Package 'BVS'

February 19, 2015

Type Package

Title Bayesian Variant Selection: Bayesian Model Uncertainty Techniques for Genetic Association Studies

Author Melanie Quintana <maw27.wilson@gmail.com>

Maintainer Melanie Quintana <maw27.wilson@gmail.com>

Description The functions in this package focus on analyzing case-control association studies involving a group of genetic variants. In particular, we are interested in modeling the outcome variable as a function of a multivariate genetic profile using Bayesian model uncertainty and variable selection techniques. The package incorporates functions to analyze data sets involving common variants as well as extensions to model rare variants via the Bayesian Risk Index (BRI) as well as haplotypes. Finally, the package also allows the incorporation of external biological information to inform the marginal inclusion probabilities via the iBMU.

Version 4.12.1

License Unlimited

Depends MASS, msm, haplo.stats, R (>= 2.14.0)

Repository CRAN

Date/Publication 2012-08-09 17:07:48

NeedsCompilation no

R topics documented:

BVS-package *Bayesian Variant Selection: Bayesian Model Uncertainty Techniques for Genetic Association Studies*

Description

The functions in this package focus on analyzing case-control association studies involving a group of genetic variants. In particular, we are interested in modeling the outcome variable as a function of a multivariate genetic profile using Bayesian model uncertainty and variable selection techniques. The package incorporates functions to analyze data sets involving common variants as well as extensions to model rare variants via the Bayesian Risk Index (BRI of Quintana and Conti (2011)) as well as haplotypes. Finally, the package also allows the incorporation of external biological information to inform the marginal inclusion probabilities via the iBMU (Quintana and Conti (submitted)).

Details

Package:	BVS
Version:	4.12.0
Date:	2012-4-17
Depends:	MASS, msm, haplo.stats
License:	GPL-2

Author(s)

Melanie Quintana <maw27.wilson@gmail.com>

References

Quintana M, Conti D (2011). *Incorporating Model Uncertainty in Detecting Rare Variants: The Bayesian Risk Index*. Genetic Epidemiology 35:638-649.

Quintana M, Conti D (Submitted). *Integrative Variable Selection via Bayesian Model Uncertainty*.

enumerateBVS *Function to Enumerate all models for Bayesian Variant Selection Methods*

Description

This function enumerates and calculates summaries for all models in the model space. Not recommended for problems where p>20.

Usage

```
enumerateBVS(data,forced=NULL,cov=NULL,a1=0,rare=FALSE,mult.regions=FALSE,
             regions=NULL,hap=FALSE,inform=FALSE)
```

Arguments

data	a *(n x (p+1))* dimensional data frame where the first column corresponds to the response variable that is presented as a factor variable corresponding to an individuals disease status (0|1),and the final *p* columns are the SNPs of interest each coded as a numeric variable that corresponds to the number of copies of minor alleles (0|1|2)
forced	an optional *(n x c)* matrix of *c* confounding variables that one wishes to adjust the analysis for and that will be forced into every model.
inform	if inform=TRUE corresponds to the iBMU algorithm of Quintana and Conti (Submitted) that incorporates user specified external predictor-level covariates into the variant selection algorithm.
cov	an optional *(p x q)* dimensional matrix of *q* predictor-level covariates that need to be specified if inform=TRUE that the user wishes to incorporate into the estimation of the marginal inclusion probabilities using the iBMU algorithm
a1	a *q* dimensional vector of specified effects of each predictor-level covariate to be used when inform=TRUE.
rare	if rare=TRUE corresponds to the Bayesian Risk index (BRI) algorithm of Quintana and Conti (2011) that constructs a risk index based on the multiple rare variants within each model. The marginal likelihood of each model is then calculated based on the corresponding risk index.
mult.regions	when rare=TRUE if mult.regions=TRUE then we include multiple region specific risk indices in each model. If mult.regions=FALSE a single risk index is computed for all variants in the model.
regions	if mult.regions=TRUE regions is a *p* dimensional character or factor vector identifying the user defined region of each variant.
hap	if hap=TRUE we estimate a set of haplotypes from the multiple variants within each model and the marginal likelihood of each model is calculated based on the set of estimated haplotypes.

