Statistical methods for frailty models: studies on old-age mortality and recurrent events

Böhnstedt, M.

1 Introduction

1.1 Introduction

The heterogeneity between individuals is a key aspect in the statistical studies of various phenomena, including mortality and recurrent infections or fertility events. The standard approach for incorporating the differences in observed individual characteristics into the analysis are regression models. However, unobserved individual characteristics often add to the variability in the data as well. Therefore, the statistical models also need to take this unobserved heterogeneity between individuals into account.

When studying duration data that describe the time until an event of interest occurs, the heterogeneity between individuals will also lead to changes in the composition of the sample over time. These selection effects can have serious implications for the statistical analysis in different contexts. In mortality studies, for instance, the sample of survivors to a certain age may have different characteristics than the initial population. Similar effects arise in longitudinal studies if the latent characteristics affect not only the longitudinal component, but also the survival time of the individuals. Thus, in settings with delayed entry, the analysis is based on the selective sample of individuals who survived to the point of entering the study. Even if there is no delayed entry, an association between the longitudinal component and the survival time will result in a dependent drop-out mechanism.

The modeling of unobserved heterogeneity has greatly advanced in recent decades. A common approach to modeling the latent information is through random effects. In the analysis of time-to-event data, such a random effect is usually referred to as frailty. In

the following section, we give a brief general introduction to frailty modeling of survival data. Then, we present the two specific frailty models that we study in this thesis to investigate recurrent events, such as repeated infections or fertility events, and old-age mortality, respectively. An outline of the thesis is provided in Section 1.5.

1.2 Frailty modeling of survival data

Time-to-event data are commonly modeled through the hazard function, which describes the instantaneous probability of the occurrence of the event of interest for individuals who have not yet experienced the event. As these risks of the occurrence of an event can naturally differ between individuals, the statistical models have to allow for heterogeneity. To account for the differences that are attributable to observed characteristics, covariates can be included in the analysis. However, in most cases, not all characteristics that cause the hazards to vary between individuals can be observed. Frailty models provide an appealing approach to accounting for such unobserved heterogeneity. For example, in the class of frailty proportional hazards models, this type of heterogeneity is incorporated through a positive random effect, the frailty, that acts multiplicatively on a common baseline hazard. For univariate survival data, such frailty models have already been proposed by Beard (1959) and Vaupel et al. (1979).

Unobserved heterogeneity can also arise in studies of multivariate survival data. When analyzing clustered survival data, such as survival times of individuals from different families, the risks of experiencing the event may vary across the clusters. Shared frailty models, in which the value of the random effect is common to all individuals in a cluster, not only take into account the heterogeneity between clusters, but also relax the assumption of independence between the individuals within a cluster. Similarly, shared frailty models are applied in recurrent event studies in which the same type of event can occur repeatedly in the same individual. Individual-specific random effects can then accommodate individual differences in the event rates, as well as the dependence between the recurrence times within an individual. An overview of frailty modeling for multivariate survival data has been provided by, among others, Hougaard (2000).

An important feature of survival data is that they are usually subject to censoring or truncation. For example, if we only know that the individual has not yet experienced the event of interest at a certain point in time, the observation is right-censored. If an individual is included in the study only if he or she has not yet experienced the event of interest by some point in time – that is, if the event time is larger than a certain value – the observation is left-truncated. The statistical methods need to be adapted to such observation schemes. But given that even with these adjustments, the loss of information caused by the incomplete observations may pose considerable challenges for the analysis, this issue should be taken into consideration already when the study is designed.

Another central issue that should be taken into account in time-to-event analysis is that selection effects can occur if there is heterogeneity in the event rates. In studies of a single failure event, the individuals who survive up to a certain point in time will tend to

have lower failure risks than the individuals who already had a failure before that point in time. In the context of frailty models, this implies that the frailty distribution in the sample of survivors changes over time. Ignoring such selection effects may have serious consequences for the inference results.

The issues outlined above also play an important role in the study of the two distinct frailty models that are investigated in this thesis. In the first part, we consider a specific proportional hazards frailty model for univariate survival data that is used as a framework for assessing a demographic phenomenon known as mortality deceleration. In the second part, we examine inference in a joint frailty model for recurrent events and a terminal event under different observation schemes. In the following sections, we briefly introduce the two models, and provide some background on the applications and the statistical approaches used in this thesis.

