Package ‘npsm’

January 25, 2019

Type Package
Title Package for Nonparametric Statistical Methods using R
Version 0.5.1
Date 2014-11-05
Author John Kloke, Joseph McKeen
Maintainer John Kloke <kloke@biostat.wisc.edu>
Description Functions and datasets used in the book Nonparametric Statistical Methods Using R.
License GPL (>= 2)
LazyLoad yes
LazyData yes
Depends Rfit
Imports methods, plyr
Suggests boot, survival, sm, HSAUR2, devtools
NeedsCompilation no
Repository CRAN
Date/Publication 2019-01-25 20:36:53 UTC

R topics documented:

aligned.test ................................................................. 2
bb2010 ................................................................. 3
blood.plasma ............................................................. 4
cancertrt ................................................................. 5
centerx ................................................................. 5
cloud ................................................................. 6
cor.boot.ci .............................................................. 7
energy ................................................................. 8
firstbase ................................................................. 8
fk.test ................................................................. 9
fkk.test .............................................................. 10
fp.test .............................................................. 11
Aligned rank test for a group/treatment effect after adjusting for covariates.

Usage

aligned.test(x, y, g, scores = Rfit::wscores)
Arguments

- x: n by p design matrix
- y: n by 1 response vector
- g: n by 1 vector denoting group/treatment membership.
- scores: Which scores should be used for the fit and the test. An object of class scores.

Value

- statistic: The value of the test statistic.
- p.value: The p-value based on a chisq(k-1) distribution where k is the number of groups/treatments.

Author(s)

John Kloke <kloke@biostat.wisc.edu>

References


Examples

```r
y<-rt(30,2)
x<-runif(30)
g<-rep(1:3,each=10)
aligned.test(x,y,g)
```

bb2010

Batting statistics for the 2010 baseball season.

Description

Batting (average, home runs, RBIs) statistics for 2010 full time players. By full time we mean that the batter had at least 450 official at bats during the season.

Usage

data(bb2010)

Format

A data frame with 122 observations on the following 3 variables.

- ave: batting average
- hr: home runs
- rbi: runs batted in
**Source**

baseballguru.com

**Examples**

```r
plot(hr~ave, data=bb2010)
```

---

**blood.plasma**

*Blood plasma measurements related to total triglyceride level*

**Description**

Data table from Table 9.11 of Hollander and Wolfe (1999). The data consists of triglyceride levels on 13 patients. Two factors, each at two levels, were recorded: Sex and Obesity. The concomitant variables are chylomicrons, age, and three lipid variables (very low-density lipoproteins (VLDL), low-density lipoproteins (LDL), and high-density lipoproteins (HDL)).

**Usage**

```r
data(blood.plasma)
```

**Source**


**References**


**Examples**

```r
data(blood.plasma)
plot(Total~Age, data=blood.plasma)
boxplot(Total~Obese, data=blood.plasma)
```
**cancertrt**

**Survival time based on two treatments**

**Description**

Survival times (in days) for undergoing standard treatment (S) and a new treatment (N).

**Usage**

```r
data("cancertrt")
```

**Format**

A data frame with 17 observations on the following 3 variables.

- **time**: Survival time in days
- **event**: Indicator for event
- **trt**: a factor with levels `N` `S`

**References**


**Examples**

```r
data(cancertrt)
with(cancertrt, gehan.test(time, event, trt))
```

**centerx**

**Center Matrix**

**Description**

Centers a matrix.

**Usage**

```r
centerx(x)
```

**Arguments**

- `x`: a matrix
Details
Returns a centered matrix, i.e., each column of the matrix is replaced by deviations from its column mean.

Value
The centered matrix.

Author(s)
John Kloke, Joseph McKean

See Also
scale

Examples
```r
x <- cbind(seq(1,5,length=5), seq(10,20,length=5))
xc <- center(x)
apply(xc, 1, mean)
```

---

### cloud

#### Cloud Dewpoint

**Description**
A regression example with response cloud point of a liquid and predictor the percent of Iodine 8 added to the liquid; see Chapter 3 of Hettmansperger and McKean (2011) or Exercise 4.9.10 of Kloke and McKean (2014).

