Package 'bspmma'

January 13, 2019

Title Bayesian Semiparametric Models for Meta-Analysis

Version 0.1-2

Date 2019-01-12

Author Deborah Burr

Maintainer Deborah Burr <burr@stat.ufl.edu>

Description The main functions carry out Gibbs' sampler routines for nonparametric and semiparametric Bayesian models for random effects meta-analysis.

License GPL-2

LazyLoad yes

NeedsCompilation no

Repository CRAN

Date/Publication 2019-01-13 20:20:03 UTC

R topics documented:

---

bspmma-package	*bspmma: Bayesian Semiparametric Models for Meta-Analysis*

Description

Two functions carry out Gibbs' sampler routines to estimate the posterior distributions from either a non-parametric Bayesian model for random effects meta-analysis, or from a semi-parametric model. A group of three functions are used to compute Bayes factors to compare the two models. Three sample datasets are included. There are routines for graphing the posteriors and computing summary statistics.

Details

Package:	bspmma
Version:	0.1-2
Date:	2019-01-19
License:	GPL-2
LazyLoad:	yes
Built:	R 2.9.2; ; 2012-07-13 19:04:37 UTC; unix

Index:

bf.c	Compute Bayes Factors for Comparing Values of the Dirichlet Precision Parameter in the Conditional Dirichlet Model
bf.c.o	Compute Bayes Factors for Conditional vs. Ordinary Dirichlet Models
bf.o	Compute Bayes Factors for Comparing Values of the Dirichlet Precision Parameter in the Ordinary Dirichlet Model
bf1	Generate Chains for Computation of Bayes Factors
bf2	Compute Constants for Multi-Chain Algorithm to Compute Bayes Factors
breast.17	Aspirin and Breast Cancer: 17 studies
bspmma-package	bspmma: Bayesian Semiparametric Models for Meta-Analysis
caprie.3grps	CAPRIE Study: Three Risk Groups
ddtm.s	Decontamination of the Digestive Tract Mortality, Short Dataset
describe.post	Brief summary statistics of the posterior for convenient comparison of several models
dirichlet.c	Mixture of Conditional Dirichlet Model
dirichlet.o	Mixture of Ordinary Dirichlet Model
draw.bf	Plot Function for Bayes Factors

```
draw.post                Overlapping Plots of Posterior Distributions
                         for Several Models
print.dir.cond           printing method for objects of class dir.cond
print.dir.ord            printing method for objects of class dir.ord
```

The main functions are explained in *Burr (2012)*, and are illustrated on the datasets `breast.17` and `ddtm.s`. The function `dirichlet.c` carries out the Markov chain Monte Carlo (MCMC) algorithm to simulate data from the posterior distribution under the conditional Dirichlet model described in *Burr and Doss (2005)*. The computation of Bayes factors is carried out in functions `bf1`, `bf2`, `bf.c`, `bf.o`, and `bf.c.o`, which implement a multi-chain algorithm described in *Doss (2012)*.

Author(s)

Deborah Burr

Maintainer: Deborah Burr <burr@stat.ufl.edu>

References

Burr, Deborah (2012). "bspmma: An R package for Bayesian semi-parametric models for meta-analysis." *Journal of Statistical Software* 50(4), 1–23. http://www.jstatsoft.org/v50/i04/.

Doss, Hani (2012). "Hyperparameter and model selection for nonparametric Bayes problems via Radon-Nikodym derivatives." *Statistica Sinica* 22, 1–26.

---

bf.c	*Compute Bayes Factors for Comparing Values of the Dirichlet Precision Parameter in the Conditional Dirichlet Model*

Description

This function carries out the final step in computing Bayes factors for comparing a sequence of values of the Dirichlet precision parameter M for the conditional Dirichlet mixing model.

Usage

```
bf.c(df=-99, from=.4, incr=.1, to, cc, mat.list)
```

Arguments

df	degrees of freedom for the t distribution in the model; $df = -99$ corresponds to a normal distribution.
from	is the starting value for the sequence of values of the precision parameter M at which to compute the Bayes factor.
incr	is the amount by which to increment the values of M.
to	is the ending value for the sequence of values of M.
cc	is the vector of nine constants computed by `bf1` and `bf2`.
mat.list	list of nine matrices of MCMC output produced by `bf1` for the final computation of the Bayes factors.

