Aims

- Multiparameters models
- Normal model with mean and variance unknown, with
  - non-informative,
  - conjugate and semi-conjugate prior
- Monte Carlo integration (example for normal data)
- Multinomial model with conjugate Dirichlet prior
  - example (Presidential election)
- Multivariate Normal model with mean and covariance unknown, with
  - non-informative prior
  - conjugate prior
- Application: analysis of bioassay experiment

Multiparameter Models

- θ_1, θ_2
- θ_1 is of interest
- θ_2 is a nuisance parameter

Let's find the marginal posterior distribution of θ_1

$$\boxed{p(\theta_1 \mid y) = \int p(\theta_1, \theta_2 \mid y) d\theta_2}$$

Using simulations

- draw $\theta_2^{(m)} \sim p(\theta_2 \mid y)$
- draw $\theta_1^{(m)} \sim p(\theta_1 \mid \theta_2^{(m)}, y)$
- find $\theta_1^{(m)}, \theta_2^{(m)}$ for $m = 1, \dots, M$

Normal model, μ and σ^2 unknown

- μ is the parameter of interest
- σ^2 is the nuisance parameter

$$\begin{aligned} p(\mu, \sigma^2 \mid y) &\propto \quad p(y \mid \mu, \sigma^2) p(\mu, \sigma^2) \\ p(\mu \mid y) &= \int p(\mu \mid \sigma^2, y) p(\sigma^2 \mid y) d\sigma^2 \end{aligned}$$

$p(\mu \mid y)$ is a mixture of the conditional posterior distributions given σ^2, where $p(\sigma^2 \mid y)$ is a weighting function for the possible values of σ^2

Normal model, $p(\mu, \sigma^2) \propto \frac{1}{\sigma^2}$

- joint posterior

$$p(\mu, \sigma^2 \mid y) \propto \sigma^{-(n+2)} \exp\left(-\frac{1}{2\sigma^2}\left[(n-1)s^2 + n(\bar{y} - \mu)^2\right]\right)$$

- conditional posterior distribution of μ

$$\mu \mid \sigma^2, y \sim N(\mu \mid \bar{y}, \frac{\sigma^2}{n})$$

- marginal posterior distribution of σ^2

$$\sigma^2 \mid y \sim IG(\frac{n-1}{2}, \frac{(n-1)s^2}{2})$$

$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})^2$

Baby Numerical Integration

$$p(\mu, \sigma^2 \mid y) = p(\mu \mid \sigma^2, y) p(\sigma^2 \mid y)$$

- draw $\sigma^{2^{(m)}} \sim IG(\frac{n-1}{2}, \frac{(n-1)s^2}{2})$
- draw $\mu^{(m)} \sim N(\mu \mid \bar{y}, \sigma^{2^{(m)}}/n)$
- $\mu^{(m)}, \sigma^{2^{(m)}}$ $m = 1, \ldots, M$ is a sample from the joint posterior distribution $p(\mu, \sigma^2 \mid y)$


```
#example of generating a sample from the joint posterior distribution of
#a normal model with uninformative prior
#we simulate a sample of size 1000 from a N(0,2)
yy_rnorm(1000,0,sqrt(4))
draw_function(w=1,NN,yy){
  par(mfrow=c(2,1))
  if (w == 1) postscript("/home/biostats/fdominic/course/babymcmc.ps")
  mu_rep(NA,NN)
  sigmasqr_rep(NA,NN)
  nn_length(yy)
  ss_var(yy)
  ybar_mean(yy)
  mu[1]_0              #initial values
  sigmasqr[1]_2
  for(m in 2:NN){
    AA_(nn-1)/2
    BB_.5*(nn-1)*ss
    sigmasqr[m]_1/(rgamma(1,AA)/BB)
    mu[m]_rnorm(1,ybar, sqrt(sigmasqr[m]/nn))
  }
  hist(mu,density=-1,yaxt="n",ylab="",xlab="mu",cex=2,nclass=50)
  hist(sqrt(sigmasqr),density=-1,yaxt="n",ylab="",
  xlab="sigma",cex=2,nclass=50)
  par(oma=c(0,0,0,0))
  par(mfrow=c(1,1))
  if (w == 1) dev.off()}
```


