Package ‘BaM’

May 18, 2016

Title Functions and Datasets for Books by Jeff Gill
Version 1.0.1
Author Jonathan Homola <homola@wustl.edu>, Danielle Korman, Jacob Metz, Miguel Pereira <m.pereira@wustl.edu>, Mauricio Vela, and Jeff Gill <jgill@wustl.edu>
Maintainer Jeff Gill <jgill@wustl.edu>
Depends R (>= 3.0.1)
Imports coda, dlm, MASS, mice, nnet
License GPL (>= 2)
LazyData true
RoxygenNote 5.0.1
NeedsCompilation no
Repository CRAN
Date/Publication 2016-05-18 17:26:11

R topics documented:

actuarial .................................................. 3
adam.jags ................................................. 3
afghan.deaths .......................................... 4
africa ...................................................... 4
asap.data.list ......................................... 5
baldus .................................................... 6
bcp .......................................................... 8
biv.norm.post .......................................... 8
cabinet.duration ....................................... 10
child ...................................................... 10
china.wars .............................................. 11
coal.mining.disasters ................................. 12
contracep .................................................. 12
DA_cwp .................................................. 13
dmultinorm .................................................. 14
dp .................................................. 14
durations.hpd .................................................. 15
elicspend .................................................. 16
ethnic.immigration .................................................. 16
executions .................................................. 17
experts .................................................. 17
expo.gibbs .................................................. 18
expo.metrop .................................................. 18
fdr .................................................. 19
hanjack .................................................. 19
hit.run .................................................. 20
iq .................................................. 21
italy.parties .................................................. 22
lunatics .................................................. 22
marriage.rates .................................................. 22
metropolis .................................................. 23
militarydf .................................................. 23
nc.sub.dat .................................................. 24
norm.known.var .................................................. 25
normal.posterior.summary .................................................. 25
norr .................................................. 26
opic .................................................. 26
pbc.vote .................................................. 27
plot_walk_G .................................................. 27
plot_walk_MH .................................................. 28
recidivism .................................................. 28
retail.sales .................................................. 29
rmultinorm .................................................. 30
romney .................................................. 30
sir .................................................. 31
socatt .................................................. 33
strikes .................................................. 34
terrorism .................................................. 35
texas .................................................. 35
t_ci_table .................................................. 36
wars .................................................. 37
Description

actuarial claims data for three groups of insurance policyholders p. 449

Usage

data(actuarial)

Format

dataset with 5 observations of 7 variables

Details

The variables included in the dataset are:

- year
- payroll for groups 1, 2, and 3
- claims for groups 1, 2, and 3

Source


Description

data from Differences in the Validity of Self-Reported Drug Use Across Five Factors in Indianapolis, Fort Lauderdale, Phoenix, and Dallas, 1994 (ICPSR Study Number 2706, Rosay and Herz (2000), from the Arrestee Drug Abuse Monitoring (ADAM) Program/Drug Use Forecasting, ICPSR Study Number 2826. The original purpose of the study was to understand the accuracy of self-reported drug use, which is a difficult problem for obvious reasons.

The variables included in the dataset are:

- AGEGRP 1 for 1,700 cases 18 through 30 years old, 2 for 1,265 cases 31 years old or over
- CASES
- CATS
- COCSELF indicating self-reported cocaine usage prior to arrest (0 for 2,220 negative responses, 1 for 745 positive responses)
• COCTEST
• COVARS
• GROUP
• ID
• MJSELF
• MJTEST a dichotomous variable indicating a positive urine test for marijuana
• OFFENSE
• RACE 1 for 1,554 black cases, 2 for 1,411 white cases
• SEX 1 for 2,213 male cases, 2 for 752 female cases
• SITE coded according to: Indianapolis = 1 (759 cases), Ft. Lauderdale = 2 (974 cases), Phoenix = 3 (646 cases), and Dallas = 4 (586 cases)

Usage

data(adam.jags)

afghan.deaths      afghan.deaths

Description

NATO Fatalities in Afghanistan, 10/01 to 1/07. see page 350

Usage

data(afghan.deaths)

Format

52 monthly periods, listed by rows

africa      africa

Description

African Coups Data, pp.562-564

Usage

data(africa)
**Format**

data frame with 33 observations of different African countries’ military coups with 7 explanatory variables

**Details**

The variables included in the dataset are:

- MILTCOUP Military Coups
- MILITARY Military Oligarchy
- POLLIB Political Liberalization: 0 for no observable civil rights for political expression, 1 for limited, and 2 for extensive
- PARTY93 number of legally registered political parties
- PCTVOTE Percent Legislative Voting
- PCTTURN Percent registered voting
- SIZE in one thousand square kilometer units
- POP Population in millions
- NUMREGIM Regime
- NUMELEC Election

**Source**


---

**Description**

The American State Administrator’s Project (ASAP) survey asks administrators about the influence of a variety of external political actors including "clientele groups" in their agencies., see page 395.

