HAL
open science

A semiparametric shock model for a pair of event-related dependent censored failure times

Frédérique Letué

To cite this version:

Frédérique Letué. A semiparametric shock model for a pair of event-related dependent censored failure times. Journal of Statistical Planning and Inference, 2008, 138 (12), pp.3869-3884. 10.1016/j.jspi.2008.02.002 . hal-00160930

HAL Id: hal-00160930

https://hal.science/hal-00160930v1

Submitted on 9 Jul 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A semiparametric shock model for a pair of event-related dependent censored failure times

Frédérique Letué

LJK
Tour IRMA
51, rue des Mathématiques, B.P. 53
38041 Grenoble cedex 9
FRANCE

Abstract

In this paper, we propose a semiparametric shock model for two dependent failure times where the risk indicator of one failure time plays the role of a time-varying covariate for the other one. Following Hougaard (2000), the dependence between the two failure times is therefore of event-related type.

To derive our results, we define a joint distribution for the pair of failure times by means of a pair of parameters and a pair of baseline hazard rates. We then construct the associated counting processes and their compensators: in particular, we exhibit a special form of the risk indicators. We then propose a maximal partial likelihood estimator for the pair of parameters and Breslow estimators for the pair of baseline cumulative hazard rates. We establish the large sample properties of our estimators, following Andersen et al. (1992) and we illustrate our results by a short simulation study and an application to a real data set in demography

Key words: Bivariate censored data, Bivariate survival analysis, Cox model

1 Introduction

The Cox model (see Cox, 1972) is a semiparametric regression model for censored observations, where the conditional hazard rate of the failure times given the covariate $Z_i \in \mathbb{R}^d$ is given by

$$\alpha_{Z_i}(t) = e^{\theta_0^T Z_i} \alpha_0(t),$$

Email address: `Frederique.Letue@imag.fr` (Frédérique Letué).

where θ_0 is a $\mathbb{R}^d$-parameter and α_0 is a baseline hazard rate. In many pratical applications, it is useful to consider time-dependent covariates: the covariates $Z_i(.)$ are then time-varying processes. Kalbfleisch and Prentice (1980) distinguish between *defined time-dependent covariates* (when the covariate paths are deterministic, or fixed up to a $\mathcal{F}_0$-measurable random variable), *ancillary covariates* (as for instance, the level of air pollution in a study of the occurence of asthma attacks), or *internal covariates* (when the covariate is generated by the individual under study, like typically, disease complications).

In some cases, while considering two non-independent failure times $\widetilde{X}_A$ and $\widetilde{X}_B$, it can be interesting to consider $\mathbf{1\!1}_{\{\widetilde{X}_A<t\}}$ as a covariate for $\widetilde{X}_B$, and simultaneously, $\mathbf{1\!1}_{\{\widetilde{X}_B<t\}}$ as a covariate for $\widetilde{X}_A$. This may happen for instance in the case of infection within a family: when a member of the family is infected, it increases the risk of infection for the other members. Another example comes in the field of reliability, when considering the case of two parallel components: as soon as one of the two components breaks down, the other one becomes instantaneously more sollicitated, and thus, its risk of failure increases. In other cases, the risk may decrease: for instance, after a car accident for one member in a family, we can guess that every driver in the family will drive more carefully and thus reduce his own risk of accident. One can also think of two kinds of events happening to one single individual: for instance, one can guess that marriage will increase the probability to have rapidly a child, and simultaneously, having a child will probably change the probability to get married.

Described in such a way, the model is of course not well defined from a mathematical point of view. Indeed, the likelihood cannot be written simply as the product of the contributions of each member of the pair, since this would amount to describe the joint distribution of the pair as the product of the two conditional distributions. The aim of this paper is to propose a semiparametric bivariate model for the failure times $\widetilde{X}_A$ and $\widetilde{X}_B$, in which the indicators $\mathbf{1\!1}_{\{\widetilde{X}_A<t\}}$ and $\mathbf{1\!1}_{\{\widetilde{X}_B<t\}}$ will play the role of covariate for the associated counting processes. The model, called a shock model, is constructed in such a way that whenever either one of both counting process jumps, the hazard rate of the other one is instantaneously multiplied by a constant. The type of dependence between the two failure times $\widetilde{X}_A$ and $\widetilde{X}_B$ can be qualified, after Hougaard (2000), of event-related dependence, since the risk of one failure time is changed by the occurence of the other one.

In the next section, we propose a joint distribution for the pair $(\widetilde{X}_A, \widetilde{X}_B)$, we contruct the associated counting processes $(\widetilde{N}_A, \widetilde{N}_B)$ and their compensators with respect to the filtration generated by $\{\widetilde{N}_A(s), \widetilde{N}_B(s), 0 \leq s \leq t\}$.

The joint distribution is described by a pair of parameters $\rho_0 = (\rho_{A,0}, \rho_{B,0})$ and a pair of baseline hazard rates $(\alpha_{A,0}, \alpha_{B,0})$. In Section 3, we develop a maximum partial likelihood estimator for ρ_0 and Breslow estimators for $\alpha_{A,0}$ and $\alpha_{B,0}$, and we study their large sample properties applying Theorems VII.2.1, VII.2.2 and VII.2.3 of Andersen et al.(1993).

In Section 4, we illustrate our estimation results through a short simulation study for small sample sizes. In Section 5, we present an application to a data set in demography. Some theoretical justifications are gathered in Section 6.

2 A shock model for a pair of dependent failure times

2.1 The joint distribution

In this section, we propose a new class of joint distributions for pairs of failure times $(\widetilde{X}_A, \widetilde{X}_B)$, that cannot be considered as independent. This joint distribution is defined by means of other random variables, namely X_A, X_B, X'_A and X'_B.

Definition 2.1 *Let $\rho_{A,0}$ and $\rho_{B,0}$ be two real numbers, $\alpha_{A,0}$ and $\alpha_{B,0}$ be two non-negative functions. Let*

- *(X_A, X_B) be a pair of independent positive random variables with respective hazard rates $\alpha_{A,0}$ and $\alpha_{B,0}$,*
- *X'_A be a positive random variable, independent of X_A, with conditional hazard rate given X_B, $e^{\rho_{A,0}}\alpha_{A,0}(X_B + .)$,*
- *X'_B be a positive random variable, independent of X_B, with conditional hazard rate given X_A, $e^{\rho_{B,0}}\alpha_{B,0}(X_A + .)$.*

The shock distribution with parameter $\rho_0 = (\rho_{A,0}, \rho_{B,0})$ and hazard rates $\alpha_{A,0}$ and $\alpha_{B,0}$ is the distribution of the pair of failure times $(\widetilde{X}_A, \widetilde{X}_B)$ defined by

$$\begin{aligned} \widetilde{X}_A &= \min(X_A, X_B) + X'_A \, \mathbf{1}\!\!\mathbf{1}_{\{X_B < X_A\}}, \\ \widetilde{X}_B &= \min(X_B, X_A) + X'_B \, \mathbf{1}\!\!\mathbf{1}_{\{X_A < X_B\}} \, . \end{aligned} \tag{2.1}$$

Another way to define this distribution consists in saying that:

- If $X_A < X_B$, then $\widetilde{X}_A = X_A$ and $\widetilde{X}_B = X_A + X'_B$ (see Fig. 1).
- If $X_B < X_A$, then $\widetilde{X}_B = X_B$ and $\widetilde{X}_A = X_B + X'_A$ (see Fig. 2).

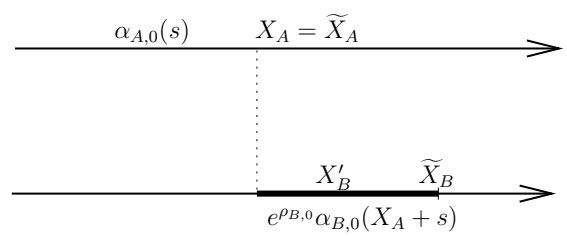

Fig. 1. Case $X_A < X_B$: $\widetilde{X}_B = X_A + X'_B$

Such a shock model can also be described as a marked point process with two jumps at times S_1 and S_2, and marks J_1 and J_2: with the notation of Definition 2.1, let us

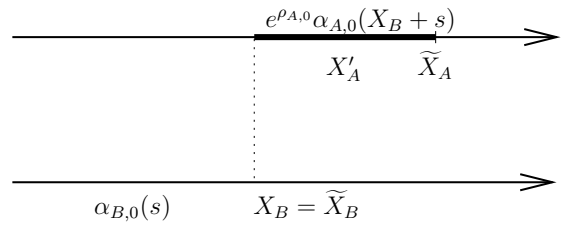

Fig. 2. Case $X_B < X_A$: $\widetilde{X}_A = X_B + X'_A$

call

- $S_1 = \min(X_A, X_B)$,
- $J_1 = A$ if $X_A < X_B$ and $J_1 = B$ if $X_B < X_A$.

