Survival analysis issues with interval–censored data

Ramon Oller Piqué

PhD Thesis directed by
Guadalupe Gómez Melis

Universitat Politècnica de Catalunya

Vic, maig del 2006

A l'Alícia,
en Martí, en Joan
i l'Albert.

Agraïments

Vull agrair aquesta tesi a la Lupe. Sense la seva qualitat humana i científica, aquest projecte no hauria arribat a bon port. He après molta estadística però també m'he sentit recolzat en tot moment. Moltes gràcies Lupe.

Vull donar les gràcies a la Malu i a tots els companys del grup GRASS, pels bons moments científics i lúdics que hem passat junts.

Thank you Thomas Scheike for your friendship during my stay in København. Also, thank you to the University of Copenhagen's Department of Biostatistics.

Thank you to the external referees whose comments have been helpful to improve this PhD thesis.

Vull recordar també a tots els companys de feina a la UVIC. El bon ambient en el departament m'ha ajudat molt durant aquests anys. Gràcies.

A la meva família i, especialment, al pare i a la mare: gràcies pel vostre suport.

Finalment, vull esmentar l'Alícia, en Martí, en Joan i l'Albert. Moltes gràcies Alícia per la teva paciència i per uns fills tan eixerits.

Summary

Survival analysis is used in various fields for analyzing data involving the duration between two events. It is also known as event history analysis, lifetime data analysis, reliability analysis or time to event analysis. One of the difficulties which arise in this area is the presence of censored data. The lifetime of an individual is censored when it cannot be exactly measured but partial information is available. Different circumstances can produce different types of censoring. Interval censoring refers to the situation when the event of interest cannot be directly observed and it is only known to have occurred during a random interval of time. This kind of censoring has produced a lot of work in the last years and typically occurs for individuals in a study being inspected or observed intermittently, so that an individual's lifetime is known only to lie between two successive observation times.

This PhD thesis is divided into two parts which handle two important issues of interval censored data. The first part is composed by Chapter 2 and Chapter 3 and it is about formal conditions which allow estimation of the lifetime distribution to be based on a well known simplified likelihood. The second part is composed by Chapter 4 and Chapter 5 and it is devoted to the study of test procedures for the k–sample problem. The present work reproduces several material which has already been published or has been already submitted.

In Chapter 1 we give the basic notation used in this PhD thesis. We also describe the nonparametric approach to estimate the distribution function of the lifetime variable. Peto (1973) and Turnbull (1976) were the first authors to propose an estimation method which is based on a simplified version of the likelihood function. Other authors have

studied the uniqueness of the solution given by this method (Gentleman and Geyer, 1994) or have improved it with new proposals (Wellner and Zhan, 1997).

Chapter 2 reproduces the paper of Oller *et al.* (2004). We prove the equivalence between different characterizations of noninformative censoring appeared in the literature and we define an analogous constant–sum condition to the one derived in the context of right censoring. We prove as well that when the noninformative condition or the constant–sum condition holds, the simplified likelihood can be used to obtain the nonparametric maximum likelihood estimator (NPMLE) of the failure time distribution function. Finally, we characterize the constant–sum property according to different types of censoring. In Chapter 3 we study the relevance of the constant–sum property in the identifiability of the lifetime distribution. We show that the lifetime distribution is not identifiable outside the class of constant–sum models. We also show that the lifetime probabilities assigned to the observable intervals are identifiable inside the class of constant–sum models. We illustrate all these notions with several examples.

Chapter 4 has partially been published in the survey paper of Gómez *et al.* (2004). It gives a general view of those procedures which have been applied in the nonparametric problem of the comparison of two or more interval–censored samples. We also develop some S–Plus routines which implement the permutational version of the Wilcoxon test, the Logrank test and the t–test for interval censored data (Fay and Shih, 1998). This part of the PhD thesis is completed in Chapter 5 by different proposals of extension of the Jonckeere's test. In order to test for an increasing trend in the k–sample problem, Abel (1986) gives one of the few generalizations of the Jonckheree's test for interval–censored data. We also suggest different Jonckheere–type tests according to the tests presented in Chapter 4. We use permutational and Monte Carlo approaches. We give computer programs for each proposal and perform a simulation study in order compare the power of each proposal under different parametric assumptions and different alternatives. We motivate both chapters with the analysis of a set of data from a study of the benefits of zidovudine in patients in the early stages of the HIV infection (Volberding *et al.*, 1995).

Finally, Chapter 6 summarizes results and address those aspects which remain to be completed.

Resum

L'anàlisi de la supervivència s'utilitza en diversos àmbits per tal d'analitzar dades que mesuren el temps transcorregut entre dos successos. També s'anomena anàlisi de la història dels esdeveniments, anàlisi de temps de vida, anàlisi de fiabilitat o anàlisi del temps fins a l'esdeveniment. Una de les dificultats que té aquesta àrea de l'estadística és la presència de dades censurades. El temps de vida d'un individu és censurat quan només és possible mesurar–lo de manera parcial o inexacta. Hi ha diverses circumstàncies que donen lloc a diversos tipus de censura. La censura en un interval fa referència a una situació on el succés d'interès no es pot observar directament i només tenim coneixement que ha tingut lloc en un interval de temps aleatori. Aquest tipus de censura ha generat molta recerca en els darrers anys i usualment té lloc en estudis on els individus són inspeccionats o observats de manera intermitent. En aquesta situació només tenim coneixement que el temps de vida de l'individu es troba entre dos temps d'inspecció consecutius.

Aquesta tesi doctoral es divideix en dues parts que tracten dues qüestions importants que fan referència a dades amb censura en un interval. La primera part la formen els capítols 2 i 3 els quals tracten sobre condicions formals que asseguren que la versemblança simplificada pot ser utilitzada en l'estimació de la distribució del temps de vida. La segona part la formen els capítols 4 i 5 que es dediquen a l'estudi de procediments estadístics pel problema de k mostres. El treball que reproduïm conté diversos materials que ja s'han publicat o ja s'han presentat per ser considerats com objecte de publicació.

En el capítol 1 introduïm la notació bàsica que s'utilitza en la tesi doctoral. També fem una descripció de l'enfocament no paramètric en l'estimació de la funció de dis-

tribució del temps de vida. Peto (1973) i Turnbull (1976) van ser els primers autors que van proposar un mètode d'estimació basat en la versió simplificada de la funció de versemblança. Altres autors han estudiat la unicitat de la solució obtinguda en aquest mètode (Gentleman i Geyer, 1994) o han millorat el mètode amb noves propostes (Wellner i Zhan, 1997).

El capítol 2 reprodueix l'article d'Oller *et al.* (2004). Demostrem l'equivalència entre les diferents caracteritzacions de censura no informativa que podem trobar a la bibliografia i definim una condició de suma constant anàloga a l'obtinguda en el context de censura per la dreta. També demostrem que si la condició de no informació o la condició de suma constant són certes, la versemblança simplificada es pot utilitzar per obtenir l'estimador de màxima versemblança no paramètric (NPMLE) de la funció de distribució del temps de vida. Finalment, caracteritzem la propietat de suma constant d'acord amb diversos tipus de censura. En el capítol 3 estudiem quina relació té la propietat de suma constant en la identificació de la distribució del temps de vida. Demostrem que la distribució del temps de vida no és identificable fora de la classe dels models de suma constant. També demostrem que la probabilitat del temps de vida en cadascun dels intervals observables és identificable dins la classe dels models de suma constant. Tots aquests conceptes els il·lustrem amb diversos exemples.

El capítol 4 s'ha publicat parcialment en l'article de revisió metodològica de Gómez *et al.* (2004). Proporciona una visió general d'aquelles tècniques que s'han aplicat en el problema no paramètric de comparació de dues o més mostres amb dades censurades en un interval. També hem desenvolupat algunes rutines amb S–Plus que implementen la versió permutacional del tests de Wilcoxon, Logrank i de la t de Student per a dades censurades en un interval (Fay and Shih, 1998). Aquesta part de la tesi doctoral es complementa en el capítol 5 amb diverses propostes d'extensió del test de Jonckeere. Amb l'objectiu de provar una tendència en el problema de k mostres, Abel (1986) va realitzar una de les poques generalitzacions del test de Jonckheere per a dades censurades en un interval. Nosaltres proposem altres generalitzacions d'acord amb els resultats presentats en el capítol 4. Utilitzem enfocaments permutacionals i de Monte Carlo. Proporcionem programes informàtics per a cada proposta i realitzem un estudi de simulació per tal de comparar la potència de cada proposta sota diferents models paramètrics i supòsits de tendència. Com a motivació de la metodologia, en els dos capítols s'analitza un conjunt de dades d'un estudi sobre els beneficis de la zidovudina en pacients en els primers estadis de la infecció del virus VIH (Volberding *et al.*, 1995).

Finalment, el capítol 6 resumeix els resultats i destaca aquells aspectes que s'han de completar en el futur.

Contents

Chapter 1

Introduction

At the beginning, this PhD thesis was motivated by the idea that existing k–sample testing methods for interval–censored data needed to be compiled and studied for further extension. The results from this research were part of the material for a seminar course that Guadalupe Gómez, M. Luz Calle and Ramon Oller did in 2001. Then, the work done in this seminar produced a survey paper about interval censoring (Gómez *et al.*, 2004) and motivated a new interest for this PhD thesis. From the interval censoring issues introduced in this seminar we questioned ourselves about the validity of the simplified likelihood function. Henceforth, this PhD thesis followed two lines of research. The first part of the present work (Chapter 2 and Chapter 3) is about theoretical foundations in the nonparametric estimation approach. The second part (Chapter 4 and Chapter 5) considers existing methods for the k–sample problem and gives new proposals.

This chapter deals with the basic concepts and notation needed for the subsequent chapters. In Section 1.1 we give the notion of interval–censored data. In the following sections we consider different aspects related to the nonparametric maximum likelihood estimator (NPMLE) of the lifetime distribution. Section 1.2 introduces the simplified likelihood function. Section 1.3 is devoted to maximization methods for the simplified likelihood. Section 1.4 describes the Turnbull's intervals, that is, the set of intervals where the nonparametric estimator concentrates its mass. Section 1.5 provides the notion of self–consistency, a concept which will play an important role in the development of this PhD thesis. Section 1.6 addresses computation of the NPMLE via the S-Plus package. Finally, we give an outline of the subsequent chapters in Section 1.7.

1.1 Interval censoring

Methods for lifetime data have been widely used in a large number of studies in medical and biological sciences. In this setting the lifetime variable of interest, T, is a positive random variable representing the time until the occurrence of a certain event $\mathcal{E}$. For instance, in the area of clinical and epidemiological studies, this event of interest is often the onset of a disease, the disappearance of disease's symptoms, or death. A key characteristic that distinguishes survival analysis from other areas in statistics is that lifetime data are usually censored. Censoring occurs when information about the lifetimes of some individuals is incomplete. Different circumstances can produce different types of censoring. It can be distinguished between right-censored data, left-censored data and interval-censored data.

Interval censoring mechanisms arise when the event of interest cannot be directly observed and it is only known to have occurred during a random interval of time. In this situation, the only information about the lifetime T is that it lies between two observed times L and R. We in fact formally observe a random censoring vector (L, R), such that $T \in \lfloor L, R \rfloor$ with probability one. We use the following bracket notation $\lfloor L, R \rfloor$ to indicate an interval that can be closed, open or half open depending on the interval censoring model. We find in the articles of Peto (1973) and Turnbull (1976) the first approach to the estimation of the distribution function when data are interval-censored. These authors consider closed intervals, $[L, R]$, so that exact observations are taking into account. We find in the literature other censoring mechanisms closely related to the concept of interval censoring as introduced by Peto and Turnbull. For example, if the event is only known to be larger or smaller than an observed monitoring time, the data conforms to the current status model or interval-censored data, case 1. In experiments with two monitoring times, U and V with $U < V$, where it is only possible to determine whether the event of interest occurs before the first monitoring time ($T \leq U$), between the two monitoring times ($U < T \leq V$), or after the last monitoring time ($T > V$), the observable data is known as interval–censored data, case 2. A natural extension of case 1 and case 2 models is the case k model, where k is a fixed number of monitoring times. Schick and Yu (2000) discuss an extended case k model where the number of monitoring times is random. In all these censoring schemes the intervals are half open and non–censored observations are not considered. Yu *et al.* (2000) generalize the case 2 model so that exact observations are allowed.

Many recent books concerning survival analysis have incorporated the interval censored data topic. We mention as references the last book editions of Kalbfleisch and Prentice (2002) and Lawless (2003) which give a comprehensive and modern approach to models and methods for lifetime data.

1.2 The simplified likelihood

A model for interval–censored data is described by the joint distribution, $F_{T,L,R}$, between the random variable T and the observables (L, R). The fact that $\lfloor L, R \rfloor$ contains T requires that the support of (T, L, R) is a subset of $\{(t, l, r) : 0 \leq l \leq t \leq r \leq +\infty\}$. We denote the lifetime distribution by

$$dW(t) = P(T \in dt),$$

and by

$$dF_{L,R}(l, r) = P(L \in dl, R \in dr, T \in \lfloor l, r \rfloor) \tag{1.1}$$

the contribution to the likelihood of an individual with observed interval $\lfloor l, r \rfloor$. Because the construction of the likelihood is not straightforward, the interval censoring problem has been generally treated via the nonparametric maximization of the simplified likelihood defined as the probability that T belongs to $\lfloor l, r \rfloor$. This likelihood considers the observed intervals as fixed in advance and ignores their randomness. If we consider a sample of n independent realizations of the observables, $(l_1, r_1), \ldots, (l_n, r_n)$, the simplified likelihood can be expressed as,

$$\mathbf{L}(W) = \prod_{i=1}^{n} \int_{\{t:t\in\lfloor l_i, r_i\rfloor\}} dW(t) = \prod_{i=1}^{n} P_W(\lfloor l_i, r_i \rfloor). \tag{1.2}$$

As uncensored observations are allowed, we assume in the sequel that $P_W(\{t\}) = dW(t)$.

The appropriateness of the simplified likelihood with interval–censored data has been based on the so-called noninformative conditions which have been introduced in the papers of Self and Grossman (1986) and Gómez *et al.* (2004). In a more general censoring framework, Heitjan and Rubin (1991), Heitjan (1993) and Gill *et al.* (1997) develop and characterize the analogous notion of coarsening at random conditions. In Oller *et al.* (2004) different characterizations for the noninformative condition are given and their equivalence is shown. They introduce a weaker condition, namely the constant–sum condition, which is sufficient for the validity of the simplified likelihood (1.2) in a nonparametric estimation of the lifetime probability distribution W. The constant–sum condition for interval censoring is an extension of the same notion in Williams and Lagakos (1977) or Ebrahimi *et al.* (2003), in the context of right censoring, and Betensky (2000), in the context of current status data.

1.3 Maximum likelihood estimation

The nonparametric likelihood estimator of $W(t)$ is a monotonically increasing function which maximizes the simplified likelihood function (1.2). The resulting estimator might not be unique because the likelihood for an interval-censored observation depends only on the difference between the survival values at the end-points of that interval and not at all on the detailed behavior within the interval. It is important to remark that computational results could be different if we treat intervals as closed, open or half open. The continuous nature of the variables would induce us to think that such a precision is not important. However, as it is exposed in Ng (2002), different interpretations of the intervals lead to different likelihood functions, which in turn could imply different nonparametric maximum likelihood estimates.

One of the first papers approaching the maximum likelihood estimation for interval–censored data is due to Peto (1973) who reports data from annual surveys on sexual maturity development of girls. Peto proposes a method based on maximizing the log–likelihood by a suitable constrained Newton–Raphson programmed search. Few years later, Turnbull (1976) approaches the more general problem of the analysis of arbitrarily grouped, censored and truncated data and derives a self–consistency method to obtain the nonparametric estimator of the distribution function. This method can be taken as a particular case of the expectation–maximization (EM) algorithm and can be applied, in particular, to deal with interval-censored situations. Few more years elapsed before these methods were applied in different setups, but these two pioneers papers are today the seed of most of the practical results. Moreover, many papers since then consider and discuss computational issues arising in the calculation of the nonparametric maximum likelihood estimation from censored data. For instance, Gentleman and Geyer (1994) provide standard convex optimization techniques to maximize the likelihood function and to check the uniqueness of the solution. Another example is the proposal in Groeneboom and Wellner (1992) who use isotonic regression theory in the interval censored–data model case 1 and case 2. This proposal implies the application of the convex minorant algorithm to determine the nonparametric maximum likelihood estimate. Wellner and Zhan (1997) improve this method with an hybrid algorithm based on a combination of the EM algorithm and a modified iterative convex minorant algorithm.

1.4 Turnbull's intervals

In what follows we introduce the set of intervals where the nonparametric estimator of the distribution function W concentrates its mass. As proposed in Turnbull (1976), we consider closed observed intervals $I_1 = [l_1, r_1], \ldots, I_n = [l_n, r_n]$. The definition below is easily modifiable to cover open or half open intervals. For instance, Gentleman and Geyer (1994) consider open intervals, Kalbfleisch and Prentice (2002) consider half open intervals and Yu *et al.* (2000) consider mixed interval censored data which include half open intervals and exact observations. From the sets $\mathcal{L} = \{l_i, 1 \leq i \leq n\}$ and $\mathcal{R} = \{r_i, 1 \leq i \leq n\}$ we can derive all the distinct closed intervals whose left and right end-points lie in the sets $\mathcal{L}$ and $\mathcal{R}$ respectively and which contain no other members of $\mathcal{L}$ or $\mathcal{R}$ other than at their left and right endpoints respectively. Let these intervals, known as Turnbull's intervals, be written in order as $[q_1, p_1], [q_2, p_2], \ldots, [q_m, p_m]$. We illustrate this construction with the following example.

Example 1.1. Suppose that the following $n = 6$ intervals have been observed $[0, 1]$, $[4, 6]$, $[2, 6]$, $[0, 3]$, $[2, 4]$, $[5, 7]$. Then, Turnbull's intervals are given by $[q_1, p_1] = [0, 1]$, $[q_2, p_2] = [2, 3]$, $[q_3, p_3] = [4, 4]$ and $[q_4, p_4] = [5, 6]$.

As noted by Peto (1973) and Turnbull (1976), any distribution function which increases outside Turnbull's intervals cannot be a maximum likelihood estimator of W. Moreover, the total likelihood is a function only of the amount that the distribution curve increases in the Turbull's intervals and is independent of how the increase actually occurs. Thus, the estimated distribution curve is unspecified in each $[q_j, p_j]$ and is well defined and flat between these intervals.

Denoting by $w_j = P_W([q_j, p_j])$ the weight of the j^{th} interval, $j = 1, \ldots, m - 1$,

$w_m = 1 - \sum_{j=1}^{m-1} w_j$, we can write down the simplified likelihood (1.2) as

$$\mathbf{L}(w_1, \ldots, w_{m-1}) = \prod_{i=1}^{n} \left(\sum_{j=1}^{m} \alpha_j^i w_j \right) \tag{1.3}$$

where the indicator $\alpha_j^i = \mathbf{1}_{\{[q_j,p_j] \subseteq [l_i,r_i]\}}$ expresses whether or not the interval $[q_j, p_j]$ is contained in $[l_i, r_i]$. The vectors $\boldsymbol{w} = (w_1, \ldots, w_m)$ define equivalence classes on the space of distribution functions W which are flat outside $\cup_{j=1}^{m}[q_j, p_j]$. Therefore, the maximum will be at best unique only up to equivalence classes and the problem of maximizing $\mathbf{L}$ has been reduced to the finite-dimensional problem of maximizing a function of $w_1, \ldots, w_{m-1}$ subject to the constraints $w_j \geq 0$ and $1 - \sum_{j=1}^{m-1} w_j \geq 0$.

The total likelihood, as a function of $w_1, \ldots, w_{m-1}$, is strictly convex (except on the boundaries of the constrained region on which the likelihood function is zero), so the values of $w_1, \ldots, w_{m-1}$ that maximize it are unique. Let $\hat{\boldsymbol{w}} = (\hat{w}_1, \ldots, \hat{w}_m)$ be the maximizing solution of (1.3). Turnbull's nonparametric estimator $\hat{W}$ for W is given by

$$\hat{W}(t) = \begin{cases} 0 & \text{if} \quad t < q_1 \\ \hat{w}_1 + \cdots + \hat{w}_k & \text{if} \quad p_k \leq t < q_{k+1}, \quad 1 \leq k \leq m-1 \\ 1 & \text{if} \quad t \geq p_m \end{cases} \tag{1.4}$$

and is not specified for $t \in [q_j, p_j]$, for $1 \leq j \leq m$.

The variances and covariances of the non zero $\hat{w}_k$ are given by the inverse of the second derivatives matrix of the logarithm of simplified likelihood (1.3) with respect to $w_1, \ldots, w_{m-1}$. However, there is no yet theoretical justification for this procedure, the problem being a violation of the usual assumption of a fixed number of unknown parameters that remains unchanged with increasing the sample size.

Example 1.2. The simplified likelihood corresponding to the previous 6 intervals in

Example 1.1 is given by

$$\begin{aligned} L_T(w_1, w_2, w_3, w_4) &= \prod_{i=1}^{6} \left(\sum_{j=1}^{4} \alpha_j^i P_W([q_j, p_j]) \right) \\ &= (w_1)(w_3 + w_4)(w_2 + w_3 + w_4)(w_1 + w_2)(w_2 + w_3)(w_4), \end{aligned}$$

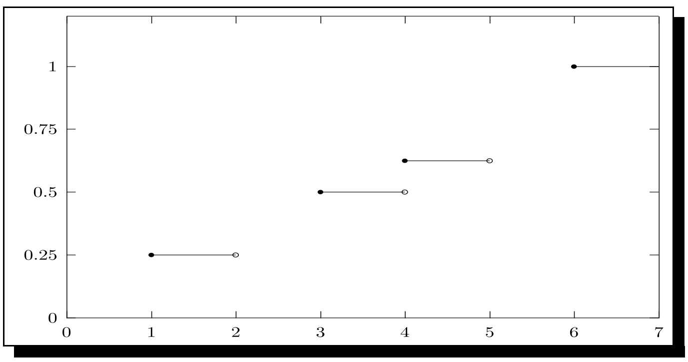
The maximizing solution is found at the point $(\hat{w}_1, \hat{w}_2, \hat{w}_3, \hat{w}_4) = (\frac{1}{4}, \frac{1}{4}, \frac{1}{8}, \frac{3}{8})$ and has estimated variance-covariance matrix equal to

$$\begin{pmatrix} 3/64 & -3/64 & 3/128 & -3/128 \\ -3/64 & 11/64 & -19/128 & 3/128 \\ 3/128 & -19/128 & 53/256 & -21/256 \\ -3/128 & 3/128 & -21/256 & 21/256 \end{pmatrix}$$

Thus Turnbull's nonparametric estimator $\hat{W}$ for W is given by

$$\hat{W}(t) = \begin{cases} 0 & \text{if} \quad t < 0 \\ \frac{1}{4} & \text{if} \quad 1 \le t < 2 \\ \frac{1}{2} = \frac{1}{4} + \frac{1}{4} & \text{if} \quad 3 \le t < 4 \\ \frac{5}{8} = \frac{1}{4} + \frac{1}{4} + \frac{1}{8} & \text{if} \quad 4 \le t < 5 \\ 1 & \text{if} \quad t \ge 6 \end{cases}$$

Figure 1.1: *Distribution function for the fictitious example. In regions* $[0,1]$, $[2,3]$, $[4,4]$, $[5,6]$ *the distribution function is not identified*



1.5 Self–consistency

We now introduce the concept of self–consistency and give its equivalence with the property of maximum likelihood. The idea of self-consistency was first used by Efron (1967) and it is applied in different fields of statistics, see Tarpey and Flury (1996). If $\hat{W}^0$ denotes the unknown empirical distribution function of the unobserved lifetimes $t_1, t_2, \ldots, t_n$,

$$\hat{W}^0(t) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{t_i \leq t\}},$$

then a distribution function $\hat{W}$ is called a self-consistent estimate of W when

$$\hat{W}(t) = E_{\hat{W}}(\hat{W}^0(t)|(l_1, r_1), \ldots, (l_n, r_n)).$$

In terms of the Turnbull's intervals, a self-consistent estimator of $\boldsymbol{w} = (w_1, \ldots, w_m)$ is defined to be any solution $\hat{\boldsymbol{w}} = (\hat{w}_1, \ldots, \hat{w}_m)$ of the following simultaneous equations:

$$\hat{w}_j = \frac{1}{n}\sum_{i=1}^{n} \frac{\alpha_j^i \hat{w}_j}{\sum_{l=1}^{m} \alpha_l^i \hat{w}_l} \quad 1 \leq j \leq m. \tag{1.5}$$

We note that the terms $\sum_{l=1}^{m} \alpha_l^i \hat{w}_l$ correspond to $P_{\hat{W}}([l_i, r_i])$, and the self–consistent equations (1.5) can also be written as

$$P_{\hat{W}}([q_j, p_j]) = \iint_{\{(l,r):[q_j,p_j]\subseteq[l,r]\}} \frac{P_{\hat{W}}([q_j, p_j])}{P_{\hat{W}}([l, r])} d\hat{F}^0_{L,R}(l, r) \quad 1 \leq j \leq m, \tag{1.6}$$

where $F^0_{L,R}$ denotes the empirical distribution function of the observed sample data, $(l_1, r_1), \ldots, (l_n, r_n)$,

$$\hat{F}^0_{L,R}(l, r) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{l_i \leq l, r_i \leq r\}}.$$

The maximization of the simplified likelihood (1.3) can be considered as a concave programming problem with linear constraints. Thus, as noted in Gentleman and Geyer

(1994), the Kuhn-Tucker conditions are necessary and sufficient for optimality. That is, $\hat{\boldsymbol{w}}$ is a maximum likelihood estimate if and only if, for every j, either

$$\iint_{\{(l,r):[q_j,p_j]\subseteq[l,r]\}} \frac{d\hat{F}^0_{L,R}(l,r)}{P_{\hat{W}}(\lfloor l,r\rfloor)} = \frac{1}{n}\sum_{i=1}^{n} \frac{\alpha^i_j}{\sum_{l=1}^{m}\alpha^i_l \hat{w}_l} = 1 \qquad \text{when } \hat{w}_j \neq 0, \tag{1.7}$$

or

$$\iint_{\{(l,r):[q_j,p_j]\subseteq[l,r]\}} \frac{d\hat{F}^0_{L,R}(l,r)}{P_{\hat{W}}(\lfloor l,r\rfloor)} = \frac{1}{n}\sum_{i=1}^{n} \frac{\alpha^i_j}{\sum_{l=1}^{m}\alpha^i_l \hat{w}_l} \leq 1 \qquad \text{when } \hat{w}_j = 0. \tag{1.8}$$

Thus, if $\hat{\boldsymbol{w}}$ is a maximum likelihood estimator for $\boldsymbol{w}$, then $\hat{\boldsymbol{w}}$ satisfies the self-consistent equations (1.6). Conversely, the solution $\hat{\boldsymbol{w}}$ of the self-consistent equations (1.6) is the nonparametric maximum likelihood estimator of $\boldsymbol{w}$ provided that condition (1.8) holds.

1.6 Computational aspects

If we first define $\mu^i_j(\boldsymbol{w}) = \frac{\alpha^i_j}{\sum_{l=1}^{m}\alpha^i_l w_l} w_j$ and $\tau_j(\boldsymbol{w}) = \frac{1}{n}\sum_{i=1}^{n}\mu^i_j(\boldsymbol{w})$, then the expectation–maximization (EM) algorithm runs as follows:

(A) Choose starting values $\boldsymbol{w^0} = (w^0_1, \ldots, w^0_m)$. This can be any set of positive numbers summing to unity.

(B) Expectation step: evaluate $\mu^i_j(\boldsymbol{w^0})$ for $i = 1, \ldots, n$ and $j = 1, \ldots, m$.

(C) Maximization step: obtain improved estimates $\boldsymbol{w^1} = (w^1_1, \ldots, w^1_m)$ by setting $w^1_j = \tau_j(\boldsymbol{w^0})$ for $j = 1, \ldots, m$.

(D) Return to step (B) with $\boldsymbol{w^1}$ replacing $\boldsymbol{w^0}$ till the required accuracy has been achieved.

(E) Denote by $\hat{\boldsymbol{w}} = (\hat{w}_1, \ldots, \hat{w}_m)$ the limiting solution and check Kuhn-Tucker condition (1.8). Finish if it is satisfied, otherwise go to step (A) and start with a different set of initial values.

The S–Plus software (version 6.0) provides a set of commands to perform survival analysis with interval–censored data. The EM algorithm used by this software considers semi–closed intervals $(L, R]$ where $L < T \leq R$ and incorporates exact, right-censored, and left-censored data. A vector `censor.codes` is first defined, it assigns a numerical value to each individual to distinguish whether the observation is exact (`censor.codes=1`), right-censored (`censor.codes=0`), left-censored (`censor.codes=2`) or interval-censored (`censor.codes=3`). Vectors `lower` and `upper` contain the lower and the upper limit, respectively, of the intervals. These are the objects that the procedure `kaplanMeier` needs in order to estimate the survival function using Turnbull's method, that is, `svf <- kaplanMeier(censor(lower, upper,censor.codes)~1)`. Plots of the estimated survival function can be obtained by either `plot(surv.est)` or `plot.kaplanMeier(surv.est)`.

It is important to note that this software occasionally returns a point that is local maximum but not is the nonparametric maximum likelihood estimator. See, for instance, Lawless (2003, page 138).

Example 1.3. We illustrate the S–Plus commands with the previous fictitious example. It is important to note that the intervals are closed and we should redefine the lower limits by subtracting a small quantity, say 0.01.

```
censor.codes <- c(3,3,3,3,3,3)
lower <- c(0,3.99,1.99,0,1.99,4.99)
upper <- c(1,6,6,3,4,7)
```

The estimated survival function we will obtain with the procedure `kaplanMeier` is the following:

```
Number  Observed:  6
Number  Censored:  6
Confidence  Type:  log
                   Survival  Std.Err  95% LCL  95% UCL
(-Inf,  0.00]      1.000     0.000    1.000        1
(   1,  1.99]      0.750     0.217    0.426        1
(   3,  3.99]      0.501     0.354    0.125        1
(   4,  4.99]      0.375     0.286    0.084        1
(   6,   Inf)      0.000     0.000       NA       NA
```

1.7 Outline of the subsequent chapters

Chapter 2 reproduces the paper of Oller *et al.* (2004). We prove the equivalence between different characterizations of noninformative censoring appeared in the literature and we define an analogous constant–sum condition to the one derived in the context of right censoring. We prove as well that when the noninformative condition or the constant–sum condition holds, the simplified likelihood can be used to obtain the nonparametric maximum likelihood estimator (NPMLE) of the failure time distribution function. Finally, we characterize the constant–sum property according to different types of censoring.

The work we introduce in Chapter 3 is a sequel of Oller *et al.* (2004) and it is under revision in an international journal. We study the relevance of the constant–sum property in the identifiability of the lifetime distribution. We show that the lifetime distribution is not identifiable outside the class of constant–sum models. We also show that the lifetime probabilities assigned to the observable intervals are identifiable inside the class of constant–sum models. We illustrate all these notions with several examples and situations.

Chapter 4 has partially been published in the survey paper of Gómez *et al.* (2004). It gives a general view of those procedures which have been applied in the nonparametric problem of the comparison of two or more interval–censored samples. We also propose a new test which generalizes the class of tests for right–censored data in Harrington and Fleming (1982). We have implemented with S–Plus functions those tests which are based on a permutational distribution method.

In Chapter 5 we propose different extensions of the Jonckeere's test. In order to test for an increasing trend in the k–sample problem, Abel (1986) gives one of the few generalizations of the Jonckheree's test for interval–censored data. We also suggest different Jonckheere–type tests according to the tests presented in Chapter 4. We use permutational and Monte Carlo approaches. We give computer programs for each proposal and perform a simulation study in order compare the power of each proposal under different parametric assumptions and different alternatives. We motivate Chapter 4 and Chapter 5 with the analysis of a set of data from a study of the benefits of zidovudine in patients in the early stages of the HIV infection (Volberding *et al.*, 1995).

Finally, Chapter 6 summarizes results and address those aspects which remain to be completed.

Chapter 2

Model characterizations for the validity of the simplified likelihood

As we have introduced in Chapter 1, inference methods with interval–censored data are mainly based on what we will refer to as the simplified likelihood, that is, the likelihood we would obtain if the censoring intervals were fixed in advance and we would ignore their randomness. Turnbull (1976), Groeneboom and Wellner (1992) and Shick and Yu (2000), among other authors, approach the estimation of the distribution function via the simplified likelihood. In this chapter we discuss different conditions under which such likelihood–based inferences are correct. Williams and Lagakos (1977), in the context of right censoring, and Betensky (2000), in the context of current status data, addressed an analogous problem. Sufficient conditions for the appropriateness of the simplified likelihood with interval–censored data are introduced in the papers of Self and Grossman (1986) and Gómez *et al.* (2004). In a more general censoring framework, Heitjan and Rubin (1991), Heitjan (1993) and Gill *et al.* (1997) develop and characterize a closely related concept, the so–called coarsening at random conditions.

This chapter adapts the paper of R. Oller, G. Gómez and M. L. Calle (2004) published in *The Canadian Journal of Statistics* to the terminology of this work. The remainder of this chapter is organized as follows. Section 2.1 introduces different non-informative censoring conditions and states their equivalences. In Section 2.2 we generalize the constant–sum condition introduced by Williams and Lagakos (1977) in the context of right–censoring. We distinguish between the constant–sum condition, which ensures that the inference process can omit the randomness of the intervals, and the noninformative conditions, which ensure that the censoring mechanism cannot affect the distribution of the time to the event of interest. We state the relationship between these two concepts. Section 2.3 reviews specific censoring schemes and states the general concepts of noninformativeness and constant–sum for these models.

2.1 Noninformative models

The goal of this work is to define conditions under which the contribution to the likelihood of an individual $dF_{L,R}(l,r) = P(L \in dl, R \in dr, T \in \lfloor l,r \rfloor)$ is proportional to $P_W(\lfloor l,r \rfloor)$, that is, the probability that T belongs to $\lfloor l,r \rfloor$ ignoring the censoring mechanism. The noninformative censoring condition is usually assumed to justify the use of the simplified likelihood. We will show later that, indeed, under a noninformative censoring mechanism, the nonparametric maximum likelihood estimator (NPMLE) of the lifetime distribution function, W, also maximizes the simplified likelihood. First, we introduce in Theorem 2.1 three definitions for noninformativeness of the interval censoring mechanism and we prove that they are equivalent. The first characterization has been proposed in Self and Grossman (1986) and it is in terms of the conditional distribution function of the lifetime variable given the observables, $F_{T|L,R}$. The second and third characterizations are in terms of the conditional distribution function of the observables given the lifetime variable, $F_{L,R|T}$. Gómez *et al.* (2004) use the second definition to derive the simplified likelihood, while the third definition follows from the coarsening at random notion used in Heitjan and Rubin (1991), Heitjan (1993) and Gill *et al.* (1997).

Theorem 2.1. *The following properties define the noninformative condition and are equivalent:*

(a) The conditional distribution of T *given* L *and* R *satisfies*

$$dF_{T|L,R}(t|l,r) = \frac{dW(t)}{P_W(\lfloor l,r \rfloor)} \; \mathbf{1}_{\{t \in \lfloor l,r \rfloor\}}$$

that is, censoring in $\lfloor l,r \rfloor$ *provides the same information as* T *being in* $\lfloor l,r \rfloor$*.*

(b) The conditional distribution of L *and* R *given* T *satisfies that*

$$dF_{L,R|T}(l,r|t) = \frac{dF_{L,R}(l,r)}{P_W(\lfloor l,r \rfloor)} \; \mathbf{1}_{\{t \in \lfloor l,r \rfloor\}} \tag{2.1}$$

that is, the observables (l,r) *are not influenced by the specific value of* T *in* $\lfloor l,r \rfloor$*.*

(c) The conditional distribution of L *and* R *given* T *satisfies that*

$$dF_{L,R|T}(l,r|t) = dF_{L,R|T}(l,r|t') \qquad \textit{on } \{(l,r) : t \in \lfloor l,r \rfloor \textit{ and } t' \in \lfloor l,r \rfloor\}$$

that is, two specific values of T *that are consistent with the observables always provide the same information.*

Proof:

(a) implies *(b)*:

If $dF_{T|L,R}(t|l,r) = \dfrac{dW(t)}{P_W(\lfloor l,r \rfloor)} \; \mathbf{1}_{\{t \in \lfloor l,r \rfloor\}}$, then for any (t,l,r) such that $t \in \lfloor l,r \rfloor$, following the usual rules for conditional distributions, we have

$$\begin{aligned} dF_{L,R|T}(l,r|t) &= \frac{dF_{T,L,R}(t,l,r)}{dW(t)} = \frac{dF_{T|L,R}(t|l,r)dF_{L,R}(l,r)}{dW(t)} \\ &= \frac{dW(t)dF_{L,R}(l,r)}{P_W(\lfloor l,r \rfloor)dW(t)} = \frac{dF_{L,R}(l,r)}{P_W(\lfloor l,r \rfloor)} \end{aligned}$$

(c) follows straightforwardly from *(b)*. The proof is omitted.

(c) implies *(a)*:

If $dF_{L,R|T}(l,r|t) = dF_{L,R|T}(l,r|t')$ on $\{(l,r) : t \in \lfloor l,r \rfloor \text{ and } t' \in \lfloor l,r \rfloor\}$, then for any fixed (t,l,r) such that $t \in \lfloor l,r \rfloor$

$$dF_{L,R}(l,r) = \int_{\{s:s\in\lfloor l,r\rfloor\}} dF_{T,L,R}(s,l,r) = \int_{\{s:s\in\lfloor l,r\rfloor\}} dF_{L,R|T}(l,r|s)dW(s)$$
$$= \int_{\{s:s\in\lfloor l,r\rfloor\}} dF_{L,R|T}(l,r|t)dW(s) = dF_{L,R|T}(l,r|t)P_W(\lfloor l,r\rfloor)$$

Then, if we use this last equality and we follow the usual rules for conditional distributions, we have

$$dF_{T|L,R}(t|l,r) = \frac{dF_{T,L,R}(t,l,r)}{dF_{L,R}(l,r)} = \frac{dF_{L,R|T}(l,r|t)\; dW(t)}{dF_{L,R|T}(l,r|t)\; P_W(\lfloor l,r\rfloor)} = \frac{dW(t)}{P_W(\lfloor l,r\rfloor)}$$

□

2.2 Constant–sum models

As mentioned before, the noninformative censoring condition allows to obtain the NPMLE of W using the simplified likelihood. That is, the observables can be treated as fixed in advance when making nonparametric inferences for the lifetime distribution function. Here we introduce a weaker condition, namely the constant–sum condition, which is sufficient for these inferences to be correct.

The following definition for the constant–sum condition extends that of Williams and Lagakos (1977) in the context of right censoring. The condition proposed here is based on the marginal laws of the censoring model, W and $F_{L,R}$.

Definition 2.2. A censoring model is constant–sum if and only if, for any $t \geq 0$ such that $dW(t) \neq 0$, the following equation holds

$$\iint_{\{(l,r):t\in\lfloor l,r\rfloor\}} \frac{dF_{L,R}(l,r)}{P_W(\lfloor l,r\rfloor)} = 1. \tag{2.2}$$

Theorem 2.3. *If a censoring model is constant–sum, the NPMLE of the lifetime distribution function also maximizes the simplified likelihood.*

Proof:

Let us start by denoting the support of the lifetime variable as $\mathcal{D}_W = \{t \geq 0\colon dW(t) \neq 0\}$. Consider a random sample $\lfloor l_1, r_1\rfloor, \ldots, \lfloor l_n, r_n\rfloor$, of (L, R). The logarithm of the full likelihood (equation (1.1)) can be written as

$$\sum_{i=1}^{n} \log P_W(\lfloor l_i, r_i\rfloor) + \sum_{i=1}^{n} \log dK(l_i, r_i) \tag{2.3}$$

where $P_W(\lfloor l, r\rfloor) = \int_{\mathcal{D}_W \cap \{t: t \in \lfloor l, r\rfloor\}} dW(t)$ and $dK(l, r) = P(L \in dl, R \in dr |\, T \in \lfloor l, r\rfloor)$.

