Package 'NADA'

March 22, 2020

Title Nondetects and Data Analysis for Environmental Data

Version 1.6-1.1

Author Lopaka Lee

Description Contains methods described by Dennis Helsel in
his book ``Nondetects And Data Analysis: Statistics
for Censored Environmental Data''.

Maintainer Lopaka Lee <rclee@usgs.gov>

Depends R (>= 2.10), methods, survival

License GPL (>= 2)

Collate All.R summary.R cen.R km.R reg.R mle.R ken.R ros.R icmle.R

NeedsCompilation no

Repository CRAN

Date/Publication 2020-03-22 09:46:09 UTC

R topics documented:

Ag	*Dissolved silver concentrations from water analyses*

Description

This dataset was used by Helsel and Cohn (1988) to verify their software. It is provided for code validation purposes.

Usage

```
data(Silver)
```

Format

A list containing 56 observations with items 'obs' and 'censored'. 'obs' is a numeric vector of all observations (both censored and uncensored). 'censored' is a logical vector indicating where an element of 'obs' is censored (a less-than value).

Source

Helsel and Cohn (1988)

References

Dennis R. Helsel and Timothy A. Cohn (1988), Estimation of descriptive statistics for multiply censored water quality data, Water Resources Research vol. 24, no. 12, pp.1997-2004

Arsenic   *Dissolved arsenic concentrations in ground water of U.S.*

Description

This dataset is a random selection of dissolved arsenic analyses taken during the U.S. Geological Survey's National Water Quality Assessment program (NAWQA).

Usage

```
data(Arsenic)
```

Format

A list containing 50 observations with items 'As', 'AsCen', 'Aquifer'. 'As' is a numeric vector of all arsenic observations (both censored and uncensored). 'AsCen' is a logical vector indicating where an element of 'As' is censored (a less-than value). 'Aquifer' is a grouping factor of hypothetical hydrologic sources for the data.

Source

U.S. Geological Survey National Water Quality Assessment Data Warehouse

References

The USGS NAWQA site at http://water.usgs.gov/nawqa

As   *Dissolved arsenic concentrations in ground water of U.S.*

Description

This dataset is a random selection of dissolved arsenic analyses taken during the U.S. Geological Survey's National Water Quality Assessment program (NAWQA).

Usage

```
data(NADA.As)
```

Format

A list containing 50 observations with items 'obs' and 'censored'. 'obs' is a numeric vector of all arsenic observations (both censored and uncensored). 'censored' is a logical vector indicating where an element of 'obs' is censored (a less-than value).

Source

U.S. Geological Survey National Water Quality Assessment Data Warehouse

References

The USGS NAWQA site at <http://water.usgs.gov/nawqa>

AsExample   *Example arsenic concentrations in drinking water*

Description

Artificial numbers representing arsenic concentrations in a drinking water supply.

Objective is to determine what can be done with data where all values are below the reporting limit. There is a detection limit at 1, and a reporting limit at 3 ug/L. Used in Chapter 8 of the NADA book

Usage

```
data(AsExample)
```

Source

None. Generated.

References

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

---

asSurv-methods   *Methods for function asSurv in Package NADA*

Description

Methods for function asSurv in package **NADA**.

asSurv converts a Cen object to a Surv object.

Usage

```
## S4 method for signature 'Cen'
asSurv(x)

## S4 method for signature 'formula'
asSurv(x)
```

Arguments

x   A Cen or formula object.

Atra	*Atrazine concentrations in Nebraska ground water*

Description

Atrazine concentrations in a series of Nebraska wells before (June) and after (September) the growing season.

Objective is to determine if concentrations increase from June to September. There is one detection limit, at 0.01 ug/L. Used in Chapters 4, 5, and 9 of the NADA book.

Usage

```
data(Atra)
```

Source

Junk et al., 1980, Journal of Environmental Quality 9, pp. 479-483.

References

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

AtraAlt	*Atrazine concentrations in Nebraska ground water – Alternative*

Description

Alternative Atrazine concentrations altered from the Atra data set so that there are more nondetects, adding a second detection limit at 0.05.

Objective is to determine if concentrations increase from June to September. There are two detection limits, at 0.01 and 0.05 ug/L. Used in Chapters 5 and 9 of the NADA book.

Usage

```
data(AtraAlt)
```

Source

Altered from the data of Junk et al., 1980, Journal of Environmental Quality 9, pp. 479-483.

References

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

| Atrazine | *Atrazine concentrations in Nebraska ground water – Another Format* |

Description

The same atrazine concentrations as in Atra, stacked into one column (col.1). Column 2 indicates the month of collection. Column 3 indicates which data are below the detection limit those with a value of 1.

Objective is to determine if concentrations increase from June to September. There is one detection limit, at 0.01 ug/L. Used in Chapter 9 of the NADA book.

Usage

```
data(Atrazine)
```

Source

Junk et al., 1980, Journal of Environmental Quality 9, pp. 479-483.

References

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

| Bloodlead | *Lead concentrations in the blood of herons in Virginia.* |

Description

Blood-lead concentrations in herons of Virginia. Objective is to compute interval estimates for lead concentrations. There is one detection limit, at 0.02 ug/g. Used in Chapter 7 of the NADA book.

Usage

```
data(Bloodlead)
```

Source

Golden et al., 2003, Environmental Toxicology and Chemistry 22, 1517-1524.

References

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

---

boxplot-methods   *Methods for function boxplot in Package NADA*

Description

Methods for box plotting objects in package **NADA**

Usage

```
    ## S4 method for signature 'ros'
boxplot(x, ...)
```

Arguments

x   An output object from a NADA function such as [ros](#).

...   Additional arguments passed to the generic boxplot method.

See Also

[boxplot](#)

Cadmium   *Cadmium concentrations in fish*

Description

Cadmium concentrations in fish for two regions of the Rocky Mountains.

Objective is to determine if concentrations are the same or different in fish livers of the two regions. There are four detection limits, at 0.2, 0.3, 0.4, and 0.6 ug/L. Used in Chapter 9 of the NADA book.

Usage

```
data(Cadmium)
```

Source

none. Data modeled after several reports.

References

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

Cen   *Create a Censored Object*

Description

Create a censored object, usually used as a response variable in a model formula.

Usage

```
Cen(obs, censored, type = "left")
```

Arguments

obs	A numeric vector of observations. This includes both censored and uncensored observations.
censored	A logical vector indicating TRUE where an observation in obs is censored (a less-than value) and FALSE otherwise.
type	character string specifying the type of censoring. Possible values are "right", "left", "counting", "interval", or "interval2". The default is "left".

Value

An object of class Cen.

details

This, and related routines, are front ends to routines in the survival package. Since the survival routines can not handle left-censored data, these routines transparently handle "flipping" input data and resultant calculations. The Cen function provides part of the necessary framework for flipping.

Author(s)

R. Lopaka Lee <rclee@usgs.gov>

Dennis Helsel <dhelsel@practicalstats.com>

References

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

See Also

cenfit, flip-methods

Examples

```
obs      = c(0.5,    0.5,   1.0,  1.5,   5.0,    10,    100)
censored = c(TRUE, FALSE, FALSE, TRUE, FALSE, FALSE, FALSE)

Cen(obs, censored)
flip(Cen(obs, censored))
```

---

cenboxplot   *Produces a censored boxplot*

Description

Draws a boxplot with the highest censoring threshold shown as a horizontal line. Any statistics below this line are invalid are must be estimated using methods for censored data.

Usage

```
cenboxplot(obs, cen, group, log=TRUE, range=0, ...)
```

Arguments

obs	A numeric vector of observations.
cen	A logical vector indicating TRUE where an observation in x is censored (a less-than value) and FALSE otherwise.
group	A factor vector used for grouping 'obs' into subsets (each group will be a separate box).
log	A TRUE/FALSE indicating if the y axis should be in log units. Default it TRUE.
range	This determines how far the plot whiskers extend out from the box. If 'range' is positive, the whiskers extend to the most extreme data point which is no more than 'range' times the interquartile range from the box. The default is zero which causes the whiskers to extend to the min and max data values.
...	Additional items that get passed to boxplot.