**Usage**
data(cloud)

**Format**
Nineteen observations on two variables.

- `cloud.point` Cloud point of the liquid
- `I8` Percent Iodine 8 added

**References**

**cor.boot.ci**

Examples

```r
cfit(cloud.point ~ I8, data=cloud)
```

---

**cor.boot.ci**

*Confidence interval for a correlation based on a bootstrap.*

---

**Description**

Returns a bootstrap confidence interval for any of the correlations available in the base R `cor` function.

**Usage**

```r
cor.boot.ci(x, y, method = "spearman", conf = 0.95, nbs = 3000)
```

**Arguments**

- **x**: n by 1 vector
- **y**: n by 1 vector
- **method**: Which correlation to use. Argument passed to `cor`.
- **conf**: Confidence level.
- **nbs**: number of bootstrap samples to base CI on.

**Details**

Obtains a percentile bootstrap confidence interval. The bootstrap samples are obtained via the function `boot`.

**Value**

A confidence interval.

**Author(s)**

John Kloke <kloke@biostat.wisc.edu>, Joseph McKean

**See Also**

See Also as `cor`

**Examples**

```r
library(boot)
with(bb2010, cor.boot.ci(ave, hr))
```
Energy as a Function of temperature difference.

Description
A regression example with response energy output in watts and the predictor temperature difference in degrees Kelvin; see Devore (2012) and Exercise 4.9.11 of Kloke and McKean (2014).

Usage
data(energy)

Format
Twenty-four observations on two variables.

output  Energy output in watts
temp.diff  Temperature difference in K

Source

References

Examples
rfit(output ~ temp.diff, data=energy)

Rounding First Base.

Description
The amount of time it took 22 baseball players to round first base for each of three methods of rounding.

Usage
data(firstbase)
A data frame with 22 observations on the following 3 variables.

- `round.out` Time when using round out method.
- `narrow.angle` Time when using narrow angle method.
- `wide.angle` Time when using wide angle method.

Details

Rounding methods are illustrated in Figure 7.1 of Hollander and Wolfe (1999).

**Source**


**References**


---

**Description**

Returns the Fligner-Kileen test for homogeneous scales for two-samples. Also estimates of ratio of scales based on the logs of folded median-aligned samples and a corresponding confidence interval is computed. `fk.test` computes the value of the statistic based on squared-normal scores following the optimal (for normal errors) such test described in Section 2.10 of Hettmansperger and McKean (2011). Hence, it will differ from the core R routine `fligner.test`; see the discussion in Section 3.3 of Kloke and McKean (2014).

**Usage**

```r
fk.test(x, y, alternative = c("two.sided", "less", "greater"), conf.level = 0.95)
```

**Arguments**

- `x` vector of first sample responses
- `y` vector of second sample responses
- `alternative` alternative indicator for hypotheses
- `conf.level` confidence coefficient for the returned confidence intervals

**Details**

Returns the Fligner-Kileen test for the two-sample scale problem.
fkk.test

Value

- statistic: chi-squared test statistic
- p.value: p-value of the test
- estimate: vector of estimates of ratio of scales
- conf.int: table of confidence intervals

Author(s)

John Kloke <kloke@biostat.wisc.edu>, Joseph McKean

References


See Also

fkk.test

Examples

```r
x <- rnorm(18)
y <- rnorm(22) * 3
fk.test(x, y)
```

fkk.test

- k-Sample version of the Fligner-Killeen test for homogeneous scales.

Description

Returns the Fligner-Killeen test for homogeneous scales for k-samples. Also estimates of ratio of scales based on the logs of folded median-aligned samples and a corresponding confidence interval is computed. The first level (sample) is referenced. See the discussion in Section 5.7 of Kloke and McKean (2014).

Usage

```r
fkk.test(y, ind, conf.level = 0.95)
```

Arguments

- y: vector of responses
- ind: vector of corresponding levels
- conf.level: confidence coefficient for the returned confidence intervals
Details

Returns the Fligner-Killeen test for the k-sample scale problem.

