Unified Mechanistic Survival Regression Models in Presence of a Cure Fraction

Francisco Louzada

(joint work with Dipak Dey, Vincente Cancho & Gladys Barriga)

University of Sao Paulo, Institute of Math. Sciences and Computing
BRAZIL

Biometrics-Biostatistics 2015

San Antonio, 16-18 November, 2015

Main Idea – Theoretical Purpose

To propose a New lifetime model for multivariate survival data with a cure fraction

- Which is developed under the presence of m types of latent competing risks and a proportion of survival individuals.
- For inferential purposes we use of Markov Chain Monte Carlo (MCMC) methods to develop a Bayesian analysis for the proposed model.
- Via a Simulation Study, we observed that the frequentist coverage probabilities of credible interval derived from the posteriors are close to the nominal.

Main Idea – Practical Purpose

The modeling is motivated by a real dataset on medical area:

- The medical data is on diabetic retinopathy

Contents

Contents

Introduction

- Survival models incorporating a cure fraction, often referred to as *cure rate model*, are becoming increasing popular in analyzing data from cancer clinical trials.
- Modelling failure time data for various types of cancers:
  - Breast cancer,
  - non-Hodgkings lymphoma,
  - leukemia,
  - prostate cancer,
  - melanoma,
  - Head and neck cancer.

Example

Example (1.1)

Kersey et al. (1987), analyzed a set of leukaemia data, which recorded the the times to recurrence of leukaemia for patients after one of the two transplants: allogeneic (Group 1, with 46 patients) or autologous (Group 2, with 44 patients).

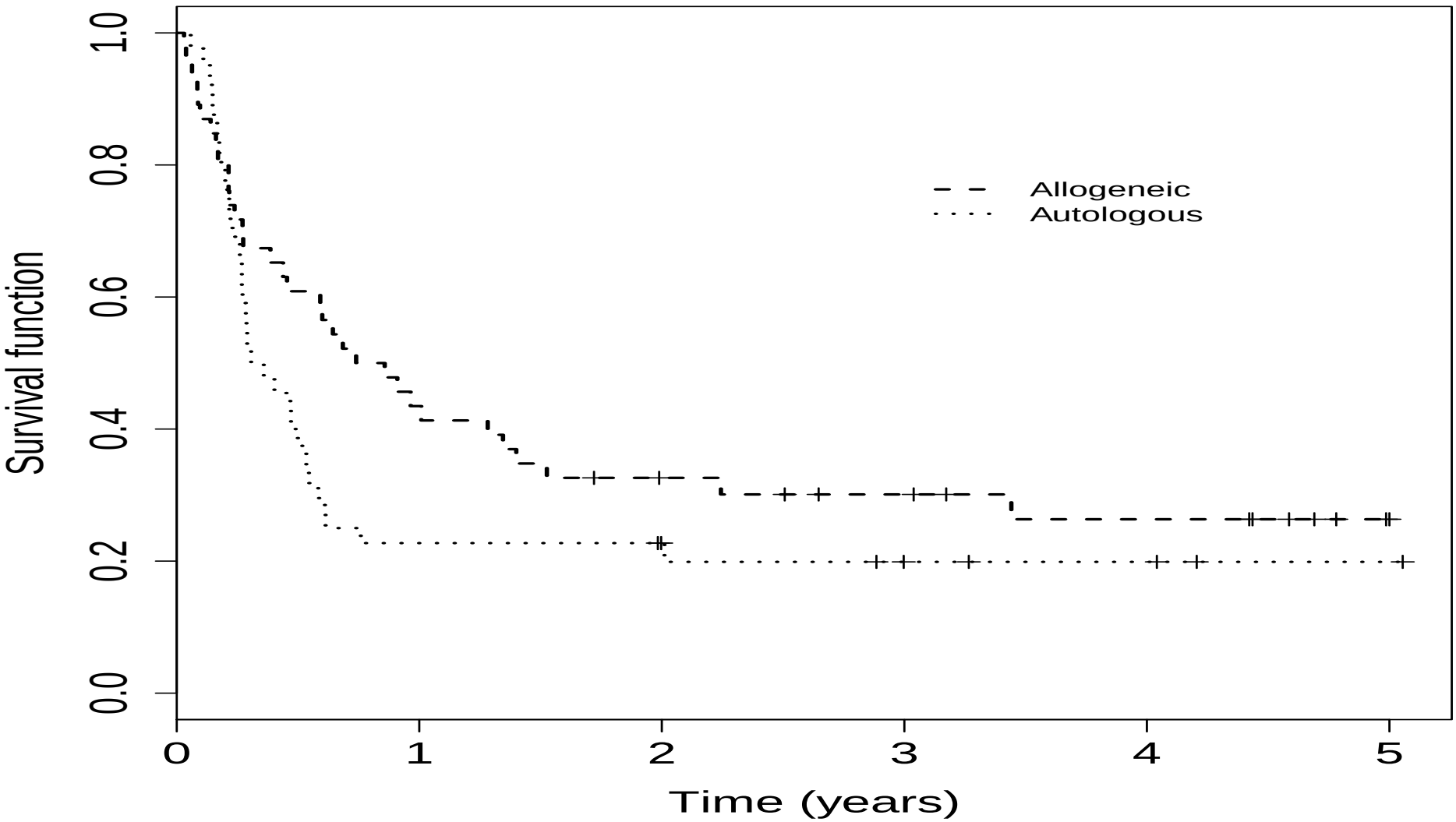


Figure: Kaplan–Meier of Leukaemia data– Exemplo-1.1 .

Example

Example (1.2)

Maller and Zhou (1996), presents a set of survival time following surgery of 45 patients with breast cancer stratified into two groups.

