Package: PScr (via r-universe)

October 3, 2026

Type Package

Title Estimation for the Power Series Cure Rate Model

Version 1.2

Date 2026-09-21

Author Diego Gallardo [aut, cre], Reza Azimi [ctb], Daniel Jana [ctb], Yolanda Gomez [ctb]

Maintainer Diego Gallardo <diego.gallardo.mateluna@gmail.com>

Description Provides estimation and simulation tools for particular cases of the power series cure rate model <doi:10.1080/03610918.2011.639971>. For the distribution of the concurrent causes the alternative models are the Poisson, logarithmic, negative binomial and Bernoulli (which are includes in the original work), the polylogarithm model <doi:10.1080/00949655.2018.1451850> and the Flory-Schulz <doi:10.3390/math10244643>. The estimation procedure is based on the EM algorithm discussed in <doi:10.1080/03610918.2016.1202276>. For the distribution of the time-to-event the alternative models are slash half-normal, Weibull, gamma and Birnbaum-Saunders distributions.

Depends R (>= 4.0.0), stats

Imports survival, pracma, VGAM

Suggests mstate

License GPL (>= 2)

NeedsCompilation no

Repository https://diegogallardo87.r-universe.dev

Date/Publication 2026-10-02 23:50:02 UTC

RemoteUrl https://github.com/cran/PScr

RemoteRef HEAD

RemoteSha 37d384db46353b23d9d01fb7a563cc9d48376aa2

Contents

EM.PScr 2
rcura 3

Index **7**

`EM.PScr` *Maximum likelihood estimation based on EM algorithm for the Power Series cure rate model*

Description

This function provides the maximum likelihood estimation based on the EM algorithm for the Power Series cure rate model

Usage

```
EM.PScr(t, delta, z, model = 1, dist = 1, max.iter = 1000,
           prec = 1e-04)
```

Arguments

`t`	observed times
`delta`	failure indicators
`z`	matrix of covariates (with n rows and r columns)
`model`	distribution to be used for the concurrent causes: 1 for Poisson, 2 for logarithmic, 3 for negative binomial, 4 for bernoulli and 5 for polylogarithm (Gallardo et al. 2018). 6 for Flory-Schulz (Azimi et al. 2022).
`dist`	distribution to be used for the time-to-event: 1 for slash half-normal (Gallardo et al., 2022), 2 for Weibull, 3 for gamma and 4 for Birnbaum-Saunders.
`max.iter`	maximum number of iterations to be used by the algorithm
`prec`	precision (in absolute value) for the parameters to stop the algorithm.

Details

The EM algorithm for the model is implemented as in Gallardo et al. (2017).

Value

`estimate`	a matrix containing the estimated parameters and their standard error
`loglike`	the estimated log-likelihood function evaluated in the maximum likelihood estimators
`AIC`	the Akaike information criterion
`BIC`	the Bayesian (also known as Schwarz) information criterion

Author(s)

Diego I. Gallardo and Reza Azimi

References

Azimi, R, Esmailian, M, Gallardo DI and Gomez HJ. (2022). A New Cure Rate Model Based on Flory-Schulz Distribution: Application to the Cancer Data. Mathematics 10, 4643

Gallardo DI, Gomez YM and De Castro M. (2018). A flexible cure rate model based on the polylogarithm distribution. Journal of Statistical Computation and Simulation 88 (11), 2137-2149

Gallardo DI, Gomez YM, Gomez HJ, Gallardo-Nelson MJ, Bourguignon M. (2022) The slash half-normal distribution applied to a cure rate model with application to bone marrow transplantation. Mathematics, Submitted.

Gallardo DI, Romeo JS and Meyer R. (2017). A simplified estimation procedure based on the EM algorithm for the power series cure rate model. Communications in Statistics-Simulation and Computation 46 (8), 6342-6359.

Examples

```
require(mstate)
data(ebmt4)
attach(ebmt4)
t = srv / 365.25 # Time in years
delta=srv.s
prophy=as.factor(proph)
year2=ifelse(year=="1985-1989",0,1)
z=t(model.matrix(~proph-1))
#Computes the estimation for Poisson-Slash half-normal cure rate model
EM.PScr(t, delta, z, model=1, dist=1)
#Computes the estimation for Flory-Schulz-Slash half-normal cure rate model
EM.PScr(t, delta, z, model=6, dist=1)
```

---

rcura — *Simulation of Right-Censored Data from a Power Series Cure Rate Model*

Description

This function generates right-censored survival data from a power series cure rate model. The number of latent competing causes follows a power series distribution (Poisson, Bernoulli, negative binomial, geometric or polylogarithm) and the time-to-event of each cause follows a Weibull, gamma or Birnbaum-Saunders distribution. The generated data can be analysed with `EM.PScr`.

