An examination of the trend-renewal process for use in recurrent events modelling in sports and medicine

by Meng Zhao

Master of Science

Department of Mathematics and Statistics

McGill University

Montréal, Québec

2016-06-25

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Masters of Science from the Department of Mathematics and Statistics at McGill University

ACKNOWLEDGEMENTS

I thank my supervisors Dr. Russell Steele and Dr. Ian Shrier for their incredible mentorship and support. The work in this thesis was supported by a discovery grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) and a Collaborative Health Research Project grant jointly funded by NSERC and the Canadian Institutes for Health Research (CIHR). I am also grateful to Dr. Diana Pietzner for helpful communication regarding the computational details of her work.

ABSTRACT

The trend-renewal model for recurrent time-to-event data is seldom used outside of the reliability literature. This thesis thoroughly discusses the foundations of the trend-renewal process, emphasizing its applicability in the fields of sports injury and medicine. It proposes ways to better utilize a popular choice of parametric framework to address research questions in practical settings, in particular, an alternative to the classical Cox proportional intensities formulation of covariate effects. Simulation studies are carried out to evaluate the finite sample inference of parametric trend renewal models with unobserved heterogeneity. Finally, an application to a medical dataset is provided.

ABRÉGÉ

Le modèle renouvellement-de-tendance pour les données récurrentes de temps de survie est rarement utilisé en dehors de la littérature de fiabilité. Cette thèse examine en profondeur les fondements du processus renouvellement-de-tendance, mettant l'accent sur son application dans les domaines des blessures sportives et de la médecine. Elle propose des moyens pour mieux utiliser un cadre paramétrique populaire pour répondre aux questions de recherche en milieu pratique, en particulier, une alternative à la méthode classique Cox d'intensités proportionnelles. Des simulations sont effectuées pour évaluer l'inférence en échantillonnage fini de modèles renouvellements-de-tendance paramétriques avec hétérogénéité non observée. Enfin, une application à un ensemble de données médicales est fournie.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
Introduction

The need for effective statistical analysis of recurrent time-to-event data has been present and is ever-growing in many fields such as the social sciences, insurance, biology, medicine, and machine reliability. Models should be carefully constructed and methods should be carefully chosen to provide valid answers to specific problems in these different fields. All models have their limitations and drawbacks, but some become more popular than others. For example, certain methods are heavily used because the literature on them is extensive, they are easy to interpret, or they are reputable in that field's research community. It might be due to a combination of these reasons that the modelling of recurrent sports injury has seen relatively slow development, such that many analysts are still defaulting to only examining the time to first event, or in the most elaborate cases, applying some variation of a modulated renewal model which incorporates covariates in a Cox proportional hazards -like fashion. These analyses can be valid for answering particular questions, but when they are invalid, it is often the case that there are at most a handful of useful but relatively unknown alternatives. One of the many reasons for this is that the literature does not contain explanation and motivation for these alternatives.

This thesis focuses a model from the reliability literature and repurposes it with a view towards the fields of sports injury and medical analyses, whose research objectives are often quite different from those of reliability.

Until recently, the trend-renewal process (TRP) has rarely been used outside of reliability analysis, i.e. only in contexts where the subjects being observed are machines or strictly mechanical processes. For example, Yang et al. (2012) analyses the failure profiles of high throughput screening (HTS) automation systems and cylinder head production processes under the TRP framework. Of its application to medical recurrent events, we found only one example: Pietzner and Wienke (2013) demonstrate the use of the TRP for a hospital readmission times dataset collected from colon cancer patients originally analyzed by Gonzalez et al. (2005). Currently in the literature there is no application of the TRP to recurrent sports injury. This thesis does not fill this gap, however, it does provide some methodology for the analysis of recurrent sports injury under the trend-renewal process framework and make suggestions about how a mathematically convenient choice of parametric TRP can be interpreted in practical settings. In a future application the TRP can potentially be used to answer specific questions about athletes' injury patterns and to help measure the efficacy of rehabilitation programs.

The fact that TRPs have not been used much in fields outside of reliability could largely be due to researchers' emphasis on their mathematical convenience rather than practical interpretability. While the statistical literature on trend-renewal models has focused on parametric estimation (Lindqvist et al. (2003); Lindqvist (2006)), semiparametric (Heggland and Lindqvist (2007); Jokiel-Rokita and Magiera (2012)), and nonparametric (Gámiz and Lindqvist (2016)), or on prediction (Franz et al. (2014)), guidelines for interpretation in a practical context are either absent or inadequate.

This thesis will highlight the convenience of the TRP and its potential to describe a variety of modulation in the failure times of repairable systems. It provides an in-depth discussion of the parametric TRP in the way of modelling and inference and offers some new considerations for how to better use the popular Weibull-power-law choice of TRP to answer questions common in medical and sports injury analyses. Chapter 2 provides a summary of the common frameworks used to model recurrent events. Chapter 3 presents useful interpretations of the Weibull-power-law TRP and then Chapter 4 contains several finite sample evaluations of inference in two common parametric TRPs. Chapter 5 provides a demonstration on the hospital readmission dataset of Gonzalez et al. (2005) and finally Chapter 6 concludes with a discussion and future work.

CHAPTER 2
Background

2.1 Overview of single-event survival analysis

The general problem in single event (or time-to-first-event) survival analysis can be described in the following way. Let Y be a random variable representing the time to first occurrence of an event of interest. For instance, the event of interest could be the failure of a machine, a patient exiting a state of remission of a disease, or an athlete getting injured during a season. To avoid unnecessary complication, we will assume the time scale – of machine operation, of remission, or of exposures on the playing field – has a well-defined initiation point.

We generally estimate the distribution of Y by sampling observations from it. Because of finite resources, the observed data is $T = \min(Y, C)$ where C is a random variable representing the *censoring time* of Y. An observation time is said to be censored if we do not know when exactly the event of interest occurred. There are different kinds of censoring but the most common form of censoring is right-censoring, i.e. when $T = \min(Y, C)$ and the only information known is that a subject had not yet experienced the event of interest by time T. Standard causes of right-censoring are from the subject staying event-free until the end of the study, in which case C is the period of study, or from the subject being lost to follow-up before the end of the study, in which case C may or may not be independent of Y.

The basic observations for m subjects are $\{(T_i, \delta_i) : \; i = 1, ..., m\}$ where δ_i is an indicator of whether subject i was censored. Given these *survival* (equivalently *failure* or *waiting*) times and censoring statuses of m subjects, it is of interest to estimate the population distribution of failure times. In addition, the effects of certain covariates $\mathbf{Z}$ are often of interest, so that the completely observed data can be written as $\{(T_i, \mathbf{Z}_i, \delta_i) : i = 1, ..., m\}$.

Let $\mathcal{F}_{t^-}$ be the complete history of all available information for subjects up to time t^- (until just before t). Then the hazard function is defined as:

$$h(t|\mathcal{F}_{t^-}) = \lim_{\Delta t \to 0} \frac{\Pr\left(t \leq Y < t + \Delta t \,|\, Y \geq t, \mathcal{F}_{t^-}\right)}{\Delta t}. \tag{2.1}$$

The hazard function is frequently used to characterize survival distributions because it can be thought of as an instantaneous rate at time t, given survival until time t, and it determines all other aspects of the survival process.

Where censored data is concerned, the common condition required for standard approaches to maximum likelihood statistical inference to be valid is that of *independent* censoring (Prentice and Kalbfleisch (1980, Section 5.2)), which is defined by (2.1) being equivalent to:

$$\lim_{\Delta t \to 0} \frac{\Pr\left(t \leq Y < t + \Delta t \,|Y \geq t, C \geq t, \mathcal{F}_{t^-}\right)}{\Delta t}.$$

In other words, the censoring mechanism must be part of the observed event history $\mathcal{F}_{t^-}$. This is because, if we can know that $\{C < t\} \in \mathcal{F}_{t^-}$, then we would also have $\{C \geq t\} = \overline{\{C < t\}} \in \mathcal{F}_{t^-}$. Independent censoring is achieved when there is no unobserved heterogeneity amongst subjects – in other words, conditional on some covariates $\mathbf{Z}$, the censoring time $C|\mathbf{Z}$ and event time $Y|\mathbf{Z}$ are independent, *and* $\mathbf{Z}$ is observed in the complete history. If we somehow could know this, then knowing a subject's censoring time would be superfluous to knowing $\mathcal{F}_{t^-}$! However, independent censoring is an unverifiable condition in practical settings (Tsiatis (1975)).

The consequences of incorrectly assuming independent censoring have been well examined and will not be explored in this thesis. Here we simply emphasize that most of statistical inference for single event and recurrent event models hinges on the validity of the independent censoring assumption.

Now, recall that if a random variable Y has a continuous distribution then its hazard function $h_Y(\cdot)$, density function $f_Y(\cdot)$, and survivor function $S_Y(\cdot)$ are related:

$$h_Y(t) = \frac{f_Y(t)}{\Pr(Y \geq t)} = \frac{f_Y(t)}{S_Y(t)}.$$

Hence, assuming the distributions of Y and C are continuous and respectively parametrized by θ_Y (the parameters of interest) and θ_C (the nuisance parameters), the full likelihood of θ_Y in hazard-survivor form for data $\{(T_i, \mathbf{Z}_i, \delta_i) : i = 1, ..., m\}$ would be, under independent and non-informative censoring:

$$L_i(\theta_Y) = \prod_{i=1}^{m} \left[h_Y(t_i)S_Y(t_i)\right]^{\delta_i} \left[S_Y(t_i)\right]^{1-\delta_i}.$$

Note that, formally, the notions of independence and non-informativeness are different; examples can be constructed where censoring is independent yet informative (Andersen et al. (1993, section III.2.2)). The subtlety is not relevant to our intents and purposes and from now on, our assumption of independent censoring will implicitly contain the assumption of non-informative censoring. All the likelihood functions in this thesis are constructed assuming that the hypothetical parameter determining the censoring mechanism contains no information about the parameter of interest determining the event process.

2.2 Overview of recurrent event survival analysis

Extending single event survival analysis to the recurrent events setting is complicated for multiple reasons. First of all, depending on the research question we must decide how to appropriately characterize the association of past event history with future event propensity in our specific research context. There are two related but different ways that researchers have modelled recurrent event processes:

1) analysing the intensity function by specifying the instantaneous probability that an event occurs, conditional on the full event history, or
2) analysing the gap times by specifying either the distributions of gap times between events or the expected event count by some time t, with either marginal or conditional assumptions.

The focus of this thesis will be on a certain class of intensity based models (approach 1 above), with minimal discussion on gap time based analysis. The role of the hazard function is integral to our approach, but requires the use of additional notation.

Suppose m subjects are under observation for an event of interest which each subject can experience multiple times. For subject i, let $T_{i1} < T_{i2} < ... < T_{in_i}$ be the times at which events occur, where T_{ij} is the time of the jth event and n_i events are observed for the i-th subject. Let the amount of time between occurrence of the $(j-1)$th event and occurrence of the jth event be W_{ij}, indexed by $j = 1, ..., n_i$; hence $T_{ik} = \sum_{j=1}^{k} W_{ij}$. Then $\{N_i(t), t \geq 0\}$ is the right-continuous counting process associated with subject i, with typically $N_i(0) = 0$ and $N(T_{ij}) = j$. Let the event history for subject i up until, but not including, time t be denoted $\mathcal{H}_i(t) = \{N_i(s) : 0 \leq s < t\}$. We use $N_i(s,t)$ to denote the number of events occurring in the interval $[s,t]$.

In practical settings subjects will not all be observed for the same period of time; when they are not observed they may or may not be *at risk* to have an observed event. If subject i is observable on the interval $[\tau_{i0}, \tau_i]$, then we use an indicator $Y_i(t) = I(\tau_{i0} \leq t \leq \tau_i)$ that equals 1 only if the subject is at risk of having an observed event at time t.

There will be further discussion in later sections about the choice of time scale (i.e. how the age of a process is measured), but common choices are either calendar time or cumulative exposure time (i.e. amount of time a subject has been observed and at risk of the event). Once the time scale is selected, the *process age* will be used synonymously with *total time*. By contrast *gap time*, the time between successive events or the time since the most recent event, is synonymous with *waiting time*.

2.3 The intensity function

The definition of the hazard function in (2.1) can be alternatively understood as the instantaneous probability of the counting process $\{N(t), t \geq 0\}$ with history $\mathcal{H}_{t^-}$ increasing from 0 to 1 at time t:

$$h(t|\mathcal{H}_{t^-}) = \lim_{\Delta t \to 0} \frac{\Pr\left(N(t+\Delta t) - N(t) = 1 | N(t^-) = 0, \mathcal{H}_{t^-}\right)}{\Delta t}.$$

If the counting process is allowed to continue past the first jump, we can still talk about the instantaneous probability of increasing by exactly 1 at time t but now that there is a process history to consider, we define this quantity *conditional* on all the known aspects of the process history up until then:

$$\lambda(t|\mathcal{H}_{t^-}) = \lim_{\Delta t \to 0} \frac{\Pr\left(N(t+\Delta t) - N(t) = 1 | \mathcal{H}_{t^-}\right)}{\Delta t}. \tag{2.2}$$

This is the *intensity function*, which identifies the event generating mechanism and determines any other characteristic of the event process. Henceforth, the *intensity* of a process will be used synonymously with *conditional intensity*; it is always implied that the intensity at time t is conditional on the process history up until t. Note: counting processes in discrete time scale are beyond the scope of this thesis; we will always assume a continuous time scale.

For a single recurrent event process observed on the time interval $[0, \tau_i]$ with n_i events at times $t_{i1}, t_{i2}, ..., t_{i,n(\tau_i)}$ the likelihood of θ which identifies the model can be shown to be: [Andersen et al. (1993, section 2.7)]

$$L_i(\theta) = \exp\left\{ - \int_0^{\tau_i} \lambda(u|\mathcal{H}_i(u))\, du \right\} \prod_{j=1}^{n_i(\tau_i)} \lambda(t_j|\mathcal{H}_i(t_{ij})) \tag{2.3}$$

provided that the time τ_i which censors the last observed event is a stopping time with respect to the process history up until τ_i. In other words, the "decision" to stop observing the process at time τ_i can depend only on aspects of the process history up until time τ_i and not on anything afterwards. Hence, by assuming independence between subjects, the full likelihood is the product of their contributions:

$$L(\theta) = \prod_{i=1}^{m} L_i(\theta)$$

and maximum likelihood procedures can be employed to estimate θ.

Extra caution must be taken in the presence of censoring. In fact, appropriately modeling dependence conditional on event history becomes extremely delicate if the last waiting time is censored, because the last waiting time is automatically a function of the censoring time and the previous gap times. In the absence of covariates, or if the measured covariates do not contain sufficient information to account for the heterogeneity among subjects in the framework of a model, the censoring can result in biased analysis (Follmann and Goldberg (1988)).

2.4 Minimal and perfect repair: two classical frameworks

The two classical frameworks in which recurrent events are modeled are the Poisson process and the renewal process. This section describes the basic assumptions and intensity functions for these two commonly used models.

A Poisson process is characterized by an intensity function that is unaffected by the occurrence of events. The event history up until time t does not affect the instantaneous probability of an event at time t. In the absence of covariates, the intensity function of a Poisson process is a function of cumulative exposure time only:

$$\lambda(t|\mathcal{H}_{t^-}) = \rho(t).$$

$\rho(t)$ is both the intensity and the *rate function* of the Poisson process (PP); as it does not depend on $\mathcal{H}_{t^-}$, it is the unconditional instantaneous probability of an event at time t. If the intensity function is a constant over time, then it describes a homogeneous Poisson process (HPP); if it varies over time, then it describes a nonhomogeneous Poisson process (NHPP).

The only restriction on $\rho(\cdot)$ is that it must be nonnegative and integrable, since the expected cumulative number of events in $(s, t]$ is given by

$$\mu(s,t) := \mathrm{E}\left[N(s,t)\right] = \int_s^t \rho(u)\, du \qquad \text{(Cook and Lawless (2007, section 2.2.1))}$$

and must be finite. The Poisson process gets its name from the following properties:

- $N(s,t)$ is distributed as Poisson with mean $\mu(s,t)$; and
- the number of events occurring in an interval $(s_1, s_2]$ is a random variable independent of the number of events that occur in an interval $(s_3, s_4]$, provided that $s_2 < s_3$ (sometimes called the property of independent increments). Note that the gap times are in general not independent unless

the process is homogeneous Poisson, in which case the gap times are i.i.d.

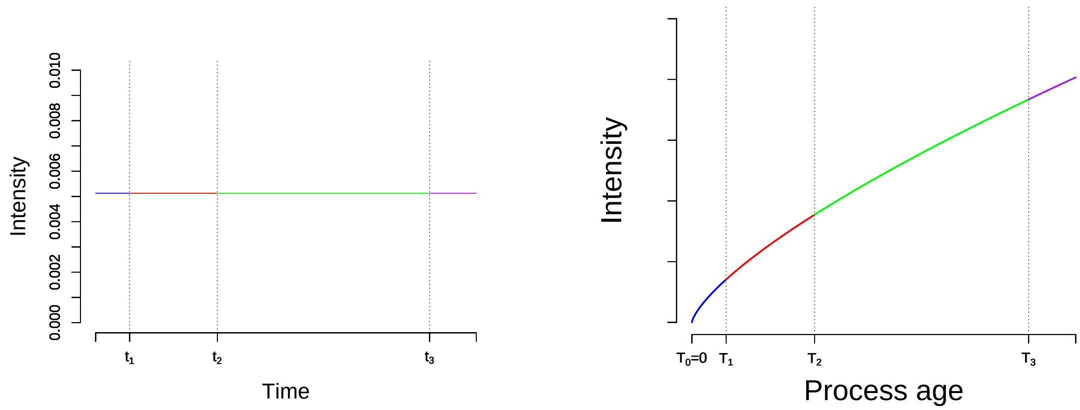

Figure 2–1: Example of a HPP intensity function (left) and a NHPP intensity function (right), emphasizing that the event history, for instance at times t_1, t_2, t_3, does not affect the intensity function of a Poisson process.

