

Stationarity in a prevalent cohort study with follow-up

Vittorio Addona

Department of Mathematics and Statistics,
McGill University,
805 rue Sherbrooke Ouest,
Montréal, Québec, Canada H3A 2K6

December 30, 2005.

A thesis submitted to McGill University in partial fulfillment of the
requirements of the degree of Doctor of Philosophy.

© Vittorio Addona 2005

Abstract

In a prevalent cohort study with follow-up, the incidence process is not directly observed since only the onset times of prevalent cases can be ascertained. Several important consequences follow if one can establish stationarity of the incidence process: (1)The useful epidemiological relationship between prevalence, incidence, and mean duration holds, (2)There is improved efficiency when estimating the underlying survivor function from a prevalent cohort study with follow-up, (3)The constancy of the incidence rate is established, and (4)The constant incidence rate can be estimated using data from a prevalent cohort study.

We propose a formal test for stationarity using data from a prevalent cohort study with follow-up, and establish new characterizations of stationarity, and of useful types of departure from stationarity.

A dual to the problem of establishing stationarity by comparing the backward and forward recurrence times is addressed. Assuming stationarity of the underlying incidence process, we use the backward and forward recurrence times to verify whether the underlying survival distribution is independent of the date of onset. In doing so, we characterize specific types of dependence of the underlying survival distribution on calendar time.

If the data are consistent with stationarity of the incidence rate, then a natural next step is to estimate the (constant) incidence rate. We derive the nonparametric maximum likelihood estimator of the constant incidence rate, prove that the estimator is weakly consistent, and show how one may construct an asymptotic confidence interval for the incidence rate. One main advantage of our procedure is that it only requires the completion of a single prevalent cohort study with follow-up.

We apply our test for stationarity to data obtained as part of the Canadian Study of Health and Aging to verify that the incidence rate of dementia amongst the elderly in Canada has remained constant. Upon concluding that this constancy is plausible, we estimate the incidence rate.

Résumé

Dans une étude de cohorte prévalente avec suivi, le processus d'incidence n'est pas observé directement car seuls les temps de début des cas prévalents peuvent être déterminés. Si nous pouvons établir la stationnarité du processus d'incidence, nous tirons profit de plusieurs conséquences: (1) La relation utile entre la prévalence, l'incidence, et la durée moyenne est valide, (2) Il y a un gain d'efficacité dans l'estimation de la fonction de survie sous-jacente, (3) La constance du taux d'incidence est établie, et (4) Le taux d'incidence constant peut être estimé à l'aide des données provenant d'une étude de cohorte prévalente.

Nous proposons un test formel pour vérifier si le processus d'incidence est stationnaire dans le cas où les données proviennent d'une étude de cohorte prévalente avec suivi, et nous établissons de nouvelles caractérisations d'un processus stationnaire et de certains types d'écarts à la stationnarité.

Nous adressons la dualité de l'établissement de la stationnarité en comparant les temps "vers l'arrière" et "vers l'avant". En supposant que le processus d'incidence est stationnaire, nous utilisons les temps "vers l'arrière" et "vers l'avant" pour vérifier si la fonction de survie est indépendante de la date de début. Ainsi, nous caractérisons certains types de dépendance entre la fonction de survie et la date de début.

Si les données sont consistantes à l'hypothèse d'un processus d'incidence stationnaire, la prochaine étape est d'estimer le taux d'incidence constant. Nous dérivons l'estimateur du maximum de vraisemblance non paramétrique du taux d'incidence constant, prouvons que cet estimateur converge en probabilité, et développons un intervalle de confiance asymptotique pour le taux d'incidence. L'avantage principal de notre méthode est qu'elle requiert la

complétion d'une seule étude de cohorte prévalente avec suivi.

Nous appliquons notre test d'établissement de stationnarité sur des données provenant de l'étude sur la santé et le vieillissement au Canada afin de vérifier la constance du taux d'incidence de démence parmi les aînés au Canada. Après avoir conclu que la constance du taux d'incidence est plausible, nous estimons le taux d'incidence.

Statement of originality

The original contributions contained in this thesis are as follows:

1. A formal test for stationarity of the incidence process, that relies on the characterization of stationarity given by Asgharian et al. (2004) and Wei's (1980) test for matched pairs of right censored data.
2. Theorem 3 from Chapter 3, which states that when the pairs arise as backward and forward recurrence times from a prevalent cohort study with follow-up, equality of the marginal distributions implies bivariate symmetry of the joint distribution.
3. Theorem 4 and Theorem 8 from Chapter 3 and Chapter 4, respectively, which provide two new characterizations of stationarity.
4. Theorem 6 from Chapter 4, which states that if the incidence intensity, $\lambda(t)$, is non-decreasing (non-increasing) and non-constant, then the forward recurrence time distribution is stochastically larger (smaller) than the backward recurrence time distribution.
5. The counterexample from Section 4.2 which shows that the converse to Theorem 6 does not hold without further restrictions, and Theorem 7 from Chapter 4, which establishes the converse to Theorem 6 when we restrict $\lambda(t)$ to the class of monotone intensity functions.
6. An alternative proof of the characterization of stationarity given by Asgharian et al. (2004), arising from (4.8).
7. Lemma 1 from Chapter 4, which gives a sufficient condition for stochastic ordering of two random variables.

8. A formal test of the assumption that the underlying survival distribution is independent of calendar time of onset.
9. Lemma 2 from Chapter 6, which provides expressions for the backward and forward recurrence time densities when the underlying survival distribution depends on calendar time of onset.
10. Theorem 9 and Theorem 10 from Chapter 6, which characterize independence of the survival distribution on calendar time of onset by equality of the backward and forward recurrence time distributions.
11. Theorem 11 from Chapter 6, which states that survival is “improving” (“worsening”) over time if and only if the forward recurrence time distribution is stochastically larger (smaller) than the backward recurrence time distribution.
12. A point estimator, $\hat{\lambda}$, for the true incidence rate, λ , based on the equation “prevalence = incidence \times mean duration”, and using data from a single prevalent cohort study with follow-up.
13. Corollary 2 from Chapter 8, which states that the asymptotic normality of the estimator $\hat{\mu}$ holds when the number of cases is random.
14. Corollary 3 from Chapter 8, which states that $\hat{\lambda}$ is weakly consistent for λ , and Theorem 12, which states that $\hat{\lambda}$ is the NPMLE for λ .
15. Lemma 4 from Chapter 8, which establishes the consistency of $\hat{\mu}$ for the true mean survival time, μ .
16. A confidence interval for λ , displayed in (8.12), that is derived from the asymptotic distribution of $\frac{1}{\lambda}$.

Acknowledgements

I would like to begin by extending my deepest gratitude to my supervisor, and friend, Dr. David Wolfson. This work would surely not have been possible without his support. Not only did he provide a wealth of ideas which led to results that are in this thesis, but he helped me improve my writing style immensely through his meticulous reading of my work. David, I may not have always been easy to work with, but I assure you that I appreciate every moment that you have taken out of your days to answer my questions, correct my mistakes, or encourage me at moments when I was feeling most down. It may not be ample compensation for all that time and energy you spent on me, but thank you!

To the love of my life, Isadora, from whom I have been apart for most of the writing process, I would like to say that this year has been long and difficult at times, but through our frequent conversations, you cheered me up the way only you can. Soon after I submit this thesis, I will be able to be with you again, and to hold you in my arms, and that is the best reward I could possibly be given.

To my brothers, Frank and Tony, and to my sister Daniela, I want you all to know how much I love you. You three, more than anyone else on the planet, have shaped me into the person I am today. I hope that you know that I will always be there for you, any time, any place. I cannot express my love for my mother in words. You have made so many sacrifices for us, that I could never thank you enough. I can only hope to be one tenth the parent to my children that you have been to us.

Contents

1	Introduction and literature review	1
1.1	General Medical Setting	1
1.2	Stationarity and estimation of the underlying survival distribution	6
1.3	Why is stationarity important?	8
1.4	Why is the assumption of a single underlying survival distribution important?	9
1.5	Verification of stationarity and of the assumption of a single survival distribution	9
1.6	Estimation of a constant incidence rate when stationarity holds	10
1.7	Overview	11
2	Notation and preliminaries	15
2.1	The setup	15
2.2	The censoring mechanism	17
2.3	Nonparametric survival estimation with left truncated, right censored data	19
2.3.1	The conditional approach	20
2.3.2	The unconditional approach	21

2.3.3	Comparison of the conditional approach and the un- conditional approach	22
3	Testing the validity of the stationarity assumption	23
3.1	Lack of formal methods to verify stationarity	23
3.1.1	Uncertainty in onset times	24
3.1.2	Informal assessments of stationarity	25
3.2	Main objective	25
3.3	Characterization of stationarity	26
3.4	Wei's test for censored, paired data	27
3.4.1	Description of Wei's procedure	27
3.5	Equivalence of (3.1) and (3.2)	30
3.6	Another characterization of stationarity	32
4	Characterizations of departures from stationarity	35
4.1	Characterization of an important alternative hypothesis	36
4.2	Restricting the class of intensities	41
4.3	One more characterization of stationarity	47
5	Detecting departure from stationarity: a power study	50
5.1	Details of the simulations	51
5.1.1	Underlying survival distribution	51
5.1.2	Sample size	52
5.1.3	Residual censoring time distribution	52

5.1.4	Onset intensity and support of the truncation time distribution	52
5.1.5	Number of replications of the procedure	54
5.2	Results of the power study	55
5.3	Discussion of the results	56
6	Verifying the assumption of a fixed survival distribution	59
6.1	Characterizing the independence of survival and calendar time of onset	60
6.1.1	The case of $K = 2$	66
6.1.2	The case of $K > 2$	69
6.2	Characterizing alternatives to a fixed survival distribution for $K = 2$	72
6.3	Utility of results	79
7	Detecting dependence of survival on calendar time: a power study	80
7.1	Details of the simulations	81
7.1.1	Underlying survival distributions	81
7.1.2	Sample size	82
7.1.3	Value of \tilde{x}	82
7.1.4	Residual censoring time distribution	83
7.1.5	Onset intensity and support of the truncation time distribution	83
7.1.6	Number of replications of the procedure	84
7.2	Results of the power study	84
7.3	Discussion of the results	86

8	Estimation of a constant incidence rate	88
8.1	Background and notation	88
8.2	A simple point estimator of λ	94
8.3	An interval estimator of λ	98
8.4	Estimation of ψ_N	100
8.5	Synopsis	101
9	Application to the CSHA	102
9.1	Overview of the CSHA	102
9.2	Testing the assumption of stationarity of the incidence rate of dementia	104
9.3	Estimating the incidence rate of dementia	108
9.4	Comparison of estimates	110
10	Concluding remarks and future directions for research	113

List of Tables

5.1	Percentage of rejections for Weibull($\gamma=2, \beta=10$)	55
5.2	Percentage of rejections for Weibull($\gamma=0.75, \beta=1.5$)	55
5.3	Percentage of rejections for lognormal($\mu=1.75, \sigma=0.4$)	55
7.1	Percentage of rejections for Weibull($\gamma=2, \beta_1=10$) with $\tilde{x} = 2$	84
7.2	Percentage of rejections for Weibull($\gamma=2, \beta_1=10$) with $\tilde{x} = 3$	85
7.3	Percentage of rejections for lognormal($\mu_1=1.75, \sigma=0.4$) with $\tilde{x} = 5$	85
7.4	Percentage of rejections for lognormal($\mu_1=1.75, \sigma=0.4$) with $\tilde{x} = 7$	85

List of Figures

1.1	Incident cohort study	3
1.2	Prevalent cohort study with follow-up	3
1.3	A biased sample obtained through cross-sectional sampling	5
2.1	Basic setup of a prevalent cohort study with follow-up	16
2.2	Backward and forward recurrence times	17
2.3	Censoring setup	18
4.1	Intuition behind Theorem 6	40
4.2	Densities satisfying a sufficient condition for stochastic ordering	42
4.3	Graph of $\lambda(t)$	43
4.4	Graph of $h(x)$	46
6.1	Survival depends on calendar time, but $Y^{fwd} >^{SL} Y^{bwd}$	61
6.2	Assume the change in survival can occur at $\tau^* - \tilde{x}$	67
6.3	Setup with $K > 2$ distinct survivor functions	69
6.4	Example with 4 distinct density functions	70
6.5	Relationship between $f_1(x)$ and $f_2(x)$ assuming (2) holds	74
9.1	Conditional (below) and Unconditional (above) NPMLE of $S(x)$	105

9.2	Kaplan-Meier estimates of the backward (solid) and forward (dotted) recurrence time survivor functions	106
9.3	Histogram of 10,000 bootstrap estimates of μ	109

Chapter 1

Introduction and literature review

1.1 General Medical Setting

Researchers are often interested in studying the length of time between some initiating event and some terminating event. For example, one may want to make inference about the length of remission duration for acute leukemia (Freireich et al., 1963). In this example, the initiating event occurs when the patient has a treatment induced remission of their leukemia, and the terminating event occurs when the patient has a relapse of leukemia. As other examples, consider the survival time of breast cancer or kidney transplant patients, where the initiating events are, respectively, the time of diagnosis of breast cancer and the time of transplant, and the terminating event in both cases is death (Sedmak et al., 1989, and Klein and Moeschberger, 1997). It is common, by convention only, to refer to the occurrence of the terminating event as “failure”. Throughout this thesis we will often assume

a medical setting since survival analysis has, perhaps, most of its applications in the medical field. There is, however, nothing particular about this setting, and the methodology carries over to a wide variety of non-medical situations (see, for example, Lancaster, 1990).

Knowledge of the natural history of a disease is often important in ascertaining the cause and finding a cure, treatment, or factors that affect survival. Survival analysis is usually one of the main statistical tools used when studying the natural history of diseases. We shall restrict ourselves to the consideration of survival from onset of a certain disease until death or some other end point of interest. For instance, a researcher may want to investigate the survival of dementia patients from onset until death (Wolfson et al., 2001). The data commonly collected for a natural-history-of-disease study, anticipating a survival analysis, consist of observing survival times in one of two ways. One way is by assembling a cohort of individuals and following them forward until some of these subjects acquire the disease under study. These incident cases are then followed for a further fixed time period until failure or censoring. This is termed an *incident cohort study*. An illustration of an incident cohort study is given in Figure 1.1.

Assembling an adequately large cohort of subjects, to ensure a reasonable number of occurrences of the disease, and following them for potentially long periods of time, to ensure that a substantial number of these cases have progressed to failure, is sometimes expensive and impractical. Thus, prevalent cases, instead of incident cases, of the disease are often recruited, over a short time period. Prevalent cases are so called because individuals who already have the disease of interest are identified. The sampling of prevalent cases is sometimes referred to as cross-sectional sampling.

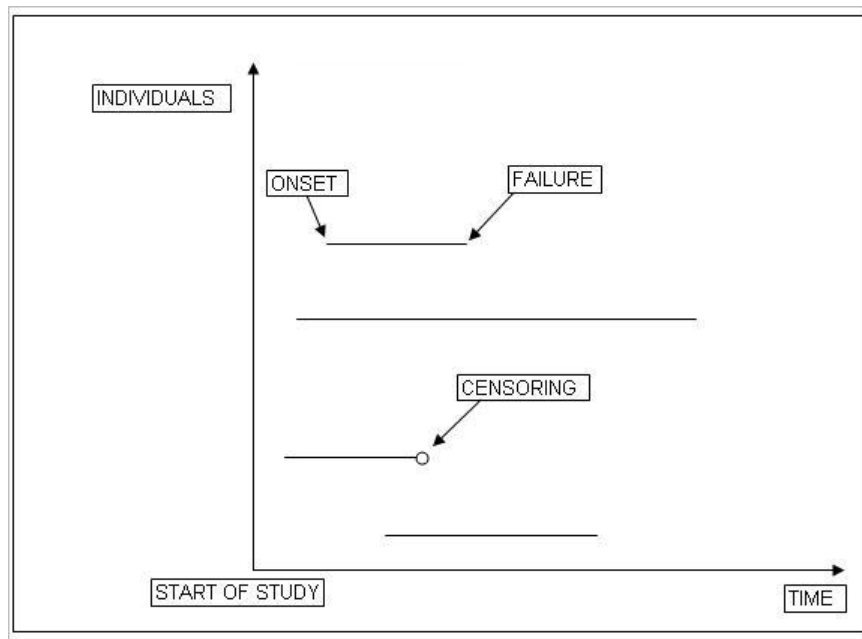


Figure 1.1: Incident cohort study

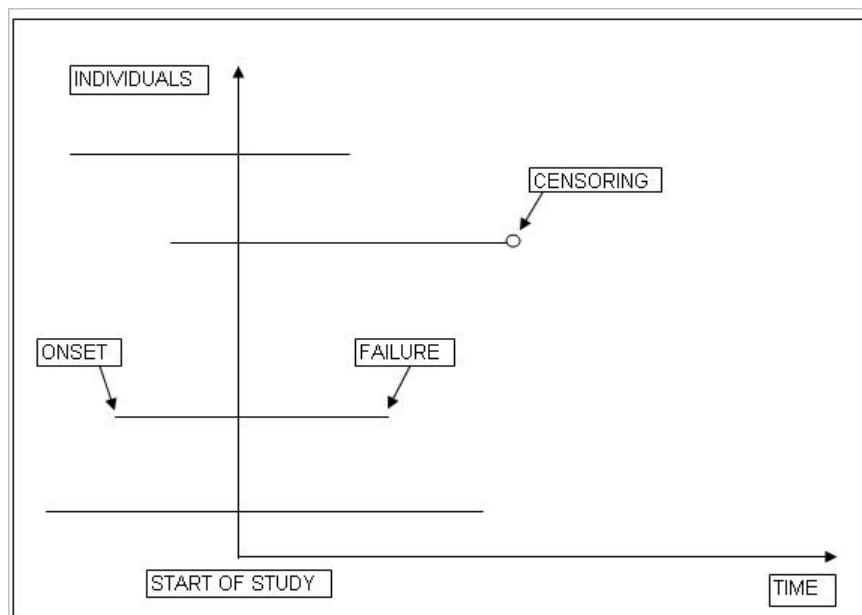


Figure 1.2: Prevalent cohort study with follow-up

The type of data collected following the identification of prevalent cases can vary from study to study. It might be possible to ascertain the onset times of each prevalent case, and then to follow these subjects for a fixed study period until failure or censoring, recording their, possibly censored, survival times. We denote these type of data *prevalent cohort survival times* or *prevalent cohort data*, and the studies that give rise to them *prevalent cohort studies with follow-up* (see Figure 1.2). The *truncation time* is defined to be the time from onset of the disease until the start of the study, even for those subjects who experience the terminating event before recruitment, and hence cannot be observed. For those subjects who survive long enough to be included in the study, we refer to their observed truncation time as the *backward recurrence time*. We denote the time from recruitment until failure as the, possibly censored, *forward recurrence time*.

Keiding (1991) discussed three other types of data that can be observed after the ascertainment of prevalent cases. The most common method for recruiting prevalent cases is through the screening of a group of individuals, of which a certain proportion are diagnosed with the condition. Two of the types of data discussed in Keiding's paper assume there is information on the "non-diseased" subjects, and that the diseased subjects are not followed after recruitment. The other type of data are prevalent cohort survival times, except without follow-up of subjects after recruitment, that is, only recording their backward recurrence times. In fact, none of the types of data discussed by Keiding (1991) include follow-up of the subjects. In this thesis, we concentrate primarily on prevalent cohort survival times, apart from Chapter 8 where, in addition, information on the "non-diseased" individuals is exploited. Therefore, we shall always assume that there is follow-up.

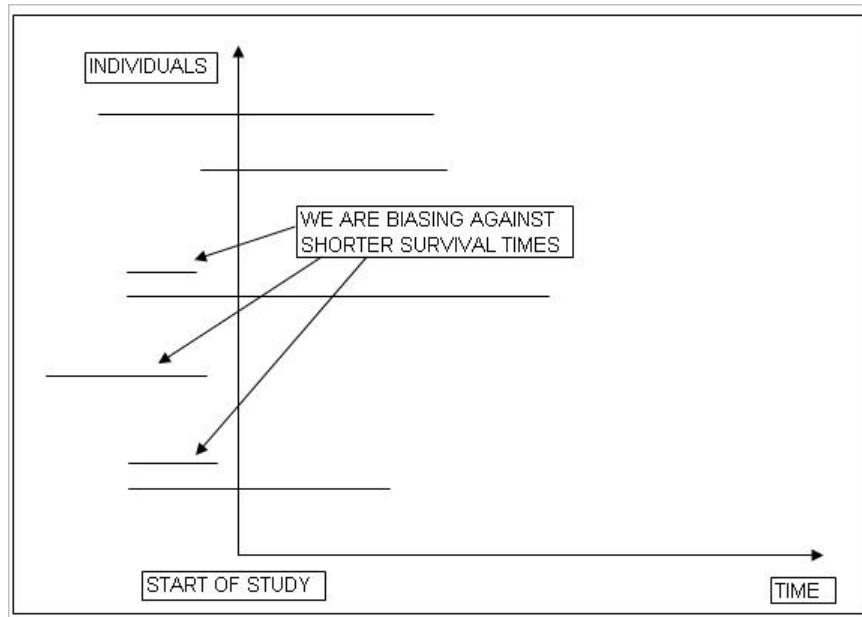


Figure 1.3: A biased sample obtained through cross-sectional sampling

Cross-sectional sampling alleviates the main practical difficulties associated with the sampling of incident cases. However, this more convenient method of sampling results in observations which are not a random sample from the target “incident” population. In particular, it is well known that cases obtained through a cross-sectional sampling scheme tend to have longer survival than cases obtained in an incident cohort study (McFadden, 1962, Blumenthal, 1967, and Cox, 1969). As we see in Figure 1.3, the sampling of prevalent cases inherently biases against shorter survival times since subjects must survive until recruitment in order to be observed. As a result, in addition to being potentially right censored, prevalent cohort survival times are subject to left truncation.

1.2 Stationarity and estimation of the underlying survival distribution

An important feature of any prevalent cohort study is that the incidence process of onset times is only partially observed, since only the onset times of prevalent cases may be ascertained. We refer to the fully observed incidence process as the underlying incidence process to differentiate it from the process of partially observed onset times of the prevalent cases. When the underlying incidence process is a stationary Poisson process, we refer simply to *stationarity*. Informally, stationarity corresponds to the situation that the incidence rate of the disease in the general population is roughly constant over time. Stationarity would not be a reasonable assumption if, for example, there was an epidemic of the disease before the start of the study. Under stationarity, data subject to left truncation are said to be *length-biased*. Stationarity will be central in this thesis and a more formal definition is given in Chapter 2.

A large portion of the statistical literature studying the natural history of a disease has focused on estimating the survival distribution for an incident case of the disease using prevalent cohort survival times. An inherent, but often tacit, assumption made in the nonparametric survival analysis literature is that there exists a single underlying survival distribution regardless of when a subject had onset of the disease. To estimate this underlying survival distribution, some authors carry out a so called “conditional” analysis, which conditions on the set of backward recurrence times in the sample (Turnbull, 1976, Wang et al., 1986, Lagakos et al., 1988, Tsai et al., 1987, and Wang, 1991). Other authors rely instead on the assumption

of stationarity and thus perform a so called “unconditional” analysis (Vardi, 1982, 1985, 1989, Gill et al., 1988, Asgharian et al., 2002, and Asgharian and Wolfson, 2005).

One proposed estimator, which we refer to as the *censoring-truncation product limit estimator* (censoring-truncation p.l.e.), was introduced by Cox and Oakes (1984). The censoring-truncation p.l.e. is analogous to the usual p.l.e. under right censoring, except with modified risk sets. The risk set at an observed failure time includes only subjects who have not failed or been censored, and who are under active follow-up. Wang et al. (1986) and Tsai et al. (1987) investigated the large sample properties of the censoring-truncation p.l.e. Wang (1987) claimed that the censoring-truncation p.l.e. is the nonparametric maximum likelihood estimator (NPMLE) of the full likelihood. The censoring-truncation p.l.e. had been derived by maximizing a function which was neither a full nor a conditional likelihood. Adopting a conditional analysis, Wang (1991) showed that this function is a conditional likelihood, and hence that the censoring-truncation p.l.e. is also the conditional NPMLE, when all the censoring times are known, even for subjects who have been observed until failure.

Assuming stationarity, Vardi (1989) showed how the unconditional NPMLE of the survival distribution associated with a “length-biased” case may be obtained via the EM algorithm. A transformation of this estimator gives the unconditional NPMLE of the underlying survival distribution for an incident case. Asgharian et al. (2002) derived the asymptotic properties of the unconditional NPMLE of the underlying survival distribution. We return to the issue of conditional analysis versus unconditional analysis in Chapter 2. For the moment, we discuss the role of stationarity in various

research problems, having already noted its role in the estimation of the survival distribution through an unconditional analysis.

1.3 Why is stationarity important?

Whether or not stationarity of the underlying incidence process holds in any particular situation is important for at least four reasons:

1. The well known epidemiological relationship, “prevalence = incidence \times mean duration”, depends, in part, on the assumption of stationarity. By building a three-state illness-death process, Keiding (1991) makes explicit all the assumptions needed for this relationship to hold.
2. When estimating the underlying survivor function from a prevalent cohort study with follow-up, there is a gain in efficiency if one can assume stationarity (Asgharian et al., 2002, and Wang, 1991).
3. A goal of a study might be to ascertain whether the underlying incidence rate of a disease, say, is constant over time, that is, to determine whether stationarity holds.
4. Assuming stationarity, the incidence rate is constant over time, and can be estimated using data from a prevalent cohort study (Keiding, 1991, and Diamond and McDonald, 1991). It is not possible, without making restrictive assumptions, to estimate the incidence rate from a prevalent cohort study if the incidence rate is non-constant.

1.4 Why is the assumption of a single underlying survival distribution important?

The assumption that the underlying survival distribution is independent of the calendar time of onset is important for at least two reasons:

1. As we mentioned in Section 1.2, the assumption of a single survival distribution is made in the nonparametric survival analysis literature in order to estimate the underlying survival distribution.
2. An assessment of whether the survival distribution has changed over time may be an end in itself.

1.5 Verification of stationarity and of the assumption of a single survival distribution

Since there are several potential benefits to be gained if one can establish stationarity of the underlying incidence process, we outline the various approaches to the problem of ascertaining stationarity of the underlying incidence process. From a single prevalent cohort study without follow-up, it is impossible to ascertain stationarity, as is pointed out by Preston (1987). Preston added that data from at least two prevalent cohort studies without follow-up are needed to verify stationarity. Clearly, if incident cases are observed, then stationarity can be assessed since, in this case, we directly observe the underlying incidence process about which we wish to make inference.

In spite of its importance in several different areas of research, there is, to our knowledge, no formal method to verify stationarity of the underlying incidence process using data from a single prevalent cohort study with follow-up. Wang (1991) proposed a method for estimating the distribution of the truncation times, which is uniform under stationarity. Her method, however, provides no goodness-of-fit test to the uniform distribution; Kolmogorov-type tests rely on the assumption that the estimator of the distribution function is an empirical distribution function derived from a set of independent and identically distributed (i.i.d.) observations, while Wang's estimator is not such an estimator. Taking a different approach, Asgharian et al. (2004) showed that stationarity holds if and only if the backward and forward recurrence time distributions are identical. They suggested a graphical method, based on plots of Kaplan-Meier estimates of the backward and forward recurrence time distributions, to check the validity of stationarity. The authors did not, however, discuss formal methods to test for stationarity.

There seems to be no literature on assessing the assumption that the underlying survival distribution is independent of calendar time of onset.

1.6 Estimation of a constant incidence rate when stationarity holds

If stationarity is deemed to be plausible, researchers will no doubt wish to estimate the constant incidence rate. This is straightforward when incident cases are observed (see, for example, Rothman, 1986). Keiding (1991) developed methods for estimating a constant incidence rate using three types

of data obtained through a cross-sectional sampling scheme, none of which include follow-up of the subjects. In addition to Keiding’s (1991) paper, others have considered this problem when prevalent cases are ascertained with no follow-up of the subjects (see, for example, Diamond and McDonald, 1991). To our knowledge, there is no literature that addresses the estimation of a constant incidence rate when one has access to prevalent cohort survival times arising from a prevalent cohort study with follow-up.

