

# Package ‘mombf’

May 26, 2017

**Version** 1.9.5

**Date** 2017-05-26

**Title** Moment and Inverse Moment Bayes Factors

**Author** David Rossell, John D. Cook, Donatello Telesca, P. Roebuck

**Maintainer** David Rossell <rosselldavid@gmail.com>

**Depends** R (>= 2.14.0), methods, mvtnorm, ncvreg, actuar, mgcv

**Imports** survival

**Suggests** parallel

**Description** Model selection and parameter estimation based on non-local and Zellner priors. Bayes factors, marginal densities and variable selection in regression setups. Routines to sample, evaluate prior densities, distribution functions and quantiles are included.

**License** GPL (>= 2)

**URL** <http://mombf.r-forge.r-project.org/>

**LazyLoad** yes

**Collate** AllClasses.R AllGenerics.R alapl.R bms\_ortho.R  
derivatives\_nlps.R msPriorSpec.R modelsearch.R modelSelection.R  
nlpMarginal.R dmom.R mombf.R pmomLM.R pmomPM.R emomLM.R  
postMode.R pplProbit.R greedyGLM.R ppmodel.R zellnerLM.R rmom.R  
cox.R

**NeedsCompilation** yes

**Repository** CRAN

**Date/Publication** 2017-05-26 09:55:27 UTC

## R topics documented:

bbPrior	2
dalapl	3
dmom	4
eprod	6
hald	7

mode2g . . . . .	7
modelSelection . . . . .	9
mombf . . . . .	13
momknown . . . . .	15
msfit-class . . . . .	17
msPriorSpec-class . . . . .	19
nlp marginals . . . . .	21
pmomLM . . . . .	23
postModeOrtho . . . . .	27
postProb . . . . .	29
rnlp . . . . .	30
<b>Index</b>	<b>33</b>

---

bbPrior	<i>Priors on model space for variable selection problems</i>
---------	--

---

## Description

unifPrior implements a uniform prior (equal a priori probability for all models). binomPrior implements a Binomial prior. bbPrior implements a Beta-Binomial prior.

## Usage

```
unifPrior(sel, logscale = TRUE)
binomPrior(sel, prob = 0.5, logscale = TRUE)
bbPrior(sel, alpha = 1, beta = 1, logscale = TRUE)
```

## Arguments

sel	Logical vector indicating which variables are included in the model
logscale	Set to TRUE to return the log-prior probability.
prob	Success probability for the Binomial prior
alpha	First parameter of the Beta-Binomial prior, which is equivalent to specifying a Beta(alpha,beta) prior on prob.
beta	First parameter of the Beta-Binomial prior, which is equivalent to specifying a Beta(alpha,beta) prior on prob.

## Value

Prior probability of the specified model

## Author(s)

David Rossell

**Examples**

```
library(mombf)
sel <- c(TRUE,TRUE,FALSE,FALSE)
unifPrior(sel,logscale=FALSE)
binomPrior(sel,prob=.5,logscale=FALSE)
bbPrior(sel,alpha=1,beta=1,logscale=FALSE)
```

---

dalapl	<i>Density and random draws from the asymmetric Laplace distribution</i>
--------	--

---

**Description**

dalapl evaluates the probability density function, palapl the cumulative probability function and ralapl generates random draws.

**Usage**

```
dalapl(x, th=0, scale=1, alpha=0, logscale=FALSE)
```

```
palapl(x, th=0, scale=1, alpha=0)
```

```
ralapl(n, th=0, scale=1, alpha=0)
```

**Arguments**

x	Vector of values at which to evaluate the pdf/cdf
n	Number of random draws
th	Location parameter (mode)
scale	Scale parameter (proportional to variance)
alpha	Asymmetry parameter, must be between -1 and 1
logscale	If TRUE the log-pdf is returned

**Details**

For  $x \leq th$  the asymmetric Laplace pdf is

$$0.5 * \exp(-\text{abs}(th-x)/(\text{sqrt}(\text{scale}) * (1+\text{alpha}))) / \text{sqrt}(\text{scale})$$

and for  $x > th$  it is

$$0.5 * \exp(-\text{abs}(th-x)/(\text{sqrt}(\text{scale}) * (1-\text{alpha}))) / \text{sqrt}(\text{scale})$$

**Value**

dalapl returns the density function, palapl the cumulative probability, ralapl random draws.

**Author(s)**

David Rossell

**Examples**

```
library(mombf)
e <- rlapl(n=10^4, th=1, scale=2, alpha=0.5)
thseq <- seq(min(e),max(e),length=1000)
hist(e, main='', breaks=30, prob=TRUE)
lines(thseq, dalapl(thseq, th=1, scale=2, alpha=0.5), col=2)
```

dmom

*Moment prior and inverse moment prior.***Description**

dmom, dimom and demom return the density for the moment, inverse moment and exponential moment priors. pmom, pimom and pemom return the distribution function for the univariate moment, inverse moment and exponential moment priors (respectively). qmom and qimom return the quantiles for the univariate moment and inverse moment priors.

**Usage**

```
dmom(x, tau, a.tau, b.tau, phi=1, r=1, V1, baseDensity='normal', nu=3,
logscale=FALSE, penalty='product')
dimom(x, tau=1, phi=1, V1, logscale=FALSE, penalty='product')
demom(x, tau, a.tau, b.tau, phi=1, logscale=FALSE)

pmom(q, V1 = 1, tau = 1)
pimom(q, V1 = 1, tau = 1, nu = 1)
pemom(q, tau, a.tau, b.tau)

qmom(p, V1 = 1, tau = 1)
qimom(p, V1 = 1, tau = 1, nu = 1)
```

**Arguments**

x	In the univariate setting, x is a vector with the values at which to evaluate the density. In the multivariate setting it is a matrix with an observation in each row.
q	Vector of quantiles.
p	Vector of probabilities.
V1	Scale matrix (ignored if penalty=='product'). Defaults to 1 in univariate setting and the identity matrix in the multivariate setting.
tau	Prior dispersion parameter is tau*phi. See details.
a.tau	If tau is left missing, an Inverse Gamma(a.tau/2,b.tau/2) is placed on tau. In this case dmom and demom return the density marginalized with respect to tau.
b.tau	See a.tau.
phi	Prior dispersion parameter is tau*phi. See details.

r	Prior power parameter for MOM prior is $2*r$
baseDensity	For baseDensity=='normal' a normal MOM prior is used, for baseDensity=='t' a T MOM prior with nu degrees of freedom is used.
nu	Prior parameter indicating the degrees of freedom for the quadratic T MOM and iMOM prior densities. The tails of the inverse moment prior are proportional to the tails of a multivariate T with nu degrees of freedom.
penalty	penalty=='product' indicates that product MOM/iMOM should be used. penalty=='quadratic' indicates quadratic iMOM. See Details.
logscale	For logscale==TRUE, dmom returns the natural log of the prior density.

### Details

For type=='quadratic' the density is as follows. Define the quadratic form  $q(\theta) = (\theta - \theta_0)' * \text{solve}(V_1) * (\theta - \theta_0) / (\tau * \phi)$ . The normal moment prior density is proportional to  $q(\theta) * \text{dmvnorm}(\theta, \theta_0, \tau * \phi * V_1)$ . The T moment prior is proportional to  $q(\theta) * \text{dmvt}(\theta, \theta_0, \tau * \phi * V_1, \nu)$ . The inverse moment prior density is proportional to  $q(\theta)^{-(\nu+d)/2} * \exp(-1/q(\theta))$ .

pmom, pimom and qimom use closed-form expressions, while qmom uses nlminb to find quantiles numerically. Only the univariate version is implemented. In this case the product MOM is equivalent to the quadratic MOM. The same happens for the iMOM.

Only the product eMOM prior is implemented.

### Value

dmom returns the value of the moment prior density. dimom returns the value of the inverse moment prior density.

### Author(s)

David Rossell

### References

Johnson V.E., Rossell D. Non-Local Prior Densities for Default Bayesian Hypothesis Tests. Journal of the Royal Statistical Society B, 2010, 72, 143-170.

Johnson V.E., Rossell D. Bayesian model selection in high-dimensional settings. Technical report. 2011

See <http://rosselldavid.googlepages.com> for technical reports.

### See Also

[g2mode](#) to find the prior mode corresponding to a given g. [mode2g](#) to find the g value corresponding to a given prior mode.

**Examples**

```
#evaluate and plot the moment and inverse moment priors
library(mombf)
tau <- 1
thseq <- seq(-3,3,length=1000)
plot(thseq,dmom(thseq,tau=tau),type='l',ylab='Prior density')
lines(thseq,dimom(thseq,tau=tau),lty=2,col=2)
```

---

eprod

---

*Expectation of a product of powers of Normal or T random variables*


---

**Description**

Compute the mean of  $\text{prod}(x)^{\text{power}}$  when  $x$  follows  $T_{\text{dof}}(\mu, \sigma)$  distribution (dof= -1 for multivariate Normal).

