Package 'MixtureRegLTIC'

August 2, 2017

Type Package

Title Mixture Regression Models for Left-Truncated and Interval-Censored Data

Version 1.0.0

Date 2017-08-02

Author Yuh-Chyuan Tsay [aut, cre], Chen-Hsin Chen [aut], Ya-Chi Wu [aut], Cheng-Fang Horng [aut]

Maintainer Yuh-Chyuan Tsay <yctsay@stat.sinica.edu.tw>

Description Fit mixture regression models with nonsusceptibility/cure for left-truncated and interval-censored (LTIC) data (see Chen et al. (2013) <doi:10.1002/sim.5845>). This package also provides the nonparametric maximum likelihood estimator (NPMLE) for the survival/event curves with LTIC data.

Depends R (>= 3.3.1)

Imports grDevices, graphics, stats, utils

License GPL-2 | GPL-3

Encoding UTF-8

LazyData Yes

NeedsCompilation yes

Repository CRAN

Date/Publication 2017-08-02 08:10:07 UTC

R topics documented:

MixtureRegLTIC-package 2
MixtureLogitAFT 3
NPMLEsurv 5
plotMmixture 7
plotNPMLEsurv 8
plotResidual 9
printMixture 10
simLTICdataA 11
simLTICdataE 12

Index **13**

`MixtureRegLTIC-package` *Mixture Regression Models for Left-Truncated and Interval-Censored Data*

Description

Fit mixture regression models with nonsusceptibility/cure for left-truncated and interval-censored (LTIC) data (see Chen et al. (2013) <doi:10.1002/sim.5845>). This package also provides the nonparametric maximum likelihood estimator (NPMLE) for the survival/event curves with LTIC data.

Details

Package:	MixtureRegLTIC
Type:	Package
Version:	1.0.0
Date:	2017-08-02
Depends:	R (>= 3.3.1)
License:	GPL-2|GPL-3
Encoding:	UTF-8
LazyData:	yes

Author(s)

Yuh-Chyuan Tsay, Chen-Hsin Chen, Ya-Chi Wu, Cheng-Fang Horng
Maintainer: Yuh-Chyuan Tsay <yctsay@stat.sinica.edu.tw>

References

Chen CH, Tsay YC, Wu YC and Horng CF. Logistic-AFT location-scale mixture regression models with nonsusceptibility for left-truncated and general interval-censored data. Statistics in Medicine, 2013; 32:4285–4305.

Frydman H. A note on nonparametric estimation of the distribution function from interval-censored and truncated observations. Journal of the Royal Statistical Society, Series B, 1994; 56:71–74.

Moore RJ. Algorithm AS. 187: derivatives of the incomplete gamma integral. Applied Statistics - Journal of the Royal Statistical Society, Series C, 1982; 31:330–333.

Turnbull BW. The empirical distribution function with arbitrarily grouped, censored and truncated data. Journal of the Royal Statistical Society, Series B, 1976; 38:290–295.

`MixtureLogitAFT` *Fit Logistic-AFT Location-Scale Mixture Regression Models*

Description

A function to fit parametric logistic-AFT location-scale mixture regression models with nonsusceptibility for LTIC data.

Usage

```
MixtureLogitAFT(formula, eventprobreg = ~1, locationreg = ~1, scalereg = ~1,
var.entry, var.mixturetype = NULL, var.weight = NULL,
data, time.origin = 0, shape = NULL)
```

