Dynamic survival analysis: modelling the hazard function via ordinary differential equations

Francisco Javier Rubio

University College London
Department of Statistical Science
Joint work with J. Andres Christen

March 2025

Overview

Motivation

Example 1: Logistic ODE

Example 2: Hazard-Response model

Real data application

Discussion and Software tools

Survival analysis

- Survival analysis methods have been applied in a number of areas, including medicine, epidemiology, genetics, engineering, and biology, to name but a few.

Survival analysis

- Survival analysis methods have been applied in a number of areas, including medicine, epidemiology, genetics, engineering, and biology, to name but a few.
- The **survival function** and the **hazard function** represent two quantities of interest in this area.

Survival analysis

- Survival analysis methods have been applied in a number of areas, including medicine, epidemiology, genetics, engineering, and biology, to name but a few.
- The **survival function** and the **hazard function** represent two quantities of interest in this area.
- The survival function provides information about the probability that an individual or population will survive beyond a certain time point: $S(t) = P(T > t)$.

Survival analysis

- Survival analysis methods have been applied in a number of areas, including medicine, epidemiology, genetics, engineering, and biology, to name but a few.
- The **survival function** and the **hazard function** represent two quantities of interest in this area.
- The survival function provides information about the probability that an individual or population will survive beyond a certain time point: $S(t) = P(T > t)$.
- Mathematically, this is interpreted in terms of the hazard function

$$h(t) = \lim_{dt \to 0} \frac{P[t \leq T < t + dt \mid T \geq t]}{dt} = \frac{f_T(t)}{S_T(t)}.$$

Survival analysis

- Survival analysis methods have been applied in a number of areas, including medicine, epidemiology, genetics, engineering, and biology, to name but a few.
- The **survival function** and the **hazard function** represent two quantities of interest in this area.
- The survival function provides information about the probability that an individual or population will survive beyond a certain time point: $S(t) = P(T > t)$.
- Mathematically, this is interpreted in terms of the hazard function

$$h(t) = \lim_{dt \to 0} \frac{P[t \leq T < t + dt \mid T \geq t]}{dt} = \frac{f_T(t)}{S_T(t)}.$$

- Appeal: qualitative and quantitative analyses.

Survival analysis

- Survival analysis methods have been applied in a number of areas, including medicine, epidemiology, genetics, engineering, and biology, to name but a few.
- The **survival function** and the **hazard function** represent two quantities of interest in this area.
- The survival function provides information about the probability that an individual or population will survive beyond a certain time point: $S(t) = P(T > t)$.
- Mathematically, this is interpreted in terms of the hazard function

$$h(t) = \lim_{dt \to 0} \frac{P[t \leq T < t + dt \mid T \geq t]}{dt} = \frac{f_T(t)}{S_T(t)}.$$

- Appeal: qualitative and quantitative analyses.
- Typically of interest for deeper analyses only.

Survival analysis

- Survival analysis methods have been applied in a number of areas, including medicine, epidemiology, genetics, engineering, and biology, to name but a few.
- The **survival function** and the **hazard function** represent two quantities of interest in this area.
- The survival function provides information about the probability that an individual or population will survive beyond a certain time point: $S(t) = P(T > t)$.
- Mathematically, this is interpreted in terms of the hazard function

$$h(t) = \lim_{dt \to 0} \frac{P[t \leq T < t + dt \mid T \geq t]}{dt} = \frac{f_T(t)}{S_T(t)}.$$

- Appeal: qualitative and quantitative analyses.
- Typically of interest for deeper analyses only. If the user only wants to report survival probabilities, they would hardly look at the hazard function.

Estimating the hazard function

- Parametric: use a distribution function + estimation method.

Estimating the hazard function

- Parametric: use a distribution function + estimation method.
- (Semi-)Parametric: splines for the log-hazard + estimation method.

Estimating the hazard function

- Parametric: use a distribution function + estimation method.
- (Semi-)Parametric: splines for the log-hazard + estimation method.
- Non-parametric: KDE-based, Bayesian non-parametrics.

Estimating the hazard function

- Parametric: use a distribution function + estimation method.
- (Semi-)Parametric: splines for the log-hazard + estimation method.
- Non-parametric: KDE-based, Bayesian non-parametrics. Flexible, harder to implement with limited software, wiggly (challenging interpretation).

Estimating the hazard function

- Parametric: use a distribution function + estimation method.
- (Semi-)Parametric: splines for the log-hazard + estimation method.
- Non-parametric: KDE-based, Bayesian non-parametrics. Flexible, harder to implement with limited software, wiggly (challenging interpretation).
- Machine learning?

