Title Page w/ ALL Author Contact Info.

Lifetime Data Analysis manuscript No.
(will be inserted by the editor)

Modeling of semi-competing risks by means of first passage times of a stochastic process

Beate Sildnes · Bo H. Lindqvist

Received: date / Accepted: date

Abstract In semi-competing risks one considers a terminal event, such as death of a person, and a non-terminal event, such as disease recurrence. We present a model where the time to the terminal event is the first passage time to a fixed level c in a stochastic process, while the time to the non-terminal event is represented by the first passage time of the same process to a stochastic threshold S, assumed to be independent of the stochastic process. In order to be explicit, we let the stochastic process be a gamma process, but other processes with independent increments may alternatively be used. For semi-competing risks this appears to be a new modeling approach, being an alternative to traditional approaches based on illness-death models and copula models. In this paper we consider a fully parametric approach. The likelihood function is derived and statistical inference in the model is illustrated on both simulated and real data.

Keywords competing risks · first passage time · gamma process · random signs censoring

Beate Sildnes
Department of Mathematical Sciences
Norwegian University of Science and Technology
N-7491 Trondheim, Norway
E-mail: beate.sildnes@gmail.com
Present address:
BearingPoint
Tjuvholmen allé 3
N-0252 Oslo, Norway

Bo H. Lindqvist
Department of Mathematical Sciences
Norwegian University of Science and Technology
N-7491 Trondheim, Norway
E-mail: bo@math.ntnu.no

Lifetime Data Analysis manuscript No.
(will be inserted by the editor)

Modeling of semi-competing risks by means of first passage times of a stochastic process

Received: date / Accepted: date

Abstract In semi-competing risks one considers a terminal event, such as death of a person, and a non-terminal event, such as disease recurrence. We present a model where the time to the terminal event is the first passage time to a fixed level c in a stochastic process, while the time to the non-terminal event is represented by the first passage time of the same process to a stochastic threshold S, assumed to be independent of the stochastic process. In order to be explicit, we let the stochastic process be a gamma process, but other processes with independent increments may alternatively be used. For semi-competing risks this appears to be a new modeling approach, being an alternative to traditional approaches based on illness-death models and copula models. In this paper we consider a fully parametric approach. The likelihood function is derived and statistical inference in the model is illustrated on both simulated and real data.

Keywords competing risks · first passage time · gamma process · random signs censoring

1 Introduction

Semi-competing risks is a variation of ordinary competing risks, apparently first introduced and named in 2001 by Fine et al (2001). In semi-competing risks, one considers two types of events: non-terminal and terminal. The difference from ordinary competing risks is that the focus is not restricted to the first event that occurs in time.While a non-terminal event may be censored by a terminal event, the non-terminal event does not prevent the occurrence of the terminal event, as it would do in ordinary competing risks problems. Thereby, more information regarding event times is obtained with semi-competing risks than with ordinary competing risks.

Semi-competing risks are of particular interest in medical applications. Here, a non-terminal event may for instance be disease recurrence, while the terminal event typically is death, see, e.g., Fine et al (2001), where the frequently studied bone marrow transplant data are analyzed. Varadhan et al (2014) give an illlustration from gerontology. Here the event 'death' will censor other events under study, such as dementia (Alzheimer's disease) or disability.

While competing risks theory is developed in fairly great detail, and has been applied in a wide range of situations, semi-competing risks studies have not yet become that common in the literature. Still, several semi-parametric models have been developed and successfully applied to semi-competing risks data, for instance in Fine et al (2001), Peng and Fine (2006) and Hsieh et al (2008). Briefly, the models assume that the joint distribution of the time to the non-terminal and terminal event times, in the present paper denoted Z and X, respectively, is given by a known copula, for example the gamma frailty copula (Clayton, 1978). The joint distribution of (X, Z) is identifiable in the upper wedge where $Z < X$ only, and the marginal distribution of the non-terminal event Z is hence not identifiable without additional assumptions. The corresponding approaches are based on latent variables, with results of interest expressed by so called *net* quantities. Xu et al (2010) argued that such models should be avoided, and they instead presented an approach based on *crude* quantities (i.e., *observable*) only, considering a class of illness-death models with shared frailty.

The semi-competing risks problem is essentially equivalent to the classical illness-death model (Fix and Neyman, 1951), which in turn is considered to be a special case of multi-state models. There is hence a considerable literature on the subject, with a theoretical foundation given in, e.g., Andersen et al (1993). A more recent review is found in Putter et al (2007). Typical assumptions are that the transitions between states are of (semi-)Markov type, whereas Meira-Machado et al (2006) considered estimation in a non-Markov illness-death model. The already cited paper by Xu et al (2010) considers a Markov model with a shared gamma frailty, and demonstrates the connection to the approach of Fine et al (2001) who use the Clayton copula.

In survival analysis, a lifetime may in many cases conveniently be modeled as a first passage time of a boundary or threshold state of a stochastic process, either observable or latent. This kind of modeling seems to become more and more popular. Aalen and Gjessing (2001) gave a review of such models, considering in particular the Wiener process as the underlying stochastic processes. While the Wiener process is often seen in medical applications, the gamma process seems to be more popular in engineering applications, see, e.g., the comprehensive review by van Noortwijk (2009).

Lee and Whitmore (2006) give a thorough review of various types of threshold models for survival analysis. Typically, an endpoint of interest occurs when a process reaches an adverse threshold state for the first time. Heterogeneous behavior among units, either measured via covariates or modeled via latent variables, may then be modeled by differentitating the speed of the process or the starting state (Aalen and Gjessing, 2001). Lee and Whitmore (2006)

show in particular how latent threshold models are used in studies involving competing risks. As an example they let the underlying stochastic process be a multidimensional Wiener process with dimension equal to the number of competing causes, and with specific threshold levels for each dimension. Dependencies between the different causes may then be modeled by considering correlated Wiener processes, and removal of causes may be considered as well. In an earlier paper, Whitmore (1986) modeled independent competing risks by considering first passage times of multi-dimensional Wiener processes.

While the above idea of using first passage times of correlated stochastic processes can be adopted to the semi-competing risks case, we will in the present paper assume a single univariate underlying stochastic process, $D(t)$, say. The main idea is then to let both the terminal event and the non-terminal event correspond to hitting times at different threshold levels of this single process. More precisely, we shall let the time to the terminal event, X, be defined as the hitting time of $D(t)$ to a fixed level c, while we let the time of the non-terminal event, Z, correspond to the hitting time of a random threshold S. Here, S is assumed to be independent of the process $D(t)$, which is sensible since it in some sense plays a similar role to the constant c. The dependency between the non-terminal and terminal events is now obtained because of their relations to the common underlying process $D(t)$.

An advantage of this way of modeling semi-competing risks is that only one underlying process is involved. Apparently this leads to a more parsimonious and simply structured model than the approach using several correlated processes, and may also imply an easier interpretation and improved insight. In particular the marginal distribution of Z, which in many applications is the interesting distribution (e.g., Fine et al (2001)), can now be identified as the hitting time distribution of level S of the process $D(t)$.

Similar models were considered by Lindqvist and Skogsrud (2008) in a system maintenance application, and Horrocks and Thompson (2004) who considered an application to health status for hospitalized patients. In the former paper, the underlying process is a Wiener process modeling the state of a degrading system, where the terminal event is a critical failure corresponding to a certain threshold of the process. In order to possibly avoid critical failures, a maintenance policy is introduced, where the system is repaired when the degradation reaches a certain state S. The latter paper considers a Wiener process model for health status for hospitalized patients, with two barriers, corresponding to discharge and death, which here correspond to two terminal events. The authors then suggested the extension where a decision to transfer the patient is considered when the process reaches a certain intermediate level. In our terminology this would correspond to the threshold S.

In threshold modeling for medical applications, the latent underlying process $D(t)$ is commonly thought of as the health condition of the person under study (Lee and Whitmore (2006)), with the endpoint of interest corresponding to the crossing of a given threshold c. Suppose then, as in semi-competing risks, that another event is of interest which may or may not occur before the terminal event. The proposed model makes the coupling of the two events

through introducing another threshold S of the process $D(t)$. The crucial assumption is hence that both events are determined by crossings of the same underlying process $D(t)$.