On the other hand, a renewal process is characterized by an intensity function that *is* affected by the occurrence of events. After each event, the intensity function resets to its value at time 0 and evolves along a fixed trajectory. In the absence of covariates, or in the presence of covariates which are known in the process history $\mathcal{H}_{t^-}$, the intensity function of a renewal process is a function of only the elapsed time since the most recent event:

$$\lambda(t|\mathcal{H}_{t^-}) = h(t - T_{N(t^-)})$$

where $T_{N(t^-)}$ is the time since $T_0 = 0$ at which the $N(t^-)$th event occurred and $t - T_{N(t^-)}$ is the *waiting time* since the most recent event.

Equivalently, the renewal process (RP) can be characterized by the length of the gaps between successive events being identically distributed, i.e. $W_1, W_2, ..., W_k$ $\overset{i.i.d.}{\sim}$ some distribution F. The function $h(\cdot)$ in the intensity is the hazard function corresponding to F.

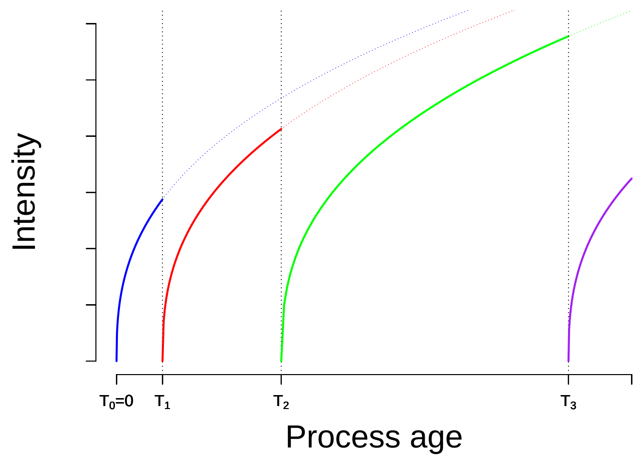

Figure 2–2: Example of a renewal process intensity function, emphasizing that the intensity function undergoes a perfect renewal after each event at times t_1, t_2, t_3

Hence, the distribution of $N(t)$ is given by

$$\Pr\left(N(t) \geq n\right) = \Pr\left(T_n \leq t\right) = \Pr\left(\sum_{j=1}^{n} W_j \leq t\right)$$

whereas the distribution of $N(s,t)$ is generally intractable.

We can view the Poisson process and the renewal process as sitting on opposite ends of the spectrum of *repair* condition. In actuality, any event of interest will take a non-zero amount of time to occur and to be observed, but in survival modelling an event is just an instantaneous jump in a counting process on some time scale. Furthermore, the concept of repair requires a consideration of how much time a repair takes and whether a subject is at risk of failure during a repair. In the field of machine reliability, it is usually realistic to assume that while a system is undergoing repair it is not at risk for the next event; the time axis for the intensity can and should be exactly exposure time. In contrast, in a medical setting this simplification can be less tenable depending on the event of interest. Whereas two machines that have the same operation history but different rest periods can be relied upon to have similar risk for the next failure, two patients with the same event history but different recovery periods can differ drastically in their risk for the next event. Still, for our purposes we will always assume the total time for a given process is its own cumulative exposure time.

Regarding subjects as repairable systems, they follow a Poisson process if they undergo *minimal* repair, i.e. after a failure the system returns to the *state* it was in immediately before the failure and the system's risk of the next failure follows the same trajectory as it would if said failure had not occurred. On the other extreme, they follow a renewal process if they undergo *perfect* repair, i.e. after a failure the system returns to the *state* it was in at time 0 and the trajectory of the system's risk of the next failure restarts and follows the same trajectory as it did from time 0, where the set trajectory does not change with respect to $N(t)$. Hence we will say that the *principal* time scale of the Poisson process is total (or cumulative) time while that of the renewal process is the gap (or waiting) time. Note that the homogeneous Poisson process is the only example of a counting process which is both Poisson and renewal. The only way that a repairable system undergoes both minimal and perfect repair is if its intensity function is constant over its operational lifespan. The homogeneous Poisson process can be specified by the gap times $W_1, W_2, ...$ being i.i.d. Exponential, or equivalently that the number of events in an interval is Poisson with mean proportional to the length of the interval.

Since humans can be thought of as repairable systems experiencing recurrent illnesses or injuries which require treatment and/or rehabilitation, when analysing recurrent injury data the intensity-based models such as the Poisson process and the renewal process are often appealing candidates because of their simplicity and interpretability. However, sometimes neither of these approaches are adequate because a repairable system might repair partially, its behaviour falling patently between the extremes. The model we will examine in the next section can be thought of as either modifying the Poisson process intensity by allowing the intensity function to undergo some sort of renewal after each event or modifying the renewal process by allowing the intensity function to evolve differently after each event.

2.5 Imperfect Repair: the trend-renewal-process

Lindqvist et al. (2003) introduced the trend-renewal process as a generalization of an idea attributed to Berman (1981), who had proposed the inhomogeneous gamma process. Berman noted that the recurrent event process with event times $\{T_1, T_2, ...\}$ obtained by taking every κ-th event of a nonhomogeneous Poisson process with intensity function $\lambda(t)$ is equivalent to the process $\{\Lambda(T_1), \Lambda(T_2), ...\}$, where $\Lambda(t) = \int_0^t \lambda(u)\, du$ is a renewal process in which the gaps $\{\Lambda(T_1), \Lambda(T_2) - \Lambda(T_1), ...\}$ are distributed i.i.d. gamma with unit scale parameter and shape parameter κ. He used this model to define $\{T_1, T_2, ...\}$ as an inhomogeneous gamma process; it reduces to a NHPP if the gaps are i.i.d. unit Exponential (i.e. $\kappa = 1$) and reduces to a RP if $\lambda(t)$ is a constant. Lindqvist extended Berman's strategy to create a more general class of models.

Given a recurrent event process that starts at time $T_0 = 0$ with subsequent event times $\{T_1, T_2, ...\}$, if there exists a monotone (order-preserving) transformation of the time scale such that the failure times on the transformed time scale follow a RP, then $\{T_1, T_2, ...\}$ is a trend-renewal process. Denote this monotone transformation $A(t)$ for $t \geq 0$. Then $V_j = A(T_j) - A(T_{j-1})$, $j = 1, 2, ...$ are i.i.d. with cdf F. Figure 2–3 gives an example of how the transformation might look.

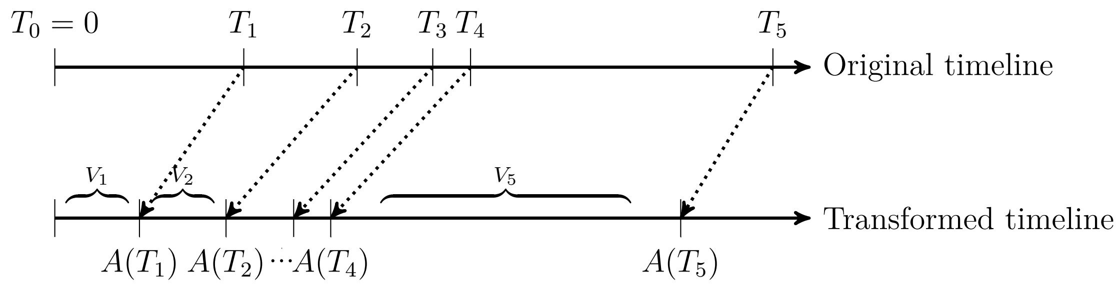

Figure 2–3: Heuristic illustration of an event process on the trend-renewal timescale and its corresponding renewal process timescale after a transformation $A(\cdot)$, for a history with events at $T_1, T_2, ..., T_5$.

Letting $f(v)$ denote the density function of F, $h(v)$ denote the hazard function $\frac{f(v)}{1-F(v)}$, and $a(t) = \dfrac{d}{dt}A(t)$, the intensity function of such a trend-renewal process (TRP) can be written:

$$\lambda(t|\mathcal{H}_{t^-}) = h(A(t) - A(t_{N(t^-)}))a(t). \tag{2.4}$$

Notice that the intensity function is a product of two factors: one which depends only on the time elapsed since the most recent event on the transformed timeline, and another which depends only on the total age of the process. In the words of Cook and Lawless (2007, section 5.2.2), a trend-renewal process seems to describe situations where "the propensity for an event changes over time while the mechanism triggering events is stationary". Pietzner and Wienke (2013) noted that the trend-renewal process is a good "alternative to other recurrent event models especially when the timescale is difficult or unclear" and that while the other common recurrent event models (such as Andersen-Gill (AG), Wei-Lin-Weissfeld (WLW), and Prentice-Williams-Peterson (PWP)) require us to commit to either the total time or the waiting time as the principal time scale of the intensity, the TRP intensity connects the two in a mathematically elegant way. Moreover, $a(\cdot)$ and F are commonly specified such that, if accounting for one of these operating time scales is superfluous in the TRP framework, its corresponding term in the intensity reduces to a constant. One then needs only to select among nested models to determine whether one or both of the operating time scales are indeed significant. For instance, if $a(t)$ is constant in (2.4), then we would have the intensity function of a simple renewal process, and if F is the unit Exponential distribution, then we would have the intensity of a Poisson process.

For a TRP, the distribution of the number of events occurring in $(0, t]$ is given by

$$\Pr\left(N(t) \geq n\right) = \Pr\left(T_n \leq t\right) = \Pr\left(A(T_n) \leq A(t)\right)$$

and is analytically tractable if, for example, the gap times on the transformed timeline are i.i.d. exponential random variables.

To specify a parametric trend-renewal model, we need only specify the monotone function $a(t)$ (called the *trend function*) and the distribution F (called the *renewal distribution*) of the gap times in the renewal process. There exists extensive literature about parametric inference in the TRP framework (Lindqvist et al. (2003); Lindqvist (2006)) and less has been established about semiparametric inference, i.e. when certain assumptions about either the trend function or the renewal distribution are relaxed (Heggland and Lindqvist (2007); Jokiel-Rokita and Magiera (2012)). Recently, Gámiz and Lindqvist (2016) developed a method for nonparametric inference in the TRP framework, using kernel smoothing.

2.5.1 Covariates and unobserved heterogeneity

In many settings a covariate may influence the propensity of having an event and it is necessary to control for that influence, where measurements are available. It is therefore common to incorporate covariates into the trend function, for example in a multiplicative way:

$$a_i(t) = a_0(t)g(\mathbf{X}_i(t)) \tag{2.5}$$

where $g(\cdot)$ is a function of the possibly time-dependent covariates $\mathbf{X}_i(t)$. For fixed, time-independent covariates the most common choice is $g(\mathbf{X}_i) = \exp(\mathbf{X}_i^\top \beta)$ for some parameter β of covariate effects. This formulation would allow, for example, the ratio of the conditional intensities of two processes from different treatments to be a constant at every time t.

When the measured covariates are insufficient to account for differences among subjects, then we have unobserved heterogeneity. In sports injury, this is more commonly referred to as "frailty". A random effect with a prespecified distribution can be incorporated into the model as a surrogate for the underlying mechanisms that cause individuals to differ with respect to their event intensities. In a TRP, this is most often done by specifying a subject i's conditional intensity in the following form:

$$\lambda(t|\mathcal{H}_{t^-}, u_i) = u_i \lambda_0(t|\mathcal{H}_{t^-}) \tag{2.6}$$

where u_i is an unobservable subject-specific realization from a distribution H and multiplicatively modifies the baseline intensity $\lambda_0(\cdot)$ defined by a trend function and renewal distribution common to all subjects. Whereas a TRP model without heterogeneity assumes that subjects with all the same measured covariates and the same time of most recent event have the same intensity going forward regardless of their history before their most recent event, the TRP with heterogeneity models dependence among all the event times within a subject.

This concludes the introduction to the trend-renewal process framework. In principal, there are many choices of trend function and renewal distribution that one can assume to build a trend-renewal model; one can also loosen the assumptions and fit the model semiparametrically. This thesis will not discuss semiparametric and nonparametric inference for TRPs, but the following subsection discusses popular choices of fully parametric TRPs where the trend function is of the "power-law" form in the reliability literature, the renewal distribution is either Weibull or gamma, and in the case of frailty, H is gamma or log-Normal.

2.5.2 Parametric inference

Let the trend function and renewal distribution be parametrically specified and, in the absence of covariates, θ be the parameter vector specifying $A(t)$ and F. Furthermore, let

- $\{t_{ij} : i = 1, ..., m;\ \ j = 1, ..., n_i\}$ be the observed event times for m subjects,
- $v_{ij} = A(t_{ij}) - A(t_{i,j-1})$ for $i = 1, ..., m,\ \ j = 1, ..., n_i$ denote the observed transformed gap times,
- τ_i independently and non-informatively censor subject i's final waiting time, and
- $\nu_i = A_i(\tau_i)$ denote subject i's transformed censoring (or end-of-study) time.

Then by independence between subjects and independence of gap times on the transformed time scale $\{V_{ij} : j = 1, 2, ...\}$, we can write the likelihood of θ as:

$$L(\theta) = \prod_{i=1}^{m} \left\{ \prod_{j=1}^{n_i} f\left(v_{ij}\right) \right\} S\left(\nu_i - A_i(t_{i,n_i})\right). \tag{2.7}$$

Applying the transformation from the renewal timescale to the original timescale, we arrive at the form:

$$L(\theta) = \prod_{i=1}^{m} \left\{ \prod_{j=1}^{n_i} f\left(A_i(t_{ij}) - A_i(t_{i,j-1})\right) a_i(t_{ij}) \right\} S(A_i(\tau_i) - A_i(t_{i,n_i})). \tag{2.8}$$

In general the solutions to the score equations have no closed form but can be found using general-purpose optimization software.

Weibull-power-law trend-renewal model

A popular choice of trend function is $a(t) = ab\,t^{b-1}$ for parameters $a, b > 0$. An equally popular choice of renewal distribution is the Weibull. For identifiability, one must restrict the form of the Weibull, for example to having unit mean or unit scale and a free shape parameter (Pietzner and Wienke (2013)). For convenience, the latter formulation is used in this thesis, so that the hazard function of the renewal distribution is $h(w) = cw^{c-1}$ for parameter $c > 0$.

The function $a(t) = ab\,t^{b-1}$ is the Jacobian of the time transformation $A(t) = at^b$, which in reliability is referred to as "power law" form and also of the same form as a Weibull hazard. Therefore, a TRP where the hazard of the renewal distribution and the trend function are both monomials is often called a Weibull-Weibull (or Weibull-power-law) TRP. In the absence of covariates, its intensity function is:

$$\lambda(t|\mathcal{H}_{t^-}) = a^c bct^{b-1}(t^b - T^b_{N(t^-)})^{c-1}. \tag{2.9}$$

a is merely a scaling parameter and has no influence on the shape of the intensity. If $b = 1$, then (2.9) does not depend on total time; it restarts after the occurrence of each event and so reduces to the intensity of a RP with Weibull gaps. If $c = 1$, then (2.9) does not depend on the wait time and is unaffected by the occurrence of previous events, thereby reducing to the intensity of a PP.

Figure 2–4 shows what the Weibull-power-law TRP intensity function might look like for a process that has had three events, illustrating why this parametric TRP can be seen as nonhomogeneous Poisson process extended to have imperfect renewals. As an example, note that the scenarios corresponding to $b = 1.72, c = 0.8$ and $b = 1.72, c = 0.95$ could reflect situations where the intensity (risk of an event) is always decreasing as time goes on, but overall, the subject is in worse and worse shape after each event compared to how it was going forward from any previous event.

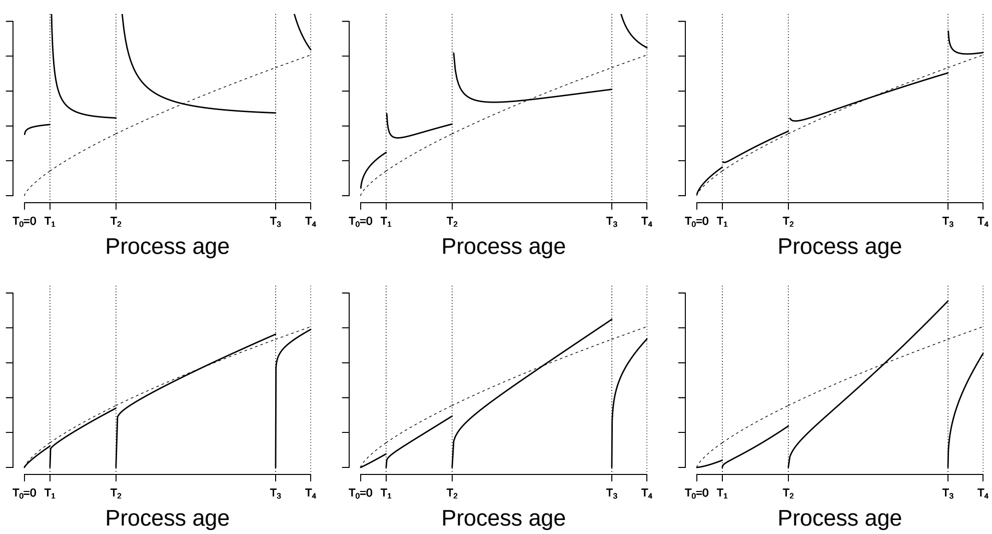

Figure 2–4: Scale-free plots of the TRP intensity function (solid lines) of the form (2.9) with fixed trend parameter $b = 1.72$ and varying renewal parameter c taking values $0.6, 0.8, 0.95, 1.05, 1.2$, and 1.4 (in order of the plots from left to right), compared with the NHPP intensity function (dashed lines) of the form (2.9) with fixed trend parameter $b = 1.72$ and $c = 1$, conditional on a hypothetical event history of event occurrences at T_1, T_2, T_3.

In a later section we discuss the interpretation of the parameters in the Weibull-power-law TRP for practical purposes - in particular, their roles in describing the association between a past event and future risk looking forward.

