ACST3061_Assignment2

Yunbae Chae

2023-05-19

Contents

Question 1

```r
dat <- read.csv("Assignment2_Dataset2023.csv")
dat <- data.frame(dat)
length(dat$Year)
```

```
## [1] 38
```

$N_{it},\ i = 1, 2, 3,\ t \in [1, 38]$ belong to Poisson distribution and refer to different types of hurricanes listed in the assignment prompt. The probability mass function of N_{it} is

$$p(N_{it}|\Theta_i) = \frac{\Theta_i^{N_{it}} e^{-\Theta_i}}{N_{it}!}$$

The likelihood function of Θ_i given the sample data $\{N_{i1}, N_{i2}, \cdots, N_{i38}\}$ is

$$L(\Theta_i|N_{i1}, N_{i2}, \cdots, N_{i38})$$

$$= \frac{\Theta_i^{\sum_{t=1}^{38} N_{it}} \cdot e^{-38\Theta_i}}{(\prod_{t=1}^{38} N_{it})!}$$

$$\propto \Theta_i^{\sum_{t=1}^{38} N_{it}} \cdot e^{-38\Theta_i}$$

The prior distribution of Θ_i is (Gamma distribution given from the prompt)

$$\pi(\Theta_i) = \frac{\beta^\alpha}{\Gamma(\alpha)} \Theta_i^{\alpha-1} exp(-\beta\Theta_i)$$

$$\propto \Theta_i^{\alpha-1} exp(-\beta\Theta_i)$$

$$\pi(\Theta_i|N_i) \propto L(\Theta_i|N_i) \times \pi(\Theta_i)$$

$$= \Theta_i^{\alpha+\sum_{t=1}^{38} N_{it}-1} exp(-(38+\beta)\Theta_i)$$

Then, $\alpha_n := \alpha + \sum_{t=1}^{38} N_{it}$ and $\beta_n := 38 + \beta$.

Now, α and β are known since $E[\Theta_i]$ and $Var[\Theta_i]$ are given. From the prior distribution (Gamma), the mean and variance are $\frac{\alpha}{\beta} = 4.5789$ and $\frac{\alpha}{\beta^2} = 11.6619$ respectively. Solving for them,

$$\alpha = 0.392638 \ \beta = 1.797848$$

```
alpha <- 4.5789/11.6619
beta <- 4.5789^2/11.6619
alphaPost1 <- alpha+sum(dat$Tropical.Storms)
alphaPost2 <- alpha+sum(dat$Hurricanes)
alphaPost3 <- alpha+sum(dat$Major.Hurricanes)
BetaPost <- length(dat$Year)+beta

mean1 <- alphaPost1/BetaPost
mean2 <- alphaPost2/BetaPost
mean3 <- alphaPost3/BetaPost

var1 <- mean1/BetaPost
var2 <- mean2/BetaPost
var3 <- mean3/BetaPost

mean1
```

```
## [1] 7.045422
```

```
mean2
```

```
## [1] 5.035263
```

```
mean3
```

```
## [1] 1.065199
```

```
var1
```

```
## [1] 0.1770302
```

```
var2
```

```
## [1] 0.126521
```

```
var3
```

```
## [1] 0.02676525
```

Question 2

There is no simple closed equation to find the median for a Gamma distribution. Using R function qgamma() instead, it can still be found. The following three values are medians for N_1, N_2, N_3 respectively.

```
qgamma(0.5,shape=alphaPost1,rate=BetaPost)
```

```
## [1] 7.037048
```

```
qgamma(0.5,shape=alphaPost2,rate=BetaPost)
```

```
## [1] 5.02689
```

```
qgamma(0.5,shape=alphaPost3,rate=BetaPost)
```

```
## [1] 1.056835
```

The mode for a Gamma distribution is $\frac{\alpha-1}{\beta}$.

```
(alphaPost1-1)/BetaPost
```

```
## [1] 7.020295
```

```
(alphaPost2-1)/BetaPost
```

```
## [1] 5.010136
```

```
(alphaPost3-1)/BetaPost
```