One might argue that in many applications it will not be clear how to think of a common underlying process governing both the terminal and non-terminal event. This may for example be the case in the above mentioned case of semi-competing risks in connection with Alzheimer's disease, where one may think of different processes that govern the onset of Alzheimer's disease and death, respectively. Furthermore, if there is one process, this process may change its behavior at the occurrence of Alzheimer's disease and hence again violate the model. Still it is believed that a relatively simple model like the proposed model may give useful insight also in such cases, although the interpretation of the latent underlying process may not be straightforward.

The present paper is to the best of our knowledge the first to use the idea of first passage times to study semi-competing risks problems. To be concrete, we have chosen to consider gamma processes as the underlying stochastic process. The most crucial property of the stochastic process in our approach appears to be the independent increment property. Thus the modification needed to replace the gamma process by, e.g., a Wiener process, appears to be essentially straightforward. Although we will not study the case of observed covariates, one may imagine modeling of the influence of covariates both on the underlying process and on the distribution of S. For easy reference we shall denote the model by *the gamma threshold model*, or simply *the threshold model*.

The rest of the paper is organized as follows. In Section 2 we first review the basic notation of competing risks, and then give an introduction to the gamma process and its first passage time distribution. The gamma threshold model and the corresponding likelihood function are presented in Section 3, while a simulated data set is analyzed in Section 4. In Section 5 we use the threshold model to analyze the bone marrow transplant data set. Some final remarks on the threshold model are given in Section 6, while we finally consider some variations and possible extensions of the model in Section 7. The article is ended by an Appendix presenting some technical derivations and formulae.

2 Preliminaries

2.1 Two competing risks

2.1.1 Notation and definitions.

We start by introducing notation and basic definitions from ordinary competing risks with two latent event times, X and Z. Later, X will be the time of the terminal event, while Z is the time of the non-terminal event. Throughout the paper we assume that (X, Z) is a pair of continuously distributed positive random variables such that $P(X = Z) = 0$. In the ordinary competing

risks situation we observe (only) the pair (T, D) where $T = \min(X, Z)$ and $D = I(X < Z)$, where $I(A)$ is the indicator function of the event A.

We let $F_X(t) = P(X \leq t)$ and $F_Z(t) = P(Z \leq t)$ be the cumulative distribution functions of X and Z, respectively. The subdistribution functions (also called cumulative incidence functions in the competing risks literature) of X and Z are defined as, respectively, $F_X^*(t) = P(X \leq t, X < Z)$ and $F_Z^*(t) = P(Z \leq t, Z < X)$. Similarly, the subsurvival functions are $K_X^*(t) = P(X > t, X < Z)$ and $K_Z^*(t) = P(Z > t, Z < X)$, while the subdensity functions are $f_X^*(t) = F_X^{*'}(t) = -K_X^{*'}(t)$ and similarly for $f_Z^*(t)$. It is clear that $K_T(t) = P(T > t) = K_X^*(t) + K_Z^*(t)$. From this we can define the so-called cause-specific hazard rates $\lambda_X^*(t) = f_X^*(t)/K_T(t)$ and $\lambda_Z^*(t) = f_Z^*(t)/K_T(t)$.

The functions F_X^* and F_Z^* are nondecreasing with $F_X^*(0) = 0$ and $F_Z^*(0) = 0$. Moreover, $F_X^*(\infty) + F_Z^*(\infty) = 1$. We will also use the notion of conditional sub-distribution functions, defined by $\tilde{F}_X(t) = P(X \leq t | X < Z) = F_X^*(t)/F_X^*(\infty)$ and $\tilde{F}_Z(t) = P(Z \leq t | Z < X) = F_Z^*(t)/F_Z^*(\infty)$.

2.1.2 Random signs censoring.

Random signs censoring was first introduced by Cooke (1993) as a notion of age-dependent censoring. Considering X as the time of failure and Z as the censoring time, the idea is that, whether a unit is censored or not, is independent of the age of the unit. A precise definition can be given as follows (Lindqvist et al, 2006): Let (X, Z) be a pair of positive random variables. Then Z is called a *random signs censoring* of X if the event $\{X < Z\}$ is stochastically independent of X.

This means that a unit is censored independently of the time X where it would have failed. As an interpretation, we can imagine that at some time before the unit fails, it will emit a signal indicating that a failure is emerging. The crucial assumption is that the event that the emitted signal is discovered, and hence censoring is executed, does not depend on the unit's age. The assumption of random signs censoring may in many cases be unreasonable. However, for certain phenomena it gives a good description of reality. For instance, if the unit in question is a machine, then typical signals may be excessive noise and/or vibration. For a human being in a medical study, the signal may be symptoms of disease.

Random signs censoring implies that the marginal distribution of X is identifiable from ordinary competing risks data (Cooke, 1993). In fact, the definition of random signs censoring leads to the following conditional subdistribution function for X,

$$\tilde{F}_X(t) = P(X \leq t | X < Z) = P(X \leq t) = F_X(t). \tag{1}$$

Thus, the marginal distribution of X is actually the same as the distribution of the observed failure times X, discarding all the observations where Z is observed. The distribution of Z is, however, in general not identifiable under random signs censoring.

Cooke (1993) showed that a *necessary* condition for a joint distribution of (X, Z) to satisfy the random signs censoring conditon, is that

$$\tilde{F}_Z(t) > \tilde{F}_X(t) \text{ for all } t > 0. \tag{2}$$

In our data application in Section 5 we will check this condition by plotting the corresponding estimated functions.

The model considered in the present paper has the random signs censoring property. Until now, the random signs censoring has been considered mainly in applications to engineering reliability, e.g., Cooke and Bedford (2002); Lindqvist et al (2006); Christen et al (2011); Lindqvist and Skogsrud (2008).

2.2 The gamma process and its first passage distribution

2.2.1 The gamma process

A continuous time stochastic process $\{D(t), t \geq 0\}$ is a gamma process with shape function $v(t) > 0$ and scale parameter $u > 0$ if

1. $D(0) = 0$ with probability 1,
2. $\{D(t), t \geq 0\}$ has independent increments,
3. $D(t) - D(s)$ is gamma distributed with shape parameter $v(t) - v(s)$ and scale parameter u for every $0 < s < t$.

It follows that the probability density function of $D(t)$ is the gamma density

$$f_{D(t)}(x) = Ga(x; v(t), u); \;\; x > 0, \tag{3}$$

where $Ga(x; v, u) = u^v(\Gamma(v))^{-1}x^{v-1}\exp(-ux)$.

In empirical studies involving the gamma process, particularly in engineering applications, it has been seen that the expected deterioration, $E(D(t)) = v(t)/u$ often follows a power function in t (van Noortwijk, 2009). This suggests the form $v(t) = \alpha t^\beta$, for constants $\alpha > 0$ and $\beta > 0$, which will be used later in this paper. The gamma process is called stationary if the expected value is linear, i.e., $\beta = 1$ and non-stationary if $\beta \neq 1$.

2.2.2 The first passage time distribution

Let $D(t)$ be a gamma process with shape function $v(t)$ and scale parameter u. The first passage time of the process $D(t)$ over a fixed threshold $d > 0$, denoted T_d, is defined as the time until the process crosses the level d, i.e., $T_d = \inf\{t : D(t) \geq d\}$. The cumulative distribution function of T_d is found from

$$P(T_d \leq t) = P(D(t) > d) = \frac{\Gamma(v(t), d \cdot u)}{\Gamma(v(t))} \tag{4}$$

where $\Gamma(a, x)$ is the upper incomplete gamma function defined by $\Gamma(a, x) = \int_x^\infty z^{a-1}e^{-z}dz$.

Since the scale parameter u only appears together with the threshold level d as the product $d \cdot u$ in (4), it follows that we may, without loss of generality, let $u = 1$. This will be done in the following, where we define the survival function and cumulative distribution function of T_d, respectively, by

$$K(t; v(t), d) = 1 - F(t; v(t), d) = 1 - \frac{\Gamma(v(t), d)}{\Gamma(v(t))}. \tag{5}$$

If $v(t)$ is differentiable, then the probability density function of T_d is (Park and Padgett, 2005; Paroissin and Salami, 2014)

$$f(t; v(t), d) = v'(t) \left[\Psi(v(t)) - \log d\right] \left(1 - \frac{\Gamma(v(t), d)}{\Gamma(v(t))}\right) + \frac{v'(t)}{v(t)^2 \Gamma(v(t))} d^{v(t)} {}_2F_2(v(t), v(t); v(t)+1, v(t)+1; -d).$$

Here, $\Psi(a) = \frac{d}{da} \ln \Gamma(a) = \frac{\Gamma'(a)}{\Gamma(a)}$ is the digamma function, and ${}_2F_2()$ is the generalized hypergeometric function of order (2,2). Recall that he generalized hypergeometric function of order (p, q) is defined as

$${}_pF_q(a_1, ..., a_p; b_1, ..., b_q; z) = \sum_{k=0}^{\infty} \frac{(a_1)_k \cdots (a_p)_k}{(b_1)_k \cdots (b_q)_k} \frac{z^k}{k!}$$

where $(x)_n = \frac{\Gamma(x+n)}{\Gamma(x)}$ is the Pochammer symbol.

