The rankreg Package

August 4, 2006

Maintainer Mai Zhou <mai@ms.uky.edu>
Version 0.2-2
Depends R (>= 2.0.1), quantreg
Author Mai Zhou, Splus original of aft.fun by Jin Zhezhen.
Description Obtain rank regression estimator for the AFT model with right censored data. Testing a given value of the regression coefficient and Re-sampling variance estimator can also be computed.
Title Rank regression for censored data AFT model.
License Free for nonprofit use.

\textbf{R topics documented:}

\begin{itemize}
  \item \texttt{RankRegV} \hspace{4cm} 1
  \item \texttt{aft.fun} \hspace{4cm} 3
  \item \texttt{myeloma} \hspace{4cm} 4
  \item \texttt{rankaft} \hspace{4cm} 4
  \item \texttt{rankreg-internal} \hspace{4cm} 6
  \item \texttt{smallcell} \hspace{4cm} 6
\end{itemize}

Index

\begin{itemize}
  \item \texttt{RankRegV} \hspace{2cm} Test and Variance estimator of rank estimating function in censored AFT model
\end{itemize}

Description

Compute the score type test, with the chi square value for testing \( \text{Ho}: \beta \text{null}\).
It also returns the variance-covariance matrix of the rank estimating function (score function) at \( \beta \) for the censored AFT model.

Usage

\begin{verbatim}
RankRegV(y, d, x, beta, betanull = beta, type="Gehan")
\end{verbatim}
Arguments

- **y**: a vector containing the censored responses in the AFT model.
- **d**: a vector of 1’s and 0’s. censoring indicator. 1(uncensor), 0(censored). Both y and d should be of length n.
- **x**: the design matrix, of size n by q. Should not have a column of 1s.
- **beta**: a vector of dimension q. Usually it should be the solution of the estimation equation, from `rankreg()`.
- **betanull**: the null value (under H0) of beta to be tested.
- **type**: either equal to "Gehan" (default) or "Logrank".

Details

It returns the chi square value and P-value which are based on $EF(\text{betanull}) \cdot V_{\text{test}}(\beta)^{-1} \cdot EF(\text{betanull})$.

The input beta usually should be the solution of the estimating function. This solution can be obtained by using `rankaft`. But other values of beta may also make sense sometime.

When `betanull` equals `beta` from `rankaft`, then the chi square statistic should be zero (since $EF(\text{betanull})=0$) and the P-value should be 1.

Value

A list with `VEF`, which is the variance estimator of the estimating function at `beta`;

- `chisqvalue` which is the quadratic form $EF(\text{betanull}) \cdot \text{VEF}^(-1) \cdot EF(\text{betanull})$;
- and `Pval` that is the P-value from the chi square quantile.

Author(s)

Mai Zhou.

References


Examples

```r
data(myeloma)
RankRegV(y=myeloma[,1],d=myeloma[,2],x=cbind(myeloma[,3],myeloma[,4]),
   beta=c(-15.01117, 1.317596), betanull=c(-15.01117, 1.317596))
### you should get a chisquare = 0.001124574 , Pval = 0.9994379
GP1 <- c(143,164,188,188,190,192,206,209,213,216,220,227,230,234,246,265,304,
216,244)
GP2 <- c(142,156,163,198,205,232,232,233,233,233,239,240,261,280,280,296,
216,244)
status1 <- c(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,0,0)
status2 <- c(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,0,0)
RankRegV(y=c(GP1,GP2),d=c(status1,status2),x=c(rep(0,19),rep(1,21)),
   beta=0,betanull=0)
```
Description

Use linear programming to solve the Gehan rank estimation equation for censored AFT model. Use iterated Gehan type solutions to solve the Logrank estimation equation. Finally, it computes the variance-covariance estimator for both rank regression estimators in the censored AFT model by re-sampling.

Usage

```r
aft.fun(x, y, delta, randomseed=10, weight="logrank", nstep=3, mcsize=100)
```

Arguments

- `x`: the design matrix, of size n by q.
- `y`: a vector containing the censored responses in the AFT model.
- `delta`: a vector of 1’s and 0’s. censoring indicator. 1(uncensor), 0(censored). Both y and d should be of length n.
- `randomseed`: an integer. The number of iterations used to compute the logrank type estimator starting from the Gehan estimator.
- `weight`: an integer. The number of resamples used to compute the variance estimator.

