The coda Package

March 27, 2006

Version 0.10-5

Date 2006-03-27

Title Output analysis and diagnostics for MCMC

Author Martyn Plummer, Nicky Best, Kate Cowles, Karen Vines

Maintainer Martyn Plummer <plummer@iarc.fr>

Depends R ($>= 2.2.0$), lattice

Description Output analysis and diagnostics for Markov Chain Monte Carlo simulations.

License GPL Version 2 or later.

R topics documented:

2 Cramer Contract of the Contr

Index [39](#page-38-0)

Cramer *The Cramer-von Mises Distribution*

Description

Distribution function of the Cramer-von Mises distribution.

Usage

pcramer(q, eps)

Arguments

Value

pcramer gives the distribution function,

References

Anderson TW. and Darling DA. Asymptotic theory of certain 'goodness of fit' criteria based on stochastic processes. *Ann. Math. Statist.*, 23, 192-212 (1952).

Csorgo S. and Faraway, JJ. The exact and asymptotic distributions of the Cramer-von Mises statistic. J. Roy. Stat. Soc. (B), 58, 221-234 (1996).

Create Highest Posterior Density (HPD) intervals for the parameters in an MCMC sample.

Usage

```
HPDinterval(obj, prob = 0.95, ...)
## S3 method for class 'mcmc':
HPDinterval(obj, prob = 0.95, ...)
## S3 method for class 'mcmc.list':
HPDinterval(obj, prob = 0.95, ...)
```
Arguments

Details

For each parameter the interval is constructed from the empirical cdf of the sample as the shortest interval for which the difference in the ecdf values of the endpoints is the nominal probability. Assuming that the distribution is not severely multimodal, this is the HPD interval.

Value

For an "mcmc" object, a matrix with columns "lower" and "upper" and rows corresponding to the parameters. The attribute "Probability" is the nominal probability content of the intervals. A list of such matrices is returned for an "mcmc.list" object.

Author(s)

Douglas Bates

Examples

```
data(line)
HPDinterval(line)
```


as.ts.mcmc will coerce an mcmc object to a time series.

Usage

as.ts.mcmc (x, \ldots)

Arguments

Author(s)

Martyn Plummer

See Also

[as.ts](#page-0-0)

autocorr *Autocorrelation function for Markov chains*

Description

autocorr calculates the autocorrelation function for the Markov chain mcmc.obj at the lags given by lags. The lag values are taken to be relative to the thinning interval if relative=TRUE.

High autocorrelations within chains indicate slow mixing and, usually, slow convergence. It may be useful to thin out a chain with high autocorrelations before calculating summary statistics: a thinned chain may contain most of the information, but take up less space in memory. Re-running the MCMC sampler with a different parameterization may help to reduce autocorrelation.

Usage

autocorr(x, lags = $c(0, 1, 5, 10, 50)$, relative=TRUE)

autocorr.diag 5

Value

A vector or array containing the autocorrelations.

Author(s)

Martyn Plummer

See Also

[acf](#page-0-0), [autocorr.plot](#page-5-1).

autocorr.diag *Autocorrelation function for Markov chains*

Description

autocorr.diag calculates the autocorrelation function for the Markov chain mcmc.obj at the lags given by lags. The lag values are taken to be relative to the thinning interval if relative=TRUE. Unlike autocorr, if mcmc.obj has many parmeters it only computes the autocorrelations with itself and not the cross correlations. In cases where autocorr would return a matrix, this function returns the diagonal of the matrix. Hence it is more useful for chains with many parameters, but may not be as helpful at spotting parameters.

If mcmc.obj is of class mcmc.list then the returned vector is the average autocorrelation across all chains.

Usage

autocorr.diag(mcmc.obj, ...)

Arguments

Value

A vector containing the autocorrelations.

Author(s)

Russell Almond

See Also

[autocorr](#page-3-1), [acf](#page-0-0), [autocorr.plot](#page-5-1).

Plots the autocorrelation function for each variable in each chain in x.

Usage

```
autocorr.plot(x, lag.max, auto.layout = TRUE,
ask=dev.interactive(), ...)
```
Arguments

See Also

[autocorr](#page-3-1).

Description

Effective standard deviation of population to produce the correct standard errors.

Usage

```
batchSE(x, batchSize=100)
```
Arguments

Details

Because of the autocorrelation, the usual method of taking var(x)/n overstates the precision of the estimate. This method works around the problem by looking at the means of batches of the parameter. If the batch size is large enough, the batch means should be approximately uncorrelated and the normal formula for computing the standard error should work.

The batch standard error procedure is usually thought to be not as accurate as the time series methods used in summary and effectiveSize. It is included here for completeness.

bugs2jags 7

Value

A vector giving the standard error for each column of x.

Author(s)

Russell Almond

References

Roberts, GO (1996) Markov chain concepts related to sampling algorithms, in Gilks, WR, Richardson, S and Spiegelhalter, DJ, *Markov Chain Monte Carlo in Practice*, Chapman and Hall, 45-58.

See Also

[spectrum0.ar](#page-29-1), [effectiveSize](#page-11-1), [summary.mcmc](#page-30-1)

bugs2jags *Convert WinBUGS data file to JAGS data file*

Description

bugs2jags converts a WinBUGS data in the format called "S-Plus" (i.e. the format created by the dput function) and writes it in dump format used by JAGS.

NB WinBUGS stores its arrays in row order. This is different from R and JAGS which both store arrays in column order. This difference is taken into account by bugs2jags which will automatically reorder the data in arrays, without changing the dimension.

Usage

```
bugs2jags(infile, outfile)
```
Arguments

Author(s)

Martyn Plummer

References

Spiegelhalter DJ, Thomas A, Best NG and Lunn D (2003). *WinBUGS version 1.4 user manual* MRC Biostatistics Unit, Cambridge, UK.

See Also

[dput](#page-0-0), [dump](#page-0-0).

coda.options is a utility function that queries and sets options for the codamenu() function. These settings affect the behaviour of the functions in the coda library only when they are called via the codamenu() interface.