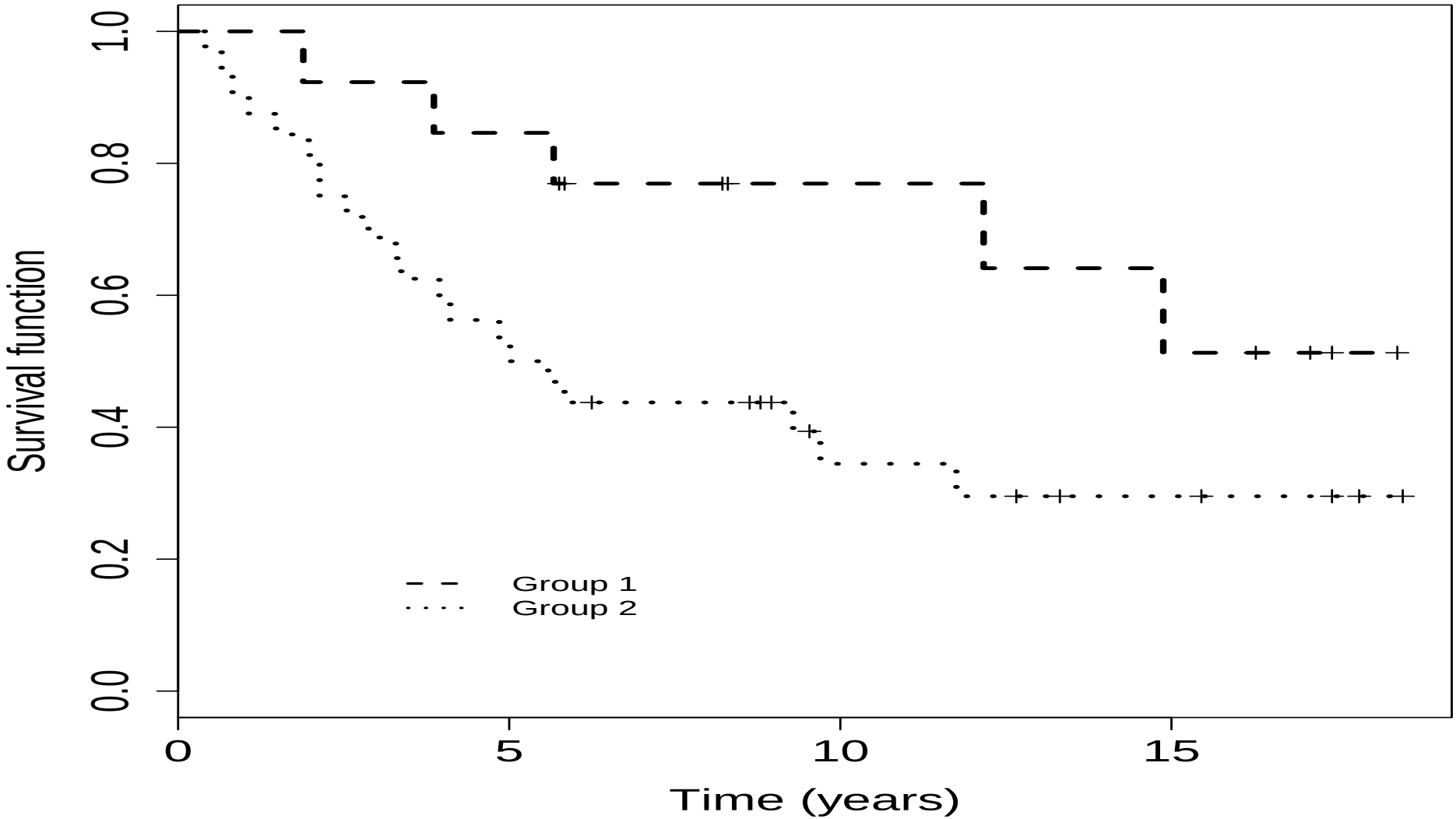


Figure: Kaplan–Meier of breast cancer data: Group 1=(negative staining),Group 2=(positive staining) – Exemplo-1.2 .

Motivation

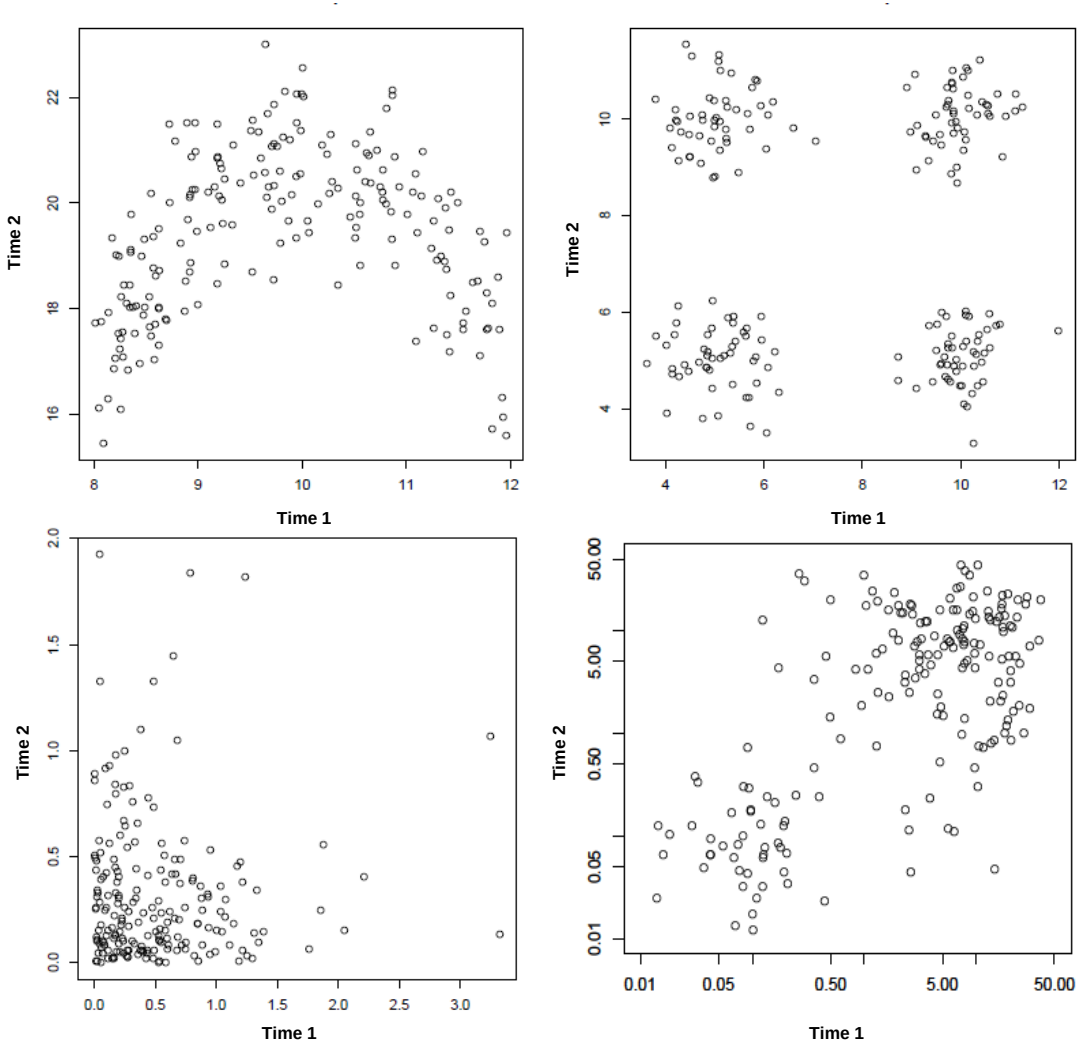
- In survival and reliability we may also observe two lifetimes (or even more) for a same subject.

  Examples:

- Medical area – lifetimes of matched human organs, as kidneys and eyes, and double recurrence of a certain disease.
- Industrial area – systems whose duration times depend on the durability of two components. Examples: damage of dual generators in a power plant or the lifetime of motors in a twin-engine airplane.
- Financial area – lifetimes of two types of insurances or two different credit products for the same client.

The Presence of Correlation

It is usual to observe dependence on bivariate survival data.



Interest

Let (T_1, T_2) be the lifetimes until the occurrence of an event of interest.

- To study the dependence between T_1 and T_2

The diabetic retinopathy study

Diabetic retinopathy is a leading cause of blindness worldwide.

It is estimated that blindness is 25 times more common in people with diabetes than in those without the disease.

Normal Vision and the same scene viewed by a person with diabetic retinopathy.

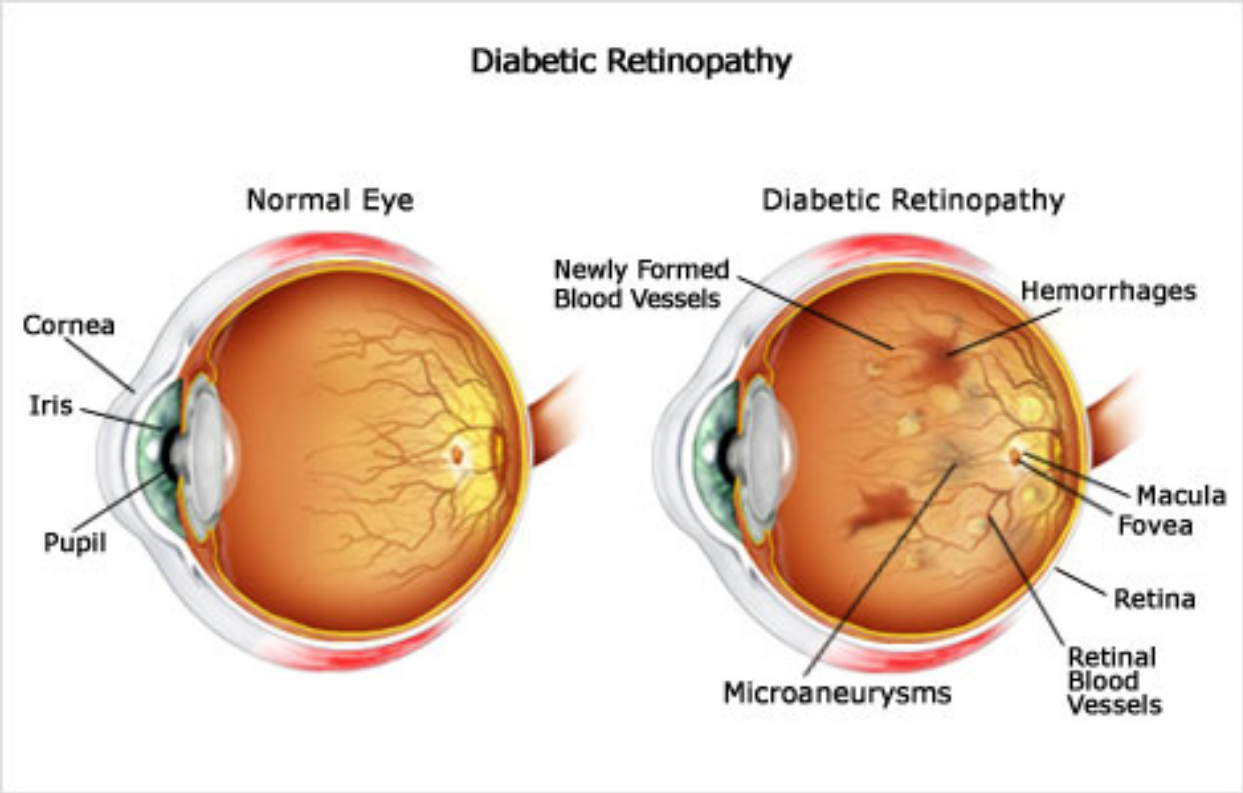
Normal vision

Same scene viewed by a person with diabetic retinopathy

The diabetic retinopathy study

- The main endpoint is severe visual loss in each eye.

