

Package ‘simexaft’

February 15, 2012

Type Package

Title simexaft

Version 1.0.6

Date 2012-01-08

Author Juan Xiong <jxiong2@uwo.ca>, Wen-qing He <whe@stats.uwo.ca>, Grace Y. Yi <yyi@uwaterloo.ca>

Maintainer Juan Xiong <jxiong2@uwo.ca>

Description Implement of the Simulation-Extrapolation (SIMEX) algorithm for the accelerated failure time (AFT) with covariates subject to measurement error.

License GPL

Depends survival, mvtnorm

LazyLoad yes

Repository CRAN

Date/Publication 2012-01-09 06:34:12

R topics documented:

simexaft-package	2
BHS	2
linearextrapolation	4
plotsimexaft	5
print.simexaft	6
quadraticextrapolation	7
rhDNase	8
simexaft	9
summary.simexaft	12

Index	14
--------------	-----------

simexaft-package	<i>SIMEX algorithm for the accelerated failure time model with mismeasured covariates</i>
------------------	---

Description

Implementation of Simulation-Extrapolation (SIMEX) algorithm for the accelerated failure time (AFT) model with mismeasured covariates.

Details

Package:	simexaft
Type:	Package
Version:	1.0.6
Date:	2012-01-08
License:	GPL
Depends:	survival, mvtnorm
LazyLoad:	yes

Author(s)

Juan Xiong <jxiong2@uwo.ca>, Wenqing He <whe@stats.uwo.ca>, Grace Y. Yi <yyi@uwaterloo.ca>
 Maintainer: Juan Xiong <jxiong2@uwo.ca>

References

Genz, A., Bretz, F., Miwa, T., Mi, X., Leisch, F., Scheipl, F. and Hothorn, T. (2011). mvtnorm: Multivariate Normal and t Distributions. R package version 0.9-9991, URL <http://CRAN.R-project.org/package=mvtnorm>.

He, W., Yi, G. Y. and Xiong, J. (2007). Accelerated Failure Time Models with Covariates Subject to Measurement Error. *Statistics in Medicine*, 26, 4817-4832.

Therneau, T. and Lumley, T. (2011). survival: Survival Analysis, Including Penalised Likelihood. R package version 2.36-10, URL <http://CRAN.R-project.org/package=survival>.

 BHS

Busselton Health Study

Description

This dataset is a subset of the Busselton Health study. The Busselton Health study was a repeated cross-sectional survey that was conducted to the community of Busselton in Western Australian.

Usage

```
data(BHS)
```

Format

A data frame with 100 observations on the following 18 variables.

PAIR spouse pair id number

AGE age at survey

SEX sex

SBP systolic blood pressure

DBP diastolic blood pressure

BMI body mass index

CHOL cholesterol level

DIABETES history of diabetes

RXHYPER on blood pressure treatment

CHID history of coronary heart disease

SMOKE smoking status

DRINKING alcohol consumption level

SURVTIME survival time from survey data to date last known alive

DTHCENS censoring indicator

CHDCENS indicator of the death from coronary heart disease

CVDCENS indicator of the death from cardiovascular disease

SMOKE1 indicator of ex-smoker

SMOKE2 indicator of current smoker

Details

This dataset is a subset of the Busselton Health study. The Busselton Health study was a repeated cross-sectional survey that was conducted to the community of Busselton in Western Australian.

Source

He, W., Yi, G. Y. and Xiong, J. (2007). Accelerated Failure Time Models with Covariates Subject to Measurement Error. *Statistics in Medicine*, 26, 4817-4832.

Knuiman, M. W., Cullent, K. J., Bulsara, M. K., Welborn, T. A. and Hobbs, M. S. T. (1994). Mortality trends, 1965 to 1989, in Busselton, The Site of Repeated Health Surveys and Interventions. *Australian Journal of Public Health*, 18, 129-135.

See Also

[simexaft](#)

linearextrapolation *Linear Extrapolation Method*

Description

Linear extrapolation step of SIMEX algorithm.

Usage

```
linearextrapolation(A1, A2, A3, lambda)
```

Arguments

A1	estimates obtained from each level of lambda.
A2	variances estimates obtained from each level of lambda.
A3	scale estimates obtained from each level of lambda.
lambda	vector of lambdas, the grids for the extrapolation step.

Value

reg1	extrapolation back to lambda=-1 yield the SIMEX estimates
reg2	extrapolation back to lambda=-1 yield the SIMEX estimates of variances
scalereg	extrapolation back to lambda=-1 yield the SIMEX estimates of scale

Author(s)

Juan Xiong, Wenqing He and Grace Y. Yi

References

Genz, A., Bretz, F., Miwa, T., Mi, X., Leisch, F., Scheipl, F. and Hothorn, T. (2011). *mvt-norm: Multivariate Normal and t Distributions*. R package version 0.9-9991, URL <http://CRAN.R-project.org/package=mvtnorm>.

