Package ‘BEST’

Type     Package
Title    Bayesian Estimation Supersedes the t-Test
Version  0.5.1
Depends  HDInterval
Imports  coda, rjags
Date     2018-04-19
Author   John K. Kruschke and Mike Meredith
Maintainer Mike Meredith <mmeredith@wcs.org>
Description An alternative to t-tests, producing posterior estimates for group means and standard deviations and their differences and effect sizes.
License  GPL (>= 3)
NeedsCompilation no
Repository CRAN
Date/Publication 2018-04-19 19:30:13 UTC

R topics documented:

<table>
<thead>
<tr>
<th>Function</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>BEST-package</td>
<td>2</td>
</tr>
<tr>
<td>BESTmcmc</td>
<td>3</td>
</tr>
<tr>
<td>BESTpower</td>
<td>6</td>
</tr>
<tr>
<td>makeData</td>
<td>9</td>
</tr>
<tr>
<td>pairs.BEST</td>
<td>11</td>
</tr>
<tr>
<td>plot.BEST</td>
<td>12</td>
</tr>
<tr>
<td>plotAll</td>
<td>13</td>
</tr>
<tr>
<td>plotAreaInROPE</td>
<td>14</td>
</tr>
<tr>
<td>plotPost</td>
<td>16</td>
</tr>
<tr>
<td>plotPostPred</td>
<td>19</td>
</tr>
<tr>
<td>postPriorOverlap</td>
<td>20</td>
</tr>
<tr>
<td>print.BEST</td>
<td>21</td>
</tr>
<tr>
<td>summary.BEST</td>
<td>22</td>
</tr>
</tbody>
</table>

Index 25
**Description**

An alternative to \( t \) tests, producing posterior estimates for groups means and standard deviations and their differences and effect sizes. Bayesian estimation provides a much richer picture of the data, and can be summarised as point estimates and credible intervals.

**Details**

The core function, `BESTmcmc`, generates posterior distributions to compare the means of two groups, or to compare the mean of one group with a standard, taking into account the standard deviation(s). It is thus similar to a \( t \) test. However, our Bayesian approach results in probability statements about the values of interest, rather than \( p \)-values and significance levels.

In addition, the procedure accounts for departures from normality by using a \( t \)-distribution to model the variable of interest and estimating a measure of normality.

Functions to summarise and to visualise the output are provided.

The function `BESTpower` allows simulation-based estimates of power, either retrospective power directly with `BESTmcmc` output or prospective power analysis with `makeData`.

**Author(s)**

Original code by John K. Kruschke, [http://www.indiana.edu/~kruschke/BEST/](http://www.indiana.edu/~kruschke/BEST/)

Maintainer: Mike Meredith <mmeredith@wcs.org>

**References**


**Examples**

```r
## Comparison of two groups:
## ===============
y1 <- c(5.77, 5.33, 4.59, 4.33, 3.66, 4.48)
y2 <- c(3.88, 3.55, 3.29, 2.59, 2.33, 3.59)

# Run an analysis, takes up to 1 min.
BESTout <- BESTmcmc(y1, y2)

# Look at the result:
BESTout
```
```r
summary(BESTout)
plot(BESTout)
plot(BESTout, "sd")
plotPostPred(BESTout)
plotAll(BESTout, credMass=0.8, ROPEm=c(-0.1,0.1),
        ROPEeff=c(-0.2,0.2), compValm=0.5)
plotAll(BESTout, credMass=0.8, ROPEm=c(-0.1,0.1),
        ROPEeff=c(-0.2,0.2), compValm=0.5, showCurve=TRUE)
summary(BESTout, credMass=0.8, ROPEm=c(-0.1,0.1), ROPEsd=c(-0.15,0.15),
        ROPEeff=c(-0.2,0.2))
pairs(BESTout)

head(BESTout$mu1)
muDiff <- BESTout$mu1 - BESTout$mu2
mean(muDiff > 1.5)
mean(BESTout$sigma1 - BESTout$sigma2)
hist(BESTout$nu)

# Retrospective power analysis
# --------------------------
# This takes time, so we do 2 simulations here; a real analysis needs several hundred

powerRet <- BESTpower(BESTout, N1=length(y1), N2=length(y2),
                       ROPEm=c(-0.1,0.1), maxHDIWm=2.0, nRep=2)
powerRet
# We only set criteria for the mean, so results for sd and effect size are all NA.

## Analysis with a single group:
## -----------------------------
y0 <- c(1.89, 1.78, 1.30, 1.74, 1.33, 0.89)

# Run an analysis, takes up to 40 secs.
BESTout1 <- BESTmcmc(y0)
BESTout1
summary(BESTout1)
plot(BESTout1)

head(BESTout1$mu)
mean(BESTout1$sigma)
```

