Solution to computer exam in Bayesian learning

Bertil Wegmann

2022-10-20

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)
```

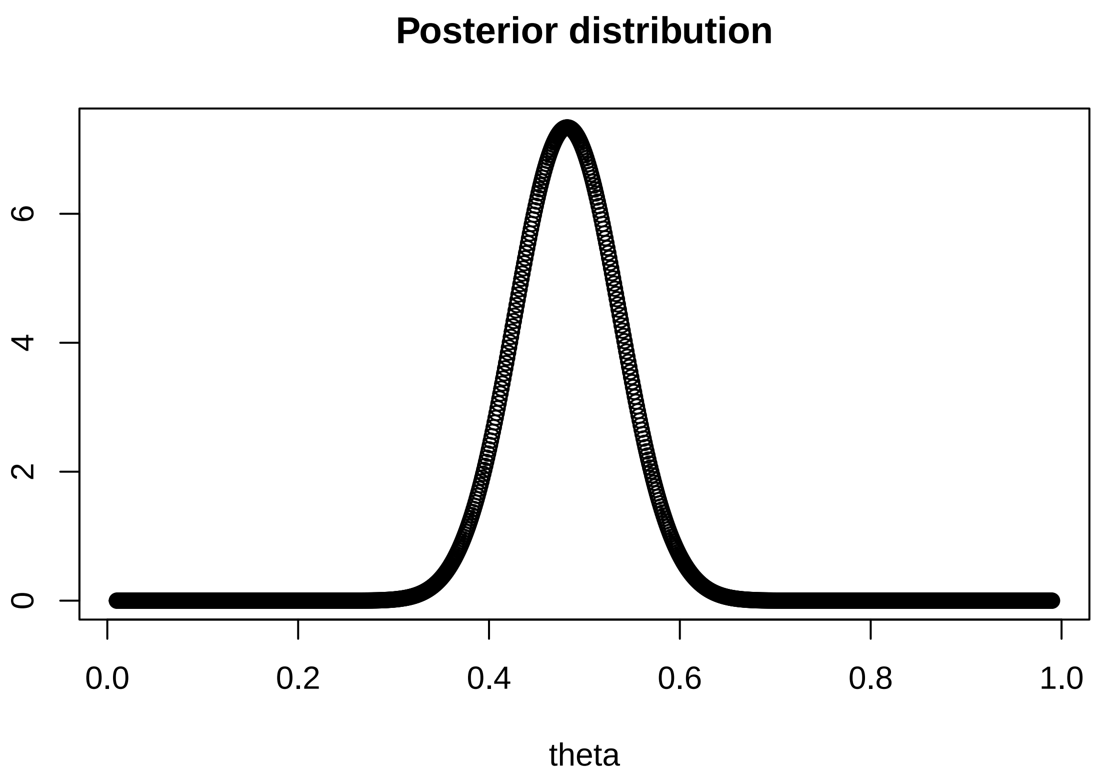
Problem 1

1d

```
LogPost <- function(theta,N,n,Sum_x){

  logLik <- Sum_x*log(theta) +  (n*N-Sum_x)*log(1-theta);
  logPrior <- log(theta) + 2*log(1-theta);