We will also need the joint density of (T_d, T_c) when $d < c$, defined as

$$f(t_1, t_2; v(t_1), v(t_2), d, c) dt_1 dt_2 = P(t_1 \leq T_d \leq t_1 + dt_1, t_2 \leq T_c \leq t_2 + dt_2).$$

Since the gamma process is a pure jump process, the calculation of this density has to take into account the fact that the threshold d is crossed by a jump, so that $D(T_d)$ is almost surely strictly above d (Kahle et al (2016)). The derivation of this density is deferred to the Appendix (Section 8.1).

3 The gamma threshold model

3.1 The ordinary competing risks setting

Consider a gamma process $D(t)$ with shape function $v(t)$ as considered above. Let X be the first passage time to a fixed level $c > 0$, i.e., $X = T_c$. Let further S be a positive random variable which is stochastically independent of the process $D(t)$, and define $Z = T_S$. Assume also that $P(S = c) = 0$.

It follows that if X, Z are the latent variables of an ordinary competing risks situation, we observe Z (and not X) if and only if $S < c$. An illustration is given in Figure 1. Here $X = T_c$, and we observe $Z = T_{s_1}$ in case $S = s_1 < c$, while we observe X if $S = s_2 > c$.

The assumption that S is independent of the gamma process $D(t)$ implies that (X, Z) satisfies the requirement of a random signs censoring (see Section

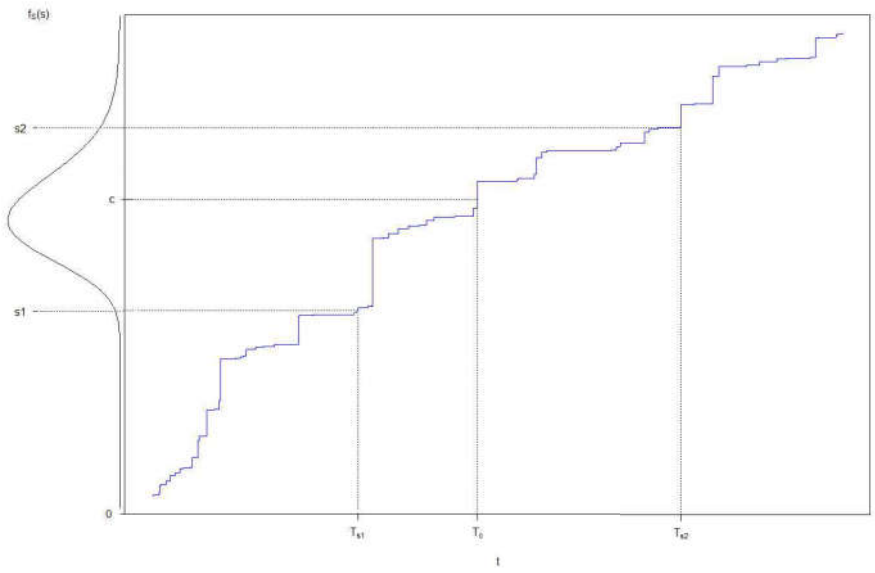

Fig. 1 Illustration of the case of a gamma process $D(t)$ with a fixed level c and a lognormally distributed S

2). In fact, the event $\{X < Z\}$ now is equivalent to $S > c$, which is independent of $X = T_c$.

The model can be viewed as an extension of the single-threshold problem considered by Paroissin and Salami (2014), who in our notation considered $Z = T_S$.

3.2 The semi-competing risks setting

The competing risks model considered in the previous subsection is easily extended to the semi-competing risks case. We then assume that if $S < c$, then after Z is observed, the gamma process $D(t)$ continues until level c has been crossed, i.e., until X is observed. On the other hand, if $S > c$, then the process is stopped when level c is crossed, so Z will not be observed. This defines a semi-competing risks situation with terminal event at time X and non-terminal event at time Z, which as indicated in the Introduction will be called the *the gamma threshold model* or *the threshold model.*

By the construction, it is clear that the random variables X and Z are dependent random variables. Moreover, since Z is a random signs censoring of X, as shown in the previous subsection, the relation (2) holds. In the data examples, we will start the analyses by plotting the corresponding estimated functions as given by (15) in order to empirically verify the possibility of an underlying threshold model.

In practice one is often interested in the marginal distributions of X and Z. While the distribution of X is identifiable in semi-competing risks, this is not the case of Z. Key quantities related to Z are hence what we in the introduction have called net quantities, and their relevance are therefore strongly connected to the assumptions we have made. Since $X = T_c$, the distribution of X has been completely described in the previous section. For Z, on the other hand, we have

$$\begin{aligned} K_Z(t) &= P(T_S > t) \\ &= \int_0^\infty P(T_s > t) f_S(s) ds \\ &= \int_0^\infty K(t; v(t), s) f_S(s) ds \end{aligned} \tag{6}$$

where $f_S(s)$ is the density of S. The marginal hazard rate for Z is from this

$$\lambda_Z(t) = \frac{\int_0^\infty f(t; v(t), s) f_S(s) ds}{\int_0^\infty K(t; v(t), s) f_S(s) ds}. \tag{7}$$

Thereby, parametric estimates of $K_Z(t)$ and $\lambda_Z(t)$ can be found by inserting parameter estimates into the expressions in (6) and (7), respectively.

We will also later consider the corresponding crude quantities under the assumed model assumptions.

$$\begin{aligned} F_Z^*(t) &= P(Z \leq t, Z < X) \\ &= P(T_S \leq t, S < c) \\ &= \int_0^c P(T_s \leq t) f_S(s) ds \\ &= \int_0^c F(t; v(t), s) f_S(s) ds \end{aligned} \tag{8}$$

The cause-specific hazard rate for Z can from this be expressed as

$$\lambda_Z^*(t) = \frac{\int_0^c f(t; v(t), s) f_S(s) ds}{K(t)}$$

where $K(t) = F_X^*(t) + F_Z^*(t)$. Here the latter term in the sum is given in (8), while the former is $P(T_c > t, S > c) = K(t; v(t), c)(1 - F_S(c))$.

3.3 Representation of data from semi-competing risks

Consider the semi-competing risks situation with (X, Z) given as above, and let $\tau > 0$ be a possible censoring time. In the present paper we shall assume that censoring is random, and that τ is independent of (X, Z).

For a unit under study we observe the following variables, $Y_1 = \min\{Z, X, \tau\}$, $Y_2 = \min\{X, \tau\}, \delta_1 = I\{Z \leq Y_2\}$ and $\delta_2 = I\{X \leq \tau\}$. With n units under observation we hence have the i.i.d. data $(Y_{1i}, Y_{2i}, \delta_{1i}, \delta_{2i})$ for $i = 1, 2, \ldots, n$.

In order to write down the likelihood function we find it convenient to consider the six possible orderings of (Z, X, τ), as given in Table 1. It follows that the essential observations will be of four different types, in the table named Case 1 to Case 4.

Table 1 Possible orderings of variables in semi-competing risks data. X = time of terminal event, Z = time of non-terminal event, τ = censoring time.

Order	$(Y_1, Y_2, \delta_1, \delta_2)$	Case
Z, X, τ	(Z, X, 1, 1)	1
X, Z , τ	(X, X, 0, 1)	2
X, τ, Z	(X, X, 0, 1)	2
Z, τ, X	(Z, τ, 1, 0)	3
τ, Z, X	(τ, τ, 0, 0)	4
τ, X, Z	(τ, τ, 0, 0)	4

The likelihood function in the case of illness-death models for semicompeting risks have been derived, e.g., by Putter et al (2007) and Xu et al (2010). As will become clear in the next subsection, the special structure of the threshold model requires a somewhat different way of deriving the likelihood.

