General Families of Cure Rate Models and Some of its Properties

Fotios S. Milienos
Department of Sociology, Panteion University of Social and Political Sciences

Athens University of Economics and Business
Athens,2020

Outline

Introduction

It was late 1940s when researchers studying the statistical aspects of the effectiveness of cancer treatments realized that a non negligible proportion of patients, after undergoing an appropriate treatment, were not faced any recurrence of the disease, no matter how long they have been followed up (Boag, 1949).

Introduction

It was late 1940s when researchers studying the statistical aspects of the effectiveness of cancer treatments realized that a non negligible proportion of patients, after undergoing an appropriate treatment, were not faced any recurrence of the disease, no matter how long they have been followed up (Boag, 1949).

The estimation of this proportion became of great importance while the existed estimation approaches which were based on the number of patients who remained symptoms free for few years (typically, three to five years), proved to be ineffective due to, for example, the long delay, the high average age of patients, etc.

Introduction

The aim of cure rate models is the study of survival times or generally, the times till the occurrence of an event.

This event may be:

- the failure of a unit;
- the occurrence/recurrence of a specific problem in a system;
- the occurrence/recurrence of a disease;
- to find the first job after graduation.

Introduction

The aim of cure rate models is the study of survival times or generally, the times till the occurrence of an event.

This event may be:

- the failure of a unit;
- the occurrence/recurrence of a specific problem in a system;
- the occurrence/recurrence of a disease;
- to find the first job after graduation.

Using the models found in traditional survival/reliability/event history analysis, every item will experience the event of interest, at some point.

Introduction

The aim of cure rate models is the study of survival times or generally, the times till the occurrence of an event.

This event may be:

- the failure of a unit;
- the occurrence/recurrence of a specific problem in a system;
- the occurrence/recurrence of a disease;
- to find the first job after graduation.

Using the models found in traditional survival/reliability/event history analysis, every item will experience the event of interest, at some point.

However, cure rate models allow for a proportion of items which will never experience the event of interest.

Introduction

Therefore, in cure rate modeling, the population survival function $S_P(t)$ tends to level off at a value greater than zero, as time t goes to infinity.

.

Introduction

Therefore, in cure rate modeling, the population survival function $S_P(t)$ tends to level off at a value **greater than zero**, as time t goes to infinity.

.

Introduction

Therefore, in cure rate modeling, the population survival function $S_P(t)$ tends to level off at a value **greater than zero**, as time t goes to infinity.

Equivalently, the corresponding population cumulative hazard function $H_P(t)$ is bounded, i.e. $\lim_{t\to\infty} H_P(t) = \theta \in (0, \infty)$.

.

Introduction

Therefore, in cure rate modeling, the population survival function $S_P(t)$ tends to level off at a value **greater than zero**, as time t goes to infinity.

Equivalently, the corresponding population cumulative hazard function $H_P(t)$ is bounded, i.e. $\lim_{t\to\infty} H_P(t) = \theta \in (0,\infty)$.

Hence,

$$S_P(t) = e^{-H_P(t)} = e^{-\theta(1-S(t))},$$

for some (proper) survival function $S(t)$. Note: $\lim_{t\to\infty} S_P(t) = \exp(-\theta)$.

.

Introduction

Therefore, in cure rate modeling, the population survival function $S_P(t)$ tends to level off at a value **greater than zero**, as time t goes to infinity.

Equivalently, the corresponding population cumulative hazard function $H_P(t)$ is bounded, i.e. $\lim_{t\to\infty} H_P(t) = \theta \in (0,\infty)$.

Hence,

$$S_P(t) = e^{-H_P(t)} = e^{-\theta(1-S(t))},$$

for some (proper) survival function $S(t)$. Note: $\lim_{t\to\infty} S_P(t) = \exp(-\theta)$.

Known as: bounded cumulative hazard (BCH) or promotion time cure rate model.

Introduction

Therefore, in cure rate modeling, the population survival function $S_P(t)$ tends to level off at a value **greater than zero**, as time t goes to infinity.

Equivalently, the corresponding population cumulative hazard function $H_P(t)$ is bounded, i.e. $\lim_{t\to\infty} H_P(t) = \theta \in (0,\infty)$.

Hence,

$$S_P(t) = e^{-H_P(t)} = e^{-\theta(1-S(t))},$$

for some (proper) survival function $S(t)$. Note: $\lim_{t\to\infty} S_P(t) = \exp(-\theta)$.

Known as: **bounded cumulative hazard (BCH)** or promotion time cure rate model.

Introduction

Therefore, in cure rate modeling, the population survival function $S_P(t)$ tends to level off at a value **greater than zero**, as time t goes to infinity.

Equivalently, the corresponding population cumulative hazard function $H_P(t)$ is bounded, i.e. $\lim_{t\to\infty} H_P(t) = \theta \in (0,\infty)$.

Hence,

$$S_P(t) = e^{-H_P(t)} = e^{-\theta(1-S(t))},$$

for some (proper) survival function $S(t)$. Note: $\lim_{t\to\infty} S_P(t) = \exp(-\theta)$.

Known as: **bounded cumulative hazard (BCH)** or **promotion time cure rate model**.

Introduction

Therefore, in cure rate modeling, the population survival function $S_P(t)$ tends to level off at a value **greater than zero**, as time t goes to infinity.

Equivalently, the corresponding population cumulative hazard function $H_P(t)$ is bounded, i.e. $\lim_{t\to\infty} H_P(t) = \theta \in (0,\infty)$.

Hence,

$$S_P(t) = e^{-H_P(t)} = e^{-\theta(1-S(t))},$$

for some (proper) survival function $S(t)$. Note: $\lim_{t\to\infty} S_P(t) = \exp(-\theta)$.

Known as: **bounded cumulative hazard (BCH)** or **promotion time cure rate model**.

e.g. Cantor and Shuster (1992, SIM), Yakovlev, Cantor and Shuster (1994, SIM), Tsodikov (1998a, BM), and Tsodikov, Ibrahim and Yakovlev (2003, JASA).

Introduction

Therefore, in cure rate modeling, the population survival function $S_P(t)$ tends to level off at a value **greater than zero**, as time t goes to infinity.

Equivalently, the corresponding population cumulative hazard function $H_P(t)$ is bounded, i.e. $\lim_{t\to\infty} H_P(t) = \theta \in (0,\infty)$.

Hence,

$$S_P(t) = e^{-H_P(t)} = e^{-\theta(1-S(t))},$$

for some (proper) survival function $S(t)$. Note: $\lim_{t\to\infty} S_P(t) = \exp(-\theta)$.

Known as: **bounded cumulative hazard (BCH)** or **promotion time cure rate model**.

e.g. Cantor and Shuster (1992, SIM), Yakovlev, Cantor and Shuster (1994, SIM), Tsodikov (1998a, BM), and Tsodikov, Ibrahim and Yakovlev (2003, JASA).

Introduction

Biological motivation of the BCH model:

- let M denote the number of carcinogenic cells (clonogens; **competing causes**) left active after a treatment;
- assume that M follows a Poisson distribution with mean θ;
- let $W_i, i = 1, 2, \ldots$ denote the time for the ith clonogen to produce a detectable cancer mass;
- assume that $W_i \sim S(t)$, are i.i.d. and independent of M;
- then, letting $T = \min\{W_0, W_1, \ldots, W_M\}$ (convention: $W_0 = \infty$ a.s.) be the population time-to-event, we have

$$T \sim S_P(t) = e^{-\theta(1-S(t))}.$$

The probability someone to be cured is defined as the probability of the event $M = 0$, i.e. no clonogens have survived by the end of the treatment.

Introduction

The BCH model can easily be generalized by letting M follow any discrete random variable with support $\{0, 1, \ldots\}$ and p.g.f. $\varphi(z)$.