Value

Returns the output of the default boxplot method.

Author(s)

R. Lopaka Lee <rclee@usgs.gov>

Dennis Helsel <dhelsel@practicalstats.com>

References

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

Examples

```
data(Golden)
with(Golden, cenboxplot(Blood, BloodCen, DosageGroup))
```

---

cendiff   *Test Censored ECDF Differences*

Description

Tests if there is a difference between two or more empirical cumulative distribution functions (ECDF) using the G^ρ family of tests, or for a single curve against a known alternative.

Usage

```
cendiff(obs, censored, groups, ...)
```

Arguments

obs	Either a numeric vector of observations or a formula. See examples below.
censored	A logical vector indicating TRUE where an observation in 'obs' is censored (a less-than value) and FALSE otherwise.
groups	A factor vector used for grouping 'obs' into subsets.
...	Additional items that are common to this function and the survdiff function from the 'survival' package. See Details.

Details

This, and related routines, are front ends to routines in the survival package. Since the survival routines can not handle left-censored data, these routines transparently handle "flipping" input data and resultant calculations.

This function shares the same arguments as survdiff. The most important of which is rho which controls the type of test. With rho = 0 this is the log-rank or Mantel-Haenszel test, and with rho = 1 it is equivalent to the Peto & Peto modification of the Gehan-Wilcoxon test. The default is rho = 1, or the Peto & Peto test. This is the most appropriate for left-censored log-normal data.

For the formula interface: if the right hand side of the formula consists only of an offset term, then a one sample test is done. To cause missing values in the predictors to be treated as a separate group, rather than being omitted, use the factor function with its exclude argument.

Value

Returns a list with the following components:

n	the number of subjects in each group.
obs	the weighted observed number of events in each group. If there are strata, this will be a matrix with one column per stratum.

exp	the weighted expected number of events in each group. If there are strata, this will be a matrix with one column per stratum.
chisq	the chisquare statistic for a test of equality.
var	the variance matrix of the test.
strata	optionally, the number of subjects contained in each stratum.

Author(s)

R. Lopaka Lee <rclee@usgs.gov>

Dennis Helsel <dhelsel@practicalstats.com>

References

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

Harrington, D. P. and Fleming, T. R. (1982). A class of rank test procedures for censored survival data. *Biometrika* **69**, 553-566.

See Also

Cen,

survdiff

Examples

```
data(Cadmium)

obs      = Cadmium$Cd
censored = Cadmium$CdCen
groups   = Cadmium$Region

# Cd differences between regions?
cendiff(obs, censored, groups)

# Same as above using formula interface
cenfit(Cen(obs, censored)~groups)
```

---

cendiff-methods   *Methods for function cendiff in Package NADA*

Description

See cendiff for all the details.

---

cenfit *Compute an ECDF for Censored Data*

Description

Computes an estimate of an empirical cumulative distribution function (ECDF) for censored data using the Kaplan-Meier method.

Usage

```
cenfit(obs, censored, groups, ...)
```

Arguments

obs   Either a numeric vector of observations or a formula. See examples below.

censored   A logical vector indicating TRUE where an observation in 'obs' is censored (a less-than value) and FALSE otherwise.

groups   A factor vector used for grouping 'obs' into subsets.

...   Additional items that are common to this function and the survfit function from the 'survival' package. See Details.

Details

This, and related routines, are front ends to routines in the survival package. Since the survival routines can not handle left-censored data, these routines transparently handle "flipping" input data and resultant calculations. Additionally provided are query and prediction methods for cenfit objects.

There are many additional options that are supported and documented in survfit. Only a few have application to the geosciences. However, the most important is 'conf.int'. This is the level for a two-sided confidence interval on the ECDF. The default is 0.95.

If you are using the formula interface: The censored and groups parameters are not specified – all information is provided via a formula as the obs parameter. The formula must have a Cen object as the response on the left of the ~ operator and, if desired, terms separated by + operators on the right.

Value

a cenfit object. Methods defined for cenfit objects are provided for print, plot, lines, predict, mean, median, sd, quantile.

If the input formula contained factoring groups (ie., cenfit(obs,censored,groups), individual ECDFs can be obtained by indexing (eg., model[1], etc.).

Author(s)

R. Lopaka Lee <rclee@usgs.gov>

Dennis Helsel <dhelsel@practicalstats.com>

References

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

Dorey, F. J. and Korn, E. L. (1987). Effective sample sizes for confidence intervals for survival probabilities. *Statistics in Medicine* 6, 679-87.

Fleming, T. H. and Harrington, D.P. (1984). Nonparametric estimation of the survival distribution in censored data. *Comm. in Statistics* 13, 2469-86.

Kalbfleisch, J. D. and Prentice, R. L. (1980). *The Statistical Analysis of Failure Time Data.* Wiley, New York.

Link, C. L. (1984). Confidence intervals for the survival function using Cox's proportional hazards model with covariates. *Biometrics* 40, 601-610.

See Also

survfit, Cen, plot-methods, mean-methods, sd-methods, median-methods, quantile-methods, predict-methods, lines-methods, summary-methods, cendiff

Examples

```
# Create a Kaplan-Meier ECDF, plot and summarize it.

data(Cadmium)

obs      = Cadmium$Cd
censored = Cadmium$CdCen

mycenfit = cenfit(obs, censored)

plot(mycenfit)
summary(mycenfit)
quantile(mycenfit, conf.int=TRUE)
median(mycenfit)
mean(mycenfit)
sd(mycenfit)
predict(mycenfit, c(10, 20, 100), conf.int=TRUE)

# With groups
groups = Cadmium$Region

cenfit(obs, censored, groups)

# Formula interface -- no groups
cenfit(Cen(obs, censored))

# Formula interface -- with groups
cenfit(Cen(obs, censored)~groups)
```

cenfit-class   *Class "cenfit"*

Description

A cenfit object is returned from the NADA cenfit function.

Slots

survfit: Object of class survfit returned from the survfit function.

Methods

[signature(x = "cenfit", i = "numeric", j = "missing"): ...

mean signature(x = "cenfit"): ...

median signature(x = "cenfit"): ...

plot signature(x = "cenfit", y = "ANY"): ...

predict signature(object = "cenfit"): ...

print signature(x = "cenfit"): ...

quantile signature(x = "cenfit"): ...

sd signature(x = "cenfit"): ...

summary signature(object = "cenfit"): ...

Author(s)

R. Lopaka Lee <rclee@usgs.gov>

Dennis Helsel <dhelsel@practicalstats.com>

See Also

cenfit

Examples

```
obs     = c(0.5,    0.5,   1.0,  1.5,  5.0,    10,   100)
censored = c(TRUE, FALSE, FALSE, TRUE, FALSE, FALSE, FALSE)

class(cenfit(Cen(obs, censored)))
```

---

cenfit-methods   *Methods for function cenfit in Package NADA*

Description

See cenfit for all the details.

Examples

```
data(Atrazine)

cenfit(Atrazine$Atra, Atrazine$AtraCen)
cenfit(Atrazine$Atra, Atrazine$AtraCen, Atrazine$Month)

cenfit(Cen(Atrazine$Atra, Atrazine$AtraCen))
cenfit(Cen(Atrazine$Atra, Atrazine$AtraCen)~Atrazine$Month)
```

---

cenken   *Compute Kendall's tau correlation coefficient and associated line for censored data. Computes the Akritas-Theil-Sen nonparametric line, with the Turnbull estimate of intercept.*

Description

Computes Kendall's tau for singly (y only) or doubly (x and y) censored data. Computes the Akritas-Theil-Sen nonparametric line, with the Turnbull estimate of intercept.

