Package 'supernova'

September 26, 2019

Type Package

Title Judd, McClelland, & Ryan Formatting for ANOVA Output

Version 2.1.1

Date 2019-09-20

Encoding UTF-8

LazyData yes

Description
Produces ANOVA tables in the format used by Judd, McClelland, and Ryan (2017, ISBN: 978-1138819832) in their introductory textbook, Data Analysis. This includes proportional reduction in error and formatting to improve ease the transition between the book and R.

License GPL-3

Depends R (>= 2.10)

Suggests car, dplyr, mosaic, Lock5withR, okcupiddata, fivethirtyeight, purrr, rlang, testthat

Imports magrittr, stringr

RoxygenNote 6.1.1

URL

BugReports

NeedsCompilation no

Author Adam Blake [aut],
Jeff Chrabaszcz [aut],
Ji Son [aut],
Jim Stigler [cre, aut]

Maintainer Jim Stigler <jstigler@gmail.com>

Repository CRAN

Date/Publication 2019-09-26 15:40:02 UTC

R topics documented:

---

estimate_extraction   *Extract Estimates/Statistics From a Model*

Description

This collection of functions is useful for extracting estimates and statistics from a fitted model. They are particularly useful when estimating many models, like when bootstrapping confidence intervals. Each function can be used with an already fitted model as an lm object, or a formula and associated data can be passed to it.

Usage

```
b0(object, data = NULL, ...)

b1(object, data = NULL, ...)

fVal(object, data = NULL, ...)

PRE(object, data = NULL, ...)

SSE(object, data = NULL, ...)

SSM(object, data = NULL, ...)

SSR(object, data = NULL, ...)
```

Arguments

object	A lm object, or formula.
data	If object is a formula, the data to fit the formula to as a data.frame.
...	Passthrough arguments to lm.

Details

b0 The intercept from the full model.

b1 The slope b1 from the full model.

fVal The F value from the full model.

PRE The Proportional Reduction in Error for the full model.

SSE The SS Error (SS Residual) from the model.

SSM The SS Model (SS Regression) for the full model.

SSR Alias for SSM.

Value

The value of the estimate as a single number.

Fingers *Data from introductory statistics students at a university.*

Description

Students at a university taking an introductory statistics course were asked to complete this survey as part of their homework.

Usage

Fingers

Format

A dataset with 157 observations on the following 16 variables.

Sex Sex of participant.

RaceEthnic Racial or ethnic background.

FamilyMembers Members of immediate family (excluding self).

SSLast Last digit of social security number (NA if no SSN).

Year Year in school: 1=First, 2=Second, 3=Third, 4=Fourth, 5=Other

Job Current employment status: 0=not working, 1=part-time job, 2=full-time job

MathAnxious Agreement with this statement "In general I tend to feel very anxious about mathematics": 2=Strongly Agree, 1=Agree, 0=Neither Agree nor Disagree, -1=Disagree, -2=Strongly Disagree

Interest Interest in statistics and the course: 2=very interested in course and statistics, 1=somewhat interested, 0=no interest, -1=dread the course.

GradePredict Prediction for final grade in the course from the university's grade points per unit: 4.0=A, 3.7=A-, 3.3=B+, 3.0=B, 2.7=B-, 2.3=C+, 2.0=C, 1.7=C-, 1.3=Below C-

Thumb Length in mm from tip of thumb to the crease between the thumb and palm.

Index Length in mm from tip of index finger to the crease between the index finger and palm.

Middle Length in mm from tip of middle finger to the crease between the middle finger and palm.

Ring Length in mm from tip of ring finger to the crease between the middle finger and palm.

Pinkie Length in mm from tip of pinkie finger to the crease between the pinkie finger and palm

Height Height in inches.

Weight Weight in pounds.

Fingers.messy *Data from introductory statistics students at a university.*

Description

Students at a university taking an introductory statistics course were asked to complete this survey as part of their homework.

Usage

```
Fingers.messy
```

Format

A dataset with 210 observations on the following 16 variables.

Sex Sex of participant: 1=female, 2=male, 3=prefer not to answer

RaceEthnic Racial or ethnic background: 1=White, 2=African American, 3=Asian, 4=Latino, 5=Other

FamilyMembers Members of immediate family (excluding self).

SSLast Last digit of social security number (NA if no SSN).

Year Year in school.

Job Current employment status.

MathAnxious Agreement with this statement "In general I tend to feel very anxious about mathematics."

Interest Interest in statistics and the course.

GradePredict Prediction for final grade in the course from the university's grade points per unit.

Thumb Length in mm from tip of thumb to the crease between the thumb and palm.

Index Length in mm from tip of index finger to the crease between the index finger and palm.

Middle Length in mm from tip of middle finger to the crease between the middle finger and palm.

Ring Length in mm from tip of ring finger to the crease between the middle finger and palm.

Pinkie Length in mm from tip of pinkie finger to the crease between the pinkie finger and palm

Height Height in inches.

Weight Weight in pounds.

---

generate_models	*Generate a List of Models for Computing Different Types of Sums of Squares*

Description

This function will return a list of lists where the top-level keys (names) of the items indicate the component of the full model (i.e. the term) that the generated models can be used to test. At each of these keys is a list with both the complex and simple models that can be compared to test the component. The complex models always include the target term, and the simple models are identical to the complex except the target term is removed. Thus, when the models are compared (e.g. using anova, except for Type III; see details below), the resulting values will show the effect of adding the target term to the model. There are three generally used approaches to determining what the appropriate comparison models should be, called Type I, II, and III. See the sections below for more information on these types.