Value

This function outputs a list of the following values:

fitness	A vector of the fitness values (log(Model likelihood) - log(Model Prior)) of each enumerated model.
logPrM	A vector of the log Model Priors of each enumerated model.
which	A vector identifying the character representation of each model indicator vector.
coef	If rare=FALSE we report a matrix where each row corresponds to the estimated coefficients for all variables within each enumerated model. If rare=TRUE we report a vector where each entry corresponds to the estimated coefficient of the risk index (or multiple risk indices if mult.regions = TRUE) corresponding to each enumerated model.
alpha	If inform=FALSE that is simply a vector of 0's. If inform=TRUE we report a matrix where each row corresponds to the specified effects (alpha's) of each predictor-level covariate for each enumerated model.

Author(s)

Melanie Quintana <maw27.wilson@gmail.com>

References

Quintana M, Conti D (2011). *Incorporating Model Uncertainty in Detecting Rare Variants: The Bayesian Risk Index*. Genetic Epidemiology 35:638-649.

Quintana M, Conti D (Submitted). *Integrative Variable Selection via Bayesian Model Uncertainty*.

Examples

```
## Load the data for Rare variant example
data(RareData)

## Enumerate model space for a subset of 5 variants and save output to BVS.out
## for rare variant example.
RareBVS.out <- enumerateBVS(data=RareData[,1:6],rare=TRUE)
```

---

fitBVS — *Function to calculate fitness for each model for Bayesian Variant Selection Methods*

Description

This function takes one of the models and calculates the fitness/cost value of the model.

Usage

```
fitBVS(Z,data,forced=NULL,cov=NULL,a1=NULL,rare=FALSE,mult.regions=FALSE,
       regions=NULL,hap=FALSE,inform=FALSE,which=NULL,which.char=NULL)
```

Arguments

Z	a *p* dimensional vector specifying a model of interest. In particular if the jth value of the vector is *0* the jth variant is not included in the model and if the jth value of the vector is *1* the jth variant is included in the model.
data	a *(n x (p+1))* dimensional data frame where the first column corresponds to the response variable that is presented as a factor variable corresponding to an individuals disease status (0|1),and the final *p* columns are the SNPs of interest each coded as a numeric variable that corresponds to the number of copies of minor alleles (0|1|2)
forced	an optional *(n x c)* matrix of *c* confounding variables that one wishes to adjust the analysis for and that will be forced into every model.
inform	if inform=TRUE corresponds to the iBMU algorithm of Quintana and Conti (Submitted) that incorporates user specified external predictor-level covariates into the variant selection algorithm.
cov	an optional *(p x q)* dimensional matrix of *q* predictor-level covariates (need when inform=TRUE) that the user wishes to incorporate into the estimation of the marginal inclusion probabilities using the iBMU algorithm
a1	a *q* dimensional vector of specified (or sampled) effects of each predictor-level covariate to be used when inform=TRUE.
rare	if rare=TRUE corresponds to the Bayesian Risk index (BRI) algorithm of Quintana and Conti (2011) that constructs a risk index based on the multiple rare variants within each model. The marginal likelihood of each model is then calculated based on the corresponding risk index.
mult.regions	when rare=TRUE if mult.regions=TRUE then we include multiple region specific risk indices in each model. If mult.regions=FALSE a single risk index is computed for all variants in the model.
regions	if mult.regions=TRUE regions is a *p* dimensional character or factor vector identifying the user defined region of each variant.
hap	if hap=TRUE we estimate a set of haplotypes from the multiple variants within each model and the marginal likelihood of each model is calculated based on the set of estimated haplotypes.
which	optional current which matrix of sampled models from sampleBVS that is used to see if a model has already been sampled so that that fitness does not have to be recalculated.
which.char	optional vector that identifies that current models that have been sampled from sampleBVS that is also used to determine if a model has already been sampled.

Details

Uses the glm function to calculate the marginal likelihood and fitness function of the model of interest. If rare = TRUE the marginal likelihood is based on the risk index produced from the subset of variants within the model of interest and if hap = TRUE the marginal likelihood is based on the estimated haplotypes produced from the subset of variants within the model of interest.