Usage

```
cenken(y, ycen, x, xcen)
```

Arguments

y	A numeric vector of observations or a formula.
ycen	A logical vector indicating TRUE where an observation in x is censored (a less-than value) and FALSE otherwise. Can be missing/omitted for the case where x is not censored.
x	A numeric vector of observations.
xcen	A logical vector indicating TRUE where an observation in y is censored (a less-than value) and FALSE otherwise.

Details

If you are using the formula interface: The ycen, x and xcen parameters are not specified – all information is provided via a formula as the y parameter. The formula must have a Cen object as the response on the left of the ~ operator and, if desired, terms separated by + operators on the right. See example below.

Kendall's tau is a nonparametric correlation coefficient measuring the monotonic association between y and x. For left-censored data, concordant and discordant directions between x and y are measured whenever possible. So with increasing x values, a change in y from <1 to 10 is an increase (concordant). A change from a <1 to a detected 0.5 is considered a tie, as is a <1 to a <5, because neither can definitively be called an increase or decrease. Tie corrections are employed for the variance of the test statistic in order to account for the many ties when computing p-values. The ATS line is the slope that results in a Kendalls tau of 0 for correlation between the residuals, y-slope*x and x. The cenken routine performs an iterative bisection search to find that slope. The intercept is the median residual, where the median for censored data is computed using the Turnbull estimate for interval censored data, as implmented in the Icens contributed package for R.

Value

Returns tau (Kendall's tau), slope, and p-value for the regression.

References

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

Akritas, M.G., S. A. Murphy, and M. P. LaValley (1995). The Theil-Sen Estimator With Doubly Censored Data and Applications to Astronomy. Journ. Amer. Statistical Assoc. 90, p. 170-177.

Examples

```
# Both y and x are censored
# (exercise 11-1 on pg 198 of the NADA book)
data(Golden)
with(Golden, cenken(Blood, BloodCen, Kidney, KidneyCen))

## Not run:
# x is not censored
# (example on pg 213 of the NADA book)
data(TCEReg)
with(TCEReg, cenken(log(TCEConc), TCECen, PopDensity))
# formula interface
with(TCEReg, cenken(Cen(log(TCEConc), TCECen)~PopDensity))

# Plotting data and the regression line
data(DFe)
# Recall x and y parameter positons are swapped in plot vs regression calls
with(DFe, cenxyplot(Year, YearCen, Summer, SummerCen))    # x vs. y
reg = with(DFe, cenken(Summer, SummerCen, Year, YearCen)) # y~x
lines(reg)

## End(Not run)
```

---

cenken-class   *Class "cenken"*

Description

A "cenken" object is returned from cenken. It extends the 'list' class.

Objects from the Class

Objects can be created by calls of the form cenken(y,ycen,x,xcen).

Slots

.Data: Object of class "list"

Extends

Class "list", from data part.

Methods

lines signature(x = "cenken"): ...

Author(s)

R. Lopaka Lee <rclee@usgs.gov>

Dennis Helsel <dhelsel@practicalstats.com>

See Also

cenken

---

cenken-methods   *Methods for function cenken in Package NADA*

Description

See cenken for all the details.

| cenmle | *Regression by Maximum Likelihood Estimation for Left-censored Data* |

Description

Regression by Maximum Likelihood (ML) Estimation for left-censored ("nondetect" or "less-than") data. This routine computes regression estimates of slope(s) and intercept by maximum likelihood when data are left-censored. It will compute ML estimates of descriptive statistics when explanatory variables following the ~ are left blank. It will compute ML tests similar in function and assumptions to two-sample t-tests and analysis of variance when groups are specified following the ~. It will compute regression equations, including multiple regression, when continuous explanatory variables are included following the ~. It will compute the ML equivalent of analysis of covariance when both group and continuous explanatory variables are specified following the ~. To avoid an appreciable loss of power with regression and group hypothesis tests, a probability plot of residuals should be checked to ensure that residuals from the regression model are approximately gaussian.

Usage

```
cenmle(obs, censored, groups, ...)
```

Arguments

obs	Either a numeric vector of observations or a formula. See examples below.
censored	A logical vector indicating TRUE where an observation in 'obs' is censored (a less-than value) and FALSE otherwise.
groups	A factor vector used for grouping 'obs' into subsets.
...	Additional items that are common to this function and the survreg function from the 'survival' package. The most important of which is 'dist' and 'conf.int'. See Details below.

Details

This routine is a front end to the survreg routine in the survival package.

There are many additional options that are supported and documented in survfit. Only a few have relevance to the evironmental sciences.

A very important option is 'dist' which specifies the distributional model to use in the regression. The default is 'lognormal'.

Another important option is 'conf.int'. This is NOT an option to survreg but is an added feature (due to some arcane details of R it can't be documented above). The 'conf.int' option specifies the level for a two-sided confidence interval on the regression. The default is 0.95. This interval will be used in when the output object is passed to other generic functions such as mean and quantile. See Examples below.

Also supported is a 'gaussian' or a normal distribution. The use of a gaussian distribution requires an interval censoring context for left-censored data. Luckily, this routine automatically does this for you – simply specify 'gaussian' and the correct manipulations are done.

If any other distribution is specified besides lognormal or gaussian, the return object is a raw survreg object – it is up to the user to 'do the right thing' with the output (and input for that matter).

If you are using the formula interface: The censored and groups parameters are not specified – all information is provided via a formula as the obs parameter. The formula must have a Cen object as the response on the left of the ~ operator and, if desired, terms separated by + operators on the right. See Examples below.

Value

a cenmle object. Methods defined for cenmle objects are provided for mean, median, sd.

Author(s)

R. Lopaka Lee <rclee@usgs.gov>

Dennis Helsel <dhelsel@practicalstats.com>

References

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

See Also

[Cen](), [cenmle-methods](), [mean-methods](), [sd-methods](), [median-methods](), [quantile-methods](), [summary-methods]()

Examples

```
# Create a MLE regression object

data(TCEReg)

tcemle = with(TCEReg, cenmle(TCEConc, TCECen))

summary(tcemle)
median(tcemle)
mean(tcemle)
sd(tcemle)
quantile(tcemle)

# This time specifiy a different confidence interval
tcemle = with(TCEReg, cenmle(TCEConc, TCECen, conf.int=0.80))

# Use the model's confidence interval with the quantile function
quantile(tcemle, conf.int=TRUE)

# With groupings
with(TCEReg, cenmle(TCEConc, TCECen, PopDensity))
```

cenmle-class *Class "cenmle"*

Description

A "cenmle" object is returned from cenmle. It extends the 'cenreg' class returned from survreg.

Objects from the Class

Objects can be created by calls of the form cenmle(obs,censored).

Slots

survreg: Object of class "survreg"

Extends

Class "list", from data part. Class "vector", by class "list".

Methods

mean signature(x = "cenmle"): ...

median signature(x = "cenmle"): ...

sd signature(x = "cenmle"): ...

summary signature(object = "cenmle"): ...

Author(s)

R. Lopaka Lee <rclee@usgs.gov>

Dennis Helsel <dhelsel@practicalstats.com>

See Also

survreg

Examples

```
    x      = c(0.5,    0.5,   1.0,  1.5,  5.0,   10,   100)
    xcen   = c(TRUE, FALSE, FALSE, TRUE, FALSE, FALSE, FALSE)

    class(cenmle(x, xcen))
```

cenmle-gaussian-class   *Class "cenmle-gaussian"*

Description

A "cenmle-gaussian" object is returned from `cenmle` when a gaussian distribution is chosen with the 'dist' option.

Objects from the Class

Objects can be created by calls of the form `cenmle(obs,censored,dist="gaussian")`.

Slots

`n`: Total number of observations associated with the model

`n.cen`: Number of censored observations

`y`: Vector of observations

`ycen`: Censoring indicator

`conf.int`: Confidence interval associated with the model

`survreg`: Object of class `"survreg"`

Extends

Class `"cenmle"`

Methods

mean `signature(x = "cenmle")`: ...

median `signature(x = "cenmle")`: ...

sd `signature(x = "cenmle")`: ...

summary `signature(object = "cenmle")`: ...