Let us now define the conditional distribution of (S_2, J_2) given (S_1, J_1) by

- S_2 has conditional hazard rate $e^{\rho_{\bar{J}_1,0}}\alpha_{\bar{J}_1,0}(S_1 + .)$ given (S_1, J_1),
- $J_2 = \bar{J}_1$,

where $\bar{A} = B$ and $\bar{B} = A$. The pair (S_1, S_2) has the same distribution as the pair $(\widetilde{X}_A, \widetilde{X}_B)$ in Definition 2.1.

Another illustration is given in Fig. 3. From a state 0 where none of both events has occurred, we can go with hazard rate $\alpha_{A,0}$ to a state A where only A's event has occurred or with hazard rate $\alpha_{B,0}$ to a state B where only B's event has occurred. From these two states, we can go to a state AB where both events have occurred, and we pass with hazard rate $e^{\rho_{A,0}}\alpha_{A,0}$ if we come from state B and with hazard rate $e^{\rho_{B,0}}\alpha_{B,0}$ if we come from state A.

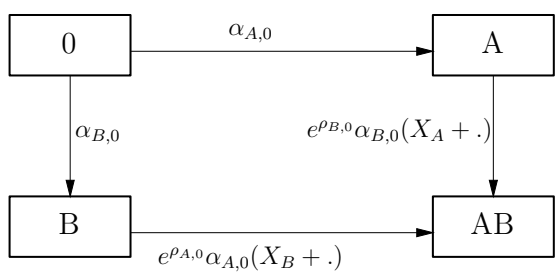

Fig. 3. A four states-model: state 0: no event has occurred, state A: only A's event has occurred, state B: only B's event has occurred, state AB: both events have occurred.

For the sake of completeness, we give in the next proposition the joint density of $(\widetilde{X}_A, \widetilde{X}_B)$.

Proposition 2.1 *The joint density of* $(\widetilde{X}_A, \widetilde{X}_B)$ *with shock distribution with pa-*

rameter $\rho_0 = (\rho_{A,0}, \rho_{B,0})$ and hazard rates $\alpha_{A,0}$ and $\alpha_{B,0}$ is given by:

$$\begin{aligned} f(s,t) &= \frac{e^{\rho_{B,0}}\alpha_{A,0}(s)\alpha_{B,0}(t)}{\exp\left[\int_0^s \alpha_{A,0} + \int_0^t e^{\rho_{B,0}\,\mathbf{1}_{\{u\geq s\}}}\alpha_{B,0}(u)du\right]}, \ \forall s < t, \\ f(s,t) &= \frac{e^{\rho_{A,0}}\alpha_{A,0}(s)\alpha_{B,0}(t)}{\exp\left[\int_0^s e^{\rho_{A,0}\,\mathbf{1}_{\{u\geq t\}}}\alpha_{A,0}(u)du + \int_0^t \alpha_{B,0}\right]}, \ \forall t < s. \end{aligned} \tag{2.2}$$

2.2 The associated counting processes

In this section, we consider the basic counting processes $\widetilde{N}_h$ for $h = A, B$, associated with the pair $(\widetilde{X}_A, \widetilde{X}_B)$ in a shock model. Let $(\mathcal{N}_t, t \geq 0)$ be the family of the σ-algebras $\mathcal{N}_t$ generated by $\{\widetilde{N}_A(u), \widetilde{N}_B(u), u \leq t, (Z_A, Z_B)\}$. We want to calculate the compensators of the counting processes with respect to the filtration $(\mathcal{N}_t, t \geq 0)$. These counting processes are defined by: $\widetilde{N}_h(t) = \mathbf{1}_{\{\widetilde{X}_h \leq t\}}$, for $t > 0$. Let us also define the functions $\widetilde{\alpha}_A$ and $\widetilde{\alpha}_B$ by

$$\begin{aligned} \widetilde{\alpha}_A(t) &= e^{\rho_{A,0}\,\mathbf{1}_{\{\widetilde{X}_B < \widetilde{X}_A, \widetilde{X}_B < t\}}}\alpha_{A,0}(t), \\ \widetilde{\alpha}_B(t) &= e^{\rho_{B,0}\,\mathbf{1}_{\{\widetilde{X}_A < \widetilde{X}_B, \widetilde{X}_A < t\}}}\alpha_{B,0}(t), \ \text{for } t > 0. \end{aligned} \tag{2.3}$$

Let us define the risk indicator $\widetilde{Y}_h$ for $h = A, B$ by $\widetilde{Y}_h(t) = \mathbf{1}_{\{\widetilde{X}_h \geq t\}}$, and

Assumption 1 *There exist some sets $\mathcal{T}_h \subset [0, \tau]$, $h = A, B$ such that*

$$\int_{\mathcal{T}_h} \alpha_{h,0}(s)ds > 0, \ h = A, B.$$

Furthermore,

$$\int_0^\tau \alpha_{h,0}(s)ds < \infty, \ h = A, B.$$

Proposition 2.2 *Under Assumption 1, the counting processes $\widetilde{N}_h$ for $h = A, B$ have predictable compensators $\widetilde{\Lambda}_h$, where $\widetilde{\Lambda}_h(t) = \int_0^t \widetilde{\alpha}_h(s)\widetilde{Y}_h(s)ds$ for any positive t, with respect to the filtration $(\mathcal{N}_t, t \geq 0)$.*

Let denote by $\widetilde{M}_h$ the difference $\widetilde{N}_h - \widetilde{\Lambda}_h$. We just need to prove that the processes $\widetilde{M}_h$ are $\mathcal{N}_t$-martingales, which is done in subsection 6.2.

3 Semi-parametric estimation in the shock model in the presence of censoring

In practice, the previous counting processes are not observed because of the presence of the censoring. We now have to define new counting processes that take account of the censoring. Then, we examine the large sample properties of the maximum log

partial likelihood of the parameters $\rho_{A,0}$ and $\rho_{B,0}$ and of the cumulative hazard rates $A_{h,0} = \int_0^{\cdot} \alpha_{h,0},\ h = A, B$.

3.1 The censored counting processes

We suppose now that the random variables $\widetilde{X}_A$ and $\widetilde{X}_B$ may be censored: we observe the random variables in some deterministic time interval $[0, \tau]$ and there exist two right-censoring times U_A and U_B verifying the following assumption:

Assumption 2 *The random vectors $(\widetilde{X}_A, \widetilde{X}_B)$ and (U_A, U_B) are independent. The distribution of (U_A, U_B) is such that:*

$$\mathbb{P}\left(U_A \geq \tau,\ U_B \geq \tau\right) > 0.$$

We allow, for instance, the case $U_A = U_B$. Instead of observing $(\widetilde{X}_A, \widetilde{X}_B)$, we actually only observe $(T_A, T_B, \delta_A, \delta_B)$, where

$$\begin{aligned} T_A &= \min(\widetilde{X}_A, U_A),\ T_B = \min(\widetilde{X}_B, U_B), \\ \delta_A &= \mathbf{1\!l}_{\{\widetilde{X}_A \leq U_A\}}, \qquad \delta_B = \mathbf{1\!l}_{\{\widetilde{X}_B \leq U_B\}}\,. \end{aligned} \tag{3.1}$$

In order to define the censored counting processes, let us consider individual A. We expect some counting process of the form $\delta_A\, \mathbf{1\!l}_{\{T_A \leq t\}}$, like in the classical Cox model. So, let us consider the case when $\delta_A = 1$. Whenever an event of type A occurs before an event of type B, that is to say whenever T_A is smaller than T_B, $\widetilde{X}_A$ is necessarily smaller than $\widetilde{X}_B$, so that process A cannot have been influenced by process B (see Fig. 4).

