1 & if\ (X_i > X_j, \delta_j = 1)\ or\ (X_i = X_j, \delta_i = 0, \delta_j = 1) \\ 0 & otherwise \\ -1 & if\ (X_i < X_j, \delta_j = 0)\ or\ (X_i = X_j, \delta_i = 1, \delta_j = 0). \end{cases} \tag{2.29}$$

This score function is equivalent to that proposed by Gehan and Mantel as a modification to generalize the Wilcoxon-Mann-Whitney statistic for two samples test with right censored data. Both Gehan (1965a) and Mantel (1967) proposed their nonparametric test for comparing two samples under the additional assumption of identical censoring distributions. Gehan and Mantel's form of the statistic for comparing the samples $\{Z_1, ..., Z_k\}$ and $\{Z_{k+1}, ..., Z_n\}$ is

$$U_k = \sum_{i=1}^{k} \sum_{j=k+1}^{n} h(Z_i, Z_j). \tag{2.30}$$

By convention, if an uncensored observation $Z_i = (X_i, \delta_i = 1)$ and a censored observation $Z_j = (X_j, \delta_j = 0)$ are tied, (i.e., $X_i = X_j$) we consider the uncensored X_i to occur just before the censored observation X_j, i.e., we break the tie by considering $X_i <$

X_j. On the other hand, since $X_1, ..., X_n$ are independent continuous random variables, the probability of ties is zero, so for simplicity and convenience in calculations, we assume no ties without loss of generality. The score function (2.29) then simplifies as

$$h(Z_i, Z_j) = I(X_i > X_j, \delta_j = 1) - I(X_i < X_j, \delta_i = 1). \tag{2.31}$$

For a kernel h of degree two, we say h is symmetric if

$$h(x, y) = h(y, x),$$

and anti-symmetric if

$$h(x, y) = -h(y, x).$$

We assume that the kernel h is of bounded variation as a function of x or of y with the other variable fixed at any value. For the sequence $X_1, ..., X_n$ of independent random variables, and with a kernel h of second degree, we assume that

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

Define

$$h_1(t) = Eh(X_1, t) - Eh(X_1, X_2)$$

and

$$h_2(t) = Eh(X_n, t) - Eh(X_{n-1}, X_n).$$

Condition $(*)$ implies that $Eh_1^2(X_1) < \infty$ and $Eh_2^2(X_n) < \infty$. The kernel h is said to be non-degenerate if

$$Eh_1^2(X_1) > 0 \text{ and } Eh_2^2(X_n) > 0, \qquad **$$

otherwise, we have the degenerate case.

$\mathcal{U}$-statistics can be used to test for change in the distribution of a sequence of random variables. This thesis considers $\mathcal{U}$-statistic based tests using anti-symmetric and non-degenerate kernels. The score function (2.31) is the kernel function of second degree and is anti-symmetric since $h(Z_i, Z_j) = -h(Z_j, Z_i)$. If Z_1, Z_2 and Z_{n-1}, Z_n have the same distribution function, then we have

$$Eh(Z_1, Z_2) = 0 \text{ and } Eh(Z_{n-1}, Z_n) = 0 \tag{2.32}$$

and

$$0 < E\{h^2(Z_i, Z_j)\} < \infty, \text{ all } i < j. \tag{2.33}$$

Define

$$h_1(Z_1) = Eh((Z_1, Z_2)|Z_1) \tag{2.34}$$

and

$$\sigma^2 = var(h_1(Z_1)) = Var(Eh(Z_1, Z_2)|Z_1) > 0. \tag{2.35}$$

The function defined in (2.34) is the projection of $\mathcal{U}$-statistics and (2.33) ensures that σ^2 is finite. Using the definitions in (2.34) and (2.35), we have

$$Eh_1(Z_1) = E\{E\{h(Z_1, Z_2)|Z_1\}\} = Eh(Z_1, Z_2) = 0 \tag{2.36}$$

and

$$E(\frac{1}{\sigma}h_1(Z_i))^2 = 1, i = 1, ..., n, \tag{2.37}$$

where $h_1(Z_1), ..., h_1(Z_n)$ are $i.i.d.$ random variables under H_o.

Following Mantel (1967), we can write

$$\mathcal{U}_i' = \sum_{j=1}^{n} h(Z_i, Z_j), \quad i = 1, ..., n, \tag{2.38}$$

and since h is anti-symmetric, we can write

$$\sum_{i=1}^{k}\sum_{j=1}^{n} h(Z_i, Z_j) = \sum_{i=1}^{k}\sum_{j=k+1}^{n} h(Z_i, Z_j),\ 1 \le k \le n, \tag{2.39}$$

where $h(Z_i, Z_j)$'s cancel each other out in the sum in the left hand side for $1 \le i, j \le k$. Therefore, from (2.30) we have

$$U_k = \sum_{i=1}^{k}\sum_{j=k+1}^{n} h(Z_i, Z_j) = \sum_{i=1}^{k} \mathcal{U}_i'. \tag{2.40}$$

The statistic U_k is a form of a $\mathcal{U}$-$\mathcal{SUM}$ statistic with an anti-symmetric kernel h. Furthermore, since $h(Z_i, Z_j)$ is anti-symmetric,

$$\sum_{i=1}^{n} \mathcal{U}_i' = \sum_{i=1}^{n}\sum_{j=1}^{n} h(Z_i, Z_j) = 0\,, \tag{2.41}$$

and

$$\sum_{i=i}^{n} \{\mathcal{U}_i'/(\sum_{j=1}^{n} [\mathcal{U}_j']^2)^{1/2}\}^2 = 1. \tag{2.42}$$

Note that (2.41) and (2.42) are the sample representatives of (2.36) and (2.37).

Let $\widetilde{\mathcal{H}}(t)$ be the sub-distribution function of $X_1, ..., X_n$ defined by

$$\begin{aligned} \widetilde{\mathcal{H}}(t) &= P_{Ho}\{X_1 \le t\} \\ &= \int_{-\infty}^{t} (1 - G(u))dF(u). \end{aligned} \tag{2.43}$$

Lemma 2.1 Under H_o,

$$\sigma^2 = Var(Eh(Z_1, Z_2)|Z_1) = \int (1 - \mathcal{H}(t))^2 d\widetilde{\mathcal{H}}(t) > 0 \tag{2.44}$$

and

$$\frac{1}{n^3} \sum_{i=1}^{n} [\mathcal{U}_i']^2 \xrightarrow{a.s.} \sigma^2, \tag{2.45}$$

where $\mathcal{H}(t)$, $\widetilde{\mathcal{H}}(t)$ and $\mathcal{U}_i'$ are defined in (2.25), (2.43) and (2.38), respectively.

For the proof of this lemma, see Miller (1981). Lemma 2.1 provides a consistent estimator of σ^2; hence we can use (2.48) to estimate σ^2 from the data using the estimator

$$\widehat{\sigma}^2 = \frac{1}{n^3} \sum_{i=1}^{n} [\mathcal{U}_i']^2. \tag{2.46}$$

Csörgő and Horváth (1997) showed that, for large samples, $n^{-3/2}\sigma^{-1}U_k$ can be approximated by a Brownian bridge under H_o; hence large sample critical values can be obtained for different functionals of this process.

Theorem 2.1. Under H_o,

$$U_{o,k} = \max_{1 \le k \le n} \frac{|U_k|}{n^{3/2}\sigma} \xrightarrow{D} \sup_{0 \le \lambda \le 1} |B(\lambda)|, \quad k = 1, 2, ..., n, \tag{2.47}$$

where $B(\lambda)$ denotes a Brownian bridge.

The distribution of the right-hand side of (2.47) is well known, so tests of significance can be performed at any specified significance level α. See Gombay and Liu (2000) for details of the proof and some applications of this theorem.

From (2.45), we write $\sum_{i=1}^{n} [\mathcal{U}_i']^2 \approx n^3\sigma^2$. Let

$$\widetilde{U}_{o,k} = \max_{1 \le k \le n} \frac{U_k}{\left(\sum_{i=1}^{n} [\mathcal{U}_i']^2 \right)^{1/2}}. \tag{2.48}$$

Then $\widetilde{U}_{o,k}$ is one-sided version of $U_{o,k}$. Here, it is proposed to reject H_o for large values of $\widetilde{U}_{o,k}$. We can estimate $U_k/(n^{3/2}\sigma)$ using

$$\widehat{U}_{o,k} = \frac{U_k}{\left(n^3\widehat{\sigma}^2\right)^{1/2}} = \frac{U_k}{n^{3/2}\widehat{\sigma}}, \tag{2.49}$$

where $\widehat{\sigma}^2$ is defined in (2.46).

Test: When we want to detect a change in the underlying stochastic mechanism of the data (Z_k, δ_k), given that all past and future observations are available, we use the statistic (2.47) or (2.49). For (2.47) the critical values are determined from the well known identity of Kolmogorov (1933):

$$P\left\{\sup_{0<\lambda<1}|B(\lambda)|>c\right\}=2\sum_{i=1}^{\infty}(-1)^{i-1}\exp(-2i^2c^2),\ c>0. \tag{2.50}$$

Let c_α denote the $(1-\alpha)-$quantile of the distribution of $\sup_{0<\lambda<1}|B(\lambda)|$. Then the test for the change-point problem with a fixed sample size n, is defined as

$$I\left(\max_{1\le k<\le n}|U_{o,k}|>c_\alpha\right)=\begin{cases}1 & reject\ H_o\\ 0 & fail\ to\ reject\ H_o.\end{cases} \tag{2.51}$$

Hence, based on the process $\{U_{o,k}\}_{1<k<n}$, conclude that change occurs in the distribution function F at k, $2\le k\le n$, if

$$I\left(\max_{1\le k<\le n}U_{o,k}>c_\alpha\right)=1. \tag{2.52}$$

Some critical values c_α, for $\alpha=0.001$, 0.01, 0.05 and 0.10, using identity (2.50) are 1.78, 1.63, 1.36, and 1.22, respectively.

Chapter 3

Sequential Change Detection

In clinical studies, it is important to follow interim results closely and continuously as data become available. In this chapter, we construct sequential analyses for experimental clinical studies allowing for hypothesis testing at each inspection point in a continuous time when the data gathered are right censored failure time data, i.e.,we formulate nonparametric sequential test statistics for the change detection problem. The tests we construct are most useful for examining the identical distribution assumption underlying the survival distribution on the random vector $Z = (X, \delta)$.

3.1 Formulation of Sequential Change Detection Problem

Sequential change detection problems involve consideration of a random sequence of lifetimes $X_1, X_2, ..., X_k$, which may not be completely observable due to the existence of corresponding censoring variables $C_1, C_2, ...C_k$, where k is a random integer. Accrued data are analyzed after each k^{th} inspection and a decision is made whether to reject the null hypothesis or continue sampling. If experimentation is terminated at the k^{th} stage, where $k = 1, 2, ...$ then we observe the data $\{(Z_i, \delta_i) : 1 \leq i \leq k\}$, where δ_i equals 0 or 1 according as Z_i is a censoring time or true lifetime. This method and censoring scheme may be used in any experimentation in which ethical or economical reasons require that observations be ceased at the earliest possible stage if the current accumulated data warrant a clear statistical decision. In this section, we look at the problem of constructing a nonparametric sequential test for change detection in the distribution of right censored data.

When data come as a sequence of independent observations $T_1, T_2, ...,$ we want to test the hypothesis

$$H_o' : T_i,\ i = 1, 2, ...,\ \text{have identical distributions } F, \tag{3.1}$$

against the alternative ,

$$H_a' :\ T_i,\ i = 1, ..., k,\ \text{have d.f } F^{(1)} \text{ and } T_i,\ i = k+1, k+2, ...,\ \text{have d.f } F^{(2)}, \tag{3.2}$$

where $F^{(1)}(t) \neq F^{(2)}(t)$ for some $t \in R$, and k is an unknown positive integer. This test advocates a serial testing decision. The difference between the sequential change detection

test and the change-point problem discussed earlier in section 2.3 is that, in the case of the latter, we have a full set of data $\{(Z_i, \delta_i) : 1 \leq i \leq n\}$ of size n and want to detect at most one change at some point k, in the underlying distribution. Under the present problem, we look at the value of the test statistic after k observations and decide whether a change has occurred in distribution at that point. In sequential testing, we may not have the full data set, but it is desirable to formulate it in terms of some truncation point, N. Observations come in one by one and a decision is taken after a random number of observations are observed. The sequential test stops after one abrupt change. We refer to Csörgő and Horváth (1997) for information on the change point detection literature, and to Gombay (2000b) for new developments on sequential change detection.