3.4 The likelihood function

The contributions to the likelihood function corresponding to each case are considered separately below. Since we have random censoring, with censoring times being independent of the process $X(t)$ and the variable S, we may by conditioning on the censoring times assume that they are given by constants.

In the derivations below we use notation and results for the gamma process from Section 2.2.

Case 1: Observe both Z and X.

The data are $Z = T_S = t_1$, $X = T_c = t_2$ with $t_1 < t_2$. Thus $S < c$ is a consequence, and we have

$$\begin{aligned} & P\left(t_1 \leq T_S \leq t_1 + dt_1, t_2 \leq T_c \leq t_2 + dt_2\right) \\ &= \int_0^c P(t_1 \leq T_s \leq t_1 + dt_1, t_2 \leq T_c \leq t_2 + dt_2) f_S(s) ds \\ &= [\int_0^c f(t_1, t_2; v(t_1), v(t_2), s, c) f_S(s) ds] dt_1 dt_2 \end{aligned} \tag{9}$$

The expression in square brackets is now the contribution to the likelihood function.

Case 2: Observe X only.

Suppose we have observed $X = T_c = t_2$, when we know that $S > c$. Since S is independent of the process, we have

$$\begin{aligned} & P\left(t_2 \leq T_c \leq t_2 + dt_2, S > c\right) \\ &= P(t_2 \leq T_c \leq t_2 + dt_2) P(S > c) \\ &= [f(t_2; v(t_2), c)(1 - F_S(c))] dt_2 \end{aligned}$$

where the expression in square brackets defines the contribution to the likelihood function for this case.

Case 3: Observe Z and a censoring time τ.

In this case we observe $Z = t_1$ and a censoring time τ with $t_1 < \tau$. Furthermore, we have necessarily $S < c$. Thus

$$\begin{aligned} P(t_1 \leq T_S \leq t_1 + dt_1, T_c > \tau) &= \int_0^c P(t_1 \leq T_s \leq t_1 + dt_1, T_c > \tau) f_S(s) ds \\ &= \int_0^c P(t_1 \leq T_s \leq t_1 + dt_1) P(T_c > \tau | T_s = t_1) f_S(s) ds \\ &= [\int_0^c f(t_1, v(t_1), s) K(\tau - t_1; v(\tau) - v(t_1), c - s) f_S(s) ds] dt_1 \end{aligned}$$

which defines the contribution to the likelihood function.

Case 4: Observe a censoring time τ only.

The contribution to the likelihood function is clearly

$$\begin{aligned} P(T_S > \tau, T_c > \tau) &= P(T_c > \tau, S > c) + P(T_S > \tau, S < c) \\ &= P(T_c > \tau)(1 - F_S(c)) + \int_0^c P(T_s > \tau) f_S(s) ds \\ &= K(\tau; v(\tau), c)(1 - F_S(c)) + \int_0^c K(\tau; v(\tau), s) f_S(s) ds. \end{aligned}$$

The complete likelihood for the data is now the product of the contributions from each of the n observations, as given above.

3.5 Identifiability of the threshold model

In the following we will parametrize $v(t)$ by introducing parameters $\alpha > 0, \beta > 0$ such that

$$v(t) = \alpha t^{\beta}, \tag{10}$$

as suggested in Section 2.2. Let c be the threshold corresponding to the terminal event at time X. As already mentioned, the distribution of X is non-parametrically identifiable in the semi-competing risks case. This is because X is either observed or is censored by an independent censoring mechanism. In Appendix (Section 8.2) we first prove identifiability of the parameters c, α, β from the distribution of X. Precisely, this means to show that the function $t \mapsto P(X > t)$ uniquely determines c, α, β. We furthermore prove in the Appendix that the distribution of S conditional on $S < c$ is identifiable when c and the parameters of the process $X(t)$ are given.

4 Analysis of a simulated data set

In order to illustrate the model, notation and method, we simulated a data set using the threshold model with $v(t) = \alpha t^{\beta}$, with $\alpha = 5, \beta = 1$, and threshold $c = 7$. We further let S be lognormally distributed with $\log S$ being normal with expectation $\mu_S = 2$ and standard deviation $\sigma_S = 0.25$. We simulated $n = 1000$ realizations of the model. The distribution of the censoring time τ was chosen to be a gamma distribution with parameter values giving approximately 10% observations with X censored.

The resulting data consisted of 385 observations from Case 1, 494 from Case 2, 16 from Case 3 and 105 from Case 4 (see Section 3.3). Maximum likelihood estimates and corresponding estimated standard errors (using a standard approach based on the Hessian matrix) are given in Table 2. As is seen from the table, the estimated parameter values are quite close to the true ones.

Table 2 Simulated data. Maximum likelihood estimates of the parameters in the model with lognormal S. In addition, the true values and estimated standard errors calculated from the Hessian matrix.

Parameter	True value	ML estimates	St. error
α	5	4.6975	0.5861
β	1	1.0109	0.0522
c	7	6.7394	0.7506
μ_S	2	1.9440	0.1113
σ_S	0.25	0.2564	0.0195

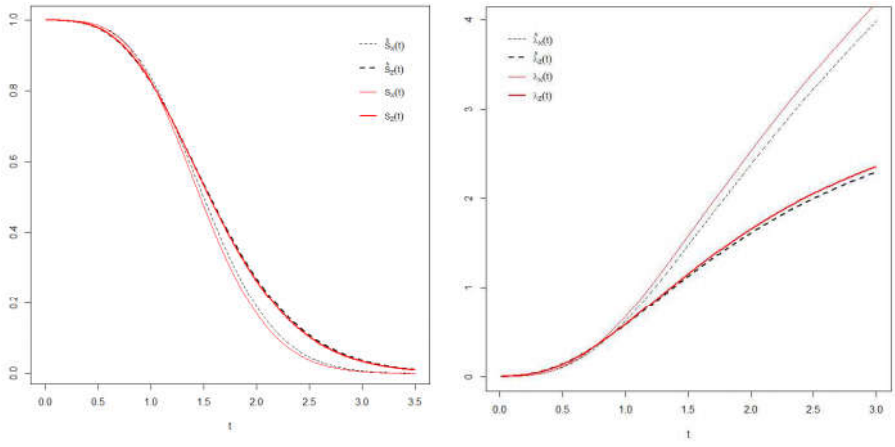

Fig. 2 Simulated data. True and estimated marginal survival functions $K_Z(t)$ and $K_X(t)$ (left) and marginal hazard functions $\lambda_Z(t)$ and $\lambda_X(t)$ (right).

In the left panel of Figure 2 we have plotted the estimated and true marginal survival functions, $K_X(t) = K(t; v(t), c)$ and $K_Z(t)$, where the latter is given by (6). The estimated curves are based on the estimated parameter values from Table 2. In the right panel of Figure 2 we plotted the estimated and true marginal hazard functions, $\lambda_X(t) = f(t; v(t), c)/K(t; v(t), c)$ and $\lambda_Z(t)$ from (7). The true curves are close to the estimated ones.

Next, we plotted the nonparametrically estimated crude quantities $F_Z^*(t)$ and $\Lambda_Z^*(t)$, i.e., the estimated sub-distribution function and cumulative sub-hazard rate of Z, in Figures 3 and 4, respectively. Here we used standard methods from the literature, as reviewed in the Appendix (Section 8.3), see (14) and (16). These curves, which do not depend on the choice of model, are compared to the true functions, (8) and (9), respectively, as well as the ones using estimated parameter values. As we can see from both figures, the curves all match each other very well.

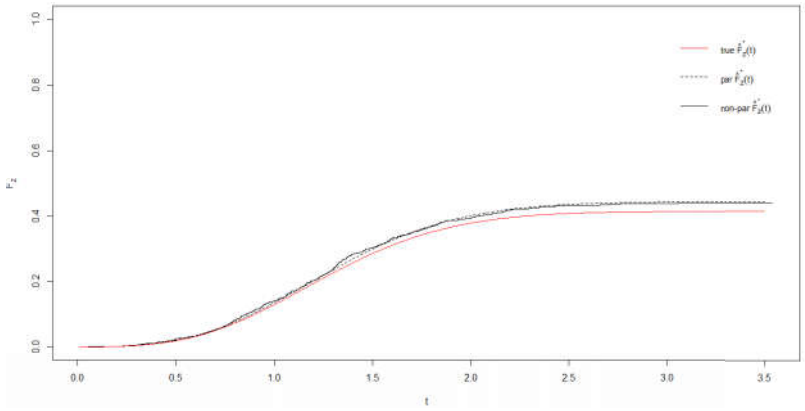

Fig. 3 Simulated data. Parametric and nonparametric estimates of the sub-distribution function for Z, $F_Z^*(t)$.

