Package 'rERR'

June 6, 2018

Title Excess Relative Risk Models

Version 0.1

Author Francesc Badia Roca (ISGlobal) <francesc.badia@isglobal.org>,
David Moriña Soler (Catalan Institute of Oncology (ICO)-IDIBELL) <david.morina@uab.cat>

Description Fits a linear excess relative risk model by maximum likelihood, possibly including several variables and allowing for lagged exposures. Allow time dependent covariates.

Depends R (>= 3.5.0)

License GPL (>= 2)

Encoding UTF-8

LazyData true

RoxygenNote 6.0.1

Imports survival, stats4, plyr, dplyr, reshape2, numDeriv, ggplot2

Maintainer Francesc Badia Roca <francesc.badia@isglobal.org>

Date 2018-06-05

Suggests testthat, knitr, rmarkdown

VignetteBuilder knitr

NeedsCompilation yes

Repository CRAN

Date/Publication 2018-06-06 10:30:02 UTC

R topics documented:

`cohort_ef` *Simulated cohort of subjects exposed to ct scans*

Description

List of scans recived by a cohort of subjects

Usage

```
cohort_ef
```

Format

A data frame with 6315 rows and 8 variables:

id subject identificator, integer

sex sex of the subject, 1-male / 2-female

entry_age age of entry at the cohort, numeric

exit_age age at exit of the cohort, numeric

outcome disease or not, 0-no disease / 1-disease

age age at scan, numeric

dose dose of the scan, numeric in mGy

country country, character

cohort_wf *Simulated cohort of subjects exposed to ct scans*

Description

List of subjects that recived at least 1 scan

Usage

```
cohort_wf
```

Format

A data frame with 1000 rows and 80 variables:

