

Package ‘bayesianova’

May 31, 2020

Type Package

Title Bayesian Inference in the Analysis of Variance via Markov Chain Monte Carlo in Gaussian Mixture Models

Version 1.4

Date 2020-05-30

Author Riko Kelter

Maintainer Riko Kelter <riko.kelter@uni-siegen.de>

Description Provides a Bayesian version of the analysis of variance based on a three-component Gaussian mixture for which a Gibbs sampler produces posterior draws. For details about the Bayesian ANOVA based on Gaussian mixtures, see Kelter (2019) <arXiv:1906.07524>.

License GPL-2

Suggests MCMCpack, coda, knitr

NeedsCompilation no

Repository CRAN

Date/Publication 2020-05-31 01:40:10 UTC

R topics documented:

bayesANOVA-package	2
anovaplot	2
assumption.check	4
bayes.anova	5
post.pred.check	6

Index	8
--------------	----------

bayesANOVA-package *Bayesian ANOVA*

Description

Provides a Bayesian version of the analysis of variance (ANOVA) based on a three-component Gaussian mixture, for which a Gibbs sampler produces the posteriors of the means and standard deviation of each component. Also, model assumptions can be checked and results visualised.

Details

The DESCRIPTION file: This package was not yet installed at build time.

Index: This package was not yet installed at build time.

The core function is `bayes.anova` which provides the Bayesian version of the ANOVA. Also, assumptions can be checked via `assumption.check`, and `anovaplot` produces visualizations of the results.

Author(s)

Riko Kelter

Maintainer: Riko Kelter

References

For details, see: <https://arxiv.org/abs/1906.07524v1>

Examples

```
set.seed(42)
x1=rnorm(75,0,1)
x2=rnorm(75,1,1)
x3=rnorm(75,2,1)
assumption.check(x1,x2,x3,conf.level = 0.95)
result=bayes.anova(n=1000,first=x1,second=x2,third=x3)
anovaplot(result)
```

anovaplot

anovaplot

Description

Plots the results of a Bayesian ANOVA

Usage

```
anovaplot(dataframe, type="rope", sd="sd", ci=0.95)
```

Arguments

dataframe	A dataframe which is the result of a Bayesian ANOVA
type	Selects the type of plot which should be produced. The default is "rope", which produces a posterior region of practical equivalence (ROPE) analysis and posterior distributions for the three effect sizes of interest. Other options are to produce posteriors for each parameter via <code>type=="pars"</code> , or to produce posteriors for the difference of means and variances via <code>type=="diff"</code> .
sd	Selects if the results include posterior draws for the standard deviation (default) or the variance. <code>sd="sd"</code> is the default. <code>sd="var"</code> assumes that the dataframe includes posterior draws of the variances of each group.
ci	The credible level used for producing credible intervals. The default is <code>ci=0.95</code> .

Value

Produces plots of the results depending on which type is selected.

Author(s)

Riko Kelter

References

For details, see: <https://arxiv.org/abs/1906.07524v1>

Examples

```
set.seed(42)
x1=rnorm(75,0,1)
x2=rnorm(75,1,1)
x3=rnorm(75,2,1)
result=bayes.anova(n=1000,first=x1,second=x2,third=x3)

anovaplot(result)
anovaplot(result, type="effect")

x4=rnorm(75,3,1)
result2=bayes.anova(n=1000,first=x1,second=x2,third=x3,fourth=x4)
anovaplot(result2)
```

assumption.check *assumption.check*

Description

This function checks the assumption of normality for each of the groups `x1`, `x2`, `x3` (and optional `x4`, `x5` and `6`) used in the Bayesian ANOVA via Shapiro-Wilk tests with confidence level `conf.level`.

Usage

```
assumption.check(x1, x2, x3, x4=NULL, x5=NULL, x6=NULL, conf.level=0.95)
```

Arguments

<code>x1</code>	Numerical vector containing the values for the first group
<code>x2</code>	Numerical vector containing the values for the second group
<code>x3</code>	Numerical vector containing the values for the third group
<code>x4</code>	Numerical vector containing the values for the fourth group. Defaults to <code>NULL</code> .
<code>x5</code>	Numerical vector containing the values for the fifth group. Defaults to <code>NULL</code> .
<code>x6</code>	Numerical vector containing the values for the sixth group. Defaults to <code>NULL</code> .
<code>conf.level</code>	Confidence level of the Shapiro-Wilk test used. Significance level equals $1 - \text{conf.level}$

Details

If a single Shapiro-Wilk test fails, the method returns a warning and recommends to use further diagnostics.