**BESTmcmc**

**Generate MCMC samples for posterior distributions**

**Description**

This function is the core of the BEST package. It calls JAGS and passes a description of the model, priors, and data, then retrieves and returns the MCMC samples for the parameters.
Usage

\[
\text{BESTmcmc}(y_1, y_2 = \text{NULL}, \text{priors} = \text{NULL}, \text{doPriorsOnly} = \text{FALSE},
\text{numSavedSteps} = 1e+05, \text{thinsteps} = 1, \text{burnInSteps} = 1000,
\text{verbose}=\text{TRUE}, \text{rnd.seed}=\text{NULL}, \text{parallel}=\text{NULL})
\]

Arguments

- **y1**: a numeric vector of data values.
- **y2**: a vector of values for a second group, or NULL if there is only one group of observations.
- **priors**: an optional list of values controlling the priors, see Details.
- **doPriorsOnly**: if TRUE, BESTmcmc returns MCMC chains representing the prior distributions, not the posterior distributions for your data set.
- **numSavedSteps**: the number of MCMC observations to be returned.
- **thinsteps**: thinning rate. If set to n > 1, n steps of the MCMC chain are calculated for each one returned. This is useful if autocorrelation is high.
- **burnInSteps**: number of steps to discard as burn-in at the beginning of the chain.
- **verbose**: if FALSE, output to the R Console is suppressed. If chains are run in parallel, the output from JAGS is not displayed in the Console, even if verbose = TRUE.
- **rnd.seed**: a positive integer (or NULL): the seed for the random number generator, used to obtain reproducible samples if required. Values generated in different versions of BEST or different versions of JAGS may differ, even with the same seed.
- **parallel**: if NULL or TRUE and > 3 cores are available, the MCMC chains are run in parallel. (If TRUE and < 4 cores are available, a warning is given.)

Details

The function generates vectors of random draws from the posterior distributions of the group means (\(\mu\)) and standard deviations (\(\sigma\)), as well as the measure of normality (\(\nu\)). The procedure uses a Bayesian MCMC process implemented in JAGS (Plummer 2003).

If **priors** = NULL, broad priors as described by Kruschke (2013) are used. For \(\mu\), \(\text{Normal(mean}(y), 1000 \times \text{sd}(y))\); for \(\sigma\), \(\text{Uniform}(\text{sd}(y)/1000, \text{sd}(y) \times 1000)\); for \(\nu\), \(\text{Exponential}(1/29) + 1\), with the constraint that \(nu >= 1\). Here \(y = c(y_1, y_2)\). Note that **priors** = NULL is not equivalent to **priors** = list().

Alternatively, **priors** can be a list with elements specifying the priors for one or more parameters: \(\mu\): population means have separate normal priors, with mean \(\text{muMean}\) and standard deviation \(\text{muSD}\); if not included in the list, default values of \(\text{muMean} = \text{mean}(y), \text{muSD} = \text{sd}(y)\times5\) are used; \(\sigma\): population standard deviations have separate gamma priors, with mode \(\text{sigmaMode}\) and standard deviation \(\text{sigmaSD}\); defaults are \(\text{sigmaMode} = \text{sd}(y), \text{sigmaSD} = \text{sd}(y)\times5\); \(\nu\): the normality parameter has a gamma prior with mean \(\text{numean}\) and standard deviation \(\text{nusd}\); defaults are \(\text{numean} = 30, \text{nusd} = 30\); versions before 0.4.0 constrained \(\nu\) to be >1.

If there are 2 groups of observations, \(\text{muM}, \text{muSD}, \text{sigmaMode}, \text{sigmaSD}\) may be vectors of length 2 or scalar; if scalar, the same value is used for each population.
The model is shown in the diagram below.