Details

This function carries out the fourth and final step in the computation of Bayes factors for the selection of M in the conditional Dirichlet mixing model. In the current version of the package, the Bayes factors for M are computed relative to the model with $M = 4$. The sequence of steps implements a multiple-chain version of Equation (2.6) of Burr (2012); the details of the multiple-chain algorithm are given in Doss (2012). Previous steps are two calls to bf1 and a call to bf2, as illustrated in the Examples section and in Burr (2012).

Value

A list with three named components, Mnew, y, and yinfinity, needed to produce the plot of Bayes factors via the function draw.bf. The vector Mnew is the sequence of (finite) values of M. The vector y is the estimates of the Bayes factors corresponding to the finite values of Mnew, and the object yinfinity is the value of the Bayes factor for M at infinity, that is, for the parametric model.

References

Burr, Deborah (2012). "bspmma: An R package for Bayesian semiparametric models for meta-analysis." *Journal of Statistical Software* 50(4), 1–23. http://www.jstatsoft.org/v50/i04/.

Doss, Hani (2012). "Hyperparameter and model selection for nonparametric Bayes problems via Radon-Nikodym derivatives." *Statistica Sinica* 22, 1–26.

Examples

```
## Not run:
##  CPU times are from runs of the R command system.time() on an
##  Intel $2.8$ GHz Q$9550$ running Linux.
## Preliminary steps

data(breast.17) # the breast cancer dataset
breast.data <- as.matrix(breast.17) # put data in matrix object
chain1.list <- bf1(breast.data) # 40.5 secs
cc <- bf2(chain1.list) # 1.6 secs
## Next get a second set of 9 chains, with a different seed
chain2.list <- bf1(breast.data,seed=2) # 40.4 secs

## Compute and plot the Bayes factors
breast.bfc <- bf.c(to=20, cc=cc, mat.list=chain2.list)
draw.bf(breast.bfc)

## End(Not run)
```

---

bf.c.o   *Compute Bayes Factors for Conditional vs. Ordinary Dirichlet Models*

Description

This function carries out the final step in computing Bayes factors for comparing conditional and ordinary Dirichlet mixing models, for a sequence of Dirichlet precision parameters M.

Usage

```
bf.c.o(df=-99, from=.4, incr=.1, to, cc, mat.list)
```

Arguments

df	degrees of freedom for the t distribution in the model; $df = -99$ corresponds to a normal distribution.
cc	is the vector of nine constants computed by bf1 and bf2.
from	is the starting value for the sequence of values of the precision parameter M at which to compute the Bayes factor.
to	is the ending value for the sequence of values of M.
incr	is the amount by which to increment the values of M.
mat.list	list of nine matrices of MCMC output produced by bf1 for the final computation of the Bayes factors.

Details

This function carries out the fourth and final step in the computation of Bayes factors for the conditional vs. ordinary Dirichlet mixing models. It implements a multiple-chain version of Equation (2.7) of Burr (2012); the details of the multiple-chain algorithm are given in Doss (2012). Previous steps are two calls to bf1 and a call to bf2, as illustrated in the Examples section and in Burr (2012).

Value

A list with two named components, Mnew and y. The vector Mnew is the sequence of (finite) values of M. The vector y is the estimates of the Bayes factors corresponding to Mnew.

References

Burr, Deborah (2012). "bspmma: An R package for Bayesian semiparametric models for meta-analysis." *Journal of Statistical Software* 50(4), 1–23. http://www.jstatsoft.org/v50/i04/.

Doss, Hani (2012). "Hyperparameter and model selection for nonparametric Bayes problems via Radon-Nikodym derivatives." *Statistica Sinica* 22, 1–26.

