Package 'sasLM'

January 16, 2021

Version 0.3.0

Title 'SAS' Linear Model

Description This is a core implementation of 'SAS' procedures for linear models - GLM, REG, and ANOVA. Some packages provide type II and type III SS. However, the results of nested and complex designs are often different from those of 'SAS.' Different results does not necessarily mean incorrectness. However, many wants the same results to SAS. This package aims to achieve that. Reference: Littell RC, Stroup WW, Freund RJ (2002, ISBN:0-471-22174-0).

Depends R (¿= 3.0.0)

Author Kyun-Seop Bae [aut]

Maintainer Kyun-Seop Bae <k@acr.kr>

License GPL-3

Repository CRAN

URL https://cran.r-project.org/package=sasLM

R topics documented:

---

sasLM-package *'SAS' Linear Model*

Description

This is a core implementation of 'SAS' procedures for linear models - GLM, REG, and ANOVA. Some packages provide type II and type III SS. However, the results of nested and complex designs are often different from those of 'SAS'. Different results does not necessarily mean incorrectness. However, many wants the same results to 'SAS'. This package aims to achieve that. Reference: Littell RC, Stroup WW, Freund RJ (2002, ISBN:0-471-22174-0).

Details

This will serve those who want SAS PROC GLM, REG, and ANOVA in R.

Author(s)

Kyun-Seop Bae k@acr.kr

Examples

```
## SAS PROC GLM Script for Typical Bioequivalence Data
# PROC GLM DATA=BEdata;
#   CLASS SEQ SUBJ PRD TRT;
#   MODEL LNCMAX = SEQ SUBJ(SEQ) PRD TRT;
#   RANDOM SUBJ(SEQ)/TEST;
#   LSMEANS TRT / DIFF=CONTROL("R") CL ALPHA=0.1;
#   ODS OUTPUT LSMeanDiffCL=LSMD;

# DATA LSMD;  SET LSMD;
#   PE = EXP(DIFFERENCE);
#   LL = EXP(LowerCL);
#   UL = EXP(UpperCL);
# PROC PRINT DATA=LSMD; RUN;
##

## SAS PROC GLM equivalent
BEdata = af(BEdata, c("SEQ", "SUBJ", "PRD", "TRT")) # Columns as factor
formula1 = log(CMAX) ~ SEQ/SUBJ + PRD + TRT # Model
GLM(formula1, BEdata) # ANOVA tables of Type I, II, III SS
T3MS(formula1, BEdata) # EMS table
T3test(formula1, BEdata, Error="SEQ:SUBJ") # Hypothesis test
exp(CIest(formula1, BEdata, "TRT", c(-1, 1), 0.10)) # 90% CI of GMR

## 'nlme' or SAS PROC MIXED is preferred for an unbalanced case
## SAS PROC MIXED equivalent
# require(nlme)
# Result = lme(log(CMAX) ~ SEQ + PRD + TRT, random=~1|SUBJ, data=BEdata)
```

```
# summary(Result)
# VarCorr(Result)
# ci = intervals(Result, 0.90) ; ci
# exp(ci$fixed["TRTT",])
##
```

`af` *Convert some columns of a data.frame to factors*

Description

Conveniently convert some columns of data.frame into factors.

Usage

```
af(DataFrame, Cols)
```

Arguments

`DataFrame`	a `data.frame`
`Cols`	column names or indices to be converted

Details

It performs conversion of some columns in a `data.frame` into factors conveniently.

Value

Returns a `data.frame` with converted columns.

Author(s)

Kyun-Seop Bae k@acr.kr

`ANOVA` *Analysis of Variance similar to SAS PROC ANOVA*

Description

Analysis of variance with type I, II, and III sum of squares.

Usage

```
ANOVA(Formula, Data, eps=1e-8)
```

Arguments

`Formula`	a conventional formula for a linear model.
`Data`	a `data.frame` to be analyzed
`eps`	Less than this value is considered as zero.

