Package ‘sporm’

February 20, 2015

Type Package

Title Semiparametric proportional odds rate model

Version 1.1

Date 2011-01-18

Author Zhong Guan <zguan@iusb.edu>; Cheng Peng <cpeng@usm.maine.edu>

Maintainer Zhong Guan <zguan@iusb.edu>


License GPL (>= 2)

LazyLoad yes

LazyData yes

Repository CRAN

Date/Publication 2012-10-29 08:59:49

NeedsCompilation no

R topics documented:

  sporm-package .................................................. 2
  confid.int.theta .............................................. 3
  dd.est .......................................................... 4
  EllTheta ....................................................... 5
  elltheta ....................................................... 6
  grad.hessinv ................................................ 7
  H.Binv ......................................................... 8
  ks.sporm ...................................................... 9
  ks.stat ....................................................... 10
  mrle.sporm .................................................. 11
  newton.theta ................................................ 12
  phi ........................................................... 13
  plotor ....................................................... 14
A rank-based empirical likelihood approach to two-sample proportional odds model and its goodness-of-fit. Let \( x_1, \ldots, x_m \) and \( y_1, \ldots, y_n \) be two independent samples from distributions \( F \) and \( G \) that satisfy

\[
\frac{G(x)}{1 - G(x)} \cdot \frac{1 - F(x)}{F(x)} = \frac{G(x)}{1 - G(x)} \cdot \frac{1 - F(x)}{F(x)} = \theta
\]

Function \texttt{mrle.sporm} returns rank-based maximum likelihood estimates of \( \theta, \hat{\theta} \), and probability masses \( p_1, \ldots, p_N \) of \( F \) at the sorted pooled sample values \( z_1 < \cdots < z_N, N = m + n \).

The most important function is \texttt{mrle.sporm} which returns the maximum rank-based likelihood estimates the proportionality parameter \( \theta \) and the baseline distribution. Function \texttt{ks.sporm} is used to do the GOF test of the model assumption using a Kolmogorov-Smirnov type test statistic; \texttt{confid.int.theta} returns a confidence interval for \( \theta \); \texttt{test.theta} does the hypothesis testing for \( \theta \); \texttt{ell.theta} calculates the profile loglikelihood \( \ell(\theta) \) on interval \((\theta_1, \theta_2)\) which contains \( \hat{\theta} \); and plotor plot the empirical odds ratio. Functions \texttt{newton.theta}, \texttt{dd.est} and \texttt{phi} can be used to calculate other initials. There are few internal functions: \texttt{V.theta}, \texttt{H.bin}, \texttt{grad.hess.inv}, \texttt{ks.stat}, and \texttt{ell.theta}. Dataset \texttt{RadarTube} contains the failure times (in days) of two types of radar tubes.

Author(s)

Zhong Guan \(<zguan@iusb.edu>\); Cheng Peng \(<cpeng@usm.maine.edu>\)
Zhong Guan \(<zguan@iusb.edu>\)

References

Description

Confidence interval of the proportionality parameter $\theta$ of the proportional odds rate model

Usage

```r
confid.int.theta(x, y, method = c("chi-sq", "simulate"),
    conf.level = 0.95, grd = 0.001, B = 1000,
    tol = 1e-07, maxit = 500)
```

Arguments

- `x, y` Vectors containing the data values of the two samples $x_1, \ldots, x_m$ and $y_1, \ldots, y_n$.
- `method` A character string specifying the alternative hypothesis, must be one of "chi-sq" (default), "simulate". You can specify just the initial letter.
- `conf.level` Confidence level of the interval.
- `grd` Increment of the grid of points for searching the end-points of the interval
- `B` Number of Monte Carlo trials for approximating the critical values using simulation approach
- `tol` Convergence tolerance used in the Newton iteration
- `maxit` The maximum number of Newton iterations.

Details

See the reference below.