1.7 Overview

Using the characterization from Asgharian et al. (2004), the problem of verifying the validity of stationarity can be restated as a test of the null hypothesis, $H_0 : F_{bwd}(x) \equiv F_{fwd}(x)$, where $F_{bwd}(x)$ and $F_{fwd}(x)$ represent the backward and forward recurrence time distribution functions, respectively. This null hypothesis is that the population ensemble of backward recurrence times has the same distribution as that of the ensemble of forward recurrence times. However, a two-sample test that allows for censoring, such as a logrank test (Mantel, 1966), is precluded since, although the pairs of backward and forward recurrence times between subjects are independent, the “within subject” backward and forward recurrence times are correlated. In a sense, the backward and forward recurrence times are matched pairs, with the forward recurrence time possibly randomly right censored. In Chapter 10, we take up the issue of an alternative two sample test based on the Kaplan-Meier estimators for the two samples.

Methods have been developed for testing certain hypotheses, including some procedures designed to test for equality in distribution, when the

data are potentially censored matched pairs data (Wei, 1980, Cheng, 1984, Woolson and Lachenbruch, 1980, Pettitt, 1983, O'Brien and Fleming, 1987, Albers, 1988, and Dabrowska, 1989, 1990). In much of this literature, the assumption is made that the within pair members are subject to the same censoring random variable. This is not a viable assumption when the pairs are backward and forward recurrence times since, in this case, only the forward recurrence times can be censored. We refer to this restrictive censoring assumption as the *equal censorship assumption*. Gehan (1965) and Gilbert (1962) independently generalized the Wilcoxon two-sample test to the case where the samples are subject to right censoring. Wei (1980) used the same scoring function as Gehan and Gilbert to construct an asymptotically distribution-free test when the data are vectors of paired observations where both components of a pair may be right censored. In his paper, Wei did not make use of the equal censorship assumption, and introduced a consistent estimator of the asymptotic variance of the test statistic. Jung (1999) proposed a method of estimating the asymptotic variance for general rank tests for paired data. It is unclear, however, whether Jung's procedure outperforms that of Wei, and we proceed by following Wei.

In Chapter 2, we outline the basic framework for the thesis, and formally define some key terms. In Chapter 3, we exploit the characterization of stationarity given by Asgharian et al. (2004) and show how Wei's (1980) test for matched pairs of right censored data may be adapted to give a formal test for stationarity of the underlying incidence process. In Chapter 4, we extend the result of Asgharian et al. (2004) by characterizing incidence rates that are monotone over calendar time. These characterizations help us to explore the power of our test for stationarity. In Chapter 5, we report

on the results of a power study of our test for stationarity.

With certain conditions, the assumption that the underlying survival distribution is independent of calendar time of onset may not be appropriate due to improvements in patient care, say, or possibly due to the introduction of a new treatment for the disease. In Chapter 6, we prove a dual to the characterization of stationarity given by Asgharian et al. (2004), that is, we show that, assuming stationarity of the incidence process of onsets, independence between the underlying survival distribution and calendar time of onset is equivalent to equality of the backward and forward recurrence time distributions. We extend this theorem by showing that survival is “improving” (“worsening”) over time if and only if the backward and forward recurrence time distributions are stochastically ordered. These results allow us to use the test for stationarity of the incidence process to also test for dependence of the underlying survival distribution on calendar time. Thus, if a researcher is not willing to, a priori, assume that survival is independent of calendar time of onset, the validity of this assumption can be verified. To our knowledge, no such assessment has been proposed to date. The results from another power study are presented in Chapter 7, where we examine the ability of our proposed test to detect a dependence between the survival distribution and calendar time of onset.

In Chapter 8, we propose a simple point estimator of a constant incidence rate that is based on the classical epidemiological equation “prevalence = incidence \times mean duration”. In doing so, we use prevalent cohort survival times and, importantly, we exploit the sampling scheme that often yields such data. When conditioning on the observed number of cases of the disease, as most authors do, estimation of the incidence rate is impossible

without further assumption (Keiding, 1991). We use the added information of knowing the fixed number screened that gave rise to the random number of cases, in order to help estimate the assumed constant incidence rate of the disease. Our point estimator is consistent for the true incidence rate. Moreover, we show that our estimator, proposed in an *ad hoc* fashion, represents the NPMLLE of the true incidence rate. We also develop an interval estimator for a constant incidence rate. One main advantage of the procedures developed in Chapter 8 is that they only require a single prevalent cohort study with follow-up, as opposed to the completion of an incident cohort study.

We apply our methods to data obtained as part of the Canadian Study of Health and Aging (CSHA) in Chapter 9. First, we verify the constancy of the incidence rate of dementia among the elderly in Canada, and, having done so, we estimate the constant incidence rate. Our estimates agree with previous estimates obtained from the CSHA.

Chapter 2

Notation and preliminaries

We begin by carefully describing the problem and presenting the notation which will be used throughout this thesis. As discussed in Chapter 1, prevalent cohort studies are useful since practical limitations, such as financial considerations and time, sometimes make it very difficult to identify and follow incident cases. In the following, we present the notation and discussion in the context of a medical study, although many of the results carry over to other settings. For example, prevalent cohort studies have been used to investigate employment patterns in the population (Lancaster, 1990).

2.1 The setup

Data from prevalent cohort studies with follow-up in the medical field arise as follows: Let X_1, X_2, \dots, X_m be m i.i.d. positive random variables representing the lifetimes of individuals from onset of a certain condition (or some other initiating event) to an end point of interest or terminating event. Let the X_i 's have survivor function $S(x) := P(X_i > x)$, with probability density

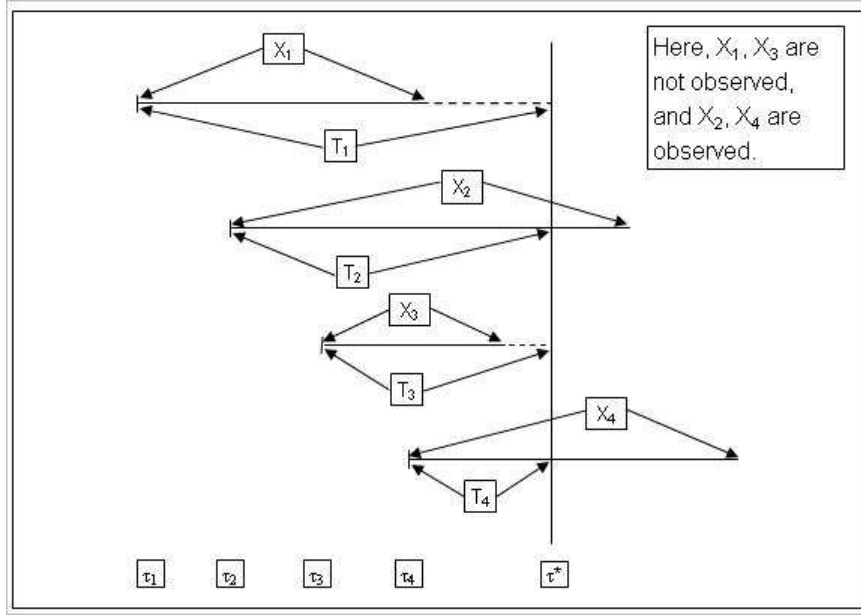


Figure 2.1: Basic setup of a prevalent cohort study with follow-up

function $f(x)$, with respect to Lebesgue measure. Let $\tau_1, \tau_2, \dots, \tau_m$ be the m calendar times of onset corresponding to X_1, X_2, \dots, X_m , and let τ^* be the calendar time of recruitment for the study. Individual i is observed in the study only if $X_i \geq \tau^* - \tau_i$, and we therefore say that X_i is left truncated, with left truncation time $T_i = \tau^* - \tau_i$ (see Figure 2.1). Since the onset times are random, the truncation times are random variables, with distribution function denoted by $G(t)$, and density $g(t)$, with respect to Lebesgue measure.

Let Y_1, Y_2, \dots, Y_n be the *observed* left truncated lifetimes, with $n \leq m$, recalling that not all subjects will survive long enough to be observed in the study. We write $Y_i = Y_i^{bwd} + Y_i^{fwd}$, where Y_i^{bwd} is the time from onset of the condition to recruitment into the study or the *backward recurrence time*, and Y_i^{fwd} is the time from recruitment to failure, or the *forward recurrence time* (see Figure 2.2). Let $Y_i^{bwd} \sim f_{bwd}$ and $Y_i^{fwd} \sim f_{fwd}$. We

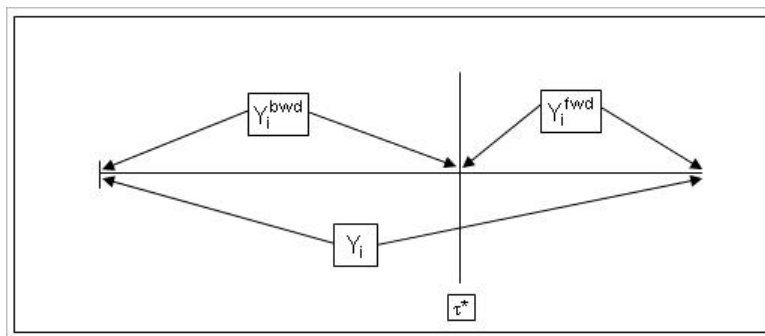


Figure 2.2: Backward and forward recurrence times

refer to $f_{bwd}(x)$ and $f_{fwd}(x)$, respectively, as the backward and forward recurrence time density. Observe that Y_i^{bwd} and Y_i^{fwd} are not independent since individuals with longer backward recurrence times will tend to have shorter forward recurrence times. That is, Y_i^{bwd} and Y_i^{fwd} are known to be negatively correlated.

2.2 The censoring mechanism

Frequently, in studies that include follow-up, not all of the subjects can be followed until the end point is reached. When an individual is not observed until failure he/she is said to be right censored, with a censoring time C corresponding to the last time the subject is observed.

In a prevalent cohort study with follow-up, suppose that individual i has censoring time $C_i^* = Y_i^{bwd} + C_i$, where C_i , which we call the *residual censoring time*, is the time from recruitment until the individual is censored (see Figure 2.3). We assume that $P(C_i^* > T_i) = 1$, and thus observe only $\min(C_i^*, Y_i)$. By definition, however, the backward recurrence times are always fully observed, and we shall suppose that we ob-

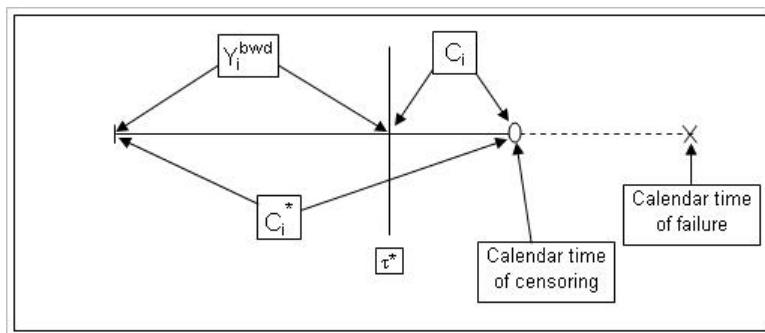


Figure 2.3: Censoring setup

serve $(Y_i^{bwd}, Y_i^{obs}, \epsilon_i), i = 1, 2, \dots, n$, where $Y_i^{obs} = \min(Y_i^{fwd}, C_i)$, and $\epsilon_i = \mathbf{1}[Y_i^{fwd} \leq C_i]$ indicates whether subject i has been followed until failure.

It is clear that C_i^* is not independent of Y_i^{bwd} . Furthermore, C_i^* and Y_i have Y_i^{bwd} in common, and we thus have *informative censoring* (Vardi, 1989). Roughly, informative censoring occurs when knowledge of the censoring time, C_i^* , provides information on the survival time, Y_i , in addition to simply knowing that $Y_i > C_i^*$. We still, however, assume that residual censoring time C_i is independent of both Y_i^{fwd} and Y_i^{bwd} . The assumed independence between C_i and Y_i^{bwd} seems reasonable in practice since, for example, knowing the magnitude of Y_i^{bwd} should not give any information on the magnitude of C_i . Furthermore, the assumed independence between C_i and Y_i^{fwd} corresponds to the usual random censoring assumption.

2.3 Nonparametric survival estimation with left truncated, right censored data

We present two approaches that have been used to obtain the NPMLE of $S(x)$, the common underlying survivor function of the data. First, we express the likelihood of the data in its most general form. Subsequently, we outline the two approaches, pointing out certain differences that exist between them. Finally, we provide a comparison of the two approaches.

The likelihood contribution of an individual who is observed until failure is proportional to: $f(y_i) \left(\int_{y_i}^{\infty} h(y_i^{bwd}, u) du \right)$ where $h(., .)$ represents the joint density function of the truncation time and the censoring time. The likelihood contribution of an individual who is censored before failure is proportional to: $S(c_i^*) h(y_i^{bwd}, c_i^*)$. The full likelihood for the observed data is thus proportional to,

$$L = \prod_{i=1}^n \left[\left(f(y_i) \int_{y_i}^{\infty} h(y_i^{bwd}, u) du \right)^{\epsilon_i} \left(S(c_i^*) h(y_i^{bwd}, c_i^*) \right)^{1-\epsilon_i} \right] \quad (2.1)$$

where the constant of proportionality is $\int_0^{\infty} S(u) g(u) du$, and represents the probability of being observed, or not being truncated, in the study.

Two general strategies have been utilized in the statistical literature to obtain the NPMLE of $S(x)$ from the likelihood displayed in (2.1). One is a “conditional” approach which has been studied by, amongst others, Wang (1991), Wang et al. (1993), Anderson et al. (1993), and Tsai et al. (1987). The other is an “unconditional” approach examined by authors including, Vardi (1982, 1985, 1989), and Vardi and Zhang (1992).

2.3.1 The conditional approach

The conditional approach is so called because we condition on the observed truncation times, that is, on the backward recurrence times in the sample. This allows us to perform the maximization of the likelihood without considering the incidence process which is assumed to generate cases of the medical condition.

First, note that by multiplying and dividing each factor of (2.1) by $S(y_i^{bwd})$, the full likelihood can be rewritten as a product of,

$$L_1 = \prod_{i=1}^n \frac{f(y_i)^{\epsilon_i} S(c_i^*)^{1-\epsilon_i}}{S(y_i^{bwd})}, \text{ and}$$

$$L_2 = \prod_{i=1}^n \left[\frac{S(y_i^{bwd}) h(y_i^{bwd}, c_i^*)^{1-\epsilon_i} (\int_{y_i}^{\infty} h(y_i^{bwd}, u) du)^{\epsilon_i}}{\int_0^{\infty} S(u) g(u) du} \right]$$

Wang (1991) maximized L_1 as a function of $S(x)$. L_1 can be viewed as the conditional likelihood of the data given the values of the backward recurrence times. In fact, L_1 is only a conditional likelihood if one assumes that all the censoring times are known, even for subjects who are observed to fail. This may, for instance, be the case if subjects could not be lost to follow-up, and could only be censored at the end of the study. In this situation, the residual censoring times, C_i , are fixed constants and, conditional on the backward recurrence times (or equivalently, on the observed onset times) we know all the censoring times even if some individuals fail before the study is terminated. The CSHA provides an example where very few patients were lost to follow-up, and almost all of the censoring was due to the termination of the study (CSHA working group, 1994, 2000).

2.3.2 The unconditional approach

When adopting an unconditional approach, we do not condition on the backward recurrence times in order to facilitate estimation of $S(x)$. Instead, the unconditional approach relies on the so called stationarity assumption described in Chapter 1. We first define an intensity function, and a stationary Poisson process, and follow with the definition of stationarity.

Definition 1 *An intensity function $\lambda(t)$ for a certain disease is the instantaneous probability of getting the disease at time t , given that one is non-diseased just before t . That is,*

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq \tau < t + \Delta t \mid \tau \geq t)}{\Delta t},$$

where τ represents the calendar time of onset of the disease.

Definition 2 *Let $N(t)$ be a Poisson process with intensity function $\lambda(t)$. We say that $N(t)$ is a stationary Poisson process if $\lambda(t) \equiv \lambda$, that is, if $\lambda(t)$ is constant.*

Definition 3 *Let $I(t)$ represent the incidence process of the onsets of some disease. Stationarity is said to hold if $I(t)$ is a stationary Poisson process.*

A relationship exists between the intensity function, $\lambda(t)$, and the truncation time density, $g(t)$. If the onsets, or initiation times, occur over an interval $(0, \tau^*)$, then an onset at calendar time $\tau^* - x$ for some $x \in (0, \tau^*)$ implies a truncation time of x . Since $g(x)$ is a density, we have that,

$$g(x) = \frac{\lambda(\tau^* - x)\mathbf{1}[0 < x < \tau^*]}{\Lambda(\tau^*)} \tag{2.2}$$

where $\Lambda(u) = \int_0^u \lambda(v)dv$ (see, for example, Ross, 2003).

Under stationarity, data subject to left truncation are referred to as length-biased, and hence “length-biasedness” is a special case of left truncation. Due to informative censoring, the NPMLE of the length-biased survivor function is not the Kaplan-Meier estimator. In turn, the NPMLE of $S(x)$ cannot be obtained by transforming the Kaplan-Meier estimator of the length-biased survivor function. The truncation time distribution is uniform under stationarity, conditional on the number of incident times in $(0, \tau^*)$. As a result, the likelihood is proportional to,

$$L_U = \prod_{i=1}^n \frac{f(y_i)^{\epsilon_i} S(c_i^*)^{1-\epsilon_i}}{\mu} ,$$

where $\mu = \int_0^\infty S(x)dx$. We maximize L_U with respect to $S(x)$ to obtain $\hat{S}(x)$, the unconditional NPMLE of $S(x)$ (Vardi, 1989).

2.3.3 Comparison of the conditional approach and the unconditional approach

A conditional approach becomes necessary if it is known that the stationarity assumption does not hold, because the model is otherwise non-identifiable due to the dependence of the data on both the truncation distribution and the underlying survival distribution (Wang et al., 1986). Although the conditional approach does not require any further assumption, it is less efficient than the unconditional approach when the stationarity assumption is, in fact, valid (Wang, 1991, and Asgharian et al., 2002). Another shortcoming of the conditional approach is that it may lead to an underestimation of survival for small x , as pointed out by Pan and Chappell (1998). As a general rule, we believe that one should use the unconditional approach whenever stationarity may be assumed.

Chapter 3

Testing the validity of the stationarity assumption

We saw in Section 1.3 that establishing stationarity of the underlying incidence process can be important for several reasons, including improved efficiency in the estimation of the underlying survivor function, and the ability to estimate the resulting constant incidence rate.

3.1 Lack of formal methods to verify stationarity

If incident cases of a disease are identified over some interval $(0, \tau^*)$, say, with onset times at $\tau_1, \tau_2, \dots, \tau_n$, an empirical distribution function (e.d.f.) using the truncation times $\tau_i - \tau^*$ for $i = 1, \dots, n$ can be used in order to assess stationarity. Under stationarity, the truncation time distribution is uniform, and thus the e.d.f. obtained from the observed truncation times can be compared against the uniform distribution using a goodness-of-fit test. If the

data is consistent with stationarity, we can estimate the constant incidence rate by computing the ratio of the observed number of cases to the total person-time observed (Rothman, 1986). Note that since the denominator in the ratio is person-time, this estimator is valid even when individuals are followed for varying amounts of time, or when some individuals are lost to follow-up before developing the condition.

3.1.1 Uncertainty in onset times

With prevalent cohort survival times, however, we ascertain the onset times of the cases retrospectively. Although this might be an unrealistic goal when the condition has an insidious onset, as is the case with dementia, there are circumstances where reliable information can be obtained on the onset dates. For example, the initiating event can be something as clear as calendar time of first stroke. Even with Alzheimer's disease, which has an insidious onset, rough onset dates may be obtained by interviewing caregivers, as was done in the CSHA (CSHA working group, 1994). One can never be absolutely certain of onset times obtained in this fashion, but there is no reason to believe that these dates will be systematically biased in one direction or another. It is important to realize that even with so called incident cases, the exact date of onset may be unclear since subjects are only screened periodically. Even with the added knowledge of the onset times of the prevalent cases, it is still difficult in practice to determine whether the stationarity assumption is reasonable because the cases being observed do not correspond to onsets which form a random sample of all the onsets of the disease.

3.1.2 Informal assessments of stationarity

Despite the importance of stationarity, but no doubt because of the above difficulty, little work has been done on formal methods to verify stationarity using prevalent cohort data. Wang (1991) proposed a method for estimating the truncation distribution function, $G(t)$. She even found the asymptotic distribution of her estimator, $\hat{G}(t)$, and uses the “obvious” method of bootstrapping to arrive at a 95% confidence band for $\hat{G}(t)$ in an example. Aside from inspection, however, there is currently no formal procedure for testing departure from the uniform distribution using her method. Kolmogorov-type tests rely on the fact that the estimator of the distribution function is derived from a set of i.i.d. observations. Unfortunately, $\hat{G}(t)$ is not obtained in such a fashion, and this eliminates the possibility for the direct application of such a test in this case. As discussed in Chapter 1, Asgharian et al. (2004) gave a characterization of stationarity and suggested an informal graphical method to check the validity of stationarity.

3.2 Main objective

A key observation is that the observed backward and forward recurrence times are, in essence, matched pairs, where one member of each pair, the forward recurrence time, is potentially right censored. Furthermore, there are existing procedures for testing certain hypotheses when the data are potentially censored matched pairs data. In this chapter, we will describe one of these procedures, and show how it is relevant, in conjunction with the characterization of Asgharian et al. (2004), in formally testing whether the stationarity assumption is reasonable.

3.3 Characterization of stationarity

Asgharian et al. (2004) showed that stationarity is, under mild assumptions, equivalent to equality of the backward and forward recurrence time distributions. In renewal theory, it is known that stationarity of the renewal process implies equality in distribution of the backward and forward recurrence time distributions (Karlin and Taylor, 1975). Asgharian et al. (2004) proved their result in a different setting which does not involve renewal processes. We now state their theorem.

Theorem 1 *Let X and T be, respectively, the true lifetime and true truncation time with corresponding density functions $f(x)$ and $g(t)$ (possibly improper), having the same support. Suppose X and T are independent. Let $\int_0^\infty g(t)S(t)dt < \infty$, where S is the distribution function of X , and suppose one of the following two conditions hold: (A) The truncation density g has an absolute maximum with unbounded support. (B) The truncation density g is continuous with bounded support. Then the truncating density, g , is constant if and only if $f_{bwd} \equiv f_{fwd}$.*

Thus stationarity is equivalent to $Y^{bwd} = Y^{fwd}$ in distribution, and a null hypothesis of stationarity can be restated as,

$$H_0 : F_{bwd}(x) = F_{fwd}(x) \quad \forall x \geq 0 . \quad (3.1)$$

With H_0 stated as in (3.1), one might initially consider a two-sample test, such as a logrank test. Now, because in the present setting the two samples are of backward and forward recurrence times, respectively, standard two-sample tests are inapplicable. However, Wei (1980), modified a two-sample

test to make it applicable when the data are matched pairs. We show how Wei's test for censored, paired data can be utilized to test H_0 in (3.1).

3.4 Wei's test for censored, paired data

Gehan (1965) and Gilbert (1962) separately generalized the Wilcoxon two-sample test to the case where the samples are subject to right censoring. This was done through the introduction of a generalized Wilcoxon scoring function. This scoring function is a natural generalization of the traditional Mann-Whitney scoring function in that it assigns non-zero values only to observed “ (X_i, Y_j) ” pairs where one member is *known* to be larger than the other (Wilcoxon, 1945, and Mann and Whitney 1947). Of course, in the presence of censoring, there may be some ambiguity about the ordering of two random variables and such pairs are assigned scores of zero. Wei (1980) used the same scoring function to construct an asymptotically distribution-free test for the hypothesis of bivariate symmetry when the observations are vectors of *paired* observations where both components of a pair may be right censored. The procedure is developed as follows:

3.4.1 Description of Wei's procedure

Suppose that $(X_1^0, Y_1^0), \dots, (X_n^0, Y_n^0)$ are i.i.d. random vectors with common joint distribution function $H^0(s, t)$. Let $F^0(s)$ and $G^0(t)$ be the marginal distribution functions of the X_i^0 's and Y_i^0 's, respectively. Assume that $H^0(s, t)$, and hence $F^0(s)$, and $G^0(t)$, are continuous distribution functions. The hypothesis of interest in Wei's paper is,

$$H_0 : H^0(s, t) = H^0(t, s) \quad \forall (s, t) \in \mathbb{R}^2 . \quad (3.2)$$

Alternative hypotheses which may be of interest are those indicating that the X_i^0 's tend to be larger than the Y_j^0 's. It is very important to note that bivariate symmetry of a joint distribution function implies equality of the marginal distributions, but, in general, the converse is not true. Thus, the hypothesis in (3.2) is not, in general, equivalent to the hypothesis in (3.1) with $F_{bwd}(x) \equiv F^0(x)$ and $F_{fwd}(x) \equiv G^0(x)$. We show later, however, that in the particular case where the random variables in each pair correspond to backward and forward recurrence times, the two hypotheses are in fact equivalent.

The observed data are $(X_1^*, Y_1^*), \dots, (X_n^*, Y_n^*)$, where $X_i^* = \min(X_i^0, U_i)$, $Y_i^* = \min(Y_i^0, V_i)$, $\delta_i = \mathbf{1}[X_i^* = X_i^0]$, and $\epsilon_i = \mathbf{1}[Y_i^* = Y_i^0]$, for $i = 1, \dots, n$. The random variables U_i and V_i are the censoring variables of the X_i^0 's and Y_i^0 's, respectively. We assume that $U_i \sim J(s)$ and $V_i \sim K(t)$. It is also assumed that the pairs (X_i^0, Y_i^0) and (U_i, V_i) are independent for all $i = 1, \dots, n$, and that U_i is independent of V_j for $i \neq j$.

Let $X_i^* \sim F(s)$, $Y_i^* \sim G(t)$, and $(X_i^*, Y_i^*) \sim H(s, t)$. Define $\tilde{F}(s) = P(X_i^* \leq s, \delta_i = 1)$, $\tilde{G}(t) = P(Y_i^* \leq t, \epsilon_i = 1)$, and $\tilde{H}_{ab}(s, t) = P(X_i^* \leq s, Y_i^* \leq t, \delta_i = a, \epsilon_i = b)$. Let the scoring function $\Psi(x_i^*, y_j^*, \delta_i, \epsilon_j) = \mathbf{1}[x_i^* > y_j^*, \epsilon_j = 1] - \mathbf{1}[x_i^* < y_j^*, \delta_i = 1]$, which reduces to the traditional Mann-Whitney scoring function when there is no censoring. Also, let,

$$W_n = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \left[\Psi(X_i^*, Y_j^*, \delta_i, \epsilon_j) - p \right], \quad (3.3)$$

where $p = E(\tilde{G}(X)) - E(\tilde{F}(Y))$. Theorem 2, proved by Wei (1980), provides a large sample approximation of the distribution of the test statistic W_n in (3.3).