Arguments

`formula`	A formula object specifies the regression response using a survival object of the same form as from the Surv function. The status indicator is 0=right censored, 1=event at time, 2=left censored, and 3=interval censored.
`eventprobreg`	The formula for covariates of the logistic regression part for the event (susceptibility/non-cure) probability in the mixture model. The eventprobreg = NULL corresponds to the one-component model.
`locationreg`	The formula for covariates of the location regression part in the AFT submodel.
`scalereg`	The formula for covariates of the scale regression part in the AFT submodel.
`var.entry`	A variable specifies each study subject's left-truncated time at the entry in the follow-up study. The var.entry = NULL corresponds to no left-truncation.
`var.mixturetype`	A variable specifies the number of fitted component(s) in the mixture regression model for each study subject. The defined value of 1 is for "1-component", and the value of 2 is for "2-component with cure". All subjects in the same stratum must have the same defined values. If all strata have either one-component or two-component models, then var.mixturetype needs not to be included in the data.frame by setting var.mixturetype = NULL.
`var.weight`	A numeric variable specifies the weight for each observation contributed differently to the log-likelihood. The default var.weight = NULL corresponds to equal contributions.
`data`	A data.frame contains the variables named in the formula, eventprobreg, locationreg, scalereg, var.entry, var.mixturetype and var.weight.
`time.origin`	A numeric value specifies the time origin in the AFT submodel. The default time.origin = 0.
`shape`	The default shape = NULL indicates the shape parameter of the generalized log-gamma distribution to be estimated from the data. Otherwise, a numeric value assigns the fixed shape parameter of a specified generalized log-gamma distribution of the error variate to be fitted in the AFT submodel, or a character string "logistic" stands for the logistic distribution of the error variate to be fitted in the AFT submodel.

Details

This function fits the logistic-AFT location-scale mixture regression models with nonsusceptibility/cure for LTIC data. The event time is assumed following either a generalized gamma distribution or a log-logistic distribution, i.e., the error variate in the AFT submodel has a generalized log-gamma distribution or a logistic distribution. The family of generalized gamma distributions for the event time includes many important distributions as its special cases: Weibull (shape = 1), lognormal (shape = 0), and reciprocal Weibull (shape = -1) distributions. Ordinarily, the shape parameter of the generalized gamma distribution is estimated jointly with the logistic, location and scale regression parameters. If a special error distribution is preferred, then the argument shape should be specified.

To obtain the maximum likelihood estimators of the parameters in the logistic-AFT location-scale mixture regression model, R function optim() is applied using the negative log-likelihood function and its first and second derivatives.

As indicated in the scenario (e) of Figure 1 in Chen et al. (2013), some strata may present with nonsusceptibility/cure and the other strata without that. This function can simultaneously handle situations in which both one-component and two-component location-scale regression models jointly emerge as in the above-mentioned scenario. The usage of var.mixturetype facilitates this kind of analysis. However, a consideration of the joint one-component and two-component mixture regression models needs special caution in selecting the covariates and interpreting the corresponding parameter estimates in the logistic regression part.

Note

To obtain the standard errors and confidence intervals of the estimated regression parameters, formulas of the associated first and second derivatives were given in Appendix A of Chen et al. (2013) by the chain rule and the Leibnitz rule. An unexported R function HGNlogL() calculates the negative log-likelihood and the corresponding first and second derivatives with several unexported R functions. In addition, a FORTRAN program with algorithm and codes in Moore (1982) were used to compute the incomplete gamma integral.

References

Chen CH, Tsay YC, Wu YC and Horng CF. Logistic-AFT location-scale mixture regression models with nonsusceptibility for left-truncated and general interval-censored data. Statistics in Medicine, 2013; 32:4285–4305.

Moore RJ. Algorithm AS. 187: derivatives of the incomplete gamma integral. Applied Statistics - Journal of the Royal Statistical Society, Series C, 1982; 31:330–333.

See Also

NPMLEsurv, plotMixture, plotNPMLEsurv, plotResidual, printMixture

Examples

```
data(simLTICdataA)

##### fit the logistic-AFT location-scale model for LTIC data
fit=MixtureLogitAFT(formula=Surv(time1,time2,status)~1,
```

```
    eventprobreg=~X1,locationreg=~X1,scalereg=~X1,
    var.entry="entry",data=simLTICdataA)

##### print regression results of the fitted regression model
printMixture(fit)

##### plot estimated survival curves
#win.graph(width=18,height=10)
#par(mfrow=c(1,2))
plot.fit=plotMixture(fit)
legend(20,0.4,legend=plot.fit$legend,col=plot.fit$col,lty=plot.fit$lty,
       title=" Strata (Case / Total)")

plotD.fit=plotMixture(fit,dist="cond")
legend(3,0.4,legend=plotD.fit$legend,col=plotD.fit$col,lty=plotD.fit$lty,
       title=" Strata (Case / Total)")

##### estimate the NPMLE
est=NPMLEsurv(formula=Surv(time1,time2,status)~X1,var.entry="entry",data=simLTICdataA)

##### plot estimated event curves with both the regression model and NPMLE
#win.graph(width=18,height=10)
#par(mfrow=c(1,2))
plot.fit=plotMixture(fit,curve="event",col=c("red","blue"))
legend(20,1,legend=plot.fit$legend,col=plot.fit$col,lty=plot.fit$lty,
       title=" Strata (Case / Total)")
par(new=TRUE)
plot.NPMLE=plotNPMLEsurv(est,curve="event",lty=c(2,2),col=c("red","blue"))

plotD.fit=plotMixture(fit,curve="event",dist="cond",col=c("red","blue"))
legend(3,1,legend=plotD.fit$legend,col=plotD.fit$col,lty=plotD.fit$lty,
       title=" Strata (Case / Total)")
par(new=TRUE)
plotD.NPMLE=plotNPMLEsurv(est,dist="cond",curve="event",lty=c(2,2),col=c("red","blue"))
```