Aims

- To propose a novel approach for **parametrically** modelling the dynamics of the hazard function using systems of first order **ODEs** (whose solution is a positive function).

Aims

- To propose a novel approach for **parametrically** modelling the dynamics of the hazard function using systems of first order **ODEs** (whose solution is a positive function).
- This framework is of particular interest to users aiming at **understanding and interpreting** the evolution of the hazard function over time.

Aims

- To propose a novel approach for **parametrically** modelling the dynamics of the hazard function using systems of first order **ODEs** (whose solution is a positive function).
- This framework is of particular interest to users aiming at **understanding and interpreting** the evolution of the hazard function over time.
- We will focus on the context without covariates, but we will conclude with a discussion on general strategies for the inclusion of covariates.

Notation

- Let $\mathbf{o} = \{o_1, \ldots, o_n\}$ be a sequence of survival times, $c_i \in \mathbb{R}_+$ be the corresponding right-censoring times, $t_i = \min\{o_i, c_i\}$ be the observed times, and $\delta_i = \mathrm{I}(o_i \leq c_i)$ be the indicator that observation i is uncensored, $i = 1, \ldots, n$.

Notation

- Let $\mathbf{o} = \{o_1, \ldots, o_n\}$ be a sequence of survival times, $c_i \in \mathbb{R}_+$ be the corresponding right-censoring times, $t_i = \min\{o_i, c_i\}$ be the observed times, and $\delta_i = \mathrm{I}(o_i \leq c_i)$ be the indicator that observation i is uncensored, $i = 1, \ldots, n$.
- $S(t) = 1 - F(t)$ is the survival function. f is the pdf, F is the CDF, h is the hazard function, and $H(t) = \int_0^t h(r)dr$ is the cumulative hazard function.

Proposed formulation

Let $q_j : \mathbb{R}^+ \to \mathbb{R}$, $j = 1, \ldots, m$, be a collection of differentiable functions, and let us denote $\mathbf{Y}(t) = (h(t), q_1(t), \ldots, q_m(t))^\top$, $t > 0$.

Proposed formulation

Let $q_j : \mathbb{R}^+ \to \mathbb{R}$, $j = 1, \ldots, m$, be a collection of differentiable functions, and let us denote $\mathbf{Y}(t) = (h(t), q_1(t), \ldots, q_m(t))^\top$, $t > 0$. Define the system of ODEs

$$\begin{cases} \mathbf{Y}'(t) = \psi_{\boldsymbol{\theta}}(\mathbf{Y}(t), t), \\ H'(t) = Y_1(t), \end{cases} \tag{1}$$

with initial conditions $\mathbf{Y}(0) = \mathbf{Y}_0$ and $H(0) = 0$.

Proposed formulation

Let $q_j : \mathbb{R}^+ \to \mathbb{R}$, $j = 1, \ldots, m$, be a collection of differentiable functions, and let us denote $\mathbf{Y}(t) = (h(t), q_1(t), \ldots, q_m(t))^\top$, $t > 0$. Define the system of ODEs

$$\begin{cases} \mathbf{Y}'(t) = \psi_{\boldsymbol{\theta}}(\mathbf{Y}(t), t), \\ H'(t) = Y_1(t), \end{cases} \tag{1}$$

with initial conditions $\mathbf{Y}(0) = \mathbf{Y}_0$ and $H(0) = 0$.

The vector field $\psi_{\boldsymbol{\theta}}$, its domain D, and the initial condition $\mathbf{Y}_0$ are such that the solution for $h(t)$ is positive for all $t > 0$, and for any parameter value $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^d$.

Details of the proposed formulation

- Equation (1) is a family of systems of ODEs for which one state variable is the hazard function h, leading to a family of hazard functions defined by θ and $\mathbf{Y}_0$.

Details of the proposed formulation

- Equation (1) is a family of systems of ODEs for which one state variable is the hazard function h, leading to a family of hazard functions defined by θ and $\mathbf{Y}_0$.
- The cumulative hazard function H is also included in the formulation (1), and is thus obtained in the solution of the system.

Details of the proposed formulation

- Equation (1) is a family of systems of ODEs for which one state variable is the hazard function h, leading to a family of hazard functions defined by θ and $\mathbf{Y}_0$.
- The cumulative hazard function H is also included in the formulation (1), and is thus obtained in the solution of the system.
- The initial conditions in (1) could include $h(0) = h_0 > 0$ indicating that the hazard function takes a non-negative, **finite**, value at $t = 0$, and that all individuals are alive at the start of the follow-up.

