Package ‘BFpack’

May 11, 2020

Type Package
Title Flexible Bayes Factor Testing of Scientific Expectations
Version 0.2.1
Date 2020-05-11
Maintainer Joris Mulder <j.mulder3@tilburguniversity.edu>
Description Implementation of various default Bayes factors for testing statistical hypotheses. The package is intended for applied quantitative researchers in the social and behavioral sciences, medical research, and related fields. The Bayes factor tests can be executed for statistical models such as univariate and multivariate normal linear models, generalized linear models, special cases of linear mixed models, survival models, relational event models. Parameters that can be tested are location parameters (e.g., regression coefficients), variances (e.g., group variances), and measures of association (e.g., bivariate correlations). The statistical underpinnings are described in
License GPL (>= 3)
URL https://github.com/jomulder/BFpack
BugReports https://github.com/jomulder/BFpack/issues
Encoding UTF-8
LazyData true
NeedsCompilation yes
RoxygenNote 6.1.1
Depends R (>= 3.0.0), bain
Imports stats, MASS, Matrix, mvtnorm, pracma, lme4
Suggests testthat, polycor, survival, pscl

Author Joris Mulder [aut, cre],
Caspar van Lissa [aut, ctb],
Xin Gu [aut],
Anton Olsson-Collentine [aut, ctb],
Florian Boeing-Messing [aut, ctb],
Donald R. Williams [aut, ctb],
Jean-Paul Fox [aut, ctb],
Janosch Menke [ctb],
Barry Brown [ctb],
James Lovato [ctb],
Kathy Russell [ctb],
Lapack 3.8 [ctb],
Jack Dongarra [ctb],
Jim Bunch [ctb],
Cleve Moler [ctb],
Gilbert Stewart [ctb],
John Burkandt [ctb],
Ashwith Rego [ctb],
Alexander Godunov [ctb],
Alan Miller [ctb],
Jean-Pierre Moreau [ctb],
The R Core Team [cph]

Repository CRAN
Date/Publication 2020-05-11 17:50:02 UTC

R topics documented:

actors ................................................................. 3
attention ............................................................. 4
bartlett_test ......................................................... 4
BF ................................................................. 6
cor_test .............................................................. 8
fmri ................................................................. 9
memory ............................................................... 10
relevents ............................................................ 10
same_building ...................................................... 11
same_division ....................................................... 12
same_hierarchy .................................................... 12
**Actors**

Information on 25 actors of a consultancy firm for which a sequence of e-mail messages is observed (can be accessed through the 'events' data object). The actor data is simulated based on information provided in Mulder & Leenders (2019). In the original data, 70 actors were involved. The current data is a random sample of 25 actors.

**Usage**

```r
data(actors)
```

**Format**

dataframe (25 rows, 4 columns)

<table>
<thead>
<tr>
<th>Field</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>actors$id</td>
<td>integer</td>
<td>ID of the employee, corresponding to the sender and receiver IDs in the events dataframe</td>
</tr>
<tr>
<td>actors$position</td>
<td>numeric</td>
<td>Hierarchical position of the employee, ranging from 1-4</td>
</tr>
<tr>
<td>actors$division</td>
<td>character</td>
<td>Categorical variable, indicating the division of the employee</td>
</tr>
<tr>
<td>actors$location</td>
<td>integer</td>
<td>Categorical variable, indicating the location of the building the employee works in</td>
</tr>
</tbody>
</table>

**Details**

The related data files 'events', 'same_building', 'same_division' and 'same_hierarchy' contain information on the event sequence and three event statistics respectively.

**Source**

doi:10.1016/j.chaos.2018.11.027

**References**

Description

Data from a psychological study comparing attentional performances of Tourette’s syndrome (TS) patients, ADHD patients, and controls. These data were simulated using the sufficient statistics from Silverstein, Como, Palumbo, West, and Osborn (1995).

Usage

data(attention)

Format

A data.frame with 51 rows and 2 columns.