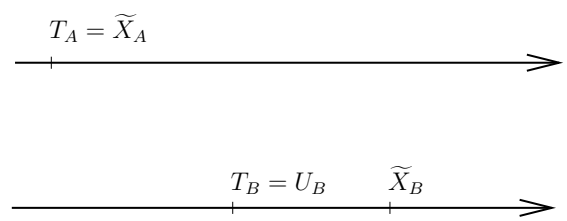

Fig. 4. Case $T_A < T_B$, $\delta_A = 1$, $\delta_B = 0$.

This situation is illustrated by the counting process N_1 defined by:

$$N_1(t) = \delta_A\, \mathbf{1\!l}_{\{T_A \leq T_B, T_A \leq t\}}\,.$$

On the contrary, whenever an event of type A occurs after one of type B, if process B is individually censored ($\delta_B = 0$), process A must also be considered as censored because in this case, we do not observe the time $\widetilde{X}_B$ when A's hazard rate is modified by the constant $e^{\rho_{A,0}}$ (see Fig. 5).

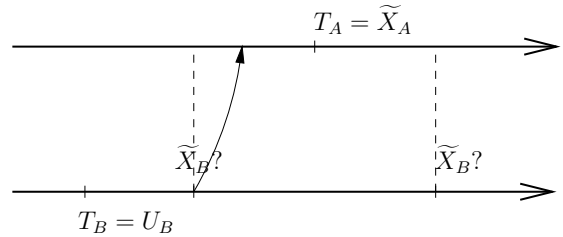

Fig. 5. Case $T_B < T_A$, $\delta_A = 1$, $\delta_B = 0$.

This leads to the counting process N_2:

$$N_2(t) = \delta_A \delta_B \, \mathbf{1\!\!1}_{\{T_B < T_A, T_A \leq t\}} \, .$$

We can now define the counting process N_A corresponding to individual A by $N_A = N_1 + N_2$ which implies formula (3.2). In the same manner, we define

$$\begin{aligned} N_3(t) &= \delta_B \, \mathbf{1\!\!1}_{\{T_B \leq T_A, T_B \leq t\}}, \\ N_4(t) &= \delta_A \delta_B \, \mathbf{1\!\!1}_{\{T_A < T_B, T_B \leq t\}}, \end{aligned}$$

and $N_B = N_3 + N_4$.

The two censored counting processes N_A and N_B resume then in:

$$\begin{aligned} N_A(t) &= \delta_A \, \mathbf{1\!\!1}_{\{T_A \leq t\}} (\delta_B \, \mathbf{1\!\!1}_{\{T_B < T_A\}} + \mathbf{1\!\!1}_{\{T_A \leq T_B\}}), \\ N_B(t) &= \delta_B \, \mathbf{1\!\!1}_{\{T_B \leq t\}} (\delta_A \, \mathbf{1\!\!1}_{\{T_A < T_B\}} + \mathbf{1\!\!1}_{\{T_B \leq T_A\}}), \text{ for any } t > 0, \end{aligned} \tag{3.2}$$

In order to calculate easily the compensator of these counting processes with respect to the filtration generated by the observations, we can try to interpret N_A and N_B as integrals of predictable processes with respect to the previous basic counting processes $\widetilde{N}_A$ and $\widetilde{N}_B$.

Recalling the definitions (3.1), we can write

$$\begin{aligned} N_A(t) &= \mathbf{1\!\!1}_{\{\widetilde{X}_A \leq U_A, \widetilde{X}_A \leq t\}} (\mathbf{1\!\!1}_{\{\widetilde{X}_B \leq U_B, \widetilde{X}_B < \widetilde{X}_A\}} + \mathbf{1\!\!1}_{\{\widetilde{X}_A \leq \widetilde{X}_B, \widetilde{X}_A \leq U_B\}}) \\ &= \int_0^t \mathbf{1\!\!1}_{\{u \leq U_A\}} (\mathbf{1\!\!1}_{\{\widetilde{X}_B \leq U_B, \widetilde{X}_B < u\}} + \mathbf{1\!\!1}_{\{u \leq \widetilde{X}_B, u \leq U_B\}}) d\widetilde{N}_A(u), \end{aligned}$$

since $T_h = \widetilde{X}_h$ as soon as $\delta_h = 1$. In the same way, we have

$$N_B(t) = \int_0^t \mathbf{1\!\!1}_{\{u \leq U_B\}} (\mathbf{1\!\!1}_{\{\widetilde{X}_A \leq U_A, \widetilde{X}_A < u\}} + \mathbf{1\!\!1}_{\{u \leq \widetilde{X}_A, u \leq U_A\}}) d\widetilde{N}_B(u).$$

Define now the censoring processes C_A and C_B by

$$\begin{aligned} C_A(u) &= \mathbf{1\!1}_{\{u\le U_A\}}(\mathbf{1\!1}_{\{\widetilde{X}_B\le U_B,\widetilde{X}_B<u\}}+\mathbf{1\!1}_{\{u\le\widetilde{X}_B,u\le U_B\}}),\\ C_B(u) &= \mathbf{1\!1}_{\{u\le U_B\}}(\mathbf{1\!1}_{\{\widetilde{X}_A\le U_A,\widetilde{X}_A<u\}}+\mathbf{1\!1}_{\{u\le\widetilde{X}_A,u\le U_A\}}). \end{aligned} \tag{3.3}$$

Integrating the censoring processes C_h with respect to the compensators $\widetilde{\Lambda}_h$, we define the processes $\Lambda_h(t) = \int_0^t C_h(u)d\widetilde{\Lambda}_h(u)$. Recalling the definitions of the $\widetilde{\Lambda}_h$ given in Proposition 2.2, we obtain:

$$\begin{aligned} \Lambda_A(t) &= \int_0^t \mathbf{1\!1}_{\{u\le U_A\}}(\mathbf{1\!1}_{\{\widetilde{X}_B\le U_B,\widetilde{X}_B<u\}}+\mathbf{1\!1}_{\{u\le\widetilde{X}_B,u\le U_B\}})\\ &\qquad e^{\rho_{A,0}\,\mathbf{1\!1}_{\{X_B<X_A,X_B<u\}}}\alpha_{A,0}(u)\,\mathbf{1\!1}_{\{\widetilde{X}_A\ge u\}}\,du\\ &= \int_0^t (e^{\rho_{A,0}}\delta_B\,\mathbf{1\!1}_{\{T_B<u\}}+\mathbf{1\!1}_{\{T_B\ge u\}})\,\mathbf{1\!1}_{\{T_A\ge u\}}\,\alpha_{A,0}(u)du,\\ \Lambda_B(t) &= \int_0^t (e^{\rho_{B,0}}\delta_A\,\mathbf{1\!1}_{\{T_A<u\}}+\mathbf{1\!1}_{\{T_A\ge u\}})\,\mathbf{1\!1}_{\{T_B\ge u\}}\,\alpha_{B,0}(u)du. \end{aligned}$$

After gathering terms, we can write the processes Λ_h as

$$\Lambda_h(t) = \int_0^t e^{\rho_{h,0}W_h(u)}Y_h(u)\alpha_{h,0}(u)du, \tag{3.4}$$

with notations

$$\begin{aligned} W_A(t) &= \mathbf{1\!1}_{\{T_B<t\}},\\ W_B(t) &= \mathbf{1\!1}_{\{T_A<t\}}\\ Y_A(t) &= \mathbf{1\!1}_{\{T_A\ge t\}}(\delta_B\,\mathbf{1\!1}_{\{T_B<t\}}+\mathbf{1\!1}_{\{T_B\ge t\}}),\\ Y_B(t) &= \mathbf{1\!1}_{\{T_B\ge t\}}(\delta_A\,\mathbf{1\!1}_{\{T_A<t\}}+\mathbf{1\!1}_{\{T_A\ge t\}}). \end{aligned} \tag{3.5}$$

Note the unusual form of the risk indicators Y_h, whereas an "intuitive" construction of the likelihood would have simply led to $Y_h(t) = \mathbf{1\!1}_{\{T_h\ge t\}}, h = A, B$. Here lies the interest of the long development of the contruction of the likelihood.

Finally, define the filtration $\mathcal{F}_t$ as the filtration generated by one observation, namely by $\{\mathbf{1\!1}_{\{T_A\le u\}}, \mathbf{1\!1}_{\{T_B\le u\}}, u\le t, \delta_A, \delta_B\}$.