1.3 Assessing mortality deceleration

The first part of this thesis was motivated by studies on old-age mortality patterns, which have frequently observed that human death rates slow down at advanced ages. This mortality deceleration can be described through a specific proportional hazards frailty model. However, non-standard conditions for the inference and limited data availability can complicate the statistical analysis. Thus, our aim is to investigate both traditional and new approaches to assessing mortality deceleration in this setting.

1.3.1 The gamma-Gompertz model

The observation that human mortality rates increase exponentially over most of the adult lifespan was already made by Gompertz (1825). However, later investigations based on improved vital registration records suggested that death rates slow down at the oldest ages (Thatcher et al., 1998). Such a downward deviation from the exponentially increasing hazard at high ages is called mortality deceleration, and can be explained by selection effects due to heterogeneous mortality risks. As individuals with higher mortality risks tend to die at younger ages, the population of survivors to advanced ages tends to be made up of individuals with lower mortality risks. This, in turn, results in a comparatively lower population hazard at these advanced ages.

Mortality deceleration can therefore be examined in the framework of a frailty proportional hazards model. The individual risks at any age x are specified in terms of the conditional hazards $h(x|Z = z) = z \cdot h_0(x)$ for given frailty value $Z = z$, with $h_0(x)$ denoting the baseline hazard. More specifically, we consider the gamma-Gompertz model in which the individual risks are defined by an exponentially increasing Gompertz baseline hazard, $h_0(x) = ae^{bx}$, that is multiplied by a gamma distributed frailty. The frailty variance σ^2 describes the amount of heterogeneity in mortality risks. In the presence of

heterogeneity, indicated by a positive value of the frailty variance, the marginal population hazard,

$$h(x) = \frac{ae^{bx}}{1 + \sigma^2 \frac{a}{b}(e^{bx} - 1)}, \quad x \geq 0,\ a, b > 0,\ \sigma^2 \geq 0,$$

decelerates at older ages, as illustrated in Figure 1.1. If there is no heterogeneity and the frailty variance is zero, the population hazard is of the same exponentially increasing Gompertz form as the individual hazards. Hence, in the setting of the gamma-Gompertz model, the frailty variance determines whether the population hazard does or does not decelerate at older ages, and the statistical assessment of the phenomenon will therefore largely depend on this variance parameter.

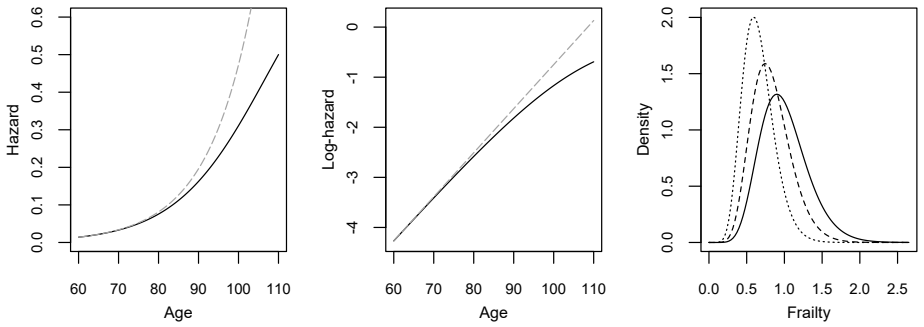

Figure 1.1: Hazard (left) and log-hazard (middle) of the gamma-Gompertz model with parameters $a = 0.014$, $b = 0.088$, and $\sigma^2 = 0.1$ (black-solid) or $\sigma^2 = 0$ (gray-dashed). Right: Frailty distribution at the starting age of 60 (solid) and among survivors to ages 90 and above (dashed), or to ages 100 and above (dotted), which illustrates that the distribution of frailty among survivors to higher ages is concentrated at lower values, and has smaller variance.

Assuming a fully parametric model for human adult mortality may seem restrictive in light of the advances in semiparametric and nonparametric survival analysis. However, the simple relation of a linear increase of the log mortality rate with age, which is imposed by the Gompertz model, has been repeatedly found to hold across a large part of the age range for human adults in various studies covering different countries and time periods. In addition to its importance in actuarial studies and demography, the Gompertz model has been applied in research on the biology of aging. While the search for an explanation for the exponential increase in mortality with age is still ongoing, it has been shown that models of damage accumulation in specific systems can produce such patterns (Gavrilov and Gavrilova, 2001; Kirkwood, 2015).

