Bayesian survival analysis with Stan

Danilo Alvares

MRC Biostatistics Unit, University of Cambridge

Reference I

Received: 11 November 2019 | Revised: 18 January 2021 | Accepted: 13 February 2021

DOI: 10.1002/sim.8933

TUTORIAL IN BIOSTATISTICS

Statistics in Medicine WILEY

Bayesian survival analysis with BUGS

Danilo Alvares[1] | Elena Lázaro[2] | Virgilio Gómez-Rubio[3] | Carmen Armero[4]

[1]Department of Statistics, Pontificia Universidad Católica de Chile, Santiago, Chile

[2]Plant Protection and Biotechnology Centre, Instituto Valenciano de Investigaciones Agrarias, Valencia, Spain

[3]Department of Mathematics, School of Industrial Engineering-Albacete, Universidad de Castilla-La Mancha, Ciudad Real, Spain

[4]Department of Statistics and Operational Research, Universitat de València, Valencia, Spain

Survival analysis is one of the most important fields of statistics in medicine and biological sciences. In addition, the computational advances in the last decades have favored the use of Bayesian methods in this context, providing a flexible and powerful alternative to the traditional frequentist approach. The objective of this article is to summarize some of the most popular Bayesian survival models, such as accelerated failure time, proportional hazards, mixture cure, competing risks, multi-state, frailty, and joint models of longitudinal and survival data. Moreover, an implementation of each presented model is provided using a BUGS syntax that can be run with JAGS from the R programming language. Reference to other Bayesian R-packages is also discussed.

KEYWORDS

Bayesian inference, JAGS, R-packages, time-to-event analysis

github.com/daniloalvares/Bayes-surv-BUGS

Reference II

Received: 13 November 2023 | Revised: 7 June 2024 | Accepted: 12 June 2024

DOI: 10.1002/sim.10160

TUTORIAL IN BIOSTATISTICS

Statistics in Medicine WILEY

Bayesian survival analysis with INLA

Danilo Alvares[1] | Janet van Niekerk[2] | Elias Teixeira Krainski[2] | Håvard Rue[2] | Denis Rustand[2]

[1]MRC Biostatistics Unit, University of Cambridge, Cambridge, UK

[2]Computer, Electrical and Mathematical Sciences and Engineering Division, King Abdullah University of Science and Technology, Thuwal, Kingdom of Saudi Arabia

This tutorial shows how various Bayesian survival models can be fitted using the integrated nested Laplace approximation in a clear, legible, and comprehensible manner using the `INLA` and `INLAjoint` R-packages. Such models include accelerated failure time, proportional hazards, mixture cure, competing risks, multi-state, frailty, and joint models of longitudinal and survival data, originally presented in the article "*Bayesian survival analysis with BUGS*." In addition, we illustrate the implementation of a new joint model for a longitudinal semi-continuous marker, recurrent events, and a terminal event. Our proposal aims to provide the reader with syntax examples for implementing survival models using a fast and accurate approximate Bayesian inferential approach.

KEYWORDS

Bayesian inference, INLA, joint modeling, R-packages, time-to-event analysis

github.com/DenisRustand/Bayes-surv-INLA

Outline

Part I

Motivation

The outcome variable of interest is time until an event occurs

- Survival analysis, sometimes referred to as failure-time analysis, is one of the most important fields of statistics, mainly in medicine, biological sciences, and engineering.

- Survival times are data that measure follow-up time from a defined starting point to the occurrence of a given event of interest or endpoint, for instance, onset of disease, cure, death, and so on.

Motivation - Example 1

Chilean Journal of Statistics
Vol. 12, No. 1, April 2021, 53–70

Bayesian Survival Analysis
Research Paper

Bayesian long-term survival model including a frailty term: Application to melanoma data

Agatha S. Rodrigues[1,*], Vinicius F. Calsavara[2,3], Eduardo Bertolli[4], Stela V. Peres[5], and Vera L.D. Tomazella[6]

[1]Department of Statistics, Federal University of Espírito Santo, Vitória-ES, Brazil
[2]Department of Epidemiology and Statistics, A.C.Camargo Cancer Center, São Paulo-SP, Brazil
[3]Biostatistics and Bioinformatics Research Center, Cedars-Sinai Medical Center, Los Angeles-CA, USA
[4]Skin Cancer Department, A.C.Camargo Cancer Center, São Paulo-SP, Brazil
[5]Department of Information and Epidemiology, Fundação Oncocentro de São Paulo, São Paulo-SP, Brazil
[6]Department of Statistics, Federal University of São Carlos, São Carlos-SP, Brazil

(Received: 04 August 2020 · Accepted in final form: 28 January 2021)

Abstract

In this paper, we propose a flexible cure rate model including a frailty term, which was obtained by incorporating a random effect in the risk function of latent competing causes. The number of competing causes of the event of interest follows a negative binomial distribution, and the frailty variable follows a power variance function distribution, which includes other frailty models such as gamma, positive stable, and inverse Gaussian frailty models as special cases. The proposed model takes into account the presence of covariates and right-censored data, which are suitable for populations with a long-term survivors. Besides, it allows quantification of the degree of unobserved heterogeneity induced by unobservable risk factors, which is important to explain the lifetime. Once the posterior density function is not expressed in the closed form, Markov chain Monte Carlo algorithms are performed for the estimation procedure. Simulation studies were considered in order to evaluate the proposed model performance, and its practical relevance was illustrated in a real medical dataset from a population-based study of incident cases of melanoma diagnosed in the state of São Paulo, Brazil.

Keywords: Competing causes · Frailty models · Markov chain Monte Carlo · Negative binomial distribution · Power variance function

Mathematics Subject Classification: Primary 62N01 · Secondary 62P10.

*Corresponding author. Email: agatha.rodrigues@ufes.br

ISSN: 0718-7912 (print)/ISSN 0718-7920 (online)

http://www.soche.cl/chjs

Motivation - Example 2

Reliability Engineering and System Safety 204 (2020) 107145

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress

Bayesian non-parametric frailty model for dependent competing risks in a repairable systems framework

Marco Pollo Almeida[a,b], Rafael S. Paixão[a,b], Pedro L. Ramos[b,*], Vera Tomazella[a], Francisco Louzada[b], Ricardo S. Ehlers[b]

[a] *Federal University of São Carlos, São Carlos, São Paulo, Brazil*
[b] *Institute of Mathematical and Computer Sciences, University of São Paulo, São Carlos, São Paulo, Brazil*

ARTICLE INFO

Keywords:
Bayesian non-parametric
Shared frailty
Repairable systems
Power law process
Dependent competing risks
Hamiltonian Monte Carlo

ABSTRACT

The aim of this article is to analyze multiple repairable systems data under the presence of dependent competing risks. It is known that the dependence effect in this scenario influences the estimates of the model parameters. Hence, under the assumption that the cause-specific intensities follow a power law process (PLP), we propose a frailty-induced dependence approach to incorporate the dependence among the cause-specific recurrent processes. Moreover, the misspecification of the frailty distribution may lead to errors when estimating the parameters of interest. Because of this, we considered a nonparametric approach to model the frailty density using a Dirichlet process mixture prior, which offers more flexibility to provide consistent estimates for the PLP model, as well as insights about heterogeneity among the systems. We proposed an orthogonal parametrization for the PLP model parameters that allowed us to specify a joint prior distribution for the parameters that returned closed-form estimators for the posterior mean. Additionally, a hybrid MCMC sampler algorithm composed by Hamiltonian Monte Carlo and Gibbs sampling was built for computing the posterior estimates with respect to the frailty distribution. A simulation study was conducted to evaluate the efficiency of our estimates. This method was used to analyze a real dataset. Algorithms, code, and data are provided in supplementary material available online.