The coda.options() function behaves just like the options() function in the base library, with the additional feature that coda.options (default=TRUE) will reset all options to the default values.

Options can be pretty-printed using the display.coda.options() function, which groups the options into sections.

Available options are

- bandwidth Bandwidth function used when smoothing samples to produce density estimates. Defaults to Silverman's "Rule of thumb".
- combine.corr Logical option that determines whether to combine multiple chains when calculating cross-correlations.
- combine.plots Logical option that determines whether to combine multiple chains when plotting.
- combine.plots Logical option that determines whether to combine multiple chains when calculating summary statistics.
- data.saved For internal use only.
- densplot Logical option that determines whether to plot a density plot when plot methods are called for mcmc objects.
- digits Number of significant digits to use when printing.
- frac1 For Geweke diagnostic, fraction to use from start of chain. Defaults to 0.1
- frac2 For Geweke diagnostic, fraction to use from end of chain. Default to 0.5.
- gr.bin For Geweke-Brooks plot, number of iterations to use per bin.
- **gr.max** For Geweke-Brooks plot, maximum number of bins to use. This option overrides $q\textbf{r}$. bin.
- halfwidth For Heidelberger and Welch diagnostic, the target value for the ratio of half width to sample mean.
- lowess Logical option that controls whether to plot a smooth line through a trace plot when plotting MCMC output.
- q For Raftery and Lewis diagnostic, the target quantile to be estimated
- r For Raftery and Lewis diagnostic, the required precision.
- s For Raftery and Lewis diagnostic, the probability of obtaining an estimate in the interval (q-r, q+r).
- quantiles Vector of quantiles to print when calculating summary statistics for MCMC output.
- trace Logical option that determines whether to plot a trace of the sampled output when plotting MCMC output.
- user.layout Logical option that determines whether current value of par("mfrow") should be used for plots (TRUE) or whether the optimal layout should be calculated (FALSE).

codamenu 99 animates and the contract of the c

Usage

```
coda.options(...)
display.coda.options(data = FALSE, stats = FALSE, plots = FALSE, diags = FALSE)
.Coda.Options
.Coda.Options.Default
```
Arguments

See Also

[options](#page-0-0)

codamenu *Main menu driver for the coda package*

Description

codamenu presents a simple menu-based interface to the functions in the coda package. It is designed for users who know nothing about the R language.

Usage

codamenu()

Author(s)

Kate Cowles, Nicky Best, Karen Vines, Martyn Plummer

crosscorr *Cross correlations for MCMC output*

Description

crosscorr calculates cross-correlations between variables in Markov Chain Monte Carlo output. If x is an mcmc.list then all chains in x are combined before calculating the correlation.

Usage

crosscorr(x)

Arguments

x an mcmc or mcmc.list object.

Value

A matrix or 3-d array containing the correlations.

See Also

[crosscorr.plot](#page-9-1), [autocorr](#page-3-1).

crosscorr.plot *Plot image of correlation matrix*

Description

crosscorr.plot provides an image of the correlation matrix for x. If x is an mcmc.list object, then all chains are combined.

The range $[-1,1]$ is divided into a number of equal-length categories given by the length of col and assigned the corresponding color. By default, topographic colours are used as this makes it easier to distinguish positive and negative correlations.

Usage

crosscorr.plot $(x, col = topo.close(10), ...)$

Arguments

See Also

[crosscorr](#page-8-1), [image](#page-0-0), [topo.colors](#page-0-0).

cumuplot *Cumulative quantile plot*

Description

Plots the evolution of the sample quantiles as a function of the number of iterations.

Usage

```
cumuplot(x, probs=c(0.025,0.5,0.975), ylab="",
         lty=c(2,1), lwd=c(1,2), type="l", ask=dev.interactive(),
         auto.layout=TRUE, col=1, ...)
```


densplot the contract of the c

Arguments

Author(s)

Arni Magnusson <arnima@u.washington.edu>

densplot *Probability density function estimate from MCMC output*

Description

Displays a plot of the density estimate for each variable in x, calculated by the density function.

Usage

```
densplot(x, show.obs=TRUE, bwf, main, ylim, ...)
```
Arguments

Note

You can call this function directly, but it is more usually called by the plot.mcmc function.

If a variable is bounded below at 0, or bounded in the interval [0,1], then the data are reflected at the boundary before being passed to the density() function. This allows correct estimation of a non-zero density at the boundary.

See Also

[density](#page-0-0), [plot.mcmc](#page-22-1).

Sample size adjusted for autocorrelation.

Usage

```
effectiveSize(x)
```
Arguments

x An mcmc or mcmc.list object.

Details

For a time series x of length N, the standard error of the mean is var(x)/n where n is the effective sample size. $n = N$ only when there is no autocorrelation.

Estimation of the effective sample size requires estimating the spectral density at frequency zero. This is done by the function spectrum0.ar

For a mcmc.list object, the effective sizes are summed across chains. To get the size for each chain individually use lapply(x,effectiveSize).

Value

A vector giving the effective sample size for each column of x.

See Also

[spectrum0.ar](#page-29-1).

gelman.diag *Gelman and Rubin's convergence diagnostic*

Description

The 'potential scale reduction factor' is calculated for each variable in x , together with upper and lower confidence limits. Approximate convergence is diagnosed when the upper limit is close to 1. For multivariate chains, a multivariate value is calculated that bounds above the potential scale reduction factor for any linear combination of the (possibly transformed) variables.

The confidence limits are based on the assumption that the stationary distribution of the variable under examination is normal. Hence the 'transform' parameter may be used to improve the normal approximation.