Examples

```
## Not run:
##  CPU times are from runs of the R command system.time() on an
##  Intel $2.8$ GHz Q$9550$ running Linux.
## Preliminary steps

data(breast.17) # the breast cancer dataset
breast.data <- as.matrix(breast.17) # put data in matrix object
```

```
chain1.list <- bf1(breast.data) # 40.5 secs
cc <- bf2(chain1.list) # 1.6 secs
## Next get a second set of 9 chains, with a different seed
chain2.list <- bf1(breast.data,seed=2) # 40.4 secs

## OR load the chains and constants saved earlier
load("breast-rdat-2lists-1000")
load("breast-rdat-2lists-1000")

## Compute and plot the Bayes factors
breast.bfco <- bf.c.o(to=20, cc=cc, mat.list=chain2.list) # 107 secs
draw.bf(breast.bfco)

## End(Not run)
```

---

bf.o	*Compute Bayes Factors for Comparing Values of the Dirichlet Precision Parameter in the Ordinary Dirichlet Model*

Description

This function carries out the final step in computing Bayes factors for comparing a sequence of values of the Dirichlet precision parameter M for the ordinary Dirichlet mixing model.

Usage

```
bf.o(df=-99, from=.4, incr=.1, to, cc, mat.list)
```

Arguments

df	degrees of freedom for the t distribution in the model; $df = -99$ corresponds to a normal distribution.
from	is the starting value for the sequence of values of the precision parameter M at which to compute the Bayes factor.
incr	is the amount by which to increment the values of M.
to	is the ending value for the sequence of values of M.
cc	is the vector of nine constants computed by bf1 and bf2 prior to this step in the algorithm.
mat.list	list of nine matrices of MCMC output produced by bf1 for the final computation of the Bayes factors.

Details

This function carries out the fourth and final step in the computation of Bayes factors for the selection of M in the ordinary Dirichlet mixing model. In the current version of the package, the Bayes factors for M are computed relative to the model with $M = 4$. The sequence of steps implements a multiple-chain version of Equation (2.7) of Burr (2012); the details of the multiple-chain algorithm are given in Doss (2012). Previous steps are calls to bf1, bf2, and bf1 again, in that order, as illustrated in the Examples section and in Burr (2012).

Value

A list with three named components, Mnew, y, and yinfinity, needed to produce the plot of Bayes factors via the function draw.bf. The vector Mnew is the sequence of (finite) values of M. The vector y is the estimates of the Bayes factors corresponding to the finite values of Mnew, and the object yinfinity is the value of the Bayes factor for M at infinity, that is, for the parametric model.

References

Burr, Deborah (2012). "bspmma: An R package for Bayesian semi-parametric models for meta-analysis." *Journal of Statistical Software* 50(4), 1–23. http://www.jstatsoft.org/v50/i04/.

Doss, Hani (2012). "Hyperparameter and model selection for nonparametric Bayes problems via Radon-Nikodym derivatives." *Statistica Sinica* 22, 1–26.

Examples

```
## Not run:
##  CPU times are from runs of the R command system.time() on an
##  Intel $2.8$ GHz Q$9550$ running Linux.
## Preliminary steps

data(breast.17) # the breast cancer dataset
breast.data <- as.matrix(breast.17) # put data in matrix object
chain1.list <- bf1(breast.data) # 40.5 secs
cc <- bf2(chain1.list) # 1.6 secs
## Next get a second set of 9 chains, with a different seed
chain2.list <- bf1(breast.data,seed=2) # 40.4 secs

## Compute and plot the Bayes factors
breast.bfo <- bf.o(to=20, cc=cc, mat.list=chain2.list) #51 secs
draw.bf(breast.bfo)

## End(Not run)
```

---

bf1   *Generate Chains for Computation of Bayes Factors*

Description

Generate nine matrices of MCMC output under the ordinary Dirichlet model, for nine fixed values of the precision parameter M. This MCMC output is needed for computing Bayes factors.