Details

It performs the core function of SAS PROC ANOVA.

Value

The result is comparable to that of SAS PROC ANOVA.

`ANOVA`	ANOVA table for the model
`Type I`	Type I sum of square table
`Type II`	Type II sum of square table
`Type III`	Type III sum of square table

Author(s)

Kyun-Seop Bae k@acr.kr

Examples

```
ANOVA(uptake ~ Plant + Type + Treatment + conc, CO2)
```

---

aov1 *ANOVA with Type I SS*

Description

ANOVA with Type I SS.

Usage

```
aov1(Formula, Data, eps=1e-8)
```

Arguments

`Formula`	a conventional formula for a linear model.
`Data`	a `data.frame` to be analyzed
`eps`	Less than this value is considered as zero.

Details

It performs the core function of SAS PROC ANOVA.

Value

The result table is comparable to that of SAS PROC ANOVA.

`Df`	degree of freedom
`Sum Sq`	sum of square for the set of contrasts
`Mean Sq`	mean square
`F value`	F value for the F distribution
`Pr(>F)`	proability of larger than F value

Author(s)

Kyun-Seop Bae k@acr.kr

Examples

```
aov1(uptake ~ Plant + Type + Treatment + conc, CO2)
```

---

aov2 *ANOVA with Type II SS*

Description

ANOVA with Type II SS.

Usage

```
aov2(Formula, Data, eps=1e-8)
```

Arguments

Formula	a conventional formula for a linear model.
Data	a data.frame to be analyzed
eps	Less than this value is considered as zero.

Details

It performs the core function of SAS PROC ANOVA.

Value

The result table is comparable to that of SAS PROC ANOVA.

Df	degree of freedom
Sum Sq	sum of square for the set of contrasts
Mean Sq	mean square
F value	F value for the F distribution
Pr(>F)	proability of larger than F value

Author(s)

Kyun-Seop Bae k@acr.kr

Examples

```
aov2(uptake ~ Plant + Type + Treatment + conc, CO2)
```

aov3 *ANOVA with Type III SS*

Description

ANOVA with Type III SS.

Usage

```
aov3(Formula, Data, eps=1e-8)
```

Arguments

Formula	a conventional formula for a linear model.
Data	a data.frame to be analyzed
eps	Less than this value is considered as zero.

Details

It performs the core function of SAS PROC ANOVA.

Value

The result table is comparable to that of SAS PROC ANOVA.

Df	degree of freedom
Sum Sq	sum of square for the set of contrasts
Mean Sq	mean square
F value	F value for the F distribution
Pr(>F)	proability of larger than F value

Author(s)

Kyun-Seop Bae k@acr.kr

Examples

```
aov3(uptake ~ Plant + Type + Treatment + conc, CO2)
```

BEdata — *An Example Data of Bioequivalence Study*

Description

Contains Cmax data from a real bioequivalence study.

Usage

```
BEdata
```

Format

A data frame with 91 observations on the following 6 variables.

`ADM` Admission or Hospitalization Group Code: 1, 2, or 3

`SEQ` Group or Sequence character code: 'RT' or 'TR"

`PRD` Period numeric value: 1 or 2

`TRT` Treatment or Drug code: 'R' or 'T'

`SUBJ` Subject ID

`CMAX` Cmax values

Details

This contains a real data of 2x2 bioequivalence study, which have three different hospitalization groups. See Bae KS, Kang SH. Bioequivalence data analysis for the case of separate hospitalization. Transl Clin Pharmacol. 2017;25(2):93-100. doi.org/10.12793/tcp.2017.25.2.93

CIest — *Confidence Interval Estimation*

Description

Get point estimate and its confidence interval with given contrast and alpha value using t distribution.