Motivation - Example 3

Analysis of Permanence Time in Emotional States: A Case Study Using Educational Software

Helena Reis[1(✉)], Danilo Alvares[2(✉)], Patricia Jaques[3(✉)], and Seiji Isotani[1(✉)]

[1] Instituto de Ciências de Computação e Matematica Computacional (ICMC), Universidade de São Paulo (USP), São Carlos, SP, Brazil
{helenamcd,sisotani}@icmc.usp.br
[2] Harvard T.H. Chan School of Public Health, Boston, MA, USA
dalvares@hsph.harvard.edu
[3] Programa de Pos-Graduação em Computação Aplicada (PPGCA), Universidade do Vale do Rio dos Sinos (UNISINOS), São Leopoldo, RS, Brazil
pjaques@unisinos.br

Abstract. This article presents the results of an experiment in which we investigated how prior algebra knowledge and personality can influence the permanence time from the confusion state to frustration/boredom state in a computer learning environment. Our experimental results indicate that people with a neurotic personality and a low level of algebra knowledge can deal with confusion for less time and can easily feel frustrated/bored when there is no intervention. Our analysis also suggest that people with an extroversion personality and a low level of algebra knowledge are able to control confusion for longer, leading to later interventions. These findings support that it is possible to detect emotions in a less invasive way and without the need of physiological sensors or complex algorithms. Furthermore, obtained median times can be incorporated into computational regulation models (e.g. adaptive interfaces) to regulate students' emotion during the teaching-learning process.

R. Nkambou et al. (Eds.): ITS 2018, LNCS 10858, pp. 180–190, 2018.
https://doi.org/10.1007/978-3-319-91464-0_18

Basic notation

- T: non-negative random variable (time-to-event)
- $\boldsymbol{\theta}$: set of unknown quantities (typically parameters)
- $h(t \mid \boldsymbol{\theta})$: hazard function
- $f(t \mid \boldsymbol{\theta})$: density function
- $S(t \mid \boldsymbol{\theta})$: survival function

$$h(t \mid \boldsymbol{\theta}) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t \mid T \geq t, \boldsymbol{\theta})}{\Delta t} = \frac{f(t \mid \boldsymbol{\theta})}{S(t \mid \boldsymbol{\theta})}, \quad t > 0$$

Definition: the hazard function of T at time t represents the instantaneous rate of the event occurrence for the population group that is still at risk at time t.

Usual relationships

$$H(t \mid \boldsymbol{\theta}) = \int_0^t h(u \mid \boldsymbol{\theta})\, \mathrm{d}u, \quad t > 0 \quad \text{[cumulative hazard]}$$

$$S(t \mid \boldsymbol{\theta}) = \exp\big\{ - H(t \mid \boldsymbol{\theta})\big\}, \quad t > 0$$

$$f(t \mid \boldsymbol{\theta}) = h(t \mid \boldsymbol{\theta})\, S(t \mid \boldsymbol{\theta}) = h(t \mid \boldsymbol{\theta}) \exp\big\{ - H(t \mid \boldsymbol{\theta})\big\}, \quad t > 0$$

$$F(t \mid \boldsymbol{\theta}) = 1 - S(t \mid \boldsymbol{\theta}), \quad t > 0 \quad \text{[cumulative distribution]}$$

Example (exponential model):

- $f(t \mid \boldsymbol{\theta}) = \lambda \exp\{-\lambda t\}$
- $h(t \mid \boldsymbol{\theta}) = \lambda$
- $S(t \mid \boldsymbol{\theta}) = \exp\{-\lambda t\} \quad \Rightarrow \quad F(t \mid \boldsymbol{\theta}) = 1 - \exp\{-\lambda t\}$
- $H(t \mid \boldsymbol{\theta}) = \lambda t$

Usual relationships

$$H(t \mid \boldsymbol{\theta}) = \int_0^t h(u \mid \boldsymbol{\theta})\, \mathrm{d}u, \quad t > 0 \quad \text{[cumulative hazard]}$$

$$S(t \mid \boldsymbol{\theta}) = \exp\big\{ - H(t \mid \boldsymbol{\theta})\big\}, \quad t > 0$$

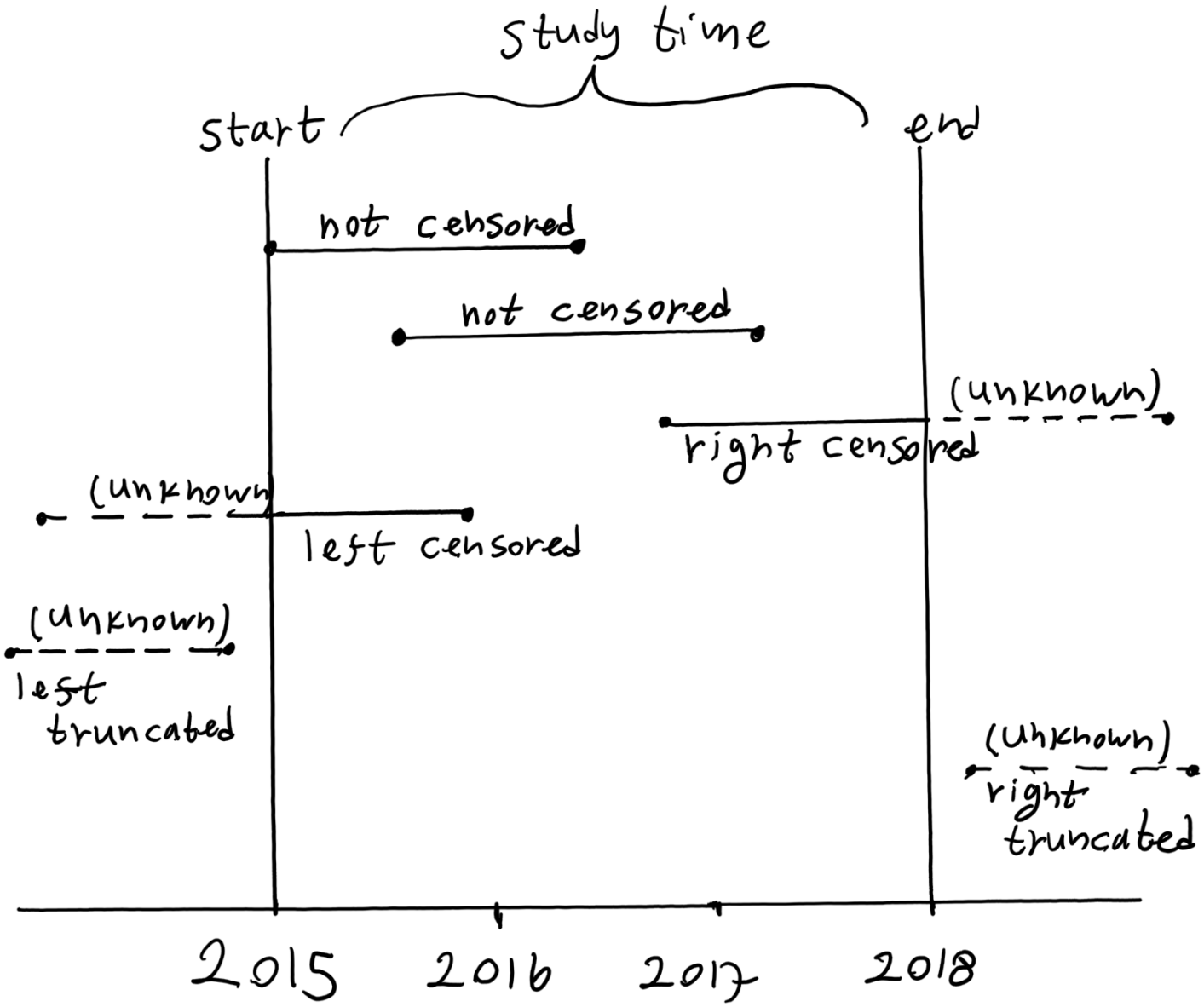
$$f(t \mid \boldsymbol{\theta}) = h(t \mid \boldsymbol{\theta})\, S(t \mid \boldsymbol{\theta}) = h(t \mid \boldsymbol{\theta}) \exp\big\{ - H(t \mid \boldsymbol{\theta})\big\}, \quad t > 0$$