id subject identificator, integer

sex sex of the subject, 1-male / 2-female

YearInit Year of entry of the subject, integer

AgeAtEntry age of entry at the cohort, numeric

age_at_event age at exit of the cohort, numeric

outcome disease or not, 0-no disease / 1-disease

end_status status at exit of the cohort, 0-disease / 1-death / 2-healthy

ses socio economic status,integer

number_of_ct number of ct's recived in all follow-up, integer

ctage1 middle age at i-th year of being in the cohort, numeric

ctage2 middle age at i-th year of being in the cohort, numeric

ctage3 middle age at i-th year of being in the cohort, numeric

ctage4 middle age at i-th year of being in the cohort, numeric

ctage5 middle age at i-th year of being in the cohort, numeric

ctage6 middle age at i-th year of being in the cohort, numeric

ctage7 middle age at i-th year of being in the cohort, numeric

ctage8 middle age at i-th year of being in the cohort, numeric

ctage9 middle age at i-th year of being in the cohort, numeric

ctage10 middle age at i-th year of being in the cohort, numeric

ctage11 middle age at i-th year of being in the cohort, numeric

ctage12 middle age at i-th year of being in the cohort, numeric

ctage13 middle age at i-th year of being in the cohort, numeric

ctage14 middle age at i-th year of being in the cohort, numeric

ctage15 middle age at i-th year of being in the cohort, numeric

ctage16 middle age at i-th year of being in the cohort, numeric

ctage17 middle age at i-th year of being in the cohort, numeric

ctage18 middle age at i-th year of being in the cohort, numeric

ctage19 middle age at i-th year of being in the cohort, numeric

ctage20 middle age at i-th year of being in the cohort, numeric

ctage21 middle age at i-th year of being in the cohort, numeric

ctage22 middle age at i-th year of being in the cohort, numeric

ctage23 middle age at i-th year of being in the cohort, numeric

ctage24 middle age at i-th year of being in the cohort, numeric

ctage25 middle age at i-th year of being in the cohort, numeric

ctage26 middle age at i-th year of being in the cohort, numeric

ctage27 middle age at i-th year of being in the cohort, numeric

ctage28 middle age at i-th year of being in the cohort, numeric

ctage29 middle age at i-th year of being in the cohort, numeric

ctage30 middle age at i-th year of being in the cohort, numeric

ctage31 middle age at i-th year of being in the cohort, numeric

ctage32 middle age at i-th year of being in the cohort, numeric

ctage33 middle age at i-th year of being in the cohort, numeric

ctage34 middle age at i-th year of being in the cohort, numeric

ctage35 middle age at i-th year of being in the cohort, numeric

dose1 grouped doses recived the i-th year of being in the cohort, numeric

dose2 grouped doses recived the i-th year of being in the cohort, numeric

dose3 grouped doses recived the i-th year of being in the cohort, numeric

dose4 grouped doses recived the i-th year of being in the cohort, numeric

dose5 grouped doses recived the i-th year of being in the cohort, numeric

dose6 grouped doses recived the i-th year of being in the cohort, numeric

dose7 grouped doses recived the i-th year of being in the cohort, numeric

dose8 grouped doses recived the i-th year of being in the cohort, numeric

dose9 grouped doses recived the i-th year of being in the cohort, numeric

dose10 grouped doses recived the i-th year of being in the cohort, numeric

dose11 grouped doses recived the i-th year of being in the cohort, numeric

dose12 grouped doses recived the i-th year of being in the cohort, numeric

dose13 grouped doses recived the i-th year of being in the cohort, numeric

dose14 grouped doses recived the i-th year of being in the cohort, numeric

dose15 grouped doses recived the i-th year of being in the cohort, numeric

dose16 grouped doses recived the i-th year of being in the cohort, numeric

dose17 grouped doses recived the i-th year of being in the cohort, numeric

dose18 grouped doses recived the i-th year of being in the cohort, numeric

dose19 grouped doses recived the i-th year of being in the cohort, numeric
dose20 grouped doses recived the i-th year of being in the cohort, numeric
dose21 grouped doses recived the i-th year of being in the cohort, numeric
dose22 grouped doses recived the i-th year of being in the cohort, numeric
dose23 grouped doses recived the i-th year of being in the cohort, numeric
dose24 grouped doses recived the i-th year of being in the cohort, numeric
dose25 grouped doses recived the i-th year of being in the cohort, numeric
dose26 grouped doses recived the i-th year of being in the cohort, numeric
dose27 grouped doses recived the i-th year of being in the cohort, numeric
dose28 grouped doses recived the i-th year of being in the cohort, numeric
dose29 grouped doses recived the i-th year of being in the cohort, numeric
dose30 grouped doses recived the i-th year of being in the cohort, numeric
dose31 grouped doses recived the i-th year of being in the cohort, numeric
dose32 grouped doses recived the i-th year of being in the cohort, numeric
dose33 grouped doses recived the i-th year of being in the cohort, numeric
dose34 grouped doses recived the i-th year of being in the cohort, numeric
dose35 grouped doses recived the i-th year of being in the cohort, numeric
country country, character

`confint.rERR` *Confidence intervals*

Description

Show the confidence intervals for each parameter of the model. The likelihood ratio test ci for linear varaibles, and the Wald ci for the loglinear terms

Usage

```
## S3 method for class 'rERR'
confint(object, ...)
```

Arguments

`object`	an object of class rERR
`...`	for future methods

Value

a list with the confidence intervals

Examples

```
ci.rERR(fit)
```

f_exclusion *Exclusion period related with the latency - lag period*

Description

Exclude subjects in the cohort that have less than lag time of follow up

Usage

```
f_exclusion(formula, data, lag)
```

Arguments

formula	Surv(entry_time,exit_time,outcome)~loglin(loglin_var1,..,loglin_varn)+ lin(lin_var1,..,lin_varm)+strata(strat_var1,...strat_varp)
data	initial data set
lag	latency period

Value

a data set with the exclusion updated

Examples

```
f_exclusion(formula,data,lag)
```

f_fit_linERR *fit Excess Relative Risk Model (internal use)*

Description

function that calls the optimization (mle from stats4 package, so use optim), and return a rERR object with the estimation and summary

Usage

```
f_fit_linERR(formula, data, rsets, n_lin_vars, n_loglin_vars, id_name,
  time_name)
```

Arguments

formula	Surv(entry_time,exit_time,outcome)~loglin(loglin_var1,..,loglin_varn)+ lin(lin_var1,..,lin_varm)+strata(strat_var1,...strat_varp)
data	data set returned from f_to_model_data
rsets	list of risksets, output of f_risksets
n_lin_vars	number of linear variables (attribute of the to_model_data)
n_loglin_vars	number of loglinear varaibles (attribute of the to_model_data)
id_name	name of variable containing the names of subjects
time_name	name of the time variable