Author(s)

R. Lopaka Lee <rclee@usgs.gov>

Dennis Helsel <dhelsel@practicalstats.com>

See Also

cenmle survreg

cenmle-lognormal-class

Class "cenmle-lognormal"

Description

A "cenmle-lognormal" object is returned from `cenmle` when a lognormal distribution is chosen with the 'dist' option.

Objects from the Class

Objects can be created by calls of the form `cenmle(obs,censored,dist="lognormal")`.

Slots

`n:` Total number of observations associated with the model

`n.cen:` Number of censored observations

`y:` Vector of observations

`ycen:` Censoring indicator

`conf.int:` Confidence interval associated with the model

`survreg:` Object of class `"survreg"`

Extends

Class `"cenmle"`

Methods

mean `signature(x = "cenmle")`: ...

median `signature(x = "cenmle")`: ...

sd `signature(x = "cenmle")`: ...

summary `signature(object = "cenmle")`: ...

Author(s)

R. Lopaka Lee <rclee@usgs.gov>

Dennis Helsel <dhelsel@practicalstats.com>

See Also

cenmle survreg

---

cenmle-methods   *Methods for function cenmle in Package NADA*

Description

See [cenmle](#) for all the details.

---

cenreg   *Compute regression equations and likelihood correlation coefficient
                                    for censored data.*

Description

Computes regression equations for singly censored data using maximum likelihood estimation. Estimates of slopes and intercept, tests for significance of parameters,and predicted quantiles (Median = points on the line) with confidence intervals can be computed.

Usage

```
cenreg(obs, censored, groups, ...)
```

Arguments

obs   Either a numeric vector of observations or a formula. See examples below.

censored   If a formula is not specified, this should be a logical vector indicating TRUE
                where an observation in obs is censored (a less-than value) and FALSE other-
                wise.

groups   If a formula is not specified, this should be a numeric or factor vector that rep-
                resents the explanatory variable.

...   Additional items that are common to this function and the survreg function
                from the 'survival' package. The most important of which is 'dist' and 'conf.int'.
                See Details below.

Details

This routine is a front end to the survreg routine in the survival package.

There are many additional options that are supported and documented in survfit. Only a few have relevance to the evironmental sciences.

A very important option is 'dist' which specifies the distributional model to use in the regression. The default is 'lognormal'.

Another important option is 'conf.int'. This is NOT an option to survreg but is an added feature (due to some arcane details of R it can't be documented above). The 'conf.int' option specifies the level for a two-sided confidence interval on the regression. The default is 0.95. This interval will

be used in when the output object is passed to other generic functions such as mean and quantile. See Examples below.

Also supported is a 'gaussian' or a normal distribution. The use of a gaussian distribution requires an interval censoring context for left-censored data. Luckily, this routine automatically does this for you – simply specify 'gaussian' and the correct manipulations are done.

If any other distribution is specified besides lognormal or gaussian, the return object is a raw survreg object – it is up to the user to 'do the right thing' with the output (and input for that matter).

If you are using the formula interface: The censored and groups parameters are not specified – all information is provided via a formula as the obs parameter. The formula must have a Cen object as the response on the left of the ~ operator and, if desired, terms separated by + operators on the right. See examples below.

The reported likelihood r correlation coefficient measures the linear association between y (groups) and x (obs), based on the difference in log likelihoods between the fitted model and the null model. Slopes and intercepts are fit by maximum likelihood. A lognormal distribution is fit by default, with a normal distribution being an option. Estimates of predicted values on the line can be obtained by specifying the values for all x variables at which y is to be predicted. Requesting the median (p=0.5) will provide estimates on the line for a lognormal distribution. Estimates of the mean are also possible, as are estimates of other percentiles. Equations for confidence intervals follow those of Meeker and Escobar (1098).

Value

Returns a summary.cenreg object.

Author(s)

R. Lopaka Lee <rclee@usgs.gov>

Dennis Helsel <dhelsel@practicalstats.com>

References

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

Meeker, W.Q. and L. A. Escobar (1998). Statistical Methods for Reliability Data. John Wiley and Sons, USA, NJ.

See Also

Cen, cenmle, predict-methods

Examples

```
# (examples in Chap 12 of the NADA book)
data(TCEReg)
```

```
# Using the formula interface
with(TCEReg, cenreg(Cen(TCEConc, TCECen)~PopDensity))

# Two or more explanatory variables requires the formula interface
tcemle2 = with(TCEReg, cenreg(Cen(TCEConc, TCECen)~PopDensity+Depth))

# Prediction of quantiles at PopDensity=5 and Depth=110
predict(tcemle2, c(5, 110))
```

---

cenreg-class   *Class "cenreg"*

Description

A "cenreg" object is returned from cenreg. It extends the 'cenreg' class returned from survreg.

Objects from the Class

Objects can be created by calls of the form cenreg(obs,censored,groups).

Slots

conf.int: Numeric value of confidence level (0.95)

n: Total number of samples

n.cen: Total censored samples

survreg: Object of class "survreg"

y: Total y samples

ycen: Total censored y samples

Extends

Class "list", from data part. Class "vector", by class "list".

Methods

predict signature(object = "cenreg"): ...

print signature(x = "cenreg"): ...

summary signature(object = "cenreg"): ...

Author(s)

R. Lopaka Lee <rclee@usgs.gov>

Dennis Helsel <dhelsel@practicalstats.com>

See Also

survreg

cenreg-gaussian-class *Class "cenreg-gaussian"*

Description

A "cenreg-gaussian" object is returned from `cenreg` when a gaussian distribution is chosen with the 'dist' option.

Objects from the Class

Objects can be created by calls of the form `cenreg(obs,censored,dist="gaussian")`.

Slots

`n:` Total number of observations associated with the model

`n.cen:` Number of censored observations

`y:` Vector of observations

`ycen:` Censoring indicator

`conf.int:` Confidence interval associated with the model

`survreg:` Object of class `"survreg"`

Extends

Class `"cenreg"`

Methods

predict `signature(object = "cenreg")`: ...

print `signature(x = "cenreg")`: ...

summary `signature(object = "cenreg")`: ...

Author(s)

R. Lopaka Lee <rclee@usgs.gov>

Dennis Helsel <dhelsel@practicalstats.com>

See Also

cenreg survreg

cenreg-lognormal-class

Class "cenreg-lognormal"

Description

A "cenreg-lognormal" object is returned from cenreg when a lognormal distribution is chosen with the 'dist' option.

Objects from the Class

Objects can be created by calls of the form cenreg(obs,censored,dist="lognormal").

Slots

n: Total number of observations associated with the model

n.cen: Number of censored observations

y: Vector of observations

ycen: Censoring indicator

conf.int: Confidence interval associated with the model

survreg: Object of class "survreg"

Extends

Class "cenreg"

Methods

predict signature(object = "cenreg"): ...

print signature(x = "cenreg"): ...

summary signature(object = "cenreg"): ...

Author(s)

R. Lopaka Lee <rclee@usgs.gov>

Dennis Helsel <dhelsel@practicalstats.com>

See Also

cenreg survreg

---

cenreg-methods *Methods for function cenreg in Package NADA*

Description

See [cenreg](#) for all the details.

---

censtats *Produces summary statistics using ROS, MLE, and K-M methods.*

Description

A convenience function that produces a comparative table of summary statistics obtained using the `cenros`, `cenmle` and `cenfit` routines. These methods are, Regression on Order Statistics (ROS), Maximum Likelihood Estimation (MLE), and Kaplan-Meier (K-M).

Usage

```
censtats(obs, censored)
```

Arguments

obs   A numeric vector of observations.

censored   A logical vector indicating TRUE where an observation in x is censored (a less-
                than value) and FALSE otherwise.