Theorem 2 As $n \rightarrow \infty$, $\sqrt{n}W_n$ converges in distribution to a normal random variable with mean 0 and variance σ^2 , where, under H_0 ,

$$\begin{aligned} \sigma^2 &= \int \left(1 - G(s)\right)^2 d\tilde{F}(s) + \int \left(1 - F(t)\right)^2 d\tilde{G}(t) \\ &+ 2 \left[\int \tilde{F}(t) \left(1 - G(s)\right) d\left(\tilde{H}_{10}(s, t) + \tilde{H}_{11}(s, t)\right) \right. \\ &\quad - \int \tilde{F}(t) \tilde{G}(s) dH(s, t) - \int \left(1 - G(s)\right) \left(1 - F(t)\right) d\tilde{H}_{11}(s, t) \\ &\quad \left. + \int \tilde{G}(s) \left(1 - F(t)\right) d\left(\tilde{H}_{01}(s, t) + \tilde{H}_{11}(s, t)\right) \right]. \end{aligned} \quad (3.4)$$

Under H_0 , a consistent estimator of σ^2 can be obtained by replacing all the distribution functions in (3.4) by the corresponding e.d.f.'s. Wei noted that another consistent estimator of σ^2 under H_0 can be found by computing,

$$\begin{aligned} \hat{\sigma}_0^2 &= n^{-3} \sum_{j=1}^n \left[\left(\sum_{i=1}^n \mathbf{1}(y_i^* > x_i^*) \right)^2 \delta_j + \left(\sum_{i=1}^n \mathbf{1}(x_i^* > y_i^*) \right)^2 \epsilon_j \right. \\ &\quad + 2 \left(\sum_{i=1}^n \mathbf{1}(y_i^* \leq x_j^*) \epsilon_i \right) \left(\sum_{k=1}^n \mathbf{1}(x_k^* > y_j^*) \epsilon_j \right) \\ &\quad + 2 \left(\sum_{i=1}^n \mathbf{1}(x_i^* \leq y_j^*) \delta_i \right) \left(\sum_{k=1}^n \mathbf{1}(y_k^* > x_j^*) \delta_j \right) \\ &\quad + 2 \left(\sum_{i=1}^n \mathbf{1}(x_i^* \leq y_j^*) \delta_i \right) \left(\sum_{k=1}^n \mathbf{1}(y_k^* \leq x_j^*) \epsilon_k \right) \\ &\quad \left. - 2 \left(\sum_{i=1}^n \mathbf{1}(x_i^* > y_j^*) \right) \left(\sum_{k=1}^n \mathbf{1}(y_k^* > x_j^*) \right) \delta_j \epsilon_j \right]. \end{aligned} \quad (3.5)$$

Thus, Wei's asymptotically nonparametric test compares observed values of $\sqrt{n}W_n/\hat{\sigma}_0$ to a standard normal distribution. We may use this procedure to test for stationarity, identifying (X_i^0, Y_i^0) with (Y_i^{bwd}, Y_i^{fwd}) and (U_i, V_i) with $(+\infty, C_i^*)$, for $i = 1, 2, \dots, n$, provided we can establish the equivalence of (3.2) and (3.1), with $F_{bwd}(x) \equiv F^0(x)$ and $F_{fwd}(x) \equiv G^0(x)$.

3.5 Equivalence of (3.1) and (3.2)

Wei's null hypothesis, shown in (3.2), is not in general equivalent to the null hypothesis in (3.1) which is of interest in this thesis. It is, however, true that the null hypothesis in (3.2) always implies the null hypothesis in (3.1), and when the paired data arise as backward and forward recurrence times from a prevalent cohort study with follow-up, the converse is also true.

Theorem 3 $Y^{bwd} = Y^{fwd}$ in distribution $\Leftrightarrow f_{Y^{bwd}, Y^{fwd}}(x, y) \equiv f_{Y^{bwd}, Y^{fwd}}(y, x)$.

Proof: It is not difficult to find the joint distribution of the backward and forward recurrence times. Let Y represent the observed left truncated lifetime. Then,

$$f_{Y^{bwd}|Y}(x|l) = \frac{g(x)\mathbf{1}[0 \leq x \leq l]}{G(l)},$$

where $g(t)$ and $G(t)$ represent the truncation time density and distribution function, respectively. Also,

$$f_Y(l) = \frac{f(l)G(l)\mathbf{1}[l \geq 0]}{\int_0^\infty g(u)S(u)du},$$

where $f(x)$ and $S(x)$ represent the underlying density and survivor function, respectively, of the incident survival time. Therefore,

$$\begin{aligned} f_{Y^{bwd}, Y}(x, l) &= f_{Y^{bwd}|Y}(x|l)f_Y(l) \\ &= \frac{g(x)f(l)\mathbf{1}[0 \leq x \leq l]\mathbf{1}[l \geq 0]}{\int_0^\infty g(u)S(u)du}. \end{aligned}$$

Now, performing the bivariate transformation $Y^{fwd} = Y - Y^{bwd}$, and $Y^{bwd} \equiv Y^{bwd}$, we have that,

$$\begin{aligned}
f_{Y^{bwd}, Y^{fwd}}(x, y) &= f_{Y^{bwd}, Y}(x, x + y) \\
&= \frac{g(x)f(x + y)\mathbf{1}[0 \leq x \leq x + y]\mathbf{1}[0 \leq x + y \leq \infty]}{\int_0^\infty g(u)S(u)du} \\
&= \frac{g(x)f(x + y)\mathbf{1}[0 \leq x \leq \infty]\mathbf{1}[0 \leq y \leq \infty]}{\int_0^\infty g(u)S(u)du} \tag{3.6}
\end{aligned}$$

$$= \frac{g(x)f(x + y)\mathbf{1}[0 \leq x \leq \infty]\mathbf{1}[0 \leq y \leq \infty]}{\int_0^\infty \int_0^\infty g(u)f(u + v)dvdu} \tag{3.7}$$

Under stationarity, we know that $g(x)$ is constant so that from (3.7) the joint density of the backward and forward recurrence times depends only on their sum, $(x + y)$. Hence, it is clear that stationarity implies that $f_{Y^{bwd}, Y^{fwd}}(x, y) = f_{Y^{bwd}, Y^{fwd}}(y, x)$, that is, that the joint density of the backward and forward recurrence times is symmetric in its arguments. However, stationarity is equivalent to the hypothesis in (3.1) by the characterization of Asgharian et al., and so we have that (3.1) \Rightarrow (3.2). ■

We thus have a procedure to formally test for the stationarity assumption. In order to explore the power of our test, we need to specify alternative hypotheses which may be of interest in practice. For example, a researcher may be interested in an alternative to stationarity which states that the onset intensity, $\lambda(t)$, is not constant, but is instead non-decreasing over time. In Chapter 4 we develop a characterization for such an alternative, and for related alternatives, which extend the results of Asgharian et al. (2004).

3.6 Another characterization of stationarity

In Section 3.3 we stated the characterization of stationarity given by Asgharian et al. (2004), who established the equivalence between stationarity and equality in distribution of the backward and forward recurrence time distributions. We now provide another characterization of stationarity by assuming that $\lambda(t)$ belongs to the class of monotone intensity functions, that is, by assuming that $\lambda(t)$ is either non-decreasing or non-increasing.

Theorem 4 *Suppose $\lambda(t)$ belongs to the class of monotone (possibly constant) intensity functions and let $W = Y^{fwd} - Y^{bwd}$. Then W is symmetrically distributed about 0 if and only if stationarity holds.*

Proof: The proof essentially requires that we find the density of W . We use the expression for the joint backward and forward recurrence time distribution in (3.6), and let $W = Y^{fwd} - Y^{bwd}$, along with the dummy transformation $T = Y^{bwd}$. This implies that,

$$\begin{aligned} f_{W,T}(w,t) &= \frac{g(t)f(2t+w)\mathbf{1}[t \geq 0]\mathbf{1}[t+w \geq 0]}{\int_0^\infty g(u)S(u)du} \\ &= \frac{g(t)f(2t+w)\mathbf{1}[w \in \mathcal{R}]\mathbf{1}[t \geq \max(0, -w)]}{\int_0^\infty g(u)S(u)du}, \end{aligned}$$

which implies that,

$$f_W(w) = \begin{cases} \frac{\int_{-w}^\infty g(t)f(2t+w)dt}{\int_0^\infty g(u)S(u)du} & \text{if } w < 0 \\ \frac{\int_0^\infty g(t)f(2t+w)dt}{\int_0^\infty g(u)S(u)du} & \text{if } w \geq 0 \end{cases}.$$

For $w \geq 0$, let $t' = t + w$. This gives that,

$$\begin{aligned} f_W(w) &= \frac{\int_w^\infty g(t' - w)f(2(t' - w) + w)dt'}{\int_0^\infty g(u)S(u)du} \\ &= \frac{\int_w^\infty g(t - w)f(2t - w)dt}{\int_0^\infty g(u)S(u)du}, \end{aligned}$$

switching back to the dummy variable t . Therefore, for all $w \in \mathbb{R}$,

$$f_W(w) = \frac{\int_{|w|}^\infty g(t + \min(0, -w))f(2t - |w|)dt}{\int_0^\infty g(u)S(u)du}. \quad (3.8)$$

Under stationarity, (3.8) reduces to,

$$f_W(w) = \frac{\int_{|w|}^\infty f(2t - |w|)dt}{\int_0^\infty g(u)S(u)du}. \quad (3.9)$$

From (3.9), it is clear that stationarity implies that W is symmetrically distributed about 0. Conversely, we next show that if W is symmetrically distributed about 0, then stationarity holds. Equivalently, we can show that non-stationarity implies that W is not symmetrically distributed about 0. Recall that non-stationarity is equivalent to $\lambda(t)$ being non-constant, which is equivalent to $g(t)$ being non-constant. By restricting the class of intensities to those that are monotone, we have that non-stationarity is equivalent to $\lambda(t)$ being either non-decreasing or non-increasing, and non-constant (i.e. $g(t)$ either non-increasing or non-decreasing, and non-constant). It is sufficient to show that there exists $w > 0$ such that $f_W(w) \neq f_W(-w)$ or $f_W(w) - f_W(-w) \neq 0$.

Therefore, we need to show that for $w > 0$,

$$\int_w^\infty [g(t-w) - g(t)]f(2t-w)dt \neq 0. \quad (3.10)$$

But if $g(t)$ is non-decreasing and non-constant, then $[g(t-w) - g(t)] \leq 0$ for all $t > w$, with a strict inequality holding for at least one t . This implies that the L.H.S. of (3.10) is strictly negative. And, if $g(t)$ is non-increasing and non-constant, then $[g(t-w) - g(t)] \geq 0$ for all $t > w$, with a strict inequality holding for at least one t . This implies that the L.H.S. of (3.10) is strictly positive. Thus, in either case, (3.10) holds for all $w > 0$, which implies that W is not symmetrically distributed about 0. ■

Note that, in general, if X and Y are random variables such that $W = Y - X$ is symmetrically distributed about 0, this does not necessarily imply that $X = Y$ in distribution. For example, X and Y can be symmetrically distributed about the same point, but have different variances.

In summary, Theorem 5 presents a series of equivalent statements of stationarity of the underlying incidence process.

Theorem 5 *Suppose $\lambda(t)$ belongs to the class of monotone (possibly constant) intensity functions. The following are then equivalent:*

- (i) *Stationarity of the underlying incidence process.*
- (ii) *$Y^{bwd} = Y^{fwd}$ in distribution.*
- (iii) *$f_{Y^{bwd}, Y^{fwd}}(x, y) \equiv f_{Y^{bwd}, Y^{fwd}}(y, x)$.*
- (iv) *$f_W(w)$ is symmetric about 0, where $W = Y^{fwd} - Y^{bwd}$.*

Chapter 4

Characterizations of departures from stationarity

In Chapter 3, we proposed a formal test of the assumption of stationarity. The test is asymptotically distribution-free, with a suitably standardized test statistic, W_n , that converges in distribution to a standard normal distribution. Two factors are instrumental in suggesting the utilization of this test. First is that the problem involves data that constitute matched pairs, that is, the backward and forward recurrence times. Second is the result which established the equivalence of stationarity to equality of the backward and forward recurrence time distributions (Asgharian et al., 2004).

The performance of a hypothesis testing procedure must be assessed through its power under specified alternative hypotheses. Until now, we have not discussed the possible alternatives to stationarity of the underlying incidence process. Since there are many possible ways in which there could be departure from stationarity, we will restrict ourselves to only certain types. In this chapter, we extend the result of Asgharian et al. (2004)

by characterizing departures from stationarity which may be of interest in practice. These characterizations will help us to explore the power of the test put forward in Chapter 3 to detect non-stationarity.

4.1 Characterization of an important alternative hypothesis

Stationarity is assumed to hold if the incidence process of initiation times is a Poisson process, with intensity $\lambda(t) \equiv \lambda$. By Theorem 1, the null hypothesis of stationarity, $H_0 : \lambda(t) \equiv \lambda$, is equivalent to $H_0 : F_{bwd}(x) \equiv F_{fwd}(x)$. The most usual alternative to stationarity is that the intensity of the incidence process is non-decreasing, and non-constant (or alternatively, non-increasing and non-constant) as a function of time. Researchers might be interested in determining whether the incidence rate of diabetes, say, has been increasing over time. For example, Keiding et al. (1989) estimated the incidence of insulin-dependent diabetes mellitus over 1933-73 in Fyn county, Denmark, for prevalent cases identified by Green et al. (1981). Commenges et al. (2004) provided another example, as they estimated the incidence of Alzheimer's disease or dementia in subjects 65 years and older living at home in southwestern France. A two-sided alternative is less common. For the moment we focus on,

$$H_0 : \lambda(t) \equiv \lambda \text{ vs. } H_1 : \lambda(t) \text{ non-decreasing and non-constant. (4.1)}$$

We wish to investigate the relationship between $F_{bwd}(x)$ and $F_{fwd}(x)$ under the alternative hypothesis in (4.1). From Asgharian et al. (2004), we have that for $x > 0$,

$$f_{bwd}(x) = \frac{g(x)S(x)}{\int_0^\infty g(u)S(u)du}, \text{ and} \quad (4.2)$$

$$f_{fwd}(x) = \frac{\int_0^\infty g(u)f(u+x)du}{\int_0^\infty g(u)S(u)du}. \quad (4.3)$$

We shall show that the backward and forward recurrence time distributions are stochastically ordered when it is assumed that $\lambda(t)$ is a non-decreasing and non-constant function. First, recall the definition of stochastic ordering,

Definition 4 *A random variable X is stochastically larger than a random variable Y if $F_X(s) \leq F_Y(s)$ for all s , and $F_X(t) < F_Y(t)$ for some t . We will write $X >^{SL} Y$ if X is stochastically larger than Y .*

Theorem 6 *Suppose that $\lambda(t)$ is non-decreasing and non-constant. Then $Y^{fwd} >^{SL} Y^{bwd}$.*

Proof: From (4.2) and (4.3) we have that,

$$F_{bwd}(x) = \frac{\int_0^x g(t)S(t)dt}{\int_0^\infty g(u)S(u)du}, \text{ and} \quad (4.4)$$

$$F_{fwd}(x) = \frac{\int_0^x \int_0^\infty g(u)f(u+t)dudt}{\int_0^\infty g(u)S(u)du}. \quad (4.5)$$

However, (4.4) and (4.5) can be expressed differently, and in a more convenient manner, for the purposes of this proof. Let X represent the true lifetime, and T the left truncation time. Then,

$$\begin{aligned}
F_{bwd}(x) &= P(Y^{bwd} \leq x | X \geq T) \\
&= \frac{P(Y^{bwd} \leq x, X \geq T)}{P(X \geq T)} \\
&= \frac{\int_0^\infty \int_0^a P(Y^{bwd} \leq x, X \geq T | X = a, T = t) g(t | X = a) f(a) dt da}{\int_0^\infty P(X \geq T | T = u) g(u) du} \\
&= \frac{\int_0^\infty \int_0^a P(Y^{bwd} \leq x, X \geq T | X = a, T = t) g(t) f(a) dt da}{\int_0^\infty P(X \geq u | T = u) g(u) du} \\
&= \frac{\int_0^\infty \int_0^a P(Y^{bwd} \leq x, X \geq T | X = a, T = t) g(t) f(a) dt da}{\int_0^\infty S(u) g(u) du} .
\end{aligned}$$

Also, $P(Y^{bwd} \leq x, X \geq T | X = a, T = t) = \mathbf{1}[t \leq x]$. Therefore,

$$F_{bwd}(x) = \frac{\int_0^x \int_0^a g(t) f(a) dt da + \int_x^\infty \int_0^x g(t) f(a) dt da}{\int_0^\infty S(u) g(u) du} . \quad (4.6)$$

Similarly, for $F_{fwd}(x)$, we have that,

$$\begin{aligned}
F_{fwd}(x) &= P(Y^{fwd} \leq x | X \geq T) \\
&= \frac{P(Y^{fwd} \leq x, X \geq T)}{P(X \geq T)} \\
&= \frac{\int_0^\infty \int_0^a P(Y^{fwd} \leq x, X \geq T | X = a, T = t) g(t) f(a) dt da}{\int_0^\infty S(u) g(u) du} .
\end{aligned}$$

Using that $P(Y^{fwd} \leq x, X \geq T | X = a, T = t) = \mathbf{1}[t \geq a - x]$ we have,

$$F_{fwd}(x) = \frac{\int_0^x \int_0^a g(t)f(a)dt da + \int_x^\infty \int_{a-x}^a g(t)f(a)dt da}{\int_0^\infty S(u)g(u)du} . \quad (4.7)$$

To see that (4.4) and (4.5) are equivalent to (4.6) and (4.7), we need only show that the numerators of the corresponding expressions are identical. We have that $F_{bwd}(x) \propto \int_0^x g(t)S(t)dt$. Using $S(t) = \int_0^\infty f(u+t)du$, we have that $F_{bwd}(x) \propto \int_0^x \int_0^\infty g(t)f(u+t)dudt$. Letting $a = u + t$, we have,

$$\begin{aligned} F_{bwd}(x) &\propto \int_0^x \int_t^\infty g(t)f(a)dadt \\ &= \int_0^x \int_0^a g(t)f(a)dt da + \int_x^\infty \int_0^x g(t)f(a)dt da . \end{aligned}$$

Similarly, $F_{fwd}(x) \propto \int_0^x \int_0^\infty g(u)f(u+t)dudt$. Letting $a = u + t$ we obtain,

$$\begin{aligned} F_{fwd}(x) &\propto \int_0^x \int_t^\infty g(a-t)f(a)dadt \\ &= \int_0^x \int_0^a g(a-t)f(a)dt da + \int_x^\infty \int_0^x g(a-t)f(a)dt da \\ &= \int_0^x \int_0^a g(t)f(a)dt da + \int_x^\infty \int_{a-x}^a g(t)f(a)dt da . \end{aligned}$$

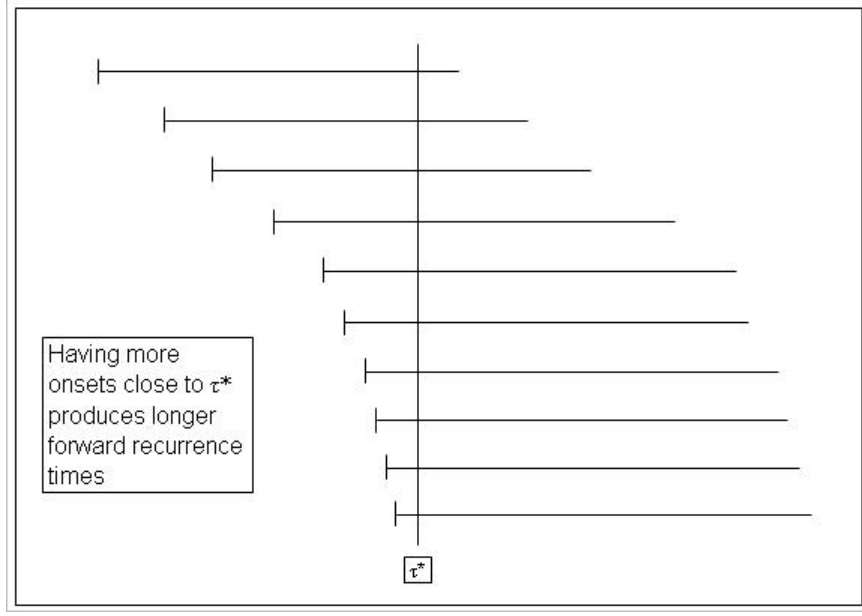


Figure 4.1: Intuition behind Theorem 6

Therefore, we have that,

$$\begin{aligned}
F_{bwd}(x) - F_{fwd}(x) &\propto \int_x^\infty \int_0^x g(t)f(a)dt da - \int_x^\infty \int_{a-x}^a g(t)f(a)dt da \\
&= \int_x^\infty \left[\int_0^x g(t)f(a)dt - \int_{a-x}^a g(t)f(a)dt \right] da \\
&= \int_x^\infty \left[\int_0^x g(t)dt - \int_{a-x}^a g(t)dt \right] f(a) da \\
&= \int_x^\infty \int_0^x \left[g(t) - g(t + (a - x)) \right] f(a)dt da \quad (4.8) \\
&\geq 0 \quad \forall x \geq 0,
\end{aligned}$$

since when $\lambda(t)$ is non-decreasing, $g(t)$ is non-increasing, as seen from (2.2), and the fact that $a > x$. This implies that $Y^{fwd} >^{SL} Y^{bwd}$. ■

The intuition behind Theorem 6 is seen in Figure 4.1. Even under stationarity, we observe more subjects who had onset closer to recruitment,

simply because these individuals do not need to survive as long in order to be observed. Hence, we observe more subjects who have “short” backward recurrence times, and who are more likely to have “long” forward recurrence times. Using the same reasoning, we observe fewer subjects who had onset a long time before recruitment, and thus have “long” backward recurrence times. Figure 4.1 emphasizes, however, that, in the case of an increasing intensity, the discrepancy between the numbers of individuals with “short” backward recurrence times and those with “long” backward recurrence times is over and above what would be expected under stationarity.

Equation (4.8) provides an alternative proof of the characterization of stationarity given by Asgharian et al. (2004). From (4.8) we see that,

$$F_{bwd}(x) = F_{fwd}(x) \quad \forall x \geq 0 \iff g(t) \equiv \text{constant},$$

which is true if and only if $\lambda(t) \equiv \lambda$ (i.e. if and only if stationarity holds).

Analogously to Theorem 6, a similar argument shows that if we suppose $\lambda(t)$ is non-increasing and non-constant then $Y^{bwd} >^{SL} Y^{fwd}$.

4.2 Restricting the class of intensities

The question now arises: Is the converse to Theorem 6 true? That is, does $Y^{fwd} >^{SL} Y^{bwd}$ imply $\lambda(t)$ is non-decreasing and non-constant? Intuitively, we can see that this is an extremely demanding implication. It would mean that if $\lambda(t)$ were non-decreasing over nearly its entire support except for a very small portion, we could not have stochastic ordering of the backward and forward recurrence times. In fact, it is not difficult to give an example where $\lambda(t)$ is strictly decreasing over a portion of its support, but where $Y^{fwd} >^{SL} Y^{bwd}$.

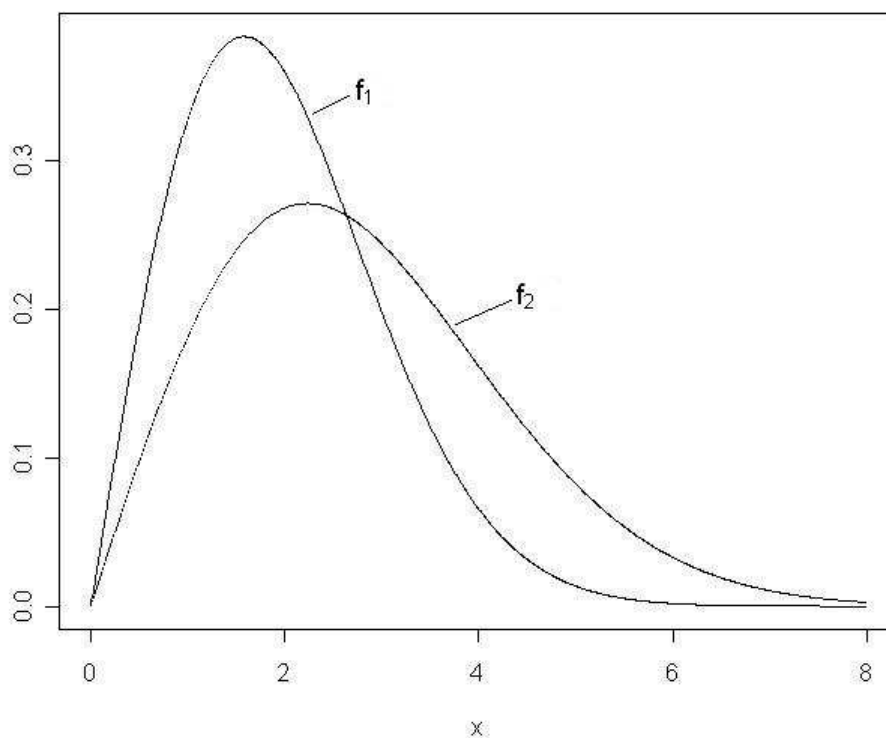


Figure 4.2: Densities satisfying a sufficient condition for stochastic ordering

Counterexample: First, we point out that if any two density functions, $f_1(x)$ and $f_2(x)$, having common support, cross each other only once over their support, then the two distributions must be stochastically ordered. For example, in Figure 4.2, if $X \sim f_1$ and $Y \sim f_2$, then $X >^{SL} Y$. We provide a more precise statement of this result in the following Lemma.

Lemma 1 *Let $f_1(x)$ and $f_2(x)$ be two density functions sharing a common support and satisfying the following property: there exists a unique x^* such that $f_1(x) > f_2(x)$ for $x \in (0, x^*)$, $f_1(x) \leq f_2(x)$ for $x > x^*$, and $f_1(x^*) = f_2(x^*)$. Let $X \sim f_1$ and $Y \sim f_2$, then $X >^{SL} Y$.*

Proof: Let $F_1(x) = \int_0^x f_1(u)du$, $S_1(x) = \int_x^\infty f_1(u)du$, $F_2(x) = \int_0^x f_2(u)du$, and $S_2(x) = \int_x^\infty f_2(u)du$. Then $F_1(x) > F_2(x)$ for all $x \in (0, x^*)$ since

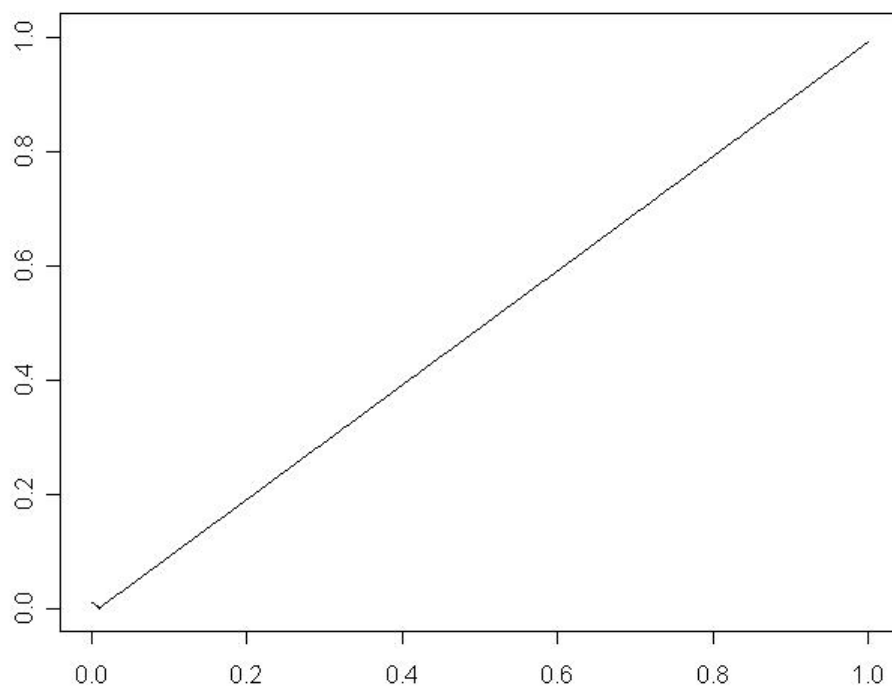


Figure 4.3: Graph of $\lambda(t)$

for these x 's we have that $f_1(x) > f_2(x)$. Also, $S_2(x) \geq S_1(x)$ for all $x \in [x^*, \infty)$ since for these x 's we have that $f_2(x) \geq f_1(x)$. This implies that $F_1(x) \geq F_2(x)$ for all $x \in [x^*, \infty)$. Therefore, $X >^{SL} Y$. ■

For our counterexample, we select an intensity $\lambda(t)$ and an underlying survival density $f(x)$ which result in $f_{bwd}(x)$ and $f_{fwd}(x)$ crossing only once, such that $Y^{fwd} >^{SL} Y^{bwd}$. Let,

$$\lambda(t) = (0.01 - t)\mathbf{1}[0 < t \leq 0.01] + (t - 0.01)\mathbf{1}[0.01 < t < 1] .$$

Using equation (2.2), this implies that,

$$g(x) \propto (0.99 - x)\mathbf{1}[0 < x \leq 0.99] + (x - 0.99)\mathbf{1}[0.99 < x < 1] .$$

A graph of $\lambda(t)$ is provided in Figure 4.3, and it is clear that $\lambda(t)$ is decreasing over $(0, 0.01)$ (see lower endpoint of the graph). Let $f(x) = \mathbf{1}[0 < x < 1]$,

that is, assume that the underlying survival distribution is uniform(0,1).