Usage

```
gelman.diag(x, confidence = 0.95, transform=FALSE, autoburnin=TRUE)
```
gelman.diag 13

Arguments

Theory

Gelman and Rubin (1992) propose a general approach to monitoring convergence of MCMC output in which two or more parallel chains are run with starting values that are overdispersed relative to the posterior distribution. Convergence is diagnosed when the chains have 'forgotten' their initial values, and the output from all chains is indistinguishable. The q elman.diag diagnostic is applied to a single variable from the chain. It is based a comparison of within-chain and between-chain variances, and is similar to a classical analysis of variance.

There are two ways to estimate the variance of the stationary distribution: the mean of the empirical variance within each chain, W, and the empirical variance from all chains combined, which can be expressed as

$$
\widehat{\sigma}^2 = \frac{(n-1)B}{n} + \frac{W}{n}
$$

where B is the empirical between-chain variance.

If the chains have converged, then both estimates are unbiased. Otherwise the first method will *underestimate* the variance, since the individual chains have not had time to range all over the stationary distribution, and the second method will *overestimate* the variance, since the starting points were chosen to be overdispersed.

The convergence diagnostic is based on the assumption that the target distribution is normal. A Bayesian credible interval can be constructed using a t-distribution with mean

 $\hat{\mu}$ = Sample mean of all chains combined

and variance

$$
\widehat{V} = \widehat{\sigma}^2 + \frac{B}{mn}
$$

where m is the number of chains, and degrees of freedom estimated by the method of moments

$$
d = \frac{2 * \widehat{V}}{\text{Var}(\widehat{V})}
$$

Use of the t-distribution accounts for the fact that the mean and variance of the posterior distribution are estimated.

The convergence diagnostic itself is

$$
R = \sqrt{\frac{(d+3)\hat{V}}{(d+1)W}}
$$

Values substantially above 1 indicate lack of convergence. If the chains have not converged, Bayesian credible intervals based on the t-distribution are too wide, and have the potential to shrink by this factor if the MCMC run is continued.

Note

The multivariate a version of Gelman and Rubin's diagnostic was proposed by Brooks and Gelman (1997).

References

Gelman, A and Rubin, DB (1992) Inference from iterative simulation using multiple sequences, *Statistical Science*, 7, 457-511.

Brooks, SP. and Gelman, A. (1997) General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, 7, 434-455.

See Also

[gelman.plot](#page-13-1).

gelman.plot *Gelman-Rubin-Brooks plot*

Description

This plot shows the evolution of Gelman and Rubin's shrink factor as the number of iterations increases.

Usage

```
gelman.plot(x, bin.width = 10, max.bins = 50,
confidence = 0.95, transform = FALSE, auto.layout = TRUE,
ask = dev.interactive(), col, lty, xlab, ylab, type, \ldots)
```


geweke.diag 15

Details

The Markov chain is divided into bins according to the arguments bin.width and max.bins. Then the Gelman-Rubin shrink factor is repeatedly calculated. The first shrink factor is calculated with observations 1:50, the second with observations 1 : $(50+n)$ where n is the bin width, the third contains samples $1:(50+2n)$ and so on.

Theory

A potential problem with q elman.diag is that it may mis-diagnose convergence if the shrink factor happens to be close to 1 by chance. By calculating the shrink factor at several points in time, gelman.plot shows if the shrink factor has really converged, or whether it is still fluctuating.

References

Brooks, S P. and Gelman, A. (1998) General Methods for Monitoring Convergence of Iterative Simulations. Journal of Computational and Graphical Statistics. 7. p434-455.

See Also

[gelman.diag](#page-11-2).

geweke.diag *Geweke's convergence diagnostic*

Description

Geweke (1992) proposed a convergence diagnostic for Markov chains based on a test for equality of the means of the first and last part of a Markov chain (by default the first 10% and the last 50%). If the samples are drawn from the stationary distribution of the chain, the two means are equal and Geweke's statistic has an asymptotically standard normal distribution.

The test statistic is a standard Z-score: the difference between the two sample means divided by its estimated standard error. The standard error is estimated from the spectral density at zero and so takes into account any autocorrelation.

The Z-score is calculated under the assumption that the two parts of the chain are asymptotically independent, which requires that the sum of frac1 and frac2 be strictly less than 1.

Usage

```
geweke.diag(x, frac1=0.1, frac2=0.5)
```
Arguments

Value

Z-scores for a test of equality of means between the first and last parts of the chain. A separate statistic is calculated for each variable in each chain.

References

Geweke, J. Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In *Bayesian Statistics 4* (ed JM Bernado, JO Berger, AP Dawid and AFM Smith). Clarendon Press, Oxford, UK.

See Also

[geweke.plot](#page-15-1).

geweke.plot *Geweke-Brooks plot*

Description

If geweke.diag indicates that the first and last part of a sample from a Markov chain are not drawn from the same distribution, it may be useful to discard the first few iterations to see if the rest of the chain has "converged". This plot shows what happens to Geweke's Z-score when successively larger numbers of iterations are discarded from the beginning of the chain. To preserve the asymptotic conditions required for Geweke's diagnostic, the plot never discards more than half the chain.

The first half of the Markov chain is divided into nbins - 1 segments, then Geweke's Z-score is repeatedly calculated. The first Z-score is calculated with all iterations in the chain, the second after discarding the first segment, the third after discarding the first two segments, and so on. The last Z-score is calculated using only the samples in the second half of the chain.

Usage

```
qeweke.plot(x, frac1 = 0.1, frac2 = 0.5, nbins = 20,
            pvalue = 0.05, auto.layout = TRUE, ask = dev.interactive(), ...)
```
Arguments

Note

The graphical implementation of Geweke's diagnostic was suggested by Steve Brooks.

See Also

[geweke.diag](#page-14-1).

heidel.diag is a run length control diagnostic based on a criterion of relative accuracy for the estimate of the mean. The default setting corresponds to a relative accuracy of two significant digits.

heidel.diag also implements a convergence diagnostic, and removes up to half the chain in order to ensure that the means are estimated from a chain that has converged.