Value

rERR object with the estimation

Examples

```
f_fit_linERR(formula,data,rsets,n_lin_vars,n_loglin_vars,id_name,time_name)
```

---

f_fit_linERR_all *fit Excess Relative Risk Model*

Description

function that calls the optimization (mle from stats4 package, so use optim), and return a rERR object with the estimation and summary

Usage

```
f_fit_linERR_all(formula, data, id_name, time_name, lag)
```

Arguments

formula	Surv(entry_time,exit_time,outcome)~loglin(loglin_var1,..,loglin_varn)+ lin(lin_var1,..,lin_varm)+strata(strat_var1,...strat_varp)
data	data set returned from f_to_model_data
id_name	name of variable containing the names of subjects
time_name	name of the time variable
lag	latency period

Value

rERR object with the estimation

Examples

```
f_fit_linERR_all(formula,data,id_name,time_name,lag)
```

f_fit_linERR_ef *fit Excess Relative Risk Model*

Description

function that calls the optimization (mle from stats4 package, so use optim) from an event format data set, and return a rERR object with the estimation and summary

Usage

```
f_fit_linERR_ef(formula, data, id_name, dose_name, time_name, covars_names, lag,
  exclusion_done = F)
```

Arguments

formula	Surv(entry_time,exit_time,outcome)~loglin(loglin_var1,..,loglin_varn)+ lin(lin_var1,..,lin_varm)+strata(strat_var1,...strat_varp)
data	data set returned from f_to_model_data
id_name	name of variable containing the names of subjects
dose_name	name of variable containing the doses at each time
time_name	name of the time variable
covars_names	a character vector with the names of the variables used as covariates in the fomrula (adjustments and stratification)
lag	latency period
exclusion_done	a logical indicating wheather the exclusion is already done or not

Value

rERR object with the estimation

Examples

```
# set the formulas for the models
formula1  <- Surv(entry_age,exit_age,outcome) ~ lin(dose_cum) + strata(sex)
formula2  <- Surv(entry_age,exit_age,outcome) ~ loglin(factor(country)) + lin(dose_cum) +
                                                strata(sex)

# fit the models
fit1 <- f_fit_linERR_ef(formula1,data=cohort_ef,id_name="id",dose_name="dose",
                        time_name="age",covars_names=c("sex"),lag=2,exclusion_done=TRUE)
fit2 <- f_fit_linERR_ef(formula2,data=cohort_ef,id_name="id",dose_name="dose",
                    time_name="age",covars_names=c("sex","country"),lag=2,exclusion_done=TRUE)

# display a summary
summary(fit1)
summary(fit2)
```

```
# confidence intervals
confint(fit1)
confint(fit2)

# likelihood ratio test between nested and nesting models
f_lrt(fit1,fit2)
```

`f_fit_linERR_wf` *fit Excess Relative Risk Model*

Description

function that calls the optimization (mle from stats4 package, so use optim) from an event format data set, and return a rERR object with the estimation and summary

Usage

```
f_fit_linERR_wf(formula, data, id_name, doses, times, covars, lag,
  exclusion_done = F)
```

Arguments

`formula`	Surv(entry_time,exit_time,outcome)~loglin(loglin_var1,..,loglin_varn)+ lin(lin_var1,..,lin_varm)+strata(strat_var1,...strat_varp)
`data`	data set returned from f_to_model_data
`id_name`	name of variable containing the names of subjects
`doses`	sub data set of grouped doses
`times`	sub data set of times relatives to doses
`covars`	sub data set of the covars that will be involved in the model (adjustments and stratification)
`lag`	latency period
`exclusion_done`	a logical indicating wheather the exclusion is already done or not

