

Package ‘cervmix’

September 5, 2022

Type Package

Title Competing Cause Mixture Models for Cervical Cancer Screening Data

Version 0.1.0

Author Kelsi Kroon

Maintainer Kelsi Kroon <k.kroon@amsterdamumc.nl>

Description This package fits Competing Cause Prevalence-Incidence mixture models to interval-censored cervical cancer screening data and obtains parameter estimates using the gradient EM algorithm.

License GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 7.1.1

Suggests knitr, rmarkdown, testthat (>= 3.0.0)

VignetteBuilder knitr

Config/testthat/edition 3

R topics documented:

cervmix.fit	1
cervmix.predict	3
cervmix.simulator	3
create.interval.data	4
Index	6

cervmix.fit

Competing Cause Prevalence-Incidence Mixture Models

Description

This function fits Competing Cause Prevalence-Incidence mixture models to interval-censored cervical cancer screening data and obtains parameter estimates. It is possible for the user to select the covariates that will be used for each parameter.

Usage

```
cervmix.fit(
  l1_x,
  l2_x,
  pi_x,
  data,
  num.runs = 30,
  short.epsilon = 0.1,
  epsilon = 1e-08,
  silent = T
)
```

Arguments

l1_x	A vector containing the names of covariates used in the λ_1 (progression rate) parameter (must match column name(s) in the input data)
l2_x	A vector containing the names of covariates used in the λ_2 (clearance rate) parameter (must match column name(s) in the input data)
pi_x	A vector containing the names of covariates used in the π parameter (probability of prevalent disease) (must match column name(s) in the input data)
data	Data used to fit the model containing columns for each term in l1_x, l2_x and pi_x. The first three columns must be (1) left interval, (2) right interval and (3) z indicator for prevalent/incident disease. The data must contain observed prevalent and incident cases, however some cases may be unknown ('NA').
num.runs	Number of runs of the 'Short' EM algorithm used to determine initial values for the EM algorithm. Defaults to 30.
short.epsilon	Convergence criteria used in the 'Short' EM algorithm to determine initial values. Defaults to 0.1.
epsilon	Convergence criteria for the change in log-likelihood value used for stopping the EM algorithm. Defaults to 1e-08.

Value

The output is a list containing the following elements:

- initial.values - initial values determined by the Short EM algorithm process
- theta.hat - optimum parameter values estimated by the EM algorithm
- num.iterations - number of iterations until the EM algorithm converged
- log.likelihood - value of the log.likelihood at the optimum parameter values
- hess - hessian matrix
- std.dev - standard deviation of parameter estimates
- summary - data frame with estimate, std.dev, and 95% CI for each parameter (useful for data set comparisons)

Author(s)

Kelsi Kroon <k.kroon@amsterdamumc.nl>

Examples

```
sim_dat <- simulator(2000, c("hvp"), c(), c(),
                    c(exp(-5), -3, 2, -2, 0.25), show_prob = 0.9, i=5)
cervmix.fit(c("hvp"), c(), c(), sim_dat, silent=F)
```

cervmix.predict

*Competing Cause Prevalence-Incidence Mixture Model Predictions***Description**

Competing Cause Prevalence-Incidence Mixture Model Predictions

Usage

```
cervmix.predict(l1_x, l2_x, pi_x, data, time.points, theta.hat)
```

Arguments

l1_x	A vector containing the names of covariates used in the λ_1 (progression rate) parameter (must match column name(s) in the input data)
l2_x	A vector containing the names of covariates used in the λ_2 (clearance rate) parameter (must match column name(s) in the input data)
pi_x	A vector containing the names of covariates used in the π parameter (probability of prevalent disease) (must match column name(s) in the input data)
data	Data set of covariates from which to make the predictions.
time.points	Numeric vector of time points used to make cumulative risk predictions
theta.hat	Parameter estimates for the model to be used (output from CCmixture.fit)

cervmix.simulator

*Simulator***Description**

Cervical cancer screening data simulator.

Usage

```
cervmix.simulator(n, l1_x, l2_x, pi_x, params, show_prob = 0.9, i = 5)
```

Arguments

n	Number of women in the simulated data set.
l1_x	A vector containing the names of covariates used in the λ_1 (progression rate) parameter. Options are "age", "HPV16" and "cytology".
l2_x	A vector containing the names of covariates used in the λ_2 (clearance rate) parameter. Options are "age", "HPV16" and "cytology".
pi_x	A vector containing the names of covariates used in the π parameter (probability of prevalent disease). Options are "age", "HPV16" and "cytology".
params	Numerical vector the parameter values to be used in the data simulation (first value is background risk, then l1, l2, pi)
show_prob	A value representing the probability of a woman showing up for the screening visit. Defaults to 0.9.
i	A value representing the interval between screening rounds (in years). Defaults to 5.

Value

A data frame containing the left and right interval of CIN2+ detection, the indicator of prevalent disease, age (younger or older than 39, 1=younger), HPV genotype (HPV16 or other, 1=HPV16), and cytology result (normal or abnormal, 1=abnormal).

Author(s)

Kelsi Kroon <k.kroon@amsterdamumc.nl>

create.interval.data *Create interval censored data*

Description

Create interval censored data

Usage

```
create.interval.data(patient_info, hist, cyt, hpv)
```

Arguments

patient_info	Data frame containing patient information such as Idwoman, Study, Danasc, Age, Darecl, Arm, DaCens, StCens
hist	Data frame containing results from histological tests for patients. Must have column names Study, Idwoman, VisitHisto, Idhisto, Datehisto, Typehisto, Is-torandom, Classhisto.
cyt	Data frame containing results from cytological tests for patients. Must have column names Study, Idwoman, VisitCyto, Idtest, HPVpres, cyto_idHPV, Date-cyto, Type, CytMed, Classcyto, Racctest, FUP
hpv	Data frame containing results from HPV tests for patients

Value

A data frame containing the left and right intervals of CIN2+ detection for women who were HPV+ at baseline.

Index

`cervmix.fit`, [1](#)
`cervmix.predict`, [3](#)
`cervmix.simulator`, [3](#)
`create.interval.data`, [4](#)