$$F(t \mid \boldsymbol{\theta}) = 1 - S(t \mid \boldsymbol{\theta}), \quad t > 0 \quad \text{[cumulative distribution]}$$

Example (exponential model):

- $f(t \mid \boldsymbol{\theta}) = \lambda \exp\{-\lambda t\}$
- $h(t \mid \boldsymbol{\theta}) = \lambda$
- $S(t \mid \boldsymbol{\theta}) = \exp\{-\lambda t\} \quad \Rightarrow \quad F(t \mid \boldsymbol{\theta}) = 1 - \exp\{-\lambda t\}$
- $H(t \mid \boldsymbol{\theta}) = \lambda t$

Censoring and truncation



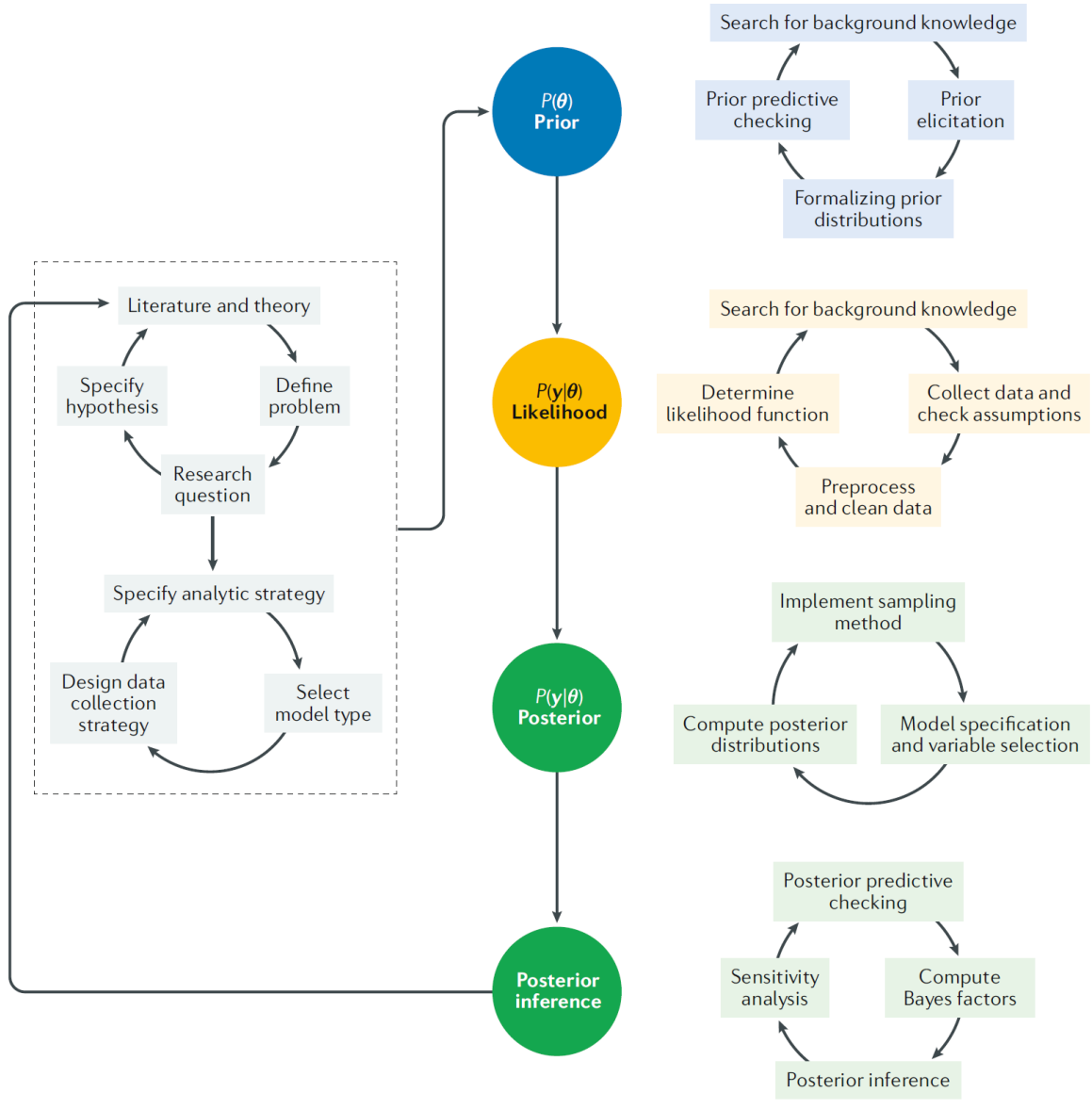
Censoring and truncation - Likelihood

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} L_i(\boldsymbol{\theta}) = \prod_{i=1}^{n} f(t_i \mid \boldsymbol{\theta})$$