Value

This function outputs a vector of the following values:

coef	If rare=FALSE we report a vector where each value corresponds to the estimated coefficients for all variables within the model of interest. If rare=TRUE we report a value corresponding to the estimated coefficient of the risk index (or risk indices if multi.regions=TRUE) corresponding to each model of interest.
fitness	The value of the fitness function (log(Model likelihood) - log(Model Prior)) of the model of interest.
logPrM	The value of the log prior on the model of interest.

Author(s)

Melanie Quintana <maw27.wilson@gmail.com>

References

Quintana M, Conti D (2011). *Incorporating Model Uncertainty in Detecting Rare Variants: The Bayesian Risk Index*. Genetic Epidemiology 35:638-649.

Quintana M, Conti D (Submitted). *Integrative Variable Selection via Bayesian Model Uncertainty*.

Examples

```
## Load the data for Rare variant example
data(RareData)
p = dim(RareData)[2] -1

## Fit the Null model
fit.null = fitBVS(rep(0,p),data=RareData,rare=TRUE)
```

---

hapBVS *Function to estimate and report a set of haplotypes given a subset of variants*

Description

This function takes a subset of variants and estimates a set of haplotypes. Only haplotypes with a frequency greater than min.Hap.freq are reported.

Usage

```
hapBVS(G,min.Hap.freq)
```

Arguments

G	an *(n x g)* matrix of a subset of *g* SNPs of interest that are each coded as a numeric variable that corresponds to the number of copies of minor alleles (0|1|2)
min.Hap.freq	the minimum haplotype frequency of which an estimated haplotype is reported

Value

This function outputs a matrix of estimated haplotypes.

Author(s)

Melanie Quintana <maw27.wilson@gmail.com>

`InformBVS.I.out` *Example Output From 100K iterations of sampleBVS with Informative Data*

Description

Output from 100K iterations of *sampleBVS* with the informative study-based data set *InformData*. This was ran with inform=TRUE and gene based predictor-level covariates so that the analysis follows iBMU framework described in Quintana and Conti (submitted) where we sample that the effects of the predictor-level covariates.

Usage

```
data(InformBVS.I.out)
```

References

Quintana M, Conti D (Submitted). *Integrative Variable Selection via Bayesian Model Uncertainty*.

`InformBVS.NI.out` *Example Output From 100K iterations of sampleBVS with Informative Data*

Description

Output from 100K iterations of *sampleBVS* with the informative study-based data set *InformData*. This was ran with inform=FALSE so that the analysis corresponds to the basic Bayesian model uncertainty framework where we assume that the effects of the predictor-level covariates are 0 (alpha=0).

Usage

```
data(InformBVS.NI.out)
```

`InformData` *PNAT Study-based Simulation: Informative Data.*

Description

PNAT study-based simulated data set of 122 variants as described in Quintana and Conti (submitted). The first column represents the disease status of the individual, the remaining columns the counts of minor alleles (0|1|2) for each variant. The simulation was created by using the genotype data from a systems-based candidate gene study of smoking cessation as part of the Pharmacogenetics of Nicotine Addiction and Treatment Consortium. In particular, data set was formed from genotypes of 122 variants within 789 individuals. The 122 variants are from 7 unique gene regions and thus are comprised of a great deal of correlation between the markers within each gene. In this simulation we assumed that the predictor-level covariate corresponding to the gene *CHRNB2* was informative with regards to which variants are associated with smoking cessation.

Usage

```
data(InformData)
```

Value

A list of the following items:

`data`	A data set with 122 variants from 789 individuals.
`cov`	A set of dummy variables indicating the gene of each variant. This set of dummy variables is used as the predictor-level covariates within an informative analysis (inform=TRUE).
`genes`	A vector indicating the gene of each variant in the data set.

References

Quintana M, Conti D (Submitted). *Integrative Variable Selection via Bayesian Model Uncertainty*.

