Solution to computer exam in Bayesian learning

Bertil Wegmann

2022-08-17

First load all the data into memory by running the R-file given at the exam

```
rm(list=ls())
source("ExamData.R")
set.seed(1)
```

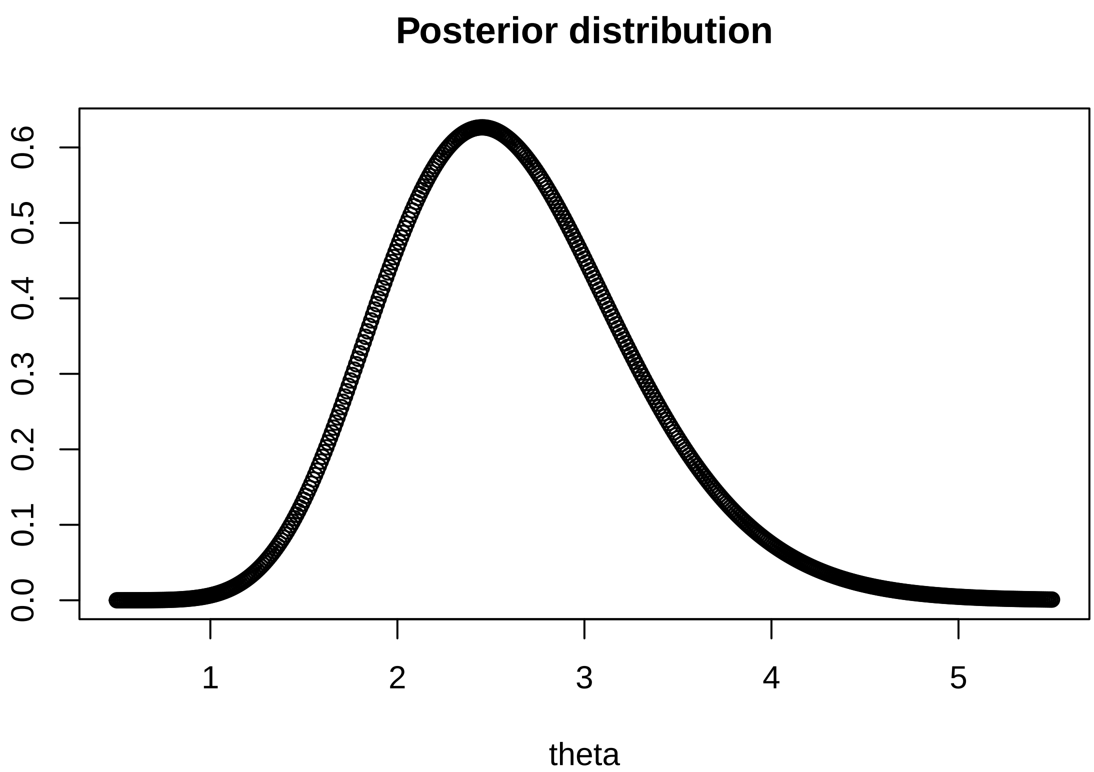
Problem 1

1d

```
LogPost <- function(theta,n,Sum_inv_x){

  logLik <- 3*n*log(theta) -  Sum_inv_x*theta;
  logPrior <- 3*log(theta) - 2*theta;

  return(logLik + logPrior)
}
theta_grid <- seq(0.5,5.5,0.01)
x_vals <- c(0.7,1.1,0.9,1.5)
Sum_x_inv <- sum(1/x_vals)
PostDens_propto <- exp(LogPost(theta_grid,4,Sum_x_inv))
PostDens <- PostDens_propto/(0.01*sum(PostDens_propto))
plot(theta_grid,PostDens,main="Posterior distribution",xlab="theta", ylab="")
```