Value

- statistic: chi-squared test statistic
- p.value: p-value of the test
- estimate: vector of estimates of ratio of scales
- conf.int: table of confidence intervals
- cwts: vector of weights based on the estimates difference in scales

Author(s)

John Kloke <kloke@biostat.wisc.edu>, Joseph McKean

References


See Also

fk.test

Examples

```r
y1 <- rnorm(10)
y2 <- rnorm(12)*3
y3 <- rnorm(15)*5
y <- c(y1, y2, y3)
ind <- rep(1:3, times=rep(10,12,15))
fk.test(y, ind)
```

fp.test

*Placement Test for the Behrens-Fisher problem.*

Description

Returns the test based on placements for the Behrens-Fisher problem. This test was developed by Fligner and Policello (1981); see, also, Section 2.11 of Hettmansperger and McKean (2011) and Section 4.4 of Hollander and Wolfe (1999). The version computed by fp.test is discussed in Section 4.4 of Kloke and McKean (2014).
Usage

fp.test(x, y, delta0 = 0, alternative = "two.sided")

Arguments

x vector of first sample responses
y vector of second sample responses
delta0 null value tested
alternative alternative indicator for hypotheses

Details

Returns the Placement Test for the Behrens-Fisher problem.

Value

statistic chi-squared test statistic
p.value p-value of the test
numerator numerator of test statistic
denominator denominator of test statistic

Author(s)

John Kloke <kloke@biostat.wisc.edu>, Joseph McKean

References


gehan.test  

Gehan generalization the Wilcoxon two-sample test

Description

Generalization of the Wilcoxon rank sum which allows for censored data.

Usage

gehan.test(time, event, trt)

Arguments

time  Time of event or of censoring

event  Indicator variable representing a event occur or not (time is censored)

trt  Variable indicating treatment group.

Value

statistic  Value of the test statistic

p.value  p-value

Author(s)

John Kloke <kloke@biostat.wisc.edu>

References


Examples

```R
n<-76
y<-rexp(n)

# about 30% censored
event<-rbinom(n,1,0.7)

trt<-sample(c(0,1),n,replace=TRUE)

gehan.test(y,event,trt)
```
getxact  
*Design Function for Robust Analysis of Covariance*

**Description**

Returns the heterogeneous slopes design matrix used in ANCOVA. It references the first level.

**Usage**

```r
getchxact(amat, bmat)
```

**Arguments**

- `amat`  
  cell mean design matrix of factor.
- `bmat`  
  matrix of covariates.

**Details**

Returns the heterogeneous slopes analysis of covariance matrix.

**Value**

- `cmat`  
  heterogeneous slopes analysis of covariance matrix

**References**


---

getxact2  
*Design Function for Robust Analysis of Covariance*

**Description**

Returns the heterogeneous slopes design matrix used in ANCOVA. It references the first level. Also, column names are supplied.

**Usage**

```r
getchxact2(amat, bmat)
```

**Arguments**

- `amat`  
  cell mean design matrix of factor.
- `bmat`  
  matrix of covariates.
hemorrhage

Details

Returns the heterogeneous slopes analysis of covariance matrix.

Value

cmat heterogeneous slopes analysis of covariance matrix eith columns named

References


Hemorrhage data from Dupont.

Description

Hemorrhage data from Dupont.

Usage

data(hemorrhage)

Format

A data frame with 71 observations on the following 3 variables.

genotype a numeric vector
time a numeric vector
recur a numeric vector

References

Dupont

Examples

data(hemorrhage)
## maybe str(hemorrhage) ; plot(hemorrhage) ...
hodgkins  

Relapse-Free Survival Times for Hodgkin’s Disease Patients

Description

These data are described in Example 11.7 of Hollander and Wolfe (1999). Results from a clinical trial in early Hodgkin’s disease. Subjects received one of two treatments: radiation of affected node (AN) or total nodal radiation (TN).