5 Case study: Analysis of the bone marrow transplant data

These data are collected from Klein and Moeschberger (1997), while they are originally from a study by Copeland et al (1991). The dataset contains observations of 137 patients that have undergone allogeneic bone marrow transplantation as treatment for acute leukemia. The terminal event in this case is death, while the non-terminal event is cancer relapse. For more information about the data and the study, see Klein and Moeschberger (1997). All times are measured in days from the time of transplantation. In the dataset there are 40 observations in Class 1, 41 in Class 2, 2 in Class 3 and 54 in Class 4. Thus this data set is also small compared to the simulated data set.

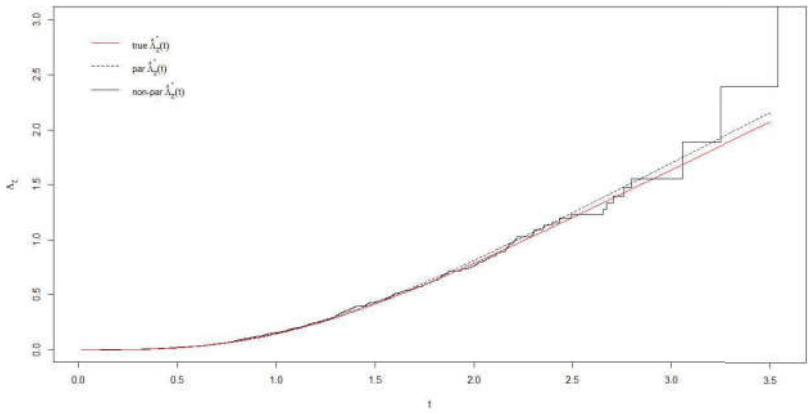

Fig. 4 Simulated data. Parametric and nonparametric estimates of the cumulative sub-hazard rate for Z, $\Lambda_Z^*(t)$.

In order to demonstrate that the threshold model may be an appropriate model, we started by checking whether the data indicate that condition (2) is reasonable. This is done in Figure 5, where the conditional survival functions (15) are plotted. It can be seen that $\hat{\tilde{S}}_Z(t)$ (thick line) $< \hat{\tilde{S}}_X(t)$ (thin line) for most values of time, except possibly for $t < 100$, where the difference is however not very large. Thus we judge this to be a sufficient indication for using the threshold model.

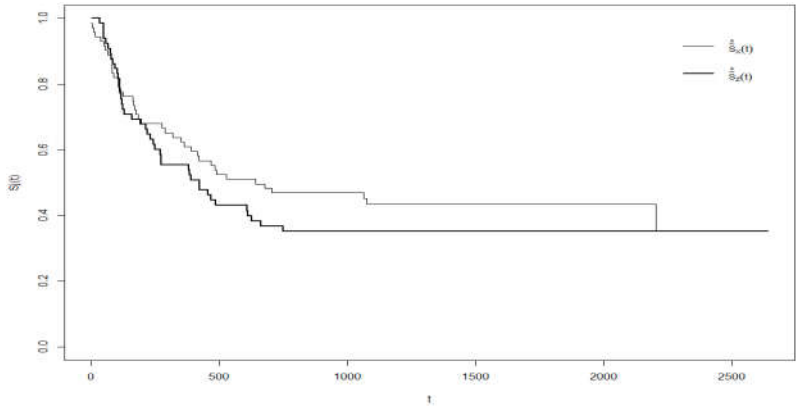

Fig. 5 Bone marrow transplant data. Nonparametric conditional sub-survival functions $\hat{\tilde{S}}_Z(t)$ and $\hat{\tilde{S}}_X(t)$

The bone marrow transplant data were also studied by Fine et al (2001), who assumed that the bivariate distribution of X and Z is a known copula, more specifically a gamma frailty copula. Among other quantities, they esti-

mated the marginal survivor function for the time to relapse. We will below compare their curve with the one obtained by our model.

We first fitted the threshold model to the data again using a lognormal S, obtaining the estimates $\hat{\mu}_S = 2.490$, $\hat{\sigma}_S = 0.0401$. Since $\hat{\sigma}_S/\hat{\mu}_S << 1$ we concluded that S might be better modeled by a normal distribution. In fact, this led to much more trustworthy estimates of standard errors when using the Hessian. The results from the maximum likelihood estimation procedure in the normal case are shown in Table 3.

Table 3 Bone marrow transplant data. Maximum likelihood estimates of the parameters in the model with normally distributed S. In addition, standard errors calculated from the Hessian matrix

Parameter	Estimate	Standard error
α	4.9218	0.6373
β	0.1382	0.006
c	12.1376	1.3475
μ_S	12.1620	1.3453
σ_S	0.5000	0.0799

In the same way as for the simulated data, we plotted in Figure 6 the estimated marginal survival functions $K_Z(t)$ and $K_X(t)$, as well as the estimated marginal hazard functions $\lambda_Z(t)$ and $\lambda_X(t)$. Unlike the plots in Figure 2, the curves for X and Z are very close, which we believe is a coincidence to be further discussed below.

The estimated density $f_S(s)$ is shown in Figure 7. From this we estimated the probability $F_S(c)$ of experiencing a relapse to be 0.4831. The fact that this value is close to 0.5, and in addition the symmetry of the normal density $f_S(s)$, may explain the strong similarity of the curves in Figure 6.

We next consider plots of the parametric and nonparametric estimates of the crude quantities $F_Z^*(t)$ and $\Lambda_Z^*(t)$, following the same procedure as for the simulated example. The plots are given in Figures 8 and 9.

In general, the parametric curves seem to fit fairly well to the nonparametric curves, although not as good as in the simulated example. We can see that the nonparametric curves for both $F_Z^*(t)$ and $\Lambda_Z^*(t)$ flatten out relatively early, since the majority of the observations are for smaller values of t. Hence, it might be difficult for the parametric estimates to match the nonparametric ones very closely.

We finally compare the estimated marginal distribution of Z to the one displayed by Fine et al (2001), see left panel of Figure 10. The survival function estimated by Fine et al. flattens out more than the curves estimated with our lognormal S model, which is probably caused by the relative lack of data with $t > 1000$. Still, our estimated function is mostly inside the estimated 95% confidence interval in the black dashed lines.

The right panel of Figure 10 is intended to give an indication of the variance of the estimated curve. The purple curves are the estimates of $K_Z(t)$ obtained

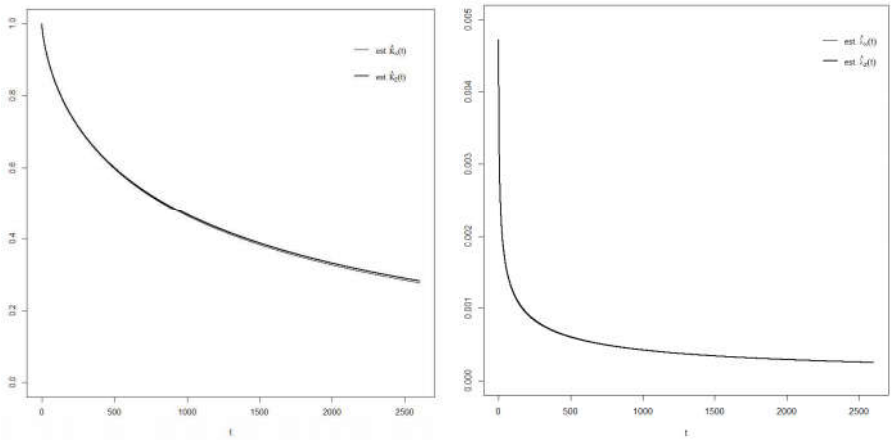

Fig. 6 Bone marrow transplant data. Estimated marginal survival functions $\hat{K}_Z(t)$ and $\hat{K}_X(t)$ (left) and hazard functions $\hat{\lambda}_Z(t)$ and $\hat{\lambda}_X(t)$ (right) with normal S