The NPMLE of the pair (W, dK) is obtained by maximizing equation (2.3) subject to the constraints: (i) W is a distribution function with support $\mathcal{D}_W$, (ii) dK takes values in $[0, 1]$ and (iii) the following link between W and dK given by the total probability of the observables

$$\int\!\!\int_{\{0 \leq l \leq r\}} dF_{L,R}(l, r) = \int\!\!\int_{\{0 \leq l \leq r\}} P_W(\lfloor l, r\rfloor) dK(l, r) = 1. \tag{2.4}$$

Equation (2.4) can be equivalently written as

$$\int_{\mathcal{D}_W} \left(\int\!\!\int_{\{(l,r): t \in \lfloor l, r\rfloor\}} dK(l, r) \right) dW(t) = 1.$$

If we assume that the model is constant-sum, then in the maximization problem we have to add equation (2.2) as a new constraint (iv)

$$\int\!\!\int_{\{(l,r): t \in \lfloor l, r\rfloor\}} dK(l, r) = 1 \qquad \text{for any } t \in \mathcal{D}_W.$$

Condition (iii) derives from condition (iv), and consequently it can be omitted in the maximization problem. This means that equation (2.4) is no longer a constraint between W and dK. Thus, the NPMLE of each component of the pair (W, dK) can be obtained separately by maximizing the left–hand side of equation (2.3) under the constraint (i) and maximizing the right–hand side of equation (2.3) under the constraint (ii) and (iv). This proves the theorem because the left–hand side of equation (2.3) is the logarithm of the simplified likelihood for the given sample.

□

For the sake of completeness, it is interesting to note that for any $t \geq 0$ the constant–sum condition (2.2) can be expressed as well as

$$\iint_{\{(l,r):t\in\lfloor l,r\rfloor\}} \frac{dW(t)}{P_W(\lfloor l,r\rfloor)}\, dF_{L,R}(l,r) = dW(t). \tag{2.5}$$

Equation (2.5) is the well-known self–consistent equation which is the basis of the nonparametric maximum likelihood estimation of W, see Turnbull (1976).

In the rest of this section we discuss the relationship between the noninformative and the constant–sum conditions. The following two propositions show that the noninformative condition is sufficient but not necessary for a model to be constant–sum.

Proposition 2.4. *If a censoring model is noninformative then the model is constant–sum.*

Proof:

Indeed, for any $t \geq 0$ such that $dW(t) \neq 0$, it follows from Equation (2.1) that

$$\iint_{\{(l,r):t\in\lfloor l,r\rfloor\}} \frac{dF_{L,R}(l,r)}{P_W(\lfloor l,r\rfloor)} = \iint_{\{(l,r):t\in\lfloor l,r\rfloor\}} dF_{L,R|T}(l,r|t) = 1$$

and, consequently, the constant–sum condition holds.

□

This proposition together with Theorem 2.3 justifies the use of the simplified likelihood when the censoring mechanism is noninformative. Next proposition ensures that if the underlying model is constant–sum, there exists a full model that satisfies the noninformative condition and has the same marginal laws of the lifetimes and the observables.

Proposition 2.5. *If a censoring model, F_{T_1,L_1,R_1}, satisfies the constant–sum condition, then there always exists a noninformative model, F_{T_2,L_2,R_2}, such that $W_2 = W_1$ and $F_{L_2,R_2} = F_{L_1,R_1}$.*

Proof:

Define F_{T_2,L_2,R_2} by

$$dF_{T_2,L_2,R_2}(t,l,r) = \frac{dW_1(t)dF_{L_1,R_1}(l,r)}{P_{W_1}(\lfloor l,r\rfloor)}\ \mathbf{1}_{\{t\in\lfloor l,r\rfloor\}}.$$

Since model F_{T_1,L_1,R_1} is constant–sum, the model F_{T_2,L_2,R_2} defines a probability measure such that $T_2 \in \lfloor L_2, R_2\rfloor$ with probability one:

$$\begin{aligned}\iiint_{\{(t,l,r):t\in\lfloor l,r\rfloor\}} dF_{T_2,L_2,R_2}(t,l,r) &= \int_0^{+\infty} dW_1(t)\left(\iint_{\{(l,r):t\in\lfloor l,r\rfloor\}} \frac{dF_{L_1,R_1}(l,r)}{P_{W_1}(\lfloor l,r\rfloor)}\right)\\ &= \int_0^{+\infty} dW_1(t) = 1.\end{aligned}$$

Furthermore, for any $t \geq 0$, one has

$$\begin{aligned}dW_2(t) &= \iint_{\{(l,r):t\in\lfloor l,r\rfloor\}} dF_{T_2,L_2,R_2}(t,l,r)\\ &= dW_1(t)\left(\iint_{\{(l,r):t\in\lfloor l,r\rfloor\}} \frac{dF_{L_1,R_1}(l,r)}{P_{W_1}(\lfloor l,r\rfloor)}\right) = dW_1(t)\end{aligned}$$

and for any (l,r) such that $0 \leq l \leq r$,

$$dF_{L_2,R_2}(l,r) = \int_{\{t:t\in\lfloor l,r\rfloor\}} dF_{T_2,L_2,R_2}(t,l,r)$$

$$= dF_{L_1,R_1}(l,r)\left(\int_{\{t:t\in\lfloor l,r\rfloor\}} \frac{dW_1(t)}{P_{W_1}(\lfloor l,r\rfloor)}\right) = dF_{L_1,R_1}(l,r).$$

Finally, it follows that $F_{L_2,R_2|T_2}$ satisfies Equation (2.1) for any (t,l,r) such that $t \in \lfloor l,r\rfloor$ and $dW_2(t) \neq 0$, viz.

$$dF_{L_2,R_2|T_2}(l,r|t) = \frac{dF_{T_2,L_2,R_2}(t,l,r)}{dW_2(t)} = \frac{dW_1(t)dF_{L_1,R_1}(l,r)}{P_{W_1}(\lfloor l,r\rfloor)dW_2(t)} = \frac{dF_{L_2,R_2}(l,r)}{P_{W_2}(\lfloor l,r\rfloor)}.$$

□

Further discussion about the relationship between the noninformative and the constant–sum conditions is given by Lawless (2004). The author consider situations where an inspection process defines the censoring observations. Then, the independence between the inspection process and T implies that the noninformative condition, and consequently the constant–sum condition, holds. Moreover, when the inspection process depends on T, Lawless (2004) proves that the constant–sum property is equivalent to the existence of an alternative inspection process which is independent of T and which gives the same distribution for the observables, $F_{L,R}$, as the underlying true inspection process.

In the following example we illustrate that the constant–sum condition does not imply the noninformative condition.

Example 2.6. Here we present two related models sharing the same marginal distributions. The first one satisfies the constant–sum condition but not the noninformative censoring condition while the second one is noninformative.

Let $\mathcal{D}_W = \{0,1,2,3\}$ be the support of the lifetime variable and $\mathcal{D}_{F_{L,R}} = \{[0,0],[0,2],[1,1],[1,3],[2,2],[3,3]\}$ the observable censoring intervals. We consider the model determined by the joint probability between the lifetime variable and the observables,

$dF_{T,L,R}(t,l,r)$, given by Table 2.1. It is easy to verify that this model holds the constant–sum condition (2.2) for each $t \in \{0,1,2,3\}$. For instance, for $t = 1$ the constant–sum condition is

$$\sum_{\{(l,r):1\in\lfloor l,r\rfloor\}} \frac{dF_{L,R}(l,r)}{P_W(\lfloor l,r\rfloor)} = \frac{dF_{L,R}(0,2)}{P_W(\lfloor 0,2\rfloor)} + \frac{dF_{L,R}(1,1)}{P_W(\lfloor 1,1\rfloor)} + \frac{dF_{L,R}(1,3)}{P_W(\lfloor 1,3\rfloor)} =$$

$$= \frac{3/16}{3/4} + \frac{2/16}{1/4} + \frac{3/16}{3/4} = 1.$$

However, this model does not hold the noninformative condition. For instance, $dF_{L,R|T}(0,2|0) = 1/4$, while $dF_{L,R|T}(0,2|1) = 0$ and $dF_{L,R|T}(0,2|2) = 1/2$, so condition (c) in Theorem 2.1 fails.

Table 2.1: *Joint probability $dF_{T,L,R}$ of a constant–sum model.*

t \ $\lfloor l,r\rfloor$	[0,0]	[0,2]	[1,1]	[1,3]	[2,2]	[3,3]	$dW(t)$
0	3/16	1/16	0	0	0	0	1/4
1	0	0	2/16	2/16	0	0	1/4
2	0	2/16	0	0	2/16	0	1/4
3	0	0	0	1/16	0	3/16	1/4
$dF_{L,R}(l,r)$	3/16	3/16	2/16	3/16	2/16	3/16	1

In Table 2.2, we used Proposition 2.5 to construct a noninformative version of the above model. Note that, since both models have the same marginal distributions, they are indistinguishable on the basis of repeated observations and inferences for the lifetime probabilities will lead to the same estimate of $(dW(0), dW(1), dW(2), dW(3)) = (1/4, 1/4, 1/4, 1/4)$.

Table 2.2: *Joint probability $dF_{T,L,R}$ of a noninformative model.*

$\lfloor l, r \rfloor$ / t	[0,0]	[0,2]	[1,1]	[1,3]	[2,2]	[3,3]	$dW(t)$
0	3/16	1/16	0	0	0	0	1/4
1	0	1/16	2/16	1/16	0	0	1/4
2	0	1/16	0	1/16	2/16	0	1/4
3	0	0	0	1/16	0	3/16	1/4
$dF_{L,R}(l,r)$	3/16	3/16	2/16	3/16	2/16	3/16	1

2.3 Censoring models

We discuss the meaning of the noninformative and constant–sum conditions for the particular cases of right–censored data, double–censored data and interval–censored data case k. The results for right–censored data and interval–censored data case 1 are similar to those in Williams and Lagakos (1977) and Betensky (2000), respectively.

2.3.1 Right–censored data

Right censored-data arise when the event of interest can only be observed if the lifetime does not exceed the value of a positive random censoring variable, C. The observed data for an individual is traditionally expressed by the pair (X, δ) where $X = \min(T, C)$ and $\delta = \mathbf{1}_{\{T \leq C\}}$. Using interval censoring notation, the vector of observables is,

$$(L, R) = (T, T) \cdot \delta + (C, +\infty) \cdot (1 - \delta)$$

and the observable intervals are defined as

$$\lfloor l, r \rfloor = \begin{cases} [l, r] & \text{if } l = r \\ (l, r) & \text{if } r = +\infty. \end{cases}$$

Thus, the joint distribution function for L, R, T is given by:

$$dF_{T,L,R}(t,l,r) = \begin{cases} P(T \in dt, C \geq t) & \text{if} \quad l = t = r \\ P(T \in dt, C \in dl) & \text{if} \quad l < t \quad \text{and} \quad r = +\infty \\ 0 & \text{otherwise.} \end{cases}$$

Following analogous steps to the proof of Proposition 2.7 given in Subsection 2.3.3, the noninformative condition and the constant–sum condition are respectively given by:

$$P(C \in dl|T = t) = P(C \in dl|T > l) \quad \text{for any } t > l > 0$$

$$P(C \geq t|T = t) + \int_0^{t^-} P(C \in dl|T > l) = 1 \qquad \text{for any } t \geq 0 \text{ such that } dW(t) \neq 0.$$

When T and C are continuous, Kalbfleisch and MacKay (1979) show that the constant–sum condition is equivalent to the following relationship between hazard functions,

$$P(T \in dt|C \geq t, T \geq t) = P(T \in dt|T \geq t) \qquad \text{for any } t > 0\ .$$

The characterization in terms of the hazard functions is easier to interpret and can be viewed as a kind of noninformative condition. However, as it is shown in Proposition 2.4, the constant–sum condition is weaker than the non–informative condition defined in this paper. Williams and Lagakos (1977) give an example of a right–censored model which is constant–sum but informative. We also note that if the variables T and C are independent, then the constant–sum condition as well as the noninformative condition are satisfied.

2.3.2 Doubly–censored data

Data is said to be doubly–censored when the event of interest can only be observed inside the window $[C_1, C_2]$, where C_1 and C_2 are positive random variables and $C_1 < C_2$ (Chang and Yang, 1987). The observed data for an individual is of the form (X, δ, γ)

where $\delta = \mathbf{1}_{\{T<C_1\}}$, $\gamma = \mathbf{1}_{\{T\leq C_2\}}$ and $X = C_1 \cdot \delta + T \cdot (1-\delta)\cdot\gamma + C_2 \cdot (1-\delta)\cdot(1-\gamma)$. In the interval censoring framework, the vector of observables can be expressed as

$$(L,R) = (0,C_1)\cdot\delta + (T,T)\cdot(1-\delta)\cdot\gamma + (C_2,+\infty)\cdot(1-\delta)\cdot(1-\gamma)$$

and intervals are defined as

$$\lfloor l,r \rfloor = \begin{cases} [l,r) & \text{if } l=0 \\ [l,r] & \text{if } l=r \\ (l,r) & \text{if } r=+\infty. \end{cases}$$

In this model the joint probability law of the lifetimes and the observables is given by,

$$dF_{T,L,R}(t,l,r) = \begin{cases} P(T\in dt, C_1\in dr) & \text{if} \quad l=0 \quad \text{and} \quad t<r \\ P(T\in dt, C_1\leq t, C_2\geq t) & \text{if} \quad l=t=r \\ P(T\in dt, C_2\in dl) & \text{if} \quad l<t \quad \text{and} \quad r=+\infty \\ 0 & \text{otherwise.} \end{cases}$$

Under a double censoring setup the noninformative condition is expressed through the following two equalities:

- $P(C_1 \in dr|T=t) = P(C_1\in dr|T<r)$ for any $0<t<r$
- $P(C_2 \in dl|T=t) = P(C_2\in dl|T>l)$ for any $t>l>0$.

Furthermore, for any $t\geq 0$ such that $dW(t)\neq 0$, the constant–sum condition reduces to

$$\int_t^{+\infty} P(C_1\in dr|T<r) + P(C_1\leq t, C_2\geq t|T=t)$$

$$+\int_0^{t^-} P(C_2\in dl|T>l) = 1.$$

We observe again that independence between T and (C_1,C_2) implies both conditions. Further details on the development are similar to those given below in Proposition 2.7.

2.3.3 Interval–censored data, case k

This interval censoring scheme has been widely studied, specially the case 1 and case 2 (Groeneboom and Wellner, 1992; Schick and Yu, 2000). In the interval–censored model, case 1 or current status data, the event is only known to be larger or smaller than an observed monitoring time. The interval–censored model, case 2, considers two monitoring times, X_1 and X_2 with $X_1 < X_2$, where it is only possible to determine whether the event of interest occurs before the first monitoring time ($T \leq X_1$), between the two monitoring times ($X_1 < T \leq X_2$), or after the last monitoring time ($T > X_2$). Although interval censoring case 2 looks like the double censoring model, it is fundamentally different because the value of T is unknown inside the window $(X_1, X_2]$. The general case k model considers k positive random monitoring times, $X_1 < \cdots < X_k$, such that the event of interest can only be determined to have occurred before, between or after those times. The vector of observables is

$$(L, R) = (0, X_1)\mathbf{1}_{\{T \leq X_1\}} + \sum_{j=2}^{k} \left\{(X_{j-1}, X_j)\mathbf{1}_{\{X_{j-1} < T \leq X_j\}}\right\} + (X_k, +\infty)\mathbf{1}_{\{T > X_k\}}.$$

Thus, the intervals are defined as,

$$\lfloor l, r \rfloor = \begin{cases} (l, r) & \text{if } r = +\infty. \\ (l, r] & \text{otherwise} \end{cases}$$

The joint distribution function for L, R, T is expressed as

$$dF_{T,L,R}(t, l, r) = \begin{cases} P(T \in dt, X_1 \in dr) & \text{if } l = 0 \text{ and } t \leq r \\ \displaystyle\sum_{j=2}^{k} P(T \in dt, X_{j-1} \in dl, X_j \in dr) & \text{if } 0 < l < t \leq r < +\infty \\ P(T \in dt, X_k \in dl) & \text{if } l < t \text{ and } r = +\infty \\ 0 & \text{otherwise.} \end{cases}$$

Proposition 2.7. *In the case k interval censoring model the noninformative condition can be written as,*

- $P(X_1 \in dr|T=t) = P(X_1 \in dr|T \le r)$ for any $0 < t \le r$

- $\sum_{j=2}^{k} P(X_{j-1}\in dl, X_j \in dr|T=t) = \sum_{j=2}^{k} P(X_{j-1}\in dl, X_j\in dr|l < T \le r)$

  for any $0 < l < t \le r$

- $P(X_k \in dl|T=t) = P(X_k \in dl|T > l)$ for any $t > l > 0$

and, for any $t \ge 0$ *such that* $dW(t) \ne 0$*, the constant–sum equation is*

$$\int_{t-}^{+\infty} P(X_1 \in dr|T \le r) + \sum_{j=2}^{k} \int_0^{t^-}\int_{t^-}^{+\infty} P(X_{j-1} \in dl, X_j \in dr|l < T \le r)$$

$$+ \int_0^{t^-} P(X_k \in dl|T > l) = 1.$$

Proof:

The definition of the case k interval censoring model, $F_{T,L,R}$, implies that

$$dF_{L,R|T}(l,r|t) = \begin{cases} P(X_1 \in dr|T=t) & \text{if } l=0 \text{ and } t \le r \\ \sum_{j=2}^{k} P(X_{j-1} \in dl, X_j \in dr|T=t) & \text{if } 0< l <t \le r<+\infty \\ P(X_k \in dl|T=t) & \text{if } l < t \text{ and } r=+\infty \\ 0 & \text{otherwise} \end{cases}$$

and

$$dF_{L,R}(l,r) = \begin{cases} P(T \le r, X_1 \in dr) & \text{if } l=0 \\ \sum_{j=2}^{k} P(l < T \le r, X_{j-1} \in dl, X_j \in dr) & \text{if } 0< l <r<+\infty \\ P(T > l, X_k \in dl) & \text{if } r=+\infty \\ 0 & \text{otherwise} \end{cases}$$

Thus,

$$\frac{dF_{L,R}(l,r)}{P_W(\lfloor l,r \rfloor)} = \begin{cases} P(X_1 \in dr | T \leq r) & \text{if } l = 0 \\ \sum_{j=2}^{k} P(X_{j-1} \in dl, X_j \in dr | l < T \leq r) & \text{if } 0 < l < r < +\infty \\ P(X_k \in dl | T > l) & \text{if } r = +\infty \\ 0 & \text{otherwise} \end{cases}$$

If we impose the second characterization of the noninformative condition in terms of the above expressions, then the following equations should be satisfied:

- If $l = 0$ and $t \leq r$, $P(X_1 \in dr | T = t) = P(X_1 \in dr | T \leq r)$.
- If $0 < l < t \leq r < +\infty$, $\sum_{j=2}^{k} P(X_{j-1} \in dl, X_j \in dr | T = t) = \sum_{j=2}^{k} P(X_{j-1} \in dl, X_j \in dr | l < T \leq r)$.
- If $t > l$ and $r = +\infty$, $P(X_k \in dl | T = t) = P(X_k \in dl | T > l)$.

Furthermore, for any $t \geq 0$ such that $dW(t) \neq 0$, the constant–sum condition can be written as:

$$\int_{t-}^{+\infty} P(X_1 \in dr | T \leq r) + \sum_{j=2}^{k} \int_{0}^{t^-} \int_{t^-}^{+\infty} P(X_{j-1} \in dl, X_j \in dr | l < T \leq r)$$

$$+ \int_{0}^{t^-} P(X_k \in dl | T > l) = 1$$

□

Proposition 2.8. *If T is a positive continuous random variable with $dW(t) \neq 0$ for any $t > 0$, then the constant–sum condition reduces to the following equality*

$$P(X_1 \in dt | T \leq t) + \sum_{j=2}^{k} \int_{\{l: l \in (0,t)\}} P(X_{j-1} \in dl, X_j \in dt | l < T \leq t)$$

$$= \sum_{j=2}^{k} \int_{\{r:r\in(t,+\infty)\}} P(X_{j-1} \in dt, X_j \in dr | t < T \leq r) + P(X_k \in dt | T > t).$$

Proof:

This result can be shown in a general interval censoring model supposing that intervals cannot be singletons and T is a positive random variable with $dW(t) \neq 0$ for any $t > 0$. Then, we have to prove that the constant–sum condition reduces to

$$\int_{\{l:l\in[0,t)\}} \frac{dF_{L,R}(l,t)}{P_W((l,t])} = \int_{\{r:r\in(t,+\infty]\}} \frac{dF_{L,R}(t,r)}{P_W((t,r])}.$$

We will suppose that $\lfloor l, r \rfloor = (l, r]$ without loss of generality.

If the constant–sum condition holds, and we define $dK(l,r) = \dfrac{dF_{L,R}(l,r)}{P_W(\lfloor l,r \rfloor)}$, then for any $t > 0$ it follows that

$$1 = \int_{[0,t)} \int_{[t,+\infty]} dK(l,r).$$

This property implies that for any $0 < a < b < +\infty$

$$\int_{[0,a)} \int_{[a,+\infty]} dK(l,r) = \int_{[0,b)} \int_{[b,+\infty]} dK(l,r).$$

By splitting the support of the intervals as $[a, +\infty] = [a, b) \cup [b, +\infty]$ and $[0, b) = [0, a) \cup [a, b)$ and simplifying terms, we see that that this equality can also be expressed as

$$\int_{[0,a)} \int_{[a,b)} dK(l,r) = \int_{[a,b)} \int_{[b,+\infty]} dK(l,r)$$

From the set equivalence $[a, r) \times [a, b) = [a, b) \times (l, b)$, the previous equation expands to

$$\int_{[0,r)} \int_{[a,b)} dK(l,r) = \int_{[a,b)} \int_{(l,+\infty]} dK(l,r)$$

Thus, we have proved that $\int_{\{l:l\in[0,t)\}} dK(l,t) = \int_{\{r:r\in(t,+\infty]\}} dK(t,r)$ for any interval $[a,b)$. Using a monotone class theorem this result extends to the σ-algebra on $(0,+\infty)$.

To prove the reciprocal suppose that for any $t > 0$

$$\int_{\{l:l\in[0,t)\}} dK(l,t) = \int_{\{r:r\in(t,+\infty]\}} dK(t,r).$$

Then, it follows, from stepping backward the above equivalences, that for any $0 < a < b < +\infty$

$$\int_{[0,a)} \int_{[a,+\infty]} dK(l,r) = \int_{[0,b)} \int_{[b,+\infty]} dK(l,r) = k,$$

which in turn means that

$$\int_0^{+\infty} \left(\iint_{\{(l,r):t\in(l,r]\}} dK(l,r) \right) dW(t) = k$$

and this equality is only possible if $k = 1$.

□

Again, when the model satisfies the usual assumption of independence between the lifetime, T, and the monitoring times, $(X_1, \ldots, X_k)$, all the equations in Proposition 2.7 and Proposition 2.8 hold.

Chapter 3

Identifiability and the constant–sum property

In Chapter 2 different characterizations for the noninformative condition are given and their equivalence is shown. We have introduced, as well, a weaker condition, namely the constant–sum condition, which is sufficient for the validity of the simplified likelihood (1.2) in a nonparametric estimation of the lifetime distribution W. The motivation of the constant–sum property could be found in situations where an inspection process defines the censoring observations. In these particular settings the independence between the inspection process and T implies that the noninformative condition, and consequently the constant–sum condition, holds. Moreover, when the inspection process depends on T, Lawless (2004) proves that the constant–sum property is equivalent to the existence of an alternative inspection process which is independent of T and which gives the same distribution for the observables, $F_{L,R}$, as the underlying true inspection process. The constant–sum property is a central concept in the development of this chapter.

The work we introduce here is a sequel of Oller *et al.* (2004) and it is under revision in an international journal. The present chapter is devoted to study the identifiability of the lifetime distribution W on the basis of the assumed support of the lifetimes $\mathcal{D}_W = \{t \geq 0 : dW(t) > 0\}$ and the distribution for the observables $F_{L,R}$. This problem was already approached in the early papers of Tsiatis (1975) and Williams and Lagakos

(1977) and in more recent papers such as Wang *et al.* (1994), Betensky (2000) and Ebrahimi *et al.* (2003). The results of this chapter emphasize the importance of the constant–sum condition to ensure identifiability of W. Section 3.1 shows that W is not identifiable outside the class of constant–sum models. On the other hand, Section 3.2 shows that inside the class of constant–sum models probabilities assigned by W to the observable intervals $\lfloor l, r \rfloor$ are identifiable.

3.1 Nonidentifiability outside the class of constant–sum models

Throughout the chapter it is assumed a known support for the lifetime variable T, $\mathcal{D}_W = \{t \geq 0 : dW(t) > 0\}$, which is not necessarily equal to the usual assumption $\mathcal{D}_W = (0, \infty)$. Definition 3.1 formally gives the notion of nonidentifiability and Proposition 3.2 gives a constructive way of obtaining censoring models with W being nonidentifiable.

Definition 3.1. Given a censoring model $F_{T,L,R}$, we say that W is nonidentifiable when there exists a censoring model having different lifetime distribution but sharing the same lifetime support $\mathcal{D}_W$ and the same distribution for the observables $F_{L,R}$.

A question which naturally arises when thinking of identifiability is whether any set $\mathcal{D} \subset [0, +\infty)$ and any bivariate distribution G with support contained in $\{(l, r) : 0 \leq l \leq r \leq +\infty\}$ could respectively be the lifetime support and the observables distribution of a censoring model. As it is seen in Proposition 3.2 below, the answer is affirmative provided the following relationship between $\mathcal{D}$ and G holds:

$$\iint_{\{(l,r):t\in\lfloor l,r\rfloor\}} dG(l, r) > 0 \qquad \forall t \in \mathcal{D}. \tag{3.1}$$

Relationship (3.1) is a necessary constraint between the lifetime support $\mathcal{D}_W$ and the distribution for the observables $F_{L,R}$ in a censoring model. That is, for any $t \in \mathcal{D}_W$,

$$dW(t) = \iint_{\{(l,r):t\in\lfloor l,r\rfloor\}} dF_{T,L,R}(t,l,r) = \iint_{\{(l,r):t\in\lfloor l,r\rfloor\}} dF_{T|L,R}(t|l,r)dF_{L,R}(l,r) > 0$$

and, consequently, the set $\{(l,r) : t \in \lfloor l,r\rfloor\}$ is a $dF_{L,R}$–measurable non null set. Intuitively, relationship (3.1) ensures that any lifetime value $t \in \mathcal{D}_W$ is contained in at least one observable interval.

Proposition 3.2. *If $(\mathcal{D}, G)$ is a pair which satisfies equation (3.1) and F is any distribution function with support $\mathcal{D}$ satisfying $P_F(\lfloor l,r\rfloor) > 0$ dG–almost surely, then*

$$dF_{T,L,R}\,(t,l,r) = \frac{dF(t)\;dG(l,r)}{P_F(\lfloor l,r\rfloor)}\,\mathbf{1}_{\{t\in\lfloor l,r\rfloor\}} \tag{3.2}$$

defines a censoring model such that $(\mathcal{D}_W, F_{L,R}) = (\mathcal{D}, G)$.

Proof:

We first show that

$$dF_{T,L,R}\,(t,l,r) = \frac{dF(t)\;dG(l,r)}{P_F(\lfloor l,r\rfloor)}\,\mathbf{1}_{\{t\in\lfloor l,r\rfloor\}}$$

defines a probability measure such that $T \in \lfloor L,R\rfloor$ with probability one,

$$\iiint_{\{(t,l,r):t\in\lfloor l,r\rfloor\}} dF_{T,L,R}(t,l,r) = 1$$

This result is easily seen since F is a distribution function which satisfies that $P_F(\lfloor l,r\rfloor) > 0$ dG–almost surely and G is a bivariate distribution with support contained in $\{(l,r) : 0 \le l \le r \le +\infty\}$,

$$\iiint_{\{(l,r,t):t\in\lfloor l,r\rfloor\}} dF_{T,L,R}(t,l,r) = \iint_{\{(l,r):\,0\le l\le r\le+\infty\}} \left(\int_{\{t:t\in\lfloor l,r\rfloor\}} \frac{dF(t)}{P_F(\lfloor l,r\rfloor)}\right)\,dG(l,r)$$

$$= \iint_{\{(l,r):\ 0 \le l \le r \le +\infty\}} dG(l,r) = 1$$

To prove that $\mathcal{D}_W = \mathcal{D}$ we note that the marginal lifetime distribution which derives from this censoring model satisfies

$$dW(t) = \iint_{\{(l,r):t \in \lfloor l,r \rfloor\}} dF_{T,L,R}(t,l,r) = dF(t) \left(\iint_{\{(l,r):t \in \lfloor l,r \rfloor\}} \frac{dG(l,r)}{P_F(\lfloor l,r \rfloor)} \right)$$

Since F is a distribution function which has support $\mathcal{D}$, equality $\mathcal{D}_W = \mathcal{D}$ is reached from the fact that the pair $(\mathcal{D}, G)$ holds equation (3.1) and, consequently, the last element of the right–hand side in equation above is not null for all $t \in \mathcal{D}$.

Finally, it follows that any pair (l,r) satisfies that

$$dF_{L,R}(l,r) = \int_{\{t:t \in \lfloor l,r \rfloor\}} dF_{T,L,R}(t,l,r) = dG(l,r) \left(\int_{\{t:t \in \lfloor l,r \rfloor\}} \frac{dF(t)}{P_F(\lfloor l,r \rfloor)} \right) = dG(l,r).$$

□

This proposition gives a constructive way of obtaining two different censoring models sharing the same pair $(\mathcal{D}_W, F_{L,R})$ but different probability lifetime distributions. In the following example, we use this result to illustrate the notion of nonidentifiability.

Example 3.3. We construct two different models with different lifetime distributions W_1 and W_2 but with the same lifetime support $\mathcal{D}_W = \{0,1,2,3,4\}$ and the same censoring intervals $\{[0,1],[0,2],[2,4],[3,4]\}$ having $dF_{L,R}$ observable probability equal to $\{1/6, 1/3, 1/3, 1/6\}$, respectively. In order to build these two models we use equation (3.2) with two different auxiliary distributions F_1 and F_2 defined as $(dF_1(0), dF_1(1), dF_1(2), dF_1(3), dF_1(4)) = (1/5, 1/5, 1/5, 1/5, 1/5)$ and $(dF_2(0), dF_2(1), dF_2(2), dF_2(3), dF_2(4)) = (1/8, 1/8, 1/2, 1/8, 1/8)$. We note that the pair $(\mathcal{D}_W, F_{L,R})$ and the auxiliary distributions F_1 and F_2 satisfy conditions in Proposition 3.2. Now, from equation (3.2), the joint probabilities $dF_{T_1,L_1,R_1}(t,l,r)$ and $dF_{T_2,L_2,R_2}(t,l,r)$ for each model are respectively given by Table 3.1 and Table 3.2.

Table 3.1: *Joint probability dF_{T_1,L_1,R_1} of a non constant–sum model.*

t \ $\lfloor l,r \rfloor$	[0,1]	[0,2]	[2,4]	[3,4]	$dW_1(t)$
0	1/12	1/9	0	0	7/36
1	1/12	1/9	0	0	7/36
2	0	1/9	1/9	0	8/36
3	0	0	1/9	1/12	7/36
4	0	0	1/9	1/12	7/36
$dF_{L_1,R_1}(l,r)$	1/6	1/3	1/3	1/6	1

Table 3.2: *Joint probability dF_{T_2,L_2,R_2} of a non constant–sum model.*

t \ $\lfloor l,r \rfloor$	[0,1]	[0,2]	[2,4]	[3,4]	$dW_2(t)$
0	1/12	1/18	0	0	5/36
1	1/12	1/18	0	0	5/36
2	0	2/9	2/9	0	16/36
3	0	0	1/18	1/12	5/36
4	0	0	1/18	1/12	5/36
$dF_{L_2,R_2}(l,r)$	1/6	1/3	1/3	1/6	1

Note that these two models F_{T_1,L_1,R_1} and F_{T_2,L_2,R_2} share the same pair $(\mathcal{D}_W, F_{L,R})$ but have different lifetime distributions W_1 and W_2. Other censoring models with the same pair $(\mathcal{D}_W, F_{L,R})$ could also be built. Thus, it is clear that the pair $(\mathcal{D}_W, F_{L,R})$ does not identify the marginal lifetime distribution without additional assumption. It can be verified that neither model holds the constant–sum condition (2.2). For instance for model F_{T_1,L_1,R_1} and $t = 1$

$$\sum_{\{(l,r):1\in\lfloor l,r\rfloor\}} \frac{dF_{L,R}(l,r)}{P_{W_1}(\lfloor l,r\rfloor)} = \frac{dF_{L,R}(0,1)}{P_{W_1}([0,1])} + \frac{dF_{L,R}(0,2)}{P_{W_1}([0,2])} = \frac{1/6}{14/36} + \frac{1/3}{22/36} = 75/77 \neq 1.$$

Moreover, each model assigns different lifetime probabilities to the observable intervals $\lfloor l,r\rfloor$. For instance, in the first censoring model $P_{W_1}([0,2]) = 22/36$ while in the second model $P_{W_2}([0,2]) = 26/36$. As we show in next section, the constant–sum property is a sufficient condition for identifying at least the lifetime probabilities assigned to the observable intervals $\lfloor l, r \rfloor$.

3.2 Identifiability inside the class of constant–sum models

The following theorem shows that the probabilities assigned by the lifetime distribution to the observable intervals $\lfloor l, r \rfloor$ can be identified from the pair $(\mathcal{D}_W, F_{L,R})$ within the class of constant–sum models. This result enlightens the importance of the assumed support of the lifetime variable. However, additional conditions on the observables support will be necessary to assure the complete identifiability of W.

Theorem 3.4. *Let $F_{T,L,R}$ and F_{T^*,L^*,R^*} be constant–sum models so that $(\mathcal{D}_W, F_{L,R}) = (\mathcal{D}_{W^*}, F_{L^*,R^*})$, then $P_W(\lfloor l, r \rfloor) = P_{W^*}(\lfloor l, r \rfloor)$ $dF_{L,R}$–almost surely.*

Proof:

If model $F_{T,L,R}$ is constant–sum and $\mathcal{D}_W = \mathcal{D}_{W^*}$ then

$$\iint_{\{(l,r): t \in \lfloor l,r \rfloor\}} \frac{dF_{L,R}(l,r)}{P_W(\lfloor l, r \rfloor)} = 1 \qquad \forall t \in \mathcal{D}_{W^*},$$

which implies that

$$\int_0^{+\infty} dW^*(t) \iint_{\{(l,r): t \in \lfloor l,r \rfloor\}} \frac{dF_{L,R}(l,r)}{P_W(\lfloor l, r \rfloor)} = 1,$$

which can be rewritten as

$$\iint_{\{0 \leq l \leq r \leq +\infty\}} \frac{P_{W^*}(\lfloor l, r \rfloor)}{P_W(\lfloor l, r \rfloor)} dF_{L,R}(l,r) = 1. \tag{3.3}$$

Analogously, it is clear that starting with model F_{T^*,L^*,R^*} it follows that

$$\iint_{\{0\leq l\leq r\leq +\infty\}} \frac{P_W(\lfloor l,r\rfloor)}{P_{W^*}(\lfloor l,r\rfloor)} dF_{L^*,R^*}(l,r) = 1. \tag{3.4}$$

Equations (3.3) and (3.4), identity $F_{L,R} = F_{L^*,R^*}$ and Lemma A.1 given in Appendix A prove the theorem.

□

Example 3.5. We illustrate Theorem 3.4 with two constant–sum models, F_{T_3,L_3,R_3} and F_{T_4,L_4,R_4}, both having the same pair $(\mathcal{D}_W, F_{L,R})$ introduced in Example 3.3. From the joint probability $dF_{T_3,L_3,R_3}(t,l,r)$ and $dF_{T_4,L_4,R_4}(t,l,r)$ respectively given by Table 3.3 and Table 3.4, it is easy to verify the constant–sum property.

Table 3.3: *Joint probability dF_{T_3,L_3,R_3} of a constant–sum model.*

t \ $\lfloor l,r\rfloor$	[0,1]	[0,2]	[2,4]	[3,4]	$dW_3(t)$
0	1/24	3/24	0	0	1/6
1	3/24	1/24	0	0	1/6
2	0	1/6	1/6	0	1/3
3	0	0	1/24	3/24	1/6
4	0	0	3/24	1/24	1/6
$dF_{L_3,R_3}(l,r)$	1/6	1/3	1/3	1/6	1

Table 3.4: *Joint probability dF_{T_4,L_4,R_4} of a constant–sum model.*

t \ $\lfloor l,r\rfloor$	[0,1]	[0,2]	[2,4]	[3,4]	$dW_4(t)$
0	1/36	3/36	0	0	1/9
1	5/36	3/36	0	0	2/9
2	0	1/6	1/6	0	1/3
3	0	0	3/36	5/36	2/9
4	0	0	3/36	1/36	1/9
$dF_{L_4,R_4}(l,r)$	1/6	1/3	1/3	1/6	1

Moreover, it follows that both models have different lifetime distribution W_3 and W_4 but they assign the same lifetime probabilities to the observable intervals: $P_{W_3}([0,1]) = P_{W_4}([0,1]) = 1/3$, $P_{W_3}([0,2]) = P_{W_4}([0,2]) = 2/3$, $P_{W_3}([2,4]) = P_{W_4}([2,4]) = 2/3$ and $P_{W_3}([3,4]) = P_{W_4}([3,4]) = 1/3$. Consequently, we also derive that both models assign the same lifetime probabilities to the sets $\{0,1\}$, $\{2\}$ and $\{3,4\}$. These sets are the so called expected Turnbull's intervals which are defined in the following subsection. This example confirms that inside the class of constant–sum models the entire distribution for the lifetimes can be nonidentifiable.

3.2.1 Expected Turnbull's intervals

In this subsection we assume a finite support of the observables, that is, $\mathcal{D}_{L,R} = \{(l,r) : 0 \leq l \leq r \leq +\infty,\ dF_{L,R}(l,r) > 0\}$ is a finite set. We define a collection of sets we call expected Turnbull's intervals, $\{A_j\}_{j=1}^{m}$, and we show that they partition $\mathcal{D}_W$ under the constant–sum property assumption. Finally, as a main result, we conclude that all that is identifiable about W are the interval probabilities $P_W(A_j)$.

The expected Turnbull's intervals are a generalization of the sample Turnbull's intervals introduced by Turnbull (1976). To build the expected Turnbull's intervals, we first derive all the distinct intervals whose left and right end–points lie in the support of L and the support of R respectively and which contain no members of the support of L or the support of R other than at their left and right endpoints respectively. The intersection of these intervals with the lifetime support gives the expected Turnbull's intervals (see Example 3.5). The notion of expected Turnbull's interval is equivalent to the notion of population innermost interval used in Yu *et al.* (2000). For technical purposes, another definition is formally given below:

Definition 3.6. A set A_j is an expected Turnbull's interval if and only if A_j is a nonempty intersection of observable intervals, $A_j = \bigcap_{\{(l,r)\in\mathcal{D}_{L,R}: A_j \subseteq \lfloor l,r \rfloor\}} (\lfloor l,r \rfloor \cap \mathcal{D}_W)$, and for any observable $(l,r) \in \mathcal{D}_{L,R}$, either $A_j \subset \lfloor l,r \rfloor$ or $A_j \cap \lfloor l,r \rfloor = \emptyset$.

Since the expected Turnbull's intervals are disjoint by construction, the following proposition demonstrates that they are a partition of $\mathcal{D}_W$.