Marginal posterior distribution of μ

$$p(\mu \mid y) \;=\; \int_0^\infty p(\mu \mid \sigma^2, y) p(\sigma^2 \mid y) d\sigma^2$$

$$\mu \mid y \qquad \sim t_{n-1}(\bar{y}, \frac{s^2}{n})$$

1. $\frac{\mu - \bar{y}}{s/\sqrt{n}} \mid y \sim t_{n-1}$
2. $\frac{\bar{y} - \mu}{s/\sqrt{n}} \mid \mu \sim t_{n-1}$

(1) the posterior distribution of the pivotal quantity doesn't depend on the data, and (2) the sampling distribution of the pivotal quantity doesn't depend on the parameter

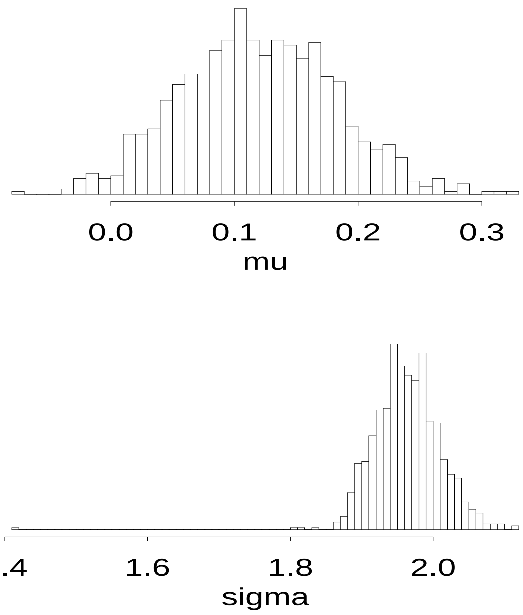


Figure 1: histograms of sample from the marginal posterior distribution of μ and σ

Posterior predictive distribution

$$
\begin{aligned}
p(\tilde{y} \mid y) &= \int p(\tilde{y} \mid \mu, \sigma^2) p(\mu, \sigma^2 \mid y) d\mu d\sigma^2 \\
&= \int p(\tilde{y} \mid \mu, \sigma^2) p(\mu \mid \sigma^2, y) p(\sigma^2 \mid y) d\mu d\sigma^2 \\
&= \int N\left(\tilde{y} \mid \bar{y}, (1 + \tfrac{1}{n})\right) p(\sigma^2 \mid y) d\sigma^2 \\
&= t_{n-1}\left(\bar{y}, (1 + \tfrac{1}{n})^{1/2} s\right)
\end{aligned}
$$

Normal data with a conjugate prior distribution

$$
\begin{aligned}
p(y \mid \mu, \sigma^2) &\sim (\tfrac{1}{\sigma})^{n+2} \\
&\quad \exp\left(-\tfrac{1}{2\sigma^2}\left[(n-1)s^2 + n(\bar{y} - \mu)^2\right]\right) \\
p(\mu, \sigma^2) &\sim (\tfrac{1}{\sigma})(\tfrac{1}{\sigma})^{\nu_0+2} \\
&\quad \exp\left(-\tfrac{1}{2\sigma^2}\left[\nu_0\sigma_0^2 + k_0(\mu_0 - \mu)^2\right]\right)
\end{aligned}
$$

- μ and σ^2 are dependent in their joint conjugate prior density

$$
\begin{aligned}
\mu \mid \sigma^2 &\sim N(\mu_0, \sigma^2/k_0) \\
\sigma^2 &\sim IG(\tfrac{\nu_0}{2}, \tfrac{\nu^2\sigma_0^2}{2})
\end{aligned}
$$

- $\mu_n = \frac{k_0}{k_0+n}\mu_0 + \frac{n}{k_0+n}\bar{y}$
- $k_n = k_0 + n$
- $\nu_n\sigma_n^2 = \nu_0\sigma_0^2 + (n-1)s^2 + \frac{k_0 n}{k_0+n}(\bar{y} - \mu_0)^2$

