Package ‘gginference’

March 21, 2020

Type Package
Title Visualise the Results of Inferential Statistics using 'ggplot2'
Version 0.1.1
Date 2020-03-21
Description Visualise the results of F test to compare two variances, Student's t-test, test of equal or given proportions, Pearson's chi-squared test for count data and test for association/correlation between paired samples.
Maintainer Kleanthis Koupidis <koupidis@okfn.gr>
URL https://github.com/okgreece/gginference
BugReports https://github.com/okgreece/gginference/issues
Depends R (>= 3.5.0)
License GPL-2 | file LICENSE
Encoding UTF-8
LazyData true
Imports ggplot2, rlang, stats, utils
RoxygenNote 7.0.2
Suggests knitr, MASS, rmarkdown
NeedsCompilation no
Author Charalampos Bratsas [aut],
       Anastasia Foudouli [aut],
       Kleanthis Koupidis [aut, cre]
Repository CRAN
Date/Publication 2020-03-21 05:40:06 UTC

R topics documented:

    accidentsData .................................................. 2
    BirthDeath .................................................... 3
    BirthDeath2000 ................................................. 3
accidentsData

Description

A data frame showing the use of seat belt and the driver status after a car accident in Greece.

Usage

accidentsData

Format

A data frame with 383 observations of 2 columns:

record  factor representing the driver status
seatBelt  factor indicating whether the driver wore a seatbelt

Source

The original data are available at Hellenic Statistical Authority
**Description**

A data frame containing the number of births and deaths along with their rates from 1932 to 2016.

**Usage**

`BirthDeath`

**Format**

A data frame with 71 observations of 5 columns:

- **Year**: years 1932-2016
- **Deaths**: number of deaths
- **DeathsRate**: number of deaths per 1000 citizen
- **Births**: number of births
- **BirthRate**: number of births per 1000 citizen

**Source**

The original data are available at Hellenic Statistical Authority

---

**Description**

A data frame containing samples with the number of births and deaths before and after 2000.

**Usage**

`BirthDeath2000`

**Format**

A data frame with 30 observations of 3 columns:

- **deaths**: number of deaths
- **births**: number of births
- **type**: factor indicating if the number of births and deaths correspond before 2000 or after 2000

**Source**

The original data are available at Hellenic Statistical Authority
**Births in Greece, 1976-1989**

**Description**

A data frame giving the number of births per 1000 people in Greece from 1976 to 1989.

**Usage**

births

**Format**

A data frame with 14 observations of 2 columns:

- year  years from 1976 to 1989
- rate  number of births per 1000 people

**Source**

The original data are available at Hellenic Statistical Authority

**Bio diesel and RON 95 consumption**

**Description**

A data frame including a sample of bio diesel and RON 95 consumption in Greece.

**Usage**

DieselbioRon95

**Format**

A data frame with 24 observations of 5 columns:

- region  factor of Greek regions
- DieselBio_consumption2006  metric tons of bio-diesel consumption in 2006
- DieselBio_consumption2016  metric tons of bio-diesel consumption in 2016
- RON95_consumption2006  metric tons of ron 95 consumption in 2006
- RON95_consumption2016  metric tons of ron 95 consumption in 2016

**Source**

The original data are available at Hellenic Statistical Authority
**Description**

A data frame containing the fuel consumption in Greece.

**Usage**

FuelConsumption

**Format**

A data frame with 50 observations of 8 columns:

 Geographic.area  factor with geographic area of Greece  
 Regions  factor with regions of Greece  
 Runits  factor with regional units of Greece  
 RON95  metric tons of ron 95 consumption  
 RON98_100  metric tons of ron 98 consumption  
 DieselBio  metric tons of bio diesel consumption  
 LPG  metric tons of liquefied petroleum gas consumption  
 DieselC  metric tons of heating oil consumption