Usage

```
bf1(data,seed=1,ncycles=2000,d=c(.1,.1,0,1000),K=10,burnin=1000)
```

Arguments

data	is a two-column matrix with a row for each study in the meta-analysis. The first column is the log of estimate of relative risk, often a log(odds ratio). The second column is the true or estimated standard error of the log(odds ratio).
seed	is the value of the seed for starting the random number generator, which will be used before each of the nine calls to the function `dirichlet.o`.
ncycles	is the number of cycles of the Markov chain.
d	is a vector of length four with the values of the hyperparameters, in order, the location and scale of the Gamma inverse prior, mean and variance multiplier for the normal prior on μ.
K	is the number of summands to include when one uses Sethuraman's (1994) representation for getting the parameter $\eta = \text{mean}(F)$. If you do not intend to use this parameter, then take K small, say $K = 10$.
burnin	is the number of Markov chain cycles to drop.

Details

Doss (2012) describes a method for estimating Bayes factors for many M values in a Dirichlet mixing model; the method requires judicious selection of multiple hyperparameter points at which to estimate the posterior distribution by MCMC under the ordinary Dirichlet model. The function bf1 is used for estimating Bayes factors for conditional vs.\ ordinary Dirichlet models, and for comparing values of M in the conditional model or in the ordinary model, for a range of the precision parameter M which cover the range of values of interest in most practical problems. The function bf1 generates the MCMC output for a hard-wired selection of hyperparameters which work well to give low-variance estimates of Bayes factors of interest in practice. Chains are generated for nine values of the Dirichlet precision parameter M: $.25, .5, 1, 2, 4, 8, 16, 32$, and 64. The rest of the Dirichlet model is specified by the parameters of the normal/inverse Gamma prior, which by default are $\vec{d} = (.1, .1, 0, 1000)$.

Value

List with nine matrix components. Each matrix has nr rows and nc columns, where $nr = \text{ncycles} - \text{burnin}$, $nc = \text{(number of studies)} + 4$ for the row label, the individual study parameter values, and the three overall parameters, μ, τ, and η.

References

Doss, Hani (2012). "Hyperparameter and model selection for nonparametric Bayes problems via Radon-Nikodym derivatives." *Statistica Sinica*, 22, 1–26.

Sethuraman, J. (1994). "A constructive definition of Dirichlet priors." *Statistica Sinica 4*, 639–650.

Examples

```
## Not run:
## Set up the data.

data(breast.17) # the breast cancer dataset
```

```
breast.data <- as.matrix(breast.17) # put data in matrix object

## Default values ncycles=2000, burnin=1000, seed=1
##  CPU time is given from a run of the R command system.time() on an
##  Intel $2.8$ GHz Q$9550$ running Linux
chain1.list <- bf1(breast.data) # 40.5 secs
## Next get a second set of 9 chains, with a different seed
chain2.list <- bf2(breast.data, seed=2) # 40.4 secs

## Perhaps save for another time.
save(chain1.list,chain2.list,file="breast-rdat-2lists-1000",compress=TRUE)

## later session
load("breast-rdat-2lists-1000")


## End(Not run)
```

---

bf2	*Compute Constants for Multi-Chain Algorithm to Compute Bayes Factors*

Description

This function computes nine constants needed in the multi-chain algorithm for Bayes factors comparing conditional and ordinary Dirichlet mixing models, and for Bayes factors comparing Dirichlet precision parameter (M) values for the conditional model, or for the ordinary model.

Usage

```
bf2(chain.list)
```

Arguments

chain.list   is a list of nine matrices of MCMC output produced by function bf1

Details

This function computes the constants needed for the denominator of the left-side of Eqn. (2.5) of Doss (2012). This is the step in which Radon-Nikodym derivatives are evaluated for each line of MCMC output and then averaged to estimate the constants. The actual algorithm is a little more complicated than that to make use of output from multiple chains. Nine constants are computed in this way.

Value

A vector of nine constants which is needed in the next step of the computation of the Bayes factors. Burr (2012) gives detailed explanations of the algorithm and illustrates the steps in the algorithm.

References

Burr, Deborah (2012). "bspmma: An R package for Bayesian semi-parametric models for meta-analysis." *Journal of Statistical Software* 50(4), 1–23. http://www.jstatsoft.org/v50/i04/.

Doss, Hani (2012). "Hyperparameter and model selection for nonparametric Bayes problems via Radon-Nikodym derivatives." *Statistica Sinica*, 22, 1–26.