Proposition 3.7. *Let $F_{T,L,R}$ be a constant–sum model with $\mathcal{D}_{L,R}$ being a finite set, then it follows that $\mathcal{D}_W = \bigcup_{j=1}^m A_j$.*

Proof:

It holds by construction that $\bigcup_{j=1}^m A_j \subseteq \mathcal{D}_W$. To prove the equality we first note that for any $t \in \mathcal{D}_W$ equation (3.1) states that $I = \bigcap_{\{(l,r)\in\mathcal{D}_{L,R}: t\in\lfloor l,r\rfloor\}} (\lfloor l,r\rfloor \cap \mathcal{D}_W)$ is nonempty. Thus, to complete the proof we should show that I is an expected Turnbull's interval. If $s \in I$, then by the constant–sum property we have

$$1 = \iint_{\{(l,r):\, s\in\lfloor l,r\rfloor\}} \frac{dF_{L,R}(l,r)}{P_W(\lfloor l,r\rfloor)}$$

$$= \iint_{\{(l,r):\, I\subset\lfloor l,r\rfloor\}} \frac{dF_{L,R}(l,r)}{P_W(\lfloor l,r\rfloor)} + \iint_{\{(l,r):\, I\not\subset\lfloor l,r\rfloor, s\in\lfloor l,r\rfloor\}} \frac{dF_{L,R}(l,r)}{P_W(\lfloor l,r\rfloor)}$$

$$= \iint_{\{(l,r):\, t\in\lfloor l,r\rfloor\}} \frac{dF_{L,R}(l,r)}{P_W(\lfloor l,r\rfloor)} + \iint_{\{(l,r):\, I\not\subset\lfloor l,r\rfloor, s\in\lfloor l,r\rfloor\}} \frac{dF_{L,R}(l,r)}{P_W(\lfloor l,r\rfloor)}$$

$$= 1 + \iint_{\{(l,r):\, I\not\subset\lfloor l,r\rfloor, s\in\lfloor l,r\rfloor\}} \frac{dF_{L,R}(l,r)}{P_W(\lfloor l,r\rfloor)} .$$

As a consequence, $\{(l,r) \in \mathcal{D}_{L,R} : \; I \not\subset \lfloor l,r\rfloor, s \in \lfloor l,r\rfloor\}$ is an empty set and I is necessarily an expected Turnbull's interval.

□

As a consequence of the previous result, the following corollary characterizes the constant–sum condition in terms of the expected Turnbull's intervals. Under the constant–sum property assumption, next theorem shows not only that the lifetime probabilities assigned to the expected Turnbull's intervals are identifiable but also that W is nonidentifiable inside the expected Turnbull's intervals.

Corollary 3.8. *A censoring model $F_{T,L,R}$ with $\mathcal{D}_{L,R}$ being a finite set is constant–sum if and only if we have for any expected Turnbull's interval A_j,*

$$\iint_{\{(l,r):A_j\subseteq\lfloor l,r\rfloor\}} \frac{dF_{L,R}(l,r)}{P_W(\lfloor l,r\rfloor)} = 1.$$

Theorem 3.9. *Let $F_{T,L,R}$ be a constant–sum model with $\mathcal{D}_{L,R}$ being a finite set.*

(a) If $F_{T*,L*,R*}$ is another constant–sum model so that $(\mathcal{D}_{W^*}, F_{L^*,R^*}) = (\mathcal{D}_W, F_{L,R})$, then $P_{W^*}(A_j) = P_W(A_j)$ for every $j = 1, \ldots, m$.

(b) If I_k is not a singleton for some $k = 1, \ldots, m$, then there exists a constant–sum model $F_{T*,L*,R*}$ so that $(\mathcal{D}_{W^*}, F_{L^*,R^*}) = (\mathcal{D}_W, F_{L,R})$ and $W^*(t) \neq W(t)$ for every $t \in I_k$.

Proof:

To show (a) we assume, without loss of generality, that the expected Turnbull's intervals are ordered, $I_1 < I_2 < \cdots < I_m$. From the definition of the first expected Turnbull's interval, there necessarily exists an observable $(l_1, r_1) \in \mathcal{D}_{L,R}$ such that $I_1 = \lfloor l_1, r_1\rfloor \cap \mathcal{D}_W$ and, consequently, Theorem 3.4 implies $P_{W^*}(I_1) = P_{W^*}(\lfloor l_1, r_1\rfloor) = P_W(\lfloor l_1, r_1\rfloor) = P_W(I_1)$. For the second expected Turnbull's interval, there also exists an observable $(l_2, r_2) \in \mathcal{D}_{L,R}$ such that either $I_2 = \lfloor l_2, r_2\rfloor \cap \mathcal{D}_W$ and Theorem 3.4 implies $P_{W^*}(I_2) = P_{W^*}(\lfloor l_2, r_2\rfloor) = P_W(\lfloor l_2, r_2\rfloor) = P_W(I_2)$, or $I_1 \cup I_2 = \lfloor l_2, r_2\rfloor \cap \mathcal{D}_W$ and Theorem 3.4 implies $P_{W^*}(I_2) = P_{W^*}(\lfloor l_2, r_2\rfloor) - P_{W^*}(I_1) = P_W(\lfloor l_2, r_2\rfloor) - P_W(I_1) = P_W(I_2)$. This process can continue indefinitely providing the identifiability of the lifetime probabilities assigned to all the expected Turnbull's intervals, $P_{W^*}(A_j) = P_W(A_j)$ for every $j = 1, \ldots, m$.

Now, we show that (b) follows straightforwardly from Proposition 3.2. If $(\mathcal{D}, G) = (\mathcal{D}_W, F_{L,R})$ and F, with support $\mathcal{D}_W$, satisfies $P_F(A_j) = P_W(A_j)$ for every $j = 1, \ldots, m$, then equation (3.2) defines a constant–sum model $F_{T*,L*,R*}$ so that $(\mathcal{D}_{W^*}, F_{L^*,R^*}) =$

$(\mathcal{D}_W, F_{L,R})$. Since it results that $W^*(t) = F(t)$ for every $t \in \mathcal{D}_W$, we complete the proof by choosing F so that $F(t) \neq W(t)$ for every $t \in I_k$.

□

3.2.2 Illustration in a finite censoring setting

In practical situations it is quite usual to assume that $\mathcal{D}_W = (0, \infty)$, and that every observable $(l, r) \in \mathcal{D}_{L,R}$ arises from a random inspection process leading to intervals which are half open $\lfloor l, r \rfloor = (l, r]$. It is also not restrictive to assume that the support of the inspection times is finite, or equivalently, that L and R lie in a set $\{a_0, a_1, \ldots, a_k\}$ with $0 = a_0 < a_1 < \cdots < a_k = +\infty$. In this case, all that is potentially identifiable about W are the interval probabilities $P_W((a_{j-1}, a_j]) = W(a_j) - W(a_{j-1})$ for $j = 1, \ldots, k$. Since the results in Subsection 3.2.1 apply, Theorem 3.9 proves that these probabilities are identifiable inside the class of constant–sum models only when there is coincidence between these intervals and the expected Turnbull's intervals, $m = k$ and $A_j = (a_{j-1}, a_j]$ for $j = 1, \ldots, k$. This identifiability condition is equivalent to the fact that the set $\{a_0, a_1, \ldots, a_k\}$ is the support of L (except for $a_k = +\infty$) and the support of R (except for $a_0 = 0$): (i) $P(L = 0) > 0$, (ii) $P(R = +\infty) > 0$ and (iii) $P(L = a_j) > 0$ and $P(R = a_j) > 0$ for $j = 1, \ldots, k-1$. However, under the constant–sum property it is easily seen that $P(L = a_j) > 0$ holds if and only if $P(R = a_j) > 0$ holds for $j = 1, \ldots, k-1$. Thus, condition (iii) simplifies either in terms of L or in terms of R.

In this framework, Lawless (2004) proves that for a possibly non–independent inspection process, the constant-sum property is equivalent to the existence of a Markovian inspection process that is independent of lifetimes, and which gives the same distribution for the observables as the underlying true inspection process. Thus, the identifiability condition above is equivalent to every value in the set $\{a_0, a_1, \ldots, a_k\}$ being visited by this Markovian process. That is, if we define the transition probabilities of this process as $\pi_{l,r} = P(\text{next inspection is at } a_r \mid \text{inspection at } a_l)$ for $l = 0, \ldots, k-1$ and

$r = l+1,\dots,k$, then (i) $\sum_{l=0}^{j-1} \pi_{lj} > 0$ for $j = 1,\dots,k$, and (ii) $\sum_{r=j+1}^{k} \pi_{jr} = 1$ for $j = 0,\dots,k-1$.

Example 3.10. We now present a simple illustration of the above identifiability condition. We consider $k = 4$ and $a_1 = 12$, $a_2 = 24$ and $a_3 = 36$. We also consider a Markovian inspection process independent of T which does not visit $a_2 = 24$ and has, for instance, the following non null transition probabilities $\pi_{01} = \pi_{03} = \pi_{13} = \pi_{14} = \frac{1}{2}$ and $\pi_{34} = 1$. Then, the observable intervals are $(0, 12]$, $(0, 36]$, $(12, 36]$, $(12, +\infty]$ and $(36, +\infty]$. Consequently, the expected Turnbull's intervals are $(0, 12]$, $(12, 36]$ and $(36, +\infty]$, and the interval probabilities $(12, 24]$ and $(24, 36]$ are not identifiable.

Chapter 4

The k–sample problem

One important question that arises in many survival studies is to establish if there are differences in the lifetimes among different groups of individuals. While many k–sample tests have been developed when data are uncensored or right–censored, research for interval–censored data is still ongoing. Most approaches to this problem try to generalize these known tests to the interval–censored framework. In Mantel (1967) we find an interval–censored data version of the Wilcoxon test. In Peto and Peto (1972) we find a different extension of the Wilcoxon test and an extension of the Log–rank test. In Fay and Shih (1998) we find an interval–censored data form of the t–test. The main characteristic of these papers is the use of permutational distributions. The difficulty of finding the distribution of the test statistic is avoided with this permutational approach. Other approaches assume that the collection of possible interval endpoints is discrete. This assumption ensures a finite number of parameters in the log–likelihood which allows to find test statistics with known asymptotic distribution, see for example Finkelstein (1986) and Petroni and Wolfe (1994). Finally, Pan (2000) proposes to use an approximate Bayesian bootstrap method to impute exact lifetimes from interval-censored observations and apply known test statistics for right–censored data.

In this chapter we survey different testing methods for interval censored data and we propose new methods. Section 4.1 is devoted to permutational tests. We introduce the permutational methodology and provide description of the permutational Wilcoxon and Log–rank tests for interval–censored data. In Section 4.2 permutational tests are revisited following the interesting approach given in Fay and Shih (1998). This approach is the basis of new proposals given in the subsequent sections and in Chapter 5. In Section 4.3, we illustrate the permutational methodology. We analyze data from an AIDS clinical trial designed to study the benefits of zidovudine therapy in patients in the early stages of the HIV infection (Volberding *et al.*, 1995). In Section 4.4, we give a new permutational test proposal which generalizes the class of tests for right–censored data in Harrington and Fleming (1982). In Section 4.5, a likelihood approach for this new proposal is also considered. We also extend the relationship given in Fay (1999) between permutational methods and likelihood methods based on score test statistics. Section 4.6 contains different tests which generalize the Weighted Kaplan–Meier class of tests for right–censored data in Pepe and Fleming (1989). Finally, in Section 4.7 we provide and describe several S–plus functions which have been implemented for the permutational methodology.

In this chapter and in the following, we consider closed observed intervals. This agrees with the interpretation of the intervals done in Chapter 1. As mentioned before, the definitions below are easily modifiable to cover open or half open intervals.

4.1 Permutational tests

We introduce now the permutational approach to the k–sample problem. Let T be the time to the event of interest. Assume that we have k groups of data, $G_1, \ldots, G_k$ with respective sample sizes $n_1, \ldots, n_k$. Define $W_1, \ldots, W_k$ the distribution functions of T under each one of these groups. The k–sample problem establishes a test between $H_0 : W_1 = \cdots = W_k$ and $H_a : W_i \neq W_j$ for some i,j. Denote by $\mathbf{z}_i$ a vector of covariates representing to which group the i^{th} observation belongs. In the two sample problem, the usual choice of the covariate is $\mathbf{z}_i = \alpha_i^{(2)}$ where $\alpha_i^{(2)}$ is an indicator function that is equal to 0 if the individual belongs to group G_1 and 1 if it belongs to group G_2. When we have k groups many choices of $\mathbf{z}_i$ are possible, for instance, we could take

$$\mathbf{z}_i = \left(\frac{\alpha_i^{(1)}}{\sqrt{n_1}}, \frac{\alpha_i^{(2)}}{\sqrt{n_2}}, \ldots, \frac{\alpha_i^{(k)}}{\sqrt{n_k}} \right)'$$

where $\alpha_i^{(j)}$ is an indicator function that is equal to 1 if the individual belongs to group G_j and 0 otherwise.

A permutational linear test statistic is of the form:

$$L_0 = \sum_{i=1}^{n} \mathbf{z}_i c_i, \tag{4.1}$$

where c_i is a scalar score associated to the i^{th} observation which is independent of the covariates. These scalar scores are often built as

$$c_i = \sum_{j=1}^{n} \Phi(i, j),$$

where $\Phi(i, j)$ represents a comparison between pairs of observations. Since usually Φ compares functions, in the sequel we refer to Φ as a functional.

The idea behind the permutational test is that, if the null hypothesis is true and the censoring mechanism does not depend on the grouping, the labels on the scores are exchangeable. Thus, the permutational distribution of L_0 is obtained by permuting the labels and recomputing the test statistic for all the possible rearranged labels. The main key for these procedures is to use scores that are sensitive to the alternative hypothesis and, in that case, the null hypothesis will be rejected if L_0 is an extreme value for the permutational distribution. This permutational distribution can be computed exactly when the sample size is small. When n is large, a version of the Central Limit theorem for exchangeable random variables allow us to rely on a normal asymptotic approximation for the permutational distribution of L_0 where $E(L_0) = n\bar{c}\bar{\mathbf{z}}'$ ($\bar{c} = 0$ in our examples) and the variance–covariance matrix is

$$V_0 = \frac{\left(\sum_{i=1}^{n} c_i^2 - n\bar{c}^2\right)\left(\sum_{i=1}^{n}(\mathbf{z}_i\mathbf{z}_i' - \bar{\mathbf{z}}\bar{\mathbf{z}}')\right)}{(n-1)}. \quad (4.2)$$

4.1.1 Scores used in permutational tests

The choice of different scores yields to different permutational tests. The well known tests are the permutational forms of the Wilcoxon–Gehan test, the Wilcoxon–Peto test and the Log–rank test.

For each observation $[l_i, r_i]$, $i = 1, \ldots, n$, the Wilcoxon–Gehan (WG) score is the difference between the number of lifetimes that are undoubtedly to its left and the number of lifetimes that are undoubtedly to its right. Intervals which overlap with the i^{th} interval do not contribute to the computation of the i^{th} score. The Wilcoxon–Gehan score for the i^{th} individual is given by

$$WGc_i = \sum_{j=1}^{n} \Phi_{WG}(i,j), \quad (4.3)$$

where

$$\Phi_{WG}(i,j) = \mathbf{1}_{\{r_j<l_i\}} - \mathbf{1}_{\{l_j>r_i\}}. \quad (4.4)$$

Gehan (1965) proposes these scores as an extension of the two sample Wilcoxon test for right–censored data. Gehan's scores are generalized by Mantel (1967) to allow the use of interval–censored data. A k–sample version of this test is proposed in Schemper (1983). In Schemper (1982, 1984) the Wilcoxon–Gehan functional (4.4) is considered to respectively extend the Kendall's correlation coefficient for two dimensional censored–data and the Friedman's test. Abel (1986) uses as well (4.4) for a test against ordered alternatives which generalizes the Jonckheere's test and which will be discussed in Chapter 5.

The Wilcoxon–Peto (WP) score for each observation is the difference between Turnbull's estimated proportion of lifetimes that are to the left and Turbull's estimated proportion of lifetimes that are to the right, that is,

$$WPc_i = \hat{W}(l_i^-) - (1 - \hat{W}(r_i)) = \hat{W}(l_i^-) + \hat{W}(r_i) - 1,$$

where $\hat{W}$ is Turnbull's estimator for the pooled sample. This proposal is introduced by Peto and Peto (1972) and it is asymptotically efficient for lifetime distributions in the logistic family.

In the same article Peto and Peto extend the Savage or Log-rank (LR) test to interval–censored data. The Log-rank scores are given by

$$LRc_i = \frac{(1 - \hat{W}(r_i))\log(1 - \hat{W}(r_i)) - (1 - \hat{W}(l_i^-))\log(1 - \hat{W}(l_i^-))}{\hat{W}(r_i) - \hat{W}(l_i^-)},$$

where again $\hat{W}$ is Turnbull's estimator for the pooled sample. This proposal is asymptotically efficient for lifetime distributions with Lehmann–type alternatives.

4.2 Permutational tests using estimated distribution functions

Fay and Shih (1998) introduce what they call distribution permutational tests, which provides another interesting approach to the k–sample problem. These are permutational tests where the scalar scores are obtained using an estimate of the distribution function for each observation and comparing it to the overall Turnbull's estimate of the distribution function. For particular ways of comparing these estimated distributions Fay and Shih obtain the Wilcoxon–Peto test, the Log–rank test and a new test called the difference in means (DiM) test. We use this methodology in Chapter 5 to build tests against ordered alternatives.

4.2.1 Estimating individual distributions

The estimate of the distribution function for each observation is based on the self–consistent equations (1.5). At convergence of the EM algorithm, the expectation step implies truncation of the Turnbull's estimator $\hat{W}$ at each observed interval $[l_i, r_i]$. This truncation defines an estimate of the distribution function for the i^{th} observation,

$$\begin{aligned} \hat{W}^i(t) = P_{\hat{W}}([0,t] \mid [l_i, r_i]) &= \frac{\hat{W}(r_i \wedge t) - \hat{W}(l_i^- \wedge t)}{\hat{W}(r_i) - \hat{W}(l_i^-)} \\ &= \begin{cases} 0 & \text{if} \quad t < q_1 \\ \mu_1^i(\hat{\boldsymbol{w}}) + \cdots + \mu_k^i(\hat{\boldsymbol{w}}) & \text{if} \quad p_k \le t < q_{k+1}, \quad 1 \le k \le m-1 \\ 1 & \text{if} \quad t \ge p_m \end{cases} . \end{aligned} \tag{4.5}$$

where

$$\mu_j^i(\hat{\boldsymbol{w}}) = P_{\hat{W}}([q_j, p_j] \mid [l_i, r_i]) = \frac{\alpha_j^i}{\sum_{l=1}^m \alpha_l^i \hat{w}_l} \hat{w}_j \qquad 1 \le j \le m.$$

and for each Turnbull's interval $[q_j, p_j]$ $j = 1, \ldots, m$, $\hat{w}_j = P_{\hat{W}}([q_j, p_j])$ and $\alpha_j^i = \mathbf{1}_{\{[q_j,p_j] \subseteq [l_i,r_i]\}}$.

The maximization step implies that Turnbull's estimator $\hat{W}$ can be rewritten as

$$\hat{W}(t) = \frac{1}{n} \sum_{i=1}^n \hat{W}^i(t).$$

We note that if the i^{th} observation is not censored, the above empirical distribution estimate (4.5) coincides with the usual one, i.e., $\hat{W}^i(t) = \mathbf{1}_{\{t_i \le t\}}(t)$.

4.2.2 L–functionals

Fay and Shih define an scalar score as a comparison between pairs of empirical distribution estimates:

$$c_i = \frac{1}{n}\sum_{j=1}^{n}\Phi(\hat{W}^i, \hat{W}^j) = \Phi(\hat{W}^i, \hat{W}) \qquad i = 1, \ldots, n\ . \tag{4.6}$$

where Φ is a L–functional which operates linearly on distribution functions. That is, Φ should satisfy the following properties for every distribution function F and G:

(a) $\Phi(F, F) = 0$,

(b) $\Phi(F, G) = -\Phi(G, F)$ and

(c) $\int \Phi(F, G)d\nu(G) = \Phi(F, \int Gd\nu(G))$ for any probability measure ν, where the integration is over the overall space associate with ν.

The authors focus their work on three L–functionals:

(1) The Mann–Whitney (or Wilcoxon–Peto) functional Φ_{MW} defined as,

$$\Phi_{MW}(F, G) = \int G(s)dF(s) - \int F(s)dG(s) = 2((P(X > Y) + \frac{1}{2}P(X = Y)) - 1$$

where X and Y are random variables with distribution functions F and G, respectively. The scores for the Mann–Whitney functional are equivalent to the Wilcoxon–Peto scores.

(2) The Weighted Mann–Whitney functional Φ_{WMW} defined as,

$$\Phi_{WMW}(F, G) = \int Q(s)G(s)dF(s) - \int Q(s)F(s)dG(s) \tag{4.7}$$

A suitable choice of the weight function gives also known scores for the Weighted Mann–Whitney functional. Fay and Shih derive the weighting scheme which is necessary to obtain the Log–rank scores. In Section 4.4 we describe these weights and we also obtain new scores which generalize the class of tests for right–censored data in Harrington and Fleming (1982).

(3) Difference in means functional Φ_{DiM} defined as,

$$\Phi_{DiM}(F,G) = \int xdF(x) - \int ydG(y) = E(X) - E(Y). \tag{4.8}$$

where X and Y are random variables with distribution functions F and G, respectively. In next subsection, we describe this L–functional in detail.

4.2.3 Difference in means test

Since $\hat{W}$ is not defined inside Turnbull's intervals (see equation (1.4)), the empirical distribution estimates $\hat{W}^i$ are not defined either (see equation (4.5)). This nonidentifiability does not affect the calculation of $\Phi(\hat{W}^i, \hat{W}^j)$ for the Mann–Whitney and the Weighted Mann–Whitney functional, though it does so for the Difference in means functional. Fay and Shih avoid this problem collapsing each Turnbull's interval $[q_j, p_j]$ to the right endpoint p_j and assigning all the probability of $[q_j, p_j]$, $\hat{w}_j$, to p_j. When $p_m = \infty$, they let $p_m = q_m$. This method produces one of the possible indistinguishable distribution functions which are flat outside $\cup_{j=1}^{m}[q_j, p_j]$ and which have vector of probabilities $\hat{\boldsymbol{w}} = (\hat{w}_1, \ldots, \hat{w}_m)$:

$$\hat{W}(t) = \begin{cases} 0 & \text{if} \quad t < p_1 \\ \hat{w}_1 + \cdots + \hat{w}_k & \text{if} \quad p_k \leq t < p_{k+1}, \quad 1 \leq k \leq m-1 \\ 1 & \text{if} \quad t \geq p_m \end{cases} . \tag{4.9}$$

In the sequel, we will take (4.9) as the definition of the Turnbull's estimator of W and, consequently, equation (4.5) reduces to

$$\hat{W}^i(t) = \begin{cases} 0 & \text{if} \quad t < p_1 \\ \mu_1^i(\hat{\boldsymbol{w}}) + \cdots + \mu_k^i(\hat{\boldsymbol{w}}) & \text{if} \quad p_k \leq t < p_{k+1}, \quad 1 \leq k \leq m-1 \\ 1 & \text{if} \quad t \geq p_m \end{cases} . \tag{4.10}$$

In what follows we describe the Difference in means test as an extension of the permutational t–test. The use of the Difference in means functional gives a score for the i^{th} individual which is the difference between its imputed mean value and the total mean of the distribution,

$$\begin{aligned} DiMc_i &= \int_0^{+\infty} t \, d\hat{W}^i(t) - \int_0^{+\infty} t \, d\hat{W}(t) \\ &= \sum_{j=1}^{m} p_j \mu_j^i(\hat{\boldsymbol{w}}) - \sum_{j=1}^{m} p_j \hat{w}_j = \frac{\sum_{j=1}^{m} p_j \hat{w}_j \alpha_j^i}{\hat{W}(r_i) - \hat{W}(l_i^-)} - \sum_{j=1}^{m} p_j \hat{w}_j . \end{aligned}$$

Note that because of the self–consistent property of Turnbull's estimate, the mean of the imputed mean of each individual is equal to the total mean of the distribution.

Example 4.1. We illustrate the computation of the scores and the permutational approach in the two–sample problem with an extension of Example 1.3 developed in Chapter 1. Let W_1 and W_2 be the distribution functions of the lifetimes in group G_1 and G_2, respectively. We are interested in testing the hypothesis $\text{H}_0 : W_1 = W_2$ versus $\text{H}_a : W_1 \neq W_2$. The covariate choice will be $\mathbf{z}_i = \alpha_i^{(2)}$ where $\alpha_i^{(2)}$ is an indicator function that is equal to 0 if the individual belongs to group G_1 and 1 if it belongs to group G_2. Then, the permutational linear test statistic given by (4.1) and its variance reduce, respectively, to

$$L_0 = \sum_{i=1}^{n} \alpha_i^{(2)} c_i \qquad \text{and} \qquad V_0 = \frac{n_1 n_2}{n(n-1)} \sum_{i=1}^{n} c_i^2 \, . \tag{4.11}$$

In particular, if the individuals in the first group are those given in Example 1.3, that is,

$$G_1 = \{[0,1], [4,6], [2,6], [0,3], [2,4], [5,7]\} \, ,$$

and we consider five new individuals in the second group with intervals,

$$G_2 = \{[0,5],[4,4],[7,8],[7,9],[10,\infty)\}\,,$$

then Turnbull's estimate of the distribution function from the pooled sample is:

$$\hat{W}(t) = \begin{cases} 0 & \text{if} \quad t<1 \\ \frac{7}{33} = 0+\frac{7}{33} & \text{if} \quad 1\le t<3 \\ \frac{7}{33} = 0+\frac{7}{33}+0 & \text{if} \quad 3\le t<4 \\ \frac{21}{33} = 0+\frac{7}{33}+0+\frac{14}{33} & \text{if} \quad 4\le t<5 \\ \frac{21}{33} = 0+\frac{7}{33}+0+\frac{14}{33}+0 & \text{if} \quad 5\le t<7 \\ \frac{30}{33} = 0+\frac{7}{33}+0+\frac{14}{33}+0+\frac{9}{33} & \text{if} \quad 7\le t<10 \\ 1 = 0+\frac{7}{33}+0+\frac{14}{33}+0+\frac{9}{33}+\frac{3}{33} & \text{if} \quad t\ge 10 \end{cases}$$

As illustration of an empirical estimate of the distribution function for one observation, we consider the interval $[0,5]$. Then $P_{\hat{W}}([0,5]) = \hat{W}(5) = \frac{21}{33}$ and, consequently, it follows that

$$\hat{W}^i(t) = \begin{cases} 0 & \text{if} \quad t<1 \\ \frac{1}{3} = 0+\frac{7/33}{21/33} & \text{if} \quad 1\le t<3 \\ \frac{1}{3} = 0+\frac{7/33}{21/33}+0 & \text{if} \quad 3\le t<4 \\ 1 = 0+\frac{7/33}{21/33}+0+\frac{14/33}{21/33} & \text{if} \quad 4\le t<5 \\ 1 & \text{if} \quad 5\le t<7 \\ 1 & \text{if} \quad 7\le t<10 \\ 1 & \text{if} \quad t\ge 10 \end{cases}$$

Table 4.1 gives the scores, statistic value, variance and p–value of the different permutational tests statistics. As we see, although neither one of the four tests reject the null hypothesis, the p-value of the tests are quite different and in particular the Wilcoxon–Gehan test is close to the 5% level of significance. Thus, we conclude that there are not significant differences at the 5% level of significance between the distributions of the two groups. The exact permutational distribution of each test is obtained computing all the values of the linear test statistic L_0 for all the possible rearrangements of the scores, that is $462 = \binom{11}{5}$ combinations. Then, the p-value of each permutational test is the percentage of rearrangements with absolute value of L_0 larger than or equal to that of our original sample. We should remark that we have also used the asymptotic approximation for the permutational distribution despite of the fact that the sample size of the pooled sample is 11.

Table 4.1: *Different scores and statistic tests for the data in the example. We use the exact permutational distribution, p–value$_1$, and the normal approximation, p–value$_2$.*

	Wilcoxon–Gehan			**Wilcoxon–Peto**			**Log–rank**	**Difference in means**	
	Left value	Right value	WGc_i	Left value	Right value	WPc_i	LRc_i	Imputed mean	$DiMc_i$
G	0	8	-8	0	26/33	-26/33	-0.88	1	-41/11
R	2	3	-1	7/33	12/33	-5/33	-0.42	4	-8/11
O	1	3	-2	7/33	12/33	-5/33	-0.42	4	-8/11
U	0	6	-6	0	26/33	-26/33	-0.88	1	-41/11
P	1	4	-3	7/33	12/33	-5/33	-0.42	4	-8/11
1	4	1	3	21/33	3/33	18/33	0.55	7	25/11
.	.....	.....	......	.....	.....	......	......	....	......
G	0	3	-3	0	12/33	-12/33	-0.58	3	-19/11
R	2	4	-2	7/33	12/33	-5/33	-0.42	4	-8/11
O	7	1	6	21/33	3/33	18/33	0.55	7	25/11
U	7	1	6	21/33	3/33	18/33	0.55	7	25/11
P	10	0	10	30/33	0	30/33	2.40	10	58/11
2									
	MEAN		0			0	0	52/11	0
	L_0		17			49/33	2.49		81/11
	V_0		84			0.87	2.53		20.78
	p–value$_1$		0.0649			0.1450	0.1234		0.1450
	p–value$_2$		0.0636			0.1111	0.1168		0.1062

4.3 Illustration

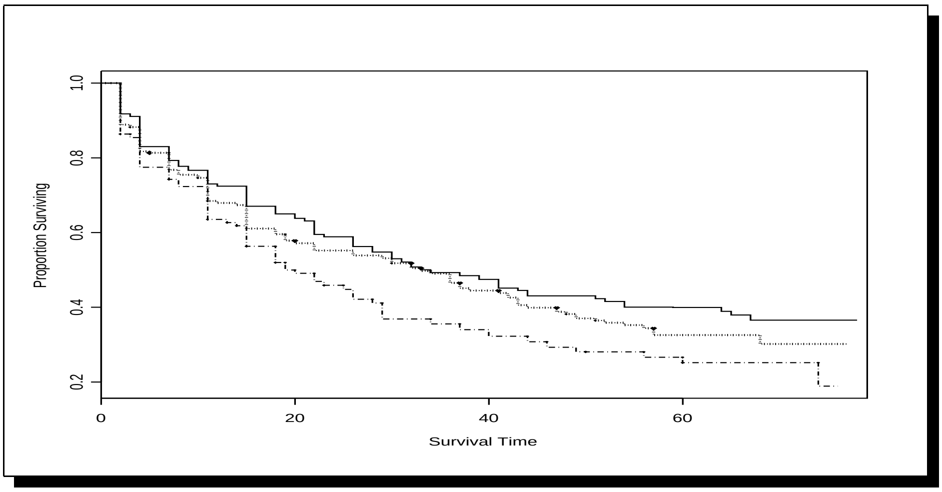
An instance of interval-censored data is found in an AIDS Clinical Trial designed to study the benefits of zidovudine therapy in patients in the early stages of the human immunodeficiency virus (HIV) infection, see Volberding *et al.* (1995). The design compares three groups. The first group, G_1, corresponds to those patients who started zidovudine monotherapy after their CD4 cell count fell below 500 per cubic millimeter. In the second and third groups, G_2 and G_3, two different dosages of zidovudine were given immediately after randomization. Among the 1607 subjects who could be evaluated, 541 were in the deferred-therapy group, 538 in the 500–mg group and 528 in the 1500–mg group. Subjects were followed prospectively until the development of AIDS or death. As a measure of the clinical progression of the disease, CD4 cell counts were periodically determined. The reported data included the times of the first count below 500 cells per cubic millimeter, as well as below 400 and below 300. We will focus on the time T, measured in months from randomization, until the CD4 count first reaches 400 cells per cubic millimeter. The random variable T is interval-censored, that is, for each individual i, we know that T_i is between l_i and r_i where r_i is the time of the first visit when CD4 was observed to be below 400 cells per cubic millimeter and l_i is defined to be the time of the preceding visit. Figure 4.1 shows the probabilities of keeping CD4 values larger than a certain number of months. The estimated survival curves suggest differences between the deferred–therapy group and the immediate–therapy groups (500-mg and 1500-mg). In particular, the immediate–therapy group for a heavier dose of zidovudine shows a better survival than the other two groups.

We illustrate now the above permutational methodology with the comparison of the survival of these three groups ($k = 3$). The choice for the $\mathbf{z}_i$ covariates is the following,

$$\mathbf{z'}_i = \left(\frac{\alpha_i^{(1)}}{\sqrt{n_1}}, \frac{\alpha_i^{(2)}}{\sqrt{n_2}}, \frac{\alpha_i^{(3)}}{\sqrt{n_3}} \right) = \left(\frac{\alpha_i^{(1)}}{23.2594}, \frac{\alpha_i^{(2)}}{23.1948}, \frac{\alpha_i^{(3)}}{22.9783} \right)$$

where $\alpha_i^{(j)}$ is an indicator function that is equal to 1 if the individual belongs to group

Figure 4.1: *Probabilities of keeping CD4 values larger than a certain number of months for the group receiving 500 mg (thick dotted curve), 1500 mg (solid curve) and deferred therapy (dashed curve)*



G_j and 0 otherwise. Then the linear permutational statistic form simplifies to the expression,

$$L_0 = \sum_{i=1}^{n} \mathbf{z}_i c_i = \begin{pmatrix} \sqrt{n_1}\ \bar{c}_{(1)} \\ \sqrt{n_2}\ \bar{c}_{(2)} \\ \sqrt{n_3}\ \bar{c}_{(3)} \end{pmatrix} = \begin{pmatrix} 23.2594\ \bar{c}_{(1)} \\ 23.1948\ \bar{c}_{(2)} \\ 22.9783\ \bar{c}_{(3)} \end{pmatrix},$$

where $\bar{c}_{(j)} = \frac{1}{n_j} \sum_{i=1}^{n} c_i \alpha_i^{(j)}$. The permutational distribution of L_0 is asymptotically distributed as a k–dimensional normal and we can use the Mahalanobis distance (Md) to obtain a $\chi^2_{k-1} = \chi^2_2$ distribution:

$$Md = L_0' V_0^- L_0 = \frac{n-1}{\sum_{i=1}^{n} c_i^2} \sum_{j=1}^{k} n_j \bar{c}^2_{(j)} = \frac{1606}{\sum_{i=1}^{n} c_i^2} (541\ \bar{c}^2_{(1)} + 538\ \bar{c}^2_{(2)} + 528\ \bar{c}^2_{(3)}),$$

where V_0^- is the generalized inverse of the variance–covariance matrix V_0. The results using each of the permutational tests (see Table 4.2) show significant evidence of the differences between the survival curves.

Table 4.2: *Permutational test statistic (L_0), Mahalanobis distance (Md) and p–values for the null hypothesis of equal distributions: $H_0 : W_1 = W_2 = W_3$ versus the alternative of some differences between the distributions $H_a : W_i \neq W_j$ for some i,j, for different score choices*

	Wilcoxon–Gehan	*Wilcoxon–Peto*	*Log–Rank*	*Difference in Means*
L_0	$\begin{pmatrix} -1804.732 \\ 337.9202 \\ 1485.709 \end{pmatrix}$	$\begin{pmatrix} -1.5351 \\ 0.2687 \\ 1.2826 \end{pmatrix}$	$\begin{pmatrix} -2.2098 \\ 0.3449 \\ 1.8887 \end{pmatrix}$	$\begin{pmatrix} -84.0323 \\ 16.4337 \\ 68.4719 \end{pmatrix}$
Md	16.3978	16.6800	17.6607	17.8151
p–value	0.000275	0.000239	0.000146	0.000135

4.4 A new permutational family of tests

In this section we propose new scores which generalize the class of tests for right–censored data given in Harrington and Fleming (1982) (see also Fleming and Harrington, 1991 and Lawless, 2003).

Definition 4.2. For any $\rho \geq 0$, we define the Harrington and Fleming scores as

$$HFc_i = \frac{(1-\hat{W}(r_i))^{\rho+1} - (1-\hat{W}(l_i^-))^{\rho+1}}{\rho\,(\hat{W}(r_i) - \hat{W}(l_i^-))} + \frac{1}{\rho}\,. \tag{4.12}$$

Analogously to Harrington and Fleming (1982), the special case $\rho = 1$ gives the Wilcoxon–Peto scores and $\rho = 0$ gives the Log–rank scores. In the sequel of this section, we base on Fay and Shih (1998) and Fay (1996) to derive the Harrington and Fleming scores (4.12). In next section, we show that this class of test statistics can be written in a Weighted Log–rank form or can be derived as a class of efficient score test statistics under an accelerated failure time model.

Proposition 4.3. *Let weights in the Weighted Mann–Whitney functional (4.7) depend on* $\hat{W}$ *as in Fay and Shih (1998),*

$$Q(t) = \frac{\gamma(\hat{W}(t)) - \gamma(\hat{W}(t^-))}{\hat{W}(t) - \hat{W}(t^-)} = \frac{d\gamma(\hat{W}(t))}{d\hat{W}(t)} \tag{4.13}$$

where $\gamma(t)$ *is a nondecreasing function. Then,*

$$\gamma(t) = \frac{(1-t)}{t} \, \frac{(1-t)^\rho - 1}{\rho} \tag{4.14}$$

gives the Harrington and Fleming scores (4.12).

Proof:

The proof of this result follows straightforward by characterizing the Weighted Mann–Whitney scores in terms of the function γ. We use our notation to derive this characterization given in Fay and Shih (1998).

The definition of $\hat{W}^i$ as a truncation of $\hat{W}$ at the observed interval $[l_i, r_i]$ provides,

$$\begin{aligned}
c_i = \Phi_{WMW}(\hat{W}^i, \hat{W}) &= \int_0^{+\infty} Q(t)\hat{W}(t)d\hat{W}^i(t) - \int_0^{+\infty} Q(t)\hat{W}^i(t)d\hat{W}(t) \\
&= \frac{\int_{l_i^-}^{r_i} Q(t)\hat{W}(t)d\hat{W}(t)}{\hat{W}(r_i) - \hat{W}(l_i^-)} - \left\{ \frac{\int_{l_i^-}^{r_i} Q(t)(\hat{W}(t) - \hat{W}(l_i^-))d\hat{W}(t)}{\hat{W}(r_i) - \hat{W}(l_i^-)} + \int_{r_i}^{+\infty} Q(t)d\hat{W}(t) \right\} \\
&= \frac{\hat{W}(l_i^-)}{\hat{W}(r_i) - \hat{W}(l_i^-)} \int_{l_i^-}^{r_i} Q(t)d\hat{W}(t) - \int_{r_i}^{+\infty} Q(t)d\hat{W}(t).
\end{aligned}$$

Now, from the weighting definition (4.13), the scores simplify as follows,

$$
\begin{aligned}
c_i &= \frac{\hat{W}(l_i^-)}{\hat{W}(r_i) - \hat{W}(l_i^-)} \int_{l_i^-}^{r_i} d\gamma(\hat{W}(t)) - \int_{r_i}^{+\infty} d\gamma(\hat{W}(t)) \\
&= \frac{\hat{W}(l_i^-)}{\hat{W}(r_i) - \hat{W}(l_i^-)} \left\{\gamma(\hat{W}(r_i)) - \gamma(\hat{W}(l_i^-))\right\} - \left\{\gamma(1) - \gamma(\hat{W}(r_i))\right\} \qquad (4.15) \\
&= \frac{\hat{W}(r_i)\,\gamma(\hat{W}(r_i)) - \hat{W}(l_i^-)\,\gamma(\hat{W}(l_i^-))}{\hat{W}(r_i) - \hat{W}(l_i^-)} - \gamma(1).
\end{aligned}
$$

Finally, substitution of (4.14) in equation (4.15) completes the proof.

