Package ‘LearnBayes’

February 19, 2015

Type Package
Title Functions for Learning Bayesian Inference
Version 2.15
Date 2014-05-28
Author Jim Albert
Maintainer Jim Albert <albert@bgsu.edu>
LazyData yes
Description LearnBayes contains a collection of functions helpful in learning the basic tenets of Bayesian statistical inference. It contains functions for summarizing basic one and two parameter posterior distributions and predictive distributions. It contains MCMC algorithms for summarizing posterior distributions defined by the user. It also contains functions for regression models, hierarchical models, Bayesian tests, and illustrations of Gibbs sampling.
License GPL (>= 2)
NeedsCompilation no
Repository CRAN
Date/Publication 2014-05-29 17:33:08

R topics documented:

  achievement .................................................. 3
  baseball.1964 .............................................. 4
  bayes.influence .......................................... 5
  bayes.model.selection .................................... 5
  bayes.probit ............................................. 6
  bayesresiduals ............................................ 7
  bermuda.grass ............................................. 8
  beta.select ............................................... 8
  betabinexch ............................................. 9
  betabinexch0 ............................................ 10
  bfexch .................................................. 11
  btindep .................................................. 11
R topics documented:

- binomial.beta.mix ................................................................. 12
- birdextinct ................................................................. 13
- birthweight ................................................................. 14
- blinreg ................................................................. 14
- blinregexpected ................................................................. 15
- blinregpred ................................................................. 16
- bprobit.probs ................................................................. 17
- bradley.terry.post ................................................................. 18
- breastcancer ................................................................. 18
- calculus.grades ................................................................. 19
- cancermortality ................................................................. 20
- careertraj.setup ................................................................. 20
- cauchyerrorpost ................................................................. 21
- chemotherapy ................................................................. 22
- ctable ................................................................. 22
- darwin ................................................................. 23
- discint ................................................................. 23
- discrete.bayes ................................................................. 24
- discrete.bayes.2 ................................................................. 25
- dmnorm ................................................................. 26
- dmt ................................................................. 26
- donner ................................................................. 27
- election ................................................................. 28
- election.2008 ................................................................. 28
- footballscores ................................................................. 29
- gibbs ................................................................. 29
- groupeddatapost ................................................................. 30
- hearttransplants ................................................................. 31
- hiergibbs ................................................................. 32
- histprior ................................................................. 32
- howardprior ................................................................. 33
- impsampling ................................................................. 34
- indepmetrop ................................................................. 35
- iowagpa ................................................................. 36
- jeter2004 ................................................................. 36
- laplace ................................................................. 37
- lbinorm ................................................................. 38
- logctablepost ................................................................. 39
- logisticonpost ................................................................. 39
- logpoissgamma ................................................................. 40
- logpoissnormal ................................................................. 41
- marathontimes ................................................................. 42
- mnormt.onesided ................................................................. 42
- mnormt.twosided ................................................................. 43
- mycontour ................................................................. 44
- normal.normal.mix ................................................................. 45
- normal.select ................................................................. 46
- normchi2post ................................................................. 46
achievement

School achievement data

Description

Achievement data for a group of Austrian school children

Usage

achievement
**Format**

A data frame with 109 observations on the following 7 variables.

- **Gen** gender of child where 0 is male and 1 is female
- **Age** age in months
- **IQ** iq score
- **math1** test score on mathematics computation
- **math2** test score on mathematics problem solving
- **read1** test score on reading speed
- **read2** test score on reading comprehension

**Source**


---

**baseball.1964**

*Team records in the 1964 National League baseball season*

**Description**

Head to head records for all teams in the 1964 National League baseball season. Teams are coded as Cincinnati (1), Chicago (2), Houston (3), Los Angeles (4), Milwaukee (5), New York (6), Philadelphia (7), Pittsburgh (8), San Francisco (9), and St. Louis (10).

**Usage**

`baseball.1964`

**Format**

A data frame with 45 observations on the following 4 variables.

- **Team.1** Number of team 1
- **Team.2** Number of team 2
- **Wins.Team1** Number of games won by team 1
- **Wins.Team2** Number of games won by team 2

**Source**

bayes.influence  Observation sensitivity analysis in beta-binomial model

Description

Computes probability intervals for the log precision parameter K in a beta-binomial model for all "leave one out" models using sampling importance resampling

Usage

bayes.influence(theta, data)

Arguments

theta  matrix of simulated draws from the posterior of (logit eta, log K)
data  matrix with columns of counts and sample sizes

Value

summary  vector of 5th, 50th, 95th percentiles of log K for complete sample posterior
summary.obs  matrix where the ith row contains the 5th, 50th, 95th percentiles of log K for posterior when the ith observation is removed

Author(s)

Jim Albert

Examples

data(cancermortality)
start=array(c(-7,6),c(1,2))
fit=laplace(betabinexch,start,cancermortality)
tpar=list(m=fit$mode,var=2*fit$var,df=4)
theta=sir(betabinexch,tpar,1000,cancermortality)
intervals=bayes.influence(theta,cancermortality)

bayes.model.selection  Bayesian regression model selection using G priors

Description

Using Zellner’s G priors, computes the log marginal density for all possible regression models

Usage

bayes.model.selection(y, X, c, constant=TRUE)
**bayes.probit**

Simulates from a probit binary response regression model using data augmentation and Gibbs sampling

**Description**

Gives a simulated sample from the joint posterior distribution of the regression vector for a binary response regression model with a probit link and a informative normal(\(\beta, P\)) prior. Also computes the log marginal likelihood when a subjective prior is used.

**Usage**

```r
bayes.probit(y, X, m, prior=list(beta=0, P=0))
```

**Arguments**

- **y**: vector of binary responses
- **X**: covariate matrix
- **m**: number of simulations desired
- **prior**: list with components beta, the prior mean, and P, the prior precision matrix
bayesresiduals

Value
- **beta**: matrix of simulated draws of regression vector beta where each row corresponds to one draw
- **log.marg**: simulation estimate at log marginal likelihood of the model

Author(s)
- Jim Albert

Examples
```r
response=c(0,1,0,0,0,1,1,1,1,1)
covariate=c(1,2,3,4,5,6,7,8,9,10)
x=cbind(1,covariate)
prior=list(beta=c(0,0),P=diag(c(.5,10)))
m=1000
s=bayes.probit(response,x,m,prior)
```

Description
Computes the posterior probabilities that Bayesian residuals exceed a cutoff value for a linear regression model with a noninformative prior

Usage
```r
bayesresiduals(lmfit,post,k)
```

Arguments
- **lmfit**: output of the regression function lm
- **post**: list with components beta, matrix of simulated draws of regression parameter, and sigma, vector of simulated draws of sampling standard deviation
- **k**: cut-off value that defines an outlier

Value
- vector of posterior outlying probabilities

Author(s)
- Jim Albert
Examples

chirps = c(20, 16, 0, 19.8, 18.4, 17.1, 15.5, 14.7, 17.1, 15.4, 16.2, 15, 17.2, 16, 17, 14.1)
 temp = c(88.6, 71.6, 93.3, 84.3, 80.6, 75.2, 69.7, 82, 69.4, 83.3, 78.6, 82.6, 80.6, 83.5, 76.3)
 X = cbind(1, chirps)
 lmf = lm(temp ~ X)
 m = 1000
 post = blinreg(temp, X, m)
 k = 2
 bayesresiduals(lmf, post, k)

bermuda.grass

Bermuda grass experiment data

Description

Yields of bermuda grass for a factorial design of nutrients nitrogen, phosphorus, and potassium.

Usage

bermuda.grass

Format

A data frame with 64 observations on the following 4 variables.

y  yield of bermuda grass in tons per acre
Nit  level of nitrogen
Phos  level of phosphorus
Pot  level of potassium

Source


beta.select

Selection of Beta Prior Given Knowledge of Two Quantiles

Description

Finds the shape parameters of a beta density that matches knowledge of two quantiles of the distribution.