The diabetic retinopathy study

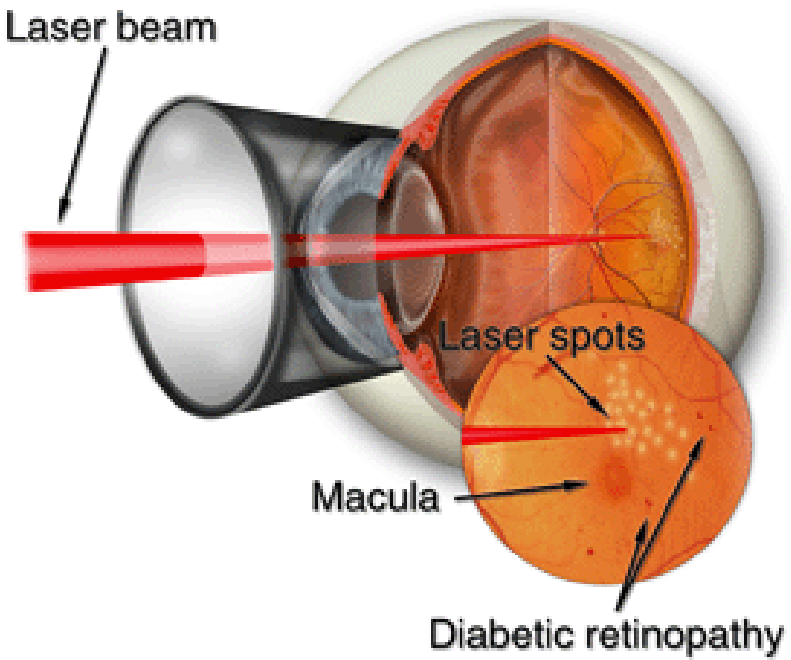


- The main endpoint is severe visual loss in each eye.

  Obstructing the vision.

The diabetic retinopathy study

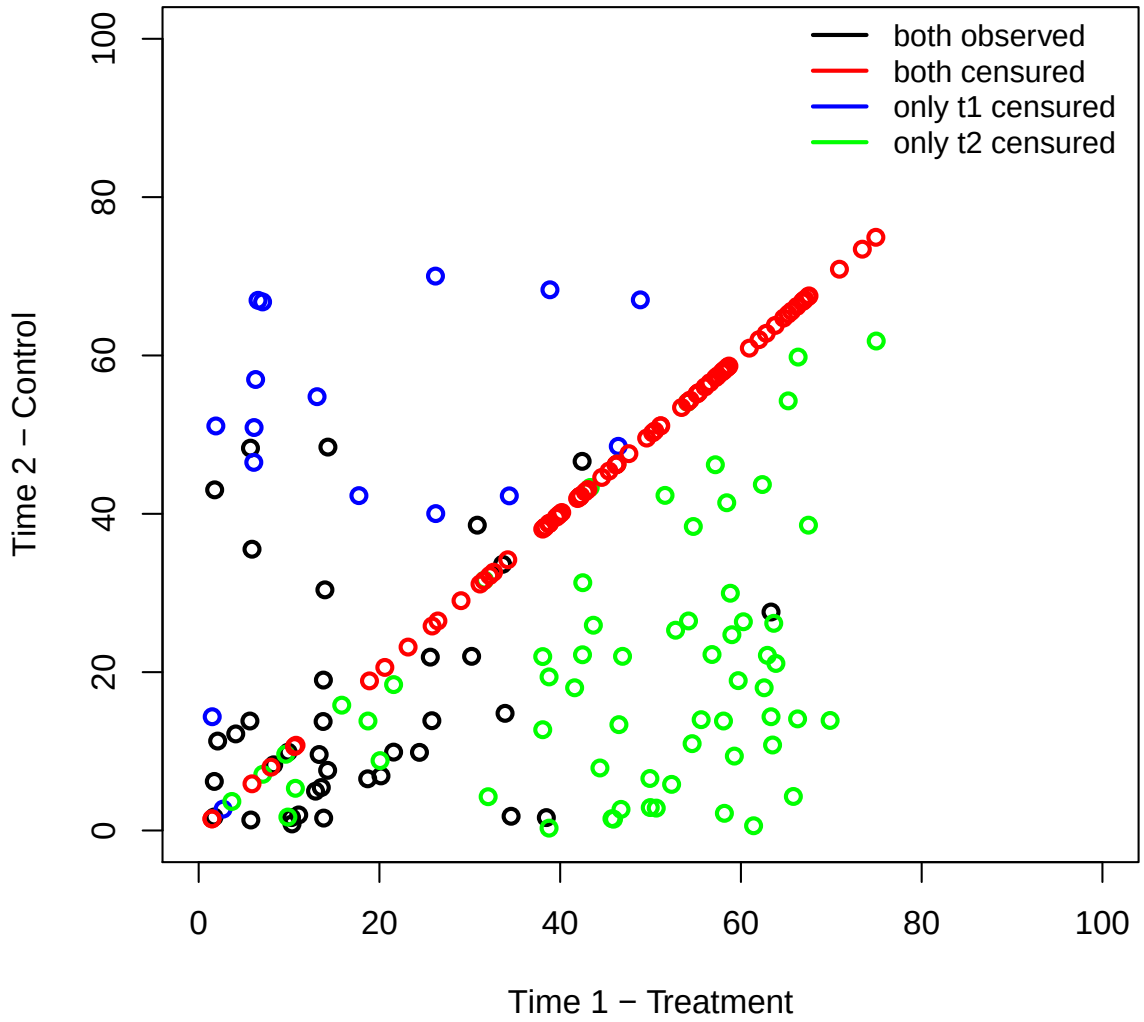
- Interest: Verify the effectiveness of the treatment (laser photo coagulation) in delaying the onset of blindness.
- The treatment was randomly assigned randomly to one eye of each patient