**Usage**

```
eprod(m, S, power = 1, dof = -1)
```

**Arguments**

m	Location parameter
S	Scale matrix. For multivariate T with dof>2 the covariance is $S \cdot \text{dof} / (\text{dof} - 2)$ . For the multivariate Normal the covariance is S.
power	Power that the product is raised to
dof	Degrees of freedom of the multivariate T. Set to -1 for the multivariate Normal.

**Details**

The calculation is based on the computationally efficient approach by Kan (2008).

**Value**

Expectation of the above-mentioned product

**Author(s)**

John Cook

**References**

Kan R. From moments of sum to moments of product. *Journal of Multivariate Analysis* 99 (2008), 542-554.

**Examples**

```
#Check easy independence case
m <- c(0,3); S <- matrix(c(2,0,0,1),ncol=2)

eprod(m, S, power=2)

(m[1]^2+S[1][1])*(m[2]^2+S[2][2])
```

hald

*Hald Data***Description**

Montgomery and Peck (1982) illustrated variable selection techniques on the Hald cement data and gave several references to other analysis. The response variable  $y$  is the *heat evolved* in a cement mix. The four explanatory variables are ingredients of the mix, i.e.,  $x_1$ : *tricalcium aluminate*,  $x_2$ : *tricalcium silicate*,  $x_3$ : *tetracalcium alumino ferrite*,  $x_4$ : *dicalcium silicate*. An important feature of these data is that the variables  $x_1$  and  $x_3$  are highly correlated ( $\text{corr}(x_1, x_3) = -0.824$ ), as well as the variables  $x_2$  and  $x_4$  (with  $\text{corr}(x_2, x_4) = -0.975$ ). Thus we should expect any subset of  $(x_1, x_2, x_3, x_4)$  that includes one variable from highly correlated pair to do as any subset that also includes the other member.

**Usage**

```
data(hald)
```

**Format**

hald is a matrix with 13 observations (rows) and 5 variables (columns), the first column is the dependent variable.  $y$ .hald and  $x$ .hald are also availables.

**Source**

Montgomery, D.C., Peck, E.A. (1982) *Introduction to linear regression analysis*, John Wiley, New York.

mode2g

*Moment and inverse moment prior elicitation.***Description**

mode2g finds the  $g$  value corresponding to a given prior mode. g2mode finds the prior mode corresponding to a given  $g$  value. priorp2g finds the  $g$  value giving priorp prior probability to the interval  $(-q, q)$ .

All routines operate in the standardized effect sizes scale.

**Usage**

```
mode2g(prior.mode, prior=c("iMom", "normalMom", "tMom"), nu=1, dim=1)
g2mode(g, prior=c("iMom", "normalMom", "tMom"), nu=1, dim=1)
priorp2g(priorp, q, nu=1, prior=c("iMom", "normalMom", "tMom"))
```

**Arguments**

prior.mode	Prior mode for the quadratic form $(\theta - \theta_0)' * \text{solve}(\Sigma) * (\theta - \theta_0) / (n * g * \sigma^2)$ , where $\sigma$ is the dispersion parameter and $\Sigma$ is given by the design matrix.
prior	prior=='normalMom' does computations for the normal moment prior, prior=='tMom' for the T moment prior, prior=='iMom' does computations for the inverse moment prior. Currently prior=='tMom' is not implemented in priorp2g.
nu	Prior degrees of freedom for the T moment prior or the iMom prior (ignored if prior=='normalMom').
dim	Dimensionality of the parameter, i.e. dim==1 for univariate, dim==2 for bivariate and so on.
g	Prior parameter. See dimom for details.
priorp	priorp2g returns g giving priorp prior probability to the interval $(-q, q)$ .
q	priorp2g returns g giving priorp prior probability to the interval $(-q, q)$ .

**Details**

See dmom and dimom for details on the meaning of the prior parameters.

**Value**

mode2g returns the value of the prior parameter g matching the given mode.

g2mode returns the prior mode for a given prior parameter g.

priorp2g returns g giving priorp prior probability to the interval  $(-q, q)$ .

**Author(s)**

David Rossell <rosselldavid@gmail.com>

**References**

See <http://rosselldavid.googlepages.com> for technical reports.

**See Also**

[dmom](#), [dimom](#), [mombf](#), [imombf](#)



**Examples**

```

#find g value giving a prior mode for (theta/(sigma*n*Sigma))^2 at 0.2^2
data(hald)
lm1 <- lm(hald[, 1] ~ hald[, 2] + hald[, 3] + hald[, 4] + hald[, 5])
prior.mode <- .2
gmom <- mode2g(prior.mode^2, prior='normalMom')
gtmom <- mode2g(prior.mode^2, prior='tMom', nu=3)
gimom <- mode2g(prior.mode^2, prior='iMom')
gmom
gimom

#find g value giving 0.05 probability to interval (-.2,.2)
priorp <- .05; q <- .2
gmom <- priorp2g(priorp=priorp, q=q, prior='normalMom')
gimom <- priorp2g(priorp=priorp, q=q, prior='iMom')
gmom
gimom

```

---

modelSelection

*Bayesian variable selection for linear models via non-local priors.*


---

**Description**

Bayesian model selection for linear, asymmetric linear, median and quantile regression under non-local or Zellner priors.  $p \gg n$  can be handled.

modelSelection enumerates all models when feasible and uses a Gibbs scheme otherwise. See rnlp for posterior samples for the coefficients.

modelsearchBlockDiag seeks the highest posterior probability model using an iterative block search.

**Usage**

```

modelSelection(y, x, center=TRUE, scale=TRUE,
  enumerate= ifelse(ncol(x)<15,TRUE,FALSE),
  includevars=rep(FALSE,ncol(x)),
  maxvars= ncol(x),
  niter=10^4, thinning=1,
  burnin=round(niter/10), family='normal', priorCoef=momprior(tau=0.348),
  priorDelta=modelbbprior(alpha.p=1,beta.p=1),
  priorVar=igprior(alpha=.01,lambd=.01),
  priorSkew=momprior(tau=0.348), phi, deltaini=rep(FALSE,ncol(x)),
  initSearch='greedy', method='auto', hess='asyp', optimMethod='CDA',
  B=10^5, verbose=TRUE)

```

```

modelsearchBlockDiag(y, x, priorCoef=momprior(tau=0.348),
  priorDelta=modelbbprior(1,1), priorVar=igprior(0.01,0.01),
  blocksize=10, maxiter=10, maxvars=100, maxlogmargdrop=20,
  maxenum=10, verbose=TRUE)

```

**Arguments**

<code>y</code>	Vector with observed responses
<code>x</code>	Design matrix with all potential predictors
<code>center</code>	If TRUE, <code>y</code> and <code>x</code> are centered to have zero mean, therefore eliminating the need to include an intercept term in <code>x</code>
<code>scale</code>	If TRUE, <code>y</code> and columns in <code>x</code> are scaled to have standard deviation 1
<code>enumerate</code>	If TRUE all models with up to <code>maxvars</code> are enumerated, else Gibbs sampling is used to explore the model space
<code>includevars</code>	Logical vector of length <code>ncol(x)</code> indicating variables that should always be included in the model, i.e. variable selection is not performed for these variables
<code>maxvars</code>	When <code>enumerate==TRUE</code> only models with up to <code>maxvars</code> variables enumerated. In <code>modelsearchBlockDiag</code> a sequence of models is defined from 1 up to <code>maxvars</code>
<code>niter</code>	Number of Gibbs sampling iterations
<code>thinning</code>	MCMC thinning factor, i.e. only one out of each thinning iterations are reported. Defaults to <code>thinning=1</code> , i.e. no thinning
<code>burnin</code>	Number of burn-in MCMC iterations. Defaults to <code>.1*niter</code> . Set to 0 for no burn-in
<code>family</code>	Residual distribution. Possible values are 'normal', 'twopiecenormal', 'laplace', 'twopiecelaplace', or 'auto'. For the latter the residual distribution is inferred from the data. 'laplace' corresponds to median regression and 'twopiecelaplace' to quantile regression. See argument <code>priorSkew</code>
<code>priorCoef</code>	Prior on coefficients, created by <code>momprior</code> , <code>imomprior</code> , <code>emomprior</code> or <code>zellnerprior</code> . Prior dispersion is on coefficients/ $\sqrt{\text{scale}}$ for Normal and two-piece Normal, and on coefficients/ $\sqrt{2*\text{scale}}$ for Laplace and two-piece Laplace.
<code>priorDelta</code>	Prior on model space. Use <code>modelbbprior()</code> for Beta-Binomial prior, <code>modelbinomprior(p)</code> for Binomial prior with prior inclusion probability <code>p</code> , or <code>modelunifprior()</code> for Uniform prior
<code>priorVar</code>	Inverse gamma prior on scale parameter, created by <code>igprior()</code> . For Normal <code>variance=scale</code> , for Laplace <code>variance=2*scale</code> .
<code>priorSkew</code>	Either a fixed value for $\tanh(\alpha)$ where $\alpha$ is the asymmetry parameter or a prior on $\tanh(\alpha)$ . For <code>family=='twopiecelaplace'</code> setting <code>alpha=a</code> is equivalent to performing quantile regression for the quantile $(1+a)/2$ . Ignored if family is 'normal' or 'laplace'.
<code>phi</code>	Residual variance. Typically this is unknown and therefore left missing. If specified argument <code>priorVar</code> is ignored
<code>deltaini</code>	Logical vector of length <code>ncol(x)</code> indicating which coefficients should be initialized to be non-zero. Defaults to all variables being excluded from the model
<code>initSearch</code>	Algorithm to refine <code>deltaini</code> . <code>initSearch=='greedy'</code> uses a greedy Gibbs sampling search. <code>initSearch=='SCAD'</code> sets <code>deltaini</code> to the non-zero elements in a SCAD fit with cross-validated regularization parameter. <code>initSearch=='none'</code> leaves <code>deltaini</code> unmodified

