Package 'BEST'

December 28, 2015

Type Package

Title Bayesian Estimation Supersedes the t-Test

Version 0.4.0

Imports coda, jagsUI (>= 1.3.6)

Date 2015-12-19

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 2015-12-28 09:01:12

R topics documented:

BEST-package *Bayesian Estimation Supersedes the t Test*

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/

Maintainer: Mike Meredith <mmeredith at wcs.org>

References

Kruschke, J. K. 2013. Bayesian estimation supersedes the *t* test. *Journal of Experimental Psychology: General* 142(2):573-603. doi: 10.1037/a0029146

Kruschke, J. K. 2011. *Doing Bayesian data analysis: a tutorial with R and BUGS.* Elsevier, Amsterdam, especially Chapter 18.

Examples

```
## 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
```

```
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

```
BESTmcmc(y1, y2 = NULL, priors = NULL, doPriorsOnly = FALSE,
  numSavedSteps = 1e+05, thinSteps = 1, burnInSteps = 1000,
verbose=TRUE, rnd.seed=NULL, parallel=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.
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 may differ.
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.)

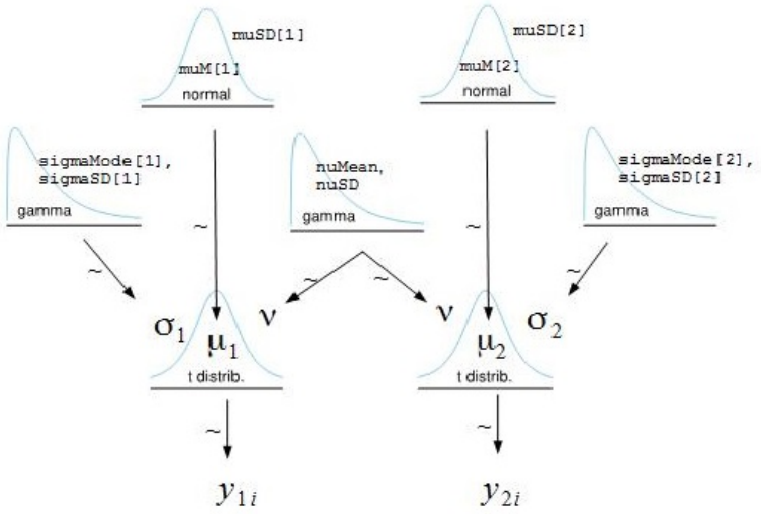
Details

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

If `priors = NULL`, broad priors as described by Kruschke (2013) are used. The user can specify priors for all parameters by including elements in the `priors` list:
μ : population means have separate normal priors, with mean `muM` and standard deviation `muSD`;
σ : population standard deviations have separate gamma priors, with *mode* `sigmaMode` and standard deviation `sigmaSD`;
ν : the normality parameter has a gamma prior with *mean* `nuMean` and standard deviation `nuSD`. (Versions before 0.4.0 constrained nu to be >1.)

If there are 2 groups of observations, `muM, muSD, sigmaMode, sigmaSD` may be vectors of length 2 or scalar, in which case the same value is used for each population. The default priors correspond to `list(muM = mean(y), muSD = sd(y)*5, sigmaMode = sd(y), sigmaSD = sd(y)*5, nuMean = 30, nuSD = 30)`, where `y = c(y1, y2)`.

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

Kruschke, J K. 2013. Bayesian estimation supersedes the *t* test. *Journal of Experimental Psychology: General* 142(2):573-603. doi: 10.1037/a0029146

Plummer, Martyn (2003). JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling, *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*, March 20-22, Vienna, Austria. ISSN 1609-395X

See Also

plot, summary, pairs for relevant methods.

Examples

```
## 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

```
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)
```

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

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

Kruschke, J. K. 2013. Bayesian estimation supersedes the *t* test. *Journal of Experimental Psychology: General* 142(2):573-603. doi: 10.1037/a0029146

Kruschke, J. K. 2011. *Doing Bayesian data analysis: a tutorial with R and BUGS.* Elsevier, Amsterdam, Chapter 13.

See Also

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

Examples

```
## 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)

# 2. Generate credible parameter values from the idealised data:
proMCMC <- BESTmcmc(proData$y1, proData$y2, numSavedSteps=2000)