Derived parameters, including the differences in means or standard deviations, and effect sizes can be obtained from the results of the BESTmcmc run.

The output from BESTmcmc has class BEST, which has print, plot and summary methods. These permit the extraction and display of credible intervals and proportions of the posterior mass above or below values of interest.

**Value**

An object of class BEST inheriting from data.frame. If two samples are compared, the output has the following columns:

- **mu1, mu2**: simulated observations of means for each population
- **sigma1, sigma2**: simulated observations of standard deviations for each population
- **nu**: simulated observations of normality parameter

while for a single sample, the columns are **mu, sigma, nu**.

The output has the following attributes:

- **call**: the call to the function.
- **Rhat**: the 'potential scale reduction factor'.
- **n.eff**: sample size adjusted for autocorrelation.
- **data**: a list with elements y1 and y2 containing the original data; y2 may be NULL.
- **priors**: a list with the priors used, if the priors argument is not NULL.
- **doPriorsOnly**: logical, the value of the doPriorsOnly argument.

The package provides print, plot and summary methods for BEST objects.
**Author(s)**

Original code by John K. Kruschke, modified by Mike Meredith.

**References**


For the informative priors, see Kruschke’s blog post at [http://doingbayesiandataanalysis.blogspot.com/2015/04/informed-priors-for-bayesian-comparison.html](http://doingbayesiandataanalysis.blogspot.com/2015/04/informed-priors-for-bayesian-comparison.html)

For the constraint on $\nu$, see the blog post at [http://doingbayesiandataanalysis.blogspot.com/2015/12/prior-on-df-normality-parameter-in-t.html](http://doingbayesiandataanalysis.blogspot.com/2015/12/prior-on-df-normality-parameter-in-t.html)


**See Also**

`plot`, `summary`, `pairs` for relevant methods.

**Examples**

```r
## See examples in BEST-package help.
```

---

**BESTpower**

*Estimating statistical power*

**Description**

Estimation of the probability of meeting the goals of a study given initial information or assumptions about the population parameters. For prospective power estimation, the sequence `makeData -> BESTmcmc -> BESTpower` is recommended: see `makeData`.

**Usage**

```r
BESTpower(BESTobj, N1, N2, credMass=0.95, ROPEm, ROPEsd, ROPEeff, maxHDIwM, maxHDIwSd, maxHDIwEff, compValm = 0, nRep = 200, mcmcLength = 10000, saveName = NULL, showFirstNrep = 0, verbose = 2, rnd.seed=NULL, parallel=NULL)
```
Arguments

**BESTobj**
- an object of class BEST produced by BESTmcmc.

**N1**
- planned sample size for the first (or only) group of observations; may be a scalar if sample size is fixed, or a vector if sample size varies; values will be recycled if necessary.

**N2**
- planned sample size for the second group of observations; ignored if BESTobj concerns only one group.

**credMass**
- the probability mass to include in HDIs when checking criteria.

**ROPEm**
- a two element vector, such as c(-1, 1), specifying the limit of the ROPE on the difference of means (for 2 groups) or the mean (for 1 group).

**ROPEsd**
- a two element vector, such as c(-1, 1), specifying the limit of the ROPE on the (difference of) standard deviations.

**ROPEeff**
- a two element vector, such as c(-1, 1), specifying the limit of the ROPE on the effect size.

**maxHDIwm**
- the maximum acceptable width for the HDI for the difference in means (for 2 groups) or for the mean (for a single group).

**maxHDIwsd**
- the maximum acceptable width for the HDI for the (difference of) standard deviation.

**maxHDIweff**
- the maximum acceptable width for the HDI for the effect size.

**compValm**
- for a single group, the value of the mean which represents no effect; used to calculate the effect size. Ignored for 2 groups.

**nRep**
- number of simulations to carry out.

**mcmcLength**
- length of the MCMC chains to use for each simulation.

**saveName**
- if required, the results may saved to a file after each iteration and saveName specifies the file name (or path relative to the current working directory) to use. The power object can be loaded with load. Set to NULL (the default) to disable saving.

**showFirstNrep**
- the number of results to display as plots at the beginning of the simulation run. (This uses dev.new(), which does not work in Rstudio. The plots will appear sequentially in the plot window and you will have to use the back arrow to review them.)