  return(logLik + logPrior)
}
theta_grid <- seq(0.01,0.99,0.001)
x_vals <- c(13,8,11,7)
Sum_x <- sum(x_vals)
PostDens_propto <- exp(LogPost(theta_grid,20,4,Sum_x))
PostDens <- PostDens_propto/(0.001*sum(PostDens_propto))
plot(theta_grid,PostDens,main="Posterior distribution",xlab="theta", ylab="")
```

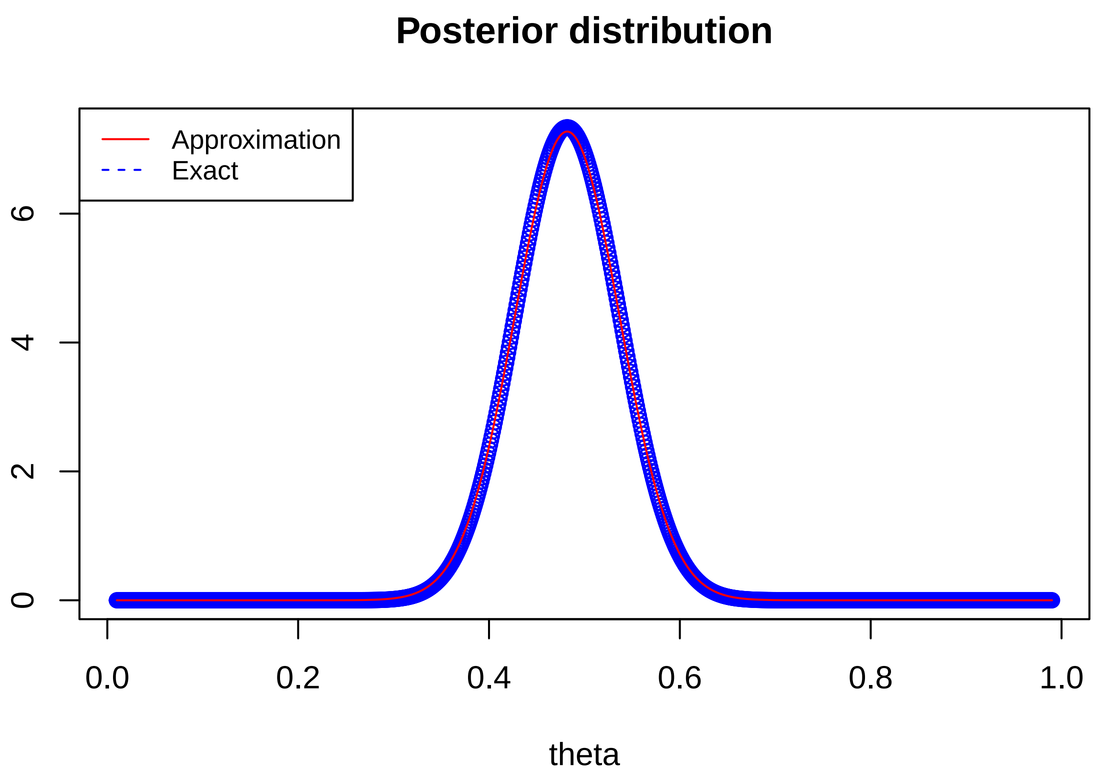


The posterior distribution is given above.

1e

```
N <- 20
n <- 4
OptRes <- optim(0.5,LogPost,gr=NULL,N,n,Sum_x,method=c("L-BFGS-B"),lower=0.2,
                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)
```

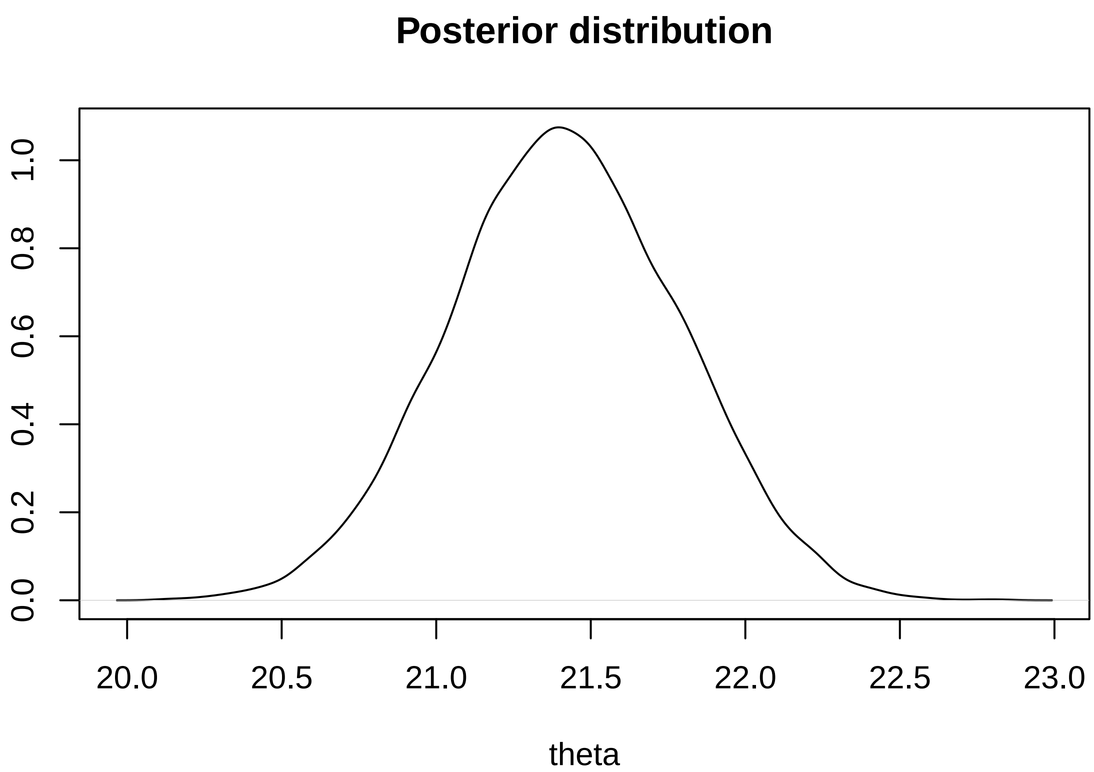


The posterior approximation is very accurate.

Problem 2

2b