Usage

data("hodgkins")

Format

A data frame with 49 observations on the following 3 variables.

- **time**  Survival time
- **relapse**  Indicator variable for relapse
- **trt**  treatment: a factor with levels AN TN

References


hogg.test  

Hogg’s Adaptive Test

Description

Based on selector statistics (Q1 & Q2) one of four score functions is choosen. A rank test and p-value is then calculated based on it.

Usage

hogg.test(x, y, ...)

Arguments

- **x**  n by 1 vector
- **y**  m by 1 vector
- **...**  additional arguments. currently not used
Value

statistic  Value of the test statistic.
p.value  p-value based on a normal approximation.
scores  Which of the score functions was chosen.

Author(s)

John Kloke <kloke@biostat.wisc.edu>, Patrick Kimes

References


Examples

hogg.test(rt(20,1),rt(22,1)+0.2)

Description

Q1 is a measure of skewness and Q2 is a measure of tail heaviness.

Usage

Q1(z)

Arguments

z  n by 1 vector

Details

Used as selector statistics in adaptive schemes. Both Q1 and Q2 are ratios. For Q1, the numerator is upper 5% mean minus the middle 50% mean, while the denominator is difference between the middle 5% mean and the lower 5% mean. For Q2, the numerator is upper 5% mean minus the lower 5% mean, while the denominator is difference between the upper 50% mean and the lower 50% mean. These statistics are not robust.

Author(s)

John Kloke <kloke@biostat.wisc.edu>
References


See Also

hogg.test, ~~~

---

**huitema496**  
*Analysis of Covariance Data Set*

**Description**

A data set presented on Page 496 of huitema (2011). The design is a 2 by 2 with one covariate.

**Usage**

```r
data(huitema496)
```

**Format**

A 16 by 4 array with the following 4 columns:

- `y` number of novel responses.
- `i` type of reinforcement (2 levels).
- `j` type of program (2 levels).
- `x` covariate, a measure of verbal fluency.

**Details**

Discussion can be found in both references listed below.

**Source**


**References**


**Insulating Fluid Data**

**Description**

Study the breakdown time of an electrical insulating fluid subject to seven different levels of voltage stress.

**Usage**

```r
data("insulation")
```

**Format**

A data frame with 76 observations on the following 2 variables.

- `log.stress` log of voltage stress
- `log.time` log of failure time

**References**


**Examples**

```r
myscores <- logGFscores
myscores@param <- c(1,5)
fit <- rfit(log.time ~ log.stress,scores=myscores,data=insulation)
summary(fit)
fit$tauhat
```

---

**Internal**

Internal functions not intended for general use.

**Description**

Internal functions not intended for general use.

**Usage**

```r
lmean(z, p)
```

**Arguments**

- `z` n by 1 vector
- `p` scalar
**Author(s)**
John Kloke, Joseph McKean

---

**Jonckheere’s Test for Ordered Alternatives**

**Description**
Computes Jonckheere’s Test for Ordered Alternatives; see Section 5.6 of Kloke and McKean (2014).

**Usage**

\[ \text{jonckheere}(y, \text{groups}) \]

**Arguments**
- \( y \) vector of responses
- \( \text{groups} \) vector of associated groups (levels)

**Details**
Computes Jonckheere’s Test for Ordered Alternatives. The main source was downloaded from the site:

smtp.biostat.wustl.edu/sympa/biostat/arc/s-news/2000-10/msg00126.html

**Value**
- \( \text{Jonckheere} \) test statistic
- \( \text{ExpJ} \) null expectation
- \( \text{VarJ} \) null variance
- \( p \) p-value

**Author(s)**
John Kloke <kloke@biostat.wisc.edu>, Joseph McKean

**References**

smtp.biostat.wustl.edu/sympa/biostat/arc/s-news/2000-10/msg00126.html

**Examples**

\[ r \leftarrow \text{rnorm}(30) \]
\[ \text{gp} \leftarrow \text{c(rep}(1,10), \text{rep}(2,10), \text{rep}(3,10)) \]
\[ \text{jonckheere}(r, \text{gp}) \]
Robust Analysis of Covariance under Heterogeneous Slopes for a k-way layout

Description

Returns a robust rank-based analysis of covariance for a k-way layout assuming heterogenous slopes; see Section 5.4 of Kloke and McKean (2014). Currently only wilcoxon scores are used.