He, W., Yi, G. Y. and Xiong, J. (2007). Accelerated Failure Time Models with Covariates Subject to Measurement Error. *Statistics in Medicine*, 26, 4817-4832.

Therneau, T. and Lumley, T. (2011). *survival: Survival Analysis, Including Penalised Likelihood*. R package version 2.36-10, URL <http://CRAN.R-project.org/package=survival>.

See Also

[quadraticextrapolation](#)

plotsimexaft *Plot the Extrapolation Curve*

Description

A function to give the plot of the extrapolation curve for any covariables of the AFT model.

Usage

```
plotsimexaft(obj, var, extrapolation=c("linear", "quadratic", "both"), ylimit)
```

Arguments

obj	an object returned by the function "simexaft".
var	a character string of any covariate used in the AFT model.
extrapolation	a character string giving the type of the extrapolation method, the default is set to be linear extrapolation.
ylimit	the y limits of the plot.

Details

The green points are the average of estimates of B iteration for each labmda.

The linear extrapolation curve is in blue, the corresponding SIMEX estimate is the solid red circle.

The quadratic extrapolation curve is in red, the corresponding SIMEX estimate is the solid blue circle.

The "both" option of the extrapolation method gives both linear and quadratic extrapolation curves.

Author(s)

Juan Xiong, Wenqing He and Grace Y. Yi

See Also

[survreg](#)

Examples

```
#####example for the dataset with known variance.#####
library("simexaft")
data("BHS")
dataset <- BHS
dataset$SBP <- log(dataset$SBP-50)

set.seed(120)
formula <- Surv(SURVTIME, DTHCENS)~SBP+CHOL+AGE+BMI+SMOKE1+SMOKE2

ind <- c("SBP", "CHOL")
```

```

err.mat <- diag(rep(0.5625, 2))

### fit an AFT model with quadratic extrapolation
out2 <- simexaft(formula = formula, data = dataset, SIMEXvariable = ind,
  repeated = FALSE, repind = list(), err.mat = err.mat, B = 50,
  lambda=seq(0, 2, 0.1),extrapolation="quadratic", dist="weibull")

summary(out2)

plotsimexaft(out2,"SBP","both",ylimit=c(-3,1))

```

```

print.simexaft      Print Method for the SIMEXAFT Class

```

Description

Printing the most important values in a clear way.

Usage

```

## S3 method for class 'simexaft'
print(x, digits = max(3, getOption("digits") - 3), ...)

```

Arguments

x	object of class SIEMXAFT.
digits	number of digits to be printed.
...	arguments passed to other functions.

Author(s)

Juan Xiong, Wenqing He and Grace Y. Yi

References

- Genz, A., Bretz, F., Miwa, T., Mi, X., Leisch, F., Scheipl, F. and Hothorn, T. (2011). *mvt-norm: Multivariate Normal and t Distributions*. R package version 0.9-9991, URL <http://CRAN.R-project.org/package=mvtnorm>.
- He, W., Yi, G. Y. and Xiong, J. (2007). Accelerated Failure Time Models with Covariates Subject to Measurement Error. *Statistics in Medicine*, 26, 4817-4832.
- Therneau, T. and Lumley, T. (2011). *survival: Survival Analysis, Including Penalised Likelihood*. R package version 2.36-10, URL <http://CRAN.R-project.org/package=survival>.

quadraticextrapolation

Quadratic Extrapolation Method

Description

Quadratic extrapolation step of SIMEX algorithm.

Usage

```
quadraticextrapolation(A1, A2, A3, lambda)
```

Arguments

A1	estimates obtained from each level of lambda.
A2	variances estimates obtained from each level of lambda.
A3	scale estimates obtained from each level of lambda.
lambda	vector of lambdas, the grids for the extrapolation step.

Value

reg1	extrapolation back to lambda=-1 yield the SIMEX estimates
reg2	extrapolation back to lambda=-1 yield the SIMEX estimates of variances
scalereg	extrapolation back to lambda=-1 yield the SIMEX estimates of scale

Author(s)

Juan Xiong, Wenqing He and Grace Y. Yi

References

Genz, A., Bretz, F., Miwa, T., Mi, X., Leisch, F., Scheipl, F. and Hothorn, T. (2011). *mvt-norm: Multivariate Normal and t Distributions*. R package version 0.9-9991, URL <http://CRAN.R-project.org/package=mvtnorm>.