Details

If the data do not fulfill the criteria for the application of any method no summary statistics will be produced.

Value

A dataframe with the summary statistics.

Author(s)

R. Lopaka Lee <rclee@usgs.gov>

Dennis Helsel <dhelsel@practicalstats.com>

References

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

Examples

```
data(DFe)
with(DFe, censtats(Summer, SummerCen))
```

---

censummary   *Produces basic summary statistics on censored data*

Description

Produces basic, and hopefully useful, summary statistics on censored data.

Usage

```
censummary(obs, censored, groups)
```

Arguments

obs	A numeric vector of observations.
censored	A logical vector indicating TRUE where an observation in x is censored (a less-than value) and FALSE otherwise.
groups	A factor vector used for grouping 'obs' into subsets.

Value

A censummary object.

Author(s)

R. Lopaka Lee <rclee@usgs.gov>

Dennis Helsel <dhelsel@practicalstats.com>

References

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

Examples

```
data(DFe)
with(DFe, censummary(Summer, SummerCen))
```

---

censummary-methods   *Methods for function censummary in Package NADA*

Description

See censummary for all the details.

---

cenxyplot   *Produces a censored x-y scatter plot*

Description

Draws a x-y scatter plot with censored values represented by dashed lines spanning the from the censored threshold to zero.

Usage

```
cenxyplot(x, xcen, y, ycen, log="", lty="dashed", ...)
```

Arguments

x	A numeric vector of observations.
xcen	A logical vector indicating TRUE where an observation in x is censored (a less-than value) and FALSE otherwise.
y	A numeric vector of observations.
ycen	A logical vector indicating TRUE where an observation in y is censored (a less-than value) and FALSE otherwise.
log	A character string which contains '"x"' if the x axis is to be logarithmic, '"y"' if the y axis is to be logarithmic and '"xy"' or '"yx"' if both axes are to be logarithmic. Default is '""', or both axis linear.
lty	The line type of the lines representing the censored-data ranges.
...	Additional items that get passed to `plot`.

Author(s)

R. Lopaka Lee <rclee@usgs.gov>

Dennis Helsel <dhelsel@practicalstats.com>

References

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

Examples

```
data(DFe)
with(DFe, cenxyplot(Year, YearCen, Summer, SummerCen))
```

ChlfmCA   *Chloroform concentrations in California groundwater.*

Description

Chloroform concentrations in groundwaters of California.

Objective is to determine if concentrations differ between urban and rural areas. There are three detection limits, at 0.05, 0.1, and 0.2 ug/L. Used in Chapter 9 of the NADA book.

Usage

```
data(ChlfmCA)
```

Source

Squillace et al., 1999, Environmental Science and Technology 33, 4176-4187.

References

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

---

coef-methods   *Methods for function coef in package NADA*

Description

Methods for extracting coefficients from MLE regression models in package **NADA**

Usage

```
## S4 method for signature 'cenreg'
coef(object, ...)
```

Arguments

object	An output object from a NADA function such as cenreg.
...	Additional parameters to subclasses – currently none

See Also

cenreg

---

cor-methods   *Methods for function cor in Package NADA*

Description

Methods for function cor in package **NADA**

Methods

x = "cenreg" Extracts the r-likelihood correlation coefficient from a cenreg object.

CuZn   *Copper and zinc concentrations in ground water*

Description

Copper and zinc concentrations in ground waters from two zones in the San Joaquin Valley of California. The zinc concentrations were used.

Objective is to determine if zinc concentrations differ between the two zones. Zinc has two detection limits, at 3 and 10 ug/L. Used in Chapters 4, 5 and 9 of the NADA book.

Usage

```
data(CuZn)
```

Source

Millard and Deverel, 1988, Water Resources Research 24, pp. 2087-2098.

References

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

CuZnAlt   *Zinc concentrations of the CuZn data set*

Description

Zinc concentrations of the CuZn data set; concentrations in the Alluvial Fan zone have been altered so that there are more nondetects. This produces a greater signal, even with more nondetects.

Objective is to determine if zinc concentrations differ between the two zones. Zinc has two detection limits, at 3 and 10 ug/L. Used in Chapter 9 of the NADA book.

Usage

```
data(CuZnAlt)
```

Source

Altered from the data of Millard and Deverel, 1988, Water Resources Research 24, pp. 2087-2098.

References

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

DFe   *Dissolved iron concentrations from the Brazos River, USA*

Description

Dissolved iron concentrations over several years in the Brazos River, Texas. Summer concentrations were used.

Objective is to determine if there is a trend over time. Iron has two detection limits, at 3 and 10 ug/L. Used in Chapters 5, 11 and 12 of the NADA book.

Usage

```
data(DFe)
```

Source

Hughes and Millard, 1988, Water Resources Bulletin 24, pp. 521-531.

References

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

DOC *DOC in ground water*

Description

Dissolved Organic Carbon (DOC) concentrations in ground waters of irrigated and non-irrigated areas.

Objective is to determine if concentrations differ between irrigated and non-irrigated areas. There is one detection limit at 0.2 ug/L. Used in Chapter 9 of the NADA book.

Usage

```
data(DOC)
```

Source

Junk et al., 1980, Journal of Environmental Quality 9, pp. 479-483.

References

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

---

flip-methods *Methods for function flip in Package NADA*

Description

Methods for function `flip` in package **NADA**.

When used in concert with `Cen`, `flip` rescales left-censored data into right-censored data for use in the survival package routines (which can only handle right-censored data sets).

Usage

```
## S4 method for signature 'Cen'
flip(x)

## S4 method for signature 'formula'
flip(x)
```

Arguments

x   A Cen or `formula` object.

Notes

Flips, or rescales a `Cen` object or a formula object.

By default, `flip` rescales the input data by subtracting a large constant that is larger than maximum input value from all observations. It then marks the data as right censored so that routines from the `survival` package can be used.

IMPORTANT: All NADA routines transparently handle flipping and re-transforming data. Thus, `flip` should almost never be used, except perhaps in the development of an extension function.

Also, flipping a Cen object results in a Surv object – which presently cannot be flipped back to a Cen object!

Flipping a formula just symbolically updates the response (which should be a Cen object). Result is like: flip(Cen(obs, cen))~groups

---

Golden   *Blood lead in organs of herons from Virginia*

Description

Lead concentrations in the blood and several organs of herons in Virginia.

Objective is to determine the relationships between lead concentrations in the blood and various organs. Do concentrations reflect environmental lead concentrations, as represented by dosing groups? There is one detection limit, at 0.02 ug/g. Used in Chapters 10 and 11 of the NADA book.

Usage

```
data(Golden)
```

Source

Golden et al., 2003, Environmental Toxicology and Chemistry 22, pp. 1517-1524.

References

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

Hatchery	*Antibiotic concentrations in fish-hatchery drainage*

Description

Proportions of detectable concentrations of antibiotics (ug/L) in drainage from fish hatcheries across the United States.

Objective is to compute confidence intervals and tests on proportions.

There is one detection limit for each compound, all at 0.05 ug/L. Used in Chapters 8 and 9 of the NADA book.

Usage

```
data(Hatchery)
```

Source

Thurman et al., 2002, Occurrence of antibiotics in water from fish hatcheries. USGS Fact Sheet FS 120-02.

References

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

---

hc.ppoints	*Helsel-Cohn style plotting positions*

Description

Helsel-Cohn style plotting positions for multiply-censored data.

Usage

```
hc.ppoints(obs, censored, na.action)
hc.ppoints.uncen(obs, censored, cn, na.action)
hc.ppoints.cen(obs, censored, cn, na.action)
```

Arguments

obs	A numeric vector of observations. This includes both censored and uncensored observations.
censored	A logical vector indicating TRUE where an observation in v is censored (a less-than value) and FALSE otherwise.
cn	An optional argument for internal-code use only. cn = a Cohn Numbers list (quantities described by Helsel and Cohn (1988) in their formulation of the problem).
na.action	A function which indicates what should happen when the data contain NAs. The default is set by the na.action setting of [options](#), and is [na.omit](#) if that is unset. Another possible value is NULL, no action.