Usage

```
CIest(Formula, Data, Term, Contrast=c(-1, 1), Alpha=0.10)
```

Arguments

`Formula`	a conventional formula for a linear model
`Data`	a `data.frame` to be analyzed
`Term`	a factor name to be estimated
`Contrast`	a level vector. Level is alphabetically ordered by default.
`Alpha`	0.05 means 95 percent and 0.10 means 90 percent confidence interval.

Details

Get point estimate and its confidence interval with given contrast and alpha value using t distribution.

Value

Returns point estimate and its confidence interval

Author(s)

Kyun-Seop Bae k@acr.kr

Examples

```
CIest(log(CMAX) ~ SEQ/SUBJ + PRD + TRT, BEdata, Term="TRT") # 90% CI
```

---

cSS *Sum of Square with a Given Contrast Set*

Description

Calculates sum of squares of a contrast from a `lfit` result.

Usage

```
cSS(K, rx, eps=1e-8)
```

Arguments

`K`	contrast matrix. Each row is a contrast.
`rx`	a result of `lfit` function
`eps`	Less than this value is considered as zero.

Details

It calculates sum of squares with given a contrast matrix and a `lfit` result. It corresponds to SAS PROC GLM CONTRAST.

Value

Returns sum of square and its F value and p-value.

`Df`	degree of freedom
`Sum Sq`	sum of square for the set of contrasts
`Mean Sq`	mean square
`F value`	F value for the F distribution
`Pr(>F)`	proability of larger than F value

Author(s)

Kyun-Seop Bae k@acr.kr

Examples

```
x = ModelMatrix(uptake ~ Type, CO2)
y = model.frame(uptake ~ Type, CO2)[,1]
rx = lfit(x, y)
cSS(t(c(0, -1, 1)), rx) # sum of square
ANOVA(uptake ~ Type, CO2) # compare with the above
```

---

e1 *Get a Contrast Matrix for Type I SS*

Description

Makes a contrast matrix for type I SS using forward Doolittle method.

Usage

```
e1(Formula, Data, eps=1e-8)
```

Arguments

Formula	a conventional formula for a linear model
Data	a data.frame to be analyzed
eps	Less than this value is considered as zero.

Details

It makes a contrast matrix for type I SS.

Value

A contrast matrix for type I SS.

Author(s)

Kyun-Seop Bae k@acr.kr

Examples

```
round(e1(uptake ~ Plant + Type + Treatment + conc, CO2), 12)
```

e2 *Get a Contrast Matrix for Type II SS*

Description

Makes a contrast matrix for type II SS.

Usage

```
e2(Formula, Data, eps=1e-8)
```

Arguments

Formula	a conventional formula for a linear model
Data	a data.frame to be analyzed
eps	Less than this value is considered as zero.

Details

It makes a contrast matrix for type II SS.

Value

Returns a contrast matrix for type II SS.

Author(s)

Kyun-Seop Bae k@acr.kr

Examples

```
round(e2(uptake ~ Plant + Type + Treatment + conc, CO2), 12)
```

e3 *Get a Contrast Matrix for Type III SS*

Description

Makes a contrast matrix for type III SS.

Usage

```
e3(Formula, Data, eps=1e-8)
```

Arguments

Formula	a conventional formula for a linear model
Data	a data.frame to be analyzed
eps	Less than this value is considered as zero.

Details

It makes a contrast matrix for type III SS.

Value

Returns a contrast matrix for type III SS.

Author(s)

Kyun-Seop Bae k@acr.kr

Examples

```
round(e3(uptake ~ Plant + Type + Treatment + conc, CO2), 12)
```

---

est *Estimate Linear Contrast*

Description

Estimates Linear Contrast(s) with a given GLM result.

Usage

```
est(L, rx)
```

Arguments

`L`	a matrix of linear contrast rows to be tested
`rx`	a result of `lfit` function

Details

It tests rows of linear contrast. It corresponds to SAS PROC GLM ESTIMATE.

Value

Returns a table of expectations, t values and p-values.