Details of the proposed formulation

- Equation (1) is a family of systems of ODEs for which one state variable is the hazard function h, leading to a family of hazard functions defined by θ and $\mathbf{Y}_0$.
- The cumulative hazard function H is also included in the formulation (1), and is thus obtained in the solution of the system.
- The initial conditions in (1) could include $h(0) = h_0 > 0$ indicating that the hazard function takes a non-negative, **finite**, value at $t = 0$, and that all individuals are alive at the start of the follow-up.
- The solution may be available in closed form (analytic solution) or may be obtained using an ODE Solver (vast literature on numerical methods).

Related works

- Tang et al. [2022, 2023] proposed modelling the cumulative hazard function via an ODE. That is,

$$\begin{cases} H'(t;\mathbf{x}) & = & \Psi(H(t;\mathbf{x}), t; \mathbf{x}), \\ H(t_0;\mathbf{x}) & = & c(\mathbf{x}), \end{cases} \tag{2}$$

  where $\mathbf{x} \in \mathbb{R}^p$ are the available covariates.

Related works

- Tang et al. [2022, 2023] proposed modelling the cumulative hazard function via an ODE. That is,

$$\begin{cases} H'(t;\mathbf{x}) & = & \Psi(H(t;\mathbf{x}), t;\mathbf{x}), \\ H(t_0;\mathbf{x}) & = & c(\mathbf{x}), \end{cases} \tag{2}$$

  where $\mathbf{x} \in \mathbb{R}^p$ are the available covariates. This formulation aims at modelling the dynamics of the cumulative hazard function, which implicitly means modelling the hazard function $h(t;\mathbf{x}) = H'(t;\mathbf{x})$, rather than obtaining it as a solution.

Related works

- Tang et al. [2022, 2023] proposed modelling the cumulative hazard function via an ODE. That is,

$$\begin{cases} H'(t;\mathbf{x}) & = & \Psi(H(t;\mathbf{x}), t; \mathbf{x}), \\ H(t_0;\mathbf{x}) & = & c(\mathbf{x}), \end{cases} \tag{2}$$

  where $\mathbf{x} \in \mathbb{R}^p$ are the available covariates. This formulation aims at modelling the dynamics of the cumulative hazard function, which implicitly means modelling the hazard function $h(t;\mathbf{x}) = H'(t;\mathbf{x})$, rather than obtaining it as a solution.
- The formulation (2) is not seen as an approach to model the (cumulative) hazard function, but simply as a device to use a neural network [Tang et al., 2022] or a semiparametric method [Tang et al., 2023] to construct it.

Example 1: Logistic ODE

- The logistic growth hazard model is defined through the following system of ODEs,

$$\begin{cases} h'(t) = \lambda h(t)\left(1 - \dfrac{h(t)}{\kappa}\right), & h(0) = h_0 \\ H'(t) = h(t), & H(0) = 0. \end{cases} \tag{3}$$

Example 1: Logistic ODE

- The logistic growth hazard model is defined through the following system of ODEs,

$$\begin{cases} h'(t) = \lambda h(t)\left(1 - \dfrac{h(t)}{\kappa}\right), & h(0) = h_0 \\ H'(t) = h(t), & H(0) = 0. \end{cases} \tag{3}$$

- $\lambda > 0$ represents the intrinsic growth rate of the hazard function, $\kappa > 0$ represents the upper or lower bound of the hazard function, and $h_0 > 0$ is the value of the hazard function at $t = 0$.

Analytic solution

- This ODE has the following analytic solution

$$h(t \mid \lambda, \kappa, h_0) = \frac{\kappa h_0 e^{\lambda t}}{\kappa + h_0(e^{\lambda t} - 1)}.$$

The solution $h(t)$, $t \in [0, \infty)$ is positive and finite, and $\lim_{t\to\infty} h(t) = \kappa$.

Analytic solution

- This ODE has the following analytic solution

$$h(t \mid \lambda, \kappa, h_0) = \frac{\kappa h_0 e^{\lambda t}}{\kappa + h_0(e^{\lambda t} - 1)}.$$

The solution $h(t)$, $t \in [0, \infty)$ is positive and finite, and $\lim_{t\to\infty} h(t) = \kappa$.

- The corresponding cumulative hazard function is

$$H(t \mid \lambda, \kappa, h_0) = \frac{\kappa}{\lambda} \log\left(\frac{\kappa + h_0\left(e^{\lambda t} - 1\right)}{\kappa}\right).$$

Analytic solution

- This ODE has the following analytic solution

$$h(t \mid \lambda, \kappa, h_0) = \frac{\kappa h_0 e^{\lambda t}}{\kappa + h_0(e^{\lambda t} - 1)}.$$

  The solution $h(t)$, $t \in [0, \infty)$ is positive and finite, and $\lim_{t\to\infty} h(t) = \kappa$.