Examples

```
## Not run:
## Get the two sets of chains saved from runs of bf1 from
## the breast cancer example in the help file for bf1.
load("breast-rdat-2lists-1000")

## Default values ncycles=2000, burnin=1000
##  CPU time is from a run of the R command system.time() on an
##  Intel $2.8$ GHz Q$9550$ running Linux.
cc <- bf2(chain1.list) #1.6 secs

## Perhaps save for another time.
save(cc,file="breast-rdat-constants",compress=TRUE)

## Next session
load("breast-rdat-constants")


## End(Not run)
```

---

breast.17   *Aspirin and Breast Cancer: 17 studies*

Description

This dataset gives log odds of breast cancer for long-term aspirin users, and its standard error, derived from 17 cohort and case-control studies.

Usage

```
data(breast.17)
```

Format

A data frame with seventeen rows, corresponding to the seventeen papers. There are two columns: psi.hat (numeric, the log odds ratio), and se.psi.hat (numeric, estimated SE of the log odds ratio). The rownames attribute gives the first author of the paper and the citation number of the study in *Harris et.\ al. (2005).*

Source

Harris, R., Beebe-Donk, J., Doss, H., and Burr, D. (2005). "Aspirin, ibuprofen, and other non-steroidal anti-inflammatory drugs in cancer prevention: A critical review of non-selective COX-2 blockade (Review)." *Oncology Reports 13* 559-583.

References

Gonzalez-Perez, A., Rodriguez, L., and Lopez-Ridaura, R. (2003). "Effects of non-steroidal anti-inflammatory drugs on cancer sites other than the colon and rectum: a meta-analysis." *BMC Cancer 3* 28.

Khuder, S. and Mutgi, A. (2001). "Breast cancer and NSAID use: a meta-analysis." *British Journal of Cancer 84*, 1188–1192.

---

caprie.3grps	*CAPRIE Study: Three Risk Groups*

Description

From the CAPRIE study comparing clopidogrel versus Aspirin, this dataset gives risk ratios and their SEs separately for patients who had stroke, heart attack (myocardial infarction or MI), and peripheral arterial disease (PAD).

Usage

```
data(caprie.3grps)
```

Format

A data frame with three rows, corresponding to the three risk groups. There are three columns: study (character, for the risk groups), psi.hat (numeric, the log odds ratio, and se.psi.hat (numeric, estimated SE of the log odds ratio).

Source

CAPRIE Steering Committee (1996), A randomized, blinded trial of clopidogrel versus aspirin in patients at risk of ischaemic events (CAPRIE), *Lancet*, 348, 1329–1339.

Burr and Doss (2005) state how the SEs for the log odds ratios for the three risk groups are obtained from data on p.\ 1334 of the CAPRIE article.

References

Burr, D. and Doss, H. (2005). A Bayesian semi-parametric model for random effects meta analysis. *The Journal of the American Statistical Association*,100, 242–251.

---

ddtm.s   *Decontamination of the Digestive Tract Mortality, Short Dataset*

Description

This dataset gives number of deaths and number of subjects in treatment vs. control groups in 14 studies from a meta-analysis of studies on antiobiotics to reduce infection in intensive-care units.

Usage

```
data(ddtm.s)
```

Format

A data frame with fourteen rows, corresponding to fourteen different, unidentified studies. There are four columns: number of deaths in the treatment group, sample size in the treatment group, number of deaths in the control group, sample size in the control group.

Burr and Doss (2005) give the background for this dataset.

Source

Selective Decontamination of the Digestive Tract Trialists' Collaborative Group (1993). "Meta-analysis of randomised controlled trials of selective decontamination of the digestive tract." *British Medical Journal 307* 525–532.

References

Burr, D. and Doss, H. (2005). "A Bayesian semiparametric model for random-effects meta-analysis." *Journal of the American Statistical Association 100* 242–251.

---

describe.post   *Brief summary statistics of the posterior for convenient comparison of several models*

Description

Compute, print posterior means and posterior P(odds ratio < 1) for the individual study parameters and hyperparameters of the model.

