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>

Description R implementation of the methods described in ``A rank-based empirical likelihood approach to two-sample proportional odds model and its goodness-of-fit" by Zhong Guan and Cheng Peng, Journal of Nonparametric Statistics, to appear.

License GPL (>= 2)

LazyLoad yes

LazyData yes

Repository CRAN

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

NeedsCompilation no

R topics documented:

`sporm-package` *Semiparametric proportional odds rate model*

Description

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

$$[G(x)/\{1-G(x)\}]/[F(x)\{1-F(x)\}] = G(x)\{1-F(x)\}/[F(x)\{1-G(x)\}] = \theta$$

Function `mele.theta.p` returns rank-based maximum likelihood estimates of θ, $\hat{\theta}$, and probability masses $p_1, \ldots, p_N$ of F at the sorted pooled sample values $z_1 < \cdots < z_N$, $N = m + n$.

Details

Package:	sporm
Type:	Package
Version:	1.0
Date:	2011-01-01
License:	GPL 2.0 or newer
LazyLoad:	yes

The most important function is `mrle.sporm` which returns the maximum rank-based likelihood estimates the proportionality paramter θ and the baseline distribution. Function `ks.sporm` is used to do the GOF test of the model assumption using a Kolmogorov-Smirnov type test statistic; `confid.int.theta` returns a confidence interval for θ; `test.theta` does the hypothesis testing for θ; `Ell.Theta` calculates the profile loglikelihood $\ell(\theta)$ on interval (θ_1, θ_2) which contains $\hat{\theta}$; and `plotor` plot the empirical odds ratio. Functions `newton.theta`, `dd.est` and `phi` can be used to calculate other initials. There are few internal functions: `V.theta`, `H.Binv`, `grad.hessinv`, `ks.stat`, and `elltheta`. Dataset `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

Zhong Guan and Cheng Peng (2011), "A rank-based empirical likelihood approach to two-sample proportional odds model and its goodness-of-fit", Journal of Nonparametric Statistics, to appear.

confid.int.theta *Confidence interval of the proportionality parameter*

Description

Confidence interval of the proportionality parameter θ of the proportional odds rate model

Usage

```
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

Zhong Guan and Cheng Peng (2011), "A rank-based empirical likelihood approach to two-sample proportional odds model and its goodness-of-fit", Journal of Nonparametric Statistics, to appear.

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 θ 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

Zhong Guan and Cheng Peng (2011), "A rank-based empirical likelihood approach to two-sample proportional odds model and its goodness-of-fit", Journal of Nonparametric Statistics, to appear.

D. M. Dabrowska and K. A. Doksum (1988), Estimation and testing in a two-sample generalized odds-rate model, J. Amer. Statist. Assoc. 83 pp. 744–749.

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 (θ_1, θ_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

Zhong Guan and Cheng Peng (2011), "A rank-based empirical likelihood approach to two-sample proportional odds model and its goodness-of-fit", Journal of Nonparametric Statistics, to appear.

See Also

elltheta, mrle.sporm.

---

elltheta *Internal function of "sporm"*

Description

Calculating profile loglikelihood for given θ

Usage

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

Arguments

theta	The given θ 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 θ value.

Author(s)

Zhong Guan <zguan@iusb.edu>

References

Zhong Guan and Cheng Peng (2011), "A rank-based empirical likelihood approach to two-sample proportional odds model and its goodness-of-fit", Journal of Nonparametric Statistics, to appear.

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 θ.
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
Ainv	the inverse of the Hessian matrix

Author(s)

Zhong Guan <zguan@iusb.edu>

References

Zhong Guan and Cheng Peng (2011), "A rank-based empirical likelihood approach to two-sample proportional odds model and its goodness-of-fit", Journal of Nonparametric Statistics, to appear.

See Also

mrle.sporm, ks.sporm.

H.Binv *Internal function of package "sporm"*

Description

Calculate gradients and the inverse of the Hessian matrix of the profile loglikelihood for a given θ.

Usage

```
H.Binv(theta, p, r)
```

Arguments

theta	Given value of the proportionality parameter θ.
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
Binv	the inverse of the Hessian matrix
ell	the profile loglikelihood for a given θ

Author(s)

Zhong Guan <zguan@iusb.edu>

References

Zhong Guan and Cheng Peng (2011), "A rank-based empirical likelihood approach to two-sample proportional odds model and its goodness-of-fit", Journal of Nonparametric Statistics, to appear.

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

```
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-vakue 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

Zhong Guan and Cheng Peng (2011), "A rank-based empirical likelihood approach to two-sample proportional odds model and its goodness-of-fit", Journal of Nonparametric Statistics, to appear.