Usage

beta.select(quantile1, quantile2)
Arguments

quantile1 list with components p, the value of the first probability, and x, the value of the first quantile
quantile2 list with components p, the value of the second probability, and x, the value of the second quantile

Value

vector of shape parameters of the matching beta distribution

Author(s)

Jim Albert

Examples

# person believes the median of the prior is 0.25
# and the 90th percentile of the prior is 0.45
quantile1=list(p=.5,x=0.25)
quantile2=list(p=.9,x=0.45)
beta.select(quantile1,quantile2)

betabinexch Log posterior of logit mean and log precision for Binomial/beta exchangeable model

Description

Computes the log posterior density of logit mean and log precision for a Binomial/beta exchangeable model

Usage

betabinexch(theta,data)

Arguments

theta vector of parameter values of logit eta and log K
data a matrix with columns y (counts) and n (sample sizes)

Value

value of the log posterior

Author(s)

Jim Albert
Examples

```r
n=c(20,20,20,20,20)
y=c(1,4,3,6,10)
data=cbind(y,n)
theta=c(-1,0)
betabinexch(theta,data)
```

---

betabinexch

*Log posterior of mean and precision for Binomial/beta exchangeable model*

Description

Computes the log posterior density of mean and precision for a Binomial/beta exchangeable model

Usage

```r
betabinexch(theta,data)
```

Arguments

- `theta` vector of parameter values of eta and K
- `data` a matrix with columns y (counts) and n (sample sizes)

Value

value of the log posterior

Author(s)

Jim Albert

Examples

```r
n=c(20,20,20,20,20)
y=c(1,4,3,6,10)
data=cbind(y,n)
theta=c(.1,10)
betabinexch0(theta,data)
```
bfexch  

Logarithm of integral of Bayes factor for testing homogeneity of proportions

Description
Computes the logarithm of the integral of the Bayes factor for testing homogeneity of a set of proportions

Usage
bfexch(theta, datapar)

Arguments
theta  value of the logit of the prior mean hyperparameter
datapar list with components data, matrix with columns y (counts) and n (sample sizes), and K, prior precision hyperparameter

Value
value of the logarithm of the integral

Author(s)
Jim Albert

Examples
```r
y=c(1,3,2,4,6,4,3)
n=c(10,10,10,10,10,10,10)
data=cbind(y,n)
K=20
datapar=list(data=data,K=K)
theta=1
bfexch(theta, datapar)
```

bfindep  Bayes factor against independence assuming alternatives close to independence

Description
Computes a Bayes factor against independence for a two-way contingency table assuming a “close to independence” alternative model
Usage

`bfindep(y,K,m)`

Arguments

- `y`: matrix of counts
- `K`: Dirichlet precision hyperparameter
- `m`: number of simulations

Value

- `bf`: value of the Bayes factor against hypothesis of independence
- `nse`: estimate of the simulation standard error of the computed Bayes factor

Author(s)

Jim Albert

Examples

```r
y = matrix(c(10,4,6,3,6,10),c(2,3))
K = 20
m = 1000
bfindep(y,K,m)
```

---

**binomial.beta.mix**

Computes the posterior for binomial sampling and a mixture of betas

Description

Computes the parameters and mixing probabilities for a binomial sampling problem where the prior is a discrete mixture of beta densities.

Usage

`binomial.beta.mix(probs,betapar,data)`

Arguments

- `probs`: vector of probabilities of the beta components of the prior
- `betapar`: matrix where each row contains the shape parameters for a beta component of the prior
- `data`: vector of number of successes and number of failures
**Value**

probs vector of probabilities of the beta components of the posterior  
betapar matrix where each row contains the shape parameters for a beta component of the posterior  

**Author(s)**

Jim Albert  

**Examples**

```r  
probs=c(.5, .5)  
beta.par1=c(15,5)  
beta.par2=c(10,10)  
betapar=rbind(beta.par1,beta.par2)  
data=c(20,15)  
binomial.beta.mix(probs,betapar,data)  
```

**birdextinct**  

*Bird measurements from British islands*

**Description**

Measurements on breedings pairs of landbird species were collected from 16 islands about Britain over several decades.

**Usage**

birdextinct

**Format**

A data frame with 62 observations on the following 5 variables.

- **species** name of bird species  
- **time** average time of extinction on the islands  
- **nesting** average number of nesting pairs  
- **size** size of the species, 1 or 0 if large or small  
- **status** status of the species, 1 or 0 if resident or migrant

**Source**

**Description**

Dobson describes a study where one is interested in predicting a baby’s birthweight based on the gestational age and the baby’s gender.

**Usage**

birthweight

**Format**

A data frame with 24 observations on the following 3 variables.

- **age** gestational age in weeks
- **gender** gender of the baby where 0 (1) is male (female)
- **weight** birthweight of baby in grams

**Source**


---

**blinreg**

**Simulation from Bayesian linear regression model**

**Description**

Gives a simulated sample from the joint posterior distribution of the regression vector and the error standard deviation for a linear regression model with a noninformative or g prior.

**Usage**

blinreg(y,X,m,prior=NULL)

**Arguments**

- **y** vector of responses
- **X** design matrix
- **m** number of simulations desired
- **prior** list with components c0 and beta0 of Zellner’s g prior
blinregeexpected

Value

- **beta**: matrix of simulated draws of beta where each row corresponds to one draw
- **sigma**: vector of simulated draws of the error standard deviation

Author(s)

Jim Albert

Examples

```r
chirps<-c(20.0,16.8,19.4,18.1,15.5,14.7,17.1,15.4,16.2,15.7,16.1,17.1)
temp<-c(88.6,71.6,93.3,84.3,80.6,75.2,69.7,82.6,69.4,83.3,78.6,82.6,80.6,83.5,76.3)
X<-cbind(1,chirps)
m=1000
s=blinrege(temp,X,m)
```

---

**blinregeexpected**  
Simulates values of expected response for linear regression model

Description

Simulates draws of the posterior distribution of an expected response for a linear regression model with a noninformative prior

Usage

```r
blinregeexpected(X1,theta.sample)
```

Arguments

- **X1**: matrix where each row corresponds to a covariate set
- **theta.sample**: list with components beta, matrix of simulated draws of regression vector, and sigma, vector of simulated draws of sampling error standard deviation

Value

matrix where a column corresponds to the simulated draws of the expected response for a given covariate set

Author(s)

Jim Albert
Examples

```r
chirps=c(20,16.0,19.8,18.4,17.1,15.5,14.7,17.1,15.4,16.2,15,17.2,16,17,14.1)
temp=c(88.6,71.6,93.3,84.3,80.6,75.2,69.7,82,69.4,83.3,78.6,82.6,80.6,83.5,76.3)
X=cbind(1,chirps)
m=1000
theta.sample=blinreg(temp,X,m)
covset1=c(1,15)
covset2=c(1,20)
X1=rbind(covset1,covset2)
blinregpred(X1,theta.sample)
```

---

**blinregpred**  
*Simulates values of predicted response for linear regression model*

**Description**

Simulates draws of the predictive distribution of a future response for a linear regression model with a noninformative prior.

**Usage**

```r
blinregpred(X1, theta.sample)
```

**Arguments**

- `X1`  
  - matrix where each row corresponds to a covariate set
- `theta.sample`  
  - list with components beta, matrix of simulated draws of regression vector, and sigma, vector of simulated draws of sampling error standard deviation

**Value**

- matrix where a column corresponds to the simulated draws of the predicted response for a given covariate set

**Author(s)**

Jim Albert

**Examples**

```r
chirps=c(20,16.0,19.8,18.4,17.1,15.5,14.7,17.1,15.4,16.2,15,17.2,16,17,14.1)
temp=c(88.6,71.6,93.3,84.3,80.6,75.2,69.7,82,69.4,83.3,78.6,82.6,80.6,83.5,76.3)
X=cbind(1,chirps)
m=1000
theta.sample=blinreg(temp,X,m)
covset1=c(1,15)
covset2=c(1,20)
X1=rbind(covset1,covset2)
blinregpred(X1,theta.sample)
```
Simulates fitted probabilities for a probit regression model

Description

Gives a simulated sample for fitted probabilities for a binary response regression model with a probit link and noninformative prior.

