

# Package ‘LearnBayes’

April 17, 2009

**Type** Package

**Title** Functions for Learning Bayesian Inference

**Version** 2.0

**Date** 2008-11-08

**Author** Jim Albert

**Maintainer** Jim Albert <albert@bgnet.bgsu.edu>

**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)

**Repository** CRAN

**Date/Publication** 2009-01-05 10:34:28

## 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
bfndep . . . . .	11

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
dmnorm . . . . .	24
dmt . . . . .	25
donner . . . . .	26
election . . . . .	26
election.2008 . . . . .	27
footballscores . . . . .	27
gibbs . . . . .	28
groupeddatapost . . . . .	29
hearttransplants . . . . .	29
hiergibbs . . . . .	30
histprior . . . . .	31
howardprior . . . . .	31
impsampling . . . . .	32
indepmetrop . . . . .	33
iowagpa . . . . .	34
jeter2004 . . . . .	34
laplace . . . . .	35
lbinorm . . . . .	36
logtablepost . . . . .	37
logisticpost . . . . .	37
logpoissgamma . . . . .	38
logpoissnormal . . . . .	39
marathontimes . . . . .	40
mnormt.onesided . . . . .	40
mnormt.twosided . . . . .	41
mycontour . . . . .	42
normal.normal.mix . . . . .	43
normal.select . . . . .	44
normchi2post . . . . .	44
normpostpred . . . . .	45
normpostsim . . . . .	46

ordergibbs . . . . .	47
pbetap . . . . .	47
pbetat . . . . .	48
pdisc . . . . .	49
pdiscp . . . . .	49
poissgamexch . . . . .	50
poisson.gamma.mix . . . . .	51
predplot . . . . .	52
puffin . . . . .	52
rdirichlet . . . . .	53
reg.gprior.post . . . . .	54
regroup . . . . .	54
rejectsampling . . . . .	55
rigamma . . . . .	56
rmnorm . . . . .	57
rmt . . . . .	57
robustt . . . . .	58
rtruncated . . . . .	59
rwmetrop . . . . .	60
schmidt . . . . .	61
simcontour . . . . .	62
sir . . . . .	62
sluggerdata . . . . .	63
soccergoals . . . . .	64
stanfordheart . . . . .	65
strikeout . . . . .	65
studentdata . . . . .	66
transplantpost . . . . .	67
triplot . . . . .	67
weibullregpost . . . . .	68
<b>Index</b>	<b>69</b>

---

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**

Abraham, B., and Ledolter, J. (2006), Introduction to Regression Modeling, Duxbury.

---

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**

www.baseball-reference.com website.

---

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 $i$ th row contains the 5th, 50th, 95th percentiles of log K for posterior when the $i$ th observation is removed

**Author(s)**

Jim Albert

**Examples**

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

---

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)
```

**Arguments**

y	vector of response values
X	matrix of covariates
c	parameter of the G prior
constant	logical variable indicating if a constant term is in the matrix X

**Value**

mod.prob	data frame specifying the model, the value of the log marginal density and the value of the posterior model probability
converge	logical vector indicating if the laplace algorithm converged for each model

**Author(s)**

Jim Albert

**Examples**

```
data(birdextinct)
logtime=log(birdextinct$time)
X=cbind(1,birdextinct$nesting,birdextinct$size,birdextinct$status)
bayes.model.selection(logtime,X,100)
```

---

bayes.probit	<i>Simulates from a probit binary response regression model using data augmentation and Gibbs sampling</i>
--------------	--

---

**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**

```
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

**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**

```
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)
```

---

bayesresiduals	<i>Computation of posterior residual outlying probabilities for a linear regression model</i>
----------------	---

---

**Description**

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

**Usage**

```
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)
lmfit=lm(temp~X)
m=1000
post=blinreg(temp,X,m)
k=2
bayesresiduals(lmfit,post,k)
```

---

bermuda.grass	<i>Bermuda grass experiment data</i>
---------------	--------------------------------------

---

**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**

McCullagh, P., and Nelder, J. (1989), *Generalized Linear Models*, Chapman and Hall.