Details

<table>
<thead>
<tr>
<th>accuracy</th>
<th>numeric</th>
<th>Participant’s accuracy in the attentional task</th>
</tr>
</thead>
<tbody>
<tr>
<td>group</td>
<td>factor</td>
<td>Participant’s group membership (TS patient, ADHD patient, or control)</td>
</tr>
</tbody>
</table>

Source

DOI:10.1037/0894-4105.9.2.157

References


Description

Performs Bartlett’s test of the null that the variances in each of the groups (samples) are the same.
**Usage**

```r
bartlett_test(x, g, ...)
```

### Default S3 method:

```r
bartlett_test(x, g, ...)
```

**Arguments**

- `x` a numeric vector of data values, or a list of numeric data vectors representing the respective samples, or fitted linear model objects (inheriting from class "lm").
- `g` a vector or factor object giving the group for the corresponding elements of `x`. Ignored if `x` is a list.
- `...` further arguments to be passed to or from methods.

**Details**

`x` must be a numeric data vector, and `g` must be a vector or factor object of the same length as `x` giving the group for the corresponding elements of `x`.

**Value**

A list with class "bartlett_h_test" containing the following components:

- `statistic` Bartlett's K-squared test statistic.
- `parameter` the degrees of freedom of the approximate chi-squared distribution of the test statistic.
- `p.value` the p-value of the test.
- `conf.int` a confidence interval for the mean appropriate to the specified alternative hypothesis.
- `method` the character string "Bartlett test of homogeneity of variances".
- `data.name` a character string giving the names of the data.
- `vars` the sample variances across groups (samples).
- `n` the number of observations per group (sample)

**Bain t_test**

In order to allow users to enjoy the functionality of bain with the familiar stats-function `bartlett.test`, we have had to make minor changes to the function `bartlett.test.default`. All rights to, and credit for, the function `bartlett.test.default` belong to the R Core Team, as indicated in the original license below. We make no claims to copyright and incur no liability with regard to the changes implemented in `bartlett_test`.

This the original copyright notice by the R core team: File src/library/stats/R/bartlett_test.R Part of the R package, https://www.R-project.org

Copyright (C) 1995-2015 The R Core Team
This program is free software; you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation; either version 2 of the License, or (at your option) any later version.

This program is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details.

A copy of the GNU General Public License is available at https://www.R-project.org/Licenses/

References


Examples

require(graphics)

plot(count ~ spray, data = InsectSprays)
bartlett_test(InsectSprays$count, InsectSprays$spray)

BF

Bayes factors for Bayesian exploratory and confirmatory hypothesis testing

Description

The BF function can be used for hypothesis testing and model selection using the Bayes factor. By default exploratory hypothesis tests are performed of whether each model parameter equals zero, is negative, or is positive. Confirmatory hypothesis tests can be executed by specifying hypotheses with equality and/or order constraints on the parameters of interest.

Usage

BF(x, hypothesis, prior, complement, ...)

Arguments

x
An R object containing the outcome of a statistical analysis.
hypothesis
A character string containing the informative hypotheses to evaluate. The default is NULL, which will result in an exploratory analysis.
prior
A vector specifying the prior probabilities of the hypotheses. The default is NULL which will specify equal prior probabilities.
complement
a logical specifying whether the complement should be added to the tested hypothesis under hypothesis.
...
Parameters passed to and from other functions.
Details

The function requires a fitted modeling object. Current analyses that are supported: `t_test`, `bartlett_test`, `aov`, `manova`, `lm`, `mlm`, `glm`, `hetcor`, `lmer`, `coxph`, `survreg`, `zeroinfl`, and `polr`.

For testing means and regression coefficients of model classes `t_test`, `aov`, `manova`, `lm`, `mlm`, adjusted fractional Bayes factors are computed. For testing group variances using `bartlett_test`, adjusted fractional Bayes factors are computed. For testing measures of association (e.g., correlations) under model class `mlm` and for testing intraclass correlations under model class `lmerMod`, default Bayes factors based on uniform priors are computed. For all other model classes an approximate Bayes factor is computed using a Gaussian approximation of the posterior, similar as a classical Wald test.