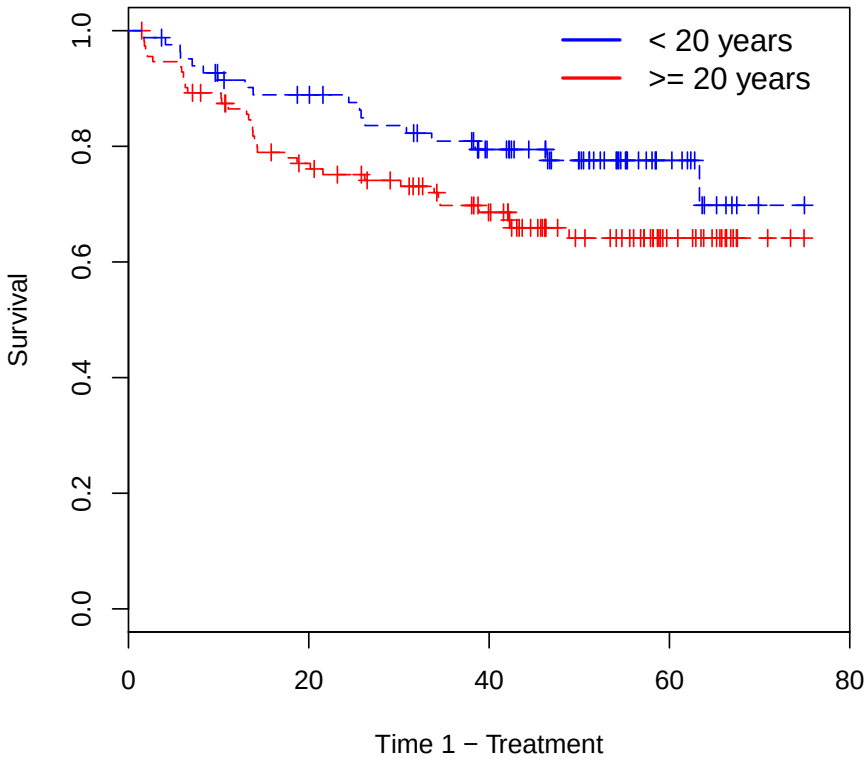
The diabetic retinopathy study

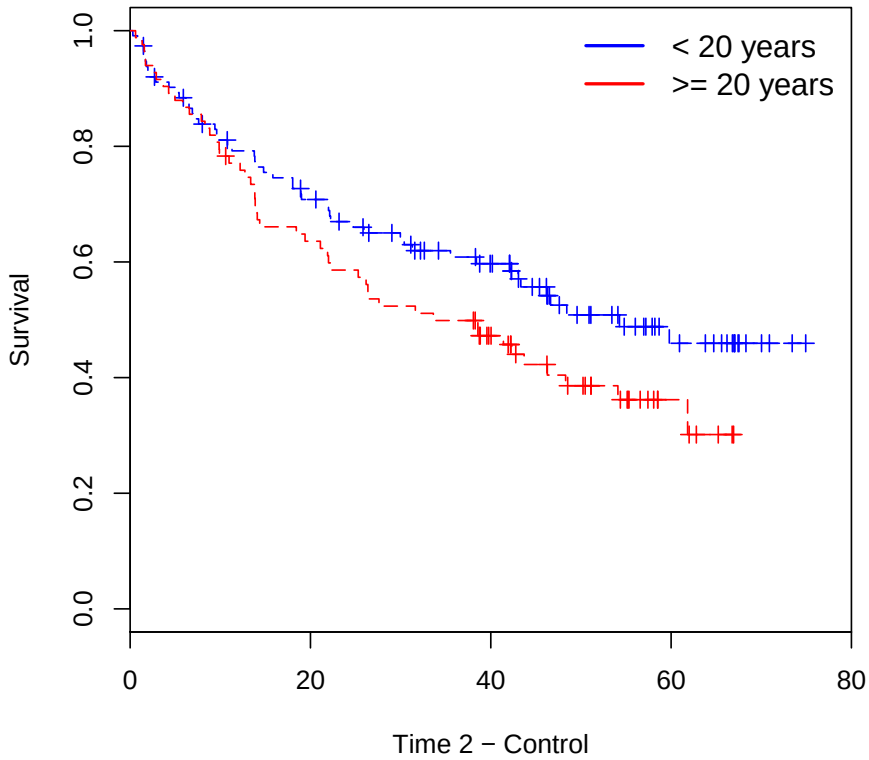
- Dataset: The Diabetic Retinopathy Study Research Group (1976)
- Nuñes, Tanaka and Pedroso de Lima (2006)
- One eye randomly received the treatment.
- The other eye was considered as a control.
- Censoring was caused by death, dropout or end of the study.
- Total of 197 patients were considered in the study.
- The subjects could be censored, which happened for 73% of the treated eyes and 49% of the untreated eyes.
- Age was considered as a covariate to create two groups, with a cutoff point of 20 years (58% of the subjects were less than 20-years-old).

The diabetic retinopathy study



The diabetic retinopathy study





- Presence of long-term survivals.

Some Review – Formulation

- There are two prevalent formulations of cure rate models in the literature.
  - The Standard Mixture Cure Model (Boag, 1949; Berkson & Gage, 1952), where the number of causes of the event of interest is a binary random variable on $\{0, 1\}$,
  - The Promotion Time Cure Model (Yakovlev & Tsodikov, 1996), where this number follows a Poisson distribution.
- Although extensions of cure rate models were developed, limited attention has been paid to the research on multivariate cure rate models.

Some Review – Approaches

- In the frequentist framework, extensions of the standard mixture cure model:
  - Chatterjee & Shih (2001) proposed a marginal approach using bivariate copula models.
  - Price & Manatunga (2001) imposed frailty to account for correlation and conducted the maximum likelihood estimation under a parametric model assumption.
- In the Bayesian approach, Chen *et al.* (2002) generalized the work of Yakovlev & Tsodikov (1996) to multivariate failure time data by introducing a positive stable frailty.
- Louzada *et al.* (2013) proposed bivariate long-term distribution based on the Farlie-Gumbel-Morgenstern copula model.

Our Proposal

- We propose a new lifetime model for multivariate survival data in presence of surviving fractions and examine some of its properties.
- Its genesis is based on situations in which there are m types of unobservable competing causes, where each cause is related to a time of occurrence of an event of interest.
- Our model is a multivariate extension of the univariate survival cure rate model proposed by Tsodikov *et al.* (2003) and Rodrigues *et al.* (2009).

Practical Motivation – Diabetic Retinopathy Study

Presence of:

- long term survivals
- competing severe visual loss causes & which are latent
- Although the cause of severe visual loss may be latent, we can conjecture some possible competitive causes (Fong, et al. 1999) for it:
  - 1- Vitreous or preretinal hemorrhage
  - 2- Macular edema or macular pigmentary changes related to macular edema
  - 3- Macular or retinal detachment
  - 4- Neovascular glaucoma
  - 5- Retinal detachment

Contents

The Model

- For an individual in the population, let N_k $(k = 1, \ldots, m)$ be the random variable that denotes the unobservable number of causes of type k of the event of interest for this individual.
- We assume that $\boldsymbol{N} = (N_1, N_2, \ldots, N_m)$ follows a multivariate Poisson distribution with probability mass function

$$P\left[N_1 = n_1, \ldots, N_m = n_m\right] = e^{-\left\{\sum_{i=1}^{m} \theta_i\right\}} \prod_{i=1}^{m} \frac{\theta_i^{n_i}}{n_i!} \sum_{i=0}^{s} \prod_{j=1}^{m} \binom{n_j}{i!} i! \left(\frac{\theta_0}{\prod_{i=1}^{m} \theta_i}\right)^i \tag{1}$$

where $n_j = 0, 1, \ldots,$ $\theta_j > 0$, $j = 0, 1, \ldots, m$ and $s = \min\{n_1, \ldots, n_m\}$.

The Model

- The multivariate distribution in (1) allows for positive dependence between the two random variables.
- Marginally each random variable follows a Poisson distribution with

$$E(N_j) = \theta_j + \theta_0$$

and

$$\text{Cov}(N_i, N_j) = \theta_0,$$

for $i \neq j = 1, \ldots, m$.

- Hence θ_0 is a measure of dependence between the two random variables.

The Model

- The time for the jth competing cause of type k to produce the event of interest is denoted by Z_{kj}, $k = 1, \ldots, m$, $j = 1, 2, \ldots$.
- Given $N_k = n_k$, the $Z_{k1} \ldots, Z_{kn_k}$ are independent and identically distributed random variables with cumulative distribution function $F_k(\cdot) = 1 - S_k(\cdot)$.
- The observed times to event are defined by the random variables

$$Y_k = \min \{Z_{k0}, Z_{k1}, \ldots Z_{kN_k}\}$$

with $P(Z_{k0} = \infty) = 1$, k=1,2,. . . ,m.

The Model

- Under this setup we can demonstrate, that the population survival function for $\boldsymbol{Y} = (Y_1, \ldots, Y_m)$ is given by

$$S_{\text{pop}}(\boldsymbol{y}) \quad = \quad \exp\left\{ -\sum_{i=1}^{m} \theta_i (1 - S_i(y_i)) - \theta_0 (1 - \prod_{i=1}^{m} S_i(y_i)) \right\}. \quad (2)$$

- The survival function $S_{\text{pop}}(\boldsymbol{y})$ in (2) is not a proper survival, that is, $\lim_{y_1,\ldots,y_m \to \infty} S_{\text{pop}}(\boldsymbol{y}) = \exp\{-\sum_{i=0}^{m} \theta_i\} > 0$ (the joint cure fraction).
- Note that when $\theta_0 = 0$ in (2), the joint survival function reduces to the product of m independent survival functions.