In this thesis, we wish to design a nonparametric sequential test for H_a', where F, $F^{(1)}$, $F^{(2)}$, and G are unknown. This test will use $\mathcal{U}$-statistics with anti-symmetric kernels $h(x, y)$ which are better suited for change detection than symmetric kernels. In Chapter 2, $\mathcal{U}$-$\mathcal{SUM}$ statistics were examined for full data when there was a change at k and the parameter n is the sample size. Change-point test and sequential change detection can be compared. The new asymptotic analysis gives the result that in many situations, a sequential test analysis is more powerful than a fixed sample test for the AMOC change-point problem.

3.2 Sequential Change Detection Procedures

Sequential analysis was originally developed by Wald (1947) and applied to clinical trials by others. For example, Armitage (1975) gave a procedure for repeated testing of

data. Armitage's method assumes that the only decision to be made is whether the trial should continue or be terminated because one of the groups is responding significantly better, or worse, than the other. This classical sequential decision rule is called an "open plan" because there is no guarantee of when a decision to terminate will be reached. Strict adherence to an open plan will mean that the study could not be terminated at a fixed sample size, N. For this reason, very few clinical trials use the "open" or classical sequential design because there is no certainty of ever reaching a point at which the trial would be stopped. The method also requires data to be paired, one observation from each group. In many instances, the pairing of participants is not appealing because the paired participants may be different and may not be well matched in important prognostic variables.

Because of these limitations of classical sequential methods, adhoc rules have been suggested by others that attempt to ensure a conservative interpretation of interim results. It is desirable to define procedures that are truncated after some number of observations (say, N) and use a constant critical value, c_α, at each interim look. Hypotheses have to be formulated in terms of these truncation points. Accrued data at the k^{th} inspection is of the form $\{(Z_i, \delta_i) : i = 1, 2, ..., k\}$ (where Z_i's are defined in (2.28)).

Using the anti-symmetric non-degenerate kernel h, denote

$$\mathcal{U}_i^* = \sum_{j=1}^{i-1} h(Z_i, Z_j), \quad i = 2, ..., k. \tag{3.3}$$

We define

$$U_k^* = \sum_{i=1}^{k} \mathcal{U}_i^* = \sum_{1 \le j < i \le k} h(Z_i, Z_j), \quad k = 2, 3, ..., \tag{3.4}$$

which is a sequence of $\mathcal{U}$-$\mathcal{SUM}$ statistics. Our sequential analysis considers the distribution of different functional of the process $\{U_k^*\}$ under H_o'. We use the test process

$\{k^{-3/2}U_k^*/\sigma\}$ whose limiting distribution is the well known Ornstein-Uhlenbeck process $\Gamma(t)$, i.e., a zero mean stationary Gaussian process with covariance function $\exp(-|s-t|)$. To ensure that the variance of the process $\{k^{-3/2}U_k^*\}$ is bounded at each inspection, assume that

$$0 < E\{h^2(Z_i, Z_j)\} < \infty \text{ for all } i < j. \tag{3.5}$$

Denote

$$\sigma^2 = var(h_1(Z_1)) = \int (1 - H(t))^2 d\tilde{H}(t) > 0, \tag{3.6}$$

where

$$h_1(Z_k) = E\{h(Z_j, Z_k)|Z_j\}, \text{ for } 1 \leq j < k \text{ and } k = 2, 3.... \tag{3.7}$$

Let

$$U_{i,k} = \sum_{j=1}^{k} h(Z_i, Z_j), \ 1 \leq i \leq k \text{ and } k = 2, 3, ... \tag{3.8}$$

From (2.41) we have $\sum_{i=1}^{k} U_{i,k} = 0$. Under Lemma 2.1 and for a reasonable size k, we can write

$$\frac{1}{k^3} \sum_{i=1}^{k} U_{i,k}^2 \xrightarrow{a.s} \sigma^2, \tag{3.9}$$

where σ^2 is the variance of our process $\{k^{-3/2}U_k^*\}$. We define our sequential statistic at each inspection point as follows:

$$stat(k) = \frac{k^{-3/2}U_k^*}{\sqrt{\frac{1}{k^3}\sum_{i=1}^{k} U_{i,k}^2}} = \frac{U_k^*}{\sqrt{\sum_{i=1}^{k} U_{i,k}^2}}. \tag{3.10}$$

Our analysis uses different modifications of (3.10). Before we look at the sequential testing procedures, let us first discuss the limiting distribution of the test statistic $\max_{1<k<N} U_k^*$ under the null hypothesis based on anti-symmetric non-degenerate kernel. The following

theorems are needed to establish the asymptotic distribution of the test statistic under the null hypothesis of no change in distribution.

Theorem 3.1. Under H_o', if (3.5) holds and

$$E|h(Z_i, Z_j)|^r < \infty \text{ for some } r > 2, \tag{3.11}$$

then

(*i*)

$$U_{o,k}^{(1)} = \max_{1 \le k \le N} \frac{\sqrt{3}U_k^*}{k^{3/2}\sigma} \xrightarrow{D} \sup_{0 \le t \le T} \Gamma(t), \tag{3.12}$$

where $\Gamma(t)$ is the Ornstein-Uhlenbeck process and $T = 3\log_e(N)$: and

(*ii*)

$$U_{o,k}^{(2)} = \max_{1 \le k \le N} \frac{\sqrt{3}|U_k^*|}{k\sigma\sqrt{N}} \xrightarrow{D} \sup_{0 \le t \le 1} |W(t)|, \tag{3.13}$$

where $\{W(t), t \ge 0\}$ is a Wiener process.

Furthermore, if we let $a(T) = (2\log(T))^{1/2}$ and $b(T) = 2\log(T) + \frac{1}{2}\log\log(T) - \frac{1}{2}\log_e(\pi)$, then

$$\lim_{N \to \infty} P\left\{ a(T)\frac{\sqrt{3}}{\sigma} \max_{1 \le k \le N} k^{-3/2}|U_k| \le t + b(T) \right\} = \exp(-2e^{-t}), \tag{3.14}$$

for all t. Sequential processes based on this theorem and its proof are discussed in Gombay (2000b). For a sequential test truncated at N, the critical value c for a level α test based on Theorem 3.1 can be obtained from the following theorem.

Theorem 3.2.

(i) (Vostrikova, 1981): For all $T > 0$,

$$P\left\{\sup_{0\le t\le T}|\Gamma(t)| > c\right\} \approx \frac{c^d \exp(-c^2/2)}{2^{d/2}\Gamma(d/2)}\left\{T - \frac{d}{c^2}T + \frac{4}{c^2} + \mathbf{O}\left(\frac{1}{c^4}\right)\right\} \tag{3.15}$$

as $x \longrightarrow \infty$.

(ii)

$$P\left\{\sup_{0\le t\le 1}|W(t)| \le c\right\} = \frac{4}{\pi}\sum_{k=0}^{\infty}\frac{(-1)^k}{2k+1}\exp(-\frac{\pi(2k+1)^2}{8c^2}), \tag{3.16}$$

where $\{W(t), t \ge 0\}$ is a Wiener process.

For the analysis based on the process $\{k^{-3/2}U_k^*\}$, the critical values of the test of H_o' can be determined using Theorem 3.2.

In our application, the critical value $c_\alpha^1(N)$ for a test based on (i) of Theorem 3.1 depends on both the level of significance α and on the truncation point N. Critical values c_α^2 of a level α test using (ii) of Theorem 3.1 can be obtained from the distribution of $\sup_{0\le t\le 1}|W(t)|$ in (ii) of Theorem 3.2.

3.3 Sequential Tests

(1) For a given data set, suppose the absolute value of the sequential statistic is denoted by $|\widehat{U}_{o,k}|$. Calculate in sequence $|\widehat{U}_{o,k}|$, $k = 5, 6, ..., N$, at each inspection, using (i) or (ii) of Theorem 3.1. We start at $k = 5$ or 10 only, because of the estimation of σ.

(2) If we use a test based on (i) of Theorem 3.1, then the critical values $c_\alpha^1(N)$ are calculated using (3.15), where $T = 3\log_e(N)$ and $d = 1$. Compute the LHS for different

values of c, so that the value of $P\left\{\sup_{0\le t\le T}|\Gamma(t)|>c\right\}$ is between 0 and 1, and then take the values c which correspond to some selected (1-α)-quantile, say, 0.90, 0.95, 0.99, and 0.999. These values of c are the critical values $c_\alpha^1(N)$ for α levels, say, 0.10, 0.05, 0.01, and 0.001. Similarly, the critical values for a test based on (ii) of Theorem 3.1 can be calculated using (ii) of Theorem 3.2.

(3) Define a sequential decision rule as follows:

$$I\left(|\widehat{U}_{o,k}|>c_\alpha\right)_{k=5,6,\dots}=\begin{cases}1 \text{ and } k\le N & \text{reject } H_o'\\ 0 \text{ and } k<N & \text{fail to reject } H_o' \text{ and continue sampling}\end{cases} \tag{3.17}$$

The stopping time of the sequential test can be defined as

$$k^*=\min\left\{k<N: I\left(\max_{1\le k\le N}|\widehat{U}_{o,k}|\ge c_\alpha\right)_{k=5,6,\dots}=1\right\}. \tag{3.18}$$

Test 1. Let

$$\widetilde{U}_{o,k}^{(1)}=\max_{1\le k\le N}\frac{\sqrt{3}U_k^*}{\left(\sum_{i=1}^{k}U_{i,k}^2\right)^{1/2}}. \tag{3.19}$$

Then $\widetilde{U}_{o,k}^{(1)}$ is equivalent to $U_{o,k}^{(1)}$. We estimate $\sqrt{3}U_k^*/(k^{3/2}\sigma)$ by

$$\widehat{U}_{o,k}^{(1)}=\frac{\sqrt{3}U_k^*}{k^{3/2}\widehat{\sigma}}, \tag{3.20}$$

or in practice, we use

$$|\widehat{U}_{o,k}^{(1)}|=\frac{\sqrt{3}|U_k^*|}{k^{3/2}\widehat{\sigma}}, \tag{3.21}$$

where $\widehat{\sigma}_k^2=\frac{1}{k^3}\sum_{i=1}^{k}U_{i,k}^2$ is the estimator of σ^2.

At the first $k=5,6,\dots$, where

$$I\left(|\widehat{U}_{o,k}^{(1)}|>c_\alpha^{(1)}(N)\right)_{k=5,6,\dots}=1, \tag{3.22}$$

stop and declare a change has occurred in distribution. This is an evidence for the alternative hypothesis.

Test 2. This test is designed with a higher sensitivity to detect small deviations from H'_o. Let

$$\widehat{U}^{(2)}_{o,k} = \frac{\sqrt{3}U^*_k}{k\widehat{\sigma}\sqrt{N}}, \tag{3.23}$$

where $\widehat{\sigma}^2$ is as defined in Test 1. Reject H'_o if

$$I\left(\max_{1\leq k\leq N} |\widehat{U}^{(2)}_{o,k}| > c^{(1)}_\alpha\right)_{k=5,6,\ldots} = 1. \tag{3.24}$$

Otherwise, do not reject H'_o.

The test using critical values from (3.14) gives a good conservative approximation for Test 1. Using this test, stop and declare a change in distribution at the first k, when

$$I\left(\frac{\sqrt{3}|U^*_k|}{k^{3/2}\widehat{\sigma}}. \geq c^{(1)}_\alpha(T)\right) = 1, \tag{3.25}$$

where

$$c^{(1)}_\alpha(T) = a^{-1}(T)(-\log_e(-\frac{1}{2}\log_e(1-\alpha)) + b(T)) \tag{3.26}$$

is obtained from (3.14), with $T = 3\log_e(N)$, $b(T)$ and $a(T)$ as defined in Theorem 3.1, and α is the level of significance. The theories to develop this test can be found in Gombay (2000b).

In Chapter 4, we apply these tests and compare their relative performances at the usual cut-off significance levels 0.1, 0.05, 0.01, and 0.001.

Chapter 4

Implementation of Sequential Method

In sequential testing, we reject H_o' the first time the value of the test statistic exceeds some critical value at a chosen significance level. In this chapter, we first calculate the critical values c_α using (3.16) and (3.26) for some selected significance levels and truncation points. We then investigate the relative performance of Test 1 and Test 2 with a series of simulations. We also compare the power of our test with that of the AMOC procedure. Finally, we demonstrate the application of the sequential procedure on two real and well-known data sets: the Stanford Heart Transplant data and the Radiation Oncology Group data. These two illustrative data sets are typical examples of prophylactic and therapeutic trials, respectively. Note that both Stanford Heart Transplant Data and Radiation Oncology Group data have been used many times in the literature, e.g., Turnbull, Brown and Hu (1974), Crowley and Hu (1977), Kalbfleisch and Prentice (1980), Liu (1998), Gombay and Liu (2000).