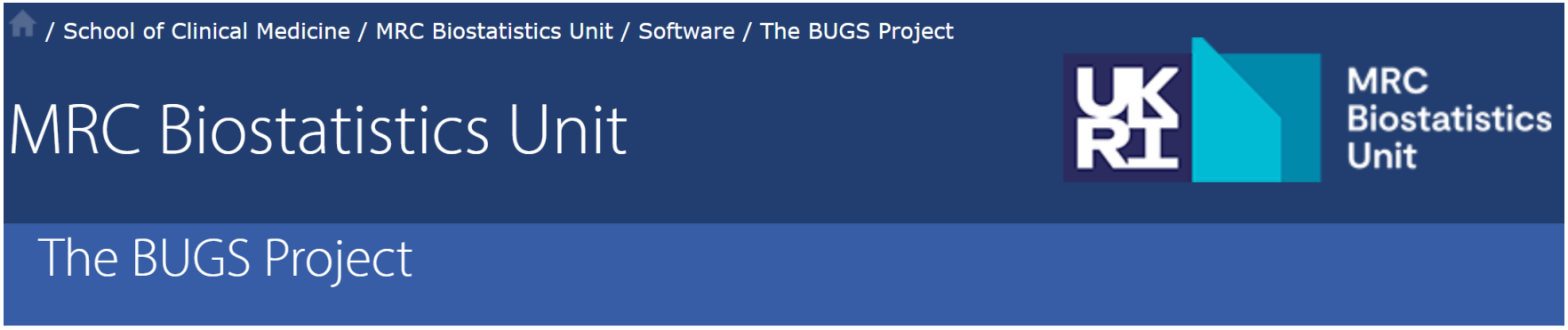
$$f^*(t_i \mid \boldsymbol{\theta}) = \begin{cases} h(t_i \mid \boldsymbol{\theta}) S(t_i \mid \boldsymbol{\theta}), & \text{if exact lifetime} \\ S(t_i \mid \boldsymbol{\theta}), & \text{if right-censoring} \\ 1 - S(t_i \mid \boldsymbol{\theta}), & \text{if left-censoring} \\ S(t_{i,L} \mid \boldsymbol{\theta}) - S(t_{i,R} \mid \boldsymbol{\theta}), & \text{if interval-censoring} \end{cases}$$

$$f(t_i \mid \boldsymbol{\theta}) = \begin{cases} f^*(t_i \mid \boldsymbol{\theta}), & \text{if no truncation} \\ f(t_i \mid T < v, \boldsymbol{\theta}) = \dfrac{f^*(t_i \mid \boldsymbol{\theta})}{1 - S(v \mid \boldsymbol{\theta})}, & \text{if right truncation} \\ f(t_i \mid T > u, \boldsymbol{\theta}) = \dfrac{f^*(t_i \mid \boldsymbol{\theta})}{S(u \mid \boldsymbol{\theta})}, & \text{if left truncation} \end{cases}$$

Bayesian inference



BUGS project - The origin

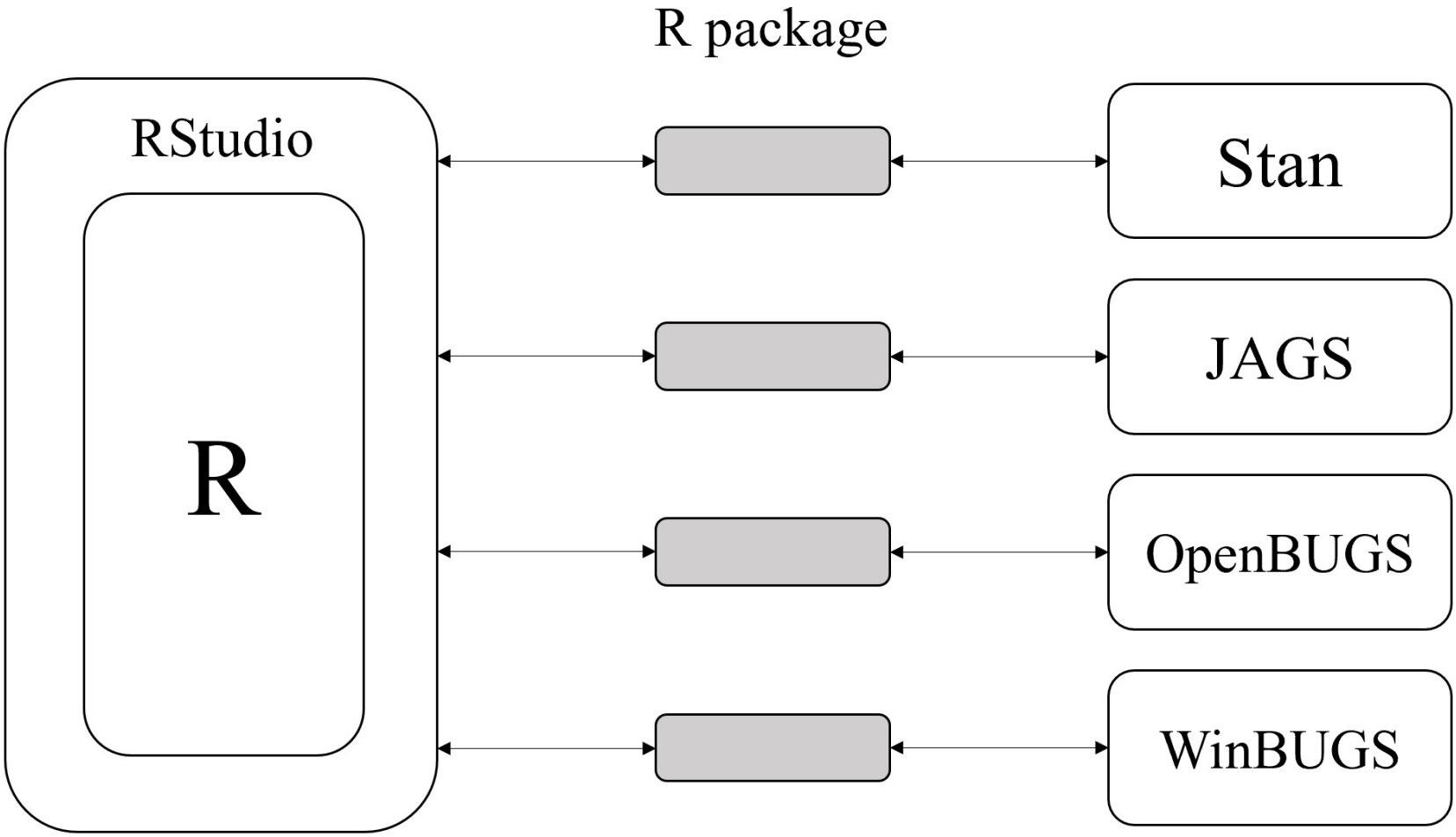


BUGS language

- WinBUGS
- OpenBUGS
- MultiBUGS
- JAGS (**J**ust **A**nother **G**ibbs **S**ampler)
- NIMBLE (**N**umerical **I**nference for statistical **M**odels using **B**ayesian and **L**ikelihood **E**stimation)
- **Stan** (named in honor of Stanislaw Ulam): inspired by BUGS and superficially similar, but it is conceptually different in many ways.

Also, there are many R-packages that implement statistical models from a Bayesian perspective.