The variables included in the dataset are:

- **contracting** scale from 0 : 6 where higher indicates more private contracting within the respondent’s agency.
- **gov.influence** respondents’ assessment of the governor’s influence on contracting in their agency.
- **leg.influence** respondents’ assessment of the legislatures’ influence on contracting in their agency, ranging from 0 : 21.
- **elect.board** dichotomous variable coded 1 if appointed by a board, a commission or elected, and 0 otherwise.
- **years.tenure** number of years that the respondent has worked at their current agency.
• education ordinal variable for level of education possessed by the respondent.
• partisan.ID a 5-point ordinal variable (1-5) for the respondent’s partisanship (strong Democrat to strong Republican).
• category categories of agency type.
• medt.time whether the respondent spent more or less than the sample median with representatives of interest groups.
• medt.contr interaction variable between medt.time and contracting.
• gov.ideology state government ideology from Berry et al. (1998) from 0 to 100.
• nonprofits provides the total number of nonprofit groups in the respondents’ state in the year 2008, divided by 10,000.

Usage

data(asap.data.list)

baldus Baldus Dataset

Description

Data from Baldus Study on death sentences in Georgia (Exercise 14.2, p. 521). To use the data in JAGS or WinBugs, see baldus.jags and baldus.winbugs, respectively.

Usage

data(baldus)

Details

The variables included in the dataset are:
• race Defendant’s race (1 = Black)
• educatn Educational level
• employm Employment status (1 = Employed)
• SES Socioeconomic status (1 = Low Wage)
• married Marital status (1 = Married)
• num.chld Number of children
• military Military experience (1 = Serving, 0 = No military service, -1 = Dishonorable Discharge)
• pr.arrst
• plea Plea to Murder Indictment
• sentence Sentenced
• defense Status of Principle Defense Council (1 = Retained, 2 = Appointed)
• dp.sgght Prosecutor Waive/Fail to Seek DP (1 = Failed/Unknown, 2 = Sought DP)
• jdge.dec Judge Took Sentence from Jury?
• pen.phse Was there a penalty trial?
• did.appl Did defendant appeal cov. or sentence?
• out.appl Outcome of appeal
• vict.sex Victim sex
• pr.incr
• vict.age Victim’s age
• vict.rel Relation of victim with defendant
• vict.st1 Victim status (0 = Non-police+judicial, 1 = Polic+judicial)
• vict.st2
• specialA Special Circumstances ()
• methodA Method of killing
• num.kill Number of persons killed by defendant
• num.prps Number of persons killed by coperpetrator
• def.age Defendant’s age
• aggravat Aggravating circumstances
• bloody Bloody crime
• fam.lov
• insane Defendant invoked insanity defense
• mitcir
• num.depr
• rape Rape involved

Source

See Also
baldus.jags baldus.winbugs
**Description**

Implementation of bcp function, see pages 362-363 (2nd Edition).

**Usage**

```r
bcp(theta.matrix, y, a, b, g, d)
```

**Arguments**

- `theta.matrix`: theta.matrix
- `y`: Counts of Coal Mining Disasters
- `a`: Alpha Value in the lambda Prior
- `b`: Beta Value in the lambda Prior
- `g`: Gamma Value in the phi Prior
- `d`: Delta Value in the phi Prior

**Author(s)**

Jeff Gill

**Examples**

```r
## Not run:
bcp(theta.matrix,y,a,b,g,d)

## End(Not run)
```

---

**Description**

A function to calculate posterior quantities of the bivariate normal. See page 94.

**Usage**

```r
biv.norm.post(data.mat, alpha, beta, m, n0=5)
```
Arguments

data.mat    A matrix with two columns of normally distributed data
alpha      Wishart first (scalar) parameter
beta       Wishart second (matrix) parameter
m          prior mean for mu
n0         prior confidence parameter

Value

Returns

mu2        posterior mean, dimension 1
sig1       posterior mean, dimension 2
sig2       posterior variance, dimension 1
rho        posterior variance, dimension 2

Author(s)