Value

The output is an object of class `BF`. The object has elements: `BFtu_exploratory`, `PHP_exploratory`, `BFtu_confirmatory`, `PHP_confirmatory`, `BFmatrix_confirmatory`, `BTable_confirmatory`, `BFtu_main`, `PHP_main`, `BFtu_interaction`, `PHP_interaction`, `prior`, `hypotheses`, `estimates`, `model`, `call`.

References


Examples

```r
# EXAMPLE 1. One-sample t test
ttest1 <- bain::t_test(therapeutic,mu=5)
print(ttest1)
# confirmatory Bayesian one sample t test
BF1 <- BF(ttest1,"mu=5")
summary(BF1)
# exploratory Bayesian one sample t test
BF(ttest1)

# EXAMPLE 2. ANOVA
aov1 <- aov(price ~ anchor*motivation,data=tvprices)
# check the names of the model parameters
names(aov1$coefficients)
BF1 <- BF(aov1,hypothesis="anchorrounded=motivationlow;
anchorrounded<motivationlow;
anchorrounded>motivationlow")
summary(BF1)

# EXAMPLE 3. Logistic regression
fit <- glm(sent ~ ztrust + zfWHR + zAfro + glasses + attract + maturity +
tattoos, family = binomial(), data = wilson)
```
BF1 <- BF(fit, hypothesis = "ztrust > (zfWHR, zAfro) > 0;  
ztrust > 0 & zfWHR = zAfro = 0")
summary(BF1)

# EXAMPLE 4. Correlation analysis
set.seed(123)
cor1 <- cor_test(memory[1:20,1:3])
BF1 <- BF(cor1)
summary(BF1)
BF2 <- BF(cor1, hypothesis="Wmn_with_Im > Wmn_with_Del > 0;  
Wmn_with_Im = Wmn_with_Del = 0")
summary(BF2)

---

**cor_test**  
*Bayesian correlation analysis*

**Description**

Estimate the unconstrained posterior for the correlations using a joint uniform prior.

**Usage**

cor_test(..., formula = NULL, iter = 5000)

**Arguments**

- `...` matrices (or data frames) of dimensions n (observations) by p (variables) for different groups (in case of multiple matrices or data frames).
- `formula` an object of class formula. This allows for including control variables in the model (e.g., ~ education).
- `iter` number of iterations from posterior (default is 5000).

**Value**

list of class cor_test:

- `meanF` posterior means of Fisher transform correlations
- `covmF` posterior covariance matrix of Fisher transformed correlations
- `correstimates` posterior estimates of correlation coefficients
- `corrdraws` list of posterior draws of correlation matrices per group
- `corrnames` names of all correlations
Examples

```r
# Bayesian correlation analysis of the 6 variables in 'memory' object
# we consider a correlation analysis of the first three variable of the memory data.
fit <- cor_test(BFpack::memory[,1:3])

# Bayesian correlation of variables in memory object in BFpack while controlling
# for the Cat variable
fit <- cor_test(BFpack::memory[,c(1:4)],formula = ~ Cat)

# Bayesian correlation analysis of first three variables in memory data
# for two different groups
HC <- subset(BFpack::memory[,c(1:3,7)], Group == "HC")[-4]
SZ <- subset(BFpack::memory[,c(1:3,7)], Group == "SZ")[-4]
fit <- cor_test(HC,SZ)
```

---

**fmri**

---

**fMRI data**

**Description**

fMRI data assessing relation between individual differences in the ability to recognize faces and cars and thickness of the superficial, middle, and deep layers of the fusiform face area, as assessed by high-resolution fMRI recognition (Williams et al, 2019, under review)

**Usage**

```r
data(fmri)
```

**Format**

A data.frame with 13 rows and 6 columns.

