Package ‘aftgee’

February 19, 2015

Type Package
Title Accelerated Failure Time Model with Generalized Estimating Equations
Version 1.0-0
Date 2014-11-12
Author Sy Han (Steven) Chiou, Sangwook Kang, Jun Yan
Maintainer Sy Han (Steven) Chiou <schiou@umn.edu>
Description This package features both rank-based estimates and least square estimates to the Accelerated Failure Time (AFT) model. For rank-based estimation, it provides approaches that include the computationally efficient Gehan's weight and the general's weight such as the logrank weight. For the least square estimation, the estimating equation is solved with Generalized Estimating Equations (GEE). Moreover, in multivariate cases, the dependence working correlation structure can be specified in GEE's setting.
License GPL (>= 3)
Imports MASS, BB, survival, geepack
Suggests copula
LazyLoad yes
NeedsCompilation yes
Repository CRAN
Date/Publication 2014-11-13 08:06:35

R topics documented:

aftgee-package ......................................................... 2
aftgee ................................................................. 3
aftgee.control ......................................................... 5
aftsrr ................................................................. 6

Index 9
Description

A package that uses Generalized Estimating Equations (GEE) to estimate Multivariate Accelerated Failure Time Model (AFT). This package implements recently developed inference procedures for AFT models with both the rank-based approach and the least squares approach. For the rank-based approach, the package allows various weight choices and uses an induced smoothing procedure that leads to much more efficient computation than the linear programming method. With the rank-based estimator as an initial value, the generalized estimating equation approach is used as an extension of the least squares approach to the multivariate case. Additional sampling weights are incorporated to handle missing data needed as in case-cohort studies or general sampling schemes.

Details

Package: aftgee Type: Package Version: 1.0-0 Date: 2014-11-12 License: GPL(>=3) LazyLoad: yes

Author(s)

Sy Han Chiou <schiou@umn.edu>, Jun Yan <jun.yan@uconn.edu>, Sangwook Kang <kanggi1@yonsei.ac.kr>

References


Description

Fits a semiparametric accelerated failure time (AFT) model with least-squares approach. Generalized estimating equation is generalized to multivariate AFT modeling to account for multivariate dependence through working correlation structures to improve efficiency.

Usage

aftgee(formula, data, subset, id = NULL, contrasts = NULL, weights = NULL, margin = NULL, corstr = "independence", binit = "srrgehan", B = 100, control = aftgee.control())

Arguments

formula a formula expression, of the form response ~ predictors. The response is a Surv object object with right censoring. In the case of no censoring, aftgee will return an ordinary least estimate when corstr = "independence". See the documentation of lm, coxph and formula for details.

data an optional data frame in which to interpret the variables occurring in the formula.

id an optional vector used to identify the clusters. If missing, then each individual row of data is presumed to represent a distinct subject. The length of id should be the same as the number of observations.

subset an optional vector specifying a subset of observations to be used in the fitting process.

corstr a character string specifying the correlation structure. The following are permitted: independence, exchangeable, ar1, unstructured, userdefined, and fixed

B a numeric value specifies the resampling number. When B = 0, only the beta estimate will be displayed.

contrasts an optional list.

binit can be either a vector or a character string specifying the initial slope estimator. When binit is a vector, its length should be the same as the number of covariates. When binit is a character string, the following are permitted: lm for simple linear regression, srrgehan for smoothed gehan weight estimator.

weights an optional vector of observation weights.

margin a sformula vector; default at 1.

control controls maxiter and tolerance.

Details
Value

An object of class "aftgee" representing the fit. An object of class "aftgee" is a list containing at least the following components:

- `coefficients`: a vector of initial value and a vector of point estimates
- `coef.res`: a vector of point estimates
- `var.res`: estimated covariance matrix
- `coef.init`: a vector of initial value
- `var.init.mat`: estimated initial covariance matrix
- `binit`: a character string specifying the initial estimator.
- `conv`: An integer code indicating type of convergence after GEE iteration. 0 indicates successful convergence; 1 indicates that the iteration limit `maxit` has been reached
- `ini.conv`: An integer code indicating type of convergence for initial value. 0 indicates successful convergence; 1 indicates that the iteration limit `maxit` has been reached
- `conv.step`: An integer code indicating the step until convergence

Author(s)

Sy Han Chiou, Sangwook Kang, Jun Yan.

References


Examples

```r
library(survival)
library(copula)
datgen <- function(n = 100, tau = 0.3, cen = 75.4, dim = 2) {
  kt <- iTau(claytonCopula(1), tau)
  copula <- claytonCopula(kt, dim = dim)
  id <- rep(1:n, rep(dim, n))
  x1 <- rbinom(dim * n, 1, 0.5)
  x2 <- rnorm(dim * n)
  ed <- mvdc(copula, rep("weibull", dim), rep(list(list(shape = 1)), dim))
  e <- c(t(Mvdc(n, ed)))
  T <- exp(2 + x1 + x2 + e)
  cstime <- runif(n, 0, cen)
}
```
\begin{verbatim}
delta <- (T < cstime) * 1
Y <- pmin(T, cstime)
out <- data.frame(T = T, Y = Y, delta = delta, x1 = x1, x2 = x2, id = rep(1:dim, n))
}
set.seed(1)
mydata <- datgen(n = 50, dim = 2)
summary(aftgee(Surv(Y, delta) ~ x1 + x2, data = mydata,
           id = id, corstr = "ind", B = 10))
summary(aftgee(Surv(Y, delta) ~ x1 + x2, data = mydata,
           id = id, corstr = "ex", B = 10))
\end{verbatim}

\section*{aftgee.control \hspace{1cm} Auxiliary for Controlling AFTGEE Fitting}

\subsection*{Description}
Auxiliary function as user interface for \texttt{aftgee} and \texttt{aftsrr} fitting.