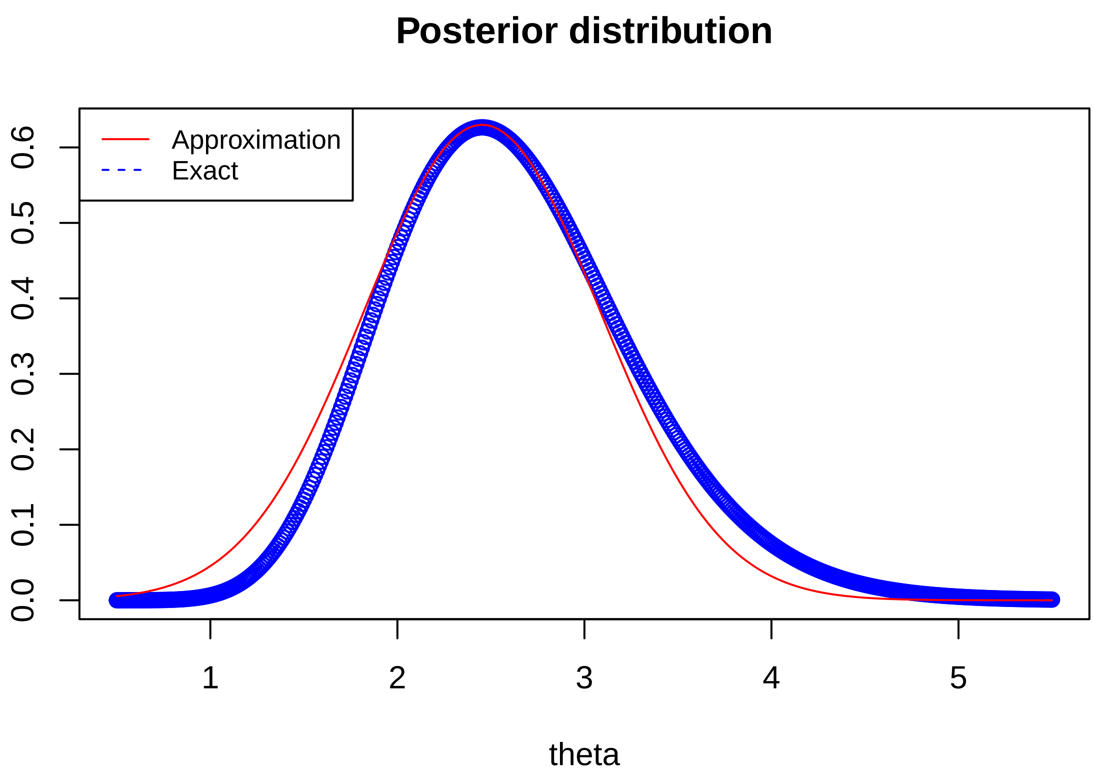


The posterior distribution is given above.

1e

```
n <- 4
OptRes <- optim(2,LogPost,gr=NULL,n,Sum_x_inv,method=c("L-BFGS-B"),lower=0.1,
                control=list(fnscale=-1),hessian=TRUE)

plot(theta_grid,PostDens,col="blue",main="Posterior distribution",xlab="theta", ylab="")
lines(theta_grid,dnorm(theta_grid,mean = OptRes$par,sd = sqrt(-1/OptRes$hessian)),col="red")
legend("topleft", legend=c("Approximation", "Exact"), col=c("red", "blue"), lty=1:2, cex=0.8)
```



The posterior approximation is quite accurate, but the exact posterior distribution is skewed to the right.

Problem 2

2a

```
mu_0 <- as.vector(rep(0,3))
Sigma_0 <- 10**2*diag(3)
nIter <- 20000

X <- as.matrix(X)
PostDraws <- BayesLogitReg(y, X, mu_0, Sigma_0, nIter)
Betas <- PostDraws$betaSample
quantile(Betas[,2],probs=c(0.025,0.975))
```

```
##        2.5%      97.5%
## 0.01429131 0.18208566
```

It is 95 % posterior probability that beta_1 is on the interval (0.014,0.182).

2b

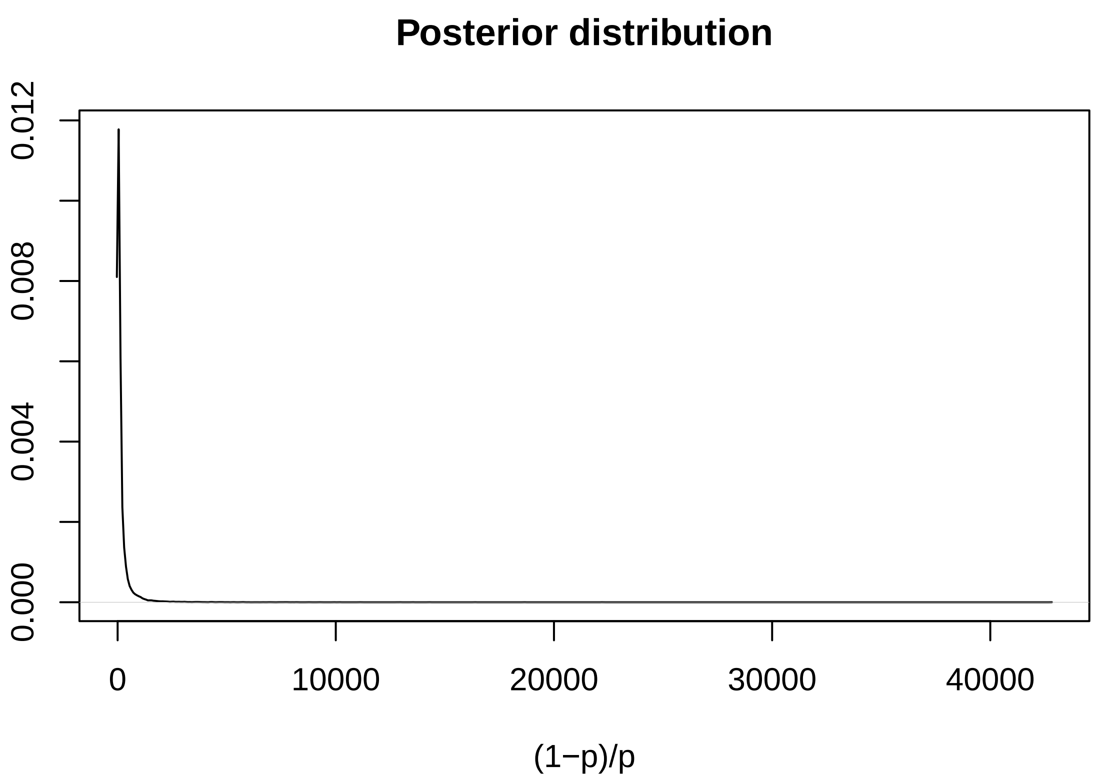
```
mean(Betas[,2]>0 & Betas[,3]>0)
```

```
## [1] 0.9145
```