Usage

kancova(levs, data, xcov, print.table=TRUE)

Arguments

levs vector of levels corresponding to the factors A, B, C, etc.
data matrix with response in column 1 and level in column 2
xcov matrix of covariates
print.table logical indicating a table should be printed

Details

Returns the analysis of covariance table assuming heterogenous slopes for a k-way layout.

Value

tab2 analysis of covariance
fint rank-based ful model (heterogenous slopes
fithomog rank-based ful model (homogeneous slopes

Author(s)

John Kloke <kloke@biostat.wisc.edu>, Joseph McKean

References


Examples

levels <- c(2,2)
y.group <- huiitema496[,c('y','i','j')]
xcov <- huiitema496[,c('x')]
kancova(levels,y.group,xcov)
**kancovarown**  
*routine used in the ANCOVA table obtained by kancova*  

**Description**  
routine used in making the display of the ANCOVA table obtained by kancova.

**Usage**  
```
kancovarown(vec)
```

**Arguments**  
vec vector to be labeled.

**Details**  
Returns the labels.

**Value**  

```
nm vector of labels
```

**References**  

---

**latour**  
*Chateau Latour Wine Data*  

**Description**  
The response variable is the quality of a vintage based on a scale of 1 to 5 over the years 1961 to 2004. The predictor is end of harvest, days between August 31st and the end of harvest for that year, and the factor of interest is whether or not it rained at harvest time.

**Usage**  
```
data(latour)
```
onecova

Format

A data frame with 44 rows and 4 columns.

year Year of harvest
quality Rating on a scale of 1-5
end.of.harvest Days August 31 and the end of harvest
rain indicator variable for rain

References


Examples

data(latour)
plot(quality~end.of.harvest,pch='.',data=latour)
points(quality~end.of.harvest,data=latour[latour$rain==0,],pch=3)
points(quality~end.of.harvest,data=latour[latour$rain==1,],pch=4)

onecova Robust Analysis of Covariance under Heterogeneous Slopes

Description

Returns tests for homogeneous slopes and also assuming homogeneous slopes a test for differences in level. Currently only wilcoxon scores are used.

Usage

onecova(levs,data,xcov,print.table=TRUE)

Arguments

levs Number of levels of the one-way design
data matrix with response in column 1 and level in column 2
xcov matrix of covariates
print.table logical indicating a table should be printed

Details

Returns the analysis of covariance table.

Value

tab analysis of covariance
References


Examples

data=latour[,c('quality','rain')]
xcov<-cbind(latour['end.of.harvest'])
onecova(2,data,xcov,print.table=TRUE)

---

onecovaheter *Robust Analysis of Covariance under Heterogeneous Slopes*

Description

Returns a robust rank-based analysis of covariance for a one-way layout assuming heterogenous slopes; see Section 5.4 of Kloke and McKean (2014). Currently only wilcoxon scores are used.

Usage

onecovaheter(levs,data,xcov,print.table=TRUE)

Arguments

levs Number of levels of the one-way design
data matrix with response in column 1 and level in column 2
cov matrix of covariates
print.table logical indicating a table should be printed

details

Returns the analysis of covariance table assuming heterogenous slopes.

Value

tab analysis of covariance
fit rank-based ful model (heterogenous slopes

References


Examples

data=latour[,c('quality','rain')]
xcov<-cbind(latour['end.of.harvest'])
onecovaheter(2,data,xcov,print.table=TRUE)
Robust Analysis of Covariance under Heterogeneous Slopes

Description

Returns a robust rank-based analysis of covariance for a one-way layout assuming homogeneous slopes; see Section 5.4 of Kloke and McKean (2014). Currently only wilcoxon scores are used.