Usage

```
generate_models(model, type)
```

Arguments

model	The model to generate the models from, of the type lm, aov, or formula.
type	The type of sums of squares to calculate:

  - Use 1, I, and sequential for Type I.
  - Use 2, II, and hierarchical for Type II.
  - Use 3, III, and orthogonal for Type III.

Value

A list of the augmented models for each term, where the associated term is the key for each model in the list.

Type I

For Type I SS, or sequential SS, each term is considered in order after the preceding terms are considered. Consider the example model Y ~ A + B + A:B, where ":" indicates an interaction. To determine the Type I effect of A, we would compare the model Y ~ A to the same model without the term: Y ~ NULL. For B, we compare Y ~ A + B to Y ~ A; and for A:B, we compare Y ~ A + B + A:B to Y ~ A + B. Incidentally, the anova function that ships with the base installation of R computes Type I statistics.

Type II

For Type II SS, or hierarchical SS, each term is considered in the presence of all of the terms that do not include it. For example, consider an example three-way factorial model Y ~ A + B + C + A:B + A:C + B:C + A:B:C, where ":" indicates an interaction. The effect of A is found by comparing Y ~ B + C + B:C + A to Y ~ B + C + B:C (the only terms included are those that do not include A). For B, the comparison models would be Y ~ A + C + A:C + B and Y ~ A + C + A:C; for A:B, the models would be Y ~ A + B + C + A:C + B:C + A:B and Y ~ A + B + C + A:C + B:C; and so on.

Type III

For Type III SS, or orthogonal SS, each term is considered in the presence of all of the other terms. For example, consider an example two-way factorial model Y ~ A + B + A:B, where ":" indicates an interaction. The effect of A is found by comparing Y ~ B + A:B + A to Y ~ B + A:B; for B, the comparison models would be Y ~ A + A:B + B and Y ~ A + A:B; and for A:B, the models would be Y ~ A + B + A:B and Y ~ A + B.

Unfortunately, anova() cannot be used to compare Type III models. anova() does not allow for violation of the principle of marginality, which is the rule that interactions should only be tested in the context of their lower order terms. When an interaction term is present in a model, anova() will automatically add in the lower-order terms, making a model like Y ~ A + A:B unable to be compared: it will add the lower-order term B,and thus use the model Y ~ A + B + A:B instead. To get the appropriate statistics for Type III comparisons, use drop1() with the full scope, i.e. drop1(model_fit,scope = . ~ .).

Examples

```
# create all type 2 comparison models
mod <- lm(Thumb ~ Height * Sex, data = Fingers)
mods_2 <- generate_models(mod, type = 2)

# compute the SS for the Height term
mod_Height <- anova(mods_2[["Height"]]$simple, mods_2[["Height"]]$complex)
mod_Height[["Sum of Sq"]][[2]]
```

Servers   *Servers data*

Description

Data about tips collected from an experiment with 44 servers.

Note: this dataset will be removed in future versions in favor of Tables.

Usage

```
Servers
```

Format

A dataset with 44 observations on the following 2 variables.

ServerID A number assigned to each server.

Tip How much the tip was.

---

supernova	*supernova*

Description

An alternative set of summary statistics for ANOVA. Sums of squares, degrees of freedom, mean squares, and F value are all computed with Type III sums of squares. This package adds proportional reduction in error, an explicit summary of the whole model, and separate formatting of p values and is intended to match the output used in Judd, McClelland, and Ryan (2017).

Usage

```
supernova(fit, type = 3)

superanova(fit, type = 3)
```

Arguments

fit	A fitted `lm` object
type	The type of sums of squares to calculate:

- 1, I, and `sequential` compute Type I SS.
- 2, II, and `hierarchical` compute Type II SS.
- 3, III, and `orthogonal` compute Type III SS.

Details

`superanova()` is an alias of `supernova()`

Value

An object of the class `supernova`, which has a clean print method for displaying the ANOVA table in the console as well as a named list:

tbl	The ANOVA table as a `data.frame`
fit	The original `lm` object being tested
models	Models created by `generate_models`

References

Judd, C. M., McClelland, G. H., & Ryan, C. S. (2017). *Data Analysis: A Model Comparison Approach to Regression, ANOVA, and Beyond* (3rd ed.). New York: Routledge. ISBN:879-1138819832

Examples

```
supernova(lm(Thumb ~ Weight, data = Fingers))
```

Survey	*Students at a university were asked to enter a random number between 1-20 into a survey.*

Description

Students at a university taking an introductory statistics course were asked to complete this survey as part of their homework.

Usage

```
Survey
```

Format

A dataset with 211 observations on the following 1 variable:

Any1_20 The random number between 1 and 20 that a student thought of.

Tables	*Tables data*

Description

Data about tips collected in an experiment from 44 different tables.

Usage

```
Tables
```

Format

A dataset with 44 observations on the following 2 variables.

TableID A number assigned to each table.

Tip How much the tip was.

Examples

```
str(Tables)
head(Tables)
```

TipExperiment   *Data from an experiment about smiley faces and tips*

Description

Tables were randomly assigned to receive checks that either included or did not include a drawing of a smiley face. Data was collected from 44 tables in an effort to examine whether the added smiley face would cause more generous tipping.

Usage

```
TipExperiment
```

Format

A dataset with 44 observations on the following 3 variables.

TableID A number assigned to each table.

Tip How much the tip was.

Condition Which experimental condition the table was randomly assigned to.

---

variables   *Extract the variables from a model*

Description

Extract the variables from a model

Usage

```
variables(object)
```

Arguments

object   An lm or supernova object

Value

A list containing the outcome and predictor variables in the model.

Index