The joint posterior probability that both beta_1 and beta_2 are positive is roughly 0.91.

2c

```
x_obs <- as.vector(c(1,5,1))
lin_pred <- Betas%*%x_obs
p_i <- exp(lin_pred)/(1+exp(lin_pred))
inv_odds <- (1-p_i)/p_i
plot(density(inv_odds),main="Posterior distribution",xlab="(1-p)/p", ylab="")
```




```
min(X[,2])
```

```
## [1] 19
```

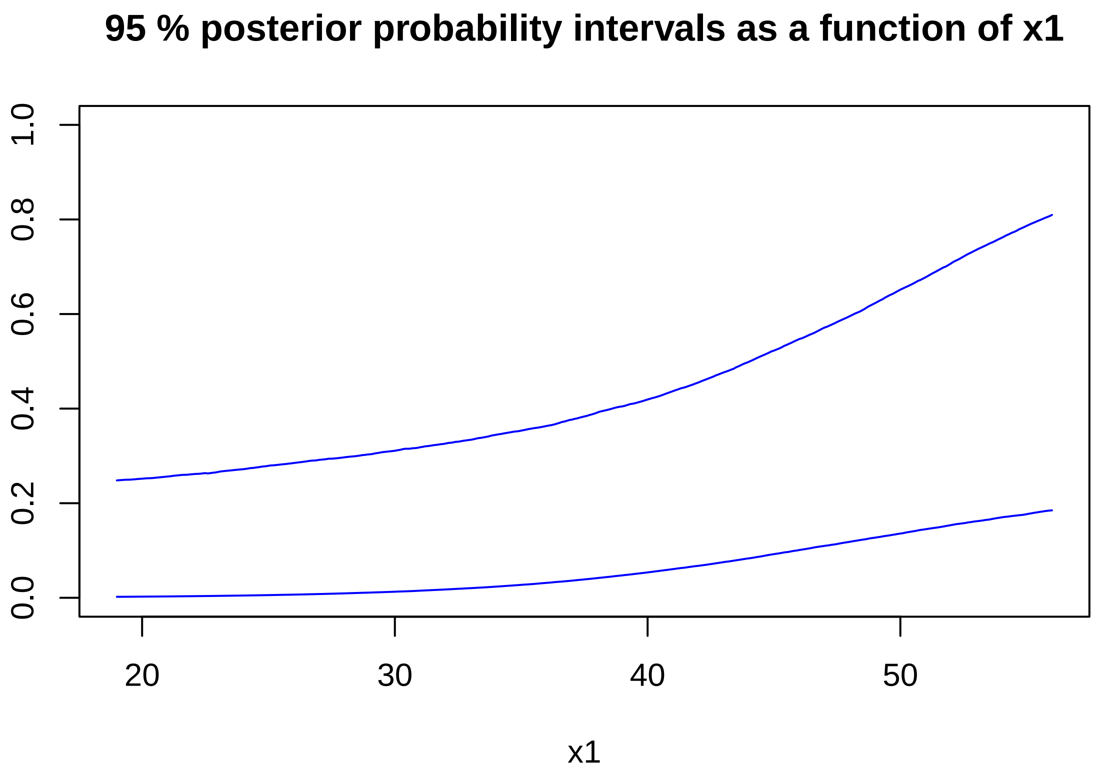
The posterior distribution of the odds of not repairing the bridge is given above. Yes, it is reasonable with very large values of this odds as this bridge is built very recently. Yes, we should question the reliability of these results because a five-year-old bridge is much newer compared to the youngest bridge of 19 years in the data.