Connection between R and external programs



Stan blocks

```
functions{
    // ... function declarations and definitions ...
}
data{
    // ... declarations ...
}
transformed data{
    // ... declarations ... statements ...
}
parameters{
    // ... declarations ...
}
transformed parameters{
    // ... declarations ... statements ...
}
model{
    // ... declarations ... statements ...
}
generated quantities{
    // ... declarations ... statements ...
}
```

Toy example - Simple linear regression in Stan

- Linear model:

$$
\begin{aligned}
(y_i \mid \mu_i, \sigma) &\sim \text{Normal}(\mu_i, \sigma^2) \\
\mu_i &= \beta_0 + \beta_1 x_i \\
i &= 1, \ldots, n
\end{aligned}
$$

- Priors:

$$
\begin{aligned}
\beta_0 &\sim \text{Normal}(0, 100^2) \\
\beta_1 &\sim \text{Normal}(0, 100^2) \\
\sigma &\sim \text{Gamma}(0.01, 0.01)
\end{aligned}
$$

Toy example - Simple linear regression in Stan

- Linear model:

$$
\begin{aligned}
(y_i \mid \mu_i, \sigma) &\sim \text{Normal}(\mu_i, \sigma^2) \\
\mu_i &= \beta_0 + \beta_1 x_i \\
i &= 1, \ldots, n
\end{aligned}
$$

- Priors:

$$
\begin{aligned}
\beta_0 &\sim \text{Normal}(0, 100^2) \\
\beta_1 &\sim \text{Normal}(0, 100^2) \\
\sigma &\sim \text{Gamma}(0.01, 0.01)
\end{aligned}
$$

Accelerated failure time models

Accelerated failure time (AFT) models

Survival time T in logarithmic scale is expressed in terms of a linear combination of covariates $\boldsymbol{x}$ with regression coefficients $\boldsymbol{\beta}$ and a measurement error ξ as follows:

$$\log(T) = \boldsymbol{x}^\top \boldsymbol{\beta} + \sigma\,\epsilon,$$

where σ is a scale parameter and ϵ an error term usually expressed via a normal, logistic or a standard Gumbel probabilistic distribution.

Example: a standard Gumbel distribution for ϵ implies a conditional (on $\boldsymbol{\beta}$ and σ) Weibull survival model for T with shape $\alpha = 1/\sigma$ and scale $\lambda(\boldsymbol{\beta}, \sigma) = \exp\left\{-\boldsymbol{x}^\top \boldsymbol{\beta}/\sigma\right\}$ parameters:

$$h(t \mid \lambda, \alpha) = \lambda\,\alpha\,t^{\alpha-1} \quad \text{[Weibull hazard]}$$

Accelerated failure time (AFT) models

Survival time T in logarithmic scale is expressed in terms of a linear combination of covariates $\boldsymbol{x}$ with regression coefficients $\boldsymbol{\beta}$ and a measurement error ξ as follows:

$$\log(T) = \boldsymbol{x}^\top \boldsymbol{\beta} + \sigma\,\epsilon,$$

where σ is a scale parameter and ϵ an error term usually expressed via a normal, logistic or a standard Gumbel probabilistic distribution.

Example: a standard Gumbel distribution for ϵ implies a conditional (on $\boldsymbol{\beta}$ and σ) Weibull survival model for T with shape $\alpha = 1/\sigma$ and scale $\lambda(\boldsymbol{\beta}, \sigma) = \exp\left\{-\boldsymbol{x}^\top \boldsymbol{\beta}/\sigma\right\}$ parameters:

$$h(t \mid \lambda, \alpha) = \lambda\,\alpha\,t^{\alpha-1} \quad \text{[Weibull hazard]}$$

Accelerated failure time (AFT) models

Survival time T in logarithmic scale is expressed in terms of a linear combination of covariates $\boldsymbol{x}$ with regression coefficients $\boldsymbol{\beta}$ and a measurement error ξ as follows:

$$\log(T) = \boldsymbol{x}^\top \boldsymbol{\beta} + \sigma\,\epsilon,$$

where σ is a scale parameter and ϵ an error term usually expressed via a normal, logistic or a standard Gumbel probabilistic distribution.

Example: a standard Gumbel distribution for ϵ implies a conditional (on $\boldsymbol{\beta}$ and σ) Weibull survival model for T with shape $\alpha = 1/\sigma$ and scale $\lambda(\boldsymbol{\beta}, \sigma) = \exp\left\{-\boldsymbol{x}^\top \boldsymbol{\beta}/\sigma\right\}$ parameters:

$$h(t \mid \lambda, \alpha) = \lambda\,\alpha\,t^{\alpha-1} \quad \text{[Weibull hazard]}$$

Outline

Part II

Proportional hazards models

Proportional hazards (PH) models

$$h(t \mid h_0, \boldsymbol{\beta}) = h_0(t) \exp\left\{\boldsymbol{x}^\top \boldsymbol{\beta}\right\}$$

- Baseline hazard functions, $h_0(t)$, can be modeled considering some of the usual probability distributions in the survival analysis framework such as exponential, Weibull, Gompertz, etc.

- $h_0(t)$ can also defined as a mixture of piecewise constant functions, $h_0(t \mid \boldsymbol{\lambda}) = \sum_{k=1}^{K} \lambda_k I_{(a_{k-1}, a_k]}(t)$, $t > 0$, where $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_K)$ and $I_{(a_{k-1}, a_k]}(t)$ is the indicator function defined as 1 when $t \in (a_{k-1}, a_k]$ and 0 otherwise.

Proportional hazards (PH) models

$$h(t \mid h_0, \boldsymbol{\beta}) = h_0(t) \exp\left\{\boldsymbol{x}^\top \boldsymbol{\beta}\right\}$$

- Baseline hazard functions, $h_0(t)$, can be modeled considering some of the usual probability distributions in the survival analysis framework such as exponential, Weibull, Gompertz, etc.

- $h_0(t)$ can also defined as a mixture of piecewise constant functions, $h_0(t \mid \boldsymbol{\lambda}) = \sum_{k=1}^{K} \lambda_k I_{(a_{k-1}, a_k]}(t)$, $t > 0$, where $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_K)$ and $I_{(a_{k-1}, a_k]}(t)$ is the indicator function defined as 1 when $t \in (a_{k-1}, a_k]$ and 0 otherwise.

Proportional hazards (PH) models

$$h(t \mid h_0, \boldsymbol{\beta}) = h_0(t) \exp\left\{\boldsymbol{x}^\top \boldsymbol{\beta}\right\}$$

- Baseline hazard functions, $h_0(t)$, can be modeled considering some of the usual probability distributions in the survival analysis framework such as exponential, Weibull, Gompertz, etc.

- $h_0(t)$ can also defined as a mixture of piecewise constant functions, $h_0(t \mid \boldsymbol{\lambda}) = \sum_{k=1}^{K} \lambda_k \, I_{(a_{k-1}, a_k]}(t)$, $t > 0$, where $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_K)$ and $I_{(a_{k-1}, a_k]}(t)$ is the indicator function defined as 1 when $t \in (a_{k-1}, a_k]$ and 0 otherwise.