Usage

```
rcura(n, beta, z = NULL, q = NULL, alpha, sigma, C, model, dist)
```

Arguments

`n`	sample size (a positive integer).
`beta`	numeric vector of regression coefficients (of length r) related to the power parameter θ of the distribution of the latent causes. Its length must be equal to the number of rows of `z`.
`z`	matrix of covariates with r rows (one per covariate, the first one usually being the intercept) and `n` columns (one per individual). If `NULL` (default), an intercept row of ones is used and, if `length(beta) > 1`, the remaining covariates are generated as independent Bernoulli(0.5) variables.
`q`	positive numeric parameter of the distribution of the latent causes. It is required for `model = "binneg"` (size parameter of the negative binomial distribution) and for `model = "polilog"` (parameter of the polylogarithm distribution). It is ignored for the other models. Default is `NULL`.
`alpha`	positive shape parameter of the distribution of the time-to-event of each latent cause.
`sigma`	positive scale parameter of the distribution of the time-to-event of each latent cause.
`C`	positive censoring time, common to all individuals.
`model`	character string with the distribution of the latent causes: `"poisson"`, `"bernoulli"`, `"binneg"` (negative binomial), `"geometrica"` (geometric) or `"polilog"` (polylogarithm).
`dist`	character string with the distribution of the time-to-event of each latent cause: `"weibull"`, `"gamma"` or `"bs"` (Birnbaum-Saunders).

Details

Let M_i be the number of latent causes of individual i, $i = 1, \ldots, n$, with power parameter $\theta_i = g(z_i^\top \beta)$. Given $M_i > 0$, the time-to-event is $T_i = \min(W_{i1}, \ldots, W_{iM_i})$, where the W_{ij} are independent and identically distributed with the distribution selected in `dist`. Individuals with $M_i = 0$ are cured and have $T_i = \infty$. The observed time is $t_i = \min(T_i, C)$ and the failure indicator is $\delta_i = I(T_i \leq C)$.

The distributions available for the latent causes, the link functions and the corresponding cure fractions $p_0 = P(M = 0)$ are:

`model`	distribution of M	link g	cure fraction p_0
`"poisson"`	Poisson(θ)	exponential	$\exp(-\theta)$
`"bernoulli"`	Bernoulli($\theta/(1+\theta)$)	exponential	$1/(1+\theta)$
`"binneg"`	negative binomial(q, $1-\theta$)	logit	$(1-\theta)^q$
`"geometrica"`	geometric (`"binneg"` with $q = 1$)	logit	$1-\theta$
`"polilog"`	$X - 1$, with $P(X = m) = \theta^m/(m^q \mathrm{Li}_q(\theta))$, $m \geq 1$	logit	$\theta/\mathrm{Li}_q(\theta)$

The time-to-event distributions are parameterized by a shape `alpha` and a scale `sigma`: `"weibull"` uses `rweibull`, `"gamma"` uses `rgamma` and `"bs"` uses `rbisa`. This is the same parameterization used by `EM.PScr`, so the simulated data can be fitted with the matching `dist` code.

The cure fraction and the percentage of censoring of a simulated sample depend on `beta`, `alpha`, `sigma` and `C`. If the time-to-event distribution has a heavy tail relative to `C`, a fraction of the susceptible individuals is censored before the event occurs, so the Kaplan-Meier curve does not level off at p_0 and the cure fraction is difficult to identify. The empirical cure fraction and the proportion of censored observations of each sample are returned as attributes.

Value

A `data.frame` with `n` rows and two columns:

`t`	observed times, $\min(T, C)$.
`delta`	failure indicators: 1 if the event was observed and 0 if the observation was censored.

The data frame has the following attributes, which can be extracted with `attr`:

`z`	the matrix of covariates used to generate the data.
`theta`	numeric vector with the power parameter θ_i of each individual.
`M`	integer vector with the simulated number of latent causes of each individual.
`p0_empirico`	empirical cure fraction, that is, the proportion of individuals with `M = 0`.
`cens_pct`	proportion of censored observations.