This implies that $S(x) = \mathbf{1}[x \leq 0] + (1 - x)\mathbf{1}[0 < x < 1]$. We define,

$$\begin{aligned} h(x) &= g(x)S(x) - \int_0^\infty g(u)f(u+x)du \\ &= \int_0^\infty (g(x) - g(u))f(u+x)du \\ &\propto f_{bwd}(x) - f_{fwd}(x) . \end{aligned}$$

We will show that $\lim_{x \rightarrow 0^+} h(x) > 0$, and that $h(x)$ crosses 0 only once on the interval (0, 1). This implies that $f_{bwd}(x)$ and $f_{fwd}(x)$ cross each other only once over (0, 1), and that $Y^{fwd} >^{SL} Y^{bwd}$, which will complete the counterexample. We have that for $0 < x < 1$,

$$\begin{aligned} g(x)S(x) &\propto \left[(0.99 - x)\mathbf{1}[0 < x \leq 0.99] + (x - 0.99)\mathbf{1}[0.99 < x < 1] \right] (1 - x) \\ &= (x^2 - 1.99x + 0.99)\mathbf{1}[0 < x \leq 0.99] \\ &\quad - (x^2 - 1.99x + 0.99)\mathbf{1}[0.99 < x < 1] . \end{aligned}$$

Also,

$$\begin{aligned} \int_0^\infty g(u)f(u+x)du &= \int_0^\infty g(u)\mathbf{1}[0 < u+x < 1]du \\ &\propto \int_0^\infty (0.99 - u)\mathbf{1}[0 < u < 0.99]\mathbf{1}[0 < u+x < 1]du \\ &\quad + \int_0^\infty (u - 0.99)\mathbf{1}[0.99 < u < 1]\mathbf{1}[0 < u+x < 1]du \\ &= \left[\int_0^{0.99} (0.99 - u)du + \int_{0.99}^{1-x} (u - 0.99)du \right] \mathbf{1}[0 < x \leq 0.01] \\ &\quad + \left[\int_0^{1-x} (0.99 - u)du \right] \mathbf{1}[0.01 < x < 1] \end{aligned}$$

$$\begin{aligned}
&= \left[\frac{x^2}{2} - 0.01x + 0.4901 \right] \mathbf{1}[0 < x \leq 0.01] \\
&\quad + \left[\frac{-x^2}{2} + 0.01x + 0.49 \right] \mathbf{1}[0.01 < x < 1] .
\end{aligned}$$

This implies that,

$$h(x) \propto \begin{cases} \frac{x^2}{2} - 1.98x + 0.4999 & \text{if } 0 < x \leq 0.01 \\ \frac{3x^2}{2} - 2x + 0.5 & \text{if } 0.01 < x \leq 0.99 \\ \frac{-x^2}{2} + 1.98x - 1.48 & \text{if } 0.99 < x < 1 \end{cases} .$$

The graph of $h(x)$, up to a constant of proportionality, on the interval $(0,1)$ is presented in Figure 4.4. We see that $h(x)$ is strictly positive on $(0, \frac{1}{3})$, and strictly negative on $(\frac{1}{3}, 1)$, with $h(\frac{1}{3}) = 0$. This proves that $Y^{fwd} >^{SL} Y^{bwd}$, in spite of the fact that $\lambda(t)$ is strictly decreasing on $(0, 0.01)$.

Hence, we do not yet have a characterization of a non-decreasing, non-constant intensity, $\lambda(t)$, in terms of $F_{bwd}(x)$ and $F_{fwd}(x)$. However, if we restrict the class of admissible intensities, we have the following theorem.

Theorem 7 *Suppose that $\lambda(t)$ belongs to the class of monotone (possibly constant) intensity functions. Then $Y^{fwd} >^{SL} Y^{bwd} \Leftrightarrow \lambda(t)$ is non-decreasing, and non-constant.*

Proof:(\Leftarrow) See Theorem 6.

(\Rightarrow) Since $Y^{fwd} >^{SL} Y^{bwd}$, we cannot have $\lambda(t)$ constant as this would violate the characterization of stationarity (Theorem 1) given by Asgharian et al. (2004). It is also not possible that $\lambda(t)$ is non-increasing since this would imply that $Y^{bwd} >^{SL} Y^{fwd}$. Hence, $\lambda(t)$ must be non-decreasing, and non-constant. ■

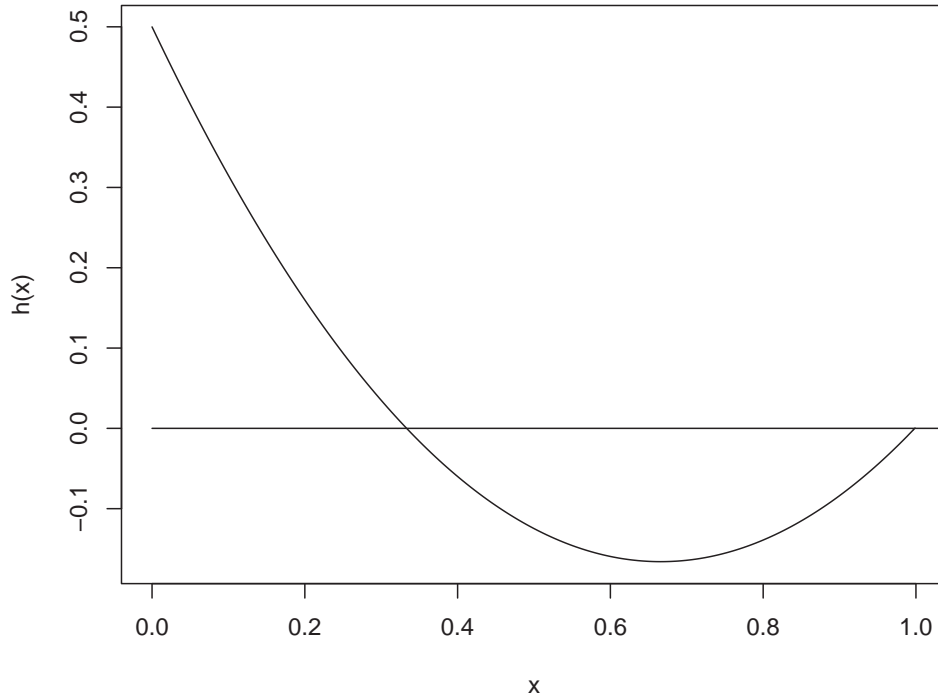


Figure 4.4: Graph of $h(x)$

Clearly, Theorem 7 also provides the characterization, $Y^{bwd} >^{SL} Y^{fwd} \Leftrightarrow \lambda(t)$ is non-increasing, and non-constant, if we again restrict $\lambda(t)$ to the class of monotone intensity functions.

In view of Theorem 6 and Theorem 7, the hypotheses

$$H_0 : \lambda(t) \equiv \lambda \text{ vs. } H_1 : \lambda(t) \text{ non-decreasing and non-constant,}$$

are equivalent to,

$$H_0^* : Y^{fwd} = Y^{bwd} \text{ in distribution vs. } H_1^* : Y^{fwd} >^{SL} Y^{bwd} .$$

In Chapter 5, we provide an examination of the power of our proposed test for detecting departure from stationarity in favor of the types of alternatives which we have characterized in Chapter 4.

4.3 One more characterization of stationarity

In Section 3.6, Theorem 4 provided a characterization of stationarity which is based on the distribution of the difference $Y^{fwd} - Y^{bwd}$. Implicitly, in Theorem 4, we refer to a pair (Y^{bwd}, Y^{fwd}) where Y^{bwd} and Y^{fwd} are correlated because they arise from the same left truncated lifetime. We can also speak of a pair consisting of a backward and a forward recurrence time, where these are independently drawn from their respective univariate distributions, $F_{bwd}(x)$ and $F_{fwd}(x)$. We label such a pair $(\tilde{Y}^{bwd}, \tilde{Y}^{fwd})$.

Theorem 8 *Suppose that $\lambda(t)$ belongs to the class of monotone (possibly constant) intensity functions. Let $\theta_1 = P(Y^{bwd} < Y^{fwd})$, and $\theta_2 = P(\tilde{Y}^{bwd} < \tilde{Y}^{fwd})$. Then, $\theta_1 = 0.5 \Leftrightarrow \theta_2 = 0.5 \Leftrightarrow$ stationarity holds.*

Proof: We have already shown that stationarity $\Rightarrow \theta_1 = 0.5$ in Theorem 4 since we show that stationarity is equivalent to $f_W(w)$ being symmetrically distributed about 0. Furthermore, in the proof of Theorem 4, we use (3.8) to show that if $\lambda(t)$ is non-increasing and non-constant then $f_W(w) \leq f_W(-w) \forall w \geq 0$. This implies that $\theta_1 < 0.5$. Similarly, if $\lambda(t)$ is non-decreasing and non-constant then $f_W(w) \geq f_W(-w) \forall w \geq 0$, which implies that $\theta_1 > 0.5$. This shows that $\theta_1 = 0.5 \Leftrightarrow$ stationarity holds.

We also have that,

$$\begin{aligned} f_{\tilde{Y}^{bwd}, \tilde{Y}^{fwd}}(x, y) &= f_{\tilde{Y}^{bwd}}(x) f_{\tilde{Y}^{fwd}}(y) \\ &= \left(\frac{g(x)S(x)}{\int_0^\infty g(u)S(u)du} \right) \left(\frac{\int_0^\infty g(u)f(u+y)du}{\int_0^\infty g(u)S(u)du} \right). \end{aligned} \quad (4.9)$$

Using (4.9), and performing the transformation $W^* = \tilde{Y}^{fwd} - \tilde{Y}^{bwd}$, we have,

$$f_{W^*}(w) = \begin{cases} \frac{\int_{-w}^{\infty} g(t)S(t) \left(\int_0^{\infty} g(u)f(u+w+t)du \right) dt}{\left(\int_0^{\infty} g(u)S(u)du \right)^2} & \text{if } w < 0 \\ \frac{\int_w^{\infty} g(t-w)S(t-w) \left(\int_0^{\infty} g(u)f(u+t)du \right) dt}{\left(\int_0^{\infty} g(u)S(u)du \right)^2} & \text{if } w \geq 0 \end{cases} .$$

For all $w \in \mathbb{R}$, this can be written as,

$$f_{W^*}(w) = \frac{\int_{|w|}^{\infty} g(t + \min(0, -w))S(t + \min(0, -w))I(t, w)dt}{\left(\int_0^{\infty} g(u)S(u)du \right)^2} ,$$

where we define $I(t, w) = \int_0^{\infty} g(u)f(u + t + \min(0, w))du$. And, under stationarity, this reduces to,

$$\begin{aligned} f_{W^*}(w) &= \frac{\int_{|w|}^{\infty} S(t + \min(0, -w)) \left(\int_0^{\infty} f(u + t + \min(0, w))du \right) dt}{\left(\int_0^{\infty} S(u)du \right)^2} \\ &= \frac{\int_{|w|}^{\infty} S(t + \min(0, -w))S(t + \min(0, w))dt}{\left(\int_0^{\infty} S(u)du \right)^2} . \end{aligned} \quad (4.10)$$

From (4.10) it is clear that stationarity implies that $\theta_2 = 0.5$, since $f_{W^*}(w)$ is symmetric about 0. If stationarity does not hold, then we have that $\lambda(t)$ is either non-decreasing and non-constant or non-increasing and non-constant. By applying Theorem 7 here, we know that $\lambda(t)$ is non-decreasing and non-constant if and only if $F_{\tilde{Y}^{bwd}}(x) \geq F_{\tilde{Y}^{fwd}}(x) \forall x$, with a strict inequality for at least one x , and an analogous statement can be made when $\lambda(t)$ is

non-increasing and non-constant. Thus, if $\lambda(t)$ is non-decreasing and non-constant,

$$\begin{aligned}
\theta_2 &= P(\tilde{Y}^{bwd} < \tilde{Y}^{fwd}) = \int_0^\infty P(\tilde{Y}^{bwd} < \tilde{Y}^{fwd} | \tilde{Y}^{fwd} = y) f_{\tilde{Y}^{fwd}}(y) dy \\
&= \int_0^\infty F_{\tilde{Y}^{bwd}}(y) f_{\tilde{Y}^{fwd}}(y) dy \\
&> \int_0^\infty F_{\tilde{Y}^{fwd}}(y) f_{\tilde{Y}^{fwd}}(y) dy \\
&= P(\tilde{Y}_1^{fwd} < \tilde{Y}_2^{fwd}) = 0.5 ,
\end{aligned}$$

since \tilde{Y}_1^{fwd} and \tilde{Y}_2^{fwd} are i.i.d. random variables with common distribution function F_{fwd} . Clearly, $\lambda(t)$ non-increasing and non-constant implies that $\theta_2 < 0.5$. Thus, we have that $\theta_2 = 0.5 \Leftrightarrow$ stationarity holds. ■

Chapter 5

Detecting departure from stationarity: a power study

In this chapter, we examine the power of our test for stationarity, introduced in Chapter 3, assuming varying degrees of non-stationarity. To carry out this assessment, we generated onsets over a pre-specified interval $(0, \tau^*)$ according to an intensity that is strictly monotone, which, by the characterizations from Chapter 4, corresponds to a situation where the backward and forward recurrence time distributions are stochastically ordered. For each onset generated, we generated a corresponding survival time. The survival times which extended beyond τ^* , originating from their respective onset times, constituted the observed prevalent cohort survival times. We constructed an α -level critical region using our test for equality of the backward and forward recurrence time distributions.

5.1 Details of the simulations

The following issues need to be considered: the specification of the underlying survival density, $f(x)$, the residual censoring time distribution, the form of the onset intensity, $\lambda(t)$, and the specific parameters used in $\lambda(t)$, the support of the truncation time distribution $(0, \tau^*)$, where τ^* represents the calendar time of recruitment into the study, the sample size, n , and the number of replications of the entire procedure, M .

5.1.1 Underlying survival distribution

Three different underlying survival distributions were simulated:

$$\text{(A) } Weibull(\gamma = 2, \beta = 10)$$

$$\text{(B) } Weibull(\gamma = 0.75, \beta = 1.5)$$

$$\text{(C) } lognormal(\mu = 1.75, \sigma = 0.4)$$

where the parametrization of the $Weibull(\gamma, \beta)$ distribution is:

$$f(x) = \frac{\gamma}{\beta} x^{\gamma-1} e^{-\frac{x^\gamma}{\beta}} \mathbf{1}[x > 0], \quad (5.1)$$

and the parametrization of the $lognormal(\mu, \sigma^2)$ distribution is:

$$f(x) = \frac{e^{-\frac{(\log x - \mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma x} \mathbf{1}[x > 0]. \quad (5.2)$$

We chose the Weibull and lognormal distributions because of their common use in analyzing failure time data. The (γ, β) pair in **(A)** represents a Weibull with an increasing hazard function, whereas the (γ, β) pair in **(B)** represents a Weibull with a decreasing hazard function.

5.1.2 Sample size

For each of (A), (B), and (C), we examined the power of our test using sample sizes of $n \approx 500$ and $n \approx 1000$.

5.1.3 Residual censoring time distribution

For each survival distribution, the residual censoring time distribution was chosen to be exponential with mean, β , selected so that approximately 25% of the forward recurrence times were censored. Specifically, we employed:

$$\text{(A) } \textit{exponential}(\beta = 6)$$

$$\text{(B) } \textit{exponential}(\beta = 8)$$

$$\text{(C) } \textit{exponential}(\beta = 12)$$

where the parametrization of the exponential(β) distribution is:

$$f(x) = \frac{1}{\beta} e^{-\frac{x}{\beta}} \mathbf{1}[x > 0]. \quad (5.3)$$

5.1.4 Onset intensity and support of the truncation time distribution

The onset process was chosen to be a Poisson process with log-linear intensity function,

$$\lambda(t) = \lambda e^{\alpha_1 t}. \quad (5.4)$$

The intensity in (5.4) is strictly increasing when α_1 is positive, and strictly decreasing when α_1 is negative. Stationarity holds if and only if $\alpha_1 = 0$. For the cases representing non-homogeneous Poisson processes ($\alpha_1 \neq 0$), we

simulated the onsets over an interval, $(0, \tau^*)$, corresponding to the support of the truncation time distribution, by following Lewis and Shedler (1976). For simulation of the stationary onsets, $\text{uniform}(0, \tau^*)$ random variables were generated and ordered. The support of the truncation time distribution was chosen to be sufficiently large to ensure that the backward recurrence time distribution was not “clipped” beyond a certain value: clearly, if onsets are generated on $(0, \tau^*)$ then no backward recurrence time can be observed larger than τ^* . If τ^* were chosen too small, we would have artificially reduced the support of the backward recurrence time distribution. Of course, since the survival distributions in **(A)**, **(B)**, and **(C)** have their support on $(0, \infty)$, we can never select τ^* large enough to guarantee that there is absolutely no “clipping” of the backward recurrence time distribution. We can, however, choose τ^* large enough to ensure that this “clipping” is negligible. The values of τ^* that were chosen for the present power study are as follows:

$$\text{(A)} \tau^* = 6 \quad \text{(B)} \tau^* = 12 \quad \text{(C)} \tau^* = 16 \quad .$$

The parameter that controls the degree of non-stationarity is α_1 , and the larger the magnitude of α_1 , the more there is departure from stationarity. The α_1 values that were chosen for the three survival distributions were:

$$\text{(A)} (1)0.05 \quad (2)0.10 \quad (3)0.15 \quad (4) - 0.05 \quad (5) - 0.10 \quad (6) - 0.15$$

$$\text{(B)} (1)0.05 \quad (2)0.10 \quad (3)0.15 \quad (4) - 0.05 \quad (5) - 0.10 \quad (6) - 0.15$$

$$\text{(C)} (1)0.03 \quad (2)0.06 \quad (3)0.09 \quad (4) - 0.03 \quad (5) - 0.06 \quad (6) - 0.09$$

These α_1 values cover a range of circumstances from “mild” non-stationarity, to “moderate” non-stationarity, to “severe” non-stationarity. The specific values of α_1 were chosen by considering the expected number of onsets that

would occur over $(0, \tau^*)$ under the non-stationary onset intensity relative to the same number under stationarity. It should be noted, however, that only a portion of these onsets survive long enough to be observed at recruitment. For a non-stationary intensity function, the expected number of onsets over $(0, \tau^*)$ is,

$$\int_0^{\tau^*} \lambda(t)dt = \int_0^{\tau^*} \lambda e^{\alpha_1 t} dt = \frac{\lambda(e^{\alpha_1 \tau^*} - 1)}{\alpha_1},$$

and under stationarity, the expected number of onsets over $(0, \tau^*)$ is, $\lambda\tau^*$. Therefore, the ratio, r , of the expected number of onsets in the non-stationary case to the stationary case is,

$$r = \frac{e^{\alpha_1 \tau^*} - 1}{\alpha_1 \tau^*},$$

where we can substitute the values $\tau^* = 6$, $\tau^* = 12$, and $\tau^* = 16$ for **(A)**, **(B)**, and **(C)**, respectively.

The λ parameter does not affect the degree of non-stationarity. The values of λ were chosen to attain the desired sample size (either $n \approx 500$ or $n \approx 1000$), and are not reported in this thesis.

We simulated stationary onsets by generating a sufficient number of uniform random variables over $(0, \tau^*)$ in order to attain the desired sample size of either $n \approx 500$ or $n \approx 1000$.

5.1.5 Number of replications of the procedure

The three survival distributions, two sample sizes, and seven different onset intensities (six non-stationary intensities and a constant intensity representing stationarity) led to 42 different simulation scenarios. For each of these 42 scenarios, we performed $M=200$ replicates, and we recorded the number

of rejections at the $\alpha = 0.05$ level. Note that we used a two-sided rejection region in assessing the properties of this test, as is commonly done. In practice, a researcher would typically have in mind, a priori, exactly one of the alternatives (i.e. either a non-decreasing or non-increasing onset intensity).

5.2 Results of the power study

The percentage of rejections of stationarity are presented in Table 5.1, Table 5.2, and Table 5.3 for **(A)**, **(B)**, and **(C)**, respectively.

α_1	0.05	0.10	0.15	-0.05	-0.10	-0.15	0 (stationarity)
$n \approx 500$	9.5	40.5	81.0	6.0	33.5	78.0	1.0
$n \approx 1000$	24.5	83.0	98.0	11.0	72.5	99.0	1.0

Table 5.1: Percentage of rejections for Weibull($\gamma=2, \beta=10$)

α_1	0.05	0.10	0.15	-0.05	-0.10	-0.15	0 (stationarity)
$n \approx 500$	10.0	49.5	90.5	2.5	46.5	95.0	0.5
$n \approx 1000$	32.0	91.5	100.0	6.0	84.0	100.0	0.5

Table 5.2: Percentage of rejections for Weibull($\gamma=0.75, \beta=1.5$)

α_1	0.03	0.06	0.09	-0.03	-0.06	-0.09	0 (stationarity)
$n \approx 500$	16.5	68.0	90.5	14.5	69.0	97.5	1.5
$n \approx 1000$	28.5	93.0	99.5	31.5	95.5	100.0	1.0

Table 5.3: Percentage of rejections for lognormal($\mu=1.75, \sigma=0.4$)

From these three tables, we see that when $n \approx 500$, our test only performed well for the most severe non-stationary onset intensities. For the increasing onset intensities the power ranged from 81% to 90.5% and for the decreasing intensities the power ranged from 78% to 97.5%. The effect of doubling the sample size was substantial in all three cases in that it considerably increased the power for the “moderate” and “mild” departures from stationarity. We can see that when $n \approx 1000$, the power of our test was quite high except for the mildest departure from stationarity. For the “moderate” departure from non-stationarity, the power ranged from 83% to 93% when the onset intensity was increasing and from 72.5% to 95.5% when the onset intensity was decreasing. A very important result to notice is that under stationarity, the rejection rate was much lower than the nominal 5% value. When $n \approx 500$, the average rejection rate was 1% for the three scenarios. As expected, increasing the sample size did not have any effect on this rejection rate. When $n \approx 1000$, the average rejection rate was slightly below 1% for the three scenarios.

5.3 Discussion of the results

This limited power study demonstrated some general properties of our formal test for stationarity of the incidence process. First, if the departure from stationarity is “severe” this will be detected, even at reasonably small sample sizes. The smallest sample size used in these simulations is $n \approx 500$, which we feel is very attainable for most prevalent cohort studies with follow-up. Also, we notice that, even for relatively large sample sizes, the “mild” departures from stationarity will not be detected. We did not use sample

sizes that were larger than $n \approx 1000$ since it is well known that, due to the nature of statistical hypothesis testing, very large sample sizes can lead to misleadingly high power.

Unfortunately, only when $n \approx 1000$ did our test have good power for detecting “moderate” departures from stationarity. Reasons for this result might be numerous. It is possible that using a high censoring proportion ($\approx 25\%$) compromised the power of the test when $n \approx 500$. Perhaps in studies with substantially lower censoring rates, smaller sample sizes would be adequate to detect these “moderate” departures from stationarity, although this needs to be verified. The power was also slightly reduced by the use of a two-sided rejection region. Another factor which most certainly affected the power of the test is that its rejection rate under the null hypothesis did not reach the nominal 5% value, but was much closer to 1%. We briefly discuss a paper by Cheng (1984), who addressed the reason for this occurrence.

Cheng (1984) generalized Wei’s (1980) test under the additional assumption that the censoring distributions are identical for the censoring variables $U_i \sim J(s)$ and $V_i \sim K(t)$ (i.e. $J(s) \equiv K(s)$). This generalization of Wei’s test is not applicable in our setting since the added assumption of an identical censoring distribution for the backward and forward recurrence times is known not to be valid. The backward recurrence times are, by definition, not censored whereas the forward recurrence times may or may not be censored. Cheng, however, also noted that Wei’s estimator, $\hat{\sigma}_0^2$, is not a good estimator of σ^2 under H_0 . He provided an alternative estimator of σ^2 under H_0 , which, as he demonstrated through examples, performed better than Wei’s estimator. Unfortunately, Cheng’s estimator is still only valid

under the assumption of identical censoring distributions. Jung (1999) also provided a method for estimating σ^2 under H_0 but it is not clear whether this method is better than Wei's estimator. Jung's procedure does not require the equal censorship assumption. Although we cannot utilize Cheng's estimator for our purposes, it is useful to examine the results presented by Cheng (1984) which showed that using Wei's estimator for σ^2 under H_0 yielded percentiles of the test statistic, $\sqrt{n}W_n/\hat{\sigma}_0$, which are not consistent with those of a normal(0,1) random variable. Cheng's results are more in accordance with those obtained in this thesis, that is, under H_0 , $\sqrt{n}W_n/\hat{\sigma}_0$ will fall outside of the interval $(-z_{0.025}, z_{0.025})$ only with a probability of roughly 0.01, where $z_\alpha =$ the $100(1 - \alpha)$ percentile of a standard normal distribution. The "conservative" nature of the test statistic surely affected the power of the test in these simulations. This may be the principal reason why "moderate" departures from stationarity were not detected with a high probability when $n \approx 500$. It is clear that obtaining a better estimator of σ^2 under H_0 , without imposing further assumptions on the censoring distributions, would be a very useful result. This, and other future directions for research are discussed in Chapter 10.

Chapter 6

Verifying the assumption of a fixed survival distribution

Asgharian et al. (2004) showed that stationarity of the initiation times is equivalent to equality of the backward and forward recurrence time distributions, under the implicit assumption that the survival distribution is independent of calendar time of onset. In our previous chapters we have investigated stationarity, under this assumption. Conversely, suppose that we are prepared to assume that the underlying incidence process is stationary, and we wish to examine the hypothesis that the survival distribution is independent of calendar time of onset. It is not possible to simultaneously examine both stationarity of the underlying incidence process and independence of the survival distribution of calendar time of onset.

6.1 Characterizing the independence of survival and calendar time of onset

Theorem 1, from Chapter 3, assumes that the underlying survivor function does not depend on the date of onset to show that,

$$f_{bwd}(x) \equiv f_{fwd}(x) \Leftrightarrow \text{the incidence process is stationary.}$$

One might therefore conjecture that, if one *assumes* stationarity of the underlying incidence process, then the following dual to Theorem 1 exists,

$$f_{bwd,\tau^*}(x) \equiv f_{fwd,\tau^*}(x) \Leftrightarrow S_t(x) \equiv S(x) ,$$

where τ^* represents the calendar time of recruitment into the study, $S_t(x)$ represents the survivor function for an individual who had onset at calendar time (τ^*-t) , and $S(x)$ represents a survivor function which is independent of calendar time of onset.

Before continuing, we note that one direction of the conjecture (\Leftarrow) is clear, since the assumption that the survival distribution is fixed over time signifies, by Theorem 1, that we are once again in the situation where stationarity is equivalent to $f_{bwd,\tau^*}(x) \equiv f_{fwd,\tau^*}(x)$.