Proportional hazards (PH) models

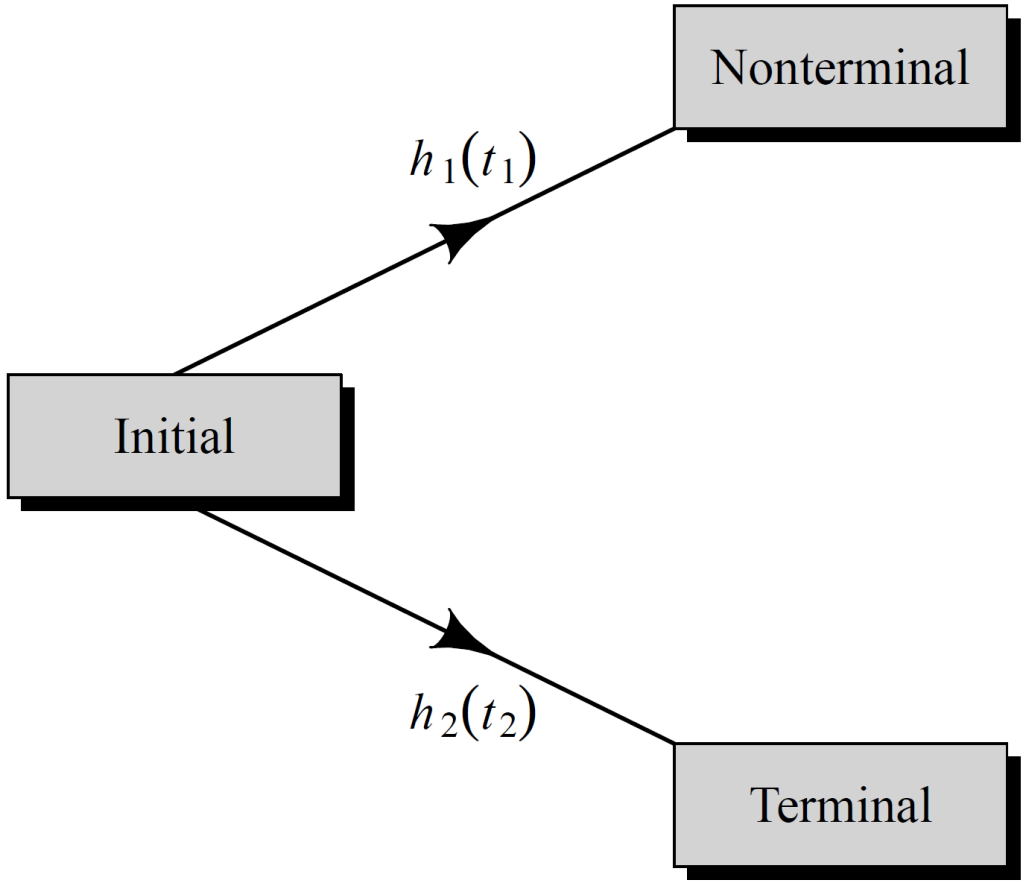
$$h(t \mid h_0, \boldsymbol{\beta}) = h_0(t) \exp\left\{\boldsymbol{x}^\top \boldsymbol{\beta}\right\}$$

- Baseline hazard functions, $h_0(t)$, can be modeled considering some of the usual probability distributions in the survival analysis framework such as exponential, Weibull, Gompertz, etc.

- $h_0(t)$ can also defined as a mixture of piecewise constant functions, $h_0(t \mid \boldsymbol{\lambda}) = \sum_{k=1}^{K} \lambda_k \, I_{(a_{k-1}, a_k]}(t)$, $t > 0$, where $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_K)$ and $I_{(a_{k-1}, a_k]}(t)$ is the indicator function defined as 1 when $t \in (a_{k-1}, a_k]$ and 0 otherwise.

Competing risks models

Competing risks (CR) models



Competing risks (CR) models

- Cause-specific hazard function for cause k:

$$h_k(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t, \delta = k \mid T \geq t)}{\Delta t}$$

- Cumulative incidence function for cause k:

$$F_k(t) = P(T \leq t, \delta = k) = \int_0^t h_k(u)\, S(u)\, \mathrm{d}u$$

- Overall survival function:

$$S(t) = P(T > t) = \exp\Big\{ -\sum_{k=1}^{K} \int_0^t h_k(u)\, \mathrm{d}u \Big\}$$

Example (Weibull hazard):

$$h_k(t \mid \lambda_k, \alpha_k, \boldsymbol{\beta}_k) = \lambda_k\, \alpha_k\, t^{\alpha_k - 1} \exp\Big\{ \boldsymbol{x}^\top \boldsymbol{\beta}_k \Big\}$$

Competing risks (CR) models

- Cause-specific hazard function for cause k:

$$h_k(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t, \delta = k \mid T \geq t)}{\Delta t}$$

- Cumulative incidence function for cause k:

$$F_k(t) = P(T \leq t, \delta = k) = \int_0^t h_k(u)\, S(u)\, \mathrm{d}u$$

- Overall survival function:

$$S(t) = P(T > t) = \exp\Big\{ - \sum_{k=1}^{K} \int_0^t h_k(u)\, \mathrm{d}u \Big\}$$

Example (Weibull hazard):

$$h_k(t \mid \lambda_k, \alpha_k, \boldsymbol{\beta}_k) = \lambda_k\, \alpha_k\, t^{\alpha_k - 1} \exp\Big\{ \boldsymbol{x}^\top \boldsymbol{\beta}_k \Big\}$$

Competing risks (CR) models

- Cause-specific hazard function for cause k:

$$h_k(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t, \delta = k \mid T \geq t)}{\Delta t}$$

- Cumulative incidence function for cause k:

$$F_k(t) = P(T \leq t, \delta = k) = \int_0^t h_k(u)\, S(u)\, \mathrm{d}u$$

- Overall survival function:

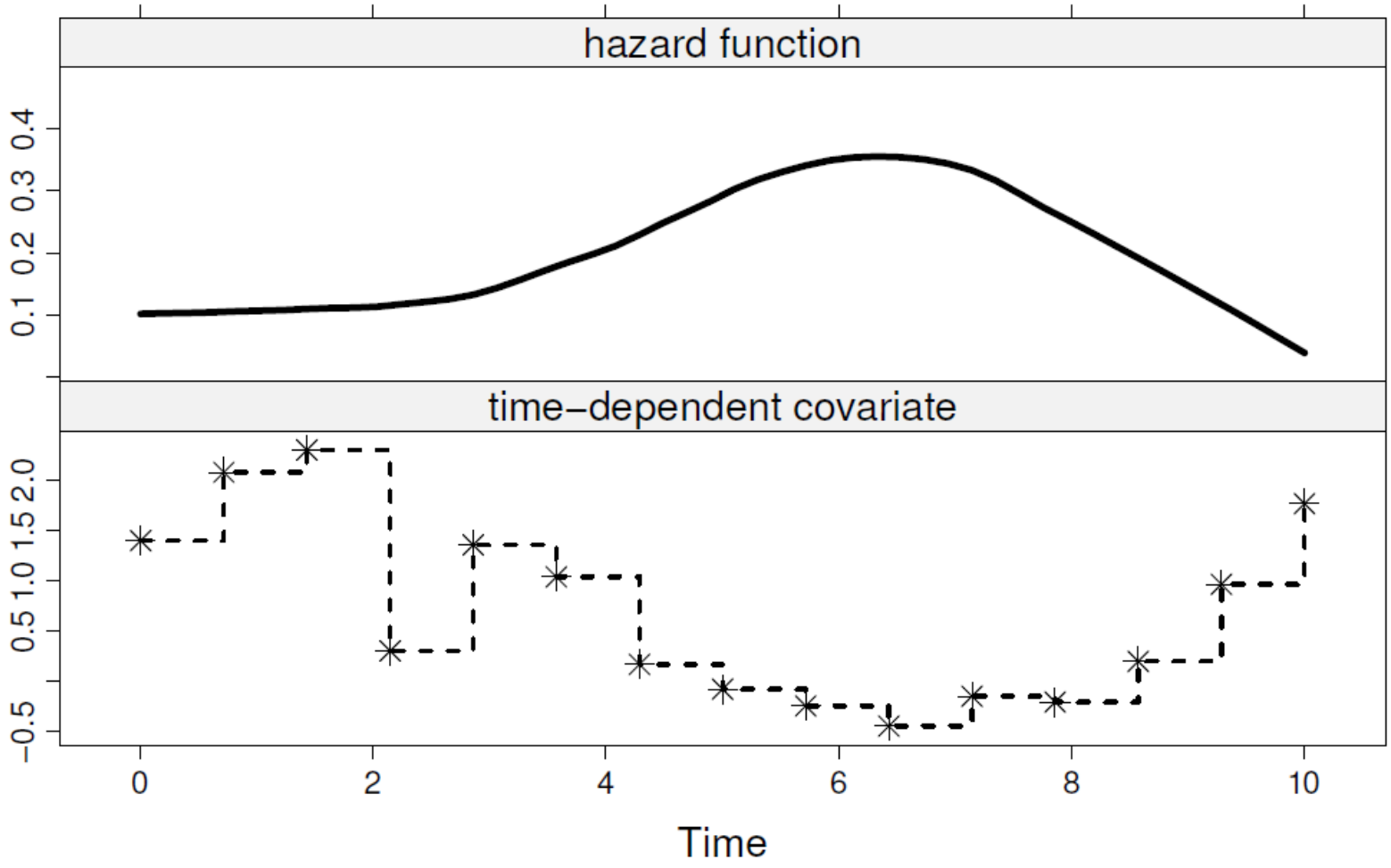
$$S(t) = P(T > t) = \exp\Big\{ -\sum_{k=1}^{K} \int_0^t h_k(u)\, \mathrm{d}u \Big\}$$

Example (Weibull hazard):

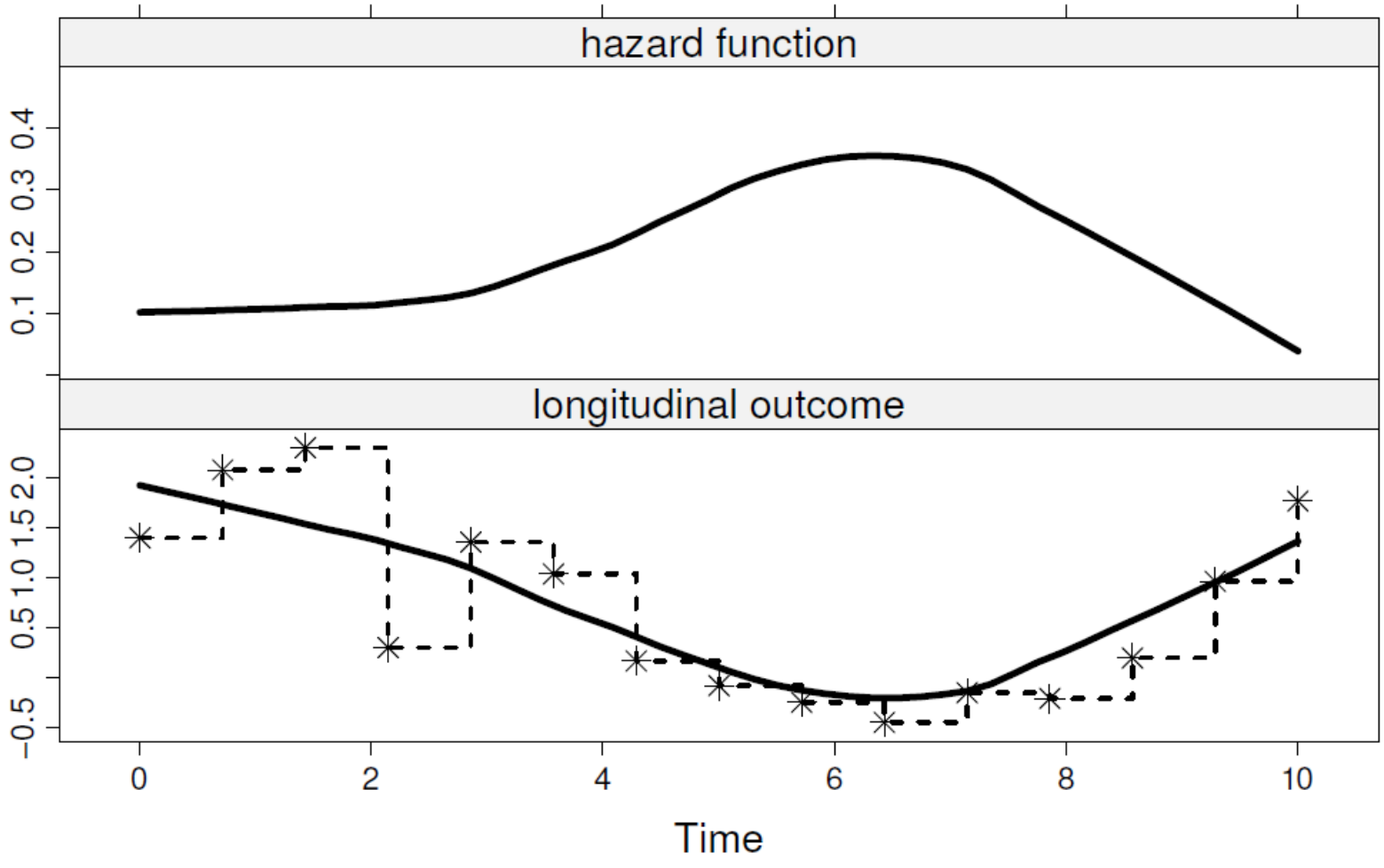
$$h_k(t \mid \lambda_k, \alpha_k, \boldsymbol{\beta}_k) = \lambda_k\, \alpha_k\, t^{\alpha_k - 1} \exp\left\{ \boldsymbol{x}^\top \boldsymbol{\beta}_k \right\}$$

Joint models of longitudinal and survival data

Joint models (JM) of longitudinal and survival data



Joint models (JM) of longitudinal and survival data



Joint models (JM) of longitudinal and survival data

Full joint distribution:

$$f(\boldsymbol{y}, \boldsymbol{s}, \boldsymbol{b}, \boldsymbol{\theta}) = f(\boldsymbol{y}, \boldsymbol{s} \mid \boldsymbol{b}, \boldsymbol{\theta})\, f(\boldsymbol{b} \mid \boldsymbol{\theta})\, \pi(\boldsymbol{\theta})$$