Proposition 3.1 *Under Assumption 1, the counting processes N_h for $h = A, B$ have predictable compensators Λ_h with respect to the filtration $(\mathcal{F}_t, t\ge 0)$.*

Define first the filtrations $\mathcal{U}_t$ as the filtration generated by $\{\mathbf{1\!1}_{\{U_A\ge u\}}, \mathbf{1\!1}_{\{U_B\ge u\}}, u\le t\}$ and $\mathcal{G}_t$ as the filtration generated by both $\mathcal{N}_t$ and $\mathcal{U}_t$. The censoring processes C_h are cag-lad and clearly $\mathcal{G}_t$-adapted, so that they are $\mathcal{G}_t$-predictable.

Since the processes $\widetilde{N}_t$ and $\widetilde{\Lambda}_t$ are independent with respect to the filtration $\mathcal{U}_t$ and since they are $\mathcal{G}_t$-adapted, the $\mathcal{N}_t$-martingales $\widetilde{M}_h$ are also $\mathcal{G}_t$-martingales.

Since the censoring processes C_h are bounded and $\mathcal{G}_t$-predictable, we can affirm that the processes $M_h(t) = \int_0^t C_h(u)d\widetilde{M}_h(u)$ are also $\mathcal{G}_t$-martingales. Each σ-algebra $\mathcal{F}_t$ is contained in the σ-algebra $\mathcal{G}_t$, and it is clear from formulas (3.2), (3.4) and (3.5) that these martingales are also $\mathcal{F}_t$-measurable. Thus, they also are $\mathcal{F}_t$-martingales.

3.2 Construction of the likelihood

Suppose now that we observe a n-sample of $\{\mathbf{1\!l}_{\{T_A\leq u\}}, \mathbf{1\!l}_{\{T_B\leq u\}}, u\leq t, \delta_A, \delta_B\}$, namely $\{\mathbf{1\!l}_{\{T_{A,i}\leq u\}}, \mathbf{1\!l}_{\{T_{B,i}\leq u\}}, u\leq t, \delta_{A,i}, \delta_{B,i}\}$. In order to estimate the model parameters, we shall write the likelihood of the model using Jacod's formula (see Jacod 1973, 1974-75 , Andersen et al.1992). Since the counting processes $N_{h,i}, h = A, B, 1\leq i\leq n$ cannot jump simultaneously, the likelihood is proportional to:

$$\begin{aligned}
L_n(\rho) &= \prod_{h=A,B}\prod_{i=1}^{n}\prod_{s\leq t}(\alpha_{h,i}(s)Y_{h,i}(s))^{\Delta N_{h,i}(s)}\exp\left(-\int_0^\tau \alpha_{h,i}(u)Y_{h,i}(u)du\right)\\
&= \prod_{i=1}^{n}\prod_{s\leq t}\left(e^{\rho_A W_{A,i}(s)}Y_{A,i}(s)\right)^{\Delta N_{A,i}(s)}\left(e^{\rho_B W_{B,i}(s)}Y_{B,i}(s)\right)^{\Delta N_{B,i}(s)}\\
&\quad (\alpha_{A,0}(s))^{\Delta N_{A,i}(s)}(\alpha_{B,0}(s))^{\Delta N_{B,i}(s)}\\
&\quad \exp\left(-\int_0^\tau nS_{n,A}(\rho,u)\alpha_{A,0}(u)du\right)\exp\left(-\int_0^\tau nS_{n,B}(\rho,u)\alpha_{B,0}(u)du\right),
\end{aligned}$$

where $S_{n,A}$ and $S_{n,B}$ are defined by :

$$S_{n,h}(\rho,s) = n^{-1}\sum_{i=1}^{n} e^{\rho_h W_{h,i}(s)}Y_{h,i}(s) \tag{3.6}$$

Maximizing with respect to $\alpha_{A,0}$ and $\alpha_{B,0}$ for a fixed $\rho = (\rho_A, \rho_B)^T$, we finally find the log partial likelihood defined by :

$$l_n(\rho) = \sum_{h=A,B}\sum_{i=1}^{n}\int_0^\tau \log\left(\frac{e^{\rho_h W_{h,i}(s)}}{S_{n,h}(\rho,s)}\right)dN_{h,i}(s) \tag{3.7}$$

A version of this likelihood with added fixed-time covariates can be found in Letué (2000).

Remark 3.1 *Note that in the case where the two individuals are subject to the same kind of events (think of the case of two twins brothers subject to the same disease), we may assume that $\alpha_{A,0} = \alpha_{B,0}$ and maximizing the likelihood leads to another form of the log partial likelihood. In that case, just replace the sums $S_{n,h}$ by $S_{n,A} + S_{n,B}$.*

Our estimator $\hat{\rho}_n$ is then defined by

$$\hat{\rho}_n = (\hat{\rho}_{A,n}, \hat{\rho}_{B,n}) = \operatorname*{argmax}_{\rho} l_n(\rho_A, \rho_B). \tag{3.8}$$

Estimators for the integrated hazard rates $A_{h,0}(t) = \int_0^t \alpha_{h,0}(s)ds,\ h = A, B$ are the Breslow estimators $\hat{A}_{h,n}(\hat{\rho}_n, t)\ h = A, B$, where $\hat{A}_{h,n}(\rho, t)\ h = A, B$ are the Nelson-Aalen estimators :

$$\hat{A}_{h,n}(\rho, t) = n^{-1} \sum_{i=1}^{n} \int_0^t \frac{J_h(s)}{S_{h,n}(\rho, s)} dN_{h,i}(s),$$

with $J_h(s) = \mathbf{1}\!\mathbf{l}_{\{\sum_{j=1}^n Y_{h,j}(s)>0\}}$.

3.3 Large sample properties for the maximum partial likelihood estimators and the cumulative hazard rates

At this point, we are exactly in the framework described by Andersen and Gill (1982) and by Andersen et al.(1992, Chapter VII). Therefore, we can apply their Theorems VII.2.1 and VII.2.2 in order to establish our results. Recalling the definition (3.8) of our estimator, we can first state the consistency of our estimator.

Theorem 3.1 *Under Assumption 2, the maximal log partial likelihood $\hat{\rho}_n$ tends to the true parameter ρ_0, when n tends to infinity.*

In the next theorem, the asymptotic behaviour of the estimator is established and an estimator of the inverse of the asymptotic covariance matrix is proposed.

Let us introduce some notation. Let us denote by DF and D^2F the vector of first derivatives and the matrix of second derivatives of the function F with respect to ρ. Let $S_h, h = A, B$ be defined by $S_h(\rho, s) = \mathbb{E}\left(S_{n,h}(\rho, s)\right)$, let $E_h, h = A, B$ be the vectors $E_h = DS_h/S_h$ and let $V_h, h = A, B$ be the matrices $V_h = D^2S_h/S_h - (E_h)^{\otimes 2}$, where $x^{\otimes 2}$ denotes the vector product xx^T. Finally, let Σ_τ be the matrix:

$$\sum_{h=A,B} \int_0^\tau V_h(\rho_0, s) S_h(\rho_0, s) \alpha_{h,0}(s) ds. \tag{3.9}$$

Theorem 3.2 *Under Assumptions (1) and (2), the random vector $n^{1/2}(\hat{\rho}_n - \rho_0)$ converges in distribution to a Gaussian vector $\mathcal{N}\left(0, (\Sigma_\tau)^{-1}\right)$, and the matrix $-n^{-1}D^2 l_n(\hat{\rho}_n)$ tends in probability to Σ_τ.*

The proofs of these theorems were first established in Andersen and Gill (1982), they can also be found in Andersen et al.(1992). In the same paper, the authors propose some simple conditions to be verified in the special case where the observations are independent and identically distributed (Theorem 4.1 in Andersen and Gill (1982)). Since this is our case, we only need checking the assumptions of this theorem (see section 6.3).