- The corresponding cumulative hazard function is

$$H(t \mid \lambda, \kappa, h_0) = \frac{\kappa}{\lambda} \log\left(\frac{\kappa + h_0\left(e^{\lambda t} - 1\right)}{\kappa}\right).$$

- h_0 can be estimated as a parameter.

Example 2: Hazard-Response model

- The hazard-response model is defined through the system of ODEs:

$$\begin{cases} h'(t) = \lambda h(t)\left(1 - \frac{h(t)}{\kappa}\right) - \alpha q(t)h(t), & h(0) = h_0 \\ q'(t) = \beta q(t)\left(1 - \frac{q(t)}{\kappa}\right) - \alpha q(t)h(t), & q(0) = q_0 \\ H'(t) = h(t), & H(0) = 0, \end{cases} \tag{4}$$

Example 2: Hazard-Response model

- The hazard-response model is defined through the system of ODEs:

$$\begin{cases} h'(t) = \lambda h(t)\left(1 - \frac{h(t)}{\kappa}\right) - \alpha q(t)h(t), & h(0) = h_0 \\ q'(t) = \beta q(t)\left(1 - \frac{q(t)}{\kappa}\right) - \alpha q(t)h(t), & q(0) = q_0 \\ H'(t) = h(t), & H(0) = 0, \end{cases} \tag{4}$$

- with $\lambda > 0$, $\alpha \geq 0$, $\beta > 0$, $\kappa > 0$, $h_0 > 0$, and $q_0 > 0$.

Example 2: Hazard-Response model

- The hazard-response model is defined through the system of ODEs:

$$\begin{cases} h'(t) = \lambda h(t)\left(1 - \frac{h(t)}{\kappa}\right) - \alpha q(t)h(t), & h(0) = h_0 \\ q'(t) = \beta q(t)\left(1 - \frac{q(t)}{\kappa}\right) - \alpha q(t)h(t), & q(0) = q_0 \\ H'(t) = h(t), & H(0) = 0, \end{cases} \tag{4}$$

- with $\lambda > 0$, $\alpha \geq 0$, $\beta > 0$, $\kappa > 0$, $h_0 > 0$, and $q_0 > 0$.
- Similar to the competitive Lotka-Volterra equations. No analytical solution.

Attractors

- Consider the values

$$h^* = \kappa \left(\frac{1 - \alpha\kappa\lambda^{-1}}{D} \right) \quad \text{and} \quad q^* = \kappa \left(\frac{1 - \alpha\kappa\beta^{-1}}{D} \right),$$

with $D = 1 - \frac{(\alpha\kappa)^2}{\lambda\beta}$.

Attractors

- Consider the values

$$h^* = \kappa \left(\frac{1 - \alpha\kappa\lambda^{-1}}{D} \right) \quad \text{and} \quad q^* = \kappa \left(\frac{1 - \alpha\kappa\beta^{-1}}{D} \right),$$

  with $D = 1 - \frac{(\alpha\kappa)^2}{\lambda\beta}$.

- If $q^* < 0$, then the hazard $h(t)$ reaches its carrying capacity as $t \to \infty$ and the response $q(t)$ “losses” the competition ($q(t) \to 0$).

Attractors

- Consider the values

$$h^* = \kappa \left(\frac{1 - \alpha\kappa\lambda^{-1}}{D} \right) \quad \text{and} \quad q^* = \kappa \left(\frac{1 - \alpha\kappa\beta^{-1}}{D} \right),$$

  with $D = 1 - \frac{(\alpha\kappa)^2}{\lambda\beta}$.
- If $q^* < 0$, then the hazard $h(t)$ reaches its carrying capacity as $t \to \infty$ and the response $q(t)$ "losses" the competition ($q(t) \to 0$).
- If $h^* < 0$, then the response $q(t)$ "wins" the competition ($q(t) \to \kappa$) and the hazard "losses" the competition ($h(t) \to 0$).

Attractors

- Consider the values

$$h^* = \kappa \left(\frac{1 - \alpha\kappa\lambda^{-1}}{D} \right) \quad \text{and} \quad q^* = \kappa \left(\frac{1 - \alpha\kappa\beta^{-1}}{D} \right),$$

  with $D = 1 - \frac{(\alpha\kappa)^2}{\lambda\beta}$.
- If $q^* < 0$, then the hazard $h(t)$ reaches its carrying capacity as $t \to \infty$ and the response $q(t)$ “losses” the competition ($q(t) \to 0$).
- If $h^* < 0$, then the response $q(t)$ “wins” the competition ($q(t) \to \kappa$) and the hazard “losses” the competition ($h(t) \to 0$).
- If both $h^* > 0$ and $q^* > 0$, the hazard remains positive ($h(t) \to h^* > 0$), although not at its maximum (or minimum), but in an equilibrium with the response ($h^* < \kappa$).