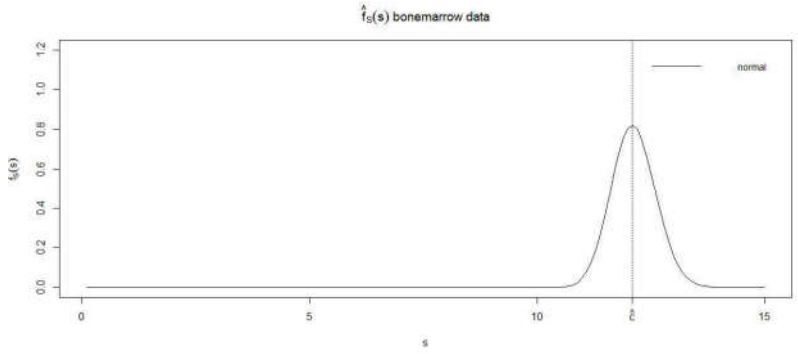

Fig. 7 Bone marrow transplant data. The estimated density $\hat{f}_S(s)$

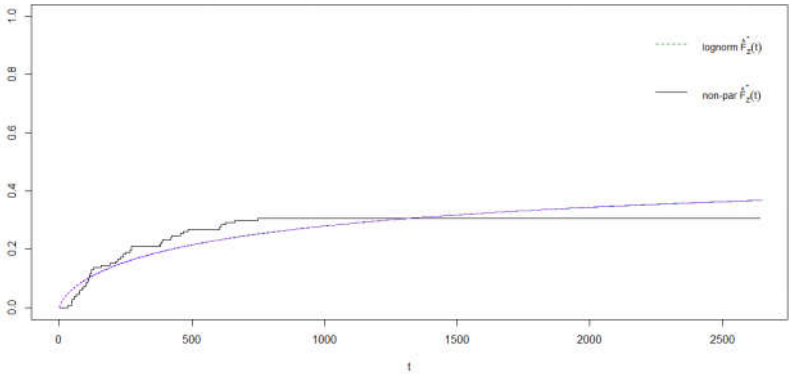

Fig. 8 Bone marrow transplant data. Parametric and nonparametric estimates of $F_Z^*(t)$

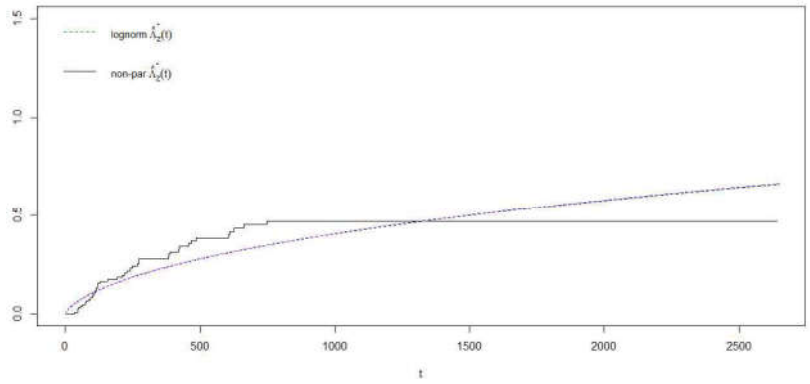

Fig. 9 Bone marrow transplant data.. Parametric and nonparametric estimates of $\Lambda_Z^*(t)$

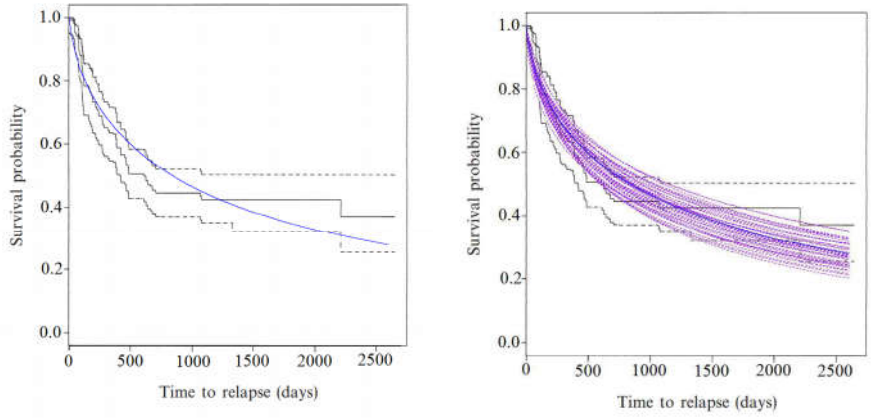

Fig. 10 Bone marrow transplant data. Estimated marginal survival function for the time to relapse, Z, compared to the nonparametric estimate and 95% confidence intervals from Fine et al (2001) (left). The same estimated curves, together with 50 estimated parametric curves obtained by nonparametric bootstrapping (right).

from 50 nonparametric bootstrap samples. As is seen, we get a bunch of curves with approximately the same width as the estimated confidence interval by Fine et al.

To summarize, the comparison of the parametric and nonparametric estimates for $F_Z^*(t)$ and $\Lambda_Z^*(t)$ gave us some indication that the fit of the parametric threshold model is satisfactory. Further, we have seen that our estimate for the marginal survival function of time to relapse, Z, was visually consistent with that obtained by Fine et al (2001). This may be interpreted as an indication of the usefulness of the suggested model.

6 Some final remarks on the threshold model

6.1 Comparison to the illness-death model

An illness-death model is given by the three states, 0 = "Healthy", 1 = "Diseased", 2 = "Dead". Consider the threshold model given by a process $D(t)$. At time 0 we have $D(0) = 0$, and we can think of the unit starting in state 0 in the illness-death model. In order to explain the threshold model, we may think that the random variable S is drawn at the start of the process. This is non-problematic since S is independent of the process by definition. After starting the process, the unit will remain in state 0 until time $\min(T_S, T_c)$, when it moves to state 1 if $S < c$ and to state 2 if $S > c$. If it moves to 1, it stays there until $D(t)$ crosses the level c. The sojourn time in state 1 (time to death) will hence depend on the value of S, as well as the process $D(t)$.

6.2 The role of the random threshold S

We may think of the threshold model as a model for the lifetime X of a unit, with the possibility of a non-terminal event happening at time $Z < X$. The occurrence or non-occurrence as well as the time of occurrence of the non-terminal event are governed by the random variable S which hence can be viewed as a way of modeling the tendency of the non-terminal event. Consider now two different distributions for S, represented by two variables S_1 and S_2. Then if S_1 is stochastically smaller than S_2 we have $P(S_1 < c) > P(S_2 < c)$, so the non-terminal event would have a larger probability of occurrence for the distribution S_1 than for S_2, whatever be the threshold c. Moreover, in this case we would have T_{S_1} stochastically smaller than T_{S_2}.

6.3 Dependence of the terminal and non-terminal event

We have earlier pointed to the dependency of X and Z. We are in fact able to show that X and Z possess the strong property of positive dependence named association (Esary et al, 1967). To see this, we use (Lindqvist, 1988, Theorem 6.1) to show that the process $\{D(t)\}$ is an associated stochastic process. The clue here is that the process has independent increments. Next, since for given values $s, c > 0$, the pair (T_s, T_c) is decreasing as a function of the process $\{D(t)\}$, we have that (T_s, T_c) is associated. Finally, since (T_s, T_c) is stochastically increasing in s, we can conclude from (Lindqvist, 1988, Theorem 4.1) that $(T_S, T_c) \equiv (Z, X)$ is an associated pair. By the definition this means that for any real functions $f(t_1, t_2), g(t_1, t_2)$ which are increasing in each argument, we have $\mathrm{Cov}(f(X, Z), g(X, Z)) \geq 0$ and hence in particular $\mathrm{Cov}(X, Z) \geq 0$.

7 Variations and extensions of the model

In this paper we have presented a new approach to modeling of semi-competing risks by means of first passage times of a stochastic process $\{D(t)\}$. The time to the non-terminal event equals to the first passage time to a stochastic level S,while the time to the terminal event is represented by the first passage time to a fixed level c. The two crucial assumptions are that S is independent of the process $\{D(t)\}$ and that the process $\{D(t)\}$ has independent increments.

To be explicit, we chose to consider a fully parametric model where the process $\{D(t)\}$ is a gamma-process and S is lognormally distributed. The approach may be successfully modified by replacing these by other choices. For example, the given model is close to the model considered in Lindqvist and Skogsrud (2008) which considers a Wiener process with drift for $\{D(t)\}$. Since the distribution of S (conditional on $S < c$) is nonparametrically identifiable (Section 11), we might in principle use any distribution for S. An interesting project for the future would be to estimate the distribution of S (conditional on $S < c$) nonparametrically.