Value

rERR object with the estimation

Examples

```
# set the formulas for the models
formula1 <- Surv(AgeAtEntry,age_at_event,outcome) ~ lin(dose_cum) + strata(sex)
formula2 <- Surv(AgeAtEntry,age_at_event,outcome) ~ loglin(factor(country)) + lin(dose_cum) +
                                                    strata(sex)
```

```
# fit the models
fit1 <- f_fit_linERR_wf(formula1,data=cohort_wf,id_name="id",doses=cohort_wf[,45:79],
                        times=cohort_wf[,10:44],covars=cohort_wf[,c("sex","country")],
                        lag=2,exclusion_done = FALSE)

fit2 <- f_fit_linERR_wf(formula2,data=cohort_wf,id_name="id",doses=cohort_wf[,45:79],
                        times=cohort_wf[,10:44],covars=cohort_wf[,c("sex","country")],
                        lag=2,exclusion_done = FALSE)

# display a summary
summary(fit1)
summary(fit2)

# confidence intervals
confint(fit1)
confint(fit2)

# likelihood ratio test between nested and nesting models#'
f_lrt(fit1,fit2)
```

`f_lrt` *likelihood ratio test*

Description

function that thtat compute the lrt test for a nested and a nesting models

Usage

```
f_lrt(fit1, fit2)
```

Arguments

`fit1`	the nested model
`fit2`	the nesting model

Value

a list containing the lrt statistic and the corresponding p_value from the Chi-square test

Examples

```
lrt(fit1,fit2)
```

f_parse_formula *Parse formula (internal use)*

Description

Return list with the terms and elements of the formula

Usage

```
f_parse_formula(formula)
```

Arguments

formula	Surv(entry_time,exit_time,outcome)~loglin(loglin_var1,..,loglin_varn)+ lin(lin_var1,..,lin_varm)+strata(strat_var1,...strat_varp)

Value

list of terms in the formula

Examples

```
f_parse_formula(formula)
```

f_plot_linERR *plot the likelihood*

Description

plot the partial log likelihood function in the case of one dimension in the linear part

Usage

```
f_plot_linERR(object, formula, data, rsets, n_lin_vars, n_loglin_vars, id_name,
  time_name)
```

Arguments

object	An rERR class object
formula	Surv(entry_time,exit_time,outcome)~loglin(loglin_var1,..,loglin_varn)+ lin(lin_var1,..,lin_varm)+strata(strat_var1,...strat_varp)
data	data set returned from f_to_model_data
rsets	list of risksets, output of f_risksets
n_lin_vars	number of linear variables (attribute of the to_model_data)
n_loglin_vars	number of loglinear varaibles (attribute of the to_model_data)
id_name	name of variable containing the names of subjects
time_name	name of the time variable

Value

rERR object with the estimation

Examples

```
f_fit_linERR(formula,data,rsets,n_lin_vars,n_loglin_vars,id_name,time_name)
```

`f_plot_linERR_ef` *plot likelihood function from ef*

Description

plot the partial log likelihood function in the case of one dimension in the linear part

Usage

```
f_plot_linERR_ef(object, formula, data, id_name, dose_name, time_name,
  covars_names, lag, exclusion_done = F)
```

Arguments

`object`	an rERR class object
`formula`	Surv(entry_time,exit_time,outcome)~loglin(loglin_var1,..,loglin_varn)+ lin(lin_var1,..,lin_varm)+strata(strat_var1,...strat_varp)
`data`	data set returned from f_to_model_data
`id_name`	name of variable containing the names of subjects
`dose_name`	name of variable containing the doses at each time
`time_name`	name of the time variable
`covars_names`	a character vector with the names of the variables used as covariates in the fomrula (adjustments and stratification)
`lag`	latency period
`exclusion_done`	a logical indicating wheather the exclusion is already done or not

Value

rERR object with the estimation

Examples

```
# set the formulas for the models
formula1  <- Surv(entry_age,exit_age,outcome) ~ lin(dose_cum) + strata(sex)

# fit the model
fit1 <- f_fit_linERR_ef(formula1,data=cohort_ef,id_name="id",dose_name="dose",
                        time_name="age",covars_names=c("sex"),lag=2,exclusion_done=TRUE)

# plot the partial loglikelihood function
f_plot_linERR_ef(fit1,formula1,data=cohort_ef,id_name="id",dose_name="dose",
                        time_name="age",covars_names=c("sex"),lag=2,exclusion_done=TRUE)
```

