BayesExam

Christophoros Spyretos

2022-10-17

Problem 1

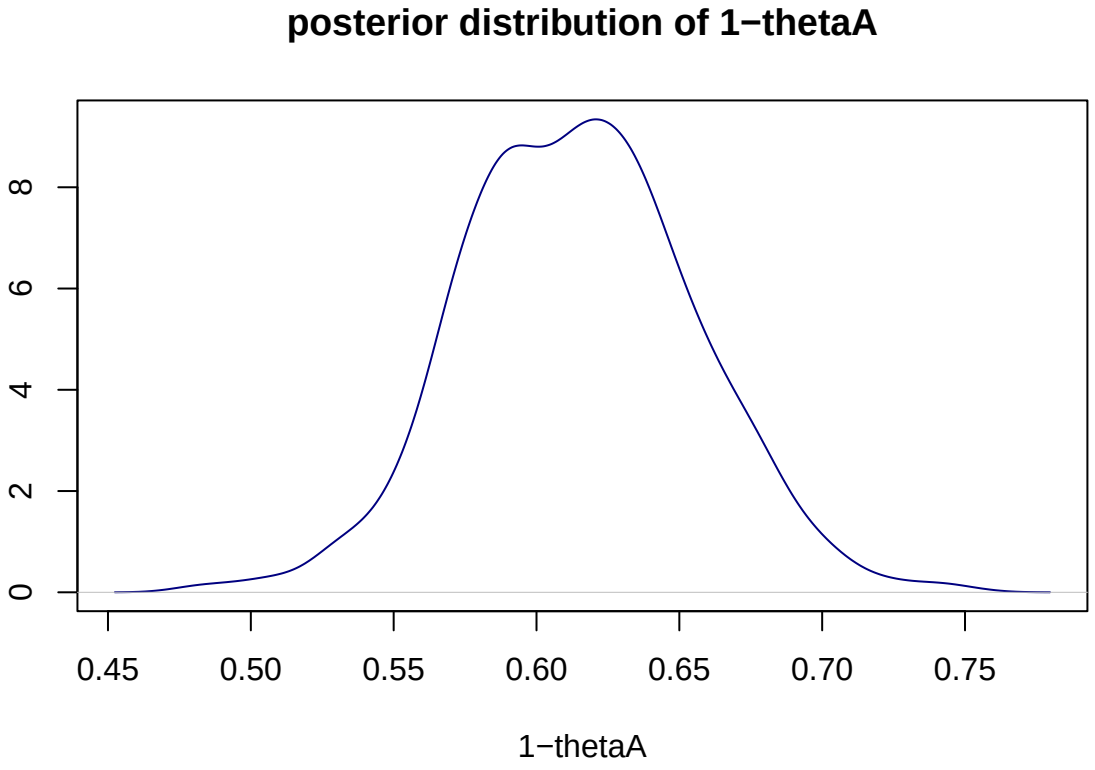
Task a

```
set.seed(12345)

n <- 100
s <- 38
f <- n-s
alpha <- 16
beta <- 24
nSim <- 1000

thetaA <- rbeta(nSim, shape1 = alpha+s, shape2 = beta+f)

plot(density(1-thetaA), col = "navy",
     main = "posterior distribution of 1-thetaA",
     xlab = "1-thetaA", ylab="")
```




```
prob1a <- pbeta(0.4, shape1 = alpha+s, shape2 = beta+f, lower.tail = FALSE)
```

The posterior probability that $\theta_A > 0.4$ is approximately 0.36.

Task b

```
ratio <- (1-thetaA)/thetaA

interval <- quantile(ratio, probs = c(0.025,0.975))

interval <- data.frame(interval = interval[1], upper_bound = interval[2])
colnames(interval) <- c("lower bound", "upper bound")
rownames(interval) <- c("95% Equal Tail Credible Interval")
knitr::kable(interval)
```

	lower bound	upper bound
95% Equal Tail Credible Interval	1.14581	2.272752

The ratio shows the odds of not selecting the brand A. The 95% posterior probability that the ratio takes values from the interval.