# 3. Compute the prospective power for planned sample sizes:
# We'll  do just 5 simulations to show it works; should be several hundred.
N1plan <- N2plan <- 50
powerPro <- BESTpower(proMCMC, N1=N1plan, N2=N2plan,
               ROPEm=c(-1.5,1.5), ROPEsd=c(-2,2), ROPEeff=c(-0.5,0.5),
               maxHDIWm=15.0, maxHDIWsd=10.0, maxHDIWeff=1.0, nRep=5)
powerPro
```

hdi *Highest Density Interval*

Description

Calculate the highest density interval (HDI) for a given probability mass, see Details. The function is generic, with methods for a range of input objects.

Usage

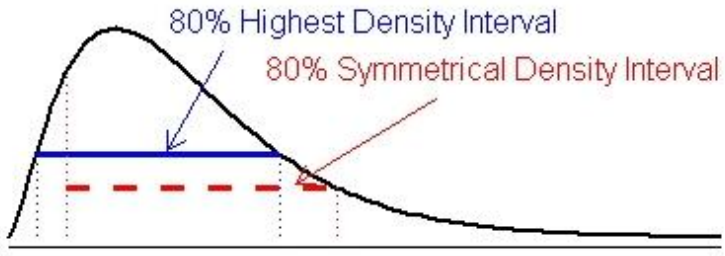
```
hdi(object, credMass = 0.95, ...)

## Default S3 method:
hdi(object, credMass = 0.95, ...)

## S3 method for class 'function'
hdi(object, credMass = 0.95, tol, ...)

## S3 method for class 'matrix'
hdi(object, credMass = 0.95, ...)

## S3 method for class 'data.frame'
hdi(object, credMass = 0.95, ...)

## S3 method for class 'density'
hdi(object, credMass = 0.95, allowSplit=FALSE, ...)

## S3 method for class 'mcmc.list'
hdi(object, credMass = 0.95, ...)
```

Arguments

`object`	an object specifying the target distribution; see Details.
`credMass`	a scalar [0, 1] specifying the mass within the credible interval.
`tol`	the desired accuracy; see `optimize`; default is 1e-8.
`allowSplit`	if TRUE and the HDI is discontinuous, the beginning and end of each segment are returned; see Value.
`...`	named parameters to be passed to other methods; see Examples.

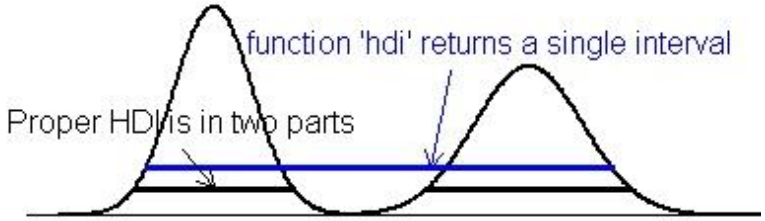
Details

The HDI is the interval which contains the required mass such that all points within the interval have a higher probability density than points outside the interval. When applied to a posterior probability density, it is often known as the Highest Posterior Density (HPD).



In contrast, a symmetric density interval defined by (eg.) the 10% and 90% quantiles may include values with lower probability than those excluded.

For a unimodal distribution, the HDI is the narrowest interval containing the specified mass, and the `hdi` function actually returns the narrowest interval. This does not always work properly for multimodel densities, where the HDI may be discontinuous (the horizontal black line in the Figure below). The single interval returned by `hdi` (the blue line) may incorrectly include values between the modes with low probability density. The `density` method with `allowSplit = TRUE` gives separate limits for discontinuous HDIs.



The default method expects a vector representing draws from the target distribution, such as is produced by an MCMC process. Missing values are silently ignored; if the vector has no non-missing values, NAs are returned.

The matrix and data frame methods expect an object with vectors of the above type for each parameter in columns. The result is a matrix with parameters in columns, and rows with the upper and lower limits of the HDI.

The mcmc.list method expects an object of type `mcmc.list` as defined in package **coda**.

None of the above use interpolation: the values returned correspond to specific values in the data object. Results thus depend on the random draws, and will be unstable if few values are provided. For a 95% HDI, 10,000 independent draws are recommended; a smaller number will be adequate for a 80% HDI, many more for a 99% HDI.

The function method requires the name for the inverse cumulative density function (ICDF) of the distribution; standard R functions for this have a `q-` prefix, eg. `qbeta`. Arguments required by the ICDF must be specified by their (abbreviated) names; see the examples.

Value

a vector of length 2 or a 2-row matrix with the lower and upper limits of the HDI, with an attribute "credMass".

The density method with allowSplit = TRUE produces a matrix with a row for each component of a discontinuous HDI and columns for begin and end. It has an additional attribute "height" giving the probability density at the limits of the HDI.

Author(s)

Mike Meredith. Code for hdi.function based on hpd by Greg Snow, corrected by John Kruschke.

References

Kruschke, J. K. 2011. *Doing Bayesian data analysis: a tutorial with R and BUGS.* Elsevier, Amsterdam, section 3.3.5.

Examples