A possibility that would flip the situation around, is to let $S = s$ be fixed while c is random. This problem would be very similar to the one considered, but the model interpretation would be different. For example, the conditional sub-survival curve of Z would now need to dominate that of X, which was not very likely in the two data examples. We have therfore not pursued this option.

A different way of incorporating a random effect in the model, would be to replace the scale parameter u by wu where w is a random effect (but which is set to 1 in our model). Actually, this would not be too different from letting one of the thresholds c or s be random, since (4) shows that u always appears together with the threshold in the distribution of the first passage time. On the other hand, the randomness would affect both $f_Z^*(z)$ and $f_X^*(x)$, not just one of them as in our case where only S is random. Thus, such a modification would only affect the speed of the underlying process $\{D(t)\}$ itself. Lawless and Crowder (2004) and Paroissin and Salami (2014) study models of this type.

For practical use it is of interest to include covariates in the model. There are several ways in which this could be done. Aalen and Gjessing (2001) distinguish between two types of covariates: (i) those that only represent measures of how far the underlying process has advanced (e.g., threshold level) and (ii) those that have causal influences on the development (e.g., drift parameter in a Wiener process). An option, studied by Lawless and Crowder (2004), is to let the scale parameter u depend on a vector of covariates $\mathbf{x}$, $u = u(\mathbf{x})$. Alternatively, Bagdonavicius and Nikulin (2001) included covariates in their model via an accelerated life test model, in our notation replacing $v(t)$ by $v(te^{\mathbf{x}^T\boldsymbol{\rho}})$. This is however computationally more demanding. By letting u be a function of covariates, the covariates affect the scaling of the degradation process only, and not the shape function. A natural choice might be $u(\mathbf{x}) = \exp(\boldsymbol{\rho}'\mathbf{x})$ where $\boldsymbol{\rho}$ is a vector of regression coefficients.

Putting the covariates multiplicatively on u is effectively the same as putting them on the critical threshold c or on s. As mentioned above in the discussion of random effects, u and c or s always appear together as the product $u \cdot c$ or $u \cdot s$ in the first passage time distribution. The scale parameter u is the same in both $f_X^*(x_i)$ and $f_Z^*(z_j)$, while the critical threshold is c in $f_X^*(x_i)$ and s in $f_Z^*(z_j)$. Therefore, the covariates have quite different meanings depending on which parameter they are assigned to.

For instance, consider including covariates in the parameter c. If we let 0 be the starting point of the underlying process for all units, we may use covariates on c to vary the threshold for the terminal event. For example it might be natural to assume that a smoking patient with a heart disease may reach the critical level faster than a non-smoker.

Covariates put on the threshold s on the other hand, provide information about the treatment policy of the item. For example in medicine this level may describe the level where a disease is diagnosed and treatment is started. Then, patients who are examined more often are more likely to be diagnosed at an early stage and have a lower level s than those that only rarely are examined.

To let u depend on covariates, has yet another implication. If the thresholds c and s are both fixed, as well as the probability $P(S < c)$, the covariates in u tell us something about the overall speed of the process, i.e., both the time until s is reached and until c is reached.

There is a large number of other possibilities that may be explored. A possibility is to experiment with different shape functions for $v(t)$. An alternative is for instance $v(t) = e^{\alpha+\beta t}$, or even let the $v(t)$ be a nonparametric function. One could also let the parameters α or β in the shape function $v(t)$ be random quantities. This is more computationally demanding and does not provide a similar intuitive interpretation as letting a threshold parameter be random.

References

Aalen OO, Gjessing HK (2001) Understanding the shape of the hazard rate: A process point of view. Statistical Science 16(1):1–22

Andersen PK, Borgan O, Gill RD, Keiding N (1993) Statistical models based on counting processes. Springer Science & Business Media, New York

Bagdonavicius V, Nikulin M (2001) Estimation in degradation models with explanatory variables. Lifetime Data Analysis 7:85–103

Borgan O (1998) Three contributions to the Encyclopedia of Biostatistics: The Nelson-Aalen, Kaplan-Meier and Aalen-Johansen estimators. Encyclopedia of Biostatistics

Casella G, Berger R (2002) Statistical Inference, 2nd edn. Duxbury, Pacific Grove.

Christen JA, Ruggeri F, Villa E (2011) Utility based maintenance analysis using a random sign censoring model. Reliability Engineering & System Safety 96(3):425–431

Clayton DG (1978) A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. Biometrika 65:141–151

Cooke RM (1993) The total time on test statistic and age-dependent censoring. Statistics & Probability Letters 18:307–312

Cooke RM, Bedford T (2002) Reliability databases in perspective. Reliability, IEEE Transactions on 51(3):294–310

Copeland EA, Biggs JC, Thompson JM, Crilley P, Szer J, Klein JP, Kapoor N, Avalos BR, Cunningham I, Atkinson K, Downs K, Harmon GS, Daly MB, Brodsky I, Bulova PJ S I amd Tutschka (1991) Treatment for acute myelocytic leukemia with allogeneic bone marrow transplantation following preparation with bucy2. Blood 78:838–843

Esary JD, Proschan F, Walkup DW, et al (1967) Association of random variables, with applications. The Annals of Mathematical Statistics 38(5):1466–1474

Fine JP, Jiang H, Chappell R (2001) On semi-competing risks data. Biometrika 88:907–919

Fix E, Neyman J (1951) A simple stochastic model of recovery, relapse, death and loss of patients. Human Biology 23:205–241

Horrocks J, Thompson ME (2004) Modeling event times with multiple outcomes using the wiener process with drift. Lifetime Data Analysis 10:29–49

Hsieh JJ, Wang W, Ding AA (2008) Regression analysis based on semicompeting risks data. Journal of the Royal Statistical Society series B 70:3–20

Kahle W, Mercier S, Paroissin C (2016) Degradation Processes in Reliability. John Wiley & Sons, Hoboken.

Klein JP, Moeschberger ML (1997) Survival Analysis: Techniques for Censored and Truncated Data, 1st edn. Springer Science+Business Media, New York

Lawless J, Crowder M (2004) Covariates and random effects in a gamma process model with application to degradation and failure. Lifetime Data Analysis 10:213–227

Lee MLT, Whitmore G (2006) Threshold regression for survival analysis: modeling event times by a stochastic process reaching a boundary. Statistical Science pp 501–513

Lindqvist BH (1988) Association of probability measures on partially ordered spaces. Journal of Multivariate Analysis 26(2):111–132

Lindqvist BH, Skogsrud G (2008) Modeling of dependent competing risks by first passage times of wiener processes. IIE Transactions 41(1):72–80

Lindqvist BH, Støve B, Langseth H (2006) Modelling of dependence between critical failure and preventive maintenance: the repair alert model. Journal of Statistical Planning and Inference 136(5):1701–1717

Meira-Machado L, de Uña-Álvarez J, Cadarso-Suárez C (2006) Nonparametric estimation of transition probabilities in a non-markov illness–death model. Lifetime Data Analysis 12(3):325–344

van Noortwijk J (2009) A survey of the application of gamma processes in maintenance. Reliability Engineering and System Safety 94:2–21

Park C, Padgett W (2005) Accelerated degradation models for failure based on geometric brownian motion and gamma processes. Lifetime Data Analysis 11:511–527

Paroissin C, Salami A (2014) Failure time of non homogeneous gamma process. Communications in Statistics - Theory and Methods 43(15):3148–3161

Peng L, Fine JP (2006) Regression modeling of semicompeting risks data. Biometrics 63:96–108

Putter H, Fiocco M, Geskus R (2007) Tutorial in biostatistics: competing risks and multi-state models. Statistics in medicine 26(11):2389–2430

Temme N (1975) Uniform asymptotic expansions of the incomplete gamma functions and the incomplete beta function. Mathematics of Computation 29(132):1109–1114

Varadhan R, Xue QL, Bandeen-Roche K (2014) Semicompeting risks in ageing research: methods, issues and needs. Lifetime Data Analysis 20:538–562

Whitmore G (1986) First-passage-time models for duration data: regression structures and competing risks. The Statistician pp 207–219

Xu J, Kalbfleisch JD, Tai B (2010) Statistical analysis of illness–death processes and semicompeting risks data. Biometrics 66(3):716–725