Suppose that stationarity holds, and $f_{bwd,\tau^*}(x) \equiv f_{fwd,\tau^*}(x)$. We would like to show that the survival distribution is independent of the calendar time of the origin. Without some restriction, however, on the manner in which the survival distributions depend on the time origin, the implication in this direction does not hold. We present a simple counterexample which demonstrates that even if the initiation process is stationary, it is possible to have the survival distribution changing over time, while $f_{bwd,\tau^*}(x) \equiv f_{fwd,\tau^*}(x)$.

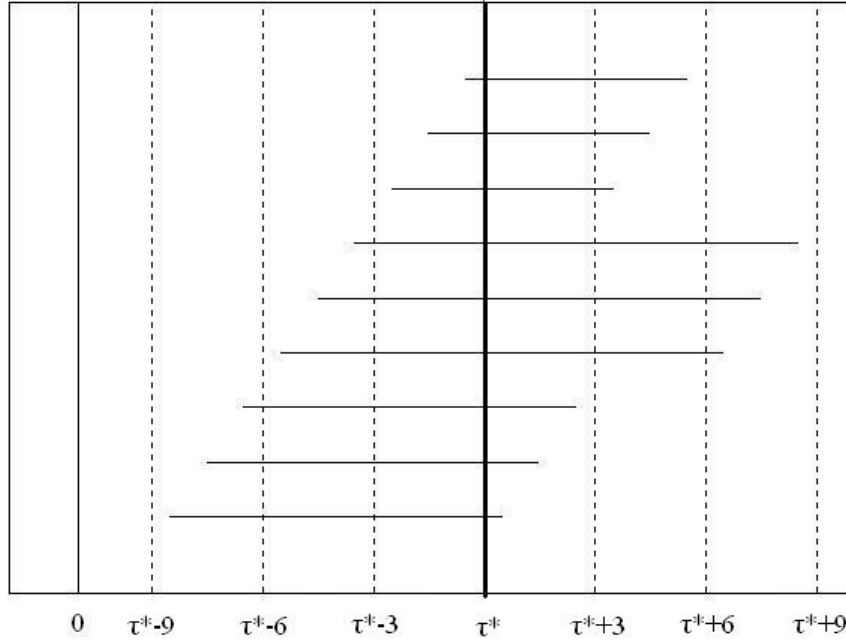


Figure 6.1: Survival depends on calendar time, but $Y^{fwd} >^{SL} Y^{bwd}$

Counterexample: Consider the following survival experience over time: all subjects who had onset at least 6 months from recruitment survive for 9 months, those who had onset between 3 and 6 months from recruitment survive for 12 months, and those who had onset within 3 months of recruitment survive for 6 months. A picture representing this scenario is displayed in Figure 6.1. In this case, the backward and forward recurrence time distributions are both $\text{uniform}(0,9)$. To see this, notice that we cannot observe any subject who had onset before $(\tau^* - 9)$, and we observe all subjects who had onset between $(\tau^* - 9)$ and τ^* . Therefore, the backward recurrence time distribution is identical to the truncation time distribution arising from stationary onsets between $(\tau^* - 9)$ and τ^* , which is $\text{uniform}(0,9)$. Since the survival distributions are degenerate, a subject's forward recurrence time is determined with certainty by their calendar time of onset, which here

also determines their backward recurrence time. This, and the fact that no subject can have a forward recurrence time greater than 9, implies that the forward recurrence time distribution is also uniform(0,9). More formally,

$$F_{fwd}(x) = \begin{cases} F_{bwd}(x+6) - F_{bwd}(6) & \text{if } 0 < x \leq 3 \\ \frac{1}{3} + F_{bwd}(x-3) - F_{bwd}(0) & \text{if } 3 < x \leq 6 \\ \frac{2}{3} + F_{bwd}(x-3) - F_{bwd}(3) & \text{if } 6 < x < 9 \end{cases}$$

This simplifies to $F_{fwd}(x) = \frac{x}{9}\mathbf{1}[0 < x < 9]$, as expected. The survival distribution, however, certainly depends on when a subject had onset, that is, on the calendar time of the origin.

We restrict the manner in which the survival distribution can depend on calendar time through the following three assumptions:

Assumption 1: If the survival distribution changes over time, then as time progresses subsequent survival distributions are stochastically ordered. That is, the survival times are stochastically increasing (respectively decreasing) so that survival improves (respectively worsens) as the date of onset advances.

Assumption 2: The survival distribution can change at, at most, $(K - 1) < \infty$ points in time, so that there are only K possible distinct survivor functions. We will comment further on the need for Assumption 2 later in this chapter.

Assumption 3: Let $f_1(x), f_2(x), \dots, f_K(x)$ denote the K survival densities. We assume that the set A , defined as, $A = \{x : f_i(x) = f_j(x) \text{ for some } i, j = 1, \dots, K\}$, has Lebesgue measure zero.

Note that Assumption 3 does not place any limit on the number of times that any two densities, f_i and f_j , can cross, as long as the stochastic ordering assumption is maintained. Two densities cannot, however, be identical on intervals of positive length, say, or any other set having positive Lebesgue measure. Assumption 3 is not as restrictive as it may initially seem. It holds, for example, if the underlying survival distributions are Weibull or gamma with fixed shape parameter but with scale parameters that change monotonically as the initiation origin advances. It also holds if the underlying survival distributions are lognormal with fixed standard deviation parameter and monotone mean parameter, in the sense just stated.

We return to our conjecture under these newly added restrictions. We begin by redeveloping expressions for $f_{bwd}(x)$ and $f_{fwd}(x)$, recalling that the expressions that were used in Chapter 4 are only valid under the assumption that the survival distribution is independent of calendar time. Also, in the sequel, we suppress the subscript τ^* in $f_{bwd}(x)$ and $f_{fwd}(x)$. Let X denote the underlying survival time, T the truncation time, and Y the left truncated survival time (Y can be called the length-biased survival time, since we are assuming stationarity of the underlying incidence process throughout this chapter). We have that,

$$\begin{aligned}
 f_Y(x) &= \int_0^x f(x, t | X \geq T) dt \\
 &= \int_0^x \frac{f(x, t)}{P(X \geq T)} dt \\
 &= \frac{\int_0^x f_t(x) g(t) dt}{P(X \geq T)} \\
 &= \frac{\int_0^x f_t(x) dt}{P(X \geq T)} ,
 \end{aligned}$$

where we define $f_t(x) = f(x|T = t)$, the survival density given that a subject has a truncation time t , or alternatively, had onset at calendar time $\tau^* - t$. The survival density for a subject now includes a subscript, since it depends on the truncation time, t . Also,

$$\begin{aligned}
f_{bwd}(x|Y = x_0) &= g(x|Y = x_0) \\
&= \frac{f_{T,Y}(x, x_0)}{f_Y(x_0)} \\
&= \frac{f_{T,X}(x, x_0|X > T)}{f_X(x_0|X > T)} \\
&= \frac{\left[\frac{P(X > T|T=x, X=x_0)f_{T,X}(x, x_0)}{P(X > T)} \right]}{\left[\frac{P(X > T|X=x_0)f(x_0)}{P(X > T)} \right]} \\
&= \frac{\mathbf{1}[x_0 > x]f_x(x_0)g(x)}{P(T < x_0|X = x_0)f(x_0)} \\
&= \frac{\mathbf{1}[x_0 > x]f_x(x_0)g(x)}{\int_0^{x_0} f_t(x_0)g(t)dt} \quad ,
\end{aligned}$$

and, under stationarity, we have that,

$$f_{bwd}(x|Y = x_0) = \frac{\mathbf{1}[x_0 > x]f_x(x_0)}{\int_0^{x_0} f_t(x_0)dt} \quad .$$

Lemma 2 *Removing the assumption of a single survival distribution, which is independent of calendar time of onset, and assuming stationarity of the underlying incidence process, we have that:*

$$(i) \quad f_{bwd}(x) \propto \int_x^\infty f_x(x_0) dx_0$$

$$(ii) \quad f_{fwd}(x) \propto \int_x^\infty f_{x_0-x}(x_0) dx_0$$

Proof: Under stationarity,

$$\begin{aligned} f_{bwd}(x) &= \int_x^\infty \left[\frac{f_x(x_0)}{\int_0^{x_0} f_t(x_0) dt} \right] \left[\frac{\int_0^{x_0} f_t(x_0) dt}{P(X \geq T)} \right] dx_0 \\ &= \frac{\int_x^\infty f_x(x_0) dx_0}{P(X \geq T)} \\ &\propto \int_x^\infty f_x(x_0) dx_0 . \end{aligned}$$

And since, $f_{fwd}(x|Y = x_0) = f_{bwd}(x_0 - x|Y = x_0)$, we also have,

$$\begin{aligned} f_{fwd}(x) &= \frac{\int_x^\infty f_{x_0-x}(x_0) dx_0}{P(X \geq T)} \\ &\propto \int_x^\infty f_{x_0-x}(x_0) dx_0 , \end{aligned}$$

which completes the proof. ■

Hence,

$$\begin{aligned}
f_{bwd}(x) &= f_{fwd}(x) \quad \forall x > 0 \quad \Leftrightarrow \\
\int_x^\infty f_x(x_0)dx_0 &= \int_x^\infty f_{x_0-x}(x_0)dx_0 \quad \forall x > 0 \quad \Leftrightarrow \\
\int_x^\infty [f_x(x_0) - f_{x_0-x}(x_0)]dx_0 &= 0 \quad \forall x > 0 \quad . \tag{6.1}
\end{aligned}$$

6.1.1 The case of $K = 2$

We want to show that (6.1) implies that the survival distribution is independent of calendar time, that is, that $f_1 \equiv f_2 \equiv f_3, \dots, \equiv f_K$. For ease of exposition, we begin with the special case $K = 2$. Suppose for contradiction that (6.1) holds and that the survivor functions are stochastically increasing over time. The proofs for the case where the survivor functions are stochastically decreasing over time are similar.

Lemma 3 *In (6.1), there exists u such that, $\forall x_0 > u$, $f_u(x_0) = f_{x_0-u}(x_0)$.*

Proof: Let $\tau^* - \tilde{x}$ be the calendar time where the change in the survival distribution can occur, for some \tilde{x} (see Figure 6.2). Since the two survivor functions must be stochastically ordered ($S_1(x) \leq S_2(x) \forall x > 0$, with a strict inequality for at least one x), then there exists u such that,

$$f_1(x) \leq f_2(x) \quad \forall x \geq u \quad .$$

For this u , we have that $f_u(x_0) \equiv f_1(x_0)$ if $u > \tilde{x}$ or that $f_u(x_0) \equiv f_2(x_0)$ if $u < \tilde{x}$ (see Figure 6.2). Also,

$$f_{x_0-u}(x_0) = f_1(x_0)\mathbf{1}(x_0 > u + \tilde{x}) + f_2(x_0)\mathbf{1}(x_0 \leq u + \tilde{x}) \quad .$$

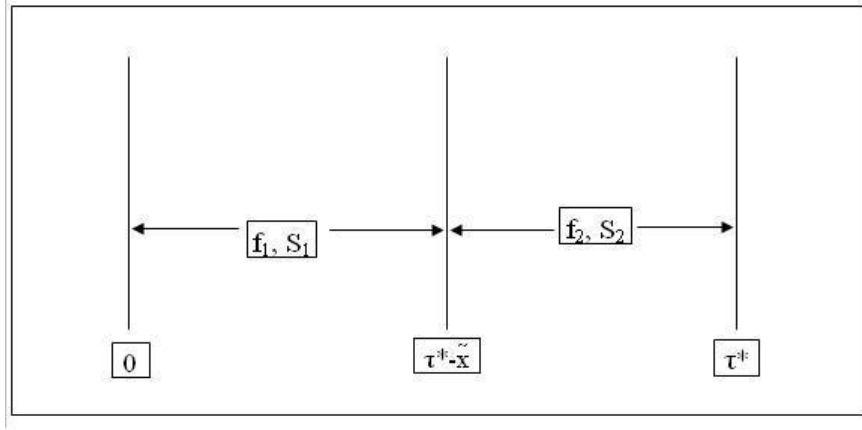


Figure 6.2: Assume the change in survival can occur at $\tau^* - \tilde{x}$

In particular, we note that $\forall x_0 > u$, we have that either: (i) $f_u(x_0) \leq f_{x_0-u}(x_0)$ or (ii) $f_u(x_0) \geq f_{x_0-u}(x_0)$. Since (6.1) must hold for all $x > 0$, it must hold for u . But this implies that $f_u(x_0) = f_{x_0-u}(x_0) \quad \forall x_0 > u$. ■

Theorem 9 *Suppose that the underlying incidence process is stationary. Then under Assumption 1, Assumption 2 (with $K=2$), and Assumption 3, equality of the backward and forward recurrence time distributions (characterized by (6.1)) $\Rightarrow f_1 \equiv f_2$ (that is, the underlying survival distribution is independent of the date of initiation).*

Proof: Let x^* be the point where $f_1(x^*) = f_2(x^*)$, and $f_1(x) < f_2(x) \quad \forall x > x^*$ (see Figure 4.2 for an example). Assumptions 1 and 3 guarantee the existence of such an x^* . We will prove the result by considering the following two cases separately: (i) $x^* \geq \tilde{x}$ and (ii) $x^* < \tilde{x}$.

Case (i): $x^* \geq \tilde{x}$

Since $x^* \geq \tilde{x}$, we have $f_{x^*}(x_0) = f_1(x_0) \quad \forall x_0 > 0$. Also, by Lemma 3,

$$f_{x^*}(x_0) = f_1(x_0) = f_{x_0-x^*}(x_0) \quad \forall x_0 > x^*.$$

For $x_0 \in (x^*, x^* + \tilde{x})$, $f_{x_0-x^*}(x_0) = f_2(x_0)$. This implies that,

$$f_1(x_0) = f_2(x_0) \quad \forall x_0 \in (x^*, x^* + \tilde{x}).$$

Now it is important to note that Lemma 3 can be applied using x^* , but also using any $u > x^*$. Let $x_2^* = x^* + \tilde{x} \geq \tilde{x}$. Then, as before, we have,

$$f_{x_2^*}(x_0) = f_1(x_0) \quad \forall x_0 > 0.$$

Hence by Lemma 3, $f_1(x_0) = f_2(x_0) \quad \forall x_0 \in (x_2^*, x_2^* + \tilde{x}) \equiv (x^* + \tilde{x}, x^* + 2\tilde{x})$.

We can proceed in this fashion and obtain a sequence $x_n^* = x_{n-1}^* + \tilde{x}$, which yields that $f_1(x_0) = f_2(x_0) \quad \forall x_0 \in (x^*, x_{n+1}^*)$, which as $n \rightarrow \infty$ gives,

$$f_1(x_0) = f_2(x_0) \quad \forall x_0 > x^* \quad . \quad (6.2)$$

Since this violates Assumption 3, we conclude that the survival distribution is independent of the date of onset.

Case (ii): $x^* < \tilde{x}$

Suppose $x_1 > \tilde{x} > x^*$. Then we have that $f_{x_1}(x_0) = f_1(x_0) \quad \forall x_0 > 0$. Applying Lemma 3, we thus have that $f_1(x_0) = f_{x_0-x_1}(x_0) \quad \forall x_0 > x_1$. Now, for $x_0 \in (x_1, x_1 + \tilde{x})$, $f_{x_0-x_1}(x_0) = f_2(x_0)$. Therefore,

$$f_1(x_0) = f_2(x_0) \quad \forall x_0 \in (x_1, x_1 + \tilde{x}) \quad .$$

Let $x_2 = x_1 + \tilde{x}$, $x_3 = x_2 + \tilde{x}$, and continue in this fashion. We proceed as in Case (i) to obtain that,

$$f_1(x_0) = f_2(x_0) \quad \forall x_0 > x_1,$$

where x_1 can be chosen arbitrarily close to \tilde{x} .

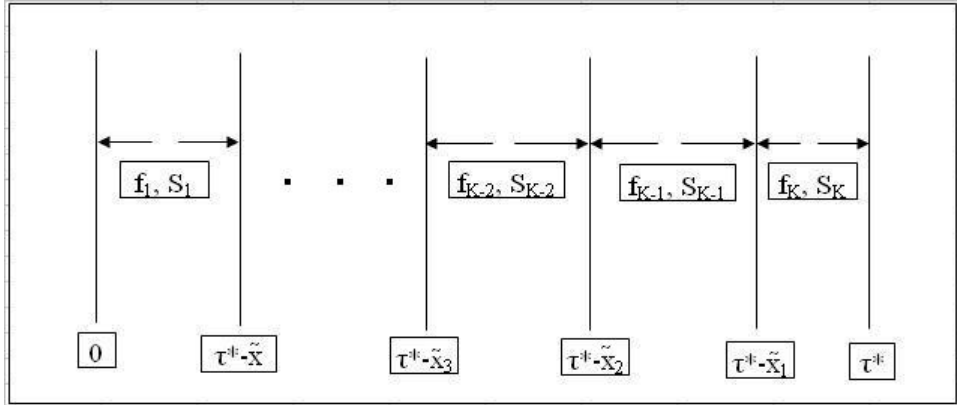


Figure 6.3: Setup with $K > 2$ distinct survivor functions

This violates Assumption 3, and we again conclude that the survival distribution is independent of the date of initiation. In both cases, we have concluded that equality of the backward and forward recurrence time distributions $\Rightarrow f_1 \equiv f_2$, as required. ■

Remark: We point out that for $K=2$, if Assumption 3 were replaced by an assumption that the two densities, f_1 and f_2 , only crossed once (that is, we assume that there exists a unique $x > 0$ such that $f_1(x) = f_2(x)$) then the result would still hold in Case (i). In this scenario, (6.2) would imply that we must have $f_1(x_0) = f_2(x_0) \forall x_0 > 0$, since f_1 and f_2 are both densities, and to ensure that Assumption 1 is not violated.

6.1.2 The case of $K > 2$

The proof of Theorem 9 extends easily to the case where $K > 2$, that is, to the case where there are more than two distinct underlying survivor functions over time. Suppose the setup is as displayed in Figure 6.3 and refer to Figure 6.4 for an example of how the survival densities can change in the case $K = 4$.

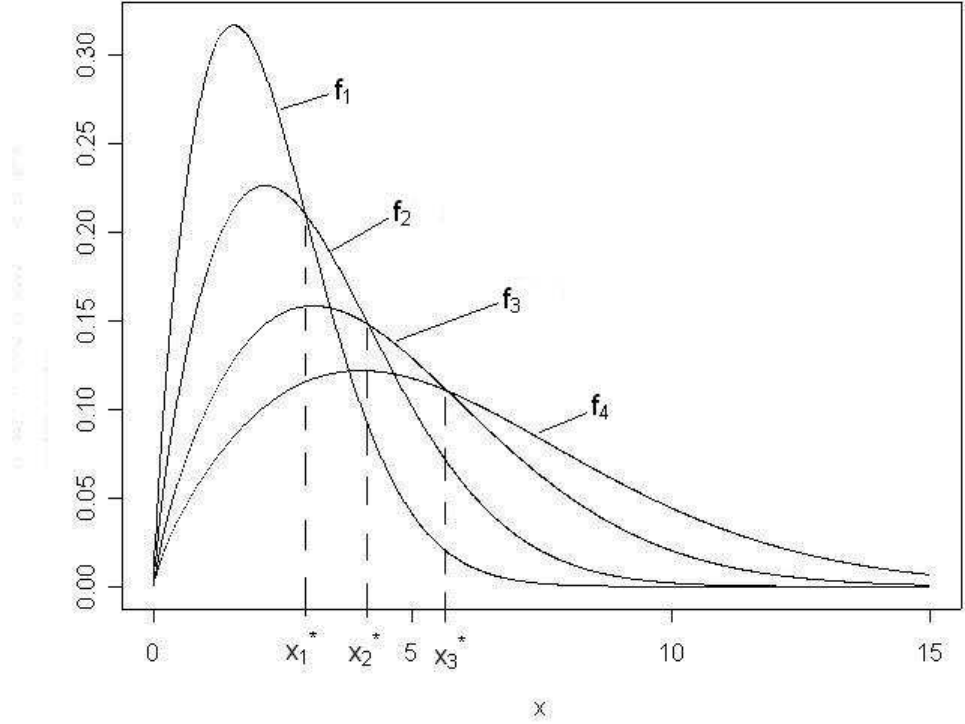


Figure 6.4: Example with 4 distinct density functions

It is not difficult to see that there exists x' such that:

$$f_1(x) \leq f_2(x) \leq f_3(x) \leq \dots \leq f_K(x) \quad \forall x \geq x',$$

$$\text{and such that } f_{x'}(x_0) = f_1(x_0) \quad \forall x_0 > 0.$$

Specifically, we could always choose $x' = \max(\tilde{x}, x_1^*, x_2^*, \dots, x_{K-1}^*)$. By restricting survival so that there are only possibly a finite number of distinct underlying survival distributions, Assumption 2 ensures the existence of such an x' .

Applying Lemma 3, we then have that,

$$f_{x'}(x_0) = f_{x_0-x'}(x_0) \quad \forall x_0 > x' .$$

Therefore, for $x_0 \in (x', x' + \tilde{x}_1)$,

$$f_1(x_0) = f_{x_0-x'}(x_0) = f_K(x_0),$$

and for $x_0 \in (x' + \tilde{x}_1, x' + \tilde{x}_1 + \tilde{x}_2)$,

$$f_1(x_0) = f_{x_0-x'}(x_0) = f_{K-1}(x_0),$$

and continuing in this fashion, we get that for $x_0 \in (x' + \tilde{x}_1 + \tilde{x}_2 + \dots + \tilde{x}_{K-2}, x' + \tilde{x}_1 + \tilde{x}_2 + \dots + \tilde{x}_{K-1})$,

$$f_1(x_0) = f_{x_0-x'}(x_0) = f_2(x_0).$$

This actually implies that,

$$f_1 \equiv f_K \text{ on } (x', x' + \tilde{x}_1),$$

$$f_1 \equiv f_{K-1} \text{ on } (x', x' + \tilde{x}_1 + \tilde{x}_2),$$

$$f_1 \equiv f_{K-2} \text{ on } (x', x' + \tilde{x}_1 + \tilde{x}_2 + \tilde{x}_3), \dots$$

$$f_1 \equiv f_2 \text{ on } (x', x' + \tilde{x}_1 + \tilde{x}_2 + \tilde{x}_3 + \dots + \tilde{x}_{K-1}) .$$

These statements contradict Assumption 3, and so we conclude that

$f_1 \equiv f_2 \equiv f_3 \equiv \dots \equiv f_K$, or that the survival distribution has not changed with calendar time of onset. Theorem 10 provides a summary of the result.

Theorem 10 *Suppose that the underlying incidence process is stationary. Then under Assumptions 1, 2, and 3, equality of the backward and forward recurrence time distributions holds if and only if $f_1 \equiv f_2 \equiv f_3 \equiv \dots \equiv f_K$, where f_1, f_2, \dots, f_K represent K , potentially distinct, underlying survival densities.*

Since K can be taken to be arbitrarily large, we conjecture that a version of Theorem 10 remains valid even if we allow survival to change *continuously* as a function of the initiation date. We have, however, not yet established a formal proof for this extension. Even in its present form, though, Theorem 10 probably captures most situations in practice since, changes in survival, if they occur, are caused by events at one or more time points (for example, the date at which a new treatment is introduced).

6.2 Characterizing alternatives to a fixed survival distribution for $K = 2$

Analogously to the results from Chapter 4, we would like to characterize “improving” or “worsening” survival over time in terms of the backward and forward recurrence time distributions, under the assumption of stationarity. Motivated by Theorem 6 and Theorem 7, we postulate the dual results:

$$Y^{fwd} >^{SL} Y^{bwd} \Leftrightarrow \text{survival is improving, and}$$

$$Y^{bwd} >^{SL} Y^{fwd} \Leftrightarrow \text{survival is worsening,}$$

where we remind the reader that the manner in which survival can “improve” or “worsen” is restricted by Assumption 1. We only consider the case $K = 2$, so the reader can refer to Figure 6.2 for the setup. Moreover, we replace Assumption 3 by the following assumption:

Assumption 3’: Let $f_1(x)$ and $f_2(x)$ denote the two underlying survival densities. We assume that f_1 and f_2 cross each other only once; that is, we assume that there exists a unique $x > 0$ such that $f_1(x) = f_2(x)$.

Theorem 11 *Suppose that the underlying incidence process is stationary. Under Assumption 1, Assumption 2 (with $K=2$), and Assumption 3', $Y^{fwd} >^{SL} Y^{bwd}$ if and only if $S_1(x) \leq S_2(x) \quad \forall x \geq 0$ (with a strict inequality holding for at least one x), where S_1 and S_2 are the survivor functions associated with f_1 and f_2 , respectively.*

Proof: Assuming stationarity of the onset times, we have that,

$$S_{bwd}(x) \propto \int_x^\infty \int_u^\infty f_u(x_0) dx_0 du, \text{ and}$$

$$S_{fwd}(x) \propto \int_x^\infty \int_u^\infty f_{x_0-u}(x_0) dx_0 du,$$

where the constant of proportionality is $P(X \geq T)$ in both expressions. We begin by assuming that $Y^{fwd} >^{SL} Y^{bwd}$, which implies that $S_{fwd}(x) - S_{bwd}(x) \geq 0 \quad \forall x \geq 0$, with strict inequality holding for at least one x . This holds if and only if,

$$\int_x^\infty \int_u^\infty [f_{x_0-u}(x_0) - f_u(x_0)] dx_0 du \geq 0 \quad \forall x \geq 0.$$

Given the restrictions on how survival can possibly change over time, there are only three possible scenarios:

- (1) $S_1(x) \equiv S_2(x)$ (survival has not changed), or
- (2) $S_1(x) \geq S_2(x) \quad \forall x$ (survival is “worsening”), or
- (3) $S_1(x) \leq S_2(x) \quad \forall x$ (survival is “improving”).

We have shown that (1) is equivalent to $S_{bwd} \equiv S_{fwd}$, and thus (1) cannot hold here. Suppose that (2) holds. Then the relationship between $f_1(x)$ and $f_2(x)$ is as depicted in Figure 6.5, where survival has changed at

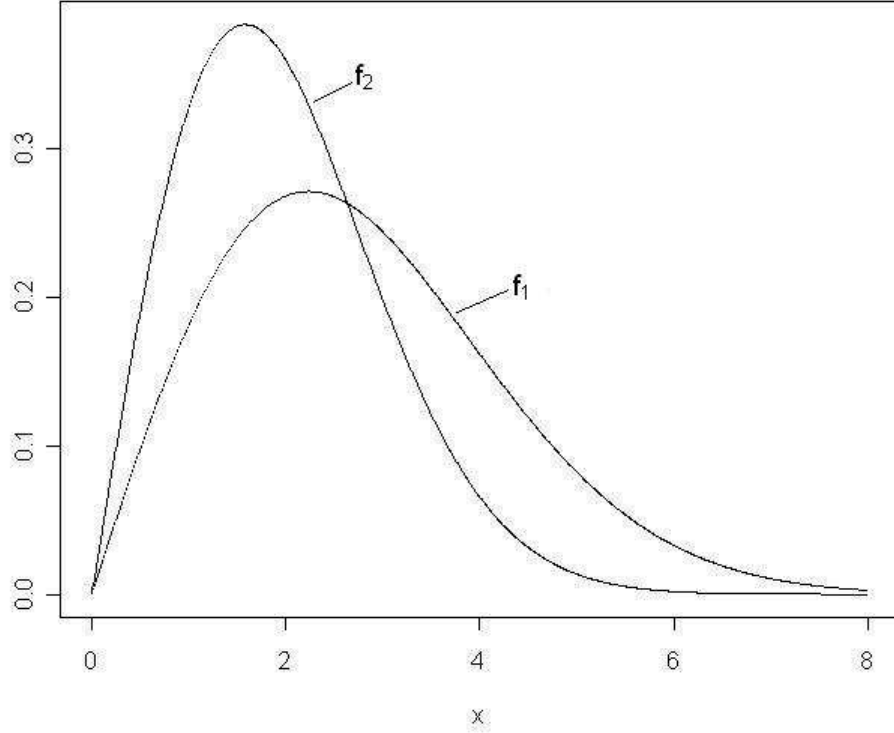


Figure 6.5: Relationship between $f_1(x)$ and $f_2(x)$ assuming (2) holds

$\tau^* - \tilde{x}$, say, as shown in Figure 6.2. Let $x' = \max(\tilde{x}, x^*)$, then $f_{x'}(x_0) = f_1(x_0) \forall x_0 > 0$. Also, we have that $f_1(x) > f_2(x) \forall x > x'$. Let

$$h(u) = \int_u^\infty [f_{x_0-u} - f_u(x_0)] dx_0 . \quad (6.3)$$

Then,

$$S_{fwd}(x') - S_{bwd}(x') \propto \int_{x'}^\infty h(u) du .$$

Note that $\forall u > x'$, $f_u(x_0) = f_1(x_0) \forall x_0 > 0$. Thus, we have that, $f_{x_0-u}(x_0) - f_1(x_0) < 0 \forall x_0 > u$, and $\forall u > x'$, which implies that,

$$h(u) = \int_u^\infty [f_{x_0-u} - f_1(x_0)] dx_0 < 0 \quad \forall u > x' .$$

This, in turn, implies that $S_{fwd}(x') - S_{bwd}(x') < 0$, which contradicts the assumption that $Y^{fwd} >^{SL} Y^{bwd}$. Hence (2) cannot hold, and (3) must hold.