If subject i has event times $t_{i1}, t_{i2}, ..., t_{i,n_i}$ and censoring time t_{i,n_i+1}, then under the Weibull-power-law TRP framework parametrized by $\theta = (a, b, c)^\top$ the gap times on the renewal timeline are given by

$$v_{i1} = a\, t_{i1}^b \;\; ; \;\; v_{ij} = a\, t_{ij}^b - a\, t_{i,j-1},\; j = 2, 3, ..., n_i + 1.$$

Under independent censoring, its likelihood contribution is given by (2.8) as:

$$L_i(\theta) = \left\{ \prod_{j=1}^{n_i} a^c b\, t_{ij}^{b-1} c\, v_{ij}^{c-1} \right\} \cdot \exp\left(- \sum_{j=1}^{n_i+1} v_{ij}^c \right) \tag{2.10}$$

Fitting a Weibull-power-law TRP to m subjects is just a matter of maximizing $\prod_{i=1}^{m} L_i(\theta)$, which standard software can handle with ease. For a derivation of the maximum likelihood estimators in the Weibull-power-law TRP, see Jokiel-Rokita and Magiera (2012). Estimated standard errors can be obtained from the square root of the diagonal of the inverse observed information matrix. Although there are concerns about making possibly incorrect assumptions on both the trend function and renewal distribution, the clear advantage is that the usual properties of maximum likelihood estimation can be applied, for example, to do hypothesis testing.

Cook and Lawless (2007, chapter 5) discuss trend-renewal models briefly in their book, demonstrating the use of a TRP in section 5.2.2. They let the trend function be $a(t) = \exp(\alpha + \beta t)$ for $\alpha, \beta > 0$ and the renewal distribution be Weibull with unit scale and shape δ. They don't refer to this model as anything other than a trend-renewal process, but it is, more specifically, a reparametrization of the Weibull-power-law TRP. If the original event times $\{T_1, T_2, ...\}$ follow a Weibull-power-law TRP, then $\{Y_1 = \log T_1, Y_2 = \log T_2, ...\}$ follow Cook and Lawless's TRP. This can be seen by reparametrizing the trend function.

From the loglinear trend function $a(y) = \exp(\alpha + \beta y)$, we get the transformation that takes the original process to a renewal process:

$$A(y) = \int_0^y a(u)\,du = e^{\alpha}(e^{\beta y} - 1)/\beta$$

Let $y(t) = \log(t)$ be an additional transformation of the original time scale which we will compose with $A(y)$. Then the trend function for the log time scale is

$$\begin{aligned} a(t) &= \frac{dA(y(t))}{dt} = \frac{dA(y(t))}{y(t)} \cdot \frac{dy(t)}{dt} = e^{\alpha} t^{\beta} \cdot \frac{1}{t} \\ &= e^{\alpha} t^{\beta - 1} \end{aligned}$$

Letting $a = \dfrac{e^{\alpha}}{\beta}$, $b = \beta$, $c = \delta$ we obtain a power-law trend function. Hence, fitting the TRP with the a log-linear trend function and Weibull renewal to event times on the log scale gives the same estimates as fitting the TRP with power-law trend and Weibull renewal to event times on their original scale. There's no clear reason to prefer one parametrization over the other except for convenience – e.g. to facilitate numerical optimization.

Frailty can be modeled multiplicatively in the intensity function (see (2.6)). Then the intensity function of a *heterogeneous Weibull-power-law TRP* is:

$$\lambda(t|\mathcal{H}_{t^-}, \mathbf{X}_i, u_i) = u_i a^c b\, t^{b-1} c(t^b - T^b_{N(t^-)})^{c-1}. \tag{2.11}$$

Note that, conditional on u_i, this is the process intensity of a TRP with trend $a(t) = ab\, t^{b-1}$ and Weibull renewal distribution with shape c and scale $u_i^{-1/c}$.

If u_i follows a distribution H parametrized by some measure of variance γ, then the heterogeneous Weibull-power-law TRP as above is parametrized by $\theta = (a, b, c, \gamma)^{\top}$. Fixing b to 1, c to 1, or γ to 0, or combinations thereof, give nested models. The complete nesting structure can be visually depicted as a cube with relevant submodels at each vertex and each edge connects two models that differ by exactly one parameter (Lindqvist et al. (2003)).

We adopt a notation similar to that of Lindqvist et al. (2003) and use HTRP$(a(\cdot), F, H)$ to denote the parametric "heterogeneous trend-renewal process" that has trend function $a(\cdot)$, renewal distribution F, and frailty distribution H. Then the following are its "closest" submodels:

1. If $a(t)$ is a constant with respect to t then the model becomes a renewal process with frailty denoted by HRP(a, F, H).
2. If F is an Exponential (i.e. with a constant hazard) then it becomes a nonhomogeneous Poisson process with frailty denoted by HNHPP$(a(\cdot), \exp, H)$.
3. If H has a variance of 0, then it becomes a TRP$(a(\cdot), F, 1)$.

If these simplifications happen in combination, then further submodels are obtained. In particular, 1 and 2 in conjunction gives the HHPP$(a, \exp, H)$, which is a well-defined model but, to avoid confusion, should be referred to by its initialism and not the expansion. The directed arrows in Figure 2–5 (similar to the figures in Lindqvist et al. (2003)) depict increases in complexity in three possible directions.

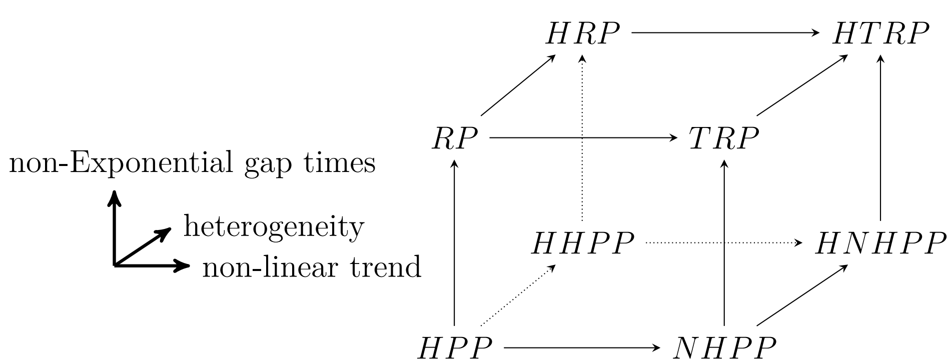

Figure 2–5: Nesting structure of the model with formulation HTRP$(F, a(\cdot), H)$ and its seven submodels, depicted as a cube, with each arrow pointing to the larger model of the two vertices.

Comparing a nested pair of models in one of the three directions corresponds to testing a null hypotheses of either the Poisson property, homogeneity across systems, or constant trend. A likelihood ratio test (LRT) will suffice for selecting among nested models on the cube if certain conditions hold; the asymptotics are valid for heterogeneous subjects only as the number of subjects tends to infinity, and for small samples of homogeneous subjects if they are observed long enough (Andersen et al. (1993, section VI.1.2)). However, it is worth noting that testing a null hypothesis of homogeneity across subjects does not yield a null distribution of the LRT statistic that is the usual chi-squared distribution. There is no heterogeneity across subjects only when the random effect distribution is degenerate, i.e. when the frailty parameter falls on the boundary of its parameter space. One way to perform this test, which a result of Chernoff (1954) provides, is to note that under the null, the LR test statistic takes 0 with probability 1/2 and chi-squared with probability 1/2. This thesis will not explore other options for testing homogeneity across subjects.

To avoid issues with multiple testing, using information criteria such as AIC or BIC may be preferable. Selection using LRT and BIC will be investigated via simulation in a later section but further discussion about appropriate model selection in TRPs is beyond the scope of this thesis. For example, there exist modifications of the LRT for heterogeneity which involve extending the parameter space at the cost of interpretability (e.g. Aalen and Husebye (1991, Appendix II) and Andersen et al. (1993, chapter IX)).

Finally, time-independent covariates can be incorporated in the Weibull-power-law HTRP in the usual Cox proportional hazards construction via the trend function as in (2.5), i.e. by letting $a_i(t) = ab\,t^{b-1}\exp(\mathbf{X}_i^\top \beta)$. The intensity function of a heterogeneous Weibull-power-law TRP with time-independent covariates is:

$$\lambda(t|\mathcal{H}_{t^-}, \mathbf{X}_i, u_i) = u_i a^c \exp(c\mathbf{X}_i^\top \beta) b\, t^{b-1} c(t^b - T^b_{N(t^-)})^{c-1}. \tag{2.12}$$

Note that, conditional on u_i, this is the process intensity of a TRP with trend $a(t) = ab\, t^{b-1} \exp(\mathbf{X}_i^\top \beta)$ and renewal distribution Weibull with shape c and scale $u_i^{-1/c}$.

Gamma renewal

If we keep the power-law trend function $a(t) = ab\, t^{b-1}$ for parameters $a, b > 0$, but define the renewal distribution to be gamma with shape c and scale 1 (again, because the a parameter accounts for the scaling) then we obtain the gamma-power-law trend-renewal model. This allows for a different class of intensity functions than the Weibull-power-law TRP while also containing the Poisson process as a special case. Neither the intensity function nor the likelihood are available in closed form, but the likelihood is easily computed using numerical methods:

$$L(\theta) = \prod_{i=1}^{m} \left\{ \prod_{j=1}^{n_i} f_G \left(at_{ij}^b - at_{i,j-1}^b \right) abt_{ij}^{b-1} \right\} \int_{a\tau_i{}^b - at_{i,n_i}^b}^{\infty} f_G(u)\, du \tag{2.13}$$

where $f_G(x) = \dfrac{1}{\Gamma(c)} x^{c-1} e^{-x}$.

If $b = 1$, then the process reduces to a RP, more specifically, a homogeneous gamma process. If $c = 1$, then the gap times on the renewal scale are again exponential and we obtain a NHPP.

Inference for the TRP with Gamma distributed frailty

General optimization software can usually fit the heterogeneous Weibull-power-law TRP with intensity (2.12) with ease. In Chapter 4, all of the finite sample inference is performed efficiently in R (R Core Team (2015)) for even moderate-sized datasets.

Consider, in the absence of covariates, a fully parametric heterogeneous trend-renewal process with trend $a(\cdot)$, continuous renewal distribution F, and subject-level random effects i.i.d. gamma with mean 1 and variance γ (i.e. shape $1/\gamma$ and scale γ). Let $h(\cdot), S(\cdot)$ denote the hazard function of the renewal distribution. In such a model, subject i's intensity function conditional on their random effect is given by:

$$\lambda(t|\mathcal{H}_{t^-}, u_i) = u_i\lambda_0(t) = a(t)u_i h(A(t) - A(T_{N(t^-)})). \tag{2.14}$$

Conditional on a fixed u_i, (2.14) is the intensity function of a TRP where $a(t)$ is the trend and the renewal distribution has hazard function $h_i^*(t) = u_i h(t)$. In other words, the renewal distribution would conveniently have survivor function $S_i^*(t) = [S(t)]^{u_i}$.

It is worth noting that when fitting parametric TRPs with gamma frailty in the above form, if the subjects are too homogeneous, the MLE of the parameter determining the random effects distribution might fall on the boundary of the parameter space. For frailty that is multiplicative on the intensity function, this corresponds to the frailty distribution being degenerate at 1. If the full unconditional likelihood for a TRP parametrized by $\theta = (a, b, c, \gamma)^\top$ with frailty density f_H is

$$L(\theta) = \prod_{i=1}^{m} \int_0^\infty L_i(\theta \,|\, u_i) f_H(u_i)\, du_i,$$

then it will converge to

$$L(\theta) = \prod_{i=1}^{m} L_i(\theta \,|\, u_i = 1)$$

as the variance γ of the frailty distribution goes to 0. However, reaching an extremely small MLE of γ might cause numerical instability, and so checking that $\hat{\gamma}_{MLE} = 0$ should be done manually, by verifying that $\left.\frac{\partial \log L(\theta)}{\partial \gamma}\right|_{\gamma=0}$ evaluated at the MLEs of the other parameters is negative (Lindqvist et al. (2003)). A negative profile likelihood derivative at the frailty parameter boundary point of interest and the MLEs of the other parameters indicates that there is indeed a local maximum; a positive value means that the MLE of the frailty parameter is strictly not on the boundary.

As a quick example, for a TRP with trend function $a(t) = \frac{dA(t)}{dt}$, renewal distribution gamma (scale 1, shape = c), and frailty distribution gamma (mean = 1, variance = γ), we have from Lindqvist et al. (2003) that the derivative (from the right) of the full log-likelihood, evaluated at $\gamma = 0$ is

$$\frac{\partial \log L(\theta)}{\partial \gamma} \approx \frac{c^2}{2} \sum_{i=1}^{m} \left\{ [n_i - A(T_{i,n_i})]^2 - \frac{n_i}{c} \right\} \tag{2.15}$$

after applying asymptotic approximations of the digamma function. Let subject i be observed from time 0 to time $t_{i,n_i+1} = \tau_i$ with events occurring at $t_{i1}, ..., t_{in_i}$. For convenience, let the transformed observed gap times be $v_{ij} = A(t_{ij}) - A(t_{i,j-1})$, $j = 1, 2, ..., n_i + 1$. Then the conditional likelihood for subject i is:

$$L_i(\theta|u_i) = \left[\prod_{j=1}^{n_i} u_i \lambda_0(t_{ij}) \left[S(v_{ij})\right]^{u_i} \right] \cdot \left[S(v_{i,n_i+1})\right]^{u_i}. \tag{2.16}$$

Letting $f_H(u) = \frac{1}{\Gamma(1/\gamma)\gamma^{1/\gamma}} u^{1/\gamma - 1} \exp(-u/\gamma)$ denote the density of the frailty distribution, the marginal likelihood for subject i is given by

$$L_i(\theta) = \int_0^\infty L_i(\theta|u_i) f_H(u_i)\, du_i,$$

which, because u is gamma distributed, is available in closed form:

$$
\begin{aligned}
L_i(\theta) &= \left[\prod_{j=1}^{n_i} \lambda_0(t_{ij})\right] \cdot \frac{1}{\Gamma(1/\gamma)\gamma^{1/\gamma}} \int_0^\infty u_i^{1/\gamma+n_i-1} \exp\left(-u_i/\gamma\right) \left[\prod_{j=1}^{n_i+1} S(v_{ij})\right]^{u_i} du_i \\
&= \left[\prod_{j=1}^{n_i} \lambda_0(t_{ij})\right] \cdot \frac{1}{\Gamma(1/\gamma)\gamma^{1/\gamma}} \int_0^\infty u_i^{1/\gamma+n_i-1} \exp\left\{-u_i \left(1/\gamma - \sum_{j=1}^{n_i+1} \log S(v_{ij})\right)\right\} du_i \\
&= \left[\prod_{j=1}^{n_i} \lambda_0(t_{ij})\right] \cdot \frac{\Gamma(1/\gamma+n_i)\left(1/\gamma - \sum_{j=1}^{n_i+1} \log S(v_{ij})\right)^{-(1/\gamma+n_i)}}{\Gamma(1/\gamma)\,\gamma^{1/\gamma}} \\
&= \left[\prod_{j=1}^{n_i} \lambda_0(t_{ij})\right] \cdot \left[\prod_{j=1}^{n_i} (1/\gamma+j)\,\gamma\right] \cdot \left(1 - \gamma \sum_{j=1}^{n_i+1} \log S(v_{ij})\right)^{-(1/\gamma+n_i)}
\end{aligned}
$$

where the latter two factors are recognizable as the gamma moment generating function $M(s) = (1-\gamma s)^{-1/\gamma}$ taken to the n_i-th derivative and evaluated at $\sum_{j=1}^{n_i+1} \log S(v_{ij})$.

One might wonder which random effect distributions besides the gamma yield tractable marginal likelihoods. Cook and Lawless (2007, section 3.5.1) note that any distribution with non-negative support and a closed-form Laplace transform (or closed-form moment generating function) will suffice. To see this, let the Laplace transform of such a density f_H be:

$$
\mathcal{L}(s) = \int_0^\infty e^{-su} f_H(u)\, du
$$

whose rth derivative is:

$$
\mathcal{L}^{(r)}(s) = \int_0^\infty (-1)^r u^r e^{-su} f_H(u)\, du.
$$

Then the marginal likelihood for one subject is given by:

$$
\begin{aligned}
L_i(\theta) &= \int_0^\infty \left[\prod_{j=1}^{n_i} \lambda_0(t_ij)\, [S(v_{ij})]^{u_i}\right] \cdot u_i^{n_i}\, [S(v_{i,n_i+1})]^{u_i}\, f_H(u_i)\, du_i \\
&= \left[\prod_{j=1}^{n_i} \lambda_0(t_{ij})\right] \int_0^\infty u_i^{n_i} \exp\left\{u_i \sum_{j=1}^{n_i+1} \log S(v_{ij})\right\} f_H(u_i)\, du_i.
\end{aligned}
$$

Now letting $s^* = -\sum_{j=1}^{n_i+1} \log S(v_{ij})$, we have that:

$$L_i(\theta) = \left[\prod_{j=1}^{n_i} \lambda_0(t_{ij})\right] \mathcal{L}^{(n_i)}(s^*)$$

which is the frailty distribution's Laplace transform taken to as many derivatives as subject i's event count, evaluated at the sum of the renewal scale cumulative hazards at the renewal scale gap times, and scaled by the product of the baseline intensity evaluated at the original event times.

Mathematical tractability is the main advantage of using gamma distributed random effects in the fashion of (2.6) to model frailty in a population but it may not always be a suitable choice. Indeed the choice of frailty distribution can have substantial impact on conclusions in survival analysis. (e.g. see Congdon (1995)) The following gives a brief discussion about an alternative distribution for the subject-level random effects, one that doesn't have a closed form moment-generating function but that does have a simpler interpretation, as is discussed in the next subsection.

Inference for the TRP with log-Normal distributed frailty

Let m subjects follow a process intensity given by (2.6), $i = 1, ..., m$, where u_i is the unobservable subject-specific realization from a log-Normal distribution with median 1. The log-intensity can be written as:

$$\log \lambda(t|\mathcal{H}_i(t), u_i) = \log u_i + \log \lambda_0(t) \tag{2.17}$$

where $\log u_i \sim$ Normal with mean 0 and variance σ^2. Then the frailty parameter σ^2 is directly interpretable as the variance of the error distribution around $\log \lambda_0(t)$, the expected value of the log-intensity at t.