Value

- `theta.L` Lower confidence bound
- `theta.U` Upper confidence bound
- `theta.hat` Maximum rank-based likelihood estimate of theta
- `C.alpha` Critical value

Author(s)

Zhong Guan <zguan@iusb.edu>

References

dd.est

See Also

mrle.sporm.

Examples

# Radar tube life data
z<-RadarTube$Days
v<-RadarTube$Type
x<-z[v==1]; y<-z[v==2]
confid.int.theta(x, y, conf.level=.95, grd = 0.01, B=100)
confid.int.theta(x, y, method= "simulate", conf.level=.95, grd = 0.01, B=100)

dd.est                  Dabrowska-Doksum's estimate of theta

Description

Returns the estimate of the proportionality parameter \( \theta \) of Dabrowska and Doksum (1988)

Usage

dd.est(x, y)

Arguments

x, y

Vectors containing the data values of the two samples \( x_1, \ldots, x_m \) and \( y_1, \ldots, y_n \).

Details

See the references below.

Value

Dabrowska-Doksum’s estimate of theta

Author(s)

Zhong Guan <zguan@iusb.edu>

References


Ell. Theta

Examples
#
# Radar tube life data
z<-RadarTube$Days
v<-RadarTube$Type
x<-z[v==1]; y<-z[v==2]
#
# Dabrowska-Doksum's estimate of theta
dd.est(x,y)
dd.est(y,x)

Ell. Theta

Profile loglikelihood of theta.

Description
Calculates the profile loglikelihood \( \ell(\theta) \) on interval \((\theta_1, \theta_2)\) which contains the maximum rank-based likelihood estimate \( \hat{\theta} \)

Usage
Ell. Theta(x, y, theta.hat, p.hat, theta1, theta2, n.theta = 40,
tol = 1e-07, maxit = 500)

Arguments

- **x, y**: Vectors containing the data values of the two samples \( x_1, \ldots, x_m \) and \( y_1, \ldots, y_n \).
- **theta.hat**: Maximum rank-based likelihood estimate of theta.
- **p.hat**: Maximum rank-based likelihood estimate of \( p \).
- **theta1, theta2**: Left and right end-points of the interval on which the profile loglikelihood is calculated.
- **n.theta**: Number of theta values in the above interval on which the profile loglikelihood is calculated.
- **tol**: Convergence tolerance used in the Newton iteration.
- **maxit**: The maximum number of Newton iterations.

Details
See the reference below.

Value

- **ell**: The profile loglikelihood.
- **theta**: The interval on which the profile loglikelihood is calculated.

Author(s)

Zhong Guan <zguan@iusb.edu>
References


See Also

elltheta, mrlsporm.

Description

Calculating profile loglikelihood for given $\theta$

Usage

elltheta(theta, p0, r, tol=1e-7, maxit=500)

Arguments

- theta: The given $\theta$ value
- p0: Initial values for probability masses $p_1, \ldots, p_N$ of the discretized baseline distribution $F$.
- r: vector of ranks of $y_1, \ldots, y_n$ in the pooled sample $x_1, \ldots, x_m, y_1, \ldots, y_n$
- tol: Convergence tolerance used in the Newton iteration
- maxit: The maximum number of Newton iterations.

Details

See the reference below.

Value

- ell: the profile loglikelihood
- p: the estimated probability masses $p_1, \ldots, p_N$ of the discretized baseline distribution $F$ for the given $\theta$ value.

Author(s)

Zhong Guan <zguan@iusb.edu>

References

grad.hessinv

See Also

Ell.Theta, mrle.sporm.

grad.hessinv  Internal function of package "sporm"

Description

Calculate gradients and the inverse of the Hessian matrix of the loglikelihood.

Usage

grad.hessinv(theta, p, r)

Arguments

theta  Initial value for proportionality parameter $\theta$.
p  Initial value for probability masses $p_1, \ldots, p_N$ of the discretized baseline distribution $F$.
r  vector of ranks of $y_1, \ldots, y_n$ in the pooled sample $x_1, \ldots, x_m, y_1, \ldots, y_n$

Details

See the reference below.

