Author: Balan, T.A.
Title: Advances in frailty models
Issue Date: 2018-09-26

Summary

Survival analysis is the study of time to event data, and it is a major topic in statistics. A prominent type of time to event data is represented by life times, which motivates much of the terminology in the field. As a convention, it is common to refer to the event of interest as *death* or *failure*. An individual that is at risk for *dying* is said to be *alive*. Probably the most distinctive feature of survival data is that the event of interest is not always observed. Rather, the only information available is that the individual had not died before a certain time point. This phenomenon is known as *right censoring* and has motivated the development of special statistical methods for this kind of data.

The probability of being alive at a given time point is given by the *survival* function. The most popular way of estimating this in the presence of right censoring is the "product-limit" estimator, better known as the Kaplan-Meier estimator (Kaplan and Meier, 1958). Their seminal paper, *Nonparametric estimation from incomplete observations*, was found to be the most cited paper in statistics in a recent article in *Nature* (Van Noorden, Maher, and Nuzzo, 2014).

The instantaneous probability of dying at a given time point, given that the individual has not died before, is known as the *hazard* function. In demographics, it is also referred to as the "instantaneous mortality rate". In survival analysis, it is more common to work with the hazard rather than the probability density function. The most popular regression model for survival data is the "proportional hazards" model, commonly referred to as the Cox model (Cox, 1972). The paper that introduced this, titled *Regression Models and Life-Tables*, is the second most cited paper in statistics, according to the same Nature article.

In *The impact of heterogeneity in individual frailty on the dynamics of mortality* (Vaupel, Manton, and Stallard, 1979), the authors refer to the effect of unobserved heterogeneity on mortality as *frailty*. The authors state that "mortality rates for individual may increase faster with age than observed mortality rates for cohorts". This implies that there is a distinction between the individual hazard ("mortality rate") and the population hazard ("mortality rate for cohorts"). Most importantly, Vaupel et al. recognize that the individual hazard cannot be directly observed in the presence of unobserved heterogeneity.

The subtle aspect of the hazard is that, by definition, it refers to the individuals still alive at a certain time point. As individuals with a high frailty tend to die faster, it is likely that individuals who survived longer are less frail, on average, as compared to the whole sample at the start of follow-up. Frailty models, which aim to model the unobserved heterogeneity with random effects, are discussed in most survival analysis monographs (Andersen, Borgan, et al., 1993; Kalbfleisch and Prentice, 2002; Klein and Moeschberger, 2005; Aalen, Borgan, and Gjessing, 2008). Several books offer an exhaustive treatment of such models (Hougaard, 2000; Duchateau and Janssen, 2007; Wienke, 2010).

This dissertation describes new statistical methodology that aims to provide more insight into different aspects of frailty models. Both theoretical properties and practical problems are addressed. Of special interest are the "shared frailty" models, that are employed when the frailty is "shared" between several observations. This is usually the case when an individual may experience more events (recurrent events) or when individuals are related (clustered survival data). In Chapter 1 we focus on the frailty effects on observable quantities in Cox models. In Chapter 2, we present a simulation study that focuses on the properties of shared frailty models for clustered survival data, when the size of the clusters is small. In Chapter 3, we discuss a proposed score test for association between a recurrent event process and a terminal event, when the frailty is shared by both processes. In Chapter 4, we discuss selection bias in the context of recurrent events, where the selection depends on the outcome and on the underlying frailty. In Chapter 5, we present the estimation procedure implemented in the **frailtyEM** R package. In what follows, we show a more detailed summary for each chapter.

Chapter 1 is the introduction to this dissertation. It follows the structure of a tutorial, providing an overview of theory and practice surrounding frailty models. In Section 1.2, we address to *univariate frailty* models. These are related to the original formulation of Vaupel, Manton, and Stallard (1979), where the outcome of interest is a singular event for individuals (death), and the individual event times are assumed to be independent of each other. Via simulated examples, we illustrate two phenomena specific to Cox models. First, the *selection* process, that describes the distribution of risk factors in the population of survivors. Second, the observed *marginal* covariate effect in the Cox model, when important explanatory variables are omitted. The same phenomena are then studied in detail with frailty models, for different frailty distributions. The chapter concludes with a discussion of the identifiability properties of frailty models in univariate survival data.

In Section 1.3, we illustrate via a simulated data example how marginal correlation between event times may arise, when covariates "shared" by related individuals are missing. This is further studied with *shared frailty* models, wherein the random effect is assumed to be shared between different individuals. We study how different correlation patterns arise from different frailty distributions and we discuss how shared frailty models may be used for modeling recurrent events. In Section 1.4 we address practical issues surrounding the estimation of frailty models. We discuss different procedures for semi-parametric and parametric models, we review the available software and describe how different data types can be accommodated by software packages. Finally, in Section 1.5 we discuss several proposed extensions of the frailty model.