Details

The function computes Wiebull-type plotting positions of data containing mixed uncensored and censored data. The formula was first described by Hirsch and Stedinger (1897) and latter reformulated by Helsel and Cohn (1988). It assumes that censoring is left-censoring (less-thans). A detailed discussion of the formulation is in Lee and Helsel (in press).

Note that if the input vector 'censored' is of zero length, then the plotting positions are calculated using ppoints. Otherwise, hc.ppoints.uncen and hc.ppoints.cen are used.

hc.ppoints.uncen calculates plotting positions for uncensored data only.

hc.ppoints.cen calculates plotting positions for censored data only.

Value

hc.ppoints returns a numeric vector of plotting positions which correspond to the observations in the input vector 'obs'.

hc.ppoints.uncen returns a numeric vector of plotting positions which correspond to the uncensored observations in the input vector 'obs'.

hc.ppoints.cen returns a numeric vector of plotting positions which correspond to the censored observations in the input vector 'obs'.

Author(s)

R. Lopaka Lee <rclee@usgs.gov>

Dennis Helsel <dhelsel@practicalstats.com>

References

Lee and Helsel (in press), Statistical analysis of environmental data containing multiple detection limits: S-language software for linear regression on order statistics, Computers in Geoscience vol. X, pp. X-X

Dennis R. Helsel and Timothy A. Cohn (1988), Estimation of descriptive statistics for multiply censored water quality data, Water Resources Research vol. 24, no. 12, pp.1997-2004

Robert M. Hirsch and Jery R. Stedinger (1987), Plotting positions for historical floods and their precision. Water Resources Research, vol. 23, no. 4, pp. 715-727.

See Also

[ros](#), [splitQual](#)

Examples

```
obs     = c(0.5,    0.5,   1.0,  1.5,   5.0,   10,   100)
censored = c(TRUE, FALSE, FALSE, TRUE, FALSE, FALSE, FALSE)

hc.ppoints(obs, censored)
```

HgFish *Mercury concentrations in fish across the United States.*

Description

Mercury concentrations in fish across the United States.

Objective is to determine if mercury concentrations differ by watershed land use. Can concentrations be related to water and sediment characteristics of the streams?

There are three detection limits, at 0.03, 0.05, and 0.10 ug/g wet weight. Used in Chapters 10, 11 and 12 of the NADA book.

Usage

```
data(HgFish)
```

Source

Brumbaugh et al., 2001, USGS Biological Science Report BSR-2001-0009.

References

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

---

lines-methods *Methods for function lines in Package NADA*

Description

Methods for adding `lines` to plots in package **NADA**

Usage

```
## S4 method for signature 'ros'
lines(x, ...)

## S4 method for signature 'cenfit'
lines(x, ...)

## S4 method for signature 'cenken'
lines(x, ...)
```

Arguments

x   An output object from a NADA function such as `ros`.

...   Additional arguments passed to the generic method.

See Also

`lines`, `plot`

MDCu   *Copper in ground water from of San Joaquin Valley, USA*

Description

Copper concentrations in ground water from the Alluvial Fan zone in the San Joaquin Valley of California. One observation was altered to become a <21, larger than all of the detected observations (the largest detected observation is a 20).

Objective is to calculate summary statistics when the largest observation is censored.

There are five detection limits, at 1, 2, 5, 10 and 20 ug/L. An additional artificial detection limit of 21 was added to illustrate a point. Used in Chapter 6 of the NADA book.

Usage

```
data(MDCu)
```

Source

Millard and Deverel, 1988, Water Resources Research 24, pp. 2087-2098.

References

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

---

mean-methods *Methods for function mean in Package NADA*

Description

Methods for computing the mean using model objects in package **NADA**

Usage

```
## S4 method for signature 'ros'
mean(x, ...)

## S4 method for signature 'cenfit'
mean(x, ...)

## S4 method for signature 'cenmle'
mean(x, ...)
```

Arguments

x   An output object from a NADA function such as [ros](#ros).

...   Additional arguments passed to the generic method.

See Also

[mean](#mean)

---

median-methods *Methods for function median in Package NADA*

Description

Methods for computing the median using model objects in package **NADA**

Usage

```
## S4 method for signature 'ros'
median(x, na.rm=FALSE)

## S4 method for signature 'cenfit'
median(x, na.rm=FALSE)

## S4 method for signature 'cenmle'
median(x, na.rm=FALSE)
```

Arguments

x   An output object from a NADA function such as ros.

na.rm   Should NAs be removed prior to computation?

See Also

median

NADAList-class   *Class "NADAList"*

Description

A "NADAList" simply extends the 'list' class.

Objects from the Class

NADAList objects are created by calls like cenken(y,ycen,x,xcen) and other functions.

Slots

.Data: Object of class "list"

Extends

Class "list", from data part.

Methods

show signature(object = "NADAList"): ...

Author(s)

R. Lopaka Lee <rclee@usgs.gov>

Dennis Helsel <dhelsel@practicalstats.com>

See Also

[cenken](#)

Oahu   *Arsenic concentrations in Manoa Stream, Oahu Hawaii*

Description

Arsenic concentrations (ug/L) in an urban stream, Manoa Stream at Kanewai Field, on Oahu, Hawaii.

Objective is to characterize conditions by computing summary statistics.

There are three detection limits, at 0.9, 1, and 2 ug/L. Uncensored values reported below the lowest detection limit indicate that informative censoring may have been used, and so the results are likely biased high. Used in Chapter 6 of the NADA book.

Usage

```
data(Oahu)
```

Source

Tomlinson, 2003, Effects of Ground-Water/Surface-Water Interactions and Land Use on Water Quality. Written communication (draft USGS report).

References

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

---

pctCen *Calculate the percentage of values censored*

Description

pctCen is a simple, but convenient, function that calculates the percentage of censored values.

Usage

```
pctCen(obs, censored, na.action)
```

Arguments

obs	A numeric vector of observations. This includes both censored and uncensored observations.
censored	A logical vector indicating TRUE where an observation in v is censored (a less-than value) and FALSE otherwise.
na.action	A function which indicates what should happen when the data contain NAs. The default is set by the na.action setting of options, and is na.omit if that is unset. Another possible value is NULL, no action.

Details

100*(length(obs[censored])/length(obs))

Value

pctCen returns a single numeric value representing the percentage of values censored in the "obs" vector.

Author(s)

R. Lopaka Lee <rclee@usgs.gov>

Dennis Helsel <dhelsel@practicalstats.com>

See Also

splitQual, [ros](ros),

Examples

```
    obs      = c(0.5,    0.5,   1.0,  1.5,   5.0,    10,   100)
    censored = c(TRUE, FALSE, FALSE, TRUE, FALSE, FALSE, FALSE)

    pctCen(obs, censored)
```

---

plot-methods *Methods for function plot in Package NADA*

Description

Methods for plotting objects in package **NADA**

Usage

```
## S4 method for signature 'ros'
plot(x, plot.censored=FALSE, lm.line=TRUE, grid=TRUE, ...)

## S4 method for signature 'cenfit'
plot(x, conf.int=FALSE, ...)

## S4 method for signature 'cenmle'
plot(x, ...)

## S4 method for signature 'cenreg'
plot(x, ...)
```

Arguments

x	An output object from a NADA function such as [ros](#).
conf.int	A logical indicating if confidence intervals should be computed. For cenfit objects, the confidence interval is set during the call to [cenfit](#). Currently not supported for ros objects.
plot.censored	ros: should censored values be plotted?
lm.line	ros: should the linear regression line be plotted?
grid	ros: should a grid be overlayed?
...	Additional arguments passed to the generic method.

See Also

[plot](#)

predict-methods   *Methods for function predict in package NADA*

Description

Functions that perform predictions using NADA model objects.