Let the TRP be parametrically specified such that the trend function is $a(t) = \frac{dA(t)}{dt}$ and the renewal distribution has density $f(\cdot)$, hazard $h(\cdot)$, and survivor function $S(\cdot)$. Then the likelihood of the ith process (again with $t_{11}, ..., t_{1,n_i}$ denoting the event times and $\tau_i = t_{i,n_i+1}$ being a stopping time) is:

$$L_i(\theta) = \left\{ \prod_{j=1}^{n_i} a(t_{ij}) h(v_{ij}) \right\} \cdot \mathrm{E}_Z \left[\exp \left\{ n_i Z + e^Z \sum_{j=1}^{n_i+1} \log S\left(v_{ij}\right) \right\} \right]$$

where $v_{ij} = A(t_{ij}) - A(t_{i,j-1}),\ j = 1, ..., n_i + 1$ and the expectation is of a function of Z which follows a Normal distribution with mean 0 and variance σ^2.

Standard numerical optimization software will, in general, be able to handle the maximum likelihood procedure. Exceptions can occur – especially for gradient-based methods – for datasets in which some subjects have extremely high event counts; this issue is relevant and will be encountered in the simulation studies of Chapter 4, but will not be explored further in this chapter.

2.6 A view towards other modulated renewal models

In the previous section, we presented the trend-renewal process as a convenient option for modelling imperfect repair; it can be obtained simply by adding a time trend to a pure renewal process. For this reason, the TRP is one way to add *modulation* to a renewal process, even though the TRP is not generally considered a modulated renewal process. The term modulated RP is most commonly used to describe a semiparametric approach to modelling recurrent events and usually refers to Cox's modulated renewal process first introduced in 1972. Other frameworks for adding modulation to a RP fall under the class of time transform models, as does the TRP itself in some contexts (see Cook and Lawless (2007, section 5.2.2)) however, they involve modelling either the expected length of gap times or the cumulative mean of the counting process itself. We address these alternatives in order to round out the literature review of the closest cousins to the TRP among well-established methods.

2.6.1 Cox modulated renewal model

The modulated renewal process is an extension of the canonical renewal process and provides a flexible way to allow dependence of the intensity function on aspects of the process history. The distribution of the jth gap time W_j is specified conditional on covariates $z(t)$, fixed and known at time $T_{j-1} = \sum_{k=1}^{j-1} W_k$. The mathematical definition of a modulated renewal model is given by the intensity function:

$$\lambda(t|\mathcal{H}_{t^-}) = h_0(t - T_{N_{t^-}})\, g(z(t)) \tag{2.18}$$

where $h_0(\cdot)$ depends only on $t - T_{N_{t^-}}$, the time elapsed since the most recent event, and g depends on arbitrary features of the event history $\mathcal{H}_{t^-}$, such as event count up to time t or previous gap times. As in a renewal process, there is a term in the intensity function that, after each event, resets to what it was at time 0. This piece is the (unspecified) baseline renewal process with hazard function $h_0(\cdot)$, and the covariates $z(t)$ modify this baseline via the function $g(\cdot)$. A flexible parametric modulated renewal process might let $W_j\,|z(t) \sim e^{Y_j}$ where Y_j is some random variable whose mean is $z(t)^\top\beta$. Semiparametric estimation would make no assumptions about the form of $h_0(\text{cot})$ and, in analogy to the semiparametric Cox PH model for the usual single-event survival analysis, a common choice of g is $g(z(t)) = \exp(z(t)^\top\beta)$ (see Cox (1972)).

Comparing the above (2.18) and the TRP intensity (2.4) with trend function (2.5) makes it clear that TRP models are not contained in the class of modulated RPs because the principal time scale of a modulated RP is still assumed to be waiting time. Neither are modulated RPs special cases of TRPs because the covariates in a TRP affect the transformation from the original time scale to the renewal time scale.

2.6.2 Time transform models

Some might argue that a more natural way of modelling times to events, at least in the single time-to-event setting, is not through the hazard function but rather through the expected length of gap times or the cumulative mean of the counting process itself. In the words of D. R. Cox (Reid (1994)), "accelerated life models are in many ways more appealing [than the proportional hazards model] because of their quite direct physical interpretation".

Recall that the accelerated failure time (AFT) model in single-event survival analysis incorporates covariate effects as "accelerating" (or decelerating) the age of a process. In general, given time-fixed covariates $\mathbf{X}$, the hazard function of the first event time is:

$$h(t|\theta) = \theta h_0(\theta t)$$

where θ is some function of $\mathbf{X}$, the most popular choice being $\theta = e^{\mathbf{X}_i^\top \beta}$. Inference on θ can be done parametrically, semiparametrically, or nonparametrically. The many ways to extend the AFT in the recurrent events context loosely form a class of *time transform models*. For instance, Strawderman (2005) presented a semiparametric model incorporating covariate effects in the conditional intensity function. In brief, it assumes that if an event process has fixed covariates $\mathbf{X}$, then conditional on $\mathbf{X}$ the waiting times $W_1, W_2, W_3, ...$ are independent random variables such that $W_j = V_j\, e^{\mathbf{X}^\top \beta}$, $j = 1, 2, 3, ...$ where $V_1, V_2, V_3, ...$ are i.i.d. from some distribution F_0. In the absence of censoring, this is equivalent to letting the hazard function of W_j be:

$$h_j(t|\mathbf{X}) = e^{\mathbf{X}_i^\top \beta} h_0 \left(e^{\mathbf{X}_i^\top \beta} t \right)$$

where $h_0(\cdot)$ is the hazard function associated with F_0. This formulation is in direct analogy to the usual AFT, and is referred to as the accelerated gap time (AGT) model.

Lin et al. (1998) present a marginal model where the counting process $\{N_i(t), t \geq 0\}$ associated with fixed covariates $\mathbf{X}_i$ has a cumulative mean function $\mu_i(t) = \mathrm{E}\left[N_i(t)|\mathbf{X}_i\right]$ of the form

$$\mu_i(t) = \mu_0 \left(e^{\mathbf{X}_i^\top \beta} t \right). \tag{2.19}$$

That is, the expected number of events experienced by a process with fixed covariates $\mathbf{X}_i$ by time t on the original time scale is a given by a (population-fixed) function evaluated at $e^{\mathbf{X}_i^\top \beta} t$. Strawderman (2005) notes that if a process follows an AGT with covariate effects β and renewal distribution F_0, then its mean function is (2.19) with the same β and $\mu_0(\cdot)$ being the renewal function associated with F_0.

Broadly speaking, one may be tempted to think of the TRP as a time transform model too, except that unlike the models mentioned, in a TRP framework the transformation from the original timescale is not limited to a linear transformation in total time t.

2.7 Motivation for studying trend-renewal processes

This chapter has reviewed the foundations of the TRP in detail and discussed briefly the alternatives for modelling recurrent events in the literature. However, outside of specialized literature on reliability modelling, there is less written about TRPs than about the models discussed in Section 2.6. As mentioned before, one of the few examples of the TRP being applied to medical recurrent events is given by Pietzner and Wienke (2013), where the Weibull-power-law TRP with log-Normal distributed frailty is fitted to a dataset of hospital readmission times of colon cancer patients (Gonzalez et al. (2005)).

The reason the TRP is not used in analysis of medical or sports injury data could be because researchers find well-established methods more appealing; it could alternatively be because TRPs are not well understood. It might seem inconvenient to have to establish what it means in a practical context for the recurrent event data to be thought of as one time-scale transform away from following a renewal process. From the lack of literature applying the TRP to data where the subjects are humans, it is unclear what a parametric (or even semiparametric) TRP would offer that the other approaches, such as parametric accelerated-time or semiparametric conditional intensity, do not. The following chapter attempts to address some of these issues by discussing how to more effectively utilize a popular choice of parametric TRP.

CHAPTER 3
Improving the utility of the Weibull-power-law TRP

While semiparametric and nonparametric inference for the TRP may be appropriate when the goal is optimal fit and prediction, parametric inference for the TRP requires relatively smaller sample sizes, can be quite flexible, and allows standard maximum likelihood theory to be applied. Furthermore, the popular choices of power-law trend and Weibull renewal allow for convenient plotting and interpretation for many research questions.

This chapter begins with an overview of standard interpretations of how the trend and renewal parameters in a Weibull-power-law TRP affect the intensity of the process. It then presents two different and targeted considerations of what the trend and renewal parameters imply about the association between past event history and future risk. Finally, it will give a new suggestion for how covariate effects can be modelled differently than in the classical approach.

3.1 Classical interpretation of the Weibull-power-law TRP

Pietzner and Wienke (2013) provide an interpretation of the parameters in a Weibull-power-law TRP in terms of how the parameters affect the behaviour of the intensity function over the interval until the first event and then over the subsequent intervals. Recall from equation (2.9) the intensity function of the Weibull-power-law TRP:

$$\lambda(t|\mathcal{H}_{t^-}) = a^c bct^{b-1} \left(t^b - T^b_{N(t^-)}\right)^{c-1}.$$

The a parameter is only a scaling parameter and does not affect the shape of the intensity function. The roles that the trend parameter b and renewal parameter c play must be addressed separately for the intensity until the first event and for the intensity until subsequent events. For the first event, if $bc > 1$, the intensity is increasing until the first event, irrespective of the shape after the first event. If $bc < 1$, it is decreasing until the first event, again irrespective of the shape after the first event (Pietzner and Wienke (2013)). Table 3–1 provides a useful summary of the behaviour of the intensity function for all subsequent events for various values of b and c.

It is an artifact of the mathematical formulation that it is possible for the shape of the intensity until the first event to have different direction of monotonicity from that of the intensity between all subsequent events. If $bc = 1$ but $b \neq 1, c \neq 1$, then the intensity until the first event is constant and its shape between subsequent events changes according to either the first or the third row of Table 3–1, depending on b.

Figure 2–4 shows some examples of the Weibull-power-law TRP intensity for a subject with three events in its history for $b > 1$ combined with various values of c, and Pietzner and Wienke (2013) provides an example intensity plot for most of the combinations in Table 3–1. However, with these interpretations, we are limited to describing the baseline intensity function for a population only in general terms, such as whether it is increasing or decreasing or whether the renewals follow a upward or downward time trend. When discussing the association between past event history and future risk in, for example, sports injury analysis, more targeted questions might be asked, such as “in what way would the risk of a subject, with a certain history, differ from the event risk of himself had he had a different but comparable history?” or “by how much would the injury risk of a player who gets injured during a season differ from the injury risk of another player who remains injury-free?” These questions are not easily answered by Table 3–1 or a simple plot of the

Table 3–1: Summary of the behaviour of the Weibull-power-law TRP intensity function after the 1st event for different values of the trend and renewal parameters (Pietzner and Wienke (2013))

Behaviour of intensity function		Renewal parameter $c < 1$	$c = 1$	$c > 1$
Trend parameter	$b < 1$	*Decreases* between successive event times; *faster* decrease after jth event than after $(j-1)$th event	*Decreases* and is unaffected by occurrence of events	*Increases* between successive event times; *slower* increase after jth event than after $(j-1)$th event
	$b = 1$	*Decreases* between successive event times in the exact same pattern after each event	Constant w.r.t. time	*Increases* between successive event times in the exact same pattern after each event
	$b > 1$	*Decreases* between successive event times; *slower* decrease after jth event than after $(j-1)$th event	*Increases* and is unaffected by occurrence of events	*Increases* between successive event times; *faster* increase after jth event than after $(j-1)$th event

fitted intensity function.

3.2 New considerations for the Weibull-power-law TRP

Existing interpretations of the trend parameter b and renewal parameter c in the Weibull-power-law TRP are limited to only describing in broad terms how they affect the intensity function underlying an event process. Here, we discuss their roles in quantifying the difference between the intensities of processes with comparable but different histories. The following two subsections each present a formulation of *relative intensity* that quantifies notions such as extent of recovery in recurrent injury data.

3.2.1 Relative intensity: comparing two processes with regard to their most recent event time

For the first formulation let us begin by considering a process that has had one event with a process that did not have that event but was exposed for the same amount of time. What does the combination of the trend and renewal parameter values tell us about the difference in intensity of these processes, going forward from the first process's event?

Let us compare two processes realized from the same Weibull-power-law TRP:

- subject 1 who has been exposed on the interval $[0, t)$ and has had an event at $s < t$;
- subject 2 who has been exposed on the interval $[0, t)$ and has not had an event.

Using the usual parametrization involving a, b, c, the ratio of intensities functions matching subject 1's and subject 2's descriptions, at time t, is:

$$\frac{\lambda_1(t|\mathcal{H}_1(t))}{\lambda_2(t|\mathcal{H}_2(t))} = \frac{ab\,t^{b-1}c(at^b - as^b)^{c-1}}{ab\,t^{b-1}cat^{b(c-1)}} = \left[1 - \left(\frac{s}{t}\right)^b\right]^{c-1}. \tag{3.1}$$

If we think of $x = t - s$ as the time elapsed since s, then we can define

$$r_{s,0}(x) = \left[1 - \left(\frac{1}{1 + x/s}\right)^b\right]^{c-1} \tag{3.2}$$

to be the function which gives the intensity ratio at time $x+s$ as a function of $\frac{x}{s}$, the elapsed time since the event scaled by the time of the event. Note that $r_{s,0}(x)$ does not depend on t, the cumulative time since process initiation.

The behaviour of $r_{s,0}(x)$ is independent of the scaling parameter a and depends only on b and c. Moreover, it suffices to consider just $r_{1,0}(x)$ (i.e. where $s = 1$) because $r_{1,0}(x)$ differs from $r_{s,0}(x)$ only by a horizontal scaling factor of s.

In practical terms, $r_{s,0}(x)$ is the relative intensity of having had a first event versus having had none in the framework of the Weibull-power-law TRP. Its properties can be described as follows, particularly with regard to the trend and renewal parameters:

1. $\lim_{x\to\infty} r_{s,0}(x) = 1$ for any values of b and c, i.e. the effect of having had the event at s "wears off" as more and more time passes.
2. If $c = 1$, then $r_{s,0}(x)$ is identically 1; indeed, as in a NHPP, having an event does not change one's risk of future event from what it would have been if the event had not occurred;
3. Only c controls the monotonicity and concavity of $r_{s,0}(x)$; both c and b control the rate of change (refer to Figure 3–1).
   - If $c < 1$, then $r_{s,0}(x)$ decreases from infinity to 1 in a convex fashion; the closer c is to 1 (for a fixed value of b) the faster the decrease, the greater b is (for a fixed value of c) the faster the decrease.
   - If $1 < c < 2$, then $r_{s,0}(x)$ increases from 0 to 1 in a concave fashion; the closer c is to 1 (for a fixed value of b) the faster the increase, the greater b is (for a fixed value of c) the faster the increase.
   - If $c > 2$, then $r_{s,0}(x)$ increases from 0 first in a convex fashion and then continues concavely to 1; the greater c is (for a fixed value of

b) the later the transition to concave, and again, the greater b is (for a fixed value of c) the faster the increase.

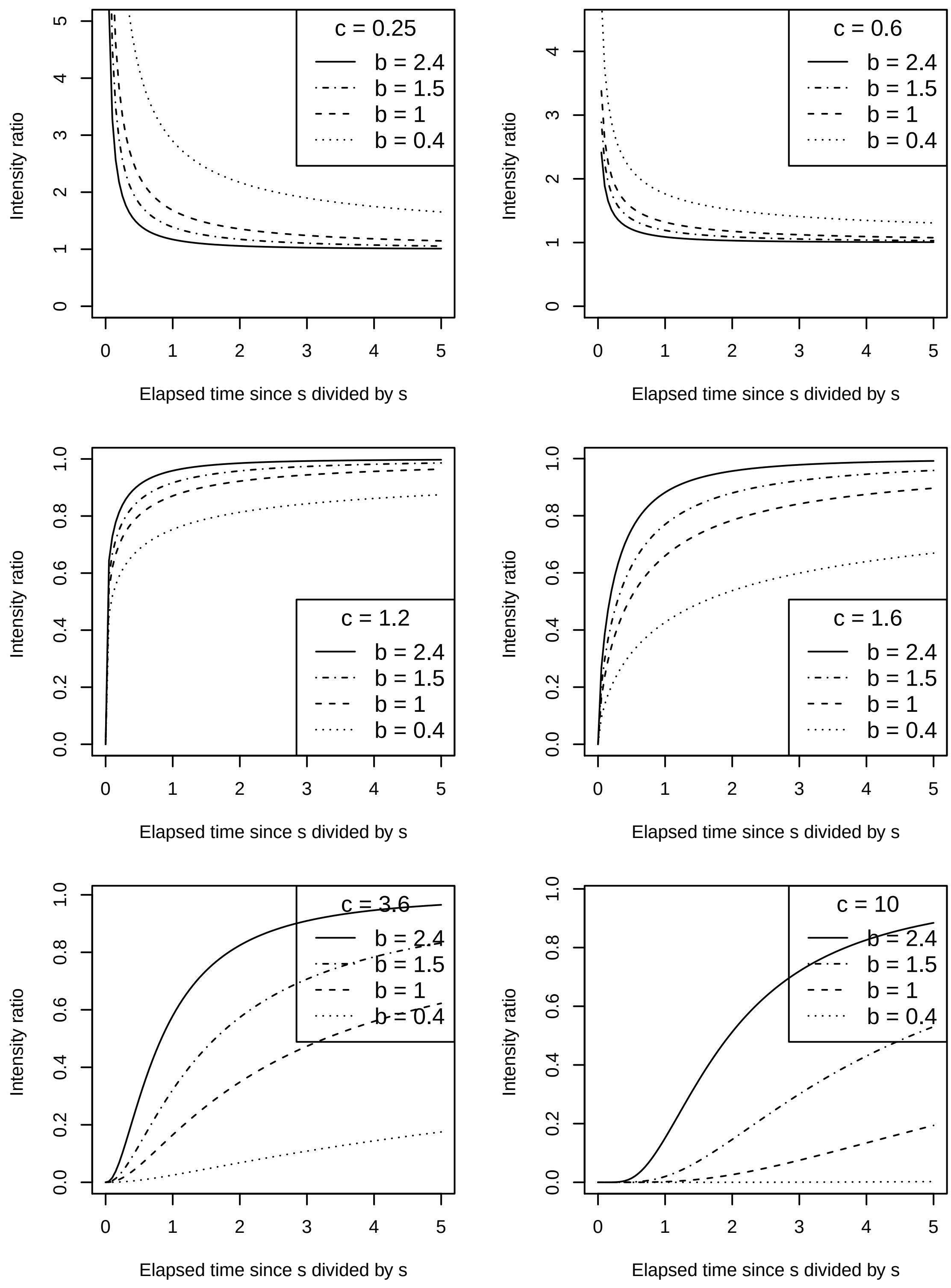

Figure 3–1: Plots of $r_{1,0}(x)$, the relative intensity of having had a first event versus having had none in the Weibull-power-law TRP framework, for various combinations of trend and renewal parameters b and c.