1.3.2 Old-age mortality studies: challenges and design

Several challenges arise in the study of mortality deceleration. In the setting of the gamma-Gompertz model, the absence of mortality deceleration corresponds to a value of zero for the frailty variance, which lies on the boundary of the parameter space. The presence of such a 'boundary parameter' introduces a non-standard condition, such that the standard asymptotic results for likelihood-based inference, which are derived under regularity conditions, will generally not apply (Self and Liang, 1987). Therefore, the statistical techniques for assessing mortality deceleration need to be adapted to the non-standard setting.

Empirical studies of mortality deceleration often face the challenge that the available data are limited. Even if the studies are based on large populations, relatively few deaths will be observed at the advanced ages at which the deceleration is most pronounced. This is an inherent feature when studying phenomena that occur at the tail of a distribution, where the data naturally become sparse. Moreover, scientific validation of the ages at death is often considered necessary in order to avoid biased mortality estimates due to age misreporting. Because of the time and the costs involved in verifying individuals' vital records, mortality data are usually validated for only parts of the population of interest. When studying old-age mortality, samples are often restricted to include only those individuals who survived beyond some advanced age, such as survivors beyond age 80. In the setting of the gamma-Gompertz model, the sample of survivors to ages 80 and above will be less heterogeneous in their mortality risks than the sample of survivors to, say, ages 70 and above, due to selection effects (see the right panel of Figure 1.1 for an illustration). Hence, the left truncation of the observations will affect the inference in this model, and, thus, the ability to detect a slowing down of the death rates.

For the purposes of quantifying the effects of the age of left truncation and the sample size on the assessment of mortality deceleration, information measures can be derived from the Fisher information matrix of the gamma-Gompertz model. This approach draws on the concept of the Fisher information as a means of measuring the amount of information the data carry about the model parameters (Lehmann, 1999). Applying ideas from the theory of optimal design (Atkinson, 1988), study designs can be evaluated based on different optimality criteria, which are defined in terms of scalar functions of the information matrix. As a result, we will be able to gauge, for instance, the amount of information that can be gained by adding the observations on the deaths at ages 85-89 to an existing data set of survivors to ages 90 and above.

1.3.3 Statistical techniques for assessing mortality deceleration

The non-standard condition of the boundary parameter and the aspects of the study design laid out above will have an impact on the statistical methods used in the study of mortality deceleration.

In the framework of the gamma-Gompertz model, one approach for assessing the phenomenon is a likelihood ratio test for a zero frailty variance. Taking into account the non-

standard condition that the frailty variance lies on the boundary of the parameter space under the null hypothesis is essential for deriving asymptotic properties of the test. Under the null hypothesis, the test statistic is not asymptotically chi-squared distributed with one degree of freedom, but instead asymptotically follows a 50:50-mixture distribution of a point mass at zero and a chi-squared distribution with one degree of freedom. We also determine a formula for the local asymptotic power of the likelihood ratio test, which involves one of the information measures derived from the Fisher information. Calculations for specific scenarios will illustrate that the test's power to detect mortality deceleration is low if the data are left-truncated at relatively high ages.

A drawback of the hypothesis testing approach is that it is designed to control for the probability of committing a type I error, which in this case is the probability of falsely rejecting the hypothesis of no mortality deceleration. However, the probability of a type II error – that is, of failing to reject the hypothesis of no mortality deceleration although mortality deceleration is present – is not directly controlled, and may be relatively large. This asymmetry of hypothesis testing suggests that alternative methods should be considered here.

A different approach to assessing mortality deceleration is model selection based on information criteria (Burnham and Anderson, 2002). The gamma-Gompertz model is simplified to the Gompertz model if the frailty variance takes the value of zero and there is no mortality deceleration. Thus, we have to choose between two nested models that differ by one parameter only. However, as the selection concerns a boundary parameter, the standard formulas for the information criteria may not be applicable. In Chapter 2, we show that the standard version of the popular Akaike information criterion (AIC; Akaike, 1974) is biased in the setting of the gamma-Gompertz model due to this non-standard condition. Moreover, even after adding a correction term to reduce the bias, model selection based on the AIC will not enable us to reliably detect mortality deceleration for samples that include only survivors to advanced ages.