- $\nu_n = \nu_0 + n$

$$
\begin{aligned}
\mu \mid \sigma^2, y &\sim N(\mu_n, \sigma^2/k_n) \\
\sigma^2 &\sim IG(\tfrac{\nu_n}{2}, \tfrac{\nu_n\sigma_n^2}{2})
\end{aligned}
$$

Normal data with a semi-conjugate prior distribution

- μ and σ^2 are independent

$$
\begin{aligned}
\mu \mid \sigma^2 &\sim N(\mu_0, \tau_0^2) \\
\sigma^2 &\sim IG(\tfrac{\nu_0}{2}, \tfrac{\sigma_0^2}{2})
\end{aligned}
$$

- $\mu_n = \frac{\mu_0/\tau_0^2 + n/\sigma^2}{1/\tau_0^2 + n/\sigma^2}$
- $\tau_n^2 = \frac{1}{1/\tau_0^2 + n/\sigma^2}$

$$
\begin{aligned}
\mu \mid \sigma^2, y &\sim N(\mu_n, \tau_n^2) \\
\sigma^2 \mid y &\sim \text{not in closed form}
\end{aligned}
$$

Multinomial model

- y_j is the number of observations for the j-th outcome category
- $y = (y_1, \ldots, y_k)$ is the vector of counts of the number of observations of each outcome
- $\sum \theta_j = 1$
- $\sum y_j = n$

$$p(y \mid \theta) \propto \prod_{j=1}^{k} \theta_j^{y_j}$$

$$p(\theta \mid \alpha) \propto \prod_{j=1}^{k} \theta_j^{(\alpha_j - 1)} \propto \mathcal{D}(\alpha)$$

$$p(\theta \mid y, \alpha) \propto \prod_{j=1}^{k} \theta_j^{(y_j + \alpha_j - 1)} \propto \mathcal{D}(y + \alpha)$$

Dirichlet prior

The Dirichlet prior $\mathcal{D}(\alpha)$ contains equivalent information to $\sum_{j=1}^{k} \alpha_j$ observations, with α_j observations of the j-th outcome category

- Uniform prior: $\alpha_j = 1$, for every j
- Improper prior: $\alpha_j = 0$, for every j i.e. uniform on $\log(\theta_j)$
- *the posterior is proper if there is at least one observation in each of the k categories, so that each component y is positive*

[NB $\theta_1, \theta_2, \theta_3 \sim \mathcal{D}(\alpha_1, \alpha_2, \alpha_3)$ then $\theta_1/(\theta_1 + \theta_2) \sim \text{Beta}(\alpha_1, \alpha_2)$ (3.1)]

Example

Survey 1447 adults in US to find out their preferences in the upcoming presidential election

- y_1=727 supported George Bush
- y_2=583 supported Michael Dukakis
- y_3=173 no opinion
- $\theta_1 - \theta_2$ is the difference between George Bush supporter and Michael Dukakis supporter