4.1 Tables of Critical Values for Selected Significance Levels

Let c_α denote the critical value at significance level α, and N, the sample size in cases where the full data are available; otherwise, it represents the truncation point of the sequential observation. By using (3.16) and (3.26), we compute the critical values c_α for Test 1 and Test 2 as discussed in Chapter 3. We will consider the truncation points $N = 50$, 100, 200, 500, and 1000. The results are reported in Table 4.1 below. It is clear from the table that, whereas Test 2 statistic depends on the truncation point, its critical values do not.

Table 4.1 *Critical Values for the Sequential Tests*

N	α	Test 1	Test 2
50	0.10	3.4908	1.9597
	0.05	3.8151	2.2414
	0.01	4.5495	2.8070
	0.001	5.5891	3.4808
100	0.10	3.5370	1.9600
	0.05	3.8510	2.2414
	0.01	4.5623	2.8070
	0.001	5.5691	3.4808
200	0.10	3.4430	1.9600
	0.05	3.7490	2.2414
	0.01	4.4421	2.8070
	0.001	5.4230	3.4808
500	0.10	3.4913	1.9600
	0.05	3.7889	2.2414
	0.01	4.4627	2.8070
	0.001	5.4165	3.4808
1000	0.10	3.5229	1.9600
	0.05	3.8153	2.2414
	0.01	4.4773	2.8070
	0.001	5.4143	3.4808

4.2 Simulation Study

In this section, we investigate the power of our proposed sequential change detection procedure. First, we carried out simulations to compare the power and stopping time of Test 1 and Test 2. We then looked at the relative performance of both tests with respect to the AMOC change-point problem (see Liu, 1998). We performed $M = 5000$ simulations in each case by Monte Carlo. For the sequence of independent observations $T_1, T_2, ...,$ the hypotheses of interest are those of (3.1) and (3.2). We assumed that the survival of each individual satisfies the memoryless property (i.e., the survival probability of each individual surviving beyond any time t in the trial is independent of entry time or how long the individual has already been in the trial). That is,

$$P(T > r + t | T > r) = P(T > t),$$

for all $t > 0$ and $r > 0$ (may be the time an individual enters the trial). Hence the simulated random variables T_i's have an exponential distribution. Assume that under the null hypothesis, T_i, $i = 1, 2, ...,$ have identical distribution $F \equiv exp(\mu)$. Under the alternative, let T_i, $i = 1, ..., k$ have $d.f.$ $F_1 \equiv exp(\mu_1)$ before change at some k and T_i, $i = k + 1, k + 2, ...,$ have $d.f.$ $F_2 \equiv exp(\mu_2)$ after change. Also, the simulated censoring variables C_i's are assumed to have exponential distribution, with $d.f.$ $G \equiv exp(\mu_c)$.

Simulation of exponential random variable T:

Note that for exponential random variable T, the cumulative distribution function is given by

$$F_T(t) = \int_0^t \lambda e^{-\lambda y} dy = 1 - e^{-\lambda t} = u(say) \in [0,1], \quad t \geq 0.$$

Generate uniform random numbers u between 0 and 1. Set $u = 1 - e^{-\lambda t}$. Then we have

$$t = -\frac{1}{\lambda} \ln(1-u),$$

where $\lambda = \frac{1}{\mu}$. We then obtained the simulated data $Z_i = (X_i, \delta_i)$ and values of the score function $h(Z_i, Z_j)$ using (2.24) and (2.31), respectively.

In all cases, we considered $\mu_1 = 1.0$ and $\mu_c = 3.0$. Note that, in case of the sequential test, the hypothesis has to be formulated in terms of some truncation point (say, N). We considered $N = 50, 100, 200, 500, 1000$, $\mu_2 = 1.5, 2.0, 2.5, 3.0, 3.5$, and $\alpha = 0.10, 0.05, 0.01, 0.001$. Suppose τ denote the point at which the change occurred in the distribution of T_i's i.e., $\tau = \min\{k > 0 : F^{(2)}(t) \neq F^{(2)}(t)\}$, and $k^* = \min\{k < N : I(\max|\widehat{U}_{o,k}| \geq c_\alpha)\}$ defines the stopping time. Then $k^* > \tau$ and the mean value function of the process $\{k^{-3/2}U_k^*\}$ takes the maximum of its absolute value at $k^* = 2\tau$ (see Gombay, 2000b). As the variance is bounded, the sequential process has the best chance to detect changes when the number of observations before and after change are approximately the same. We first fixed the change at $\tau = N/2$. The simulation results are reported in the following tables.

Suppose $\theta = |\mu_2 - \mu_1|$. Let $\mathcal{P}(\theta|\tau, N)$ denote the power function of θ, given τ and N. Let $\mathcal{P}_1$ and $\mathcal{P}_2$ denote the power function for Test 1 and Test 2, respectively. From the

simulated results, $\mathcal{P}(\theta|\tau = N/2, N)$ is a monotonic increasing function of θ for both Test1

Table 4.2 *Simulated Power of Test 1*

N	τ	μ_2	$\alpha = 0.10$ *power*	$\alpha = 0.05$ *power*	$\alpha = 0.01$ *power*	$\alpha = 0.001$ *power*
50	25	1.5	0.1980	0.1380	0.0612	0.0258
		2.0	0.2382	0.1860	0.0858	0.0366
		2.5	0.3402	0.2326	0.0942	0.0428
		3.0	0.4404	0.3226	0.1002	0.0554
		3.5	0.5654	0.4066	0.1528	0.0668
100	50	1.5	0.2454	01718	0.0684	0.0196
		2.0	0.4116	0.2968	0.1182	0.0234
		2.5	0.6178	0.4966	0.2470	0.0448
		3.0	0.7712	0.6626	0.4114	0.1182
		3.5	0.7638	0.6756	0.4138	0.1104
200	100	1.5	0.3946	0.2906	0.1306	0.0300
		2.0	0.6906	0.5870	0.3534	0.1080
		2.5	0.8896	0.8188	0.6504	0.3186
		3.0	0.9712	0.9442	0.8580	0.5728
		3.5	0.9928	0.9862	0.9492	0.8088
500	250	1.5	0.6234	0.5460	0.4429	0.3783
		2.0	0.9624	0.9332	0.8382	0.7644
		2.5	0.9986	0.9970	0.9912	0.9320
		3.0	0.9998	0.9998	0.9990	0.9952
		3.5	1.0000	1.0000	1.0000	0.9998
1000	500	1.5	0.7842	0.6582	0.5551	0.4523
		2.0	0.9942	0.9451	0.9010	0.8467
		2.5	1.0000	0.9998	0.9945	0.9736
		3.0	1.0000	1.0000	1.0000	1.0000
		3.5	1.0000	1.0000	1.0000	1.0000

Table 4.3 *Estimated Average Stopping (EAS) Times of Test* 1

N	τ	μ_2	$\alpha = 0.10$ k^*	$\alpha = 0.05$ k^*	$\alpha = 0.01$ k^*	$\alpha = 0.001$ k^*
50	25	1.5	47	48	49	49
		2.0	47	48	49	49
		2.5	46	48	49	49
		3.0	45	47	48	49
		3.5	44	46	48	49
100	50	1.5	92	94	97	99
		2.0	89	92	97	99
		2.5	83	88	95	99
		3.0	79	88	91	98
		3.5	78	83	90	98
200	100	1.5	175	182	192	197
		2.0	160	168	183	195
		2.5	144	153	169	188
		3.0	133	140	154	177
		3.5	127	132	144	163
500	250	1.5	409	426	441	457
		2.0	336	351	394	402
		2.5	304	313	334	372
		3.0	291	296	311	335
		3.5	283	288	298	317
1000	500	1.5	821	843	870	930
		2.0	734	756	792	854
		2.5	679	702	743	823
		3.0	602	687	701	795
		3.5	562	584	659	712

Table 4.4 *Simulated Power of Test 2*

N	τ	μ_2	$\alpha = 0.10$ *power*	$\alpha = 0.05$ *power*	$\alpha = 0.01$ *power*	$\alpha = 0.001$ *power*
50	25	1.5	0.6524	0.5266	0.3166	0.1654
		2.0	0.7282	0.6026	0.3680	0.1698
		2.5	0.7822	0.6848	0.4496	0.2020
		3.0	0.8508	0.7610	0.5380	0.2914
		3.5	0.8890	0.8266	0.6482	0.3784
100	50	1.5	0.7168	0.6010	0.3634	0.1656
		2.0	0.8330	0.7498	0.5394	0.2992
		2.5	0.9194	0.8586	0.6994	0.4702
		3.0	0.9698	0.9426	0.8254	0.6540
		3.5	0.9886	0.9702	0.9136	0.7834
200	100	1.5	0.8006	0.7016	0.4930	0.2468
		2.0	0.9364	0.8918	0.7680	0.5622
		2.5	0.9902	0.9758	0.9322	0.8342
		3.0	0.9984	0.9964	0.9822	0.9460
		3.5	0.9994	0.9992	0.9964	0.9838
500	250	1.5	0.9230	0.8658	0.6392	0.5643
		2.0	1.0000	0.9942	0.9460	0.9726
		2.5	1.0000	1.0000	0.9999	0.9920
		3.0	1.0000	1.0000	1.0000	1.0000
		3.5	1.0000	1.0000	1.0000	1.0000
1000	500	1.5	0.9942	0.9245	0.8555	0.7381
		2.0	1.0000	0.9998	0.9824	0.8999
		2.5	1.0000	1.0000	0.9945	0.9899
		3.0	1.0000	1.0000	1.0000	1.0000
		3.5	1.0000	1.0000	1.0000	1.0000

Table 4.5 *Estimated Average Stopping (EAS) Times of Test 2*

N	τ	μ_2	$\alpha = 0.10$ k^*	$\alpha = 0.05$ k^*	$\alpha = 0.01$ k^*	$\alpha = 0.001$ k^*
50	25	1.5	39	42	45	47
		2.0	38	41	45	47
		2.5	37	40	44	47
		3.0	36	39	43	47
		3.5	36	38	42	46
100	50	1.5	76	81	90	95
		2.0	73	77	86	93
		2.5	69	74	82	91
		3.0	65	69	78	87
		3.5	63	67	74	82
200	100	1.5	148	158	175	189
		2.0	134	142	159	177
		2.5	124	131	145	161
		3.0	119	124	135	149
		3.5	117	120	129	140
500	250	1.5	338	360	385	392
		2.0	283	311	339	342
		2.5	276	291	308	325
		3.0	272	282	296	304
		3.5	268	277	287	317
1000	500	1.5	625	742	794	821
		2.0	561	610	693	854
		2.5	542	560	664	823
		3.0	515	537	621	795
		3.5	505	511	520	523

and Test 2 and $\mathcal{P} \rightarrow 1$ as $\theta \rightarrow \infty$. When both θ and N are small (say, $\theta < 2.5$ and $N \leq 500$), we note that $\mathcal{P}_1(\theta|\tau, N) \ll \mathcal{P}_2(\theta|\tau, N)$ and $k_2^* < k_1^*$. However, $\mathcal{P}_1(\theta|\tau, N) \approx \mathcal{P}_2(\theta|\tau, N)$ as $N \rightarrow \infty$. Comparing our results with those of the AMOC procedure, we note the greater power and sensitivity of our proposed sequential method. For the same set of parameters θ, τ, and N, the power of Test 2 in detecting a change in the sequence increases on average by about 50% that of the AMOC procedure. Both tests yeilded considerable reductions in the stopping time as compared with the AMOC procedure. Our proposed test is really an impressive decision maker.

Next we carried out simulations to investigate the power of Test 1 and Test 2 given different locations of change. One way of dealing with different locations given the truncation point N is to let $\tau = [\upsilon N]$, where $\upsilon \in [0, 1]$ is the location fraction. We considered $N = 100$, $\mu_2 = 2.5, 3.0, 3.5$, $\alpha = 0.1, 0.05, 0.01$, and $\upsilon = 0.1, 0.2, ..., 0.9$, in each simulation for both Test 1 and Test 2. The results are reported in the following tables.

From Table 4.6 and Table 4.7, we note that the power of both tests reaches its greatest value when the change is fixed at the middle time. That is, when change occurs near the middle of a sequence of observation, it will be easier for both tests to detect this change than when it occurs at the tails of the sequence. The power decreases very quickly when the change occurs towards the upper tail (i.e., when $\upsilon > 0.5$) than the lower tail for both Test 1 and Test 2. Also, we can see that for each τ, the power is an increasing function of θ. In all cases, we note the greater sensitivity of Test 2.