□

The Harrington and Fleming scores (4.12) can be seen as a special case of the scores proposed in Fay (1996) under a grouped continuous accelerated time model. For these general scores Fay and Shih (1998) provide that

$$
\gamma(t) = -\frac{G'(G^{-1}(t))}{t}, \qquad (4.16)
$$

where G is the distribution function of the error term in the model. A logistic distribution for G gives the Wilcoxon–Peto scores and a extreme value distribution gives the Log–rank scores. The scores we propose in (4.12) arise from a family of distributions G which depends on a parameter $\rho \geq 0$ and extends the logistic ($\rho = 1$) and the extreme value ($\rho = 0$) distributions,

$$
G(t) = 1 - (1 + \rho \exp(t))^{\frac{-1}{\rho}} . \qquad (4.17)
$$

By substitution of (4.17) in (4.16), we obtain $\gamma(t) = -\frac{G'(G^{-1}(t))}{t} = \frac{(1-t)}{t}\,\frac{(1-t)^{\rho}-1}{\rho}$ and, consequently, the scalar scores (4.12). When data are right–censored, it is well known that the Harrington and Fleming class of tests can be derived as an efficient score statistic under an accelerated failure time model with error term distribution (4.17). In the likelihood approach of next section, we also derive this result with interval–censored data.

4.5 Weighted Log–rank tests

Finkelstein (1986) proposes an extension to interval–censored data of the proportional hazards model. Finkelstein assumes a discrete lifetime distribution and derives, from the likelihood function, the score statistic that results for testing the hypothesis of a null regression coefficient. This statistic has the form $\sum(O - E)$ and it can be seen as the Log–rank test proposed by Peto and Peto. Because of the discrete nature of the data, Finkelstein uses the Fisher information matrix to derive the asymptotic distribution of the statistic instead of the permutational distribution. Their approach, however, produces numerical problems when applied to a large group of patients because the calculation of the variance–covariance matrix involves dealing with high dimension matrices or because sometimes there are parameters on the boundary of the parameter space. Fay (1996) extends Finkelstein's work to a grouped continuous model. The score statistic for testing the null hypothesis that the failure times are unrelated to the covariates, reduces to the Wilcoxon-Peto or the Log–rank tests as special cases. Fay (1999) shows the equivalence between the permutational linear form of these two tests, see equation (4.1), and a Weighted Log–rank form given by $\sum\ Q^* \cdot (O - E)$. Some particular cases of these Weighted Log–rank tests are also considered in Sun (1996) and Sun *et al.* (2005).

In the remainder of this section we extend the results in Fay (1999). We give a framework which can be used in a natural way to generalize known tests for right–censored data as permutational tests for interval–censored data. In Subsection 4.5.1 we write the Weighted Log–rank tests in terms of the individual distribution function estimates defined by Fay and Shih (1998). Then, we relate the weights in the Weighted Mann–Whitney functional and the weights in the Weighted Log–rank tests. We use this connection to obtain the Weighted Log–rank form of the Harrington and Fleming class of tests. In Subsection 4.5.2 we derive the Harrington and Fleming class of tests as an efficient score statistic. Finally, in Subsection 4.5.3 we describe the generalized Log–

rank tests proposed in Sun *et al.* (2005). We also address the formulation differences between the Harrington and Fleming class of tests considered by Sun *et al.* (2005) and our proposal.

4.5.1 Equivalence between test statistic forms

In this subsection we extend the equivalence between test statistic forms given in Fay (1999). First, we use estimated individual distribution functions $\hat{W}^i$ to write the general form of the Weighted Log–rank tests considered in Fay (1999):

$$\begin{aligned} U &= \int_0^{+\infty} Q^*(t)\ \{O(t) - E(t)\} \\ &= \int_0^{+\infty} Q^*(t)\ \sum_{i=1}^{n} \mathbf{z}_i \left\{ d\hat{W}^i(t) - \frac{1-\hat{W}^i(t^-)}{1-\hat{W}(t^-)} d\hat{W}(t) \right\} \\ &= \sum_{i=1}^{n} \mathbf{z}_i \int_0^{+\infty} Q^*(t) \left\{ d\hat{W}^i(t) - \frac{1-\hat{W}^i(t^-)}{1-\hat{W}(t^-)} d\hat{W}(t) \right\} \end{aligned} \tag{4.18}$$

where $\sum_{i=1}^{n} \mathbf{z}_i d\hat{W}^i(t)$ represents the expected value of the number of deaths in t for the group determined by the covariate $\mathbf{z}_i$, $n \cdot d\hat{W}(t)$ represents the expected value of the total number of deaths in t and, similarly, $\sum_{i=1}^{n} \mathbf{z}_i(1-\hat{W}^i(t^-))$ and $n \cdot (1-\hat{W}(t^-))$ represent the expected number at risk.

Next, we give the relationship between weights in the Weighted Mann–Whitney functional and weights in the Weighted Log–rank tests. From this result, we derive the Weighted Log–rank form for the Harrington and Fleming class of tests given in Section 4.4.

Proposition 4.4. *A Weighted Log–rank statistic U (4.18) with the following weights*

$$Q^*(t) = (1-\hat{W}(t^-))\, \frac{\xi(\hat{W}(t)) - \xi(\hat{W}(t^-))}{\hat{W}(t) - \hat{W}(t^-)} = (1-\hat{W}(t^-))\, \frac{d\xi(\hat{W}(t))}{d\hat{W}(t)}\ ,$$

and a permutational linear statistic L_0 (4.1) with weights in the Weighted Mann–Whitney functional as in (4.13)

$$Q(t) = \frac{\gamma(\hat{W}(t)) - \gamma(\hat{W}(t^-))}{\hat{W}(t) - \hat{W}(t^-)} = \frac{d\gamma(\hat{W}(t))}{d\hat{W}(t)} \ ,$$

are equivalent when

$$\gamma(t) = \frac{1-t}{t}\,\xi(t) \qquad (\xi(0) = \gamma(1) = 0) \ .$$

Proof:

From definitions of the test statistics U and L_0, equivalence between both formulas

$$U = \sum_{i=1}^{n} \mathbf{z}_i \int_0^{+\infty} Q^*(t) \left\{ d\hat{W}^i(t) - \frac{1-\hat{W}^i(t^-)}{1-\hat{W}(t^-)} d\hat{W}(t) \right\} = \sum_{i=1}^{n} \mathbf{z}_i c_i = L_0,$$

imply the definition of the following score values

$$\begin{aligned}
c_i &= \int_0^{+\infty} Q^*(t) \left(d\hat{W}^i(t) - \frac{1-\hat{W}^i(t^-)}{1-\hat{W}(t^-)} d\hat{W}(t) \right) \\
&= \int_0^{+\infty} \frac{d\xi(\hat{W}(t))}{d\hat{W}(t)} \Big((1-\hat{W}(t^-))\, d\hat{W}^i(t) - (1-\hat{W}^i(t^-))\, d\hat{W}(t) \Big) \\
&= \int_0^{+\infty} \frac{d\xi(\hat{W}(t))}{d\hat{W}(t)} \Big((1+d\hat{W}(t)-\hat{W}(t))\, d\hat{W}^i(t) - (1+d\hat{W}^i(t)-\hat{W}^i(t))\, d\hat{W}(t) \Big) \\
&= \int_0^{+\infty} \frac{d\xi(\hat{W}(t))}{d\hat{W}(t)} \Big((1-\hat{W}(t))\, d\hat{W}^i(t) - (1-\hat{W}^i(t))\, d\hat{W}(t) \Big) \ .
\end{aligned}$$

Since of $\hat{W}^i$ is a truncation of $\hat{W}$ at the observed interval $[l_i, r_i]$, following analogous steps to the development in equation (4.15), these scores reduce to

$$c_i = \frac{(1-\hat{W}(r_i))\,\xi(\hat{W}(r_i)) - (1-\hat{W}(l_i^-))\,\xi(\hat{W}(l_i^-))}{\hat{W}(r_i) - \hat{W}(l_i^-)} + \xi(0) \ .$$

Now, if we compare these scores and the scores in equation (4.15) derived from the Weighted Mann–Whitney functional

$$c_i = \frac{\hat{W}(r_i)\,\gamma(\hat{W}(r_i)) - \hat{W}(l_i^-)\,\gamma(\hat{W}(l_i^-))}{\hat{W}(r_i) - \hat{W}(l_i^-)} - \gamma(1)$$

it is clear that

$$\xi(t) = \frac{t}{1-t}\,\gamma(t) \qquad (\xi(0) = \gamma(1) = 0)\ .$$

□

Corollary 4.5. *The Harrington and Fleming class of tests proposed in equation (4.12) admits a Weighted Log–rank form such that,*

$$Q^*(t) = (1 - \hat{W}(t^-))\,\frac{(1-\hat{W}(t))^\rho - (1-\hat{W}(t^-))^\rho}{\rho\,(\hat{W}(t) - \hat{W}(t^-))}\ .$$

When $d\hat{W}(t) \longrightarrow 0$, we obtain a characterization of the Harrington and Fleming weights in Corollary 4.5 similar to the one in the right–censored data framework, $Q^*(t) \longrightarrow -(1-\hat{W}(t^-))^\rho$. If $\rho \neq 0$, the magnitude of the weight function decreases monotonically and the parameter ρ determines the rate of this diminution. As ρ increases, earlier differences are emphasized stronger than late differences.

Remark 4.6. We note that these weights are negative. In accordance with literature, these weights should be consider positive and this chapter should be rewritten in order to have positive weights. The problem is that the sign of these weights is opposite to the sign of the weights in the Weighted Mann–Whitney functional. Thus, we have negative weights in the Weighted Log–rank form of the test and positive weights in the Weighted Mann–Whitney functional. A possible solution is to change the order of the difference between integrals in the definition of Weighted Mann–Whitney functional, see equation (4.7). Nevertheless, we have decided to keep the definition given in Fay and Shih (1998) because this fact does not essentially affect the interpretation of the weights. This weighting scheme implies that $\Phi_{WMW}(\hat{W}^i, \hat{W}^j) > 0$ can be related to the i^{th} individual having higher probability of survival than the j^{th} individual. Similarly, in the two sample problem, a positive value in the test statistic can be related to individuals in the second group having higher probabilities of survival than those in the first group. This interpretation also applies in the k–sample trend problem in Chapter 5.

4.5.2 Score test statistics

When $\mathbf{z}_i = \left(\alpha_i^{(1)}, \alpha_i^{(2)}, \ldots, \alpha_i^{(k)}\right)'$ is a vector of indicators of k treatments and data are discrete or grouped continuous, Fay (1999) explicitly gives the efficient score statistic U and its likelihood based variance–covariance matrix V for a proportional odds model (Wilcoxon–Peto test), a proportional hazards model (Log–rank test) and a logistic model. Then, under the null hypothesis and regularity conditions, UV^-U' is asymptotically chi–squared with $k-1$ degrees of freedom, where V^- is the generalized inverse of the variance–covariance matrix V. We also extend these results to lifetimes following a regression model with hazard function,

$$\lambda(t|\mathbf{z}_i'\beta) = \lambda_0(t)\exp(\mathbf{z}_i'\beta)\{(1-W_0(t))^\rho + [1-(1-W_0(t))^\rho]\exp(\mathbf{z}_i'\beta)\}^{-1},$$

and distribution function,

$$W(t|\mathbf{z}_i'\beta) = 1-(1-W_0(t))\left\{(1-W_0(t))^\rho + [1-(1-W_0(t))^\rho]\exp(\mathbf{z}_i'\beta)\right\}^{-\frac{1}{\rho}}. \quad (4.19)$$

As showed in Fay (1999), the efficient score statistic for a regression model with discrete data or grouped continuous interval censored data can be written in a linear form like the one used with permutational tests, see equation (4.1). The scalar scores of this linear form are derived by Fay as

$$c_i = \frac{\hat{W}'(r_i) - \hat{W}'(l_i^-)}{\hat{W}(r_i) - \hat{W}(l_i^-)} \tag{4.20}$$

where

$$\hat{W}(p_j) = [W(p_j|\eta = \mathbf{z}_i'\beta, \gamma)]_{\eta=0,\gamma=\hat{\gamma}}$$

$$\hat{W}'(p_j) = \left[\frac{\partial W(p_j|\eta = \mathbf{z}_i'\beta, \gamma)}{\partial \eta}\right]_{\eta=0,\gamma=\hat{\gamma}}$$

and γ is a vector of nuisance parameters.

In the present regression model (4.19), we consider the parameterization given by $\gamma_j = \log\left(\frac{1}{\rho}[1-(1-W_0(p_j))^\rho]/(1-W_0(p_j))^\rho\right)$. Then,

$$\hat{W}(p_j) = 1 - [1 + \rho\exp(\hat{\gamma}_j)]^{-\frac{1}{\rho}}$$

and

$$\hat{W}'(p_j) = \frac{1}{\rho}\{1 - \hat{W}(p_j) - (1 - \hat{W}(p_j))^{\rho+1}\}.$$

Substitution of these results in equation (4.20) provides the Harrington and Fleming scores (4.12) except for the sign which is just the opposite. We note that this question about the sign assigned to the scores has already been discussed in Remark 4.6. In next proposition we give the form of V. The proof of this result is omitted because it follows from standard statistical theory and it is analogous to Fay (1999).

Proposition 4.7. *Under the regression model (4.19), the likelihood based variance–covariance matrix V of the efficient score statistic for β is given by*

$$V = -\left[\left(\frac{\partial^2 \log(\mathbf{L})}{\partial\beta\partial\beta'}\right) - \left(\frac{\partial^2 \log(\mathbf{L})}{\partial\beta\partial\gamma'}\right)\left(\frac{\partial^2 \log(\mathbf{L})}{\partial\gamma\partial\gamma'}\right)^{-1}\left(\frac{\partial^2 \log(\mathbf{L})}{\partial\beta\partial\gamma'}\right)\right]_{\beta=0,\gamma=\hat{\gamma}},$$

where for arbitrary parameters ψ_u and ψ_v $(u, v = 1, 2, \ldots, k)$

$$\frac{\partial^2 \log(\mathbf{L})}{\partial \psi_u \partial \psi_v} = \sum_{i=1}^{n} \frac{1}{g_i} \left[\left(\frac{\partial^2 g_i}{\partial \psi_u \partial \psi_v} \right) - \frac{1}{g_i} \left(\frac{\partial g_i}{\partial \psi_u} \right) \left(\frac{\partial g_i}{\partial \psi_v} \right) \right],$$

$$g_i = P_W([l_i, r_i] | \mathbf{z}_i' \beta, \gamma)) = W(r_i | \mathbf{z}_i' \beta, \gamma)) - W(l_i^- | \mathbf{z}_i' \beta, \gamma))$$

and

$$\left[\frac{\partial W(p_j | \mathbf{z}_i' \beta, \gamma)}{\partial \beta_u} \right]_{\beta=0, \gamma=\hat{\gamma}} = \frac{1}{\rho} \{1 - \hat{W}(p_j)\} \{1 - (1 - \hat{W}(p_j))^\rho\} \, \alpha_i^{(u)}$$

$$\left[\frac{\partial W(p_j | \mathbf{z}_i' \beta, \gamma)}{\partial \gamma_u} \right]_{\beta=0, \gamma=\hat{\gamma}} = \frac{1}{\rho} \{1 - \hat{W}(p_j)\} \{1 - (1 - \hat{W}(p_j))^\rho\} \, \mathbf{1}_{\{j=u\}}$$

$$\left[\frac{\partial^2 W(p_j | \mathbf{z}_i' \beta, \gamma)}{\partial \beta_u \partial \beta_v} \right]_{\beta=0, \gamma=\hat{\gamma}} = -\frac{1}{\rho^2} \{1 - \hat{W}(p_j)\} \{1 - (1 - \hat{W}(p_j))^\rho\} \cdot \{1 - (\rho + 1)(1 - \hat{W}(p_j))^\rho\} \, \alpha_i^{(u)} \alpha_i^{(v)}$$

$$\left[\frac{\partial^2 W(p_j | \mathbf{z}_i' \beta, \gamma)}{\partial \beta_u \partial \gamma_v} \right]_{\beta=0, \gamma=\hat{\gamma}} = -\frac{1}{\rho^2} \{1 - \hat{W}(p_j)\} \{1 - (1 - \hat{W}(p_j))^\rho\} \cdot \{1 - (\rho + 1)(1 - \hat{W}(p_j))^\rho\} \, \alpha_i^{(u)} \mathbf{1}_{\{j=v\}}$$

$$\left[\frac{\partial^2 W(p_j | \mathbf{z}_i' \beta, \gamma)}{\partial \gamma_u \partial \gamma_v} \right]_{\beta=0, \gamma=\hat{\gamma}} = -\frac{1}{\rho^2} \{1 - \hat{W}(p_j)\} \{1 - (1 - \hat{W}(p_j))^\rho\} \cdot \{1 - (\rho + 1)(1 - \hat{W}(p_j))^\rho\} \mathbf{1}_{\{j=u=v\}}$$

4.5.3 Generalized Log–rank tests in Sun *et al.* (2005)

When dealing with interval–censored data case 2, Sun *et al.* (2005) propose another approach to derive the asymptotic distribution of the weighted Log–rank tests through the use of the empirical process theory developed in Groeneboom and Wellner (1992) and Groeneboom (1996). As an advantage to the previous methodology in

Subsection 4.5.2, this approach applies to interval-censored data measured on continuous scale. Moreover, a simulation study shows that it performs well when interval–censored data is not strictly a case 2 and arise from periodic follow–up studies. As a disadvantage, however, it does not apply to situations where data include uncensored observations. Finally, it is important to remark that in the simulation study Sun *et al.* (2005) consider a generalization of the Harrington and Fleming class of tests which implies the following scalar scores in the linear permutational form,

$$c_i = \frac{(1-\hat{W}(r_i))^{\rho+1}\log(1-\hat{W}(r_i)) - (1-\hat{W}(l_i^-))^{\rho+1}\log(1-\hat{W}(l_i^-))}{(\hat{W}(r_i)-\hat{W}(l_i^-))} \,. \tag{4.21}$$

Now, following the proof of Proposition 4.4, we have that the weights in the Weighted Log–rank form are given by

$$Q^*(t) = (1-\hat{W}(t^-))\,\frac{(1-\hat{W}(t))^{\rho}\log(1-\hat{W}(t)) - (1-\hat{W}(t^-))^{\rho}\log(1-\hat{W}(t^-))}{\hat{W}(t)-\hat{W}(t^-)} \tag{4.22}$$

where $Q^*(t) \longrightarrow -(1-\hat{W}(t^-))^{\rho}\,(\rho\log(1-\hat{W}(t^-))+1)$ as $d\hat{W}(t) \longrightarrow 0$. Note that the scores in equation (4.21) and the weights in equation (4.22) do not coincide with our proposal of scores in equation (4.12) and weights in Corollary 4.5. It would be interesting to compare the efficiency of the related tests both in the permutational framework and in the Sun *et al.* (2005) framework.

4.6 Generalization of weighted Kaplan–Meier statistics

The approach by Petroni and Wolfe (1994) is different from all the above methods. Their proposal is a class of two sample tests based on Turnbull's estimated survival function from each group and requires a finite pre–specified number of intervals. These tests are based on the integrated weighted difference in Turnbull's estimators and extend the Weighted Kaplan–Meier class developed by Pepe and Fleming (1989) for right–censored data. Under the null hypothesis of no difference between the distributions, the distribution of these tests is asymptotically normal and the variance is obtained via information matrices. This approach is specially indicated under crossing hazard alternatives.

Recently, Fang *et al.* (2002) and Lim and Sun (2003) discuss other generalizations of the Weighted Kaplan–Meier statistics that do not require discrete interval–censored data. Fang *et al.* (2002) derive an asymptotic variance estimate which has a very complex form and they suggest an alternative bootstrap procedure. Lim and Sun (2003) consider test statistics which are based on the integrated weighted difference in estimates of the survival probability function, the hazard function or the cumulative hazard function and they suggest a bootstrap procedure. Via a simulation study, their method is shown to perform quite well for nonmonotone departures from the null hypothesis of equality of survival or hazard functions.

4.7 Computational aspects

The scores, functionals and permutational tests given in this chapter have been implemented with the S–Plus functions given in Appendix B. The function `cdf.data(·,·,·)` uses the output of the `kaplanMeier()` S–plus procedure and it computes the estimated lifetime distribution function at every left and right endpoint of the interval data sample. The following three functions, `WGsc(·,·)`, `HFsc(·,·)` and `DiMsc(·,·,·)` implement, respectively, the Wilcoxon–Gehan scores, the Harrington–Fleming scores and the Difference in Means scores. The test statistic can be computed from each set of scores using either the two sample methodology (`w2test(·,·)`), or the k–sample methodology, (`wktest(·,·)`). These functions assume that the intervals are semi–closed because they use the `kaplanMeier()` procedure which considers semi–closed intervals. If the intervals are closed as it is the case in this chapter, we can replace each interval $[l_i, r_i]$ by

$(l_i - \epsilon, r_i]$ where ϵ is a small quantity.

Let `lower` and `upper` be two vectors containing the left and right endpoints of an interval data sample (when a right endpoint is $+\infty$, we write down `1e+029`), then the steps needed to compute the permutational tests are the following:

A. Estimate the survival function from the pooled sample using Turnbull's method,

```
svf <- kaplanMeier(censor(lower,upper,censor.codes)~1)
```

B. Compute the estimated lifetime distribution function at every left and right endpoint of the interval data sample,

```
cdf <- cdfdata(lower,upper,surv.est)
```

C. Compute the scores values.

   C1. Wilcoxon–Gehan scores: `scores1 <- WGsc(lower,upper)`

   C2. Wilcoxon–Peto scores ($\rho = 1$): `scores2 <- HFsc(cdf,1)`

   C3. Log–rank scores ($\rho = 0$): `scores3 <- HFsc(cdf,0)`

   C4. Difference in means scores: `scores4 <- DiMsc(lower,upper,svf)`

D. Create a vector of covariates `covar` which assigns a numerical value to each individual to distinguish whether the observation belongs to one group or another.

   D1. Two sample problem: assign the value 0 for individuals in the first group and 1 to individuals in the second group.

   D2. k–sample problem: assign the value 1 for individuals in the first group, the value 2 for individuals in the second group and likewise until the k^{th} group. The `wktest(·,·)` routine would transform each covariate value s in a k–vector whose s–component is $1/\sqrt{n_s}$ and the rest of components are 0.

E. Compute the permutational test statistic with the two sample method,

```
w2test(scores,covar)
```

or the k–sample method,

```
wktest(scores,covar).
```

Chapter 5

The Jonckheere's test

An important issue that arises in survival studies is to establish an increasing or decreasing trend in the k–sample problem. In medical and epidemiological studies, survival of the groups is expected to follow an order given by the covariates. For instance, the effect of increasing dose levels of a drug is expected to increase survival. Alternatively, the effect of increasing levels of exposure to a risk factor is expected to decrease survival. Trends like these can be examined with the the permutational linear test statistic considered in Chapter 4. If groups are monotonically ordered according to a covariate $\mathbf{z}_i$, the use of this covariate in equation (4.1) gives a test for detecting this trend. Without the requirement of a covariate specification, Jonckheere (1954) and Terpstra (1952) were among the first to develop a nonparametric statistic to test for monotonically ordered alternatives. This test for trend has received much attention in the uncensored and right–censored data literature. Abel (1986) proposal is perhaps one of the few generalizations of the Jonckheere's test for interval–censored data.

In the uncensored data framework, the Jonckheere's test is an extension of the Mann–Whitney test. As an alternative, Puri (1965) presents a modification of the Jonckheere's test which is an extension of the Chernoff–Savage test. Some other modifications for the Jonckheere's test are recently proposed in Neuhäuser *et al.* (1998), Büning and Kössler (1999) and Terpstra and Magel (2003). The proposal in Neuhäuser *et al.* (1998) shows substantially more power in the detection of trend than the Jonckheere's test. Büning and Kössler (1999) define what they call the class of Jonckheere–type tests, a proposal which is similar to the one in Puri (1965). Since the Jonckheere's test is based on a sum of two–sample Mann–Whitney statistics, Büning and Kössler (1999) study the modification of the test with other two-sample statistics and obtain a general expression for the asymptotic power function. Moreover, the authors prove that this proposal is asymptotically equivalent to Puri's proposal. Terpstra and Magel (2003) raise the following important issue: a test for trend should have low power for any alternative that does not fit the profile given in the alternative hypothesis. The authors propose a new trend test in order to tackle this problem. When the alternative hypothesis holds, the Terpstra and Magel (2003) proposal has similar power to the Jonckheere's test and the modification in Neuhäuser *et al.* (1998). When the assumed ordering in the alternative hypothesis is uncorrect, it has lower power. For further reading on the Jonckheere's test and possible extensions, see Hollander and Wolfe (1999), Robertson *et al.* (1988) and Barlow *et al.* (1972).

The Log–rank trend tests in Liu *et al.* (1993), Liu *et al.* (1998) and Liu and Tsai (1999) are similar to the Jonckheere's test and are commonly used in the right–censored data framework. In Jones (2001) it is showed that these Jonckheere–type test statistics are special cases of the class of single–covariate nonparametric test statistics introduced by Jones and Crowley (1989) and Jones and Crowley (1990). Thus, each Jonckheere–type test statistic is seeking for the trend defined by a non–explicit covariate. The author explicitly gives these time–dependent covariates and shows that they depend on the initial group sizes and censoring distributions. As an alternative approach not requiring specification of a covariate, we mention the order–restricted inference method used in Mau (1988) and the ordered test based on two–sample weighted Kaplan–Meier statistics proposed in Chi (2002).

In the literature, as far as we know and except for Abel (1986), there are not extensions of the Jonckheere's test for interval–censored data. This chapter deals with new proposals which extend this test for interval–censored data. In Section 5.1 we introduce the Jonckheere's test for uncensored data as a sum of two–sample Mann–Whitney test statistics. In Section 5.2 and Section 5.3, we suggest different test statistics which extend in a natural way the Jonckheere's test for interval–censored data. These Jonckheere–type tests are sum of the two–sample test statistics studied in Chapter 4, such as the Wilcoxon–Gehan test statistics, the Harrington and Fleming class of test statistics and the Difference in Means test statistic. We use permutational and bootstrap methods to obtain an asymptotic distribution for these tests. In Section 5.4 we give computer programs for each proposal. In Section 5.5 we perform a simulation study in order compare the power of each proposal under different parametric assumptions and different alternatives. In Section 5.6, we end this chapter with the analysis of a set of data.

5.1 Uncensored data

A nonparametric test for trend considers the hypothesis $H_0 : W_1 = \cdots = W_k$ against the alternative of stochastic order $H_a : W_1 \geq \cdots \geq W_k$. In terms of survival probabilities, the alternative hypothesis imply that individuals in the first groups have lower probabilities of survival that those in the last groups. To investigate asymptotic efficiencies with ordered alternatives, Tryon and Hettmansperger (1973) or Büning and Kössler (1999) consider the location parameter problem. That is, the unknown distributions of T under each group are assumed to be of the same type and only differ in location $W_i(t) = W(t - \theta_i)$, if at all. In this situation, the hypothesis of interest are $H_0 : \theta_1 = \cdots = \theta_k$ versus $H_a : \theta_1 \leq \cdots \leq \theta_k$. In a lifetime data analysis framework, the same simplification of the hypothesis hold when we consider a location shift in an accelerated failure time regression model, $\log(T) = \theta + \lambda Z$. That is, the distribution function W for the random variable Z and the scale parameter λ are assumed to be equal for all lifetimes but the location parameter θ may differ among groups. In this model the distribution function of T under each group is of the type $W_i(t) = W(\frac{\log(t) - \theta_i}{\lambda})$.

When data are uncensored, the Jonckheere's statistic for testing the nondecreasing ordered alternative is a sum of two–sample Mann–Whitney statistics,

$$J = \sum_{\substack{r,s=1 \\ r<s}}^{k} M_{r;s} = \sum_{\substack{r,s=1 \\ r<s}}^{k} \sum_{i,j=1}^{n} \alpha_i^{(s)} \, \alpha_j^{(r)} \, \Phi_M(i,j) \tag{5.1}$$

where $\Phi_M(i,j) = \mathbf{1}_{\{t_j<t_i\}}$. Equivalently, the Jonckheere's test can also be expressed as

$$J = \sum_{s=2}^{k} M_{1,\ldots,s-1;s} = \sum_{s=2}^{k} \sum_{i,j=1}^{n} \alpha_i^{(s)} \left(\sum_{r=1}^{s-1} \alpha_j^{(r)} \right) \Phi_M(i,j) \tag{5.2}$$

or

$$J = \sum_{r=1}^{k-1} M_{r;r+1,\ldots,k} = \sum_{r=1}^{k-1} \sum_{i,j=1}^{n} \left(\sum_{s=r+1}^{k} \alpha_i^{(s)} \right) \alpha_j^{(r)} \, \Phi_M(i,j) \tag{5.3}$$

where $M_{1,\ldots,s-1;s}$ denotes the Mann–Whitney statistic computed for the pooled group $G_1 \cup \cdots \cup G_{s-1}$ versus G_s, and $M_{r;r+1,\ldots,k}$ denotes the Mann–Whitney statistic computed for G_r versus the pooled group $G_{r+1} \cup \cdots \cup G_k$.

The Jonckheere's test was proposed independently by Terpstra (1952) and Jonckheere (1954). The statistic J coincides with the Mann–Whitney test when $k = 2$. Moreover, it can be viewed as a sum of two–sample Wilcoxon tests or as a Kendall's correlation coefficient. The trend sought by the alternative hypothesis is evidenced by larger J. Although critical values for J have been tabulated from the exact probability distribution, a large–sample normal approximation is usually applied. Under the null hypothesis, the expected value and variance of J are

$$E(J) = \frac{1}{4} \left[n^2 - \sum_{r=1}^{k} n_r^2 \right]$$

and

$$V(J) = \frac{1}{72} \left[n^2(2n+3) - \sum_{r=1}^{k} n_r^2 (2n_r + 3) \right] \tag{5.4}$$

If there are ties (Hollander and Wolfe, 1999), then

$$V(J) = \tfrac{1}{72}\left[n(n-1)(2n+5) - \sum_{r=1}^{k} n_r(n_r-1)(2n_r+5) - \sum_{s=1}^{e} d_s(d_s-1)(2d_s+5)\right]$$
$$+ \tfrac{1}{36n(n-1)(n-2)}\left[\sum_{r=1}^{k} n_r(n_r-1)(n_r-2)\right]\left[\sum_{s=1}^{e} d_s(d_s-1)(d_s-2)\right]$$
$$+ \tfrac{1}{8n(n-1)}\left[\sum_{r=1}^{k} n_r(n_r-1)\right]\left[\sum_{s=1}^{e} d_s(d_s-1)\right] \quad (5.5)$$

where e is the number of different values, d_1 is the number observations which are equal to the smallest value, d_2 is the number observations which are equal to the next smallest, and so on.

5.2 Jonckheere–type tests

When data are interval–censored, the Jonckheere's test allow different ways to be extended by means of any two–sample statistic other than the Mann–Whitney statistic. In this section we begin introducing the proposal in Abel (1986) which is a sum of two–sample Wilcoxon–Gehan statistics. Next, in Subsection 5.2.1 we suggest other Jonckheere–type tests which are sum of the Weighted Log–rank statistics introduced in Chapter 4. Finally, Subsection 5.2.2 is devoted to the permutational distribution of these Jonckheere–type tests. We discuss some problems of this permutational approach which motivate the modification we propose in Section 5.3.

The Abel's test statistic is a sum of the two–sample Wilcoxon–Gehan statistics L^{WG} given in equations (4.11) and (4.3). If we consider J in equation (5.1), the analogous form of the Abel's test is the following:

$$JA = \sum_{\substack{r,s=1 \\ r<s}}^{k} L^{WG}_{r;s} = \sum_{\substack{r,s=1 \\ r<s}}^{k} \sum_{i,j=1}^{n} \alpha_i^{(s)}\, \alpha_j^{(r)}\, \Phi_{WG}(i,j) \quad (5.6)$$

where $\Phi_{WG}(i,j) = \mathbf{1}_{\{r_j<l_i\}} - \mathbf{1}_{\{l_j>r_i\}}$. Equivalently, the Abel's test can also be expressed a generalization of equation (5.2)

$$JA = \sum_{s=2}^{k} L^{WG}_{1,\ldots,s-1;s} = \sum_{s=2}^{k} \sum_{i,j=1}^{n} \alpha_i^{(s)} \left(\sum_{r=1}^{s-1} \alpha_j^{(r)} \right) \Phi_{WG}(i,j) \tag{5.7}$$

or equation (5.3)

$$JA = \sum_{r=1}^{k-1} L^{WG}_{r;r+1,\ldots,k} = \sum_{r=1}^{k-1} \sum_{i,j=1}^{n} \left(\sum_{s=r+1}^{k} \alpha_i^{(s)} \right) \alpha_j^{(r)} \Phi_{WG}(i,j).$$

where $L^{WG}_{1,\ldots,s-1;s}$ denotes the Wilcoxon–Gehan statistic computed for the pooled group $G_1 \cup \cdots \cup G_{s-1}$ versus G_s, and $L^{WG}_{r;r+1,\ldots,k}$ denotes the Wilcoxon–Gehan statistic computed for G_r versus the pooled group $G_{r+1} \cup \cdots \cup G_k$.

For the distribution of JA, the author considers a permutational approach. Under the null hypothesis the permutational distribution of JA is asymptotically normal with mean equal to zero. The permutational variance given in Abel (1986) will be presented later in Subsection 5.3.3.

5.2.1 Weighted Log–rank tests under order restrictions

As a new extension of the Jonckheere's test we propose to use the two–sample Weighted Log–rank statistics U given in equations (4.11), (4.6), (4.7) and (4.18). In this case, the three forms of the Jonckheere's statistic in equations (5.1), (5.2) and (5.3) provide three extensions which are not equivalent. The first Jonckheere–type test statistic is,

$$\begin{aligned} J_1 &= \sum_{\substack{r,s=1 \\ r<s}}^{k} U_{r;s} = \sum_{\substack{r,s=1 \\ r<s}}^{k} \sum_{i=1}^{n} \alpha_i^{(s)} \Phi^{WMW}_{r,s}(\hat{W}^i_{r,s}, \hat{W}_{r,s}) \\ &= \sum_{\substack{r,s=1 \\ r<s}}^{k} \sum_{i=1}^{n} \alpha_i^{(s)} \int_0^{+\infty} Q^*_{r,s}(t) \left\{ d\hat{W}^i_{r,s}(t) - \frac{1-\hat{W}^i_{r,s}(t^-)}{1-\hat{W}_{r,s}(t^-)} d\hat{W}_{r,s}(t) \right\} \end{aligned} \tag{5.8}$$

where $\hat{W}_{r,s}$ is the Turnbull's estimate of the distribution function in the pooled group $G_r \cup G_s$, which is used to derive the estimate of the distribution function for the i^{th} observation, $\hat{W}^i_{r,s}$, the Weigthed Mann–Withney functional $\Phi^{WMW}_{r,s}$ and the weights $Q^*_{r,s}(t)$. Alternatively, the other Jonckheere–type test statistics are given by

$$\begin{aligned} J_2 &= \sum_{s=2}^{k} U_{1,\ldots,s-1;s} = \sum_{s=2}^{k}\sum_{i=1}^{n} \alpha_i^{(s)}\ \Phi^{WMW}_{1,\ldots,s}(\hat{W}^i_{1,\ldots,s}, \hat{W}_{1,\ldots,s}) \\ &= \sum_{s=2}^{k}\sum_{i=1}^{n} \alpha_i^{(s)} \int_0^{+\infty} Q^*_{1,\ldots,s}(t) \left\{ d\hat{W}^i_{1,\ldots,s}(t) - \frac{1-\hat{W}^i_{1,\ldots,s}(t^-)}{1-\hat{W}_{1,\ldots,s}(t^-)} d\hat{W}_{1,\ldots,s}(t) \right\} \end{aligned} \tag{5.9}$$

and

$$\begin{aligned} J_3 &= \sum_{r=1}^{k-1} U_{r;r+1,\ldots,k} = \sum_{r=1}^{k-1}\sum_{i=1}^{n} \left(\sum_{s=r+1}^{k} \alpha_i^{(s)} \right) \Phi^{WMW}_{r,\ldots,k}(\hat{W}^i_{r,\ldots,k}, \hat{W}_{r,\ldots,k}) \\ &= \sum_{r=1}^{k-1}\sum_{i=1}^{n} \left(\sum_{s=r+1}^{k} \alpha_i^{(s)} \right) \int_0^{+\infty} Q^*_{r,\ldots,k}(t) \left\{ d\hat{W}^i_{r,\ldots,k}(t) - \frac{1-\hat{W}^i_{r,\ldots,k}(t^-)}{1-\hat{W}_{r,\ldots,k}(t^-)} d\hat{W}_{r,\ldots,k}(t) \right\} \end{aligned} \tag{5.10}$$

The test statistics J_1, J_2 and J_3 are at the same time an extension of the Jonckheere's test modifications for right censored–data given in Liu *et al.* (1993) and Liu and Tsai (1999). As these authors note for the right–censored data framework, the way each test statistic is built determine the power to possible alternative configurations. For instance, in an extreme case like $H_a : W_1 = \cdots = W_{k-1} \geq W_k$ the test statistics J_3 may have more power than J_1 and J_2. This is because in this situation $J_3 = \sum_{r=1}^{k-1} U_{r;r+1,\ldots,k}$ will detect differences in every $U_{r;r+1,\ldots,k}$, however $J_1 = \sum_{\substack{r,s=1 \\ r<s}}^{k} U_{r;s}$ will only detect differences in the two–sample statistics $U_{r;k}$ $(r = 1, \ldots, k-1)$ and $J_2 = \sum_{s=2}^{k} U_{1,\ldots,s-1;s}$ only in $U_{1,\ldots,k-1;k}$. On the other hand, in a case like $H_a : W_1 \geq W_2 = \cdots = W_k$ the test statistic J_2 may have more power than J_1 and J_3. We will explore this power features in the simulation study of Section 5.5.

5.2.2 Permutational distribution

In the case of right censored–data, counting process approaches can be applied to derive the asymptotic distribution of the Jonkheere's test modifications given in Liu *et al.* (1993) and Liu and Tsai (1999). In the case of interval censored–data, we could have tried a likelihood approach for tests J_1, J_2 and J_3. However, this approach needs a discrete nature of the data and it surely would have involved complicated variance estimation. As in Abel (1986) for the test JA, we propose to use the permutational approach for tests J_1, J_2 and J_3 because it is less restrictive and more simple. If the null hypothesis is true and the censoring mechanism does not depend on the grouping, the data labels are exchangeable. Thus, the permutational distribution of the test is obtained by permuting the data labels or, equivalently, regrouping the data in k groups where the size of the i^{th} group is n_i. Corresponding to this, we have to calculate the test statistic for each of the $n!/(n_1! \ldots n_k!)$ partitions of the numbers $1, \ldots, n$. The null hypothesis will be rejected if the observed statistic value is extreme for the permutational distribution. Unfortunately, as the number of observations increase the number of partitions becomes less manageable and the exact permutational distribution is computationally intensive. One way of solving this problem is to find a normal approximation of the permutational distribution. Another way is to use a Monte Carlo approach and consider a random sample of all possible partitions.

In Abel (1986) a normal approximation of the permutational distribution is used. For proposals in equations (5.8), (5.9) and (5.10), we use a Monte Carlo approach. However, since each resampling step needs of the computation of the Turnbull's estimate for the lifetime distribution of the $k(k-1)/2$ in J_1 and $k-1$ pooled groups in J_2 and J_3, this method does not alleviate the computations needed to obtain the permutational distribution. In the following sections, we give a new modification of the Jonckheere's test which reduce this computational work. We base this proposal on the estimate of the lifetime distribution function for each observation introduced in Fay and Shih

(1998), see equation (4.10).