Jeff Gill

Examples

```
rwishart <- function(df, p = nrow(SqrtSigma), SqrtSigma = diag(p)) {
  if((Ident <- missing(SqrtSigma)) && missing(p)) stop("either p or SqrtSigma must be specified")
  Z <- matrix(0, p, p)
  diag(Z) <- sqrt(rchisq(p, df:(df-p+1)))
  if(p > 1) {
    pseq <- 1:(p-1)
    Z[rep(p*pseq, pseq) + unlist(lapply(pseq, seq))] <- rnorm(p*(p-1)/2)
  }
  if(Ident) crossprod(Z)
  else crossprod(Z %*% SqrtSigma)
}

data.n10 <- rmultinorm(10, c(1,3), matrix(c(1,0.7,0.7,3.0),2,2))
rep.mat <- NULL; reps <- 1000
for (i in 1:reps){
  rep.mat <- rbind(rep.mat, biv.norm.post(data.n10, matrix(c(10,5,5,10),2,2),c(2,2)))
}
round(normal.posterior.summary(rep.mat),3)
```
Description

Cabinet duration (constitutional inter-election period) for eleven Western European countries from 1945 to 1980, page 65

Usage

cabinet.duration

Format

cabinet duration of 11 countries

Details

The variables included in the dataset are:

- N number of cabinets
- dur average length of duration

Note

Row names indicate country.

References


Description


Usage

child

Format

observations of 8 variables for 50 states
Details

The variables included in the dataset are:

- **SCCOLL** Change in Child Support collections
- **ACES** Chapters per Population
- **INSTABIL** Policy Instability
- **AAMBIG** Policy Ambiguity
- **CSTAFF** Change in Agency Staffing
- **ARD** State Divorce Rate
- **ASLACK** Organizational Slack
- **AEXPEND** State Level Expenditures

Source


---

### china.wars

**Description**

Modeling code for the example of ancient Chinese wars. See page 163-165

**Usage**

```
china.wars()
```

**Author(s)**

Jeff Gill

**Source**


Henry A. Murray Research Archive
Center for International Relations, Department of Political Science, University of Colorado, Boulder, USA
Description
A vector of British Coal Mining Disasters, see page 549-550

Usage
coal.mining.disasters

Format
vector of length 111

Source

Description
Contraception Data by country. See page 446

Usage
data(contracep)

Format
4 variables for 15 countries

Details
The variables included in the dataset are:

- Country Developing countries by size
- URC Rural Childhood
- WED Years of Education for the Woman
- FPE Exposure to Family Planning Efforts
- WED.FPE Interaction term specified by Wong and Mason
Source


Description

Data on ancient Chinese wars

Details

The variables included in the dataset are:

- X1
- CHLEG010
- LEGHUANG
- X . 2697
- X . 2697.1
- X2
- X1.1
- X1.2
- X0
- X0.1
- X2.1
- X3
- X2.2
- X3.1
- X2.3
**dmultinorm**

**Description**

dmultinorm function, see page 376.

**Usage**

dmultinorm(xval,yval,mu.vector,sigma.matrix)

**Arguments**

- **xval** Vector of X Random Variables
- **yval** Vector of Y Random Variables
- **mu.vector** Mean Vector
- **sigma.matrix** Matrix of Standard Deviations

**Author(s)**

Jeff Gill

---

**dp**

**Description**

Death Penalty Data, See Page 142.

**Usage**

data(dp)

**Format**

7 variables for 17 states
Details

The variables included in the dataset are:

- X State
- EXECUTIONS Number of capital punishments at state level in 1997
- INCOME Median per capita income in dollars
- PERPOVERTY Percent classified as living in poverty
- PERBLACK Percent of black citizens in population
- VC100k96 Rate of violent crime per 100,000 residents for 1996
- SOUTH Is the state in the South?
- PROPDegree Proportion of population with college degree

Source


Description

Simple HPD calculator from Chapter 2 (page 51, 2nd Edition).

Usage

durations.hpd(support, fn.eval, start, stop, target=0.90, tol=0.01)

Arguments

- support  x-axis values
- fn.eval   function values at x-axis points
- start     starting point in the vectors
- stop      stopping point in the vectors
- target    Desired X Level
- tol       Tolerance for round-off

Author(s)

Jeff Gill
Examples

```r
## Not run:
get("cabinet.duration")
ruler <- seq(0.45, 0.75, length=10000)
g.vals <- round(dgamma(ruler, shape=sum(cabinet.duration$N),
  rate=sum(cabinet.duration$N*cabinet.duration$dur))/2)
start.point <- 1000; stop.point <- length(g.vals)
durations.hpd(ruler, g.vals, start.point, stop.point)
```

# End(Not run)

elixspend

Description

Eliciting expected campaign spending data. Eight campaign experts are queried for quantiles at levels \( m = [0.1, 0.5, 0.9] \), and they provide the following values reflecting the national range of expected total intake by Senate candidates (in thousands). See page 120

Usage

```r
data(elixspend)
```

ethnic.immigration

Description

1990-1993 W.Europe Ethnic/Minority Populations. see page 280.