`NPMLEsurv` *Estimate Overall and Conditional Survival/Event Curve(s) Based on the NPMLE for LTIC Data*

Description

A function to estimate the nonparametric maximum likelihood estimator (NPMLE) of the distribution function for LTIC data based on the Turnbull (1976) method corrected by Frydman (1994).

Usage

```
NPMLEsurv(formula, var.entry, var.weight = NULL, data, time.origin = 0)
```

Arguments

formula	A formula object specifies the regression response on the left of a ~ operator using a survival object of the same form as from the Surv function, and fitted covariates on the right. The status indicator is 0=right censored, 1=event at time, 2=left censored, and 3=interval censored.
var.entry	A variable specifies each study subject's left-truncated time at the entry in the follow-up study. The var.entry = NULL corresponds to no left-truncation.
var.weight	A numeric variable specifies the weight for each observation contributed differently to the log-likelihood. The default var.weight = NULL corresponds to equal contributions.
data	A data.frame contains the variables named in the formula, var.entry and var.weight.
time.origin	A numeric value specifies the time origin. The default time.origin = 0.

References

Frydman H. A note on nonparametric estimation of the distribution function from interval-censored and truncated observations. Journal of the Royal Statistical Society, Series B, 1994; 56:71–74.

Turnbull BW. The empirical distribution function with arbitrarily grouped, censored and truncated data. Journal of the Royal Statistical Society, Series B, 1976; 38:290–295.

See Also

MixtureLogitAFT, plotNPMLEsurv

Examples

```
data(simLTICdataA)

##### estimate the NPMLE
est=NPMLEsurv(formula=Surv(time1,time2,status)~X1,var.entry="entry",data=simLTICdataA)

### plot estimated survival curves with NPMLE
#win.graph(width=18,height=10)
#par(mfrow=c(1,2))
plot.NPMLE=plotNPMLEsurv(est,lty=c(2,2),col=c("red","blue"))
legend(20,0.4,legend=plot.NPMLE$legend,col=plot.NPMLE$col,lty=plot.NPMLE$lty,
       title=" Strata (Case / Total)")

plotD.NPMLE=plotNPMLEsurv(est,dist="cond",lty=c(2,2),col=c("red","blue"))
legend(3,0.4,legend=plotD.NPMLE$legend,col=plotD.NPMLE$col,lty=plotD.NPMLE$lty,
       title=" Strata (Case / Total)")
```

plotMmixture *Plot Survival/Event Curve(s) from the Fitted Regression Model*

Description

A function to plot estimated overall and conditional survival/event curve(s) from the fitted regression models.

Usage

```
plotMixture(fit, dist = "overall", curve = "survival",
xlab = NULL, ylab = NULL, main = NULL, col = NULL, lty = NULL, lwd = 1, axes = T)
```

Arguments

fit	the output object from the fitted mixture regression model.
dist	a character string specified with either "overall" or "cond". The default dist="overall" plots the overall distribution of the event time, and dist="cond" plots the conditional distribution for the logarithm of the event time for susceptible/non-cured subjects.
curve	a character string specifies the type of desired curves to be plotted. The default curve="survival" plots the survival curves, and curve="event" plots event curves.
xlab	the title for x axis.
ylab	the title for y axis.
main	the main title of the plot.
col	a vector of colors.
lty	a vector of line types.
lwd	a numeric value specifies the line width.
axes	a logical value specifies whether axes should be drawn. If axes=FALSE, both x and y axes are not shown .