Details

For data sets with more than 400 observations, this function is slow. The reason is that it needs to solve linear programming problems of size n square. So 400 becomes 160000.

Value

A list with the following components. `beta`: first column is the Gehan estimator, the rest are logrank type estimators. `betaw`: the estimates from re-sampling. `covw`: variance-covariance estimator of `beta` from resampling.

Author(s)

Original Splus code by Z. Jin. Adapted to R by Mai Zhou.

References


Description

Krall, Uthoff, and Harley (1975) analyzed data from a study on multiple myeloma in which researchers treated 65 patients with alkylating agents. Of those patients, 48 died during the study and 17 survived. In the data set MYELOMA, the variable TIME represents the survival time in months from diagnosis. The variable VSTATUS consists of two values, 0 and 1, indicating whether the patient was alive or dead, respectively, at the end of the study. If the value of VSTATUS is 0, the corresponding value of TIME is censored. The variables thought to be related to survival are LOGBUN (log BUN at diagnosis), HGB (hemoglobin at diagnosis), PLATELET (platelets at diagnosis: 0=abnormal, 1=normal), AGE (age at diagnosis in years), LOGWBC (log WBC at diagnosis), FRAC (fractures at diagnosis: 0=none, 1=present), LOGPBM (log percentage of plasma cells in bone marrow), PROTEIN (proteinuria at diagnosis), and SCALC (serum calcium at diagnosis).


Usage

data(myloma)

Format

A data frame containing 65 observations on 11 variables:

[1,] "time"
[2,] "vstatus"
[3,] "logBUN"
[4,] "HGB"
[5,] "platelet"
[6,] "age"
[7,] "logWBC"
[8,] "FRAC"
[9,] "logPBM"
[10,] "protein"
[11,] "SCALC"

References


rankaft

Compute Rank estimator in right censored data AFT model
**rankaft**

**Description**

Compute the Gehan and Logrank type rank regression estimators in the censored AFT model, using linear programming. This function is similar to `aft.fun()` except we strip away the re-sampling part in order to speedup things.

**Usage**

```r
rankaft(x, y, delta)
```

**Arguments**

- `x` the design matrix, of size n by q.
- `y` a vector containing the censored responses in the AFT model.
- `delta` a vector of 1’s and 0’s. censoring indicator. 1(uncensor), 0(censored). Both y and delta should be of length n.

**Details**

This program is memory hungry. Caution: at least 1G of RAM needed for sample size 1000; at least 512MB RAM for sample size 400.

We cut the re-sampling part (from `aft.fun()`) to save computing time/memory, and concentrate on the bottleneck.

For statistical inference, there are three options: (1) re-sampling method to estimate the var-cov matrix (available in `aft.fun`) (2) score type test available from function `RankRegV()` and (3) by empirical likelihood (see the reference).

**Value**

A list with `beta` which is the Gehan (betag) and Logrank type (betal) estimate rbinded together; and `residuals`.

**Author(s)**

Original Splus code by Jin Z. Adapted to R by Mai Zhou.

**References**


**Examples**

```r
data(myeloma)
rankaft(x=cbind(myeloma[,3],myeloma[,4]),y=myeloma[,1],delta=myeloma[,2])```
**rankreg-internal**  
*Internal rankreg functions*

**Description**

Internal rankreg functions

**Usage**

```
RankRegEF(y, d, x, beta, type = "Gehan")
```

**Details**

These are not intended to be called by the user.

`RankRegEF` computes the value of the rank estimating function at given `beta`, and is used by `RRV`.

---

**smallcell**  
*Smallcell Lung Cancer Data*

**Description**

There are 121 observations on 4 variables. Arm is the indication of two treatments. Entry is the age of the patient at entry. Survival is the survival time and indicator is the censoring indicator (right censoring). For more details please see the reference below.


**Usage**

```
data(smallcell)
```

**Format**

A data frame containing 121 observations on 4 variables:

```
[,1] "arm"
[,2] "entry"
[,3] "survival"
[,4] "indicator"
```

**References**

Index

*Topic datasets
   myeloma, 4
   smallcell, 6

*Topic htest
   aft.fun, 3
   rankaft, 4
   RankRegV, 1

*Topic internal
   rankreg-internal, 6

*Topic nonparametric
   aft.fun, 3
   rankaft, 4
   RankRegV, 1

aft.fun, 3

myeloma, 4

rankaft, 4
rankreg-internal, 6
RankRegEF(rankreg-internal), 6
RankRegV, 1

smallcell, 6