`Informresults.I` *Example Summary From 100K iterations of sampleBVS with Informative Data*

Description

Summary from 100K iterations of *sampleBVS* with the informative study-based data set *InformData* using *summaryBVS*. This was ran with inform=TRUE and gene based predictor-level covariates so that the analysis follows iBMU framework described in Quintana and Conti (submitted) where we sample that the effects of the predictor-level covariates.

Usage

```
data(Informresults.I)
```

References

Quintana M, Conti D (Submitted). *Integrative Variable Selection via Bayesian Model Uncertainty*.

`Informresults.NI` *Example Summary From 100K iterations of sampleBVS with Informative Data*

Description

Summary from 100K iterations of *sampleBVS* with the informative study-based data set *InformData* using *summaryBVS*. This was ran with inform=FALSE so that the analysis corresponds to the basic Bayesian model uncertainty framework where we assume that the effects of the predictor-level covariates are 0 (alpha=0).

Usage

```
data(Informresults.NI)
```

`plotBVS` *Image Plots for top Variant and Region Inclusions*

Description

This function allows the user to create image plots of the top variants and top Regions (any user specified set of variants such as pathways or genes) included in the top models. Variants and Regions are ordered based on marginal BF and regional BF which are plotted on the right axis. The width of the inclusion blocks are proportional to the posterior model probability that the variant or region is included in.

Usage

```
plotBVS(results, num.models=100, num.snps=20, num.regions=20, plot.coef=FALSE,
        true.coef=NULL,main=NULL, regions=NULL, type="s",prop.cases=NULL,...)
```

Arguments

results	output list from summaryBVS.
num.models	the number of the top models to place on the x-axis.
num.snps	If type="s", the number of the top variants to place on the y-axis.
num.regions	If type="r", the number of the top regions to place on the y-axis.
plot.coef	Only to be used for rare variant analysis when rare=TRUE and mult.regions = FALSE. When plot.coef=TRUE, the log(OR) of the risk indices specified by each of the top models are plotted on the x axis
type	specifies if we want to plot the variant inclusion ("s") or region inclusion ("r")
true.coef	optional vector giving the true odds ratios of each of the variants (if results are from a simulation)
main	optional vector giving the title of the plot
regions	an optional vector of character strings giving the names of the regions for each of the variants in data set needed when plotting type is "r" or can be added to include the region names of each variant on the y axis when plotting type is "s".
prop.cases	an optional *(p x 2)* dimensional matrix giving the number of cases that have the variant in column 1 and the number of controls with the variant in column 2. If specified, these counts will be reported on the right axis under each variants marginal BF
...	General parameters for plotting functions

Author(s)

Melanie Quintana <maw27.wilson@gmail.com>

Examples

```
## RARE VARIANT BRI EXAMPLE
  ## Load the data for Rare variant example
  data(RareData)

  ## Load the results from running sampleBVS on rare variant data for 100K iterations
  data(RareBVS.out)

  ## Load summary results
  data(RareResults)

  ## Plot the variant inclusions in the top 100 models for the top 10 variants
  plotBVS(RareResults,num.models=100,num.snps=10)

  ##Include the estimated log(OR) of the risk indices for the top models
  plotBVS(RareResults,num.models=100,num.snps=10,plot.coef=TRUE)


## INFORMATIVE iBMU EXAMPLE
  ##Load the data for the informative example
  data(InformData)
```

```
## Load the results from running sampleBVS with inform=FALSE for 100K iteration
data(InformBVS.NI.out)

## Load summary results
data(Informresults.NI)

## Make SNP and Gene inclusion plots
plotBVS(Informresults.NI,num.models=50,num.snps=10,regions=InformData$genes)
plotBVS(Informresults.NI,num.models=50,num.regions=10,regions=InformData$genes,type="r")

## Load the results from running sampleBVS with inform=TRUE for 100K iterations
data(InformBVS.I.out)

## load summary results
data(Informresults.I)

## Make SNP and Gene inclusion plots
plotBVS(Informresults.I,num.models=50,num.snps=10,regions=InformData$genes)
plotBVS(Informresults.I,num.models=50,num.regions=10,regions=InformData$genes,type="r")
```

RareBVS.out *Example Output From 100K iterations of sampleBVS with Rare Data*

Description

Output from 100K iterations of *sampleBVS* with the Rare variant data set *RareData*. This was ran with rare=TRUE to correspond to the BRI analysis of Quintana and Conti (2011).