Usage

onecovahomog(levs, data, xcov, print.table=TRUE)

Arguments

levs Number of levels of the one-way design
data matrix with response in column 1 and level in column 2
xcov matrix of covariates
print.table logical indicating a table should be printed

Details

Returns the analysis of covariance table assuming homogeneous slopes.

Value

tab analysis of covariance
fit rank-based ful model (homogeneous slopes)

References


Examples

data=latour[,c('quality','rain')]
xcov<-cbind(latour['end.of.harvest'])
onecovahomog(2, data, xcov, print.table=TRUE)
place

Placements.

Description

Returns the placements of the first vector in terms of the second vector used the R function fp.test; see Section 2.11 of Hettmansperger and McKean (2011) and Section 4.4 of Hollander and Wolfe (1999). The version computed by fp.test is discussed in Section 4.4 of Kloke and McKea

Usage

place(x,y)

Arguments

x  
first vector

y  
second vector of second sample responses

Details

Returns the Placements for the routine fp.test.

Value

ic  
vector of placements.

Author(s)

John Kloke <kloke@biostat.wisc.edu>, Joseph McKean

References


plank

Plank data

Description

Ask Joe

Usage

data(plank)

Format

A data frame with 64 observations on the following 4 variables.

response a numeric vector
strain a factor with levels 1 2
gender a factor with levels 1 2
age a factor with levels 1 2 3

References


Examples

data(plank)
boxplot(response~strain,data=plank)

poly

A Simulated Polynomial Data Set.

Description

A simulated polynomial (3rd degree) model discussed in Section 4.7.1 of Kloke and McKean (2014).

Usage

data(poly)
Format

One-hundred observations on two variables.

y response variable
x predictor

References


Examples

plot(y ~ x, data=poly)

---

polydeg  

Description

Tests for the degree of a polynomial. This test was suggested by Graybill (1976) and is discussed from a robust point-of-view in Section 4.7.1. of Kloke and McKean (2014).

Usage

polydeg(y, x, P, alpha = 0.05)

Arguments

y vector of responses
x Predictor
P Super degree of polynomial which provides a satisfactory fit
alpha Level of the testing

details

Returns the degree of the polynomial based on the algorithm.

Value

deg The determined degree
coll Matrix of step information
fitf Fit of the polynomial based on the determined degree
References


Examples

```r
x <- 1:20
xc <- x - mean(x)
y <- .2*xc + xc^3 + rt(20, 3)*90
plot(y~x)
polydeg(y, xc, 6)
```

print

*Internal print functions*

Description

Internal print functions

Usage

```r
## S3 method for class 'hogg.test'
print(x, digits = max(5, .Options$digits - 2), ...)
## S3 method for class 'rank.test'
print(x,...)
## S3 method for class 'fkk.test'
print(x,...)
```

Arguments

- `x` Object to be printed.
- `digits` Number of digits to present. Passed to print function.
- `...` Additional arguments.

Author(s)

John Kloke, Joseph McKean
**prostate**

*DES for treatment of prostate cancer.*

---

**Description**

Under investigation in this clinical trial was the pharmaceutical agent diethylstilbestrol DES; subjects were assigned treatment to 1.0 mg DES (treatment = 2) or to placebo (treatment = 1).

**Usage**

data(prostate)

**Format**

A data frame with 38 observations on the following 8 variables.

- patient: a numeric vector
- treatment: a numeric vector
- time: a numeric vector
- status: a numeric vector
- age: a numeric vector
- shb: a numeric vector
- size: a numeric vector
- index: a numeric vector

**Source**

http://www.crcpress.com/product/isbn/9781584883258

**References**


**Examples**

data(prostate)

boxplot(size~treatment, data=prostate)
Description

A regression example with response yearly upkeep of a home and the predictor value of home; see Bowerman et al. (2005) and Exercise 4.9.8 of Kloke and McKean (2014).