**verbose**
- controls output to the R Console: 0 suppresses all output; 1 gives just a progress bar; 2 gives maximum detail.

**rndSeed**
- a positive integer (or NULL): the seed for the random number generator, used to obtain reproducible samples if required.

**parallel**
- if NULL or TRUE and > 3 cores are available, the MCMC chains are run in parallel. (If TRUE and < 4 cores are available, a warning is given.)

Details

For each of the parameters of interest - (difference in) mean, (difference in) standard deviation and effect size - we consider 4 criteria and the probability that each will be met:
1. The HDI of the posterior density of the parameter lies entirely outside the ROPE and is greater than the ROPE.
2. The HDI of the posterior density of the parameter lies entirely outside the ROPE and is less than the ROPE.
3. The HDI of the posterior density of the parameter lies entirely inside the ROPE.
4. The width of the HDI is less than the specified `maxHDIWx`.

The mass inside the above HDIs depends on the `credMass` argument. A uniform beta prior is used for each of these probabilities and combined with the results of the simulations to give a conjugate beta posterior distribution. The means and 95% HDI credible intervals are returned.

**Value**

A matrix with a row for each criterion and columns for the mean and lower and upper limits of a 95% credible interval for the posterior probability of meeting the criterion.

Note that this matrix always has 12 rows. Rows corresponding to criteria which are not specified will have NAs.

**Note**

At least 1000 simulations are needed to get good estimates of power and these can take a long time. If the run is interrupted, the results so far can be recovered from the file specified in `saveName`.

The chains in `BESTobj` must have at least `nRep` values. To allow for some degree of autocorrelation among values, it would be prudent to make these chains at least 10 * `nRep` in length.

**Author(s)**

Original code by John Kruschke, modified by Mike Meredith.

**References**


**See Also**

`makedata` for details of preparing a `BESTobj` for a prospective power analysis.

**Examples**

```r
## For retrospective power analysis, see the example in BEST-package.

# 1. Generate idealised data set:
proData <- makedata(mu1=108, sd1=17, mu2=100, sd2=15, nPerGrp=20,
                   pcntOut=10, sdOutMult=2.0, rnd.seed=NULL)
```
makeData

Population parameter specification for a power analysis

Description

The function allows the analyst to prepare an idealized data set which exactly matches selected point values, and incorporates uncertainty in these values in terms of sample size.

Usage

makeData(mu1, sd1, mu2 = NULL, sd2 = NULL, nPerGrp,
          pcntOut = 0, sdOutMult = 2, rnd.seed = NULL, showPlot = TRUE)

Arguments

mu1 the mean for the first (or only) population.
sd1 the standard deviation for the main part of the first population, excluding outliers.
mu2 the mean for the second population; NULL if only one population is involved.
sd2 the standard deviation for the main part of the second population; NULL if only one population is involved.
nPerGrp sample size per group; large sample size reflects a high degree of precision in the values for the means and standard deviations.
pcntOut the percentage of outliers in each population.
sdOutMult the standard deviation of the outliers as a multiple of the standard deviation of the main part of the population.
rnd.seed a seed for the random number generator, used to obtain reproducible samples if required.
showPlot if TRUE, displays the results as a plot (see Details).
Details

The arguments to this function provide a framework to specify the hypothesised values of the parameters of the populations under study, while the sample size is chosen to reflect the confidence in the values specified.

The function produces idealized samples, i.e., samples which exactly match the specified means and standard deviations. If showPlot = TRUE, the results are displayed as a plot:

![Simulated Data](image)

*Histograms: actual sample values; red dashed line: distribution of the outliers; blue dashed line: distribution of the non-outliers; black line: combined distribution.*

These idealised samples are passed to BESTmcmc, which generates a series of sets of credible values for the parameters, including the normality parameter, taking account of correlations among them.

The sets of credible parameter values which constitute the BESTmcmc output are used by BESTpower to simulate new data sets which might arise during a subsequent experiment.

Value

A list with two components:

- `y1` A vector of simulated values for the first (or only) group.
- `y2` A vector of simulated values for the second group or NULL.

Author(s)

John Kruschke

References

pairs.BEST

See Also

BESTpower for examples.