In **Chapter 2** we analyze situations where it is difficult to tell the difference between non-proportional hazards and unobserved heterogeneity. This chapter builds on the results discussed in Chapter 1, especially those regarding the identifiability of frailty models. A well known result is that the frailty model is identifiable if covariates are present and the frailty distribution has finite moments. We argue that this is problematic, because the frailty may falsely explain a time dependent covariate effect as evidence for unobserved heterogeneity. While generally thought that this is not a problem for shared frailty models, we show that it may be, especially if the cluster size is small.

In Section 2.2, we review the proportional hazards models and the conditional proportional hazards assumption commonly made for frailty models. Next, we discuss how marginal non-proportional hazards may arise from different frailty models. In Section 2.3, we present the simulation study. We study the effect of the cluster size (in fact, how "multivariate" the outcome is) on detecting frailty models, when there is no real unobserved heterogeneity. We analyze the results for different quantities of interest: the likelihood ratio test, the score test for heterogeneity and estimated parameters. Our main conclusion is that time dependent covariate effects may falsely appear as evidence for frailty, when the path of the effect is somewhat similar to the marginal hazard ratio implied by the frailty model. Although this problem is mitigated with larger sample sizes, when the cluster size is small (e.g. 2, 3) the distinction between unobserved heterogeneity and time-dependent covariate effects is subtle. The results are extended to recurrent events, and a combination of time dependent covariate effects in the presence of frailty. Finally, the phenomenons analyzed in this chapter are illustrated with a data analysis of a well known data set on recurrent kidney infections.

In **Chapter 3**, we introduce a score test for association between recurrent events and a terminal event. If frailty is present and high frailty individuals are associated both with a higher rate of recurrent events and a higher mortality, then the two event processes must be jointly analyzed. This is complicated in practice, especially for semiparametric models. We propose a simple score test for association testing the null hypothesis that the two models are independent. If this is not rejected, simpler separate analyses may be carried out.

In Section 3.2, a joint model for recurrent events and a terminal event is introduced, employing a gamma distributed frailty. This model includes an association parameter that may be estimated, for which different inference methods are compared. In Section 3.3, the "robust score test" is introduced, together with other well known statistical tests, for the null hypothesis of no association. In Section 3.4, we show via a simulation study that the proposed test behaves well and, in terms of power, is comparable to more complicated alternatives. In Section 3.5, the proposed methodology is illustrated on a data set comprising recurrent skin tumors.

In **Chapter 4**, the problem of selection bias (or "ascertainment" bias) in recurrent events is analyzed. The motivating example is a data set comprising recurrent pneumothoraces. The data was collected only for individuals that had at least one recorded event during a certain accrual time window. For the selected individuals, the whole event history was collected. The problem is that, by design, individuals with a higher rate of events will be over represented in this sample. If unobserved heterogeneity is present, high frailty patients are over represented. In this chapter, we study the estimation of frailty parameters and covariate effects in this type of scenarios.

In Section 4.2, several selection schemes and a general adjusted likelihood approach are introduced. We discuss the effects of the ascertainment on the estimates from a model without frailty and from a model with frailty. For the latter, a pseudo maximum likelihood estimation algorithm is presented. In Section 4.3, the performance of the adjusted likelihood approach is studied for different selection scenarios, and it is shown to work well in general. Finally, in Section 4.4, the proposed methodology is illustrated on the original motivating data set.

In **Chapter 5**, we study the estimation of semiparametric shared frailty models in practice, with a focus on the **frailtyEM** package (Balan and Putter, 2017) for the R programming language. This software is meant to combine the flexibility of semiparametric models with a large choice of frailty distributions. A major motivation behind writing this package was to provide well documented user level features. In Section 5.1, we present an overview of the currently available software for the estimation of frailty models.

In Section 5.2, the likelihood construction and the effect of left truncation and ascertainment are discussed in the context of frailty models. Next, we make an overview of related results regarding practical problems: hypothesis testing, marginal and conditional quantities and goodness of fit. In Section 5.3, the software implementation of a profile expectation maximization algorithm is discussed. The proposed estimation method and the calculations required to obtain standard errors are presented. From a practical point of view, the functions provided by the package are presented, together with their corresponding syntax. Finally, the features of the package are illustrated with examples involving three well known data sets, covering three important scenarios: recurrent events in calendar time, recurrent events in gap time and clustered failures.