Usage

heidel.diag(x, eps=0.1, pvalue=0.05)

Arguments x

Details

The convergence test uses the Cramer-von-Mises statistic to test the null hypothesis that the sampled values come from a stationary distribution. The test is successively applied, firstly to the whole chain, then after discarding the first 10%, 20%, ... of the chain until either the null hypothesis is accepted, or 50% of the chain has been discarded. The latter outcome constitutes 'failure' of the stationarity test and indicates that a longer MCMC run is needed. If the stationarity test is passed, the number of iterations to keep and the number to discard are reported.

The half-width test calculates a 95% confidence interval for the mean, using the portion of the chain which passed the stationarity test. Half the width of this interval is compared with the estimate of the mean. If the ratio between the half-width and the mean is lower than eps, the halfwidth test is passed. Otherwise the length of the sample is deemed not long enough to estimate the mean with sufficient accuracy.

Theory

The heidel.diag diagnostic is based on the work of Heidelberger and Welch (1983), who combined their earlier work on simulation run length control (Heidelberger and Welch, 1981) with the work of Schruben (1982) on detecting initial transients using Brownian bridge theory.

Note

If the half-width test fails then the run should be extended. In order to avoid problems caused by sequential testing, the test should not be repeated too frequently. Heidelberger and Welch (1981) suggest increasing the run length by a factor $I > 1.5$, each time, so that estimate has the same, reasonably large, proportion of new data.

References

Heidelberger P and Welch PD. A spectral method for confidence interval generation and run length control in simulations. Comm. ACM. 24, 233-245 (1981)

Heidelberger P and Welch PD. Simulation run length control in the presence of an initial transient. *Opns Res.*, 31, 1109-44 (1983)

Schruben LW. Detecting initialization bias in simulation experiments. *Opns. Res.*, 30, 569-590 (1982).

line *Simple linear regression example*

Description

Sample MCMC output from a simple linear regression model given in the BUGS manual.

Usage

data(line)

Format

An mcmc object

Source

Spiegelhalter, D.J., Thomas, A., Best, N.G. and Gilks, W.R. (1995) BUGS: Bayesian inference using Gibbs Sampling, Version 0.5, MRC Biostatistics Unit, Cambridge.

mcmc *Markov Chain Monte Carlo Objects*

Description

The function 'mcmc' is used to create a Markov Chain Monte Carlo object. The data are taken to be a vector, or a matrix with one column per variable.

An mcmc object may be summarized by the summary function and visualized with the plot function.

MCMC objects resemble time series (ts) objects and have methods for the generic functions $time$, start, end, frequency and window.

Usage

```
mcmc(data= NA, start = 1, end = numeric(0), thin = 1)
as.mcmc(x)
is.mcmc(x)
```


mcmc.convert 19

Arguments

Note

The format of the mcmc class has changed between coda version 0.3 and 0.4. Older mcmc objects will now cause is.mcmc to fail with an appropriate warning message. Obsolete mcmc objects can be upgraded with the mcmcUpgrade function.

Author(s)

Martyn Plummer

See Also

[mcmc.list](#page-19-1), [mcmcUpgrade](#page-20-1), [thin](#page-31-1), [window.mcmc](#page-37-1), [summary.mcmc](#page-30-1), [plot.mcmc](#page-22-1).

mcmc.convert *Conversions of MCMC objects*

Description

These are methods for the generic functions as $matrix($, as.array() and as.mcmc().

as.matrix() strips the MCMC attributes from an mcmc object and returns a matrix. If iters = TRUE then a column is added with the iteration number. For mcmc.list objects, the rows of multiple chains are concatenated and, if chains = TRUE a column is added with the chain number.

mcmc.list objects can be coerced to 3-dimensional arrays with the as.array() function.

An mcmc.list object with a single chain can be coerced to an mcmc object with $as.mcmc()$. If the argument has multiple chains, this causes an error.

Usage

```
as.matrix.mcmc(x, iters = FALSE)
as.matrix.mcmc.list(x, iters = FALSE, chains = FALSE)
as.array.mcmc.list(x, drop, ...)
```


See Also

[as.matrix](#page-0-0), [as.array](#page-0-0), [as.mcmc](#page-17-1),

mcmc.list *Replicated Markov Chain Monte Carlo Objects*

Description

The function 'mcmc.list' is used to represent parallel runs of the same chain, with different starting values and random seeds. The list must be balanced: each chain in the list must have the same iterations and the same variables.

Diagnostic functions which act on mcmc objects may also be applied to mcmc.list objects. In general, the chains will be combined, if this makes sense, otherwise the diagnostic function will be applied separately to each chain in the list.

Since all the chains in the list have the same iterations, a single time dimension can be ascribed to the list. Hence there are time series methods time, window, start, end, frequency and thin for mcmc.list objects.

An mcmc. List can be indexed as if it were a single mcmc object using the [operator (see examples below). The [[operator selects a single mcmc object from the list.

Usage

```
mcmc.list(...)
as.mcmc.list(x, \ldots)is.mcmc.list(x)
```
Arguments

Author(s)

Martyn Plummer

See Also

[mcmc](#page-17-2).

Examples

```
data(line)
x1 <- line[[1]] #Select first chain
x2 <- line[,1, drop=FALSE] #Select first var from all chains
varnames(x2) == varnames(line)[1] #TRUE
```


These are methods for subsetting mcmc objects. You can select iterations using the first dimension and variables using the second dimension. Selecting iterations will return a vector or matrix, not an mcmc object. If you want to do row-subsetting of an mcmc object and preserve its dimensions, use the window function.

Subsetting applied to an mcmc.list object will simultaneously affect all the parallel chains in the object.

Usage

x[i,j]

Arguments

See Also

[\[](#page-0-0), [window.mcmc](#page-37-1)

mcmcUpgrade *Upgrade mcmc objects in obsolete format*

Description

In previous releases of CODA, an mcmc object could be a single or multiple chains. A new class mcmc.list has now been introduced to deal with multiple chains and mcmc objects can only have data from a single chain.

Objects stored in the old format are now obsolete and must be upgraded.