In Liu (1998), extensive simulation studies were done to see how the change-point test performs when the change occurs close to the beginning or close to the end of the

sequence. The result is that the test is not very powerful, as the power decreases quickly towards the tail ends for the change-point problem than we observed in our sequential change detection test.

In conclusion, comparing our simulation results with those of the AMOC procedure, we note the superiority of our sequential method in terms of both power and stopping time.

Table 4.6 *Power and EAS Times of Test* 1 *for* $N = 100$

μ	τ	$\alpha = 0.10$		0.05		0.01	
		power	k^*	*power*	k^*	*power*	k^*
2.5	0.1	0.5762	60	0.4678	68	0.2622	83
	0.2	0.4030	75	0.3074	81	0.1654	89
	0.3	0.3062	82	0.2390	86	0.1270	92
	0.4	0.5170	72	0.4392	76	0.2972	83
	0.5	0.5946	84	0.4680	88	0.2252	95
	0.6	0.4990	89	0.3842	92	0.1674	97
	0.7	0.3724	94	0.2640	96	0.1114	98
	0.8	0.2150	97	0.1388	98	0.0570	99
	0.9	0.1094	98	0.0620	99	0.0266	99
3.0	0.1	0.5662	61	0.4590	69	0.2454	84
	0.2	0.4008	75	0.3042	81	0.1534	90
	0.3	0.2954	83	0.2314	86	0.1154	93
	0.4	0.5870	68	0.4936	72	0.3534	80
	0.5	0.7334	80	0.6218	84	0.3720	92
	0.6	0.6448	87	0.5284	90	0.2924	95
	0.7	0.5002	93	0.3860	94	0.1896	97
	0.8	0.2814	97	0.1918	98	0.0706	99
	0.9	0.1056	99	0.0706	99	0.0286	99
3.5	0.1	0.5552	62	0.4426	70	0.2364	85
	0.2	0.3820	76	0.2946	82	0.1470	90
	0.3	0.3076	82	0.2274	86	0.1108	93
	0.4	0.6460	64	0.5774	67	0.4028	97
	0.5	0.8360	76	0.7670	80	0.5134	89
	0.6	0.7772	84	0.6862	87	0.4324	93
	0.7	0.6194	91	0.5058	93	0.2636	97
	0.8	0.3574	96	0.2588	97	0.1018	99
	0.9	0.1158	99	0.0818	99	0.0268	99

Table 4.7 *Power and EAS Times of Test 2 for* $N = 100$

μ	τ	$\alpha = 0.1$		0.05		0.01	
		power	k^*	*power*	k^*	*power*	k^*
2.5	0.1	0.7572	65	0.6348	74	0.3864	87
	0.2	0.6960	69	0.5650	77	0.3134	90
	0.3	0.6236	71	0.4656	78	0.2312	89
	0.4	0.7924	61	0.6802	67	0.4366	79
	0.5	0.9224	69	0.8700	73	0.7052	82
	0.6	0.8978	76	0.8316	80	0.6556	86
	0.7	0.8360	84	0.7456	86	0.5484	92
	0.8	0.7054	91	0.5826	93	0.3898	95
	0.9	0.4658	97	0.3690	97	0.2112	98
3.0	0.1	0.7552	65	0.6270	75	0.3822	88
	0.2	0.6940	69	0.5468	78	0.3014	90
	0.3	0.6250	70	0.4542	79	0.2208	90
	0.4	0.8336	58	0.7362	63	0.5112	75
	0.5	0.9644	66	0.9358	70	0.8436	78
	0.6	0.9480	74	0.9144	77	0.8000	83
	0.7	0.9132	82	0.8428	85	0.6846	90
	0.8	0.7758	90	0.6886	92	0.4770	95
	0.9	0.4994	97	0.4124	97	0.2278	98
3.5	0.1	0.7564	65	0.6186	76	0.3864	88
	0.2	0.6836	69	0.5356	79	0.2816	91
	0.3	0.6090	71	0.4770	78	0.2124	90
	0.4	0.8774	54	0.7754	60	0.5594	72
	0.5	0.9850	64	0.9738	67	0.9208	74
	0.6	0.9780	72	0.9616	75	0.8826	81
	0.7	0.9496	81	0.9094	83	0.7992	88
	0.8	0.8534	89	0.9658	91	0.5728	94
	0.9	0.5406	96	0.4362	97	0.2564	98

4.3 Application of the Sequential Procedures

Example 4.1 Stanford Heart Transplant Survival Study

We reproduce below the data provided by Crowley and Hu (1977) and Kalbfleisch and Prentice (1980) on 103 potential heart transplant recipients from their date of acceptance into the Stanford Heart Transplant study. The patients entered the study randomly between 1967 and 1974 and received a heart transplant when a donor heart became available. Upon transplantation, it is assumed that the patient has migrated from the nontransplant population (considered as placebo) to the transplant population, and the covariate that indicates transplant changes from 0 to 1. Although this study was observational, it was generally thought that no systematic bias is introduced by selection of patients for the transplant operation. It was believed that no patient receives a heart transplant preferentially. Hence, it is assumed that the availability of a donor heart is random enough with respect to a patient receiving a heart transplant so that prognosis dissimilarity between the transplant and placebo groups in terms of other covariates such as age, waiting time to transplantation, calendar time of entry, etc., is reasonably zero. Table 4.8 is the sequential list in order of patient acceptance in the Heart Transplant study. X_i denotes the i^{th} patient's survival time recorded in days starting at the date of acceptance in the program to the date of death, or lost to follow up, or censored by the date prior to the closing date of the study (April 1, 1974). See Crowley and Hu (1977) and Kalbfleisch and Prentice (1980) for more detailed descriptions of this data set.

The survival time is said to be uncensored or censored depending on whether

Table 4.8 *Stanford Heart Transplant Data*

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

Table 4.8 *Stanford Heart Transplant Data (continued)*

i	101	102	103
X_i	39	31	11
δ_i	0	0	0
S_i	1	0	0

Sources : Crowley and Hu (1977), Kalbfleisch and Prentice (1980) and Liu (1998)

the date last seen is the date of death or the closing date of the program. The censoring indicator δ_i equals 0 or 1 depending on whether an observation is censored or not. S_i indicates the i^{th} patient's transplant status (1 = received transplant, 0 =no transplant) and i denotes the sequential order based on the date of patient's acceptance.

We initially perform some interim fixed sample size survival analyses on this data set by comparing the mortality in the transplant and non-transplant participants using the standard methods in Chapter 2. The question of interest in this study was, do heart transplant patients survive longer than heart disease patients who did not receive heart transplants? The cumulative mortality curves of the two groups are shown in Figure 4.1. This figure presents the mortality comparison over the follow-up time of the trial. There is an observable consistent mortality difference in the two groups. Slower mortality rates are apparent for the patients who received heart transplants than for the nontransplant patients. There were 71 heart transplant recipients with 24 censored observations by the end of the study. Note that two transplant patients were deselected during the study and their survival times were censored by dates prior to the closing date. The mean (limited to 1799 days) and median of the transplant group are 630 and 285 days with 95% confidence intervals (455, 806) and (125, 445), respectively.

There were 4 censored observations among the 32 non-migrating patients.

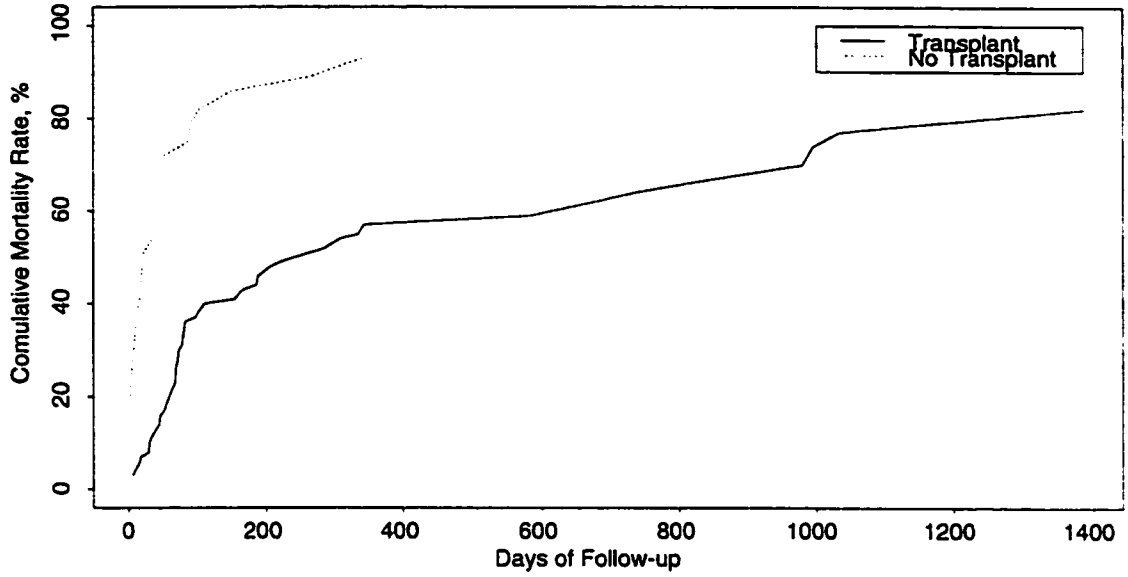


Figure 4.1 *Cumulative mortality curves comparing transplant and nontransplant*

The mean (limited to 1400 days) and median of this group are 147 and 21 days with 95% confidence intervals (18, 276) and (0, 44), respectively. Clearly, there is substantial evidence that transplant increases survival significantly.

Using the original Gehan score for comparing the two groups or Mantel's computationally simpler score function U_{kl}^{*} $(k, l = 1, \ldots, 103)$ of (2.17), by (2.19), we have

$$U^{*} = \sum_{k=1}^{103} U_{k}^{*} I(k \in S_1) = 1359,$$

where S_1 is the set of observations from the heart transplant population and U_k^* is defined in (2.18). From (2.20), we have

$$var_P(U^{*} | H_0^{*}, n_1, n_2) = 71510.2578.$$

From these, we obtain the standardized statistic,

$$Z_G = \frac{U^*}{\sqrt{var_p(U^*|H_0, n_1, n_2)}} = 5.0820,$$

with p-value$\approx$ 0, which is highly significant. Hence, we conclude from the Gehan test that the groups really differ in survival experience and that the transplant group has a significantly better prognosis.

Next, our interest is to see whether there was a change in survival times over the years. The covariates were disregarded, and we look at the variables X_is, in order of patient's entry into the study. Sequentially, we are testing for only one abrupt change in the sequence of the Stanford Heart Transplant data set of size $N = 103$, (we assume N is the truncation point of the sequence). We first apply our sequential test based on Theorem 3.1 by inspecting the sequential plot of (3.21) at each observation. The plot is shown in Figure 4.2. The horizontal line indicates the critical level. The change is detected at the first intersection of the sequential plot with the chosen critical line.

Our sequential plot suggests change at the 21^{st} observation. Test 1 at the 5% level of significance detected this change the first time at the 27^{th} inspection, and the value of the test statistic at this point is

$$|\widehat{U}_{o,k}^{(1)}| = 3.9694, \; k = 27.$$

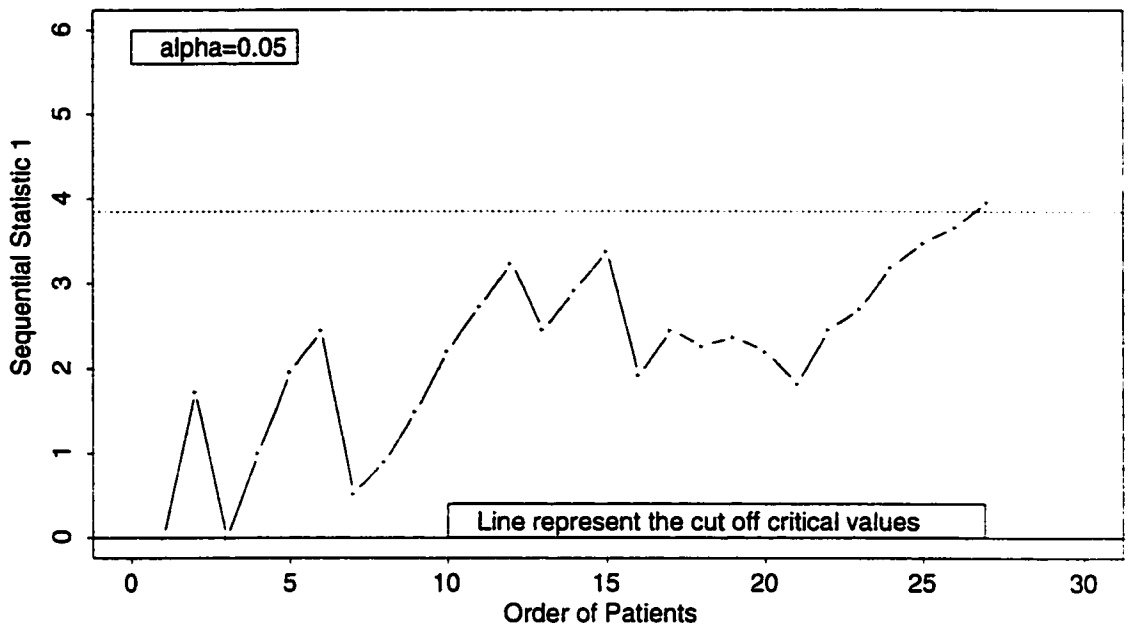


Figure 4.2 *Plot of* $|\widehat{U}_{o,k}^{(1)}|$, $k = 1, 2, ...,$ *for the Stanford Heart Transplant Data*

The p-value for Test 1 at this point is 0.0412, so the null hypothesis of identical distribution of survival times is rejected. Here, the sequential test stopped after the 27^{th} observation at $\alpha = 0.05$, and we conclude that there is a change in the survival distribution over the study period at this point. This concurs with Kalbfleisch and Prentice's findings that the year of acceptance to the study is significant. Hence, thier analyses uses time dependent covariates in the proportional hazard model, which is justified.