$$p(y \mid \theta) \propto \theta_1^{727} \theta_2^{583} \theta_3^{173}$$

$$p(\alpha) \propto \mathcal{D}(1, 1, 1)$$

$$p(\theta \mid y, \alpha) \propto \mathcal{D}(728, 584, 174)$$


```
#multinomial dirichlet model
# pre-election polling sec 3.5 pag 76 Gelman's book
#S-plus function for generating sample from a Dirichlet distribution
#GENERATE SAMPLE FROM DIRICHLET--------------------------
rdirichlet _ function(n,p) {
# return n random samples from a Dirichlet distribution with parameter p
  if (    !is.vector(n, "numeric")
      | length(n) != 1
      | !is.vector(p, "numeric")
      ) { stop("error in call to rdirichlet") }
  mat _ matrix ( NA, n, length(p) )
  mat[,1] _ rbeta ( n, p[1], sum(p[-1]) )
  for ( i in 2:(length(p)-1) ) {
    mat[,i] _ (     rbeta ( n, p[i], sum(p[(i+1):length(p)]) )
                * ( 1 - apply ((mat[,1:(i-1),drop=F]), 1, sum) ))}
  mat[,length(p)] _  1 - apply ( (mat[,-length(p),drop=F]), 1, sum )
  return ( mat )
}
```

```
#######EXAMPLE: PRE-ELECTION POLLING
multdirichlet_function(w=0,NN=1000,y=c(727,583,173),alpha=c(1,1,1)){
  if (w == 1) postscript("/home/biostats/fdominic/course/multdirichlet.ps")
  theta_rdirichlet(NN,alpha+y)
  hist(theta[,1]-theta[,2],xlab="theta1-theta2",ylab="",yaxt="n",nclass=50,
       density=-1)
  par(oma=c(0,0,0,0))
  par(mfrow=c(1,1))
  if (w == 1) dev.off()
}
```

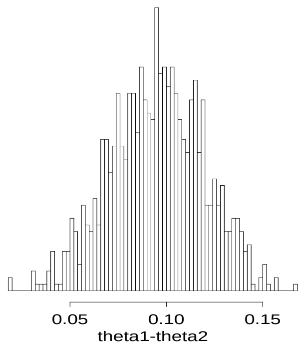



Figure 2: Histogram of values of $\theta_1 - \theta_2$ for 1000 simulations from the posterior distribution for the election polling example

Multivariate Normal Model

y is a vector of d components

$$
\begin{aligned}
y \mid \mu, \Sigma \quad &\sim N_d(y \mid \mu, \Sigma) \\
p(y \mid \mu, \Sigma) &\propto |\Sigma|^{-1/2} \\
&\qquad \exp\left(-\tfrac{1}{2}(y-\mu)^t \Sigma^{-1}(y-\mu)\right)
\end{aligned}
$$

for a sample of n iid obs. $y_1, \ldots, y_n$

$$
\begin{aligned}
p(y_1, \ldots, y_n \mid \mu, \Sigma) &\propto |\Sigma|^{-n/2} \exp\left(-\tfrac{1}{2} tr(\Sigma^{-1} S_0)\right) \\
S_0 &= \textstyle\sum_{i=1}^n (y_i - \mu)(y_i - \mu)^t
\end{aligned}
$$

where S_0 is the matrix of "sum of squares"


```
#GENERATE MULTIVARIATE NORMAL SAMPLE---------------------------
simulate.multnorm _ function(mu0,Q,M=1)
{# return M random samples from a Multivariate Normal distribution
 #with vector mean mu0 and covariance matrix Q
  s _ length(mu0)
  Z _  matrix(rnorm(s*M),nrow=s)
  V <- t(chol(Q) )
  x _ (V%*%Z)+mu0
  return(x)
}
#sample from a multivariate normal distribution, with plot of all the joint
samplemultnorm_function(w=1,NN){
 if (w == 1) postscript("/home/biostats/fdominic/course/samplemultnorm.ps")
  mu0_c(0,0,0)
  Q_cbind(c(1,.9,.1),c(.9,1,0),c(.1,0,1))
  YY_simulate.multnorm(c(0,0,0),Q,M=NN)
  pairs(t(YY),labels=c("y1","y2","y3"))
par(oma=c(0,0,0,0))
par(mfrow=c(1,1))
if (w == 1) dev.off()
}
```

Multivariate Normal with known variance

- Conjugate prior distribution for μ with Σ known

$$\begin{aligned} y \mid \mu, \Sigma &\sim N(y \mid \mu, \Sigma) \\ \mu \mid \mu_0, \Lambda_0 &\sim N(\mu \mid \mu_0, \Lambda_0) \\ \mu \mid y, \mu_0, \Lambda_0 &\sim N(\mu \mid \mu_n, \Lambda_n) \end{aligned}$$