**Details**

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Subject</strong></td>
<td>numeric</td>
<td>Participant ID number</td>
</tr>
<tr>
<td><strong>Face</strong></td>
<td>numeric</td>
<td>Standardized score on face recognition battery</td>
</tr>
<tr>
<td><strong>Vehicle</strong></td>
<td>numeric</td>
<td>Standardized score on vehicle recognition battery</td>
</tr>
<tr>
<td><strong>Superficial</strong></td>
<td>numeric</td>
<td>Depth in mm of superficial layer of FFA</td>
</tr>
<tr>
<td><strong>Middle</strong></td>
<td>numeric</td>
<td>Depth in mm of middle layer of FFA</td>
</tr>
<tr>
<td><strong>Bform</strong></td>
<td>numeric</td>
<td>Depth in mm of deep layer of FFA</td>
</tr>
</tbody>
</table>
References


| memory | Memory data on health and schizophrenic patients |

Description

Data set from study assessing differences between schizophrenic patients and healthy control participants in patterns of correlations among 6 verbal memory tasks (Ichinose et al., 2019).

| Im   | numeric | Percent correct on immediate recall of 3 word lists |
| Del  | numeric | Percent correct on delayed recall of 3 word lists   |
| Wmn  | numeric | Number correct on letter-number span test of auditory working memory |
| Cat  | numeric | Number correct on category fluency task             |
| Fas  | numeric | Number correct on letter fluency task               |
| Rat  | numeric | Number correct on remote associates task            |
| Group| factor  | Participant Group (HC = Healthy Control; SZ = Schizophrenia) |

Usage

data(memory)

Format

A data.frame with 40 rows and 8 columns.

References


| relevents | A sequence of innovation-related e-mail messages |

Description

A time-ordered sequence of e-mail messages between employees of a consultancy firm and information on the actors in the relational event sequence. The data is originally analyzed by Mulder & Leenders (2019), to find drivers of innovation-related e-mail messages exchanged between employees of a large consultancy firm. Originally, the data consist of 2081 e-mail messages exchanged between 70 employees over the course of a year. The current data is a sample of a simulated data set, based on estimates of the model parameters in Mulder & Leenders (2019).
same_building

Usage

data(relevents)

Format

dataframe (227 rows, 3 columns)

relevents$time numeric Time of the e-mail message, in seconds since onset of the observation
relevents$sender integer ID of the sender, corresponding to the employee IDs in the actors dataframe
relevents$receiver integer ID of the receiver

Details

The related data files actors’, ‘same_building’, ‘same_division’ and ‘same_hierarchy’ contain information on the actors and three event statistics respectively.

Source

doi:10.1016/j.chaos.2018.11.027

References


<table>
<thead>
<tr>
<th>same_building</th>
<th>Same building event statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Description

A matrix coding whether senders of events (in the rows) and receivers of events (in the column) work in the same building. Related to the ’events’ data object, that contains a relational event sequence, and the ’actors’ object, that contains information on the 25 actors involved in the relational event sequence.

Usage

data(same_building)

Format

dataframe (25 rows, 4 columns)

<table>
<thead>
<tr>
<th>same_building</th>
<th>integer</th>
<th>Event statistic. Matrix with senders in the rows and receivers in the columns. The event statistic is</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Source

doi:10.1016/j.chaos.2018.11.027

References


<table>
<thead>
<tr>
<th>same_division</th>
<th>Same division event statistic</th>
</tr>
</thead>
</table>

Description

A matrix coding whether senders of events (in the rows) and receivers of events (in the column) work in the same division. Related to the `events` data object, that contains a relational event sequence, and the `actors` object, that contains information on the 25 actors involved in the relational event sequence.

Usage

data(same_division)

Format

dataframe (25 rows, 4 columns)

<table>
<thead>
<tr>
<th>same_division</th>
<th>integer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event statistic. Matrix with senders in the rows and receivers in the columns. The event statistic is</td>
<td></td>
</tr>
</tbody>
</table>

Source

doi:10.1016/j.chaos.2018.11.027

References


<table>
<thead>
<tr>
<th>same_hierarchy</th>
<th>Same hierarchical position event statistic</th>
</tr>
</thead>
</table>

---

same_hierarchy

---
Description
A matrix coding whether senders of events (in the rows) and receivers of events (in the column) work in the same hierarchical position. Related to the 'events' data object, that contains a relational event sequence, and the 'actors' object, that contains information on the 25 actors involved in the relational event sequence.