**Source**

The original data are available at Hellenic Statistical Authority

---

### ggaov

**Description**

Visualise anova F-test to determine whether group means are equal

**Usage**

```r
  ggaov(t, alpha=0.05, colaccept="lightsteelblue1",  
  colreject="grey84", colstat="navyblue")
```
Arguments

- **t**: an object of class aov
- **alpha**: alpha level for finding critical F value
- **colaccept**: color for the acceptance region of the test
- **colreject**: color for the area of rejection of the test
- **colstat**: color of the statistic of the test line

Examples

```r
# 21-th day
chick21 <- ChickWeight[ChickWeight$Time == 21,]
chickaov <- aov(weight~Diet, data = chick21)
summary(chickaov)

ggaov(chickaov, colaccept = "grey89", colreject = "black")
```

---

**ggchisqtest**  
*Plot for Pearson’s Chi-squared Test for Count Data*

Description

Visualise chi-squared contingency table tests and goodness-of-fit tests.

Usage

```r
ggchisqtest(t, colaccept="lightsteelblue1", colreject="gray84", colstat="navyblue", alpha=0.05)
```

Arguments

- **t**: a list result of `chisq.test` of "htest" class
- **colaccept**: color the acceptance area of the test
- **colreject**: color for the rejection area of the test
- **colstat**: color for the test statistic vline
- **alpha**: default set to 0.05, choose confidence level for the plot as it is not stated in `chisqtest`
Examples

## Chi-squared test for given probabilities

```r
x <- c(A = 20, B = 15, C = 25)
chisq_test <- chisq.test(x)
chisq_test
ggchisqtest(chisq_test)
```

```r
x <- c(10, 86, 45, 38, 10)
p <- c(0.10, 0.40, 0.20, 0.20, 0.10)
chisq_test2 <- chisq.test(x, p = p)
chisq_test2
ggchisqtest(chisq_test2)
```

## Pearson's Chi-squared test

```r
library(MASS)
sex_smoke <- table(survey$Sex, survey$Smoke)
chisq_test3 <- chisq.test(sex_smoke)
chisq_test3
ggchisqtest(chisq_test3)
```

---

**ggcortest**  
*Plot test for association between paired samples*

**Description**

Visualise test for association between paired samples, using Pearson’s product moment correlation coefficient.

**Usage**

```r
ggcortest(t, colaccept="lightskyblue1", colreject="grey94", colstat="navy")
```

**Arguments**

- **t**: a list result of `cor.test` of "htest" class
- **colaccept**: color the acceptance area of the test
- **colreject**: color for the rejection area of the test
- **colstat**: color for the test statistic vline
Examples

corr_test <- cor.test(iris$Sepal.Length, iris$Sepal.Width)
corr_test
ggcortest(corr_test)

---

ggproptest  

Plot test of Equal or Given Proportions

Description

Visualise prop.test for testing the null that the proportions (probabilities of success) in several groups are the same, or that they equal certain given values.

Usage

ggproptest(t, alpha=0.05, colaccept="lightsteelblue1", colreject="gray84", colstat="navyblue")

Arguments

t  a list result of prop.test of "htest" class
alpha  alpha level for plotting distribution, when prop.test is used on more than 2 samples
colaccept  color the acceptance area of the test
colreject  color the rejection area of the test
colstat  color for the test statistic vline

Examples

x <- c(5, 8, 12)
y <- c(8, 9, 13)
pr_test <- prop.test(x, y)
pr_test  
ggproptest(pr_test)
**ggttest**  
*Student's t-test plot*

**Description**

Visualise one and/or two sample t-tests on vectors of data.