- $\mu_n = \left(\Lambda_0^{-1} + n\Sigma^{-1}\right)^{-1} \left(\Lambda_0^{-1}\mu_0 + n\Sigma^{-1}\bar{y}\right)$
- $\Lambda_n^{-1} = \left(\Lambda_0^{-1} + n\Sigma^{-1}\right)$

$$\mu = \begin{bmatrix} \mu^{(1)} \\ \mu^{(2)} \end{bmatrix} \Lambda = \begin{bmatrix} \Lambda^{(11)} & \Lambda^{(12)} \\ \Lambda^{(21)} & \Lambda^{(22)} \end{bmatrix}$$

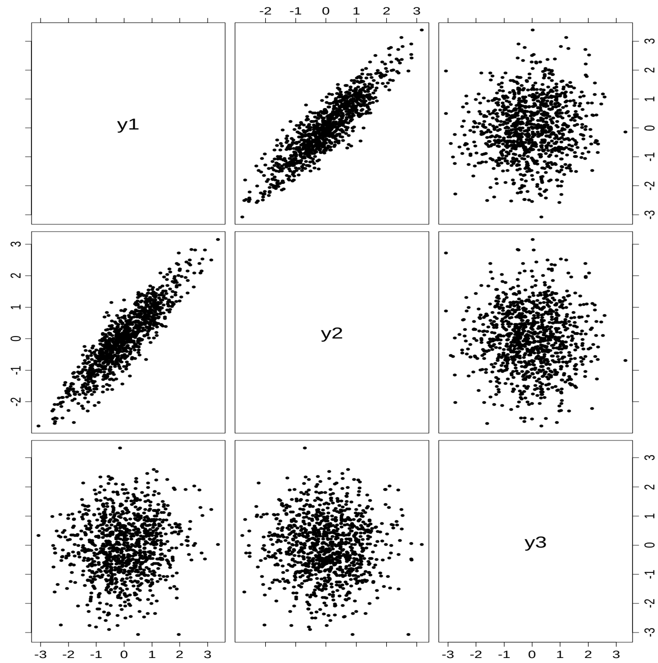


Figure 3: joint likelihoods $p(y_1, y_2 \mid \mu, \Sigma)$, $p(y_1, y_3 \mid \mu, \Sigma)$, and $p(y_2, y_3 \mid \mu, \Sigma)$

Posterior conditional distribution of a sub-vector $\mu^{(1)}$ with Σ known

$$\begin{aligned} \mu^{(1)} \mid \mu_0^{(1)}, \Lambda_0^{(11)} &\sim N(\mu^{(1)} \mid \mu_0^{(1)}, \Lambda_0^{(11)}) \\ \mu^{(1)} \mid \mu^{(2)} &\sim N(\mu_0^{(1)} + B_0(\mu^{(2)} - \mu_0^{(2)}), \Lambda_0^{11.2}) \\ \mu^{(1)} \mid y, \mu^{(2)} &\sim N(\mu_n^{(1)} + B_n(\mu^{(2)} - \mu_n^{(2)}), \Lambda_n^{11.2}) \end{aligned}$$

- $B_0 = \Lambda_0^{(12)} \left[\Lambda_0^{(22)}\right]^{-1}$
- $\Lambda_0^{11.2} = \Lambda_0^{(22)} - \Lambda_0^{(12)} \left[\Lambda_0^{(22)}\right]^{-1} \Lambda_0^{(21)}$
- $B_n = \Lambda_n^{(12)} \left[\Lambda_n^{(22)}\right]^{-1}$
- $\Lambda_n^{11.2} = \Lambda_n^{(22)} - \Lambda_n^{(12)} \left[\Lambda_n^{(22)}\right]^{-1} \Lambda_n^{(21)}$

Posterior predictive distribution for a new data

$$p(\tilde{y}, \mu \mid y) = N(\tilde{y} \mid \mu, \Sigma) N(\mu \mid \mu_n, \Lambda_n)$$

- $\tilde{y}, \mu \mid y$ has a multivariate normal distribution
- $\tilde{y} \mid y$ has a multivariate normal distribution
- $E(\tilde{y} \mid y) = \mu_n$
- $Var(\tilde{y} \mid y) = \Lambda_n + \Sigma$