Usage

bprobit.probs(X1, fit)

Arguments

X1                     matrix where each row corresponds to a covariate set
fit                    simulated matrix of draws of the regression vector

Value

matrix of simulated draws of the fitted probabilities, where a column corresponds to a particular covariate set

Author(s)

Jim Albert

Examples

response=c(0,1,0,0,1,1,1,1,1,1)
covariate=c(1,2,3,4,5,6,7,8,9,10)
x=cbind(1,covariate)
m=1000
fit=bayes.probit(response,X,m)
x1=c(1,3)
x2=c(1,8)
X1=rbind(x1,x2)
fittedprobs=bprobit.probs(X1,fit$beta)
bradley.terry.post  

Log posterior of a Bradley Terry random effects model

Description

Computes the log posterior density of the talent parameters and the log standard deviation for a Bradley Terry model with normal random effects

Usage

bradley.terry.post(theta, data)

Arguments

theta  
vector of talent parameters and log standard deviation

data  
data matrix with columns team1, team2, wins by team1, and wins by team2

Value

value of the log posterior

Author(s)

Jim Albert

Examples

data(baseball.1964)
team.strengths=rep(0,10)
log.sigma=0
bradley.terry.post(c(team.strengths,log.sigma),baseball.1964)

breastcancer  

Survival experience of women with breast cancer under treatment

Description

Collett (1994) describes a study to evaluate the effectiveness of a histochemical marker in predicting the survival experience of women with breast cancer.

Usage

breastcancer
calculus.grades

Format

A data frame with 45 observations on the following 3 variables.

- **time**: survival time in months
- **status**: censoring indicator where 1 (0) indicates a complete (censored) survival time
- **stain**: indicates by a 0 (1) if tumor was negatively (positively) stained

Source


calculus.grades

Calculus grades dataset

Description

Grades and other variables collected for a sample of calculus students.

Usage

calculus.grades

Format

A data frame with 100 observations on the following 3 variables.

- **grade**: indicates if student received a A or B in class
- **prev.grade**: indicates if student received a A in prerequisite math class
- **act**: score on the ACT math test

Source

Collected by a colleague of the author at his university.
cancermortality  
\textit{Cancer mortality data}

**Description**

Number of cancer deaths and number at risk for 20 cities in Missouri.

**Usage**

cancermortality

**Format**

A data frame with 20 observations on the following 2 variables.

- \textit{y} number of cancer deaths
- \textit{n} number at risk

**Source**


careertraj.setup  
\textit{Setup for Career Trajectory Application}

**Description**

Sets up the data matrices for the use of WinBUGS in the career trajectory application.

**Usage**

careertraj.setup(data)

**Arguments**

data  
data matrix for ballplayers with variables Player, Year, Age, G, AB, R, H, X2B, X3B, HR, RBI, BB, SO

**Value**

- \textit{player.names}  
  vector of player names
- \textit{y}  
  matrix of home runs for players where a row corresponds to the home runs for a player during all the years of his career
- \textit{n}  
  matrix of AB-SO for all players
- \textit{x}  
  matrix of ages for all players for all years of their careers
- \textit{T}  
  vector of number of seasons for all players
- \textit{N}  
  number of players
cauchyerrorpost

Author(s)
Jim Albert

Examples
data(sluggerdata)
careertraj.setup(sluggerdata)

cauchyerrorpost  

Description
Computes the log posterior density of (M, log S) when a sample is taken from a Cauchy density with location M and scale S and a uniform prior distribution is taken on (M, log S)

Usage
cauchyerrorpost(theta, data)

Arguments
theta  vector of parameter values of M and log S
data  vector containing sample of observations

Value
value of the log posterior

Author(s)
Jim Albert

Examples
data=c(108, 51, 7, 43, 52, 54, 53, 49, 21, 48)
theta=c(40, 1)
cauchyerrorpost(theta, data)
chemotherapy

*Chemotherapy treatment effects on ovarian cancer*

**Description**

Edmunson et al (1979) studied the effect of different chemotherapy treatments following surgical treatment of ovarian cancer.

**Usage**

chemotherapy

**Format**

A data frame with 26 observations on the following 5 variables.

- **patient**  patient number
- **time**  survival time in days following treatment
- **status**  indicates if time is censored (0) or actually observed (1)
- **treat**  control group (0) or treatment group (1)
- **age**  age of the patient

**Source**


catable

*Bayes factor against independence using uniform priors*

**Description**

Computes a Bayes factor against independence for a two-way contingency table assuming uniform prior distributions

**Usage**

catable(y,a)

**Arguments**

- **y**  matrix of counts
- **a**  matrix of prior hyperparameters
**Value**

value of the Bayes factor against independence

**Author(s)**

Jim Albert

**Examples**

```r
y = matrix(c(10,4,6,3,6,10),c(3,2))
a = matrix(rep(1,6),c(2,3))
ctable(y,a)
```

---

### darrow

**Darwin’s data on plants**

---

**Description**

Fifteen differences of the heights of cross and self fertilized plants quoted by Fisher (1960)

**Usage**

darrow

**Format**

A data frame with 15 observations on the following 1 variable.

- **difference** difference of heights of two types of plants

**Source**


---

### discint

**Highest probability interval for a discrete distribution**

**Description**

Computes a highest probability interval for a discrete probability distribution

**Usage**

discint(dist, prob)
### discrete.bayes

**Posterior distribution with discrete priors**

Computes the posterior distribution for an arbitrary one parameter distribution for a discrete prior distribution.

**Usage**

`discrete.bayes(df,prior,y,...)`

**Arguments**

- `df` name of the function defining the sampling density
- `prior` vector defining the prior density; names of the vector define the parameter values and entries of the vector define the prior probabilities
- `y` vector of data values
- `...` any further fixed parameter values used in the sampling density function

**Value**

- `prob` vector of posterior probabilities
- `pred` scalar with prior predictive probability
Author(s)
Jim Albert

Examples

```r
class=prior=c(.25,.25,.25,.25) names(prior)=c(.2,.25,.3,.35) y=5 n=10 discrete.bayes(dbinom,prior,y,size=n)
```

Description
Computes the posterior distribution for an arbitrary two parameter distribution for a discrete prior distribution.

Usage

```r
discrete.bayes.2(df,prior,y=NULL,...)
```

Arguments

- `df`: name of the function defining the sampling density of two parameters
- `prior`: matrix defining the prior density; the row names and column names of the matrix define respectively the values of parameter 1 and values of parameter 2 and the entries of the matrix give the prior probabilities
- `y`: `y` is a matrix of data values, where each row corresponds to a single observation
- `...`: any further fixed parameter values used in the sampling density function

Value

- `prob`: matrix of posterior probabilities
- `pred`: scalar with prior predictive probability

Author(s)
Jim Albert

Examples

```r
p1 = seq(0.1, 0.9, length = 9) p2 = p1 prior = matrix(1/81, 9, 9) dimnames(prior)[[1]] = p1 dimnames(prior)[[2]] = p2 discrete.bayes.2(twoproplike,prior)
```
dmnorm

The probability density function for the multivariate normal (Gaussian) probability distribution

Description
Computes the density of a multivariate normal distribution

Usage
dmnorm(x, mean = rep(0, d), varcov, log = FALSE)

Arguments
- **x**: vector of length d or matrix with d columns, giving the coordinates of points where density is to evaluated
- **mean**: numeric vector giving the location parameter of the distribution
- **varcov**: a positive definite matrix representing the scale matrix of the distribution
- **log**: a logical value; if TRUE, the logarithm of the density is to be computed

Value
vector of density values

Author(s)
Jim Albert

Examples
mu <- c(1,12,2)
Sigma <- matrix(c(1,2,0,2,5,0.5,0,0.5,3, 3, 3)
x <- c(2,14,0)
f <- dmnorm(x, mu, Sigma)

dmt

Probability density function for multivariate t

Description
Computes the density of a multivariate t distribution

Usage
dmt(x, mean = rep(0, d), S, df = Inf, log=FALSE)
Arguments

- **x**: vector of length d or matrix with d columns, giving the coordinates of points where density is to be evaluated
- **mean**: numeric vector giving the location parameter of the distribution
- **S**: a positive definite matrix representing the scale matrix of the distribution
- **df**: degrees of freedom
- **log**: a logical value; if TRUE, the logarithm of the density is to be computed

Value

- vector of density values

Author(s)

Jim Albert

Examples

```r
mu <- c(1,12,2)
Sigma <- matrix(c(1,2,0,2,5,0.5,0,0.5,3), 3, 3)
df <- 4
x <- c(2,14,0)
f <- dmt(x, mu, Sigma, df)
```

Description

Data contains the age, gender and survival status for 45 members of the Donner Party who experienced difficulties in crossing the Sierra Nevada mountains in California.