\subsection*{Usage}
aftgee.control(maxiter = 50,
               reltol = 0.001,
               trace = FALSE)

\subsection*{Arguments}
\begin{itemize}
  \item maxiter \hspace{1cm} max number of iteration.
  \item reltol \hspace{1cm} relative error tolerance.
  \item trace \hspace{1cm} binary variable, determine whether to display output for each iteration.
\end{itemize}

\subsection*{Details}
When \texttt{trace} is true, output for each iteration is printed to the screen.

\subsection*{Value}
A list with the arguments as components.

\subsection*{Author(s)}
Sy Han Chiou, Sangwook Kang, Jun Yan.

\subsection*{See Also}
aftgee, aftgee.fit and aftgee.est.
Accelerated Failure Time with Smooth Rank Regression

Description

Fits a semiparametric accelerated failure time (AFT) model with rank-based approach. General weights, additional sampling weights and fast sandwich variance estimations are also incorporated. Estimating equations are solved with Barzilai-Borwein spectral method implemented as BBsolve in package BB.

Usage

aftsrr(formula, data, subset, id = NULL, contrasts = NULL, strata = NULL, weights = NULL, rankWeights = "gehan", method = "sm", variance = "ISMB", B = 100, SigmaInit = NULL, control = aftgee.control())

Arguments

- **formula**: a formula expression, of the form response ~ predictors. The response is a Surv object object with right censoring. See the documentation of lm, coxph and formula for details.
- **data**: an optional data frame in which to interpret the variables occurring in the formula.
- **id**: an optional vector used to identify the clusters. If missing, then each individual row of data is presumed to represent a distinct subject. The length of id should be the same as the number of observation.
- **subset**: an optional vector specifying a subset of observations to be used in the fitting process.
- **contrasts**: an optional list.
- **strata**: a vector which identifies strata. This can also be used to distinct case-cohort sampling and stratified sampling.
- **weights**: an optional vector of observation weights.
- **rankWeights**: a character string specifying the type of general weights. The following are permitted: logrank: logrank weight, gehan: Gehan’s weight, PW: Prentice-Wilcoxon weight, GP: GP class weight.
- **method**: a character string specifying the methods to estimate the regression parameter. The following are permitted: nonsm: Regression parameters are estimated by directly solving the nonsmooth estimating equations. sm: Regression parameters are estimated by directly solving the smoothed estimating equations. monosm: Regression parameters are estimated by iterating the monotonic smoothed estimating equations. This is typical when rankWeights = "PW" and rankWeights = "GP".
aftsrr

variance a character string specifying the covariance estimating method. The following are permitted: \texttt{MB}: multiplier resampling, \texttt{ZLCF}: Zeng and Lin's approach with closed form Si, \texttt{ZLMB}: Zeng and Lin's approach with empirical Si, \texttt{SHCF}: Huang's approach with closed form Si, \texttt{SHMB}: Huang's approach with empirical Si, \texttt{ISCF}: Johnson and Strawderman's sandwich variance estimates with closed form Si, \texttt{ISMB}: Johnson and Strawderman's sandwich variance estimates with empirical Si, \texttt{JS}: Johnson and Strawderman's iterating approach.

\texttt{B} a numeric value specifies the resampling number. When \texttt{B} = 0, only the beta estimate will be displayed.

\texttt{SigmaInit} the initial covariance matrix; default at identity matrix.

\texttt{control} controls maxiter and tolerance.

\texttt{Value}

\texttt{aftsrr} returns an object of class "aftsrr" representing the fit. An object of class "aftsrr" is a list containing at least the following components:

\texttt{beta} A vector of beta estimates

\texttt{covmat} A list of covariance estimates

\texttt{convergence} An integer code indicating type of convergence. 0 indicates successful convergence. Error codes are 1 indicates that the iteration limit \texttt{maxit} has been reached; 2 is failure due to stagnation; 3 indicates error in function evaluation; 4 is failure due to exceeding 100 step length reductions in line-search; and 5 indicates lack of improvement in objective function.

\texttt{bhist} When variance = "\texttt{MB}" , \texttt{bhist} gives the bootstrap samples.

\texttt{Author(s)}

Sy Han Chiou, Sangwook Kang, Jun Yan.

\texttt{References}


Examples

```r
### kidney data
library(survival)
data(kidney)
foo <- aftsrr(Surv(time, status) ~ age + sex - 1, id = id,
data = kidney, variance = c("ISMB", "ZLMB"), B = 10)
foo

### nwtco data
library(survival)
data(nwtco)
subinx <- sample(1:nrow(nwtco), 668, replace = FALSE)
nwtco$subcohort <- 0
nwtco$subcohort[subinx] <- 1
pn <- table(nwtco$subcohort)[[2]] / sum(table(nwtco$subcohort))
nwtco$hi <- nwtco$rel + (1 - nwtco$rel) * nwtco$subcohort / pn
nwtco$age12 <- nwtco$age / 12
nwtco$study <- nwtco$study - 3
nwtco$histol <- nwtco$histol - 1
sub <- nwtco[subinx]
fit <- aftsrr(Surv(edrel, rel) ~ histol + age12 + study, id = seqno,
weights = hi, data = sub, B = 10, variance = c("ISMB", "ZLMB"),
subset = stage == 4)
summary(fit)
```
Index

*Topic **aftgee**
  - aftgee, 3
  - aftgee-package, 2

*Topic **aftsrr**
  - aftsrr, 6

aftgee, 3
aftgee-package, 2
aftgee.control, 5
aftsrr, 6

summary.aftgee (aftgee), 3
summary.aftsrr (aftsrr), 6