Figure 3–1 illustrates the behaviour of $r_{1,0}(x)$ for various combinations of b and c. These plots are useful because one unit on the time axis represents s amount of time. For example, let the time axis be exposure time and consider two athletes hypothetically following the same Weibull-power-law TRP: one player got injured after playing for $s = 5$ weeks into a season and returned to play for another 5 injury-free weeks while the other had an injury-free 10 weeks. To compare the first player at 10 weeks of his playing time to the second player at 10 weeks of his playing time, we would compute $x = (10 - 5)/5$ as the elapsed time since the more recent event scaled by the time of the more recent event. The ratio of the first player's intensity at 10 weeks of his playing time to the second player's intensity at 10 weeks of his playing time would thus be given by $r_{1,0}(1)$.

The impact of time-independent covariates on the intensity ratio is simple to see if the covariates are incorporated in the power-law trend function as:

$$a(t) = a_0(t)\exp(\mathbf{X}^\top \beta)$$

where $a_0(t) = abt^{b-1}$ is the baseline trend, $\mathbf{X} = (x_1, ..., x_p)^\top$ is the covariate vector, and $\beta = (\beta_1, ..., \beta_p)^\top$ denotes the corresponding effects. With this formulation, we have the familiar interpretation that, fixing all other covariates, every 1 unit increase in x_j corresponds to a change in the baseline trend by a factor of e^{β_j} (i.e. a change in the Weibull-power-law TRP baseline intensity by a factor of $e^{c\beta}$). More generally, if subject 1 with covariates $\mathbf{X}_1$ and subject 2 with covariates $\mathbf{X}_2$ are observed on $[0, t)$ and subject 1 has had an event at time $s < t$ while subject 2 is event-free, then the ratio at time $t = s + x$ of their intensities is given by

$$\frac{\lambda_1(t|\mathcal{H}_1(t), \mathbf{X}_1)}{\lambda_2(t|\mathcal{H}_2(t), \mathbf{X}_2)} = r_{s,0}(x)\exp(c\,(\mathbf{X}_1 - \mathbf{X}_2)^\top \beta)$$

which reduces to (3.1) if the two subjects have no measured differences. Note that this formulation of the baseline intensity incorporates covariate effects merely as scaling constants on the baseline intensity; they are absorbed into the overall scaling parameter a, and cannot effect the shape of the intensity. In a later subsection we discuss how to compare the trend and renewal behaviour between, say, two treatment groups.

One of the reasons why the Weibull-power-law TRP is so easy to work with is the availability of the intensity function in closed form. The fact that the ratio $r_{s,0}(x)$ is easy to interpret is another. Now we make this quantity a bit more general.

Compare two subjects following the same baseline intensity:

- subject 1 with covariates $\mathbf{X}_1$ who has been exposed on the interval $[0, t)$ and whose most recent event was at $s_1 < t$;
- subject 2 with covariates $\mathbf{X}_2$ who has been exposed on the interval $[0, t)$ and whose most recent event was at $s_2 < s_1 < t$.

At time t, the ratio of the intensity functions accounting for covariates and given these histories is:

$$\begin{aligned}\frac{\lambda_1(t|\mathcal{H}_1(t), \mathbf{X}_1)}{\lambda_2(t|\mathcal{H}_2(t), \mathbf{X}_2)} &= \frac{ab\, t^{b-1} c(at^b - as_1^b)^{c-1} \exp(c\mathbf{X}_1^\top \beta)}{ab\, t^{b-1} c(at^b - as_2^b)^{c-1} \exp(c\mathbf{X}_2^\top \beta)} \\ &= \left[\frac{(t/s_1)^b - 1}{(t/s_1)^b - (s_2/s_1)^b}\right]^{c-1} \exp(c(\mathbf{X}_1 - \mathbf{X}_2)^\top \beta).\end{aligned}$$

Letting $p = s_2/s_1 \in [0, 1]$ and $x = t - s_1 > 0$ be the elapsed time since the more recent event, define

$$r_{s_1,s_2}(x) = \left[\frac{(1 + x/s_1)^b - 1}{(1 + x/s_1)^b - (s_2/s_1)^b}\right]^{c-1} \tag{3.3}$$

and hence we have that the ratio of the subjects' intensities depends on the ratio of the two subjects' most recent event times and is a function of x/s_1, the elapsed time since s_1 scaled by s_1. Moreover, similar to before, we can define a scaled quantity

$$r_{1,p}(x) = \left[\frac{(1+x)^b - 1}{(1+x)^b - p^b}\right]^{c-1}. \tag{3.4}$$

(3.4) only differs from (3.3) by a horizontal scaling of s_1 and therefore captures the behaviour of (3.3) where b and c are concerned. And indeed, if $p = 0$ i.e. subject 2 never had an event, then $r_{1,p}(x)$ reduces to the $r_{1,0}(x)$ discussed previously.

A natural question to ask at this point is, how would we construct a confidence interval for $r_{1,p}(x)$ for a fitted Weibull-power-law TRP? It may be sensible to use the parametric bootstrap. After fitting a model, one could easily simulate bootstrap datasets from the TRP specified by the point estimates (for more on the simulation procedure, see section 4.1) and hence approximate a distribution for the point estimates. Then the lower and upper percentiles of the bootstrap parameter estimates can be used to compute the corresponding confidence bounds on $r_{1,0}(x)$. However, at this point constructing confidence bounds on this intensity ratio has not been well developed enough to merit further discussion in this thesis.

In practical terms, $r_{s_1,s_2}(x)$ is the relative intensity of two processes following the same Weibull-power-law TRP which are the same age but differ in the times of their most recent events. It is clear from inspection of (3.4) that, for a fixed p, the properties of $r_{1,p}(x)$ are identical to the three properties of $r_{s_1,0}(x)$ detailed previously. For fixed b and c, the closer p is to 1 (i.e. the closer the two subjects' most recent event times) the faster $r_{1,p}(x)$ approaches 1, and the closer p is to 0 (i.e. the further apart the two subjects' most recent event times) the slower $r_{1,p}(x)$ approaches 1. Figure 3–2 gives plots of $r_{1,p}(x)$ for $c = 0.6, 1.6$ and $b = 0.4, 1.5$, which have similar behaviour to the corresponding plots in Figure 3–1 and show the extra dimension of p where increasing p corresponds to faster convergence of $r_{1,p}(x)$ to 1.

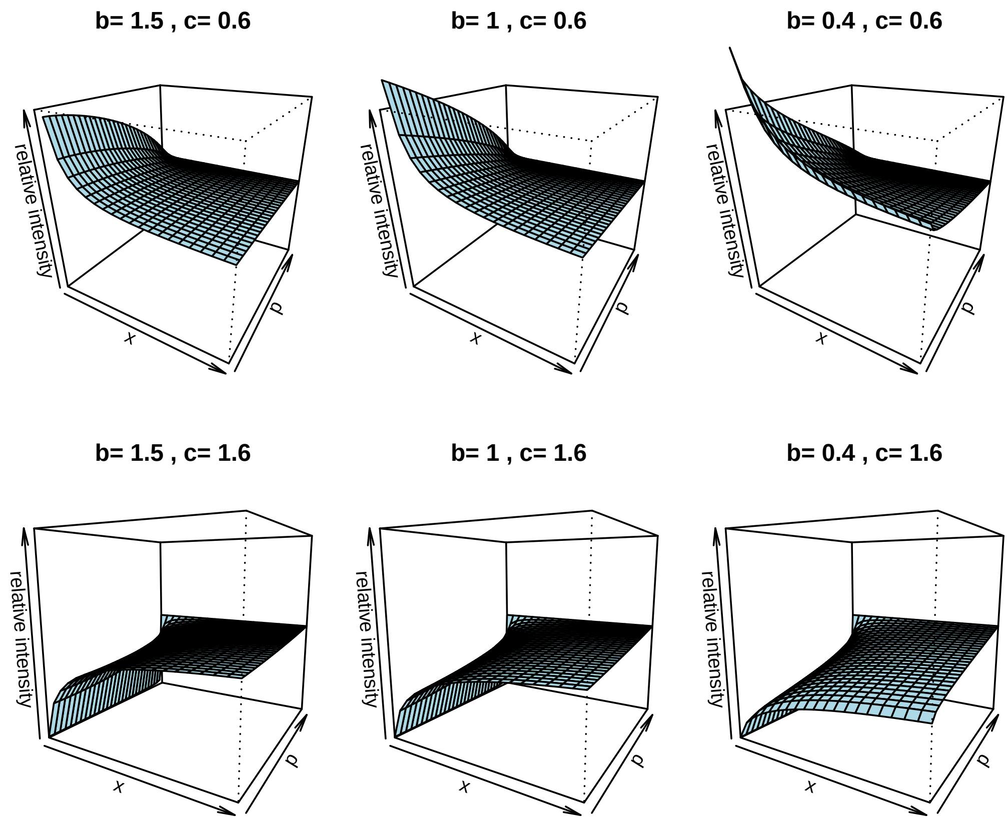

Figure 3–2: Plots of $r_{1,p}(x)$, the relative intensity of having had a most recent event at s_1 versus having had one at $s_2 < s_1$ in the Weibull-power-law TRP framework, for various combinations of trend and renewal parameters b and c.

Finally, we conclude this subsection with an extension of the previous example of what $r_{1,p}(x)$ practically quantifies. Let the time axis be exposure time and consider two athletes hypothetically following the same Weibull-power-law TRP: one player got injured after playing for $s_1 = 8$ weeks into a season and returned to play for another 4 injury-free weeks while the other got injured after playing for $s_2 = 6$ weeks and returned to play for another 4 injury-free weeks. To compare the first player at 10 weeks of his playing time to the second player at 10 weeks of his playing time, we would compute

$p = s_2/s_1 = 0.75$ as the ratio of the event times and $x = (10-8)/8$ as the elapsed time since the more recent event scaled by the time of the more recent event. The ratio of the first player's intensity at 10 weeks of his playing time to the second player's intensity at 10 weeks of his playing time would thus be given by $r_{1,0.75}(0.25)$. We could also go the other way and ask, for instance, how long would the first player play after returning from getting injured at s_1 before his intensity gets to $100q\%$ of the second player's intensity? Estimating this would mean solving for x in $r_{1,0.75}(x) = q$.

3.2.2 Relative intensity: comparing going forward from the most recent event and going forward from a previous event

For the second formulation let us consider a subject following a Weibull-power-law TRP parametrized by a, b, c who has been exposed in the interval $[0, t)$, with a most recent event at $s_1 < t$. How would the process intensity going forward from t compare with the process intensity going forward from an event that occurred prior to the one at time s_1? Let $x = t - s_1$ denote the elapsed time since his most recent event and $s_2 < s_1$ be the time of an earlier event (where $s_2 = 0$ means that the comparison is with the subject's intensity going forward from time 0). Then the comparison can be made using the intensity ratio

$$\frac{\lambda(x+s_1|\mathcal{H}(x+s_1))}{\lambda(x+s_2|\mathcal{H}(x+s_2))} = \frac{a^c\, b\, (x+s_1)^{b-1} c\, [(x+s_1)^b - s_1^b]^{c-1}}{a^c\, b\, (x+s_2)^{b-1} c\, [(x+s_2)^b - s_2^b]^{c-1}}.$$

In a familiar fashion as before, we can define this as the intensity ratio $\mathfrak{r}_{s_1,s_2}(x)$ and write it as a function of $\dfrac{x}{s_1}$, the elapsed since s_1 scaled by s_1, and of

$p = s_2/s_1$ the ratio of the two relevant event times:

$$\begin{aligned}\mathfrak{r}_{s_1,s_2}(x) &= \left(\frac{x/s_1+1}{x/s_1+s_2/s_1}\right)^{b-1}\left[\frac{(x/s_1+1)^b-1}{(x/s_1+s_2/s_1)^b-(s_2/s_1)^b}\right]^{c-1}\\ &= \left(\frac{x/s_1+1}{x/s_1+p}\right)^{b-1}\left[\frac{(x/s_1+1)^b-1}{(x/s_1+p)^b-p^b}\right]^{c-1}.\end{aligned}$$

And hence we can define

$$\mathfrak{r}_{1,p}(x) = \left(\frac{x+1}{x+p}\right)^{b-1}\left[\frac{(x+1)^b-1}{(x+p)^b-p^b}\right]^{c-1} \tag{3.5}$$

which differs from $\mathfrak{r}_{s_1,s_2}(x)$ only by a horizontal scaling of s_1 and whose behaviour is independent of the scaling parameter a.

In practical terms, $\mathfrak{r}_{s_1,s_2}(x)$ is the relative intensity of going forward from the most recent event versus going forward from another earlier event, in the framework of the Weibull-power-law TRP. Inspection of (3.5) tells us two straightforward properties:

1. $\lim_{x\to\infty} \mathfrak{r}_{s_2,s_1}(x) = 1$ for any values of b and c, i.e. in the long run a subject's intensity becomes more and more like how it would be if the most recent event had not occurred;
2. If $b = 1$, then $\mathfrak{r}_{s_2,s_1}(x)$ is identically 1; indeed, as in a RP, after each event occurrence the intensity simply restarts the same fixed trajectory.

If $p = 0$, then the relative intensity quantifies the difference between the intensity going forward from some event and the intensity going forward from process initiation; for example, in sports injury, it may be of interest to compare an athlete after getting injured with either himself or a similar athlete at the start of the season. Hence $\mathfrak{r}_{1,0}(x)$ quantifies how far from perfectly renewed a subject is, after having had an event, if perfectly renewed is the state of the subject going forward from process initiation. Refer to Figure 3–3 for plots of $\mathfrak{r}_{1,0}(x)$ for various combinations of trend and renewal parameter values. While the roles of b and c in the behaviour of $\mathfrak{r}_{1,0}(x)$ are not easily described, but in general:

- If $b > 1$, the ratio decreases from infinity and if $b < 1$ the ratio increases from 0.
- If c is small enough there exists a finite point where the ratio equals 1.
  - If $b > 1$, it means that for a period of time after having had an event at s_1 the subject has a higher intensity than he did going forward from time 0; at some point his risk is equal to what it was at an earlier time in the process but after this point his intensity will be lower than what it was when he was event-free;
  - If $b < 1$, it means that for a period of time after having had an event at s_1 the subject has a lower intensity than he did going forward from time 0; at some point his risk is equal to what it was at an earlier time in the process, but after this point his intensity will be higher than what it was when he was event-free.

Once again, the following is an example of what the intensity ratio $\mathfrak{r}_{1,0}(x)$ practically quantifies. Let us consider just one athlete assumed to follow the Weibull-power-law TRP whose process age is his exposure time. He gets injured after playing for $s_2 = 5$ weeks into a season and returns to play for another 5 injury-free weeks before getting injured again at $s_1 = 10$ weeks of playing time; after he returns he remains injury-free for the rest of the season. To compare his risk at 12 weeks of his cumulative playing time (or 2 weeks after his second injury) with his risk when he had played 2 weeks into the season, we would have $p = 0/s_1 = 0$ and $x = (12 - s_1)/s_1 = 0.2$ as the elapsed time since the more recent event scaled by the time of the more recent event. The ratio of his intensity at 2 weeks after coming back from his second injury

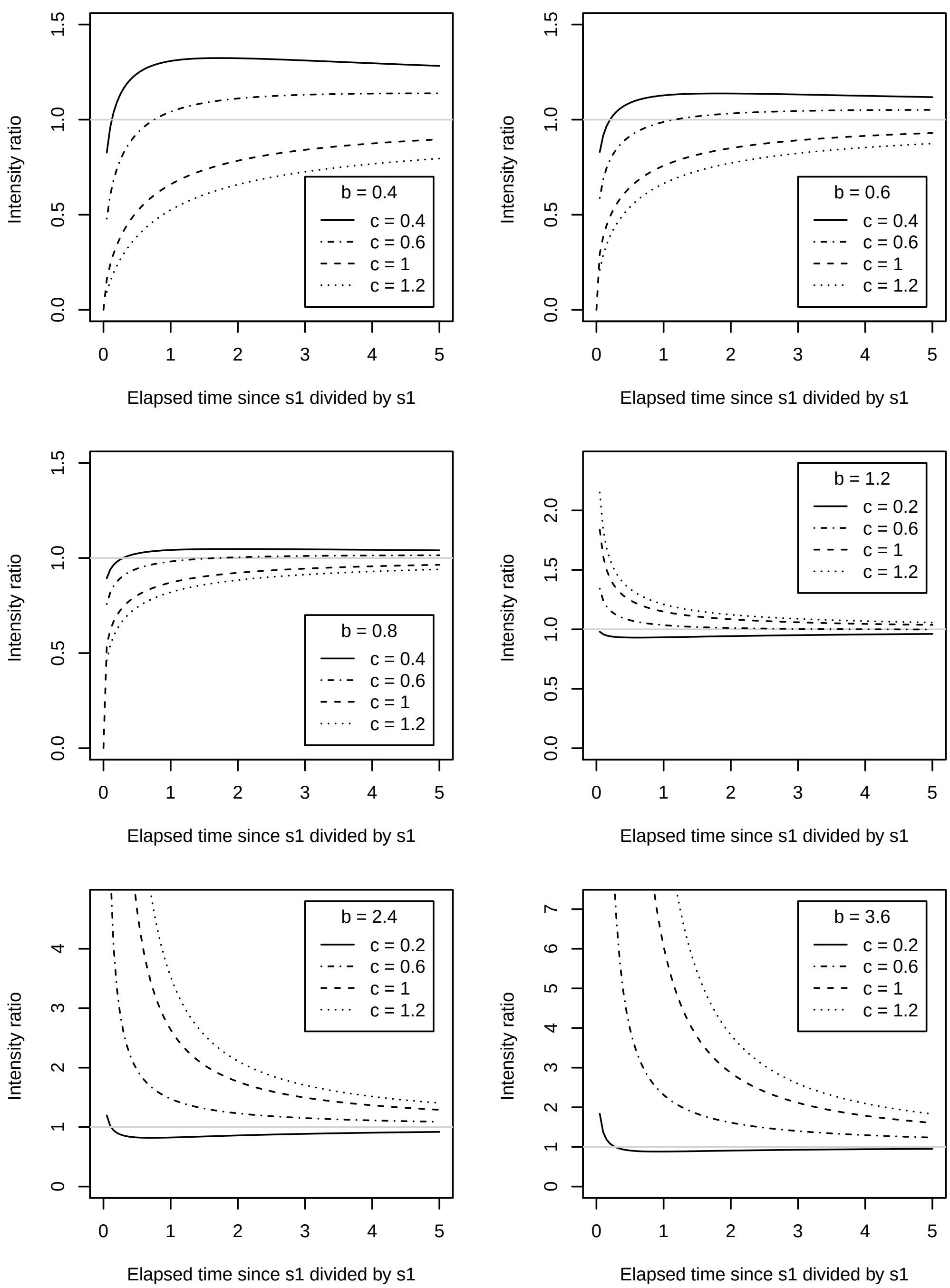

Figure 3–3: Plots of $\mathfrak{r}_{1,0}(x)$, the relative intensity comparing going forward from an event at s_1 and going forward from process initiation in the Weibull-power-law TRP framework, for various combinations of trend and renewal parameters b and c, with a gray line at 1 for reference

to his intensity at 2 weeks into the season would thus be given by $\mathfrak{r}_{1,0}(0.2)$. Alternatively, we could compare his risk at 7 weeks of his cumulative playing time (or 2 weeks after his first injury) with his risk at 2 weeks into the season; we would have $p = 0/s_2$ and $x = (7 - s_2)/s_2 = 0.4$ and compute the relative intensity $\mathfrak{r}_{1,0}(0.4)$. And in certain cases, we could also ask: how long would the first player need to play after returning from getting injured at s_1 before his intensity is equal to what it would be if he were playing that long from the beginning of the season? Estimating this quantity would require solving for x in $\mathfrak{r}_{1,0}(x) = 1$.