Non informative prior for μ

- $p(\mu) =$ constant, $|\Lambda_0|^{-1} \to 0$
- $\mu \mid y, \Sigma \sim N(\mu \mid \bar{y}, \Sigma/n)$

$p(\mu \mid y, \Sigma)$ *is a proper posterior only if* $n \geq d$

Multivariate Normal with unknown mean and variance

- Conjugate family of prior distribution
- Univariate case $\mu, \sigma^2 \sim N\chi^{-2}_{inv}(\mu_0, \sigma_0^2/k_0, \nu_0, \sigma_0^2)$
- Multivariate case $\mu, \Sigma \sim NIW(\mu_0, \Lambda_0/k_0, \nu_0, \Lambda_0)$

$$p(\mu, \Sigma) \propto |\Sigma|^{-[(\nu_0+d)/2]+1} \exp\left(-\tfrac{1}{2}tr(\Lambda_0\Sigma^{-1}) - \tfrac{k_0}{2}(\mu-\mu_0)^t\Sigma^{-1}(\mu-\mu_0)\right)$$

- $\mu, \Sigma \mid y \sim NIW(\mu_n, \Lambda_n/k_n, \nu_n, \Lambda_n)$
- $\mu_n = \frac{k_0}{(k_0+n)}\mu_0 + \frac{n}{(k_0+n)}\bar{y}$
- $\Lambda_n = \Lambda_0 + S + \frac{k_0 n}{k_0+n}(\bar{y}-\mu_0)(\bar{y}-\mu_0)^t$
- $S = \sum(y_i - \bar{y})(y_i - \bar{y})^t$


```
#GENERATE WISHART------------------------------------------------------------
"rwish" <- function(s,nu,Cov)
{
#sxs Wishart matrix, nu degree of freedom, var/covar Cov based on
#P.L.Odell & A.H. Feiveson(JASA 1966 p.199-203). Returns w=(RU)'RU
#where Cov=U'U (U is upper triang) and where upper-tri R is
# R_ij~N(0,1), i<j ; (R_ii)^2~Chisq(nu-s+i)
if (nu<=s-1) stop ("Wishart algorithm requires nu>s-1")
R<- diag(sqrt(2*rgamma(s,(nu+1 - 1:s)/2)))
R[outer(1:s, 1:s,  "<")] <- rnorm (s*(s-1)/2)
R <- R%*% chol(Cov)
return(t(R)%*%R)
}

#GENERATE INVERSE WISHART----------------------------------------------------
"riwish" <- function(s,df,Prec)
{
#sxs Inverse Wishart matrix, df degree of freedom, precision matrix
#Prec. Distribution of W^{-1} for Wishart W with nu=df+s-1 degree of
# freedoom, covar martix Prec^{-1}.
# NOTE  mean of riwish is proportional to Prec
if (df<=0) stop ("Inverse Wishart algorithm requires df>0")
R <- diag(sqrt(2*rgamma(s,(df + s  - 1:s)/2)))
R[outer(1:s, 1:s,  "<")] <- rnorm (s*(s-1)/2)
S <- t(solve(R))%*% chol(Prec)
return(t(S)%*%S)
}
```


Marginal posterior of μ and posterior predictive distributions

- $\mu \mid y \sim t_{\nu_n-d+1}\left(\mu_n, \Lambda_n/(k_n(\nu_n-d+1))\right)$
- $\tilde{y} \mid y \sim t_{\nu_n-d+1}\left(\mu_n, (\Lambda_n + k_n + 1)/(k_n(\nu_n-d+1))\right)$

Non informative prior distribution

- Multivariate Jeffreys prior density

$$p(\mu, \Sigma) \propto |\Sigma|^{-(d+1)/2}$$

- limit of the conjugate prior for

  $k_0 \to 0, \mid \Lambda_0 \mid \to 0$ and $\nu_0 \to -1$

- $\Sigma \mid y \sim IW_{n-1}(S)$
- $\mu \mid \Sigma, y \sim N(\bar{y}, \Sigma/n)$