Usage

`donner`

Format

A data frame with 45 observations on the following 3 variables.

- **age**: age of person
- **male**: gender that is 1 (0) if person is male (female)
- **survival**: survival status, 1 or 0 if person survived or died

Source

Florida election data

Description
For each of the Florida counties in the 2000 presidential election, the number of votes for George Bush, Al Gore, and Pat Buchanan is recorded. Also the number of votes for the minority candidate Ross Perot in the 1996 presidential election is recorded.

Usage
election

Format
A data frame with 67 observations on the following 5 variables.

- county: name of Florida county
- perot: number of votes for Ross Perot in 1996 election
- gore: number of votes for Al Gore in 2000 election
- bush: number of votes for George Bush in 2000 election
- buchanan: number of votes for Pat Buchanan in 2000 election

Poll data from 2008 U.S. Presidential Election

Description
Results of recent state polls in the 2008 United States Presidential Election between Barack Obama and John McCain.

Usage
election.2008

Format
A data frame with 51 observations on the following 4 variables.

- State: name of the state
- M.pct: percentage of poll survey for McCain
- O.pct: percentage of poll survey for Obama
- EV: number of electoral votes

Source
**footballscores**

*Game outcomes and point spreads for American football*

**Description**

Game outcomes and point spreads for 672 professional American football games.

**Usage**

`footballscores`

**Format**

A data frame with 672 observations on the following 8 variables.

- **year** year of game
- **home** indicates if favorite is the home team
- **favorite** score of favorite team
- **underdog** score of underdog team
- **spread** point spread
- **favorite.name** name of favorite team
- **underdog.name** name of underdog team
- **week** week number of the season

**Source**


---

**gibbs**

*Metropolis within Gibbs sampling algorithm of a posterior distribution*

**Description**

Implements a Metropolis-within-Gibbs sampling algorithm for an arbitrary real-valued posterior density defined by the user.

**Usage**

`gibbs(logpost,start,m,scale,...)`
groupeddatapost

Description

Computes the log posterior density of \((M, \log S)\) for normal sampling where the data is observed in grouped form.

Usage

groupeddatapost(theta, data)

Arguments

- **theta**: vector of parameter values \(M\) and \(\log S\)
- **data**: list with components int.lo, a vector of left endpoints, int.hi, a vector of right endpoints, and f, a vector of bin frequencies

Value

value of the log posterior
hearttransplants

Author(s)

Jim Albert

Examples

```r
int.lo=c(-Inf,10,15,20,25)
int.hi=c(10,15,20,25,Inf)
f=c(2,5,8,4,2)
data=list(int.lo=int.lo,int.hi=int.hi,f=f)
theta=c(20,1)
groupeddatapost(theta,data)
```

hearttransplants  Heart transplant mortality data

Description

The number of deaths within 30 days of heart transplant surgery for 94 U.S. hospitals that performed at least 10 heart transplant surgeries. Also the exposure, the expected number of deaths, is recorded for each hospital.

Usage

hearttransplants

Format

A data frame with 94 observations on the following 2 variables.

- **e** expected number of deaths (the exposure)
- **y** observed number of deaths within 30 days of heart transplant surgery

Source

hiergibbs  
*Gibbs sampling for a hierarchical regression model*

**Description**

Implements Gibbs sampling for estimating a two-way table of means under a hierarchical regression model.

**Usage**

```r
hiergibbs(data,m)
```

**Arguments**

- `data` : data matrix with columns observed sample means, sample sizes, and values of two covariates
- `m` : number of cycles of Gibbs sampling

**Value**

- `beta` : matrix of simulated values of regression vector
- `mu` : matrix of simulated values of cell means
- `var` : vector of simulated values of second-stage prior variance

**Author(s)**

Jim Albert

**Examples**

```r
data(iowagpa)
m=1000
s=hiergibbs(iowagpa,m)
```

---

histprior  
*Density function of a histogram distribution*

**Description**

Computes the density of a probability distribution defined on a set of equal-width intervals

**Usage**

```r
histprior(p,midpts,prob)
```
**howardprior**

**Arguments**

- `p` vector of values for which density is to be computed
- `midpts` vector of midpoints of the intervals
- `prob` vector of probabilities of the intervals

**Value**

vector of values of the probability density

**Author(s)**

Jim Albert

**Examples**

```r
midpts=c(.1,.3,.5,.7,.9)
prob=c(.2,.2,.4,.1,.1)
p=seq(.01,.99,by=.01)
plot(p,histprior(p,midpts,prob),type="l")
```

---

**Description**

Computes the logarithm of a dependent prior on two proportions proposed by Howard in a Statistical Science paper in 1998.

**Usage**

```r
howardprior(xy,par)
```

**Arguments**

- `xy` vector of proportions p1 and p2
- `par` vector containing parameter values alpha, beta, gamma, delta, sigma

**Value**

value of the log posterior

**Author(s)**

Jim Albert
Impsampling  

Importance sampling using a t proposal density

Description

Implements importance sampling to compute the posterior mean of a function using a multivariate t proposal density

Usage

impsampling(logf, tpar, h, n, data)

Arguments

- `logf`: function that defines the logarithm of the density of interest
- `tpar`: list of parameters of t proposal density including the mean m, scale matrix var, and degrees of freedom df
- `h`: function that defines h(theta)
- `n`: number of simulated draws from proposal density
- `data`: data and or parameters used in the function logf

Value

- `est`: estimate at the posterior mean
- `se`: simulation standard error of estimate
- `theta`: matrix of simulated draws from proposal density
- `wt`: vector of importance sampling weights

Author(s)

Jim Albert

Examples

data(cancermortality)
start=c(-7,6)
fit=laplace(betabinchx, start, cancermortality)
tpar=list(m=fit$mode, var=2*fit$var, df=4)
myfunc=function(theta) return(theta[2])
theta=impsampling(betabinchx, tpar, myfunc, 1000, cancermortality)
indepmetrop

| indepmetrop | Independence Metropolis independence chain of a posterior distribution |

Description

Simulates iterates of an independence Metropolis chain with a normal proposal density for an arbitrary real-valued posterior density defined by the user

Usage

indepmetrop(logpost, proposal, start, m, ...)

Arguments

logpost function defining the log posterior density

proposal a list containing mu, an estimated mean and var, an estimated variance-covariance matrix, of the normal proposal density

start vector containing the starting value of the parameter

m the number of iterations of the chain

... data that is used in the function logpost

Value

par a matrix of simulated values where each row corresponds to a value of the vector parameter

accept the acceptance rate of the algorithm

Author(s)

Jim Albert

Examples

data=c(6,2,3,10)
proposal=list(mu=array(c(2.3,-1),c(2,1)), var=diag(c(1,1)))
start=array(c(0,0),c(1,2))
m=1000
fit=indepmetrop(logctablepost, proposal, start, m, data)
iowagpa  
*Admissions data for an university*

**Description**

Students at a major university are categorized with respect to their high school rank and their ACT score. For each combination of high school rank and ACT score, one records the mean grade point average (GPA).

**Usage**

iowagpa

**Format**

A data frame with 40 observations on the following 4 variables.

- **gpa**  mean grade point average
- **n**  sample size
- **HSR**  high school rank
- **ACT**  act score

**Source**


---

**jeter2004**  
*Hitting data for Derek Jeter*

**Description**

Batting data for the baseball player Derek Jeter for all 154 games in the 2004 season.

**Usage**

jeter2004
Format

A data frame with 154 observations on the following 10 variables.

- **Game**: the game number
- **AB**: the number of at-bats
- **R**: the number of runs scored
- **H**: the number of hits
- **X2B**: the number of doubles
- **X3B**: the number of triples
- **HR**: the number of home runs
- **RBI**: the number of runs batted in
- **BB**: the number of walks
- **SO**: the number of strikeouts

Source


---

**laplace**

*Summarization of a posterior density by the Laplace method*

Description

For a general posterior density, computes the posterior mode, the associated variance-covariance matrix, and an estimate at the logarithm at the normalizing constant.