In such a case, the distribution of $T = \min\{W_0, W_1, \ldots, W_M\}$ is given by

$$S_P(t) = \varphi(S(t)).$$

Introduction

The BCH model can easily be generalized by letting M follow any discrete random variable with support $\{0, 1, \ldots\}$ and p.g.f. $\varphi(z)$.

In such a case, the distribution of $T = \min\{W_0, W_1, \ldots, W_M\}$ is given by

$$S_P(t) = \varphi(S(t)).$$

M may follow a **negative binomial** (Castro, Cancho and Rodrigues, 2009, BMJ; Ortega et al., 2014, JDS), **geometric** (Gu, Sinha and Banerjee, 2011, LDA), **COM-Poisson** (Rodrigues, de Castro, et al., 2009, JSPI; Balakrishnan and Pal, 2013, AR-ERA), **weighted Poisson** (Rodrigues et al., 2011, LDA).

Introduction

The BCH model can easily be generalized by letting M follow any discrete random variable with support $\{0,1,\ldots\}$ and p.g.f. $\varphi(z)$.

In such a case, the distribution of $T = \min\{W_0, W_1, \ldots, W_M\}$ is given by

$$S_P(t) = \varphi(S(t)).$$

If M is a Bernoulli r.v. with $p = P(M = 0)$, then

$$S_P(t) = p + (1-p)S(t) : \textbf{standard/mixture cure rate model}$$

which can be traced back at least to the works of Boag (1948a,b, 1949, JRSS) and Berkson and Gage (1952, JASA).

Introduction

The BCH model can easily be generalized by letting M follow any discrete random variable with support $\{0,1,\ldots\}$ and p.g.f. $\varphi(z)$.

In such a case, the distribution of $T = \min\{W_0, W_1, \ldots, W_M\}$ is given by

$$S_P(t) = \varphi(S(t)).$$

If M is a Bernoulli r.v. with $p = P(M = 0)$, then

$$S_P(t) = p + (1-p)S(t) : \textbf{standard/mixture cure rate model}$$

In sociology mixture cure rate model is also referred to as *split population model* (Schmidt and Witte, 1988, Ch. 5). In engineering, is known as *limited-failure population model* (Meeker, 1987; Meeker and LuValle, 1995).

Introduction

The BCH model can easily be generalized by letting M follow any discrete random variable with support $\{0,1,\ldots\}$ and p.g.f. $\varphi(z)$.

In such a case, the distribution of $T = \min\{W_0, W_1, \ldots, W_M\}$ is given by

$$S_P(t) = \varphi(S(t)).$$

If M is a Bernoulli r.v. with $p = P(M = 0)$, then

$$S_P(t) = p + (1-p)S(t) : \textbf{standard/mixture cure rate model}$$

Motivated by the existence of a zero-inflated distribution on the number of causes, we get (Balakrishnan and Milienos, 2020, BMJ) a generalization, given by,

$$S_P(t) = p + (1-p)\varphi(S(t)) : \textbf{zero-inflated mixture model}$$

Introduction

Zeng, Yin and Ibrahim (2006, JASA) followed a frailty approach assuming that M follows a Poisson distribution with mean $\Xi\theta$ (where Ξ is a positive r.v.; $W_i's$ assumed i.i.d. given M and Ξ), and then

$$S_P(t) = E_\Xi\left[\exp(-\theta\Xi F(t))\right], \text{ with } F(t) = 1 - S(t).$$

If Ξ follows a gamma distribution with mean 1, then

$$S_P(t) = (1 + \gamma\theta F(t))^{-1/\gamma}, \gamma \geq 0, \theta > 0,$$

where γ is the scale parameter of gamma distribution .

Introduction

Koutras and Milienos (2017, SIM) motivated by the previous biological application, introduced a more flexible transformation cure rate model.

> Assume that the jth metastasis-competent tumor cell produces a detectable tumor mass only when λ distinct biological latent factors affect the cell.

Introduction

Koutras and Milienos (2017, SIM) motivated by the previous biological application, introduced a more flexible transformation cure rate model.

> Assume that the jth metastasis-competent tumor cell produces a detectable tumor mass only when λ distinct biological latent factors affect the cell.

Under this scenario, the promotion time W_j can be viewed as a maximum of λ random variables, say $W_{jk}, k = 1, \ldots, \lambda$, with $W_{jk} \sim F(t)$. Then, the population survival function reads

$$S_P(t) = (1 + \gamma\theta F(t)^{\lambda})^{-1/\gamma}, \gamma \geq 0, \theta > 0.$$

Introduction

Koutras and Milienos (2017, SIM) motivated by the previous biological application, introduced a more flexible transformation cure rate model.

> Assume that the jth metastasis-competent tumor cell produces a detectable tumor mass only when λ distinct biological latent factors affect the cell.

Under this scenario, the promotion time W_j can be viewed as a maximum of λ random variables, say $W_{jk}, k = 1, \ldots, \lambda$, with $W_{jk} \sim F(t)$. Then, the population survival function reads

$$S_P(t) = (1 + \gamma\theta F(t)^{\lambda})^{-1/\gamma}, \gamma \geq 0, \theta > 0.$$

This motivation is similar to the last-activation scheme discussed in Cooner, Banerjee, Carlin and Sinha (2007, JASA), wherein the event occurs only when a number of latent factors have been activated.

Introduction

The last model admits two other interesting alternative interpretations:

a) let M, the number of competing causes, under the same motivation with the BCH model, follow the generalized Linnik distribution with probability generating function

$$\varphi(z) = (1 + \theta\gamma(1 - z)^{\lambda})^{-1/\gamma}, \text{ with } \lambda \in (0, 1), \gamma \geq 0, \theta > 0.$$

b) let M follow a Poisson distribution with random parameter $\Xi = Y^{1/\lambda} V$, where Y, V are independent random variables with Y following a Gamma distribution with scale and shape parameter equal to $\gamma > 0$, and V being a positive random variable with Laplace transform $e^{-\theta z^{\lambda}}$.

Introduction

Yin and Ibrahim (2005, CJS), by imposing a Box-Cox transformation on the population survival function, studied a model of the same nature with that of Zeng et al. (2006) and Koutras and Milienos (2017); in their work, the population survival function was given by

$$S_P(t) = \left(1 + \gamma\theta F(t)\right)^{-1/\gamma}, \gamma \in [-1, 0],$$

where

$$\theta = \theta(\gamma; \boldsymbol{X}) = \exp(\boldsymbol{\beta}'\boldsymbol{X}) / (1 - \gamma \exp(\boldsymbol{\beta}'\boldsymbol{X})),$$

with $\boldsymbol{\beta}$ denoting the vector of regression coefficients and $\boldsymbol{X}$ is the vector of covariates.

The new model

Yin and Ibrahim's model,

$$S_P(t) = \left(1 + \gamma\theta F(t)\right)^{-1/\gamma}, \gamma \in [-1, 0],$$

has a nice property: **the mixture model is one of its special cases** (for $\gamma = -1$ and $\theta = \exp(\boldsymbol{\beta}'\boldsymbol{X}) \,/(1 - \gamma \exp(\boldsymbol{\beta}'\boldsymbol{X}))$).

Note that the most well studied cure rate models are also special cases of

$$S_P(t) = (1 + \gamma\theta F(t)^{\lambda})^{-1/\gamma}, \gamma \geq 0$$

but not the mixture model.

Using the above θ, could we also include the interval $[-1, 0]$ to the parameter space of γ, and getting the binary model as a special case?

The new model

Yin and Ibrahim's model,

$$S_P(t) = \left(1 + \gamma\theta F(t)\right)^{-1/\gamma}, \gamma \in [-1, 0],$$

has a nice property: **the mixture model is one of its special cases** (for $\gamma = -1$ and $\theta = \exp(\boldsymbol{\beta}'\boldsymbol{X}) / (1 - \gamma \exp(\boldsymbol{\beta}'\boldsymbol{X}))$).