Examples

## See examples for BESTpower.

---

pairs.BEST        Scatterplot matrix for a BEST object

Description

Function to produce a scatterplot matrix of a BEST object produced by BESTmcmc, with correlation coefficients in the lower triangle.

Usage

## S3 method for class 'BEST'
pairs(x, nPtToPlot = 1000, col = "skyblue", ...)

Arguments

x an object of class BEST
nPtToPlot number of points to plot
col colour to use for the points plotted.
... other graphical parameters passed to plot.default.

Value

None; used for its side effect.

Author(s)

Original code by John Kruschke, adapted as a pairs method by Mike Meredith

See Also

pairs in package graphics.

Examples

# See examples in BEST-package
plot.BEST

A plot method for objects of class 'BEST'

Description
Displays a plot showing the posterior probability distribution of one of the parameters of interest, the difference in means by default.

Usage
```r
# S3 method for class 'BEST'
plot(x, which = c("mean", "sd", "effect", "nu"), credMass = 0.95,
     ROPE = NULL, compVal = 0, showCurve = FALSE, ...)
```

Arguments
- `x`: an object of class BEST, as produced by the function `BESTmcmc`.
- `which`: one of "mean", "sd", "effect" or "nu" or an abbreviation of one of these; indicates which estimate to plot. For a comparison of two groups, "mean" and "sd" display the difference in means or standard deviation.
- `credMass`: the probability mass to include in credible intervals, or NULL to suppress plotting of the credible interval.
- `ROPE`: a two element vector, such as c(-1, 1), specifying the limit of the ROPE on the estimate; see Details.
- `showCurve`: logical: if TRUE, the posterior density will be represented by a kernel density function instead of a histogram.
- `compVal`: a value for comparison with the (difference of) parameters.
- `...`: other graphical parameters.

Details
The posterior distribution is shown as a histogram or density curve (if `showCurve = TRUE`), together with the Highest Density Interval. A ROPE and comparison value are also shown if appropriate.

The probability that the difference in means (or effect size, etc.) is precisely zero is zero. More interesting is the probability that the difference may be too small to matter. We can define a region of practical equivalence (ROPE) around zero, and obtain the posterior probability that the true value lies therein.

Value
Returns an object of class `histogram` invisibly. Used mainly for the side effect.

Author(s)
Mike Meredith, adapted from code by John Kruschke.
References


See Also

`plotAll` for a single plot showing all the parameters, `summary` for values of the corresponding summary statistics and `pairs` for a scatterplot matrix plot and correlation coefficients.

Examples

```r
# See examples in BEST-package.
```

`plotAll`  A series of plots displaying the results of 'BEST' analysis.

Description

Displays a series of plots showing the posterior probability distributions of the parameters of interest.

Usage

```r
plotAll(BESTobj, credMass = 0.95,  
ROPEm = NULL, ROPEsd = NULL, ROPEeff = NULL,  
compValm = 0, compValsd = NULL, compValeff = 0,  
showCurve = FALSE, ...)
```

Arguments

- **BESTobj**: an object of class BEST, as produced by the function `BESTmcmc`.
- **credMass**: the probability mass to include in credible intervals, or NULL to suppress plotting of the credible interval.
- **ROPEm**: a two element vector, such as c(-1, 1), specifying the limit of the ROPE on the difference of means (for 2 groups) or the mean (for 1 group).
- **ROPEsd**: a two element vector, such as c(-1, 1), specifying the limit of the ROPE on the (difference of) standard deviations.
- **ROPEeff**: a two element vector, such as c(-1, 1), specifying the limit of the ROPE on the effect size.
- **showCurve**: logical: if TRUE, the posterior density will be represented by a kernel density function instead of a histogram.
- **compValm**: a value for comparison with the (difference of) means.
- **compValsd**: a value for comparison with the (difference of) standard deviations.
- **compValeff**: a value for comparison with the effect size.
- **...**: other graphical parameters (currently ignored).
The display has a series of panels displaying the posterior distributions of each of the parameters (and differences between groups) together with summary statistics; see `plotPost` for details. Also a chart showing approx. 30 plots of posterior predictive distributions, together with histograms of the original data.

Returns NULL invisibly. Used for the side effect.