Usage

```r
laplace(logpost, mode, ...) 
```

Arguments

- `logpost`: function that defines the logarithm of the posterior density
- `mode`: vector that is a guess at the posterior mode
- `...`: vector or list of parameters associated with the function logpost

Value

- `mode`: current estimate at the posterior mode
- `var`: current estimate at the associated variance-covariance matrix
- `int`: estimate at the logarithm of the normalizing constant
- `converge`: indication (TRUE or FALSE) if the algorithm converged
**Author(s)**

Jim Albert

**Examples**

```r
logpost=function(theta,data)
{
  s=5
  sum(-log(1+(data-theta)^2/s^2))
}
data=c(10,12,14,13,12,15)
start=10
laplace(logpost,start,data)
```

---

**lbinorm**

*Logarithm of bivariate normal density*

**Description**

Computes the logarithm of a bivariate normal density

**Usage**

```r
lbinorm(xy,par)
```

**Arguments**

- `xy`: vector of values of two variables x and y
- `par`: list with components m, a vector of means, and v, a variance-covariance matrix

**Value**

value of the kernel of the log density

**Author(s)**

Jim Albert

**Examples**

```r
mean=c(0,0)
varcov=diag(c(1,1))
value=c(1,1)
param=list(m=mean,v=varcov)
lbinorm(value,param)
```
logctablepost

Log posterior of difference and sum of logits in a 2x2 table

Description
Computes the log posterior density for the difference and sum of logits in a 2x2 contingency table for independent binomial samples and uniform prior placed on the logits.

Usage
logctablepost(theta, data)

Arguments
theta vector of parameter values "difference of logits" and "sum of logits"
data vector containing number of successes and failures for first sample, and then second sample

Value
value of the log posterior

Author(s)
Jim Albert

Examples
s1=6; f1=2; s2=3; f2=10
data=c(s1,f1,s2,f2)
theta=c(2,4)
logctablepost(theta, data)

logisticpost

Log posterior for a binary response model with a logistic link and a uniform prior

Description
Computes the log posterior density of (beta0, beta1) when yi are independent binomial(ni, pi) and logit(pi)=beta0+beta1*xi and a uniform prior is placed on (beta0, beta1)

Usage
logisticpost(beta, data)
logpoissgamma

Description

Computes the logarithm of the posterior density of a Poisson log mean with a gamma prior

Usage

logpoissgamma(theta, datapar)

Arguments

theta vector of values of the log mean parameter
datapar list with components data, vector of observations, and par, vector of parameters of the gamma prior

Value

vector of values of the log posterior for all values in theta

Author(s)

Jim Albert
logpoissnormal

Examples

data=c(2,4,3,6,1,0,4,3,10,2)
par=c(1,1)
datapar=list(data=data,par=par)
theta=c(-1,0,1,2)
logpoissnormal(theta,datapar)

logpoissnormal, Log posterior with Poisson sampling and normal prior

Description

Computes the logarithm of the posterior density of a Poisson log mean with a normal prior

Usage

logpoissnormal(theta,datapar)

Arguments

theta vector of values of the log mean parameter
datapar list with components data, vector of observations, and par, vector of parameters of the normal prior

Value

vector of values of the log posterior for all values in theta

Author(s)

Jim Albert

Examples

data=c(2,4,3,6,1,0,4,3,10,2)
par=c(0,1)
datapar=list(data=data,par=par)
theta=c(-1,0,1,2)
logpoissnormal(theta,datapar)
marathontimes  Marathon running times

Description
Running times in minutes for twenty male runners between the ages 20 and 29 who ran the New York Marathon.

Usage
marathontimes

Format
A data frame with 20 observations on the following 1 variable.

time  running time

Source
www.nycmarathon.org website.

mnormt.onesided  Bayesian test of one-sided hypothesis about a normal mean

Description
Computes a Bayesian test of the hypothesis that a normal mean is less than or equal to a specified value.

Usage
mnormt.onesided(m0,normpar,data)

Arguments

m0  value of the normal mean to be tested
normpar  vector of mean and standard deviation of the normal prior distribution
data  vector of sample mean, sample size, and known value of the population standard deviation

Value
BF  Bayes factor in support of the null hypothesis
prior.odds  prior odds of the null hypothesis
post.odds  posterior odds of the null hypothesis
postH  posterior probability of the null hypothesis
**mnormt.twosided**

**Author(s)**

Jim Albert

**Examples**

```r
y = c(182, 172, 173, 176, 176, 180, 173, 174, 179, 175)
pop.s = 3
data = c(mean(y), length(data), pop.s)
m0 = 175
normpar = c(170, 1000)
mnormt.onesided(m0, normpar, data)
```

---

**mnormt.twosided**  
Bayesian test of a two-sided hypothesis about a normal mean

**Description**

Bayesian test that a normal mean is equal to a specified value using a normal prior

**Usage**

```r
mnormt.twosided(m0, prob, t, data)
```

**Arguments**

- `m0`: value of the mean to be tested
- `prob`: prior probability of the hypothesis
- `t`: vector of values of the prior standard deviation under the alternative hypothesis
- `data`: vector containing the sample mean, the sample size, and the known value of the population standard deviation

**Value**

- `bf`: vector of values of the Bayes factor in support of the null hypothesis
- `post`: vector of posterior probabilities of the null hypothesis

**Author(s)**

Jim Albert
**mycontour**

### Description
For a general two parameter density, draws a contour graph where the contour lines are drawn at 10 percent, 1 percent, and .1 percent of the height at the mode.

### Usage
```
mycontour(logf, limits, data, ...)
```

### Arguments
- `logf` function that defines the logarithm of the density
- `limits` limits (xlo, xhi, ylo, yhi) where the graph is to be drawn
- `data` vector or list of parameters associated with the function logpost
- `...` further arguments to pass to contour

### Value
A contour graph of the density is drawn

### Author(s)
Jim Albert

### Examples
```
# Example for normal distribution
m = array(c(0,0),c(2,1))
v = array(c(1.6,1),c(2,2))
normpar = list(m=m, v=v)
mycontour(lbinorm, c(-4,-4), normpar)
```
normal.normal.mix  Computes the posterior for normal sampling and a mixture of normals prior

Description

Computes the parameters and mixing probabilities for a normal sampling problem, variance known, where the prior is a discrete mixture of normal densities.

Usage

normal.normal.mix(probs, normalpar, data)

Arguments

probs  vector of probabilities of the normal components of the prior
normalpar  matrix where each row contains the mean and variance parameters for a normal component of the prior
data  vector of observation and sampling variance

Value

probs  vector of probabilities of the normal components of the posterior
normalpar  matrix where each row contains the mean and variance parameters for a normal component of the posterior

Author(s)

Jim Albert

Examples

probs=c(.5, .5)
normal.par1=c(0,1)
normal.par2=c(2,.5)
normalpar=rbind(normal.par1,normal.par2)
y=1; sigma2=.5
data=c(y,sigma2)
normal.normal.mix(probs,normalpar,data)
normal.select  

*Selection of Normal Prior Given Knowledge of Two Quantiles*

**Description**

Finds the mean and standard deviation of a normal density that matches knowledge of two quantiles of the distribution.

**Usage**

```r
normal.select(quantile1, quantile2)
```

**Arguments**

- `quantile1`: list with components `p`, the value of the first probability, and `x`, the value of the first quantile.
- `quantile2`: list with components `p`, the value of the second probability, and `x`, the value of the second quantile.

**Value**

- `mean`: mean of the matching normal distribution.
- `sigma`: standard deviation of the matching normal distribution.

**Author(s)**

Jim Albert

**Examples**

```r
# person believes the 15th percentile of the prior is 100
# and the 70th percentile of the prior is 150
quantile1=list(p=.15,x=100)
quantile2=list(p=.7,x=150)
normal.select(quantile1,quantile2)
```

---

Log posterior density for mean and variance for normal sampling

**Description**

Computes the log of the posterior density of a mean `M` and a variance `S^2` when a sample is taken from a normal density and a standard noninformative prior is used.
normnormexch

Usage

normchi2post(theta, data)

Arguments

theta vector of parameter values M and S2
data vector containing the sample observations

Value

value of the log posterior

Author(s)

Jim Albert

Examples

parameter = c(25, 5)
data = c(20, 32, 21, 43, 33, 21, 32)
normchi2post(parameter, data)

normnormexch  Log posterior of mean and log standard deviation for Normal/Normal exchangeable model

Description

Computes the log posterior density of mean and log standard deviation for a Normal/Normal exchangeable model where (mean, log sd) is given a uniform prior.