Note that the most well studied cure rate models are also special cases of

$$S_P(t) = (1 + \gamma\theta F(t)^{\lambda})^{-1/\gamma}, \gamma \geq 0$$

but not the mixture model.

Using the above θ, could we also include the interval $[-1, 0]$ to the parameter space of γ, and getting the binary model as a special case?

No, because θ (and $S_p(t)$) could be negative when γ is positive.

The new model

Yin and Ibrahim's model,

$$S_P(t) = \left(1 + \gamma\theta F(t)\right)^{-1/\gamma}, \gamma \in [-1, 0],$$

has a nice property: **the mixture model is one of its special cases** (for $\gamma = -1$ and $\theta = \exp(\boldsymbol{\beta}'\boldsymbol{X}) / (1 - \gamma \exp(\boldsymbol{\beta}'\boldsymbol{X}))$).

Note that the most well studied cure rate models are also special cases of

$$S_P(t) = (1 + \gamma\theta F(t)^{\lambda})^{-1/\gamma}, \gamma \geq 0$$

but not the mixture model.

Using the above θ, could we also include the interval $[-1, 0]$ to the parameter space of γ, and getting the binary model as a special case?

No, because θ (and $S_p(t)$) could be negative when γ is positive.

In this work, in order to solve this issue, we propose a re-parametrization the last model.

The new model

Firstly, note that the model

$$S_P(t) = (1 + \gamma\theta F(t)^{\lambda})^{-1/\gamma}, \gamma \geq 0, \theta, \lambda > 0$$

can be equivalently written as

$$S_P(t) = (1 + \gamma F(t)^{\lambda})^{-\theta/\gamma}, \gamma \geq 0, \theta, \lambda > 0.$$

The new model

Firstly, note that the model

$$S_P(t) = (1 + \gamma\theta F(t)^{\lambda})^{-1/\gamma}, \gamma \geq 0, \theta, \lambda > 0$$

can be equivalently written as

$$S_P(t) = (1 + \gamma F(t)^{\lambda})^{-\theta/\gamma}, \gamma \geq 0, \theta, \lambda > 0.$$

Problem: find a function $g(\gamma)$ such that the parameter space of γ could be extended to the whole real line

$$S_P(t) = (1 + g(\gamma) F(t)^{\lambda})^{-\theta/\gamma}, \gamma \in \Re, \lambda, \theta > 0$$

The new model

Specifically, we study re-parametrization of the form $g(\gamma) = \gamma c^{\gamma}$, i.e.

$$S_P(t) = (1 + \gamma c^{\gamma} F(t)^{\lambda})^{-\theta/\gamma}, \gamma \in \Re,$$

with $\theta > 0$ and $\lambda > 0$.

The constant c must be chosen such that the function $g(\gamma) = \gamma c^{\gamma}$ has the following two properties:

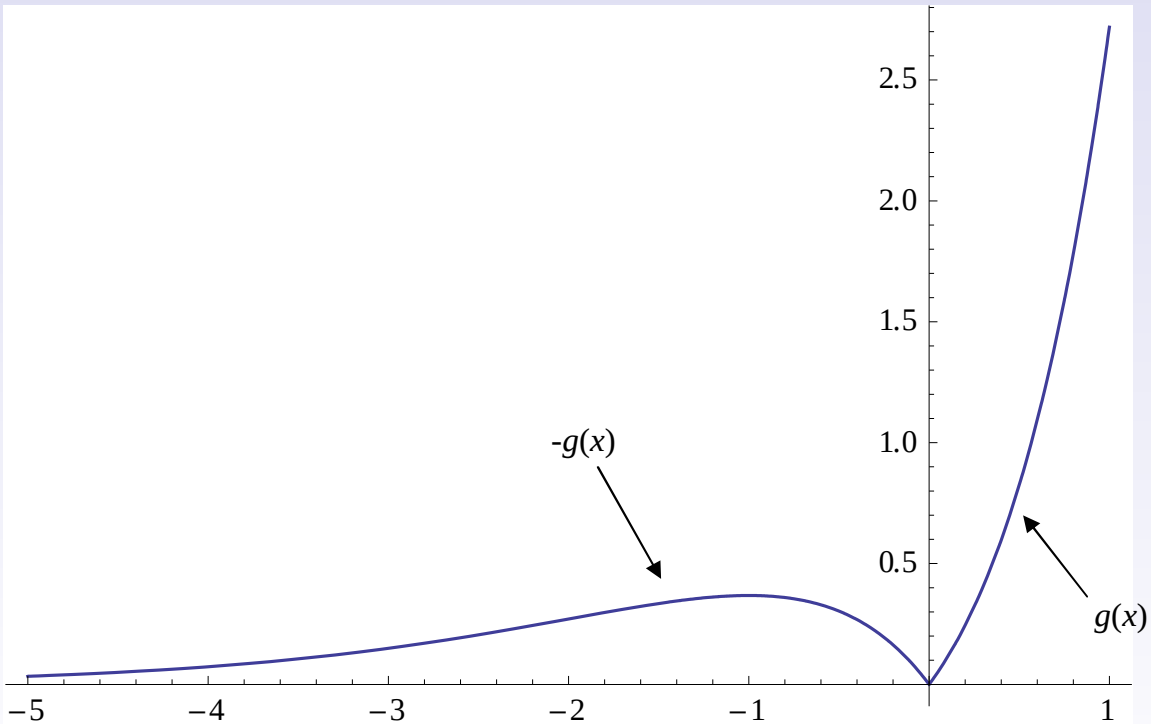
(a) $g(\gamma)$ is positive and surjective, for $\gamma \geq 0$;

(b) $g(\gamma)$ is surjective and $g(\gamma) \in [-1, 0]$, for $\gamma < 0$.

The new model

Why this form of re-parametrization?

The motivation comes from the function $g(x) = xe^x$, which has all the required properties but $-g(x)$ is not surjective on $[0,1]$, when $\gamma < 0$.



The new model

Suppose that $\gamma \geq 0$: then, the condition (a) is satisfied for every $c > 1$; this is true since $g(\gamma)$ is continuous, with $g(0) = 0$, $g'(\gamma) > 0$, and $\lim_{\gamma \to \infty} g(\gamma) = \infty$.

The new model

Suppose that $\gamma \geq 0$: then, the condition (a) is satisfied for every $c > 1$; this is true since $g(\gamma)$ is continuous, with $g(0) = 0$, $g'(\gamma) > 0$, and $\lim_{\gamma \to \infty} g(\gamma) = \infty$.

(a) $g(\gamma)$ is positive and surjective, for $\gamma \geq 0$;

The new model

Suppose that $\gamma \geq 0$: then, the condition (a) is satisfied for every $c > 1$; this is true since $g(\gamma)$ is continuous, with $g(0) = 0$, $g'(\gamma) > 0$, and $\lim_{\gamma \to \infty} g(\gamma) = \infty$.

Suppose that $\gamma < 0$: it can be seen that the condition (b) is satisfied when c is such that $g(-\ln(c)^{-1}) = 1$; the solution of this equation is $c = e^{e^{-1}}$.

The new model

Suppose that $\gamma \geq 0$: then, the condition (a) is satisfied for every $c > 1$; this is true since $g(\gamma)$ is continuous, with $g(0) = 0$, $g'(\gamma) > 0$, and $\lim_{\gamma \to \infty} g(\gamma) = \infty$.