Code by John Kruschke, modified by Mike Meredith.


See Also

- `plot` for plots of individual parameters, `summary` for values of the corresponding summary statistics and `pairs` for a scatterplot matrix plot and correlation coefficients.

Examples

```r
# See examples in BEST-package.

plotAreaInROPE
```

Area of the posterior density in the ROPE as a function of its width.

Calculates and (optionally) plots the posterior probability mass included in the Region of Practical Equivalence (ROPE: see `plot.BEST`) as a function of the width of the ROPE.

```r
plotAreaInROPE(paramSampleVec, credMass = 0.95, compVal = 0, maxROPEradius, n = 201, plot = TRUE, ...)
```
**Arguments**

- `paramSampleVec` A vector of samples drawn from the target distribution; see Examples.
- `credMass` The probability mass to include in credible intervals.
- `compVal` a value for comparison with those plotted.
- `maxROPEradius` The maximum value of the ROPE radius (ie. half-width) to include in the plot.
- `n` The number of equally spaced points at which the area in the ROPE is to be estimated.
- `plot` If FALSE, the plot will be suppressed but the values will be returned.
- `...` Other graphical parameters.

**Details**

Defining a Region of Practical Equivalence (ROPE) allows decisions on whether a parameter is, for practical purposes, equivalent to a hypothetical null value, given a posterior probability density for the parameter. The null value may be considered credible if (A) 95% (say) of the probability mass lies within the ROPE, or (B) the 95% highest density interval (95% HDI) lies entirely within the ROPE.

How wide should the ROPE be? Different people at different times will have different ideas on the range of values equivalent to the null. The function `plotAreaInROPE` plots the probability mass lying within the ROPE for a range of widths (or rather radii or half-widths). It also shows the radius at which the HDI falls entirely within the ROPE.

**Value**

Returns invisibly a list with elements:

- `x` A vector of ROPE radii from 0 to `maxROPEradius`.
- `y` The corresponding proportion of the posterior density included in the ROPE.

**Author(s)**

John K. Kruschke, with minor modifications by Mike Meredith.

**References**


**Examples**

```r
# Generate a fake MCMC posterior for effect size and plot it:
mcmcChain <- rnorm(50000, 0.03, 0.025)
plotPost(mcmcChain, compVal=0, ROPE=c(-0.1, 0.1))

# How does the mass within the ROPE vary with ROPE radius?
plotAreaInROPE(mcmcChain, credMass = 0.95, compVal = 0,
               maxROPEradius = 0.15)
```
# Generate real MCMC chains, takes up to 1 min:
y1 <- c(4.77, 4.33, 3.59, 3.33, 2.66, 3.48)
y2 <- c(3.88, 3.55, 3.29, 2.59, 2.33, 3.59)
BESTout <- BESTmcmc(y1, y2)
plot(BESTout)

meanDiff <- BESTout$mu1 - BESTout$mu2
plotAreaInROPE(meanDiff, credMass = 0.95, compVal = 0,
               maxROPeradius = 3)

---

**plotPost**

*Graphic display of a posterior probability distribution*

**Description**

Plot the posterior probability distribution for a single parameter from a vector of samples, typically from an MCMC process, with appropriate summary statistics.

**Usage**

```r
plotPost(paramSampleVec, credMass = 0.95, compVal = NULL, ROPE = NULL,
         HDItextPlace = 0.7, showMode = FALSE, showCurve = FALSE, ...)
```

**Arguments**

- **paramSampleVec**: A vector of samples drawn from the target distribution.
- **credMass**: the probability mass to include in credible intervals, or NULL to suppress plotting of credible intervals.
- **compVal**: a value for comparison with those plotted.
- **ROPE**: a two element vector, such as `c(-1, 1)`, specifying the limits of the Region Of Practical Equivalence.
- **HDItextPlace**: a value in [0,1] that controls the horizontal position of the labels at the ends of the HDI bar.
- **showMode**: logical: if TRUE, the mode is displayed instead of the mean.
- **showCurve**: logical: if TRUE, the posterior density will be represented by a kernel density function instead of a histogram.
- **...**: graphical parameters and the breaks parameter for the histogram.
The data are plotted either as a histogram (above) or, if `showCurve = TRUE`, as a fitted kernel density curve (below). Either the mean or the mode of the distribution is displayed, depending on the parameter `showMode`. The Highest Density Interval (HDI) is shown as a horizontal bar, with labels for the ends of the interval.
If values for a ROPE are supplied, these are shown as dark red vertical dashed lines, together with the percentage of probability mass within the ROPE. If a comparison value (compVal) is supplied, this is shown as a vertical green dotted line, together with the probability mass below and above this value.