---

beta.select	<i>Selection of Beta Prior Given Knowledge of Two Quantiles</i>
-------------	---

---

**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	<i>Log posterior of logit mean and log precision for Binomial/beta exchangeable model</i>
-------------	---

---

**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**

```
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)
```

---

betabinexch0	<i>Log posterior of mean and precision for Binomial/beta exchangeable model</i>
--------------	---

---

**Description**

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

**Usage**

```
betabinexch0(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**

```
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	<i>Logarithm of integral of Bayes factor for testing homogeneity of proportions</i>
--------	---

---

**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**

```
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	<i>Bayes factor against independence assuming alternatives close to independence</i>
---------	--

---

**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**

```
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 prior*

---

**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**

```
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 breeding 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**

Pimm, S., Jones, H., and Diamond, J. (1988), On the risk of extinction, *American Naturalists*, 132, 757-785.

---

birthweight	<i>Birthweight regression study</i>
-------------	-------------------------------------

---

**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**

Dobson, A. (2001), An Introduction to Generalized Linear Models, New York: Chapman and Hall.

---

blinreg	<i>Simulation from Bayesian linear regression model</i>
---------	---

---

**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

**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**

```
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
s=blinreg(temp,X,m)
```

---

blinregexpected      *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**

```
blinregexpected(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**

```

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)
blinregexpected(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**

```
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**

```

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)

```

---

bprobit.probs	<i>Simulates fitted probabilities for a probit regression model</i>
---------------	---

---

### 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(0,1,0,0,0,1,1,1,1,1)
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**

<code>theta</code>	vector of talent parameters and log standard deviation
<code>data</code>	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
```

**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**

Collett, D. (1994), *Modelling Survival Data in Medical Research*, London: Chapman and Hall.

---

`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      *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.

**y** number of cancer deaths

**n** number at risk

**Source**

Tsutakawa, R., Shoop, G., and Marienfeld, C. (1985), Empirical Bayes Estimation of Cancer Mortality Rates, *Statistics in Medicine*, 4, 201-212.

---

careertraj.setup      *Setup for Career Trajectory Application*

---

**Description**

Setups 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**

**player.names**      vector of player names

**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

**n**                      matrix of AB-SO for all players

**x**                      matrix of ages for all players for all years of their careers

**T**                      vector of number of seasons for all players

**N**                      number of players

**Author(s)**

Jim Albert

**Examples**

```
data(sluggerdata)
careertraj.setup(sluggerdata)
```

---

cauchyerrorpost	<i>Log posterior of median and log scale parameters for Cauchy sampling</i>
-----------------	---

---

**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	<i>Chemotherapy treatment effects on ovarian cancer</i>
--------------	---

---

**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**

Edmonson, J., Felming, T., Decker, D., Malkasian, G., Jorgensen, E., Jefferies, J., Webb, M., and Kvols, L. (1979), Different chemotherapeutic sensitivities and host factors affecting prognosis in advanced ovarian carcinoma versus minimal residual disease, *Cancer Treatment Reports*, 63, 241-247.

---

ctable	<i>Bayes factor against independence using uniform priors</i>
--------	---

---

**Description**

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

**Usage**

```
ctable(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**

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

---

darwin

*Darwin's data on plants*

---

**Description**

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

**Usage**

darwin

**Format**

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

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

**Source**

Fisher, R. (1960), Statistical Methods for Research Workers, Edinburgh: Oliver and Boyd.

---

discint

*Highest probability interval for a discrete distribution*

---

**Description**

Computes a highest probability interval for a discrete probability distribution

**Usage**

```
discint(dist, prob)
```

**Arguments**

dist	probability distribution written as a matrix where the first column contain the values and the second column the probabilities
prob	probability content of interest

**Value**

prob	exact probability content of interval
set	set of values of the probability interval

**Author(s)**

Jim Albert

**Examples**

```
x=0:10
probs=dbinom(x, size=10, prob=.3)
dist=cbind(x, probs)
pcontent=.8
discint(dist, pcontent)
```

---

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 be 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**

```
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)
```

---

 donner

*Donner survival study*


---

**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**

Grayson, D. (1960), Donner party deaths: a demographic assessment, *Journal of Anthropological Assessment*, 46, 223-242.

---

 election

*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

---

election.2008      *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**

Data collected by author in November 2008 from [www.cnn.com](http://www.cnn.com) website.

---

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

Gelman, A., Carlin, J., Stern, H., and Rubin, D. (2003), Bayesian Data Analysis, Chapman and Hall.