- $\boldsymbol{y}$: longitudinal process
- $\boldsymbol{s}$: survival process
- $\boldsymbol{b}$: random effects
- $\boldsymbol{\theta}$: parameters and hyperparameters

Shared-parameter specification:

$$f(\boldsymbol{y}, \boldsymbol{s} \mid \boldsymbol{b}, \boldsymbol{\theta}) = f(\boldsymbol{y} \mid \boldsymbol{b}, \boldsymbol{\theta}) f(\boldsymbol{s} \mid \boldsymbol{b}, \boldsymbol{\theta})$$

Joint models (JM) of longitudinal and survival data

Full joint distribution:

$$f(\boldsymbol{y}, \boldsymbol{s}, \boldsymbol{b}, \boldsymbol{\theta}) = f(\boldsymbol{y}, \boldsymbol{s} \mid \boldsymbol{b}, \boldsymbol{\theta})\, f(\boldsymbol{b} \mid \boldsymbol{\theta})\, \pi(\boldsymbol{\theta})$$

- $\boldsymbol{y}$: longitudinal process
- $\boldsymbol{s}$: survival process
- $\boldsymbol{b}$: random effects
- $\boldsymbol{\theta}$: parameters and hyperparameters

Shared-parameter specification:

$$f(\boldsymbol{y}, \boldsymbol{s} \mid \boldsymbol{b}, \boldsymbol{\theta}) = f(\boldsymbol{y} \mid \boldsymbol{b}, \boldsymbol{\theta}) f(\boldsymbol{s} \mid \boldsymbol{b}, \boldsymbol{\theta})$$

Joint models (JM) of longitudinal and survival data

- Linear mixed-effects submodel:

$$
\begin{aligned}
y_i(t \mid \mu_i, \sigma) &\sim \text{Normal}\big(\mu_i(t), \sigma^2\big) \\
\mu_i(t) &= \beta_{L1} + b_{i1} + (\beta_{L2} + b_{i2})t + \beta_{L3} x_i \\
\boldsymbol{b}_i &\sim \text{Normal}_2\big(\mathbf{0}, \boldsymbol{\Sigma}\big)
\end{aligned}
$$

- Survival submodel:

$$
h_i(t \mid \alpha, \boldsymbol{\beta}_S, \gamma, \boldsymbol{b}_i) = \alpha\, t^{\alpha-1} \exp\left\{\beta_{S1} + \beta_{S2} x_i + \gamma\,(b_{i1} + b_{i2} t)\right\},
$$

where α and $\lambda = \exp(\beta_{S1})$ are the Weibull shape and scale parameters, respectively.

Survival function - A complicated integral

$$S_i(t \mid \boldsymbol{\theta}_i) = \exp\left\{ - H_i(t \mid \boldsymbol{\theta}_i)\right\} = \exp\left\{ - \int_0^t h_i(u \mid \boldsymbol{\theta}_i)\, \mathrm{d}u \right\}$$

$$= \exp\left\{ - \int_0^t \alpha\, u^{\alpha-1} \exp\left\{\beta_{S1} + \beta_{S2} x_i + \gamma\left(b_{i1} + b_{i2} u\right)\right\} \mathrm{d}u \right\}$$

where $\boldsymbol{\theta}_i = (\alpha, \boldsymbol{\beta}_S, \gamma, \boldsymbol{b}_i)$.

Survival function - A complicated integral

$$S_i(t \mid \boldsymbol{\theta}_i) = \exp\left\{ - H_i(t \mid \boldsymbol{\theta}_i)\right\} = \exp\left\{ - \int_0^t h_i(u \mid \boldsymbol{\theta}_i)\, \mathrm{d}u \right\}$$

$$= \exp\left\{ - \int_0^t \alpha\, u^{\alpha-1} \exp\left\{\beta_{S1} + \beta_{S2} x_i + \gamma\,(b_{i1} + b_{i2} u)\right\} \mathrm{d}u \right\}$$

where $\boldsymbol{\theta}_i = (\alpha, \boldsymbol{\beta}_S, \gamma, \boldsymbol{b}_i)$.

Solution: numerical integration!

Gauss-Legendre quadrature

$$\int_{-1}^{1} f(z)\,\mathrm{d}z \;\approx\; \sum_{j=1}^{K} w_j f(x_j)$$

- K: number of nodes
- x_j: quadrature nodes
- w_j: quadrature weights

Gauss-Legendre quadrature

$$\int_{-1}^{1} f(z)\,\mathrm{d}z \;\approx\; \sum_{j=1}^{K} w_j f(x_j)$$

- K: number of nodes
- x_j: quadrature nodes
- w_j: quadrature weights

Change of interval:

$$\int_a^b f(z)\,\mathrm{d}z = \int_{-1}^{1} f\left(\frac{b-a}{2}\xi + \frac{a+b}{2}\right)\frac{\mathrm{d}x}{\mathrm{d}\xi}\,d\xi \quad \text{with} \quad \frac{\mathrm{d}z}{\mathrm{d}\xi} = \frac{b-a}{2}$$

$$\Longrightarrow \int_a^b f(z)\,\mathrm{d}z \;\approx\; \frac{b-a}{2}\sum_{j=1}^{K} w_j f\left(\frac{b-a}{2}x_j + \frac{a+b}{2}\right)$$

Gauss-Legendre quadrature

$$\int_{-1}^{1} f(z)\,\mathrm{d}z \;\approx\; \sum_{j=1}^{K} w_j f(x_j)$$

- K: number of nodes
- x_j: quadrature nodes
- w_j: quadrature weights

Change of interval:

$$\int_a^b f(z)\,\mathrm{d}z = \int_{-1}^{1} f\left(\frac{b-a}{2}\xi + \frac{a+b}{2}\right)\frac{\mathrm{d}x}{\mathrm{d}\xi}\,d\xi \quad \text{with} \quad \frac{\mathrm{d}z}{\mathrm{d}\xi} = \frac{b-a}{2}$$

$$\Longrightarrow \int_a^b f(z)\,\mathrm{d}z \;\approx\; \frac{b-a}{2}\sum_{j=1}^{K} w_j f\left(\frac{b-a}{2}x_j + \frac{a+b}{2}\right)$$

Conclusion

Conclusion

Stan offers the opportunity of easily use and adapt Bayesian hierarchical models without the need to manually implement MCMC methods.

We hope this short-course will encourage researchers who use survival models make their analyzes based on the Bayesian paradigm from the Stan codes we have provided and easily adapt them to novel settings.

Implementations of other Bayesian survival models not covered in this short-course, such as **mixture cure**, **multi-state**, and **frailty**, can be found at

- Stan: `github.com/daniloalvares/Bayes-surv-BUGS/tree/main/STAN`
- INLA: `github.com/DenisRustand/Bayes-surv-INLA`
- JAGS: `github.com/daniloalvares/Bayes-surv-BUGS/tree/main/JAGS`

Thank you for your attention

Contact information:

Email: danilo.alvares@mrc-bsu.cam.ac.uk

Webpage: `https://sites.google.com/site/daniloalvares`