We next look at the performance of Test 2 using (3.23). The graph of the sequential process is shown in Figure 4.3. The plot based on Test 2 also suggested change at the 21^{st} observation and the test detected this change the first time at the 27^{th} inspection with 5% significance giving $|\widehat{U}_{o,k}^{(2)}| = 2.3574$, $k = 27$. Based on (3.16) of Theorem 3.2, the p-value

is 0.0401 and the conclusion is similar to that of Test 1, i.e., we reject the null hypothesis of identical distributions for the Stanford Heart Transplant data.

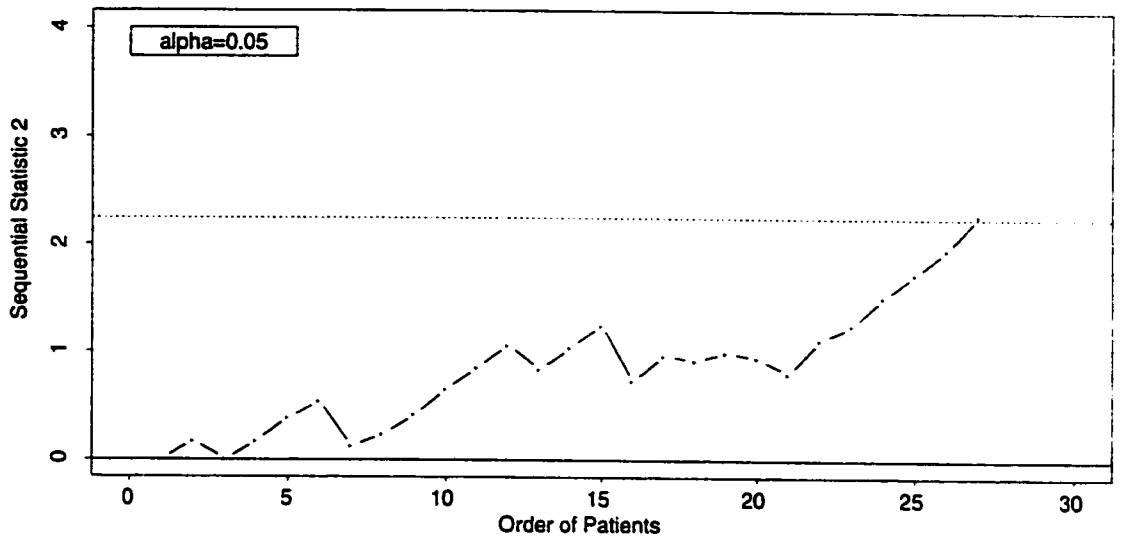


Figure 4.3 *Plot of* $|\widehat{U}_{o,k}^{(2)}|$, $k = 1, 2, \ldots$, *for the Stanford Heart Transplant Data*

The following figures are the sequential paired plots for both tests at significance levels $\alpha = 0.1$, 0.01, and 0.001, respectively, for the Stanford Heart Transplant data.

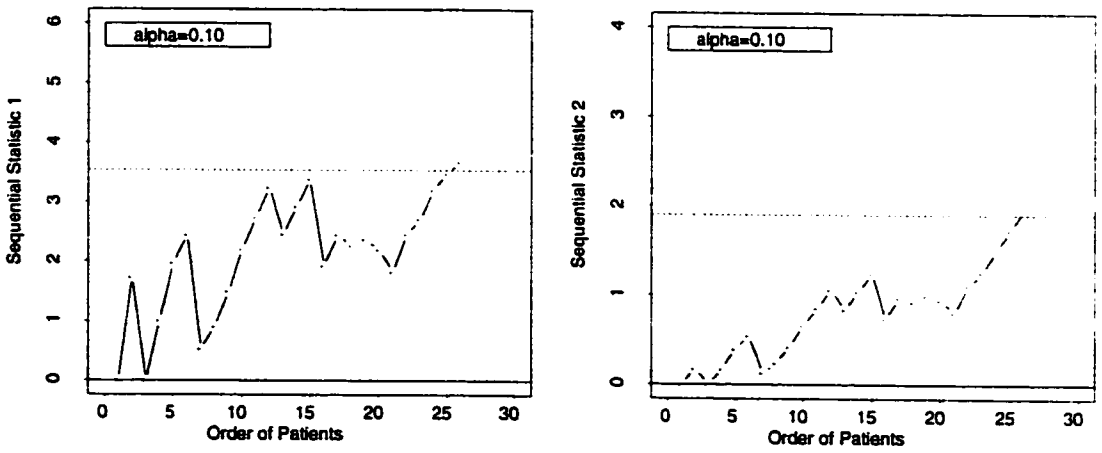


Figure 4.4 *Plots of* $|\widehat{U}_{o,k}^{(1)}|$ *and* $|\widehat{U}_{o,k}^{(2)}|$ at $\alpha = 0.10$ *for Heart Transplant Data*

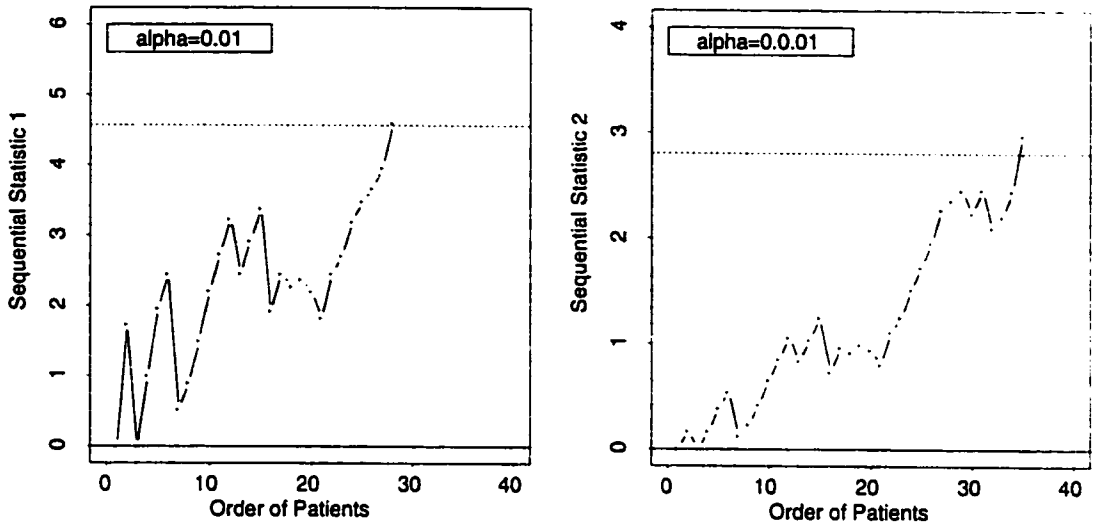


Figure 4.5 *Plots of* $|\widehat{U}_{o,k}^{(1)}|$ *and* $|\widehat{U}_{o,k}^{(2)}|$ at $\alpha = 0.01$ *for Heart Transplant Data*

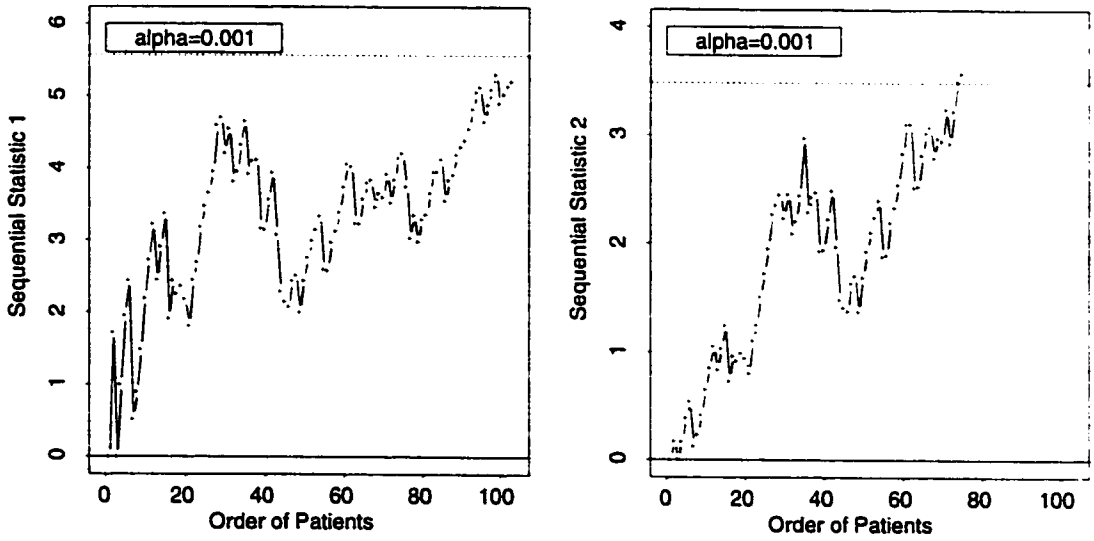


Figure 4.6 *Plots of* $|\widehat{U}_{o,k}^{(1)}|$ *and* $|\widehat{U}_{o,k}^{(2)}|$ at $\alpha = 0.001$ *for Heart Transplant Data*

Table 4.9 gives the various stopping times at their corresponding significance levels and truncation point $N = 103$.

Table 4.9 *Table of α-level Stopping Times for Test1*

Stopping Time k^*	α	$\lvert\widehat{U}^{(1)}_{o,k^*}\rvert$	*P-value*
26	0.10	3.5813	0.0972
28	0.01	4.5593	0.0089
∞	0.001	NA	NA

Table 4.10 lists the various stopping times for Test 2 using the Stanford Heart Transplant study.

Table 4.10 *Table of α-level Stopping Times for Test2*

Stopping Time k^*	α	$\lvert\widehat{U}^{(2)}_{o,k^*}\rvert$	*P-value*
26	0.10	1.9656	0.0672
36	0.01	2.5491	0.0043
73	0.001	3.5327	0

We note from the tables that the sensitivities of both tests are approximately the same for detecting early changes in the sequence. By inspection of the plots, Test 1 could not detect any change at $\alpha = 0.001$ which occurred later in the sequence but was detected by Test 2 at the 73^{rd} inspection. Hence, Test 2 is more sensitive in detecting changes than Test 1.

Considering data from only the transplant group and using Test 1, the sequential test gave $\max_{1<k\leq 71} |\widehat{U}^{(1)}_{o,k}| = 3.441$ (with a p-value of 0.1407). The result is not statistically significant implying there are no statistical differences in methods of transplantation over the study period. Hence, the observed significant change in the distribution of the data is not due to changes in the transplant techniques over the time period.

We note the power of the sequential statistics over that of the fixed sample method. Allowing for sequential testing, an initially $6\frac{1}{2}$ year planned study could be stopped after

an average period of only $2\frac{1}{2}$ years.