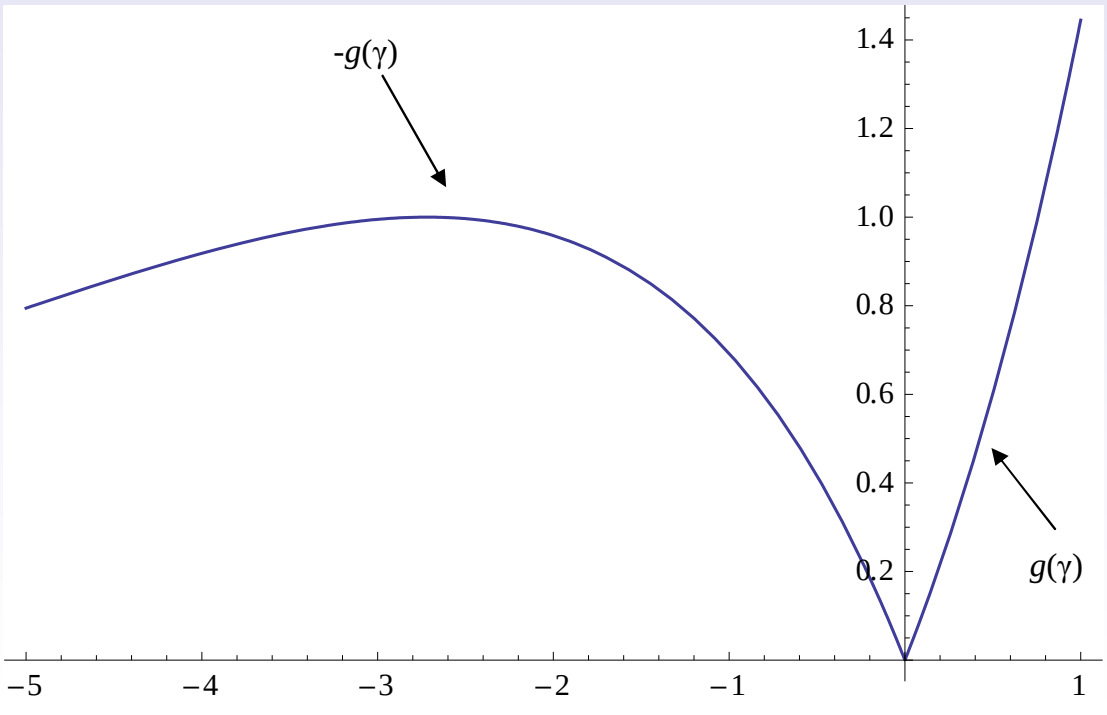
Suppose that $\gamma < 0$: it can be seen that the condition (b) is satisfied when c is such that $g(-\ln(c)^{-1}) = 1$; the solution of this equation is $c = e^{e^{-1}}$.

(b) $g(\gamma)$ is surjective and $g(\gamma) \in [-1, 0]$, for $\gamma < 0$;

The new model

Suppose that $\gamma \geq 0$: then, the condition (a) is satisfied for every $c > 1$; this is true since $g(\gamma)$ is continuous, with $g(0) = 0$, $g'(\gamma) > 0$, and $\lim_{\gamma \to \infty} g(\gamma) = \infty$.

Suppose that $\gamma < 0$: it can be seen that the condition (b) is satisfied when c is such that $g(-\ln(c)^{-1}) = 1$; the solution of this equation is $c = e^{e^{-1}}$.



Special cases

Therefore, the model we are going to study is given by

$$S_P(t) = (1 + \gamma e^{\gamma e^{-1}} F(t)^{\lambda})^{-\theta/\gamma}, \gamma \in \Re,$$

with $\theta > 0$ and $\lambda > 0$.

Some well known special cases of the proposed model are the:

Special cases

Therefore, the model we are going to study is given by

$$S_P(t) = (1 + \gamma e^{\gamma e^{-1}} F(t)^{\lambda})^{-\theta/\gamma}, \gamma \in \Re,$$

with $\theta > 0$ and $\lambda > 0$.

Some well known special cases of the proposed model are the:

- binary cure rate model: $\gamma = -\theta$ and $\lambda = 1$;

Special cases

Therefore, the model we are going to study is given by

$$S_P(t) = (1 + \gamma e^{\gamma e^{-1}} F(t)^{\lambda})^{-\theta/\gamma}, \gamma \in \Re,$$

with $\theta > 0$ and $\lambda > 0$.

Some well known special cases of the proposed model are the:

- binary cure rate model: $\gamma = -\theta$ and $\lambda = 1$;
- BCH model: $\gamma \to 0$ and $\lambda = 1$;

Special cases

Therefore, the model we are going to study is given by

$$S_P(t) = (1 + \gamma e^{\gamma e^{-1}} F(t)^{\lambda})^{-\theta/\gamma}, \gamma \in \Re,$$

with $\theta > 0$ and $\lambda > 0$.

Some well known special cases of the proposed model are the:

- binary cure rate model: $\gamma = -\theta$ and $\lambda = 1$;
- BCH model: $\gamma \to 0$ and $\lambda = 1$;
- negative binomial cure rate model: $\gamma > 0$ and $\lambda = 1$;

Special cases

Therefore, the model we are going to study is given by

$$S_P(t) = (1 + \gamma e^{\gamma e^{-1}} F(t)^{\lambda})^{-\theta/\gamma}, \gamma \in \Re,$$

with $\theta > 0$ and $\lambda > 0$.

Some well known special cases of the proposed model are the:

- binary cure rate model: $\gamma = -\theta$ and $\lambda = 1$;
- BCH model: $\gamma \to 0$ and $\lambda = 1$;
- negative binomial cure rate model: $\gamma > 0$ and $\lambda = 1$;
- geometric cure rate model: $\gamma = \lambda = 1$;

Special cases

Therefore, the model we are going to study is given by

$$S_P(t) = (1 + \gamma e^{\gamma e^{-1} F(t)^{\lambda}})^{-\theta/\gamma}, \gamma \in \Re,$$

with $\theta > 0$ and $\lambda > 0$.

Some well known special cases of the proposed model are the:

- binary cure rate model: $\gamma = -\theta$ and $\lambda = 1$;
- BCH model: $\gamma \to 0$ and $\lambda = 1$;
- negative binomial cure rate model: $\gamma > 0$ and $\lambda = 1$;
- geometric cure rate model: $\gamma = \lambda = 1$;
- some well studied destructive cure rate models, such as the geometric and negative binomial destructive cure rate model;

Special cases

Therefore, the model we are going to study is given by

$$S_P(t) = (1 + \gamma e^{\gamma e^{-1} F(t)^{\lambda}})^{-\theta/\gamma}, \gamma \in \Re,$$

with $\theta > 0$ and $\lambda > 0$.

Some well known special cases of the proposed model are the:

- binary cure rate model: $\gamma = -\theta$ and $\lambda = 1$;
- BCH model: $\gamma \to 0$ and $\lambda = 1$;
- negative binomial cure rate model: $\gamma > 0$ and $\lambda = 1$;
- geometric cure rate model: $\gamma = \lambda = 1$;
- some well studied destructive cure rate models, such as the geometric and negative binomial destructive cure rate model;
- the discrete stable, the Mittag-Leffler cure rate models and others.

Identifiability

The model introduced by Zeng, Yin and Ibrahim (2006, JASA) had some identifiability issues (this was also the case for the model studied by Koutras and Milienos, 2017, SIM).

Assuming that the parameters are independent of any set of covariates and $\gamma_0 \neq \gamma_1$, then we can always find $\theta_0 \neq \theta_1$ such that $(1+\theta_0 g(\gamma_0))^{-1/\gamma_0} = (1+\theta_1 g(\gamma_1))^{-1/\gamma_1}$.

Identifiability

The model introduced by Zeng, Yin and Ibrahim (2006, JASA) had some identifiability issues (this was also the case for the model studied by Koutras and Milienos, 2017, SIM).