```
# for a vector:
tst <- rgamma(1e5, 2.5, 2)
hdi(tst)
hdi(tst, credMass=0.8)
# For comparison, the symmetrical 80% CrI:
quantile(tst, c(0.1,0.9))

# Now a data frame:
tst <- data.frame(mu = rnorm(1e4, 4, 1), sigma = rlnorm(1e4))
hdi(tst, 0.8)
apply(tst, 2, quantile, c(0.1,0.9))

# For a function:
hdi(qgamma, 0.8, shape=2.5, rate=2)
# and the symmetrical 80% CrI:
qgamma(c(0.1, 0.9), 2.5, 2)

# A severely bimodal distribution:
tst2 <- c(rnorm(1e5), rnorm(5e4, 7))
hist(tst2, freq=FALSE)
(hdiMC <- hdi(tst2))
segments(hdiMC[1], 0, hdiMC[2], 0, lwd=3, col='red')
# This is a valid 95% CrI, but not a Highest Density Interval

dens2 <- density(tst2)
lines(dens2, lwd=2, col='blue')
(hdiD1 <- hdi(dens2))  # default allowSplit = FALSE; note the warning
segments(hdiD1[1], 0.01, hdiD1[2], 0.01, lty=3, col='blue')
# This is a valid 95% CrI, but not an HDI
(hdiD2 <- hdi(dens2, allowSplit=TRUE))
(ht <- attr(hdiD2, "height"))
segments(hdiD2[, 1], ht, hdiD2[, 2], ht, lwd=3, col='blue')
# This is the correct 95% HDI.
```

`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)
```

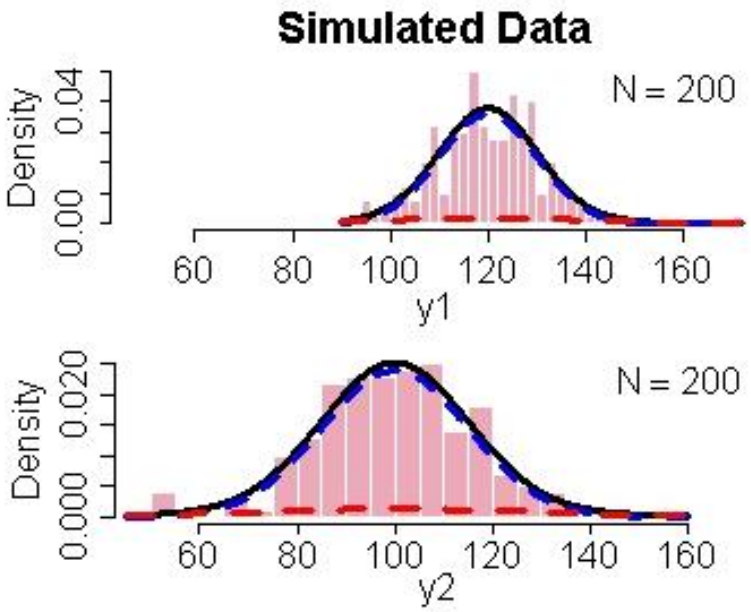
Arguments

`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, ie. samples which exactly match the specified means and standard deviations. If showPlot = TRUE, the results are displayed as a plot:



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

Kruschke, J. K. 2013. Bayesian estimation supersedes the *t* test. *Journal of Experimental Psychology: General* 142(2):573-603. doi: 10.1037/a0029146

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

```
## 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

Kruschke, J. K. 2013. Bayesian estimation supersedes the *t* test. *Journal of Experimental Psychology: General* 142(2):573-603. doi: 10.1037/a0029146

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

```
# 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

```
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).

Details

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.

Value

Returns NULL invisibly. Used for the side effect.

Author(s)

Code by John Kruschke, modified by Mike Meredith.

References

Kruschke, J. K. 2013. Bayesian estimation supersedes the *t* test. *Journal of Experimental Psychology: General* 142(2):573-603. doi: 10.1037/a0029146

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

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

---

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

Description

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.

Usage

```
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

http://doingbayesiandataanalysis.blogspot.com/2013/08/how-much-of-bayesian-posterior.html