Usage

```r
data(ethnic.immigration)
```

Format

- total number of ethnic immigrants living in Western Europe from 22 countries

Details

The variables included in the dataset are:

- `Country.of.Origin` Country of origin of immigrants
- `Estimated.Total.K` Estimated total ethnic minority population in Western European Countries
- `Percent.of.Total` Percent of Total
Source


Description

Execution data.
The variables included in the dataset are:

- State
- executions Number of capital punishments at state level in 1997
- Median.Income Median per capita income in dollars
- Percent.Poverty Percent classified as living in poverty
- Percent.Black Percent of black citizens in population
- Violent.Crime Rate of violent crime per 100,000 residents for 1996

Usage

data(executions)

Format

explanatory variables for 17 states

Description

Fabricated data on campaign fundraising elicitations. See page 120

Usage

experts(q1,q2,q3)

Arguments

q1 the 0.1 quantile
q2 the 0.5 quantile
q3 the 0.9 quantile
**Description**

Simple Gibbs sampler demonstration on conditional exponentials from Chapter 1 (pages 25-27).

**Usage**

`expo.gibbs(B, k, m)`

**Arguments**

- `B` an upper bound
- `k` length of the subchains
- `m` number of iterations

**Author(s)**

Jeff Gill

---

**Description**

Simple Metropolis algorithm demonstration using a bivariate exponential target from Chapter 1 (pages 27-30).

**Usage**

`expo.metrop(m, x, y, L1, L2, L, B)`

**Arguments**

- `m` number of iterations
- `x` starting point for the x vector
- `y` starting point for the y vector
- `L1` event intensity for the x dimension
- `L2` event intensity for the y dimension
- `L` shared event intensity
- `B` upper bound
Author(s)
Jeff Gill

Examples

dep.m(5000, x=0.5, y=0.5, L1=0.5, L2=0.1, L=0.01, B=8)

data(fdr)

Description
FDR election data. See page 576

The variables included in the dataset are:

- State State name
- FDR Whether or not FDR won the state in 1932 election, 1 = won, 0 = lost
- PRE.DEP Mean income per state before the Great Depression (1929), in dollars
- POST.DEP Mean income per state after the Great Depression (1932), in dollars
- FARM Total farm wage and salary disbursements in thousands of dollars per state in 1932

Usage
data(fdr)

Description
1964 presidential election data. See page 221

Usage
depack(N,F,L,W,K,IND,DEM,WR,WD,SD)
hit.run

Arguments

- **N**: number of cases in the group
- **F**: Observed cell proportion voting for Johnson
- **L**: log-ratio of this proportion, see p. 246
- **W**: collects the inverse of the diagonal of the matrix for the group-weighting from $[N_iP_i(1-P_i)]$
- **K**: constant
- **IND**: indifference to the election
- **DEM**: stated preference for Democratic party issues
- **WR**: Weak Republican
- **WD**: Weak Democrat
- **SD**: Strong Democrat

References


Description

Implementation of hit.run algorithm, p. 361.

Usage

hit.run(theta.mat, reps, I.mat)

Arguments

- **theta.mat**: theta.mat
- **reps**: reps
- **I.mat**: I.mat

Author(s)

Jeff Gill
Examples

```r
## Not run:
# code to implement graph on p. 362, see page 376.
	num.sims <- 10000
Sig.mat <- matrix(c(1,0,0.95,0.95,1,0,2,2)
walks<-rbind(c(-3,-3),matrix(NA,nrow=(num.sims-1),ncol=2))
walks <- hit.run(walks,num.sims,Sig.mat)
z.grid <- outer(seq(-3,3,length=100),seq(-3,3,length=100),
    FUN=dmultinorm,c(0,0),Sig.mat)
contour(seq(-3,3,length=100),seq(-3,3,length=100),z.grid,
levels=c(0.05,0.1,0.2))
points(walks[5001:num.sims,],pch=".")
```

---

### iq

**iq data frame**

---

### Description

IQ data for 80 countries. See pages 85-87

### Usage

```r
data(iq)
```

### Source


### Examples

```r
## Not run:
{
    data(iq)
    n <- length(iq[,1])
t.iq <- (iq[,1]-mean(as.numeric(iq)))/(sd(iq[,1])/sqrt(n))
    r.t <- (rt(100000, n-1)*(sd(iq)/sqrt(n)) + mean(as.numeric(iq))
    quantile(r.t,c(0.01,0.10,0.25,0.5,0.75,0.90,0.99))
    r.sigma.sq <- 1/rgamma(100000,shape=(n-2)/2, rate=var(as.numeric(iq))*(n-1)/2)
    quantile(sqrt(r.sigma.sq), c(0.01,0.10,0.25,0.5,0.75,0.90,0.99))
}
## End(Not run)
```
### italy.parties

**Description**


**Usage**

```r
data(italy.parties)
```

### lunatics

**Description**

An 1854 study on mental health in the fourteen counties of Massachusetts yields data on 14 cases. This study was performed by Edward Jarvis (then president of the American Statistical Association).