Assuming that the parameters are independent of any set of covariates and $\gamma_0 \neq \gamma_1$, then we can always find $\theta_0 \neq \theta_1$ such that $(1+\theta_0 g(\gamma_0))^{-1/\gamma_0} = (1+\theta_1 g(\gamma_1))^{-1/\gamma_1}$.

However, assuming the existence of a continuous covariate with nonzero effect on θ (i.e. $\theta = \theta(\boldsymbol{X}) = \exp(\boldsymbol{\beta}'\boldsymbol{X})$), it can be proved that the parameters $\boldsymbol{\beta}$, γ and λ are identifiable.

The proof is carried out by following similar steps with Zeng, Yin and Ibrahim (2006, JASA).

Estimation

In this work, we consider the scenario wherein the time-to-event is subject to non-informative random right censoring.

Estimation

In this work, we consider the scenario wherein the time-to-event is subject to non-informative random right censoring.

We adopted a profile likelihood approach for γ since this method turned out to be quite effective even for small sample sizes; specifically, we fix a set of distinct (admissible) values of γ, and for each case we estimate (by a direct maximization of the log-likelihood function) the rest of model parameters.

Finally, our estimates are those which return the maximum value of the likelihood function.

Data and Estimation

Denoting with C_i and T_i the censoring time and lifetime of the ith individual, respectively, we then observe

$$Y_i = \min\{T_i, C_i\}$$

and $\delta_i = I(T_i \leq C_i)$, i.e.

$$\delta_i = \begin{cases} 1, & \text{if } Y_i \text{ is a time-to-event} \\ 0, & \text{if } Y_i \text{ is a censoring time,} \end{cases}, i = 1, 2, \ldots, n.$$

Data and Estimation: likelihood function

From n pairs of times and censoring indicators $(y_1, \delta_1), \ldots, (y_n, \delta_n)$, the likelihood function can be written as

$$L = L(\boldsymbol{\varphi}; \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\delta}) \propto \prod_{i=1}^{n} f_P(y_i, \boldsymbol{x}_i; \boldsymbol{\varphi})^{\delta_i} S_P(y_i, \boldsymbol{x}_i; \boldsymbol{\varphi})^{1-\delta_i},$$

where $\boldsymbol{x}_i$ is the vector of covariates for the ith individual, $\boldsymbol{x} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$, $\boldsymbol{y} = (y_1, \ldots, y_n)$, $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_n)$ and $\boldsymbol{\varphi}$ is the set of model parameters.

Data and Estimation: likelihood function

Thus, the likelihood becomes

$$L(\boldsymbol{\varphi}; \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\delta}) \propto$$

$$\prod_{i=1}^{n} (1 - p_0(\boldsymbol{x}_i; \boldsymbol{\varphi}))^{\delta_i} f_U(y_i, \boldsymbol{x}_i; \boldsymbol{\varphi})^{\delta_i} [p_0(\boldsymbol{x}_i; \boldsymbol{\varphi}) + (1 - p_0(\boldsymbol{x}_i; \boldsymbol{\varphi})) S_U(y_i, \boldsymbol{x}_i; \boldsymbol{\varphi})]^{1-\delta_i},$$

where $p_0(\boldsymbol{x}_i; \boldsymbol{\varphi})$ is the probability someone to be cured, i.e.

$$p_0(\boldsymbol{x}_i; \boldsymbol{\varphi}) = (1 + \gamma e^{\gamma e^{-1}})^{-\exp(\boldsymbol{\beta}' \boldsymbol{x}_i)/\gamma}$$

and $\theta = \exp(\boldsymbol{\beta}' \boldsymbol{x}_i)$, with $\boldsymbol{\beta}$ denoting the vector of regression coefficients

Data and Estimation: likelihood function

Thus, the likelihood becomes

$$L(\boldsymbol{\varphi}; \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\delta}) \propto$$

$$\prod_{i=1}^{n} (1 - p_0(\boldsymbol{x}_i; \boldsymbol{\varphi}))^{\delta_i} f_U(y_i, \boldsymbol{x}_i; \boldsymbol{\varphi})^{\delta_i} [p_0(\boldsymbol{x}_i; \boldsymbol{\varphi}) + (1 - p_0(\boldsymbol{x}_i; \boldsymbol{\varphi})) S_U(y_i, \boldsymbol{x}_i; \boldsymbol{\varphi})]^{1-\delta_i},$$

where $p_0(\boldsymbol{x}_i; \boldsymbol{\varphi})$ is the probability someone to be cured, i.e.

$$p_0(\boldsymbol{x}_i; \boldsymbol{\varphi}) = (1 + \gamma e^{\gamma e^{-1}})^{-\exp(\boldsymbol{\beta}' \boldsymbol{x}_i)/\gamma}$$

and $\theta = \exp(\boldsymbol{\beta}' \boldsymbol{x}_i)$, with $\boldsymbol{\beta}$ denoting the vector of regression coefficients

$$S_U(y_i, \boldsymbol{x}_i; \boldsymbol{\varphi}) = \frac{S_P(y_i, \boldsymbol{x}_i; \boldsymbol{\varphi}) - p_0(\boldsymbol{x}_i; \boldsymbol{\varphi})}{1 - p_0(\boldsymbol{x}_i; \boldsymbol{\varphi})}, f_U(y_i, \boldsymbol{x}_i; \boldsymbol{\varphi}) = \frac{f_P(y_i, \boldsymbol{x}_i; \boldsymbol{\varphi})}{1 - p_0(\boldsymbol{x}_i; \boldsymbol{\varphi})}$$

are the survival and probability density functions of the susceptibles, respectively.

Simulation

We assume that:

- W_i follows an exponential distribution with $\mu = 1$ (for every i);
- we have two covariates: X_1 being a symmetric Bernoulli r.v. and X_2 being a continuous uniformly distributed r.v. on $[0, 1]$;
- we have two sets of data of size 400 and 600;

The number of replications used was $r = 200$.

Simulation

We present results for some well known special cases of our model, i.e. the BCH (Poisson), the binary, the geometric, the negative binomial and the Mittag- Leffler cure rate model.

The (unobserved) cured proportion of our data-set ranges from 10% to 22%, while the (unobserved) censored proportion among the non-cured items ranges from 2% to 12%

BCH model

		β_0	β_1	β_2	λ	μ	γ
	true	1	-0.3	-0.3	1	1	0
$n = 400$	est	1.001	-0.295	-0.324	1.007	1.019	-0.024
	s.e.	0.133	0.120	0.216	0.076	0.151	0.273
	RMSE	0.018	0.014	0.047	0.006	0.023	0.075
	cp	0.950	0.940	0.930	0.895	0.905	-
$n = 600$	est	1.002	-0.304	-0.311	1.004	1.016	-0.031
	s.e.	0.115	0.095	0.164	0.061	0.131	0.268
	RMSE	0.013	0.009	0.027	0.004	0.017	0.072
	cp	0.960	0.935	0.955	0.895	0.870	-

- Grid search area: from -1 to 1 (step=0.2), 200 replications
- Overall (unobserved) sample cured proportion equals: 14%;
- Overall (unobserved) sample censored proportion among non-cured items equals: 10%.

BCH model

		β_0	β_1	β_2	λ	μ	γ
	true	1	-0.3	-0.3	1	1	0
$n = 400$	est	1.001	-0.295	-0.324	1.007	1.019	-0.024
	s.e.	0.133	0.120	0.216	0.076	0.151	0.273
	RMSE	0.018	0.014	0.047	0.006	0.023	0.075
	cp	0.950	0.940	0.930	0.895	0.905	-
$n = 600$	est	1.002	-0.304	-0.311	1.004	1.016	-0.031
	s.e.	0.115	0.095	0.164	0.061	0.131	0.268
	RMSE	0.013	0.009	0.027	0.004	0.017	0.072
	cp	0.960	0.935	0.955	0.895	0.870	-

- Grid search area: from -1 to 1 (step=0.2), 200 replications
- Overall (unobserved) sample cured proportion equals: 14%;
- Overall (unobserved) sample censored proportion among non-cured items equals: 10%.