```
alpha <- sum(Wolves$y) + 40
n <- length(Wolves$y)
beta <- n + 2
Thetas <- rgamma(1e4,alpha,beta)
plot(density(Thetas),main="Posterior distribution",xlab="theta", ylab="")
```




```
mean(Thetas<21)
```

```
## [1] 0.1319
```

The posterior probability that theta is smaller than 21 is roughly 0.13.

2c

```
Wolves_A <- Wolves$y[Wolves$x==1]
alpha_A <- sum(Wolves_A) + 40
n_A <- length(Wolves_A)
beta_A <- n_A + 2
Thetas_A <- rgamma(1e4,alpha_A,beta_A)
Wolves_B <- Wolves$y[Wolves$x==0]
alpha_B <- sum(Wolves_B) + 40
n_B <- length(Wolves_B)
beta_B <- n_B + 2
Thetas_B <- rgamma(1e4,alpha_B,beta_B)
y_A <- rpois(1e4,Thetas_A)
y_B <- rpois(1e4,Thetas_B)
mean(y_B>y_A)
```

```
## [1] 0.6885
```

The posterior probability is roughly 0.68.

2d

```
mean(Thetas_B>(1.10*Thetas_A))
```

```
## [1] 0.9859
```

Yes, I find the statement reasonable because the probability is roughly 99 % that the average number of wolves per week in region B is at least 10 % more than the average number of wolves per week in region A.

Problem 3

3a

```
# Prior
nCovs = dim(X)[2]
mu_0 = rep(0,nCovs)
Omega_0 = (1/100)*diag(nCovs)
v_0 = 1
sigma2_0 = 4^2

BostonRes <-  BayesLinReg(y, X, mu_0, Omega_0, v_0, sigma2_0, nIter = 10000)
Bmedian = apply(BostonRes$betaSample,2,median)
Bq025 = apply(BostonRes$betaSample,2,quantile,.025)
Bq975 = apply(BostonRes$betaSample,2,quantile,.975)
print(data.frame(round(cbind(Bmedian,Bq025,Bq975),3)),row.names=covNames)
```

```
##           Bmedian   Bq025   Bq975
## intercept  36.000  26.091  45.683
## crim       -0.108  -0.173  -0.044
## zn          0.047   0.020   0.074
## indus       0.020  -0.099   0.139
## chas        2.699   1.026   4.389
## nox       -17.544 -24.847 -10.146
## rm          3.821   3.025   4.631
## age         0.001  -0.026   0.026
## dis        -1.468  -1.860  -1.072
## rad         0.305   0.176   0.433
## tax        -0.012  -0.019  -0.005
## ptratio    -0.944  -1.200  -0.692
## black       0.009   0.004   0.015
## lstat      -0.524  -0.623  -0.426
```

It is 95 % posterior probability that the regression coefficient beta_1 for the variable crim is on the interval (-0.17,-0.04). It is therefore a high probability that crim affects the house prices negatively.

3b