He, W., Yi, G. Y. and Xiong, J. (2007). Accelerated Failure Time Models with Covariates Subject to Measurement Error. *Statistics in Medicine*, 26, 4817-4832.

Therneau, T. and Lumley, T. (2011). *survival: Survival Analysis, Including Penalised Likelihood*. R package version 2.36-10, URL <http://CRAN.R-project.org/package=survival>.

See Also

[linearextrapolation](#)

rhDNase

*rhDNase Data Set***Description**

This is a dataset reported by Fuchs et al. (1994) for a double-blind randomized multicenter clinical trial designed to evaluate the effect of rhDNase, a recombinant deoxyribonuclease I enzyme, versus placebo on the occurrence of respiratory exacerbations among patients with cystic fibrosis. Data on the occurrence and resolution of all exacerbations were recorded for 645 patients in this trial. For more details about the dataset feature, see Cook and Lawless (2007). Here we only include the first record of the patients that have `etype=1`.

Usage

```
data(rhDNase)
```

Format

A data frame with 641 observations on the following 11 variables.

`id` patient identifier

`trt` the treatment assignment, `trt=1` if patient receive rhDNase and 0 if patient receive placebo

`fev` baseline measurement of forced expiratory volume

`fev2` baseline measurement of forced expiratory volume

`time1` the start of a period indicating when subjects become " at risk" for a transition

`time2` if `etype=1` then `time2` corresponds the onset of an exacerbation (or censoring) and if `etype=2` then `time2` corresponds to the time of a resolution of an exacerbation (or censoring)

`status` status equals 1 if `time2` is a transition time and equals 0 if it is a censoring time

`etype` the indicator of the nature of the event time recorded in `time2`

`enum` the cumulative number of lines in the data frame for each individual

`enum1` the cumulative number of exacerbation-free periods

`enum2` a numeric vector

Source

Cook, R. J. and Lawless, J. F. (2007). The Statistical Analysis of Recurrent Events. Springer, New York.

See Also

[simexaft](#)

simexaft	<i>SIMEX Algorithm for Accelerated Failure Time Model with Covariates Subject to Measurement Error</i>
----------	--

Description

Implementation of the SIMEX algorithm for Accelerated Failure Time model with covariates subject to measurement error.

Usage

```
simexaft(formula = formula(data), data = parent.frame(),
         SIMEXvariable, repeated = FALSE, repind = list(),
         err.mat = err.mat, B = 50, lambda = seq(0, 2, 0.1),
         extrapolation = "quadratic", dist = "weibull")
```

Arguments

formula	specifies the model to be fitted, with the variables coming with data. This argument has the same format as the formula argument in the existing R function "survreg".
data	optional data frame in which to interpret the variables occurring in the formula.
SIMEXvariable	the index of the covariate variables that are subject to measurement error.
repeated	set to TRUE or FALSE to indicate if there are repeated measurements for the mis-measured variables.
repind	the index of the repeated measurement variables for each mis-measured variable. It has an R list form. If repeated = TRUE, repind must be specify.
err.mat	specifies the variables with measurement error, If repeated = FALSE, err.mat must be specify.
B	the number of simulated samples for the simulation step. The default is set to be 50.
lambda	the vector of lambdas, the grids for the extrapolation step.
extrapolation	specifies the function form for the extrapolation step. The options are linear, quadratic and both. The default is set to be quadratic.(first 4 letters are enough)
dist	specifies a parametric distribution that is assumed in AFT model. This argument is the same as the dist option in the existing R function "survreg". These include "weibull", "exponential", "gaussian", "logistic", "lognormal", and "loglogistic".

Details

If the SIMEXvariable is repeated measured then you only need to use arguments repeated and repind without mention err.mat. The summary.simex will contain repind.

Value

coefficient	the corrected coefficients of the AFT model
se	the standard deviation of each coefficient
pvalue	the p-value for the hypothesis of that coefficient equal zero
scalreg	the estimate of the scale
theta	the estimates for every B and lambda
lambda	the vector of lambdas for which the simulation step should be done
B	the number of simulated samples for the simulation step.
formula	the model to be fitted in the survreg function
err.mat	the covariance matrix of the variables with measurement error
repind	the list contains the names of the repeat measurement variables
extrapolation	the extrapolation method: linear ,quadratic are implemented (first 4 letters are enough)
SIMEXvariable	the vector contains the names of the variables with measurement error

Author(s)

Juan Xiong, Wenqing He and Grace Y. Yi

References

- Genz, A., Bretz, F., Miwa, T., Mi, X., Leisch, F., Scheipl, F. and Hothorn, T. (2011). *mvt-norm: Multivariate Normal and t Distributions*. R package version 0.9-9991, URL <http://CRAN.R-project.org/package=mvtnorm>.
- He, W., Yi, G. Y. and Xiong, J. (2007). Accelerated Failure Time Models with Covariates Subject to Measurement Error. *Statistics in Medicine*, 26, 4817-4832.
- Therneau, T. and Lumley, T. (2011). *survival: Survival Analysis, Including Penalised Likelihood*. R package version 2.36-10, URL <http://CRAN.R-project.org/package=survival>.