Usage
data(same_hierarchy)

Format
dataframe (25 rows, 4 columns)
same_hierarchy integer Event statistic. Matrix with senders in the rows and receivers in the columns. The event statistic is 1 if sender and receiver work in the same hierarchical position and 0 otherwise.

Source
doi:10.1016/j.chaos.2018.11.027

References

Description
Data from an experimental study, using the Wason selection task (Wason 1968) to examine whether humans have cognitive adaptations for detecting violations of rules in multiple moral domains. Moral domains are operationalized in terms of the five domains of the Moral Foundations Questionnaire (Graham et al. 2011). These data were simulated using the R-package synthpop, based on the characteristics of the original data.

Usage
data(sivan)

Format
A data.frame with 887 rows and 12 columns.

Details
Data come from an experimental study (Rosa, Rosa, Sarner, and Barrett, 1998) that were also used in Howell (2012, p.196). An experiment was conducted to investigate if Therapeutic Touch practitioners who were blindfolded can effectively identify which of their hands is below the experimenter’s. Twenty-eight practitioners were involved and tested 10 times in the experiment. Researchers expected an average of 5 correct answers from each practitioner as it is the number by chance if they do not outperform others.
timssICC

Format
A data.frame with 22 rows and 1 column.

References

timssICC  Trends in International Mathematics and Science Study (TIMSS) 2011-2015

Description
A stratified sample was drawn by country and school to obtain a balanced sample of $p = 15$ grade-4 students per school for each of four countries (The Netherlands (NL), Croatia (HR), Germany (DE), and Denmark (DK)) and two measurement occasions (2011, 2015). Achievement scores (first plausible value) of overall mathematics were considered. Performances of fourth and eight graders from more than 50 participating countries around the world can be found at (https://www.iea.nl/timss) The TIMSS achievement scale is centered at 500 and the standard deviation is equal to 100 scale score points. The TIMSS data set has a three-level structure, where students are nested within classrooms/schools, and the classrooms/schools are nested within countries. Only one classroom was sampled per school. Changes in the mathematics achievement can be investigated by examining the grouping of students in schools across countries. Changes in country-specific intraclass correlation coefficient from 2011 to 2015, representing heterogeneity in mathematic achievements within and between schools across years, can be tested. When detecting a decrease in average performance together with an increase of the intraclass correlation, a subset of schools performed worse. For a constant intraclass correlation across years the drop in performance applied to the entire population of schools. For different countries, changes in the intraclass correlation across years can be tested concurrently to examine also differences across countries.

Usage
data(timssICC)

Format
A data.frame with 16770 rows and 15 columns.

Details
<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>math</td>
<td>numeric</td>
<td>math score child</td>
</tr>
<tr>
<td>groupNL11</td>
<td>numeric</td>
<td>Indicator for child from NL in 2011</td>
</tr>
<tr>
<td>groupNL15</td>
<td>numeric</td>
<td>Indicator for child from NL in 2015</td>
</tr>
<tr>
<td>groupHR11</td>
<td>numeric</td>
<td>Indicator for child from HR in 2011</td>
</tr>
<tr>
<td>groupHR15</td>
<td>numeric</td>
<td>Indicator for child from HR in 2015</td>
</tr>
<tr>
<td>groupDE11</td>
<td>numeric</td>
<td>Indicator for child from DE in 2011</td>
</tr>
<tr>
<td>groupDE15</td>
<td>numeric</td>
<td>Indicator for child from DE in 2015</td>
</tr>
</tbody>
</table>
tvprices

groupDR11 numeric Indicator for child from DK in 2011
groupDR15 numeric Indicator for child from DK in 2015
gender numeric Female=0, Male=1
weight numeric Child sampling weight
yeargender numeric Interaction for occasion and gender
ln numeric total number of children in school-class
groupschool factor Nested indicator for school in country
schoolID factor Unique indicator for school

References

Description
Data from an experimental study where participants have to guess the price of a plasma tv. There were two experimental conditions. These data were simulated using the sufficient statistics from Janiszewski & Uy (2008).