The Model

- From (2) the marginal survival functions are

$$S_{pop}(y_k) = \exp\left\{-(\theta_k + \theta_0)F_k(y_k)\right\}, \ k = 1, \ldots, m. \qquad (3)$$

- Equation (3) indicates that the marginal survival function has a cure rate structure with probability of cure $p_{0k} = e^{-\theta_k - \theta_0}$ for Y_k, $k = 1, \ldots, m$.
- It is important to note in (3) that each marginal survival function has the structure of the promotion time cure model (Yakovlev & Tsodikov, 1996; Chen *et al.*, 1999).

The model

- In (3) that each marginal distribution has a proportional hazards structure as long as the covariates, only enter through θ_k and θ_0.
- The marginal hazard function is given by, $(\theta_k + \theta_0) f_k(y_k)$ which satisfies the conditions for the proportional hazards model (Cox & Oakes, 1984).
- This is a desirable feature of the proposed model that leads to attractive theoretical properties.

The Model – The Bivariate Case – Local measure of association

- Without loss of generality, considering the bivariate distribution of (Y_1, Y_2), then joint survival function is given by

$$S_{\text{pop}}(y_1, y_2) = \exp\left\{-\theta_1(1 - S_1(y_1)) - \theta_2(1 - S_2(y_2)) - \theta_0(1 - S_1(y_1)S_2(y_2))\right\}. \tag{4}$$

- The parameter θ_0 is a measure of association between (Y_1, Y_2).
- As $\theta_0 \to 0$, this implies less association between (Y_1, Y_2) which can be seen from (4).

The Model – The Bivariate Case

- Following Clayton (1978) and Oakes (1982), we can compute a local measure of association, denoted by $\vartheta^*(Y_1, Y_2)$, as a function of θ_0.
- This measure of association is defined as

$$\vartheta^*(Y_1, Y_2) = \frac{S_{\text{pop}}(y_1, y_2)\frac{\partial^2}{\partial y_1 \partial y_2} S_{\text{pop}}(y_1, y_2)}{\left(\frac{\partial S_{\text{pop}}(y_1, y_2)}{\partial y_1}\right)\left(\frac{\partial S_{\text{pop}}(y_1, y_2)}{\partial y_2}\right)}. \tag{5}$$

The model – The Bivariate Case

- For the bivariate cure rate model in (3), $\vartheta^*(y_1, y_2)$ is well defined and is given by

$$\vartheta^*(y_1, y_2) = 1 + \theta_0 \left\{ [\theta_1 + \theta_0 S_2(y_2)] [\theta_2 + \theta_0 S_1(y_1))] \right\}^{-1}. \quad (6)$$

- We see that $\vartheta^*(y_1, y_2)$ in (6) increases in (y_1, y_2). That is, the association between (y_1, y_2) is less when (y_1, y_2) are small and the association increases over time.

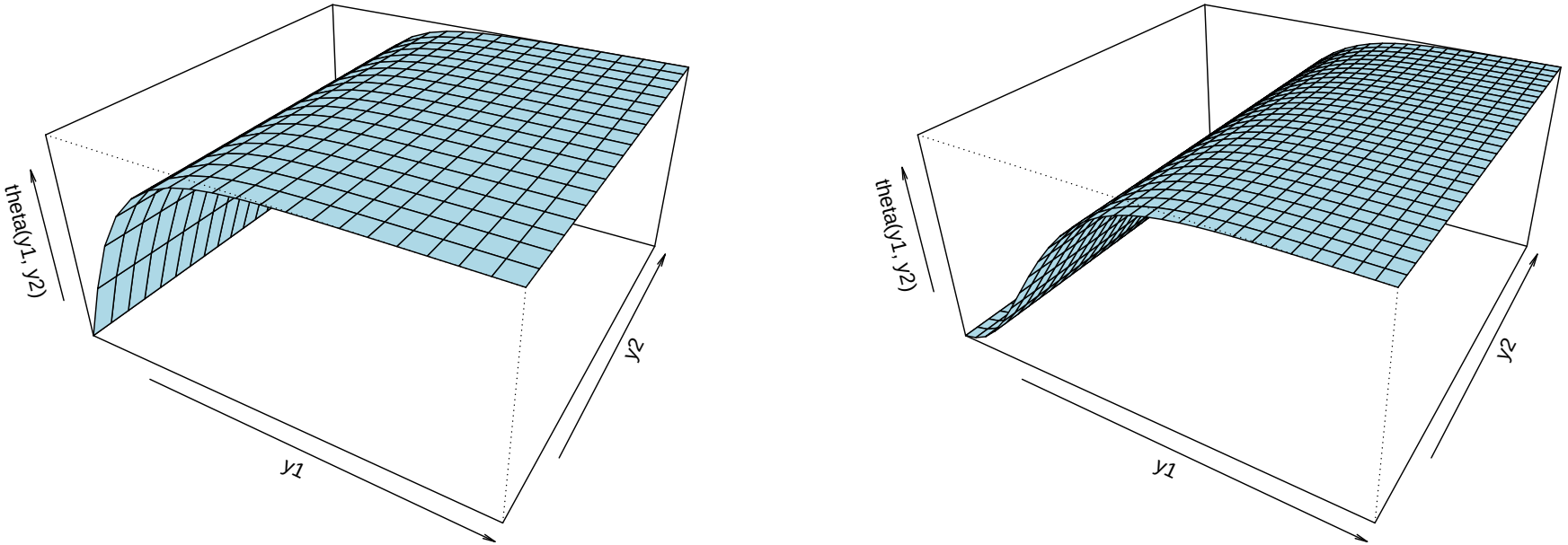


Figure: Local measure of association for model with $\theta_1 = 0.2$, $\theta_2 = 0.3$, $\theta_0 = 0.2$ (left panel) and $\theta_0 = 2$ (right panel)

- The following Theorem is a generalization of the proposed model in (2) and provides a natural extension of the univariate survival cure rate models.

Theorem (1)

Let $\boldsymbol{N} = (N_1, N_2, \ldots, N_m)$ be a random vector with probability generating function, $\varphi_{\boldsymbol{N}}(w_1, \ldots, w_m)$ and random vector $\boldsymbol{Y} = (Y_1, \ldots, Y_m)$ defined above. Then the joint survival function of $\boldsymbol{Y}$ is given by

$$S_{\text{pop}}(\boldsymbol{y}) = \varphi_{\boldsymbol{N}}(S_1(y_1), \ldots, S_m(y_m)). \tag{7}$$

Contents

Inference for the bivariate survival cure rate model

- Let us consider the situation when the failure times (Y_1, Y_2) are not completely observed and are subject to right censoring.
- Let C_{ki} denote the censoring time of k component, $k = 1, 2$.
- Suppose that $(Y_{1i}; Y_{2i})$ and $(C_{1i}; C_{2i})$ are independent.
- For each individual i, observed quantities are represented by the random variables $t_{ki} = \min\{Y_{ki}, C_{ki}\}$ and $\delta_{ki} = I(Y_{ki} < C_{ki})$, which denotes a censorship indicator, $k = 1, 2$, $i = 1, \ldots, n$.

- Let $\boldsymbol{x}_{ki}$ (k=1,2,) denote the vectors of covariates for the i^{th} individual.
- Extending our model, we propose to relate the parameters θ_{1i} and θ_{2i} of the bivariate Poisson distribution to the covariates by the logarithmic link

$$\log(\theta_{1i}) = \boldsymbol{x}_{1i}^{\top}\boldsymbol{\beta_1} \quad \text{and} \quad \log(\theta_{2i}) = \boldsymbol{x}_{2i}^{\top}\boldsymbol{\beta_2}, \tag{8}$$

- where $\boldsymbol{\beta}_k = (\beta_{k1}, \ldots, \beta_{kp_k})^{\top}$ is the vector of regression coefficients associated with the covariates $\boldsymbol{x}_{ki}$.