Usage

data(qhic)

Format

Forty observations on two variables.

upkeep annual upkeep expenditure of home (y)
value value of the home (x)

References


Examples

plot(upkeep~value,data=qhic,xlab='Value (in $1000s)',ylab='Annual upkeep (in $10s)')

Description

Quail from a two-factor experiment.

Usage

data(quail2)

Format

A data frame with 30 observations on the following 2 variables.

treat indicator variable for treatment
ldl ldl measurement
rank.test

References
New York: Chapman-Hall.

Examples
data(quail2)
boxplot(ldl~treat,data=quail2)

---

rank.test General scores rank test for two sample problem

Description
A generalization of the Wilcoxon rank-sum test where a score function is applied to the ranks. Any
scores from Rfit can be used as well as user defined. Default is to perform a Wilcoxon analysis.

Usage
rank.test(x, y, alternative = "two.sided", scores = Rfit::wscores,
          conf.int = FALSE, conf.level = 0.95)

Arguments
x m x 1 vector
y n x 1 vector
alternative one of 'two.sided', 'less', or 'greater'
scores an object of class scores
conf.int logical indicating if a confidence interval should be estimated
conf.level desired level of confidence for interval

Details
Test is based on $T = \sum_{i} a(R(y_{i}))$ where $R$ is the rank based on the combined sample and $a(t) = \varphi(t/(N+1))$. Confidence interval, if requested, is based on call to Rfit.

Value
statistic Standardized value of test statistics
Sphi test statistic
p.value p-value
conf.int confidence interval for shift in location
estimate point estimate for shift in location
Description

Generate a random sample from a contaminated normal distribution.

Usage

`rcn(n, eps, sigmac)`

Arguments

- `n` sample size
- `eps` proportion of proportion of contamination
- `sigmac` standard deviation of contaminated component

Details

With probability (1-eps) a deviates are drawn from a standard normal distribution. With probability eps deviates are drawn from a normal distribution with mean 0 and standard deviation sigmac

Value

n x 1 numeric vector containing the random deviates.

Author(s)

John Kloe <kloke@biostat.wisc.edu>, Joseph McKean
References


See Also

rnorm

Examples

qqnorm(rgen(100,.25,10))

---

rs Simulated Regression Model

Description

A simulated regression model with one response and one predictor. It is discussed in Exercise 6.5.6 of Kloke and McKean (2014).

Usage

data(rs)

Format

Fifty observations on two variables.

y simulated response
x simulated predictor

References


Examples

rfit(y ~ x, data=rs)
**SCUD**  
* Cyclone Data

**Description**

A data set discussed in Hollander and Wolfe (1999) and Exercise 5.8.9 of Kloke and McKean (2014). It contains part of a study on the effects of cloud seeding of cyclones.

**Usage**

```r
data(SCUD)
```

**Format**

Twenty-one observations on three variables.

- **trt**: treatment indicator (1) is Seeded and (2) is control
- **M**: predictor M, the geostrophic meridional circulation index
- **RI**: measure of precipitation

**References**


**Examples**

```r
plot(RI ~ M, data=SCUD)
```

---

**sievers**

* Doksum and Sievers rat data

**Description**

Doksum and Sievers (1976) describe an experiment involving the effect of ozone on weight gain of rats. The experimental group consisted of 22 rats which were placed in an ozone environment for seven days, while the control group contained 21 rats which were placed in an ozone-free environment for the same amount of time. The response was the weight gain in a rat over the time period.

**Usage**

```r
data(sievers)
```
Format

A data frame with 45 observations on the following 2 variables.

group  indicator variable for treatment
weight.gain  response variable of weight gain

References


Examples

data(sievers)
boxplot(weight.gain~group,data=sievers)

---

simon  

*Simon (the memory game) dataset*

Description

A experiment in which the members of two groups of students each played the game Simon twice.

Usage

data("simon")

Format

A data frame with 31 observations on the following 3 variables.

game1  score on first trial
game2  score on second trial
class  group variable

Details

Demonstrates the concept of regression toward the mean. Simulated data to represent a realistic realization of the experiment. See problem 4.9.20 of Kloke and McKeane (2014)

References

sincos

**Examples**

data(simon)
plot(game2~game1, data=simon)
rfit(game2~game1, data=simon)

---

**sincos**

* Sine Cosine Model

**Description**

Simulated dataset

**Usage**

data("sincos")

**Format**

A data frame with 197 observations on the following 2 variables.