Note

The function uses the random number generator, so results are reproducible only after calling `set.seed`. The computing time grows linearly with `n`.

Correspondence with the codes of `EM.PScr`: `model = "poisson"` corresponds to `model = 1`, `"binneg"` and `"geometrica"` to `model = 3` (with $q = 1$ in the geometric case), `"bernoulli"` to `model = 4` and `"polilog"` to `model = 5`; `dist = "weibull"`, `"gamma"` and `"bs"` correspond to `dist = 2`, `3` and `4`, respectively. The logarithmic and Flory-Schulz models and the slash half-normal and log-normal time-to-event distributions available in `EM.PScr` are not implemented in `rcura`.

Author(s)

Daniel Jana, Yolanda Gomez and Diego Gallardo

References

Birnbaum ZW and Saunders SC. (1969). A new family of life distributions. Journal of Applied Probability 6 (2), 319-327.

Cancho VG, Louzada F and Ortega EMM. (2013). The power series cure rate model: an application to a cutaneous melanoma data. Communications in Statistics - Simulation and Computation 42 (3), 586-602. doi:10.1080/03610918.2011.639971

Gallardo DI, Gomez YM and De Castro M. (2018). A flexible cure rate model based on the polylogarithm distribution. Journal of Statistical Computation and Simulation 88 (11), 2137-2149.

Gallardo DI, Romeo JS and Meyer R. (2017). A simplified estimation procedure based on the EM algorithm for the power series cure rate model. Communications in Statistics - Simulation and Computation 46 (8), 6342-6359. doi:10.1080/03610918.2016.1202276

Yakovlev AY and Tsodikov AD. (1996). Stochastic Models of Tumor Latency and Their Biostatistical Applications. World Scientific, Singapore.

See Also

EM.PScr

Examples

```
set.seed(2026)

## Poisson cure rate model with Weibull times and no covariates.
## With theta = -log(0.7) the cure fraction is p0 = exp(-theta) = 0.7
theta <- -log(0.7)
dat <- rcura(n = 500, beta = log(theta), alpha = 2, sigma = 3,
             C = 8, model = "poisson", dist = "weibull")
head(dat[dat$delta == 1, ])  # first observed events
attr(dat, "p0_empirico")     # empirical cure fraction (approximately 0.7)
attr(dat, "cens_pct")        # proportion of censored observations

## The Kaplan-Meier curve levels off near the cure fraction
fit <- survival::survfit(survival::Surv(t, delta) ~ 1, data = dat)
plot(fit, conf.int = FALSE, xlab = "Time", ylab = "Survival probability")
abline(h = 0.7, lty = 2)

## Negative binomial model (q = 2) with gamma times and cure fraction 0.4.
## Since p0 = (1 - theta)^q, we have theta = 1 - p0^(1 / q)
q <- 2
theta <- 1 - 0.4^(1 / q)
dat2 <- rcura(n = 200, beta = qlogis(theta), q = q, alpha = 2, sigma = 1.4,
              C = 8, model = "binneg", dist = "gamma")
mean(dat2$delta == 0)

## Intercept and a binary covariate (covariates in rows, individuals in columns)
n <- 200
z <- rbind(1, rbinom(n, 1, 0.5))
dat3 <- rcura(n = n, beta = c(-0.5, 0.7), z = z, alpha = 0.5, sigma = 2.5,
              C = 8, model = "poisson", dist = "bs")
table(dat3$delta)


## Fit the first simulated data set with EM.PScr (model = 1: Poisson, dist = 2: Weibull)
fit.em <- EM.PScr(t = dat$t, delta = dat$delta,
                  z = matrix(1, nrow = 1, ncol = nrow(dat)), model = 1, dist = 2)
fit.em$estimate

## Polylogarithm model
dat4 <- rcura(n = 200, beta = 0.5, q = 2, alpha = 2, sigma = 3,
              C = 8, model = "polilog", dist = "weibull")
attr(dat4, "p0_empirico")
```

Index

∗ **datagen**
    rcura, 3
∗ **survival**
    rcura, 3

attr, *5*

EM.PScr, 2, *3–6*

rbisa, *4*
rcura, 3
rgamma, *4*
rweibull, *4*

set.seed, *5*