3.2.3 New consideration of how to incorporate categorical covariates

In the previous section, we discussed how a given Weibull-power-law TRP can be interpreted in the context of two formulations of relative intensity, where covariates enter the intensity function in a manner analogous to the classical Cox proportional hazards formulation in standard survival analysis. A review of the literature on TRPs reveals only examples where covariates are incorporated in the trend-renewal intensity function in a multiplicative way, i.e.

$$\lambda(t|\mathcal{H}_i(t), \mathbf{X}_i) = \lambda_0(t|\mathcal{H}_i(t)) \exp\left(\mathbf{X}_i^\top \beta\right).$$

For example, Lindqvist et al. (2003) suggests incorporating the effect of covariates in the trend function by letting the trend function for the ith subject be $a_i(t) = g(\mathbf{X}_i)a_0(t)$ so that $g(\mathbf{X}_i)$ scales the "baseline" trend function $a_0(t)$ common to all subjects. The example he gives is $g(\mathbf{X}_i) = \exp(\mathbf{X}_i^\top \beta)$. In the Weibull-power-law TRP, this allows for what we will henceforth call a *proportional intensities* model with a parametric baseline because the intensity takes the form:

$$\lambda(t|\mathcal{H}_i(t), \mathbf{X}_i) = a^c b\, t^{b-1} c(t^b - T_{N(t^-)})^{c-1} \exp(c\mathbf{X}_i^\top \beta). \tag{3.6}$$

Pietzner and Wienke (2013) also use this structure for their analysis of covariate effects.

If the covariates are categorical, then defining the intensity in this manner for a Weibull-power-law TRP simply creates a different scaling factor a for each subgroup, i.e.

$$\begin{aligned}\lambda(t|\mathcal{H}_i(t), \mathbf{X}_i) &= a^c b\, t^{b-1} c \left(t^b - T_{N(t^-)}\right)^{c-1} \exp(c\mathbf{X}_i^\top \beta) \\ &= \exp\left\{c\left(\log a + \mathbf{X}_i^\top \beta\right)\right\} b\, t^{b-1} c \left(t^b - T_{N(t^-)}\right)^{c-1}.\end{aligned}$$

However, it can be limiting to incorporate covariates in the model in such a way that their multiplicative effects on the intensity are merely constant over time. A proportional intensity ratio between two groups means that a process in group A with the same history as a process in group B has the same shape of intensity function. But there are many settings where this might not be the case. Take the hospital readmission data of Gonzales et al. (2005) as an example. If a question of interest is "do women have shorter readmission times on average than men?", then it is appropriate to fix the trend parameter b and the renewal parameter c to be the same for the two groups and allow only the scaling parameter to depend on sex. However, answering the question "does the risk of readmission for women renew differently than the risk of readmission for men?" is more complex. If it is reasonable to assume that the shape of the intensity function is the same for a man and a woman with the same process history, then we can proceed as before. If instead the intensity for males increases over time whereas the intensity for females decreases, then we would run into a familiar problem: just as the Cox proportional hazards model in single-event survival analysis is inappropriate if the hazard function for one group crosses with that of another, the very realistic possibility of crossing intensity functions in the recurrent event setting undermines the suitability of the proportional intensities -type TRP model.

Consider now a second example. It is conceivable that late stage cancer patients would be far more vulnerable to readmission after a hospital visit (e.g. due to surgery or other complications) than patients in early stage. Due to their frailty, the sojourn times of late stage patients could have higher variance than those of early stage patients. This would translate to an increasing intensity for early stage patients and a decreasing intensity for late stage. The next subsection explains how the TRP framework can perhaps be utilized to address this.

Allowing a different renewal parameter for each level

If the research question pertaining to categorical covariate effects is "what do different values of this covariate say about the association between past event history and future risk?" then it might actually make more sense to model the covariates in a Cox-like fashion not on the a parameter, but on the c parameter in the Weibull-power-law TRP.

If there is reason to believe that the shape of the intensity function could be significantly different for the p levels of a predictor, then let the intensity be parametrized by the usual a and b as well as a vector β_c in the following way:

$$\lambda(t|\mathcal{H}_i(t), \mathbf{X}_i) = a^{c_i} b\, t^{b-1} c_i \left(t^b - T^b_{N(t^-)}\right)^{c_i - 1} \tag{3.7}$$

where $\mathbf{X}_i = (X_1, X_2, ..., X_{p-1})^\top$ are indicators and $c_i = \exp\left\{\mathbf{X}_i^\top \beta_c\right\}$ is the renewal parameter corresponding to the subgroup in which subject i belongs. This formulation assumes the same scaling and trend parameters for everyone, to maintain comparability and might be suitable if it is believed that how the underlying propensity for an event changes over time does not vary much among subjects but the evolution of the process triggering events does. For example, athletes in different positions hypothetically differ greatly in style of play but can be expected to become weary over a season at mostly the same pace.

As an illustration, take a hypothetical sample such that each process is a realization of a Weibull-power-law TRP with scale parameter $a = 0.002$ and trend parameter $b = 1.2$ but the renewal parameter c equals $c_1 = 0.6$ for a subject in the untreated group and equals $c_2 = 2$ for a subject in the treated group. We set the censoring time to 730 and simulate a dataset of 100 subjects with half treated and half untreated with median event counts of 4 and 5 respectively. Then, fit to this dataset two models: the true model allowing for different renewal parameters for the two groups but assuming the same scaling and trend, and the alternative model which assumes "proportional intensities", allowing different scaling parameters but assuming the same trend and renewal parameters for everyone. Now consider a process history of one event at time 200. In Figure (3–4) the black solid lines show the true intensities from which the data are simulated, the blue lines show the estimated intensities from the correct model, and the red lines show the estimated intensities from the incorrect model. If the proportional intensities model is fitted to subpopulations that exhibit this kind of disparity, it may not capture the difference between them adequately or accurately. The red solid and dashed red lines respectively illustrate the intensity ratio between the treatment groups if that ratio *has to be constant* over time. But clearly, with such different renewal patterns between the subgroups, the intensity ratio must *vary over time*.

Fitting the wrong model might lead to the inference that one group is consistently at lower risk of readmission than the other, when actually the

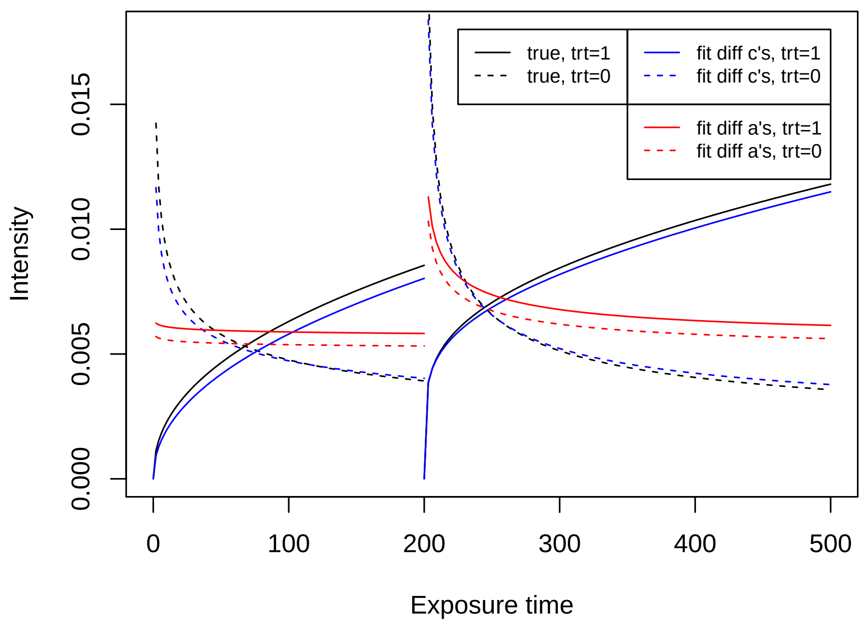

Figure 3–4: True intensity functions (conditional on same event history) for levels of a binary covariate depicted by the black lines with solid for treated and dashed for untreated. The estimated intensity functions from the correctly specified model (different renewal parameter c for each group) are depicted by the blue lines with solid for treated and dashed for untreated. When the incorrectly specified model allowing for a different scaling parameter a for each group is fitted, the estimated intensity functions are depicted by the red lines with solid for treated and dashed for untreated.

difference can be as subtle as the “strong” group’s risk of readmission starting out lower than but eventually surpassing the “weaker” group’s risk. Of course, if we had simulated data from a model with covariate effects modelled by proportional intensities, then the model with covariate effects modelled on the renewal parameter would yield poor fit.

In a practical context, there is no known true model to speak of for any given dataset. Therefore, if one opts to perform analysis in the TRP framework where the data has subpopulations that have to be directly compared and could be reasonably described by significantly different shapes of intensity, one could consider fitting both the models fitted above as well as the larger model which allows different scaling and different renewal in combination. For instance, if one is interested in comparing the p levels of a predictor encoded as above, one could consider letting the intensity function be parametrized by the fixed b as well as the vectors β_a and β_c and defining it as

$$\lambda(t|\mathcal{H}_i(t), \mathbf{X}_i) = a_i^{c_i} b\, t^{b-1} c_i \left(t^b - T_{N(t^-)}^b\right)^{c_i - 1} \tag{3.8}$$

where $a_i = \exp\left\{\mathbf{X}_i^\top \beta_a\right\}$ and $c = \exp\left\{\mathbf{X}_i^\top \beta_c\right\}$ are the scaling and renewal parameters in the intensity function associated with covariate vector $\mathbf{X}_i$. The models (3.6) and (3.7) would be nested in (3.8) and hence can be compared with it using the LRT. If the extra parameters are deemed insignificant, then the two simpler models could reasonably be compared using BIC.

It is conceivable to allow a different c for each subgroup, so why not also consider the case where a different b is allowed for each subgroup? On one hand, Figure 3–4 illustrates an extreme example showing the consequences of fitting a proportional intensities model when the assumption is violated. On the other hand, it is unclear whether there would ever be a practical situation where there would be reason to believe that fixing the scaling parameter and allowing treatment groups to have different trend parameters would yield much better fit or significantly different conclusions than fixing the trend parameter and allowing treatment groups to have different scaling parameters. Even with moderate observed event counts, if the shape of the intensity functions for different subgroups is assumed to be the same (i.e. to maintain comparability by having fixed c) then it is reasonable and also suffices for interpretation's sake to assume that the time trends for two subgroups differ by a constant factor over time.

3.3 Analogous interpretations in other parametric TRPs

The previously introduced ideas for interpretation of parametric TRPs using notions of relative intensities of comparable event histories would also be feasible for different choices of trend function and/or renewal distribution; the Weibull-power-law TRP is only special because the relative intensities we examined have closed form and have behaviour dependent only on two parameters.

For example, in the popular but less convenient gamma-power-law TRP the relative intensity for the pair of comparable event histories defined in section 3.2.1 and in section 3.2.2 is not available in closed form and depends on all the parameters. Nevertheless, the relative intensities can still be graphed and used similarly to our previous demonstration. In fact, preliminary investigation reveals that, if the shape parameter c defining the gamma renewal distribution is the renewal parameter of the gamma-power-law TRP, then the overall effect of the trend parameter b and renewal parameter c on the relative intensity in the gamma-power-law framework is the same as the overall effect of the corresponding parameters on the relative intensity in the Weibull-power-law framework. In other words, the effect of $b > 1$ or $b < 1$ in the gamma-power-law TRP goes in the same direction as the effect of $b > 1$ or $b < 1$ in the Weibull-power-law TRP, and same goes for c. We leave to future work the explicit computation of each relative intensity quantity (and the quantification of the trend and renewal parameter effects) in parametric contexts beyond the Weibull-power-law TRP.

Finally, we also note that it would be straightforward to incorporate categorical covariates in the gamma-power-law TRP by allowing different renewal parameters for each level of a covariate. However, explicit exploration and implementation of this idea to contexts beyond the Weibull-power-law case is beyond the scope of this thesis.

CHAPTER 4
Finite sample evaluation of inference

Maximum likelihood inference is known to work well in large samples, however, in a practical application, we are often faced with the reality of datasets that are potentially not large enough to yield useful or trustworthy parameter estimates.

In the context of analysis on recurrent machine failures, large sample approximations are often appropriate because of the high number of events as well as the given that the operating environment and "subjects" in a sample are made to be as similar as possible. In the contexts of recurrent illness or recurrent sports event, however, we often must make due with limited numbers of subjects and even more limited event counts per subject; this is especially a problem if athletes are sufficiently different that frailty should be accounted for but events within subjects are infrequent. In this chapter, simulation studies are used to gauge how well parameters are estimated in correctly specified models for moderate sample sizes of subjects and low event counts when the data are generated from the Weibull-power-law and gamma-power-law frameworks.

All of the simulation studies in this chapter are conducted in R and optimization is performed with the optim function. Note that in order to have an unrestricted parameter space for optim to search, the log-likelihood function is always reparametrized with the parameters on the log-scale. The stargazer package by Hlavac (2015) streamlined the reporting of results.

We are concerned only with fitting models under the assumption of independent censoring, where the likelihood in (2.8) is valid. This is commonly achieved by either simulating censoring times from a distribution independent of that from which the renewal process gap times are drawn or prespecifying the observed event count. We set censoring times equal to a constant such that the desired average event count n is observed in the subjects. For example, for a TRP without unobserved heterogeneity, the censoring time in the RP timeline for all subjects will be $\nu = n\mathrm{E}[V_1]$, making $\tau = A^{-1}(\nu)$ the censoring time on the TRP timeline.

Finite sample evaluation of inference for categorical covariates in the Weibull-power-law TRP – both the framework where each treatment group has a different scaling parameter (i.e. proportional-intensities covariate effects) and the framework where each treatment group has a different renewal parameter – were conducted but we omit the results from this thesis for two reasons. Firstly, note that to simulate from such frameworks involves merely adjoining smaller datasets simulated from the TRPs defined by the parameter for the respective treatment group and baseline parameters shared by everyone. Hence, the quality of estimation for the baseline parameters depends on the total subject count and observed event counts, while the quality of estimation for each of the subgroup-specific parameters depends on the subgroup's subject count and observed event counts. The simulation studies in this chapter suffice to report equivalent conclusions as would the omitted simulation studies. Another reason is that Pietzner and Wienke (2013) already conducted simulation studies for the heterogeneous Weibull-power-law TRP with two continuous covariates and log-Normal distributed subject-level random effects. They showed that the estimation of all parameters is quite good for moderate subject counts and very good for low event counts.

Pietzner and Wienke (2013) is the only example we found in the literature which evaluates finite sample inference in the TRP framework; whereas they simulated from the heterogeneous Weibull-power-law TRP with log-Normal random effects, we simulate from the heterogeneous Weibull-power-law TRP with gamma random effects and the gamma-power-law TRP with gamma random effects.

We hypothesize that estimation of the baseline parameters in a parametric TRP with random effects is not robust to misspecification of the distribution of the random effects, but this has not yet been investigated in the literature nor in our work.

4.1 Simulation study: Estimation of parameters in the Weibull-power-law TRP

When the true data generating process is the Weibull-power-law TRP without heterogeneity and the correct model is fitted, we expect maximum likelihood estimation of the parameters in the correctly specified model to do extremely well for moderate sample sizes. For this first study, we evaluate how well the scaling, trend, and renewal parameters in the Weibull-power-law TRP can be estimated if the average observed event count is low to moderate.

It is very straightforward to simulate events for one subject. Once the time transformation $A(\cdot)$ and renewal distribution F are chosen, one needs only to:

1. generate $V_1, V_2, V_3, ...$ i.i.d. from F to make gap times from a RP; censor the RP at some fixed time ν such that n events are observed; and
2. compute $T_j = A^{-1}\left(\sum_{k=1}^{i} V_k\right)$ for $j = 1, 2, 3, ..., n$ as the event times on the original timeline.