Attractors

- Consider the values

$$h^* = \kappa \left(\frac{1 - \alpha\kappa\lambda^{-1}}{D} \right) \quad \text{and} \quad q^* = \kappa \left(\frac{1 - \alpha\kappa\beta^{-1}}{D} \right),$$

  with $D = 1 - \frac{(\alpha\kappa)^2}{\lambda\beta}$.
- If $q^* < 0$, then the hazard $h(t)$ reaches its carrying capacity as $t \to \infty$ and the response $q(t)$ "losses" the competition ($q(t) \to 0$).
- If $h^* < 0$, then the response $q(t)$ "wins" the competition ($q(t) \to \kappa$) and the hazard "losses" the competition ($h(t) \to 0$).
- If both $h^* > 0$ and $q^* > 0$, the hazard remains positive ($h(t) \to h^* > 0$), although not at its maximum (or minimum), but in an equilibrium with the response ($h^* < \kappa$).
- No oscillations (possible for a more general version).

Real data application: Data

- We analyse the `rotterdam` data set [Royston and Altman, 2013] from the `survival` R package.

Real data application: Data

- We analyse the `rotterdam` data set [Royston and Altman, 2013] from the `survival` R package.
- This data set contains information about the survival of $n = 2982$ breast cancer patients, from which 1272 died within the maximum follow-up period (19.3 years).

Real data application: Data

- We analyse the `rotterdam` data set [Royston and Altman, 2013] from the `survival` R package.
- This data set contains information about the survival of $n = 2982$ breast cancer patients, from which 1272 died within the maximum follow-up period (19.3 years).
- It is known that some of these patients received hormonal treatment, chemotherapy, and/or surgical treatment.

Model

- We expect that the evolution of the population hazard function over time to depend on the response to the treatments and the natural immunological response. Thus, we model the survival times using the hazard-response model (4).

Model

- We expect that the evolution of the population hazard function over time to depend on the response to the treatments and the natural immunological response. Thus, we model the survival times using the hazard-response model (4).
- The minimum survival time in this data set is 45 days, thus, we do not expect to have information about the initial conditions h_0 and q_0.

Model

- We expect that the evolution of the population hazard function over time to depend on the response to the treatments and the natural immunological response. Thus, we model the survival times using the hazard-response model (4).
- The minimum survival time in this data set is 45 days, thus, we do not expect to have information about the initial conditions h_0 and q_0.
- We have verified this assumption by fitting the model where the initial conditions are assumed to be unknown parameters, and we found that the posteriors are virtually the same as the priors, indicating weak identifiability [Cole, 2020] of h_0 and q_0.

Likelihood function

Since the proposed hazard models are parametric, the log-likelihood function can be written in terms of the hazard function and the cumulative hazard function, as usual:

$$\ell(\boldsymbol{\theta}, \mathbf{Y}_0) = \sum_{i=1}^{n} \delta_i \log h(t_i \mid \boldsymbol{\theta}, \mathbf{Y}_0) - \sum_{i=1}^{n} H(t_i \mid \boldsymbol{\theta}, \mathbf{Y}_0),$$

(MLE also feasible)

Priors

We have used generic weakly informative priors,

- For all positive parameters, we adopt gamma priors with scale parameter 2 and shape parameter 2.
- This prior has mean 4, variance 8, it accumulates 95% of the probability in the interval $(0.5, 11.2)$, and it vanishes at 0 (which helps repelling the MCMC samplers from visiting regions near zero, that may cause numerical problems).

Initial conditions

- It is known that the prognosis of breast cancer is relatively good compared to other cancers, and that the mortality rate during the first few months is very low.

Initial conditions

- It is known that the prognosis of breast cancer is relatively good compared to other cancers, and that the mortality rate during the first few months is very low.
- We assume that the survival probability at one month $\Delta t = 1/12$ is $S(\Delta t) \approx 0.999$.

Initial conditions

- It is known that the prognosis of breast cancer is relatively good compared to other cancers, and that the mortality rate during the first few months is very low.
- We assume that the survival probability at one month $\Delta t = 1/12$ is $S(\Delta t) \approx 0.999$.
- Then, we use the approximation

$$h_0 = h(0) = -\frac{S'(0)}{S(0)} \approx -\frac{S'(\Delta t)}{S(\Delta t)} \approx -\frac{S(\Delta t) - S(0)}{\Delta t S(\Delta t)} \approx 0.01.$$

Initial conditions

- It is known that the prognosis of breast cancer is relatively good compared to other cancers, and that the mortality rate during the first few months is very low.
- We assume that the survival probability at one month $\Delta t = 1/12$ is $S(\Delta t) \approx 0.999$.
- Then, we use the approximation

$$h_0 = h(0) = -\frac{S'(0)}{S(0)} \approx -\frac{S'(\Delta t)}{S(\Delta t)} \approx -\frac{S(\Delta t) - S(0)}{\Delta t S(\Delta t)} \approx 0.01.$$

- To define the initial condition on q, we use that treatment does not usually start at the beginning of follow-up.