---

gibbs	<i>Metropolis within Gibbs sampling algorithm of a posterior distribution</i>
-------	---

---

### 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, data)
```

### Arguments

logpost	function defining the log posterior density
start	array with a single row that gives the starting value of the parameter vector
m	the number of iterations of the chain
scale	vector of scale parameters for the random walk Metropolis steps
data	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	vector of acceptance rates of the Metropolis steps of the algorithm

### Author(s)

Jim Albert

### Examples

```
data=c(6, 2, 3, 10)
start=array(c(1, 1), c(1, 2))
m=1000
scale=c(2, 2)
s=gibbs(logctablepost, start, m, scale, data)
```

---

groupeddatapost      *Log posterior of normal parameters when data is in grouped form*

---

**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

**Author(s)**

Jim Albert

**Examples**

```
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**

Christiansen, C. and Morris, C. (1995), Fitting and checking a two-level Poisson model: modeling patient mortality rates in heart transplant patients, in Berry, D. and Stangl, D., eds, Bayesian Biostatistics, Marcel Dekker.

---

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**

```
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**

```
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**

```
histprior(p, midpts, prob)
```

**Arguments**

<code>p</code>	vector of values for which density is to computed
<code>midpts</code>	vector of midpoints of the intervals
<code>prob</code>	vector of probabilities of the intervals

**Value**

vector of values of the probability density

**Author(s)**

Jim Albert

**Examples**

```
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")
```

---

`howardprior`*Logarithm of Howard's dependent prior for two proportions*

---

**Description**

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

**Usage**

```
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

**Examples**

```
param=c(1,1,1,1,2)
p=c(.1,.5)
howardprior(p,param)
```

---

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(betabinexch,start,cancermortality)
tpar=list(m=fit$mode,var=2*fit$var,df=4)
myfunc=function(theta) return(theta[2])
theta=impsampling(betabinexch,tpar,myfunc,1000,cancermortality)

```

---

indepmetrop	<i>Independence Metropolis independence chain of a posterior distribution</i>
-------------	---

---

**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,data)
```

**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	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**

Albert, J. (1994), A Bayesian approach to estimation of GPA's of University of Iowa freshmen under order restrictions, *Journal of Educational Statistics*, 19, 1-22.

---

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**

Collected from game log data from [www.retrosheet.org](http://www.retrosheet.org).

---

 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**

```
laplace(logpost, mode, par)
```

**Arguments**

logpost	function that defines the logarithm of the posterior density
mode	vector that is a guess at the posterior mode
par	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**

```
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**`lbinorm(xy,par)`**Arguments**

<code>xy</code>	vector of values of two variables x and y
<code>par</code>	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**

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

---

logctablepost	<i>Log posterior of difference and sum of logits in a 2x2 table</i>
---------------	---

---

**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	<i>Log posterior for a binary response model with a logistic link and a uniform prior</i>
--------------	---

---

**Description**

Computes the log posterior density of (beta0, beta1) when yi are independent binomial(ni, pi) and  $\text{logit}(\pi) = \beta_0 + \beta_1 \cdot x_i$  and a uniform prior is placed on (beta0, beta1)

**Usage**

```
logisticpost(beta, data)
```

**Arguments**

beta                vector of parameter values beta0 and beta1  
 data                matrix of columns of covariate values x, sample sizes n, and number of successes y

**Value**

value of the log posterior

**Author(s)**

Jim Albert

**Examples**

```
x = c(-0.86, -0.3, -0.05, 0.73)
n = c(5, 5, 5, 5)
y = c(0, 1, 3, 5)
data = cbind(x, n, y)
beta=c(2, 10)
logisticpost(beta, data)
```

---

logpoissgamma

*Log posterior with Poisson sampling and gamma prior*

---

**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

**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)
logpoissgamma(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	<i>Marathon running times</i>
---------------	-------------------------------

---

**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](http://www.nycmarathon.org) website.

---

mnormt.onesided	<i>Bayesian test of one-sided hypothesis about a normal mean</i>
-----------------	--

---

**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

**Author(s)**

Jim Albert

**Examples**

```
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**

```
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

**Examples**

```

m0=170
prob=.5
tau=c(.5,1,2,4,8)
samplesize=10
samplemean=176
popstd=3
data=c(samplemean,samplesize,popstd)
mnormt.twosided(m0,prob,tau,data)

```

---

mycontour

*Contour plot of a bivariate density function*


---

**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**

```

m=array(c(0,0),c(2,1))
v=array(c(1,.6,.6,1),c(2,2))
normpar=list(m=m,v=v)
mycontour(lbinorm,c(-4,4,-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

```
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

```
# 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)
```

---

normchi2post      *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.