- Now with the expression (8) we can express the likelihood of $\boldsymbol{\vartheta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \theta_0)$ under non-informative censoring as,

$$
\begin{aligned}
L(\boldsymbol{\vartheta}|\boldsymbol{\mathcal{D}}) = & \prod_{i=1}^{n} S_{\text{pop}}(t_{1i}, t_{2i}) \prod_{k=1}^{2} [f_k(t_{ki}|\boldsymbol{\gamma}_k)]^{\delta_{ki}} \\
& \times [\theta_0 + (\theta_{2i} + \theta_0 S_1(t_{1i}|\boldsymbol{\gamma}_1))(\theta_{1i} + \theta_0 S_2(t_{2i}|\boldsymbol{\gamma}_2))]^{\delta_{1i}\delta_{2i}} \\
& \times (\theta_{1i} + \theta_0 S_2(t_{2i}|\boldsymbol{\gamma}_2))^{\delta_{1i}(1-\delta_{2i})} (\theta_{2i} + \theta_0 S_1(t_{1i}|\boldsymbol{\gamma}_1))^{\delta_{2i}(1-\delta_{1i})},
\end{aligned}
\tag{9}
$$

where $S_{\text{pop}}(t_1, t_2)$ is survival function given in (4) and $f_k(t_{ki}|\boldsymbol{\gamma}_k)$ and $S_k(t_{ki}|\boldsymbol{\gamma}_k)$, $k = 1, 2$ are density and survival functions of the Weibull distribution.

Prior and Posterior

- The normal distribution and gamma distribution with a as shape and b as scale (and mean a/b) are denoted by $N(\mu, \sigma^2)$ and $G(a, b)$.
- In this context we assume that $\boldsymbol{\beta}_k$, $k = 1, 2$, $\boldsymbol{\gamma}_{k1}$, γ_{k2} and θ_0 are *a priori* independent, that is,

$$\pi(\boldsymbol{\vartheta}) = \prod_{k=1}^{2} \pi(\boldsymbol{\beta}_k)\pi(\boldsymbol{\gamma}_{k1})\pi(\gamma_{k2})\pi(\theta_0), \tag{10}$$

- where $\boldsymbol{\beta}_k \sim N_{p_k}(\mathbf{0}, \Sigma_{k0})$, $\boldsymbol{\gamma}_{k1} \sim \mathrm{G}(a_{k0}, b_{k0})$, $\gamma_{k2} \sim N_1(0, \sigma^2_{\gamma_{k2}})$ $k = 1, 2$ and $\theta_0 \sim \mathrm{G}(a_0, b_0)$.

Prior and Posterior

- Combining the likelihood function (9) and the prior distribution in (10), the joint posterior distribution for $\boldsymbol{\vartheta}$ is obtained as

$$\pi(\boldsymbol{\vartheta}|\boldsymbol{\mathcal{D}}) \propto L(\boldsymbol{\vartheta}|\boldsymbol{\mathcal{D}}) \prod_{k=1}^{2} \pi(\boldsymbol{\beta}_k)\pi(\gamma_{k1})\pi(\gamma_{k2})\phi(\theta_0).$$

- This joint posterior density is analytically intractable.
- Thus the computational problem can be easily handled by using MCMC methods for sampling from the posterior distribution.

Prior and Posterior

- The full conditional distributions of $\boldsymbol{\beta}_k$, $\boldsymbol{\gamma}_k = (\gamma_{1k}, \gamma_{2k})$ and θ_0 as

$$\pi(\boldsymbol{\beta}_k \mid \cdot) \propto \exp[-\sum_{i=1}^{n} \theta_{ki} F_k(t_{ki}|\boldsymbol{\gamma}_k)] \prod_{i=1}^{n} \Delta_i \Lambda_{ki} \, \pi(\boldsymbol{\beta}_k), \quad k = 1, 2, \tag{11}$$

$$\begin{aligned} \pi(\boldsymbol{\gamma}_1 \mid \cdot) \propto \exp[-\sum_{i=1}^{n} (\theta_{1i} F_1(t_{1i}|\boldsymbol{\gamma}_1) - \theta_0 S_1(t_{1i}|\boldsymbol{\gamma}_1) S_2(t_{2i}|\boldsymbol{\gamma}_2))] \\ \times \prod_{i=1}^{n} f_1(t_{1i}|\boldsymbol{\gamma}_1) \Delta_i \, \Lambda_{2i} \pi(\boldsymbol{\gamma}_1), \end{aligned} \tag{12}$$

$$\pi(\gamma_2 \mid \cdot) \propto \exp[-\sum_{i=1}^{n}(\theta_{2i}F_2(t_{2i}|\gamma_2) - \theta_0 S_1(t_{1i}|\gamma_1)S_2(t_{2i}|\gamma_2))] \times \prod_{i=1}^{n} f_2(t_{2i}|\gamma_2)\Delta_i \, \Lambda_{1i}\pi(\gamma_2), \tag{13}$$

and

$$\pi(\theta_0 \mid \cdot) \propto \exp[-\theta_0 S_1(t_{1i}|\gamma_1)S_2(t_{2i}|\gamma_2))] \prod_{i=1}^{n} \Lambda_{1i} \, \Delta_i \, \Lambda_{2i} \, \pi(\theta_0), \tag{14}$$

where $\Delta_i = [\theta_0 + (\theta_{2i} + \theta_0 S_1(t_{1i}|\gamma_1))(\theta_{1i} + \theta_0 S_2(t_{2i}|\gamma_2))]^{\delta_{1i}\delta_{2i}}$,
$\Lambda_{1i} = (\theta_{1i} + \theta_0 S_2(t_{2i}|\gamma_2))^{\delta_{1i}(1-\delta_{2i})}$ and $\Lambda_{2i} = (\theta_{2i} + \theta_0 S_1(t_{1i}|\gamma_1))^{\delta_{2i}(1-\delta_{1i})}$.

- We then implement a Metropolis-Hasting algorithm within Gibbs iterations (Chib & Greenberg, 1995).
- For example, to implement the Metropolis-Hastings algorithm for the parameter θ_0, we consider a target distribution $g_{\theta_0}(\theta_0) = \pi(\theta_0|\cdot)$, and under given model, $\theta_0 > 0$ we consider the transformation $\theta_0 = \exp(\eta)$, where, $-\infty < \eta < \infty$.
- Then, $g_\eta(\eta) = g_{\theta_0}(\eta)e^\eta$.
- Instead of directly sampling θ_0, we generate η by choosing a normal proposal $N(\hat{\eta}, \sigma^2_{\hat{\eta}})$ where $\hat{\eta}$ is the maximizer of the logarithm of $g_\eta(\eta)$ and $\sigma^2_{\hat{\eta}}$ is the minus of the inverse of the second derivative of logarithm of $g_\eta(\eta)$ evaluated in $\eta = \hat{\eta}$.

The algorithm

- The algorithm to generate η operates as follows:
  (1) let η be the current value;
  (2) generate a point η^* from $N(\hat{\eta}, \sigma^2_{\hat{\eta}})$;
  (3) a move from η to η^* is made with probability

$$\min\left\{1, \frac{g_\eta(\eta^*)\phi\left(\frac{\eta-\hat{\eta}}{\sigma_{\hat{\eta}}}\right)}{g_\eta(\eta)\phi\left(\frac{\eta^*-\hat{\eta}}{\sigma_{\hat{\eta}}}\right)}\right\},$$

  where $\phi(\cdot)$ is the standard normal probability density function.

- After we sample η we obtain $\theta_0 = e^\eta$.

Contents

Simulation – General Specifications

- Aiming 1: To evaluate the asymptotic properties of the parameter estimates for the proposed cure rate model.
- Aiming 2: To analyse the frequentist coverage probabilities of credible interval derived from the posteriors

Simulation – General Specifications

- In this study we consider the bivariate cure rate model with the Weibull distribution for the event time (Z_{kj}, $k = 1, 2$, $j = 1, 2, \ldots$.), with parameter, $\gamma_{k1} = 1.4$ and $\gamma_{k2} = 2.0$.
- For each individual i, $i = 1, \ldots, n$, the number of causes of the event of interest for this individual (N_1, N_2) is generated from the bivariate Poisson distribution with parameter $\theta_0 = 0.5$, $\theta_{1i} = \exp(\beta_{10} + \beta_{11} x_i)$ and
- $\theta_{2i} = \exp(\beta_{20} + \beta_{21} x_i)$, where $\beta_{10} = -1.6$, $\beta_{11} = 1.3$, $\beta_{20} = -1.5$, $\beta_{21} = 1.1$ and the covariates x_i are generated from a Bernoull distribution with parameter 0.5.

Simulation – General Specifications

- The censoring times C_{ki} are sampled from the uniform distribution on the interval $(0, \tau_k\ )$, where τ_k is set in order to control the proportion of censored observations.
- In this study the proportion of censored observations was on an average approximately equal to 55% and 50% respectively.