Initial conditions

- It is known that the prognosis of breast cancer is relatively good compared to other cancers, and that the mortality rate during the first few months is very low.
- We assume that the survival probability at one month $\Delta t = 1/12$ is $S(\Delta t) \approx 0.999$.
- Then, we use the approximation

$$h_0 = h(0) = -\frac{S'(0)}{S(0)} \approx -\frac{S'(\Delta t)}{S(\Delta t)} \approx -\frac{S(\Delta t) - S(0)}{\Delta t S(\Delta t)} \approx 0.01.$$

- To define the initial condition on q, we use that treatment does not usually start at the beginning of follow-up. Thus, the response on reducing the hazard function should be small at the beginning of follow-up, and we fix this value at $q_0 = 10^{-6}$.

Computations

- Implementation in R and Python, using the Livermore Solver for Ordinary Differential Equations (LSODE). `deSolve`

Computations

- Implementation in R and Python, using the Livermore Solver for Ordinary Differential Equations (LSODE). `deSolve`
- The Jacobian used in the implementation of LSODE is calculated explicitly (see the Appendix) for a better performance of the ODE solver.

Computations

- Implementation in R and Python, using the Livermore Solver for Ordinary Differential Equations (LSODE). `deSolve`
- The Jacobian used in the implementation of LSODE is calculated explicitly (see the Appendix) for a better performance of the ODE solver.
- The implementation was done in terms of $\log h(t)$ for a better performance of the solver.

Computations

- Implementation in R and Python, using the Livermore Solver for Ordinary Differential Equations (LSODE). `deSolve`
- The Jacobian used in the implementation of LSODE is calculated explicitly (see the Appendix) for a better performance of the ODE solver.
- The implementation was done in terms of $\log h(t)$ for a better performance of the solver.
- Posterior samples were obtained using the Python library `twalk`, and the R package `spBayes` (adaptive Metropolis within Gibbs). Several initial points, including the MLE.

Computations

- Implementation in R and Python, using the Livermore Solver for Ordinary Differential Equations (LSODE). `deSolve`
- The Jacobian used in the implementation of LSODE is calculated explicitly (see the Appendix) for a better performance of the ODE solver.
- The implementation was done in terms of $\log h(t)$ for a better performance of the solver.
- Posterior samples were obtained using the Python library `twalk`, and the R package `spBayes` (adaptive Metropolis within Gibbs). Several initial points, including the MLE.
- Posterior sample size 1000, with thinning 100 and burn-in $10,000$.

Posterior predictive hazard and response

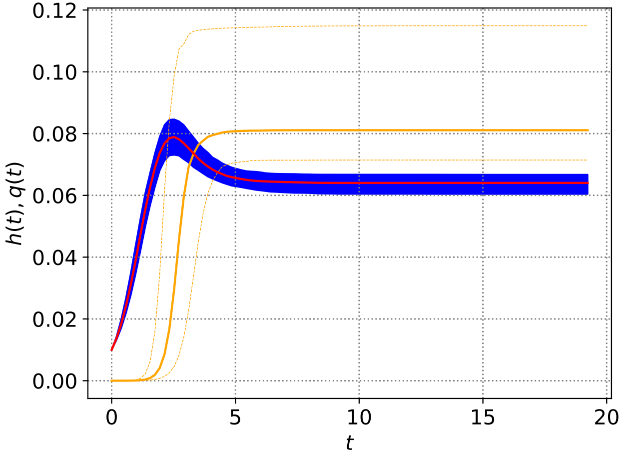

Figure: `rotterdam` data: posterior predictive hazard function (red line) and response function (yellow line).

Posterior predictive survival

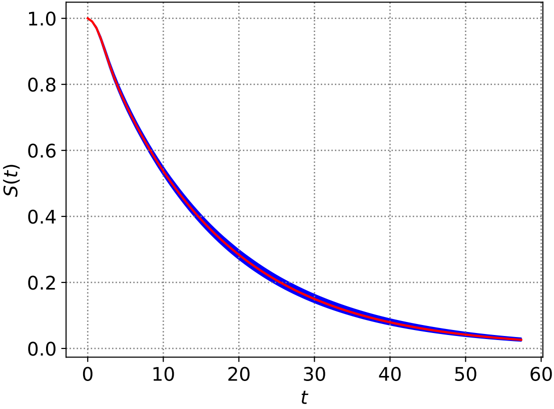

Figure: `rotterdam` data: survival function (red line) for the Hazard-Response model (4), 0.1 to 0.9 quantiles and the median.