---

f_plot_linERR_wf *plot likelihood function from wf*

Description

plot the partial log likelihood function in the case of one dimension in the linear part

Usage

```
f_plot_linERR_wf(object, formula, data, id_name, doses, times, covars, lag,
  exclusion_done = F)
```

Arguments

object	an rERR class object
formula	Surv(entry_time,exit_time,outcome)~loglin(loglin_var1,..,loglin_varn)+ lin(lin_var1,..,lin_varm)+strata(strat_var1,...strat_varp)
data	data set returned from f_to_model_data
id_name	name of variable containing the names of subjects
doses	sub data set of grouped doses
times	sub data set of times relatives to doses
covars	sub data set of the covars that will be involved in the model (adjustments and stratification)
lag	latency period
exclusion_done	a logical indicating wheather the exclusion is already done or not

Value

rERR object with the estimation

Examples

```
# set the formulas for the models
formula1  <- Surv(AgeAtEntry,age_at_event,outcome) ~ lin(dose_cum) + strata(sex)

# fit the model
fit1 <- f_fit_linERR_wf(formula1,data=cohort_wf,id_name="id",doses=cohort_wf[,45:79],
                        times=cohort_wf[,10:44],covars=cohort_wf[,c("sex","country")],
                        lag=2,exclusion_done = FALSE)

# plot the partial loglikelihood function
f_plot_linERR_wf(fit1,formula1,data=cohort_wf,id_name="id",doses=cohort_wf[,45:79],
                 times=cohort_wf[,10:44],covars=cohort_wf[,c("sex","country")],
                 lag=2,exclusion_done = FALSE)
```

`f_risksets` *Rsiksets*

Description

Computes the riskset for each case with the relevant variables in the formula and the stratification vars specified in strata() part of the formula. The riskset of a case inculde the subjects that are in the cohort when the case ocurs: so a subject S belongs to the riskset R of the case that have a 'fail' at time ft, if S_entry_time < ft <= S_exit_time.

Usage

```
f_risksets(formula, data, lag, id_name, time_name)
```

Arguments

`formula`	Surv(entry_time,exit_time,outcome)~loglin(loglin_var1,..,loglin_varn)+ lin(lin_var1,..,lin_varm)+strata(strat_var1,...strat_varp)
`data`	event format data set than is ouput of the functions f_to_event...
`lag`	latency period
`id_name`	name of variable containing the names of subjects
`time_name`	name of the time variable

Value

a named list with integer vectors containing the number of rows that are in each the riskset (relevant person-time)

Examples

```
f_riskset(formula,data,lag=2,id_name='patientids',time_name='time')
```

`f_to_event_table_ef_all` *Data transformation: Event format -> Event format required for the model*

Description

This function organize an input data set ef (event format), to the required ef data set for the model. It appends an index of person event 1,2..n,0 if a subject has n doses and being the 0-row the exit conditions: exit time, exit dose and the outcome. Also creates the cumulated dose.

Usage

```
f_to_event_table_ef_all(formula, data, id_name, dose_name, time_name,
  covars_names)
```

Arguments

`formula`	Surv(entry_time,exit_time,outcome)~loglin(loglin_var1,..,loglin_varn)+ (lin_var1,..,lin_varm)+strata(strat_var1,...strat_varp)
`data`	input data set - event format data set
`id_name`	name of variable containing the names of subjects
`dose_name`	name of the dose variable
`time_name`	name of the time variable
`covars_names`	names of the covars required later in the model

Value

The data set with the event-row format, including the event of exit of the cohort as a row where the outcome is set

Examples

```
f_to_event_table_ef_all(formula,data,id_name='patientids',
                                        dose_name='dose',time_name='time',
                                        covars=c('sex','country','birthcohort'))
```

f_to_event_table_ef_v2 *Data transformation: Event format -> Event format required for the model (internal use)*

Description

This function organize an input data set ef, same as f_to_event_table_ef_all but with start,stop and outcome from the formula

Usage

```
f_to_event_table_ef_v2(id, start, stop, outcome, data, times, doses, covars)
```