Task d

```
beta(alpha+s,beta+f)/beta(alpha,beta)
```

```
## [1] 7.556771e-30
```

Task d

```
y <- c(38,27,35)
alpha <- c(1,1,1)
constant <- 10
alpha <- alpha*constant
x <- matrix(0,nSim,3)
thetas <- matrix(0,nSim,3)

for (i in 1:3) {
  x[,i] <- rgamma(nSim,shape = alpha[i] + y[i], rate = 1 )
}

for (i in 1:nSim) {
  thetas[i,] <- x[i,]/sum(x[i,])
}

prob1d <- mean(thetas[,1] > thetas[,3])
```

The posterior probability that $\theta_A > theta_C$ is approximately 0.642.

Problem 2

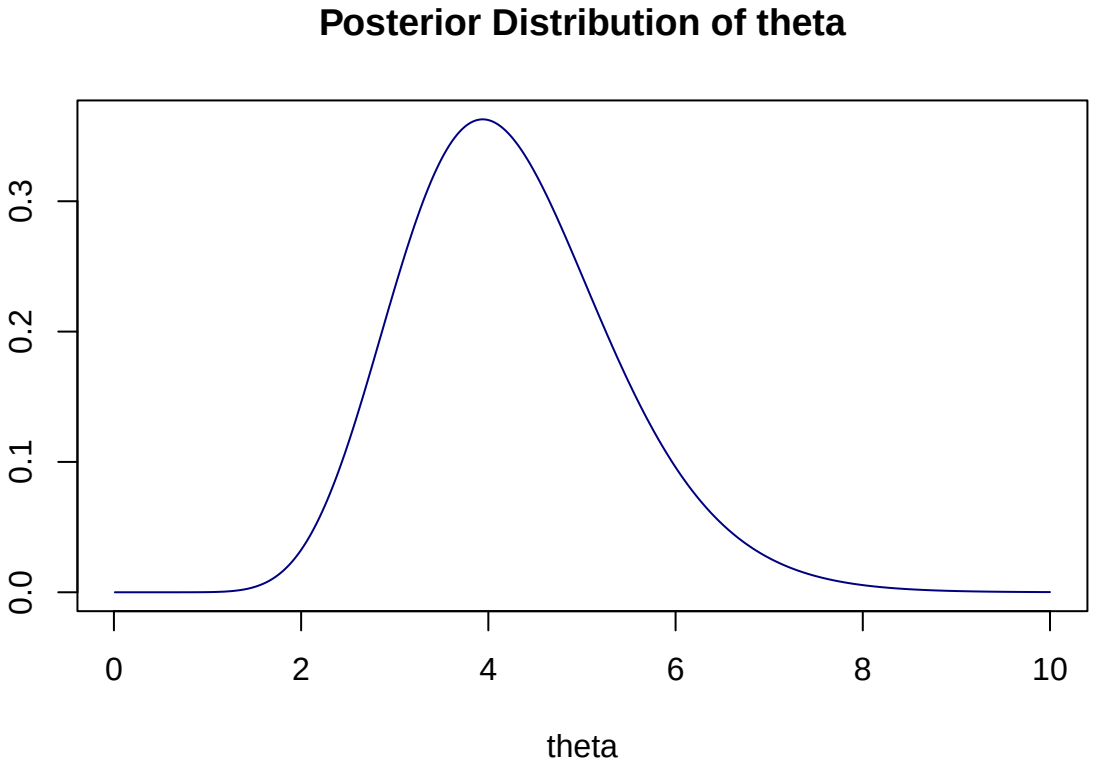
Task d

```
LogPost <- function(theta,n,sumx2){
  res <- n*log(theta) - theta*(0.5 + sumx2)
  return(res)
}

thetaGrid <- seq(0.01,10,0.01)
n <- 13
sumx2 <- 2.8

PostDens_propto <- exp(LogPost(thetaGrid,n,sumx2))
PostDens <- PostDens_propto/(0.01*sum(PostDens_propto))

plot(thetaGrid, PostDens, type = "l", col = "navy",
     main = "Posterior Distribution of theta",
     xlab = "theta", ylab = "")
```