Usage

```
mcmcUpgrade(x)
```
Arguments

x an obsolete mcmc object.

Author(s)

Martyn Plummer

See Also

[mcmc](#page-17-2).

The 'mcpar' attribute of an MCMC object gives the start iteration the end iteration and the thinning interval of the chain.

It resembles the 'tsp' attribute of time series (ts) objects.

Usage

mcpar(x)

Arguments

x An mcmcm or mcmc.list object

See Also

[ts](#page-0-0), [mcmc](#page-17-2), [mcmc.list](#page-19-1),

multi.menu *Choose multiple options from a menu*

Description

multi.menu presents the user with a menu of choices labelled from 1 to the number of choices. The user may choose one or more options by entering a comma separated list. A range of values may also be specified using the ":" operator. Mixed expressions such as "1,3:5, 6" are permitted.

If allow.zero is set to TRUE, one can select '0' to exit without choosing an item.

Usage

multi.menu(choices, title, header, allow.zero = TRUE)

Arguments

Value

Numeric vector giving the numbers of the options selected, or 0 if no selection is made.

Author(s)

Martyn Plummer

nchain 23

See Also

[menu](#page-0-0).

nchain *Dimensions of MCMC objects*

Description

These functions give the dimensions of an MCMC object

niter(x) returns the number of iterations.

 $\mathbf{nvar}(\mathbf{x})$ returns the number of variables.

chain(x) returns the number of parallel chains.

Usage

niter(x) nvar(x) nchain(x)

Arguments

x An mcmc or mcmc.list object

Value

A numeric vector of length 1:

See Also

[mcmc](#page-17-2), [mcmc.list](#page-19-1),

plot.mcmc *Summary plots of mcmc objects*

Description

plot.mcmc summarizes an mcmc or mcmc.list object with a trace of the sampled output and a density estimate for each variable in the chain.

Usage

```
## S3 method for class 'mcmc':
plot(x, trace = TRUE, density = TRUE, smooth = TRUE, bwf,
        auto.layout = TRUE, ask = dev.interactive(), \dots)
```
Arguments

Author(s)

Martyn Plummer

See Also

[densplot](#page-10-1), [traceplot](#page-32-1).

raftery.diag *Raftery and Lewis's diagnostic*

Description

raftery.diag is a run length control diagnostic based on a criterion of accuracy of estimation of the quantile q. It is intended for use on a short pilot run of a Markov chain. The number of iterations required to estimate the quantile q to within an accuracy of $+/- r$ with probability p is calculated. Separate calculations are performed for each variable within each chain.

If the number of iterations in data is too small, an error message is printed indicating the minimum length of pilot run. The minimum length is the required sample size for a chain with no correlation between consecutive samples. Positive autocorrelation will increase the required sample size above this minimum value. An estimate I (the 'dependence factor') of the extent to which autocorrelation inflates the required sample size is also provided. Values of I larger than 5 indicate strong autocorrelation which may be due to a poor choice of starting value, high posterior correlations or 'stickiness' of the MCMC algorithm.

The number of 'burn in' iterations to be discarded at the beginning of the chain is also calculated.

Usage

raftery.diag(data, $q=0.025$, $r=0.005$, $s=0.95$, converge.eps=0.001)

read.and.check 25

Value

A list with class raftery.diag. A print method is available for objects of this class. the contents of the list are

Theory

The estimated sample size for variable U is based on the process $Z_t = d(U_t \leq u)$ where d is the indicator function and u is the qth quantile of U. The process Z_t is derived from the Markov chain data by marginalization and truncation, but is not itself a Markov chain. However, Z_t may behave as a Markov chain if it is sufficiently thinned out. raftery.diag calculates the smallest value of thinning interval k which makes the thinned chain Z_t^k behave as a Markov chain. The required sample size is calculated from this thinned sequence. Since some data is 'thrown away' the sample size estimates are conservative.

The criterion for the number of 'burn in' iterations m to be discarded, is that the conditional distribution of Z_m^k given Z_0 should be within <code>converge.eps</code> of the equilibrium distribution of the chain Z_t^k .

Note

raftery.diag is based on the FORTRAN program 'gibbsit' written by Steven Lewis, and available from the Statlib archive.

References

Raftery, A.E. and Lewis, S.M. (1992). One long run with diagnostics: Implementation strategies for Markov chain Monte Carlo. *Statistical Science*, 7, 493-497.

Raftery, A.E. and Lewis, S.M. (1995). The number of iterations, convergence diagnostics and generic Metropolis algorithms. *In* Practical Markov Chain Monte Carlo (W.R. Gilks, D.J. Spiegelhalter and S. Richardson, eds.). London, U.K.: Chapman and Hall.

read.and.check *Read data interactively and check that it satisfies conditions*

Description

Input is read interactively and checked against conditions specified by the arguments what, lower, upper and answer.in. If the input does not satisfy all the conditions, an appropriate error message is produced and the user is prompted to provide input. This process is repeated until a valid input value is entered.

Usage

```
read.and.check(message = "", what = numeric(), lower, upper, answer.in,
default)
```
Arguments

Value

The value of the valid input. When the default argument is specified, a blank line is accepted as valid input and in this case read.and.check returns the value of default.

Note

Since the function does not return a value until it receives valid input, it extensively checks the conditions for consistency before prompting the user for input. Inconsistent conditions will cause an error.

Author(s)

Martyn Plummer

read.coda *Read output files in CODA format*

Description

read.coda reads Markov Chain Monte Carlo output in the CODA format produced by Open-BUGS and JAGS. By default, all of the data in the file is read, but the arguments start, end and thin may be used to read a subset of the data. If the arguments given to start, end or thin are incompatible with the data, they are ignored.

Usage

```
read.coda(output.file, index.file, start, end, thin, quiet=FALSE)
read.jags(file = "jags.out", start, end, thin, quiet=FALSE)
```


Value

An object of class mcmc containing a representation of the data in the file.