See Also

MixtureLogitAFT, plotNPMLEsurv

Examples

```
data(simLTICdataE)

##### fit the logistic-AFT location-scale model for LTIC data
fit=MixtureLogitAFT(formula=Surv(time1,time2,status)~1,
    eventprobreg=~X1,locationreg=~X1+X2,scalereg=~X1+X2,
    var.entry="entry",var.mixturetype="mtype",data=simLTICdataE)

##### print regression results of the fitted model
```

```
printMixture(fit)

##### plot estimated event curves of the fitted model
#win.graph(width=18,height=10)
#par(mfrow=c(1,2))
plot.fit=plotMixture(fit,curve="event",col=c("red","blue","deeppink"))
legend(55,0.95,legend=plot.fit$legend,col=plot.fit$col,lty=plot.fit$lty,
       title=" Strata (Case / Total)",cex=0.85)

plotD.fit=plotMixture(fit,dist="cond",curve="event",col=c("red","blue","deeppink"))
legend(3,0.95,legend=plotD.fit$legend,col=plotD.fit$col,lty=plotD.fit$lty,
       title=" Strata (Case / Total)",cex=0.85)
```

---

plotNPMLEsurv *Plot Survival/Event Curve(s) from the NPMLE*

Description

A function to plot overall and conditional survival/event curve(s) from nonparametric maximum likelihood estimators (NPMLEs) of distribution functions.

Usage

```
plotNPMLEsurv(est, dist = "overall", curve = "survival", type = "s",
xlab = NULL, ylab = NULL, main = NULL, col = NULL, lty = NULL, lwd = 1, axes = T)
```

Arguments

est	the output object from the estimated NPMLE
dist	a character string specified with either "overall" or "cond". The default dist="overall" plots the overall distribution of the event time, and dist="cond" plots the conditional distribution for the logarithm of the event time for susceptible/non-cured subjects.
curve	a character string specifies the type of desired curves to be plotted. The default curve="survival" plots the survival curves, and. curve="event" plots event curves.
type	a character string specifies the type of plot, as in plot.
xlab	the title for x axis.
ylab	the title for y axis.
main	the main title of the plot.
col	a vector of colors.
lty	a vector of line types.
lwd	a numeric value specifies the line width.
axes	a logical value specifies whether axes should be drawn. If axes=FALSE, both x and y axes are not shown .

See Also

NPMLEsurv, plotMixture

Examples

```
data(simLTICdataA)

##### estimate the NPMLE #####
est=NPMLEsurv(formula=Surv(time1,time2,status)~X1,var.entry="entry",data=simLTICdataA)

### plot estimated survival curves with NPMLE
#win.graph(width=18,height=10)
#par(mfrow=c(1,2))
plot.NPMLE=plotNPMLEsurv(est,lty=c(2,2),col=c("red","blue"))
legend(20,0.4,legend=plot.NPMLE$legend,col=plot.NPMLE$col,lty=plot.NPMLE$lty,
       title=" Strata (Case / Total)")

plotD.NPMLE=plotNPMLEsurv(est,dist="cond",lty=c(2,2),col=c("red","blue"))
legend(3,0.4,legend=plotD.NPMLE$legend,col=plotD.NPMLE$col,lty=plotD.NPMLE$lty,
       title=" Strata (Case / Total)")
```

---

plotResidual *Estimate and Plot Empirical Distributions of Residuals in the Fitted Regression Model*

Description

A function to estimate and plot estimated empirical distribution functions of the residuals in the fitted regression model.

Usage

```
plotResidual(fit, xlab = NULL, ylab = NULL, main = NULL,
col = NULL, lty = NULL, lwd = 1, axes = T)
```

Arguments

fit	the output object from the fitted mixture regression model.
xlab	the title for x axis.
ylab	the title for y axis.
main	the main title of the plot.
col	a vector of colors.
lty	a vector of line types.
lwd	a numeric value specifies the line width.
axes	a logical value specifies whether axes should be drawn. If axes=FALSE, both x and y axes are not shown .