Usage

normnormexch(theta, data)

Arguments

theta vector of parameter values of mu and log tau
data a matrix with columns y (observations) and v (sampling variances)

Value

value of the log posterior

Author(s)

Jim Albert
Examples

```r
s.var <- c(0.05, 0.05, 0.05, 0.05, 0.05)
y.means <- c(1, 4, 3, 6, 10)
data=cbind(y.means, s.var)
theta=c(-1, 0)
normnormexch(theta, data)
```

**normpostpred**

*Posterior predictive simulation from Bayesian normal sampling model*

**Description**

Given simulated draws from the posterior from a normal sampling model, outputs simulated draws from the posterior predictive distribution of a statistic of interest.

**Usage**

```r
normpostpred(parameters, sample.size, f=min)
```

**Arguments**

- `parameters` : list of simulated draws from the posterior where mu contains the normal mean and sigma2 contains the normal variance
- `sample.size` : size of sample of future sample
- `f` : function defining the statistic

**Value**

simulated sample of the posterior predictive distribution of the statistic

**Author(s)**

Jim Albert

**Examples**

```r
# finds posterior predictive distribution of the min statistic of a future sample of size 15
data(darwin)
s=normpostsim(darwin$difference)
sample.size=15
sim.stats=normpostpred(s, sample.size, min)
```
normpostsim  

Simulation from Bayesian normal sampling model

Description

Gives a simulated sample from the joint posterior distribution of the mean and variance for a normal sampling prior with a noninformative or informative prior. The prior assumes mu and sigma2 are independent with mu assigned a normal prior with mean mu0 and variance tau2, and sigma2 is assigned a inverse gamma prior with parameters a and b.

Usage

normpostsim(data,prior=NULL,m=1000)

Arguments

- data: vector of observations
- prior: list with components mu, a vector with the prior mean and variance, and sigma2, a vector of the inverse gamma parameters
- m: number of simulations desired

Value

- mu: vector of simulated draws of normal mean
- sigma2: vector of simulated draws of normal variance

Author(s)

Jim Albert

Examples

data(darwin)
s=normpostsim(darwin$difference)

ordergibbs  

Gibbs sampling for a hierarchical regression model

Description

Implements Gibbs sampling for estimating a two-way table of means under a order restriction.

Usage

ordergibbs(data,m)
Arguments
data  data matrix with first two columns observed sample means and sample sizes
m  number of cycles of Gibbs sampling

Value
matrix of simulated draws of the normal means where each row represents one simulated draw

Author(s)
Jim Albert

Examples
```
data(iowagpa)
m=1000
s=ordergibbs(iowagpa,m)
```

---

**pbetap**  
*Predictive distribution for a binomial sample with a beta prior*

Description
Computes predictive distribution for number of successes of future binomial experiment with a beta prior distribution for the proportion.

Usage
```
pbetap(ab, n, s)
```

Arguments
- `ab`  vector of parameters of the beta prior
- `n`  size of future binomial sample
- `s`  vector of number of successes for future binomial experiment

Value
vector of predictive probabilities for the values in the vector `s`

Author(s)
Jim Albert
**pbetat**

**Examples**

```r
ab = c(3, 12)
n = 10
s = 0:10
pbetap(ab, n, s)
```

**Description**

Bayesian test that a proportion is equal to a specified value using a beta prior

**Usage**

```r
pbetat(p0, prob, ab, data)
```

**Arguments**

- `p0`: value of the proportion to be tested
- `prob`: prior probability of the hypothesis
- `ab`: vector of parameter values of the beta prior under the alternative hypothesis
- `data`: vector containing the number of successes and number of failures

**Value**

- `bf`: the Bayes factor in support of the null hypothesis
- `post`: the posterior probability of the null hypothesis

**Author(s)**

Jim Albert

**Examples**

```r
p0 = .5
prob = .5
ab = c(10, 10)
data = c(5, 15)
pbetat(p0, prob, ab, data)
```
pdiscp

Posterior distribution for a proportion with discrete priors

Description

Computes the posterior distribution for a proportion for a discrete prior distribution.

Usage

pdiscHpL priorL dataI

Arguments

p vector of proportion values
prior vector of prior probabilities
data vector consisting of number of successes and number of failures

Value

vector of posterior probabilities

Author(s)

Jim Albert

Examples

p=c(.2,.25,.3,.35)
prior=c(.25,.25,.25,.25)
data=c(5,10)
pdisc(p,prior,data)

pdiscp

Predictive distribution for a binomial sample with a discrete prior

Description

Computes predictive distribution for number of successes of future binomial experiment with a discrete distribution for the proportion.

Usage

pdiscp(p, probs, n, s)
poissgamexch

Arguments

- p: vector of proportion values
- probs: vector of probabilities
- n: size of future binomial sample
- s: vector of number of successes for future binomial experiment

Value

vector of predictive probabilities for the values in the vector s

Author(s)

Jim Albert

Examples

```r
p = c(.1, .2, .3, .4, .5, .6, .7, .8, .9)
prob = c(0.05, 0.10, 0.10, 0.15, 0.20, 0.15, 0.10, 0.10, 0.05)
n = 10
s = 0:10
pdiscp(p, prob, n, s)
```

---

**poissgamexch** *Log posterior of Poisson/gamma exchangeable model*

Description

Computes the log posterior density of log alpha and log mu for a Poisson/gamma exchangeable model

Usage

```r
poissgamexch(theta, datapar)
```

Arguments

- theta: vector of parameter values of log alpha and log mu
- datapar: list with components data, a matrix with columns e and y, and z0, prior hyper-parameter

Value

value of the log posterior

Author(s)

Jim Albert
Examples

```r
e=c(532,584,672,722,904)
y=c(0,0,2,1,1)
data=cbind(e,y)
theta=c(-4,0)
z0=.5
datapar=list(data=data,z0=z0)
poissgamexch(theta,datapar)
```

---

**poisson.gamma.mix**  
*Computes the posterior for Poisson sampling and a mixture of gammas prior*

Description

Computes the parameters and mixing probabilities for a Poisson sampling problem where the prior is a discrete mixture of gamma densities.

Usage

```r
poisson.gamma.mix(probs,gammapar,data)
```

Arguments

- **probs**  
  vector of probabilities of the gamma components of the prior

- **gammapar**  
  matrix where each row contains the shape and rate parameters for a gamma component of the prior

- **data**  
  list with components `y`, vector of counts, and `t`, vector of time intervals

Value

- **probs**  
  vector of probabilities of the gamma components of the posterior

- **gammapar**  
  matrix where each row contains the shape and rate parameters for a gamma component of the posterior

Author(s)

Jim Albert

Examples

```r
probs=c(.5,.5)
gamma.par1=c(1,1)
gamma.par2=c(10,2)
gammapar=rbind(gamma.par1,gamma.par2)
y=c(1,3,2,4,10); t=c(1,1,1,1,1)
data=list(y=y,t=t)
poisson.gamma.mix(probs,gammapar,data)
```
**predplot**

*Plot of predictive distribution for binomial sampling with a beta prior*

**Description**

For a proportion problem with a beta prior, plots the prior predictive distribution of the number of successes in n trials and displays the observed number of successes.

**Usage**

```r
predplot(prior, n, yobs)
```

**Arguments**

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>prior</td>
<td>vector of parameters for beta prior</td>
</tr>
<tr>
<td>n</td>
<td>sample size</td>
</tr>
<tr>
<td>yobs</td>
<td>observed number of successes</td>
</tr>
</tbody>
</table>

**Author(s)**

Jim Albert

**Examples**

```r
prior = c(3, 10)  # proportion has a beta(3, 10) prior
n = 20            # sample size
yobs = 10         # observed number of successes
predplot(prior, n, yobs)
```

---

**prior.two.parameters**

*Construct discrete uniform prior for two parameters*

**Description**

Constructs a discrete uniform prior distribution for two parameters.