5.3 Kendall–type tests

As defined in Terpstra (1952), the Jonckheere's test for uncensored data is a Kendall's correlation coefficient. In a right–censored data framework, Jones and Crowley (1989) also show this relationship with the modification proposed in Abel (1986). Thus, the idea in this section is to introduce a new Jonckheere–type test such that the normal approximation of the permutational distribution could be easily derived from rank correlation theory, see Kendall and Gibbons (1990). In the following definition we replace the Wilcoxon–Gehan functional used in equation (5.6) by the Weighted Mann–Whitney functional defined in equation (4.7) and proposed in Fay and Shih (1998)

$$JK = \sum_{\substack{r,s=1 \\ r<s}}^{k} \sum_{i,j=1}^{n} \alpha_i^{(s)} \, \alpha_j^{(r)} \, \Phi_{WMW}(\hat{W}^i, \hat{W}^j) = \sum_{s=2}^{k} \sum_{i,j=1}^{n} \alpha_i^{(s)} \left(\sum_{r=1}^{s-1} \alpha_j^{(r)} \right) \Phi_{WMW}(\hat{W}^i, \hat{W}^j)$$
$$= \sum_{r=1}^{k-1} \sum_{i,j=1}^{n} \left(\sum_{s=r+1}^{k} \alpha_i^{(s)} \right) \alpha_j^{(r)} \, \Phi_{WMW}(\hat{W}^i, \hat{W}^j) \tag{5.11}$$

When $k = 2$, this proposal is equivalent to J_1, J_2 and J_3 given in equations (5.8), (5.9) and (5.10). However, that is not true when $k > 2$. In this statistic we use the overall estimation of the lifetime distribution $\hat{W}$ to derive the estimate of the lifetime distribution for the i^{th} observation, $\hat{W}^i$, and the weights in the functional, Φ_{WMW}, independently of the grouping. This fact simplifies the permutational distribution of the test and it is the main difference to J_1, J_2 and J_3.

Proposition 5.1. *The Jonckheere's test in equation (5.11) is equivalent to a Kendall's* *correlation coefficient,*

$$JK = \frac{1}{2} \sum_{i,j=1}^{n} a_{ij}\, b_{ij}$$

where $a_{ij} = \Phi_{WMW}(\hat{W}^i, \hat{W}^j)$ *and* $b_{ij} = \sum_{\substack{r,s=1 \\ r<s}}^{k} \alpha_i^{(s)}\, \alpha_j^{(r)} - \sum_{\substack{r,s=1 \\ r>s}}^{k} \alpha_i^{(s)}\, \alpha_j^{(r)}$.

Proof:

Since the necessary antisymmetric properties of a Kendall's correlation coefficient hold, $a_{ij} = -a_{ji}$ and $b_{ij} = -b_{ji}$, we only need to show the equivalence of both expressions. We use the antisymmetric property of the functional Φ_{WMW} and a rearrangement of indices to proof this result:

$$\begin{aligned}
\sum_{i,j=1}^{n} a_{ij}\, b_{ij} &= \sum_{i,j=1}^{n} \Phi_{WMW}(\hat{W}^i, \hat{W}^j) \left(\sum_{\substack{r,s=1 \\ r<s}}^{k} \alpha_i^{(s)}\, \alpha_j^{(r)} - \sum_{\substack{r,s=1 \\ r>s}}^{k} \alpha_i^{(s)}\, \alpha_j^{(r)} \right) \\
&= \sum_{\substack{r,s=1 \\ r<s}}^{k} \sum_{i,j=1}^{n} \alpha_i^{(s)}\, \alpha_j^{(r)} \Phi_{WMW}(\hat{W}^i, \hat{W}^j) - \sum_{\substack{r,s=1 \\ r>s}}^{k} \sum_{i,j=1}^{n} \alpha_i^{(s)}\, \alpha_j^{(r)} \Phi_{WMW}(\hat{W}^i, \hat{W}^j) \\
&= \sum_{\substack{r,s=1 \\ r<s}}^{k} \sum_{i,j=1}^{n} \alpha_i^{(s)}\, \alpha_j^{(r)} \Phi_{WMW}(\hat{W}^i, \hat{W}^j) + \sum_{\substack{r,s=1 \\ r>s}}^{k} \sum_{i,j=1}^{n} \alpha_i^{(s)}\, \alpha_j^{(r)} \Phi_{WMW}(\hat{W}^j, \hat{W}^i) \\
&= 2 \sum_{\substack{r,s=1 \\ r<s}}^{k} \sum_{i,j=1}^{n} \alpha_i^{(s)}\, \alpha_j^{(r)} \Phi_{WMW}(\hat{W}^i, \hat{W}^j)
\end{aligned}$$

□

Theorem 5.2. *Under the null hypothesis, the permutational distribution of the* JK *test statistic is asymptotically normal with zero mean and variance given by*

$$
\begin{aligned}
V(JK) =& \tfrac{n^3-3n^2-\sum_{r=1}^{k} n_r^3+3\sum_{r=1}^{k} n_r^2}{3n(n-1)(n-2)} \sum_{i=1}^{n}\left[\sum_{j=1}^{n}\Phi_{WMW}(\hat{W}^i,\hat{W}^j)\right]^2 \\
&+ \tfrac{n^3+2\sum_{r=1}^{k} n_r^3-3n\sum_{r=1}^{k} n_r^2}{6n(n-1)(n-2)} \sum_{i,j=1}^{n}\left[\Phi_{WMW}(\hat{W}^i,\hat{W}^j)\right]^2
\end{aligned} \tag{5.12}
$$

Proof:

From Proposition 5.1, the permutational distribution of the JK test statistic coincides with the distribution of a Kendall's correlation coefficient. Thus, see Kendall and Gibbons (1990), the permutational distribution is asymptotically normal with zero mean and variance given by

$$
\begin{aligned}
V(JK) =& \tfrac{1}{n(n-1)(n-2)}\left[\sum_{i,j_1,j_2=1}^{n} a_{ij_1}a_{ij_2} - \sum_{i,j=1}^{n} a_{ij}^2\right]\left[\sum_{i,j_1,j_2=1}^{n} b_{ij_1}b_{ij_2} - \sum_{i,j=1}^{n} b_{ij}^2\right] \\
&+ \tfrac{1}{2n(n-1)}\left[\sum_{i,j=1}^{n} a_{ij}^2\right]\left[\sum_{i,j=1}^{n} b_{ij}^2\right]
\end{aligned} \tag{5.13}
$$

Since $a_{ij} = \Phi_{WMW}(\hat{W}^i,\hat{W}^j)$, it immediately follows that

$$
\sum_{i,j=1}^{n} a_{ij}^2 = \sum_{i,j=1}^{n}\left[\Phi_{WMW}(\hat{W}^i,\hat{W}^j)\right]^2
$$

and

$$
\sum_{i,j_1,j_2=1}^{n} a_{ij_1}\, a_{ij_2} = \sum_{i=1}^{n}\left[\sum_{j=1}^{n}\Phi_{WMW}(\hat{W}^i,\hat{W}^j)\right]^2
$$

For the coefficient $b_{ij} = \sum_{\substack{r,s=1 \\ r<s}}^{k} \alpha_i^{(s)} \alpha_j^{(r)} - \sum_{\substack{r,s=1 \\ r>s}}^{k} \alpha_i^{(s)} \alpha_j^{(r)}$, Kendall and Gibbons (1990) show that

$$\sum_{i,j=1}^{n} b_{ij}^2 = n^2 - \sum_{r=1}^{k} n_r^2$$

and

$$\sum_{i,j_1,j_2=1}^{n} b_{ij_1} b_{ij_2} = \frac{1}{3}\left[n^3 - \sum_{r=1}^{k} n_r^3\right].$$

We also derive these two sums in Appendix C.

Finally, if we replace all these results in equation (5.13) and we do some algebraic manipulation, we obtain the stated expression for $V(JK)$.

□

5.3.1 Comparison with a k–sample test

In this subsection we compare JK with the permutational linear test statistic L_0 given in equation (4.1). First, we note that if we define $z_i = \sum_{r=1}^{k} \alpha_i^{(r)} r$ as a covariate indicating the group ordering of the i^{th} individual, then in Proposition 5.1 it follows that $b_{ij} = sign(z_i - z_j)$. Thus, the coefficient b_{ij} is equal to 0 if the i^{th} and j^{th} individuals belong to the same group, is equal to 1 if the group ordering of the i^{th} individual is bigger than the group ordering of the j^{th} individual, and -1 otherwise. Intuitively, for each pair of observations, the JK test statistic is measuring the agreement of the group ordering and the ordering given by the estimates of the lifetime distribution, $\hat{W}^i$ and $\hat{W}^j$. We also note that if we define a new Kendall's correlation coefficient with $b_{ij} = z_i - z_j = \sum_{r,s=1}^{k} \alpha_i^{(r)} \alpha_j^{(s)} (r - s)$ and $a_{ij} = \Phi_{WMW}(\hat{W}^i, \hat{W}^j)$, or alternatively $a_{ij} = c_i - c_j = \Phi_{WMW}(\hat{W}^i, \hat{W}) - \Phi_{WMW}(\hat{W}^j, \hat{W})$, then the statistic simplifies to L_0.

This result is important because illustrates the differences between using the JK test statistic and the L_0 test statistic with these particular covariates. Since the group ordering covariates $z_i = \sum_{r=1}^{k} \alpha_i^{(r)} r$ are linear, it suggests that JK may be more efficient than L_0 when the true trend is not linear. In the simulations of Section 5.5 we study the behavior of JK and L_0. We consider accelerated failure models with equal spacings in the location parameters (linear trend) and models with different spacings (nonlinear trend).

5.3.2 Dependence on group sizes

Another useful interpretation of the JK test statistic is given by the definition of $\hat{W}_r^* = \frac{1}{n_r} \sum_{i=1}^{n} \alpha_i^{(r)} \hat{W}^i$, $\hat{W}^*_{1,\ldots,s-1} = \frac{1}{n_1+\cdots+n_{s-1}} \sum_{i=1}^{n} \left(\sum_{r=1}^{s-1} \alpha_j^{(r)} \right) \hat{W}^i$ and $\hat{W}^*_{r+1,\ldots,k} = \frac{1}{n_{r+1}+\cdots+n_k} \sum_{i=1}^{n} \left(\sum_{s=r+1}^{k} \alpha_j^{(s)} \right) \hat{W}^i$ as respective estimates of the lifetime distribution in group G_r, in the pooled group $G_1 \cup \cdots \cup G_{s-1}$ and in the pooled group $G_{r+1} \cup \cdots \cup G_k$. Then, by linearity of the Weighted Mann–Whitney functional,

$$JK = \sum_{\substack{r,\, s\, =\, 1 \\ r\, <\, s}}^{k} n_r \, n_s \; \Phi_{WMW}(\hat{W}_s^*, \hat{W}_r^*)$$

or equivalently,

$$\begin{aligned} JK &= \sum_{s=2}^{k} (n_1 + \cdots + n_{s-1}) \, n_s \; \Phi_{WMW}(\hat{W}_s^*, \hat{W}^*_{1,\ldots,s-1}) \\ &= \sum_{r=1}^{k-1} n_r \; (n_{r+1} + \cdots + n_k) \, \Phi_{WMW}(\hat{W}^*_{r+1,\ldots,k}, \hat{W}_r^*) \end{aligned}$$

These expressions show that the test statistic compares lifetime distributions between groups but is sensitive to differences in group sample sizes. Those groups with larger sample size play a more important role in the detection of the trend. Consequently, it is possible that the set of alternatives against which JK is consistent may depend on the group sample sizes. This problem is mentioned in Barlow *et al.* (1972) for the Jonckheere's test and in Jones (2001) for Jonckheere–type tests with right–censored data. As an alternative, a weighted form of the JK test statistic can be used to remove this feature,

$$\begin{aligned} WJK &= \sum_{\substack{r,s=1 \\ r<s}}^{k} \Omega_{r,s} \sum_{i,j=1}^{n} \alpha_i^{(s)}\, \alpha_j^{(r)}\, \Phi_{WMW}(\hat{W}^i, \hat{W}^j) \\ &= \sum_{\substack{r,s=1 \\ r<s}}^{k} \Omega_{r,s}\, n_r\, n_s\, \Phi_{WMW}(\hat{W}_s^*, \hat{W}_r^*) \\ &= \frac{1}{2} \sum_{i,j=1}^{n} a_{ij}\, b_{ij} \end{aligned}$$

where $a_{ij} = \Phi_{WMW}(\hat{W}^i, \hat{W}^j)$, $b_{ij} = \sum_{\substack{r,s=1 \\ r<s}}^{k} \Omega_{r,s}\, \alpha_i^{(s)}\, \alpha_j^{(r)} - \sum_{\substack{r,s=1 \\ r>s}}^{k} \Omega_{r,s}\, \alpha_i^{(s)}\, \alpha_j^{(r)}$ and $\Omega_{r,s} = \Omega_{s,r} > 0$.

For $\Omega_{r,s} = 1/(n_r n_s)$, the statistic WJK compare groups in pairs and it does not depend on the group sample sizes. This is the solution suggested in Barlow *et al.* (1972) for the Jonckheere's test. For $\Omega_{r,s} = 1/(n_r + n_s)$, the statistic WJK and J_1 are equivalent when data are not censored and coincide with the Jonckheere–type test proposed in Puri (1965). Different weight schemes in Puri's proposal are studied in Tryon and Hettmansperger (1973). The authors assume equal group sample sizes and prove maximum Pitman efficiency of Puri's proposal when the alternative specifies equal spacings in the location parameters. They also give the optimum weight scheme for location alternatives without equal spacings.

As a Kendall's correlation coefficient, the permutational distribution of WJK is asymptotically normal. The main problem of this statistic is that the permutational variance does not have a simple form and we must use the general formula given in equation (5.13). For this reason, we are not going to study further this WJK statistic. In the simulation study in Section 5.5 we will pay attention to the performance of the JK test statistic when group sample sizes are not equal or when the true trend is not linear.

5.3.3 Abel's test

In this subsection we consider again the test statistic JA given in equation (5.6). Note that if we replace the Weighted Mann–Whitney functional by the Wilcoxon–Gehan functional, the JA test statistic is a particular case of the JK test statistic in equation (5.11). Thus, the permutational variance given in equation (5.12) applies. However, this is not the permutational variance given in Abel (1986). Here, we compare both variances and show that they are asymptotically equivalent.

The permutational variance derived in Abel (1986) follows from the decomposition of the JA given in equation (5.7). Since the author assure that the statistics $L^{WG}_{1,\dots,s-1;s}$ are independent, it follows that

$$V(JA) = \sum_{s=2}^{k} V(L^{WG}_{1,\dots,s-1;s}) \tag{5.14}$$

Moreover, the author obtain the permutational variance of each statistic from the two–sample method presented in Chapter 4 (see equation (4.4) and equation (4.11)),

$$V_{Abel}(L^{WG}_{1,\dots,s-1;s}) = \frac{\left(\sum_{r=1}^{s-1} n_r\right) n_s}{\left(\sum_{r=1}^{s} n_r\right)\left(\sum_{r=1}^{s} n_r - 1\right)} \sum_{i=1}^{n} \left[\sum_{j=1}^{n} \sum_{r_1,r_2=1}^{s} \alpha_i^{(r_1)} \alpha_j^{(r_2)} \Phi_{WG}(i,j) \right]^2 \tag{5.15}$$

This permutational variance is, however, incorrect because it is based on the permutation of data labels from individuals within the groups $G_1, \dots, G_s$ and neglect the permutation of data labels from individuals outside these groups. The permutational distribution of the JA test statistic, and consequently the $L^{WG}_{1,\dots,s-1;s}$ statistics, must be a rearrangement of data labels from the overall sample. We prove in Appendix D that each statistic $L^{WG}_{1,\dots,s-1;s}$ can be rewritten as a Kendall's correlation coefficient which uses the overall sample. Thus, the permutational distribution of $L^{WG}_{1,\dots,s-1;s}$ coincides with the distribution of a Kendall's correlation coefficient. In Appendix D, we derive the correct permutational variance as

$$V(L^{WG}_{1,\dots,s-1;s}) = \frac{(\sum_{r=1}^{s-1} n_r)\; n_s\; (\sum_{r=1}^{s} n_r\; -2)}{n\;(n-1)\;(n-2)} \sum_{i=1}^{n} \left[\sum_{j=1}^{n} \Phi_{WG}(i,j) \right]^2 + \frac{(\sum_{r=1}^{s-1} n_r)\; n_s\; (n - \sum_{r=1}^{s} n_r)}{n\;(n-1)\;(n-2)} \sum_{i,j=1}^{n} \left[\Phi_{WG}(i,j)\right]^2$$

We also show that the $L^{WG}_{1,\dots,s-1;s}$ statistics are uncorrelated under the permutational distribution and, consequently, the variance decomposition in equation (5.14) is correct. From the sum of variances, we obtain for $V(JA)$ an equivalent expression to $V(JK)$ (see equation (5.12)),

$$V(JA) = \frac{n^3 - 3n^2 - \sum_{r=1}^{k} n_r^3 + 3\sum_{r=1}^{k} n_r^2}{3n(n-1)(n-2)} \sum_{i=1}^{n} \left[\sum_{j=1}^{n} \Phi_{WG}(i,j) \right]^2 + \frac{n^3 + 2\sum_{r=1}^{k} n_r^3 - 3n\sum_{r=1}^{k} n_r^2}{6n(n-1)(n-2)} \sum_{i,j=1}^{n} \left[\Phi_{WG}(i,j)\right]^2 . \tag{5.16}$$

When data are not censored and there are not ties, both permutational variances (5.15) and (5.16) coincide with the variance of the Jonckheere's test (5.4). In this situation, the Jonckheere's test statistic can indistinctly be written as a function of the uncensored data, the ranks of the data in the overall sample or the ranks of the data within each of the pooled groups $G_1 \cup \cdots \cup G_s$ $(s = 2, \dots, k)$. In presence of ties, however, only the permutational variance (5.16) coincide with the variance of the Jonckheere's test (5.5).

A question which naturally arises is how different are the correct permutational variance of JA and the variance given in Abel (1986). As noted in Heimann and Neuhaus (1998), when the α–critical value of the permutational distribution of a test statistic converges in probability to the α–critical value of the non–permutational distribution, both distributions are asymptotically equivalent. In this situation, $V(JA)$ will converge to the non–permutational variance. Moreover, $V_{Abel}(JA)$ will also converge to the non–permutational variance because each $V_{Abel}(L^{WG}_{1,\ldots,s-1;s})$ will do. Consequently, $V(JA)$ and $V_{Abel}(JA)$ will be asymptotically equivalent. In a general asymptotic situation, however, the permutational distribution of a test statistic might not coincide with the non–permutational distribution. In that case, as the following proposition shows, both permutational variances are still asymptotically equivalent.

Proposition 5.3. *The variances* $V(JA)$ *and* $V_{Abel}(JA)$ *are asymptotically equivalent.*

Proof:

To prove the proposition, we show that $V(L^{WG}_{1,\ldots,s-1;s})$ and $V_{Abel}(L^{WG}_{1,\ldots,s-1;s})$ are asymptotically equivalent. First we note that

$$\sum_{i,j=1}^{n} [\Phi_{WG}(i,j)]^2 = \sum_{\substack{i,j=1\\ i<j}}^{n} \left([\Phi_{WG}(i,j)]^2 + [\Phi_{WG}(j,i)]^2\right) = 2 \sum_{\substack{i,j=1\\ i<j}}^{n} \mathbf{1}_{\left\{\substack{I_i, I_j \text{ present} \\ \text{no superposition}}\right\}}$$

and

$$\sum_{i=1}^{n}\left[\sum_{j=1}^{n}\Phi_{WG}(i,j)\right]^2 = \sum_{i,j_1,j_2=1}^{n}\Phi_{WG}(i,j_1)\ \Phi_{WG}(i,j_2)$$

$$= \sum_{\substack{i,j_1,j_2=1\\ i<j_1<j_2}}^{n}\Big[\ \Phi_{WG}(i,j_1)\Phi_{WG}(i,j_2)+\Phi_{WG}(i,j_2)\Phi_{WG}(i,j_1)$$

$$+\Phi_{WG}(j_1,i)\Phi_{WG}(j_1,j_2)+\Phi_{WG}(j_1,j_2)\Phi_{WG}(j_1,i)$$

$$+\Phi_{WG}(j_2,i)\Phi_{WG}(j_2,j_1)+\Phi_{WG}(j_2,j_1)\Phi_{WG}(j_2,i)\ \Big]$$

$$+\sum_{\substack{i,j_1,j_2=1\\ j_1=j_2}}^{n}\Phi_{WG}(i,j_1)\ \Phi_{WG}(i,j_2)$$

$$=2\sum_{\substack{i,j_1,j_2=1\\ i<j_1<j_2}}^{n}\mathbf{1}_{\left\{\substack{I_i,I_{j_1},I_{j_2}\text{ present less}\\ \text{than two superpositions}}\right\}}+\sum_{i,j=1}^{n}\left[\Phi_{WG}(i,j)\right]^2$$

$$=2\sum_{\substack{i,j_1,j_2=1\\ i<j_1<j_2}}^{n}\mathbf{1}_{\left\{\substack{I_i,I_{j_1},I_{j_2}\text{ present less}\\ \text{than two superpositions}}\right\}}+\ 2\sum_{\substack{i,j=1\\ i<j}}^{n}\mathbf{1}_{\left\{\substack{I_i,I_j\text{ present}\\ \text{no superposition}}\right\}}$$

Thus,

$$V(L^{WG}_{1,\dots,s-1;s})=\frac{2n_s(n_1+\cdots+n_{s-1})(n_1+\cdots+n_s-2)}{n(n-1)(n-2)}\sum_{\substack{i,j_1,j_2=1\\ i<j_1<j_2}}^{n}\mathbf{1}_{\left\{\substack{I_i,I_{j_1},I_{j_2}\text{ present less}\\ \text{than two superpositions}}\right\}}$$

$$+\frac{2n_s(n_1+\cdots+n_{s-1})(n-n_1-\cdots-n_s)}{n(n-1)}\sum_{\substack{i,j=1\\ i<j}}^{n}\mathbf{1}_{\left\{\substack{I_i,I_j\text{ present}\\ \text{no superposition}}\right\}}$$

and

$$V_{Abel}(L^{WG}_{1,\dots,s-1;s})=\frac{2n_s(n_1+\cdots+n_{s-1})}{(n_1+\cdots+n_s)(n_1+\cdots+n_s-1)}\sum_{\substack{i,j=1\\ i<j}}^{n}\sum_{r_1,r_2=1}^{s}\alpha_i^{(r_1)}\alpha_j^{(r_2)}\mathbf{1}_{\left\{\substack{I_i,I_j\text{ present}\\ \text{no superposition}}\right\}}$$

$$+\frac{2n_s(n_1+\cdots+n_{s-1})}{(n_1+\cdots+n_s)(n_1+\cdots+n_s-1)}\sum_{\substack{i,j_1,j_2=1\\ i<j_1<j_2}}^{n}\sum_{r_1,r_2,r_3=1}^{s}\alpha_i^{(r_1)}\alpha_{j_1}^{(r_2)}\alpha_{j_2}^{(r_3)}\mathbf{1}_{\left\{\substack{I_i,I_{j_1},I_{j_2}\text{ present less}\\ \text{than two superpositions}}\right\}}$$

Now, if the null hypothesis is true and sample sizes for all groups are large enough, it follows that

$$\sum_{\substack{i,j=1\\ i<j}}^{n}\sum_{r_1,r_2=1}^{s}\alpha_i^{(r_1)}\alpha_j^{(r_2)}\mathbf{1}_{\left\{\substack{I_i,I_j\text{ present}\\ \text{no superposition}}\right\}}\approx\frac{\binom{n_1+\cdots+n_s}{2}}{\binom{n}{2}}\sum_{\substack{i,j=1\\ i<j}}^{n}\mathbf{1}_{\left\{\substack{I_i,I_j\text{ present}\\ \text{no superposition}}\right\}}$$

and

$$\sum_{\substack{i,j_1,j_2=1\\ i<j_1<j_2}}^{n}\sum_{r_1,r_2,r_3=1}^{s}\alpha_i^{(r_1)}\alpha_{j_1}^{(r_2)}\alpha_{j_2}^{(r_3)}\mathbf{1}_{\left\{\substack{I_i,I_{j_1},I_{j_2}\text{ present less}\\ \text{than two superpositions}}\right\}}$$

$$\approx\frac{\binom{n_1+\cdots+n_s}{3}}{\binom{n}{3}}\sum_{\substack{i,j_1,j_2=1\\ i<j_1<j_2}}^{n}\mathbf{1}_{\left\{\substack{I_i,I_{j_1},I_{j_2}\text{ present less}\\ \text{than two superpositions}}\right\}}$$

This result completes the proof because the substitution of these expressions in the $V(L^{WG}_{1,\ldots,s-1;s})$ formula gives the $V_{Abel}(L^{WG}_{1,\ldots,s-1;s})$ formula.

□

5.3.4 Difference in means test for trend

In this subsection we consider a modification of JK which considers the Difference in means functional given in equation (4.8) instead of the Weighted Mann–Whitney functional. The Difference in means functional holds that $\Phi_{DiM}(\hat{W}^i,\hat{W}^j)=\Phi_{DiM}(\hat{W}^i,\hat{W})-\Phi_{DiM}(\hat{W}^j,\hat{W})$. This property implies that $\Phi_{DiM}(\hat{W}^i,\hat{W}^j)=DiMc_i-DiMc_j$ and that JK_{DiM} reduces to the permutational linear test statistic L_0 given in equation (4.1):

$$
\begin{aligned}
JK_{DiM} &= \sum_{\substack{r,\, s\, =\, 1 \\ r\, <\, s}}^{k} \sum_{i,j=1}^{n} \alpha_i^{(s)} \, \alpha_j^{(r)} \, \Phi_{DiM}(\hat{W}^i, \hat{W}^j) \\
&= \sum_{r=1}^{k-1} n_r \sum_{i=1}^{n} \left(\sum_{s=r+1}^{k} \alpha_i^{(s)} \right) DiMc_i - \sum_{s=2}^{k} n_s \sum_{j=1}^{n} \left(\sum_{r=1}^{s-1} \alpha_i^{(r)} \right) DiMc_j \\
&= \sum_{i=1}^{n} \left(\sum_{r=1}^{k} \left[\sum_{s=1}^{r-1} n_s - \sum_{s=r+1}^{k} n_s \right] \alpha_i^{(r)} \right) DiMc_i \\
&= \sum_{i=1}^{n} z_i \cdot DiMc_i
\end{aligned}
$$

where

$$
z_i = \sum_{r=1}^{k} \left[\sum_{s=1}^{r-1} n_s - \sum_{s=r+1}^{k} n_s \right] \alpha_i^{(r)}.
$$

A little algebra in equation (4.2) provides the following simplification of the permutational variance,

$$
V(JK_{DiM}) = \tfrac{n^3 - \sum_{r=1}^{k} n_r^3}{3(n-1)} \cdot \sum_{i=1}^{n} (DiMc_i)^2
$$

Note that the covariates depend on the sample size of the groups, a property which may seem undesirable. Moreover, if groups have equal sample size $n_r = \frac{k}{n}$, then the trend determined by the covariates

$$
\mathbf{z}_i = \sum_{r=1}^{k} \left[\sum_{s=1}^{r-1} n_s - \sum_{s=r+1}^{k} n_s \right] \alpha_i^{(r)} = \sum_{r=1}^{k} \tfrac{(2r-1-k)\, n}{k} \, \alpha_i^{(r)},
$$

is linear. Thus, this particular functional, or any other which holds that $\Phi(\hat{W}^i, \hat{W}^j) = c_i - c_j$, gives a JK test statistic which is equivalent to a L_0 test statistic with linear covariates.

5.4 Computational aspects

The permutational and Monte Carlo methods of this chapter have been implemented with the S–Plus functions given in Appendix E. In this section we describe each of these S-plus functions and illustrate how they work. All these functions assume that the intervals are semi–closed because several of them use the `kaplanMeier()` procedure and the S–plus functions described in Section 4.7. If the intervals are closed, as it is the case in this chapter, we can replace each interval $[l_i, r_i]$ by $(l_i - \epsilon, r_i]$ where ϵ is a small quantity.

Let `lower` and `upper` be two vectors containing the left and right endpoints of an interval data sample (when a right endpoint is $+\infty$, we write down `1e+029`). First, it is necessary to estimate the survival function from the pooled sample using Turnbull's method,

```
svf <- kaplanMeier(censor(lower,upper,censor.codes)~1)
```

and compute the estimated lifetime distribution function at every left and right endpoint of the interval data sample,

```
cdf <- cdfdata(lower,upper,surv.est)
```

It is also necessary to create a vector of covariates, `covar`, which assigns the value 1 for individuals in the first group, the value 2 for individuals in the second group and likewise until the k^{th} group.

Now, each S–plus function in Appendix E works as follows:

A. Function `JK(·,·,·,·)` computes the standardized value of the test JK. In the Weighted Mann–Whitney functional, we have implemented the weighting scheme described in equation (4.13). Thus, as it is proved in Appendix F, the Weighted Mann–Whitney functional simplifies as follows:

$$\Phi_{WMW}(\hat{W}^i, \hat{W}^j) = \frac{\hat{W}(l_i^-)\left[\gamma(\hat{W}(l_i^- \vee r_j)) - \gamma(\hat{W}(l_i^- \vee l_j^-))\right]}{\{\hat{W}(r_i) - \hat{W}(l_i^-)\}\{\hat{W}(r_j) - \hat{W}(l_j^-)\}}$$
$$- \frac{\hat{W}(l_j^-)\left[\gamma(\hat{W}(l_j^- \vee r_i)) - \gamma(\hat{W}(l_j^- \vee l_i^-))\right]}{\{\hat{W}(r_i) - \hat{W}(l_i^-)\}\{\hat{W}(r_j) - \hat{W}(l_j^-)\}}$$
$$+ \frac{\hat{W}(r_j)\left[\gamma(\hat{W}(r_j \vee r_i)) - \gamma(\hat{W}(r_j \vee l_i^-))\right]}{\{\hat{W}(r_i) - \hat{W}(l_i^-)\}\{\hat{W}(r_j) - \hat{W}(l_j^-)\}}$$
$$- \frac{\hat{W}(r_i)\left[\gamma(\hat{W}(r_i \vee r_j)) - \gamma(\hat{W}(r_i \vee l_j^-))\right]}{\{\hat{W}(r_i) - \hat{W}(l_i^-)\}\{\hat{W}(r_j) - \hat{W}(l_j^-)\}}$$

We use $\gamma(t) = \frac{(1-t)}{t}\frac{(1-t)^\rho - 1}{\rho}$, the function related to the Harrington and Fleming scores given in equation (4.12). Function `JK(·,·,·,·)` works for any $\rho \geq 0$. For instance when $\rho = 0$ and $\rho = 1$, expressions

```
JK(cdf,covar,0)
```

and

```
JK(cdf,covar,1)
```

compute respectively the Log–rank and Wilcoxon–Peto extension of the Jonckheere's test.

For the sake of time efficiency, we have used matrix algebra in the implementation of this function. Hence, there is a fourth input parameter which corresponds to the number of files of the matrices used in the computations. The default value of this parameter is equal to 25. When the sample size is large it could appear memory allocation problems, then it is recommended to reduce the value of this parameter.

B. Function `Ltrend(·,·)` computes the standardized value of the test L_0 with a trend covariate. The first input parameter corresponds to a vector of scores values, for instance, the Log–rank scores given by the function `HFsc(·,·)` introduced in Section 4.7:

```
scores <- HFsc(cdf,0)
```

The second input parameter corresponds to the vector of covariates, `trendcovar`, which specify the trend sought by the test stastistic L_0. Then,

```
Ltrend(scores,trendcovar)
```

computes the desired result.

C. Functions `J1boot(·,·,·,·)`, `J2boot(·,·,·,·)` and `J3boot(·,·,·,·)` compute the Jonckheere–type statistics J_1, J_2 and J_3 given in equations (5.8), (5.9) and (5.10) respectively. These statistics have been implemented as a sum of two–sample test statistics in the Harrington and Fleming class. The fourth input parameter corresponds to any $\rho \geq 0$. For implementing a Monte Carlo approach, these functions should be jointly used with the `bootstrap()` S–plus procedure. For instance, a Monte Carlo resampling with the statistic J_1 and $\rho = 0.5$ is performed by,

```
bootstrap(covar,J1boot(lower,upper,covar,0.5),sampler=samp.permute)
```

We note that these functions use internally the function `HFsc(·,·)` introduced in Section 4.7.

D. Functions `JA(·,·,·,·)` and `JA2(·,·,·)` compute the standardized values of the test JA considering respectively the permutational variance derived from rank correlation theory, see equation (5.16), and the variance derived in Abel (1986), see equation (5.15). They apply as follows:

```
JA(lower,upper,covar)
```

and

```
JA2(lower,upper,covar)
```

Function JA(·,·,·,·) has a fourth input parameter which corresponds to the number of files of the matrices used in the computations. The default value of this parameter is equal to 25. We also note that function JA2(·,·,·) uses internally the function WGsc(·,·) introduced in Section 4.7.

E. Function JKDiM(·,·) computes the standardized value of the Difference in means test for trend. The first input parameter corresponds to the Difference in means scores given by the function DiMsc(·,·,·) introduced in Section 4.7:

```
scores <- DiMsc(lower,upper,svf)
```

Then, this function applies as follows,

```
JKDiM(scores,covar)
```

5.5 Simulation study

This section tries to elucidate the behavior of the trend tests introduced in this chapter. A first purpose of the simulation study is to compare the powers of these tests under various linear and nonlinear trend alternatives. We are interested in the power differences between a test like L_0 which requires that the covariates should be specified according to the trend sought in the alternative hypothesis and tests like J_1, J_2, J_3, JK and JA which do not have this requirement. A second purpose of the simulation study is to confirm that the nominal significance level is roughly reached when the null hypothesis is true. We are specially interested in the Monte Carlo approach introduced for J_1, J_2 and J_3, and in the two options for JA established in terms of the permutational variance.

5.5.1 Data generation

In the generation of the simulated data for the lifetime variable T we have considered a discretization of a variable T^* under an accelerated lifetime model $\log(T^*) = \theta + \lambda Z$. We restrict ourselves to those models under which the Harrington and Fleming class of tests is efficient. Hence, the distribution function for T^* is given by

$$W_{\theta,\lambda,\rho}(t) = W_\rho(\frac{\log(t) - \theta}{\lambda}) = 1 - \big(1 + \rho\, t^{1/\lambda} \exp(-\theta/\lambda)\big)^{-1/\rho},$$

where W_ρ is the error term distribution given in equation (4.17). The discrete lifetime variable T is established to take values $1, 2, \ldots, 10$ and it is defined as $T = [T^*] + 1$ for $T^* \leq 10$ and $T = 10$ for $T^* > 10$.

The censoring mechanism of T mimics a longitudinal study where there is a periodical follow–up with scheduled visits but patients might miss some of the appointments. Specifically, there are assumed potential monitoring times $t_j = j$ for $j = 1, \ldots, 10$. The patients would assist to each of these scheduled visits with probability p. Then, for an individual i, the observed censoring interval $[L_i, R_i]$ is constructed by defining R_i as the first visit where the event of interest is observed and L_i as the previous visit of the patient. That is, $L_i = \max\{t_j | t_j \leq T_i, \delta_j^i = 1\}$ and $R_i = \min\{t_j | t_j > T_i, \delta_j^i = 1\}$, where δ_j^i is the indicator of whether the visit at time t_j occurs ($\delta_j^i = 1$) or is missed ($\delta_j^i = 0$).

The S–plus function `gendata(·,·,·,·,·)` given in Appendix G implements the generation data process. For instance,

```
gendata(100,0.9,0.6,1,0.5)
```

generates a random sample of $n = 100$ censoring intervals where the accelerated lifetime model for T^* has parameters $(\theta, \lambda, \rho) = (0.9, 0.6, 1)$ and the probability of assistance to each scheduled visit in the periodical follow–up is $p = 0.5$.

5.5.2 Simulation scenarios

In the simulation study we have considered a location shift accelerated lifetime model for the variable T^*. That is, the error distribution function W_ρ and the scale parameter λ are assumed to be equal for all the lifetimes, but the location parameter θ may differ among groups. We define scenarios where the distribution $W_{\theta,\lambda,\rho}$ has parameters $(\lambda,\rho) = (1,0), (0.75,0.5), (0.6,1), (0.5,1.5)$. For a situation of differences between groups, we study models with equal spacings in the location parameters (linear trend) and models with different spacings (nonlinear trend). Under each of the four settings fixed for (λ,ρ), we regard an scenario with equal spacings $(\theta_2-\theta_1=\cdots=\theta_k-\theta_{k-1})$ and two scenarios where all the groups have the same location parameters except one $(\theta_1<\theta_2=\cdots=\theta_{k-1}=\theta_k$ and $\theta_1=\theta_2=\cdots=\theta_{k-1}<\theta_k)$. In Table 5.1 and Table 5.2 we give the parameter settings of the scenarios when $k=3$ and $k=4$.

Table 5.1: *Parameters for each trend scenario when* $k=3$

(λ,ρ)	$(1,0)$			$(0.75,0.5)$			$(0.6,1)$			$(0.5,1.5)$		
θ_1	1.6	1.6	1.6	1.25	1.25	1.25	0.9	0.9	0.9	0.5	0.5	0.5
θ_2	1.85	2.1	1.6	1.5	1.75	1.25	1.15	1.4	0.9	0.75	1	0.5
θ_3	2.1	2.1	2.1	1.75	1.75	1.75	1.4	1.4	1.4	1	1	1

Table 5.2: *Parameters for each trend scenario when* $k=4$

(λ,ρ)	$(1,0)$			$(0.75,0.5)$			$(0.6,1)$			$(0.5,1.5)$		
θ_1	1.65	1.65	1.65	1.3	1.3	1.3	0.95	0.95	0.95	0.55	0.55	0.5
θ_2	1.8	2.1	1.65	1.45	1.75	1.3	1.1	1.4	0.95	0.6	0.55	0.5
θ_3	1.95	2.1	1.65	1.6	1.75	1.3	1.25	1.4	0.95	0.75	0.55	1
θ_4	2.1	2.1	2.1	1.75	1.75	1.75	1.4	1.4	1.4	1	1	1

The location parameters have been chosen so that the set of means of T^* roughly coincides in those scenarios with the same trend configuration (columns $1,4,7,10$, columns

$2, 5, 8, 11$ and columns $3, 6, 9, 12$ in Table 5.1 and Table 5.2). For all the parameter settings, Appendix G give the mean and median of T^*, the mean and median of T and the plots of the survival probabilities of T.

Under the null hypothesis, the parameters (θ, λ, ρ) are identical in each group. In this situation, we have also considered some scenarios with some of the parameters (θ, λ, ρ) given above.

The sample simulation of each scenario has been performed with the S–plus function `simu()` given in Appendix G. All the simulation results are based on $D = 500$ replications of the data samples and a probability of assistance to each scheduled visit of $p = 0.5$. For each of the following tests we have computed the percentage of rejections of the null hypothesis under a nominal significance level $\alpha = 0.05$.

- L^0_{trend}, $L^{0.5}_{trend}$, L^1_{trend} and $L^{1.5}_{trend}$: they are specific cases in the Harrington and Fleming class of k–sample tests. We consider the parameters $\rho = 0, 0.5, 1, 1.5$ and a trend covariate given by $z_i = \sum_{r=1}^{k} \alpha_i^{(r)}\, r$.

- JK^0, $JK^{0.5}$, JK^1 and $JK^{1.5}$: they are specific cases of the JK test. We consider the Harrington and Fleming functional and parameters $\rho = 0, 0.5, 1, 1.5$.

- J_1^ρ, J_2^ρ and J_3^ρ: they are specific cases of the J_1, J_2 and J_3 tests. We consider the Harrington and Fleming functional and the same parameter ρ as in the distribution of T^*. We use $M = 1000$ resamples in the Monte Carlo approach.

- JA: it is the JA test with the permutational variance derived from rank correlation theory.