The variables included in the dataset are:

- `nbr` the number of "lunatics" per county.
- `dist` distance to the nearest mental healthcare center
- `pop` population in the county by thousands
- `pden` population per square county mile
- `phome` the percent of "lunatics" cared for in the home

**Usage**

```r
data(lunatics)
```

### marriage.rates

**Description**

Italian Marriage Rates. See page 430

**Usage**

```r
data(marriage.rates)
```
**Format**

a vector containing 16 numbers

**Source**


---

**metropolis**

**Description**

Implementation of metropolis function, p. 359.

**Usage**

```r
metropolis(theta.matrix,reps,I.mat)
```

**Arguments**

- `theta.matrix`
- `reps`
- `I.mat`

**Author(s)**

Jeff Gill

---

**militarydf**

**Description**

A dataset of two variables. The proportional changes in military personnel for the named countries. See page 483-484

The variables included in the dataset are:

- **Year** The year selected to evaluate
- **Yugoslavia** The proportion change in the size of Yugoslavia's military
- **Albania** The proportion change in the size of Albania's military
- **Bulgaria** The proportion change in the size of Bulgaria's military
- **Czechoslovakia** The proportion change in the size of Czechoslovakia's military
• German Dem. Republic The proportion change in the size of the German Democratic Republic's military
• Hungary The proportion change in the size of Hungary's military
• Poland The proportion change in the size of Poland's military
• Romania The proportion change in the size of Romania's military
• USSR The proportion change in the size of the Soviet Union's military

Usage

data(militarydf)

Format

a data frame with 35 observations of years from 1949 to 1983 with 10 explanatory variables

Source


Description

North Carolina county level health data from the 2000 U.S. census and North Carolina public records, see page 78.

The variables included in the dataset are:

• Substantiated Abuse within family documented abuse for the county
• Percent Poverty percent within the county living in poverty, U.S. definition
• Total Population county population/1000

Usage

nc.sub.dat

Format

data frame with 100 observations of different counties in North Carolina with 3 explanatory variables

Source

data from 2000 US census and North Carolina Division of Public Health, Women's and Children's Health Section in Conjunction with State Center for Health Statistics
**norm.known.var**

---

**Description**

A function to calculate posterior quantities for a normal-normal model with known variance (pages 70-72). It produces the posterior mean, variance, and 95% credible interval for user-specified prior.

**Usage**

```
norm.known.var(data.vec, pop.var, prior.mean, prior.var)
```

**Arguments**

- `data.vec`: a vector of assumed normally distributed data
- `pop.var`: known population variance
- `prior.mean`: mean of specified prior distribution for mu
- `prior.var`: variance of specified prior distribution for mu

**Author(s)**

Jeff Gill

---

**normal.posterior.summary**

---

**Description**

A function to calculate posterior quantities of bivariate normals. See pages 74-80.

**Usage**

```
normal.posterior.summary(reps)
```

**Arguments**

- `reps`: a matrix where the columns are defined as in the output of `biv.norm.post`

**Author(s)**

Jeff Gill

**See Also**

`biv.norm.post`
Description

An 1854 study on mental health in the fourteen counties of Massachusetts yields data on 14 cases. This study was performed by Edward Jarvis (then president of the American Statistical Association).

The variables included in the dataset are:

- Current.policy Current sentencing policy
- Past.execution.rate Past execution rate
- Politicla.Culture Political culture
- Current.opinion Current opinion
- Citizen.ideology Citizen ideology
- Murder.Rate Murder rate
- Catholic Catholic
- Black Black
- Urban Urban
- Past.laws Past laws
- Past.opinion Past opinion

Usage

data(norr)

Description

Private capital investment data. See Page 390.