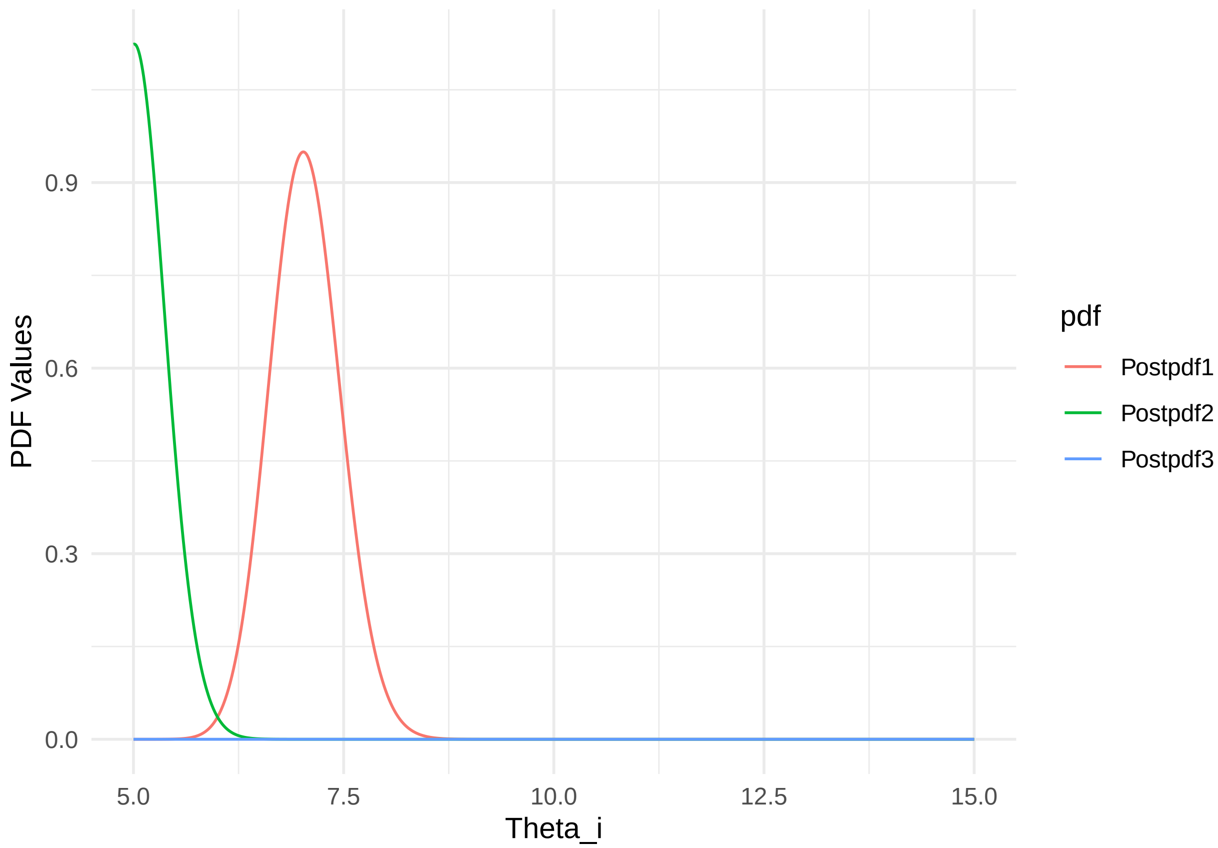
```
## [1] 1.040072
```

Question 3

```r
library(ggplot2)
library(reshape2)
```

```
## Warning: package 'reshape2' was built under R version 4.2.3
```

```r
xvalue <- seq(5,15,0.01)
Postpdf1 <- dgamma(xvalue,alphaPost1,BetaPost)
Postpdf2 <- dgamma(xvalue,alphaPost2,BetaPost)
Postpdf3 <- dgamma(xvalue,alphaPost3,BetaPost)
df <- data.frame(x = xvalue, Postpdf1 = Postpdf1, Postpdf2 = Postpdf2, Postpdf3 = Postpdf3)
df_long <- reshape2::melt(df, id.vars = "x", variable.name = "pdf", value.name = "value")
ggplot(df_long, aes(x = x, y = value, color = pdf)) +
  geom_line() +
  labs(x = "Theta_i", y = "PDF Values") +
  theme_minimal()
```