For `ros` models, predict the normal quantile of a value.

For `cenfit` objects, predict the probabilities of new observations.

Usage

```
## S4 method for signature 'ros'
predict(object, newdata, ...)

## S4 method for signature 'cenfit'
predict(object, newdata, conf.int=FALSE, ...)

## S4 method for signature 'cenreg'
predict(object, newdata, conf.int=FALSE, ...)

## S4 method for signature 'cenfit'
pexceed(object, newdata, conf.int=FALSE, ...)

## S4 method for signature 'ros'
pexceed(object, newdata, conf.int=FALSE, conf.level=0.95, ...)
```

Arguments

object	An output object from a NADA function such as [ros](#).
newdata	Numeric vector of data for which to predict model values. For `ros` objects this will be new normalized quantiles of plotting positions. For `cenfit` objects this will be new observations for which you desire the modeled probabilities.
conf.int	A logical indicating if confidence intervals should be computed. For `cenfit` objects, the confidence interval is set during the call to [cenfit](#). Currently not supported for `ros` objects.
conf.level	The actual confidence level to which to bracket the prediction. Default is 0.95
...	Additional arguments passed to the generic method.

---

print-methods   *Methods for function print in Package NADA*

Description

Methods for function print in package **NADA**

Methods

x = "ANY" Default print method

x = "Cen" Displays a [Cen](#) object.

x = "cenfit" Displays a [cenfit](#) object.

x = "cenmle" Displays a [cenmle](#) object.

x = "cenreg" Displays a [cenreg](#) object.

x = "summary.cenreg" Displays a [summary.cenreg](#) object.

x = "ros" Displays a [ros](#) object.

x = "censummary" Displays a [censummary](#) object.

x = "NADAList" Displays a NADAList object. This is an internal method and should rarely be used from the command line.

---

quantile-methods   *Methods for function quantile in Package NADA*

Description

Methods for the function quantile in package **NADA**

Compute the modeled values of quantiles or probabilities using a model object.

Usage

```
## S4 method for signature 'ros'
quantile(x, probs=NADAprobs, ...)

## S4 method for signature 'cenfit'
quantile(x, probs=NADAprobs, conf.int=FALSE, ...)

## S4 method for signature 'cenmle'
quantile(x, probs=NADAprobs, conf.int=FALSE, ...)
```

Arguments

x	An output object from a NADA fuction such as ros.
probs	Numeric vector of probabilities for which to calculate model values. The default is the global variable NADAprobs = c(0.05, 0.10, 0.25, 0.50, 0.75, 0.90, 0.95).
conf.int	A logical indicating if confidence intervals should be computed. For cenfit and cenmle objects, the confidence interval is set during the call to cenfit. Currently not supported for ros objects.
...	Additional arguments passed to the generic method.

Examples

```
data(Cadmium)

mymodel = cenfit(Cadmium$Cd, Cadmium$CdCen, Cadmium$Region)

quantile(mymodel, conf.int=TRUE)
```

Recon   *Atrazine in streams of the Midwestern U.S.*

Description

Atrazine concentrations in streams throughout the Midwestern United States.

Objective is to develop a regression of model for atrazine concentrations using explanatory variables.

There is one detection limit, at 0.05 ug/L. Used in Chapter 12 of the NADA book.

Usage

```
data(Recon)
```

Source

Mueller et al., 1997, Journal of Environmental Quality 26, pp. 1223-1230.

References

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

---

residuals-methods   *Methods for function residuals in package NADA*

Description

Methods for extracting residuals from MLE regression models in package **NADA**

Usage

```
## S4 method for signature 'cenreg'
residuals(object, ...)
```

Arguments

object   An output object from a NADA function such as cenreg.

...   Additional parameters to subclasses – currently none

See Also

cenreg

Roach   *Lindane in fish from tributaries of the Thames River, UK*

Description

Lindane concentrations in fish from tributaries of the Thames River, England.

Objective is to determine whether lindane concentrations are the same at all sites.

There is one detection limit at 0.08 ug/kg. Used in Chapter 9 of the NADA book.

Usage

```
data(Roach)
```

Source

Yamaguchi et al., 2003, Chemosphere 50, 265-273.

References

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

ros	*Regression on Order Statistics*

Description

`ros` is an implementation of a Regression on Order Statistics (ROS) designed for multiply censored analytical chemistry data.

The method assumes data contains zero to many left censored (less-than) values.

Usage

```
ros(obs, censored, forwardT="log", reverseT="exp", na.action)
```

Arguments

`obs`	A numeric vector of observations. This includes both censored and uncensored observations.
`censored`	A logical vector indicating TRUE where an observation in `obs` is censored (a less-than value) and FALSE otherwise.
`forwardT`	A name of a function to use for transformation prior to performing the ROS fit. Defaults to `log`.
`reverseT`	A name of a function to use for reversing the transformation after performing the ROS fit. Defaults to `exp`.
`na.action`	A function which indicates what should happen when the data contain NAs. The default is set by the `na.action` setting of [options](#), and is [na.omit](#) if that is unset. Another possible value is NULL, no action.

Details

By default, `ros` performs a log transformation prior to, and after operations over the data. This can be changed by specifying a forward and reverse transformation function using the `forwardT` and `reverseT` parameters. No transformation will be performed if either `forwardT` or `reverseT` are set to NULL.

The procedure first computes the Weibull-type plotting positions of the combined uncensored and censored observations using a formula designed for multiply-censored data (see `hc.ppoints`). A linear regression is formed using the plotting positions of the uncensored observations and their normal quantiles. This model is then used to estimate the concentration of the censored observations as a function of their normal quantiles. Finally, the observed uncensored values are combined with modeled censored values to corporately estimate summary statistics of the entire population. By combining the uncensored values with modeled censored values, this method is more resistant of any non-normality of errors, and reduces any transformation errors that may be incurred.

Value

`ros` returns an object of class c("ros", "lm").

`print` displays a simple summary of the ROS model. `as.data.frame` converts the modeled data in a ROS model to a data frame. Note that this discards all linear-model information from the object.

Author(s)

R. Lopaka Lee <rclee@usgs.gov>

Dennis Helsel <dhelsel@practicalstats.com>

References

Lee and Helsel (2005) Statistical analysis of environmental data containing multiple detection limits: S-language software for regression on order statistics, Computers in Geoscience vol. 31, pp. 1241-1248.

Lee and Helsel (2005) Baseline models of trace elements in major aquifers of the United States. Applied Geochemistry vol. 20, pp. 1560-1570.

Dennis R. Helsel (2005), Nondetects And Data Analysis: John Wiley and Sons, New York.

Dennis R. Helsel (1990), Less Than Obvious: Statistical Methods for, Environmental Science and Technology, vol.24, no. 12, pp. 1767-1774

Dennis R. Helsel and Timothy A. Cohn (1988), Estimation of descriptive statistics for multiply censored water quality data, Water Resources Research vol. 24, no. 12, pp.1997-2004

See Also

splitQual, predict, plot, ros-class, ros-methods, plot-methods, mean-methods, sd-methods, quantile-methods, median-methods, predict-methods, summary-methods

Examples

```
obs     = c(0.5,    0.5,   1.0,  1.5,   5.0,    10,   100)
censored = c(TRUE, FALSE, FALSE, TRUE, FALSE, FALSE, FALSE)

myros = ros(obs, censored)

plot(myros)
summary(myros)
mean(myros); sd(myros)
quantile(myros); median(myros)
as.data.frame(myros)
```

---

ros-class   *Class "ros"*

Description

A "ros" object is returned from ros. It extends the "lm" class returned from lm.

Objects from the Class

Objects can be created by calls of the form ros(obs,censored).

Slots

.Data: Object of class "list"

Extends

Class "list", from data part. Class "vector", by class "list".