Posterior of h^*

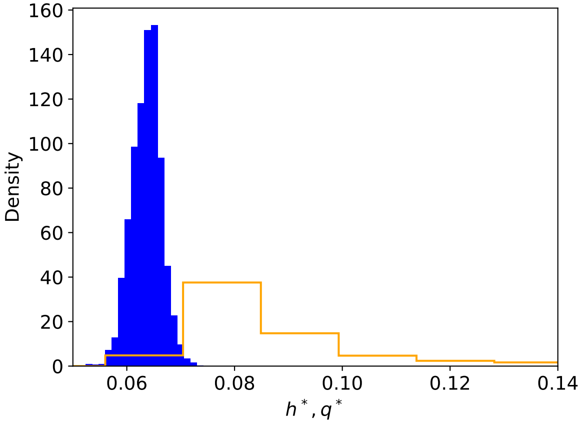

Figure: `rotterdam` data: Posterior distribution of h^* (left) and q^* (right). The system is in equilibrium with probability close to 1, and h tends to the asymptotic point h^*.

Interpretation

- Note how the hazard function exhibits an increasing behaviour during the first 2 to 3 years and then decreases and stabilises at a constant value.

Interpretation

- Note how the hazard function exhibits an increasing behaviour during the first 2 to 3 years and then decreases and stabilises at a constant value.
- Interestingly, the equilibrium state is inferred and neither h nor q go to zero.

Interpretation

- Note how the hazard function exhibits an increasing behaviour during the first 2 to 3 years and then decreases and stabilises at a constant value.
- Interestingly, the equilibrium state is inferred and neither h nor q go to zero.
- That is, a constant asymptotic hazard h^* is predicted for this data set.

Interpretation

- Note how the hazard function exhibits an increasing behaviour during the first 2 to 3 years and then decreases and stabilises at a constant value.
- Interestingly, the equilibrium state is inferred and neither h nor q go to zero.
- That is, a constant asymptotic hazard h^* is predicted for this data set.
- This coincides nicely with previous studies on the evolution of the population hazard function associated with breast cancer patients [Royston and Parmar, 2002].

Discussion

- We proposed a novel methodology for modelling the dynamics of the hazard function in survival analysis via systems of ODEs.

Discussion

- We proposed a novel methodology for modelling the dynamics of the hazard function in survival analysis via systems of ODEs.
- This framework is particularly useful for researchers interested in qualitative and quantitative analyses of the evolution of the hazard function over time.

Discussion

- We proposed a novel methodology for modelling the dynamics of the hazard function in survival analysis via systems of ODEs.
- This framework is particularly useful for researchers interested in qualitative and quantitative analyses of the evolution of the hazard function over time.
- R, Python, and Julia (`DifferentialEquations.jl`, `Turing.jl`) code and data can be found at: `https://github.com/FJRubio67/ODESurv`

Discussion

- We proposed a novel methodology for modelling the dynamics of the hazard function in survival analysis via systems of ODEs.
- This framework is particularly useful for researchers interested in qualitative and quantitative analyses of the evolution of the hazard function over time.
- R, Python, and Julia (`DifferentialEquations.jl`, `Turing.jl`) code and data can be found at: `https://github.com/FJRubio67/ODESurv`
- Extensions: others ODE systems, delay equations, integro-differential systems, second order ODEs [Christen and Rubio, 2025].

Adding covariates (in progress)

- One could use the solution to (1) as a baseline hazard in any hazard-based regression model. This includes the proportional hazards, accelerated failure time, accelerated hazards, or general/extended hazards models [Rubio et al., 2019]. ✗

Adding covariates (in progress)

- One could use the solution to (1) as a baseline hazard in any hazard-based regression model. This includes the proportional hazards, accelerated failure time, accelerated hazards, or general/extended hazards models [Rubio et al., 2019]. ✗
- Another approach, that takes advantage of the interpretability of the model parameters, consists of modelling the parameters using a transformed linear (additive) predictor.
  $\varphi_i(\theta_k) = \beta_{k0} + \mathbf{x}_{ki}^{\top}\boldsymbol{\beta}_k,$

Adding covariates (in progress)

- One could use the solution to (1) as a baseline hazard in any hazard-based regression model. This includes the proportional hazards, accelerated failure time, accelerated hazards, or general/extended hazards models [Rubio et al., 2019]. ✗
- Another approach, that takes advantage of the interpretability of the model parameters, consists of modelling the parameters using a transformed linear (additive) predictor. $\varphi_i(\theta_k) = \beta_{k0} + \mathbf{x}_{ki}^\top \boldsymbol{\beta}_k,$
- Requires building a model.