Author(s)

Karen Vines, Martyn Plummer

References

Spiegelhalter DJ, Thomas A, Best NG and Gilks WR (1995). *BUGS: Bayesian inference Using Gibbs Sampling, Version 0.50.* MRC Biostatistics Unit, Cambridge.

See Also

[mcmc](#page-17-2), [read.coda.interactive](#page-26-1), [read.openbugs](#page-27-1).

read.coda.interactive *Read CODA output files interactively*

Description

read.coda.interactive reads Markov Chain Monte Carlo output in the format produced by the classic BUGS program. No arguments are required. Instead, the user is prompted for the required information.

Usage

```
read.coda.interactive()
```
Value

An object of class mcmc.list containing a representation of the data in one or more BUGS output files.

Note

This function is normally called by the codamenu function, but can also be used on a stand-alone basis.

Author(s)

Nicky Best, Martyn Plummer

References

Spiegelhalter DJ, Thomas A, Best NG and Gilks WR (1995). *BUGS: Bayesian inference Using Gibbs Sampling, Version 0.50.* MRC Biostatistics Unit, Cambridge.

See Also

[mcmc](#page-17-2), [mcmc.list](#page-19-1), [read.coda](#page-25-1), [codamenu](#page-8-2).

read.openbugs reads Markov Chain Monte Carlo output in the CODA format produced by OpenBUGS.

This is a convenience wrapper around the function read.coda which allows you to read all the data output by OpenBUGS by specifying only the file stem.

Usage

read.openbugs(stem="", start, end, thin, quiet=FALSE)

Arguments

Value

An object of class mcmc.list containing output from all chains.

Author(s)

Martyn Plummer

References

Spiegelhalter DJ, Thomas A, Best NG and Lunn D (2004). *WinBUGS User Manual, Version 2.0, June 2004*, MRC Biostatistics Unit, Cambridge.

See Also

[read.coda](#page-25-1).

rejectionRate *Rejection Rate for Metropolis–Hastings chains*

Description

rejectionRate calculates the fraction of time that a Metropolis–Hastings type chain rejected a proposed move. The rejection rate is calculates separately for each variable in the mcmc.obj argument, irregardless of whether the variables were drawn separately or in a block. In the latter case, the values returned should be the same.

Usage

```
rejectionRate(x)
```
Arguments

x An mcmc or mcmc.list object.

Details

For the purposes of this function, a "rejection" has occurred if the value of the time series is the same at two successive time points. This test is done naively using == and may produce problems due to rounding error.

Value

A vector containing the rejection rates, one for each variable.

Author(s)

Russell Almond

spectrum0 *Estimate spectral density at zero*

Description

The spectral density at frequency zero is estimated by fitting a glm to the low-frequency end of the periodogram. spectrum0(x)/length(x) estimates the variance of mean(x).

Usage

```
spectrum0(x, max.freq = 0.5, order = 1, max.length = 200)
```


Details

The raw periodogram is calculated for the series x and a generalized linear model with family Gamma and log link is fitted to the periodogram.

The linear predictor is a polynomial in terms of the frequency. The degree of the polynomial is determined by the parameter order.

Value

A list with the following values

spec The predicted value of the spectral density at frequency zero.

Theory

Heidelberger and Welch (1991) observed that the usual non-parametric estimator of the spectral density, obtained by smoothing the periodogram, is not appropriate for frequency zero. They proposed an alternative parametric method which consisted of fitting a linear model to the log periodogram of the batched time series. Some technical problems with model fitting in their original proposal can be overcome by using a generalized linear model.

Batching of the data, originally proposed in order to save space, has the side effect of flattening the spectral density and making a polynomial fit more reasonable. Fitting a polynomial of degree zero is equivalent to using the 'batched means' method.

Note

The definition of the spectral density used here differs from that used by $spec$, $pqram$. We consider the frequency range to be between 0 and 0.5, not between 0 and f requency(x)/2.

The model fitting may fail on chains with very high autocorrelation.

References

Heidelberger, P and Welch, P.D. A spectral method for confidence interval generation and run length control in simulations. Communications of the ACM, Vol 24, pp233-245, 1981.

See Also

[spectrum](#page-0-0), [spectrum0.ar](#page-29-1), [glm](#page-0-0).

spectrum0.ar *Estimate spectral density at zero*

Description

The spectral density at frequency zero is estimated by fitting an autoregressive model. spectrum0(x)/length(x) estimates the variance of mean (x) .

Usage

spectrum0.ar(x)

summary.mcmc 31

Arguments

x A time series.

Details

The $ar($) function to fit an autoregressive model to the time series x. For multivariate time series, separate models are fitted for each column. The value of the spectral density at zero is then given by a well-known formula.

Value

A list with the following values

Note

The definition of the spectral density used here differs from that used by $spec.pgram$. We consider the frequency range to be between 0 and 0.5, not between 0 and f requency(x)/2.

See Also

[spectrum](#page-0-0), [spectrum0](#page-28-1), [glm](#page-0-0).

summary.mcmc *Summary statistics for Markov Chain Monte Carlo chains*

Description

summary.mcmc produces two sets of summary statistics for each variable:

Mean, standard deviation, naive standard error of the mean (ignoring autocorrelation of the chain) and time-series standard error based on an estimate of the spectral density at 0.

Quantiles of the sample distribution using the quantiles argument.

Usage

```
## S3 method for class 'mcmc':
summary(object, quantiles = c(0.025, 0.25, 0.5, 0.75, 0.975), ...)
```
Arguments

Author(s)

Martyn Plummer

See Also

[mcmc](#page-17-2), [mcmc.list](#page-19-1).

thin returns the interval between successive values of a time series. thin(x) is equivalent to 1/frequency(x).

This is a generic function. Methods have been implemented for mcmc objects.

Usage

thin(x, \ldots)

Arguments

Author(s)

Martyn Plummer

See Also

[time](#page-0-0).

time.mcmc *Time attributes for mcmc objects*

Description

These are methods for mcmc objects for the generic time series functions.