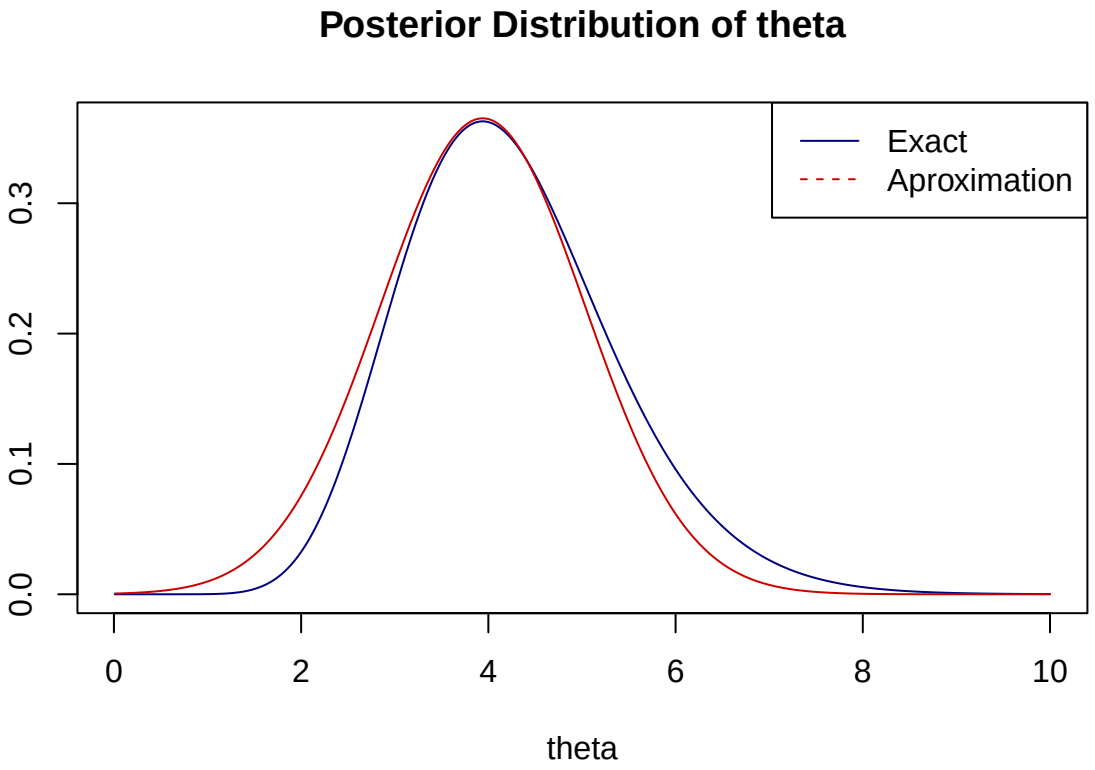


Task e

```
OptimRes <- optim(3, LogPost, gr = NULL, n, sumx2,
                  method = c("L-BFGS-B"),lower = 0.1,
                  control = list(fnscale = -1),
                  hessian = TRUE)

approx <- dnorm(thetaGrid, mean = OptimRes$par, sd = sqrt(diag(-solve(OptimRes$hessian))))

plot(thetaGrid, PostDens, type = "l", col = "navy",
     main = "Posterior Distribution of theta",
     xlab = "theta", ylab = "")
lines(thetaGrid,approx,type = "l", col = "red3")
legend("topright", legend = c("Exact","Aproximation"),
       col = c("navy","red3"), lty = 1:2)
```



The posterior approximation is accurated, but the exact posterior is slightly skewed to the right.