`Estimate`	point estimate of the input linear constrast
`Std. Error`	standard error of the point estimate
`t value`	value for t distribution
`Pr(>|t|`	probability of larger than absolute t value from t distribution with residual's degree of freedom

Author(s)

Kyun-Seop Bae k@acr.kr

Examples

```
x = ModelMatrix(uptake ~ Type, CO2)
y = model.frame(uptake ~ Type, CO2)[,1]
rx = lfit(x, y)
est(t(c(0, -1, 1)), rx) # Quevec - Mississippi
t.test(uptake ~ Type, CO2) # compare with the above
```

GLM *General Linear Model similar to SAS PROC GLM*

Description

GLM is the main function of this package.

Usage

```
GLM(Formula, Data, eps=1e-8)
```

Arguments

Formula	a conventional formula for a linear model.
Data	a data.frame to be analyzed
eps	Less than this value is considered as zero.

Details

It performs the core function of SAS PROC GLM.

Value

The result is comparable to that of SAS PROC GLM.

ANOVA	ANOVA table for the model
Type I	Type I sum of square table
Type II	Type II sum of square table
Type III	Type III sum of square table
Parameter	Parameter table with standard error, t value, p value

Author(s)

Kyun-Seop Bae k@acr.kr

Examples

```
GLM(uptake ~ Plant + Type + Treatment + conc, CO2)
```

lfit *Linear Fit*

Description

Fits a least square linear model.

Usage

```
lfit(x, y, eps=1e-8)
```

Arguments

x	a result of ModelMatrix
y	a column vector of response, dependent variable
eps	Less than this value is considered as zero.

Details

Minimum version of least square fit of a linear model

Value

coeffcients	beta coefficients
g2	g2 inverse
rank	rank of the model matrix
DFr	degree of freedom for the residual
SSE	sum of square error

Author(s)

Kyun-Seop Bae k@acr.kr

See Also

ModelMatrix

ModelMatrix *Model Matrix*

Description

This model matrix is similar to `model.matrix`. But it does not omit unnecessary columns.

Usage

```
ModelMatrix(Formula, Data, NOINT=FALSE, KeepOrder=FALSE)
```

Arguments

Formula	a conventional formula for a linear model
Data	a data.frame to be analyzed
NOINT	If NOINT is TRUE, no intercept model will be used. Always -1 or +0 will be ignored in the formula.
KeepOrder	If KeepOrder is TRUE, terms in Formula will be kept. This is for Type I SS.

Details

It makes the model(design) matrix for GLM.

Value

Model matrix and attributes similar to the output of model.matrix.

X	design matrix, i.e. model matrix
terms	detailed information about terms such as formula and labels
termsIndices	term indices
assign	assignemnt of columns for each terms in order, different way of expressing term indices

Author(s)

Kyun-Seop Bae k@acr.kr

REG *Regression of Linear Least Square, similar to SAS PROC REG*

Description

REG is similar to SAS PROC REG.

Usage

```
REG(Formula, Data, NOINT=FALSE, eps=1e-8, summarize=TRUE)
```

Arguments

Formula	a conventional formula for a linear model
Data	a data.frame to be analyzed
NOINT	If NOINT is TRUE, no intercept model will be used. Always -1 or +0 will be ignored in the formula.
eps	Less than this value is considered as zero.
summarize	If this is FALSE, REG returns just lfit result.

Details

It performs the core function of SAS PROC REG.

Value

The result is comparable to that of SAS PROC REG.

`Estimate`	point estimate of parameters, coefficients
`Std. Error`	standard error of the point estimate
`t value`	value for t distribution
`Pr(>|t|`	probability of larger than absolute t value from t distribution with residual's degree of freedom

If `summarize=FALSE`, `REG` returns;

`coeffcients`	beta coefficients
`g2`	g2 inverse
`rank`	rank of the model matrix
`DFr`	degree of freedom for the residual
`SSE`	sum of square error

Author(s)

Kyun-Seop Bae k@acr.kr

Examples

```
REG(uptake ~ Plant + Type + Treatment + conc, CO2)
```

---

regD *Regression of Conventional Way*

Description

`regD` provides rich diagnostics such as student residual, leverage(hat), Cook's D, studentized deleted residual, DFFITS, and DFBETAS.