**Usage**

```r
prior.two.parameters(parameter1, parameter2)
```

**Arguments**

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>parameter1</td>
<td>vector of values of first parameter</td>
</tr>
<tr>
<td>parameter2</td>
<td>vector of values of second parameter</td>
</tr>
</tbody>
</table>
Value

matrix of uniform probabilities where the rows and columns are labelled with the parameter values

Author(s)

Jim Albert

Examples

prior.two.parameters(c(1,2,3,4),c(2,4,7))

puffin

Bird measurements from British islands

Description

Measurements on breedings of the common puffin on different habits at Great Island, Newfoundland.

Usage

puffin

Format

A data frame with 38 observations on the following 5 variables.

Nest  nesting frequency (burrows per 9 square meters)
Grass  grass cover (percentage)
Soil  mean soil depth (in centimeters)
Angle  angle of slope (in degrees)
Distance  distance from cliff edge (in meters)

Source

**rdirichlet**

*Random draws from a Dirichlet distribution*

**Description**

Simulates a sample from a Dirichlet distribution

**Usage**

```
rdirichlet(n, par)
```

**Arguments**

- `n`: number of simulations required
- `par`: vector of parameters of the Dirichlet distribution

**Value**

matrix of simulated draws where each row corresponds to a single draw

**Author(s)**

Jim Albert

**Examples**

```r
c <- c(2, 4, 10)
n <- 10
rdirichlet(n, c)
```

---

**reg.gprior.post**

*Computes the log posterior of a normal regression model with a g prior.*

**Description**

Computes the log posterior of (beta, log sigma) for a normal regression model with a g prior with parameters beta0 and c0.

**Usage**

```
reg.gprior.post(theta, dataprior)
```
Arguments

theta vector of components of beta and log sigma
dataprior list with components data and prior; data is a list with components y and X, prior is a list with components b0 and c0

Value

value of the log posterior

Author(s)

Jim Albert

Examples

data(puffin)
data=list(y=puffin$Nest, X=cbind(1, puffin$Distance))
prior=list(b0=c(0,0), c0=10)
reg gprior.post(c(20,-.5,1), list(data=data, prior=prior))

regroup

Collapses a matrix by summing over rows

Description

Collapses a matrix by summing over a specific number of rows

Usage

regroup(data, g)

Arguments

data a matrix
g a positive integer between 1 and the number of rows of data

Value

reduced matrix found by summing over rows

Author(s)

Jim Albert

Examples

data=matrix(c(1:20), nrow=4, ncol=5)
g=2
regroup(data, 2)
rejectsampling

Rejecting sampling using a t proposal density

Description

Implements a rejection sampling algorithm for a probability density using a multivariate t proposal density

Usage

rejectsampling(logf, tpar, dmax, n, data)

Arguments

- **logf**: function that defines the logarithm of the density of interest
- **tpar**: list of parameters of t proposal density including the mean m, scale matrix var, and degrees of freedom df
- **dmax**: logarithm of the rejection sampling constant
- **n**: number of simulated draws from proposal density
- **data**: data and or parameters used in the function logf

Value

matrix of simulated draws from density of interest

Author(s)

Jim Albert

Examples

data(cancermortality)
start<-c(-7,6)
fit=laplace(betabinexch, start, cancermortality)
tpar=list(m=fit$mode, var=2*fit$var, df=4)
theta=rejectsampling(betabinexch, tpar, -569.2813, 1000, cancermortality)
**rigamma**  
*Random number generation for inverse gamma distribution*

**Description**

Simulates from a inverse gamma (a, b) distribution with density proportional to $y^{(-a-1)} \exp(-b/y)$

**Usage**

`rigamma(n, a, b)`

**Arguments**

- `n`: number of random numbers to be generated
- `a`: inverse gamma shape parameter
- `b`: inverse gamma rate parameter

**Value**

vector of n simulated draws

**Author(s)**

Jim Albert

**Examples**

```r
a = 10
b = 5
n = 20
rigamma(n, a, b)
```

---

**rmnorm**  
*Random number generation for multivariate normal*

**Description**

Simulates from a multivariate normal distribution

**Usage**

`rmnorm(n = 1, mean = rep(0, d), varcov)`
Arguments

- **n**: number of random numbers to be generated
- **mean**: numeric vector giving the mean of the distribution
- **varcov**: a positive definite matrix representing the variance-covariance matrix of the distribution

Value

matrix of n rows of random vectors

Author(s)

Jim Albert

Examples

```r
mu <- c(1,12,2)
Sigma <- matrix(c(1,2,0,2,5,0.5,0,0,5,3), 3, 3)
x <- rmnorm(10, mu, Sigma)
```

Description

Simulates from a multivariate t distribution

Usage

```r
rmt(n = 1, mean = rep(0, d), S, df = Inf)
```
Examples

mu <- c(1,12,2)
Sigma <- matrix(c(1.2,0,2,5,0.5,0,0.5,3), 3, 3)
df <- 4
x <- rmt(10, mu, Sigma, df)

robustt

Gibbs sampling for a robust regression model

Description

Implements Gibbs sampling for a robust t sampling model with location mu, scale sigma, and degrees of freedom v.

Usage

robustt(y, v, m)

Arguments

y
vector of data values

v
degrees of freedom for t model

m
the number of cycles of the Gibbs sampler

Value

mu
vector of simulated values of mu

s2
vector of simulated values of sigma2

lam
matrix of simulated draws of lambda, where each row corresponds to a single draw

Author(s)

Jim Albert

Examples

data=c(-67,-48,6,8,14,16,23,24,28,29,41,49,67,60,75)
fit=robustt(data,4,1000)
**rtruncated**

Simulates a truncated probability distribution

---

**Description**

Simulates a sample from a truncated distribution where the functions for the cdf and inverse cdf are available.

**Usage**

```
rtruncated(n, lo, hi, pf, qf, ...)```

**Arguments**

- `n`  
  size of simulated sample
- `lo`  
  low truncation point
- `hi`  
  high truncation point
- `pf`  
  function containing cdf of untruncated distribution
- `qf`  
  function containing inverse cdf of untruncated distribution
- `...`  
  parameters used in the functions `pf` and `qf`

**Value**

vector of simulated draws from distribution

**Author(s)**

Jim Albert

**Examples**

```
# want a sample of 10 from normal(2, 1) distribution truncated below by 3
n=10
lo=3
hi=Inf
rtruncated(n, lo, hi, pnorm, qnorm, mean=2, sd=1)
# want a sample of 20 from beta(2, 5) distribution truncated to (.3, .8)
n=20
lo=0.3
hi=0.8
rtruncated(n, lo, hi, pbeta, qbeta, 2, 5)
```
**Random walk Metropolis algorithm of a posterior distribution**

**Description**

Simulates iterates of a random walk Metropolis chain for an arbitrary real-valued posterior density defined by the user.

**Usage**

```r
rwmetrop(logpost, proposal, start, m, ...)
```

**Arguments**

- `logpost`: function defining the log posterior density
- `proposal`: a list containing `var`, an estimated variance-covariance matrix, and `scale`, the Metropolis scale factor
- `start`: vector containing the starting value of the parameter
- `m`: the number of iterations of the chain
- `...`: data that is used in the function `logpost`

**Value**

- `par`: a matrix of simulated values where each row corresponds to a value of the vector parameter
- `accept`: the acceptance rate of the algorithm

**Author(s)**

Jim Albert

**Examples**

```r
data=c(6,2,3,10)
varcov=diag(c(1,1))
proposal=list(var=varcov, scale=2)
start=array(c(1,1),c(1,2))
m=1000
s=rwmetrop(logctablepost, proposal, start, m, data)
```
Batting data for Mike Schmidt

Description
Batting statistics for the baseball player Mike Schmidt during all the seasons of his career.

Usage
schmidt

Format
A data frame with 18 observations on the following 14 variables.

Year year of the season
Age Schmidt’s age that season
G games played
AB at-bats
R runs scored
H number of hits
X2B number of doubles
X3B number of triples
HR number of home runs
RBI number of runs batted in
SB number of stolen bases
CS number of times caught stealing
BB number of walks
SO number of strikeouts

Source
simcontour  

*Simulated draws from a bivariate density function on a grid*

**Description**

For a general two parameter density defined on a grid, simulates a random sample.

**Usage**