Under the same hypotheses, Andersen and Gill (1982) proved the following results for the asymptotic joint distribution of $\hat{\rho}_n$ and of the Breslow estimators:

Theorem 3.3 *Under Assumptions (1) and (2), the random vector* $n^{1/2}(\hat{\rho}_n - \rho_0)$ *and the processes*

$$Z_{h,0}(.) = n^{1/2}(\hat{A}_{h,0}(\hat{\rho}_n, .) - A_{h,0}(.)) + n^{1/2}(\hat{\rho}_n - \rho_0)^T \int_0^{\cdot} E_h(\rho_0, s)\alpha_{h,0}(s)ds, \ h = A, B$$

are asymptotically independent. The random process $Z_h(.)$ *converges in distribution to a zero-mean Gaussian martingale with variance function*

$$\omega_h^2(.) = \int_0^{\cdot} \frac{\alpha_{h,0}(s)}{S_h(\rho_0, s)} ds.$$

The following corollary will enable us to construct confidence interval on $A_{h,0}(t)$ for a given h and a given t:

Corollary 3.1 *Under Assumptions (1) and (2), the process* $(n^{1/2}(\hat{A}_{h,0}(\hat{\rho}_n, .) - A_{h,0}(.)), h = A, B)$ *converges weakly to a bivariate Gaussian process with mean zero and covariance function*

$$\delta_{h,h'}\omega_h^2(s \wedge t) + \int_0^s E_h^T(\rho_0, u)\alpha_{h,0}(u)du\Sigma_\tau^{-1} \int_0^t E_{h'}(\rho_0, u)\alpha_{h',0}(u)du, h, h' = A, B,$$

which can be estimated uniformly consistently by

$$n\left\{\delta_{h,h'} \int_0^{s\wedge t} \hat{S}_{n,h}(\hat{\rho}_n, u)^{-2} dN_h(u) \right.$$
$$\left. + \int_0^s DS_{n,h}^T(\hat{\rho}_n, u)S_{h,n}^{-1}(u)dN_h(u)(-D^2 l_n(\hat{\rho}_n))^{-1} \int_0^t DS_{n,h'}(\hat{\rho}_n, u)S_{h',n}^{-1}(u)dN_{h'}(u)\right\},$$
$$h, h' = A, B.$$

4 A simulation study

We carry out a simulation study in order to illustrate the theoretical properties of our estimator $\hat{\rho}_n$ and to investigate its finite sample properties.

For different sample sizes ($n = 30, 50, 100$), we simulate a sample of independent and identically distributed variables $(T_{A,i}, T_{B,i}, \delta_{A,i}, \delta_{B,i})_{1\leq i\leq n}$ of size n, following the method indicated in Section 2.1: we begin with simulating a n-sample of independent and identically distributed variables $(X_{A,i}, X_{B,i})_{1\leq i\leq n}$ such that each pair has the distribution of two independent exponential random variables with parameter 1. That corresponds to constant hazard rates $\alpha_{A,0} = \alpha_{B,0} = 1$. Then, we simulate a n-sample of independent and identically distributed variables $(X'_{A,i}, X'_{B,i})_{1\leq i\leq n}$ such that each pair has the distribution of two independent exponential random variables with parameters $e^{\rho_{A,0}}$ and $e^{\rho_{B,0}}$. Here, both $\rho_{A,0}$ and $\rho_{B,0}$ equal 1. Random variables $(\widetilde{X}_{A,i}, \widetilde{X}_{B,i})_{1\leq i\leq n}$ result from formula (2.1). Right-censoring random variables $(U_{A,i}, U_{B,i})_{1\leq i\leq n}$ are also generated as independent exponential random

variables with parameter μ taking successively the values 0.0001, 0.1, 0.4 and 1. The approximate rate of censoring in each case is given in Tab. 1. The duration of study τ is taken to be 10.

For each experiment, we calculate the estimators following Remark 3.1 (recall that $\alpha_{A,0} = \alpha_{B,0}$), we calculate the estimator of the asymptotic covariance matrix and a 95%-confidence interval. We run $L = 500$ replications of such an experiment.

Table 1
Approximate rate of censoring with respect to the parameter of the censoring random variable μ

μ	0.0001	0.1	0.4	1
Approximate rate of censoring	0	0.06	0.21	0.42

For the different values of the sample size n and of the censoring parameter μ, we give the mean of the estimates over all replications, the approximate bias, the mean of asymptotic standard deviation, the coverage probabilities of the 95%-confidence interval calculated with the previous estimates. The results are given in Tab. 2.

Note that in all experiments, the bias is small. Not surprisingly, the asymptotic standard deviation decreases with the sample size for a given censoring parameter, and it increases with the censoring parameter for a given sample size. The coverage probabilities are close to the ones expected (0.025%, 0.950%, 0.025%).

To examine the performance of the estimators, we draw the empirical distribution functions of the estimates conveniently renormalized and we compare them to the distribution function of a standard Gaussian random variable. This is shown in Fig. 6 for a sample size equal to 50 and a censoring parameter equal to 0.1. We note that even with small sample sizes the distribution of the estimators are very close to the asymptotic distribution.

5 Application to a real data set in demography

We present here the application of our model to a simple demographic data set. The data were furnished by INED (Institut National d' Études Démographiques, Paris, France). They consist on the age at first child's birth (X_A) and age at marriage (X_B) of $n = 2830$ individuals, born between 1930 and 1950 in the Paris area (see Lelièvre and Vivier (2001)). The data were collected in 2001, so we observe censored data for individuals who did not get married or did not have any child at that time, and the censoring times are then their age in 2001. The censoring variables are the same for the two processes "marriage" and "first child's birth". Among the population, 332 individuals never got married in 2001, 317 individuals never had any child and 191 both never got married and had no child. The final time of study was taken as $\tau = 70$.

Table 2
Small sample performance of the estimator for the shock parameters for various values of the censoring parameter μ and of the sample size n.

μ	n	$\hat{\rho}_A$	$\hat{\rho}_A - \rho_{A,0}$	$\hat{\sigma}_A$	$<$ IC	$\in$ IC	$>$ IC
		$\hat{\rho}_B$	$\hat{\rho}_B - \rho_{B,0}$	$\hat{\sigma}_B$			
0.0001	30	0.9862	-0.0138	0.3553	0.024	0.952	0.024
		1.0019	0.0019	0.3550	0.026	0.950	0.024
	50	1.0135	0.0135	0.2684	0.028	0.948	0.024
		0.9885	-0.0115	0.2688	0.028	0.0936	0.036
	100	1.0012	0.0012	0.1864	0.030	0.950	0.020
		0.9949	-0.0051	0.1868	0.032	0.956	0.012
0.1	30	0.9651	-0.0349	0.3790	0.024	0.950	0.026
		0.9879	-0.0121	0.3809	0.028	0.950	0.022
	50	0.9875	-0.0125	0.2871	0.040	0.932	0.028
		1.0094	0.0094	0.2876	0.026	0.950	0.024
	100	0.9857	-0.0143	0.1990	0.032	0.0940	0.028
		0.9925	-0.0075	0.1992	0.022	0.944	0.034
0.4	30	1.0155	0.0155	0.4583	0.040	0.938	0.022
		0.9785	-0.0215	0.4588	0.030	0.944	0.026
	50	0.9985	-0.0015	0.3411	0.026	0.944	0.030
		1.0094	0.0094	0.3405	0.038	0.944	0.018
	100	0.9974	-0.0026	0.2355	0.042	0.930	0.028
		0.9961	-0.0039	0.2360	0.030	0.930	0.040
1	30	0.9669	-0.0331	1.5212	0.040	0.934	0.026
		0.9826	-0.0174	1.1174	0.044	0.940	0.016
	50	1.0191	0.0191	0.4526	0.022	0.958	0.020
		1.0349	0.0349	0.4531	0.034	0.954	0.012
	100	1.0017	0.0017	0.3019	0.032	0.948	0.020
		0.9993	-0.0007	0.3068	0.034	0.948	0.018

For this data set, we obtain $\hat{\rho}_A = 2.6465$ and $\hat{\rho}_B = -0.4248$ with respective estimated asymptotic standard deviation 0.0547 and 0.0997. This leads to respective p-values 0 and 2.0204×10^{-5}. Hence, these values are significant and seem to indicate that marriage increases the hazard rate of the age at the first child's birth whereas having