The AIC was constructed to evaluate overall model performance, whereas in certain applications, model performance for a specific parameter of interest might be more relevant. Claeskens and Hjort (2003) introduced a focused information criterion (FIC) for selecting a model that performs best for a given focus parameter. From a set of nested candidate models, the model with the smallest limiting risk of the estimator of the focus parameter is chosen. This approach is appealing when studying mortality deceleration, because attention can be directed to the quantities that determine old-age mortality, such as the frailty variance or the log-hazard function at some advanced age. As the original FIC was developed under standard regularity conditions, the non-standard condition of the boundary parameter in the present application requires us to derive a new version of the FIC. The performance of this new criterion for detecting mortality deceleration will be assessed in comparison with the performance of the modified AIC.

1.4 Joint modeling of recurrent events and a terminal event

The second part of this thesis is devoted to the study of recurrent events in the presence of a terminal event. In particular, methods for inference in a joint frailty model are developed under two specific observation schemes.

1.4.1 The joint frailty model

Recurrent events are found in various areas of application, including medical studies on tumor occurrences, demographic studies on fertility, and actuarial studies on insurance claims. Hence, recurrent events have received considerable attention in the recent years; an overview is provided by Cook and Lawless (2007).

In many cases, a terminal event such as death might stop the repeated occurrences of the event of interest in an individual. This potentially dependent censoring of the recurrence process requires that the two event processes are modeled jointly. In an extension of the shared frailty model, Liu et al. (2004) proposed a joint frailty model for recurrent events and a terminal event. As in the shared frailty model for recurrent events, the individual-specific frailties induce heterogeneity in the recurrence rates, and account for the dependence between the recurrences within one individual. But in the joint model, the frailties also affect the hazard of the terminal event, thereby introducing dependence between the recurrence process and the terminal event. The direction and the strength of the association can be modified through an additional dependence parameter.

Consequently, the model not only provides a suitable framework for handling the dependent censoring of the recurrent event process; but it also allows us to assess how the recurrent events and the terminal event are related. The question of whether individuals with a higher recurrence rate simultaneously have a higher or even a lower risk of experiencing the terminal event is of relevance in several applications. This issue is illustrated with a study on the fertility and mortality of a marine organism in Chapter 5. Moreover, analyses based on the joint frailty model can provide insights into how the risks of experiencing recurrent events or the terminal event evolve with time.

1.4.2 Different observational settings

The joint frailty model is usually studied in a setting in which each individual is observed from the start of the event processes and the exact times of event recurrence are available. In this thesis, we consider two variations of this observation scheme. First, we deal with a situation in which the recurrence process is observed intermittently, so that only the numbers of recurrent events that occurred between successive observation times are known. This scheme of interval counts of recurrent events can arise in different contexts, such as in a medical study in which the number of epileptic seizures of a patient is recorded at scheduled visits to the doctor, or in a laboratory experiment in which rats are examined for newly developed tumors at fixed inspection times.

Second, we consider a situation in which the recurrence process is continuously observed, but the observation does not begin until sometime after the start of the event processes. Such cases of delayed entry can, for instance, occur in studies of certain diseases in which time is measured from the diagnosis onwards, but patients are enrolled at various points in time after their diagnosis. Another example is that of studies in which age is the main time scale, but individuals enter the study at different ages. In these settings, individuals are included in the study only if they are still at risk of experiencing the terminal event; that is, if the event has not yet occurred for them. The resulting left-truncated study sample may not be representative of the underlying population, as it could consist of individuals who tend to have a lower hazard of experiencing the terminal event. These selection effects can also cause the frailty distribution in the study sample to differ from the initial distribution of frailties in the target population. Hence, inference methods that take the left truncation properly into account must be developed.

1.4.3 Inference in the joint frailty model

Several estimation procedures for the joint frailty model have been proposed for settings in which the recurrence process is observed from the start, and the exact recurrence times are known. The methods are frequently based on the marginal likelihood, which can be approximated using Gaussian quadrature. As the likelihood can be easily adjusted to different observation schemes, we can also adopt this approach for estimating the joint frailty model in the above situations with intermittent observations or delayed entry. More precisely, we build on the work of Liu and Huang (2008), who used Gauss-Hermite quadrature to approximate the likelihood and specified the event rates as piecewise-constant functions.

Given the complexity of the joint frailty model, it may be of interest to test a priori whether the recurrence process and the terminal event are associated. For that purpose, Balan et al. (2016) developed a score test for the association between the two event processes in the setting in which recurrence times are exactly observed. We adapt this test to the situation in which there are interval counts of the recurrent events.