- $JA2$: it is the JA test with the variance derived in Abel (1986).

- JK_{DiM}: it is the Difference in means test for trend.

5.5.3 Simulation results

In this subsection we report some of the simulation results given in Appendix H. First we give a description of each studied issue. Then, we give a discussion about open questions which need of further research.

1) In most of the simulation scenarios given in Table 5.1 and Table 5.2, the tests L^{ρ}_{trend} and JK^{ρ} show higher estimated power when ρ coincides with the analogous parameter in the distribution of T^*. This is an important feature which confirms the Harrington and Fleming class of tests given in Chapter 4 as a generalization of the class introduced in Harrington and Fleming (1982).

2) Under linear trend alternatives, we suggested in Subsection 5.3.1 that L^{ρ}_{trend} would have higher power than JK^{ρ}. However, the estimated power of JK^{ρ}, L^{ρ}_{trend} and J^{ρ}_1 is similar and it is generally as high or higher than those of the other tests. Table 5.3 shows these power performance when $k = 3$.

Table 5.3: *Power sizes under linear trend alternatives when $k = 3$*

(n_1, n_2, n_3)	$(\theta_1, \theta_2, \theta_3, \lambda, \rho)$					
	$(1.6, 1.85, 2.1, 1, 0)$			$(1.25, 1.5, 1.75, 0.75, 0.5)$		
	JK^0	L^0_{trend}	J^0_1	$JK^{0.5}$	$L^{0.5}_{trend}$	$J^{0.5}_1$
$(50, 50, 50)$	0.656	0.656	0.66	0.682	0.682	0.686
$(100, 100, 100)$	0.882	0.886	0.876	0.928	0.926	0.926
$(50, 100, 150)$	0.812	0.816	0.818	0.896	0.902	0.892
$(150, 100, 50)$	0.86	0.884	0.888	0.882	0.87	0.866
	$(0.9, 1.15, 1.4, 0.6, 1)$			$(0.5, 0.75, 1, 0.5, 1.5)$		
	JK^1	L^1_{trend}	J^1_1	$JK^{1.5}$	$L^{1.5}_{trend}$	$J^{1.5}_1$
$(50, 50, 50)$	0.736	0.742	0.73	0.68	0.676	0.676
$(100, 100, 100)$	0.94	0.94	0.946	0.926	0.926	0.922
$(50, 100, 150)$	0.872	0.872	0.868	0.884	0.886	0.884
$(150, 100, 50)$	0.888	0.888	0.884	0.872	0.884	0.874

3) Under nonlinear trend alternatives of the type $\theta_1 < \theta_2 = \cdots = \theta_{k-1} = \theta_k$, the estimated power of J_2^ρ is higher than those of the other tests. In Subsection 5.2.1 we have given the intuition of this result. Moreover, in Subsection 5.3.1 we suggested that JK^ρ would higher power than L_{trend}^ρ under nonlinear trends. The present trends are clearly nonlinear and, however, the estimated power of JK^ρ, L_{trend}^ρ and J_1^ρ is roughly similar. Table 5.4 shows these results when $k = 3$.

Table 5.4: *Power sizes under nonlinear trend alternatives when $k = 3$ and $\theta_1 < \theta_2 = \theta_3$*

(n_1, n_2, n_3)	$(\theta_1, \theta_2, \theta_3, \lambda, \rho)$							
	(1.6, 2.1, 2.1, 1, 0)				(1.25, 1.75, 1.75, 0.75, 0.5)			
	JK^0	L_{trend}^0	J_1^0	J_2^0	$JK^{0.5}$	$L_{trend}^{0.5}$	$J_1^{0.5}$	$J_2^{0.5}$
(50, 50, 50)	0.636	0.636	0.664	0.71	0.718	0.718	0.728	0.786
(100, 100, 100)	0.886	0.886	0.90	0.936	0.924	0.924	0.928	0.956
	(0.9, 1.4, 1.4, 0.6, 1)				(0.5, 1, 1, 0.5, 1.5)			
	JK^1	L_{trend}^1	J_1^1	J_2^1	$JK^{1.5}$	$L_{trend}^{1.5}$	$J_1^{1.5}$	$J_2^{1.5}$
(50, 50, 50)	0.746	0.74	0.746	0.798	0.732	0.736	0.724	0.782
(100, 100, 100)	0.928	0.93	0.926	0.966	0.912	0.914	0.912	0.952

4) Under nonlinear trend alternatives of the type $\theta_1 = \theta_2 = \cdots = \theta_{k-1} < \theta_k$, the estimated power of J_3^ρ is higher than those of the other tests. As we have noted in the last item, the intuition of this result was given in Subsection 5.2.1. Again, the present trends are clearly nonlinear and, contrarily to the expectancies given in 5.3.1, the estimated power of JK^ρ, L_{trend}^ρ and J_1^ρ is roughly similar. Table 5.5 shows these results when $k = 3$.

5) The nominal significance level $\alpha = 0.05$ is roughly reached in the scenarios where the null hypothesis is true. There are some exceptions, but these cases can be due to the random variability of the replicated samples. Since $D = 500$ is not a large number of replications, the percentage of rejections can produce several strange

Table 5.5: *Power sizes under nonlinear trend alternatives when $k = 3$ and $\theta_1 = \theta_2 < \theta_3$*

(n_1, n_2, n_3)	$(\theta_1, \theta_2, \theta_3, \lambda, \rho)$							
	$(1.6, 1.6, 2.1, 1, 0)$				$(1.25, 1.25, 1.75, 0.75, 0.5)$			
	JK^0	L^0_{trend}	J^0_1	J^0_3	$JK^{0.5}$	$L^{0.5}_{trend}$	$J^{0.5}_1$	$J^{0.5}_3$
$(50, 50, 50)$	0.672	0.67	0.646	0.734	0.724	0.726	0.72	0.786
$(100, 100, 100)$	0.892	0.892	0.882	0.93	0.938	0.938	0.934	0.96
	$(0.9, 0.9, 1.4, 0.6, 1)$				$(0.5, 0.5, 1, 0.5, 1.5)$			
	JK^1	L^1_{trend}	J^1_1	J^1_3	$JK^{1.5}$	$L^{1.5}_{trend}$	$J^{1.5}_1$	$J^{1.5}_3$
$(50, 50, 50)$	0.708	0.712	0.706	0.754	0.722	0.72	0.728	0.764
$(100, 100, 100)$	0.956	0.96	0.954	0.974	0.914	0.912	0.918	0.95

results. For instance, when $k = 3$, $n_1 = n_2 = n_3 = 100$ and $(\theta, \lambda, \rho) = (1.4, 0.6, 1)$, most of the tests have a percentage of rejection lower than 0.03. For this particular case, we repeated the simulation process with a different seed and the results became better.

6) The tests JA and $JA2$ do not show differences in the percentage of rejections neither under the null hypothesis nor under a trend alternative. Furthermore, they show a similar power performance to JK^1.

7) As noted by Jones (2001) in the right–censored data framework, tests like J_1, J_2 and J_3 may seek for a trend which depends on group sizes. In Subsection 5.3.2 we also suggested that the test statistic JK would depend on group sizes. This do not immediately mean that the trend sought by JK will depend on group sizes. In the simulation study, this possible dependence was checked by introducing different group sample sizes in the linear trend scenarios. Once more, the results do not show the power differences we expected between L^ρ_{trend}, JK^ρ and J^ρ_1. See, for instance, Table 5.3.

8) The power of JK_{DiM} is generally lower than the power of L^ρ_{trend}, JK^ρ and J^ρ_1 when

ρ coincides with the analogous parameter in the distribution of T^*. However, it is high or higher than the power of L^{ρ}_{trend} and JK^{ρ} when the value of ρ is far away from the analogous parameter in the distribution of T^*.

In general, the tests perform as we expected before running the simulations. However, it has been somewhat surprising to not observe clear differences in the behavior of L^{ρ}_{trend}, JK^{ρ} and J^{ρ}_1. Our main goal in this simulation study has not been attained. We pretended to find those situations where tests like JK and J_1, which do not precise of covariate specification, are preferable to a test like L_0, which necessarily does. We think that this goal could be attained if we consider a continuous model for T or, at least, a thinner discretization of the variable T^*. This would increase the variability of the data and, consequently, it would clarify the differences between the tests. Since the Monte Carlo approach for J_1, J_2 and J_3 is computationally intensive, we think that at least JK and L_0 should be compared. We will carry out this new simulation study in future work.

5.6 Illustration

The methodology of this chapter is illustrated with data from the AIDS Clinical Trial described in Section 4.3. The variable of interest is the time T, measured in months from randomization, until the CD4 count first reaches 400 cells per cubic millimeter and is interval-censored. The estimated survival curves suggest an increasing trend according to group G_1 (the deferred–therapy group), group G_2 (immediate–therapy group with 500–mg) and group G_3 (immediate–therapy group with 1500–mg).

The analysis of this data set shows the results given in Table 5.6 and Table 5.7. In Table 5.6 we give the standardized test statistics and p–values of the tests L_0, JK, JA, $JA2$ and JK_{DiM}. All these tests clearly show a significant increasing trend of the survival probabilities in each group. Since the p–values of these tests are low, in the Monte Carlo approach for the tests J_1, J_2 and J_3 we have used $M = 100000$ resamples.

The results are given in Table 5.7 and they also clearly show a significant increasing trend.

Table 5.6: *Standardized test statistics and p–values with the normal approximation of the permutational distribution*

Test	L^0_{trend}	$L^{0.5}_{trend}$	L^1_{trend}	$L^{1.5}_{trend}$	JK^0	$JK^{0.5}$
statistic:	4.150754	4.123759	4.028795	3.896602	4.153457	4.125687
p–value:	0.00002	0.00002	0.00003	0.00005	0.00002	0.00002
	JK^1	$JK^{1.5}$	JA	$JA2$	JK_{DiM}	
	4.029249	3.895257	4.011888	3.992121	4.156854	
	0.00003	0.00005	0.00003	0.00003	0.00002	

Table 5.7: *Test statistics and p–values in the Monte Carlo approach*

Test	J^0_1	$J^{0.5}_1$	J^1_1	$J^{1.5}_1$	J^0_2	$J^{0.5}_2$
statistic:	97.41890	79.44099	66.15175	56.15465	75.47007	61.33287
p–value:	0.00001	0.00002	0.00002	0.00005	0.00001	0.00001
	J^1_2	$J^{1.5}_2$	J^0_3	$J^{0.5}_3$	J^1_3	$J^{1.5}_3$
	50.90874	43.07878	70.16334	53.44330	48.06883	41.01986
	0.00001	0.00004	0.00001	0.00003	0.00004	0.00007

Chapter 6

Conclusions and future research

This thesis is divided into two parts which aimed to investigate two aspects of survival analysis with interval–censored data. The first part of this thesis is devoted to give a constant–sum condition for interval censoring models and to explore the consequences of this assumption. The second part deals with the analysis and hypothesis testing of k samples. In the following sections we summarize the results of the dissertation while suggest other possible topics or new approaches of investigation.

6.1 Constant–sum models

The simplified likelihood has been widely used with interval censored data. Nevertheless, in some cases the mechanisms leading to censoring could be related to the event time process and in this case the simplified likelihood would not be appropriate. Thus, a natural problem arises elucidating whether the simplified likelihood is the correct one to get valid inferences for a given data set. Chapter 2 specifies the conditions which ensure that the simplified likelihood is a proper basis for inferences. We describe the constant–sum condition and give the relationship to the noninformative conditions. The attractiveness of the theoretical development presented here stems from the fact that the framework is general enough to cover most of the censoring models in the literature. Henceforth, further application of the ideas pointed out in this chapter, and in particular Proposition 2.7 and Proposition 2.8, might be useful in characterizing the noninformative and the constant-sum conditions for different censoring models.

When the censoring observations occur through a longitudinal inspection process, an intuitive interpretation of the constant–sum condition is given in Lawless (2004). However, further research is needed to give to this condition more intuition and practical use. For instance, the equivalent condition given in Kalbfleisch and MacKay (1979) for right–censored data is more intuitive. We plan to continue our research along this aspect.

In Chapter 3, we have investigated which role does the constant–sum property play in the identifiability of the lifetime distribution. We have shown that the lifetime distribution is not identifiable outside the class of constant–sum models. We have also shown that the lifetime probabilities assigned to the observable intervals are identifiable inside the class of constant–sum models. In case that the observables have finite support, we have completely elucidated the issue of the identifiability of W. In a general censoring setting, it is still an open question to give sufficient conditions which ensure identification of the entire lifetime distribution.

There are specific situations where it is possible to ensure complete identifiability, for instance, when uncensored data are allowed for the whole support of the lifetime variable, that is, when $dF_{L,R}(t,t) > 0$ for any $t \in \mathcal{D}_W$. This identifiability assumption is rather mild and it is typically satisfied in right censored data and doubly censored data applications. For instance, Chang and Yang (1987) use this assumption to prove the consistency of the NPMLE with doubly censored data.

Further research is needed for the inspection model discussed in Subsection 3.2.2 when the support of the inspection times is not finite. In this setting, the characterization given in Lawless (2004) does not necessarily apply. The interval censoring model studied in Wang *et al.* (1994), which includes the well known case k interval censoring model, and the mixture of case k interval censoring models presented in Schick and Yu (2000) are examples of such inspection models. Both examples assume that each individual is inspected a countable number of times and that this inspection process is independent of the lifetime variable. As a consequence, the constant–sum condition holds and we have a particular structure for the observable intervals which, jointly with Theorem 3.4, derives in the identifiability of $1 - W(l)$ and $W(r)$ for any observable $(l, r) \in \mathcal{D}_{L,R}$. In this case, the assumption that the support of L or R covers $\mathcal{D}_W = (0, +\infty)$ would ensure identifiability of W. At this point, it might be interesting to investigate an identifiability assumption when each individual is inspected an uncountable number of times or the constant–sum condition holds but the inspection process depends on the lifetime variable.

Another point of interest is to develop a formal test to examine, from the observed data, whether the constant–sum condition does hold or not. The literature on coarsening data suggest the existence of a noninformative model, and consequently a constant–sum model, for any given observables structure, $F_{L,R}$. However, we think that there are situations where this noninformative model cannot hold the assumed lifetime support $\mathcal{D}_W$. A disagreement between $\mathcal{D}_W$ and the lifetime support of this noninformative model could open a way to detect that the constant–sum condition does not hold. For instance, this problem is considered by Gill *et al.* (1997) and Betensky (2000) in the right censoring and current status data settings, respectively. These authors establish conditions under which it is possible to state that the constant–sum property does not hold.

6.2 k–sample problem

The nonparametric tests introduced in Chapter 4 and Chapter 5 give new methods to compare survival functions among groups when data are interval–censored. Chapter 4 gives a review of k–sample tests provided in the literature. We specially focus on permutational tests. We develop links between different forms of the tests, propose new extensions and implement S–Plus functions. Chapter 5 considers tests for ordered alternatives. We propose several new generalizations of the Jonckheere's test based on the permutational tests given in Chapter 4. We provide software and simulate data to study the performance of each Joncheere–type test.

The permutational approach in Chapter 4 and Chapter 5 requires that the censoring mechanisms are identical across the groups. The testability of the equality of censoring mechanisms is an aspect which needs further research. Another point of interest is how much the non–equality of the censoring mechanism could affect the tests. For instance, in situations where an inspection process defines the censoring observations, it seems possible that having different frequencies of inspection might not be relevant for the inference.

Our plans for the future include the implementation of the likelihood score test statistic given in Subsection 4.5.2 and a simulation study to learn about the power of the different tests proposed in Chapter 4. We plan to compare the permutational approach and the likelihood approach for discrete data. For continuous data, it is of interest to compare the permutational approach and the approach introduced in Sun *et al.* (2005).

Finally, and since the simulation study in Chapter 5 has not shown substantial differences between the powers of the test statistics L_0, JK and J_1, we would like to explore the cause of these results. The similar powers might be due to the discrete nature of the the simulated data or an asymptotic equivalence between the tests. It will be interesting to explore the performance of the tests under continuous data or under small sample sizes.

References

U. Abel (1986). A nonparametric test against ordered alternatives for data defined by intervals. *Statistica Neerlandica*, 40, 87–91.

R. E. Barlow, D. J. Bartholomew, D. J. Bremner and H. D. Brunk (1972). *Order restricted statistical inference*. John Wiley & Sons, New York.

R. A. Betensky (2000). On nonidentifiability and noninformative censoring for current status data. *Biometrika*, 87, 218–221.

H. Büning and W. Kössler (1999). The asymptotic power of Jonckheere–type tests for ordered alternatives. *Australian and New Zealand Journal of Statistics*, 41, 67–77.

M. N. Chang and G. L. Yang (1987). Strong consistency of a nonparametric estimator of the survival function with doubly censored data. *The Annals of Statistics*, 16, 1536–1547.

Y. Chi (2002). Ordered tests for right–censored survival data. *Australian and New Zealand Journal of Statistics*, 44, 367–380.

N. Ebrahimi, D. Molefe and Z. Ying (2003). Identifiability and censored data. *Biometrika*, 90, 724–727.

B. Efron (1967). The two sample problem with censored data. *Proceedings of Fifth Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, 831–851.

H–B. Fang, J. Sun and M–LT. Lee (2002). Nonparametric survival comparisons for interval–censored continuous data. *Statistica Sinica*, 12, 1073–1083.

M. P. Fay (1996). Rank invariant tests for interval-censored data under the grouped continuous model. *Biometrics*, 52, 811–822.

M. P. Fay (1999). Comparing several score tests for interval-censored data. *Statistics in Medicine*, 18, 273–285.

M. P. Fay and J. H. Shih (1998). Permutation tests using estimated distribution functions. *Journal of the American Statistical Association*, 93, 387–396.

D. M. Finkelstein (1986). A proportional hazards models for interval-censored failure time data. *Biometrics*, 42, 845–854.

T. R. Fleming and D. P. Harrington (1991). *Counting processes and survival analysis.* John Wiley & Sons, New York.

E. A. Gehan (1965). A generalized Wilcoxon test for comparing arbitrarily singly censored samples. *Biometrika*, 52, 203–223.

R. Gentleman and C. J. Geyer (1994). Maximum–likelihood for interval–censored data: Consistency and computation. *Biometrika*, 81, 618–623.

R. D. Gill, M. J. van der Laan and J. M. Robins (1997). Coarsening at random: characterizations, conjectures, counter–examples. *Proceedings First Seattle Symposium on Biostatistics: Survival Analysis*, Springer–Verlag, 255–294.

G. Gómez, M. L. Calle and R. Oller (2004). Frequentist and bayesian approaches for interval–censored data. *Statistical Papers*, 45, 139–173.

P. Groeneboom (1996). Lectures on inverse problems. In *Lecture notes in Mathematics*, 1648. Springer–Verlag, Berlin.

P. Groeneboom and J. A. Wellner (1992). *Information bounds and nonparametric maximum likelihood estimation.* Birkhäuser Verlag, Basel.

D. P. Harrington and T. R. Fleming (1982). A class of rank test procedures for censored survival data. *Biometrika*, 69, 553–566.

G. Heimann and G. Neuhauss (1998). Permutational distribution of the log-rank statistic under random censorship with applications to carcinogenicity assays. *Biometrics*, 54, 168–184.

D. F. Heitjan (1993). Ignorability and coarse data: Some biomedical examples. *Biometrics*, 49, 1099–1109.

D. F. Heitjan and D. B. Rubin (1991). Ignorability and coarse data. *The Annals of Statistics*, 19, 2244–2253.

M. Hollander and D. A. Wolfe (1999). *Nonparametric statistical methods.* John Wiley & Sons, New York.

A. R. Jonckheere (1954). A distribution–free k–sample test against ordered alternatives. *Biometrika*, 41, 132–145.

M. P. Jones (2001). Unmasking the trend sought by Jonckheere–type tests for right–censored data. *Scandinavian Journal of Statistics*, 28, 527–535.

M. P. Jones and J. Crowley (1989). A general class of nonparametric tests for survival analysis. *Biometrics*, 45, 157–170.

M. P. Jones and J. Crowley (1989). Asymptotic properties of a general class of nonparametric tests for survival analysis. *The Annals of Statistics*, 18, 1203–1220.

J. D. Kalbfleisch and R. J. MacKay (1979). On constant–sum models for censored survival data. *Biometrika*, 66, 87–90.

J. D. Kalbfleisch and R. L. Prentice (2002). *The statistical analysis of failure time data.* John Wiley & Sons, New York.

M. G. Kendall and J. D. Gibbons (1990). *Rank correlation methods.* Oxford University Press, New York.

J. F. Lawless (2003). *Statistical models and methods for lifetime data.* John Wiley & Sons, New York.

J. F. Lawless (2004). A note on interval–censored lifetime data and the constant–sum condition of Oller, Gómez and Calle (2004). *The Canadian Journal of Statistics*, 32, 327–331.

H–J. Lim and J. Sun (2003). Nonparametric tests for interval–censored failure time data. *Biometrical Journal*, 45, 263–276.

P. Y. Liu, S. Green, M. Wolf and J. Crowley (1993). Testing against ordered alternatives for censored survival data. *Journal of the American Statistical Association*, 88, 153–160.

P. Y. Liu and W. Y. Tsai (1999). A modified logrank test for censored survival data under order restrictions. *Statistics and Probability Letters*, 41, 57–63.

P. Y. Liu, W. Y. Tsai and M. Wolf (1998). Design and analysis for survival data under order restrictions with a modified logrank test. *Statistics in Medicine*, 17, 1469–1479.

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

J. Mau (1988). A generalization of a nonparametric test for stochastically ordered distributions to censored survival data. *Journal of the Royal Statistical Society, Series B*, 50, 403–412.

M. Neuhäuser, P–Y. Liu and L. Hothorn (1998). Nonparametric tests for trend: Jonckheere's test, a modification and a maximum test. *Biometrical Journal*, 40, 899–909.

M. P. Ng (2002). A modification of Peto's nonparametric estimation of survival curves for interval–censored data. *Biometrics*, 58, 439–442.

R. Oller, G. Gómez and M. L. Calle (2004). Interval censoring: model characterizations for the validity of the simplified likelihood. *The Canadian Journal of Statistics*, 32, 315–326.

W. Pan (2000). A two–sample test with interval censored data via multiple imputation. *Statistics in Medicine*, 19, 1–11.

M. S. Pepe and T. R. Fleming (1989). Weighted Kaplan-Meier statistics: a class of distance tests for censored survival data. *Biometrics*, 45, 497–507.

R. Peto (1973). Experimental survival curves for interval-censored data. *Applied Statistics*, 22, 86–91.

R. Peto and J. Peto (1972). Asymptotically efficient rank invariant test procedures. *Journal of the Royal Statistical Society, Series A*, 135, 185–207.

G. R. Petroni and A. Wolfe (1994). A two sample test for stochastic ordering with interval-censored data. *Biometrics* 50, 77–87.

M. L. Puri (1965). Some distribution–free k–sample rank tests of homogeneity against ordered alternatives. *Communications on Pure and Applied Mathematics* 18, 51–63.

T. Robertson, F. T. Wright and R. L. Dykstra (1988). *Order restricted statistical inference.* John Wiley & Sons, New York.

S. G. Self and E. A. Grossman (1986). Linear rank tests for interval–censored data with application to PCB levels in adipose tissue of transformer repair workers. *Biometrics*, 42, 521–530.

M. Schemper (1982). A nonparametric test of association for data defined by intervals. *EDV in Medizin und Biologie*, 13, 44–46.

M. Schemper (1983). A nonparametric k–sample test for data defined by intervals. *Statistica Neerlandica*, 37, 69–71.

M. Schemper (1984). A generalized Friedman test for data defined by intervals. *Biometrical Journal*, 26, 305–308.

A. Schick and Q. Yu (2000). Consistency of the GMLE with mixed case interval–censored data. *Scandinavian Journal of Statistics*, 27, 45–55.

J. Sun (2005). A non–parametric test for interval–censored failure time data with application to AIDS studies. *Statistics in Medicine*, 15, 1387–1395.

J. Sun, Q. Zhao and X. Zhao (2005). Generalized log–rank tests for interval–censored failure time data. *Scandinavian Journal of Statistics*, 32, 49–57.

T. Tapey and B. Flury (1996). Self–consistency: a fundamental concept in statistics. *Statistical Science*, 11, 229–243.

T. J. Terpstra (1952). The asymptotic normality and consistency of Kendall's test against trend, when ties are present in one ranking. *Indagationes Mathematicae*, 14, 327–333.

J. T. Terpstra and R. C. Magel (2003). A new nonparametric test for the ordered alternative problem. *Nonparametric Statistics*, 15, 289–301.

P. V. Tryon and T. P. Hettmansperger (1973). A class of non–parametric tests for homogeneity against ordered alternatives. *The Annals of Statistics*, 1, 1061–1070.

A. Tsiatis (1975). A nonidentifiability aspect of the problem of competing risks. *Proceedings of the National Academy of Sciences*, 72, 20–22.

B. W. Turnbull (1976). The empirical distribution function with arbitrarily grouped, censored and truncated data. *Journal of the Royal Statistical Society, Series B*, 38, 290–295.

P. A. Volberding, S. W. Lagakos, J. M. Grimes, *et al.* (1995). A comparison of immediate with deferred zidovudine therapy for asymptomatic HIV–infected adults with CD4 cell counts of 500 or more per cubic millimiter. *New England Journal of Medicine*, 333, 401–451.

Z. Wang, J. C. Gardiner and R. V. Ramamoorthi (1994). Identifiability in interval censorship models. *Statistics and Probability Letters*, 21, 215–221.

J. A. Wellner and Y. Zhan (1997). A hybrid algorithm for computation of the nonparametric maximum likelihood estimator from censored data. *Journal of the American Statistical Association*, 92, 945–959.

J. S. Williams and S. W. Lagakos (1977). Models for censored survival analysis: constant–sum and variable–sum models. *Biometrika*, 64, 215–224.

Q. Yu, L. Li and G. Y. C. Wong (2000). On consistency of the self–consistent estimator of survival functions with interval–censored data. *Scandinavian Journal of Statistics*, 27, 35–44.

Appendix

A A probability measure result

Lemma A.1. *Let μ be a probability measure and f a μ–measurable positive function such that*

$$\int f \; d\mu = 1 \qquad \text{and} \qquad \int \frac{1}{f} \; d\mu = 1$$

then $f = 1$ μ–almost surely.

Proof:

If we sum the two integrals,

$$\int \left(f + \frac{1}{f} \right) \; d\mu = 2$$

Then we can rewrite this result as,

$$\int \left\{ \left(f + \frac{1}{f} \right) \mathbf{1}_{\{f \neq 1\}} + 2 \cdot \mathbf{1}_{\{f = 1\}} \right\} \; d\mu = 2$$

$$\Longrightarrow \int \left(f + \frac{1}{f} - 2 \right) \mathbf{1}_{\{f \neq 1\}} \; d\mu = 0$$

$$\Longrightarrow \mu \left\{ \left(f + \frac{1}{f} - 2 \right) \mathbf{1}_{\{f \neq 1\}} = 0 \right\} = 1 \Longrightarrow \mu \left\{ f = 1 \right\} = 1$$

□

B S–Plus functions for the permutational tests

cdfdata

```
function(l,r,svf){
    n <- length(l)
    q <- c(t(svf$fits[[1]][,2]))
    p <- c(t(svf$fits[[1]][,1]))
    w <- c(t(svf$fits[[1]][,3]))
    p[1:(length(q)-1)] <- p[2:length(q)]
    w[1:(length(q)-1)] <- w[1:(length(q)-1)]-w[2:length(q)]
    q <- q[1:(length(q)-1)]
    p <- p[1:(length(p)-1)]
    w <- w[1:(length(w)-1)]
    while(p[length(p)]>10^28){
        q <- q[1:(length(q)-1)]
        p <- p[1:(length(p)-1)]
        w <- w[1:(length(w)-1)]
    }
    cdf.right <- numeric(n)
    cdf.right[r>10^28] <- 1
    cdf.right[r<=10^28] <- c(w%*%matrix(p<=rep(r[r<=10^28],
    each=length(p)),length(p),))
    cdf.left <- c(w%*%matrix(p<=rep(l,each=length(p)),length(p),))
    cbind(cdf.left,cdf.right)
}
```

WGsc

```
function(l, r){
    left <- numeric(length(l))
    right <- numeric(length(l))
    cc <- numeric(length(l))
    for(i in 1:length(l)) cc[i] <- length(r[r <= l[i]])-
    length(l[l >= r[i]])
    cc
}
```

HFsc

```
function(cdf,rho){
    l <- cdf[,1]
    r <- cdf[,2]
    if(rho==0){
        cc <- numeric(length(l))
        cc[r==1] <- -(1-l[r==1])*log(1-l[r==1])/(1-l[r==1])
        cc[r<1] <- ((1-r[r<1])*log(1-r[r<1])-(1-l[r<1])*log(1-l[r<1]))/
        (r[r<1]-l[r<1])
    }
    else{
        cc <- ((1-r)^(rho+1)-(1-l)^(rho+1))/(rho*(r-l))+(1/rho)
    }
    cc
}
```

DiMsc

```
function(l,r,svf){
    n <- length(l)
    q <- c(t(svf$fits[[1]][,2]))
    p <- c(t(svf$fits[[1]][,1]))
    w <- c(t(svf$fits[[1]][,3]))
    p[1:(length(q)-1)] <- p[2:length(q)]
    w[1:(length(q)-1)] <- w[1:(length(q)-1)]-w[2:length(q)]
    q <- q[1:(length(q)-1)]
    p <- p[1:(length(p)-1)]
    w <- w[1:(length(w)-1)]
    while(p[length(p)]>10^28){
        q <- q[1:(length(q)-1)]
        p <- p[1:(length(p)-1)]
        w <- w[1:(length(w)-1)]
    }
    cdf.right <- numeric(n)
    cdf.right[r>10^28] <- 1
    cdf.right[r<=10^28] <- c(w%*%matrix(p<=rep(r[r<=10^28],
    each=length(p)),length(p),))
    cdf.left <- c(w%*%matrix(p<=rep(l,each=length(p)),length(p),))
    m <- (1-sum(w))*max(l)+sum(p*w)
    m.right <- numeric(n)
    m.right[r>10^28] <- m
    m.right[r<=10^28] <- c((p*w)%*%matrix(p<=rep(r[r<=10^28],
    each=length(p)),length(p),))
    m.left <- c((p*w)%*%matrix(p<=rep(l,each=length(p)),length(p),))
    mm <- (m.right-m.left)/(cdf.right-cdf.left)
```

```
    mm-mean(mm)
}
```

w2test

```
function(cc,z){
    L <- sum(cc[z==1])
    m <- mean(cc)*length(z[z==1])
    v <- ((length(z[z==0])*length(z[z==1]))/length(z))*var(cc)
    list(c("The L statistic"),L,c("The permutation mean"),m,c("The
    permutation variance"),v,c("The standardized value"),(L-m)/sqrt(v))
}
```

wktest

```
function(cc,z){
    zz <- matrix(data=NA,length(z),max(z))
    for(i in 1:max(z)) zz[,i] <- (z==i)/sqrt(length(z[z==i]))
    L <- t(cc)%*%zz
    m <- t(length(z)*mean(cc)*colMeans(zz))
    v <- var(cc)*((t(zz)%*%zz)-length(z)*(t(t(colMeans(zz)))%*%
    t(colMeans(zz))))
    W <- c((L-m)%*%ginverse(v)%*%t(L-m))
    list(c("The L statistic"),L,c("The permutation mean"),m,c("The
    permutation variance"),v,c("The Mahalanobis distance"),W)
}
```

C Sums in the permutational variance of the Jonckheere's test

In the JK statistic $b_{ij} = \sum_{\substack{r,s=1 \\ r<s}}^{k} \alpha_i^{(s)} \alpha_j^{(r)} - \sum_{\substack{r,s=1 \\ r>s}}^{k} \alpha_i^{(s)} \alpha_j^{(r)}$, thus

$$\sum_{i,j=1}^{n} b_{ij}^2 = \sum_{i,j=1}^{n} \left[\sum_{\substack{r,s=1 \\ r<s}}^{k} \alpha_i^{(s)} \alpha_j^{(r)} - \sum_{\substack{r,s=1 \\ r>s}}^{k} \alpha_i^{(s)} \alpha_j^{(r)} \right]^2 = \sum_{i,j=1}^{n} \sum_{\substack{r,s=1 \\ r<s}}^{k} \alpha_i^{(s)} \alpha_j^{(r)} + \sum_{i,j=1}^{n} \sum_{\substack{r,s=1 \\ r>s}}^{k} \alpha_i^{(s)} \alpha_j^{(r)}$$

$$= \sum_{\substack{r,s=1 \\ r<s}}^{k} n_s\, n_r + \sum_{\substack{r,s=1 \\ r>s}}^{k} n_s\, n_r = \sum_{r,s=1}^{k} n_s\, n_r - \sum_{\substack{r,s=1 \\ r=s}}^{k} n_s\, n_r = n^2 - \sum_{r=1}^{k} n_r^2.$$

For the other sum in the permutational variance, we derive

$$\sum_{i,j_1,j_2=1}^{n} b_{ij_1} b_{ij_2} = \sum_{i,j_1,j_2=1}^{n} \left[\sum_{\substack{r,s=1 \\ r<s}}^{k} \alpha_i^{(s)} \alpha_{j_1}^{(r)} - \sum_{\substack{r,s=1 \\ r>s}}^{k} \alpha_i^{(s)} \alpha_{j_1}^{(r)} \right] \left[\sum_{\substack{r,s=1 \\ r<s}}^{k} \alpha_i^{(s)} \alpha_{j_2}^{(r)} - \sum_{\substack{r,s=1 \\ r>s}}^{k} \alpha_i^{(s)} \alpha_{j_2}^{(r)} \right]$$

$$= \sum_{i,j_1,j_2=1}^{n} \sum_{\substack{r_1,r_2,s=1 \\ r_1,r_2<s}}^{k} \alpha_i^{(s)} \alpha_{j_1}^{(r_1)} \alpha_{j_2}^{(r_2)} - \sum_{i,j_1,j_2=1}^{n} \sum_{\substack{r_1,r_2,s=1 \\ r_2<s<r_1}}^{k} \alpha_i^{(s)} \alpha_{j_1}^{(r_1)} \alpha_{j_2}^{(r_2)}$$

$$- \sum_{i,j_1,j_2=1}^{n} \sum_{\substack{r_1,r_2,s=1 \\ r_1<s<r_2}}^{k} \alpha_i^{(s)} \alpha_{j_1}^{(r_1)} \alpha_{j_2}^{(r_2)} + \sum_{i,j_1,j_2=1}^{n} \sum_{\substack{r_1,r_2,s=1 \\ r_1,r_2>s}}^{k} \alpha_i^{(s)} \alpha_{j_1}^{(r_1)} \alpha_{j_2}^{(r_2)}$$

$$= \sum_{\substack{r_1,r_2,s=1 \\ r_1,r_2<s}}^{k} n_s n_{r_1} n_{r_2} - \sum_{\substack{r_1,r_2,s=1 \\ r_2<s<r_1}}^{k} n_s n_{r_1} n_{r_2} - \sum_{\substack{r_1,r_2,s=1 \\ r_1<s<r_2}}^{k} n_s n_{r_1} n_{r_2} + \sum_{\substack{r_1,r_2,s=1 \\ r_1,r_2>s}}^{k} n_s n_{r_1} n_{r_2}.$$

Now, we split the following sums,

$$\sum_{\substack{r_1,r_2,s=1 \\ r_1,r_2<s}}^{k} n_s n_{r_1} n_{r_2} = \sum_{\substack{r_1,r_2,s=1 \\ r_1<r_2<s}}^{k} n_s n_{r_1} n_{r_2} + \sum_{\substack{r_1,r_2,s=1 \\ r_2<r_1<s}}^{k} n_s n_{r_1} n_{r_2} + \sum_{\substack{r_1,r_2,s=1 \\ r_1=r_2<s}}^{k} n_s n_{r_1} n_{r_2},$$

$$\sum_{\substack{r_1,r_2,s=1 \\ r_1,r_2>s}}^{k} n_s n_{r_1} n_{r_2} = \sum_{\substack{r_1,r_2,s=1 \\ s<r_1<r_2}}^{k} n_s n_{r_1} n_{r_2} + \sum_{\substack{r_1,r_2,s=1 \\ s<r_2<r_2}}^{k} n_s n_{r_1} n_{r_2} + \sum_{\substack{r_1,r_2,s=1 \\ s<r_1=r_2}}^{k} n_s n_{r_1} n_{r_2},$$

and we use the following property of symmetry,

$$\sum_{\substack{r_1, r_2, s = 1 \\ r_1 < r_2 < s}}^{k} n_s n_{r_1} n_{r_2} = \sum_{\substack{r_1, r_2, s = 1 \\ r_2 < r_1 < s}}^{k} n_s n_{r_1} n_{r_2} = \sum_{\substack{r_1, r_2, s = 1 \\ r_1 < s < r_2}}^{k} n_s n_{r_1} n_{r_2} = \sum_{\substack{r_1, r_2, s = 1 \\ r_2 < s < r_1}}^{k} n_s n_{r_1} n_{r_2}$$

$$= \sum_{\substack{r_1, r_2, s = 1 \\ s < r_1 < r_2}}^{k} n_s n_{r_1} n_{r_2} = \sum_{\substack{r_1, r_2, s = 1 \\ s < r_2 < r_1}}^{k} n_s n_{r_1} n_{r_2},$$

to obtain

$$\sum_{i,j_1,j_2=1}^{n} b_{ij_1} \, b_{ij_2} = 2 \cdot \sum_{\substack{r_1, r_2, s = 1 \\ r_1 < r_2 < s}}^{k} n_s n_{r_1} n_{r_2} + \sum_{\substack{r_1, r_2, s = 1 \\ r_1 = r_2 < s}}^{k} n_s n_{r_1} n_{r_2} + \sum_{\substack{r_1, r_2, s = 1 \\ s < r_1 = r_2}}^{k} n_s n_{r_1} n_{r_2}.$$

Finally, since

$$\sum_{r_1,r_2,s=1}^{k} n_s n_{r_1} n_{r_2} = 6 \cdot \sum_{\substack{r_1, r_2, s = 1 \\ r_1 < r_2 < s}}^{k} n_s n_{r_1} n_{r_2} + 3 \cdot \sum_{\substack{r_1, r_2, s = 1 \\ r_1 = r_2 < s}}^{k} n_s n_{r_1} n_{r_2} + 3 \cdot \sum_{\substack{r_1, r_2, s = 1 \\ s < r_1 = r_2}}^{k} n_s n_{r_1} n_{r_2} + \sum_{\substack{r_1, r_2, s = 1 \\ r_1 = r_2 = s}}^{k} n_s n_{r_1} n_{r_2},$$

it follows that

$$\sum_{i,j_1,j_2=1}^{n} b_{ij_1} \, b_{ij_2} = \frac{1}{3} \left[\sum_{r_1,r_2,s=1}^{k} n_s n_{r_1} n_{r_2} - \sum_{\substack{r_1, r_2, s = 1 \\ r_1 = r_2 = s}}^{k} n_s n_{r_1} n_{r_2} \right] = \frac{1}{3} \left[n^3 - \sum_{r=1}^{k} n_r^3 \right].$$

D Permutational distribution of the Abel's test

Proposition D.1. *The statistic $L^{WG}_{1,\dots,s-1;s}$ is equivalent to a Kendall's correlation coefficient,*

$$L^{WG}_{1,\dots,s-1;s} = \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}\, b_{ij}$$

where $a_{ij} = \Phi_{WG}(i,j)$ *and* $b_{ij} = \alpha_i^{(s)} \sum_{r=1}^{s-1} \alpha_j^{(r)} - \alpha_j^{(s)} \sum_{r=1}^{s-1} \alpha_i^{(r)}.$

Proof:

Since the necessary antisymmetric properties of a Kendall's correlation coefficient hold, $a_{ij} = -a_{ji}$ and $b_{ij} = -b_{ji}$, we only need to show the equivalence of both expressions. We use the antisymmetric property of the functional Φ_{WG} and a rearrangement of indices to proof this result:

$$\begin{aligned}
\sum_{i,j=1}^{n} a_{ij}\, b_{ij} &= \sum_{i,j=1}^{n} \Phi_{WG}(i,j) \left[\alpha_i^{(s)} \sum_{r=1}^{s-1} \alpha_j^{(r)} - \alpha_j^{(s)} \sum_{r=1}^{s-1} \alpha_i^{(r)} \right] \\
&= \sum_{i,j=1}^{n} \alpha_i^{(s)} \left(\sum_{r=1}^{s-1} \alpha_j^{(r)} \right) \Phi_{WG}(i,j) - \sum_{i,j=1}^{n} \left(\sum_{r=1}^{s-1} \alpha_i^{(r)} \right) \alpha_j^{(s)}\, \Phi_{WG}(i,j) \\
&= \sum_{i,j=1}^{n} \alpha_i^{(s)} \left(\sum_{r=1}^{s-1} \alpha_j^{(r)} \right) \Phi_{WG}(i,j) + \sum_{i,j=1}^{n} \left(\sum_{r=1}^{s-1} \alpha_i^{(r)} \right) \alpha_j^{(s)}\, \Phi_{WG}(j,i) \\
&= 2 \sum_{i,j=1}^{n} \alpha_i^{(s)} \left(\sum_{r=1}^{s-1} \alpha_j^{(r)} \right) \Phi_{WG}(i,j)
\end{aligned}$$

□

Proposition D.2. *Under the null hypothesis, the permutational distribution of the statistic $L^{WG}_{1,\dots,s-1;s}$ is asymptotically normal with zero mean and variance given by*

$$V(L^{WG}_{1,\ldots,s-1;s}) = \frac{(\sum_{r=1}^{s-1} n_r)\; n_s\; (\sum_{r=1}^{s} n_r\; -2)}{n\;(n-1)\;(n-2)} \sum_{i=1}^{n}\left[\sum_{j=1}^{n}\Phi_{WG}(i,j)\right]^2$$

$$+\frac{(\sum_{r=1}^{s-1} n_r)\; n_s\; (n-\sum_{r=1}^{s} n_r)}{n\;(n-1)\;(n-2)} \sum_{i,j=1}^{n}\left[\Phi_{WG}(i,j)\right]^2$$

Proof:

The permutational distribution of the $L^{WG}_{1,\ldots,s-1;s}$ statistic coincides with the distribution of a Kendall's correlation coefficient and, consequently, is asymptotically normal with zero mean and variance given by

$$V(L^{WG}_{1,\ldots,s-1;s}) = \frac{1}{n(n-1)(n-2)}\left[\sum_{i,j_1,j_2=1}^{n} a_{ij_1}a_{ij_2} - \sum_{i,j=1}^{n} a_{ij}^2\right]\left[\sum_{i,j_1,j_2=1}^{n} b_{ij_1}b_{ij_2} - \sum_{i,j=1}^{n} b_{ij}^2\right]$$

$$+\frac{1}{2n(n-1)}\left[\sum_{i,j=1}^{n} a_{ij}^2\right]\left[\sum_{i,j=1}^{n} b_{ij}^2\right] = \left(\frac{\sum_{i,j_1,j_2=1}^{n} b_{ij_1}b_{ij_2} - \sum_{i,j=1}^{n} b_{ij}^2}{n(n-1)(n-2)}\right)\sum_{i,j_1,j_2=1}^{n} a_{ij_1}a_{ij_2}$$

$$+\left(\frac{n\sum_{i,j=1}^{n} b_{ij}^2 - 2\sum_{i,j_1,j_2=1}^{n} b_{ij_1}b_{ij_2}}{2n(n-1)(n-2)}\right)\sum_{i,j=1}^{n} a_{ij}^2$$

Since $a_{ij} = \Phi_{WG}(i,j)$, it immediately follows that

$$V(L^{WG}_{1,\ldots,s-1;s}) = \left(\frac{\sum_{i,j_1,j_2=1}^{n} b_{ij_1}b_{ij_2} - \sum_{i,j=1}^{n} b_{ij}^2}{n(n-1)(n-2)}\right)\sum_{i=1}^{n}\left[\sum_{j=1}^{n}\Phi_{WG}(i,j)\right]^2$$

$$+\left(\frac{n\sum_{i,j=1}^{n} b_{ij}^2 - 2\sum_{i,j_1,j_2=1}^{n} b_{ij_1}b_{ij_2}}{2n(n-1)(n-2)}\right)\sum_{i,j=1}^{n}\left[\Phi_{WG}(i,j)\right]^2$$

Now, we do some algebraic manipulation and we obtain

$$\sum_{i,j=1}^{n} b_{ij}^2 = \sum_{i,j=1}^{n}\left[\alpha_i^{(s)}\sum_{r=1}^{s-1}\alpha_j^{(r)} - \alpha_j^{(s)}\sum_{r=1}^{s-1}\alpha_i^{(r)}\right]^2 = \sum_{i,j=1}^{n}\alpha_i^{(s)}\sum_{r=1}^{s-1}\alpha_j^{(r)} + \sum_{i,j=1}^{n}\alpha_j^{(s)}\sum_{r=1}^{s-1}\alpha_i^{(r)}$$

$$= \sum_{r=1}^{s-1}\sum_{i,j=1}^{n}\alpha_i^{(s)}\alpha_j^{(r)} + \sum_{r=1}^{s-1}\sum_{i,j=1}^{n}\alpha_i^{(r)}\alpha_j^{(s)} = 2\; n_s\sum_{r=1}^{s-1} n_r$$

and

$$\sum_{i,j_1,j_2=1}^{n} b_{ij_1}\, b_{ij_2} = \sum_{i,j_1,j_2=1}^{n} \left[\alpha_i^{(s)} \sum_{r=1}^{s-1} \alpha_{j_1}^{(r)} - \alpha_{j_1}^{(s)} \sum_{r=1}^{s-1} \alpha_i^{(r)}\right]\left[\alpha_i^{(s)} \sum_{r=1}^{s-1} \alpha_{j_2}^{(r)} - \alpha_{j_2}^{(s)} \sum_{r=1}^{s-1} \alpha_i^{(r)}\right]$$

$$= \sum_{i,j_1,j_2=1}^{n} \sum_{r_1,r_2=1}^{s-1} \alpha_i^{(s)}\, \alpha_{j_1}^{(r_1)}\, \alpha_{j_2}^{(r_2)} + \sum_{i,j_1,j_2=1}^{n} \sum_{r=1}^{s-1} \alpha_i^{(r)}\, \alpha_{j_1}^{(s)}\, \alpha_{j_2}^{(s)}$$

$$= n_s \left(\sum_{r=1}^{s-1} n_r\right)^2 + n_s^2 \sum_{r=1}^{s-1} n_r = n_s \left(\sum_{r=1}^{s-1} n_r\right)\left(\sum_{r=1}^{s} n_r\right)$$

We complete the proof replacing these results in the permutational variance given in the previous equation.

□

Proposition D.3. *The $L^{WG}_{1,\dots,s-1;s}$ statistics are uncorrelated with the permutational distribution.*

Proof:

The mean value of two Kendall's correlation coefficients product is given by

$$E(\tau^{(1)}\tau^{(2)}) = \tfrac{4}{n(n-1)(n-2)} \left[\sum_{i,j_1,j_2=1}^{n} a_{ij_1}^{(1)} a_{ij_2}^{(2)} - \sum_{i,j=1}^{n} a_{ij}^{(1)} a_{ij}^{(2)}\right] \cdot \left[\sum_{i,j_1,j_2=1}^{n} b_{ij_1}^{(1)} b_{ij_2}^{(2)} - \sum_{i,j=1}^{n} b_{ij}^{(1)} b_{ij}^{(2)}\right]$$

$$+ \tfrac{2}{n(n-1)} \left[\sum_{i,j=1}^{n} a_{ij}^{(1)} a_{ij}^{(2)}\right]\left[\sum_{i,j=1}^{n} b_{ij}^{(1)} b_{ij}^{(2)}\right]$$

In the particular case when $\tau^{(1)} = L^{WG}_{1,\dots,s_1-1;s_1}$ and $\tau^{(2)} = L^{WG}_{1,\dots,s_2-1;s_2}$ with $s_1 < s_2$, we prove the result showing that

$$\sum_{i,j=1}^{n} b_{ij}^{(1)} b_{ij}^{(2)} = \sum_{i,j=1}^{n} \left[\alpha_i^{(s_1)} \sum_{r=1}^{s_1-1} \alpha_j^{(r)} - \alpha_j^{(s_1)} \sum_{r=1}^{s_1-1} \alpha_i^{(r)}\right] \cdot \left[\alpha_i^{(s_2)} \sum_{r=1}^{s_2-1} \alpha_j^{(r)} - \alpha_j^{(s_2)} \sum_{r=1}^{s_2-1} \alpha_i^{(r)}\right] = 0$$

and

$$\sum_{i,j_1,j_2=1}^{n} b_{ij_1}^{(1)} b_{ij_2}^{(2)} = \sum_{i,j_1,j_2=1}^{n} \left[\alpha_i^{(s_1)} \sum_{r=1}^{s_1-1} \alpha_{j_1}^{(r)} - \alpha_{j_1}^{(s_1)} \sum_{r=1}^{s_1-1} \alpha_i^{(r)} \right] \cdot \left[\alpha_i^{(s_2)} \sum_{r=1}^{s_2-1} \alpha_{j_2}^{(r)} - \alpha_{j_2}^{(s_2)} \sum_{r=1}^{s_2-1} \alpha_i^{(r)} \right]$$

$$= - \sum_{i,j_1,j_2=1}^{n} \alpha_i^{(s_1)} \, \alpha_{j_2}^{(s_2)} \sum_{r=1}^{s_1-1} \alpha_{j_1}^{(r)} + \sum_{i,j_1,j_2=1}^{n} \alpha_{j_1}^{(s_1)} \, \alpha_{j_2}^{(s_2)} \sum_{r=1}^{s_1-1} \alpha_i^{(r)} = 0$$

□

Corollary D.4. *As a sum of the permutational variances of the $L_{1,\ldots,s-1;s}^{WG}$ statistics, the permutational variance of the JA test statistic is given by*

$$V(JA) = \tfrac{n^3 - 3n^2 - \sum_{r=1}^{k} n_r^3 + 3 \sum_{r=1}^{k} n_r^2}{3n(n-1)(n-2)} \sum_{i=1}^{n} \left[\sum_{j=1}^{n} \Phi_{WG}(i,j) \right]^2$$

$$+ \tfrac{n^3 + 2 \sum_{r=1}^{k} n_r^3 - 3n \sum_{r=1}^{k} n_r^2}{6n(n-1)(n-2)} \sum_{i,j=1}^{n} \left[\Phi_{WG}(i,j) \right]^2 .$$

Proof:

To prove this result it is enough to show that:

$$\sum_{s=2}^{k} n_s \left(\sum_{r=1}^{s-1} n_r \right) \left(\sum_{r=1}^{s} n_r - 2 \right) = \frac{1}{3} \left[n^3 - 3n^2 - \sum_{r=1}^{k} n_r^3 + 3 \sum_{r=1}^{k} n_r^2 \right]$$

and

$$\sum_{s=2}^{k} n_s \left(\sum_{r=1}^{s-1} n_r \right) \left(n - \sum_{r=1}^{s} n_r \right) = \frac{1}{6} \left[n^3 + 2 \sum_{r=1}^{k} n_r^3 - 3n \sum_{r=1}^{k} n_r^2 \right]$$

These equalities follow from the same sum arguments given in Appendix C.

□

E S–Plus functions for the Jonckeere–type tests

JK

```
function(cdf,z,rho,nblock=25){
    l <- cdf[,1]
    r <- cdf[,2]
    n <- length(l)
    nn <- ceiling(n/nblock)
    aux3 <- numeric(n)
    jk <- 0
    var1 <- 0
    var2 <- 0
    for(i in 1:nn) {
        n1 <- nblock*(i-1)+1
        n2 <- ifelse(i<nn,nblock*i,n)
        n3 <- ifelse(i<nn,nblock,n2-n1+1)
        n4 <- n-n1+1
        if(rho==0){
            aux1 <- pmax(matrix(l[n1:n2],n3,n4),matrix(r[n1:n],n3,n4,
            byrow=T))
            aux2 <- pmax(matrix(l[n1:n2],n3,n4),matrix(l[n1:n],n3,n4,
            byrow=T))
            aux1[aux1==1] <- 0
            aux1[aux1!=0] <- ((1-aux1[aux1!=0])*log(1-aux1[aux1!=0]))/
            aux1[aux1!=0]
            aux2[aux2!=0] <- ((1-aux2[aux2!=0])*log(1-aux2[aux2!=0]))/
            aux2[aux2!=0]
            phi <- l[n1:n2]*(aux1-aux2)
            aux1 <- pmax(matrix(l[n1:n],n3,n4,byrow=T),matrix(r[n1:n2],
```

```
n3,n4))
aux2 <- pmax(matrix(l[n1:n],n3,n4,byrow=T),matrix(l[n1:n2],
n3,n4))
aux1[aux1==1] <- 0
aux1[aux1!=0] <- ((1-aux1[aux1!=0])*log(1-aux1[aux1!=0]))/
aux1[aux1!=0]
aux2[aux2!=0] <- ((1-aux2[aux2!=0])*log(1-aux2[aux2!=0]))/
aux2[aux2!=0]
phi <- phi-t(l[n1:n]*t(aux1-aux2))
aux1 <- pmax(matrix(r[n1:n],n3,n4,byrow=T),matrix(r[n1:n2],
n3,n4))
aux2 <- pmax(matrix(r[n1:n],n3,n4,byrow=T),matrix(l[n1:n2],
n3,n4))
aux1[aux1==1] <- 0
aux1[aux1!=0] <- ((1-aux1[aux1!=0])*log(1-aux1[aux1!=0]))/
aux1[aux1!=0]
aux2[aux2==1] <- 0
aux2[aux2!=0] <- ((1-aux2[aux2!=0])*log(1-aux2[aux2!=0]))/
aux2[aux2!=0]
phi <- phi+t(r[n1:n]*t(aux1-aux2))
aux1 <- pmax(matrix(r[n1:n2],n3,n4),matrix(r[n1:n],n3,n4,
byrow=T))
aux2 <- pmax(matrix(r[n1:n2],n3,n4),matrix(l[n1:n],n3,n4,
byrow=T))
aux1[aux1==1] <- 0
aux1[aux1!=0] <- ((1-aux1[aux1!=0])*log(1-aux1[aux1!=0]))/
aux1[aux1!=0]
aux2[aux2==1] <- 0
aux2[aux2!=0] <- ((1-aux2[aux2!=0])*log(1-aux2[aux2!=0]))/
```

```
    aux2[aux2!=0]
    phi <- phi-r[n1:n2]*(aux1-aux2)
}
else{
    aux1 <- pmax(matrix(l[n1:n2],n3,n4),matrix(r[n1:n],n3,n4,
    byrow=T))
    aux2 <- pmax(matrix(l[n1:n2],n3,n4),matrix(l[n1:n],n3,n4,
    byrow=T))
    aux1[aux1!=0] <- ((1-aux1[aux1!=0])^(rho+1)-
    (1-aux1[aux1!=0]))/(rho*aux1[aux1!=0])
    aux2[aux2!=0] <- ((1-aux2[aux2!=0])^(rho+1)-
    (1-aux2[aux2!=0]))/(rho*aux2[aux2!=0])
    phi <- l[n1:n2]*(aux1-aux2)
    aux1 <- pmax(matrix(l[n1:n],n3,n4,byrow=T),matrix(r[n1:n2],
    n3,n4))
    aux2 <- pmax(matrix(l[n1:n],n3,n4,byrow=T),matrix(l[n1:n2],
    n3,n4))
    aux1[aux1!=0] <- ((1-aux1[aux1!=0])^(rho+1)-
    (1-aux1[aux1!=0]))/(rho*aux1[aux1!=0])
    aux2[aux2!=0] <- ((1-aux2[aux2!=0])^(rho+1)-
    (1-aux2[aux2!=0]))/(rho*aux2[aux2!=0])
    phi <- phi-t(l[n1:n]*t(aux1-aux2))
    aux1 <- pmax(matrix(r[n1:n],n3,n4,byrow=T),matrix(r[n1:n2],
    n3,n4))
    aux2 <- pmax(matrix(r[n1:n],n3,n4,byrow=T),matrix(l[n1:n2],
    n3,n4))
    aux1[aux1!=0] <- ((1-aux1[aux1!=0])^(rho+1)-
    (1-aux1[aux1!=0]))/(rho*aux1[aux1!=0])
    aux2[aux2!=0] <- ((1-aux2[aux2!=0])^(rho+1)-
```

```
            (1-aux2[aux2!=0]))/(rho*aux2[aux2!=0])
            phi <- phi+t(r[n1:n]*t(aux1-aux2))
            aux1 <- pmax(matrix(r[n1:n2],n3,n4),matrix(r[n1:n],n3,n4,
            byrow=T))
            aux2 <- pmax(matrix(r[n1:n2],n3,n4),matrix(l[n1:n],n3,n4,
            byrow=T))
            aux1[aux1!=0] <- ((1-aux1[aux1!=0])^(rho+1)-
            (1-aux1[aux1!=0]))/(rho*aux1[aux1!=0])
            aux2[aux2!=0] <- ((1-aux2[aux2!=0])^(rho+1)-
            (1-aux2[aux2!=0]))/(rho*aux2[aux2!=0])
            phi <- phi-r[n1:n2]*(aux1-aux2)
        }
        phi <- phi/((r[n1:n2]-l[n1:n2])%*%t(r[n1:n]-l[n1:n]))
        jk <- jk+sum(phi*sign(z[n1:n2]*(1-lower.tri(phi))-t(z[n1:n]*
        t(1-lower.tri(phi)))))
        var1 <- var1+sum((aux3[n1:n2]+rowSums(phi))^2)
        var2 <- var2+sum((phi*(1-lower.tri(phi)))^2)
        if((n2!=n)&(nblock!=1)) aux3[(n2+1):n] <- aux3[(n2+1):n]-
        colSums(phi[,(nblock+1):n4])
        if((n2!=n)&(nblock==1)) aux3[(n2+1):n] <- aux3[(n2+1):n]-
        c(phi[,(nblock+1):n4])
    }
    ngroup <- c(table(z))
    aux1 <- n^2-3*n-(sum((ngroup^3-3*ngroup^2)/n))
    var1 <- (aux1*var1)/(3*(n-1)*(n-2))
    aux2 <- n^2+2*(sum(ngroup^3/n))-3*(sum(ngroup^2))
    var2 <- (aux2*var2)/(3*(n-1)*(n-2))
    jk/sqrt(var1+var2)
}
```

Ltrend

```
function(cc,z){
    L <- sum(cc*z)
    m <- length(z)*mean(cc)*mean(z)
    v <- (length(z)-1)*var(cc)*var(z)
    (L-m)/sqrt(v)
}
```

J1boot

```
function(l,r,z,rho){
    n <- length(z)
    k <- max(z)
    npair <- choose(k,2)
    aux <- matrix(0,k,k)
    a <- rep(z,npair)
    b <- c(t(row(aux)*(1-lower.tri(aux,diag=T))))
    b <- b[b!=0]
    b <- rep(b,each=n)
    c <- c(t(col(aux)*(1-lower.tri(aux,diag=T))))
    c <- c[c!=0]
    c <- rep(c,each=n)
    zcontrol <- numeric(n*npair)
    zcontrol[a==b] <- 1
    zcontrol[a==c] <- 2
    zl <- rep(l,npair)[zcontrol!=0]
    zr <- rep(r,npair)[zcontrol!=0]
    zz <- rep(1:npair,each=n)[zcontrol!=0]
    j1 <- 0
```

```
    for(i in 1:npair) {
        zzl <- zl[zz==i]
        zzr <- zr[zz==i]
        svf <- kaplanMeier(censor(zzl,zzr,rep(3,length(zzl)))~1,
        data.frame(zzl,zzr,rep(3,length(zzl))),se.fit=F,
        conf.interval="none")
        j1 <- j1+c((zcontrol[zcontrol!=0][zz==i]-1) %*%
        HFsc(cdfdata(zzl,zzr,svf),rho))
        }
    j1
}
```

J2boot

```
function(l,r,z,rho){
    n <- length(z)
    npair <- max(z)-1
    a <- rep(z,npair)
    b <- rep((1+npair):2,each=n)
    zcontrol <- numeric(n*npair)
    zcontrol[a<b] <- 1
    zcontrol[a==b] <- 2
    zl <- rep(l,npair)[zcontrol!=0]
    zr <- rep(r,npair)[zcontrol!=0]
    zz <- rep(1:npair,each=n)[zcontrol!=0]
    j2 <- 0
    for(i in 1:npair) {
        zzl <- zl[zz==i]
        zzr <- zr[zz==i]
```

```
        svf <- kaplanMeier(censor(zzl,zzr,rep(3,length(zzl)))~1,
        data.frame(zzl,zzr,rep(3,length(zzl))),se.fit=F,
        conf.interval="none")
        j2 <- j2+c((zcontrol[zcontrol!=0][zz==i]-1) %*%
        HFsc(cdfdata(zzl,zzr,svf),rho))
        }
    j2
}
```

J3boot

```
function(l,r,z,rho){
    n <- length(z)
    npair <- max(z)-1
    a <- rep(z,npair)
    b <- rep(1:npair,each=n)
    zcontrol <- numeric(n*npair)
    zcontrol[a==b] <- 1
    zcontrol[a>b] <- 2
    zl <- rep(l,npair)[zcontrol!=0]
    zr <- rep(r,npair)[zcontrol!=0]
    zz <- b[zcontrol!=0]
    j3 <- 0
    for(i in 1:npair) {
        zzl <- zl[zz==i]
        zzr <- zr[zz==i]
        code <-rep(3,length(zl[zz==i]))
        svf<- kaplanMeier(censor(zzl,zzr,code)~1,
        data=data.frame(zzl,zzr,code),se.fit=F,
```

```
        conf.interval="none")
        j3 <- j3+c((zcontrol[zcontrol!=0][zz==i]-1) %*%
        HFsc(cdfdata(zzl,zzr,svf),rho))
        }
    j3
}
```

JA

```
function(l,r,z,nblock=25){
    n <- length(l)
    nn <- ceiling(n/nblock)
    aux3 <- numeric(n)
    ja <- 0
    var1 <- 0
    var2 <- 0
    for(i in 1:nn) {
        n1 <- nblock*(i-1)+1
        n2 <- ifelse(i<nn,nblock*i,n)
        n3 <- ifelse(i<nn,nblock,n2-n1+1)
        n4 <- n-n1+1
        phi <- (matrix(l[n1:n2],n3,n4)>=matrix(r[n1:n],n3,n4,byrow=T))-
        (matrix(r[n1:n2],n3,n4)<=matrix(l[n1:n],n3,n4,byrow=T))
        ja <- ja+sum(phi*sign(z[n1:n2]*(1-lower.tri(phi))-
        t(z[n1:n]*t(1-lower.tri(phi)))))
        var1 <- var1+sum((aux3[n1:n2]+rowSums(phi))^2)
        var2 <- var2+sum((phi*(1-lower.tri(phi)))^2)
        if((n2!=n)&(nblock!=1)) aux3[(n2+1):n] <- aux3[(n2+1):n]-
        colSums(phi[,(nblock+1):n4])
        if((n2!=n)&(nblock==1)) aux3[(n2+1):n] <- aux3[(n2+1):n]-
```

```
        c(phi[,(nblock+1):n4])
    }
    ngroup <- c(table(z))
    aux1 <- n^2-3*n-(sum((ngroup^3-3*ngroup^2)/n))
    var1 <- (aux1*var1)/(3*(n-1)*(n-2))
    aux2 <- n^2+2*(sum(ngroup^3/n))-3*(sum(ngroup^2))
    var2 <- (aux2*var2)/(3*(n-1)*(n-2))
    ja/sqrt(var1+var2)
}
```

JA2

```
function(l,r,z){
    n <- length(z)
    npair <- max(z)-1
    a <- rep(z,npair)
    b <- rep((1+npair):2,each=n)
    zcontrol <- numeric(n*npair)
    zcontrol[a<b] <- 1
    zcontrol[a==b] <- 2
    zl <- rep(l,npair)[zcontrol!=0]
    zr <- rep(r,npair)[zcontrol!=0]
    zz <- rep(1:npair,each=n)[zcontrol!=0]
    ja2 <- 0
    var2 <- 0
    for(i in 1:npair) {
        cc <- WGsc(zl[zz==i],zr[zz==i])
        zaux <- zcontrol[zcontrol!=0][zz==i]-1
        ja2 <- ja2 + sum(cc*zaux)
```

```
        var2 <- var2 + ((length(zaux[zaux==0])*
        length(zaux[zaux==1]))/length(zaux))*var(cc)
        }
    ja2/sqrt(var2)
}
```

JKDiM

```
function(cc,z){
    ngroup <- c(table(z))
    k <- max(z) zz <- c(0,cumsum(ngroup[1:(k-1)]))-
    (c(0,cumsum(ngroup[k:2]))[k:1])
    zz <- zz[z]
    L <- sum(cc*zz)
    m <- length(z)*mean(cc)*mean(zz)
    v <- ((length(z)^3-sum(ngroup^3))/3)*var(cc) (L-m)/sqrt(v)
}
```

F Simplification of the Weighted Mann–Whitney functional

When

$$\Phi_{WMW}(F,G) = \int Q(s)G(s)dF(s) - \int Q(s)F(s)dG(s)$$

and

$$Q(t) = \frac{\gamma(\hat{W}(t)) - \gamma(\hat{W}(t^-))}{\hat{W}(t) - \hat{W}(t^-)}$$

we derive the following simplifications for the comparison between empirical distribution estimates $\hat{W}^i$ and $\hat{W}^j$.

$$\Phi_{WMW}(\hat{W}^i, \hat{W}^j) =$$

$$= \int_0^{+\infty} Q(t)\hat{W}^j(t)d\hat{W}^i(t) - \int_0^{+\infty} Q(t)\hat{W}^i(t)d\hat{W}^j(t)$$

$$= \frac{\int_{[l_i,r_i]\cap[l_j,r_j]} Q(t)(\hat{W}(t) - \hat{W}(l_j^-))d\hat{W}(t)}{(\hat{W}(r_i) - \hat{W}(l_i^-))(\hat{W}(r_j) - \hat{W}(l_j^-))} + \frac{\int_{[l_i,r_i]\cap(r_j,+\infty)} Q(t)d\hat{W}(t)}{\hat{W}(r_i) - \hat{W}(l_i^-)}$$

$$- \frac{\int_{[l_i,r_i]\cap[l_j,r_j]} Q(t)(\hat{W}(t) - \hat{W}(l_i^-))d\hat{W}(t)}{(\hat{W}(r_i) - \hat{W}(l_i^-))(\hat{W}(r_j) - \hat{W}(l_j^-))} - \frac{\int_{(r_i,+\infty)\cap[l_j,r_j]} Q(t)d\hat{W}(t)}{\hat{W}(r_j) - \hat{W}(l_j^-)}$$

$$= \frac{\int_{[l_i,r_i]\cap[l_j,r_j]} Q(t)(\hat{W}(l_i^-) - \hat{W}(l_j^-))d\hat{W}(t)}{(\hat{W}(r_i) - \hat{W}(l_i^-))(\hat{W}(r_j) - \hat{W}(l_j^-))} + \frac{\int_{[l_i,r_i]\cap(r_j,+\infty)} Q(t)d\hat{W}(t)}{\hat{W}(r_i) - \hat{W}(l_i^-)}$$

$$- \frac{\int_{(r_i,+\infty)\cap[l_j,r_j]} Q(t)d\hat{W}(t)}{\hat{W}(r_j) - \hat{W}(l_j^-)}$$

$$= \frac{\int_{[l_i,r_i]\cap[l_j,+\infty)} Q(t)(\hat{W}(l_i^-) - \hat{W}(l_j^-))d\hat{W}(t)}{(\hat{W}(r_i) - \hat{W}(l_i^-))(\hat{W}(r_j) - \hat{W}(l_j^-))} + \frac{\int_{[l_i,r_i]\cap(r_j,+\infty)} Q(t)d\hat{W}(t)}{\hat{W}(r_i) - \hat{W}(l_i^-)}$$

$$- \frac{\int_{[l_i,r_i]\cap(r_j,+\infty)} Q(t)(\hat{W}(l_i^-) - \hat{W}(l_j^-))d\hat{W}(t)}{(\hat{W}(r_i) - \hat{W}(l_i^-))(\hat{W}(r_j) - \hat{W}(l_j^-))} - \frac{\int_{(r_i,+\infty)\cap[l_j,r_j]} Q(t)d\hat{W}(t)}{\hat{W}(r_j) - \hat{W}(l_j^-)}$$

$$= \frac{\int_{[l_i,r_i]\cap[l_j,+\infty)} Q(t)(\hat{W}(l_i^-) - \hat{W}(l_j^-))d\hat{W}(t)}{(\hat{W}(r_i) - \hat{W}(l_i^-))(\hat{W}(r_j) - \hat{W}(l_j^-))}$$

$$+ \frac{\int_{[l_i,r_i]\cap(r_j,+\infty)} Q(t)(\hat{W}(r_j) - \hat{W}(l_i^-))d\hat{W}(t)}{(\hat{W}(r_i) - \hat{W}(l_i^-))(\hat{W}(r_j) - \hat{W}(l_j^-))} - \frac{\int_{(r_i,+\infty)\cap[l_j,r_j]} Q(t)d\hat{W}(t)}{\hat{W}(r_j) - \hat{W}(l_j^-)}$$

$$= \frac{(\hat{W}(l_i^-) - \hat{W}(l_j^-))(\gamma(\hat{W}(r_i \vee l_j^-)) - \gamma(\hat{W}(l_i^- \vee l_j^-)))}{(\hat{W}(r_i) - \hat{W}(l_i^-))(\hat{W}(r_j) - \hat{W}(l_j^-))}$$

$$+ \frac{(\hat{W}(r_j) - \hat{W}(l_i^-))(\gamma(\hat{W}(r_i \vee r_j)) - \gamma(\hat{W}(l_i^- \vee r_j)))}{(\hat{W}(r_i) - \hat{W}(l_i^-))(\hat{W}(r_j) - \hat{W}(l_j^-))}$$

$$- \frac{(\hat{W}(r_i) - \hat{W}(l_i^-))(\gamma(\hat{W}(r_j \vee r_i)) - \gamma(\hat{W}(l_j^- \vee r_i)))}{(\hat{W}(r_i) - \hat{W}(l_i^-))(\hat{W}(r_j) - \hat{W}(l_j^-))}$$

$$= \frac{\hat{W}(l_i^-)\left[\gamma(\hat{W}(l_i^- \vee r_j)) - \gamma(\hat{W}(l_i^- \vee l_j^-))\right]}{\{\hat{W}(r_i) - \hat{W}(l_i^-)\}\{\hat{W}(r_j) - \hat{W}(l_j^-)\}}$$

$$- \frac{\hat{W}(l_j^-)\left[\gamma(\hat{W}(l_j^- \vee r_i)) - \gamma(\hat{W}(l_j^- \vee l_i^-))\right]}{\{\hat{W}(r_i) - \hat{W}(l_i^-)\}\{\hat{W}(r_j) - \hat{W}(l_j^-)\}}$$

$$+ \frac{\hat{W}(r_j)\left[\gamma(\hat{W}(r_j \vee r_i)) - \gamma(\hat{W}(r_j \vee l_i^-))\right]}{\{\hat{W}(r_i) - \hat{W}(l_i^-)\}\{\hat{W}(r_j) - \hat{W}(l_j^-)\}}$$

$$- \frac{\hat{W}(r_i)\left[\gamma(\hat{W}(r_i \vee r_j)) - \gamma(\hat{W}(r_i \vee l_j^-))\right]}{\{\hat{W}(r_i) - \hat{W}(l_i^-)\}\{\hat{W}(r_j) - \hat{W}(l_j^-)\}}$$

G Simulation stuff

G.1 S–plus function for generating data

gendata

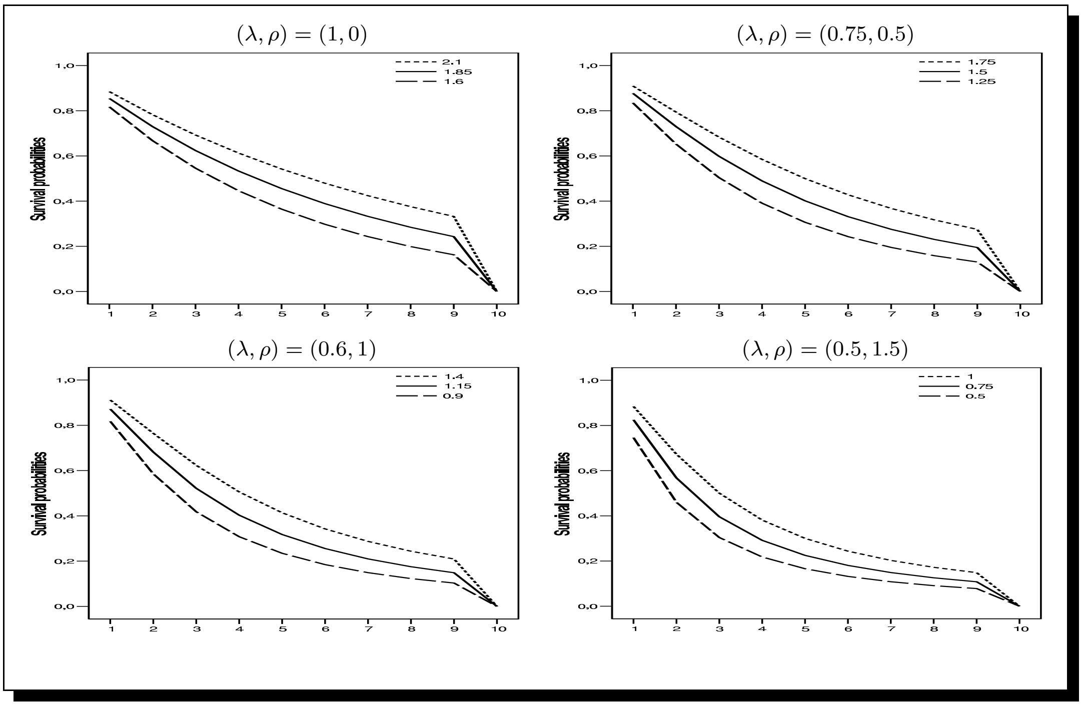
```
function(n,local,scale,rho,p){
    if(rho==0) t <- rweibull(n,1/scale,exp(local))
    else t <- exp(local)*((((1-runif(n,0,1))^-rho)-1)/rho)^scale
    t[t>=10] <- 10
    t[t<10] <- ceiling(t[t<10])
    delta <- matrix(c(rbinom(9*n,1,p),rep(1,n)),n,10)
    insp <- col(delta)>=t
    right<- row(t(delta))[t(pmin(delta,insp)*t(cumsum(t(pmin(delta,insp))))
    ==c(1,1+cumsum(rowSums(pmin(delta,insp)))[1:(n-1)]))]
    delta <- cbind(rep(1,n),delta[,1:9])
    insp <- col(delta)<=t
    left <- (row(t(delta))[t(pmin(delta,insp)*t(cumsum(t(pmin(delta,insp))))
    ==cumsum(rowSums(pmin(delta,insp))))])-1
    cbind(left,right)
}
```

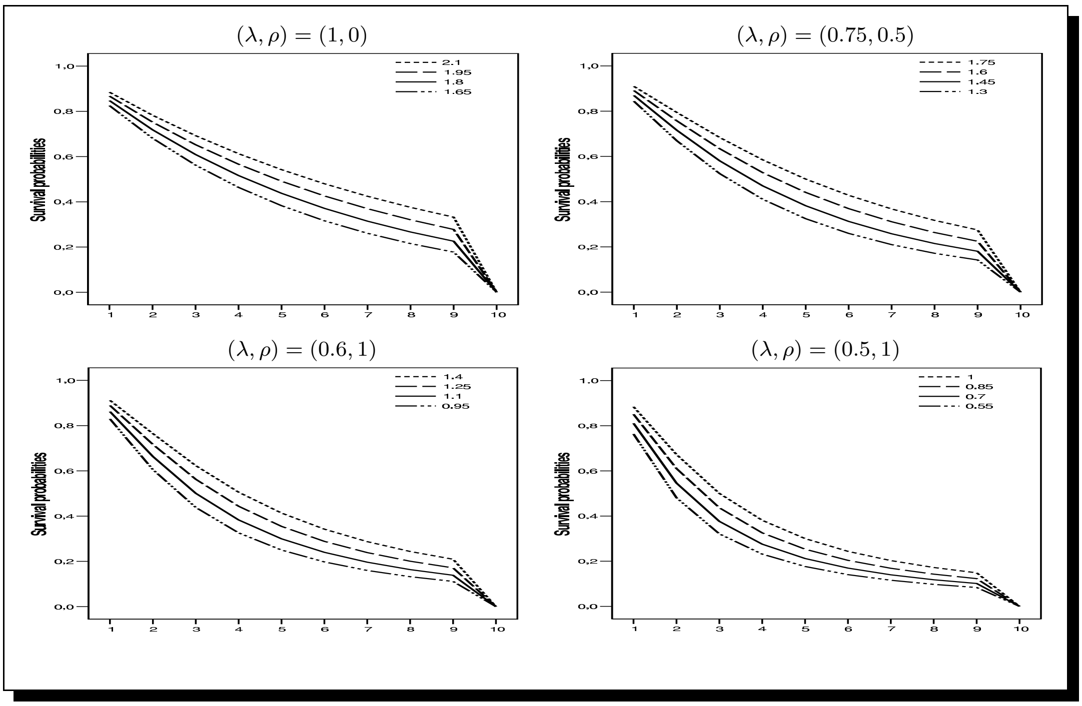
G.2 Tables of means and medians of the lifetimes models

(λ, ρ)	θ	Mean of T^*	Mean of T	Median of T^*	Median of T
$(1, 0)$	1.6	4.95	4.74	3.43	4
	1.85	6.36	5.45	4.41	5
	2.1	8.17	6.13	5.66	6
$(0.75, 0.5)$	1.25	4.89	4.42	3.03	4
	1	6.28	5.13	3.89	4
	1.75	8.06	5.86	5.00	5
$(0.6, 1)$	0.9	4.87	3.92	2.46	3
	1.15	6.26	4.58	3.16	4
	1.4	8.04	5.30	4.06	5
$(0.5, 1.5)$	0.5	4.90	3.30	1.82	2
	0.75	6.30	3.87	2.34	3
	1	8.09	4.51	3.00	4

(λ, ρ)	θ	Mean of T^*	Mean of T	Median of T^*	Median of T
$(1, 0)$	1.65	5.21	4.88	3.61	4
	1.8	6.05	5.31	4.19	5
	1.95	7.03	5.72	4.87	5
	2.1	8.17	6.13	5.66	6
$(0.75, 0.5)$	1.3	5.14	4.55	3.19	4
	1.45	5.97	4.98	3.70	4
	1.6	6.94	5.42	4.30	5
	1.75	8.06	5.86	5.00	5
$(0.6, 1)$	0.95	5.12	4.05	2.59	3
	1.1	5.95	4.45	3.00	4
	1.25	6.92	4.87	3.49	4
	1.4	8.04	5.30	4.06	5
$(0.5, 1.5)$	0.55	5.16	3.41	1.91	2
	0.7	5.99	3.75	2.22	3
	0.85	6.96	4.12	2.58	3
	1	8.09	4.51	3.00	4