Usage

```
data(RareBVS.out)
```

References

Quintana M, Conti D (2011). *Incorporating Model Uncertainty in Detecting Rare Variants: The Bayesian Risk Index*. Genetic Epidemiology 35:638-649.

RareData *Simulated Example Rare Variant data set.*

Description

Simulated data set of 134 rare variants. The first column represents the disease status of the individual, the remaining columns the counts of minor alleles (0|1|2) for each variant.

Usage

```
data(RareData)
```

Format

A data frame with 1912 observations on the following 135 variables (case, rare variants 1:134).

RareResults *Example Summary From 100K iterations of sampleBVS with Rare Data*

Description

Summary from 100K iterations of *sampleBVS* with the Rare variant data set *RareData* using *summaryBVS*. This was ran with rare=TRUE to correspond to the BRI analysis of Quintana and Conti (2011) and with a burnin of 1000 iterations.

Usage

```
data(RareResults)
```

References

Quintana M, Conti D (2011). *Incorporating Model Uncertainty in Detecting Rare Variants: The Bayesian Risk Index*. Genetic Epidemiology 35:638-649.

sampleBVS *Sampling Algorithm for Bayesian Variant Selection Methods*

Description

This function performs a basic MH Sampling algorithm to sample models from the model space when enumeration is not possible. For informative marginal inclusion probabilities the algorithm also performs a basic MCMC algorithm to sample the effects of the predictor-level covariates (alpha).

Usage

```
sampleBVS(data,forced=NULL,inform=FALSE,cov=NULL,rare=FALSE,mult.regions=FALSE,
          regions=NULL,hap=FALSE,iter=10000,save.iter=0,outfile=NULL,
          status.file=NULL,old.results=NULL)
```

Arguments

data	an *(n x (p+1))* dimensional data frame where the first column corresponds to the response variable that is presented as a factor variable corresponding to an individuals disease status (0|1),and the final p columns are the SNPs of interest each coded as a numeric variable that corresponds to the number of copies of minor alleles (0|1|2)
forced	an optional *(n x c)* dimensional matrix of c confounding variables that one wishes to adjust the analysis for and that will be forced into every model.
inform	if inform=TRUE corresponds to the iBMU algorithm of Quintana and Conti (Submitted) that incorporates user specified external predictor-level covariates into the variant selection algorithm.
cov	an optional *(p x q)* dimensional matrix of q predictor-level covariates (needed when inform=TRUE) that the user wishes to incorporate into the estimation of the marginal inclusion probabilities using the iBMU algorithm
rare	if rare=TRUE corresponds to the Bayesian Risk index (BRI) algorithm of Quintana and Conti (2011) that constructs a risk index based on the multiple rare variants within each model. The marginal likelihood of each model is then calculated based on the corresponding risk index.
mult.regions	when rare=TRUE if mult.regions=TRUE then we include multiple region specific risk indices in each model. If mult.regions=FALSE a single risk index is computed for all variants in the model.
regions	if mult.regions=TRUE regions is a p dimensional character or factor vector identifying the user defined region of each variant.
hap	if hap=TRUE we estimate a set of haplotypes from the multiple variants within each model and the marginal likelihood of each model is calculated based on the set of estimated haplotypes.
iter	the number of iterations to run the algorithm.

save.iter	the number of iterations between each checkpoint. A checkpoint file is written every save.iter iterations.
outfile	character string giving the pathname of the checkpoint file to save the output of the algorithm to.
status.file	character string giving the pathname of the file to write the status of the algorithm.
old.results	old output from sampleBVS that has been run for a subset of the total number of iterations that the user wanted to run. if specified the sampling algorithm will start from the last sampled model in old.results. To be used if sampleBVS has been interrupted for some reason.