Usage

```
## S3 method for class 'mcmc':
time(x, ...)
## S3 method for class 'mcmc':
start(x, \ldots)## S3 method for class 'mcmc':
end(x, \ldots)## S3 method for class 'mcmc':
thin(x, \ldots)
```


traceplot 33

See Also

[time](#page-0-0), [start](#page-0-0), [frequency](#page-0-0), [thin](#page-31-1).

traceplot *Trace plot of MCMC output*

Description

Displays a plot of iterations *vs.* sampled values for each variable in the chain, with a separate plot per variable.

Usage

traceplot(x, smooth = TRUE, col, type, ylab, ...)

Arguments

Note

You can call this function directly, but it is more usually called by the plot. mcmc function.

See Also

[densplot](#page-10-1), [plot.mcmc](#page-22-1).

trellisplots *Trellis plots for mcmc objects*

Description

These methods use the Trellis framework as implemented in the lattice package to produce space-conserving diagnostic plots from "mcmc" and "mcmc.list" objects. The xyplot methods produce trace plots. The densityplot methods and qqmath methods produce empirical density and probability plots. The levelplot method depicts the correlation of the series. The acfplot methods plot the auto-correlation in the series.

Usage

```
## S3 method for class 'mcmc':
densityplot(x, data,
             outer, aspect = "xy",
             default.scales = list(relation = "free"),
             start = 1, thin = 1,
             main = attr(x, "title"),xlab = "",
             plot.points = "rug",
             ...,
             subset)
## S3 method for class 'mcmc.list':
densityplot(x, data,
             outer = FALSE, groups = !outer,
             aspect = "xy",
             default.scales = list(relation = "free"),
             start = 1, thin = 1,
             main = attr(x, "title"),
             xlab = ",
             plot.points = "rug",
             ...,
             subset)
## S3 method for class 'mcmc':
levelplot(x, data, main = attr(x, "title"),start = 1, thin = 1,
             ...,
             xlab = "", ylab = "",
             cuts = 10, at,col.regions = topo.colors(100),
             subset)
## S3 method for class 'mcmc':
qqmath(x, data,
             outer, aspect = "xy",
             default.scales = list(y = list(relation = "free")),
             prepanel = prepanel.qqmathline,
             start = 1, thin = 1,
             main = attr(x, "title"),ylab = "",
             ...,
             subset)
## S3 method for class 'mcmc.list':
qqmath(x, data,
             outer = FALSE, groups = !outer,
             aspect = "xy",
             default.scales = list(y = list(relation = "free")),
             prepanel = prepanel.qqmathline,
             start = 1, thin = 1,
             main = attr(x, "title"),
             ylab = ",
             ...,
             subset)
## S3 method for class 'mcmc':
```

```
xyplot(x, data,
             outer, layout = c(1, \text{ ncol}(x)),
             default.scales = list(y = list(relation = "free")),
             type = '1',start = 1, thin = 1,
             ylab = "",
             xlab = "Iteration number",
             main = attr(x, "title"),...,
             subset)
## S3 method for class 'mcmc.list':
xyplot(x, data, outer = FALSE, groups = !outer,aspect = "xy", layout = c(1, \text{ ncol}(x[[1]])),
             default.scales = list(y = list(relation = "free")),
             type = '1',start = 1, thin = 1,
             main = attr(x, "title"),ylab = "",
             ...,
             subset)
\alphacfplot(x, data, ...)
## S3 method for class 'mcmc':
acfplot(x, data, outer,
             prepanel, panel,
             type = 'h',
             aspect = "xy",
             start = 1, thin = 1,
             lag.max = NULL,ylab = "Autocorrelation",
             xlab = "Lag",
             main = attr(x, "title"),...,
             subset)
## S3 method for class 'mcmc.list':
acfplot(x, data, outer = FALSE, groups = !outer,prepanel, panel,
             type = if (groups) 'b' else 'h',
             aspect = "xy",
             start = 1, thin = 1,
             lag.max = NULL,ylab = "Autocorrelation",
             xlab = "Lag",
             main = attr(x, "title"),...,
             subset)
```


Value

An object of class "trellis". The relevant [update](#page-0-0) method can be used to update components of the object and the [print](#page-0-0) method (usually called by default) will plot it on an appropriate plotting device.

Author(s)

Deepayan Sarkar (Deepayan.Sarkar@R-project.org)

varnames 37

See Also

[Lattice](#page-0-0) for a brief introduction to lattice displays and links to further documentation.

Examples

```
data(line)
xyplot(line)
xyplot(line[[1]], start = 10)densityplot(line, start = 10)
qqmath(line, start = 10)
levelplot(line[[2]])
acfplot(line, outer = TRUE)
```
varnames *Named dimensions of MCMC objects*

Description

varnames() returns the variable names and chanames returns the chain names, or NULL if these are not set.

If $allow null = FALSE$ then NULL values will be replaced with canonical names.

Usage

```
varnames(x, allow.null=TRUE)
chanames(x, allow.null=TRUE)
varnames(x) <- value
chanames(x) <- value
```
Arguments

Value

A character vector , or NULL.

See Also

[mcmc](#page-17-2), [mcmc.list](#page-19-1).

window.mcmc is a method for mcmc objects which is normally called by the generic function window

In addition to the generic parameters, start and end the additional parameter thin may be used to thin out the Markov chain. Setting thin=k selects every kth iteration starting with the first. Note that the value of thin is *absolute* not relative. The value supplied given to the parameter thin must be a multiple of thin (x) .

Values of start, end and thin which are inconsistent with x are ignored, but a warning message is issued.

Usage

```
## S3 method for class 'mcmc':
window(x, start, end, thin, ...)
```
Arguments

See Also

[window](#page-0-0), [thin](#page-31-1).