We carry out this procedure for the parameter configuration $a = 0.004, b = 1.2, c = 0.9$. The trend and renewal parameters are set such that the intensity is increasing until the first event and henceforth decreasing, which is a scenario that has not been simulated from before; the simulation studies reported in Pietzner and Wienke (2013) used a true parameter configuration of $b = 1.5, c = 0.5$. (See the end of section 4.2 for more details about their simulation studies.) Furthermore, b and c are deliberately close to the values that would reduce the TRP to a PP and/or a RP; we want to see how many total observed counts in the data would be adequate to distinguish the fitted model as being significantly different from the PP and RP.

For each combination of censoring time ($\tau = 185, 320, 645$, which respectively yield approximate average observed event counts of 2, 4, and 9) and sample size ($m = 25, 50, 100$) 1000 datasets were generated. To fit the (correctly specified) model, we find the maximizers of the log of the likelihood function (2.10) and the estimated standard errors of the parameter estimates are obtained by taking the square root of the diagonal of the inverse of the observed information matrix. Since all the observed gap times are assumed to be independent within as well as between subjects in a TRP without frailty, we expect the usual 95% confidence intervals to yield adequate coverage.

Tables 4–1 to 4–3 report the results for each censoring time, respectively. Each table displays, for each sample size, the true parameters on the log scale, the mean of the point estimates, the standard error of the point estimates, the mean of the estimated standard errors, and the coverage of the 95% confidence intervals.

In general, we note that at least 200 observed events are needed to distinguish the trend parameter from 1, and at least 400 observed events needed to distinguish the renewal parameter from 1. The gain in precision by doubling the average event count is roughly the same as that of doubling the subject count.

Table 4–1: Results of simulation study 4.1 where the average number of events per subject is approx. 2. Columns refer to subject count, parameter name, true parameter value, mean of the point estimates, standard error of the estimates, mean of the estimated standard errors of the point estimates, and coverage percentage of the 95% confidence intervals.

No. subjects	Parameter	True	Mean	S.e. estimates	Mean s.e.	Coverage
	log(a)	-5.521	-5.604	1.158	1.115	0.939
25	log(b)	0.182	0.178	0.177	0.172	0.937
	log(c)	-0.105	-0.072	0.14	0.139	0.929
	log(a)	-5.521	-5.594	0.745	0.787	0.95
50	log(b)	0.182	0.186	0.117	0.122	0.958
	log(c)	-0.105	-0.096	0.098	0.098	0.950
	log(a)	-5.521	-5.546	0.552	0.549	0.948
100	log(b)	0.182	0.182	0.086	0.086	0.947
	log(c)	-0.105	-0.098	0.071	0.069	0.939

Table 4–2: Results of simulation study 4.1 where the average number of events per subject is approx. 4. Columns refer to subject count, parameter name, true parameter value, mean of the point estimates, standard error of the estimates, mean of the estimated standard errors of the point estimates, and coverage percentage of the 95% confidence intervals.

No. subjects	Parameter	True	Mean	S.e. estimates	Mean s.e.	Coverage
	log(a)	-5.521	-5.661	0.851	0.826	0.953
25	log(b)	0.182	0.194	0.118	0.117	0.939
	log(c)	-0.105	-0.096	0.09	0.089	0.950
	log(a)	-5.521	-5.56	0.553	0.574	0.960
50	log(b)	0.182	0.185	0.079	0.082	0.965
	log(c)	-0.105	-0.099	0.063	0.063	0.951
	log(a)	-5.521	-5.531	0.401	0.403	0.938
100	log(b)	0.182	0.182	0.058	0.058	0.940
	log(c)	-0.105	-0.100	0.044	0.044	0.948

4.2 Simulation study: Estimation of parameters in the Weibull-power-law TRP with unobserved heterogeneity

Now we evaluate the finite sample parameter estimation in the heterogeneous Weibull-power-law TRP in the absence of covariates but with subject-level random effects following a gamma distribution with mean 1 and scale

Table 4–3: Results of simulation study 4.1 where the average number of events per subject is approx. 9. Columns refer to subject count, parameter name, true parameter value, mean of the point estimates, standard error of the estimates, mean of the estimated standard errors of the point estimates, and coverage percentage of the 95% confidence intervals.

No. subjects	Parameter	True	Mean	S.e. estimates	Mean s.e.	Coverage
25	log(a)	-5.521	-5.589	0.568	0.578	0.966
	log(b)	0.182	0.188	0.072	0.074	0.959
	log(c)	-0.105	-0.100	0.054	0.054	0.952
50	log(a)	-5.521	-5.537	0.405	0.406	0.949
	log(b)	0.182	0.182	0.052	0.052	0.954
	log(c)	-0.105	-0.103	0.038	0.038	0.946
100	log(a)	-5.521	-5.522	0.292	0.286	0.942
	log(b)	0.182	0.182	0.037	0.037	0.946
	log(c)	-0.105	-0.104	0.027	0.027	0.950

parameter γ. We assess how well the parameters can be estimated in a correctly specified model for datasets with low to moderate event counts.

It is also straightforward to simulate events for a subject following the heterogeneous Weibull-power-law TRP with subject-level frailty whose intensity function is given by (2.12). Subject i's event times can be generated as follows:

1. Sample u_i from the gamma distribution with shape $1/\gamma$ and scale γ.
2. Generate $V_{i1}, V_{i2}, V_{i3}, ...$ i.i.d. Weibull with shape c and scale $u_i^{1/c}$.
3. Compute $T_{ij} = \left(\frac{1}{a}\sum_{k=1}^{i} V_{ik}\right)^{1/b}$ for $j = 1, 2, 3, ...$ as the event times on the original timeline.

We carry out this procedure for the parameter configuration $a = \times 10^{-5}, b = 1.8, c = 2.4, \gamma = 0.5$. For each combination of censoring time ($\tau = 380, 560, 730$, which respectively yield approximate average observed event counts of 2, 4, and 9) and sample size ($m = 25, 50, 100, 200$) 1000 datasets were generated.

Tables 4–4 to 4–6 summarize, in order, the results for each censoring time. The columns in each table give the number of subjects in each dataset, the reparametrized parameters, the true values, the mean of the point estimates, the standard error of the point estimates, the mean of the estimated standard errors, and the 95% CI coverage, based on 1000 datasets.

The results show that when the heterogeneous model is correctly specified and the subject count is moderately high, the estimation of a, b, and c is excellent even when most subjects have a low event count. However, for some data sets when the mean observed event count is 2, numerical maximization of the likelihood resulted in either non-convergence or convergence to extremely small estimates of the frailty variance γ; this is reflected in the footnote and otherwise large estimated standard errors in the fourth row of Table 4–4.

4.3 Simulation study: Estimation of parameters in the gamma-power-law TRP with unobserved heterogeneity

Now we evaluate finite sample parameter estimation in the heterogeneous gamma-power-law TRP in the absence of covariates but with a random effect (frailty) at the subject level which follows a gamma distribution with mean 1 and scale parameter γ.

As before, let $h(\cdot)$ and $S(\cdot)$ denote the hazard and survivor functions of the renewal distribution. Recall from section 2.1.3 that, conditional on the random effect u_i, subject i's intensity function is that of a TRP with power law trend and renewal distribution with survivor function $S_i^*(t) = [S(t)]^{u_i}$. Since the density, survivor, and quantile functions for the gamma distribution are

[1] 8 of 1000 datasets had an estimate of γ that was too close to 0 for the standard error to be estimable.

Table 4–4: Results of simulation study 4.2 where the average number of events per subject is approx. 2. Columns refer to subject count, parameter name, true parameter value, mean of the point estimates, standard error of the estimates, mean of the estimated standard errors of the point estimates, and coverage percentage of the 95% confidence intervals.

No. subjects	Parameter	True	Mean	S.e. estimates	Mean s.e.	Coverage
25	log(a)	-9.721	-9.782	0.837	0.829	0.944
	log(b)	0.588	0.591	0.076	0.076	0.948
	log(c)	0.875	0.901	0.128	0.128	0.946
	log(γ)[1]	-0.693	-0.954	1.503	3.972	0.947
50	log(a)	-9.721	-9.705	0.578	0.584	0.945
	log(b)	0.588	0.585	0.054	0.054	0.951
	log(c)	0.875	0.885	0.086	0.090	0.955
	log(γ)	-0.693	-0.800	0.542	0.704	0.971
100	log(a)	-9.721	-9.751	0.411	0.415	0.950
	log(b)	0.588	0.590	0.038	0.038	0.947
	log(c)	0.875	0.880	0.065	0.063	0.953
	log(γ)	-0.693	-0.733	0.308	0.298	0.976
200	log(a)	-9.721	-9.704	0.290	0.292	0.943
	log(b)	0.588	0.586	0.027	0.027	0.949
	log(c)	0.875	0.880	0.046	0.045	0.934
	log(γ)	-0.693	-0.712	0.202	0.204	0.954

readily available in base statistical programs, the inverse probability transform can be exploited to generate event times. In other words, conditional on u_i, the random variable $S^{-1}(1 - Z^{1/u_i})$ (where $Z \sim Unif(0,1)$) has survivor function $S^*(\cdot)$.

Now subject i's times can be simulated by the following procedure:

1. Generate u_i from the gamma distribution with shape $1/\gamma$ and scale γ.
2. Generate $Z_{i1}, Z_{i2}, ...$ i.i.d. standard Uniform.
3. Compute $V_{i1}, V_{i2}, V_{i3}, ...$ as the $(1-Z_{i1}^{1/u_i})th, (1-Z_{i2}^{1/u_i})th, (1-Z_{i3}^{1/u_i})th, ...$ quantiles of the gamma with unit scale and shape c.
4. Compute $T_{ij} = \left(\frac{1}{a}\sum_{k=1}^{i} V_{ik}\right)^{1/b}$ for $j = 1, 2, 3, ...$ be the event times on the original timeline.

Table 4–5: Results of simulation study 4.2 where the average number of events per subject is approx. 4. Columns refer to subject count, parameter name, true parameter value, mean of the point estimates, standard error of the estimates, mean of the estimated standard errors of the point estimates, and coverage percentage of the 95% confidence intervals.

No. subjects	Parameter	True	Mean	S.e. estimates	Mean s.e.	Coverage
25	log(a)	-9.721	-9.732	0.503	0.507	0.952
	log(b)	0.588	0.587	0.044	0.044	0.954
	log(c)	0.875	0.886	0.079	0.076	0.943
	log(γ)	-0.693	-0.799	0.456	0.417	0.972
50	log(a)	-9.721	-9.726	0.370	0.360	0.941
	log(b)	0.588	0.587	0.032	0.031	0.940
	log(c)	0.875	0.878	0.053	0.054	0.955
	log(γ)	-0.693	-0.742	0.283	0.281	0.960
100	log(a)	-9.721	-9.726	0.259	0.254	0.945
	log(b)	0.588	0.588	0.023	0.022	0.948
	log(c)	0.875	0.878	0.038	0.038	0.952
	log(γ)	-0.693	-0.711	0.193	0.195	0.962
200	log(a)	-9.721	-9.730	0.182	0.180	0.951
	log(b)	0.588	0.588	0.016	0.016	0.937
	log(c)	0.875	0.876	0.027	0.027	0.944
	log(γ)	-0.693	-0.708	0.136	0.137	0.956

The above procedure is carried out for the parameter configuration $a = 7 \times 10^{-4}, b = 1.6, c = 4.2, \gamma = 0.5$. For each combination of censoring time $\tau = 600, 1000$ (respectively yielding average observed event counts of approximately 4 and 9) and sample size $m = 25, 50, 100, 200$, 1000 datasets were generated.

Tables 4–7 to 4–9 report the results for each respective censoring time. The columns are, in order: the number of subjects in each dataset, the reparametrized parameters, the true values, the mean of the point estimates, the standard error of the point estimates, the mean of the estimated standard errors, and the 95% CI coverage, based on 1000 datasets.

Table 4–6: Results of simulation study 4.2 where the average number of events per subject is approx. 9. Columns refer to subject count, parameter name, true parameter value, mean of the point estimates, standard error of the estimates, mean of the estimated standard errors of the point estimates, and coverage percentage of the 95% confidence intervals.

No. subjects	Parameter	True	Mean	S.e. estimates	Mean s.e.	Coverage
25	log(a)	-9.721	-9.734	0.390	0.388	0.947
	log(b)	0.588	0.588	0.033	0.032	0.943
	log(c)	0.875	0.882	0.058	0.057	0.943
	log(γ)	-0.693	-0.791	0.362	0.347	0.960
50	log(a)	-9.721	-9.727	0.282	0.274	0.942
	log(b)	0.588	0.588	0.024	0.023	0.940
	log(c)	0.875	0.879	0.04	0.040	0.945
	log(γ)	-0.693	-0.745	0.241	0.240	0.954
100	log(a)	-9.721	-9.719	0.199	0.194	0.946
	log(b)	0.588	0.587	0.016	0.016	0.944
	log(c)	0.875	0.877	0.028	0.028	0.948
	log(γ)	-0.693	-0.715	0.172	0.168	0.950
200	log(a)	-9.721	-9.72	0.143	0.137	0.938
	log(b)	0.588	0.588	0.012	0.011	0.940
	log(c)	0.875	0.876	0.019	0.020	0.958
	log(γ)	-0.693	-0.703	0.120	0.118	0.953

Once again, the results show that when the heterogeneous model is correctly specified and the subject count is moderately high, the estimation of a, b, and c is excellent even when most subjects have a low event count. However, for some data sets when the mean observed event count is 2, numerical maximization of the likelihood resulted in either non-convergence or convergence to extremely small estimates of the frailty variance γ; this is reflected in the footnote and otherwise large average estimated standard error in the fourth row of Table 4–7.

Table 4–7: Results of simulation study 4.3 where the average number of events per subject is approx. 2

No. subjects	Parameter	True	Mean	S.e. estimates	Mean s.e.	Coverage
25	log(a)	-7.264	-7.217	0.975	0.964	0.934
	log(b)	0.470	0.468	0.088	0.087	0.944
	log(c)	1.435	1.499	0.252	0.244	0.937
	$\log(\gamma)$[2]	-0.693	-1.123	1.955	10.385	0.939
50	log(a)	-7.264	-7.247	0.672	0.682	0.944
	log(b)	0.470	0.469	0.061	0.062	0.945
	log(c)	1.435	1.465	0.166	0.171	0.948
	$\log(\gamma)$	-0.693	-0.801	0.643	0.950	0.961
100	log(a)	-7.264	-7.261	0.481	0.482	0.940
	log(b)	0.470	0.470	0.044	0.044	0.941
	log(c)	1.435	1.449	0.120	0.120	0.953
	$\log(\gamma)$	-0.693	-0.734	0.360	0.338	0.956
200	log(a)	-7.264	-7.260	0.343	0.340	0.944
	log(b)	0.470	0.470	0.031	0.031	0.943
	log(c)	1.435	1.443	0.083	0.085	0.947
	$\log(\gamma)$	-0.693	-0.716	0.233	0.230	0.960

[2] 5 of 1000 datasets had an estimate of γ that was too close to 0 for the standard error to be estimable.

Table 4–8: Results of simulation study 4.3 where the average number of events per subject is approx. 4

No. subjects	Parameter	True	Mean	S.e. estimates	Mean s.e.	Coverage
25	log(a)	-7.264	-7.276	0.620	0.616	0.943
	log(b)	0.470	0.471	0.054	0.054	0.949
	log(c)	1.435	1.457	0.160	0.156	0.941
	log(γ)	-0.693	-0.830	0.669	0.485	0.974
50	log(a)	-7.264	-7.276	0.455	0.436	0.938
	log(b)	0.470	0.471	0.040	0.038	0.938
	log(c)	1.435	1.442	0.107	0.110	0.955
	log(γ)	-0.693	-0.743	0.315	0.310	0.964
75	log(a)	-7.264	-7.271	0.327	0.308	0.928
	log(b)	0.470	0.470	0.028	0.027	0.936
	log(c)	1.435	1.439	0.080	0.078	0.943
	log(γ)	-0.693	-0.716	0.215	0.214	0.955
100	log(a)	-7.264	-7.272	0.220	0.217	0.948
	log(b)	0.470	0.471	0.020	0.019	0.939
	log(c)	1.435	1.437	0.056	0.055	0.942
	log(γ)	-0.693	-0.709	0.148	0.150	0.961

Table 4–9: Results of simulation study 4.3 where the average number of events per subject is approx. 9

No. subjects	Parameter	True	Mean	S.e. estimates	Mean s.e.	Coverage
25	log(a)	-7.264	-7.265	0.412	0.402	0.931
	log(b)	0.470	0.470	0.033	0.033	0.940
	log(c)	1.435	1.441	0.104	0.101	0.938
	log(γ)	-0.693	-0.773	0.354	0.336	0.947
50	log(a)	-7.264	-7.278	0.288	0.284	0.941
	log(b)	0.470	0.471	0.024	0.023	0.936
	log(c)	1.435	1.439	0.069	0.071	0.959
	log(γ)	-0.693	-0.738	0.235	0.234	0.949
75	log(a)	-7.264	-7.267	0.195	0.201	0.953
	log(b)	0.470	0.470	0.016	0.016	0.953
	log(c)	1.435	1.436	0.051	0.050	0.945
	log(γ)	-0.693	-0.718	0.161	0.164	0.958
100	log(a)	-7.264	-7.265	0.143	0.142	0.942
	log(b)	0.470	0.470	0.012	0.012	0.956
	log(c)	1.435	1.436	0.036	0.036	0.941
	log(γ)	-0.693	-0.704	0.119	0.116	0.949

CHAPTER 5
Application to medical data

The trend-renewal process has potential for application in many medical settings; it is often of interest to analyze, for example, the rate of recurrence of a disease or condition within a population or to quantify the difference in propensity for the event of interest between treatment groups. In this section we extend the analysis of Pietzner and Wienke (2013) on the dataset from Gonzalez et al. (2005). Available in the R library frailtypack, this dataset gives the hospital readmission times of 403 colorectal cancer patients, i.e. the time between the first admission and the second admission, and the time between all subsequent readmissions. The first admission is defined as time point 0. The time-independent covariates that are of interest are: sex (male or female), chemotherapy (yes or no), and Duke's stage (3 groups: combined stage A-B, stage C, or stage D).