Arguments

id	name of variable containing the names of subjects
start	name of the variable containing the start time
stop	name of the variable containing the stop time
outcome	name of the variable containging the outcome
data	input data set
times	name of the time variable
doses	name of the dose variable
covars	names of the covars required later in the model

Value

The data set with the event-row format, including the event of exit of the cohort

Examples

```
f_to_event_table_ef_v2(id='patientids',start='entry_age',stop='exit_age',
                                outcome='leukaemia',data,times='age',doses='ActMar_med',
                                       covars=c('sex','country','birthcohort'))
```

`f_to_event_table_wf_all` *Data transformation: Wide format -> Event format required for the model*

Description

This function organize an input data set wf (wide format, same input as in Epicrue Peanuts), to the required ef data set for the model. It appends an index of person event 1,2..n,0 if a subject has n doses (or grouped doses), and being the 0-row the exit conditions: exit time, exit dose and the outcome. Also creates the cumulated dose.

Usage

```
f_to_event_table_wf_all(formula, data, id_name, doses, times, covars)
```

Arguments

`formula`	Surv(entry_time,exit_time,outcome)~loglin(loglin_var1,..,loglin_varn)+ lin(lin_var1,..,lin_varm)+strata(strat_var1,...strat_varp)
`data`	input data set - event format data set
`id_name`	name of variable containing the names of subjects
`doses`	sub data set of grouped doses
`times`	sub data set of times relatives to doses
`covars`	sub data set of the covars that will be involved in the model

Value

The data set with the event-row format, including the event of exit of the cohort

Examples

```
f_to_event_table_wf_all(formula,data,id_name='patientids',doses=data[,31:50],
                        times=data[,11:30],covars=data[,c('sex','country','birthcohort')])
```

`f_to_event_table_wf_v2` *ata transformation: Wide format -> Event format required for the model (internal use)*

Description

This function organize an input data set wf (wide format) to ef(event format), same as f_to_event_table_wf_all but with the start, stop and outcome from the formula

Usage

```
f_to_event_table_wf_v2(id, start, stop, outcome, data, times, doses, covars)
```

Arguments

id	name of variable containing the names of subjects
start	names of the variable containing the start time
stop	names of the variable containing the stop time
outcome	name of the variable containing the outcome
data	input data set - wide format data set
times	sub data set of times relatives to doses
doses	sub data set of grouped doses
covars	sub data set of the covars that will be involved in the model

Value

The data set with the event-row format, including the event of exit of the cohort

Examples

```
f_to_event_table_wf_v2(id='patientids',start='entry_age',stop='exit_age',
                       outcome='leukaemia',data,times=data[,11:30],doses=data[,31:50],
                            covars=data[,c('sex','country','birthcohort')])
```

---

f_to_model_data *Data transformation: keep model variables and expand categorical variables (internal use)*

Description

Transform the data set in a closed form n_row | id_name | n_pe | entry_name | exit_name | outcome | time |
linear_covariates | loglinear_covariates.
Expand if a variable is categorical to pure logical n_categories variables (excluding the reference category)

Usage

```
f_to_model_data(formula, data, id_name, time_name)
```

Arguments

formula	Surv(entry_time,exit_time,outcome)~loglin(loglin_var1,..,loglin_varn)+ lin(lin_var1,..,lin_varm)+strata(strat_var1,...strat_varp)
data	data set
id_name	name of variable containing the names of subjects
time_name	name of the time variable

Value

data set described below

Examples

```
f_to_model_data(formula,data,id_name='patientids',time_name='time')
```

`print.ci.rERR` *Print ci rERR*

Description

Print the cofnidence intervals of rERR fit

Usage

```
## S3 method for class 'ci.rERR'
print(x, ...)
```

Arguments

x	a ci.ERR object
...	for future methods

`print.summary.rERR` *Print summary rERR*

Description

Print the summary of rERR fit

Usage

```
## S3 method for class 'summary.rERR'
print(x, ...)
```

Arguments

x	a summary.ERR object
...	for future methods

summary.rERR *summary of a ERR fit*

Description

display summary of the parameter and statistics of the model

Usage

```
## S3 method for class 'rERR'
summary(object, ...)
```

Arguments

object	an object of class rERR
...	for future methods

Value

a list with the summary elements

Examples

```
summary(fit)
```

Index