Value

$H$  gradients of the loglikelihood
$A^{-1}$  the inverse of the Hessian matrix

Author(s)

Zhong Guan <zguan@iusb.edu>

References


See Also

mrle.sporm, ks.sporm.
Description

Calculate gradients and the inverse of the Hessian matrix of the profile loglikelihood for a given \( \theta \).

Usage

\[
\text{H.Binv}(\text{theta}, \ p, \ r)
\]

Arguments

- \( \text{theta} \): Given value of the proportionality parameter \( \theta \).
- \( \ p \): Given value of the probability masses \( p_1, \ldots, p_N \) of the discretized baseline distribution \( F \).
- \( \ r \): Vector of ranks of \( y_1, \ldots, y_n \) in the pooled sample \( x_1, \ldots, x_m, y_1, \ldots, y_n \).

Details

See the reference below.

Value

- \( H \): Gradients of the profile loglikelihood
- \( \text{binv} \): The inverse of the Hessian matrix
- \( \text{ell} \): The profile loglikelihood for a given \( \theta \)

Author(s)

Zhong Guan <zguan@iusb.edu>

References


See Also

test.theta, elltheta, Ell Theta.
**ks.sporm**  
*KS test for the semiparametric proportional odds rate model*

---

### Description

Goodness-of-fit test of Kolmogorov-Smirnov type for the semiparametric proportional odds rate model

### Usage

```r
ks.sporm(x, y, B = 1000)
```

### Arguments

- `x`, `y`  
  Vectors containing the data values of the two samples $x_1, \ldots, x_m$ and $y_1, \ldots, y_n$.
- `B`  
  The number of Monte Carlo trials for simulation approach.

### Details

Using the Monte Carlo simulation method to approximate the $p$-value of the KS test statistic which is distribution-free and is calculate by internal function `ks.stat`.

### Value

- `ks`  
  The Kolmogorov-Smirnov type test statistic.
- `pval`  
  The $p$-value of the KS test.

### Author(s)

Zhong Guan <zguan@iusb.edu>

### References


### See Also

- `ks.stat`, `mrle.sporm`.

### Examples

```r
# Radar tube life data  
z<-RadarTube$Days  
v<-RadarTube$Type  
x<-z[v==1]; y<-z[v==2]  
# K-S goodness-of-fit test  
ks.sporm(x,y, B=100)
```
ks.stat

KS statistic for proportional odds rate model

Description
Kolmogorov-Smirnov type test statistic for the goodness-of-fit test of the proportional odds rate model

Usage
ks.stat(x, y)

Arguments
x, y Vectors containing the data values of the two samples \( x_1, \ldots, x_m \) and \( y_1, \ldots, y_n \).

Details
See the reference below.

Value
ks KS statistic for proportional odds rate model
theta estimated proportionality parameter \( \theta \) by mrle.sporm

Author(s)
Zhong Guan <zguan@iusb.edu>

References

See Also
ks.sporm

Examples
# Use radar tube life data
z <- RadarTube$Days
v <- RadarTube$Type
x <- z[v == 1]; y <- z[v == 2]
ks.stat(x, y)
mrle.sorm

Semiparametric proportional odds rate model.

Description

Maximum rank-based likelihood estimates of the proportionality parameter $\theta$ and probability masses of the discretized baseline distribution $F$.

Usage

```r
mrle.sorm(x, y, theta = 1, p = rep(1/(length(x) + length(y)),
    length(x) + length(y)), tol = 1e-07, maxit = 50)
```

Arguments

- `x, y` Vectors containing the data values of the two samples $x_1, \ldots, x_m$ and $y_1, \ldots, y_n$.
- `theta` Initial value for proportionality parameter $\theta$.
- `p` Initial value for probability masses $p_1, \ldots, p_N$ of the discretized baseline distribution $F$.
- `tol` Convergence tolerance used in the Newton iteration.
- `maxit` The maximum number of Newton iterations.