Example 4.2 Radiation Therapy Oncology Group Study

The Radiation Therapy Oncology Trial was a randomized, multicenter clinical trial comparing two levels of treatment therapy in people with primary tumors at any four sites in the head and neck. We reproduce in Table 4.11 part of the data set for this clinical trial carried out by the Radiation Therapy Oncology Group in the United States. This is Data Set II of Kalbfleisch and Prentice (1980) which is reproduced and listed sequentially in order of patient entry date in Liu (1998). Patients entering the study were randomly assigned to one of the treatment groups: radiation therapy alone (test) or radiation therapy together with a chemotherapy agent (standard). One objective of the study was to compare the two treatment policies with respect to patient's survival and the primary question was whether the combined treatment mode is preferable to conventional radiation therapy. 195 patients with squamous carcinoma entered the study randomly between 1968 and 1972. The need to accommodate censoring was an important point throughout the study, since some patients were lost to follow-up because of change of residency, though these cases were rare. Although this was a randomized study, considerable lack of homogeneity between the individuals being studied became primarily important. To easily accommodate this in any analysis, nonparametric or robust procedures need to be considered.

The censoring indicator δ_i equals 0 or 1 depending on whether an observation is censored or not. S_i indicates the i^{th} patient's treatment group (1 = test and 0 =standard) and i denotes the sequential order based on the date of patient's acceptance.

Table 4.11 *Radiation Therapy Oncology Data*

i	1	2	3	4	5	6	7	8	9	10
X_i	631	270	327	243	916	1823	637	235	255	184
δ_i	1	1	1	1	1	0	1	1	1	1
S_i	1	0	1	1	0	0	1	0	1	1
i	11	12	13	14	15	16	17	18	19	20
X_i	1064	414	216	324	480	245	1565	560	376	911
δ_i	1	1	1	1	1	1	0	1	1	1
S_i	1	1	0	0	1	0	1	0	0	1
i	21	22	23	24	25	26	27	28	29	30
X_i	279	144	1092	94	177	1472	526	173	575	222
δ_i	1	1	1	1	1	0	1	1	1	1
S_i	0	1	1	1	0	1	0	0	1	1
i	31	32	33	34	35	36	37	38	39	40
X_i	167	1565	134	256	404	1495	162	262	307	782
δ_i	1	1	1	1	1	0	1	1	1	1
S_i	0	1	0	0	0	0	1	1	1	0
i	41	42	43	44	45	46	47	48	49	50
X_i	661	1609	546	1766	374	1489	1446	74	301	328
δ_i	1	0	1	0	1	0	0	1	1	1
S_i	0	1	1	0	1	1	1	0	1	1
i	51	52	53	54	55	56	57	58	59	60
X_i	459	446	1644	494	279	915	228	127	1574	561
δ_i	1	1	0	1	1	1	1	1	1	1
S_i	1	1	0	0	0	1	1	1	1	1
i	61	62	63	64	65	66	67	68	69	70
X_i	370	805	192	273	1377	407	929	548	1317	1317
δ_i	1	1	1	1	0	1	1	1	0	0
S_i	1	0	0	0	0	0	1	0	0	0
i	71	72	73	74	75	76	77	78	79	80
X_i	517	1307	230	763	172	1455	1234	544	800	1460
δ_i	1	0	1	1	1	0	0	1	1	0
S_i	1	1	0	1	1	0	0	0	1	1
i	81	82	83	84	85	86	87	88	89	90
X_i	785	714	338	432	1312	351	205	1219	11	666
δ_i	1	1	1	1	0	1	1	0	1	1
S_i	1	1	0	0	0	0	1	1	0	1
i	91	92	93	94	95	96	97	98	99	100
X_i	147	1060	477	1058	1312	696	112	308	15	130
δ_i	1	0	1	0	0	1	1	1	1	1
S_i	1	0	1	1	0	0	1	0	0	0

Sources : Kalbfleisch and Prentice (1980) and Liu (1998)

Table 4.11 *Radiation Therapy Oncology Data (continued)*

i	101	102	103	104	105	106	107	108	109	110
X_i	296	293	545	1086	1250	147	726	310	599	998
δ_i	1	1	1	0	0	1	1	1	1	0
S_i	1	1	0	1	1	0	1	0	1	1
i	111	112	113	114	115	116	117	118	119	120
X_i	1089	382	932	264	11	911	89	525	532	637
δ_i	0	1	0	1	1	0	1	1	0	1
S_i	0	1	1	0	0	1	1	1	0	1
i	121	122	123	124	125	126	127	128	129	130
X_i	112	1095	170	943	191	928	918	825	99	99
δ_i	1	0	1	0	1	0	0	0	1	1
S_i	1	1	1	1	1	0	1	1	1	1
i	131	132	133	134	135	136	137	138	139	140
X_i	933	461	347	372	731	363	238	593	219	465
δ_i	0	1	1	1	0	1	1	0	1	1
S_i	1	0	0	1	0	1	0	0	1	1
i	141	142	143	144	145	146	147	148	149	150
X_i	446	553	274	723	532	154	369	541	107	854
δ_i	1	1	1	0	1	1	1	1	1	0
S_i	0	0	0	0	0	0	0	1	0	0
i	151	152	153	154	155	156	157	158	159	160
X_i	822	775	336	513	914	757	794	105	733	600
δ_i	0	1	1	1	0	1	0	1	0	0
S_i	1	1	1	0	0	1	1	0	1	0
i	161	162	163	164	165	166	167	168	169	170
X_i	266	317	407	346	518	395	81	608	760	343
δ_i	1	1	1	1	1	1	1	0	0	1
S_i	1	0	0	0	0	0	1	0	1	1
i	171	172	173	174	175	176	177	178	179	180
X_i	324	254	751	334	275	546	112	182	209	208
δ_i	1	1	0	1	1	0	1	0	1	1
S_i	0	1	1	1	0	0	0	0	0	0
i	181	182	183	184	185	186	187	188	189	190
X_i	174	651	672	291	498	276	90	213	38	128
δ_i	1	0	0	1	1	0	0	1	1	1
S_i	0	1	0	0	1	0	1	1	0	1
i	191	192	193	194	195					
X_i	445	159	219	173	413					
δ_i	0	1	1	1	0					
S_i	1	1	0	0	0					

Sources : Kalbfleisch and Prentice (1980) and Liu (1998)

We begin a nonparametric discussion of these data by looking at the Kaplan Meier survival curves. The curves are shown in Figure 4.4. It is notable that the survival curves do not suggest any wide swings. The curves indicate relatively weak difference between the two groups.

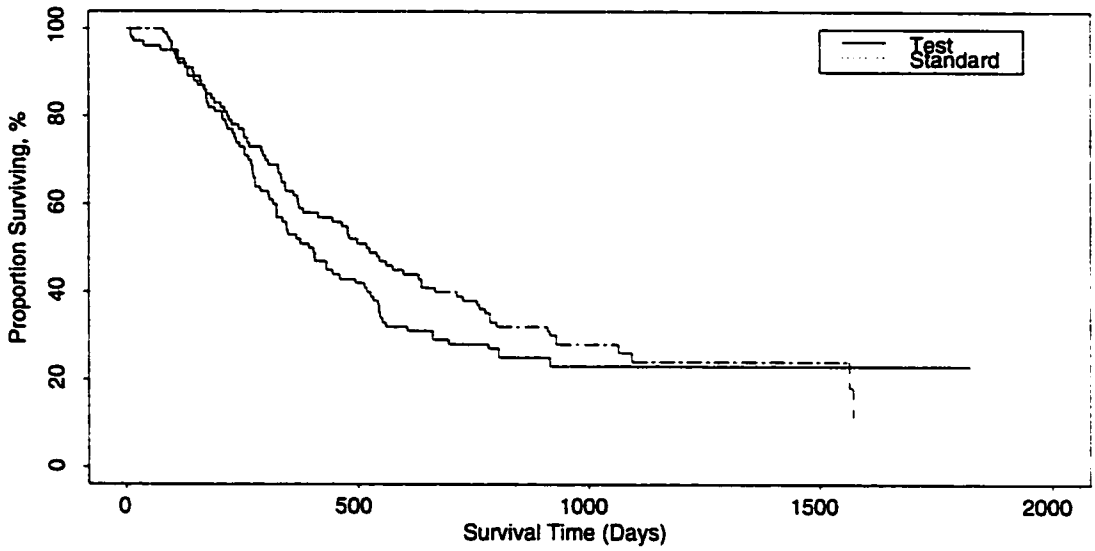


Figure 4.7 *Kaplan – Meier baseline survival curves*

The mean (limited to 1823 days) and the median of the test group are 692 and 595 days with 95% confidence intervals (555, 828) and (302, 488), respectively. For the standard group, the mean (limited to 1609 days) is 717 days with 95% confidence interval (605, 828) and the median survival time is 525 with 95% confidence interval (392, 659). It is of considerable interest to determine the relative extent to which chemotherapy relates to subsequent survival. The estimated hazard ratio is 1.1722 (p-value= 0.3458) with 95% confidence interval (0.8425, 1.6309).

Using the Gehan test (2.16) for comparing the two treatments, the normalized statistic was $Z_G = 1.3118$ (two sided p-value= 0.0948). These results indicate no statistical significant differences between the two treatment groups at the 5% level of significance, i.e., the standard method and the new test have the same survival experience. Hence, at 5% level of significance, we expect lifetime observations from the combined sample of the two groups to be identically distributed.

Now we turn to the problem of determining whether the survival time distribution changes over the period of study. The graph in Figure 4.8 is a plot of sequential statistics based on Test 1. It is evident from the graph that a change is detected after the 83^{rd} inspection at the 5% significance level. Based on the standard interpretation of cumulative sum plots, we estimate the time of change to be around 64.

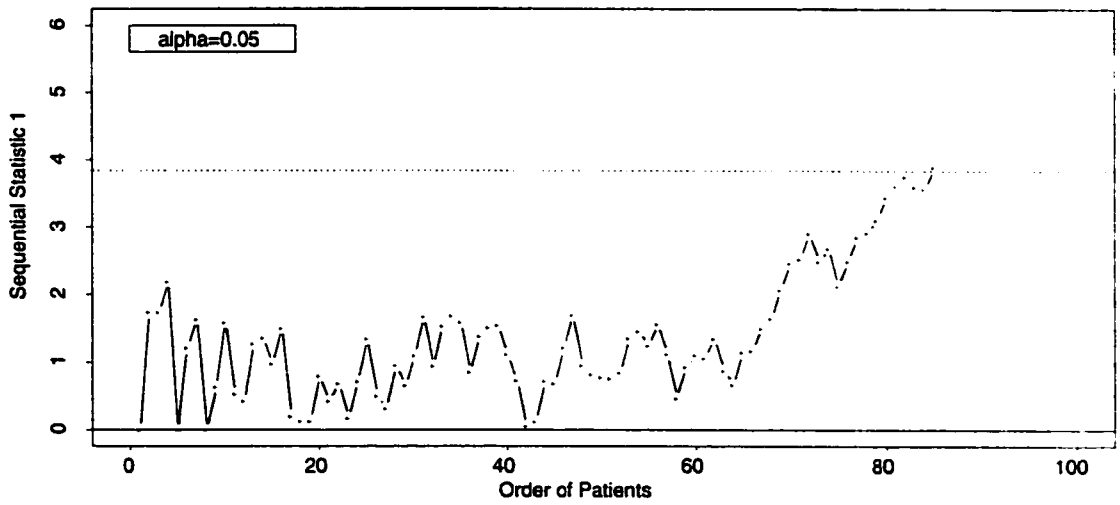


Figure 4.8 *Plot of* $|\widehat{U}_{o,k}^{(1)}|$, $k = 1, 2...$, *for the Radiation Therapy Oncology Data*

The plot of Test 2 process is shown in Figure 4.9. The results are similar and consistent with Test 1 but slightly more sensitive.

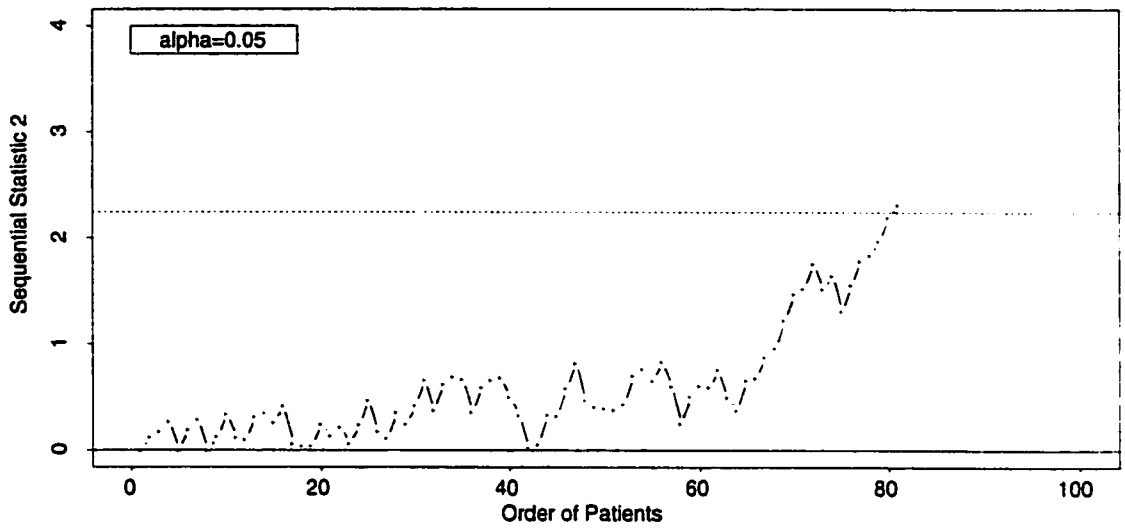


Figure 4.9 *Plot of* $|\widehat{U}_{o,k}^{(2)}|$, $k = 1, 2...$, *for the Radiation Therapy Oncology Data*

The following figures show the plots at the significance levels $\alpha = 0.1$, 0.01, and 0.001, respectively.