method	Method to compute marginal likelihood. The default is to use closed-form expressions whenever possible (only available for pMOM priors with up to 15 covariates) and Laplace approximations otherwise. <code>method=='Laplace'</code> for Laplace approx, <code>method=='plugin'</code> for BIC-type approximation, <code>method=='MC'</code> for Importance Sampling, <code>method=='Hybrid'</code> for Hybrid Laplace-IS (only available for piMOM prior). See Details.
hess	Method to estimate the hessian in the Laplace approximation to the integrated likelihood under Laplace or asymmetric Laplace errors. When <code>hess=='asymp'</code> the asymptotic hessian is used, <code>hess=='asympDiagAdj'</code> a diagonal adjustment is applied (see Rossell and Rubio for details).
optimMethod	Algorithm to maximize objective function when <code>method=='Laplace'</code> or <code>method=='MC'</code> . <code>optimMethod=='LMA'</code> uses modified Newton-Raphson algorithm, 'CDA' coordinate descent algorithm
B	Number of samples to use in Importance Sampling scheme. Ignored if <code>method=='Laplace'</code>
verbose	Set <code>verbose==TRUE</code> to print iteration progress
blocksize	Maximum number of variables in a block. Careful, the cost of the algorithm is of order $2^{\text{blocksize}}$
maxiter	Maximum number of iterations, each iteration includes a screening pass to add and subtract variables
maxlogmargdrop	Stop the sequence of models when the drop in $\log p(y \text{model})$ is greater than <code>maxlogmargdrop</code> . This option avoids spending unnecessary time exploring overly large models
maxenum	If the posterior mode found has less than <code>maxenum</code> variables then do a full enumeration of all its submodels

## Details

Let  $\delta$  be the vector indicating inclusion/exclusion of each column of  $x$  in the model. The Gibbs algorithm sequentially samples from the posterior of each element in  $\delta$  conditional on all the remaining elements in  $\delta$  and the data. To do this it is necessary to evaluate the marginal likelihood for any given model. These have closed-form expression for the MOM prior, but for models with  $>15$  variables these are expensive to compute and Laplace approximations are used instead (for the residual variance a log change of variables is used, which improves the approximation). For other priors closed forms are not available, so by default Laplace approximations are used. For the iMOM prior we also implement a Hybrid Laplace-IS which uses a Laplace approximation to evaluate the integral wrt  $\beta$  and integrates wrt  $\phi$  (residual variance) numerically.

It should be noted that Laplace approximations tend to under-estimate the marginal densities when the MLE for some parameter is very close to 0. That is, it tends to be conservative in the sense of excluding more variables from the model than an exact calculation would.

Finally, `method=='plugin'` provides a BIC-type approximation that is faster than exact or Laplace methods, at the expense of some accuracy. In non-sparse situations where models with many variables have large posterior probability `method=='plugin'` can be substantially faster.

For more details on the methods used to compute marginal densities see Johnson & Rossell (2012). `modelsearchBlockDiag` uses the block search method described in Papaspiliopoulos & Rossell. Briefly, spectral clustering is run on  $X'X$  to cluster variables into blocks of `blocksize` and subsequently the Coolblock algorithm is used to define a sequence of models of increasing size. The

exact integrated likelihood is evaluated for all models in this path, the best model chosen, and the scheme iteratively repeated to add and drop variables until convergence.

### Value

Object of class `msfit`, which extends a list with elements

<code>postSample</code>	matrix with posterior samples for the model indicator. <code>postSample[i, j]==1</code> indicates that variable <code>j</code> was included in the model in the MCMC iteration <code>i</code>
<code>postOther</code>	<code>postOther</code> returns posterior samples for parameters other than the model indicator, i.e. basically hyper-parameters. If hyper-parameters were fixed in the model specification, <code>postOther</code> will be empty.
<code>margpp</code>	Marginal posterior probability for inclusion of each covariate. This is computed by averaging marginal post prob for inclusion in each Gibbs iteration, which is much more accurate than simply taking <code>colMeans(postSample)</code>
.	.
<code>postMode</code>	Model with highest posterior probability amongst all those visited
<code>postModeProb</code>	Unnormalized posterior prob of posterior mode (log scale)
<code>postProb</code>	Unnormalized posterior prob of each visited model (log scale)
<code>coef</code>	Estimated coefficients (via posterior mode) for highest posterior probability model

### Author(s)

David Rossell

### References

Johnson V.E., Rossell D. Non-Local Prior Densities for Default Bayesian Hypothesis Tests. *Journal of the Royal Statistical Society B*, 2010, 72, 143-170.

Johnson V.E., Rossell D. Bayesian model selection in high-dimensional settings. *Journal of the American Statistical Association*, 2012, 107, 649-660.

Papaspiliopoulos O., Rossell, D. Scalable Bayesian variable selection and model averaging under block orthogonal design. 2016

Rossell D., Rubio F.J. Tractable Bayesian variable selection: beyond normality. 2016

### See Also

[msfit-class](#) for details on the output. [postProb](#) to obtain posterior model probabilities. [rnlp](#) to obtain posterior samples for the coefficients. [nlpMarginal](#) to compute marginal densities for a given model.

**Examples**

```

#Simulate data
x <- matrix(rnorm(100*3),nrow=100,ncol=3)
theta <- matrix(c(1,1,0),ncol=1)
y <- x %%% theta + rnorm(100)

#Specify prior parameters
priorCoef <- imomprior(tau=1)
priorDelta <- modelunifprior()
priorVar <- igprior(alpha=.01,lambda=.01)

#Alternative prior for model space: 0.5 prior prob for including any covariate
priorDelta <- modelbinomprior(p=0.5)

#Alternative prior for model space: Beta hyper-prior for prob of inclusion
priorDelta <- modelbbprior(alpha.p=1,beta.p=1)

#Model selection
fit1 <- modelSelection(y=y, x=x, center=FALSE, scale=FALSE, niter=10^2,
priorCoef=priorCoef, priorDelta=priorDelta, priorVar=priorVar, phi=1,
method='Laplace')
fit1$postMode
fit1$margpp
postProb(fit1)

```

---

mombf

*Moment and inverse moment Bayes factors for linear models.*


---

**Description**

mombf computes moment Bayes factors to test whether a subset of regression coefficients are equal to some user-specified value. imombf computes inverse moment Bayes factors. zellnerbf computes Bayes factors based on the Zellner-Siow prior (used to build the moment prior).

**Usage**

```

mombf(lm1, coef, g, prior.mode, baseDensity='normal', nu=3, theta0,
logbf=FALSE, B=10^5)
imombf(lm1, coef, g, prior.mode, nu = 1, theta0, method='adapt',
nquant=100, B = 10^5)

```

**Arguments**

lm1	Linear model fit, as returned by lm1.
coef	Vector with indexes of coefficients to be tested. e.g. coef==c(2, 3) and theta0==c(0, 0) tests coef(lm1)[2]=coef(lm1)[3]=0.
g	Vector with prior parameter values. See dmom and dimom for details.

prior.mode	If specified, g is determined by calling g2mode.
baseDensity	Density upon which the Mom prior is based. baseDensity=='normal' results in the normal Mom prior, baseDensity=='t' in the t Mom prior with nu degrees of freedom.
nu	For mombf, nu specifies the degrees of freedom of the t Mom prior. It is ignored unless baseDensity=='t'. nu defaults to 3. For imombf, nu specifies the degrees of freedom for the inverse moment prior (see dimom for details). Defaults to nu=1, which Cauchy-like tails.
theta0	Null value for the regression coefficients. Defaults to 0.
logbf	If logbf==TRUE the natural logarithm of the Bayes factor is returned.
method	Numerical integration method to compute the bivariate integral (only used by imombf). For method=='adapt', the inner integral is evaluated (via integrate) at a series of nquant quantiles of the residual variance posterior distribution, and then averaged as described in Johnson (1992). Set method=='MC' to use Monte Carlo integration.
nquant	Number of quantiles at which to evaluate the integral for known sigma. Only used if method=='adapt'.
B	Number of Monte Carlo samples to estimate the T Mom and the inverse moment Bayes factor. Only used in mombf if baseDensity=='t'. Only used in imombf if method=='MC'.