- x independent variable
- y dependent variable

**Details**

The data were generated using

```r
x <- seq(1, 50, by=.25) ; y <- 5*sin(3*x) + 6*cos(x/4) + rnorm(length(x), 0, 10)
```

**Examples**

data(sincos)
plot(y~x,sincos)

---

**speed**

* Predict top speed based on miles per gallon

**Description**

A sample of 82 cars with variables speed and miles per gallon collected.

**Usage**

data("speed")
vanElteren.test

Format

A data frame with 82 observations on the following 2 variables.

- mpg  Miles per gallon
- sp   a numeric vector

Source


References


Examples

```r
data(speed)
plot(sp~mpg, data=speed)
rfit(sp~mpg+I(mpg^2), data=speed)
```

Description

Performs the vanElteren extension of the Wilcoxon rank sum test for stratified experiments.

Usage

```r
vanElteren.test(g, y, b)
```

Arguments

- **g**  n x 1 vector: treatment/group indicator
- **y**  n x 1 vector: responses
- **b**  n x 1 vector: denotes strata

Value

- **statistic**  Value of the test statistic.
- **p.value**  p-value based on a normal approximation.
Description

January weather data for Kalamazoo, MI for the years 1900 to 1995. It is discussed in Example 4.7.4, page 106, of Kloke and McKean (2014).

Usage

data(weather)

Format

Ninety-six observations (1900-1995) for twelve weather variables.

avemax avemax
avemin avemin
coldestmax coldestmax
hihest hihest
lowest lowest
maxdayprec maxdayprec
maxdaysnowfall maxdaysnowfall
meantmp meantmp
totprec totalprec
totalsnow totalsnow
warmest warmest
year year

Source

http://weather-warehouse.com/WeatherHistory/

References


Examples

plot(avemax ~ year,data=weather)
Index

*Topic **datasets, survival**
  - hodgkins, 16
*Topic **datasets**
  - bb2010, 3
  - blood.plasma, 4
  - cancertrt, 5
  - cloud, 6
  - energy, 8
  - firstbase, 8
  - hemorrhage, 15
  - huitema496, 18
  - insulation, 19
  - latour, 22
  - plank, 27
  - poly, 27
  - prostate, 30
  - qhic, 31
  - quai2, 31
  - rs, 34
  - SCUD, 35
  - sievers, 35
  - simon, 36
  - sincos, 37
  - speed, 37
  - weather, 39
*Topic **nonparametrics**
  - aligned.test, 2
  - gehan.test, 13
  - hogg.test, 16
  - HoggQs, 17
  - rank.test, 32
  - vanElteren.test, 38
*Topic **survival**
  - gehan.test, 13

aligned.test, 2

bb2010, 3
blood.plasma, 4
cancertrt, 5
centerx, 5
cor.boot.ci, 7
energy, 8
firstbase, 8
fk.test, 9
fkk.test, 10
fp.test, 11
gehan.test, 13
gexact, 14
gexact2, 14
hemorrhage, 15
hodgkins, 16
hogg.test, 16, 18
HoggQs, 17
huitema496, 18
insulation, 19
internal, 19
jonckheere, 20
kancova, 21
kancovarown, 22
latour, 22
lmean (internal), 19
onecova, 23
onecovaheter, 24
onecovahomog, 25
place, 26
plank, 27
poly, 27
INDEX

polydeg, 28
print, 29
prostate, 30

Q1 (HoggQs), 17
Q2 (HoggQs), 17
qhic, 31
quail2, 31
rank.test, 32
rcn, 33
rnorm, 34
rs, 34

SCUD, 35
sievers, 35
simon, 36
sincos, 37
speed, 37

umean (internal), 19

vanElteren.test, 38

weather, 39
wilcox.test, 33