For the converse, we assume that (3) holds, and we show that $S_{fwd}(x) - S_{bwd}(x) \geq 0 \ \forall x \geq 0$ (with strict inequality holding for at least one x). We proceed, as in the proof of Theorem 9, by considering two cases: (i) $x^* < \tilde{x}$, and (ii) $x^* \geq \tilde{x}$.

Case (i): $x^* < \tilde{x}$

For all $u > \tilde{x}$, we have $f_u(x_0) = f_1(x_0)$, and for all $x_0 > \tilde{x} > x^*$, $f_2(x_0) > f_1(x_0)$. With $h(u)$ defined as in (6.3), this implies that,

$$h(u) = \int_u^\infty [f_{x_0-u}(x_0) - f_1(x_0)] dx_0 > 0 \quad \forall u > \tilde{x} .$$

Thus,

$$S_{fwd}(x) - S_{bwd}(x) \propto \int_x^\infty h(u) du \geq 0 \quad \forall x \geq \tilde{x} .$$

Now, consider $F_{fwd}(x) - F_{bwd}(x)$. We know that,

$$F_{fwd}(x) - F_{bwd}(x) \propto \int_0^x \int_u^\infty [f_{x_0-u}(x_0) - f_u(x_0)] dx_0 du .$$

Let $x \in (0, \tilde{x})$. We then have that,

$$\begin{aligned} h(u) &= \int_u^\infty [f_{x_0-u}(x_0) - f_2(x_0)] dx_0 \\ &= \int_u^{u+\tilde{x}} [f_2(x_0) - f_2(x_0)] dx_0 + \int_{u+\tilde{x}}^\infty [f_1(x_0) - f_2(x_0)] dx_0 \\ &< 0 \quad \forall u \in (0, \tilde{x}) , \end{aligned}$$

which implies that $F_{fwd}(x) - F_{bwd}(x) < 0 \ \forall x \in (0, \tilde{x})$, and this holds if and

only if $S_{fwd}(x) - S_{bwd} > 0 \quad \forall x \in (0, \tilde{x})$. Therefore, we have shown that,

$$S_{fwd}(x) - S_{bwd}(x) \geq 0 \quad \forall x > 0 ,$$

which completes the proof in Case (i).

Case (ii): $x^* \geq \tilde{x}$

For all $u \geq x^* \geq \tilde{x}$, we have that $f_u(x_0) = f_1(x_0) \quad \forall x_0 > 0$. And, for all $x_0 \geq x^*$, $f_2(x_0) \geq f_1(x_0)$. Therefore, for all $x \in (x^*, \infty)$, we have that,

$$h(u) = \int_u^\infty [f_{x_0-u}(x_0) - f_1(x_0)] dx_0 \geq 0 .$$

This implies that for all $x \in (x^*, \infty)$,

$$S_{fwd}(x) - S_{bwd}(x) \propto \int_x^\infty h(u) du \geq 0 .$$

Now, let $0 < x < \tilde{x}$, and consider $F_{fwd}(x) - F_{bwd}(x)$. For $u \in (0, x)$, $f_u \equiv f_2$, since $x < \tilde{x}$. Therefore,

$$\begin{aligned} F_{fwd}(x) - F_{bwd}(x) &\propto \int_0^x \int_u^\infty [f_{x_0-u}(x_0) - f_2(x_0)] dx_0 du \\ &= \int_0^x \int_u^{u+\tilde{x}} [f_2(x_0) - f_2(x_0)] dx_0 du \\ &\quad + \int_0^x \int_{u+\tilde{x}}^\infty [f_1(x_0) - f_2(x_0)] dx_0 du \\ &= 0 + \int_0^x \int_{u+\tilde{x}}^\infty [f_1(x_0) - f_2(x_0)] dx_0 du \\ &\leq 0 . \end{aligned}$$

The last inequality follows from the fact that $\int_{u+\tilde{x}}^\infty [f_1(x_0) - f_2(x_0)] dx_0 \leq 0$, since this integral represents the area between f_1 and f_2 beyond $(u + \tilde{x})$.

Therefore, we have that $F_{fwd}(x) - F_{bwd}(x) \leq 0 \quad \forall \quad 0 < x < \tilde{x}$, or equivalently, $S_{fwd}(x) - S_{bwd}(x) \geq 0 \quad \forall \quad 0 < x < \tilde{x}$.

Finally, let $\tilde{x} < x < x^*$, which implies that $f_u \equiv f_1$ for all $u > x$.

Hence,

$$\begin{aligned}
S_{fwd}(x) - S_{bwd}(x) &\propto \int_x^\infty \int_u^\infty [f_{x_0-u}(x_0) - f_1(x_0)] dx_0 du \\
&= \int_x^\infty \int_u^{u+\tilde{x}} [f_2(x_0) - f_1(x_0)] dx_0 du \\
&= \int_x^{x+\tilde{x}} \int_x^{x_0} [f_2(x_0) - f_1(x_0)] du dx_0 \\
&\quad + \int_{x+\tilde{x}}^\infty \int_{x_0-\tilde{x}}^{x_0} [f_2(x_0) - f_1(x_0)] du dx_0 \\
&= \int_x^{x+\tilde{x}} (x_0 - x) [f_2(x_0) - f_1(x_0)] dx_0 \quad (6.4) \\
&\quad + \int_{x+\tilde{x}}^\infty \tilde{x} [f_2(x_0) - f_1(x_0)] dx_0 .
\end{aligned}$$

We now digress to show that the right hand side of (6.4) is non-negative.

We can re-write the right hand side of (6.4) more generally as,

$$\int_x^{x^*} w_1(x_0) [f_2(x_0) - f_1(x_0)] dx_0 + \int_{x^*}^\infty w_2(x_0) [f_2(x_0) - f_1(x_0)] dx_0 ,$$

where $w_1(x_0)$ and $w_2(x_0)$ are weight functions whose form depends on whether $x^* < x + \tilde{x}$ or $x^* \geq x + \tilde{x}$. Let $I_1 = \int_x^{x^*} w_1(x_0) [f_2(x_0) - f_1(x_0)] dx_0$ and $I_2 = \int_{x^*}^\infty w_2(x_0) [f_2(x_0) - f_1(x_0)] dx_0$. It is clear that $I_1 < 0$ and that $I_2 > 0$ by examining Figure 4.2. Note that $I_2 > 0$ whether Assumption 3 or Assumption 3' were employed. However, under Assumption 3, it is not necessarily

true that $I_1 < 0$ since it is possible to have $f_2 > f_1$ over an interval arbitrarily close to x^* . Assumption 3' eliminates this possibility and guarantees that $I_1 < 0$. Now, let $I_3 = \int_0^{x^*} w_1(x_0)[f_2(x_0) - f_1(x_0)]dx_0$, and observe that $I_3 < I_1 < 0$.

Suppose that $x + \tilde{x} > x^*$, then $w_1(x_0) = x_0 - x \forall x_0 \in (0, x^*)$, and $w_2(x_0) = (x_0 - x)\mathbf{1}[x^*, x + \tilde{x}] + \tilde{x}\mathbf{1}[x + \tilde{x}, \infty]$. On the other hand, if $x + \tilde{x} \leq x^*$, then $w_1(x_0) = (x_0 - x)\mathbf{1}[0, x + \tilde{x}] + \tilde{x}\mathbf{1}[x + \tilde{x}, x^*]$, and $w_2(x_0) = \tilde{x} \forall x_0 \in (x^*, \infty)$. In either situation, three statements hold: (a) the functions w_1 and w_2 are non-decreasing, (b) $w_1(y) \leq w_2(z) \forall (y, z) \in (0, x^*) \times (x^*, \infty)$, and (c) $w_1(x^*) = w_2(x^*)$. Moreover, we have that,

$$\begin{aligned} |I_3| &\leq \int_0^{x^*} w_1(x_0)|f_2(x_0) - f_1(x_0)|dx_0 \\ &\leq w_1(x^*) \int_0^{x^*} |f_2(x_0) - f_1(x_0)|dx_0 \\ &= Aw_1(x^*) , \end{aligned}$$

where $A = |\text{area between } f_1 \text{ and } f_2 \text{ in } (0, x^*)|$. Furthermore,

$$\begin{aligned} |I_2| = I_2 &\geq \int_{x^*}^{\infty} w_2(x^*)[f_2(x_0) - f_1(x_0)]dx_0 \\ &= w_2(x^*) \int_{x^*}^{\infty} [f_2(x_0) - f_1(x_0)]dx_0 \\ &= Aw_2(x^*) , \end{aligned}$$

where $A = |\text{area between } f_1 \text{ and } f_2 \text{ in } (x^*, \infty)| = |\text{area between } f_1 \text{ and } f_2 \text{ in } (0, x^*)|$. But, since $w_1(x^*) = w_2(x^*)$, we have that, $I_2 \geq Aw_1(x^*) \geq |I_3| \geq |I_1|$, which implies that $I_2 + I_1 \geq 0$, as desired.

This, in turn, implies that $S_{fwd}(x) - S_{bwd}(x) \geq 0 \quad \forall x > 0$, which completes the proof in Case (ii). Hence, we have shown that $S_1(x) \leq S_2(x) \quad \forall x \geq 0 \Rightarrow Y^{fwd} >^{SL} Y^{bwd}$, which completes the proof of Theorem 11. ■

The extension of Theorem 11 to the case $K > 2$ does not hold under the current set of assumptions. Furthermore, it is unclear what additional assumptions are needed to ensure the validity of this extension.

6.3 Utility of results

Theorem 9 and Theorem 10 provide a characterization of independence of the underlying survival distribution on calendar time of onset by equality of the backward and forward recurrence time distributions, assuming stationarity of the onset times. This implies that the test proposed for stationarity of the onset times in Chapter 3, can also be applied to test for independence of the underlying survival distribution on calendar time of onset.

In Theorem 11, we have also characterized a certain type of dependence of the survival distribution on calendar time, when there are only two distinct survival distributions. In Chapter 7, we explore the power of our test for detecting such a departure from the null hypothesis of independence of the survival distribution on calendar time of onset.

Chapter 7

Detecting dependence of survival on calendar time: a power study

In this chapter, we assume stationarity, and examine the power of our test to detect whether the underlying survival distribution depends on the calendar time of onset. To perform this power study, we generated onsets over a pre-specified interval, $(0, \tau^*)$, assuming a constant onset intensity. We then generated a survival time for each onset in the following fashion: if the onset time was in the interval $(0, \tau^* - \tilde{x})$, for some \tilde{x} , then the survival time was generated from a certain survival density (f_1), and if the onset was in $(\tau^* - \tilde{x}, \tau^*)$, then the survival time was generated from a different survival density (f_2), where f_1 and f_2 conformed to the assumptions made in Chapter 6. The observed sample consisted of the survival times which extended beyond τ^* . We constructed an α -level critical region using our test for equality of the backward and forward recurrence time distributions.

7.1 Details of the simulations

The following issues need to be considered: the specification of the underlying survival densities, f_1 and f_2 , the residual censoring time distribution, the onset intensity and the support of the truncation time distribution, $(0, \tau^*)$, where τ^* represents the calendar time of recruitment, the value of \tilde{x} , the sample size, n , and the number of replications of the entire procedure, M .

7.1.1 Underlying survival distributions

For $f_1(x)$, two scenarios were simulated:

$$\text{(A) } Weibull(\gamma = 2, \beta_1 = 10)$$

$$\text{(B) } lognormal(\mu_1 = 1.75, \sigma = 0.4)$$

where the parametrizations of the Weibull(γ, β) and lognormal(μ, σ^2) distributions are as in (5.1) and (5.2), respectively.

In determining how survival changed at $\tau^* - \tilde{x}$, we wanted to ensure that a varying degree of “improving” and “worsening” survival distributions were investigated. Thus, for both **(A)** and **(B)**, six different possibilities were chosen for $f_2(x)$. In **(A)**, $f_2(x)$ was selected as a Weibull distribution with γ fixed as in $f_1(x)$, and by changing the value of β_1 from $f_1(x)$ to some β_2 . Similarly, in **(B)**, $f_2(x)$ was chosen to be lognormal with σ as in $f_1(x)$, and by altering the value of μ_1 from $f_1(x)$ to some μ_2 . Specifically, the values of β_2 chosen in **(A)** were:

$$(1) 13.25 \quad (2) 15 \quad (3) 17 \quad (4) 7.25 \quad (5) 6 \quad (6) 5 \quad ,$$

and the values of μ_2 chosen in **(B)** were:

$$(1) 1.89 \quad (2) 1.95 \quad (3) 2.01 \quad (4) 1.59 \quad (5) 1.50 \quad (6) 1.39 \quad .$$

The values in (1), (2), and (3) for both **(A)** and **(B)** represent an improvement in survival, and correspond, respectively, to approximately a 15%, 22.5%, and 30% increase in mean survival after $\tau^* - \tilde{x}$. The values in (4), (5), and (6) represent a decline in survival, and correspond, respectively, to approximately a 15%, 22.5%, and 30% decrease in mean survival after $\tau^* - \tilde{x}$. We also simulated a seventh situation for both **(A)** and **(B)** where survival remains the same over the entire interval $(0, \tau^*)$.

7.1.2 Sample size

We used sample sizes of $n \approx 500$ and $n \approx 1000$ in the power study.

7.1.3 Value of \tilde{x}

The value of \tilde{x} determines how far from τ^* the change in survival occurred. Of course, if \tilde{x} is very small it will be difficult to detect the change in survival since very few subjects will be observed who have experienced the “new” survival distribution. Similarly, if \tilde{x} is very large it will be difficult to detect the change in survival since very few of the subjects who have experienced the “old” survival distribution will survive long enough to be observed at recruitment. Moreover, cases where \tilde{x} is very large are not interesting since this would represent a situation where the survival distribution changed a long time before recruitment. For these reasons, we chose two values of \tilde{x} which we felt were not too large or too small. For **(A)** we chose:

$$(i) \tilde{x} = 2 \quad (ii) \tilde{x} = 3 \quad ,$$

and for **(B)** we chose:

$$(i) \tilde{x} = 5 \quad (ii) \tilde{x} = 7 \quad .$$

7.1.4 Residual censoring time distribution

For each scenario, the residual censoring time distribution was chosen to be *exponential* with mean, β , selected so that approximately 25% of the forward recurrence times were censored. A different value of β was required for each distinct $f_2(x)$ in both **(A)** and **(B)**. Specifically, for **(A)** we used:

$$(1) \beta = 6.5 \quad (2) \beta = 7 \quad (3) \beta = 7.5 \quad (4) \beta = 5 \quad (5) \beta = 4.5 \quad (6) \beta = 4 \quad ,$$

and a value of $\beta = 6$ for the case where survival remained the same over the interval $(0, \tau^*)$. For **(B)** we used:

$$(1) \beta = 14 \quad (2) \beta = 14.5 \quad (3) \beta = 15 \quad (4) \beta = 10 \quad (5) \beta = 9 \quad (6) \beta = 8 \quad ,$$

and a value of $\beta = 12$ for the case where survival did not change.

7.1.5 Onset intensity and support of the truncation time distribution

All onsets were generated assuming stationarity. Thus, the onset process was chosen to be a Poisson process with a constant intensity function, $\lambda(t) \equiv \lambda$. For simulation of the stationary onsets, $\text{uniform}(0, \tau^*)$ random variables were generated and ordered. The support of the truncation time distribution was chosen to be sufficiently large to ensure that the backward recurrence time distribution was not “clipped”, as discussed in Chapter 5. The values of τ^* that were chosen are:

$$\text{(A)} \tau^* = 10 \quad \text{(B)} \tau^* = 20$$

For both **(A)** and **(B)**, we generated a sufficient number of uniform random variables over the interval $(0, \tau^*)$ in order to attain the desired sample sizes of either $n \approx 500$ or $n \approx 1000$.

7.1.6 Number of replications of the procedure

The two choices for $f_1(x)$, two sample sizes, two values for \tilde{x} , and seven different possibilities for $f_2(x)$ (six representing either improving or worsening survival after $\tau^* - \tilde{x}$, and the null case where $f_2(x) \equiv f_1(x)$) led to 56 different simulation scenarios. For each of these 56 scenarios, we performed $M=200$ replicates, and we recorded the number of rejections at the $\alpha = 0.05$ level. A two-sided rejection region was used.

7.2 Results of the power study

The percentage of rejections of the null hypothesis that the survival distribution is independent of calendar time are presented in Table 7.1-7.4. For **(A)**, the results for $\tilde{x} = 2$ and $\tilde{x} = 3$ are presented in Table 7.1 and Table 7.2, respectively. For **(B)**, the results for $\tilde{x} = 5$ and $\tilde{x} = 7$ are presented in Table 7.3 and Table 7.4, respectively. The last column of these tables represents the scenario where there was no change in survival at $\tau^* - \tilde{x}$.

β_2	13.25	15	17	7.25	6	5	10
$n \approx 500$	17.5	36.5	63.5	29.0	74.0	91.0	0.5
$n \approx 1000$	38.0	78.0	95.0	60.0	91.5	100.0	1.0

Table 7.1: Percentage of rejections for Weibull($\gamma=2, \beta_1=10$) with $\tilde{x} = 2$

β_2	13.25	15	17	7.25	6	5	10
$n \approx 500$	22.0	56.5	83.5	23.5	60.5	84.5	1.0
$n \approx 1000$	49.0	91.0	99.0	52.0	94.5	99.5	1.5

Table 7.2: Percentage of rejections for Weibull($\gamma=2, \beta_1=10$) with $\tilde{x} = 3$

μ_2	1.89	1.95	2.01	1.59	1.50	1.39	1.75
$n \approx 500$	30.5	78.0	92.0	34.5	75.5	99.0	2.0
$n \approx 1000$	65.5	96.0	100.0	72.0	99.0	100.0	2.0

Table 7.3: Percentage of rejections for lognormal($\mu_1=1.75, \sigma=0.4$) with $\tilde{x} = 5$

μ_2	1.89	1.95	2.01	1.59	1.50	1.39	1.75
$n \approx 500$	20.0	43.5	83.5	13.0	37.5	58.0	1.0
$n \approx 1000$	45.5	85.5	99.5	38.0	69.0	90.0	2.0

Table 7.4: Percentage of rejections for lognormal($\mu_1=1.75, \sigma=0.4$) with $\tilde{x} = 7$

Table 7.1 shows that, when $n \approx 1000$, our test performed well, except for the situations where the smallest change in mean survival occurred (i.e. $\beta_2 = 13.25$ and $\beta_2 = 7.25$). When $n \approx 500$, however, only the two largest drops in mean survival were adequately detected. For the scenarios displayed in Table 7.1, the performance of our test was noticeably better for the cases of worsening survival. From Table 7.2, we see that when $n \approx 500$, only the biggest changes in mean survival were detected, but increasing the sample size to $n \approx 1000$ substantially improved the power. Over 90% power was achieved for all but the smallest changes in mean survival.

Table 7.3 displays by far the highest rejection rates. For both $n \approx 500$ and $n \approx 1000$, adequate power was attained, except for the smallest changes in mean survival. For $n \approx 1000$, even the two situations with the smallest changes in mean survival had satisfactory power (65.5% for $\mu_2 = 1.89$ and 72% for $\mu_2 = 1.59$). Table 7.4 shows that our test had poor power when $n \approx 500$, with only one case, the largest increase in mean survival, being detected with a probability over 0.80. When $n \approx 1000$, the power improved substantially, but for the cases of worsening survival, only the largest drop in mean survival was adequately detected. For the scenarios displayed in Table 7.4, the performance of our test was markedly better for the cases of improving survival.

As in Chapter 5, we notice that under the null hypothesis of no change in survival over time, the rejection rate was much lower than the nominal 5% value. When $n \approx 500$, the average rejection rate was just over 1% for the four scenarios. When $n \approx 1000$, the average rejection rate was slightly higher, at approximately 1.6%, for the four scenarios.

7.3 Discussion of the results

The results reported in this chapter for testing for a change in survival over time reveal some interesting features intrinsic to this problem. At the heart of the matter is the calendar time at which the change in the survival distribution occurs. In our simulations, this was controlled through the value of \tilde{x} . We have already discussed the fact that we attempted to select values of \tilde{x} which are both not too small and not too large because in either case the ability to detect a departure from the null hypothesis would

rapidly diminish. Even within a range of “acceptable” \tilde{x} 's, however, there is a spectrum of factors which will affect the power of any test.

We believe that the results presented in Table 7.1 and Table 7.4 were dampened because of the values of \tilde{x} , with $\tilde{x} = 2$ and $\tilde{x} = 7$ being, respectively, slightly smaller and larger than “optimal”. The results also suggest that, for changes in survival which occur closer to recruitment, it is easier to detect a decline in survival over time and, for changes in survival which occur further from recruitment, it is easier to detect an improvement in survival. It is somewhat comforting that when $n \approx 1000$, this difficulty was alleviated, although in Table 7.4 we see that there was still only 69% power for detecting a “moderate” reduction in mean survival.

Table 7.2 and Table 7.3 present results which are generally similar to those obtained in Chapter 5 for testing for stationarity. Postulated reasons for the inability to consistently detect “mild” departures from the null hypothesis or for the adequate detection of “moderate” departures only when $n \approx 1000$ are the same as those identified in Chapter 5. Amongst these, we again point out that the rejection rate under the null hypothesis of no change in survival did not attain the 5% nominal value.

Chapter 8

Estimation of a constant incidence rate

8.1 Background and notation

In this chapter, we propose a method of estimating the underlying incidence intensity when it may be assumed to be constant. We have already proposed a test for *assessing* stationarity (i.e. constancy of the incidence intensity) if we are prepared to assume that the survival distribution remains the same regardless of the time origin. If we obtain insufficient evidence in favor of non-stationarity, the next step would often be to assume stationarity, and then to estimate the constant onset intensity. The data, as before, are prevalent cohort survival times having been collected as part of a prevalent cohort study with follow-up. We exploit the sampling scheme, however, that frequently gives rise to the cases in practice, and this, as we shall see, plays an important role in the estimation procedure. We assume throughout this chapter that the survival distribution is independent of calendar time of

onset and, of course, that stationarity holds. We begin by introducing some notation and defining certain key quantities.

Definition 5 *Prevalence proportion (P): The proportion of a population with a condition at a given point in time.*

Definition 6 *True mean survival time (μ): The mean survival time for an incident case, with the condition, that is, $\mu = \int_0^\infty S(x)dx$, where $S(x)$ represents the underlying survivor function.*

Also, recall that the definition of an incidence intensity function, $\lambda(t)$, for a certain disease is the instantaneous probability of getting the disease at time t , given that one is non-diseased just before t (see Definition 1 from Chapter 2). This chapter is concerned with the estimation of $\lambda(t) \equiv \lambda$, which is assumed to be constant as a function of time. Although it would be entirely plausible, we do not allow $\lambda(t)$ to depend on the age of the subject in Definition 1. Age specific incidence rates can be estimated by performing a stratified analysis. In the sequel, we omit the potential dependence of the intensities on age.

Estimation of $S(x)$, when using prevalent cohort data, is usually achieved by adopting either a conditional or unconditional approach. We remind the reader that the conditional approach is so called because it is carried out conditional on the observed truncation times (backward recurrence times), and that the unconditional approach relies instead on the assumption of stationarity (refer to Chapter 2 for more details). Irrespective of whether one conditions on the observed truncation times or not, most authors implicitly condition on the number of cases of the disease, n ; the number of cases is almost always random, having been obtained by screening a fixed

number of subjects. In the statistical literature that uses the unconditional approach, the likelihood, which is still conditional on n , for these data can be expressed as follows:

$$L = \prod_{i=1}^n \left(dF_{LB}(y_i) \right)^{\epsilon_i} \left(\int_{c_i^*}^{\infty} \frac{dF_{LB}(z)}{z} \right)^{1-\epsilon_i}, \quad (8.1)$$

where $F_{LB}(x)$ represents the density of the left truncated lifetimes, y_i , and the remaining notation is as in Chapter 2. A development of (8.1) can be found, for example, in Section 2.2 of Asgharian and Wolfson (2005). We notice that there is no consideration of the sampling scheme in (8.1). That is, we do not account for how the n cases are identified. As a result, simultaneous inference on $F_{LB}(x)$ (and hence $S(x)$) and on λ is not possible. Intuitively, this non-identifiability can be explained as follows: observing many cases, say, could be due to subjects experiencing “long” survival with the condition, or due to a “high” incidence rate. Since the analysis is performed by conditioning on the number of cases there is no way of separating out the effect of these two potential explanations.

Wang (1991) carried out a conditional analysis, in the sense described in Chapter 2, and showed how one can simultaneously estimate $S(x)$ and $G(t)$, where $G(t)$ is the truncation time distribution function. In general, (2.2) leads to the following relation between $G(t)$ and $\lambda(t)$, for $0 < t < \tau^*$,

$$\begin{aligned} G(t) &= \frac{\int_0^t \lambda(\tau^* - x) dx}{\Lambda(\tau^*)} \\ &= \frac{\int_{\tau^*-t}^{\tau^*} \lambda(x) dx}{\Lambda(\tau^*)} \end{aligned}$$

$$\begin{aligned}
&= \frac{\Lambda(\tau^*) - \Lambda(\tau^* - t)}{\Lambda(\tau^*)} \\
&= 1 - \frac{\Lambda(\tau^* - t)}{\Lambda(\tau^*)} \\
&= 1 - \frac{\int_0^{\tau^* - t} \lambda(x) dx}{\int_0^{\tau^*} \lambda(x) dx} . \tag{8.2}
\end{aligned}$$

Under stationarity, (8.2) reduces to,

$$G(t) = 1 - \frac{(\tau^* - t)\lambda}{\tau^*\lambda} \tag{8.3}$$

$$= \frac{t}{\tau^*} . \tag{8.4}$$

From (8.2), we see that $G(t)$ is not altered if we replace $\lambda(x)$ by $\lambda^*(x) = c\lambda(x)$, where c is an arbitrary constant. Moreover, from (8.4), we see that $G(t)$ corresponds to the distribution function of a uniform random variable under stationarity, but more importantly, (8.3) and (8.4) demonstrate that this result holds independently of the value of λ .

We have demonstrated that, when conditioning on the observed number of cases, n , estimation of λ cannot be carried out. In most applications, the number of cases is not fixed a priori. Instead, we often fix the number of subjects screened, s , and the number of cases is then random. Denoting the random number of cases by N , we ascertain the respective onset times of the N cases, and follow them until failure or censoring. In this chapter, we use the added information available from knowing s and N , in order to estimate the assumed constant incidence rate, λ .

Keiding (1991) investigated nonparametric estimators of the incidence intensity using three types of cross-sectional data. Two of the types of data discussed in Keiding’s paper involved information on the $s - N$ (in our notation) “non-diseased” subjects, but none of these types of data include follow-up of the subjects, as we do in this thesis. Furthermore, Keiding made use of two assumptions which we are not prepared to adopt. Before presenting these two assumptions, we give three definitions that will aid in our discussion.