### Details

These functions actually call momunknown and imomunknown, but they have a simpler interface. See dmom and dimom for details on the moment and inverse moment priors. The Zellner-Siow g-prior is given by  $dmvnorm(\theta, \theta_0, n * g * V_1)$ .

### Value

mombf returns the moment Bayes factor to compare the model where  $\theta \neq \theta_0$  with the null model where  $\theta = \theta_0$ . Large values favor the alternative model; small values favor the null. imombf returns inverse moment Bayes factors. zellnerbf returns Bayes factors based on the Zellner-Siow g-prior.

### Author(s)

David Rossell

### References

See <http://rosselldavid.googlepages.com> for technical reports. For details on the quantile integration, see Johnson, V.E. A Technique for Estimating Marginal Posterior Densities in Hierarchical Models Using Mixtures of Conditional Densities. Journal of the American Statistical Association, Vol. 87, No. 419. (Sep., 1992), pp. 852-860.

**See Also**

[momunknown](#), [imomunknown](#) and [zbfunknown](#) for another interface to compute Bayes factors. [momknown](#), [imomknown](#) and [zbfknown](#) to compute Bayes factors assuming that the dispersion parameter is known, and for approximate Bayes factors for GLMs. [mode2g](#) for prior elicitation.

**Examples**

```
##compute Bayes factor for Hald's data
data(hald)
lm1 <- lm(hald[,1] ~ hald[,2] + hald[,3] + hald[,4] + hald[,5])

# Set g so that prior mode for standardized effect size is at 0.2
prior.mode <- .2^2
V <- summary(lm1)$cov.unscaled
gmom <- mode2g(prior.mode,prior='normalMom')
gimom <- mode2g(prior.mode,prior='iMom')

# Set g so that interval (-0.2,0.2) has 5% prior probability
# (in standardized effect size scale)
priorp <- .05; q <- .2
gmom <- c(gmom,priorp2g(priorp=priorp,q=q,prior='normalMom'))
gimom <- c(gmom,priorp2g(priorp=priorp,q=q,prior='iMom'))

mombf(lm1,coef=2,g=gmom) #moment BF
imombf(lm1,coef=2,g=gimom,B=10^5) #inverse moment BF
zellnerbf(lm1,coef=2,g=1) #BF based on Zellner's g-prior
```

---

momknown

*Bayes factors for moment, inverse moment and Zellner-Siow g-prior.*


---

**Description**

`momknown` and `momunknown` compute moment Bayes factors for linear models when  $\sigma^2$  is known and unknown, respectively. The functions can also be used to compute approximate Bayes factors for generalized linear models and other settings. `imomknown`, `imomunknown` compute inverse moment Bayes factors. `zbfknown`, `zbfunknown` compute Bayes factors based on the Zellner-Siow g-prior.

**Usage**

```
momknown(theta1hat, V1, n, g = 1, theta0, sigma, logbf = FALSE)
momunknown(theta1hat, V1, n, nuisance.theta, g = 1, theta0, ssr, logbf =
FALSE)
imomknown(theta1hat, V1, n, nuisance.theta, g = 1, nu = 1, theta0,
sigma, method='adapt', B=10^5)
imomunknown(theta1hat, V1, n, nuisance.theta, g = 1, nu = 1, theta0,
ssr, method='adapt', nquant = 100, B = 10^5)
```

```
zbfknown(theta1hat, V1, n, g = 1, theta0, sigma, logbf = FALSE)
zbfunknown(theta1hat, V1, n, nuisance.theta, g = 1, theta0, ssr, logbf =
FALSE)
```

### Arguments

theta1hat	Vector with regression coefficients estimates.
V1	Matrix proportional to the covariance of theta1hat. For linear models, the covariance is $\sigma^2 \cdot V1$ .
n	Sample size.
nuisance.theta	Number of nuisance regression coefficients, i.e. coefficients that we do not wish to test for.
ssr	Sum of squared residuals from a linear model call.
g	Prior parameter. See <code>dmom</code> and <code>dimom</code> for details.
theta0	Null value for the regression coefficients. Defaults to 0.
sigma	Dispersion parameter is $\sigma^2$ .
logbf	If <code>logbf==TRUE</code> the natural logarithm of the Bayes factor is returned.
nu	Prior parameter for the inverse moment prior. See <code>dimom</code> for details. Defaults to <code>nu=1</code> , which Cauchy-like tails.
method	Numerical integration method (only used by <code>imomknown</code> and <code>imomunknown</code> ). Set <code>method=='adapt'</code> in <code>imomknown</code> to integrate using adaptive quadrature of functions as implemented in the function <code>integrate</code> . In <code>imomunknown</code> the integral is evaluated as in <code>imomknown</code> at a series of <code>nquant</code> quantiles of the posterior for <code>sigma</code> , and then averaged as described in Johnson (1992). Set <code>method=='MC'</code> to use Monte Carlo integration.
nquant	Number of quantiles at which to evaluate the integral for known <code>sigma</code> .
B	Number of Monte Carlo samples to estimate the inverse moment Bayes factor. Ignored if <code>method!='MC'</code> .

### Details

See `dmom` and `dimom` for details on the moment and inverse moment priors. The Zellner-Siow g-prior is given by `dmvnorm(theta,theta0,n*g*V1)`.

### Value

`momknown` and `momunknown` return the moment Bayes factor to compare the model where `theta!=theta0` with the null model where `theta==theta0`. Large values favor the alternative model; small values favor the null. `imomknown` and `imomunknown` return inverse moment Bayes factors. `zbfknown` and `zbfunknown` return Bayes factors based on the Zellner-Siow g-prior.

### Author(s)

David Rossell



## References

See <http://rosselldavid.googlepages.com> for technical reports.

For details on the quantile integration, see Johnson, V.E. A Technique for Estimating Marginal Posterior Densities in Hierarchical Models Using Mixtures of Conditional Densities. *Journal of the American Statistical Association*, Vol. 87, No. 419. (Sep., 1992), pp. 852-860.

## See Also

[mombf](#) and [imombf](#) for a simpler interface to compute Bayes factors in linear regression. [mode2g](#) for prior elicitation.

## Examples

```
#simulate data from probit regression
set.seed(4*2*2008)
n <- 50; theta <- c(log(2),0)
x <- matrix(NA,nrow=n,ncol=2)
x[,1] <- rnorm(n,0,1); x[,2] <- rnorm(n,.5*x[,1],1)
p <- pnorm(x[,1]*theta[1]+x[,2]+theta[2])
y <- rbinom(n,1,p)

#fit model
glm1 <- glm(y~x[,1]+x[,2],family=binomial(link = "probit"))
thetahat <- coef(glm1)
V <- summary(glm1)$cov.scaled

#compute Bayes factors to test whether x[,1] can be dropped from the model
g <- .5
bfmom.1 <- momknown(thetahat[2],V[2,2],n=n,g=g,sigma=1)
bfimom.1 <- imomknown(thetahat[2],V[2,2],n=n,nuisance.theta=2,g=g,sigma=1)
bfmom.1
bfimom.1
```

---

msfit-class

*Class "msfit"*

---

## Description

Stores the output of Bayesian variable selection, as produced by function `modelSelection`. The class extends a list, so all usual methods for lists also work for `msfit` objects, e.g. accessing elements, retrieving names etc.

Some additional methods are provided for printing information on screen, computing posterior probabilities or sampling from the posterior of regression coefficients, as indicated below.

## Objects from the Class

Typically objects are automatically created by a call to `modelSelection`. Alternatively, objects can be created by calls of the form `new("msfit", x)` where `x` is a list with the adequate elements (see slots).

## Slots

The class extends a list with elements:

**postSample** matrix with posterior samples for the model indicator. `postSample[i, j]==1` indicates that variable `j` was included in the model in the MCMC iteration `i`

**postOther** `postOther` returns posterior samples for parameters other than the model indicator, i.e. basically hyper-parameters. If hyper-parameters were fixed in the model specification, `postOther` will be empty.

**margpp** Marginal posterior probability for inclusion of each covariate. This is computed by averaging marginal post prob for inclusion in each Gibbs iteration, which is much more accurate than simply taking `colMeans(postSample)`.