Examples

```
# 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

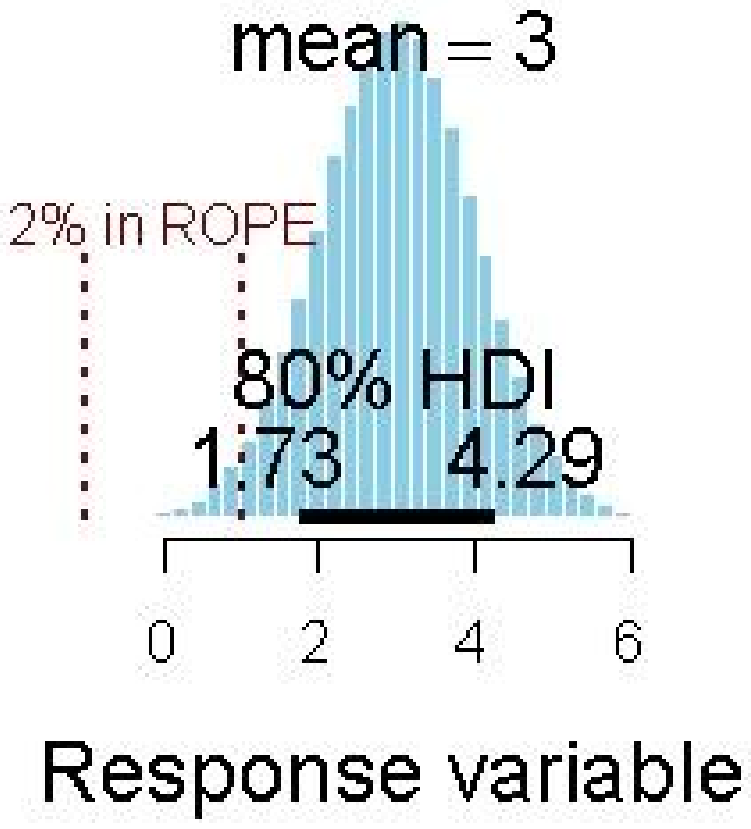
```
plotPost(paramSampleVec, credMass = 0.95, compVal = NULL, ROPE = NULL,
  HDItextPlace = 0.7, showMode = FALSE, showCurve = FALSE, ...)
```

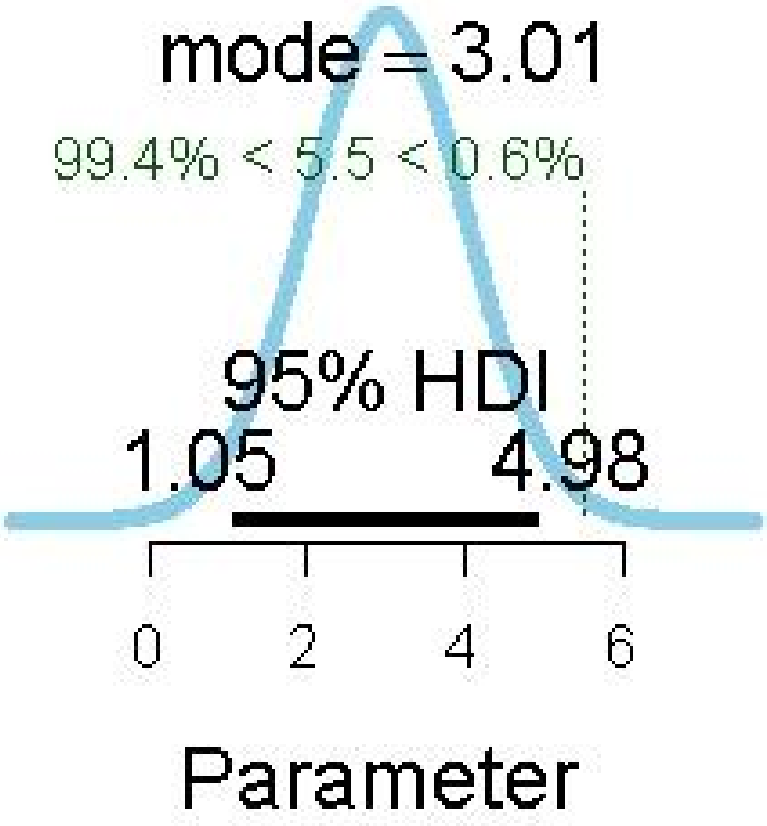
Arguments

`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.

Details

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

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

```
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
```

`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

Kruschke, J. K. 2013. Bayesian estimation supersedes the *t* test. *Journal of Experimental Psychology: General* 142(2):573-603. doi: 10.1037/a0029146

Examples

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

postPriorOverlap *Overlap between posterior and prior probability distributions.*

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

References

Gimenez, Morgan and Brooks (2009) Weak identifiability in models for mark-recapture-recovery data. pp.1055-1068 in Thomson, Cooch and Conroy (eds) *Modeling demographic processes in marked populations* Springer

Examples

```
# 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

```
## 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

```
## 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

```
## 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

Kruschke, J K. 2013. Bayesian estimation supersedes the *t* test. *Journal of Experimental Psychology: General* 142(2):573-603. doi: 10.1037/a0029146

Macmillan, N. A., & Creelman, C. D. (1991). *Detection Theory: A User's Guide*. New York, Cambridge University Press

See Also

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

Examples

```
## see "BEST-package"
```

Index