Definition 7 *Death intensity in the non-diseased ($\nu_{ND}(t)$): The instantaneous probability of dying at time t , given that one is alive and non-diseased just before t . That is,*

$$\nu_{ND}(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq A < t + \Delta t \mid A \geq t, \tau \geq t)}{\Delta t},$$

where A represents the calendar time of failure and τ represents the calendar time of onset of the disease.

In Definition 7, the intensity $\nu_{ND}(t)$ depends only on calendar time, ‘ t ’.

Definition 8 *Death intensity in the diseased ($\nu_D(t, d)$): The instantaneous probability of dying at time t , given that one is alive just before t and has been diseased since $t - d$. That is,*

$$\nu_D(t, d) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq A < t + \Delta t \mid A \geq t, \tau = t - d)}{\Delta t},$$

where A represents the calendar time of failure and τ represents the calendar time of onset of the disease.

In Definition 8, the intensity $\nu_D(t, d)$ depends on calendar time, ‘ t ’, and duration of the condition, ‘ d ’.

Definition 9 *Non-differential mortality: An assumption that the death intensity in the diseased and in the non-diseased are identical, that is, $\nu_{ND}(t) \equiv \nu_D(t, d)$.*

Keiding presented an analysis assuming that non-differential mortality holds. If we are not prepared to make the assumption of non-differential mortality, then Keiding proposes another analysis which relies on the assumption that the functions $\nu_{ND}(t)$ and $\nu_D(t, d)$ are known. We believe that, frequently, neither of these two scenarios is realistic, and propose a different estimation procedure for λ .

Keiding (1991) also discussed the relationship between the incidence rate, prevalence, and mean duration of some condition. He made explicit the assumptions needed to validate the classical epidemiological equation,

$$\textit{“prevalence} = \textit{incidence} \times \textit{mean duration”}. \quad (8.5)$$

Assuming a three state illness-death process for the life history of an individual, Keiding showed that (8.5) holds under the assumption of calendar time homogeneity of $\nu_{ND}(t)$, $\nu_D(t, d)$, and of $\lambda(t)$. That is, assuming that $\nu_{ND}(t) \equiv \nu_{ND}$, $\nu_D(t, d) \equiv \nu_D(d)$, and $\lambda(t) \equiv \lambda$.

Since we are assuming in this chapter that survival with the condition is independent of calendar time of onset and that stationarity holds, we need only further assume that $\nu_{ND}(t)$ is independent of calendar time to ensure that (8.5) is valid. A simple estimator, $\hat{\lambda}$, of λ , which we propose in the next section, is based on (8.5). We subsequently show that, in fact, $\hat{\lambda}$ represents the NPMLE of λ when the common sampling scheme of fixing s and allowing N to be random is incorporated into the likelihood.

8.2 A simple point estimator of λ

Using our notation, (8.5) can be rewritten as,

$$P = \lambda\mu \tag{8.6}$$

The idea behind our *ad hoc* estimator is to use the prevalent cohort survival times to estimate μ , and to use knowledge of the number of subjects screened, s , to estimate P , and hence to obtain $\hat{\lambda} = \frac{\hat{P}}{\hat{\mu}}$.

Suppose that a random sample of s subjects is screened at calendar time τ^* , where s is a fixed constant. Let N represent the random number of cases ascertained amongst the s subjects. An estimator, \hat{P} , of P , is obtained from the observed prevalence,

$$\hat{P} = \frac{N}{s} . \tag{8.7}$$

Let $\hat{S}(x)$ be the unconditional, in the sense described in Chapter 2, NPMLE of the underlying survivor function, $S(x)$, conditional on the number of cases, $N = n$, identified from the s screened individuals. Let $\hat{\mu}$ be the estimator of μ obtained from $\hat{S}(x)$. That is,

$$\hat{\mu} = \int_0^\infty \hat{S}(x)dx . \tag{8.8}$$

Define the estimator, $\hat{\lambda}$, of λ to be:

$$\hat{\lambda} = \frac{\hat{P}}{\hat{\mu}} . \tag{8.9}$$

We show that $\hat{\lambda}$ is consistent for λ and that, although it is proposed in an *ad hoc* fashion, $\hat{\lambda}$ is the NPMLE for λ . We first present two corollaries and a lemma which aid in our proof of the consistency of $\hat{\lambda}$. Conditional on the observation of $N = n$ cases, Asgharian et al. (2002) proved the following corollary:

Corollary 1 *As $n \rightarrow \infty$, $\sqrt{n}(\hat{\mu} - \mu) \rightarrow X$ in distribution, where $X \sim \text{normal}(0, \psi^2)$,*

and where an expression for ψ^2 is provided by Asgharian et al. (2002), and later corrected by Asgharian and Wolfson (2005).

We assume, however, that as $s \rightarrow \infty$, $N \rightarrow \infty$ with probability 1. Using a result from Richter (1965), we can thus extend Corollary 1 as follows:

Corollary 2 *As $s \rightarrow \infty$, $\sqrt{N}(\hat{\mu} - \mu) \rightarrow X$ in distribution, where X is exactly as in Corollary 1.*

As a consequence of Corollary 2, we have the following lemma.

Lemma 4 *As $s \rightarrow \infty$, $\hat{\mu} \rightarrow \mu$ in probability. That is, $\hat{\mu}$ is weakly consistent for μ .*

Proof: For all $x \in \mathbb{R}$, $\lim_{s \rightarrow \infty} P\left(\frac{\sqrt{N}}{\psi}(\hat{\mu} - \mu) \leq x\right) = \Phi(x)$, where $\Phi(x)$ represents the standard normal distribution function. Alternatively, given $\varepsilon > 0$ and x , there exists s_0 such that for all $s > s_0$,

$$\left| P\left(\frac{\sqrt{N}}{\psi}(\hat{\mu} - \mu) \leq x\right) - \Phi(x) \right| < \varepsilon .$$

Also, given $\varepsilon > 0$ there exists $x_1 > 0$ and $x_2 < 0$ such that $\Phi(x_1) > 1 - \varepsilon$ and $\Phi(x_2) < \varepsilon$.

We want to show that for all $\varepsilon > 0$, $\lim_{s \rightarrow \infty} P(|\hat{\mu} - \mu| > \varepsilon) = 0$, or, given $\varepsilon > 0$ and $\delta > 0$, there exists s_0 such that for all $s > s_0$,

$$P(|\hat{\mu} - \mu| > \varepsilon) < \delta .$$

We have that $P(|\hat{\mu} - \mu| > \varepsilon) = P(\hat{\mu} - \mu > \varepsilon) + P(\hat{\mu} - \mu < -\varepsilon)$. We show that $P(\hat{\mu} - \mu > \varepsilon)$ can be made arbitrarily small, and the argument for $P(\hat{\mu} - \mu < -\varepsilon)$ is similar.

Given $\varepsilon > 0$ and $\delta > 0$, choose $x_1 > 0$ large enough to ensure that $\Phi(x_1) > 1 - \frac{\delta}{2}$. Then,

$$\begin{aligned} P(\hat{\mu} - \mu > \varepsilon) &= 1 - P(\hat{\mu} - \mu \leq \varepsilon) \\ &= 1 - P\left(\frac{\sqrt{N}}{\psi}(\hat{\mu} - \mu) \leq \frac{\sqrt{N}}{\psi}\varepsilon\right) . \end{aligned}$$

For large enough s , $\frac{\sqrt{N}}{\psi}\varepsilon > x_1$ and $\left|P\left(\frac{\sqrt{N}}{\psi}(\hat{\mu} - \mu) \leq x_1\right) - \Phi(x_1)\right| < \frac{\delta}{2}$.

Thus,

$$\begin{aligned} P\left(\frac{\sqrt{N}}{\psi}(\hat{\mu} - \mu) \leq \frac{\sqrt{N}}{\psi}\varepsilon\right) &\geq P\left(\frac{\sqrt{N}}{\psi}(\hat{\mu} - \mu) \leq x_1\right) \\ &> \Phi(x_1) - \frac{\delta}{2} \\ &> \left(1 - \frac{\delta}{2}\right) - \frac{\delta}{2} = 1 - \delta . \end{aligned}$$

Hence, $P(\hat{\mu} - \mu > \varepsilon) < 1 - (1 - \delta) = \delta$, which implies that it can be made arbitrarily small for sufficiently large values of s . We therefore have that $\lim_{s \rightarrow \infty} P(|\hat{\mu} - \mu| > \varepsilon) = 0$ for all $\varepsilon > 0$, as required. ■

From Lemma 4, we obtain the following corollary.

Corollary 3 *As $s \rightarrow \infty$, $\hat{\lambda} \rightarrow \lambda$ in probability. That is, $\hat{\lambda}$ is weakly consistent for λ .*

Proof: By the Strong Law of Large Numbers (SLLN), we have that $\hat{P} \rightarrow P$ almost surely (and hence in probability) as $s \rightarrow \infty$. Moreover, $\hat{\mu}$ converges in probability to μ by Lemma 4. Since $h(x, y) = \frac{x}{y}$ is a continuous function from $\mathbb{R}^+ \times \mathbb{R}^+$ to \mathbb{R}^+ , it follows that $\hat{\lambda} \equiv \frac{\hat{P}}{\hat{\mu}} \rightarrow \frac{P}{\mu} \equiv \lambda$ in probability as $s \rightarrow \infty$. ■

We assume that the number of cases, N , is random having been identified through the screening of a fixed number of subjects, s . By accounting for this sampling mechanism we establish the following theorem.

Theorem 12 *$\hat{\lambda}$ is the NPMLE for the incidence rate, λ .*

Proof: Adopting an unconditional approach, in the sense described in Chapter 2, the likelihood displayed in (8.1) is a conditional likelihood, conditional on observing $N = n$ cases. The full likelihood for the data under consideration here, which includes the observation of N cases out of s subjects screened, is thus a product of the likelihood from (8.1), and the binomial probability mass function,

$$\binom{s}{N} P^N (1 - P)^{s-N} . \quad (8.10)$$

Since the likelihood in (8.1) is independent of P , the joint NPMLE of $(F_{LB}(x), P)$, which maximizes the full likelihood, can be obtained by independently maximizing (8.1) and (8.10). That is, we obtain the NPMLE

of $F_{LB}(x)$ from the conditional likelihood in (8.1), and the NPMLE of P from the marginal likelihood in (8.10).

Since $\hat{S}(x)$, the unconditional NPMLE of $S(x)$, is obtained through a transformation of the NPMLE of $F_{LB}(x)$, and since $\hat{\mu}$, defined in (8.8), is a function of $\hat{S}(x)$, this implies that $\hat{\mu}$ is the NPMLE of μ . Furthermore, (8.10) is maximized at $\frac{N}{s}$, which signifies that \hat{P} , defined in (8.7), is the NPMLE of P . Since $\hat{\lambda}$, defined in (8.9), is a function of the NPMLE's $\hat{\mu}$ and \hat{P} , this implies that $\hat{\lambda}$ is the NPMLE of $\frac{P}{\mu} \equiv \lambda$. ■

8.3 An interval estimator of λ

We develop an interval estimator of λ that is based on the asymptotic distribution of $\frac{1}{\lambda}$. The asymptotic distribution of $\frac{1}{\lambda}$ is easily obtained as a further corollary of Corollary 2 from Section 8.2, which provides the asymptotic distribution of $\hat{\mu}$.

Corollary 4 *As $s \rightarrow \infty$, $\sqrt{N}\left(\frac{1}{\hat{\lambda}} - \frac{1}{\lambda}\right) \rightarrow \tilde{X}$ in distribution, where $\tilde{X} \sim normal\left(0, \left(\frac{\psi}{P}\right)^2\right)$.*

Proof: Since $\frac{1}{\hat{P}} \rightarrow \frac{1}{P}$ in probability as $s \rightarrow \infty$, we have by Slutsky's theorem that, as $s \rightarrow \infty$,

$$\sqrt{N}\left(\frac{\hat{\mu} - \mu}{\hat{P}}\right) \equiv \sqrt{N}\left(\frac{1}{\hat{\lambda}} - \frac{1}{\lambda}\right) \rightarrow \frac{X}{P} \text{ in distribution, as required. } \blacksquare$$

Thus, asymptotically,

$$\sqrt{N}\frac{\left(\frac{1}{\hat{\lambda}} - \frac{1}{\lambda}\right)}{\frac{\psi}{P}} \sim normal(0, 1) . \quad (8.11)$$

From (8.11), we have that for large s ,

$$P\left(-z_{\alpha/2} < \sqrt{N} \frac{\left(\frac{1}{\hat{\lambda}} - \frac{1}{\lambda}\right)}{\frac{\psi}{P}} < z_{\alpha/2}\right) \approx 1 - \alpha,$$

which implies that an approximate $100(1 - \alpha)\%$ confidence interval for $\frac{1}{\lambda}$ is:

$$\left(\frac{1}{\hat{\lambda}} - \frac{\psi_N}{P} z_{\alpha/2}, \frac{1}{\hat{\lambda}} + \frac{\psi_N}{P} z_{\alpha/2}\right),$$

where we define $\psi_N = \frac{\psi}{\sqrt{N}}$. Solving for λ yields the following interval:

$$\left(\frac{1}{\frac{1}{\hat{\lambda}} + \frac{\psi_N}{P} z_{\alpha/2}}, \frac{1}{\frac{1}{\hat{\lambda}} - \frac{\psi_N}{P} z_{\alpha/2}}\right),$$

which depends on the unknown quantities ψ_N and P . We replace P by \hat{P} and ψ_N by a consistent estimator $\hat{\psi}_N$, which we obtain by performing a nonparametric bootstrap as described in Section 8.4. Finally, we substitute $\hat{\lambda} = \frac{\hat{P}}{\hat{\mu}}$ to obtain an approximate $100(1 - \alpha)\%$ confidence interval for λ , which is displayed in (8.12),

$$\left(\frac{\hat{P}}{\hat{\mu} + \hat{\psi}_N z_{\alpha/2}}, \frac{\hat{P}}{\hat{\mu} - \hat{\psi}_N z_{\alpha/2}}\right). \quad (8.12)$$

8.4 Estimation of ψ_N

Asgharian et al. (2002) provided an expression for ψ^2 , which is incorrect because it is based on an incorrect operator, \mathcal{G}_2 . The operator \mathcal{G}_2 is corrected in Asgharian and Wolfson (2005) and although it has not yet been formally established, the correct expression for ψ^2 is of the same form as in Asgharian et al. (2002), except for the modification to \mathcal{G}_2 . The complex form for the expression, however, precludes the possibility of its direct estimation, and in practice we recommend the use of a bootstrap estimator, $\hat{\psi}_N$, of ψ_N . While the appropriateness of the bootstrap method to obtain a consistent estimator of ψ_N has yet to be established, we assume consistency of $\hat{\psi}_N$.

Deciding on an appropriate bootstrap procedure to obtain an estimate of ψ_N is somewhat delicate. Although $\psi_N = \frac{\psi}{\sqrt{N}}$, which only depends on the N identified cases, is the approximate large sample standard deviation of $\hat{\mu}$, recall that N is random, having been obtained from the screening of s subjects. Since the bootstrapping procedure should mimic the experiment as closely as possible, we recommend bootstrapping as follows to obtain $\hat{\psi}_N$:

1. Obtain a bootstrap sample of size s by performing simple random sampling with replacement from the s subjects.
2. Identify the number, N^* , in the bootstrap sample that correspond to cases of the condition. Recall that it is entirely plausible that $N^* \neq N$.
3. Obtain $\hat{S}(x)$, the unconditional NPMLE of the unbiased, or underlying, survivor function, $S(x)$, from the N^* cases identified in 2.
4. Calculate $\hat{\mu}$ from $\hat{S}(x)$ using (8.8).

5. Repeat steps 1., 2., 3., and 4. W times, and label the W estimates of μ by $\hat{\mu}_1, \hat{\mu}_2, \dots, \hat{\mu}_W$.
6. Define $\hat{\psi}_N$ to be the sample standard deviation of $\hat{\mu}_1, \hat{\mu}_2, \dots, \hat{\mu}_W$.

Having obtained $\hat{\psi}_N$, an approximate $100(1 - \alpha)\%$ confidence interval for λ can be calculated using (8.12).

8.5 Synopsis

Keiding (1991) showed that (8.5) holds under calendar time homogeneity of the intensities $\nu_{ND}(t)$, $\nu_D(t, d)$, and $\lambda(t)$. In particular, if one is willing to assume that $S(x)$ has not changed over time and that the underlying incidence process is stationary, then (8.5) is valid. Estimation of the constant incidence intensity, λ , then becomes a question of interest. In Section 8.2 we proposed a point estimator, $\hat{\lambda}$, of λ , which originates from (8.5). We showed that $\hat{\lambda}$ is consistent for λ , and that it transpires that $\hat{\lambda}$, which is introduced in an *ad hoc* fashion, represents the NPMLE of λ . We developed an interval estimator for λ as well.

In Chapter 9, we apply some of the methods from this thesis to data collected as part of the Canadian Study of Health and Aging (CSHA). In particular, we will use the formal test for stationarity, proposed in Chapter 3, to assess the stationarity of the incidence process of onsets of dementia. We show that there is insufficient evidence from the data to conclude that the incidence process of onsets of dementia is non-stationary. Point and interval estimates of the incidence rate, λ , using the estimators developed in Chapter 8, are computed.

Chapter 9

Application to the CSHA

9.1 Overview of the CSHA

The CSHA is a nationwide longitudinal study of dementia among the elderly in Canada and remains one of the largest epidemiological studies of dementia ever conducted. To date, the CSHA has completed three phases. Data collected as part of the first two phases (CSHA-1 and CSHA-2) are of interest here. One of the main objectives of CSHA-1 was to estimate the prevalence of dementia in Canada in an elderly population (at least 65 years of age). In CSHA-1, which was carried out in 1991, 10,263 subjects over the age of 65 were cross-sectionally recruited from across Canada. These subjects were selected in the following age groups: 65-74 years, 75-84 years, and 85 years and over. Since the size of the population and the prevalence of dementia vary by age, oversampling of older subjects was carried out to optimize the yield of cases with dementia. The sampling fraction in the 75-84 age group was twice that in the 65-74 age group and the fraction in the 85 and over age group was 2.5 times that in the 65-74 age group (CSHA

working group, 1994 and Wolfson et al., 2001). Those subjects who were identified with dementia at CSHA-1 were followed until death, if it occurred before 1996, or censoring. Censoring occurred if an individual identified with dementia at CSHA-1 was still alive when CSHA-2 was completed in 1996. In addition to the ascertainment of death or right censoring for those cases of dementia identified at CSHA-1, CSHA-2 included a reevaluation for dementia for those subjects who were deemed not to have dementia at CSHA-1. Amongst the objectives of CSHA-2 was the estimation of the incidence rate of dementia in Canada in an elderly cohort. The researchers assumed that the incidence rate had remained constant over time, and estimated it using the incident cases observed between 1991 and 1996 (CSHA working group, 2000). Since the estimation procedure used for the incidence rate was predicated on the assumption that it was constant, it is important to formally examine the validity of this assumption, using *only* the prevalent cohort data obtained at CSHA-1 (and their follow-up failure/censoring times until the end of CSHA-2 in 1996). In this chapter, we show that the data are consistent with the assumption of stationarity of the incidence process of onsets of dementia. Furthermore, we estimate the constant incidence rate of dementia in Canada amongst those 65 years of age or older using the same prevalent cohort data available from CSHA-1 and CSHA-2.

Clinical examinations at CSHA-1 resulted in 1132 subjects being diagnosed with some form of dementia. Following Wolfson et al. (2001), we restrict the definition of dementia to a diagnosis of one of, “possible Alzheimer’s disease”, “probable Alzheimer’s disease”, or “vascular dementia”. The 175 subjects who were considered to have “other” or “unclassified dementia” were excluded from the analysis. The date of onset of dementia

was ascertained from the answers to the following three questions:

1. When did the subject first see a doctor about memory problems?
2. When did memory problems first affect the subject's life?
3. What is the duration of the memory problems?

The answer to 1. was used as the date of onset, as long as it was not missing. If the answer to 1. was missing, then the answer to 2. was taken as the date of onset, and if the answer to 2. was also missing, then the answer to 3. was taken as the date of onset. If all three questions were unanswered, the date of onset was defined to be missing. The date of onset was missing for 185 subjects, of which 51 were also in the group of 175 with other or unclassified dementia. Moreover, 2 subjects were excluded from the analysis since they were identified as unlikely to have had one of the above three dementias, having survived over 50 years with dementia. To summarize, the CSHA-1 phase produced 821 subjects, identified as having either possible Alzheimer's disease, probable Alzheimer's disease, or vascular dementia, for whom the date of onset was reliably ascertained. This group of 821 subjects is used for the assessment of stationarity presented in Section 9.2.

9.2 Testing the assumption of stationarity of the incidence rate of dementia

We computed the test statistic $\sqrt{n}W_n/\hat{\sigma}_0$, where $n = 821$, and W_n and $\hat{\sigma}_0^2$ are respectively defined in (3.3) and (3.5) from Chapter 3. The observed value of the test statistic was approximately 0.98, which is to be compared

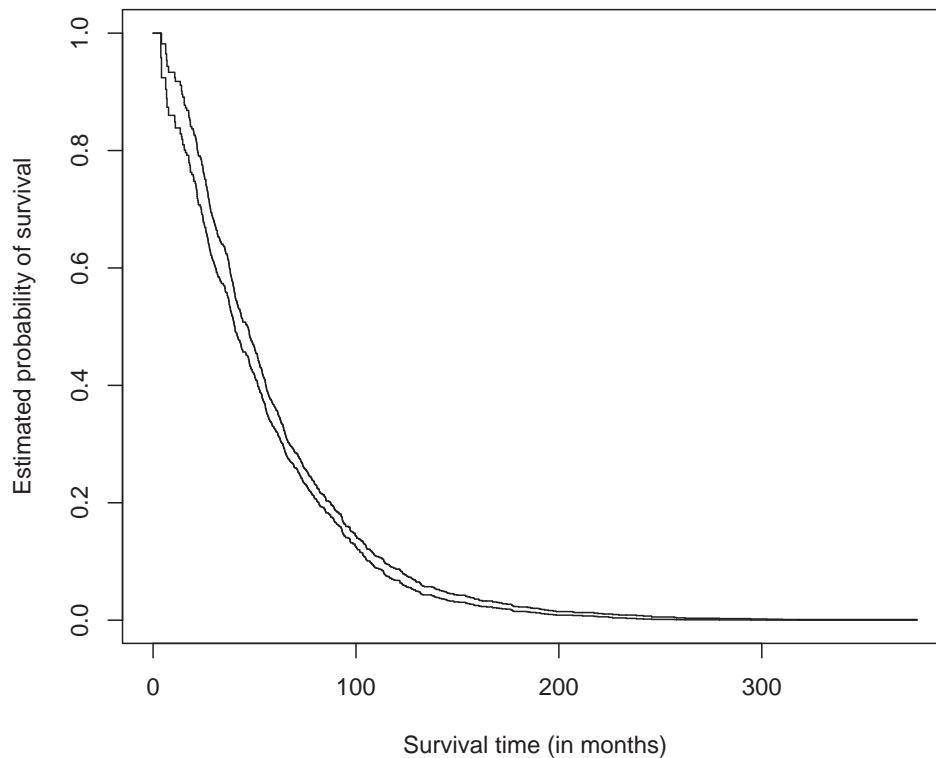


Figure 9.1: Conditional (below) and Unconditional (above) NPMLE of $S(x)$ to a standard normal distribution, yielding a two-sided p-value of roughly 0.33. This is consistent with a constant incidence rate of dementia.

Our conclusion agrees with the informal checks of stationarity carried out by Asgharian et al. (2004) and Asgharian et al. (2002). Asgharian et al. (2002) compared the conditional and unconditional NPMLE of $S(x)$. The conditional NPMLE does not rely of the assumption of stationarity, while stationarity is explicitly assumed to obtain the unconditional NPMLE. Since the two curves (which are reproduced in Figure 9.1) lie close together, this suggests that stationarity is reasonable. According to Theorem 1 from Chapter 3, due to Asgharian et al. (2004), the backward and forward recurrence time distributions are identical under stationarity. Thus, the authors

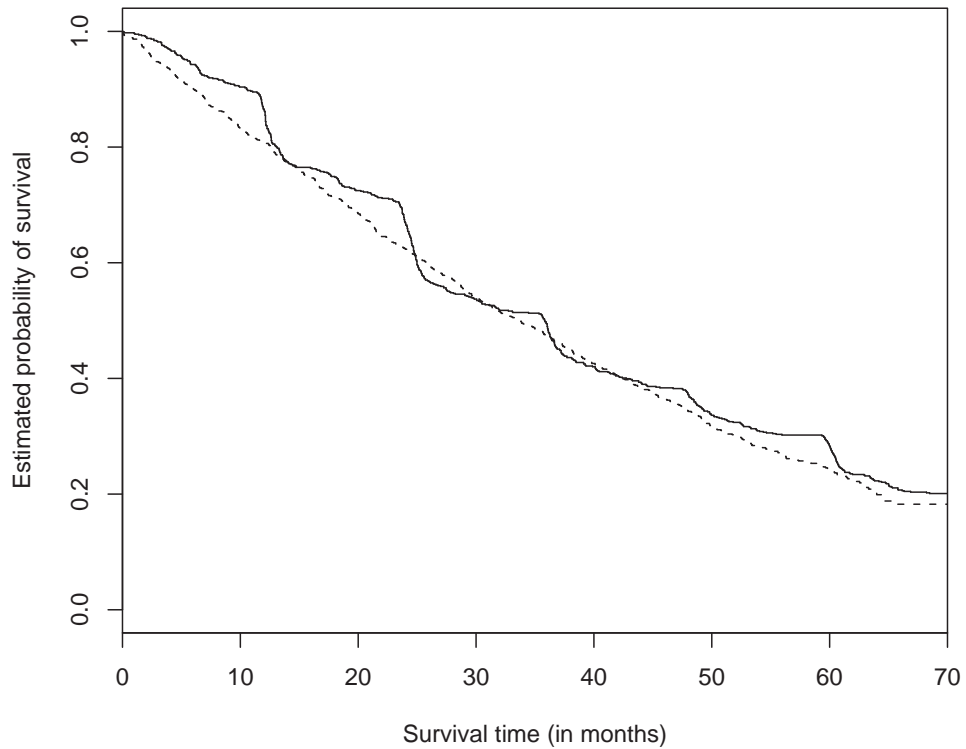


Figure 9.2: Kaplan-Meier estimates of the backward (solid) and forward (dotted) recurrence time survivor functions

compared the Kaplan-Meier estimates of the backward and forward recurrence time survivor functions and the proximity of the two curves provided evidence in favor of stationarity. The two Kaplan-Meier curves are displayed in Figure 9.2. Commenges et al. (2004) analyzed data from an incident cohort study and found that incidence rates for Alzheimer’s disease had increased over time. The authors provide reasons why this claimed increase in the incidence of dementia might be spurious.

Interestingly, Figure 9.2 reveals that the Kaplan-Meier estimate of the backward recurrence time survivor function is not as “smooth” as the Kaplan-Meier estimate for the forward recurrence time survivor function,

and behaves almost like a step function. This is most certainly due to the fact that caregivers tended to recall the onset dates of dementia in multiples of one year. Upon closer inspection of Figure 9.2, it is clear that the “steps” occur roughly every twelve months.

Although the observed test statistic, $\sqrt{n}W_n/\hat{\sigma}_0 \approx 0.98$, was not statistically significant, its positive value suggests slightly longer backward recurrence times in comparison to the forward recurrence times. This result concurs with the observation made by Asgharian et al. (2004) that the Kaplan-Meier estimate of the backward recurrence time survivor function lies slightly above that of the forward recurrence time survivor function for times close to the origin (see Figure 9.2). The authors provide one plausible reason for this occurrence, stating that tests screening for Alzheimer’s disease may not be sensitive enough to detect the condition when onset occurred close to recruitment at CSHA-1. We add another potential explanation for this result. As noted in Wolfson et al. (2001), stratified sampling was carried out with over 9,000 subjects living in the community and over 1,200 living in institutions. This represented an oversampling of subjects living in institutions. Moreover, it is likely that the cases of dementia for subjects living in institutions were more advanced than for those living at home. This provides another reason for the observation of an excess of longer backward recurrence times near the origin.