**postMode** Model with highest posterior probability amongst all those visited

**postModeProb** Unnormalized posterior prob of posterior mode (log scale)

**postProb** Unnormalized posterior prob of each visited model (log scale)

**coef** Estimated coefficients (via posterior mode) for highest posterior probability model

**family** Residual distribution, i.e. argument family when calling `modelSelection`

**p** Number of variables

## Methods

**show** `signature(object = "msfit")`: Displays general information about the object.

**postProb** `signature(object = "msfit")`: Extracts posterior model probabilities.

**rnlp** `signature(object = "msfit")`: Obtain posterior samples for regression coefficients.

## Author(s)

David Rossell

## References

Johnson VE, Rossell D. Non-Local Prior Densities for Default Bayesian Hypothesis Tests. *Journal of the Royal Statistical Society B*, 2010, 72, 143-170

Johnson VE, Rossell D. Bayesian model selection in high-dimensional settings. *Journal of the American Statistical Association*, 107, 498:649-660.

## See Also

See also [modelSelection](#) and [rnlp](#).

## Examples

```
showClass("msfit")
```

---

 msPriorSpec-class      *Class "msPriorSpec"*


---

### Description

Stores the prior distributions to be used for Bayesian variable selection in normal regression models. This class can be used to specify the prior on non-zero regression coefficients, the model indicator or the nuisance parameters.

### Usage

```

momprior(tau, tau.adj=10^6, r=1)
imomprior(tau, tau.adj=10^6)
emomprior(tau, tau.adj=10^6)
zellnerprior(tau, tau.adj=10^6)

modelunifprior()
modelbinomprior(p=0.5)
modelbbprior(alpha.p=1, beta.p=1)

igprior(alpha=.01, lambda=.01)

```

### Arguments

tau	Prior dispersion parameter for covariates undergoing selection
tau.adj	Prior variance in Normal prior for covariates not undergoing selection
r	MOM prior parameter is $2*r$
p	Prior inclusion probability for binomial prior on model space
alpha.p	Beta-binomial prior on model space has parameters alpha.p, beta.p
beta.p	Beta-binomial prior on model space has parameters alpha.p, beta.p
alpha	Inverse gamma prior has parameters alpha/2, lambda/2
lambda	Inverse gamma prior has parameters alpha/2, lambda/2

### Objects from the Class

Objects can be created by calls of the form `new("msPriorSpec", ...)`, but it is easier to use creator functions.

For priors on regression coefficients use `momprior`, `imomprior` or `emomprior`. For prior on model space `modelunifprior`, `modelbinomprior` or `modelbbprior`. For prior on residual variance use `igprior`.

**Slots**

**priorType:** Object of class "character". "coefficients" indicates that the prior is for the non-zero regression coefficients. "modelIndicator" that it is for the model indicator, and "nuisancePars" that it is for the nuisance parameters. Several prior distributions are available for each choice of priorType, and these can be specified in the slot priorDistr.

**priorDistr:** Object of class "character". If priorType=="coefficients", priorDistr can be equal to "pMOM", "piMOM", "peMOM" or "zellner" (product moment, product inverse moment, product exponential moment or Zellner prior, respectively). If priorType=="modelIndicator", priorDistr can be equal to "uniform" or "binomial" to specify a uniform prior (all models equally likely a priori) or a binomial prior. For a binomial prior, the prior inclusion probability for any single variable must be specified in slot priorPars['p']. For a beta-binomial prior, the Beta hyper-prior parameters must be in priorPars['alpha.p'] and priorPars['beta.p']. If priorType=="nuisancePars", priorDistr must be equal to "invgamma". This corresponds to an inverse gamma distribution for the residual variance, with parameters specified in the slot priorPars.

**priorPars:** Object of class "vector", where each element must be named. For priorDistr=='pMOM', there must be an element "r" (MOM power is 2r). For any priorDistr there must be either an element "tau" indicating the prior dispersion or elements "a.tau" and "b.tau" specifying an inverse gamma hyper-prior for "tau". Optionally, there may be an element "tau.adj" indicating the prior dispersion for the adjustment variables (i.e. not undergoing variable selection). If not defined, "tau.adj" is set to 0.001 by default. For priorDistr=='binomial', there must be either an element "p" specifying the prior inclusion probability for any single covariate, or a vector with elements "alpha.p" and "beta.p" specifying a Beta(alpha.p,beta.p) hyper-prior on p. For priorDistr=='invgamma' there must be elements "alpha" and "lambda". The prior for the residual variance is an inverse gamma with parameters  $.5 \times \alpha$  and  $.5 \times \lambda$ .

**Methods**

No methods defined with class "msPriorSpec" in the signature.

**Note**

When new instances of the class are created a series of check are performed to ensure that a valid prior specification is produced.

**Author(s)**

David Rossell

**References**

Johnson VE, Rossell D. Non-Local Prior Densities for Default Bayesian Hypothesis Tests. *Journal of the Royal Statistical Society B*, 2010, 72, 143-170

Johnson VE, Rossell D. Bayesian model selection in high-dimensional settings. *Journal of the American Statistical Association*, 107, 498:649-660.

**See Also**

See also [modelSelection](#) for an example of defining an instance of the class and perform Bayesian model selection.

**Examples**

```
showClass("msPriorSpec")
```

---

nlpmarginals	<i>Marginal density of the observed data for linear regression with Normal, two-piece Normal, Laplace or two-piece Laplace residuals under non-local and Zellner priors</i>
--------------	---

---

**Description**

The marginal density of the data, i.e. the likelihood integrated with respect to the given prior distribution on the regression coefficients of the variables included in the model and an inverse gamma prior on the residual variance.

nlpMarginal is the general function, the remaining ones correspond to particular cases and are kept for backwards compatibility with old code, and will be deprecated in the future.

**Usage**

```
nlpMarginal(sel, y, x, family="normal", priorCoef=momprior(tau=0.348),
  priorVar=igprior(alpha=0.01,lambda=0.01), priorSkew=momprior(tau=0.348),
  method='auto', hess='asymp', optimMethod='CDA', B=10^5, logscale=TRUE, XtX, ytX)
```

```
pimomMarginalK(sel, y, x, phi, tau=1, method='Laplace', B=10^5, logscale=TRUE, XtX, ytX)
```

```
pimomMarginalU(sel, y, x, alpha=0.001, lambda=0.001, tau=1,
  method='Laplace', B=10^5, logscale=TRUE, XtX, ytX)
```

```
pmomMarginalK(sel, y, x, phi, tau, r=1, method='auto', B=10^5,
  logscale=TRUE, XtX, ytX)
```

```
pmomMarginalU(sel, y, x, alpha=0.001, lambda=0.001, tau=1,
  r=1, method='auto', B=10^5, logscale=TRUE, XtX, ytX)
```

**Arguments**

sel	Vector with indexes of columns in x to be included in the model
y	Vector with observed responses
x	Design matrix with covariates. Only the columns specified in sel are included in the model, the rest are disregarded
family	Residual distribution. Possible values are 'normal','twopiecenormal','laplace', 'twopiecelaplace'
priorCoef	Prior on coefficients, created by momprior, imomprior, emomprior or zellnerprior. Prior dispersion is on coefficients/sqrt(scale) for Normal and two-piece Normal, and on coefficients/sqrt(2*scale) for Laplace and two-piece Laplace.

priorVar	Inverse gamma prior on scale parameter, created by <code>igprior()</code> . For Normal <code>variance=scale</code> , for La
priorSkew	Either a number fixing $\tanh(\alpha)$ where $\alpha$ is the asymmetry parameter or a prior on residual skewness parameter, assumed to be of the same family as <code>priorCoef</code> . Ignored if family is 'normal' or 'laplace'.
method	Method to approximate the integral. <code>method=='auto'</code> uses closed-form expressions whenever possible and Laplace approximations otherwise. <code>method=='Laplace'</code> for Laplace approx. <code>method=='MC'</code> for Monte Carlo
hess	Method to estimat the hessian in the Laplace approximation to the integrated likelihood under Laplace or asymmetric Laplace errors. When <code>hess=='asymp'</code> the asymptotic hessian is used, <code>hess=='asympDiagAdj'</code> a diagonal adjustment is applied (see Rossell and Rubio for details).
optimMethod	Algorithm to maximize objective function when <code>method=='Laplace'</code> or <code>method=='MC'</code> . <code>optimMethod=='LMA'</code> uses modified Newton-Raphson algorithm, 'CDA' coordinate descent algorithm
B	Number of Monte Carlo samples to use (ignored unless <code>method=='MC'</code> )
logscale	If <code>logscale==TRUE</code> the log marginal density is returned.
XtX	Optionally, specify the matrix $X'X$ . Useful when the function must be called a large number of times.
ytX	Optionally, specify the vector $y'X$ . Useful when the function must be called a large number of times.
phi	Residual variance, assumed to be known by <code>pimomMarginalK</code> and <code>pmomMarginalK</code>
alpha	Prior for phi is inverse gamma $\alpha/2, \lambda/2$
lambda	Prior for phi is inverse gamma $\alpha/2, \lambda/2$
tau	Prior dispersion parameter for MOM and iMOM priors (see details)
r	Prior power parameter for MOM prior is $2*r$

### Details

The marginal density of the data is equal to the integral of  $N(y;x[,sel]*\theta,\phi*I) * \pi(\theta|\phi,\tau) * IG(\phi;\alpha/2,\lambda/2)$  with respect to  $\theta$ , where  $\pi(\theta|\phi,\tau)$  is a non-local prior and  $IG$  denotes the density of an inverse gamma.