Value

Histograms and Quantile-Quantile plots for all groups are produced, either a warning or a confirmation of normality in all three groups is printed to the console.

Author(s)

Riko Kelter

References

For details, see: <https://arxiv.org/abs/1906.07524v1>

Examples

```
set.seed(42)
x1=rnorm(75,0,1)
x2=rnorm(75,1,1)
x3=rnorm(75,2,1)

assumption.check(x1,x2,x3,conf.level = 0.95)
```

bayes.anova	<i>bayes.anova</i>
-------------	--------------------

Description

This function runs a Bayesian analysis of variance (ANOVA) on the data. The Bayesian ANOVA model assumes normally distributed data in all three groups, and conducts inference based on a Gibbs sampler in a three-component Gaussian-mixture with unknown parameters.

Usage

```
bayes.anova(n=10000, first, second, third, fourth=NULL, fifth=NULL, sixth=NULL,
hyperpars="custom", burnin=n/2, sd="sd", q=0.1, ci=0.95)
```

Arguments

n	Number of posterior draws the Gibbs sampler produces. The default is n=10000.
first	Numerical vector containing the values of the first group
second	Numerical vector containing the values of the second group
third	Numerical vector containing the values of the third group
fourth	Numerical vector containing the values of the fourth group. Default value is NULL.
fifth	Numerical vector containing the values of the fifth group. Default value is NULL.
sixth	Numerical vector containing the values of the sixth group. Default value is NULL.
hyperpars	Sets the hyperparameters on the prior distributions. Two options are provided. The default is "custom", and the other is "rafterys". For details, see the references.
burnin	Burn-in samples for the Gibbs sampler
sd	Selects if posterior draws should be produced for the standard deviation (default) or the variance. The two options are "sd" and "var" respectively.
q	Tuning parameter for the hyperparameters. The default is q=0.1 and it is recommended not to change this.
ci	The credible level for the credible intervals produced. Default is ci=0.95.

Details

The Gibbs sampler is run with four Markov chains to run convergence diagnostics.

Value

Returns a dataframe which includes four columns for each parameter of interest. Each column corresponds to the posterior draws of a single Markov chain obtained by the Gibbs sampling algorithm.

Author(s)

Riko Kelter

ReferencesFor details, see: <https://arxiv.org/abs/1906.07524v1>**Examples**

```

set.seed(42)
x1=rnorm(75,0,1)
x2=rnorm(75,1,1)
x3=rnorm(75,2,1)
x4=rnorm(75,-1,1)
result=bayes.anova(first=x1,second=x2,third=x3)
result=bayes.anova(n=1000,first=x1,second=x2,third=x3,
hyperpars="custom",burnin=750,ci=0.99,sd="sd")
result2=bayes.anova(n=1000,first=x1,second=x2,third=x3,
fourth=x4)

```

`post.pred.check`*post.pred.check*

Description

Provides a posterior predictive check for a fitted Bayesian ANOVA model.

Usage`post.pred.check(anovafit, ngroups, out, reps = 50, eta)`**Arguments**

<code>anovafit</code>	A dataframe returned by <code>bayes.anova</code>
<code>ngroups</code>	An integer which is the number of groups used in the ANOVA
<code>out</code>	A numerical vector containing the originally observed data in all groups
<code>reps</code>	An integer which is the number of posterior predictive distributions sampled from the ANOVA models posterior distribution. Defaults to 50 sampled parameters.
<code>eta</code>	A numerical vector containing the weight values of the mixture.

Details

Provides a posterior predictive check for a fitted Bayesian ANOVA model.

Value

Produces a plot consisting of a density estimate of the original data and posterior predictive distributions sampled from the posterior of the Bayesian ANOVA model as density overlays.

Author(s)

Riko Kelter

Examples

```
set.seed(700)
x1=rnorm(1000,0,1)
x2=rnorm(1000,1,1)
x3=rnorm(1000,2,2)

result=bayes.anova(n=1000,first = x1, second=x2, third=x3)
post.pred.check(result, ngroups = 3, out = c(x1,x2,x3), reps = 25, eta = c(1/3,1/3,1/3))
```

Index

*Topic **package**

 bayesANOVA-package, [2](#)

anovaplot, [2](#)

assumption.check, [4](#)

bayes.anova, [5](#)

bayesANOVA (bayesANOVA-package), [2](#)

bayesANOVA-package, [2](#)

post.pred.check, [6](#)