Problem 3

```
source("ExamData.R")
```

Task a

```
set.seed(12345)

library("mvtnorm")

mu_0 <- as.vector(rep(0,7))
Omega_0 <- (1/25)*diag(7)
v_0 <- 1
sigma2_0 <- 4
nIter <- 10000

PostDraws <- BayesLinReg(y, X, mu_0, Omega_0, v_0, sigma2_0, nIter)
```

```
Betas <- PostDraws$betaSample

MeanBetas <- colMeans(Betas)

MeanBetas <- as.data.frame(MeanBetas)
colnames(MeanBetas) <- c("Mean Value")
rownames(MeanBetas) <- c("Beta0","Beta1","Beta2", "Beta3",
                         "Beta4","Beta5","Beta6")
```

```
knitr::kable(MeanBetas)
```

	Mean Value
Beta0	1.3082564
Beta1	0.7010702
Beta2	0.1574445
Beta3	0.4236033
Beta4	-0.1626155
Beta5	0.0742226
Beta6	-0.2399414

```
BetasIntervals <- matrix(0,7,2)

for (i in 1:7) {
  BetasIntervals[i,] <- quantile(Betas[i,], probs = c(0.025,0.975))
}

MeanBetas <- as.data.frame(BetasIntervals)
colnames(BetasIntervals) <- c("Lower Bound","Upper Bound")
rownames(BetasIntervals) <- c("Beta0","Beta1","Beta2", "Beta3",
                          "Beta4","Beta5","Beta6")
knitr::kable(BetasIntervals)
```

	Lower Bound	Upper Bound
Beta0	-0.3338755	1.301628
Beta1	-0.4234871	1.316765
Beta2	-0.4111521	1.203100
Beta3	-0.2674153	1.306349
Beta4	-0.1292668	1.206165
Beta5	-0.1660699	1.162340
Beta6	-0.2076466	1.325194

The 95% posterior probability that β_1 takes values from the interval (-0.423,1.31).

Task b

```
SD <- sqrt(PostDraws$sigma2Sample)
MedianSD <- median(SD)
```

The posterior median of the standard deviation is approximately 0.639.

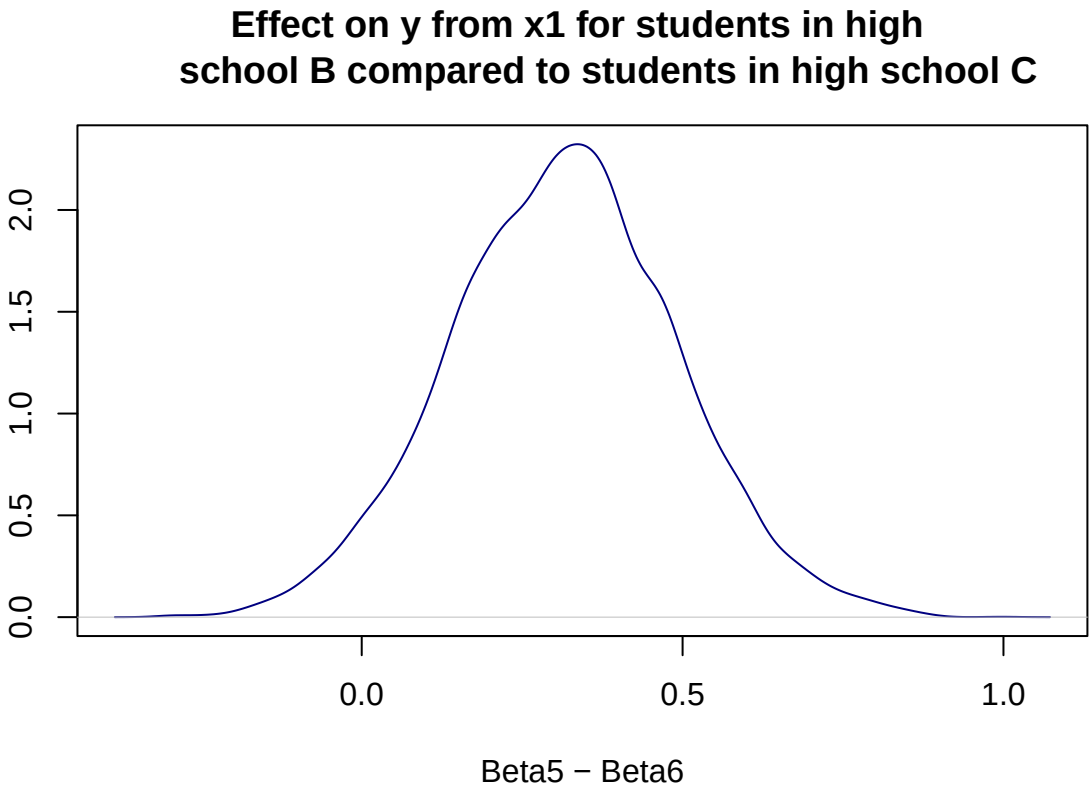
Task c

```
Effect_B <- Betas[,6]
Effect_C <- Betas[,7]

Diff <- Effect_B - Effect_C

plot(density(Diff), type = "l", col = "navy",
     main = "Effect on y from x1 for students in high
```

```
     school B compared to students in high school C",
     xlab = "Beta5 - Beta6", ylab = "")
```




```
EffectInterval <- quantile(Diff, probs = c(0.025,0.975))

EffectInterval <- as.data.frame(t(EffectInterval))
colnames(EffectInterval) <- c("Lower Bound","Upper Bound")
rownames(EffectInterval) <- c("95% Equal Tail Credible Interval")
knitr::kable(EffectInterval)
```