`pmomMarginalK` and `pimomMarginalK` assume that the residual variance is known and therefore the inverse-gamma term in the integrand can be ommitted.

The product MOM and iMOM densities can be evaluated using the functions `dmom` and `dimom`.

### Value

Marginal density of the observed data under the specified prior.

### Author(s)

David Rossell

## References

Johnson V.E., Rossell D. Non-Local Prior Densities for Default Bayesian Hypothesis Tests. Journal of the Royal Statistical Society B, 2010, 72, 143-170. See <http://rosselldavid.googlepages.com> for technical reports.

## See Also

[modelSelection](#) to perform model selection based on product non-local priors. [momunknown](#), [imomunknown](#), [momknown](#), [imomknown](#) to compute Bayes factors for additive MOM and iMOM priors. [mode2g](#) for prior elicitation.

## Examples

```
x <- matrix(rnorm(100*2),ncol=2)
y <- x %*% matrix(c(.5,1),ncol=1) + rnorm(nrow(x))
pmomMarginalK(sel=1, y=y, x=x, phi=1, tau=1, method='Laplace')
pmomMarginalK(sel=1:2, y=y, x=x, phi=1, tau=1, method='Laplace')
```

---

pmomLM

*Bayesian variable selection and model averaging for linear and probit models via non-local priors.*

---

## Description

Variable selection for linear and probit models, providing a sample from the joint posterior of the model and regression coefficients. pmomLM and pmomPM implement product Normal MOM and heavy-tailed product MOM as prior distribution for linear and probit model coefficients (respectively). emomLM and emomPM set an eMOM prior.

pp1PM finds the value of the prior dispersion parameter tau minimizing posterior expected predictive loss (Gelfand and Ghosh, 1998) for the Probit model, i.e. can be used to automatically set up tau.

ppmodel returns the proportion of visits to each model.

## Usage

```
pmomLM(y, x, xadj, center=FALSE, scale=FALSE, niter=10^4, thinning=1,
burnin=round(niter/10), priorCoef, priorDelta, priorVar, initSearch='greedy',
verbose=TRUE)
```

```
pmomPM(y, x, xadj, niter=10^4, thinning=1, burnin=round(niter/10),
priorCoef, priorDelta, initSearch='greedy', verbose=TRUE)
```

```
emomLM(y, x, xadj, center=FALSE, scale=FALSE, niter=10^4, thinning=1,
burnin=round(niter/10), priorCoef, priorDelta, priorVar, initSearch='greedy',
verbose=TRUE)
```

```
emomPM(y, x, xadj, niter=10^4, thinning=1, burnin =round(niter/10),
priorCoef, priorDelta, initSearch='greedy', verbose=TRUE)
```

```
pplPM(tauseq=exp(seq(log(.01),log(2),length=20)), kPen=1, y, x, xadj, niter=10^4,
thinning=1, burnin=round(niter/10), priorCoef, priorDelta, priorVar,
initSearch='greedy', mc.cores=1)
```

```
ppmodel(nlpfit)
```

### Arguments

y	Vector with observed responses. For pmomLM this must be a numeric vector. For pmomPM it can either be a logical vector, a factor with 2 levels or a numeric vector taking only two distinct values.
x	Design matrix with all potential predictors which are to undergo variable selection.
xadj	Design matrix for adjustment covariates, i.e. variables which are included in the model with probability 1. For instance, xadj can be used to force the inclusion of an intercept in the model.
center	If center==TRUE, y and x are centered to have zero mean, therefore eliminating the need to include an intercept term in x
scale	If scale==TRUE, y and columns in x are scaled to have standard deviation 1
niter	Number of MCMC sampling iterations
thinning	MCMC thinning factor, i.e. only one out of each thinning iterations are reported. Defaults to thinning=1, i.e. no thinning
burnin	Number of burn-in MCMC iterations. Defaults to .1*niter. Set to 0 for no burn-in
priorCoef	Prior distribution for the coefficients. Must be object of class msPriorSpec with slot priorType set to 'coefficients'. Possible values for slot priorDistr are 'pMOM', 'piMOM' and 'peMOM'
priorDelta	Prior on model indicator space. Must be object of class msPriorSpec with slot priorType set to 'modelIndicator'. Possible values for slot priorDistr are 'uniform' and 'binomial'. For 'binomial', you can either set the prior probability 'p' or specify a Beta-binomial prior by specifying the parameters 'alpha.p','beta.p'.
priorVar	Prior on residual variance. Must be object of class msPriorSpec with slot priorType set to 'nuisancePars'. Slot priorDistr must be equal to 'invgamma'
initSearch	Algorithm to refine deltaini. initSearch=='greedy' uses a greedy Gibbs sampling search. initSearch=='SCAD' sets deltaini to the non-zero elements in a SCAD fit with cross-validated regularization parameter. initSearch=='none' initializes to the null model with no variables in x included.
verbose	Set verbose==TRUE to print iteration progress
tauseq	Grid of tau values for which the posterior predictive loss should be evaluated.
kPen	Penalty term specifying the relative importance of deviations from the observed data vs deviation from the posterior predictive. kPen can be set either to a numeric value or to 'msize' to set penalty equal to the average model size. Loss is $Dev(y_p, \hat{y}) + kPen * Dev(\hat{y}, y_{obs})$ , where $y_p$ : draw from post predictive, $y_{obs}$ : observed data and $\hat{y}$ is $E(y y_{obs})$ .



<code>mc.cores</code>	Allows for parallel computing. <code>mc.cores</code> is the number of processors to use. Setting <code>mc.cores&gt;1</code> requires the <code>parallel</code> package.
<code>nlpfit</code>	Non-local prior model fit, as returned by <code>pmomLM</code> , <code>pmomPM</code> , <code>emomLM</code> or <code>emomPM</code> .

### Details

The implemented MCMC scheme makes proposals from the joint posterior of  $(\delta[i], \theta[i])$  given all other parameters and the data, where  $\delta[i]$  is the indicator for inclusion/exclusion of covariate  $i$  and  $\theta[i]$  is the coefficient value. In contrast with some model fitting options implemented in `modelSelection`, here the scheme is exact. However, sampling the coefficients can adversely affect the mixing when covariates are very highly correlated. In practice, the mixing seems to be reasonably good for correlations up to 0.9.

`pmomPM` uses the scheme of Albert & Chib (1993) for probit models.

### Value

`pmomLM` and `pmomPM` returns a list with elements

<code>postModel</code>	matrix with posterior samples for the model indicator. <code>postModel[i, j]==1</code> indicates that variable $j$ was included in the model in the MCMC iteration $i$
<code>postCoef1</code>	matrix with posterior samples for coefficients associated to $x$
<code>postCoef2</code>	matrix with posterior samples for coefficients associated to $x_{adj}$
<code>postPhi</code>	vector with posterior samples for residual variance
<code>postOther</code>	<code>postOther</code> returns posterior samples for other parameters, i.e. basically hyper-parameters. Currently the prior precision parameter $\tau$
<code>margpp</code>	Marginal posterior probability for inclusion of each covariate. This is computed by averaging marginal post prob for inclusion in each MCMC iteration, which is much more accurate than simply taking <code>colMeans(postModel)</code>

.

`pp1PM` returns a list with elements

<code>optfit</code>	Probit model fit using <code>tauopt</code> . It is the result of a call to <code>pmomPM</code> .
<code>PPL</code>	<code>data.frame</code> indicating for each value in <code>tauseq</code> the posterior predictive loss ( $PPL=G+P$ ), the goodness-of-fit ( $G$ ) and penalty terms ( $P$ )

, the average number of covariates in the model (`msize`) including  $x_{adj}$  and the smoothed `sPPL` obtained via a `gam` fit.

<code>tautopt</code>	Value of $\tau$ minimizing the PPL
----------------------	------------------------------------

### Author(s)

David Rossell, Donatello Telesca

## References

Johnson V.E., Rossell D. Non-Local Prior Densities for Default Bayesian Hypothesis Tests. *Journal of the Royal Statistical Society B*, 2010, 72, 143-170.