2d

```
x1_grid <- seq(min(X[,2]),max(X[,2]),0.1)
p_i_draws <- matrix(0,length(x1_grid),2)
for (ii in 1:length(x1_grid)){
  lin_pred <- Betas%*%as.vector(c(1,x1_grid[ii],0))
  Curr_p_i <- exp(lin_pred)/(1+exp(lin_pred))
  p_i_draws[ii,] <- quantile(Curr_p_i,probs=c(0.025,0.975))
}

plot(x1_grid,p_i_draws[,1],"n",
     main="95 % posterior probability intervals as a function of x1",
     xlab="x1", ylab="",ylim=c(0,1))
lines(x1_grid,p_i_draws[,1],col="blue")
```

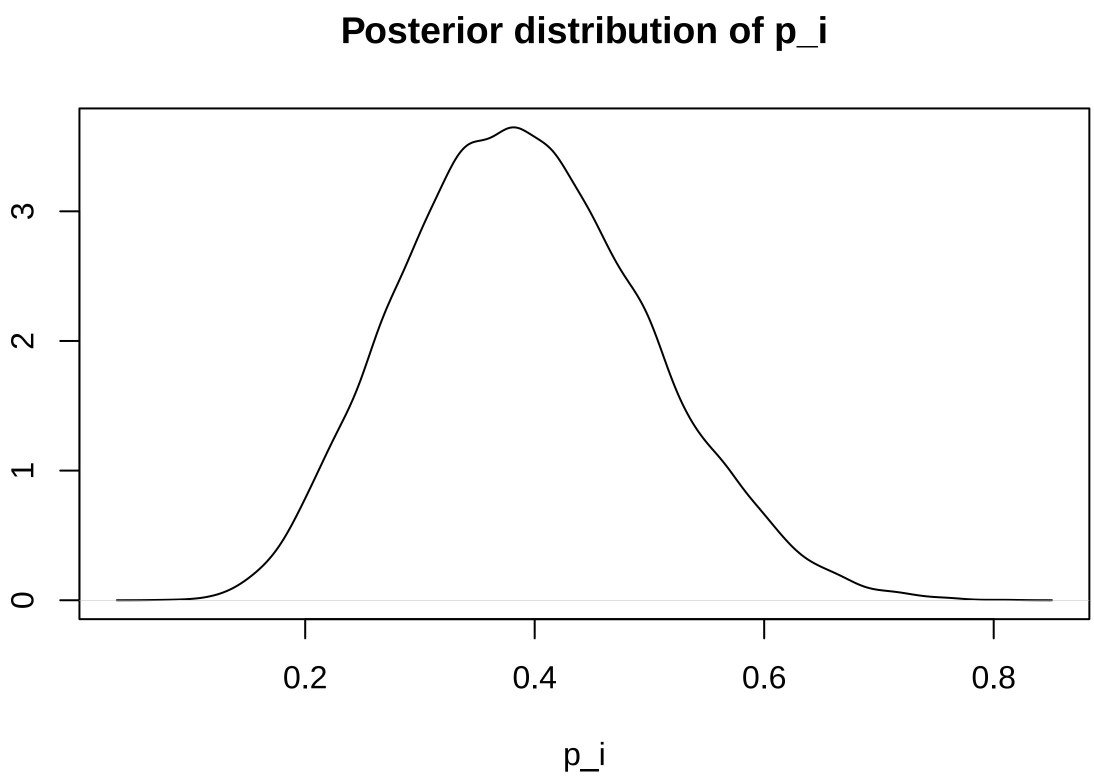
```
lines(x1_grid,p_i_draws[,2],col="blue")
```



The posterior probability intervals as a function of x1 are plotted above.

2e

```
x_obs <- as.vector(c(1,40,1))
lin_pred <- Betas%*%x_obs
p_i <- exp(lin_pred)/(1+exp(lin_pred))
plot(density(p_i),type="l",main="Posterior distribution of p_i",xlab="p_i",ylab="")
```




```
mean(p_i>0.5)
```

```
## [1] 0.15845
```

The posterior distribution of p_i is plotted above. The posterior probability is roughly 0.16.

Problem 3

3c

The expected utility when buying the option is

```
theta_51 <- 19/30
EUbuy = theta_51*60 + (1-theta_51)*(-20)
print(EUbuy)
```

```
## [1] 30.66667
```

The expected utility when not buying the option is

```
EUnotbuy = theta_51*180 + (1-theta_51)*(-240)
print(EUnotbuy)
```

```
## [1] 26
```

Since the expected utility when buying the option is higher (30.7 compared to 26), the bank should buy the option.