Normal model with conjugate priors

Case study: Nissan Maxima data-set

Georgios P. Karagiannis @ MATH3341/4031 Bayesian statistics III/IV (practical implementation)

Back to the main document

```
rm(list=ls())
```

Aim

Students will become able to:

- produce Monte Carlo approximations of posterior quantities required for Bayesian analysis with the RJAGS R package
- implement Bayesian posterior analysis in R with RJAGS package

Students are not required to learn by heart any of the concepts discussed

---

Reading material

The material about RJAGS package is not examinable material, but it is provided for the interested student. It contains references that students can follow if they want to further explore the concepts introdced.

- Lecture notes:
  - the examples related to the Bernoulli model with conjugate prior
- References for *RJAGS*:
  - JAGS homepage
  - JAGS R CRAN Repository
  - JAGS Reference Manual
  - JAGS user manual
- Reference for *R*:
  - Cheat sheet with basic commands
- Reference of *rmarkdown* (optional):
  - R Markdown cheatsheet
  - R Markdown Reference Guide
  - knitr options
- Reference for *Latex* (optional):
  - Latex Cheat Sheet

New software

- R package `rjags` functions:

    - jags.model{rjags}

    - jags.samples{rjags}
    - update{rjags}

Application:

The applications below is 'reproduced' from DASL, and

- De Veaux, R. D., Velleman, P. F., Bock, D. E., Vukov, A. M., Augustine, C. W., & Burkett, C. (2005). Stats: data and models. Boston: Pearson/Addison Wesley.

Richard DeVeaux owned a Nissan Maxima for 8 years. Being a statistician, He recorded the car's fuel efficiency (in mpg) each time he filled the tank. He wanted to know what fuel efficiency to expect as "ordinary" for his car. Knowing this, he was able to predict when he'd need to fill the tank again and to notice if the fuel efficiency suddenly got worse, which could be a sign of trouble.

General scientific question:

```
# load the data
mydata <- read.csv("./nissan.csv")
#mydata$mpg <- rnorm(100, 20, 5)
# print data
cat(mydata$mpg)
```

```
## 21.964 23.694 18.824 20.851 26.37 22.81 25.785 24.353 23.385 23.381 28.175 21.232 25.603 21.064 22.0
```

Preliminary analysis

Task

Examine graphically whether there is any substantial evidence that the above data-set is not generated from the Normal distribution.

Are there any substaintial evivence against Normality?

... your answer

Below try to address the task by creating R chunks, and running them, and adding comments tot he document.

```
# write your R code
```

Model

We specifiy the Normal distribution with uncertain mean and variance parameters μ and σ^2 as a parametric model.

We assign a Normal-Inverse Gamma conjugate prior with fixed hyper-parameters $\mu_0 = 0.0$, $\lambda_0 = 1.0$, $a_0 = 1.0$, and $b_0 = 1.0$.

The Bayesian hierarchical model under consideration is:

$$\begin{cases} x_i|\mu,\sigma^2 & \overset{\text{iid}}{\sim} \text{N}(\mu,\sigma^2),\ \forall i=1,...,n \\ \mu|\sigma^2 & \sim \text{N}(\mu_0,\frac{\sigma^2}{\lambda_0}) \\ \sigma^2 & \sim \text{G}(a,b) \end{cases}$$

for some fixed values of the hyper-parameters λ_0, a, and b.

Here, we just set:

- $\mu_0 = 0.0$,
- $\lambda_0 = 1.0$,
- $a_0 = 0.01$,
- $b_0 = 0.01$

Task

1. Write the Stan program to draw a sample from the posterior distribution of the parameters μ, and σ^2
2. Generate a sample of size $N = 1000$ from the posterior distribution of the parameters.
3. Draw trace plots of the samples desived

... your answer

Below try to address the task by creating R chunks, and running them, and adding comments tot he document.

```
# write your R code
```

Task

Produce and plot the exact joint PDF of (μ,σ^2) as a 2D plot, and as a 2D contour plot

To draw 3D plots like those required to plot the exact PDF you can use the functions `persp{graphics}`, or `contour{graphics}` as `persp(x, y , z)` or `contour(x, y, z)` , where x, y are locations of grid lines at which the values in z are measured; and z is a matrix containing the values to be plotted ;