Table 1: *Bioassay data from Racine et al. (1986)*

Dose, x_i	Number of animals, n_i	Number of deaths, y_i
-.863	5	0
-.296	5	1
-.053	5	3
.727	5	5

Analysis of Bioassay experiment

- (x_i, n_i, y_i)
- $x_i = i$-th of k dose level
- $n_i =$ number of animals treated with dose i
- $y_i =$ number of deaths
- $\theta_i =$ probability of death given dose i
- $y_i \mid \theta_i \sim \text{Bin}(y_i \mid \theta_i, n_i)$
- $\text{logit}(\theta_i) = \alpha + \beta x_i$
- $\theta_i \sim \text{Beta}(1,1)$, iid, $i = 1, 2, 3, 4.$

Posterior inference on α and β

$$p(y_i \mid \alpha, \beta) \propto \left[\text{logit}^{-1}(\alpha + \beta x_i)\right]^{y_i} \times \left[1 - \text{logit}^{-1}(\alpha + \beta x_i)\right]^{n_i - y_i}$$

$$p(\alpha, \beta) \propto \text{constant}$$

$$p(\alpha, \beta \mid y) \propto p(\alpha, \beta) \prod_{i=1}^{n} p(y_i \mid \alpha, \beta, n_i, x_i)$$

- $(\hat{\alpha}, \hat{\beta}) = (.1, 2.9)$
- $(std(\hat{\alpha}), std(\hat{\beta})) = (.3, .5)$
- contour plot of the joint posterior density
- posterior distribution of $LD50 = \alpha/\beta$

Sampling from the joint posterior

- compute the unnormalized posterior density at a grid of values that cover the range of α and β
- normalize it by approximating the distribution as a step function over the grid and setting the total probability of the grid equal to 1
- compute the marginal posterior of α by summing over β in the step-function computed on the grid
- for $l = 1, \ldots, 1000$

  draw α from $p(\alpha \mid y)$ using the inverse cdF methods

  draw β from $p(\beta \mid \alpha, y)$ given the last sampled value of α using the inverse cdf method
- $\alpha^{(l)}, \beta^{(l)},\ l = 1, \ldots 100$ scatter-plot