The variables included in the dataset are:

- Fund Name of the private company
- Age Years the company has been in existence
- Status Whether the company is investing or divesting
- Size Maximum fund size in millions

Usage

data(opic)
Description
Precinct level data for Palm Beach County, Florida from the 2000 U.S. Presidential Election, see page 149
The variables included in the dataset are:

- `badballots` Total number of spoiled ballots
- `technology` Voting Technology used, 0 for a datapunch machine or a butterfly ballot, 1 for votomatic
- `new` Number of "new" voters, as in those who have not voted in the precinct for previous 6 years
- `size` Total number of precinct voters
- `republican` The number of voters registered as Republican
- `white` The number of white nonminority voters in a given precinct

Usage
data(pbc.vote)

Format
data frame with 516 observations of each precinct in Palm Beach County with 11 explanatory variables

Source
Palm Beach Post collected data from state and federal sources about precinct level data in Palm Beach County for the 2000 US presidential election

Description
`plot_walk_G` code used to produce figure 10.2

Usage
`plot_walk_G(walk.mat,sim.rm,X=1,Y=2)`
Arguments

walk.mat  walk.mat
sim.rm  sim.rm
X  X
Y  Y

Author(s)

Jeff Gill

Description

plot_walk_MH code used to produce figure 10.4

Usage

plot_walk_MH(walk.mat)

Arguments

walk.mat  walk.mat

Author(s)

Jeff Gill

Description

Recidivism Rates. See page 188

The variables included in the dataset are:

- Crime.Type The type of crime committed
- Released The number of individuals released from a facility
- Returned The number of individuals returned to a facility
- Percentage (The number of individuals returned to a facility)/(The number of individuals released from a facility)
retail.sales

Usage

data(recidivism)

Format

data frame with 27 observations of different crime types with 5 explanatory variables

Source

state-level recidivism data as collected by the Oklahoma Department of Corrections from January 1, 1985 to June 30, 1999

retail.sales  retail.sales

Description

Retail sales from 1979 through 1989 based on data provided by the U.S. Department of Commerce through the Survey of Current Business, see page 439

The variables included in the dataset are:

• TIME the economic quarter specified, starting from the first quarter of 1979 where j=1 to the fourth quarter of 1989 where j=44
• DSB national income wage and salary disbursements (in billions of dollars)
• EMP employees on non-agricultural payrolls (in thousands)
• BDG building material dealer sales (in millions of dollars)
• CAR retail automotive dealer sales (in millions of dollars)
• FRN home furnishings dealer sales (in millions of dollars)
• GMR general merchandise dealer sales (in millions of dollars)

Usage

data(retail.sales)

Format

data frame with 44 observations of statistics for different economic quarters with 7 explanatory variables

Source

U.S. Department of Commerce data from first quarter of 1979 to fourth quarter of 1989
**Description**

A function to generate random multivariate Gaussians.

**Usage**

```r
rmultinorm(n, mu, vmat, tol = 1e-07)
```

**Arguments**

- `n`: number of observations
- `mu`: vector of mean
- `vmat`: variance-covariance matrix
- `tol`: tolerance

**Author(s)**

Jeff Gill

**See Also**

- `biv.norm.post`

---

**Description**

Analysis of cultural consensus data using binomial likelihood and beta prior.

**Usage**

```r
romney()
```

**Format**

See for yourself. Modify as desired.

**Author(s)**

Jeff Gill
Source


Description

Implementation of Rubin’s SIR, see pages 338-341 (2nd Edition)

Usage

sir(data.mat,theta.vector,theta.mat,M,m,tol=1e-06,ll.func,df=0)

Arguments

data.mat  A matrix with two columns of normally distributed data
theta.vector  The initial coefficient estimates
theta.mat  The initial vc matrix
M  The number of draws
m  The desired number of accepted values
tol  The rounding/truncing tolerance
ll.func  loglike function for empirical posterior
df  The df for using the t distribution as the approx distribution

Author(s)