Usage

```
% To plot z=f(x,y) for a specific function f(.,.) at vectors x and y, then
% for known x, y vectors of length 100

z <- matrix(rep(NaN,100*100),100,100)
for (i in 1:100)
  for (j in 1:100)
    z[i,j] <-f( x[i],z[j] )

persp(x, y , z)

% or

contour(x, y, z)
```

Hint-1

The joint posterior distribution (μ, σ^2), is:

$$\begin{cases} \mu | x_{1:n}, \sigma^2 & \sim \mathrm{N}(\mu_n, \frac{\sigma^2}{\lambda_n}) \\ \sigma^2 | x_{1:n} & \sim \mathrm{IG}(a_n, b_n) \end{cases}$$

where

$$\begin{aligned} \mu_n =& \frac{\lambda_0 \mu_n + n\bar{x}}{\lambda_0 + n} \\ \lambda_n =& \lambda_0 + n, \\ a_n =& a_0 + \frac{n}{2} \\ b_n =& b_0 + \frac{1}{2} n s^2 + \frac{1}{2} (\lambda_0 + n)^{-1} \lambda_0 n (\mu_0 - \bar{x})^2 \end{aligned}$$

Hence the PDF is

$$\pi(\mu, \sigma^2 | x_{1:n}) = \mathrm{N}(\mu | \mu_n, \frac{\sigma^2}{\lambda_n}) \mathrm{IG}(\sigma^2 | a_n, b_n)$$

Hint:

Additional functions for PDFs and CDFs are given

```
# Inverse Gamma PDF
invgamma_pdf <- function (x,a,b) {
  return(exp(a*log(b)-lgamma(a)-(a+1)*log(x)-b/x))
}
# Inverse Gamma CDF
invgamma_cdf <- function (x,a,b) {
  return(1.0-pgamma(1.0/x,shape=a,scale=1.0/b) )
}
```

. . . your answer

Below try to address the task by creating R chunks, and running them, and adding comments tot he document.

```
# write your R code
```

Task

Produce and plot the MC approximation of the joint PDF of (μ, σ^2), as a 2D histogram, and as a 2D contour plot.

To draw histograms of 2D samples like required by the MC approximation of the PDF you can use the functions `hist2d{gplots}` and `persp{graphics}`, or `contour{graphics}` as:

```
% known vectors x and y of length 100 having the 2D samples:

h2d <- hist2d(x,y,nbins=c(20,30),show=FALSE)
```

```
persp( h2d$x, h2d$y, h2d$counts, ticktype="detailed", theta=30, phi=30,expand=0.5, shade=0.5, col="cyan

contour( h2d$x, h2d$y, h2d$counts, ticktype="detailed", theta=30, phi=30,expand=0.5, shade=0.5, col="cya
```

... your answer

Below try to address the task by creating R chunks, and running them, and adding comments tot he document.

```
# write your R code
```

```
library(gplots)
```

Task

Produce and plot the Exact marginal PDF of μ.

Produce and plotthe MC approximation of the marginal PDF of μ.

You can use the functions `hist {graphics}` and `lines {graphics}` or `plots {graphics}`

Hint

The marginal distribution of μ, is:

$$\mu|x_{1:n} \sim \text{St}_k(\mu_n, \frac{b_n}{\lambda_n a_n}, 2a_n)$$

where

$$\begin{aligned}
\mu_n =& \frac{\lambda_0 \mu_n + n\bar{x}}{\lambda_0 + n} \\
\lambda_n =& \lambda_0 + n, \\
a_n =& a_0 + \frac{n}{2} \\
b_n =& b_0 + \frac{1}{2} n s^2 + \frac{1}{2}(\lambda_0 + n)^{-1} \lambda_0 n (\mu_0 - \bar{x})^2
\end{aligned}$$

Hint:

Additional PDFs and CDFs are given

```
# Student T PDF
studentT_pdf <- function(x,m,s,v) {
  return(dt((x-m)/sqrt(s),df=v)/sqrt(s))
}
# Student T CDF
studentT_cdf <- function(x,m,s,v) {
  return(pt((x-m)/sqrt(s),df=v))
}
```

... your answer

Below try to address the task by creating R chunks, and running them, and adding comments tot he document.

```r
# write your R code
```

Task

Produce and plot the Exact marginal PDF of σ^2.