See Also

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

Examples

```
# 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 θ by `mrle.sporm`

Author(s)

Zhong Guan <zguan@iusb.edu>

References

Zhong Guan and Cheng Peng (2011), "A rank-based empirical likelihood approach to two-sample proportional odds model and its goodness-of-fit", Journal of Nonparametric Statistics, to appear.

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.sporm` *Semiparametric proportional odds rate model.*

Description

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

Usage

```
mrle.sporm(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 θ.
`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 θ 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 θ.
`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

Zhong Guan and Cheng Peng (2011), "A rank-based empirical likelihood approach to two-sample proportional odds model and its goodness-of-fit", Journal of Nonparametric Statistics, to appear.

Examples

```
# 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 θ value for `mrle.sporm` based on uniform (0, 1) baseline distrbution and calculated by Newton method.

Usage

```
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 distrbution. If baseline F is not uniform (0,1) or unknown, use $y_i^* = F_n(y_j)$, $j = 1, \ldots, n$, where F_n is the empirical cdf of $x_1, \ldots, x_m$.
`theta0`	an initial value of θ
`maxit`	The maximum number of Newton iterations.
`eps`	Convergence tolerance used in the Newton iteration

Details

See the reference below.

Value

Returns the proportionality parameter θ of the parametric proportional odds rate model with Uniform(0,1) baseline by Newtom method.

Author(s)

Zhong Guan <zguan@iusb.edu>

References

Zhong Guan and Cheng Peng (2011), "A rank-based empirical likelihood approach to two-sample proportional odds model and its goodness-of-fit", Journal of Nonparametric Statistics, to appear.

See Also

`mrle.sporm`.

Examples

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

`phi` *Function phi(t)*

Description

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

Usage

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

Arguments

`N`	integer N=m+n, the sum of the two smple sizes
`theta`	the value of the proportionality parameter θ
`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

Zhong Guan and Cheng Peng (2011), "A rank-based empirical likelihood approach to two-sample proportional odds model and its goodness-of-fit", Journal of Nonparametric Statistics, to appear.

See Also

mrle.sporm.

Examples

```
# 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)
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	Vectirs containing the data values of the two samples $x_1, \ldots, x_m$ and $y_1, \ldots, y_n$.
...	other arguments for plot.

Details

See the reference below.

Author(s)

Zhong Guan <zguan@iusb.edu>

References

Zhong Guan and Cheng Peng (2011), "A rank-based empirical likelihood approach to two-sample proportional odds model and its goodness-of-fit", Journal of Nonparametric Statistics, to appear.

Examples

```
# 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.sporm(x, y, theta0.hat, phat0)$theta
## Empirical Odds Raio Plot
plotor(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

K. A. Doksum (1975), Measures of differences in reliability, in Reliability and fault tree analysiseds, J. B. F. R. E. Barlow and N. D. Singpurwalla, eds., Society for Industrial and Applied Mathematics, Philadelphia, pp. 427–449,

D. M. Dabrowska and K.~A. Doksum (1988), Estimation and testing in a two-sample generalized odds-rate model, J. Amer. Statist. Assoc. 83 pp. 744–749.

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 θ.
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 θ
p-value	The p-vakue of the test statistic

Author(s)

Zhong Guan <zguan@iusb.edu>

References

Zhong Guan and Cheng Peng (2011), "A rank-based empirical likelihood approach to two-sample proportional odds model and its goodness-of-fit", Journal of Nonparametric Statistics, to appear.

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)
```

V.theta *Internal function of "sporm"*

Description

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

Usage

```
V.theta(t, N, theta, lambda)
```

Arguments

t	vector of numbers
N	integer N=m+n
theta	the value of the proportinality parameter θ
lambda	mixture proportion $\lambda = m/N$

Value

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

Author(s)

Zhong Guan <zguan@iusb.edu>

References

Zhong Guan and Cheng Peng (2011), "A rank-based empirical likelihood approach to two-sample proportional odds model and its goodness-of-fit", Journal of Nonparametric Statistics, to appear.

Index