**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)
```

---

normpostpred	<i>Posterior predictive simulation from Bayesian normal sampling model</i>
--------------	--

---

**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**

```
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**

```
# 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 prior.

**Usage**

```
normpostsim(data, m=1000)
```

**Arguments**

<code>data</code>	vector of observations
<code>m</code>	number of simulations desired

**Value**

<code>mu</code>	vector of simulated draws of normal mean
<code>sigma2</code>	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 discrete 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

**Examples**

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

---

pbetat

*Bayesian test of a proportion*

---

**Description**

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

**Usage**

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

**Arguments**

<code>p0</code>	value of the proportion to be tested
<code>prob</code>	prior probability of the hypothesis
<code>ab</code>	vector of parameter values of the beta prior under the alternative hypothesis
<code>data</code>	vector containing the number of successes and number of failures

**Value**

<code>bf</code>	the Bayes factor in support of the null hypothesis
<code>post</code>	the posterior probability of the null hypothesis

**Author(s)**

Jim Albert

**Examples**

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

---

pdisc

*Posterior distribution for a proportion with discrete priors*

---

**Description**

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

**Usage**

```
pdisc(p, prior, data)
```

**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)
```

**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**

```
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**

```
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 hyperparameter

**Value**

value of the log posterior

**Author(s)**

Jim Albert

**Examples**

```
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**

```
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**

```
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**

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

**Arguments**

<code>prior</code>	vector of parameters for beta prior
<code>n</code>	sample size
<code>yobs</code>	observed number of successes

**Author(s)**

Jim Albert

**Examples**

```
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)
```

---

`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**

Peck, R., Devore, J., and Olsen, C. (2005), Introduction to Statistics And Data Analysis, Thomson Learning.

---

`rdirichlet`*Random draws from a Dirichlet distribution*

---

**Description**

Simulates a sample from a Dirichlet distribution

**Usage**

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

**Arguments**

<code>n</code>	number of simulations required
<code>par</code>	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**

```
par=c(2, 5, 4, 10)
n=10
rdirichlet(n, par)
```

---

`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**

```
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**

<code>n</code>	number of random numbers to be generated
<code>mean</code>	numeric vector giving the mean of the distribution
<code>varcov</code>	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**

```
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)
```

---

`rmt`*Random number generation for multivariate t*

---

**Description**

Simulates from a multivariate t distribution

**Usage**

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

**Arguments**

n	number of random numbers to be generated
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

**Value**

matrix of n rows of random vectors

**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)
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 sigma <sup>2</sup>
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	<i>Simulates from a truncated probability distribution</i>
------------	--

---

**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)
```

---

`rwmetrop`*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**

```
rwmetrop(logpost, proposal, start, m, par)
```

**Arguments**

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

**Value**

<code>par</code>	a matrix of simulated values where each row corresponds to a value of the vector parameter
<code>accept</code>	the acceptance rate of the algorithm

**Author(s)**

Jim Albert

**Examples**

```
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)
```

---

`schmidt`*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**

Sean Lahman's baseball database from [www.baseball1.com](http://www.baseball1.com).

---

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

```
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

```
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

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

**Arguments**

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

**Value**

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

**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=sir(betabinexch,tpar,1000,cancermortality)
```

---

sluggerdata

*Hitting statistics for ten great baseball players*


---

**Description**

Career hitting statistics for ten great baseball players

**Usage**

```
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

**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

Sean Lahman's baseball database from [www.baseball1.com](http://www.baseball1.com).

---

soccergoals

*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

```
soccergoals
```

### 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](http://www.espn.com) website.

---

`stanfordheart`*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**

Turnbull, B., Brown, B. and Hu, M. (1974), Survivorship analysis of heart transplant data, Journal of the American Statistical Association, 69, 74-80.

---

`strikeout`*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 supertime 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**

```
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**

```
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**

```
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
tripplot(prior,data)
```

---

weibullregpost            *Log posterior of a Weibull proportional odds model for survival 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(-.6,11,.6,0)
weibullregpost(theta,d)
```

# Index

## \*Topic **datasets**

- achievement, 1
- baseball.1964, 2
- bermuda.grass, 6
- birdextinct, 11
- birthweight, 12
- breastcancer, 16
- calculus.grades, 17
- cancermortality, 18
- chemotherapy, 20
- darwin, 21
- donner, 24
- election, 24
- election.2008, 25
- footballscores, 25
- heartttransplants, 27
- iowagpa, 32
- jeter2004, 32
- marathontimes, 38
- puffin, 50
- schmidt, 59
- sluggerdata, 61
- soccergoals, 62
- stanfordheart, 63
- strikeout, 63
- studentdata, 64