```r
simcontour(logf, limits, data, m)
```

**Arguments**

- `logf` function that defines the logarithm of the density
- `limits` limits (xlo, xhi, ylo, yhi) that cover the joint probability density
- `data` vector or list of parameters associated with the function logpost
- `m` size of simulated sample

**Value**

- `x` vector of simulated draws of the first parameter
- `y` vector of simulated draws of the second parameter

**Author(s)**

Jim Albert

**Examples**

```r
m = array(c(0,0),c(2,1))
v = array(c(1,6,6,1),c(2,2))
normpar = list(m=m, v=v)
s = simcontour(lbinorm, c(-4,4,-4,4), normpar, 1000)
plot(s$x, s$y)
```

sir  

*Sampling importance resampling*

**Description**

Implements sampling importance resampling for a multivariate t proposal density.

**Usage**

```r
sir(logf, tpar, n, data)
```
Arguments

- \texttt{logf} function defining logarithm of density of interest
- \texttt{tpar} list of parameters of multivariate t proposal density including the mean \( m \), the scale matrix \( \text{var} \), and the degrees of freedom \( \text{df} \)
- \( n \) number of simulated draws from the posterior
- \texttt{data} data and parameters used in the function \texttt{logf}

Value

matrix of simulated draws from the posterior where each row corresponds to a single draw

Author(s)

Jim Albert

Examples

```r
data(cancermortality)
start=c(-7,6)
fit=laplace(betabinexch,start,cancermortality)
tpar=list(m=fit$mode,var=2*fit$var,df=4)
theta=sir(betabinexch,tpar,1000,cancermortality)
```

sluggerdata  

\textit{Hitting statistics for ten great baseball players}

Description

Career hitting statistics for ten great baseball players

Usage

\texttt{sluggerdata}

Format

A data frame with 199 observations on the following 13 variables.

- **Player** names of the ballplayer
- **Year** season played
- **Age** age of the player during the season
- **G** games played
- **AB** number of at-bats
- **R** number of runs scored
- **H** number of hits
socccergoals

**X2B**  number of doubles
**X3B**  number of triples
**HR**   number of home runs
**RBI**  runs batted in
**BB**   number of base on balls
**SO**   number of strikeouts

**Source**


---

**socccergoals**  
*Goals scored by professional soccer team*

**Description**

Number of goals scored by a single professional soccer team during the 2006 Major League Soccer season

**Usage**

socccergoals

**Format**

A data frame with 35 observations on the following 1 variable.

**goals**  number of goals scored

**Source**

Collected by author from the www.espn.com website.
Data from Stanford Heart Transplantation Program

Description
Heart transplant data for 82 patients from Stanford Heart Transplantation Program

Usage
stanfordheart

Format
A data frame with 82 observations on the following 4 variables.
- survtime  survival time in months
- transplant variable that is 1 or 0 if patient had transplant or not
- timetotransplant  time a transplant patient waits for operation
- state  variable that is 1 or 0 if time is censored or not

Source

Baseball strikeout data

Description
For all professional baseball players in the 2004 season, dataset gives the number of strikeouts and at-bats when runners are in scoring position and when runners are not in scoring position.

Usage
strikeout

Format
A data frame with 438 observations on the following 4 variables.
- r  number of strikeouts of player when runners are not in scoring position
- n  number of at-bats of player when runners are not in scoring position
- s  number of strikeouts of player when runners are in scoring position
- m  number of at-bats of player when runners are in scoring position
Source

Collected from www.espn.com website.

studentdata  Student dataset

Description

Answers to a sheet of questions given to a large number of students in introductory statistics classes

Usage

studentdata

Format

A data frame with 657 observations on the following 11 variables.

Student  student number
Height  height in inches
Gender  gender
Shoes  number of pairs of shoes owned
Number  number chosen between 1 and 10
Dvds  name of movie dvds owned
ToSleep  time the person went to sleep the previous night (hours past midnight)
WakeUp  time the person woke up the next morning
Haircut  cost of last haircut including tip
Job  number of hours working on a job per week
Drink  usual drink at suppertime among milk, water, and pop

Source

Collected by the author during the Fall 2006 semester.
**transplantpost**

*Log posterior of a Pareto model for survival data*

### Description

Computes the log posterior density of (log tau, log lambda, log p) for a Pareto model for survival data.

### Usage

```r
transplantpost(theta, data)
```

### Arguments

- `theta`: vector of parameter values of log tau, log lambda, and log p
- `data`: data matrix with columns survival time, transplant indicator, time to transplant, and censoring indicator

### Value

value of the log posterior

### Author(s)

Jim Albert

### Examples

```r
data(stanfordheart)
theta<-c(0,3,-1)
transplantpost(theta, stanfordheart)
```

---

**triplot**

*Plot of prior, likelihood and posterior for a proportion*

### Description

For a proportion problem with a beta prior, plots the prior, likelihood and posterior on one graph.

### Usage

```r
triplot(prior, data, where="topright")
```
Arguments

prior vector of parameters for beta prior
data vector consisting of number of successes and number of failures
where the location of the legend for the plot

Author(s)

Jim Albert

Examples

prior=c(3,10) # proportion has a beta(3, 10) prior
data=c(10,6) # observe 10 successes and 6 failures
triplot(prior,data)

Description

Computes the log posterior density of (log sigma, mu, beta) for a Weibull proportional odds regression model

Usage

weibullregpost(theta,data)

Arguments

theta vector of parameter values log sigma, mu, and beta
data data matrix with columns survival time, censoring variable, and covariate matrix

Value

value of the log posterior

Author(s)

Jim Albert

Examples

data(chemotherapy)
attach(chemotherapy)
d=cbind(time,status,treat-1,age)
theta=c(-.6,.1,.6,0)
weibullregpost(theta,d)
Index

*Topic datasets
  achievement, 3
  baseball.1964, 4
  bermuda.grass, 8
  bird.extinct, 13
  birthweight, 14
  breastcancer, 18
  calculus.grades, 19
  cancermortality, 20
  chemotherapy, 22
  darwin, 23
  donner, 27
  election, 28
  election.2008, 28
  footballscores, 29
  hearttransplants, 31
  iowagpa, 36
  jeter2004, 36
  marathontimes, 42
  puffin, 56
  schmidt, 65
  sluggerdata, 67
  soccergoals, 68
  stanfordheart, 69
  strikeout, 69
  studentdata, 70

*Topic models
  bayes.influence, 5
  bayes.model.selection, 5
  bayes.probit, 6
  bayesresiduals, 7
  beta.select, 8
  betabinexch, 9
  betabinexch0, 10
  bfexch, 11
  bfindep, 11
  binomial.beta.mix, 12
  blinreg, 14
  blinregpred, 16
  bprobit.probs, 17
  bradley.terry.post, 18
  careertraj.setup, 20
  cauchyerrorpost, 21
  ctable, 22
  discint, 23
  discrete.bayes, 24
  discrete.bayes.2, 25
  dmnorm, 26
  dmt, 26
  gibbs, 29
  groupeddatapost, 30
  hiergibbs, 32
  histprior, 32
  howardprior, 33
  impsampling, 34
  indepmetrop, 35
  laplace, 37
  lbinorm, 38
  logctablepost, 39
  logisticpost, 39
  logpoissgamma, 40
  logpoissnormal, 41
  mnormt.onesided, 42
  mnormt.twosided, 43
  mycontour, 44
  normal.normal.mix, 45
  normal.select, 46
  normchi2post, 46
  normnormexch, 47
  normpostpred, 48
  normpostsim, 49
  ordergibbs, 49
  pbetap, 50
  pbetat, 51
  pdisc, 52
  pdiscp, 52
  poissgamexch, 53
INDEX

plot.bayes (discrete.bayes), 24
plot.bayes2 (discrete.bayes.2), 25
poissgamexch, 53
poisson.gamma.mix, 54
predplot, 55
print.bayes (discrete.bayes), 24
prior.two.parameters, 55
puffin, 56
rdirichlet, 57
reg.gprior.post, 57
regroup, 58
rejectsampling, 59
rigamma, 60
rmnrm, 60
rmt, 61
robust, 62
rtruncated, 63
rwmnetrop, 64
schmidt, 65
simcontour, 66
sir, 66
sluggerdata, 67
soccergoals, 68
stanfordheart, 69
strikeout, 69
studentdata, 70
summary.bayes (discrete.bayes), 24

transplantpost, 71
triplot, 71
weibullregpost, 72