8 Appendix

8.1 Joint density of (T_d, T_c) for $d < c$

We will first calculate the joint density of $(T_d, D(T_d))$. This is done by integration of the joint density of $(T_d, D(T_d^-), D(T_d))$, which is given in (Kahle et al, 2016, Theorem 2.37). The result is, in our notation,

$$\begin{aligned}\tilde{f}(t,z) &= \frac{e^{-z}v'(t)}{\Gamma(v(t))}\int_0^d \frac{x^{v(t)-1}}{z-x}dx \\ &= \frac{e^{-z}v'(t)}{\Gamma(v(t))}\cdot\frac{d^{v(t)}\left(v(t)d\,{}_2F_1(1,v(t)+1;v(t)+2;d/z)+v(t)z+z\right)}{v(t)(v(t)+1)z^2}\end{aligned}$$

From this we get the joint density of (T_d, T_c) for $d < c$,

$$\begin{aligned}f(t_1,t_2\ ;\ v(t_1),v(t_2),d,c)dt_1dt_2) &= P(t_1 \le T_d \le t_1+dt_1, t_2 \le T_c \le t_2+dt_2) \\ &= \int_{z=d}^{c} P(t_1 \le T_d \le t_1+dt_1, z \le D(T_d) \le z+dz, t_2 \le T_c \le t_2+dt_2) \\ &= \left[\int_{z=d}^{c} \tilde{f}(t,z)P(t_2 \le T_c \le t_2+dt_2 | D(T_d)=z)dz\right]dt_1 \\ &= \left[\int_{z=d}^{c} \tilde{f}(t,z)f(t_2-t_1; v(t_2)-v(t_1), c-z)dz\right]dt_1dt_2\end{aligned}$$

There is, however, also a possibility that the process crosses both d and c at the same time, giving a tie between T_d and T_c. In this case, the relevant

density is

$$\begin{aligned} f(t_1, t_1 \ ; \ v(t_1), v(t_1), d, c)dt_1dt_2) &= P(t_1 \le T_d \le t_1 + dt_1, t_1 \le T_c \le t_1 + dt_1) \\ &= P(t_1 \le T_d \le t_1 + dt_1, D(T_d) > c) \\ &= \left[\int_c^\infty \tilde{f}(t_1, z)dz\right] dt_1 \end{aligned}$$

In our computer calculations we have used the simplifying approximation of the joint density of (T_d, T_c), which assumes that $D(T_d) = d$ and $D(T_c) = c$. In this case we have, for $d < c$,

$$\begin{aligned} f(t_1, t_2 \ ; \ v(t_1), v(t_2), d, c)dt_1dt_2) &= P(t_1 \le T_d \le t_1 + dt_1, t_2 \le T_c \le t_2 + dt_2) \\ &= P(t_1 \le T_d \le t_1 + dt_1)P(t_2 \le T_c \le t_2 + dt_2 | T_d = t_1) \\ &= f(t_1; v(t_1), d)dt_1 f(t_2 - t_1; v(t_2) - v(t_1), c - d)dt_1dt_2 \end{aligned}$$

8.2 Identifiability of the model

We first prove identifiability of the parameters c, α, β from the distribution of X. Note that we have $X = T_c$. Thus, from (4) we have,

$$P(X > t) = P(T_c > t) = \gamma(v(t), c)/\Gamma(v(t)) \tag{11}$$

where $\gamma(a, c) = \int_0^c z^{a-1}e^{-z}dz$. We first show as a digression that if c is unknown, then the function $v(t)$ is not nonparametrically identifiable. This follows since in the right hand expression of (11) we may for any given $c > 0$ solve for $v(t)$ for each fixed t. To see this, note from (11) that $P(X > t)$ equals $P(W > c)$ where $W \sim$ gamma$(v(t), 1)$. Since the gamma distribution is stochastically increasing in the shape parameter, here $v(t)$, we may always adjust the $v(t)$ to get a given value for $P(W > c)$.

Thus suppose instead that (10) holds. Now we use a result from Temme (1975) to see that as $a \to \infty$ we have

$$\gamma(a, c)/\Gamma(a) \sim \frac{c^a e^{-c}}{\Gamma(1 + a)} \sim (2\pi a)^{-1/2} e^{a-c} \left(\frac{c}{a}\right)^a.$$

Here the last expression is obtained by using Stirling's formula. Letting $a = \alpha t^\beta$ and taking the logarithm we get for large t,

$$\begin{aligned} \log P(X > t) \sim &-(1/2)\log 2\pi - (1/2)\log\alpha - (1/2)\beta\log t + \alpha t^\beta - c + \alpha t^\beta \log c \\ &- \alpha t^\beta \log\alpha - \alpha t^\beta \beta \log t \end{aligned} \tag{12}$$

Suppose now there is another combination of c, α, β, denoted c^*, α^*, β^*, for which the same $P(X > t)$ is obtained for all t. Then letting $t \to \infty$, the dominant term in (12) is $\alpha t^\beta \log t$ which hence must be equal for the two parametrizations, implying $\beta = \beta^*$ and hence also $\alpha = \alpha^*$. Finally, this clearly implies $c = c^*$ and we are done.

For identifiability of the full threshold model, it remains to show that the distribution of S conditional on $S < c$ is identifiable when c and the parameters of the process $D(t)$ are given. We are in fact able to show that this distribution is nonparametrically identifiable for any given $v(t)$ and c, which we for simplicity will assume to be strictly increasing and continuous, with $v(0) = 0$, $v(\infty) = \infty$. Suppose first that $c = \infty$, so that T_S is always observed. We will show that the distribution of T_S uniquely determines the distribution of S. Now

$$P(T_S > t) = P(S > D(t)) = E[P(S > D(t)|D(t))] = E[\bar{F}_S(W)] \tag{13}$$

where $W \sim \text{gamma}(v(t), 1)$ and $\bar{F}_S(s) = P(S > s)$. Since this is to hold for all $t > 0$, from the fact that the family of $W \sim \text{gamma}(\theta, 1)$ is a complete family of distributions, it follows that $\bar{F}_S$ is uniquely given and hence that the distribution of T_S uniquely determines the distribution of S (Casella and Berger, 2002, Chapter 6.2).

Next, for a given $c < \infty$ we need to show that the (observable) distribution $P(T_S > t|S < c)$ uniquely determines the distribution $P(S > s|S < c)$. This follows directly from the above argument which had $c = \infty$ by considering only distributions for S with support in $(0, c)$.

Note finally that (13) is a result of interest in itself if the distribution of the threshold S is given and one wants the distribution of T_S. Paroissin and Salami (2014) consider the cases where S is, respectively, exponentially and gamma distributed.

8.3 Nonparametric estimation of crude quantities.

Consider competing risks with latent variables X and Z. Suppose that n units are observed, either until (independent) right censoring or until time $T = \min(X, Z)$, whatever comes first. Let $t_1 < ... < t_k$ be the sorted event times, i.e., observations of T. Let further $\hat{S}(\cdot)$ be the Kaplan-Meier estimator of the survival function of T. Then the so-called Aalen-Johansen estimator of the sub-distribution functions are (Borgan, 1998):

$$\hat{F}^*_X(t) = \sum_{i;t_i \le t} \hat{S}(t_i)\frac{\delta_{iX}}{n_i},$$

$$\hat{F}^*_Z(t) = \sum_{i;t_i \le t} \hat{S}(t_i)\frac{\delta_{iZ}}{n_i}. \tag{14}$$

Here n_i is the number at risk at time t_i while $\delta_{iX} = 1$ ($\delta_{iZ} = 1$) if the observation at time t_i is an X (Z). The natural estimates of the conditional sub-distribution functions $\tilde{F}_X(t)$ and $\tilde{F}_Z(t)$ are hence

$$\hat{\tilde{F}}_X(t) = \frac{\hat{F}^*_X(t)}{\hat{F}^*_X(\infty)} \quad \text{and} \quad \hat{\tilde{F}}_Z(t) = \frac{\hat{F}^*_Z(t)}{\hat{F}^*_Z(\infty)}. \tag{15}$$

With the same notation we have nonparametric estimates of the cumulative cause-specific hazard functions for the two risks given by

$$\hat{\Lambda}^*_X(t) = \sum_{i;t_i \leq t} \frac{\delta_{iX}}{n_i},$$

$$\hat{\Lambda}^*_Z(t) = \sum_{i;t_i \leq t} \frac{\delta_{iZ}}{n_i}. \tag{16}$$