Having found that the observed data are consistent with a constant incidence of dementia, the next step is to estimate this constant incidence rate. Recall, from Chapter 1, that it is not possible to estimate the incidence rate from a prevalent cohort study with follow-up if the incidence rate is non-constant. We emphasize that the CSHA working group (2000) estimated

the incidence rate, using incident cohort data, under the assumption that it was constant, whereas, we have formally corroborated the constancy of the incidence rate using only prevalent cohort data and now estimate the incidence rate, again using the same data, and the added knowledge of the fixed number screened to identify the cases.

9.3 Estimating the incidence rate of dementia

We calculate point and interval estimates for the incidence rate of dementia, as defined in this thesis, amongst the elderly in Canada, using the estimators introduced in Chapter 8.

An estimate, $\hat{\mu}$, of μ , the mean survival time of an incident case of dementia in Canada amongst those 65 years of age or older, can be obtained using (8.8). From the unconditional NPMLE, $\hat{S}(x)$, of $S(x)$, displayed in Figure 9.1, we obtained $\hat{\mu} \approx 4.75$ years or 57.0 months. Of the 1132 cases identified at CSHA-1, 957 were diagnosed as having possible Alzheimer's, probable Alzheimer's, or vascular dementia. For the purposes of estimating P , the prevalence of dementia in Canada amongst those 65 years of age or older, we cannot simply use (8.7) since the CSHA data did not constitute a random sample of all subjects over the age of 65 in Canada. Hence, using $\frac{957}{10,263} \approx 0.093$ as our estimate would represent an overestimate of P . Instead, we use the age-standardized prevalence of 0.066 or 6.6% (CSHA working group, 1994). For possible Alzheimer's or probable Alzheimer's, the estimated prevalence was 5.1% and for vascular dementia the estimated prevalence was 1.5%, yielding the prevalence estimate $\hat{P} \approx 0.066$. The CSHA working group (1994) also calculated an age standardized prevalence

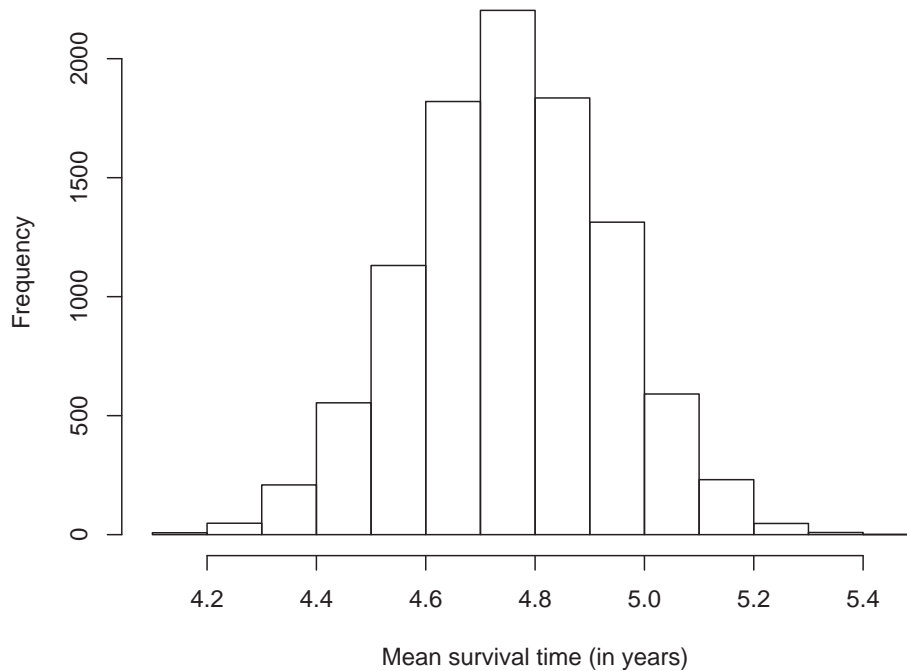


Figure 9.3: Histogram of 10,000 bootstrap estimates of μ

of 8.0% when all forms of dementia are included. We ignore the oversampling of institutionalized subjects to arrive at our estimate of prevalence.

Using $\hat{\mu} \approx 4.75$ years and $P \approx 0.066$, we obtain an estimate of λ , the incidence of dementia in Canada amongst those 65 years of age or older, from (8.9). We have that $\hat{\lambda} \approx \frac{0.066}{4.75} \approx 0.0139$, or 13.9 per 1,000 person-years. In order to use (8.12) to obtain a confidence interval for λ , we need $\hat{\psi}_N$, an estimate of the asymptotic standard deviation of $\hat{\mu}$. Using the bootstrap procedure described in Chapter 8, we obtained 10,000 bootstrap estimates of μ . A histogram of these estimates is displayed in Figure 9.3, where the asymptotic normality of $\hat{\mu}$ is evident, and in accordance with Corollary 1 (Asgharian et al., 2002) and Corollary 2 from Chapter 8. We arrived at an estimate, $\hat{\psi}_N \approx 0.181$, of ψ_N , by calculating the standard deviation

of the bootstrap sample. This yields a 95% confidence interval for λ of [0.0129, 0.0150], or [12.9, 15.0] per 1,000 person-years.

9.4 Comparison of estimates

Incidence estimates of overall dementia and Alzheimer's disease have varied in the literature depending on the population sampled and the study design (Kukull et al., 2002). Fitzpatrick et al. (2004) reported an estimated incidence rate of all forms of dementia of over 25 per 1,000 person-years for White participants, and over 30 per 1,000 person-years for African-American participants. Kukull et al. (2002) reported a substantially lower rate of overall dementia (20.30 per 1,000 person-years). The estimates of the incidence rate of Alzheimer's disease from these two U.S. studies were, however, similar. Di Carlo et al. (2002) provided results from the Italian Longitudinal Study on Aging (ILSA), and the estimated incidence rates of both overall dementia and Alzheimer's disease were much lower than in the aforementioned U.S. studies. Furthermore, Di Carlo et al. (2002) state that the estimated Alzheimer's disease incidence rate obtained from the ILSA is comparable to those reported in other studies from the U.S., Canada, and Europe (Rocca et al., 1998, Bickel and Cooper, 1994, Anderson et al., 1999, Hagnell et al., 1992, Brayne et al., 1995, and CSHA working group, 2000), which indicates that the estimates from these studies are also in contrast to those from Fitzpatrick et al. (2004) and Kukull et al. (2002).

The reported estimates of the incidence rate of Alzheimer's and of vascular dementia from the CSHA were similar to those from the ILSA (Hébert et al., 2000 and CSHA working group, 2000). The Alzheimer's

disease incidence rates reported from the CSHA, however, were underestimates of the true incidence rates amongst elderly Canadians since they were based only on subjects who survived until the end of CSHA-2 in 1996, as researchers were unable to form a differential diagnosis of Alzheimer's disease for the decedents (CSHA working group, 2000). Still more evidence of the disparity of results in the literature is seen from the higher estimates of the incidence rate of overall dementia obtained from the CSHA in comparison to those obtained from the ILSA.

Our definition of dementia in this thesis corresponds to the diagnosis of one of possible Alzheimer's disease, probable Alzheimer's disease, and vascular dementia. Other authors have grouped those with possible and probable Alzheimer's disease together. In order to compare our estimates from Section 9.3, we consider the incidence rate estimate of vascular dementia given by Hébert et al. (2000) and that of Alzheimer's disease given by the CSHA working group (2000), both based on data collected as part of the CSHA.

The incidence rate estimates of Alzheimer's disease were 7.4 and 5.9 per 1,000 person-years for women and men respectively (CSHA working group, 2000). Hébert et al. (2000) reported an estimated incidence rate of vascular dementia of 3.79 per 1,000 person-years. Since the incidence rate estimates of Alzheimer's disease reported by the CSHA working group (2000) represent an underestimate of the true incidence rates for men and women, we feel that our estimate of $\hat{\lambda} \approx 13.9$ per 1,000 person-years is in agreement with the previous estimates obtained from the CSHA.

An advantage of our analysis is that it does not start with an assumption about the constancy of the incidence rate. We began by assessing the

reliability of stationarity and, having determined that stationarity was a reasonable assumption, we estimated the constant incidence rate, using only the prevalent cohort survival times for both purposes. Hébert et al. (2000) and the CSHA working group (2000) could have performed a check of the validity of stationarity since they had access to incident cases observed between 1991 and 1996. The key point, however, is that, using our method, there is no need to carry out an incident cohort study in order to obtain estimates of the incidence rate. This represents an important economization of time and resources. Furthermore, our procedure yielded a much tighter confidence interval for λ than those reported elsewhere. Whether this will occur more generally remains unclear. It is straightforward to apply the methods developed in Chapter 8, and illustrated in Chapter 9, to arrive at estimates of age specific incidence rates. This is achieved simply by carrying out a stratified analysis by age group. Similarly, one could separately estimate the incidence rate for each level of any categorical variable.

Chapter 10

Concluding remarks and future directions for research

In this thesis, we proposed a formal test for stationarity of the incidence process, which, if we are prepared to assume stationarity, could also be used to test for independence between the underlying survival distribution and calendar time of onset. Our test for stationarity makes use of the equivalence of stationarity with equality of the backward and forward recurrence time distributions (Asgharian et al., 2004). This characterization of stationarity was proved in the setting of a prevalent cohort study with follow-up, which differs from that of renewal theory. We stress that in a prevalent cohort study with follow-up, no single renewal process exists.

In Chapter 5, we pointed out that, although it is a consistent estimator, $\hat{\sigma}_0$, given in Wei's (1980) paper, seems to overestimate σ under H_0 even for relatively large sample sizes. Asymptotically, the theory dictates that, under H_0 , $\frac{\sqrt{n}W_n}{\hat{\sigma}_0} \approx \text{normal}(0,1)$. As a result of the poor estimation of σ under H_0 , however, the density of the test statistic under H_0 has "thinner"

tails than those of a normal(0,1) density. In turn, the power of the test is adversely affected. Cheng (1984) proposed an alternative estimator of σ under H_0 which is only valid under the equal censorship assumption described in Chapter 1. This assumption is common in the censored matched pairs literature, but is not tenable when the pairs arise as backward and forward recurrence times from a prevalent cohort study with follow-up. Finding an improved estimator of σ under H_0 would be valuable. A thorough evaluation of Jung's (1999) procedure is necessary to determine whether it is superior to Wei's estimator, since Jung's limited simulations are difficult to compare with those in this thesis. Another potential solution is to follow the method used to find an estimate of the asymptotic standard deviation, ψ_N , of $\hat{\mu}$ in Chapter 8, that is, to perform a nonparametric bootstrap. A simulation study comparing the bootstrap estimator of σ under H_0 to $\hat{\sigma}_0$ would be very useful, as would a comparison of the distribution of the test statistic under H_0 , using the bootstrap estimator, against a normal(0,1) distribution. As we discussed in Chapter 8, consistency of the bootstrap estimator of ψ_N , and of the bootstrap estimator of σ under H_0 , need to be formally established.

In Chapter 6, we established Theorem 11, which states that the underlying survival distributions are stochastically ordered as a function of calendar time of onset if and only if the backward and forward recurrence time distributions are stochastically ordered. Theorem 11 was only proved for the case where at most two distinct survival distributions exist. An extension to the case which allows for more than two distinct survival distributions needs to be established.

It is interesting to note that the question of whether or not the survival distribution is independent of calendar time of onset can be viewed as a

special type of change point problem. We are trying to assess whether, at some point in time, the survival times begin arising from a different distribution than those with earlier onset times. In our case, we restrict the possible distributions by restricting ourselves only to successive distributions that are stochastically ordered. In a classical change point problem, we ask whether the distribution of a given sequence of observations has changed at some index, often time. It is possible, in our situation, to impose a parametric form on the underlying survival distribution, say, Weibull with a fixed shape parameter and varying scale parameter, and reduce the problem to one resembling a classical change point problem. The simulations in Chapter 7 adopt this approach, but, clearly, more general situations are possible.

Another issue worth investigating is the effect of misspecifying stationarity, that is, assuming stationarity and proceeding with an unconditional analysis when in fact stationarity does not hold. Some questions that arise are: 1) What does the unconditional NPMLE of $S(x)$ converge to, if anything, when stationarity does not hold? 2) If the unconditional NPMLE converges to some function, $S^*(x)$, is there a relationship between $S^*(x)$ and $S(x)$? 3) By misspecifying stationarity, do we systematically overestimate (underestimate) survival?

We speculate that, when there is no censoring, the unconditional NPMLE converges to an $S^*(x)$ where $S^*(x) \geq S(x)$ for all x if the true incidence intensity, $\lambda(t)$, is decreasing over time, and $S^*(x) \leq S(x)$ for all x if $\lambda(t)$ is increasing over time. Thus, our conjecture is that we overestimate survival if we assume stationarity when, in reality, the incidence rate is decreasing, and we underestimate survival if we assume stationarity when the incidence

rate is increasing. The situation is less clear in the presence of right censoring. If stationarity is deemed to be untenable, we recommend employing a conditional approach if a nonparametric analysis is desired.

Zhang (2001) developed nonparametric estimators of $S(x)$ using only the backward recurrence times or only the forward recurrence times. By using only the backward recurrence times one can avoid the cost and effort of following subjects. If we are worried about the reliability of onset times ascertained from prevalent cases, we can use the forward recurrence times alone. Consequently, analogous methods to those described in Chapter 8 can be used for estimating λ when only the backward or forward recurrence times are available.

It is well known that, assuming stationarity, $f_{fwd}(x) \equiv f_{bwd}(x) \equiv \frac{S(x)}{\mu}$. Thus, assuming stationarity, we can define an estimator of the backward (forward) recurrence time density as $\hat{f}_{bwd}(x) = \frac{\hat{S}(x)}{\hat{\mu}}$, where $\hat{S}(x)$ is a nonparametric estimator obtained from only the backward (forward) recurrence times and $\hat{\mu}$ is obtained from $\hat{S}(x)$. From $\hat{f}_{bwd}(x)$ we can obtain an estimator, $\hat{F}_{bwd}(x)$, for $F_{bwd}(x)$ (analogously for $F_{fwd}(x)$). If stationarity holds, we expect that $\hat{F}_{bwd}(x)$ will lie very close to the Kaplan-Meier estimator of $F_{bwd}(x)$. Thus, another informal check of stationarity can be performed by comparing $\hat{F}_{bwd}(x)$ to the Kaplan-Meier estimator of $F_{bwd}(x)$. An advantage of this method is that it provides a check of stationarity, albeit informal, which can be performed using only the backward (or forward) recurrence times. In particular, an informal check for stationarity exists which does not even require follow-up of the subjects in a prevalent cohort study.

If we are prepared to impose a parametric form on $S(x)$, we can even formalize this method since, in this case, the data can be used to estimate the parameters of $S(x)$, and hence to estimate μ , and a goodness-of-fit test can be performed. The “classical” Kolmogorov goodness-of-fit test can be used, although it is conservative when the data are used to estimate unknown parameters of the hypothesized distribution function. We can, however, perform a chi-squared goodness-of-fit test, which allows parameters to be estimated from the data. Moreover, there are “modified” Kolmogorov goodness-of-fit tests, designed for specific parametric families of distributions which allow parameters to be estimated from the data, and that have power that is favorable to the “classical” Kolmogorov goodness-of-fit test.

The external examiner brought to the author’s attention a paper by Murray (2001) which presents nonparametric tests using weighted integrated survival differences in the context of paired censored survival data. These tests may be used, as an alternative to the rank based test proposed in this thesis, to test for stationarity.

Bibliography

- [1] Albers, W. (1988) Combined rank tests for randomly censored paired data. *Journal of the American Statistical Association* **83**, 1159-1162.
- [2] Anderson, P.K., Borgan, Ø., Gill, R.D., and Keiding, N. (1993) Statistical models based on counting processes. Springer Series in Statistics, Springer-Verlag.
- [3] Anderson, K., Launer, L.J., Dewey, M.E., Letenneur, L., Ott, A., Copeland, J.R.M., Dartigues, J-F., Kragh-Sorensen, P., Baldereschi, M., Brayne, C., Lobo, A., Martinez-Lage, J.M., Stijnen, T., Hofman, A., and the EURODEM Incidence Research Group. (1999) Gender differences in the incidence of AD and vascular dementia. *Neurology* **53**, 1992-1997.
- [4] Asgharian, M., M'Lan, C.E., and Wolfson, D. (2002) Length-biased sampling with right censoring: an unconditional approach. *Journal of the American Statistical Association* **97**, No.457, 201-209.
- [5] Asgharian, M., Wolfson, D.B., and Zhang, X. (2004) Checking for stationarity of incidence using prevalent cases. Department of Mathematics and Statistics, McGill University, Technical report 2004-01.

- [6] Asgharian, M., and Wolfson, D.B. (To appear, 2005) Asymptotic behaviour of the unconditional NPMLE of the length-biased survivor function from right-censored prevalent cohort data. *The Annals of Statistics*.
- [7] Bickel, H., and Cooper, B. (1994) Incidence and relative risk of dementia in an urban elderly population: Findings of a prospective field study. *Psychological Medicine* **24**, 179-192.
- [8] Blumenthal, S. (1967) Proportional sampling in life length studies. *Technometrics* **9**, 205-218.
- [9] Brayne, C., Gill, C., Huppert, F.A., Barkley, C., Gehlhaar, E., Girling, D.M., O'Connor, D.W., and Paykel, E.S. (1995) Incidence of clinically diagnosed subtypes of dementia in an elderly population: Cambridge Project for Later Life. *The British Journal of Psychiatry* **167**, 255-262.
- [10] Cheng, K.F. (1984) Asymptotically nonparametric tests with censored paired data. *Communications in Statistics - Theory and Methods* **13**, 1453-1470.
- [11] Commenges, D., Joly, P., Leternneur, L., and Dartigues, J.F. (2004) Incidence and mortality of Alzheimer's disease or dementia using an illness-death model. *Statistics in Medicine* **23**, No.2, 199-210.
- [12] Cox, D.R. (1969) Some sampling problems in technology. In *New Developments in Survey Sampling*. Edited by Johnson and Smith. Wiley, 506-527.

- [13] Cox, D.R., and Oakes, D. (1984) Analysis of survival data. Chapman and Hall, London.
- [14] CSHA working group. (1994) Canadian study of health and aging: study methods and prevalence of dementia. *Journal of the Canadian Medical Association* **150**, 899-913.
- [15] CSHA working group. (2000) The incidence of dementia in Canada. *Neurology* **55**, 66-73.
- [16] Dabrowska, D.M. (1989) Rank tests for matched pairs experiments with censored data. *Journal of Multivariate Analysis* **28**, 88-114.
- [17] Dabrowska, D.M. (1990) Signed-rank tests for censored matched pairs. *Journal of the American Statistical Association* **85** No.410, 478-485.
- [18] Diamond, I.D., and McDonald, J.W. (1991) Analysis of current-status data. In *Demographic Applications of Event History Analysis* Edited by Trussell, J., Hankinson, R., and Tilton, J. chapter 12. Oxford: Oxford University Press.
- [19] Di Carlo, A., Baldereschi, M., Amaducci, L., Lepore, V., Bracco, L., Maggi, S., Bonaiuto, S., Perissinotto, E., Scarlatto, G., Farchi, G., and Inzitari, D. (2002) Incidence of dementia, Alzheimer's disease, and vascular dementia in Italy. The ILSA Study. *Journal of the American Geriatrics Society* **50**, 41-48.

- [20] Fitzpatrick, A., Kuller, L.H., Ives, D.G., Lopez, O.L., Jagust, W., Breitner, J.C.S., Jones, B., Lyketsos, C., and Dulberg, C. (2004) Incidence and prevalence of dementia in the Cardiovascular Health Study. *Journal of the American Geriatrics Society* **52**, 195-204.
- [21] Freireich, E.J., Gehan, E., Frei, E., Schroeder, L.R., Wolman, I.J., Anbari, R., Burgert, E.O., Mills, S.D., Pinkel, D., Selawry, O.S., Moon, J.H., Gendel, B.R., Spurr, C.L., Storrs, R., Haurani, F., Hoogstraten, B., and Lee, S. (1963) The effect of 6-mercaptopurine on the duration of steroid-induced remissions in acute leukemia: A model for evaluation of other potentially useful therapy. *Blood* **21**, 699-716.
- [22] Gehan, E.A. (1965) A generalized Wilcoxon test for comparing arbitrarily singly-censored samples. *Biometrika* **52**, No.1 and 2, 203-223.
- [23] Gilbert, J.P. (1962) Random censorship. Unpublished Ph.D. thesis, University of Chicago, Department of Statistics.
- [24] Gill, R.D., Vardi, Y., and Wellner, J.A. (1988) Large-sample theory of empirical distributions in biased sampling models. *The Annals of Statistics* **16**, 1069-1112.
- [25] Green, A., Hauge, M., Holm, N.V., and Rasch, L.L. (1981) Epidemiological studies of diabetes mellitus in Denmark: II, A prevalence study based on insulin prescriptions. *Diabetologia* **20**, 468-470.
- [26] Hagnell, O., Öjesjö, L., and Rorsman, B. (1992) Incidence of dementia in the Lundby Study. *Neuroepidemiology* **11**, Suppl.1, 61-66.

- [27] Hébert, R., Lindsay, J., Verreault, R., Rockwood, K., Hill, G., and Dubois, M-F. (2000) Vascular dementia: Incidence and risk factors in the Canadian Study of Health and Aging. *Stroke* **31**, 1487-1493.
- [28] Jung, S.H. (1999) Rank tests for matched survival data. *Lifetime Data Analysis* **5**, 67-79.
- [29] Karlin, S., and Taylor, H.M. (1975) A first course in stochastic processes. Academic, New York.
- [30] Keiding, N., Holst, C., and Green, A. (1989) Retrospective estimation of diabetes incidence from information in a current prevalent population and historical mortality. *American Journal of Epidemiology* **130**, 588-600.
- [31] Keiding, N. (1991) Age-specific incidence and prevalence: a statistical perspective. With discussion *Journal of the Royal Statistical Society, Series A* **154**, No.3, 371-412.
- [32] Klein, J.P., and Moeschberger, M.L. (1997) Survival analysis: Techniques for censored and truncated data. Springer-Verlag, New York.
- [33] Kukull, W.A., Higdon, R., Bowen, J.D., McCormick, W.C., Teri, L., Schellenberg, G.D., van Belle, G., Jolley, L., and Larson, E.B. (2002) Dementia and Alzheimer's disease incidence: A prospective cohort study. *Archives of Neurology* **59**, No.11, 1737-1746.
- [34] Lagakos, S.W., Barraj, L.M., and De Gruttola, V. (1988) Nonparametric analysis of truncated survival data, with application to AIDS. *Biometrika* **75**, No.3, 515-523.

- [35] Lancaster, T. (1990) The econometric analysis of transition data. Cambridge University Press, Cambridge.
- [36] Lewis, P.A.W., and Shedler, G.S. (1976) Simulation of nonhomogeneous Poisson processes with log linear rate function. *Biometrika* **63**, No.3, 501-505.
- [37] Mann, H.B., and Whitney, D.R. (1947) On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics* **18**, 50-60.
- [38] Mantel, N. (1966) Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemotherapy Reports* **50**, No.3, 163-170.
- [39] McFadden, J.A. (1962) On the lengths of intervals in a stationary point process. *Journal of the Royal Statistical Society, Series B* **24**, 364-382.
- [40] Murray, S. (2001) Using weighted Kaplan-Meier statistics in nonparametric comparisons of paired censored survival outcomes. *Biometrics* **57**, 361-368.
- [41] O'Brien, P.C., and Fleming, T.R. (1987) A paired Prentice-Wilcoxon test for censored paired data. *Biometrics* **43** No.1, 169-180.
- [42] Pan, W., and Chappell, R. (1998) A nonparametric estimator of survival functions for arbitrarily truncated and censored data. *Lifetime Data Analysis* **4**, 187-202.
- [43] Pettitt, A.N. (1983) Inference for censored matched pairs using signed ranks. *Australian Journal of Statistics* **25**, No.3, 487-495.

- [44] Preston, S.H. (1987) Relations among standard epidemiologic measures in a population. *American Journal of Epidemiology* **126**, 336-345.
- [45] Richter, W. (translated by Brunswick, N.A.) (1965) Limit theorems for sequences of random variables with sequences of random indices. *Theory of probability and its applications* **10**, No.1, 74-84.
- [46] Rocca, W.A., Cha, R.H., Waring, S.C., and Kokmen, E. (1998) Incidence of dementia and Alzheimer's disease: A reanalysis of data from Rochester, Minnesota, 1975-1984. *American Journal of Epidemiology* **148**, No.1, 51-62.
- [47] Ross, S.M. (2003) Introduction to probability models. Academic, San Diego.
- [48] Rothman, K.J. (1986) Modern epidemiology. Little, Brown and Company, Boston.
- [49] Sedmak, D.D., Meineke, T.A., Knechtges, D.S., and Anderson, J. (1989) Prognostic significance of cytokeratin-positive breast cancer metastases. *Modern Pathology* **2**, 516-520.
- [50] Tsai, W-Y., Jewell, N.P., and Wang, M-C. (1987) A note on the product-limit estimator under right censoring and left truncation. *Biometrika* **74**, No.4, 883-886.
- [51] Turnbull, B. (1976) The empirical distribution function with arbitrarily grouped, censored and truncated data. *Journal of the Royal Statistical Society, Series B* **38**, No.3, 290-295.

- [52] Vardi, Y. (1982) Nonparametric estimation in the presence of length bias. *The Annals of Statistics* **10**, 616-620.
- [53] Vardi, Y. (1985) Empirical distributions in selection bias models. *The Annals of Statistics* **13**, 178-205.
- [54] Vardi, Y. (1989) Multiplicative censoring, renewal processes, deconvolution and decreasing density: nonparametric estimation. *Biometrika* **76**, No.4, 751-761.
- [55] Vardi, Y, and Zhang, C.-H. (1992) Large sample study of empirical distributions in a random-multiplicative censoring model. *The Annals of Statistics* **25**, 1022-1039.
- [56] Wang, M-C., Jewell, N.P., and Tsai, W-Y. (1986) Asymptotic properties of the product limit estimate under random truncation. *The Annals of Statistics* **14**, No.4, 1597-1605.
- [57] Wang, M-C. (1987) Product-limit estimates: A generalized maximum likelihood study. *Communication in Statistics, Part A - Theory and Methods* **16**, 3117-3132.
- [58] Wang, M-C. (1991) Nonparametric estimation from cross-sectional survival data. *Journal of the American Statistical Association* **86**, No.413, 130-143.
- [59] Wang, M-C., Brookmeyer, R., and Jewell, N.P. (1993) Statistical models for prevalent cohort data. *Biometrics* **49**, No.1, 1-11.

- [60] Wei, L.J. (1980) A generalized Gehan and Gilbert test for paired observations that are subject to arbitrary right censorship. *Journal of the American Statistical Association* **75**, No.371, 634-637.
- [61] Wilcoxon, F. (1945) Individual comparisons by ranking methods. *Biometrika* **1**, 80-83.
- [62] Wolfson, C., Wolfson, D.B., Asgharian, M., M'Lan, C.E., Østbye, T., Rockwood, K., and Hogan, D.B. (2001) A reevaluation of the duration of survival after the onset of dementia. *New England Journal of Medicine* **344**, No.15, 1111-1116.
- [63] Woolson, R.F., and Lachenbruch, P.A. (1980) Rank tests for censored matched pairs. *Biometrika* **67**, 597-606.
- [64] Zhang, X. (2001) Estimating survival from partially observed data. Unpublished Master's thesis, McGill University, Department of Mathematics and Statistics.