Methods

lines signature(x = "ros"): ...
mean signature(x = "ros"): ...
median signature(x = "ros"): ...
plot signature(x = "ros", y = "missing"): ...
predict signature(object = "ros"): ...
print signature(x = "ros"): ...
quantile signature(x = "ros"): ...
sd signature(x = "ros"): ...
summary signature(object = "ros"): ...

Author(s)

R. Lopaka Lee <rclee@usgs.gov>

Dennis Helsel <dhelsel@practicalstats.com>

See Also

ros

Examples

```
obs      = c(0.5,    0.5,   1.0,  1.5,   5.0,    10,   100)
censored = c(TRUE, FALSE, FALSE, TRUE, FALSE, FALSE, FALSE)

class(ros(obs, censored))
```

---

ros-methods   *Methods for function ros in Package NADA*

Description

Methods for constructing ROS models in package **NADA**

Methods

obs = "numeric", censored = "logical" Compute and return a ROS model given a numeric vector
    of observations and a logical vector indicating TRUE or FALSE where the observations are
    not censored or censored respectively.

sd-methods   *Methods for function sd in Package NADA*

Description

Methods for computing standard deviations in package **NADA**

Usage

```
## S4 method for signature 'ros'
sd(x, na.rm=FALSE)

## S4 method for signature 'cenfit'
sd(x, na.rm=FALSE)

## S4 method for signature 'cenmle'
sd(x, na.rm=FALSE)
```

Arguments

x	An output object from a NADA function such as [ros](#).
na.rm	Should NAs be removed prior to computation?

See Also

[sd](#)

SedPb   *Lead in stream sediments before and after wildfires*

Description

Lead concentrations in stream sediments before and after wildfires.

Objective is to determine whether lead concentrations are the same pre- and post-fire.

There is one detection limit at 4 ug/L. Used in Chapter 9 of the NADA book.

Usage

```
data(SedPb)
```

Source

Eppinger et al., 2003, USGS Open-File Report 03-152.

References

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

ShePyrene   *Pyrene concentrations in water from Puget Sound, WA USA*

Description

Pyrene concentrations in milligrams per liter from 20 water-quality monitoring stations in the Puget Sound of Washington State, USA.

Used for characterizing priority pollutant concentrations in sediments of Puget Sound by computing summary statisitics. Contains eight detection limits with 11 nondetects out of 56 total measurements.

Usage

```
data(ShePyrene)
```

Source

She, N., 1997, Analyzing censored water quality data using a nonparametric approach. Journal of the American Water Resources Association, 33, pp615–624.

References

She, N., 1997, Analyzing censored water quality data using a nonparametric approach. Journal of the American Water Resources Association, 33, pp615–624.

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

---

show-methods   *Methods for function show in Package NADA*

Description

Methods for showting objects in package **NADA**

Usage

```
## S4 method for signature 'ros'
show(object)

## S4 method for signature 'cenfit'
show(object)

## S4 method for signature 'cenmle'
show(object)

## S4 method for signature 'cenreg'
show(object)

## S4 method for signature 'summary.cenreg'
show(object)

## S4 method for signature 'cenken'
show(object)

## S4 method for signature 'censummary'
show(object)

## S4 method for signature 'NADAList'
show(object)
```

Arguments

`object`   An output object from a NADA function such as [cenfit](#).

See Also

[show](#)

Silver   *Silver-standard concentrations*

Description

Silver concentrations in a standard solution sent to 56 laboratories as part of a quality assurance program.

Objective is to estimate summary statistics for the standard solution. The median or mean might be considered the most likel estimate of the concentration.

Contains twelve detection limits, the largest at 25 ug/L. Used in Chapter 6 of the NADA book.

Usage

```
data(Silver)
```

Source

Helsel and Cohn, 1988, Water Resources Research 24, pp. 1997-2004.

References

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

---

splitQual   *Split character qualifiers and numeric values from qualified data*

Description

splitQual extracts qualified and unqualified vectors from a character vector containing concatenated numeric and qualifying characters.

Typically used to split "less-thans" in qualifier-numeric concatenations like "<0.5".

Usage

```
 splitQual(v, qual.symbol= ”<”)
```

Arguments

v   A character vector.

qual.symbol   The qualifier symbol to split from the characters in v. Defaults to "<".

Value

splitQual returns a list of three vectors.

qual   A numeric vector of values associated with qualified input.

unqual   A numeric vector of values associated with unqualified input

qual.index   Indexes of qualified values (ie., where qual.symbol was matched)

unqual.index   Indexes of unqualified values (ie., where qual.symbol was not matched)

Author(s)

R. Lopaka Lee <rclee@usgs.gov>

Dennis Helsel <dhelsel@practicalstats.com>

References

Lee and Helsel (2005), Statistical analysis of environmental data containing multiple detection limits: S-language software for regression on order statistics, Computers in Geoscience vol. 31, pp. 1241-1248

Examples

```
v = c('<1', 1, '<1', 1, 2)
splitQual(v)
```

---

summary-methods   *Methods for function summary in Package 'NADA'*

Description

Methods for summarizing objects in package **NADA**

Usage

```
## S4 method for signature 'ros'
summary(object, plot=FALSE, ...)

## S4 method for signature 'cenfit'
summary(object, ...)

## S4 method for signature 'cenreg'
summary(object, ...)
```

Arguments

object	An output object from a NADA function such as ros.
plot	Logical indicating if summary graphs be generated?
...	Additional arguments passed to the generic method.

---

summary.cenreg-class   *Class "summary.cenreg"*

Description

A "summary.cenreg" object is returned from summary.

Objects from the Class

Objects can be created by calls of the form summary(cenreg(obs,censored,groups)).

Slots

.Data: Object of class "list"

Extends

Class "list". Class "vector", by class "list".

Methods

summary signature(object = "cenreg"): ...

Author(s)

R. Lopaka Lee <rclee@usgs.gov>

Dennis Helsel <dhelsel@practicalstats.com>

See Also

cenreg

Tbl1one   *Contaminant concentrations in test and a control group*

Description

Contaminant concentrations in test and a control group.

Objective is to determine whether a test group has higher concentrations than a control group.

There are three detection limits, at 1, 2, and 5 ug/L. Used in Chapter 1, Table 1.1 of the NADA book.

Usage

```
data(Tbl1one)
```

Source

None. Generated data.

References

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

TCE   *TCE in ground waters of Long Island, New York*

Description

TCE concentrations (ug/L) in ground waters of Long Island, New York. Categorized by the dominant land use type (low, medium, or high density residential) surrounding the wells.

Objective determine if concentrations are the same for the three land use types. There are Four detection limits, at 1,2,4 and 5 ug/L. Used in Chapter 10 of the NADA book.

Usage

```
data(TCE)
```

Source

Eckhardt et al., 1989, USGS Water Resources Investigations Report 86-4142.

References

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

TCEReg   *TCE ground waters of Long Island – with explanatory variables*

Description

TCE concentrations (ug/L) in ground waters of Long Island, New York, along with several possible explanatory variables.

Objective is to determine if concentrations are related to one or more explanatory variables.

There are four detection limits, at 1,2,4 and 5 ug/L. One column indicates whether concentrations are above or below 5. Used in Chapter 12 of the NADA book.

Usage

```
data(TCEReg)
```

Source

Eckhardt et al., 1989, USGS Water Resources Investigations Report 86-4142.

References

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

| Thames | *Dieldrin, lindane and PCB in fish of the Thames River, UK* |

Description

Dieldrin, lindane and PCB concentrations in fish of the Thames River and tributaries, England.

Objective is to determine if concentrations differ among sampling sites. Are dieldrin and lindane concentrations correlated? There is one detection limit per compound. Used in Chapters 11 and 12 of the NADA book.

Usage

```
data(Thames)
```

Source

Yamaguchi et al., 2003, Chemosphere 50, 265-273.

References

Helsel, Dennis R. (2005). Nondectects and Data Analysis; Statistics for censored environmental data. John Wiley and Sons, USA, NJ.

Index