Value

Returns an object of class `histogram` invisibly. Used for its plotting side-effect.

Author(s)

John Kruschke, modified by Mike Meredith

See Also

For details of the HDI calculation, see `hdi`.

Examples

```r
# Generate some data
tst <- rnorm(1e5, 3, 1)
plotPost(tst)
plotPost(tst, col='wheat', border='magenta')
```
plotPostPred

Plots for Posterior Predictive checks.

Description

Plots a number (default 30) of credible t-distributions based on posterior values of the mean, standard deviation, and normality for each group, together with histograms of the data.

Usage

plotPostPred(BESTobj, nCurvesToPlot = 30)

Arguments

BESTobj
an object of class BEST, as produced by the function BESTmcmc.
nCurvesToPlot
the number of posterior predictive curves to plot.

Value

Nothing, used for its side effect.

Author(s)

John Kruschke, modified by Mike Meredith.

References


Examples

## See examples in BEST-package.

```r
plotPost(tst, credMass=0.8, ROPE=c(-1,1), xlab="Response variable")
plotPost(tst, showMode=TRUE, showCurve=TRUE, compVal=5.5)

# For integers:
tst <- rpois(1e5, 12)
plotPost(tst)

# A severely bimodal distribution:
tst2 <- c(rnorm(1e5), rnorm(5e4, 7))
plotPost(tst2)  # A valid 95% CrI, but not HDI
plotPost(tst2, showCurve=TRUE)  # Correct 95% HDI
```
Description

Calculates and displays the overlap between a posterior distribution (as a vector of samples, typically from an MCMC process) and a prior distribution (as a vector of samples or as a function). Unidentifiable parameters will have high overlap: Gimenez et al (2009) suggest that overlap greater than 35% indicates weak identifiability.

Usage

postPriorOverlap(paramSampleVec, prior, ..., yaxt="n", ylab="", xlab="Parameter", main="", cex.lab=1.5, cex=1.4, xlim=range(paramSampleVec), breaks=NULL)

Arguments

paramSampleVec a vector of samples drawn from the target distribution.
prior either a vector of samples drawn from the prior distribution or the name for the density function of the distribution; standard R functions for this have a d- prefix, eg. dbeta. Arguments required by the function must be specified by their (abbreviated) names in the . . . argument; see the examples.

... named parameters to be passed to prior when it is a function.
yaxt a character which specifies the y axis type; the default, "n", suppresses plotting.
ylab text to use as the label of the y axis.
xlab text to use as the label of the x axis.
cex.lab the magnification to be used for x and y labels relative to the current setting of cex
cex a numerical value giving the amount by which plotting text and symbols should be magnified relative to the default
xlim a vector of length 2 giving the limits for the x axis.
main text to use as the main title of the plot
breaks controls the histogram break points or the number of bars; see hist.

Value

Returns the overlap, the area lying under the lower of the two density curves.

Author(s)

Mike Meredith
print.BEST

References


Examples

```r
# Generate some data
tst <- rbeta(1e6, 5, 7)

# check overlap with a Beta(0.2, 0.2) prior:
postPriorOverlap(tst, dbeta, shape1=0.2, shape2=0.2)

# check overlap with a Uniform(0, 1) prior:
postPriorOverlap(tst, runif(1e6))
```

---

print.BEST

*Printing a BEST object*

Description

Print method for objects of class BEST, such as produced by `BESTmcmc`.

Usage

```r
## S3 method for class 'BEST'
print(x, digits = 4, ...)
```

Arguments

- `x` an object of class BEST, as produced by `BESTmcmc`.
- `digits` the number of digits to print.
- `...` further arguments for the `print` function.