Produce and plot the MC approximation of the marginal PDF of σ^2.

You can use the functions `hist {graphics}` and `lines {graphics}` or `plot {graphics}`

Hint

The marginal distributions of σ^2, as:

$$\sigma^2|x_{1:n} \sim \text{IG}(\sigma^2|a_n, b_n)$$

where

$$
\begin{aligned}
\mu_n =& \frac{\lambda_0\mu_n + n\bar{x}}{\lambda_0 + n} \\
\lambda_n =& \lambda_0 + n, \\
a_n =& a_0 + \frac{n}{2} \\
b_n =& b_0 + \frac{1}{2}ns^2 + \frac{1}{2}(\lambda_0 + n)^{-1}\lambda_0 n(\mu_0 - \bar{x})^2
\end{aligned}
$$

Hint:

Additional PDFs and CDFs are given

```r
# Inverse Gamma PDF
invgamma_pdf <- function (x,a,b) {
  return(exp(a*log(b)-lgamma(a)-(a+1)*log(x)-b/x))
}
# Inverse Gamma CDF
invgamma_cdf <- function (x,a,b) {
  return(1.0-pgamma(1.0/x,shape=a,scale=1.0/b) )
}
```

... your answer

Below try to address the task by creating R chunks, and running them, and adding comments tot he document.

```r
# write your R code
```

Task

- Compute the MC approximate of the posterior mean of μ, and σ^2:

$$\mathrm{E}_{\pi}(\mu|y_{1:n}) \approx \frac{1}{N}\sum_{j=1}^{N}\mu^{(j)}$$

and

$$\mathrm{E}_{\pi}(\sigma^2|y_{1:n}) \approx \frac{1}{N}\sum_{j=1}^{N}\left(\sigma^{(j)}\right)^2$$

- Compute their exact values which are

$$\mathrm{E}_{\pi}(\mu|y_{1:n}) = \mu_n$$

and

$$\mathrm{E}_{\pi}(\sigma^2|y_{1:n}) = \frac{b_n}{a_n - 1}$$

... your answer

Below try to address the task by creating R chunks, and running them, and adding comments tot he document.

```
# write your R code
```

Task

- Compute the MC approximate of the posterior prabability that the mean fuel efficiency (in mpg) of car is greater or equal to 22.5 mpg.

$$\begin{aligned}
\mathrm{Pr}_{\pi}(\mu \geq 22.5|y_{1:n}) &= 1 - \mathrm{Pr}_{\pi}(\mu < 22.5|y_{1:n}) &(1)\\
&= 1 - \mathrm{E}_{\pi}(1(\mu < 22.5)|y_{1:n}) &(2)\\
&\approx 1 - \frac{1}{N}\sum_{j=1}^{N} 1(\mu^{(j)} < 22.5) &(3)
\end{aligned}$$

- Compute the exact value which is

$$\begin{aligned}
\mathrm{Pr}_{\pi}(\mu \geq 22.5|y_{1:n}) =& 1 - \mathrm{Pr}_{\mathrm{St}_k(\mu_n, \frac{b_n}{\lambda_n a_n}, 2a_n)}(\mu \leq 22.5|y_{1:n}) &(4)\\
=& 1 - \int_{-\infty}^{22.5} \mathrm{St}_k(\mu|\mu_n, \frac{b_n}{\lambda_n a_n}, 2a_n)\mathrm{d}\mu &(5)
\end{aligned}$$

Hint

Functions for the PDFs and CDF are given

```r
# Student T PDF
studentT_pdf <- function(x,m,s,v) {
  return(dt((x-m)/sqrt(s),df=v)/sqrt(s))
}
# Student T CDF
studentT_cdf <- function(x,m,s,v) {
  return(pt((x-m)/sqrt(s),df=v))
}
```

... your answer

Below try to address the task by creating R chunks, and running them, and adding comments tot he document.

```r
# write your R code
```

Task

Compute the exact 95% equal tail posterior credible interval of μ.