```
Sigma2 <- BostonRes$sigma2Sample
mean(sqrt(Sigma2))
```

```
## [1] 4.689099
```

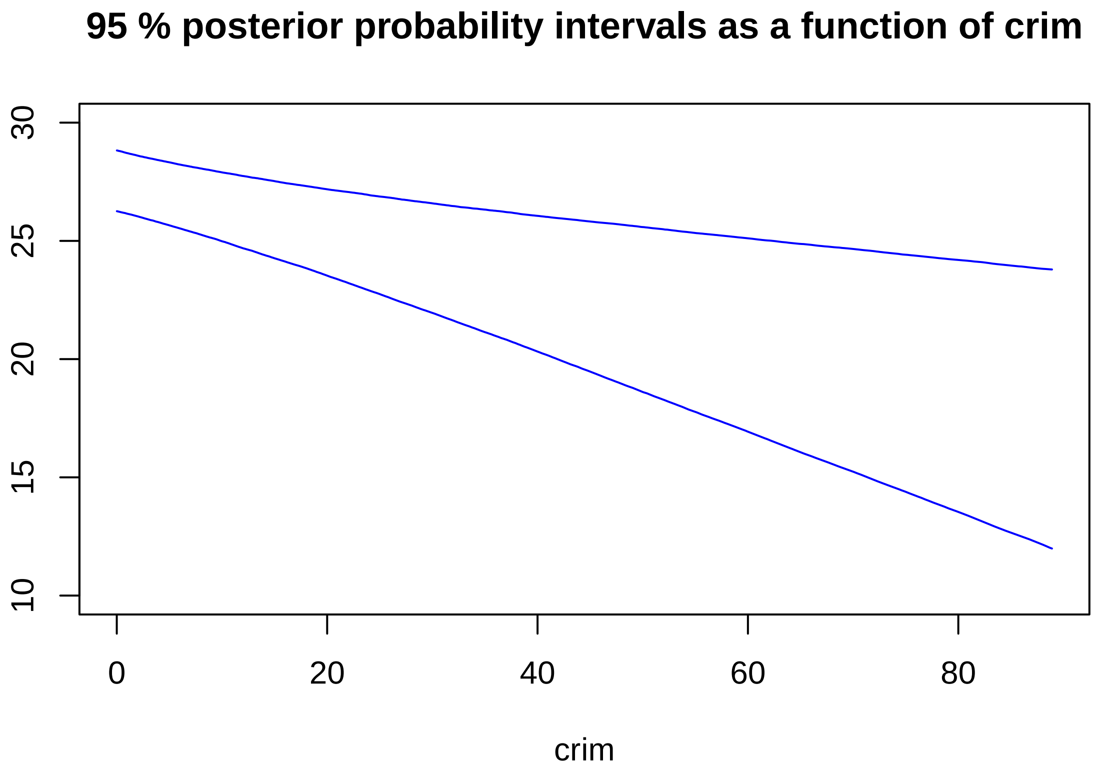
```
median(sqrt(Sigma2))
```

```
## [1] 4.685504
```

3c

```
x1_grid <- seq(min(X[,2]),max(X[,2]),0.1)
Mu_draws <- matrix(0,length(x1_grid),2)
for (ii in 1:length(x1_grid)){
```