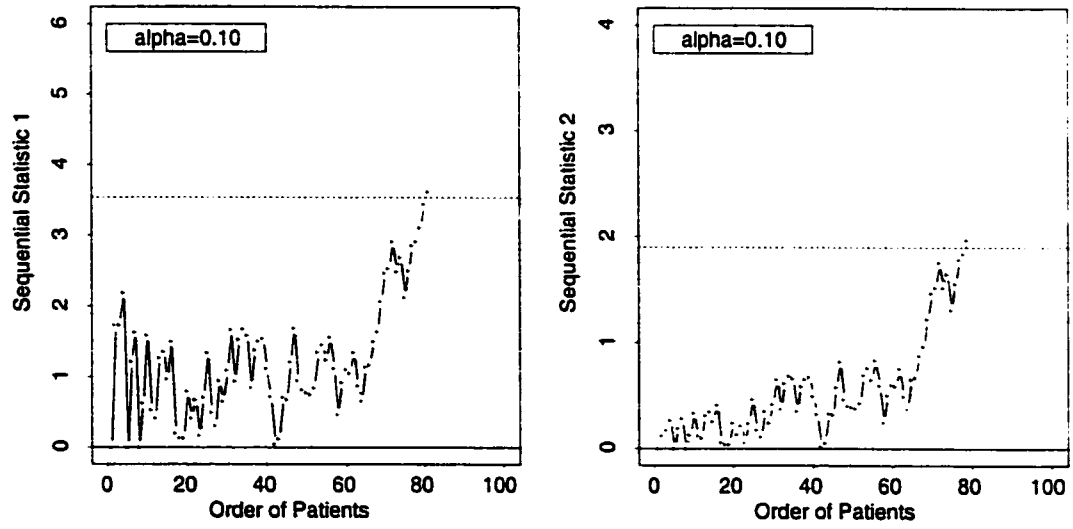


Figure 4.10 *Plots of* $|\widehat{U}_{o,k}^{(1)}|$ and $|\widehat{U}_{o,k}^{(2)}|$ *at* $\alpha = 0.10$ *for the Radiation Therapy Data*

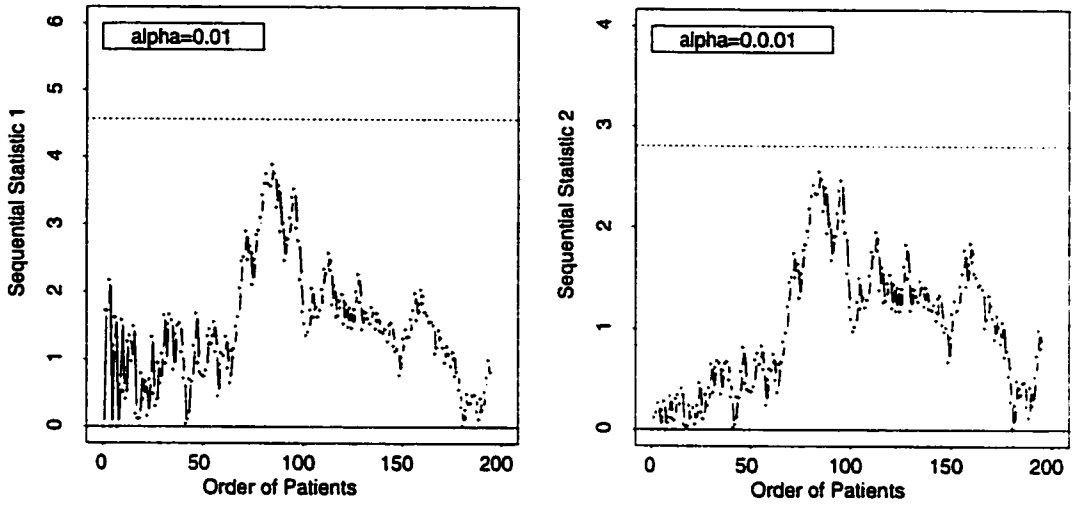


Figure 4.11 *Plots of* $|\widehat{U}_{o,k}^{(1)}|$ and $|\widehat{U}_{o,k}^{(2)}|$ at $\alpha = 0.01$ *for the Radiation Therapy Data*

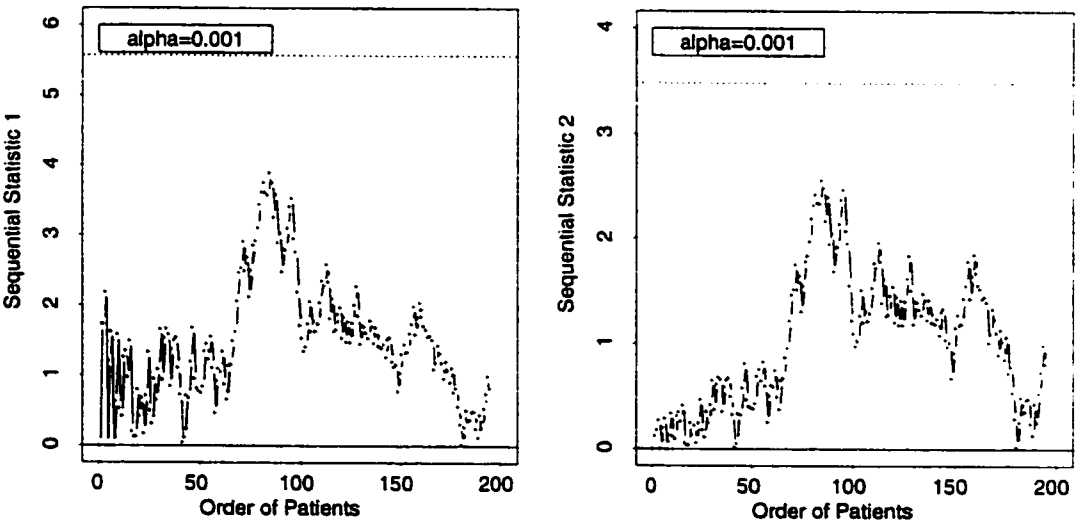


Figure 4.12 *Plots of* $|\widehat{U}_{o,k}^{(1)}|$ and $|\widehat{U}_{o,k}^{(2)}|$ at $\alpha = 0.001$ *for the Radiation Therapy Data*

It is also clear that the change is not significant at the $\alpha = 0.01$ and $\alpha = 0.001$.

Both the data sets of the Stanford Heart Transplant program and Radiation Therapy Oncology study have been extensively studied under the assumption of identical distribution of the survival time. In Gombay and Liu (2000c), using a test based on (2.48), they

show that this assumption is not reasonable for the Stanford Heart Transplant data. Their $\mathcal{U}$-statistic change-point test detected a 5% level significant change at the 49^{th} observation. Our sequential test at the same level detected a change earlier in the sequence $\{k^{-3/2}U_k^*, k = 1, ..., 103\}$, at the 21^{st} observation and stopped at the 27^{th} observation.

Their change-point test failed to detect any significant change in the process of the Radiation Therapy Oncology data, giving support to the null hypothesis. However their plot for the process suggested a test for a change alternative H_a'. They shield this deficiency by concluding that the purpose of the change-point test is to see that H_o' is not too wildly violated before a user does some statistical inference, so the use of some powerful test is not important. With this, they justify Kalbfleisch and Prentice's use of time independent covariates in the proportional hazard model. This visualized change in the change-point test plot of $\left\{\sum_{i=1}^{k} U_i, k = 1, ..., 195\right\}$ was easily detected by our sequential test, indicating the greater power and sensitivity of our sequential test. So, even if the full sequence of observations $X_1, ..., X_n$ is available, the sequential test may be preferable to the change-point test.

4.4 Some Problems in Sequential Analysis

When issues of ethics, feasibility, and cost all have to be addressed satisfactorily in a randomized biomedical or clinical research, sequential analysis represents the "gold standard". However, there are a number of issues to consider in sequential analysis to ensure valid conclusions. Sequential testing of accumulating data is essential to monitoring, but it does have statistical implications. If the null hypothesis, H_o', is in fact, true, and sequential

tests of that hypothesis are made at the same level of significance using accumulating data, the probability that, at some time, the test will be called significant by chance alone will be greater than the significance level selected. That is, the rate of incorrectly rejecting the null hypothesis will be larger than what is normally considered to be acceptable. The reason for this is that the tests are based on asymptotic results, that is, when $N \rightarrow \infty$, where N is the truncation point. So when $N = 50$, or 100, then caution must be used.

4.5 Conclusion and Future Work

The results of this thesis are based on the current research of Prof. E. Gombay. To our knowledge, no studies have been done to accommodate such change detection procedures for empirical data. In this thesis, we have demonstrated our findings and gave practical examples where our research will be useful. The results from the analysis show that our sequential procedure is an impressive decision maker. As pointed out in the simulation study, Test 2 is more powerful than Test 1. However both tests suceeded in providing extreme reduction in sample size ranging from about 20% to 80%. Comparing our results with those of the AMOC procedure, this reduction is really impressive. Because of the simplicity of applying our sequential procedure, we recommend its use whenever change detection problems are considered.

It will be interesting to investigate the characteristic behavior of the test at the early stages of the analysis. It is conjectured that the variance estimate of the process at the early stage of the process will be reasonably stable enough not to affect our procedure. Analysis of the process shows that if k is small, the process has early fluctuation even under the no

change hypothesis. This means the probability of detecting false early changes increases and the sampling process could often be truncated at some premature limit at the early stages. In most situations, the level of "safety" obtained by employing our sequential test would far exceed the level of any "harm" incurred by truncating too soon. However, the conjecture remains to be investigated later through further simulation studies.