Johnson V.E., Rossell D. Bayesian model selection in high-dimensional settings. Technical report. 2011 See <http://rosselldavid.googlepages.com> for technical reports.

Albert, J. and Chib, S. (1993) Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association*, 88, p669-679

Gelfand, A. and Ghosh, S. (1998) Model choice: A minimum posterior predictive loss approach. *Biometrika*, 85, p1-11.

## See Also

For more details on the prior specification see [msPriorSpec-class](#) To compute marginal densities for a given model see [pmomMarginalK](#), [pmomMarginalU](#), [pimomMarginalK](#), [pimomMarginalU](#).

## Examples

```
#Simulate data
x <- matrix(rnorm(100*3),nrow=100,ncol=3)
xadj <- rep(1,nrow(x))
theta <- matrix(c(1,1,0),ncol=1)
y <- 10*xadj + x %*% theta + rnorm(100)

#Beta-binomial prior on model space
priorDelta <- modelbbprior(alpha.p=1,beta.p=1)

#Non-informative prior on residual variance
priorVar <- igprior(alpha=.01,lambda=.01)

#Product MOM prior with tau=0.3 on x coefficients
#Non-informative prior on xadj coefficients
priorCoef <- momprior(tau=0.3, tau.adj=10^6)

mom0 <- pmomLM(y=y,x=x,xadj=xadj,center=FALSE,scale=FALSE,niter=1000,
priorCoef=priorCoef,priorDelta=priorDelta,priorVar=priorVar)
round(colMeans(mom0$postModel),2)
round(colMeans(mom0$postCoef1),2)
round(colMeans(mom0$postCoef2),2)

#Alternative prior: hyper-prior on tau
priorCoef <- new("msPriorSpec",priorType='coefficients',priorDistr='pMOM',
priorPars=c(a.tau=1,b.tau=.135,tau.adj=10^6,r=1)) #hyper-prior
mom1 <-
pmomLM(y=y,x=x,xadj=xadj,center=FALSE,scale=FALSE,niter=1000,
priorCoef=priorCoef,priorDelta=priorDelta,priorVar=priorVar)
mean(mom1$postOther) #posterior mean for tau

#Probit model
n <- 500; rho <- .25; niter <- 1000
theta <- c(.4,.6,0); theta.adj <- 0
```

```
V <- diag(length(theta)); V[upper.tri(V)] <- V[lower.tri(V)] <- rho
x <- rmvnorm(n,rep(0,length(theta)),V); xadj <- matrix(1,nrow=nrow(x),ncol=1)
lpred <- as.vector(x %>% matrix(theta,ncol=1) + xadj %>% matrix(theta.adj,ncol=1))
p <- pnorm(lpred)
y <- runif(n)<p

mom2 <- pmomPM(y=y,x=x,xadj=xadj,niter=1000,priorCoef=priorCoef,
priorDelta=priorDelta,initSearch='greedy')
colMeans(mom2$postCoef1)
coef(glm(y ~ x + xadj -1, family=binomial(link='probit')))
```

---

postModeOrtho	<i>Bayesian model selection and averaging under block-diagonal <math>X'X</math> for linear models.</i>
---------------	--

---

### Description

postModeOrtho is for diagonal  $X'X$ , postModeBlockDiag for the more general block-diagonal  $X'X$ , where  $X$  is the matrix with predictors.

Both functions return the model of highest posterior probability of any given size using an efficient search algorithm. This sequence of models includes the highest posterior probability model (HPM). Posterior model probabilities, marginal variable inclusion probabilities and Bayesian model averaging estimates are also provided. The unknown residual variance is integrated out using an exact deterministic algorithm of low computational cost (see details in reference).

### Usage

```
postModeOrtho(y, x, priorCoef=momprior(tau=0.348), priorDelta=modelbbprior(1,1),
priorVar=igprior(0.01,0.01), bma=FALSE, maxvars=100)
```

```
postModeBlockDiag(y, x, blocks, priorCoef=zellnerprior(tau=nrow(x)),
priorDelta=modelbinomprior(p=1/ncol(x)),priorVar=igprior(0.01,0.01), bma=FALSE,
maxvars=100, momcoef)
```

### Arguments

y	Outcome
x	Matrix with predictors. If an intercept is desired x should include a column of 1's.
blocks	Factor or integer vector of length ncol(x) indicating the block that each column in x belongs to.
priorCoef	Prior distribution for the coefficients. Object created with momprior, imomprior, emomprior or zellnerprior.
priorDelta	Prior on model space. Use modelbbprior() for Beta-Binomial prior, modelbinomprior(p) for Binomial prior with prior inclusion probability p, or modelunifprior() for Uniform prior

<code>priorVar</code>	Inverse gamma prior on residual variance, created with <code>igprior()</code>
<code>bma</code>	Set to TRUE to obtain marginal inclusion probabilities and Bayesian model averaging parameter estimates for each column of <code>x</code> .
<code>maxvars</code>	The search for the HPM is restricted to models with up to <code>maxvars</code> variables (note: posterior model probabilities and BMA are valid regardless of <code>maxvars</code> )
<code>momcoef</code>	optional argument containing pre-computed coefficients needed to obtain the marginal likelihood under the pMOM prior. A first call to <code>postModeBlockDiag</code> returns these coefficients, thus this argument is useful to speed up successive calls.

### Details

The first step is to list a sequence of models with  $0, \dots, \text{maxvars}$  variables which, under fairly general conditions listed in Papaspiliopoulos & Rossell (2016), is guaranteed to include the HPM. Then posterior model probabilities are computed for all these models to determine the HPM, evaluate the marginal posterior of the residual variance on a grid, and subsequently compute the marginal density  $p(y)$  via adaptive quadrature. Finally this adaptive grid is used to compute marginal inclusion probabilities and Bayesian model averaging estimates. For more details see Papaspiliopoulos & Rossell (2016).

### Value

List with elements

<code>models</code>	data.frame indicating the variables included in the sequence of models found during the search of the HPM, and their posterior probabilities. The model with highest posterior probability in this list is guaranteed to be the HPM.
<code>phi</code>	data.frame containing an adaptive grid of <code>phi</code> (residual variance) values and their marginal posterior density $p(\text{phily})$ .
<code>logpy</code>	log-marginal density $p(y)$ , i.e. normalization constant of $p(\text{phily})$ .
<code>bma</code>	Marginal posterior inclusion probabilities and Bayesian model averaging estimates for each column in <code>x</code> .
<code>postmean.model</code>	Coefficient estimates conditional on each of the models in <code>models</code>
<code>momcoef</code>	If a MOM prior was specified in <code>priorCoef</code> , <code>momcoef</code> stores some coefficients needed to compute its marginal likelihood

### Author(s)

David Rossell

### References

Papaspiliopoulos O., Rossell D. Scalable Bayesian variable selection and model averaging under block-orthogonal design. 2016

**Examples**

```

#Simulate data
set.seed(1)
p <- 500; n <- 510
x <- scale(matrix(rnorm(n*p),nrow=n,ncol=p),center=TRUE,scale=TRUE)
S <- cov(x)
e <- eigen(cov(x))
x <- t(t(x %*% e$vectors)/sqrt(e$values))
th <- c(rep(0,p-3),c(.5,.75,1)); phi <- 1
y <- x %*% matrix(th,ncol=1) + rnorm(n,sd=sqrt(phi))

#Fit
priorCoef=zellnerprior(tau=n); priorDelta=modelbinomprior(p=1/p); priorVar=igprior(0.01,0.01)
pm.zell <- postModeOrtho(y,x=x,priorCoef=priorCoef,priorDelta=priorDelta,priorVar=priorVar,
bma=TRUE)

#Best models
head(pm.zell$models)

#Posterior probabilities for sequence of models
nvars <- sapply(strsplit(as.character(pm.zell$models$modelid),split=','),length)
plot(nvars,pm.zell$models$pp,ylab='post prob',xlab='number of vars',ylim=0:1,xlim=c(0,50))

#Marginal posterior of phi
plot(pm.zell$phi,type='l',xlab='phi',ylab='p(phi|y)')

#Marginal inclusion prob & BMA estimates
plot(pm.zell$bma$margpp,ylab='Marginal inclusion prob')
plot(pm.zell$bma$coef,ylab='BMA estimate')

```

---

postProb

*Obtain posterior model probabilities*


---

**Description**

Obtain posterior model probabilities after running Bayesian model selection

**Usage**

```
postProb(object, nmax, method='norm')
```

**Arguments**

object	Object of class msfit, e.g. as returned by modelSelection
nmax	Maximum number of models to report (defaults to no max)
method	For 'norm' probabilities are obtained by renormalizing the stored integrated likelihoods, for 'exact' they are given by the proportion of MCMC visits to each model. 'norm' has less variability but can be biased if the chain has not converged.