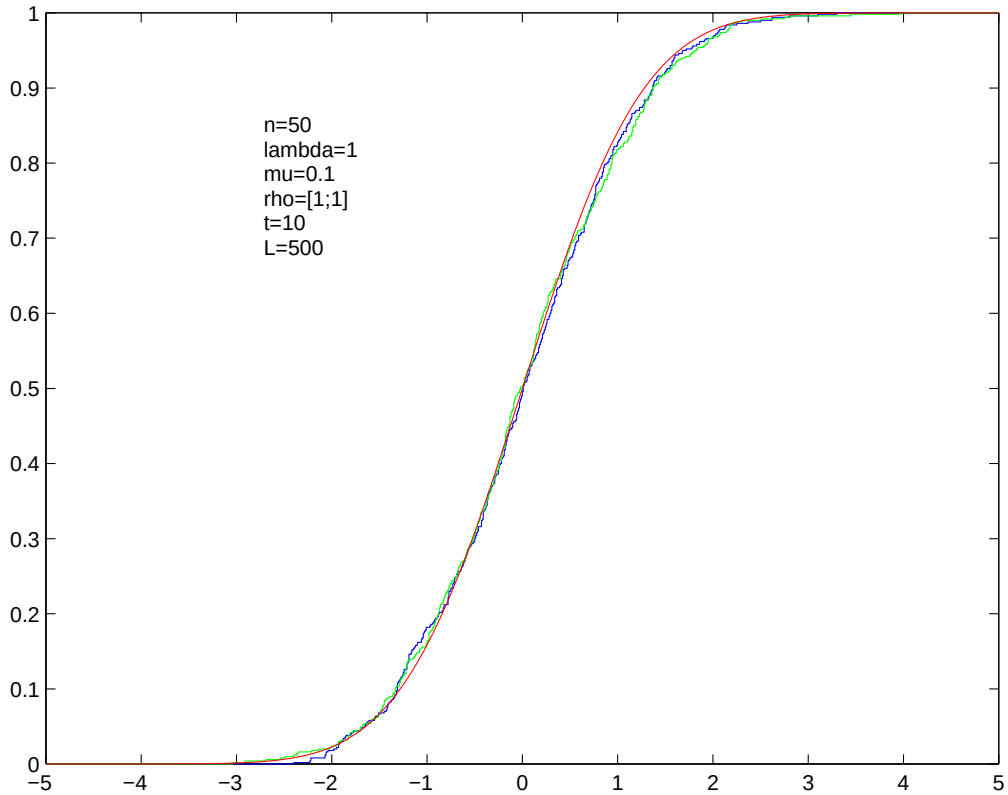

Fig. 6. Empirical distribution functions of the estimators in comparison with the standard Gaussian distribution function

a child decreases the hazard rate of the age at marriage.

We display also in Fig. 7 the plot of the Breslow estimators $\hat{A}_h, h = A, B$ of the cumulated hazard rates $\int \alpha_h, h = A, B$ (truncated to 0) and the associated estimated survival functions $exp(-\hat{A}_h), h = A, B$ (truncated to 1) . Recall that these functions should be interpreted as conditional survival functions given that the other event did not happen.

As for claiming that the model fits correctly the data, one would need goodness-of-fit tests, but this is still under investigation and left for future research.

6 Proofs

We regroup in this Section, the proofs of the various assertions and propositions stated in the main body of the paper.

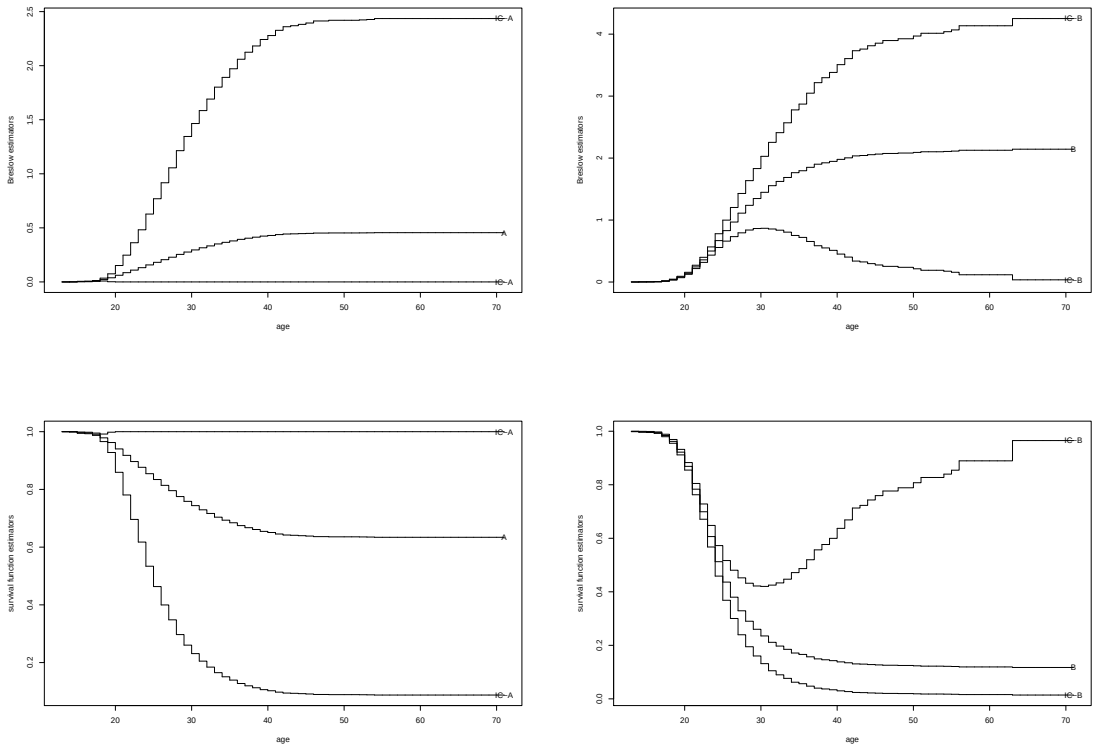

Fig. 7. Breslow estimator (above) and estimated survival functions (below) for events A (left) and B (right).

6.1 Proof of Proposition 2.1

Let us first calculate the bivariate survival function of $(\tilde{X}_A, \tilde{X}_B)$, say F, defined by:

$$F(t,s) = \mathbb{P}\left(\tilde{X}_A \geq t, \tilde{X}_B \geq s\right).$$

For $t < s$, we have on the one hand:

$$\begin{aligned}
\mathbb{P}\left(\tilde{X}_A \geq t, \tilde{X}_B \geq s, \tilde{X}_A < \tilde{X}_B\right) &= \mathbb{P}\left(X_A \geq t, X_A + X'_B \geq s, X_A < X_B\right) \\
&= \mathbb{E}\left(\mathbf{1}\!\!\mathbf{1}_{\{X_A \geq t\}} \mathbb{P}\left(X'_B \geq s - X_A, X_B > X_A | X_A\right)\right) \\
&= \mathbb{E}\left(\mathbf{1}\!\!\mathbf{1}_{\{X_A \geq t\}} e^{-\int_{X_A}^{s} e^{\rho_B} \alpha_B \mathbf{1}\!\!\mathbf{1}_{\{X_A \leq s\}}} e^{-\int_0^{X_A} \alpha_B}\right) \\
&= \int_t^{+\infty} e^{-\int_x^s e^{\rho_B} \alpha_B \mathbf{1}\!\!\mathbf{1}_{\{x \leq s\}}} e^{-\int_0^x \alpha_B} \alpha_A(x) e^{-\int_0^x \alpha_A} dx \\
&= \int_t^{s} e^{-\int_0^x \alpha_A + \alpha_B - \int_x^s e^{\rho_B} \alpha_B} \alpha_A(x) dx \\
&\quad + \int_s^{+\infty} e^{-\int_0^x \alpha_A + \alpha_B} \alpha_A(x) dx.
\end{aligned}$$

On the other hand,

$$\begin{aligned}\mathbb{P}\left(\tilde{X}_A \geq t, \tilde{X}_B \geq s, \tilde{X}_B < \tilde{X}_A\right) &= \mathbb{P}\left(\tilde{X}_B \geq s, \tilde{X}_B < \tilde{X}_A\right) = \mathbb{P}\left(X_B \geq s, X_B < X_A\right) \\ &= \mathbb{E}\left(\mathbf{1}\!\!\mathbf{1}_{\{X_B \geq s\}} \mathbb{P}\left(X_A > X_B | X_B\right)\right) \\ &= \mathbb{E}\left(\mathbf{1}\!\!\mathbf{1}_{\{X_B \geq s\}} e^{-\int_0^{X_B} \alpha_A}\right) \\ &= \int_s^{+\infty} e^{-\int_0^x \alpha_A + \alpha_B} \alpha_B(x) dx.\end{aligned}$$

Hence, we get, for $t < s$

$$F(t,s) = \int_t^s e^{-\int_0^x \alpha_A + \alpha_B - \int_x^s e^{\rho_B} \alpha_B} \alpha_A(x) dx + \int_s^{+\infty} e^{-\int_0^x \alpha_A + \alpha_B} (\alpha_A + \alpha_B)(x) dx,$$

and similarly for $s < t$

$$F(t,s) = \int_s^t e^{-\int_0^x \alpha_A + \alpha_B - \int_x^t e^{\rho_A} \alpha_A} \alpha_B(x) dx + \int_t^{+\infty} e^{-\int_0^x \alpha_A + \alpha_B} (\alpha_A + \alpha_B)(x) dx.$$

We retrieve the joint density by deriving once with respect to t and once with respect to s: $f(t,s) = \frac{\partial^2 F}{\partial t \partial s}$.