Appendix

C codes for the Simulation

```
#include <stdio.h>
#include <stdlib.h>
#include <math.h>
#include <unistd.h>
#define OUT_FILE "Result_file.dat"

int  MM; /* Number of simulation */
int  N1; /* Change point */
double  L2; /* Mean after fixed change point*/

double        L1 = 1.0;
double        L  = 3.0;
int           T  = 4 ;
int           NN;

float *sum, *sq, *stat1, *avrstop, *stat2, *sdev;
float *v, *r, *t1, *t2, *c, *count, *x, *sumsq, *u;;
int *stop

char  *d;

float TS2[4], tmpfloat, tmpfloat1, tmpfloat2;
FILE *out_putfile;

int i,l,j,a,k;

char  ProgramName[50] = "site2nn";
char  LastChange[100] = "(Last Change : December 2000)";

/*-------------------------------------------------*/

/*
  register functions called by self/main other
~
*/

void    ShowUsage(void);

void    ProcessInput(int pbv_argc, char *pbv_argv[]);

void    CreateArrays(void);
```

```
/*-------------------------------------------------*/


main(int argc, char *argv[])

{

  ProcessInput(argc,argv);
   CreateArrays();

 out_putfile=fopen(OUT_FILE, "w+");

 srand48(getpid());


 TS2[0]=1.96;
 TS2[1]=2.24;
 TS2[2]=2.81;
 TS2[3]=3.48;


 for(a=0; a<T; a++)
 {

  stop[0]=0;
  count=0.0;
  TS=TS2[a];

  for(l=1;l<=MM;l++)
  {
    u[0]=0.0;
    v[0]=0.0;
    sumsq[0]=0.0;
    sum[0]=0.0;
    sdev[0]=0.0;

      stat1[0]=0.0;
      stat2[0]=0.0;


/* generating t1, exponential time for group 1 */

    for(i=0;i<N1;i++)
    {
      tmpfloat1 = 0.0001+drand48();
      t1[i] = (-1/L1)*log(tmpfloat1);
```

```
    }

    /* generating t2, exponetial time for group 2  */
    for(i=0;i<N2;i++)
    {
      tmpfloat2 =0.0001+ drand48();
      t2[i] = (-1/L2)*log(tmpfloat2);
    }

    /* generating exponenetial censoring variable */
    for(i=0;i<NN;i++)
    {
      tmpfloat =0.0001+ drand48();
      c[i] = (-1/L)*log(tmpfloat);


    /* combine group 1 and 2 element  */
    for(i=0;i<NN;i++)
    {
      if(i<N1)
      { x[i]=t1[i]; }
      else
      { x[i]=t2[i-N1]; }
    }

    /* compute r */
    for(i=0;i<=NN;i++)
    {
      if(c[i]<x[i])
      {
        r[i] = c[i];
        d[i] = 0;
      }
      else
      {
        r[i] = x[i];
        d[i] = 1;
      }

    }


for(i=1;i<NN ;i++)
    {
       u[i]=0.0;

       for (j=0;j<=i;j++)
```

```
        {
          if (((r[i]>r[j]) && (d[j]==1))||((r[i]==r[j]) &&
           (d[i]==0) && d[j]==1))
          { u[i]=u[i]+1.0; }
          else
          {
            if ((r[i]<r[j] && d[i]==1)||(r[i]==r[j] &&
             d[i]==1 && d[j]==0))
            { u[i]=u[i]-1.0; }
            else
            { u[i]=u[i]+0.0; }
          }
        }

        sum[i]=sum[i-1]+u[i];
        sumsq[i]=0.0;

        for (k=1;k<=i;k++)
        {
          v[k]=0.0;
          for (j=0;(j<=i)&&(j!=k);j++)
          {

if(((r[k]>r[j])&&(d[j]==1))||((r[k]==r[j])&&(d[k]==0)
        &&(d[j]==1)))
            {
              v[k]=v[k]+1.0;
            }
            else
            {

if(((r[k]<r[j])&&(d[k]==1))||((r[k]==r[j])&&(d[k]==1)
        &&(d[j]==0)))
              {
                v[k]=v[k]-1.0;
              }
              else
              {
                v[k]=v[k]+0.0;
              }
            }
          }
          sumsq[i]=sumsq[i]+v[k]*v[k];
        }

        sdev[i]=sqrt(sumsq[i]);
```

```
/*Eqtn 3.19*/
stat1[i]= sqrt(3.0)*abs(sum[i])/sdev[i];
/*Eqtn 3.25*/
stat2[i]=sqrt(3.0*i)*abs(sum[i])/(sqrt(NN)*sdev[i]);


 /*-------------------------------------------------*/

void  ProcessInput(int pbv_argc, char *pbv_argv[])

{

 int CmdLinePtr;

 if (pbv_argc==1)
  ShowUsage();

 CmdLinePtr=1;

 MM  = atoi(pbv_argv[CmdLinePtr++]);
 N1  = atoi(pbv_argv[CmdLinePtr++]);
 L2  = atof(pbv_argv[CmdLinePtr++]);

 /*
   report command line unput to user
 */

 fprintf(stderr,"\n\n          User          Input          to:
\"%s\"\n\n",ProgramName);
 fprintf(stderr,"  MM    : %u\n",MM);
 fprintf(stderr,"  N1    : %u\n",N1);
 fprintf(stderr,"  L2    : %f\n",L2);
 fprintf(stderr,"\n\n");

}

/*-------------------------------------------------*/

void ShowUsage(void)

{

 fprintf(stderr,"\nUSAGE: %s  MM N1 L2\n\n",ProgramName);

 fprintf(stderr,"  MM     :  Number of Simulation\n");
 fprintf(stderr,"  N1     :  Sample Size \n");
 fprintf(stderr,"  L2     :  L2 (exp parameter)\n\n");
```

```
 fprintf(stderr,"%s\n\n",LastChange);

 exit(-1);

}

/*-------------------------------------------------*/

void CreateArrays(void)

{

 fprintf(stderr,"Dynamically allocating arrays...");


 if ((sum=(float *)malloc(NN*sizeof(float)))==NULL)
  {
   fprintf(stderr,"\nERROR : Could not malloc array for
size N : %d\n\n",NN);
   exit(-1);
  }

 if ((sq=(float *)malloc(NN*sizeof(float)))==NULL)
  {
   fprintf(stderr,"\nERROR : Could not malloc array for
size N : %d\n\n",NN);
   exit(-1);
  }

 if ((stat1=(float *)malloc(NN*sizeof(float)))==NULL)
  {
   fprintf(stderr,"\nERROR : Could not malloc array for
size N : %d\n\n",NN);
   exit(-1);
  }

if ((stat2=(float *)malloc(NN*sizeof(float)))==NULL)
  {
   fprintf(stderr,"\nERROR : Could not malloc array for
size N : %d\n\n",NN);
   exit(-1);
  }
```

```
 if ((sdev=(float *)malloc(NN*sizeof(float)))==NULL)
  {
   fprintf(stderr,"\nERROR : Could not malloc array for
size N : %d\n\n",NN);
   exit(-1);
  }

 if ((sumsq=(float *)malloc(NN*sizeof(float)))==NULL)
  {
   fprintf(stderr,"\nERROR : Could not malloc array for
size N : %d\n\n",NN);
   exit(-1);
  }

 if ((u=(float *)malloc(NN*sizeof(float)))==NULL)
  {
   fprintf(stderr,"\nERROR : Could not malloc array for
size N : %d\n\n",NN);
   exit(-1);
  }

 if ((v=(float *)malloc(NN*sizeof(float)))==NULL)
  {
   fprintf(stderr,"\nERROR : Could not malloc array for
size N : %d\n\n",NN);
   exit(-1);
  }

if ((r=(float *)malloc(NN*sizeof(float)))==NULL)
  {
   fprintf(stderr,"\nERROR : Could not malloc array for
size N : %d\n\n",NN);
   exit(-1);
  }


 if ((c=(float *)malloc(NN*sizeof(float)))==NULL)
  {
   fprintf(stderr,"\nERROR : Could not malloc array for
size N : %d\n\n",NN);        exit(-1);
  }

 if ((x=(float *)malloc(NN*sizeof(float)))==NULL)
  {
   fprintf(stderr,"\nERROR : Could not malloc array for
size N : %d\n\n",NN);
   exit(-1);
```

```
  }

 if ((t1=(float *)malloc(N1*sizeof(float)))==NULL)
  {
   fprintf(stderr,"\nERROR : Could not malloc array for
size N1 : %d\n\n",N1);
   exit(-1);
  }

 if ((t2=(float *)malloc(N1*sizeof(float)))==NULL)
  {
   fprintf(stderr,"\nERROR : Could not malloc array for
size N1 : %d\n\n",N1);
   exit(-1);
  }

 if ((count=(float *)malloc(MM*sizeof(float)))==NULL)
  {
   fprintf(stderr,"\nERROR : Could not malloc array for
size N1 : %d\n\n",MM);
   exit(-1);
  }

 if ((stop=(int *)malloc(MM*sizeof(int)))==NULL)
  {
   fprintf(stderr,"\nERROR : Could not malloc array for
size MM : %d\n\n",MM);
   exit(-1);
  }

 if ((d=(char *)malloc(NN*sizeof(char)))==NULL)
  {
   fprintf(stderr,"\nERROR : Could not malloc array for
size NN : %d\n\n",NN);
   exit(-1);
  }

 fprintf(stderr,"Done\n\n");

}
if (i>=NN-1)
        {
        i=i+1;
        count=count+0.0;
        stop[l]=stop[l-1]+NN;

        }
```

```
      else if (i>N1&&stat2[i-1]>TS)

       {

        count=count+1.0;
        stop[l]=stop[l-1]+i+1;


         break;
        }


    }

     }

  fprintf(out_putfile,"No of simmlation=%3d\n\n",MM);
  fprintf(out_putfile,"sample size=%3d\n\n",NN);
  fprintf(out_putfile,"Change point=%3d\n\n",N1);
  fprintf(out_putfile,"s-Size after Change=%3d\n\n",N2);
  fprintf(out_putfile,"Critical boundary =%5.2f\n\n",TS);
  fprintf(out_putfile,"Mean after Change=%6.1f\n\n",L2);
  fprintf(out_putfile,"simu-power=%6.4f\n\n",count/5000.0);
  fprintf(out_putfile,"stop time=%6d\n\n" , stop[MM]/5000);
}
 fclose(out_putfile);
 return 0;
}
/*-------------------------------------------------*/
End of program.
```

Bibliography

Armitage, P. (1975). *Sequential Medical Trials* (2nd Ed). Wiley & Sons, New York.

Breslow, N. (1969). On large sample sequential analysis with application to survival data. *Journal of Applied Probability*, **6**, 261-274.

Breslow, N. (1970). A generalized Kruskal-Wallis test for comparing K sample subject to unequal pattern of censorship. *Biometrika*, **57**, 579-594.

Carlstein, E. (1988). Nonparametric change-point estimation. *Annals of Statistics*, **16**, 188-197.

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. (1997). *Limit Theorems in change-point analysis*. Wiley and Sons, New York.

Csörgő, M. and Horváth, L. (1988). Nonparametric methods for change-point problems. In: *Handbook of statistics* 7 (Krishnaiah, P. R. and Rao, C. R.) North-Holland, pp. 403-425.

Fergar, D.(1994). Nonparametric change-point test of the Kolmogorov-Smirnov type. *IMS-Lecture Notes-Monograph Series (Volume 23)*, California.

Gehan, E. A. (1965a). A generalized Wilcoxon test for comparing arbitrarily single censored sample. *Biometrika*, **52**, 203-223.

Gehan, E. A. (1965b). A generalized two-sample Wilcoxon test for doubly censored data. *Biometrika*, **52**, 650-6533.

Gilbert, J. P. (1962). Random censorship. *Unpublished Ph.D. Thesis*, University of Chicago.

Gombay, E. (1995). Nonparametric truncated sequential change-point detection. *Statistics & Decisions*, **13**, 71-82

Gombay, E. (2000a). Sequential change detection with likelihood ratios. *Statistics and Probability Letters*.

Gombay, E. (2000b). U-statistics for Sequential Change Detection. *Metrika*. To appear.

Gombay, E. and Liu, S. (2000c). A nonparametric test for change in randomly censored data, *Canadian Journal of Statistics*, **28**, 113-121.

Gombay, E. (2000). Comparison of *U*-statistics in the change-point problem and in sequential change detection. To appear.

Hoeffding, W.(1948).A class of statistics with asymptotically normal distribution. *Annals of Mathematices and Statistics*, **19**, 293-325.

Jones, D. and Whitehead J. (1979). Sequential form of the logrank and modified – Wilcoxon test for censored data. *Biometrika*, **66**, 105-113

Kalbfleisch, J. D. and Prentice, R. L. (1980). *The Statistical Analysis of Failure Time Data*, John Wiley and Sons, New York.

Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, **53**, 457-481.

Lawless, J. F. (1982). *Statistical Models and Methods for Lifetime Data*. Wiley & Sons, New York.

Liu, S. (1998). Nonparametric test for change-point problems with random censorship. *Ph.D. Thesis*. University of Alberta.

Lee, A. J. (1990). *U-Statistics*. Marcel Dekker Inc., New York.

Lorden, G. (1971). Procedures for reacting to a change in distribution. *Annals of Mathematical Statistics*, **42**, 1897-1908.

Mantel, N. (1967). Ranking procedures for arbitrarily restricted observations. *Biometrics*, **23**, 65-78.

Miller, JR. and Rupert, G. (1981). *Survival Analysis*.

Peto, R.and Peto, J. (1972). Asymptotic efficient rank invariant test procedures (with discussion). *Journal of the Royal Statistical Society* Series A, **135**, 185-206.

Pollak, M. (1985). Optimal detection of change in distribution. *Annals of Statistics*, **13**, 206-277.

Prentice, R. L. (1978). Linear rank tests with right censored data. *Biometrika*, **65**, 167-179.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley and Sons, New York.

Slud, E. V. and Wei, L. J. (1982). Two-sample significance test based on modified-Wilcoxon statistics. *Journal of the American Statistical Association*.

Tarone, R. E. (1981). On the distribution of the maximum of the logrank statistics and the modified-Wilcoxon statistics. *Biometrics*, **37**, 79-86.

Tarone, R. E., and Ware, J. (1977). On distribution free tests for equality of survival distributions. *Biometrika*, **64**, 157-160.

Tsiatis, A. A. (1982) Repeated significance testing for a general class of score statistics used in censored survival analysis. *Journal of the American Statistical Association.*

Vostrikova, L. J. (1981). Detection of a "disorder" in a winner process. *Theory of Probability Appl.*, **26**, 356-362.