References

Chen CH, Tsay YC, Wu YC and Horng CF. Logistic-AFT location-scale mixture regression models with nonsusceptibility for left-truncated and general interval-censored data. Statistics in Medicine, 2013; 32:4285–4305.

See Also

MixtureLogitAFT

Examples

```
data(simLTICdataA)

##### fit the logistic-AFT location-scale model for LTIC data
fit=MixtureLogitAFT(formula=Surv(time1,time2,status)~1,
    eventprobreg=~X1,locationreg=~X1,scalereg=~X1,
    var.entry="entry",data=simLTICdataA)

##### plot the empirical distribution of residuals
plot.res=plotResidual(fit)
legend(-9.5,1,legend=plot.res$legend,col=plot.res$col,lty=plot.res$lty,
       title=" Strata (Case / Total)")
```

---

printMixture *Print the Summary Table of Regression Results*

Description

Print the summary table of regression results from the output of mixture regression models

Usage

```
printMixture(fit, digits = 3, file = NULL)
```

Arguments

fit	the output object from the fitted mixture regression model.
digits	an integer indicating the number of decimal places after rounding.
file	a character string of output filename for saving the table of regression results.

See Also

MixtureLogitAFT

Examples

```
data(simLTICdataA)

##### fit the logistic-AFT location-scale model for LTIC data
fit=MixtureLogitAFT(formula=Surv(time1,time2,status)~1,
    eventprobreg=~X1,locationreg=~X1,scalereg=~X1,
    var.entry="entry",data=simLTICdataA)

##### print regression results of the fitted model
printMixture(fit)
```

`simLTICdataA` *A Generated Set with LTIC Data in Simulation Scenario (a) of Chen et al. (2013)*

Description

A generated set with left-truncated and interval-censored data in simulation scenario (a) of Chen et al. (2013) can be used in the package MixtureRegLTIC for fitting mixture regression models.

Usage

```
data(simLTICdataA)
```

Format

There are 200 simulated observations with the following variables:

`time1` left time-endpoint of interval censored data

`time2` right time-endpoint of interval censored data

`status` status indicator with 0=right censored, 1=exact event, 2=left censored, and 3=interval censored.

`entry` time at study entry

`X1` a binary covariate

References

Chen CH, Tsay YC, Wu YC and Horng CF. Logistic-AFT location-scale mixture regression models with nonsusceptibility for left-truncated and general interval-censored data. Statistics in Medicine, 2013; 32:4285–4305.

See Also

`NPMLEsurv`, `MixtureLogitAFT`

`simLTICdataE` *A Generated Set with LTIC Data in Simulation Scenario (e) of Chen et al. (2013)*

Description

A generated set with left-truncated and interval-censored data in simulation scenario (e) of Chen et al. (2013) can be used in the package MixtureRegLTIC for fitting mixture regression models.

Usage

```
data(simLTICdataE)
```

Format

There are 200 simulated observations with the following variables:

`time1` left time-endpoint of interval censored data

`time2` right time-endpoint of interval censored data

`status` status indicator with 0=right censored, 1=exact event, 2=left censored, and 3=interval censored.

`entry` time at study entry

`X1` the first dummy variable of strata with (X1,X2)=(0,0), (0,1) and (1,0).

`X2` the second dummy variable of strata with (X1,X2)=(0,0), (0,1) and (1,0).

References

Chen CH, Tsay YC, Wu YC and Horng CF. Logistic-AFT location-scale mixture regression models with nonsusceptibility for left-truncated and general interval-censored data. Statistics in Medicine, 2013; 32:4285–4305.

See Also

`NPMLEsurv`, `MixtureLogitAFT`

Index

∗Topic **datasets**
simLTICdataA, 11
simLTICdataE, 12

MixtureLogitAFT, 3, *6, 7*, *10–12*
MixtureRegLTIC-package, 2

NPMLEsurv, *4*, 5, *9*, *11, 12*

plotMixture, *4*, *9*
plotMixture *(*plotMmixture*)*, 7
plotMmixture, 7
plotNPMLEsurv, *4*, *6, 7*, 8
plotResidual, *4*, 9
printMixture, *4*, 10

simLTICdataA, 11
simLTICdataE, 12