6.2 Proof of Proposition 2.2

The processes $\widetilde{\Lambda}_h$ are nondecreasing and $\mathcal{N}_t$-adapted, thus they are $\mathcal{N}_t$-predictable. It suffices to prove that the processes $\widetilde{M}_h$ defined just after Proposition 2.2 are $\mathcal{N}_t$-martingales, that is to say, for $\widetilde{M}_A$, that for every $s < t$, $\mathbb{E}\left(\widetilde{M}_A(t) - \widetilde{M}_B(s) \mid \mathcal{N}_s\right) = 0$. In fact, we have to prove the five following points: for every $u \leq s < t,\ v \leq s < t$,

i) $\mathbb{E}\left(|\widetilde{M}_A(t)|\right) < \infty$
ii) $\mathbb{E}\left(\left(\widetilde{M}_A(t) - \widetilde{M}_A(s)\right) \mathbf{1}\!\!\mathbf{1}_{\{\widetilde{X}_A \leq u\}}\right) = 0$
iii) $\mathbb{E}\left(\left(\widetilde{M}_A(t) - \widetilde{M}_A(s)\right) \mathbf{1}\!\!\mathbf{1}_{\{\widetilde{X}_B \leq v\}}\right) = 0$
iv) $\mathbb{E}\left(\left(\widetilde{M}_A(t) - \widetilde{M}_A(s)\right) \mathbf{1}\!\!\mathbf{1}_{\{\widetilde{X}_A \leq u\}} \mathbf{1}\!\!\mathbf{1}_{\{\widetilde{X}_B \leq v\}}\right) = 0$
v) $\mathbb{E}\left(\widetilde{M}_A(t) - \widetilde{M}_A(s)\right) = 0$

First of all, note the equality of the following events appearing in the definition (2.3) of the functions $\tilde{\alpha}_H, h = A, B$:

$$\begin{aligned}\{X_B < X_A, X_B < t\} &= \{\widetilde{X}_B < \widetilde{X}_A, \widetilde{X}_B < t\}, \\ \{X_A < X_B, X_A < t\} &= \{\widetilde{X}_A < \widetilde{X}_B, \widetilde{X}_A < t\}.\end{aligned}$$

The direct implication is clear. For the reverse implication, just recall the definition (2.1). If $\widetilde{X}_B < \widetilde{X}_A$, then $\widetilde{X}_B \mathbf{1}\!\!\mathbf{1}_{\{X_A < X_B\}} < \widetilde{X}_A \mathbf{1}\!\!\mathbf{1}_{\{X_B < X_A\}}$, and since only one of both indicators can equal 1, necessarily $X_B < X_A$, and the conclusion follows.

Point i) is fulfilled with Assumption 1.

Points ii) and iv) result from the equalities $(\widetilde{M}_A(t) - \widetilde{M}_A(s))\, \mathbf{1\!l}_{\{\widetilde{X}_A \le u\}} = 0$ and $(\widetilde{M}_A(t) - \widetilde{M}_A(s))\, \mathbf{1\!l}_{\{\widetilde{X}_A \le u\}}\, \mathbf{1\!l}_{\{\widetilde{X}_B \le v\}} = 0$.

To prove point iii), by definition of $\widetilde{X}_A$ and $\widetilde{X}_B$, we have:

$$\left(\widetilde{M}_A(t) - \widetilde{M}_A(s)\right) \mathbf{1\!l}_{\{\widetilde{X}_B \le v\}} = \mathbf{1\!l}_{\{s < X_B + X'_A \le t, X_B \le v\}} - \int_s^t e^{\rho_{A,0}} \alpha_A(x)\, \mathbf{1\!l}_{\{X_B + X_A\prime \ge x, X_B \le v\}}\, dx.$$

The expectation of the first term is $\mathbb{P}\left(s < X_B + X'_A \le t, X_B \le v\right)$ that we calculate conditioning on X_B:

$$\mathbb{P}\left(s < X_B + X'_A \le t, X_B \le v\right) = \mathbb{E}\left(\int_s^t e^{\rho_{A,0}} \alpha_A(x) \exp\left(-\int_{X_B}^x e^{\rho_{A,0}} \alpha_A\right) dx\, \mathbf{1\!l}_{\{X_B \le v\}}\right).$$

For the expectation of the second term, we also condition on X_B, and we find:

$$\mathbb{E}\left(\int_s^t e^{\rho_{A,0}} \alpha_A(x)\, \mathbf{1\!l}_{\{X_B + X_A\prime \ge x, X_B \le v\}}\, dx\right) = \mathbb{E}\left(\int_s^t e^{\rho_{A,0}} \alpha_A(x) \exp\left(-\int_{X_B}^x e^{\rho_{A,0}} \alpha_A\right) dx\, \mathbf{1\!l}_{\{X_B \le v\}}\right),$$

so that $\mathbb{E}\left(\left(\widetilde{M}_A(t) - \widetilde{M}_A(s)\right) \mathbf{1\!l}_{\{\widetilde{X}_B \le v\}}\right) = 0$.

We handle in the same way for point v).

6.3 Verification of Conditions VII.2.1 and VII.2.2 of Andersen et al. (1993)

Since the observations $(T_{A,i}, T_{B,i}, \delta_{A,i}, \delta_{B,i})_{1 \le i \le n}$ are independent and identically distributed, we are typically in the framework given by Theorem 4.1 of Andersen and Gill (1982) (see also Example VII.2.7 of Andersen et al.(1992)). We only have to verify the following points:

i) $\int_0^\tau \alpha_{h,0}(s) ds,\ h = A, B$ are finite
ii) There exists a neighbourhood $\mathcal{R}$ of ρ_0 such that for any $h = A, B$,

$$\mathbb{E}\left(\sup_{\rho \in \mathcal{R}, 0 \le s \le t} Y_h(s) |W_h(s)|^2 \exp(\rho_h^T W_h(s))\right)$$

is finite
iii) For any $h = A, B$, $\mathbb{P}\left(Y_h(s) = 1,\ \text{for any } 0 \le s \le \tau\right)$ is positive
iv) The asymptotic covariance matrix Σ_τ is positive definite

- Point i) is exactly Assumption 1.