Hint

The marginal distribution of μ, is:

$$\mu|x_{1:n} \sim \mathrm{St}_k(\mu_n, \frac{b_n}{\lambda_n a_n}, 2a_n)$$

where

$$\begin{aligned}
\mu_n =& \frac{\lambda_0 \mu_n + n\bar{x}}{\lambda_0 + n} \\
\lambda_n =& \lambda_0 + n, \\
a_n =& a_0 + \frac{n}{2} \\
b_n =& b_0 + \frac{1}{2} n s^2 + \frac{1}{2}(\lambda_0 + n)^{-1} \lambda_0 n (\mu_0 - \bar{x})^2
\end{aligned}$$

Compute the MC approximation of the 95% equal tail posterior credible interval of μ.

Hint:

Additional PDFs and CDFs are given

```r
# Student T INVERSE CDF
studentT_inv <- function(prob,m,s,v) {
  q = qt( prob, df=v )
  return( m+sqrt(s*v/(v-2))*q )
}
```

. . . your answer

Below try to address the task by creating R chunks, and running them, and adding comments tot he document.

```
# write your R code
```

Task

Compute the exact 95% equal tail posterior credible interval of μ.

Hint

The marginal distribution of μ, is:

$$\mu|x_{1:n} \sim \text{St}_k(\mu_n, \frac{b_n}{\lambda_n a_n}, 2a_n)$$

where

$$\begin{aligned}
\mu_n =& \frac{\lambda_0 \mu_n + n\bar{x}}{\lambda_0 + n} \\
\lambda_n =& \lambda_0 + n, \\
a_n =& a_0 + \frac{n}{2} \\
b_n =& b_0 + \frac{1}{2} n s^2 + \frac{1}{2}(\lambda_0 + n)^{-1} \lambda_0 n (\mu_0 - \bar{x})^2
\end{aligned}$$

Compute the MC approximation of the 95% equal tail posterior credible interval of μ.

Hint:

Additional PDFs and CDFs are given

```
# Student T INVERSE CDF
invgamma_inv <- function (prob, a, b){
    return(qgamma(1 - prob, a, b)^(-1))
}
```

. . . your answer

Below try to address the task by creating R chunks, and running them, and adding comments tot he document.

```
# write your R code
```

Task

- Compute and plot the MC approximate of the predictive PDF of the next outcome y_{n+1}

$$
\begin{aligned}
f(y_{n+1}|y_{1:n}) =& \int f(y_{n+1}|\mu,\sigma^2)\pi(\mu,\sigma^2|y_{1:n})\mathrm{d}\mu\mathrm{d}\sigma^2 \\
=& \mathrm{E}_{\pi}(f(y_{n+1}|\mu,\sigma^2)|y_{1:n}) \\
=& \mathrm{E}_{\pi}(\mathrm{N}(y_{n+1}|\mu,\sigma^2)|y_{1:n}) \\
\approx& \frac{1}{N}\sum_{j=1}^{N}\mathrm{N}\left(y_{n+1}|\mu^{(j)},\left(\sigma^{(j)}\right)^2\right)
\end{aligned}
$$

  for $y_{n+1} \in (5,40)$ where $f(y_{n+1}|\mu,\sigma^2)$ is a Normal PDF with mean μ and variance σ^2.

- Compute and plot the exact predictive PDF of the next outcome y_{n+1}, which is the PDF of

$$
y_{n+1}|y_{1:n} \sim \mathrm{St}(\mu_n, \frac{\lambda_n b_n}{(\lambda_n+1)a_n}, 2a_n)
$$

  where

$$
\begin{aligned}
\mu_n =& \frac{\lambda_0\mu_n + n\bar{x}}{\lambda_0+n} \\
\lambda_n =& \lambda_0 + n, \\
a_n =& a_0 + \frac{n}{2} \\
b_n =& b_0 + \frac{1}{2}ns^2 + \frac{1}{2}(\lambda_0+n)^{-1}\lambda_0 n(\mu_0-\bar{x})^2
\end{aligned}
$$

  for $y_{n+1} \in (5,40)$.