G.3 Survival functions of the lifetimes models





G.4 S–plus function for performing the simulation study

simu

```
function(){
    seed <- as.numeric(readline("Initial seed for the simulation? "))
    set.seed(seed) M <- as.numeric(readline("Number of
    replications? "))
    k <- as.numeric(readline("Number of groups? "))
    kn <- numeric(k) for (i in 1:k)
    kn[i] <- as.numeric(readline(paste("Number of observations
    in group",i,"? ")))
    local <- numeric(k) for (i in 1:k)
    local[i] <- as.numeric(readline(paste("Location parameter in
    the accelerated failure time model for group",i,"? ")))
    scale <- as.numeric(readline("Scale parameter in the
    accelerated failure time model? "))
    rho <- as.numeric(readline("Parameter in the
    Harrington&Fleming distribution of the error? "))
    p <- as.numeric(readline("Percentage of attendance to the
    scheduled visits? "))
    dput(list(seed,M,k,kn,local,scale,rho,p),"input")
    n <- sum(kn)
    code <-rep(3,n)
    data <- array(0,c(2,n,M))
    aux1 <- cumsum(c(1,kn[1:(k-1)]))
    aux2 <- cumsum(kn) z <- numeric(n)
    for (i in 1:k){
        data[,aux1[i]:aux2[i],] <- array(t(
        gendata(M*kn[i],local[i],scale,rho,p)),c(2,kn[i],M))
```

```
        z[aux1[i]:aux2[i]] <- i
    }
    pvalue0 <- 0
    pvalue1 <- 0
    pvalue2 <- 0
    pvalue3 <- 0
    rho.boot <<- rho
    pvalue4 <- 0
    pvalue5 <- 0
    pvalue6 <- 0
    pvalue7 <- 0
    pvalue8 <- 0
    pvalue9 <- 0
    pvalue10 <- 0
    pvalue11 <- 0
    pvalue12 <- 0
    pvalue13 <- 0
    for (i in 1:M){
        l <- data[1,,i]
        r <- data[2,,i]
        svf<- kaplanMeier(censor(l,r,code)~1,data=
        data.frame(l,r,code),se.fit=F, conf.interval="none")
        cdf <- cdfdata(l,r,svf)
        pvalue0 <- pvalue0+(pnorm(JK(cdf,z,0),0,1)>0.95)
        pvalue1 <- pvalue1+(pnorm(JK(cdf,z,0.5),0,1)>0.95)
        pvalue2 <- pvalue2+(pnorm(JK(cdf,z,1),0,1)>0.95)
        pvalue3 <- pvalue3+(pnorm(JK(cdf,z,1.5),0,1)>0.95)
        left.boot <<- data[1,,i]
        right.boot <<- data[2,,i]
```

```
    seedb <- .Random.seed
    boot1 <- bootstrap(z,J1boot(left.boot,right.boot,z,rho.boot),
    sampler=samp.permute,block.size=1000)
    pvalue4 <- pvalue4+(mean(boot1$replicates>=boot1$observed)
    <0.05)
    boot2 <- bootstrap(z,J2boot(left.boot,right.boot,z,rho.boot),
    sampler=samp.permute,seed=seedb,block.size=1000)
    pvalue5 <- pvalue5+(mean(boot2$replicates>=boot2$observed)
    <0.05)
    boot3 <- bootstrap(z,J3boot(left.boot,right.boot,z,rho.boot),
    sampler=samp.permute,seed=seedb,block.size=1000)
    pvalue6 <- pvalue6+(mean(boot3$replicates>=boot3$observed)
    <0.05)
    pvalue7 <- pvalue7+(pnorm(Ltrend(HFsc(cdf,0),z),0,1)>0.95)
    pvalue8 <- pvalue8+(pnorm(Ltrend(HFsc(cdf,0.5),z),0,1)>0.95)
    pvalue9 <- pvalue9+(pnorm(Ltrend(HFsc(cdf,1),z),0,1)>0.95)
    pvalue10 <- pvalue10+(pnorm(Ltrend(HFsc(cdf,1.5),z),0,1)>0.95)
    pvalue11 <- pvalue11+(pnorm(JKDiM(DiMsc(l,r,svf),z),0,1)>0.95)
    pvalue12 <- pvalue12+(pnorm(JA(l,r,z),0,1)>0.95)
    pvalue13 <- pvalue13+(pnorm(JA2(l,r,z),0,1)>0.95)
    print(i)
}
rm(left.boot)
rm(right.boot)
rm(rho.boot)
dput(c("JK0:",pvalue0/M,"JK0.5:",pvalue1/M,"JK1:",pvalue2/M,
"JK1.5:",pvalue3/M,"J1boot:",pvalue4/M,"J2boot:",pvalue5/M,
"J3boot:",pvalue6/M,"L0:",pvalue7/M,"L0.5:",pvalue8/M,
```

```
    "L1:",pvalue9/M,"L1.5:",pvalue10/M,"JKDiM:",pvalue11/M,
    "JA:",pvalue12/M,"JA2:",pvalue13/M),"output")
    c(pvalue0,pvalue1,pvalue2,pvalue3,pvalue4,pvalue5,pvalue6,
    pvalue7,pvalue8,pvalue9,pvalue10,pvalue11,pvalue12,pvalue13)/M
}
```

H Tables of simulation results

H.1 Percentage of rejection under ordered alternatives

$$(\lambda, \rho) = (1, 0)$$

$(\theta_1, \theta_2, \theta_3) = (1.6, 1.85, 2.1)$

(n_1, n_2, n_3)	Seed	JK^0	L^0_{trend}	J^0_1	J^0_2	J^0_3	JK_{DiM}
$(50, 50, 50)$	1227	0.656	0.656	0.66	0.638	0.638	0.642
$(100, 100, 100)$	1804	0.882	0.886	0.876	0.878	0.856	0.854
$(50, 100, 150)$	128	0.812	0.816	0.818	0.822	0.796	0.776
$(150, 100, 50)$	2895	0.86	0.884	0.888	0.854	0.876	0.838

$JK^{0.5}$	$L^{0.5}_{trend}$	JK^1	L^1_{trend}	JA	$JA2$	$JK^{1.5}$	$L^{1.5}_{trend}$
0.656	0.656	0.636	0.642	0.644	0.646	0.602	0.6
0.872	0.872	0.84	0.84	0.842	0.84	0.818	0.816
0.804	0.818	0.778	0.79	0.778	0.778	0.75	0.76
0.842	0.864	0.83	0.848	0.836	0.838	0.8	0.818

$(\theta_1, \theta_2, \theta_3) = (1.6, 2.1, 2.1)$

(n_1, n_2, n_3)	Seed	JK^0	L^0_{trend}	J^0_1	J^0_2	J^0_3	JK_{DiM}
$(50, 50, 50)$	235	0.636	0.636	0.664	0.71	0.518	0.618
$(100, 100, 100)$	279	0.886	0.886	0.90	0.936	0.812	0.88

$JK^{0.5}$	$L^{0.5}_{trend}$	JK^1	L^1_{trend}	JA	$JA2$	$JK^{1.5}$	$L^{1.5}_{trend}$
0.622	0.626	0.6	0.596	0.606	0.606	0.584	0.58
0.888	0.886	0.874	0.876	0.874	0.874	0.852	0.85

$(\theta_1, \theta_2, \theta_3) = (1.6, 1.6, 2.1)$

(n_1, n_2, n_3)	Seed	JK^0	L^0_{trend}	J^0_1	J^0_2	J^0_3	JK_{DiM}
$(50, 50, 50)$	1017	0.672	0.67	0.646	0.56	0.734	0.654
$(100, 100, 100)$	325	0.892	0.892	0.882	0.82	0.93	0.888

$JK^{0.5}$	$L^{0.5}_{trend}$	JK^1	L^1_{trend}	JA	$JA2$	$JK^{1.5}$	$L^{1.5}_{trend}$
0.664	0.668	0.62	0.622	0.638	0.64	0.592	0.586
0.89	0.888	0.88	0.874	0.88	0.882	0.844	0.84

$(\theta_1, \theta_2, \theta_3, \theta_4) = (1.65, 1.8, 1.95, 2.1)$

(n_1, n_2, n_3, n_4)	Seed	JK^0	L^0_{trend}	J^0_1	J^0_2	J^0_3	JK_{DiM}
$(50, 50, 50, 50)$	2575	0.642	0.642	0.628	0.61	0.64	0.644
$(100, 100, 100, 100)$	1955	0.86	0.858	0.86	0.844	0.84	0.84
$(40, 80, 120, 160)$	1109	0.786	0.81	0.81	0.794	0.762	0.756
$(160, 120, 80, 40)$	2525	0.81	0.828	0.826	0.778	0.814	0.774

$JK^{0.5}$	$L^{0.5}_{trend}$	JK^1	L^1_{trend}	JA	$JA2$	$JK^{1.5}$	$L^{1.5}_{trend}$
0.654	0.652	0.632	0.636	0.634	0.634	0.602	0.604
0.854	0.854	0.83	0.83	0.834	0.834	0.8	0.802
0.78	0.792	0.75	0.774	0.758	0.758	0.712	0.734
0.796	0.814	0.772	0.788	0.768	0.768	0.724	0.742

$(\theta_1, \theta_2, \theta_3, \theta_4) = (1.65, 2.1, 2.1, 2.1)$

(n_1, n_2, n_3, n_4)	Seed	JK^0	L^0_{trend}	J^0_1	J^0_2	J^0_3	JK_{DiM}
$(50, 50, 50, 50)$	8	0.522	0.518	0.552	0.632	0.386	0.508
$(100, 100, 100, 100)$	593	0.776	0.776	0.788	0.896	0.606	0.758

$JK^{0.5}$	$L^{0.5}_{trend}$	JK^1	L^1_{trend}	JA	$JA2$	$JK^{1.5}$	$L^{1.5}_{trend}$
0.518	0.52	0.516	0.51	0.506	0.506	0.496	0.496
0.768	0.772	0.76	0.764	0.756	0.754	0.746	0.746

$(\theta_1, \theta_2, \theta_3, \theta_4) = (1.65, 1.65, 1.65, 2.1)$

(n_1, n_2, n_3, n_4)	Seed	JK^0	L^0_{trend}	J^0_1	J^0_2	J^0_3	JK_{DiM}
$(50, 50, 50, 50)$	2038	0.544	0.542	0.486	0.382	0.676	0.508
$(100, 100, 100, 100)$	2213	0.842	0.84	0.802	0.644	0.902	0.792

$JK^{0.5}$	$L^{0.5}_{trend}$	JK^1	L^1_{trend}	JA	$JA2$	$JK^{1.5}$	$L^{1.5}_{trend}$
0.526	0.524	0.494	0.496	0.492	0.498	0.458	0.458
0.802	0.802	0.778	0.778	0.782	0.784	0.744	0.748

$$(\lambda, \rho) = (0.75, 0.5)$$

$(\theta_1, \theta_2, \theta_3) = (1.25, 1.5, 1.75)$

(n_1, n_2, n_3)	Seed	$JK^{0.5}$	$L^{0.5}_{trend}$	$J^{0.5}_1$	$J^{0.5}_2$	$J^{0.5}_3$	JK_{DiM}
(50, 50, 50)	27	0.682	0.682	0.686	0.688	0.676	0.678
(100, 100, 100)	297	0.928	0.926	0.926	0.914	0.922	0.928
(50, 100, 150)	1846	0.896	0.902	0.892	0.89	0.886	0.894
(150, 100, 50)	2965	0.882	0.87	0.866	0.874	0.864	0.888

JK^0	L^0_{trend}	JK^1	L^1_{trend}	JA	$JA2$	$JK^{1.5}$	$L^{1.5}_{trend}$
0.662	0.66	0.674	0.672	0.682	0.682	0.67	0.672
0.912	0.91	0.928	0.926	0.922	0.924	0.916	0.914
0.876	0.89	0.894	0.898	0.892	0.892	0.872	0.87
0.884	0.878	0.874	0.88	0.88	0.882	0.858	0.864

$(\theta_1, \theta_2, \theta_3) = (1.25, 1.75, 1.75)$

(n_1, n_2, n_3)	Seed	$JK^{0.5}$	$L^{0.5}_{trend}$	$J^{0.5}_1$	$J^{0.5}_2$	$J^{0.5}_3$	JK_{DiM}
(50, 50, 50)	110	0.718	0.718	0.728	0.786	0.594	0.704
(100, 100, 100)	1739	0.924	0.924	0.928	0.956	0.848	0.918

JK^0	L^0_{trend}	JK^1	L^1_{trend}	JA	$JA2$	$JK^{1.5}$	$L^{1.5}_{trend}$
0.694	0.696	0.726	0.722	0.71	0.71	0.712	0.714
0.912	0.91	0.914	0.908	0.912	0.912	0.902	0.902

$(\theta_1, \theta_2, \theta_3) = (1.25, 1.25, 1.75)$

(n_1, n_2, n_3)	Seed	$JK^{0.5}$	$L^{0.5}_{trend}$	$J^{0.5}_1$	$J^{0.5}_2$	$J^{0.5}_3$	JK_{DiM}
(50, 50, 50)	2411	0.724	0.726	0.72	0.63	0.786	0.738
(100, 100, 100)	2566	0.938	0.938	0.934	0.87	0.96	0.932

JK^0	L^0_{trend}	JK^1	L^1_{trend}	JA	$JA2$	$JK^{1.5}$	$L^{1.5}_{trend}$
0.732	0.732	0.704	0.708	0.72	0.722	0.662	0.668
0.936	0.934	0.924	0.926	0.934	0.934	0.918	0.916

$(\theta_1, \theta_2, \theta_3, \theta_4) = (1.3, 1.45, 1.6, 1.75)$

(n_1, n_2, n_3, n_4)	Seed	$JK^{0.5}$	$L^{0.5}_{trend}$	$J^{0.5}_1$	$J^{0.5}_2$	$J^{0.5}_3$	JK_{DiM}
$(50, 50, 50, 50)$	174	0.696	0.692	0.69	0.68	0.652	0.678
$(100, 100, 100, 100)$	1550	0.902	0.904	0.902	0.884	0.884	0.906
$(40, 80, 120, 160)$	2658	0.84	0.846	0.844	0.834	0.826	0.846
$(160, 120, 80, 40)$	93	0.826	0.842	0.828	0.788	0.828	0.826

JK^0	L^0_{trend}	JK^1	L^1_{trend}	JA	$JA2$	$JK^{1.5}$	$L^{1.5}_{trend}$
0.678	0.672	0.686	0.69	0.682	0.686	0.674	0.676
0.892	0.892	0.892	0.894	0.9	0.9	0.882	0.882
0.842	0.85	0.826	0.832	0.836	0.84	0.806	0.808
0.804	0.834	0.818	0.826	0.822	0.828	0.796	0.812

$(\theta_1, \theta_2, \theta_3, \theta_4) = (1.3, 1.75, 1.75, 1.75)$

(n_1, n_2, n_3, n_4)	Seed	$JK^{0.5}$	$L^{0.5}_{trend}$	$J^{0.5}_1$	$J^{0.5}_2$	$J^{0.5}_3$	JK_{DiM}
$(50, 50, 50, 50)$	1430	0.566	0.57	0.59	0.678	0.434	0.566
$(100, 100, 100, 100)$	2287	0.864	0.868	0.876	0.94	0.73	0.874

JK^0	L^0_{trend}	JK^1	L^1_{trend}	JA	$JA2$	$JK^{1.5}$	$L^{1.5}_{trend}$
0.554	0.556	0.572	0.578	0.572	0.572	0.558	0.554
0.848	0.848	0.878	0.878	0.88	0.878	0.872	0.872

$(\theta_1, \theta_2, \theta_3, \theta_4) = (1.3, 1.3, 1.3, 1.75)$

(n_1, n_2, n_3, n_4)	Seed	$JK^{0.5}$	$L^{0.5}_{trend}$	$J^{0.5}_1$	$J^{0.5}_2$	$J^{0.5}_3$	JK_{DiM}
$(50, 50, 50, 50)$	1655	0.63	0.634	0.604	0.468	0.718	0.634
$(100, 100, 100, 100)$	2172	0.84	0.84	0.828	0.674	0.946	0.84

JK^0	L^0_{trend}	JK^1	L^1_{trend}	JA	$JA2$	$JK^{1.5}$	$L^{1.5}_{trend}$
0.636	0.634	0.604	0.602	0.606	0.612	0.572	0.572
0.832	0.834	0.824	0.826	0.838	0.842	0.798	0.794

$(\lambda, \rho) = (0.6, 1)$

$(\theta_1, \theta_2, \theta_3) = (0.9, 1.15, 1.4)$

(n_1, n_2, n_3)	Seed	JK^1	L^1_{trend}	J^1_1	J^1_2	J^1_3	JK_{DiM}
$(50, 50, 50)$	2276	0.736	0.742	0.73	0.724	0.738	0.724
$(100, 100, 100)$	2634	0.94	0.94	0.946	0.938	0.922	0.926
$(50, 100, 150)$	1870	0.872	0.872	0.868	0.86	0.852	0.846
$(150, 100, 50)$	1993	0.888	0.888	0.884	0.872	0.878	0.854

JK^0	L^0_{trend}	$JK^{0.5}$	$L^{0.5}_{trend}$	JA	$JA2$	$JK^{1.5}$	$L^{1.5}_{trend}$
0.692	0.694	0.74	0.742	0.738	0.74	0.736	0.724
0.912	0.912	0.932	0.932	0.946	0.946	0.94	0.936
0.81	0.808	0.846	0.852	0.856	0.856	0.874	0.878
0.838	0.846	0.872	0.884	0.878	0.878	0.886	0.894

$(\theta_1, \theta_2, \theta_3) = (0.9, 1.4, 1.4)$

(n_1, n_2, n_3)	Seed	JK^1	L^1_{trend}	J^1_1	J^1_2	J^1_3	JK_{DiM}
$(50, 50, 50)$	428	0.746	0.74	0.746	0.798	0.634	0.722
$(100, 100, 100)$	869	0.928	0.93	0.926	0.966	0.828	0.914

JK^0	L^0_{trend}	$JK^{0.5}$	$L^{0.5}_{trend}$	JA	$JA2$	$JK^{1.5}$	$L^{1.5}_{trend}$
0.682	0.682	0.73	0.728	0.74	0.738	0.744	0.744
0.886	0.886	0.916	0.916	0.924	0.924	0.93	0.93

$(\theta_1, \theta_2, \theta_3) = (0.9, 0.9, 1.4)$

(n_1, n_2, n_3)	Seed	JK^1	L^1_{trend}	J^1_1	J^1_2	J^1_3	JK_{DiM}
$(50, 50, 50)$	1319	0.708	0.712	0.706	0.618	0.754	0.71
$(100, 100, 100)$	2214	0.956	0.96	0.954	0.864	0.974	0.942

JK^0	L^0_{trend}	$JK^{0.5}$	$L^{0.5}_{trend}$	JA	$JA2$	$JK^{1.5}$	$L^{1.5}_{trend}$
0.678	0.678	0.72	0.722	0.72	0.72	0.7	0.692
0.932	0.93	0.944	0.944	0.952	0.952	0.948	0.946

$(\theta_1, \theta_2, \theta_3, \theta_4) = (0.95, 1.1, 1.25, 1.4)$

(n_1, n_2, n_3, n_4)	Seed	JK^1	L^1_{trend}	J^1_1	J^1_2	J^1_3	JK_{DiM}
(50, 50, 50, 50)	2800	0.698	0.7	0.696	0.678	0.68	0.686
(100, 100, 100, 100)	1791	0.938	0.938	0.94	0.91	0.912	0.908
(40, 80, 120, 160)	671	0.86	0.854	0.846	0.842	0.824	0.83
(160, 120, 80, 40)	2483	0.844	0.846	0.832	0.814	0.832	0.822

JK^0	L^0_{trend}	$JK^{0.5}$	$L^{0.5}_{trend}$	JA	$JA2$	$JK^{1.5}$	$L^{1.5}_{trend}$
0.658	0.656	0.688	0.686	0.7	0.702	0.686	0.688
0.878	0.876	0.916	0.912	0.922	0.928	0.94	0.938
0.798	0.802	0.83	0.846	0.844	0.844	0.846	0.854
0.796	0.804	0.834	0.84	0.832	0.834	0.84	0.85

$(\theta_1, \theta_2, \theta_3, \theta_4) = (0.95, 1.4, 1.4, 1.4)$

(n_1, n_2, n_3, n_4)	Seed	JK^1	L^1_{trend}	J^1_1	J^1_2	J^1_3	JK_{DiM}
(50, 50, 50, 50)	837	0.638	0.634	0.63	0.738	0.45	0.608
(100, 100, 100, 100)	1110	0.864	0.864	0.854	0.94	0.704	0.838

JK^0	L^0_{trend}	$JK^{0.5}$	$L^{0.5}_{trend}$	JA	$JA2$	$JK^{1.5}$	$L^{1.5}_{trend}$
0.566	0.566	0.612	0.608	0.628	0.626	0.632	0.634
0.804	0.806	0.84	0.842	0.866	0.866	0.86	0.864

$(\theta_1, \theta_2, \theta_3, \theta_4) = (0.95, 0.95, 0.95, 1.4)$

(n_1, n_2, n_3, n_4)	Seed	JK^1	L^1_{trend}	J^1_1	J^1_2	J^1_3	JK_{DiM}
(50, 50, 50, 50)	1783	0.596	0.6	0.596	0.466	0.706	0.596
(100, 100, 100, 100)	137	0.862	0.862	0.858	0.698	0.942	0.87

JK^0	L^0_{trend}	$JK^{0.5}$	$L^{0.5}_{trend}$	JA	$JA2$	$JK^{1.5}$	$L^{1.5}_{trend}$
0.574	0.572	0.6	0.6	0.596	0.606	0.578	0.58
0.85	0.852	0.872	0.87	0.864	0.866	0.84	0.84

$(\lambda, \rho) = (0.5, 1.5)$

$(\theta_1, \theta_2, \theta_3) = (0.5, 0.75, 1)$

(n_1, n_2, n_3)	Seed	$JK^{1.5}$	$L^{1.5}_{trend}$	$J^{1.5}_1$	$J^{1.5}_2$	$J^{1.5}_3$	JK_{DiM}
(50, 50, 50)	2728	0.68	0.676	0.676	0.674	0.654	0.604
(100, 100, 100)	1159	0.926	0.926	0.922	0.92	0.93	0.876
(50, 100, 150)	2574	0.884	0.886	0.884	0.876	0.858	0.782
(150, 100, 50)	396	0.872	0.884	0.874	0.85	0.872	0.79

JK^0	L^0_{trend}	$JK^{0.5}$	$L^{0.5}_{trend}$	JK^1	L^1_{trend}	JA	$JA2$
0.592	0.594	0.652	0.656	0.68	0.676	0.666	0.668
0.862	0.86	0.906	0.906	0.926	0.928	0.916	0.918
0.76	0.764	0.82	0.832	0.872	0.878	0.846	0.848
0.792	0.802	0.852	0.87	0.874	0.878	0.854	0.854

$(\theta_1, \theta_2, \theta_3) = (0.5, 1, 1)$

(n_1, n_2, n_3)	Seed	$JK^{1.5}$	$L^{1.5}_{trend}$	$J^{1.5}_1$	$J^{1.5}_2$	$J^{1.5}_3$	JK_{DiM}
(50, 50, 50)	244	0.732	0.736	0.724	0.782	0.604	0.62
(100, 100, 100)	1202	0.912	0.914	0.912	0.952	0.85	0.85

JK^0	L^0_{trend}	$JK^{0.5}$	$L^{0.5}_{trend}$	JK^1	L^1_{trend}	JA	$JA2$
0.602	0.602	0.688	0.692	0.726	0.726	0.69	0.692
0.83	0.832	0.888	0.888	0.908	0.908	0.898	0.896

$(\theta_1, \theta_2, \theta_3) = (0.5, 0.5, 1)$

(n_1, n_2, n_3)	Seed	$JK^{1.5}$	$L^{1.5}_{trend}$	$J^{1.5}_1$	$J^{1.5}_2$	$J^{1.5}_3$	JK_{DiM}
(50, 50, 50)	489	0.722	0.72	0.728	0.62	0.764	0.672
(100, 100, 100)	2532	0.914	0.912	0.918	0.844	0.95	0.878

JK^0	L^0_{trend}	$JK^{0.5}$	$L^{0.5}_{trend}$	JK^1	L^1_{trend}	JA	$JA2$
0.658	0.658	0.708	0.708	0.734	0.732	0.72	0.728
0.876	0.876	0.91	0.908	0.91	0.908	0.908	0.908

$(\theta_1, \theta_2, \theta_3, \theta_4) = (0.55, 0.7, 0.85, 1)$

(n_1, n_2, n_3, n_4)	Seed	$JK^{1.5}$	$L^{1.5}_{trend}$	$J^{1.5}_1$	$J^{1.5}_2$	$J^{1.5}_3$	JK_{DiM}
(50, 50, 50, 50)	2712	0.666	0.664	0.662	0.65	0.652	0.618
(100, 100, 100, 100)	1858	0.928	0.928	0.926	0.906	0.918	0.866
(40, 80, 120, 160)	2135	0.856	0.86	0.85	0.84	0.82	0.746
(160, 120, 80, 40)	310	0.85	0.85	0.842	0.822	0.842	0.746

JK^0	L^0_{trend}	$JK^{0.5}$	$L^{0.5}_{trend}$	JK^1	L^1_{trend}	JA	$JA2$
0.604	0.602	0.65	0.65	0.668	0.668	0.658	0.66
0.848	0.848	0.91	0.912	0.928	0.926	0.922	0.922
0.732	0.736	0.806	0.806	0.842	0.852	0.812	0.814
0.738	0.752	0.81	0.816	0.844	0.842	0.816	0.814

$(\theta_1, \theta_2, \theta_3, \theta_4) = (0.55, 1, 1, 1)$

(n_1, n_2, n_3, n_4)	Seed	$JK^{1.5}$	$L^{1.5}_{trend}$	$J^{1.5}_1$	$J^{1.5}_2$	$J^{1.5}_3$	JK_{DiM}
(50, 50, 50, 50)	1297	0.592	0.596	0.59	0.696	0.436	0.478
(100, 100, 100, 100)	53	0.85	0.848	0.842	0.924	0.692	0.776

JK^0	L^0_{trend}	$JK^{0.5}$	$L^{0.5}_{trend}$	JK^1	L^1_{trend}	JA	$JA2$
0.454	0.45	0.536	0.536	0.568	0.57	0.546	0.55
0.722	0.722	0.818	0.818	0.844	0.846	0.83	0.83

$(\theta_1, \theta_2, \theta_3, \theta_4) = (0.55, 0.55, 0.55, 1)$

(n_1, n_2, n_3, n_4)	Seed	$JK^{1.5}$	$L^{1.5}_{trend}$	$J^{1.5}_1$	$J^{1.5}_2$	$J^{1.5}_3$	JK_{DiM}
(50, 50, 50, 50)	451	0.588	0.588	0.598	0.466	0.704	0.552
(100, 100, 100, 100)	1462	0.848	0.85	0.85	0.708	0.914	0.804

JK^0	L^0_{trend}	$JK^{0.5}$	$L^{0.5}_{trend}$	JK^1	L^1_{trend}	JA	$JA2$
0.548	0.544	0.594	0.594	0.592	0.598	0.596	0.606
0.794	0.792	0.848	0.85	0.858	0.858	0.858	0.862

H.2 Percentage of rejection under the null hypothesis

$$(\lambda, \rho) = (1, 0)$$

$(\theta_1, \theta_2, \theta_3) = (1.6, 1.6, 1.6)$

(n_1, n_2, n_3)	Seed	JK^0	L^0_{trend}	J^0_1	J^0_2	J^0_3	JK_{DiM}
$(50, 50, 50)$	232	0.044	0.044	0.046	0.044	0.036	0.044
$(100, 100, 100)$	2808	0.054	0.054	0.052	0.054	0.054	0.054

$JK^{0.5}$	$L^{0.5}_{trend}$	JK^1	L^1_{trend}	JA	$JA2$	$JK^{1.5}$	$L^{1.5}_{trend}$
0.048	0.048	0.048	0.048	0.048	0.048	0.046	0.046
0.052	0.052	0.05	0.05	0.056	0.058	0.054	0.056

$(\theta_1, \theta_2, \theta_3) = (2.1, 2.1, 2.1)$

(n_1, n_2, n_3)	Seed	JK^0	L^0_{trend}	J^0_1	J^0_2	J^0_3	JK_{DiM}
$(50, 50, 50)$	2222	0.036	0.036	0.038	0.034	0.044	0.04
$(100, 100, 100)$	1132	0.048	0.048	0.052	0.05	0.054	0.04

$JK^{0.5}$	$L^{0.5}_{trend}$	JK^1	L^1_{trend}	JA	$JA2$	$JK^{1.5}$	$L^{1.5}_{trend}$
0.042	0.042	0.04	0.04	0.04	0.04	0.038	0.04
0.046	0.046	0.046	0.046	0.042	0.042	0.046	0.048

$(\theta_1, \theta_2, \theta_3, \theta_4) = (1.65, 1.65, 1.65, 1.65)$

(n_1, n_2, n_3, n_4)	Seed	JK^0	L^0_{trend}	J^0_1	J^0_2	J^0_3	JK_{DiM}
$(50, 50, 50, 50)$	208	0.032	0.034	0.036	0.03	0.028	0.036
$(100, 100, 100, 100)$	2148	0.052	0.052	0.052	0.058	0.046	0.052

$JK^{0.5}$	$L^{0.5}_{trend}$	JK^1	L^1_{trend}	JA	$JA2$	$JK^{1.5}$	$L^{1.5}_{trend}$
0.042	0.044	0.042	0.042	0.034	0.034	0.048	0.048
0.05	0.05	0.042	0.044	0.048	0.048	0.052	0.05

$(\theta_1, \theta_2, \theta_3, \theta_4) = (2.1, 2.1, 2.1, 2.1)$

(n_1, n_2, n_3, n_4)	Seed	JK^0	L^0_{trend}	J^0_1	J^0_2	J^0_3	JK_{DiM}
$(50, 50, 50, 50)$	2997	0.05	0.052	0.054	0.048	0.044	0.048
$(100, 100, 100, 100)$	1957	0.058	0.058	0.054	0.054	0.054	0.056

$JK^{0.5}$	$L^{0.5}_{trend}$	JK^1	L^1_{trend}	JA	$JA2$	$JK^{1.5}$	$L^{1.5}_{trend}$
0.05	0.05	0.05	0.05	0.052	0.05	0.05	0.052
0.05	0.052	0.052	0.054	0.052	0.052	0.06	0.058

$(\lambda, \rho) = (0.75, 0.5)$

$(\theta_1, \theta_2, \theta_3) = (1.25, 1.25, 1.25)$

(n_1, n_2, n_3)	Seed	$JK^{0.5}$	$L^{0.5}_{trend}$	$J^{0.5}_1$	$J^{0.5}_2$	$J^{0.5}_3$	JK_{DiM}
$(50, 50, 50)$	2279	0.04	0.04	0.042	0.046	0.046	0.042
$(100, 100, 100)$	1707	0.054	0.054	0.058	0.056	0.052	0.058

JK^0	L^0_{trend}	JK^1	L^1_{trend}	JA	$JA2$	$JK^{1.5}$	$L^{1.5}_{trend}$
0.044	0.042	0.044	0.044	0.044	0.044	0.04	0.04
0.06	0.06	0.048	0.05	0.05	0.052	0.048	0.05

$(\theta_1, \theta_2, \theta_3, \theta_4) = (1.3, 1.3, 1.3, 1.3)$

(n_1, n_2, n_3, n_4)	Seed	$JK^{0.5}$	$L^{0.5}_{trend}$	$J^{0.5}_1$	$J^{0.5}_2$	$J^{0.5}_3$	JK_{DiM}
$(50, 50, 50, 50)$	2186	0.054	0.056	0.054	0.056	0.06	0.056
$(100, 100, 100, 100)$	166	0.058	0.058	0.054	0.05	0.058	0.052

JK^0	L^0_{trend}	JK^1	L^1_{trend}	JA	$JA2$	$JK^{1.5}$	$L^{1.5}_{trend}$
0.058	0.064	0.05	0.052	0.052	0.052	0.056	0.056
0.058	0.058	0.058	0.058	0.058	0.058	0.058	0.058

$(\theta_1, \theta_2, \theta_3) = (1.75, 1.75, 1.75)$

(n_1, n_2, n_3)	Seed	$JK^{0.5}$	$L^{0.5}_{trend}$	$J^{0.5}_1$	$J^{0.5}_2$	$J^{0.5}_3$	JK_{DiM}
(50, 50, 50)	2877	0.046	0.046	0.04	0.04	0.04	0.04
(100, 100, 100)	236	0.048	0.048	0.054	0.052	0.05	0.05

JK^0	L^0_{trend}	JK^1	L^1_{trend}	JA	$JA2$	$JK^{1.5}$	$L^{1.5}_{trend}$
0.048	0.048	0.036	0.038	0.034	0.034	0.034	0.034
0.052	0.052	0.044	0.044	0.048	0.048	0.04	0.04

$(\theta_1, \theta_2, \theta_3, \theta_4) = (1.75, 1.75, 1.75, 1.75)$

(n_1, n_2, n_3, n_4)	Seed	$JK^{0.5}$	$L^{0.5}_{trend}$	$J^{0.5}_1$	$J^{0.5}_2$	$J^{0.5}_3$	JK_{DiM}
(50, 50, 50, 50)	1149	0.052	0.05	0.044	0.054	0.048	0.058
(100, 100, 100, 100)	1109	0.042	0.044	0.042	0.042	0.046	0.048

JK^0	L^0_{trend}	JK^1	L^1_{trend}	JA	$JA2$	$JK^{1.5}$	$L^{1.5}_{trend}$
0.05	0.048	0.058	0.056	0.062	0.062	0.056	0.054
0.044	0.044	0.048	0.048	0.046	0.046	0.044	0.044

$(\lambda, \rho) = (0.6, 1)$

$(\theta_1, \theta_2, \theta_3) = (0.9, 0.9, 0.9)$

(n_1, n_2, n_3)	Seed	JK^1	L^1_{trend}	J^1_1	J^1_2	J^1_3	JK_{DiM}
(50, 50, 50)	1353	0.064	0.064	0.064	0.066	0.056	0.056
(100, 100, 100)	994	0.052	0.052	0.048	0.048	0.044	0.046

JK^0	L^0_{trend}	$JK^{0.5}$	$L^{0.5}_{trend}$	JA	$JA2$	$JK^{1.5}$	$L^{1.5}_{trend}$
0.05	0.052	0.058	0.056	0.06	0.06	0.06	0.062
0.042	0.04	0.048	0.046	0.05	0.05	0.052	0.052

$(\theta_1, \theta_2, \theta_3, \theta_4) = (0.95, 0.95, 0.95, 0.95)$

(n_1, n_2, n_3, n_4)	Seed	JK^1	L^1_{trend}	J^1_1	J^1_2	J^1_3	JK_{DiM}
$(50, 50, 50, 50)$	2462	0.044	0.042	0.042	0.046	0.04	0.034
$(100, 100, 100, 100)$	1265	0.052	0.052	0.052	0.042	0.054	0.036

JK^0	L^0_{trend}	$JK^{0.5}$	$L^{0.5}_{trend}$	JA	$JA2$	$JK^{1.5}$	$L^{1.5}_{trend}$
0.036	0.034	0.032	0.032	0.034	0.036	0.05	0.048
0.036	0.036	0.04	0.038	0.046	0.048	0.048	0.048

$(\theta_1, \theta_2, \theta_3) = (1.4, 1.4, 1.4)$

(n_1, n_2, n_3)	Seed	JK^1	L^1_{trend}	J^1_1	J^1_2	J^1_3	JK_{DiM}
$(50, 50, 50)$	1901	0.032	0.034	0.036	0.038	0.03	0.028
$(100, 100, 100)$	835	0.03	0.03	0.03	0.036	0.028	0.032

JK^0	L^0_{trend}	$JK^{0.5}$	$L^{0.5}_{trend}$	JA	$JA2$	$JK^{1.5}$	$L^{1.5}_{trend}$
0.03	0.028	0.026	0.026	0.034	0.034	0.038	0.04
0.036	0.036	0.028	0.028	0.03	0.03	0.028	0.028

$(\theta_1, \theta_2, \theta_3, \theta_4) = (1.4, 1.4, 1.4, 1.4)$

(n_1, n_2, n_3, n_4)	Seed	JK^1	L^1_{trend}	J^1_1	J^1_2	J^1_3	JK_{DiM}
$(50, 50, 50, 50)$	217	0.048	0.048	0.054	0.054	0.046	0.044
$(100, 100, 100, 100)$	2364	0.048	0.046	0.05	0.048	0.048	0.052

JK^0	L^0_{trend}	$JK^{0.5}$	$L^{0.5}_{trend}$	JA	$JA2$	$JK^{1.5}$	$L^{1.5}_{trend}$
0.044	0.044	0.046	0.046	0.044	0.044	0.048	0.05
0.05	0.048	0.054	0.052	0.054	0.054	0.048	0.048

$$(\lambda, \rho) = (0.5, 1.5)$$

$(\theta_1, \theta_2, \theta_3) = (0.5, 0.5, 0.5)$

(n_1, n_2, n_3)	Seed	$JK^{1.5}$	$L^{1.5}_{trend}$	$J^{1.5}_1$	$J^{1.5}_2$	$J^{1.5}_3$	JK_{DiM}
(50, 50, 50)	170	0.054	0.054	0.05	0.058	0.06	0.05
(100, 100, 100)	933	0.046	0.046	0.05	0.054	0.046	0.078

JK^0	L^0_{trend}	$JK^{0.5}$	$L^{0.5}_{trend}$	JK^1	L^1_{trend}	JA	$JA2$
0.052	0.05	0.052	0.054	0.048	0.052	0.052	0.054
0.078	0.078	0.066	0.064	0.05	0.05	0.058	0.058

$(\theta_1, \theta_2, \theta_3, \theta_4) = (0.55, 0.55, 0.55, 0.55)$

(n_1, n_2, n_3, n_4)	Seed	$JK^{1.5}$	$L^{1.5}_{trend}$	$J^{1.5}_1$	$J^{1.5}_2$	$J^{1.5}_3$	JK_{DiM}
(50, 50, 50, 50)	2646	0.038	0.038	0.042	0.048	0.04	0.04
(100, 100, 100, 100)	2218	0.05	0.048	0.048	0.05	0.056	0.052

JK^0	L^0_{trend}	$JK^{0.5}$	$L^{0.5}_{trend}$	JK^1	L^1_{trend}	JA	$JA2$
0.042	0.044	0.036	0.038	0.044	0.044	0.04	0.042
0.056	0.056	0.05	0.05	0.046	0.048	0.05	0.052

$(\theta_1, \theta_2, \theta_3) = (1, 1, 1)$

(n_1, n_2, n_3)	Seed	$JK^{1.5}$	$L^{1.5}_{trend}$	$J^{1.5}_1$	$J^{1.5}_2$	$J^{1.5}_3$	JK_{DiM}
(50, 50, 50)	1286	0.034	0.034	0.04	0.028	0.04	0.04
(100, 100, 100)	376	0.054	0.052	0.056	0.05	0.046	0.054

JK^0	L^0_{trend}	$JK^{0.5}$	$L^{0.5}_{trend}$	JK^1	L^1_{trend}	JA	$JA2$
0.048	0.048	0.036	0.036	0.036	0.034	0.034	0.036
0.054	0.054	0.056	0.056	0.048	0.048	0.052	0.052

$(\theta_1, \theta_2, \theta_3, \theta_4) = (1, 1, 1, 1)$

(n_1, n_2, n_3, n_4)	Seed	$JK^{1.5}$	$L^{1.5}_{trend}$	$J^{1.5}_1$	$J^{1.5}_2$	$J^{1.5}_3$	JK_{DiM}
$(50, 50, 50, 50)$	1305	0.06	0.06	0.054	0.058	0.054	0.058
$(100, 100, 100, 100)$	2007	0.044	0.044	0.046	0.044	0.044	0.048

JK^0	L^0_{trend}	$JK^{0.5}$	$L^{0.5}_{trend}$	JK^1	L^1_{trend}	JA	$JA2$
0.054	0.054	0.058	0.06	0.064	0.066	0.06	0.06
0.05	0.05	0.048	0.05	0.048	0.048	0.05	0.05