See Also

[survreg](#), [plotsimexaft](#)

Examples

```
library("simexaft")

data("BHS")

dataset <- BHS

dataset$SBP <- log(dataset$SBP - 50)

###Naive AFT approach
formula <- Surv(SURVTIME,DTHCENS) ~ SBP + CHOL + AGE + BMI + SMOKE1 + SMOKE2
```

```

out1 <- survreg(formula = formula, data = dataset, dist = "weibull")

summary(out1)

###fit a AFT model with quadratic extrapolation
set.seed(120)

ind <- c("SBP", "CHOL")

err.mat <- diag(rep(0.5625, 2))

out2 <- simexaft(formula = formula, data = dataset, SIMEXvariable = ind,
  repeated = FALSE, repind = list(), err.mat = err.mat, B = 50,
  lambda = seq(0, 2, 0.1), extrapolation = "quadratic", dist = "weibull")

summary(out2)

##### repeated measurements #####
data("rhDNase")

###true model
rhDNase$fev.ave <- (rhDNase$fev + rhDNase$fev2)/2

output1 <- survreg(Surv(time2, status) ~ trt + fev.ave, data = rhDNase,
  dist = "weibull")

summary(output1)

####sensitive analysis####
set.seed(120)

fev.error <- rhDNase$fev + rnorm(length(rhDNase$fev), mean = 0,
  sd = 0.15 * sd(rhDNase$fev))

fev.error2 <- rhDNase$fev2 + rnorm(length(rhDNase$fev2), mean = 0,
  sd = 0.15 * sd(rhDNase$fev2))

dataset2 <- cbind(rhDNase[, c("time2", "status", "trt")], fev.error, fev.error2)

formula <- Surv(time2, status) ~ trt + fev.error

ind <- "fev.error"

#####naive model using the average FEV value#####

```

```

fev.error.c <- (fev.error + fev.error2)/2

output2 <- survreg(Surv(time2, status) ~ trt + fev.error.c, data = rhDNase,
                  dist = "weibull")

summary(output2)

#####use simexaft and apply the quadratic extrapolation#####
formula <- Surv(time2, status) ~ trt + fev.error

output3 <- simexaft(formula = formula, data = dataset2, SIMEXvariable = ind,
                   repeated=TRUE,repind=list(c("fev.error", "fev.error2")), err.mat=NULL,
                   B=50, lambda=seq(0,2, 0.1), extrapolation="quadratic", dist="weibull")

summary(output3)

```

summary.simexaft

Summarizing Model fits for the AFT model by SIMEX method

Description

Summary method for the class SIMEXAFF.

Usage

```

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

```

Arguments

object	object of class SIMEXAFT.
...	further arguments.

Value

coefficients	a p x 3 matrix with columns for the estimated coefficient its standard error, corresponding(two-sided) p-value
scalereg	estimate of the scale
extrapolation	the extrapolation method
SIMEXvariable	character vector of the SIMEXvariable

Author(s)

Juan Xiong, Wenqing He and Grace Y. Yi

References

Genz, A., Bretz, F., Miwa, T., Mi, X., Leisch, F., Scheipl, F. and Hothorn, T. (2011). mvt-norm: Multivariate Normal and t Distributions. R package version 0.9-9991, URL <http://CRAN.R-project.org/package=mvtnorm>.

He, W., Yi, G. Y. and Xiong, J. (2007). Accelerated Failure Time Models with Covariates Subject to Measurement Error. *Statistics in Medicine*, 26, 4817-4832.

Therneau, T. and Lumley, T. (2011). survival: Survival Analysis, Including Penalised Likelihood. R package version 2.36-10, URL <http://CRAN.R-project.org/package=survival>.

See Also

[simexaft](#)

Index

*Topic **datasets**

BHS, [2](#)

rhDNase, [8](#)

*Topic **package**

simexaft-package, [2](#)

BHS, [2](#)

linearextrapolation, [4](#), [7](#)

plotsimexaft, [5](#), [10](#)

print.simexaft, [6](#)

quadraticextrapolation, [4](#), [7](#)

rhDNase, [8](#)

simexaft, [3](#), [8](#), [9](#), [13](#)

simexaft-package, [2](#)

summary.simexaft, [12](#)

survreg, [5](#), [10](#)