Details

The algorithm is run for a chosen number of iterations where we randomly add and remove variants from the current model based on a basic MH algorithm. If inform = TRUE we also incorporate a set of predictor-level covariates that are provided by the user and use a MCMC algorithm to sample the effects of the covariates on the marginal inclusion probabilities. Convergence of the algorithm can be determined by running two independent runs of the algorithm with different starting values and examining the marginal Bayes factors for each variant under each independent run.

Value

This function outputs a list of the following values to the file write.out if this file is specified for every save.iter number of iterations:

fitness	A vector of the fitness values (log(Model likelihood) - log(Model Prior)) of each model sampled at each iteration of the algorithm.
logPrM	A vector of the Model Priors of each model sampled at each iteration of the algorithm.
which	A vector identifying the character representation of each model sampled.
coef	If rare=FALSE we report a matrix where each row corresponds to the estimated coefficients for all variables within each model sampled at each iteration of the algorithm. If rare=TRUE we report a vector where each entry corresponds to the estimated coefficient of the risk index (or multiple risk indices if mult.regions=TRUE) corresponding to each enumerated model.
alpha	If inform=FALSE that is simply a vector of 0's. If inform=TRUE we report a matrix where each row corresponds to the estimated effects (alpha's) of each predictor-level covariate for each model sampled at each iteration of the algorithm.

Author(s)

Melanie Quintana <maw27.wilson@gmail.com>

References

Quintana M, Conti D (2011). *Incorporating Model Uncertainty in Detecting Rare Variants: The Bayesian Risk Index*. Genetic Epidemiology 35:638-649.

Quintana M, Conti D (Submitted). *Informing Variable Selection via Bayesian Model Uncertainty*.

Examples

```
## Rare Variant BRI example
  ## Load the data for Rare variant example
  data(RareData)

  ## Run algorithm for 100 iterations for rare variant example.
  ## NOTE: Results from a more realistic run with 100K
  ## iterations can be found in data(RareBVS.out).
  RareBVS.out <- sampleBVS(data=RareData,iter=100,rare=TRUE)

  ## Run algorithm for 100 iterations for multiple region rare
  ## variant example.
  p = dim(RareData)[2]-1
  regions = c(rep("Region1",(p/2)),rep("Region2",(p/2)))
  RareBVS.out <- sampleBVS(data=RareData,iter=100,rare=TRUE,mult.regions=TRUE,regions=regions)

## Informative iBMU Example
  ##Load the data for the informative example
  data(InformData)

  ## Run algorithm for 100 iterations for informative data example.
  ## This run is the basic Bayes model uncertainty algorithm with inform=FALSE
  ## NOTE: Results from a more realistic run with 100K
  ## iterations can be found in data(InformBVS.NI.out).
  InformBVS.NI.out = sampleBVS(InformData$data,inform=FALSE,iter=100)

  ## Run algorithm for 100 iterations for informative data example.
  ## This run corresponds to the iBMU algorithm with inform=TRUE
  ## and dichotomous predictor-level covariates indicating the gene of each variant.
  ## NOTE: Results from a more realistic run with 100K
  ## iterations can be found in data(InformBVS.I.out).
  InformBVS.I.out = sampleBVS(InformData$data,inform=TRUE,
                              cov=as.matrix(InformData$cov),iter=100)
```

---

summaryBVS *Calculates Posterior Summaries for BVS Methods*

Description

This function calculates the global and marginal Bayes Factors that give the strength of evidence of there being an association in the overall set of variants of interest, the individual genes of interest (if specified) and the individual variants of interest.

Usage

```
summaryBVS(BVS.out,data=data,forced=NULL,cov=NULL,burnin=1000,regions=NULL,
           rare=FALSE,mult.regions=FALSE,inform=FALSE)
```