Among the patients, 199 had no readmissions, and 150 had one or two readmissions. The patient with the highest event count had 22 readmissions. 458 were censoring times out of a total of 861 observed gap times. While the median event count is 1, this is partially offset by the large sample size of subjects; this is relevant because – as the simulation studies demonstrated – if events are too scarce then the heterogeneity is impossible to estimate.

The TRP, RP, and NHPP fitted to the data has the intensity function:

$$\lambda(t|\mathcal{H}(t), \mathbf{X}) = a^c b\, t^{b-1} c(t^b - T^b_{N(t^-)})^{c-1} \exp(\mathbf{X}^\top \beta)$$

where $\mathbf{X}_i = (X_{chemo}, X_{female}, X_{DukeC}, X_{DukeD})^\top$ is the covariates vector of indicators and b, respectively c, are constrained to 1 for the latter two models.

Table 5–1: Estimated parameter values (standard error) for Weibull-power-law models with proportional intensities covariate effects and no heterogeneity for the hospital readmissions dataset

Parameter	NHPP		RP		TRP	
log(a)	-5.22	(0.263)	-7.12	(0.204)	-10.1	(0.791)
log(b)	-0.278	(0.043)	0	–	0.339	(0.072)
log(c)	0	–	-0.607	(0.103)	-0.769	(0.056)
Chemotherapy	-0.249	(0.104)	-0.235	(0.103)	-0.18	(0.104)
Female	-0.503	(0.100)	-0.441	(0.100)	-0.439	(0.101)
Duke's stage C	0.394	(0.120)	0.347	(0.119)	0.366	(0.119)
Duke's stage D	1.57	(0.128)	1.21	(0.129)	1.26	(0.130)
BIC	6912.3		6647.1		6630.0	

The HTRP, HRP, and HNHPP fitted to the data has the intensity function:

$$\lambda(t|\mathcal{H}_i(t), \mathbf{X}_i, u_i) = u_i\, a^c b\, t^{b-1} c(t^b - T^b_{N(t^-)})^{c-1} \exp(\mathbf{X}_i \beta)$$

with again b, respectively c, constrained to 1 for the latter two models, and either $u_i \sim \exp(Y_i)$ where $Y_i \sim \mathcal{N}(0, \sigma^2)$ (see Table 5–2) or $u_i \sim$ Gamma with mean 1 and variance γ (see Table 5–3).

Table 5–2: Estimated parameter values (standard error) for Weibull-power-law models with proportional intensities covariate affects and log-Normal distributed unobserved heterogeneity fitted to the hospital readmissions dataset

Parameter	HNHPP		HRP		HTRP	
log(a)	-6.23	(0.255)	-7.58	(0.023)	-9.48	(0.709)
log(b)	-0.182	(0.027)	0	–	0.225	(0.068)
log(c)	0	–	-0.448	(0.020)	-0.554	(0.055)
Chemotherapy	-0.214	(0.206)	-0.194	(0.112)	-0.155	(0.148)
Female	-0.552	(0.170)	-0.471	(0.135)	-0.473	(0.142)
Duke's stage C	0.399	(0.196)	0.342	(0.123)	0.356	(0.166)
Duke's stage D	1.66	(0.300)	1.3	(0.158)	1.37	(0.195)
$\log(\sigma^2)$	0.192	(0.130)	-0.456	(0.023)	-0.497	(0.214)
BIC	6696.1		6566.6		6561.3	

Table 5–3: Estimated parameter values (standard error) for Weibull-power-law models with proportional intensities covariate affects and gamma distributed unobserved heterogeneity fitted to the hospital readmissions dataset

Parameter	NHPP		RP		TRP	
log(a)	-5.510	(0.316)	-7.05	(0.245)	-8.96	(0.698)
log(b)	-0.194	(0.045)	0	–	0.230	(0.069)
log(c)	0	–	-0.207	(0.144)	-0.568	(0.055)
Chemotherapy	-0.242	(0.171)	-0.207	(0.144)	-0.164	(0.143)
Female	-0.643	(0.164)	-0.516	(0.140)	-0.514	(0.139)
Duke's stage C	0.391	(0.190)	0.354	(0.161)	0.365	(0.159)
Duke's stage D	1.57	(0.225)	1.250	(0.188)	1.31	(0.188)
log(γ)	0.284	(0.147)	-0.343	(0.205)	-0.413	(0.216)
BIC	6706.3		6572.8		6567.3	

Note: the results for the NHPP, HNHPP, and HRP in Tables 5–1 and 5–2 differ from those published in Tables III and IV of Pietzner and Wienke's paper. We had performed all our optimization with R whereas they had used the NLMIXED package in SAS. We wrote to Dr. Pietzner about a discrepancy between our results and theirs and she then accepted our request for the some of the SAS code they used. They probably mis-reported the fitted values and AIC for the NHPP, HNHPP, and HRP because, after running their code, we arrived at the results shown here.

The best model out of the above, according to BIC, is the HTRP with log-Normal frailty, though it gives very similar parameter estimates to those in the HTRP with gamma frailty. The estimated trend and renewal parameters, $\hat{b}$ and $\hat{c}$ describe the behaviour of the intensity function associated with a "typical" (baseline) subject. In the framework of this model, there is evidence of heterogeneity in the patients' intensity functions around the baseline, as indicated by the estimate of $\hat{\gamma} = 0.608$ (95% CI [0.400, 0.925]).

Conclusions about the covariate effects arrived at by Pietzner and Wienke align with the conclusions of Gonzalez et al. (2005): that women have a lower risk of readmission to the hospital after treatment than men, that chemotherapy had no significant effect, and that patients at Duke's stage D have significantly higher risk compared with those in the reference group combining Duke's stage A and B.

As a demonstration we also fit the Weibull-power-law model with covariate effects modelled on the renewal parameter to this dataset. More specifically, we fit the Weibull-power-law HTRP parametrized by parameters a, b, c_0,γ, and the vector $\beta_c = (\beta_{chemo}, \beta_{female}, \beta_{DukeC}, \beta_{DukeD})^\top$, with intensity function:

$$\lambda(t|\mathcal{H}_i(t), \mathbf{X}_i, u_i) = u_i\, a_i^c b\, t^{b-1} c_i (t^b - T^b_{N(t^-)})^{c_i - 1}$$

where $c_i = \exp(c_0 + \mathbf{X}_i^\top \beta_c)$ is the renewal parameter for the subgroup with covariates $\mathbf{X}_i$ and $u_i \sim \exp(Y_i)$ where $Y_i \sim \mathcal{N}(0, \sigma^2)$. Table 5–4 displays the results for this random effect model as well as the results for the corresponding TRP without the random effect. We can see that there is no significant difference in the renewal parameters between the two levels of any covariate except for a barely significant effect of being female as opposed to male. The BIC values are also poor; this framework for modelling covariates clearly provides an ill fit to this dataset.

Note that the popular Cox-based models such as Andersen-Gill (AG), Wei-Lin-Weissfeld (WLW), and Prentice-Williams-Peterson (PWP) are used when estimating the effects of covariates is of interest and interpretation of the baseline intensity for the typical subject is not. With regard to covariate effects, since the model with best fit according to BIC out of the ones fitted above is still the Weibull-power-law with log-Normal random effects fitted by Pietzner and Wienke, we direct the reader to Pietzner and Wienke (2013) for a comparison of the estimated covariate effects in this parametric TRP framework with those reached by fitting the existing popular recurrent events

Table 5–4: Estimated parameter values (standard error) for Weibull-power-law models with covariate effects on the renewal parameter and either no unobserved heterogeneity (left column) or log-Normal distributed unobserved heterogeneity (right column) fitted to the hospital readmissions dataset

Parameter	TRP		HTRP	
log(a)	-9.392	(0.734)	-8.846	(0.191)
log(b)	0.278	(0.076)	0.155	(0.030)
β_{c0}	-0.867	(0.103)	-0.516	(0.086)
β_{c1}	0.077	(0.086)	-0.018	(0.077)
β_{c2}	0.262	(0.083)	0.191	(0.076)
β_{c3}	-0.069	(0.101)	-0.110	(0.088)
β_{c4}	-0.191	(0.107)	-0.097	(0.091)
$\log(\gamma)$	–	–	-0.101	(0.163)
BIC	6741.4		6615.1	

models such as Andersen-Gill (AG), Wei-Lin-Weissfeld (WLW), and Prentice-Williams-Peterson (PWP), and frailty models.

We explained in previous Chapter 3 how to interpret the baseline intensity parameters in a parametric TRP model to describe the behaviour of the typical subject in a simple, meaningful manner. Interpreting the trend and renewal parameters in a parametric TRP has not been done before, so we proceed by demonstrating the interpretation of the baseline intensity in the Weibull-power-law TRP with log-Normal random effects for this data set. Referring to the third column of Table 5–2: an estimate of c below 1 means that the intensity is decreasing until each hospital readmission, and the occurrence of an event then causes what looks like a “spike” in risk of readmission immediately afterwards, as the intensity renews. Over time, until the next readmission, the risk decreases to look more similar to what it would have been if the most recent readmission hadn't happened. Estimates of b above 1 and c below 1 mean that a patient's intensity function going forward from the jth readmission is heightened relative to how it was going forward from the $(j-1)$th readmission. According to $\hat{b} = 1.252$ and $\hat{c} < 1$ everyone strictly renews "worse than perfectly".

Figure 5–1 plots the intensity ratio (equation (3.2)) given the fitted $\hat{b}$ and $\hat{c}$. It shows the relative intensity, of the typical patient exposed in the interval $[0, t]$ with a first readmission at time $s < t$ to the intensity of a typical patient exposed in the interval $[0, t]$ with no readmission, as a function of x/s, the time elapsed since s scaled by s. For instance, one can tell from the graph that, if the typical patient's risk follows a Weibull-power-law TRP with trend parameter 1.252 and renewal parameter 0.575, then it would take another $0.475\,s$ amount of time after the patient's readmission at time s before his risk of readmission is 150% of what it would be if the first readmission had not occurred. If he remains event free for longer, say until $1s$ time after his first readmission, his risk will further decrease to 126% of what it would be if he were still event-free.

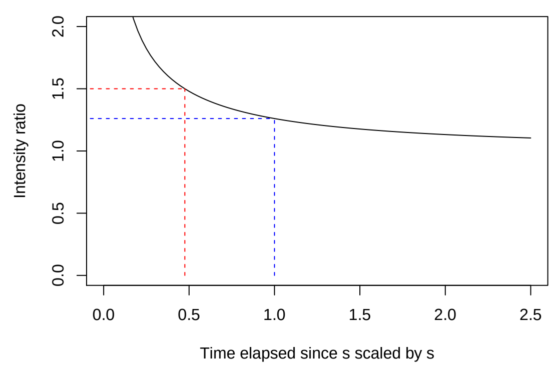

Figure 5–1: The plot of $r_{1,0}(x)$ for the baseline process in the Weibull-power-law HTRP fitted to the hospital readmission data, where $\hat{b} = 1.252$, $\hat{c} = 0.575$. The blue and red lines respectively indicate where the the relative intensity of having had an event at s to having had no event equals 1.5 and 1.26.

Figure 5–2 plots the intensity ratio (equation (3.5)), the relative intensity of going forward from the most recent event at time s versus going forward from process initiation for a typical subject. It illustrates how close to completely renewed a subject is going forward from an event, when compared with going forward from the first admission. For example, one can see from the graph that, for a subject with a readmission at time s, once $0.1s$ time has passed since that readmission his risk is 29% higher than what it was at time $0.1s$ from the first admission. At time $0.5s$ after the readmission at time s, his risk is only 9% higher than what it was at time $0.5s$ from the first admission.

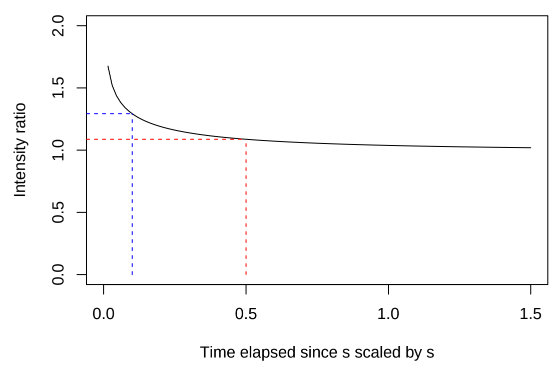

Figure 5–2: The plot of $\mathfrak{r}_{1,0}(x)$ for the baseline process in the Weibull-power-law HTRP fitted to the hospital readmission data, where $\hat{b} = 1.252$, $\hat{c} = 0.575$. The blue and red lines respectively indicate where the relative intensity equals 1.29 and 1.09.

It is worth noting that, for a dataset of this nature with low event counts for almost all subjects and high proportion of censoring, it might not make much sense to discuss interpretations of the fitted model in such precise terms. However, for another dataset with less censoring and research objectives specifically about relating past event history and future risk, the procedure to arrive at such quantitative conclusions as the above would be the same.

CHAPTER 6
Conclusion and future work

This thesis provides an in-depth introduction to the trend-renewal process for modelling recurrent time-to-event data, highlighting some attractive properties such as its mathematically elegant manner of combining the cumulative time scale and the gap time scale in the intensity function and the ease of parametric inference. One of its objectives is to make the TRP more accessible as a tool of analysis in sports injury and medical analyses.

The main contributions are new notions of relative intensity to more effectively utilize the popular Weibull-power-law choice of TRP, as well as a caution and a suggestion about potential violation of the classical Cox proportional intensities formulation of covariate effects. One avenue for future work could be to recompose these ideas in other parametric frameworks or when inference is performed semiparametrically or nonparametrically.

The only simulation studies we found in the literature evaluating finite sample inference for the TRP framework involved the Weibull-power-law TRP with log-Normal distributed frailty (Pietzner and Wienke (2013)). We conducted and reported the results of analogous simulation studies for the Weibull-power-law TRP with gamma distributed frailty and the gamma-power-law TRP with gamma distributed frailty.

In the way of statistical inference, the literature is well developed. This thesis did not discuss nonparametric inference for the TRP at all, though fully nonparametric methods of inference for multiple event processes is certainly a useful direction to pursue, as the only paper on fully nonparametric inference for the TRP (Gámiz and Lindqvist (2016)) was published very recently and proposes a method applicable only when a single event process is being observed. Along a different vein, there has been little to no work done on Bayesian inference for the TRP.

For purposes of improving interpretability in practical settings, one might build a causal model from the TRP framework. Finally, future work to extend the general applicability of the TRP might pursue the modelling of time-varying covariates or the modelling of competing risks where multiple types of events are of interest.

References

O. O. Aalen and E. Husebye. Statistical analysis of repeated events forming renewal processes. *Statistics in Medicine*, 10(8):1227–1240, 1991.

P. K. Andersen, Ø. Borgan, R. D. Gill, and N. Keiding. *Statistical models based on counting processes*. Springer-Verlag New York, New York, 1993.

M. Berman. Inhomogeneous and modulated gamma processes. *Biometrika*, 68(1):143–152, 1981.

H. Chernoff. On the distribution of the likelihood ratio. *The Annals of Mathematical Statistics*, 25(3):573–578, 1954.

P. Congdon. Modelling frailty in area mortality. *Statistics in Medicine*, 14 (17):1859–1874, 1995.

R. J. Cook and J. Lawless. *The statistical analysis of recurrent events*. Springer Science & Business Media, New York, 2007.

D. Cox. The statistical analysis of dependencies in point processes. In P. A. W. Lewis, editor, *Stochastic Point Processes: Statistical Analysis, Theory, and Applications*, pages 55–66. John Wiley & Sons, New York, 1972.

D. A. Follmann and M. S. Goldberg. Distinguishing heterogeneity from decreasing hazard rates. *Technometrics*, 30(4):389–396, 1988.

J. Franz, A. Jokiel-Rokita, and R. Magiera. Prediction in trend-renewal processes for repairable systems. *Statistics and Computing*, 24(4):633–649, 2014.

M. L. Gámiz and B. H. Lindqvist. Nonparametric estimation in trend-renewal processes. *Reliability Engineering & System Safety*, 145:38–46, 2016.

J. R. Gonzalez, E. Fernandez, V. Moreno, J. Ribes, M. Peris, M. Navarro, M. Cambray, and J. M. BorrÃăs. Sex differences in hospital readmission among colorectal cancer patients. *Journal of Epidemiology and Community Health*, 59(6):506–511, 2005.

K. Heggland and B. H. Lindqvist. A non-parametric monotone maximum likelihood estimator of time trend for repairable system data. *Reliability Engineering & System Safety*, 92(5):575–584, 2007.

M. Hlavac. *stargazer: Well-Formatted Regression and Summary Statistics Tables*. Harvard University, Cambridge, USA, 2015. URL `http://CRAN.R-project.org/package=stargazer`. R package version 5.2.

A. Jokiel-Rokita and R. Magiera. Estimation of parameters for trend-renewal processes. *Statistics and Computing*, 22(2):625–637, 2012.

D. Y. Lin, L. J. Wei, and Z. Ying. Accelerated failure time models for counting processes. *Biometrika*, 85(3):605–618, 1998.

B. Lindqvist, G. Elvebakk, and K. Heggland. The trend-renewal process for statistical analysis of repairable systems. *Technometrics*, 45(1):31–44, 2003.

B. H. Lindqvist. On the statistical modeling and analysis of repairable systems. *Statistical Science*, 21(4):532–551, 2006.

D. Pietzner and A. Wienke. The trend-renewal process: a useful model for medical recurrence data. *Statistics in Medicine*, 32(1):142–152, 2013.

R. L. Prentice and J. D. Kalbfleisch. *The statistical analysis of failure time data*. John Wiley & Sons, New York, 1980.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2015. URL `https://www.R-project.org/`.

N. Reid. A conversation with Sir David Cox. *Statistical Science*, 9(3):439–455, 1994.

R. L. Strawderman. The accelerated gap times model. *Biometrika*, 92(3): 647–666, 2005.

A. Tsiatis. A nonidentifiability aspect of the problem of competing risks. *Proceedings of the National Academy of Sciences*, 72(1):20–22, 1975.

Q. Yang, Y. Hong, Y. Chen, and J. Shi. Failure profile analysis of complex repairable systems with multiple failure modes. *IEEE Transactions on Reliability*, 61(1):180–191, 2012.