BCH model

		β_0	β_1	β_2	λ	μ	γ
	true	1	-0.3	-0.3	1	1	0
$n = 400$	est	1.001	-0.295	-0.324	1.007	1.019	-0.024
	s.e.	0.133	0.120	0.216	0.076	0.151	0.273
	RMSE	0.018	0.014	0.047	0.006	0.023	0.075
	cp	0.950	0.940	0.930	0.895	0.905	-
$n = 600$	est	1.002	-0.304	-0.311	1.004	1.016	-0.031
	s.e.	0.115	0.095	0.164	0.061	0.131	0.268
	RMSE	0.013	0.009	0.027	0.004	0.017	0.072
	cp	0.960	0.935	0.955	0.895	0.870	-

- Grid search area: from -1 to 1 (step=0.2), 200 replications
- Overall (unobserved) sample cured proportion equals: 14%;
- Overall (unobserved) sample censored proportion among non-cured items equals: 10%.

Binary cure rate model

		β_0	β_1	β_2	λ	μ	γ
	true	0.1	0.1	0.1	1	1	-1
$n = 400$	est	0.166	0.085	0.096	1.019	1.015	-1.122
	s.e.	0.196	0.122	0.178	0.094	0.142	0.299
	RMSE	0.043	0.015	0.032	0.009	0.020	0.104
	cp	0.875	0.930	0.945	0.900	0.885	-
$n = 600$	est	0.147	0.092	0.105	1.018	0.994	-1.135
	s.e.	0.165	0.097	0.152	0.072	0.121	0.278
	RMSE	0.029	0.009	0.023	0.006	0.015	0.095
	cp	0.905	0.925	0.960	0.915	0.900	-

- Grid search area: from -2 to 0 (step=0.2), 200 replications;
- Overall (unobserved) sample cured proportion equals: 22%;
- Overall (unobserved) sample censored proportion among non-cured items equals: 5%.

Binary cure rate model

		β_0	β_1	β_2	λ	μ	γ
	true	0.1	0.1	0.1	1	1	-1
$n = 400$	est	0.166	0.085	0.096	1.019	1.015	-1.122
	s.e.	0.196	0.122	0.178	0.094	0.142	0.299
	RMSE	0.043	0.015	0.032	0.009	0.020	0.104
	cp	0.875	0.930	0.945	0.900	0.885	-
$n = 600$	est	0.147	0.092	0.105	1.018	0.994	-1.135
	s.e.	0.165	0.097	0.152	0.072	0.121	0.278
	RMSE	0.029	0.009	0.023	0.006	0.015	0.095
	cp	0.905	0.925	0.960	0.915	0.900	-

- Grid search area: from -2 to 0 (step=0.2), 200 replications;
- Overall (unobserved) sample cured proportion equals: 22%;
- Overall (unobserved) sample censored proportion among non-cured items equals: 5%.

Binary cure rate model

		β_0	β_1	β_2	λ	μ	γ
	true	0.1	0.1	0.1	1	1	-1
$n = 400$	est	0.166	0.085	0.096	1.019	1.015	-1.122
	s.e.	0.196	0.122	0.178	0.094	0.142	0.299
	RMSE	0.043	0.015	0.032	0.009	0.020	0.104
	cp	0.875	0.930	0.945	0.900	0.885	-
$n = 600$	est	0.147	0.092	0.105	1.018	0.994	-1.135
	s.e.	0.165	0.097	0.152	0.072	0.121	0.278
	RMSE	0.029	0.009	0.023	0.006	0.015	0.095
	cp	0.905	0.925	0.960	0.915	0.900	-

- Grid search area: from -2 to 0 (step=0.2), 200 replications;
- Overall (unobserved) sample cured proportion equals: 22%;
- Overall (unobserved) sample censored proportion among non-cured items equals: 5%.

Geometric cure rate model

		β_0	β_1	β_2	λ	μ	γ
	true	0.5	0.5	0.5	1.	1.	1.
$n = 400$	est	0.492	0.506	0.497	1.004	1.042	0.935
	s.e.	0.138	0.109	0.188	0.057	0.154	0.426
	RMSE	0.019	0.012	0.035	0.003	0.025	0.185
	cp	0.935	0.945	0.940	0.950	0.930	-
$n = 600$	est	0.493	0.508	0.494	0.999	1.028	0.932
	s.e.	0.118	0.095	0.154	0.052	0.132	0.424
	RMSE	0.014	0.009	0.024	0.003	0.018	0.184
	cp	0.925	0.935	0.945	0.900	0.870	-

- Grid search area: from 0 to 2 (step=0.2), 200 replications;
- Overall (unobserved) sample cured proportion equals: 10%;
- Overall (unobserved) sample censored proportion among non-cured items equals: 2%.

Negative binomial cure rate model

		β_0	β_1	β_2	λ	μ	γ
	true	0.5	0.5	0.5	1.	1.	2.
$n = 400$	est	0.497	0.485	0.516	1.010	1.037	1.990
	s.e.	0.133	0.114	0.201	0.063	0.159	0.454
	RMSE	0.018	0.013	0.040	0.004	0.027	0.205
	cp	0.950	0.945	0.945	0.890	0.935	-
$n = 600$	est	0.488	0.488	0.530	1.007	1.026	2.000
	s.e.	0.106	0.094	0.154	0.057	0.137	0.446
	RMSE	0.011	0.009	0.025	0.003	0.019	0.198
	cp	0.950	0.930	0.970	0.835	0.905	-

- Grid search area: from 1 to 3 (step=0.2), 200 replications;
- Overall (unobserved) sample cured proportion equals: 12%;
- Overall (unobserved) sample censored proportion among non-cured items equals: 2%.

Mittag-Leffler cure rate model

		β_0	β_1	β_2	λ	μ	γ
	true	0.1	0.4	0.6	3.	1.	1.
$n = 400$	est	0.081	0.401	0.639	3.032	1.004	0.995
	s.e.	0.138	0.122	0.198	0.251	0.097	0.459
	RMSE	0.019	0.015	0.041	0.064	0.009	0.210
	cp	0.945	0.950	0.950	0.945	0.910	-
$n = 600$	est	0.082	0.396	0.628	3.026	1.013	0.958
	s.e.	0.121	0.102	0.167	0.191	0.087	0.453
	RMSE	0.015	0.010	0.029	0.037	0.008	0.206
	cp	0.940	0.925	0.950	0.950	0.895	-

- Grid search area: from 0 to 2 (step=0.2), 200 replications;
- Overall (unobserved) sample cured proportion equals: 20%;
- Overall (unobserved) sample censored proportion among non-cured items equals: 6%.

Recidivism for Offenders Released from Prison

The proposed model is illustrated by a data-set on Recidivism for Offenders Released from Prison. The data-set is provided by Iowa Department of Corrections, available for public use (https://data.iowa.gov/).

Recidivism for Offenders Released from Prison

The proposed model is illustrated by a data-set on Recidivism for Offenders Released from Prison. The data-set is provided by Iowa Department of Corrections, available for public use (https://data.iowa.gov/).

Every person was followed for three years (as studies have shown if an offender relapses into criminal behavior it is most likely to happen within three years of being released).

Recidivism for Offenders Released from Prison

The proposed model is illustrated by a data-set on Recidivism for Offenders Released from Prison. The data-set is provided by Iowa Department of Corrections, available for public use (https://data.iowa.gov/).

Every person was followed for three years (as studies have shown if an offender relapses into criminal behavior it is most likely to happen within three years of being released).