Details

The Newton iteration method is applied to find the maximum rank-based likelihood estimates of the proportionality parameter $\theta$ and probability masses $p_1, \ldots, p_N$ of the discretized baseline distribution $F$. If the default initial values for `theta` and/or `p` do not work, functions `newton.theta`, `dd.est` and `phi` can be used to calculate other initials.

Value

- `theta` The maximum rank-based likelihood estimate of the proportionality parameter $\theta$.
- `p` The maximum rank-based likelihood estimate of probability masses $p_1, \ldots, p_N$ of the discretized baseline distribution $F$.
- `ell` The maximum rank-based loglikelihood.
- `del` Convergent tolerance which is sum of the absolute scores, and absolute changes of the parameters `theta` and `p`.

Author(s)

Zhong Guan <zguan@iusb.edu>
References


Examples

```r
# Use radar tube life data
z <- RadarTube$Days
v <- RadarTube$Type
x <- z[v == 1]; y <- z[v == 2]
# Dabrowska-Doksum's estimate of theta
theta0.hat <- dd.est(x, y)
theta0.hat
vartheta0.hat <- dd.est(y, x)
vartheta0.hat
## mrle
m <- length(x)
n <- length(y)
N <- m + n
lambda <- m/N
phat0 <- phi(N, theta0.hat, lambda)/N
mrle.sporm(x, y, theta0.hat, phat0)
```

newton.theta  

Initial theta value by Newton method

Description

Optional initial \( \theta \) value for mrle.sporm based on uniform \((0, 1)\) baseline distribution and calculated by Newton method.

Usage

```r
newton.theta(y, theta0 = 1, maxit = 100, eps = 1e-10)
```

Arguments

- **y**  
The \( y \)-sample in proportional odds rate model with uniform \((0, 1)\) baseline distribution. If baseline \( F \) is not uniform \((0, 1)\) or unknown, use \( y^*_j = F_n(y_j) \), \( j = 1, \ldots, n \), where \( F_n \) is the empirical cdf of \( x_1, \ldots, x_m \).
- **theta0**  
an initial value of \( \theta \)
- **maxit**  
The maximum number of Newton iterations.
- **eps**  
Convergence tolerance used in the Newton iteration

Details

See the reference below.
phi

Value

Returns the proportionality parameter $\theta$ of the parametric proportional odds rate model with Uniform(0,1) baseline by Newton method.

Author(s)

Zhong Guan <zguan@iusb.edu>

References


See Also

mrle.sporm.

Examples

```r
theta<-2
u<-runif(30)
y<-u/(theta-(theta-1)*u)
newton.theta(y)
```

Description

A function can be used to calculate the approximate probability masses $p_1, \ldots, p_N$ of the discretized baseline distribution $F$.

Usage

```r
phi(N, theta, lambda)
```

Arguments

- `N` integer $N=m+n$, the sum of the two sample sizes
- `theta` the value of the proportionality parameter $\theta$
- `lambda` $m/N$

Details

Returns approximation of probability masses $p=phi(N, theta, lambda)/N$ of the discretized baseline distribution $F$. 
Author(s)
Zhong Guan <zguan@iusb.edu>

References

See Also
mrle.sporm.

Examples

```r
# Use radar tube life data
z<-radarTube$Days
v<-radarTube$Type
x<-z[v==1]; y<-z[v==2]
# Dabrowska-Doksum's estimate of theta
theta0.hat<dod.est(x,y)
m<-length(x)
n<-length(y)
N<-m+n
lambda<-m/N
phat0<-phi(N, theta0.hat, lambda)/N
```

---

**plotor**  
*Empirical odds rate plot*

Description
Plot the empirical odds rate based on empirical distributions of the two samples

Usage
```
plotor(x, y, ...)
```

Arguments
```
x, y  
...  
```

other arguments for `plot`.

Details
See the reference below.