**Usage**

```r
ggttest(t, colaccept="lightsteelblue1", colreject="grey84", colstat="navyblue")
```

**Arguments**

- `t`: a list result of `t.test` of "htest" class
- `colaccept`: color the acceptance area of the test
- `colreject`: color for the rejection area of the test
- `colstat`: color for the test statistic vline

**Examples**

```r
t_test <- t.test(sleep$extra ~ sleep$group)
t_test
ggttest(t_test)

t_test2 <- t.test(x = 1:10, y = c(7:20))
t_test2
ggttest(t_test2)
```

---

**ggvartest**  
*F test plot*

**Description**

Visualise F test to compare two variances

**Usage**

```r
ggvartest(t, colaccept = "lightsteelblue1",
        colreject = "gray84", colstat = "navyblue")
```
Arguments

- `t`: a list result of `var.test` of "htest" class
- `colaccept`: color the acceptance area of the test, see `colors`
- `colreject`: color for the rejection area of the test
- `colstat`: color for the test statistic

Examples

```r
x <- rnorm(50, mean = 0, sd = 2)
y <- rnorm(30, mean = 1, sd = 1)
var_test <- var.test(x, y)
var_test
ggvartest(var_test)
```

Description

A data frame that contains sample ratings of 18 laptops, by three experts.

Usage

LaptopRates

Format

A data frame with 54 observations of 3 columns:

- `laptop`: laptop id, 1-18
- `expert`: a character of expert1,expert2,expert3 values
- `rating`: ratings-5 likert scale, 5 indicates a very good rate
**LivLab**

**Description**

A data frame containing a sample with the results of neuropsychological assessment before and after serious game intervention in the living lab, Thess-AHALL (Thessaloniki Active and Healthy Aging Living Lab) of Medical Physics Laboratory of Aristotle University of Thessaloniki.

**Usage**

LivLab

**Format**

A data frame with 10 observations of 2 columns:

- before  score in a neuropsychological test before serious game intervention
- after   score in a neuropsychological test after serious game intervention

**Source**

<http://aha-livinglabs.com/>

---

**MilkConsumption**

**Description**

A data frame containing a sample of the number of cow, sheep and goat milk bottles sold.

**Usage**

MilkConsumption

**Format**

A data frame with 13 observations of 3 columns:

- Cow.Milk  number of cow milk bottles
- Sheep.Milk number of sheep milk bottles
- Goat.Milk  number of goat milk bottles

**Source**

The original data are available at Hellenic Statistical Authority
Questionnaire

Description
A data frame with the volume of new stores by category for urban, suburban and rural areas.

Usage
m_anova

Format
A data frame with 54 observations of 4 columns:
Categories integer representing three categories
UrbanAreas coding for urban, suburban and rural areas
Month integer representing three months
Volume volume in cubic meters

Source
The original data are available at Hellenic Statistical Authority

Questionnaire

Description
A data frame containing a sample with the answers of students.

Usage
questionnaire

Format
A data frame with 50 observations of 8 columns:
gender A factor with the student gender
writing.hand A factor with the writing hand of the students(left, right)
fold.arm A factor with the top hand when the students fold their arms
pulse Integer with the pulse rate of students (beats per minute)
exercise A factor with the frequency the students exercises (none, some,frequently)
smoke A factor with the frequency the students smokes (heavy, regularly, occasionally, never)
height Integer with the height of the students (in centimeters)
age Integer with the age of the students
**Description**

A data frame with 128 sample results of a repeated experiment. Success is noted with 1 and failure with 0.

**Usage**

`randexperiment`

**Format**

A data frame with 128 observations of 1 column.


---

**Salary_Gender**

*Female and male salaries*


---

**Description**

A sample data frame with female and male salaries of a company.

**Usage**

`Salary_Gender`

**Format**

A data frame with 100 observations of 2 columns:

- `Male_sal` male salaries
- `Female_sal` female salaries
Index

accidentsData, 2
BirthDeath, 3
BirthDeath2000, 3
births, 4
chisq.test, 6
colors, 10
cor.test, 7
DieselbioRon95, 4
FuelConsumption, 5
ggaov, 5
ggchisqtest, 6
ggcortest, 7
ggproptest, 8
ggttest, 9
ggvartest, 9
LaptopRates, 10
LivLab, 11
m_anova, 12
MilkConsumption, 11
prop.test, 8
questionnaire, 12
randexperiment, 13
Salary_Gender, 13
t.test, 9
var.test, 10