Usage

```
describe.post(mcout,burnin=1000)
```

Arguments

mcout   is a list. Each item in the list is a matrix of MCMC output, corresponding to
                 different values of M, the precision parameter of the Dirichlet model. If the
                 matrices are output from `dirichlet.c`, each matrix has $ncycles + 1$ rows and
                 $m + 2$ columns, where m is the number of studies in the meta-analysis and
                 $ncycles$ is the number of runs of the Markov chain. The matrix output from the
                 ordinary Dirichlet model function `dirichlet.o` is similar but has an additional
                 column. The rows hold output from separate Markov chain runs (the first row is
                 the initial values.) Columns 1 through m hold the individual study parameters,
                 the ψ_i's. The next two columns hold the mean and standard deviation parameters
                 of the centering normal distribution of the Dirichlet prior, μ and τ. In the case of
                 the ordinary Dirichlet model, an additional column is added to hold the values
                 of η.

burnin   is the number of initial chains to omit from the estimates.

Value

List with two named components, means.table and probs.table, returned invisibly.

Examples

```
## Not run:
## Set up the data.

data(breast.17) # the breast cancer dataset
breast.data <- as.matrix(breast.17) # put data in matrix object

## Generate at least two chains, from models which are the same except
## for different \eqn{M} values.

set.seed(1) # initialize the seed at 1
breast.c1 <- dirichlet.c(breast.data, ncycles=4000, M=5)
breast.c2 <- dirichlet.c(breast.data,ncycles=4000, M=1000)

## Create list object.
breast.c1c2 <- list("5"=breast.c1$chain, "1000"= breast.c2$chain)

## Decide on some number of initial runs to omit from the analysis.
describe.post(breast.c1c2, burnin=100)

## End(Not run)
```

---

dirichlet.c   *Mixture of Conditional Dirichlet Model*

Description

MCMC generation of posterior distributions for the conditional Dirichlet mixing distribution model,
using $m + 1$-cycle Gibbs sampler

Usage

```
dirichlet.c(data, ncycles=10, M=1,d=c(.1,.1,0,1000),
            start=NULL)
```

Arguments

`data`	is a two-column matrix with a row for each study in the meta-analysis. The first column is the log of estimate of relative risk, often a log(odds ratio). The second column is the true or estimated standard error of the log(odds ratio).
`ncycles`	is the number of cycles of the Markov chain.
`M`	is the precision parameter of the Dirichlet process prior.
`d`	is a vector of length four with the values of the hyperparameters, in order, the location and scale of the Gamma inverse prior, mean and variance multiplier for the normal prior on μ.
`start`	is an optional vector containing starting values for the parameters, $\psi_i, i = 1, \ldots, m$ where m is the number of studies in the meta-analysis, μ and τ.

Details

This function generates MCMC output for the posterior distribution for the parameters $\psi_i, i = 1, \ldots, m$ where m is the number of studies in the meta-analysis, μ and τ, in the conditional Dirichlet mixing model for random-effects meta-analysis. Notation is taken from *Burr (2012)*, Model 4.

The MCMC algorithm for estimating the posterior under this model is given in *Burr and Doss (2005)*. The chain is a $(m+1)$-cycle Gibbs sampler which cycles through the vector of ψ_i's and the pair μ, τ, and the main part of the computational burden is in the first part of the cycle, the generation of the vector of ψ_i's.

If starting values are not specified via the argument `start`, the default values are used, which are based on the data. The study estimates are the starting values for the $\psi_i, i = 1, \ldots, m$, and the mean and standard deviation of the study estimates are the starting values for μ and τ, respectively.

Value

`call`	the call that resulted in this object
`ncycles`	the number of cycles in the Markov chain
`M`	the value of the precision parameter for the conditional Dirichlet model
`prior`	the vector length four of hyperparameters
`chain`	A matrix with `ncycles` $+1$ rows and $m+2$ columns, where m is the number of studies in the meta-analysis. The rows hold output from the Markov chain runs (the first row is the initial values). Columns 1 through m hold the individual study parameters, the ψ_i's. The final two columns hold the mean and standard deviation parameters of the centering normal distribution of the Dirichlet prior, μ and τ.
`start.user`	logical, TRUE if the user supplied initial values of the parameter vector, FALSE if input argument `start` was not specified by the user.
`start`	vector of initial parameter values used in the MCMC algorithm, whether this was the default or was user-supplied

References

Burr, Deborah (2012). "bspmma: An R package for Bayesian semi-parametric models for meta-analysis." *Journal of Statistical Software* 50(4), 1–23. http://www.jstatsoft.org/v50/i04/.