**Value**

A `data.frame` with posterior model probabilities in column `pp`. Column `modelid` indicates the indexes of the selected covariates (empty for the null model with no covariates).

**Author(s)**

David Rossell

**See Also**

[modelSelection](#) to perform model selection

**Examples**

```
#See help(modelSelection)
```

---

rnlp

*Posterior sampling for Non-Local Priors*

---

**Description**

Gibbs sampler for linear and Cox proportional hazards model under product non-local priors and Zellner's prior. Both sampling conditional on a model and Bayesian model averaging are implemented (see Details).

If  $x$  and  $y$  not specified samples from non-local priors/posteriors with density proportional to  $d(\theta) N(\theta; m, V)$  are produced, where  $d(\theta)$  is the non-local penalty term.

**Usage**

```
rnlp(y, x, m, V, msfit, priorCoef, priorVar=igprior(alpha=0.01,lambda=0.01), niter=10^3,
burnin=round(niter/10), thinning=1, pp='norm')
```

**Arguments**

<code>y</code>	Vector with observed responses. When <code>class(y)=='Surv'</code> sampling is based on the Cox partial likelihood, else a linear model is assumed.
<code>x</code>	Design matrix with all potential predictors
<code>m</code>	Mean for the Normal kernel
<code>V</code>	Covariance for the Normal kernel
<code>msfit</code>	Object of class <code>msfit</code> , e.g. as returned by <code>modelSelection</code> . If left missing sampling under the full model $y \sim x$ is performed, otherwise posterior model samples in <code>msfit</code> are used.
<code>priorCoef</code>	Prior distribution for the coefficients. Must be object of class <code>msPriorSpec</code> with slot <code>priorType</code> set to <code>'coefficients'</code> . Possible values for slot <code>priorDistr</code> are <code>'pMOM'</code> , <code>'piMOM'</code> , <code>'peMOM'</code> and <code>'zellner'</code>

priorVar	Prior on residual variance. Must be object of class <code>msPriorSpec</code> with slot <code>priorType</code> set to 'nuisancePars'. Slot <code>priorDistr</code> must be equal to 'invgamma'
niter	Number of MCMC iterations
burnin	Number of burn-in MCMC iterations. Defaults to <code>.1*niter</code> . Set to 0 for no burn-in
thinning	MCMC thinning factor, i.e. only one out of each thinning iterations are reported. Defaults to no thinning
pp	When <code>msfit</code> is provided this is the method to compute posterior model probabilities, which determine the sampled models. Can be 'norm' or 'exact', see <code>postProb</code> for details.

### Details

The algorithm is implemented for product MOM (pMOM), product iMOM (piMOM) and product eMOM (peMOM) priors. The algorithm combines an orthogonalization that provides low serial correlation with a latent truncation representation that allows fast sampling.

When `y` and `x` are specified sampling is for the linear regression posterior. When argument `msfit` is left missing, posterior sampling is for the full model regressing `y` on all covariates in `x`. When `msfit` is specified each model is drawn with probability given by `postProb(msfit)`. In this case, a Bayesian Model Averaging estimate of the regression coefficients can be obtained by applying `colMeans` to the `rnlp` output matrix.

When `y` and `x` are left missing, sampling is from a density proportional to  $d(\theta) N(\theta; m, V)$ , where  $d(\theta)$  is the non-local penalty (e.g.  $d(\theta) = \prod(\theta^{2r})$  for the pMOM prior).

### Value

Matrix with posterior samples

### Author(s)

David Rossell

### References

D. Rossell and D. Telesca. Non-local priors for high-dimensional estimation, 2014. <http://arxiv.org/pdf/1402.5107v2.pdf>

### See Also

[modelSelection](#) to perform model selection and compute posterior model probabilities. For more details on prior specification see [msPriorSpec-class](#).

### Examples

```
#Generate data
set.seed(2)
n <- 10^3; tau <- 0.133; x <- rmvnorm(n, sigma=matrix(c(2,1,1,2), nrow=2))
thtrue <- c(0.5, 1); phittrue <- 1
y <- thtrue[1]*x[,1] + thtrue[2]*x[,2] + rnorm(n, sd=sqrt(phittrue))
```

```
#Specify prior parameters
priorCoef <- imomprior(tau=1)
priorVar <- igprior(alpha=.01,lambda=.01)

th <- rnlp(y=y, x=x, niter=100, priorCoef=priorCoef, priorVar=priorVar)
colMeans(th)
acf(th[,1])[1]
```



# Index

- \*Topic **classes**
  - msfit-class, 17
  - msPriorSpec-class, 19
- \*Topic **datasets**
  - hald, 7
- \*Topic **distribution**
  - bbPrior, 2
  - dalapl, 3
  - dmom, 4
  - mode2g, 7
  - postProb, 29
  - rnlp, 30
- \*Topic **distrib**
  - eprod, 6
- \*Topic **htest**
  - dmom, 4
  - mode2g, 7
  - modelSelection, 9
  - mombf, 13
  - momknown, 15
  - nlpmarginals, 21
  - pmomLM, 23
  - postModeOrtho, 27
- \*Topic **models**
  - eprod, 6
  - modelSelection, 9
  - mombf, 13
  - momknown, 15
  - nlpmarginals, 21
  - pmomLM, 23
  - postModeOrtho, 27
  - postProb, 29
  - rnlp, 30
- bbPrior, 2
- binomPrior (bbPrior), 2
- dalapl, 3
- demom (dmom), 4
- demom, data.frame-method (dmom), 4
- demom, matrix-method (dmom), 4
- demom, vector-method (dmom), 4
- demom-methods (dmom), 4
- dimom, 8
- dimom (dmom), 4
- dmom, 4, 8
- emomLM (pmomLM), 23
- emomPM (pmomLM), 23
- emomprior (msPriorSpec-class), 19
- eprod, 6
- g2mode, 5
- g2mode (mode2g), 7
- hald, 7
- igprior (msPriorSpec-class), 19
- imombf, 8, 17
- imombf (mombf), 13
- imomknown, 15, 23
- imomknown (momknown), 15
- imomprior (msPriorSpec-class), 19
- imomunknown, 15, 23
- imomunknown (momknown), 15
- mode2g, 5, 7, 15, 17, 23
- modelbbprior (msPriorSpec-class), 19
- modelbinomprior (msPriorSpec-class), 19
- modelsearchBlockDiag (modelSelection), 9
- modelSelection, 9, 18, 21, 23, 30, 31
- modelunifprior (msPriorSpec-class), 19
- mombf, 8, 13, 17
- momknown, 15, 15, 23
- momprior (msPriorSpec-class), 19
- momunknown, 15, 23
- momunknown (momknown), 15
- msfit (msfit-class), 17
- msfit-class, 17
- msPriorSpec (msPriorSpec-class), 19
- msPriorSpec-class, 19

- nlpMarginal, [12](#)
- nlpMarginal (nlpmarginals), [21](#)
- nlpmarginals, [21](#)
  
- palapl (dalapl), [3](#)
- pemom (dmom), [4](#)
- pimom (dmom), [4](#)
- pimomMarginalK, [26](#)
- pimomMarginalK (nlpmarginals), [21](#)
- pimomMarginalU, [26](#)
- pimomMarginalU (nlpmarginals), [21](#)
- pmom (dmom), [4](#)
- pmomLM, [23](#)
- pmomMarginalK, [26](#)
- pmomMarginalK (nlpmarginals), [21](#)
- pmomMarginalU, [26](#)
- pmomMarginalU (nlpmarginals), [21](#)
- pmomPM (pmomLM), [23](#)
- postModeBlockDiag (postModeOrtho), [27](#)
- postModeOrtho, [27](#)
- postProb, [12](#), [29](#)
- postProb, msfit-method (postProb), [29](#)
- postProb-methods (postProb), [29](#)
- pp1PM (pmomLM), [23](#)
- ppmodel (pmomLM), [23](#)
- priorp2g (mode2g), [7](#)
  
- qimom (dmom), [4](#)
- qmom (dmom), [4](#)
  
- ralapl (dalapl), [3](#)
- rnlp, [12](#), [18](#), [30](#)
- rnlp, ANY, matrix, missing, missing, missing-method (rnlp), [30](#)
- rnlp, ANY, matrix, missing, missing, msfit-method (rnlp), [30](#)
- rnlp, missing, missing, numeric, matrix, missing-method (rnlp), [30](#)
- rnlp-methods (rnlp), [30](#)
  
- show, msfit-method (msfit-class), [17](#)
  
- unifPrior (bbPrior), [2](#)
  
- x.hald (hald), [7](#)
- y.hald (hald), [7](#)
  
- zbfknown, [15](#)
- zbfknown (momknown), [15](#)
- zbfunknown, [15](#)
- zbfunknown (momknown), [15](#)
- zellnerbf (mombf), [13](#)
- zellnerprior (msPriorSpec-class), [19](#)