Simulation – General Specifications

- The sample sizes are taken as $n = 200$ and 400.
- For each simulated data set, the posterior summaries and 95% HPD intervals of the model parameters were obtained.
- 80,000 MCMC posterior samples are generated for each paramenter, from which 20,000 iterations are eliminated for obtaining a sample of size 50,000.
- The autocorrelation of these sampled values are reduced by taking a spacing of size 10, thus resulting in a final sample of size 6,000.

Simulation – General Specifications

- For each configuration, we conduct 500 replicates and then we determine from the estimates of each parameter, the average (AE), the root mean square error (RMSE) and the coverage probability (PC).

Simulation – Results

Table: Simulation results with 500 replications of the posterior summaries.

		Parameter								
n		β_{10}	β_{11}	γ_{11}	γ_{12}	β_{20}	β_{21}	γ_{21}	γ_{12}	θ_0
200	AE	-1.489	1.201	2.018	1.409	-1.482	1.060	2.030	1.410	0.489
	SD	0.348	0.381	0.155	0.161	0.368	0.345	0.154	0.153	0.068
	RMSE	0.368	0.401	0.156	0.163	0.373	0.361	0.157	0.1549	0.072
	PC	0.942	0.943	0.949	0.942	0.953	0.950	0.948	0.949	0.945
400	AE	-1.592	1.283	2.008	1.391	-1.501	1.010	2.021	1.403	0.502
	SD	0.268	0.252	0.133	0.130	0.368	0.256	0.143	0.139	0.034
	RMSE	0.269	0.254	0.137	0.131	0.373	0.258	0.145	0.141	0.034
	PC	0.951	0.948	0.952	0.948	0.948	0.951	0.949	0.951	0.952

Contents

The diabetic retinopathy data

- The diabetic retinopathy study (Huster *et al.*, 1989) of time of blindnes in each eye of the 197 diabetic patients with diabetic Retinopathy.
- One eye of each patients was randomly selected for treatment (the effectiveness of laser photo coagulation in delaying the onsets of blindness) and other eye was observed without treatment.
- A binary age covariate (0 for juvenile and 1 for adult) is available.
- The first component of the bivariate survival time is the time of the blindness on the treated eye (Y_1) and the second component is the similar time for the untreated eye (Y_2).

- For model fitting the following independent priors are adopted in the Bayesian computations $\beta_{kj} \sim N(0, 10^4)$, $k = 1, 2$ and $j = 0, 1$, $\gamma_{k1} \sim \mathrm{G}(1, 0.01)$ and $\gamma_{k2} \sim N(0, 10^4)$.
- 80,000 MCMC posterior samples are generated for each parameter, from which the 20,000 iterations are eliminated for obtaining a sample of size 60,000.
- The autocorrelation of these sampled values are reduced by taking a spacing of size 10 thus resulting in a final sample of size 6,000.

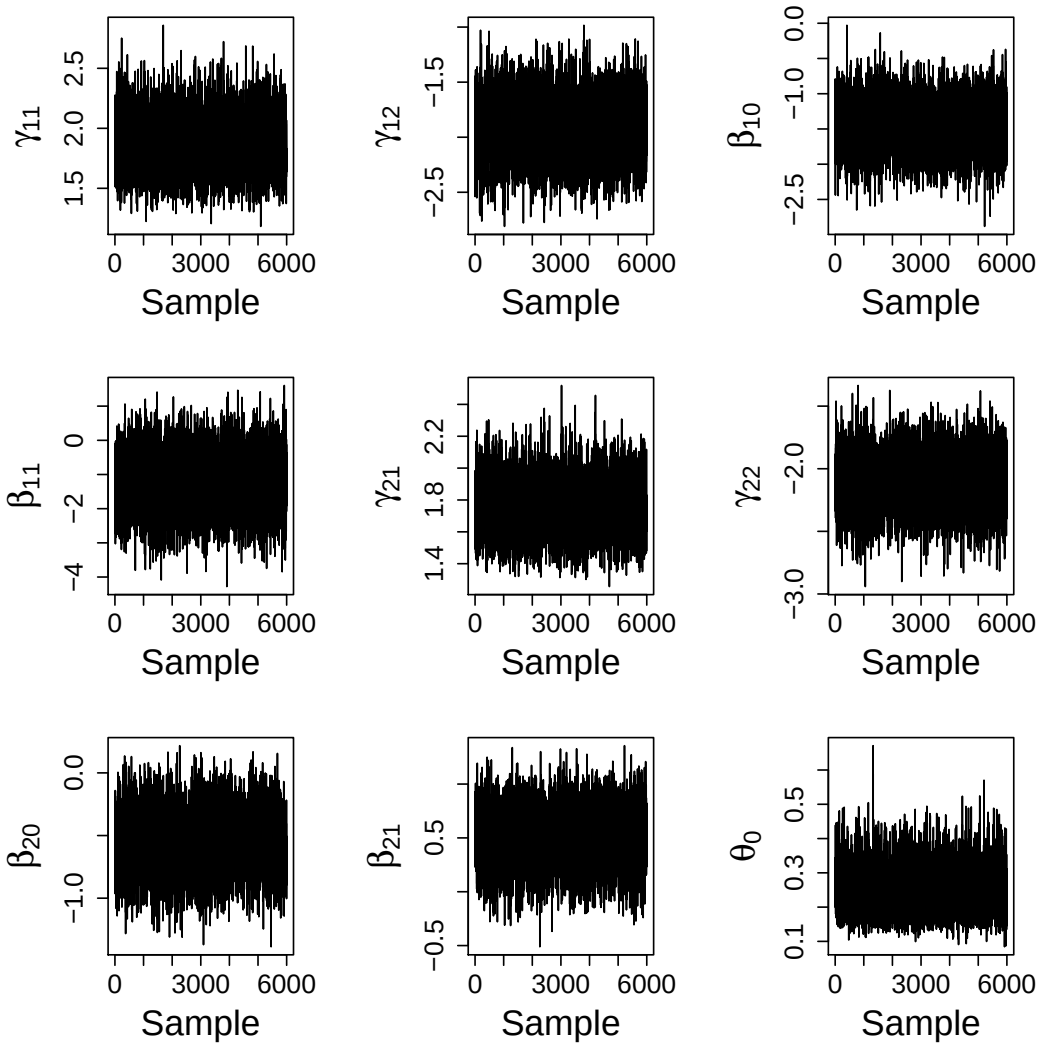


Figure: Sequence plots of the chains

Table: Posterior summaries of the parameters for the bivariate cure rate model for the diabetic retinopathy data set.

				95%HPD	
Time to blindness	Parameter	Mean	Standard Desviation	LI	LS
Treated eye	γ_{11}	1.869	0.2217	1.473	2.335
	γ_{12}	-1.934	0.2633	-2.440	-1.420
	β_{10}	-1.449	0.347	-2.129	-0.776
	β_{11}	-1.216	0.8207	-2.854	0.380
Untreated eye	γ_{21}	1.746	0.165	1.443	2.101
	γ_{22}	-2.112	0.2113	-2.529	-1.699
	β_{20}	-0.563	0.221	-0.997	-0.129
	β_{21}	0.488	0.251	-0.012	0.976
	θ_0	0.247	0.064	0.145	0.393