Author(s)
Zhong Guan <zguan@iusb.edu>
References


Examples

```R
# Use radar tube life data
z <- RadarTube$Days
v <- RadarTube$Type
x <- z[v == 1]; y <- z[v == 2]
# Dabrowska-Doksum's estimate of theta
theta0.hat <- dd.est(x, y)
vartheta0.hat <- dd.est(y, x)
# MRLE of theta
m <- length(x); n <- length(y)
N <- m+n; lambda <- m/N
phat0 <- phi(N, theta0.hat, lambda)/N
theta.hat <- mrle.sorm(x, y, theta0.hat, phat0)
## Empirical Odds Ratio Plot
plot(x, y, main = "Empirical Odds Ratio Plot", lwd = 2, ylim = c(0, 2))
abline(h = theta.hat, lwd = 2, lty = 2, col = 2)
abline(h = 1/vartheta0.hat, lwd = 2, lty = 3, col = 3)
abline(h = theta0.hat, lwd = 2, lty = 4, col = 4)
```

RadarTube  
---

Radar Tube Life data

Description

Failure times (in days) of two types of radar tubes

Usage

`data(RadarTube)`

Format

A data frame with 44 observations on the following 2 variables.

Days  a numeric vector
Type  a numeric vector

Details

The dataset contains failure times in days of two types of radar tubes. The sample sizes are m = 25 (Type 1) and n = 19 (Type 2).
Source

The dataset is from Doksum (1975) and Dabrowska and Doksum (1988).

References


Examples

data(RadarTube)
plot(RadarTube)

test.theta

Hypothesis test for proportionality parameter

description

Hypothesis test for the proportionality parameter of the semiparametric proportional odds rate model.

Usage

test.theta(x, y, alternative = c("two.sided", "less", "greater"),
theta = 1, B = 1000, conf.level = 0.95)

Arguments

x, y Vectors containing the data values of the two samples $x_1, \ldots, x_m$ and $y_1, \ldots, y_n$.
alternative A character string specifying the alternative hypothesis, must be one of "two.sided" (default), "greater" or "less". You can specify just the initial letter.
theta The hypothesized values of $\theta$.
B The number of number Monte Carlo trials for simulation approach.
conf.level Confidence level of the test.

Details

Using the Monte Carlo simulation method to approximate the p-value of the test statistic which is distribution-free.

Value

theta Maximum rank-based likelihood estimate of $\theta$
p-value The p-value of the test statistic
V.\theta

Author(s)
Zhong Guan <zguan@iusb.edu>

References

See Also
mrle.sporm.

Examples
# Radar tube life data
z<-RadarTube$Days
v<-RadarTube$Type
x<-z[v==1]; y<-z[v==2]
test.theta(x, y, B=100)

\begin{verbatim}
V.\theta
\end{verbatim}

Internal function of "sporm"

Description
Internal function of sporm for calculating \( V_\theta(t) \).

Usage
V.\theta(t, N, theta, lambda)

Arguments
\begin{itemize}
\item \( t \) vector of numbers
\item \( N \) integer \( N=m+n \)
\item \( \theta \) the value of the proportinality parameter \( \theta \)
\item \( \lambda \) mixture proportion \( \lambda = m/N \)
\end{itemize}

Value
Returns the value of \( V_\theta(t) \).

Author(s)
Zhong Guan <zguan@iusb.edu>
References

Index

*Topic **datasets**
  RadarTube, 15
*Topic **package**
  sporm-package, 2

confid.int.theta, 3

dd.est, 4, 11

Ell.Theta, 5, 7, 8
elltheta, 6, 6, 8

grad.hessinv, 7

H.Binv, 8

ks.sporm, 7, 9, 10
ks.stat, 9, 10

mrle.sporm, 4, 6, 7, 9, 10, 11, 13, 14, 17

newton.theta, 11, 12

phi, 11, 13
plot, 14
plotor, 14

RadarTube, 15

sporm(sporm-package), 2
sporm-package, 2

test.theta, 8, 16

V.theta, 17