	Lower Bound	Upper Bound
95% Equal Tail Credible Interval	-0.0339968	0.665147

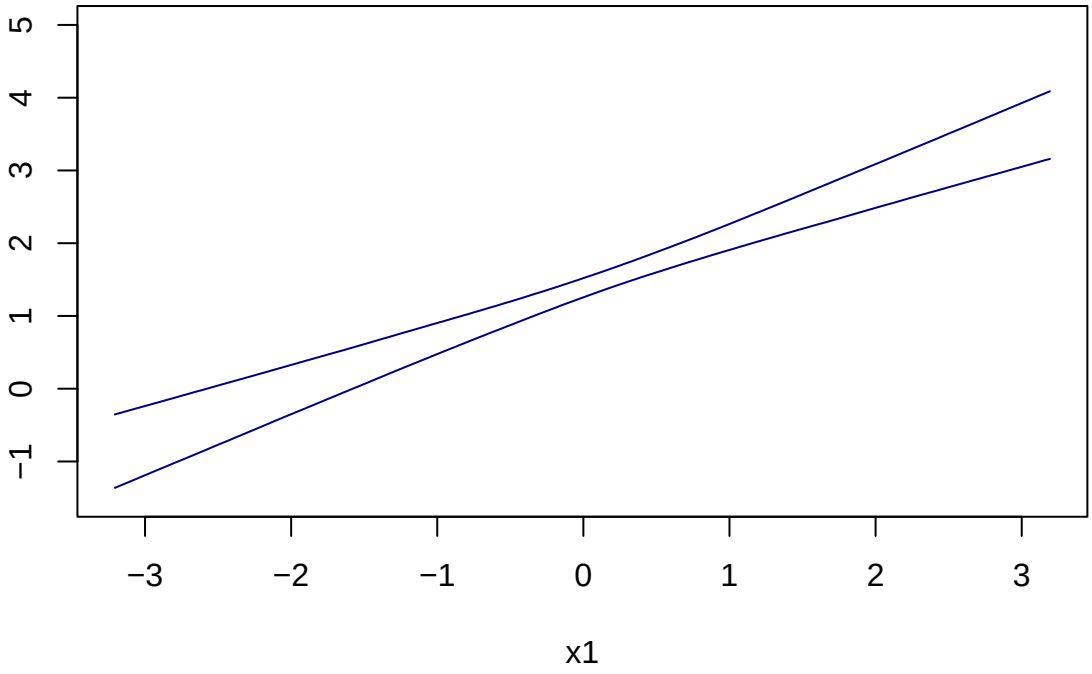
There is substantial mass probability that the effect on y from x1 for students in high school B is higher compared to students in high school C. However, the 95% equal tail credible interval shows that the difference takes negative or positive values. Hence, the probability is not that high that this effect in high school B is larger than in high school C