Jeff Gill

Examples

## Not run:
sir <- function(data.mat,theta.vector,theta.mat,M,m,tol=1e-06,ll.func,df=0) {
  importance.ratio <- rep(NA,M)
  rand.draw <- rmultinorm(M,theta.vector,theta.mat,tol = 1e-04)
  if (df > 0)
    rand.draw <- rand.draw/(sqrt(rchisq(M,df)/df))
  empirical.draw.vector <- apply(rand.draw,1,ll.func,data.mat)
  if (sum(is.na(empirical.draw.vector)) == 0) {
    print("SIR: finished generating from posterior density function")
    print(summary(empirical.draw.vector))
  } else {
    print(paste("SIR: found",sum(is.na(empirical.draw.vector)),
               "NA(s) in generating from posterior density function, quitting"))
  }
}


```
return()
)
if (df == 0) {
  normal.draw.vector <- apply(rand.draw,,1,normal.posterior.ll,data.mat)
}
else {
  theta.mat <- ((df-2)/(df))*theta.mat
  normal.draw.vector <- apply(rand.draw,,1,t.posterior.ll,data.mat,df)
}
if (sum(is.na(normal.draw.vector)) == 0) {
  print("SIR: finished generating from approximation distribution")
  print(summary(normal.draw.vector))
}
else {
  print(paste("SIR: found",sum(is.na(normal.draw.vector)),
             "NA(s) in generating from approximation distribution, quitting"))
  return()
}
importance.ratio <- exp(empirical.draw.vector - normal.draw.vector)
importance.ratio[is.finite=F] <- 0
importance.ratio <- importance.ratio/max(importance.ratio)
if (sum(is.na(importance.ratio)) == 0) {
  print("SIR: finished calculating importance weights")
  print(summary(importance.ratio))
}
else {
  print(paste("SIR: found",sum(is.na(importance.ratio)),
             "NA(s) in calculating importance weights, quitting"))
  return()
}
accepted.mat <- rand.draw[1:2,]
while(nrow(accepted.mat) < m+2) {
  rand.unif <- runif(length(importance.ratio))
  accepted.loc <- seq(along=importance.ratio)[(rand.unif-tol) <= importance.ratio]
  rejected.loc <- seq(along=importance.ratio)[(rand.unif-tol) > importance.ratio]
  accepted.mat <- rbind(accepted.mat,rand.draw[accepted.loc,])
  rand.draw <- rand.draw[rejected.loc,]
  importance.ratio <- importance.ratio[rejected.loc]
  print(paste("SIR: cycle complete,",(nrow(accepted.mat)-2),"now accepted"))
}
accepted.mat[3:nrow(accepted.mat),]
}
# The following are log likelihood functions that can be plugged into the sir function above.

logit.posterior.ll <- function(theta.vector,X) {
  Y <- X[,1]
  X[,1] <- rep(1,nrow(X))
  sum( -log(1+exp(-X
       -log(1+exp(X)))))
}

normal.posterior.ll <- function(coef.vector,X) {
  dimnames(coef.vector) <- NULL
```
Description

Data from the British Social Attitudes (BSA) Survey 1983-1986.

The variables included in the dataset are:

- District identifying for geographic district.
- Respondent.Code respondent identifier
- Num.Answers number of positive answers to seven questions
- Party 1 = Conservative, 2 = Labour, 3 = Lib/SDP/Alliance, 4 = others
• Social Class 1 = middle, 2 = upper working, 3 = lower working
• Gender 1 = male, 2 = female.
• Age age in years 18-80
• Religion 1 = Roman Catholic, 2 = Protestant/Church of England, 3 = others, 4 = none.

Usage

data(socatt)

Description

French Coal Strikes, see page 212 and 213
The variables included in the dataset are:
• Year The year the labor strikes in France occurred
• Counts The number of labor strikes that occurred in France per year

Usage

data(strikes)

Format

data frame with 11 observations of strikes that occurred in different years with 1 explanatory variable

Source


Examples

n <- length(strikes)
r <- 1
s.y <- sum(strikes)
p.posterior.1000000 <- rbeta(1000000,n+r,s.y+0.5)
length(p.posterior.1000000[p.posterior.1000000<0.05])/1000000
par(mar=c(3,3,3,3))
ruler <- seq(0,1,length=1000)
beta.vals <- dbeta(ruler,n+r,s.y+0.5)
plot(ruler[1:200],beta.vals[1:200],yaxt="n",main="",ylab="",type="l")
Description

Dataset comparing incidents of terrorism to car accidents, suicide, and murder, see page 140

The variables included in the dataset are:

- **Year** The given year in which the statistics occurred
- **X.Terrorism** The number of terrorist attacks that would occur per 100000 in the given year
- **X.Car.Accidents** The number of car accidents that would occur per 100000 in the given year
- **X.Suicide** The number of suicide that would occur per 100000 in the given year

Usage

data(terrorism)

Format

data frame with 14 observations of death rates for different years with 5 explanatory variables

Source


Description

Poverty in Texas, see page 299

The variables included in the dataset are:

- **POV** a dichotomous outcome variable indicates whether 20% or more of the county’s residents live in poverty
- **BLK** the proportion of Black residents in the county
- **LAT** the proportion of Latino residents in the county
- GVT a dichotomous variable indicating whether government activities contributed a weighted annual average of 25
- SVC a dichotomous variable indicating whether service activities contributed a weighted annual average of 50
- FED a dichotomous variable indicating whether federally owned lands make up 30
- XFR a dichotomous factor indicating whether income from transfer payments (federal, state, and local) contributed a weighted annual average of 25 percent or more of total personal income over the past three years
- POP the log of the county population total for 1989

Usage

data(texas)

t_ci_table coefs,cov.mat,level=0.95,degrees=Inf,quantiles=c(0.025,0.500,0.975))