Burr, D. and Doss, H. (2005). "A Bayesian semiparametric model for random-effects meta-analysis." *Journal of the American Statistical Association 100* 242–251.

Sethuraman, J. (1994). "A constructive definition of Dirichlet priors." *Statistica Sinica 4*, 639–650.

Examples

```
## Not run:
data(breast.17) # the dataset
breast.data <- as.matrix(breast.17) # put data in matrix object
set.seed(1) # initialize the seed at 1 for test purposes
breast.c1 <- dirichlet.c(breast.data, ncycles=4000, M=5)
breast.c2 <- dirichlet.c(breast.data,ncycles=4000, M=1000)

## End(Not run)
```

---

dirichlet.o	*Mixture of Ordinary Dirichlet Model*

Description

MCMC generation of posterior distributions for the usual (unconditional) Dirichlet mixing distribution model, using an $m + 1$-cycle Gibbs sampler

Usage

```
dirichlet.o(data, ncycles=10, M=1,d=c(.1,.1,0,1000),
            start=NULL,K=100)
```

Arguments

data
: is a two-column matrix with a row for each study in the meta-analysis. The first column is the log of estimate of relative risk, often a log(odds ratio). The second column is the true or estimated standard error of the log(odds ratio).

ncycles
: is the number of cycles of the Markov chain.

M
: is the precision parameter of the Dirichlet process prior.

d
: is a vector of length four with the values of the hyperparameters, in order, the location and scale of the Gamma inverse prior, mean and variance multiplier for the normal prior on mu.

start
: is an optional vector containing starting values for the parameters, $\psi_i, i = 1, \ldots, m$ where m is the number of studies in the meta-analysis, μ and τ.

K
: is the number of summands to include when one uses Sethuraman's (1994) representation for getting the parameter $\eta = \text{mean}(F)$. If you do not intend to use this parameter, then take K small, say $K = 10$.

Details

This function generates MCMC output for the posterior distribution for the parameters $\psi_i, i = 1, ..., m$ where m is the number of studies in the meta-analysis, μ, τ, and η in the ordinary Dirichlet mixing model for random-effects meta-analysis. Notation is taken from *Burr (2012)*, Model 2 and 3.

The MCMC algorithm for estimating the posterior under this model is given in *Burr and Doss (2005)*. The chain is a $(m + 1)$-cycle Gibbs sampler which cycles through the vector of ψ_i's and the triple μ, τ, η, and the main part of the computational burden is in the first part of the cycle, the generation of the vector of ψ_i's.

Value

`call`	the call that resulted in this object
`ncycles`	the number of cycles in the Markov chain
`M`	the value of the precision parameter for the conditional Dirichlet model
`prior`	the vector length four of hyperparameters
`chain`	A matrix with `ncycles` $+1$ rows and $m + 3$ columns, where m is the number of studies in the meta-analysis. The rows hold output from the Markov chain runs (the first row is the initial values). Columns 1 through m hold the individual study parameters, the ψ_i's. The next two columns hold the mean and standard deviation parameters of the centering normal distribution of the Dirichlet prior, μ and τ, and the final column holds the parameter η. In the ordinary Dirichlet mixing model, the parameter μ does not equal the mean of the distribution F of the ψ_i's; we denote this mean η and estimate it by Sethuraman's (1994) method.
`start.user`	logical, TRUE if the user supplied initial values of the parameter vector, FALSE if input argument `start` was not specified by the user.
`start`	vector of initial parameter values, whether default or user-supplied

References

Burr, Deborah (2012). "bspmma: An R package for Bayesian semi-parametric models for meta-analysis." *Journal of Statistical Software* 50(4), 1–23. http://www.jstatsoft.org/v50/i04/.