Question 4

```
library(actuar)
```

```
## Warning: package 'actuar' was built under R version 4.2.3

##
## Attaching package: 'actuar'

## The following objects are masked from 'package:stats':
##
##     sd, var

## The following object is masked from 'package:grDevices':
##
##     cm
```

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
dat <- select(dat,Tropical.Storms,Hurricanes,Major.Hurricanes)
tdat <- t(dat)
tdat <- data.frame(tdat)
tdat
```

```
##                  X1 X2 X3 X4 X5 X6 X7 X8 X9 X10 X11 X12 X13 X14 X15 X16 X17 X18
## Tropical.Storms   6  5  8  5  5  6  4  6  8   7   8   6   9   5   7   7   9   4
## Hurricanes        3  5  4  3  4  4  3  6  7   6   6   3   5   3   3   6   7   3
## Major.Hurricanes  1  1  2  1  1  2  0  0  1   1   0   0   0   0   0   1   1   0
##                  X19 X20 X21 X22 X23 X24 X25 X26 X27 X28 X29 X30 X31 X32 X33
## Tropical.Storms   10  11   8   5   5   7   6   5   8  12   8  11   7   6   4
## Hurricanes         7  10   6   4   3   4   5   4   3  10   6   9   4   4   3
## Major.Hurricanes   1   2   2   0   2   0   1   2   1   2   2   2   0   2   2
##                  X34 X35 X36 X37 X38
## Tropical.Storms    4   8  12  19   9
## Hurricanes         4   6  10  11   6
## Major.Hurricanes   1   0   4   2   2
```

```r
cm <- cm(formula="bayes",data=tdat,likelihood="poisson",shape=alpha,rate=beta)
summary(cm)
```

```
## Call:
## cm(formula = "bayes", data = tdat, likelihood = "poisson", shape = alpha,
##     rate = beta)
##
## Structure Parameters Estimators
##
##   Collective premium: 0.2183931
##
##   Between variance: 0.1214747
##   Within variance: 0.2183931
##
## Detailed premiums
##
##     Indiv. mean Weight Cred. factor Bayes premium
##     7.368421    38     0.9548255    7.045422
##     5.263158    38     0.9548255    5.035263
##     1.105263    38     0.9548255    1.065199
```

As shown in the results, the Bayesian credibility estimate for Θ_i is 7.045422, 5.035263 and 1.065199 for $i = 1, 2, 3$ respectively. It happens these numbers match the results shown in Question 1.

Question 5

According to the lecture notes, a Bayesian confidence interval can be computed using a Bayesian estimate and the standard deviation of the Bayesian posterior distribution. These are already computed and displayed in Question 1. The code below uses the already existing posterior variances computed for Question 1, with the cm() results for the Bayesian estimates.

```
Theta1 <- summary(cm)[[7]][[1]][[1]]
Theta2 <- summary(cm)[[7]][[1]][[2]]
Theta3 <- summary(cm)[[7]][[1]][[3]]

BCILower1 <- Theta1-qnorm(0.975)*sqrt(var1)
BCIUpper1 <- Theta1+qnorm(0.975)*sqrt(var1)
BCI1 <- c(BCILower1,BCIUpper1)

BCILower2 <- Theta2-qnorm(0.975)*sqrt(var2)
BCIUpper2 <- Theta2+qnorm(0.975)*sqrt(var2)
BCI2 <- c(BCILower2,BCIUpper2)

BCILower3 <- Theta3-qnorm(0.975)*sqrt(var3)
BCIUpper3 <- Theta3+qnorm(0.975)*sqrt(var3)
BCI3 <- c(BCILower3,BCIUpper3)

BCI1
```

```
## [1] 6.220768 7.870076
```

```
BCI2
```

```
## [1] 4.338108 5.732418
```

```
BCI3
```

```
## [1] 0.7445474 1.3858511
```