Usage

```
regD(formula, data)
```

Arguments

`formula`	a conventional formula for a linear model
`data`	a `data.frame` to be analyzed

Details

It performs the conventional regression analysis. This does not use g2 inverse, therefore it cannot handle singular matrix. If the model(design) matrix is not full rank, use `REG` or less parameters.

Value

Coefficients	conventional coefficients summary with Wald statistics
Diagnostics	Diagnostics table for detecting outlier or influential/leverage points. This includes fitted (Predicted), residual (Residual), standard error of residual(se_resid), studentized residual(RStudent), hat(Leverage), Cook's D, studentized deleted residual(sdResid), DIFFITS, and COVRATIO.
DFBETAS	Column names are the names of coefficients. Each row shows how much each coefficient is affected by deleting the coressponding row of observation.

Author(s)

Kyun-Seop Bae k@acr.kr

Examples

```
regD(uptake ~ conc, CO2)
```

---

satt *Satterthwaite Approximation of Pooled Variance and Degree of Freedom*

Description

Calculates pooled variance and degree of freedom using Satterthwaite equation.

Usage

```
satt(ws, vars, dfs)
```

Arguments

ws	a vector of weights
vars	a vector of variances
dfs	a vector of degree of freedoms

Details

The input can be more than two variances.

Value

Variance	pooled variance
Df	degree of freedom

Author(s)

Kyun-Seop Bae k@acr.kr

SS *Sum of Square*

Description

Sum of squares with ANOVA.

Usage

```
SS(x, rx, L, eps=1e-8)
```

Arguments

x	a result of ModelMatrix containing design information
rx	a result of lfit
L	linear hypothesis, a full matrix matching the information in x
eps	Less than this value is considered as zero.

Details

It calculates sum of squares and completes the ANOVA table.

Value

ANOVA table	a classical ANOVA table without the residual(Error) part.

Author(s)

Kyun-Seop Bae k@acr.kr

See Also

ModelMatrix, lfit

T3MS *Type III Expected Mean Square Formula*

Description

Calculates a formula table for expected mean square of Type III SS.

Usage

```
T3MS(Formula, Data, L0, eps=1e-8)
```

Arguments

Formula	a conventional formula for a linear model
Data	a data.frame to be analyzed
L0	a matrix of row linear contrasts, if missed, e3 is used
eps	Less than this value is considered as zero.

Details

This is necessary for further hypothesis test of nesting factors.

Value

A coefficient matrix for Type III expected mean square

Author(s)

Kyun-Seop Bae k@acr.kr

Examples

```
T3MS(log(CMAX) ~ SEQ/SUBJ + PRD + TRT, BEdata)
```

T3test *Test Type III SS using error term other than MSE*

Description

Hypothesis test of Type III SS using an error term other than MSE. This corresponds to SAS PROC GLM's RANDOM /TEST clause.

Usage

```
T3test(Formula, Data, Error="", eps=1e-8)
```

Arguments

Formula	a conventional formula for a linear model
Data	a data.frame to be analyzed
Error	an error term. Term name should be exactly same one listed the ANOVA output.
eps	Less than this value is considered as zero.

Details

It tests a factor of type III SS using some other term as an error term. Here the error term should not be MSE.

Value

Returns one or more ANOVA table(s) of type III SS.

Author(s)

Kyun-Seop Bae k@acr.kr

Examples

```
T3test(log(CMAX) ~ SEQ/SUBJ + PRD + TRT, BEdata, "SEQ:SUBJ")
```

Index