1.5 Outline of the thesis

The two parts of this thesis are structured as follows. The first part encompasses three chapters that cover the different aspects of assessing mortality deceleration in the framework of the gamma-Gompertz model.

In Chapter 2, we present asymptotic results for likelihood-based inference in the gamma-Gompertz model in a local misspecification setting. We determine the limiting distribution of the maximum likelihood estimator and an approximation of the local asymptotic power of the likelihood ratio test for a zero frailty variance under the non-standard condition of the boundary parameter. In addition, we derive the bias of the standard AIC in this model with a boundary parameter.

To address questions relevant to empirical studies of mortality deceleration, we investigate in Chapter 3 how the sample size and the age range covered by a data set affect the assessment of mortality deceleration. The corresponding information measures based on the Fisher information matrix of the gamma-Gompertz model are described. One of the criteria is used in calculating the approximate local power of the likelihood ratio test according to the formula derived in Chapter 2. We demonstrate that samples that are relatively small in size or that cover only the most advanced ages can make drawing reliable inferences difficult.

Chapter 4 discusses the approach of using model selection to assess mortality deceleration. The new version of the FIC is introduced and its performance is evaluated, especially for more demanding study designs. Comparisons between this criterion and a modified AIC with a bias correction term, which is motivated by the result from Chapter 2, are also included.

The second part of the thesis is comprised of two chapters about the inference in the joint frailty model for recurrent events and a terminal event under two different observational schemes. Chapter 5 focuses on the situation in which only interval counts of recurrent events are available. We present the method for estimating the joint frailty model by direct maximization of the approximate marginal likelihood. The question of whether the recurrence process and the terminal event process are associated is addressed by means of a score test.

In Chapter 6, the joint frailty model is studied in the situation with delayed entry. We outline the construction of the likelihood, and propose an estimation procedure based on an approximated marginal likelihood. We demonstrate the importance of taking into account the selection effects on the frailty distribution in the left-truncated sample.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control 19*(6), 716–723.

Atkinson, A. C. (1988). Recent developments in the methods of optimum and related experimental designs. *International Statistical Review 56*(2), 99–115.

Balan, T. A., S. E. Boonk, M. H. Vermeer, and H. Putter (2016). Score test for association between recurrent events and a terminal event. *Statistics in Medicine 35*, 3037–3048.

Beard, R. E. (1959). Note on some mathematical mortality models. In G. E. W. Wolstenholme and M. O'Connor (Eds.), *The Lifespan of Animals*, Ciba Foundation Colloquia on Ageing, pp. 302–311. Boston: Little, Brown.

Burnham, K. P. and D. R. Anderson (2002). *Model Selection and Multimodel Inference. A Practical Information-Theoretic Approach*. Springer.

Claeskens, G. and N. L. Hjort (2003). The focused information criterion. *Journal of the American Statistical Association 98*, 900–916.

Cook, R. J. and J. F. Lawless (2007). *The Statistical Analysis of Recurrent Events*. Statistics for Biology and Health. New York: Springer.

Gavrilov, L. A. and N. S. Gavrilova (2001). The reliability theory of aging and longevity. *Journal of Theoretical Biology 213*, 527 – 545.

Gompertz, B. (1825). On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philosophical Transactions of the Royal Society of London 115*, 513–583.

Hougaard, P. (2000). *Analysis of Multivariate Survival Data*. Statistics for Biology and Health. New York: Springer.

Kirkwood, T. B. L. (2015). Deciphering death: a commentary on Gompertz (1825) 'On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies'. *Philosophical Transactions of the Royal Society of London B 370*. doi: 10.1098/rstb.2014.0379.

Lehmann, E. L. (1999). *Elements of Large-Sample Theory*. Springer texts in statistics. New York: Springer.

Liu, L. and X. Huang (2008). The use of Gaussian quadrature for estimation in frailty proportional hazards models. *Statistics in Medicine 27*, 2665–2683.

Liu, L., R. A. Wolfe, and X. Huang (2004). Shared frailty models for recurrent events and a terminal event. *Biometrics 60*, 747–756.

Self, S. G. and K.-Y. Liang (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association 82*(398), 605–610.

Thatcher, A. R., V. Kannisto, and J. W. Vaupel (1998). *The force of mortality at ages 80 to 120*. Odense Monographs on Population Aging 5. Odense University Press.

Vaupel, J. W., K. G. Manton, and E. Stallard (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography 16*(3), 439–454.