Arguments

coeufs vector of coefficient estimates, usually posterior means
cov.mat variance-covariance matrix
level desired coverage level
degrees degrees of freedom parameter for students-t distribution assumption
quantiles vector of desired CDF points (quantiles) to return

Value

quantile.mat matrix of quantiles

Author(s)

Jeff Gill
Description

Data for Chinese wars example, see page 163

The variables included in the dataset are:

- **ONSET** ratio-level variable measuring the epochal (whether historical or calendar) time of event occurrence, measured in calendar year
- **TERM** ratio-level variable measuring the epochal (historical) time of event conclusion, measured in calendar year
- **EXTENT** number of belligerents involved on all sides of the war
- **ETHNIC** intra-group or inter-group conflict
- **DIVERSE** number of ethnic groups participating as belligerents
- **ALLIANCE** total number of alliances among belligerents
- **DYADS** number of alliance pairs
- **POL.LEV** nominal-level variable measuring the political level of belligerent involvement regarding domestic and foreign belligerents, with a 1 for internal war, 2 for interstate war
- **COMPLEX** governmental level of the warring parties, where the first variable is multiplied by ten for scale purposes
- **POLAR** number of relatively major or great powers at the time of onset
- **BALANCE** the difference in military capabilities: minor-minor, minor-major, major-major
- **TEMOR** type of war: protracted rivalry, integrative conquest, disintegrative/fracturing conflict, sporadic event
- **SCOPE** political scope of conflicts in terms of governmental units affected
- **DURATION** duration of conflict, measured in years

Usage

data(wars)

Format

a data frame of 104 observations of different China wars with 15 explanatory variables

Source

Index

actuarial, 3  
adam.jags, 3  
afghan.deaths, 4  
africa, 4  
AGEGRP (adam.jags), 3  
aggrevat (baldus), 6  
asap.data.list, 5  
baldus, 6  
bcp, 8  
biv.norm.post, 8, 25, 30  
bloody (baldus), 6  
cabinet.duration, 10  
CASES (adam.jags), 3  
CATS (adam.jags), 3  
child, 10  
china.wars, 11  
coal.mining.disasters, 12  
COCSELF (adam.jags), 3  
COCTEST (adam.jags), 3  
contracep, 12  
COVARS (adam.jags), 3  
DA_cwp, 13  
def.age (baldus), 6  
defense (baldus), 6  
did.app (baldus), 6  
dmultinorm, 14  
dp, 14  
dp.sght (baldus), 6  
durations.hpd, 15  
educatn (baldus), 6  
elleispnd, 16  
employ (baldus), 6  
ethnic.immigration, 16  
exeuctions, 17  
exteps, 17  
expo.gibbs, 18  
expo.metrop, 18  
fam.lov (baldus), 6  
fdr, 19  
GROUP (adam.jags), 3  
hanjack, 19  
hit.run, 20  
ID (adam.jags), 3  
insane (baldus), 6  
iq, 21  
italy.parties, 22  
jdge.dec (baldus), 6  
logit.posterior.ll (sir), 31  
lunatics, 22  
marrriage.rates, 22  
marrried (baldus), 6  
methodA (baldus), 6  
military (militarydf), 23  
militarydf, 23  
mitcir (baldus), 6  
MJSELF (adam.jags), 3  
MJTEST (adam.jags), 3  
nc.sub.dat, 24  
norm.known.var, 25  
normal.posterior.ll (sir), 31  
normal.posterior.summary, 25  
norr, 26  
um.chld (baldus), 6  
um.depr (baldus), 6  
um.kill (baldus), 6  
um.prps (baldus), 6  
OFFENSE (adam.jags), 3
opic, 26
out.appl(baldus), 6

pbc.vote, 27
pen.phse(baldus), 6
plea(baldus), 6
plot_walk_G, 27
plot_walk_MH, 28
pr.arrst(baldus), 6
pr.incr(baldus), 6
probit.posterior.ll(sir), 31

RACE(adam.jags), 3
race(baldus), 6
rape(baldus), 6
recidivism, 28
retail.sales, 29
rmultinorm, 30
romney, 30

sentence(baldus), 6
SES(baldus), 6
SEX(adam.jags), 3
sir, 31
SITE(adam.jags), 3
socatt, 33
specialA(baldus), 6
strikes, 34

t_ci_table, 36
t_posterior.ll(sir), 31
terrorism, 35
texas, 35

vict.age(baldus), 6
vict.rel(baldus), 6
vict.sex(baldus), 6
vict.st1(baldus), 6
vict.st2(baldus), 6

wars, 37