Details

The print method displays summary statistics for the parameters and two MCMC diagnostic measures:

- `rhat` is the 'potential scale reduction factor', which is 1 on convergence; if any parameter has a value > 1.05, rerun with increased `burnInSteps`. See `gelman.diag`.
- `n.eff` is the sample size adjusted for autocorrelation; for stable estimates of credible intervals this should be > 10,000. Rerun with increased `numSavedSteps` or increased `thinSteps`. See `effectiveSize`.
### Value
Returns \( x \) invisibly.

### Author(s)
Mike Meredith

### See Also
`BESTmcmc`.

### Examples
```r
## See examples in BEST-package help.
```

---

**summary.BEST**  
*Extract summary statistics from an object of class BEST.*

### Description
Provides summary statistics for each of the parameters (mean and standard deviation) of the group(s) of observations and their differences.

### Usage
```r
## S3 method for class 'BEST'
summary(object, credMass = 0.95,
ROPEm = NULL, ROPEsd = NULL, ROPEeff = NULL,
compValm = 0, compValsd = NULL, compValeff = 0, ...)
```

### Arguments
- `object`: an object of class BEST, as produced by the function `BESTmcmc`.
- `credMass`: the probability mass to include in credible intervals.
- `ROPEm`: a two element vector, such as `c(-1, 1)`, specifying the limit of the ROPE on the difference of means (for 2 groups) or the mean (for 1 group). See `plot.BEST` for an explanation of ROPE.
- `ROPEsd`: a two element vector, such as `c(-1, 1)`, specifying the limit of the ROPE on the (difference of) standard deviations.
- `ROPEeff`: a two element vector, such as `c(-1, 1)`, specifying the limit of the ROPE on the effect size.
- `compValm`: a value for comparison with the (difference of) means.
- `compValsd`: a value for comparison with the (difference of) standard deviations.
- `compValeff`: a value for comparison with the effect size.
- `...`: additional arguments for the summary or print function.
Value

Returns a matrix with the parameters in rows and the following columns:

- **mean**, **median**, **mode**: the mean, median and mode of the MCMC samples for the corresponding parameter.
- **hdi%**, **hdiLow**, **hdiHigh**: the percentage of posterior probability mass included in the highest density interval and the lower and upper limits.
- **compVal**, **%>compVal**: the value for comparison and the percentage of the posterior probability mass above that value.
- **ROPElow**, **ROPEhigh**, **%InROPE**: the lower and upper limits of the Region Of Practical Equivalence (ROPE) and the percentage of the posterior probability mass within the region.

If the analysis concerns a comparison of two groups, the matrix will have rows for:

- **mu1**, **mu2**, **muDiff**: the means of each group and the difference in means
- **sigma1**, **sigma2**, **sigmaDiff**: the standard deviations of each group and the difference in standard deviations
- **nu**, **log10nu**: the normality parameter and its log
- **effSz**: the effect size; \( d[a] \) from Macmillan & Creelman (1991).

For a single group, the rows will be:

- **mu**: the mean
- **sigma**: the standard deviation
- **nu**, **log10nu**: the normality parameter and its log
- **effSz**: the effect size.

Many of the elements of the matrix will be NA. The print method for the summary attempts to print this nicely.

Author(s)

Mike Meredith, based on code by John K. Kruschke.

References


See Also

Use the `plotAll` function for a graphical display of these same values.
Examples

## see "BEST-package"
Index

*Topic hplot
  plotAreaInROPE, 14
  plotPost, 16
  plotPostPred, 19
*Topic htest
  BEST-package, 2
*Topic methods
  summary.BEST, 22
*Topic package
  BEST-package, 2
*Topic print
  print.BEST, 21

BEST (BEST-package), 2
BEST-package, 2
BESTmcmc, 2, 3, 10–13, 19, 21, 22
BESTpower, 2, 6, 10, 11
effectiveSize, 21
gelman.diag, 21

dhi, 18
hist, 20

load, 7

makeData, 2, 6, 8, 9

pairs, 6, 11, 13, 14
pairs.BEST, 11
plot, 6, 14
plot.BEST, 12, 14, 22
plotAll, 13, 13, 23
plotAreaInROPE, 14
plotPost, 14, 16
plotPostPred, 19
postPriorOverlap, 20
print.BEST, 21

summary, 6, 13, 14
summary.BEST, 22