## \*Topic **models**

- bayes.influence, 3
- bayes.model.selection, 3
- bayes.probit, 4
- bayesresiduals, 5
- beta.select, 6
- betabinexch, 7
- betabinexch0, 8
- bfexch, 9
- bfindep, 9
- binomial.beta.mix, 10
- blinreg, 12
- blinregexpected, 13

- blinregpred, 14
- bprobit.probs, 15
- bradley.terry.post, 16
- careertraj.setup, 18
- cauchyerrorpost, 19
- ctable, 20
- discint, 21
- dmnorm, 22
- dmt, 23
- gibbs, 26
- groupeddatapost, 27
- hiergibbs, 28
- histprior, 29
- howardprior, 29
- impsampling, 30
- indepmetrop, 31
- laplace, 33
- lbinorm, 34
- logctablepost, 35
- logisticpost, 35
- logpoissgamma, 36
- logpoissnormal, 37
- mnormt.onesided, 38
- mnormt.twosided, 39
- mycontour, 40
- normal.normal.mix, 41
- normal.select, 42
- normchi2post, 42
- normpostpred, 43
- normpostsim, 44
- ordergibbs, 45
- pbetap, 45
- pbetat, 46
- pdisc, 47
- pdiscp, 47
- poissgamexch, 48
- poisson.gamma.mix, 49
- predplot, 50
- rdirichlet, 51

- reg.gprior.post, 52
  - regroup, 52
  - rejectsampling, 53
  - rigamma, 54
  - rmnorm, 55
  - rmt, 55
  - robustt, 56
  - rtruncated, 57
  - rwmetrop, 58
  - simcontour, 60
  - sir, 60
  - transplantpost, 65
  - triplot, 65
  - weibullregpost, 66
- achievement, 1
- baseball.1964, 2
- bayes.influence, 3
- bayes.model.selection, 3
- bayes.probit, 4
- bayesresiduals, 5
- bermuda.grass, 6
- beta.select, 6
- betabinexch, 7
- betabinexch0, 8
- bfexch, 9
- bfindep, 9
- binomial.beta.mix, 10
- birdextinct, 11
- birthweight, 12
- blinreg, 12
- blinregexpected, 13
- blinregpred, 14
- bprobit.probs, 15
- bradley.terry.post, 16
- breastcancer, 16
- calculus.grades, 17
- cancermortality, 18
- careertraj.setup, 18
- cauchyerrorpost, 19
- chemotherapy, 20
- ctable, 20
- darwin, 21
- discint, 21
- dmnorm, 22
- dmt, 23
- donner, 24
- election, 24
- election.2008, 25
- footballscores, 25
- gibbs, 26
- groupeddatapost, 27
- hearttransplants, 27
- hiergibbs, 28
- histprior, 29
- howardprior, 29
- impsampling, 30
- indepmetrop, 31
- iowagpa, 32
- jeter2004, 32
- laplace, 33
- lbinorm, 34
- logctablepost, 35
- logisticpost, 35
- logpoissgamma, 36
- logpoissnormal, 37
- marathontimes, 38
- mnormt.onesided, 38
- mnormt.twosided, 39
- mycontour, 40
- normal.normal.mix, 41
- normal.select, 42
- normchi2post, 42
- normpostpred, 43
- normpostsim, 44
- ordergibbs, 45
- pbetap, 45
- pbetat, 46
- pdisc, 47
- pdiscp, 47
- poissgamexch, 48
- poisson.gamma.mix, 49
- predplot, 50
- puffin, 50
- rdirichlet, 51

reg.gprior.post, [52](#)  
regroup, [52](#)  
rejectsampling, [53](#)  
rigamma, [54](#)  
rmnorm, [55](#)  
rmt, [55](#)  
robustt, [56](#)  
rtruncated, [57](#)  
rwmetrop, [58](#)

schmidt, [59](#)  
simcontour, [60](#)  
sir, [60](#)  
sluggerdata, [61](#)  
soccergoals, [62](#)  
stanfordheart, [63](#)  
strikeout, [63](#)  
studentdata, [64](#)

transplantpost, [65](#)  
triplot, [65](#)

weibullregpost, [66](#)