Usage
data(tvprices)

Format
A data.frame with 59 rows and 3 columns.

Details

<table>
<thead>
<tr>
<th>price</th>
<th>numeric</th>
<th>Participant z-scores of price</th>
</tr>
</thead>
<tbody>
<tr>
<td>anchor</td>
<td>factor</td>
<td>Participant anchor</td>
</tr>
<tr>
<td>motivation</td>
<td>factor</td>
<td>motivation to change</td>
</tr>
</tbody>
</table>

Source
DOI:10.1111/j.1467-9280.2008.02057.x
References


Facial trustworthiness and criminal sentencing

Description

Data from a correlational study in which the correlation between ratings of facial trustworthiness of inmates was correlated with whether they had received the death penalty or not (Wilson and Rule, 2015). These data were simulated using the R-package synthpop, based on the characteristics of the original data.

Usage

data(wilson)

Format

A data.frame with 742 rows and 13 columns.

Details

<table>
<thead>
<tr>
<th>Variable</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>stim</td>
<td>integer</td>
<td>Stimulus Number</td>
</tr>
<tr>
<td>sent</td>
<td>integer</td>
<td>Sentence: 1 = Death, 0 = Life</td>
</tr>
<tr>
<td>race</td>
<td>integer</td>
<td>Race: 1 = White, -1 = Black</td>
</tr>
<tr>
<td>glasses</td>
<td>integer</td>
<td>Glasses: 1 = Yes, 0 = No</td>
</tr>
<tr>
<td>tattoos</td>
<td>integer</td>
<td>Tattoos: 1 = Yes, 0 = No</td>
</tr>
<tr>
<td>ztrust</td>
<td>numeric</td>
<td>Trustworthiness</td>
</tr>
<tr>
<td>trust_2nd</td>
<td>numeric</td>
<td>Trustworthiness ratings with 2nd control group; Death targets are same as in primary analysis, Life targets are different.</td>
</tr>
<tr>
<td>afro</td>
<td>numeric</td>
<td>Raw Afrocentricity ratings.</td>
</tr>
<tr>
<td>zAfro</td>
<td>numeric</td>
<td>Afrocentricity ratings normalized within target race. Analyses in paper were done with this variable.</td>
</tr>
<tr>
<td>attract</td>
<td>numeric</td>
<td>Attractiveness</td>
</tr>
<tr>
<td>fWHR</td>
<td>numeric</td>
<td>Facial width-to-height</td>
</tr>
<tr>
<td>afWHR</td>
<td>numeric</td>
<td>fWHR normalized within target race. Analyses in paper were done with this variable</td>
</tr>
<tr>
<td>maturity</td>
<td>numeric</td>
<td>Maturity</td>
</tr>
</tbody>
</table>

Source

DOI:10.1177/0956797615590992
References

DOI:10.1177/0956797615590992
Index

*Topic datasets
  actors, 3
  attention, 4
  fmri, 9
  memory, 10
  relevents, 10
  same_building, 11
  same_division, 12
  same_hierarchy, 12
  sivan, 13
  therapeutic, 14
  timssICC, 15
  tvprices, 16
  wilson, 17

actors, 3
aov, 7
attention, 4

bartlett_test, 4, 7
BF, 6
cor_test, 8
coxph, 7

fmri, 9
formula, 8
glm, 7
hetcor, 7
lm, 7
lmer, 7

manova, 7
memory, 10
polr, 7
relevents, 10

same_building, 11
same_division, 12
same_hierarchy, 12
sivan, 13
survreg, 7
t_test, 7
therapeutic, 14
timssICC, 15
tvprices, 16
wilson, 17
zeroinfl, 7