Hint

Additional PDF and CDF functions are given

```
# Student T PDF
studentT_pdf <- function(x,m,s,v) {
  return(dt((x-m)/sqrt(s),df=v)/sqrt(s))
}
# Student T CDF
studentT_cdf <- function(x,m,s,v) {
  return(pt((x-m)/sqrt(s),df=v))
}
```

... your answer

Below try to address the task by creating R chunks, and running them, and adding comments tot he document.

```
# write your R code
```

Task

- Compute the MC approximate of the predictive expected value of the next outcome y_{n+1}

$$\begin{aligned}
\mathrm{E}_f(y_{n+1}|y_{1:n}) =& \int y_{n+1} f(y_{n+1}|y_{1:n}) \mathrm{d}y_{n+1} \\
=& \int y_{n+1} \left(\int f(y_{n+1}|\mu,\sigma^2)\pi(\mu,\sigma^2|y_{1:n})\mathrm{d}\mu\mathrm{d}\sigma^2 \right) \mathrm{d}y_{n+1} \\
=& \int \left(\int y_{n+1} f(y_{n+1}|\mu,\sigma^2)\mathrm{d}y_{n+1} \right) \pi(\mu,\sigma^2|y_{1:n})\mathrm{d}\mu\mathrm{d}\sigma^2 \\
=& \mathrm{E}_\pi(\mathrm{E}_f(y_{n+1}|\mu,\sigma^2)|y_{1:n}) \\
=& \mathrm{E}_\pi(\mu|y_{1:n}) \\
\approx& \frac{1}{N}\sum_{i=1}^{N} \mu^{(i)}
\end{aligned}$$

- Compute the exact predictive expected value of the next outcome y_{n+1} which is

$$\begin{aligned}
\mathrm{E}_f(y_{n+1}|y_{1:n}) =& \mathrm{E}_\pi(\mathrm{E}_f(y_{n+1}|\mu,\sigma^2)|y_{1:n}) \\
=& \mathrm{E}_\pi(\mu|y_{1:n}) \\
=& \mu_n
\end{aligned}$$

. . . your answer

Below try to address the task by creating R chunks, and running them, and adding comments tot he document.

```r
# write your R code
```

Task

- Compute the MC approximate of the predictive variance of the next outcome y_{n+1}

$$\begin{aligned}
\mathrm{Var}_f(y_{n+1}|y_{1:n}) &= \mathrm{E}_\pi\left(\mathrm{Var}_f\left(y_{n+1}|\mu,\sigma^2\right)\right) + \mathrm{Var}_\pi\left(\mathrm{E}_f\left(y_{n+1}|\mu,\sigma^2\right)\right) \\
&= \mathrm{E}_\pi\left(\sigma^2|y_{1:n}\right) + \mathrm{Var}_\pi\left(\mu|y_{1:n}\right) \\
&\approx \frac{1}{N}\sum_{j=1}^{N}\left(\sigma^{(j)}\right)^2 + \frac{1}{N}\sum_{j=1}^{N}\left(\left(\mu^{(j)}\right)^2\right) - \left(\frac{1}{N}\sum_{j=1}^{N}\mu^{(j)}\right)^2
\end{aligned}$$

- Compute the exact predictive expected value of the next outcome y_{n+1} which is

$$\mathrm{Var}_f(y_{n+1}|y_{1:n}) = \frac{\lambda_n b_n}{(\lambda_n+1)(a_n-1)}$$

where

$$
\begin{aligned}
\mu_n =& \frac{\lambda_0 \mu_n + n\bar{x}}{\lambda_0 + n} \\
\lambda_n =& \lambda_0 + n, \\
a_n =& a_0 + \frac{n}{2} \\
b_n =& b_0 + \frac{1}{2} n s^2 + \frac{1}{2}(\lambda_0 + n)^{-1} \lambda_0 n (\mu_0 - \bar{x})^2
\end{aligned}
$$

because

$$
y_{n+1} | y_{1:n} \sim \mathrm{St}(\mu_n, \frac{\lambda_n b_n}{(\lambda_n + 1) a_n}, 2a_n)
$$

… your answer

Below try to address the task by creating R chunks, and running them, and adding comments tot he document.

```
# write your R code
```