Task d

```
x1Grid <- seq(min(X[,2]),max(X[,2]),0.01)
intervals <- matrix(0, nrow = length(x1Grid), ncol = 2)

for (i in 1:length(x1Grid)) {
  mu <- Betas[,1] + Betas[,2]*x1Grid[i] + Betas[,3]*0.5
  intervals[i,] <- quantile(mu, probs = c(0.05,0.95))
}
```

```
plot(x1Grid,intervals[,1],type = "l", col = "navy",
     main = "90% Equal Tail Posterior Probability Intervals as a Function of x1",
     xlab = "x1", ylab = "", ylim = c(-1.5,5))
lines(x1Grid,intervals[,2],type = "l", col = "navy")
```

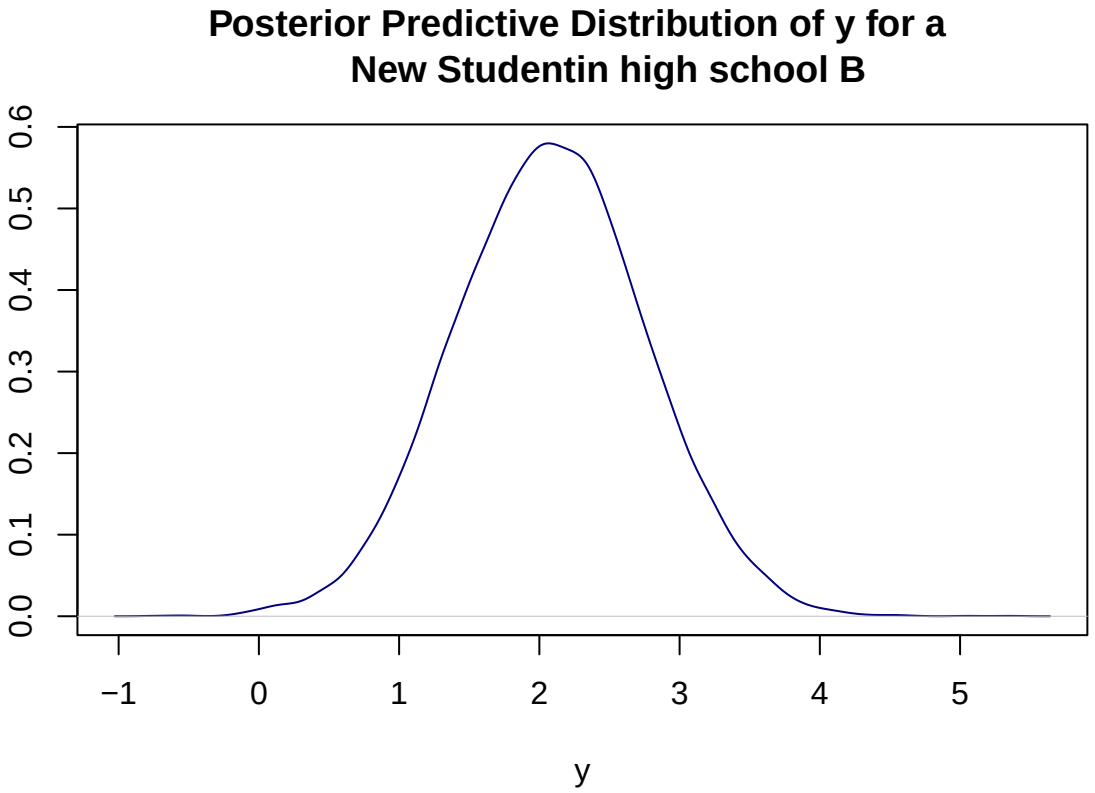


Task e

```
mu <- Betas[,1] + Betas[,2]*0.4 + Betas[,3]*1 + Betas[,4] + Betas[,7]*0.4

y_values <- rnorm(nIter, mean = mu, sd = SD)

plot(density(y_values), col = "navy",
     main = "Posterior Predictive Distribution of y for a
     New Studentin high school B",
     xlab = "y", ylab = "")
```


Posterior Predictive Distribution of y for a
New Studentin high school B
0.0
0.1
0.2
0.3
0.4
0.5
0.6
−1
0
1
2
3
4
5
y