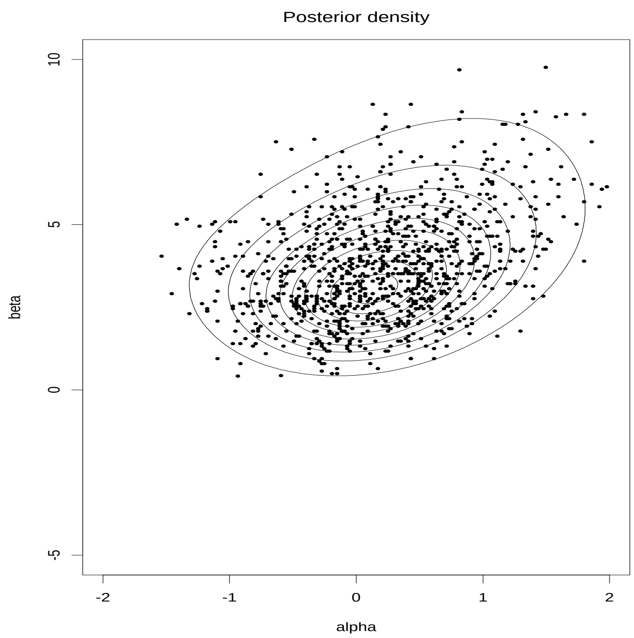

```
#Example:analysis o bioassay experiment
#Gelman and Rubin book, pag 82
#DD_input()
        input_function(LL=200){
        biossay _ data.frame(doses=c(-.863,-.296,-.053,.727),
        rats=c(5,5,5,5),deaths=c(0.5,1,3,4.5),freq=c(0.5,1,3,4.5)/5)
        DD        _ data .frame(y=log((biossay$freq)/(1-biossay$freq)),x=biossay$doses)
        estimates _ lm(y~x,data=DD)
        alpha.hat _ summary(estimates)$coeff[1,1]
        std.alpha _ summary(estimates)$coeff[1,2]
        beta.hat  _ summary(estimates)$coeff[2,1]
        std.beta  _ summary(estimates)$coeff[2,2]
        alpha     _ seq(-2,2,length=LL)
        beta      _  seq(-5,10,length=LL)
        return(biossay,estimates,alpha,beta)}

        log.post_function(alpha,beta,data=DD$biossay){
        ldens _ 0
        for (i in 1:length(data$doses)){
        theta _ 1/(1+exp(-alpha-beta*data$doses[i]))
        ldens  _ ldens + data$deaths[i]*log(theta)+
        (data$rats[i]-data$deaths[i])* log(1-theta)} ldens}

        plot.joint.post_function(w=0,data,draws){
        if (w == 1) postscript("/home/biostats/fdominic/course/jposterior.ps")
        contours _ seq(.05,.95,.1)
        logdens _outer(DD$alpha,DD$beta,log.post,data)
        dens_exp(logdens-max(logdens))
        contour(DD$alpha,DD$beta,dens,levels=contours,xlab="alpha",
        ylab="beta",labex=0,cex=1)
        points(draws$aaa,draws$bbb)
        mtext("Posterior density",3,line=1,cex=1.2)}

      grid.value_function(data=DD$biossay,alpha,beta){
        ll    _ length(alpha)
        PP    _ matrix(NA,ll,ll)
        for(i in 1:ll){
```



```
        for(j in 1:ll){
        PP[i,j]_exp(log.post(alpha[i],beta[j],data))
        }}
        ccc _ sum(PP)
        PP  _ PP/ccc
        return(PP)
}
  sampling_function(M=100,PP,alpha,beta){
  alpha.mar_apply(PP,1,sum)
  alpha.cdf  <-  cumsum(alpha.mar)
  post.alpha  <-  post.beta  <-  rep(0,M)
  for( m in 1:M){
  uuu_runif(1,0,1)
  Fhat.alpha  <-  max( alpha.cdf[ alpha.cdf <= uuu])
  post.alpha[m]  <- alpha[(1:length(alpha.cdf))[alpha.cdf == Fhat.alpha]]
  junk  <- length(alpha[alpha <= post.alpha[m]])
  PP[junk, ]  <-  PP[junk,]/sum(PP[junk,])
  beta.cond.cdf  <-   cumsum(PP[junk,])
  uuu_runif(1,0,1)
  Fhat.beta  <-  max( beta.cond.cdf[ beta.cond.cdf < uuu])
  post.beta[m]  <-  beta[(1:length(beta.cond.cdf))[beta.cond.cdf == Fhat.beta]]
}
  return(post.alpha,post.beta)
}
```

Summary

1. write the likelihood $p(y \mid \theta)$ ignoring any factor free of θ
2. write the posterior $p(\theta \mid y) \propto p(\theta)p(y \mid \theta)$
3. find a crude estimate of the parameter θ for use as starting point
4. draw sample $\theta^{(l)};\ l = 1, ..., L$ from $p(\theta \mid y)$
5. use the sample draws to compute the posterior density of any function of θ that may be of interest
6. if any predictive quantity, $\tilde{y}$, are of interest, draw sample $\tilde{y}^{(l)};\ l = 1, ..., L$ from $p(\tilde{y} \mid \theta^{(l)})$



Sampling using the inverse cdF method

$$F(v_*) = P(v \leq v_*) = \begin{cases} \sum_{v < v_*} p(v) \\ \int_{-\infty}^{v_*} p(v)dv \end{cases}$$

1. draw $u \sim U[0,1]$
2. calculate $v = F^{-1}(u)$ then $v \sim p$

Example:

- $v \sim \text{Exp}(\lambda)$
- $F(v) = 1 - e^{-\lambda v}$
- $v = F^{-1}(u)$ iff $v = -\log(1-u)/\lambda$