A small part of this data-set of size $n = 3000$ was analyzed.

Recidivism for Offenders Released from Prison

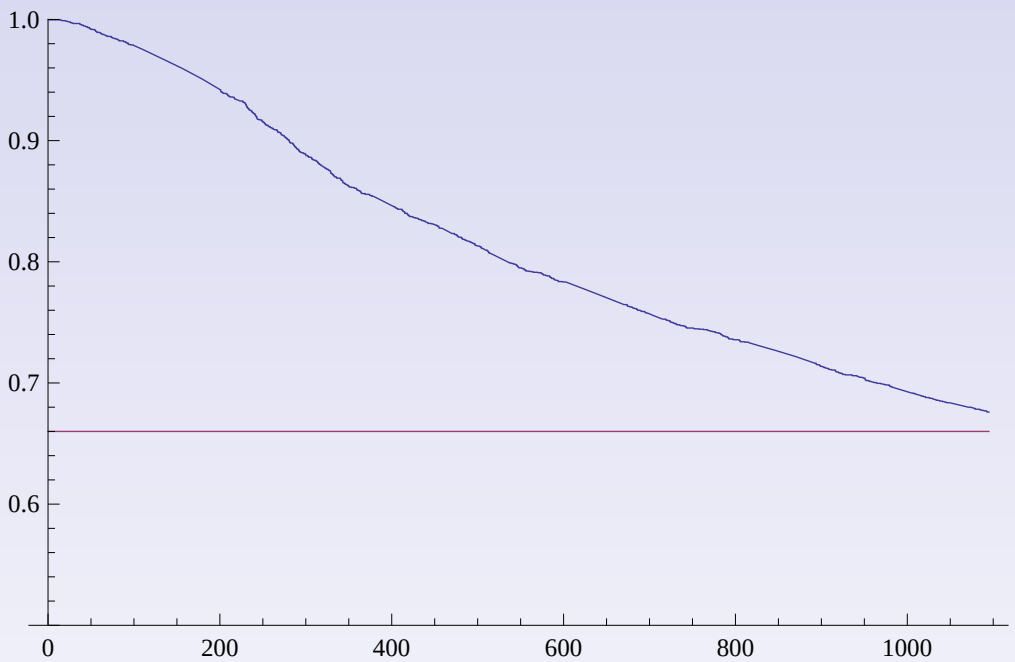


Figure: KM estimator of the survival function (data-set on recidivism)

Recidivism for Offenders Released from Prison

The covariates included in our analysis were:

- gender: male (87%), female (13%);
- age (with 5 categories): <25(18%), [25,34] (36%),[35,44] (24%),[45,54] (17%),≥ 55 (5%)

A Weibull distribution was assumed for $W_i's$.

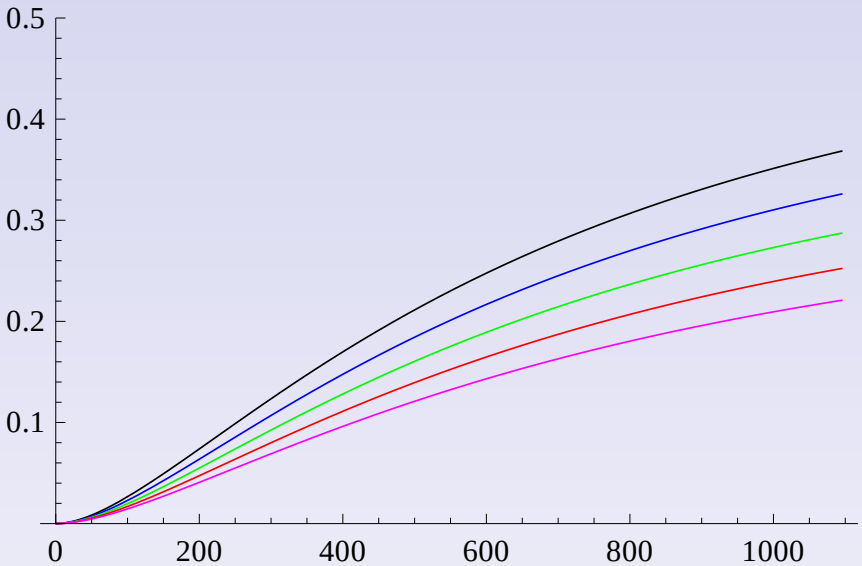
Recidivism for Offenders Released from Prison

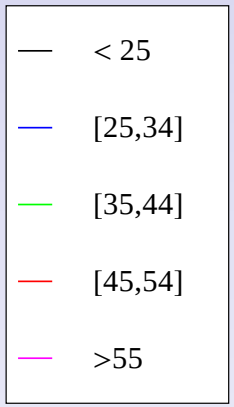
parameter	estimate	s.e.
Intercept	4.069	0.009
Gender	-0.248	0.011
Age	-0.153	0.001
λ	3.181	0.010
γ	-11.5	-
α_0 (scale)	0.003	0.0001
α_1 (shape)	0.636	0.0004

The grid search: γ on [-12, -10] (with step 0.025); several initial values for the maximization problem were taken into account and the results were quite robust.

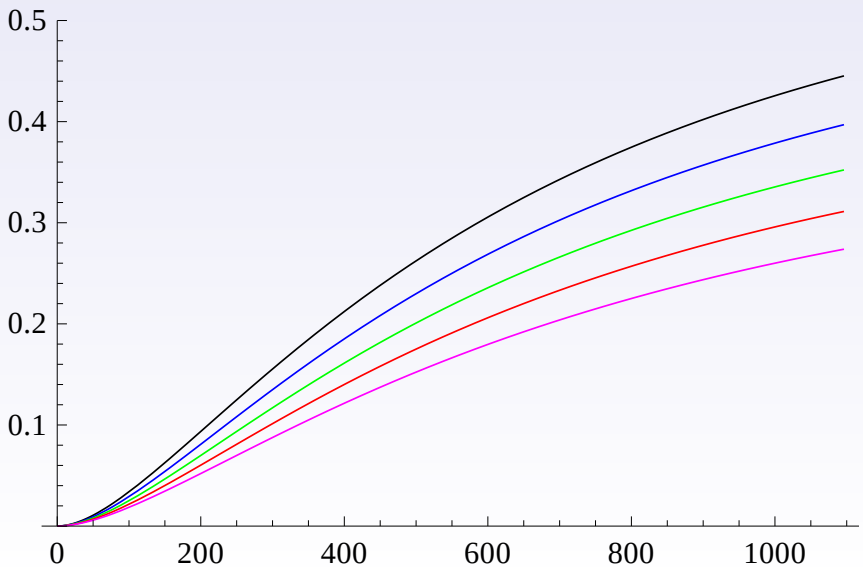
Recidivism for Offenders Released from Prison