```r
  CurrMu <- BostonRes$betaSample %*% c(1,x1_grid[ii],XNewHouse[-1:-2])
  Mu_draws[ii,] <- quantile(CurrMu,probs=c(0.025,0.975))
}
plot(x1_grid,Mu_draws[,1],"n",main="95 % posterior probability intervals as a function of crim",
     xlab="crim", ylab="",ylim=c(10,30))
lines(x1_grid,Mu_draws[,1],col="blue")
lines(x1_grid,Mu_draws[,2],col="blue")
```

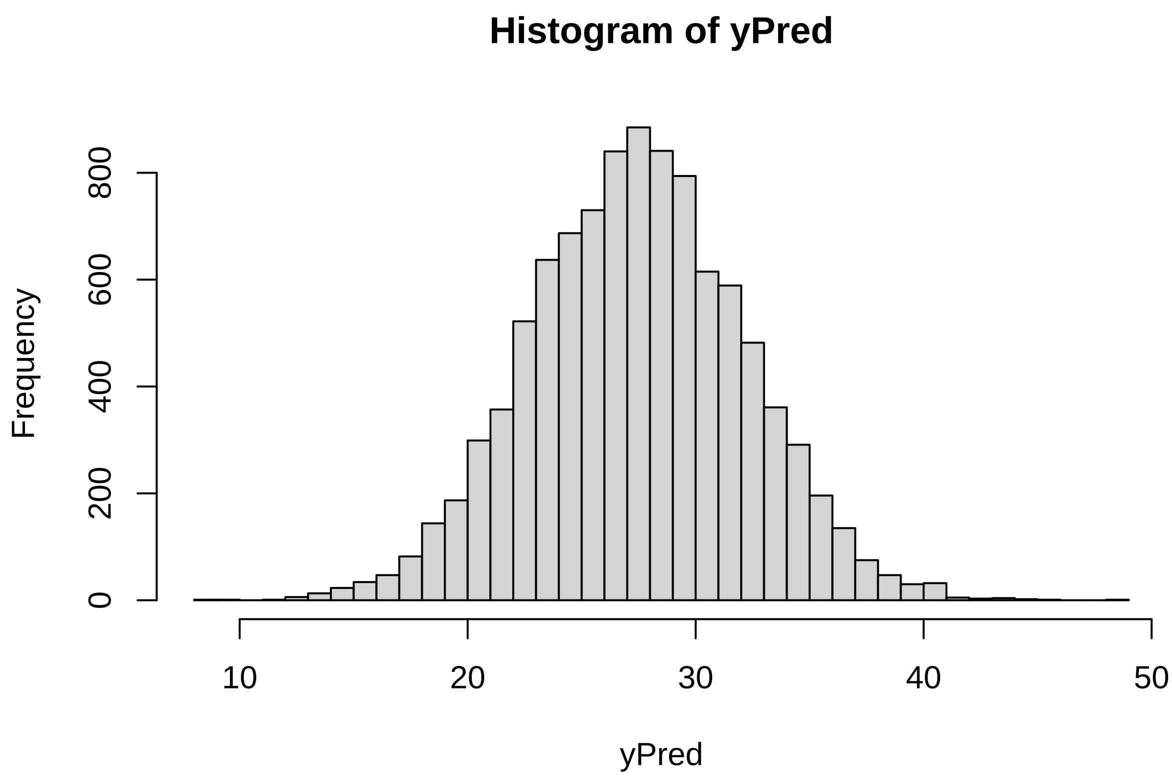


The limits of the posterior probability intervals as a function of crim are plotted above.

3d

Here, we need the predictive distribution for the price, given the explanatory variables. We can obtain the predictive distribution by simulation.

```r
nSim <- dim(BostonRes$betaSample)[1] # One predictive draw for each posterior parameter draw
yPred <- matrix(0,nSim,1)
for (i in 1:nSim){
  Mu_curr <- XNewHouse%*%BostonRes$betaSample[i,]
  yPred[i] <- Mu_curr + rnorm(n = 1, mean = 0, sd = sqrt(BostonRes$sigma2Sample[i]))
}
par(mfrow=c(1,1))
hist(yPred,50)
```




```
sum(yPred>=20)/nSim
```

```
## [1] 0.9461
```

Probability of getting \$20000 is quite large (0.94), so the construction project is probably a good idea.

3e

```
T_y <- max(y)
T_y_rep <- matrix(0,1e4,1)
Mu <- BostonRes$betaSample %*% t(X)
Sigma <- sqrt(BostonRes$sigma2Sample)
for (ii in 1:1e4){
  y_Vals <- rnorm(length(y),Mu[ii,],Sigma[ii])
  T_y_rep[ii,1] <- max(y_Vals)
}
mean(T_y_rep >= T_y)
```

```
## [1] 0.4637
```

The posterior predictive p-value is 0.47, which is quite close to 0.5. Hence, the model can replicate the value of the most expensive house well in this data.