Adding covariates (in progress)

- One could use the solution to (1) as a baseline hazard in any hazard-based regression model. This includes the proportional hazards, accelerated failure time, accelerated hazards, or general/extended hazards models [Rubio et al., 2019]. ✗
- Another approach, that takes advantage of the interpretability of the model parameters, consists of modelling the parameters using a transformed linear (additive) predictor.
  $\varphi_i(\theta_k) = \beta_{k0} + \mathbf{x}_{ki}^\top \boldsymbol{\beta}_k,$
- Requires building a model. Evaluating the likelihood function requires solving one ODE system per individual.

Adding covariates (in progress)

- One could use the solution to (1) as a baseline hazard in any hazard-based regression model. This includes the proportional hazards, accelerated failure time, accelerated hazards, or general/extended hazards models [Rubio et al., 2019]. ✗
- Another approach, that takes advantage of the interpretability of the model parameters, consists of modelling the parameters using a transformed linear (additive) predictor. $\varphi_i(\theta_k) = \beta_{k0} + \mathbf{x}_{ki}^\top \boldsymbol{\beta}_k$,
- Requires building a model. Evaluating the likelihood function requires solving one ODE system per individual.
- Possible solutions: parallel computing / multi-threading (embarrasingly parallelisable, GPU-accelerated ODE embeddings),

Adding covariates (in progress)

- One could use the solution to (1) as a baseline hazard in any hazard-based regression model. This includes the proportional hazards, accelerated failure time, accelerated hazards, or general/extended hazards models [Rubio et al., 2019]. ✗
- Another approach, that takes advantage of the interpretability of the model parameters, consists of modelling the parameters using a transformed linear (additive) predictor. $\varphi_i(\theta_k) = \beta_{k0} + \mathbf{x}_{ki}^\top \boldsymbol{\beta}_k$,
- Requires building a model. Evaluating the likelihood function requires solving one ODE system per individual.
- Possible solutions: parallel computing / multi-threading (embarrasingly parallelisable, GPU-accelerated ODE embeddings), or a fast approximation (Gaussian, Skew-Normal, Amortised inference, ...).

Adding covariates (in progress)

- One could use the solution to (1) as a baseline hazard in any hazard-based regression model. This includes the proportional hazards, accelerated failure time, accelerated hazards, or general/extended hazards models [Rubio et al., 2019]. ✗
- Another approach, that takes advantage of the interpretability of the model parameters, consists of modelling the parameters using a transformed linear (additive) predictor. $\varphi_i(\theta_k) = \beta_{k0} + \mathbf{x}_{ki}^\top \boldsymbol{\beta}_k,$
- Requires building a model. Evaluating the likelihood function requires solving one ODE system per individual.
- Possible solutions: parallel computing / multi-threading (embarrasingly parallelisable, GPU-accelerated ODE embeddings), or a fast approximation (Gaussian, Skew-Normal, Amortised inference, ...).
- Cliché: Allowing extensions to other models of interest in survival analysis such as cure models, spatial survival models, relative survival models, and etcetera.

References

J.A. Christen and F.J. Rubio. Dynamic survival analysis: modelling the hazard function via ordinary differential equations. *Statistical Methods in Medical Research*, 33(10): 1768–1782, 2024.

J.A. Christen and F.J. Rubio. On harmonic oscillator hazard functions. *Statistics & Probability Letters*, 217:110304, 2025.

D. Cole. *Parameter Redundancy and Identifiability*. CRC Press, Boca Raton, FL, 2020.

P. Royston and D.G. Altman. External validation of a cox prognostic model: principles and methods. *BMC Medical Research Methodology*, 13:1–15, 2013.

P. Royston and M.K.B. Parmar. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. *Statistics in Medicine*, 21(15): 2175–2197, 2002.

F.J. Rubio, L. Remontet, N.P. Jewell, and A. Belot. On a general structure for hazard-based regression models: an application to population-based cancer research. *Statistical Methods in Medical Research*, 28:2404–2417, 2019.

W. Tang, J. Ma, Q. Mei, and J. Zhu. SODEN: A scalable continuous-time survival model through ordinary differential equation networks. *Journal of Machine Learning Research*, 23(34):1–29, 2022.

W. Tang, K. He, G. Xu, and J. Zhu. Survival analysis via ordinary differential equations. *Journal of the American Statistical Association*, NA:1–16, 2023.