- For point ii), note that

$$Y_h(s)|W_h(s)|^2 \exp(\rho_h^T W_h(s)) \leq e^{\rho_h},$$

so that any bounded neighbourhood of ρ_0 works.
- For point iii), note that in our definition (3.5), the functions $Y_{h,i}$ are nonincreasing with respect to s, so that the event in the probability is the same as $Y_h(\tau) = 1$.

$$\begin{aligned}\mathbb{P}\left(Y_A(\tau) = 1\right) &= \mathbb{P}\left(\widetilde{X}_A \geq \tau, U_A \geq \tau, \widetilde{X}_B \leq U_B, \widetilde{X}_B < \tau\right) \\ &\quad + \mathbb{P}\left(\widetilde{X}_A \geq \tau, U_A \geq \tau, \widetilde{X}_B \geq \tau, U_B \geq \tau\right) \\ &\geq \mathbb{P}\left(\widetilde{X}_A \geq \tau, U_A \geq \tau, \widetilde{X}_B \geq \tau, U_B \geq \tau\right).\end{aligned}$$

Thanks to Assumption 2, the latter probability also equals

$$\mathbb{P}\left(U_A \geq \tau, U_B \geq \tau\right) \mathbb{P}\left(\widetilde{X}_A \geq \tau, \widetilde{X}_B \geq \tau\right).$$

The first probability is positive from Assumption 2. For the second, we can write that

$$\mathbb{P}\left(\widetilde{X}_A \geq \tau, \widetilde{X}_B \geq \tau\right) = \mathbb{P}\left(\min(\widetilde{X}_A, \widetilde{X}_B) \geq \tau\right).$$

Since $\min(\widetilde{X}_A, \widetilde{X}_B) = \min(X_A, X_B)$, this probability also equals

$$\mathbb{P}\left(X_A \geq \tau, X_B \geq \tau\right) = \exp(-\sum_{h=A,B} \int_0^\tau \alpha_{h,0}(s) ds),$$

which is positive from Assumption 1. The same arguments hold for Y_B.
- For point iv), let us recall the expression of Σ_τ (3.9):

$$\Sigma_\tau = \sum_{h=A,B} \int_0^\tau V_h(\rho_0, s) S_h(\rho_0, s) \alpha_{h,0}(s) ds,$$

with $S_h(\rho, s) = \mathbb{E}\left(e^{\rho_h W_h(s)} Y_h(s)\right)$ $h = A, B$, $E_h = DS_h/S_h$ $h = A, B$ and $V_h = D^2 S_h/S_h - (E_h)^{\otimes 2}$ $h = A, B$. Let $x = (x_A, x_B)$ be a two-dimensional vector. We have :

$$x^T \Sigma_\tau x = \sum_{h=A,B} \int_0^\tau x^T V_h(\rho_0, s) x S_h(\rho_0, s) \alpha_{h,0}(s) ds.$$

Calculating the derivatives of S_h, $h = A, B$, we find

$$DS_A(\rho,s)=\begin{pmatrix}\mathbb{E}\left(e^{\rho_A W_A(s)}Y_A(s)W_A(s)\right)\\ 0\end{pmatrix},$$

$$DS_B(\rho,s)=\begin{pmatrix}0\\ \mathbb{E}\left(e^{\rho_B W_B(s)}Y_B(s)W_B(s)\right)\end{pmatrix},$$

$$D^2S_A(\rho,s)=\begin{pmatrix}\mathbb{E}\left(e^{\rho_A W_A(s)}Y_A(s)W_A(s)^2\right) & 0\\ 0 & 0\end{pmatrix},$$

$$D^2S_B(\rho,s)=\begin{pmatrix}0 & 0\\ 0 & \mathbb{E}\left(e^{\rho_B W_B(s)}Y_B(s)W_B(s)^2\right)\end{pmatrix}.$$

and therefore

$$V_A(\rho,s)=\begin{pmatrix}\frac{\mathbb{E}\left(e^{\rho_A W_A(s)}Y_A(s)W_A(s)^2\right)}{\mathbb{E}\left(e^{\rho_A W_A(s)}Y_A(s)\right)}-\left(\frac{\mathbb{E}\left(e^{\rho_A W_A(s)}Y_A(s)W_A(s)\right)}{\mathbb{E}\left(e^{\rho_A W_A(s)}Y_A(s)\right)}\right)^2 & 0\\ 0 & 0\end{pmatrix},$$

$$V_B(\rho,s)=\begin{pmatrix}0 & 0\\ 0 & \frac{\mathbb{E}\left(e^{\rho_B W_B(s)}Y_B(s)W_B(s)^2\right)}{\mathbb{E}\left(e^{\rho_B W_B(s)}Y_B(s)\right)}-\left(\frac{\mathbb{E}\left(e^{\rho_B W_B(s)}Y_B(s)W_B(s)\right)}{\mathbb{E}\left(e^{\rho_B W_B(s)}Y_B(s)\right)}\right)^2\end{pmatrix}.$$

Thus,

$$x^T\Sigma_\tau x=\sum_{h=A,B}\int_0^\tau \mathcal{V}_h(\rho_{h,0},s)x_h^2 S_h(\rho_0,s)\alpha_{h,0}(s)ds,$$

where

$$\begin{aligned}\mathcal{V}_h(\rho_h,s)&=\frac{\mathbb{E}\left(e^{\rho_h W_h(s)}Y_h(s)W_h(s)^2\right)}{\mathbb{E}\left(e^{\rho_h W_h(s)}Y_h(s)\right)}-\left(\frac{\mathbb{E}\left(e^{\rho_h W_h(s)}Y_h(s)W_h(s)\right)}{\mathbb{E}\left(e^{\rho_h W_h(s)}Y_h(s)\right)}\right)^2\\ &=\mathbb{E}\left(\frac{e^{\rho_h W_h(s)}Y_h(s)}{\mathbb{E}\left(e^{\rho_h W_h(s)}Y_h(s)\right)}\left(W_h(s)-\mathbb{E}\left(\frac{e^{\rho_h W_h(s)}Y_h(s)}{\mathbb{E}\left(e^{\rho_h W_h(s)}Y_h(s)\right)}W_h(s)\right)\right)^2\right).\end{aligned}$$

Therefore, $\mathcal{V}_h(\rho_h,s)$ is the variance of $W_h(s)$ with respect to the probability measure $dP_{h,s}$ defined by

$$dP_{h,s}=\frac{e^{\rho_{h,0}W_h(s)}Y_h(s)}{\mathbb{E}\left(e^{\rho_{h,0}W_h(s)}Y_h(s)\right)}dP,$$

where dP is the reference probability measure. Thus, $\mathcal{V}_h(\rho_h,s)$ is clearly positive. The equation $S_h(\rho_0,s)=0,\ s\in[0,\tau]$ is equivalent to $Y_h(s)=0$ a.s., $s\in[0,\tau]$, which cannot be true because of iii). From Assumption 1, we deduce that the integrals $\int_0^\tau \mathcal{V}_h(\rho_{h,0},s)S_h(\rho_0,s)\alpha_{h,0}(s)ds$ cannot equal zero, so that we necessarily have $x_h=0,\ h=A,B$, and the matrix Σ_τ is positive definite.

Acknowledgement

The author would like to thank Eva Lelièvre and Valérie Golaz (INED) for giving access to the demographic data set presented in this paper.

References

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

Andersen, P. K., Gill, R. D., 1982. Cox's regression model for counting processes: a large sample study. Ann. Statist. 10 (4), 1100–1120.

Cox, D. R., 1972. Regression models and life-tables. J. Roy. Statist. Soc. Ser. B 34, 187–220, with discussion by F. Downton, Richard Peto, D. J. Bartholomew, D. V. Lindley, P. W. Glassborow, D. E. Barton, Susannah Howard, B. Benjamin, John J. Gart, L. D. Meshalkin, A. R. Kagan, M. Zelen, R. E. Barlow, Jack Kalbfleisch, R. L. Prentice and Norman Breslow, and a reply by D. R. Cox.

Hougaard, P., 2000. Analysis of multivariate survival data. Springer, New York.

Jacod, J., 1973. On the stochastic intensity of a random point process over the half line. Technical report 15, Department of Statistics, Princeton University.

Jacod, J., 1974/75. Multivariate point processes: predictable projection, Radon-Nikodým derivatives, representation of martingales. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 31, 235–253.

Kalbfleisch, J. D., Prentice, R. L., 1980. The statistical analysis of failure time data. John Wiley and Sons, New York-Chichester-Brisbane, wiley Series in Probability and Mathematical Statistics.

Lelièvre, E., Vivier, G., 2001. Évaluation d'une collecte à la croisée du quantitatif et du qualitatif, l'enquête " Biographies et entourage ". Population 56 (6), 1043–1074.

Letué, F., 2000a. Modèle de Cox : estimation par sélection de modèle et modèle de chocs bivarié. Thèse 6414, Université de Paris-Sud, http://ljk.imag.fr/membres/Frederique.Letue/.

Letué, F., 2000b. A semiparametric shock model in censoring bivariate survival analysis. Preprint 00.63, Université de Paris-Sud, http://www.math.u-psud.fr/ biblio/ppo/2000/ppo2000-63.html.