Index

∗Topic array crosscorr, [8](#page-7-0) mcmc.convert, [18](#page-17-0) ∗Topic datasets line, [17](#page-16-0) ∗Topic distribution Cramer, [1](#page-0-1) ∗Topic file bugs2jags, [6](#page-5-0) read.coda, [25](#page-24-0) read.coda.interactive, [26](#page-25-0) read.openbugs, [27](#page-26-0) ∗Topic hplot autocorr.plot, [5](#page-4-0) crosscorr.plot, [9](#page-8-0) cumuplot, [9](#page-8-0) densplot, [10](#page-9-0) gelman.plot, [13](#page-12-0) geweke.plot, [15](#page-14-0) plot.mcmc, [22](#page-21-0) traceplot, [32](#page-31-0) trellisplots, [32](#page-31-0) ∗Topic htest gelman.diag, [11](#page-10-0) geweke.diag, [14](#page-13-0) heidel.diag, [16](#page-15-0) HPDinterval, [2](#page-1-0) raftery.diag, [23](#page-22-0) ∗Topic manip varnames, [36](#page-35-0) ∗Topic multivariate crosscorr, [8](#page-7-0) ∗Topic ts as.ts.mcmc, [3](#page-2-0) autocorr, [3](#page-2-0) autocorr.diag, [4](#page-3-0) batchSE, [5](#page-4-0) effectiveSize, [11](#page-10-0) mcmc, [17](#page-16-0) mcmc.list, [19](#page-18-0) mcmc.subset, [20](#page-19-0) mcmcUpgrade, [20](#page-19-0) mcpar, [21](#page-20-0)

nchain, [22](#page-21-0) rejectionRate, [28](#page-27-0) spectrum0, [28](#page-27-0) spectrum0.ar, [29](#page-28-0) thin, [31](#page-30-0) time.mcmc, [31](#page-30-0) window.mcmc, [37](#page-36-0) ∗Topic univar HPDinterval, [2](#page-1-0) summary.mcmc, [30](#page-29-0) ∗Topic utilities as.ts.mcmc, [3](#page-2-0) coda.options, [7](#page-6-0) codamenu, [8](#page-7-0) multi.menu, [21](#page-20-0) read.and.check, [24](#page-23-0) .Coda.Options *(*coda.options*)*, [7](#page-6-0) [, *20* [.mcmc *(*mcmc.subset*)*, [20](#page-19-0) acf, *4*, *35* acfplot *(*trellisplots*)*, [32](#page-31-0) acfplot.mcmc.list *(*trellisplots*)*, [32](#page-31-0) as.array, *19* as.array.mcmc.list *(*mcmc.convert*)*, [18](#page-17-0) as.matrix, *19* as.matrix.mcmc *(*mcmc.convert*)*, [18](#page-17-0) as.mcmc, *19* as.mcmc *(*mcmc*)*, [17](#page-16-0) as.mcmc.list *(*mcmc.list*)*, [19](#page-18-0) as.mcmc.mcmc.list *(*mcmc.convert*)*, [18](#page-17-0) as.ts, *3* as.ts.mcmc, [3](#page-2-0) autocorr, [3,](#page-2-0) *4, 5*, *9* autocorr.diag, [4](#page-3-0) autocorr.diag.mcmc.list *(*autocorr.diag*)*, [4](#page-3-0) autocorr.plot, *4*, [5](#page-4-0) batchSE, [5](#page-4-0)

bugs2jags, [6](#page-5-0)

```
40 INDEX
chanames (varnames), 36
chanames<- (varnames), 36
coda.options, 7
codamenu, 8, 26
Cramer, 1
crosscorr, 8, 9
crosscorr.plot, 9, 9
cumuplot, 9
density, 10
densityplot.mcmc (trellisplots),
       32
densplot, 10, 23, 32
display.coda.options
       (coda.options), 7
dput, 6
dump, 6
effectiveSize, 6, 11
end.mcmc (time.mcmc), 31
frequency, 32
frequency.mcmc (time.mcmc), 31
gelman.diag, 11, 14
gelman.plot, 13, 13
geweke.diag, 14, 15
geweke.plot, 15, 15
glm, 29, 30
heidel.diag, 16
HPDinterval, 2
HPDinterval.mcmc.list
       (HPDinterval), 2
image, 9
is.mcmc (mcmc), 17
is.mcmc.list (mcmc.list), 19
Lattice, 36
levelplot.mcmc (trellisplots), 32
line, 17
mcmc, 17, 19–22, 26, 30, 36
mcmc.convert, 18
mcmc.list, 18, 19, 21, 22, 26, 30, 36
mcmc.subset, 20
mcmcUpgrade, 18, 20
mcpar, 21
menu, 22
```
nchain, [22](#page-21-0) niter *(*nchain*)*, [22](#page-21-0)

multi.menu, [21](#page-20-0)

```
nvar (nchain), 22
options, 8
panel.densityplot, 35
panel.xyplot, 35
pcramer (Cramer), 1
plot.mcmc, 10, 18, 22, 32
plot.mcmc.list (mcmc.list), 19
print, 35
print.mcmc (mcmc), 17
qqmath.mcmc (trellisplots), 32
raftery.diag, 23
read.and.check, 24
read.coda, 25, 26, 27
read.coda.interactive, 26, 26
read.jags (read.coda), 25
read.openbugs, 26, 27
rejectionRate, 28
rejectionRate.mcmc.list
        (rejectionRate), 28
spectrum, 29, 30
spectrum0, 28, 30
spectrum0.ar, 6, 11, 29, 29
start, 32
start.mcmc (time.mcmc), 31
summary.mcmc, 6, 18, 30
thin, 18, 31, 32, 37
thin.mcmc (time.mcmc), 31
time, 31, 32
time.mcmc, 31
topo.colors, 9
traceplot, 23, 32
trellisplots, 32
ts, 21
update, 35
varnames, 36
varnames<- (varnames), 36
window, 37
window.mcmc, 18, 20, 37
xyplot, 35
xyplot.mcmc (trellisplots), 32
```