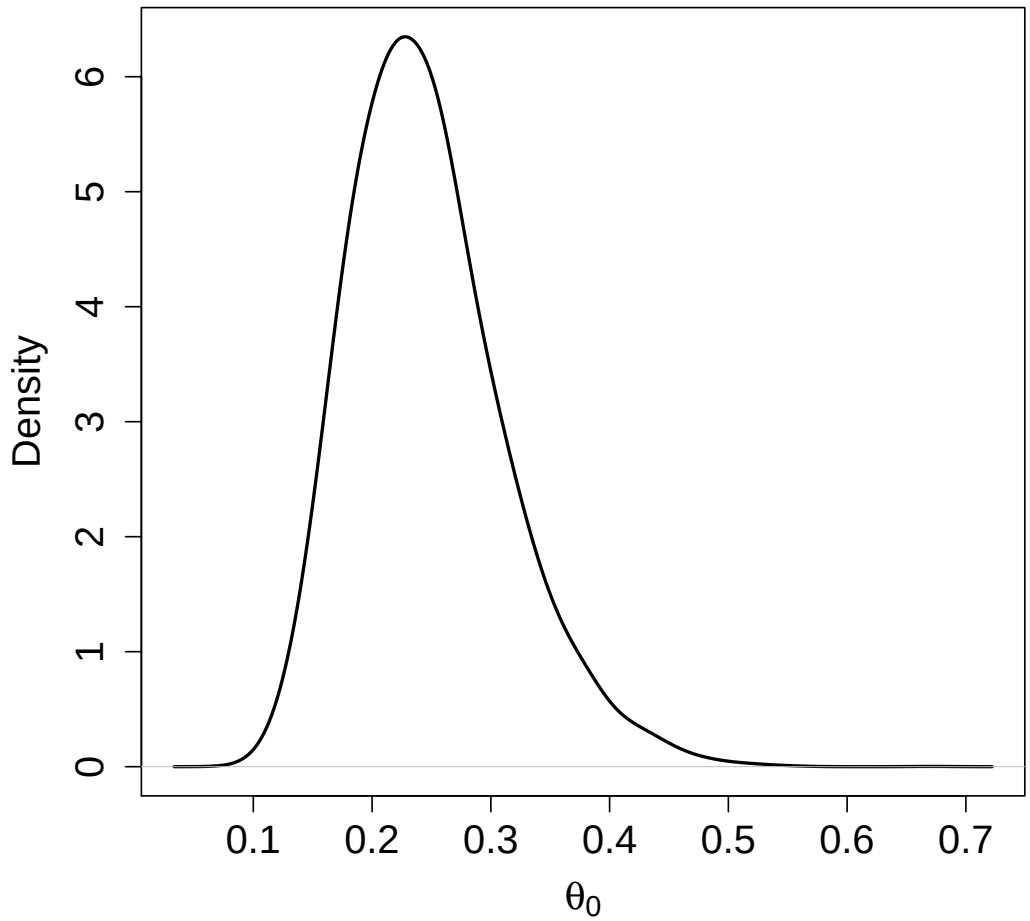


Figure: The marginal posterior density of θ_0

Table: Posterior summaries of the cured fraction stratified by juvenile and adult patients.

				95%HPD	
Eye	Patient	Mean	Standard Desviation	LI	LS
Treated	Juvenile	0.611	0.053	0.498	0.704
	Adult	0.710	0.063	0.565	0.805
Untreated	Juvenile	0.430	0.059	0.320,	0.551
	Adult	0.309	0.060	0.195	0.428

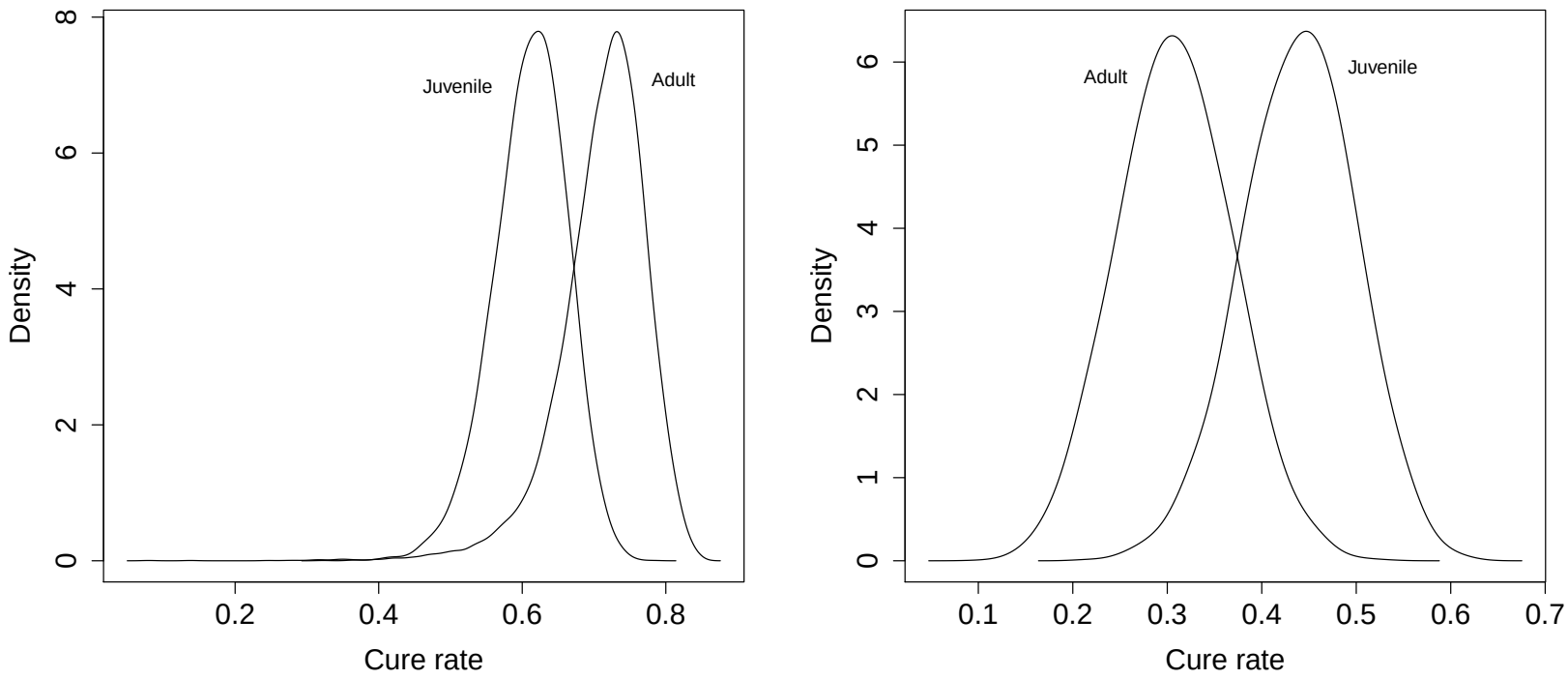


Figure: Posterior density estimates for the cure rates corresponding to treated eye (left panel) and untreated eye (rigth panel) patients.

Contents

Final Comments

The new lifetime model for multivariate survival data

- Incorporates in to the analysis important characteristics of the problem:
  - long term survivals
  - latent competing default causes
- The estimation procedure of the proposed model is straightforward via Bayesian Inference.
- The results of this study have been condensed in a paper accepted for publication in the Journal of Statistical Computation and Simulation (JSCS).

THANK YOU FOR YOUR ATTENTION !!!

Contents

References

Chatterjee, N. & Shih, J. (2001). A bivariate cure-mixture approach for modeling familial association in diseases. *Biometrics*, **57**(3), 779–786.

Chen, M.-H., Ibrahim, J. G. & Sinha, D. (1999). A new Bayesian model for survival data with a surviving fraction. *Journal of the American Statistical Association*, **94**, 909–919.

Chen, M.-H., Ibrahim, J. G. & Sinha, D. (2002). Bayesian inference for multivariate survival data with a cure fraction. *Journal of Multivariate Analysis*, **80**(1), 101–126.

Chib, S. & Greenberg, E. (1995). Hierarchical analysis of sur models with extensions to correlated serial errors and time-varying parameter models. *Journal of Econometrics*, **68**(2), 339–360.

Clayton, D. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, **65**(1), 141–151.

Cox, D. & Oakes, D. (1984). *Analysis of Survival Data*. Chapman & Hall, London.

Fong DS1, Ferris FL 3rd, Davis MD, Chew EY.

Fong DS, Ferris FL, Davis MD & Chew EY. (1999). Causes of severe visual loss in the early treatment diabetic retinopathy study: ETDRS report no. 24. Early Treatment Diabetic Retinopathy Study Research Group. *Am J Ophthalmol*, 127(2): 137–141.

Huster, W., Brookmeyer, R. & Self, S. (1989). Modelling paired survival data with covariates. *Biometrics*, pages 145–156.

Louzada, F., Suzuki, A. & Cancho, V. (2013). The FGM long-term bivariate survival copula model: Modeling, Bayesian estimation, and case influence diagnostics. *Communications in Statistics–Theory and Methods*, **42**(4), 673–691.

Oakes, D. (1982). A model for association in bivariate survival data. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 414–422.

Price, D. & Manatunga, A. (2001). Modelling survival data with a cured fraction using frailty models. *Statistics in medicine*, **20**(9-10), 1515–1527.

Rodrigues, J., Cancho, V. G., de Castro, M. & Louzada-Neto, F. (2009). On the unification of the long-term survival models.