Burr, D. and Doss, H. (2005). "A Bayesian semiparametric model for random-effects meta-analysis." *Journal of the American Statistical Association 100* 242–251.

Sethuraman, J. (1994). "A constructive definition of Dirichlet priors." *Statistica Sinica 4*, 639–650.

Examples

```
## Not run:
data(breast.17) # the dataset
breast.data <- as.matrix(breast.17) # put data in matrix object
set.seed(1) # initialize the seed at 1
diro <- dirichlet.o(breast.data, ncycles=4000, M=5)

## a short description of the model and Markov chain
print(diro1)
```

```
## the last mcmc cycle
diro$mcmc[4001,]

## End(Not run)
```

---

draw.bf   *Plot Function for Bayes Factors*

Description

This function plots the output from function bf.c.o.

Usage

```
draw.bf(obj,line.lwd=2, ...)
```

Arguments

obj   is a list with two elements, the vectors x and y to be plotted in a scatterplot,
               which are produced by function bf.c.o. The first element is the values of M
               to go on the x-axis, and the second element is the Bayes factors calculated by
               bf.c.o.

line.lwd   graphical parameter; the relative thickness of the plotted line

...   additional graphical parameters for the overall plot

---

draw.post   *Overlapping Plots of Posterior Distributions for Several Models*

Description

Draw overlapping kernel density estimates of the posterior distributions of the parameters of the
conditional or ordinary Dirichlet model, where the posteriors arise from different values of the
Dirichlet precision parameter M.

Usage

```
draw.post(mcout,burnin=1000,ind.par=NULL,adjust=1,...)
```

Arguments

mcout	is a list. Each item in the list is a matrix of MCMC output, corresponding to different values of M, the precision parameter of the Dirichlet model. If the matrices are output from dirichlet.c, each matrix has $ncycles + 1$ rows and $m + 2$ columns, where m is the number of studies in the meta-analysis and $ncycles$ is the number of runs of the Markov chain. The matrix output from the ordinary Dirichlet model function dirichlet.o is similar but has an additional column. The rows hold output from separate Markov chain runs (the first row is the initial values.) Columns 1 through m hold the individual study parameters, the ψ_i's. The next two columns hold the mean and standard deviation parameters of the centering normal distribution of the Dirichlet prior, μ and τ. In the case of the ordinary Dirichlet model, an additional column is added to hold the values of η.
burnin	is the number of initial chains to omit from the estimates, must be no larger than $ncycles - 10$.
ind.par	an integer vector, containing indices of which columns of mcout to include in the plots. By default, only the posterior density estimates for the hyperparameters μ and τ
adjust	is the bin width argument for the call to the R base package function density.
...	additional arguments to plot may be supplied.

Examples

```
## Not run:
## Conditional Dirichlet model

## Set up the breast cancer dataset.

data(breast.17)
breast.data <- as.matrix(breast.17) # Data must be a matrix object.

##  Generate at least two chains, from models which are the same except
## for different \eqn{M}{M} values.

set.seed(1) # initialize the seed at 1 for test purposes
breast.c1 <- dirichlet.c(breast.data, ncycles=4000, M=5)
breast.c2 <- dirichlet.c(breast.data,ncycles=4000, M=1000)

##  Create list object.

breast.c1c2 <- list("5"=breast.c1$chain, "1000"= breast.c2$chain)

##  Decide on some number of initial runs to omit from the analysis.

draw.post(breast.c1c2, burnin=100) # plots for hyperparameters only

## End(Not run)
```

---

print.dir.cond *printing method for objects of class dir.cond*

Description

gives limited info

Usage

```
## S3 method for class 'dir.cond'
print(x,digits=max(3,getOption("digits") -3),...)
```

Arguments

x	object of class dir.cond
digits	positive integer, number of digits for printing
...	place holder

---

print.dir.ord *printing method for objects of class dir.ord*

Description

gives limited info

Usage

```
## S3 method for class 'dir.ord'
print(x,digits=max(3,getOption("digits") -3),...)
```

Arguments

x	object of class dir.ord
digits	positive integer, number of digits for printing
...	place holder

Index