Arguments

BVS.out	Output from sampleBVS or enumerateBVS
data	an *(n x (p+1))* dimensional data frame where the first column corresponds to the response variable that is presented as a factor variable corresponding to an individuals disease status (0|1),and the final p columns are the SNPs of interest each coded as a numeric variable that corresponds to the number of copies of minor alleles (0|1|2)
forced	an optional *(n x c)* matrix of c confounding variables that one wishes to adjust the analysis for and that will be forced into every model.
burnin	an integer indicating the length of the burnin.
regions	an optional p dimensional vector of character strings giving the names of the regions (example can be gene names or pathway names) for each of the variants in data set. If a region vector is given, the function will report regional BF.
inform	if inform=TRUE corresponds to iBMU algorithm of Quintana and Conti (Submitted) that incorporates user specified external predictor-level covariates into the variant selection algorithm.
cov	an optional *(p x q)* dimensional matrix of q predictor-level covariates (needed when inform=TRUE) that the user wishes to incorporate into the estimation of the marginal inclusion probabilities using the iBMU algorithm
rare	if rare=TRUE corresponds to the Bayesian Risk index (BRI) algorithm of Quintana and Conti (2011) that constructs a risk index based on the multiple rare variants within each model. The marginal likelihood of each model is then calculated based on the corresponding risk index.
mult.regions	when rare=TRUE if mult.regions=TRUE then we include multiple region specific risk indices in each model. If mult.regions=FALSE a single risk index is computed for all variants in the model.

Details

Global and marginal Bayes factors (BF) are computed based on calculating the posterior probabilities of each of the unique models that were visited in sampleBVS or all models that were enumerated in enumerateBVS. The global BF tests the hypothesis that there is an association in the overall set of variants. BF's are also calculated at the regional (if regions are specified) and the variant level. At the regional level, BF are computed for the overall evidence of at least one of the variants within the region of interest being associated. Posterior estimates for the coefficients are also reported. Finally, if inform=TRUE posterior estimates of the effects of the posterior-level covariates on the marginal inclusion probabilities are reported.

Value

This function outputs a list of the following values:

Global	Global Bayes Factor giving the strength of evidence that at least one variant within the analysis is associated with the outcome of interest
MargBF	Marginal variant specific Bayes Factors giving the strength of evidence that each one of the variants are associated with the outcome of interest
Marg.RBF	Regional level Bayes Factors giving the strength of evidence that at least one variant within the region is associated with the outcome of interest
PostAlpha	If inform=TRUE gives that posterior estimates of the effects of the posterior-level covariates on the marginal inclusion probabilities.
PostCoef	Posterior estimates for the coefficients of each variant if rare=FALSE and of the risk index if rare=TRUE
Which	Matrix of the unique models as well as their prior probability and posterior probability
Which.r	Matrix indicating which regions are included in each of the unique models given in Which
Coef	Matrix indicating the coefficients of the variants (or risk index) included in each unique model

Author(s)

Melanie Quintana <maw27.wilson@gmail.com>

References

Quintana M, Conti D (2011). *Incorporating Model Uncertainty in Detecting Rare Variants: The Bayesian Risk Index*. Genetic Epidemiology 35:638-649.

Quintana M, Conti D (Submitted). *Integrative Variable Selection via Bayesian Model Uncertainty*.

Examples

```
## RARE VARIANT BRI EXAMPLE
  ## Load the data for Rare variant example
  data(RareData)

  ## Load the results from running sampleBVS on rare variant data for 100K iterations
  data(RareBVS.out)

  ## Summarize output with a burn in of 1000 iterations
  ## Results from summary found in data(RareResults)
  RareResults = summaryBVS(RareBVS.out,data=RareData,burnin=1000,rare=TRUE)

## INFORMATIVE iBMU EXAMPLE
  ##Load the data for the informative example
  data(InformData)

  ## Load the results from running sampleBVS with inform=FALSE for 100K iterations
  data(InformBVS.NI.out)

  ## Summarize output
```

```
## Results from summary found in data(Informresults.NI)

Informresults.NI = summaryBVS(InformBVS.NI.out,data=InformData$data,burnin=1000,
                              regions=InformData$genes,inform=FALSE)

## Load the results from running sampleBVS with inform=TRUE for 100K iterations
data(InformBVS.I.out)

## Summarize output
## Results from summary found in data(Informresults.I)

Informresults.I = summaryBVS(InformBVS.I.out,data=InformData$data,
                             cov=as.matrix(InformData$cov),burnin=1000,
                             regions=InformData$genes,inform=TRUE)
```

Index