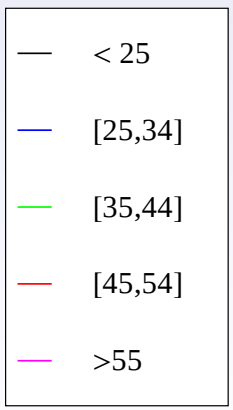
CDF for Females





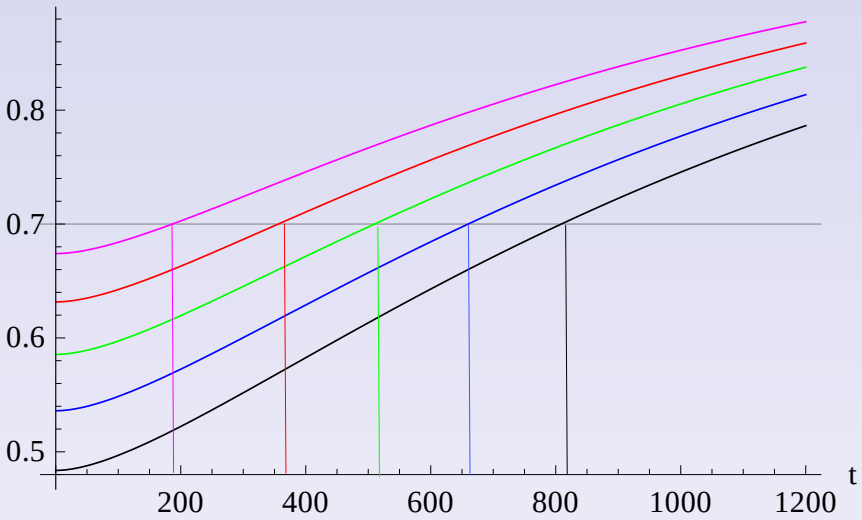
CDF for Males





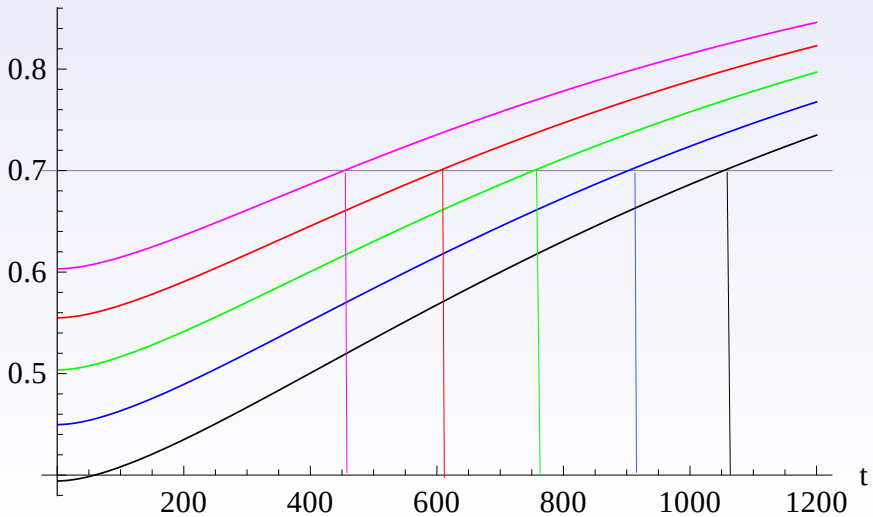
Recidivism for Offenders Released from Prison

P(cured|T>t) for Females



P(cured|T>t) for Males

The time beyond which an individual may be considered as cured (not relapsed into criminal behavior) with specific probability.



Conclusions

- A re-parametrization of a recently studied family of cure rate models was introduced.
- The new model has also as a special case the binary cure rate model, among many other well known models (the Poisson, the geometric, the negative binomial model, and the models studied by Zeng, Yin and Ibrahim, 2006, Yin and Ibrahim, 2005 and a class of models studied by Tsodikov, 2002).
- It can also handle the existence of a destructive mechanism on the initial number of clonogens.
- The suggested inferential method (profile likelihood), exhibits a high accuracy.

Conclusions

- Although not presented here, the first results for the model discrimination (based on the likelihood ratio test) are promising; EM-algorithm works quite well also.
- Interval censored data, non-parametric estimation and asymptotic properties of the estimators, are among the future directions of this study.

References

1. Balakrishnan, N. and Milienos, F.S. (2020). On a class of non-linear transformation cure rate models. *Biometrical Journal* , 62(5), 1208-1222.
2. Balakrishnan, N. and Pal, S. (2013). Expectation maximization-based likelihood inference for flexible cure rate models with Weibull distribution. *Statistical Methods in Medical Research*, available online.
3. Balakrishnan, N., Koutras, M.V., Milienos, F.S. and Pal, S. (2016). Piecewise linear approximations for cure rate models and associated inferential issues. *Methodology and Computing in Applied Probability, to appear*.
4. Berkson, J. and Gage, R.P. (1952), Survival curves for cancer patients following treatment. *Journal of the American Statistical Association*, *47*, 501-515.
5. Boag, J. W. (1948a). The presentation and analysis of the results of radiotherapy. Part I. *The British Journal of Radiology, 21(243)*, 128-138.
6. Boag, J. W. (1948b). The presentation and analysis of the results of radiotherapy. Part II. *The British Journal of Radiology, 21(244)*, 189-203.
7. Boag, J.M. (1949), Maximum likelihood estimates of the proportion of patients cured by cancer therapy. *Journal of the Royal Statistical Society, Ser. B, 11*, 15-44.
8. Cancho, V. G., Bandyopadhyay, D., Louzada, F. and Yiqi, B. (2013). The destructive negative binomial cure rate model with a latent activation scheme. *Statistical Methodology, 13*, 48-68.
9. Cantor, A.B. and Shuster, J.J. (1992). Parametric versus non-parametric methods for estimating cure rates based on censored survival data. *Statistics in Medicine*, *11*, 931-937.
10. Castro, M.D., Cancho, V.G. and Rodrigues, J. (2009). A bayesian long-term survival model parametrized in the cured fraction. *Biometrical Journal*, *51*, 443-455.
11. Christoph, G. and Schreiber, K. (1998). The generalized discrete Linnik distributions. In: Kahle, W. et al. (Eds.), *Advances in Stochastic Models for Reliability, Quality and Safety*, pp 3-18, Birkhäuser, Basel.
12. Gu, Y., Sinha, D. and Banerjee, S. (2011). Analysis of cure rate survival data under proportional odds model. *Lifetime Data Analysis*, *17*, 123-134.
13. Koutras, M.V. and Milienos, F.S. (2017). A flexible family of transformation cure rate models. *Statistics in Medicine, 36*, 2559-2575.

References

14 Meeker, W. Q. and LuValle, M. J. (1995). An accelerated life test model based on reliability kinetics. *Technometrics*, 37(2):133-146.

15 Rodrigues, J., de Castro, M., Cancho, V. G. and Balakrishnan, N. (2009). COM-Poisson cure rate survival models and an application to a cutaneous melanoma data. *Journal of Statistical Planning and Inference, 139*, 3605-3611.

16 Rodrigues, J., de Castro, M., Balakrishnan, N. and Cancho, V.G. (2011). Destructive weighted Poisson cure rate models. *Lifetime Data Analysis, 17*, 333-346.

17 Schmidt, P. and Witte, A. D. (1988). P*redicting Recidivism Using Survival Models*. Springer Science and Business Media, New York.

18 Tsodikov, A. (1998a). A proportional hazards model taking account of long-term survivors. *Biometrics*, 1508-1516.

19 Tsodikov, A. (2002). Semi-parametric models of long-and short-term survival: an application to the analysis of breast cancer survival in Utah by age and stage. *Statistics in medicine, 21*, 895-920.

20 Tsodikov, A.D., Ibrahim, J.G. and Yakovlev, A.Y. (2003). Estimating cure rates from survival data. *Journal of the American Statistical Association, 98*, 1063-1078.

21 Yakovlev, A.Y., Cantor, A.B. and Shuster, J.J. (1994). Parametric versus non-parametric methods for estimating cure rates based on censored survival data. *Statistics in Medicine, 13*, 983-986.

22 Yang, G.L. and Chen, C. W. (1991). A stochastic two-stage carcinogenesis model: a new approach to computing the probability of observing tumor in animal bioassays. *Mathematical Biosciences, 104*, 247-258.

23 Yin, G. and Ibrahim, J.G. (2005). Cure rate models: a unified approach. *Canadian Journal of Statistics, 33*, 559-570.

24 Zeng, D., Yin, G. and Ibrahim, J.G. (2006). Semiparametric transformation models for survival data with a cure fraction. *Journal of the American Statistical Association, 101*, 670-684.