베이지안통계학 과제5 solution

Jieun Shin

Spring 2022

문제 6-1. (10점)

- θ의 사후분포 $Gamma(5.5, 6)$ 유도 (5점)
- 사후분포의 밀도함수 그리기 (5점)

6 페이지 중 각 페이지의 평균 오타의 수 $X_i, i = 1, \ldots, 6$은 모수가 θ인 포아송분포를 따른다. 즉, $X_1, i = 1, \ldots, 6 \sim Poi(\theta)$. 그리고 θ는 제프리 사전분포를 따른다. 즉, $\pi(\theta) = \theta^{-\frac{1}{2}}$.

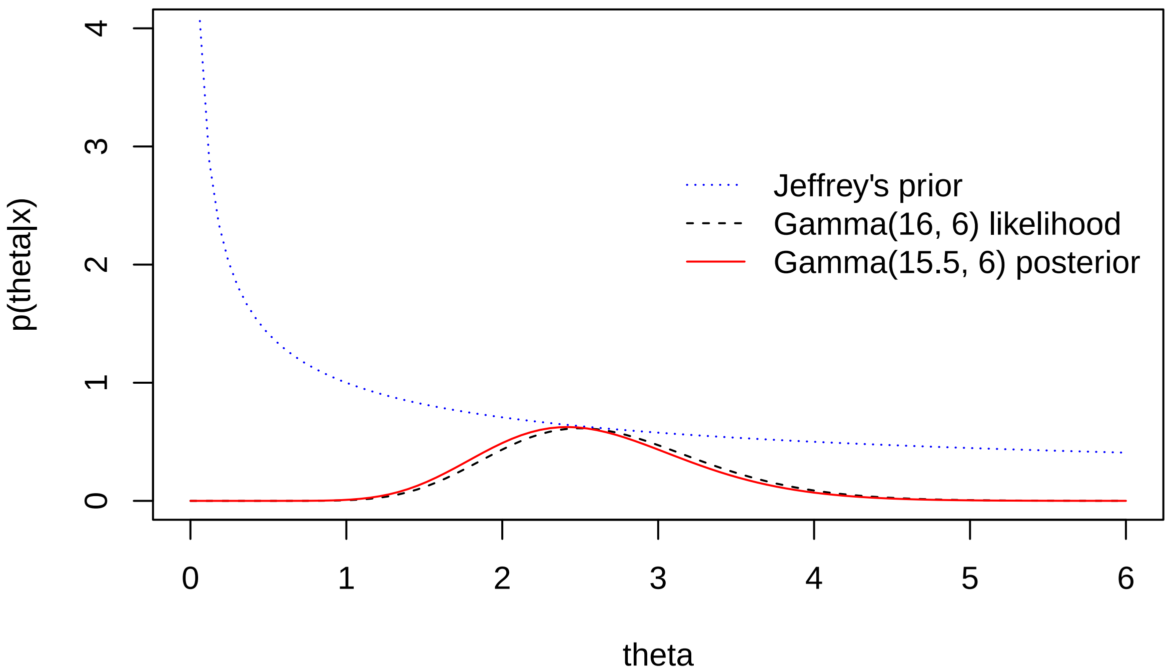
이로부터 θ의 사후분포를 다음과 같이 구할 수 있다.

$$
\begin{aligned}
\pi(\theta) &= f(x_1, \ldots, x_6; \theta)\pi(\theta) \\
&= \prod_{i=1}^{6} \frac{e^{-\theta}\theta^{x_i}}{x_i!} \cdot \theta^{-\frac{1}{2}} \\
&\propto e^{-\theta}\theta^{x_i} \cdot \theta^{-\frac{1}{2}} \\
&= e^{-6\theta} \cdot \theta^{(\sum_{i=1}^{6} x_i + \frac{1}{2}) - 1}
\end{aligned}
$$

이는 $Gamma(\sum_{i=1}^{6} x_i + \frac{1}{2} = 5.5, 6)$분포이며, 밀도함수는 다음과 같이 그려진다.

```r
theta = seq(0, 6, length = 100)
plot(theta, theta^{-1/2}, lty = 3, type = "l", xlab = "theta", ylab = "p(theta|x)",
     ylim = c(0, 4), col = "blue") # 제프리 사전분포
lines(theta, dgamma(theta, 16, 6), lty = 2, col = "black") # 가능도함수
lines(theta, dgamma(theta, 15.5, 6), col = 'red') # 사후분포

legend(3, 3, legend = c("Jeffrey's prior", "Gamma(16, 6) likelihood", "Gamma(15.5, 6) posterior"),
       lty = c(3, 2, 1), bty = "n", col = c("blue", "black", "red"))
```



문제 6-3. (15점)

(1)

- θ의 사후분포인 $Gamma(14, 11)$의 밀도함수 그리기 (그래프가 있어야 5점, 그래프가 없다면 부분점수)

하루 방문자 수 X가 평균이 θ인 포아송 분포를 따르므로 가능도함수는 $f(x_1, \ldots, x_{10}) = \prod_{i=1}^{10} \frac{e^{-\theta}\theta^{x_i}}{x_i!}$ 이고, θ의 사전분포는 $Gamma(2, 1)$을 따른다.

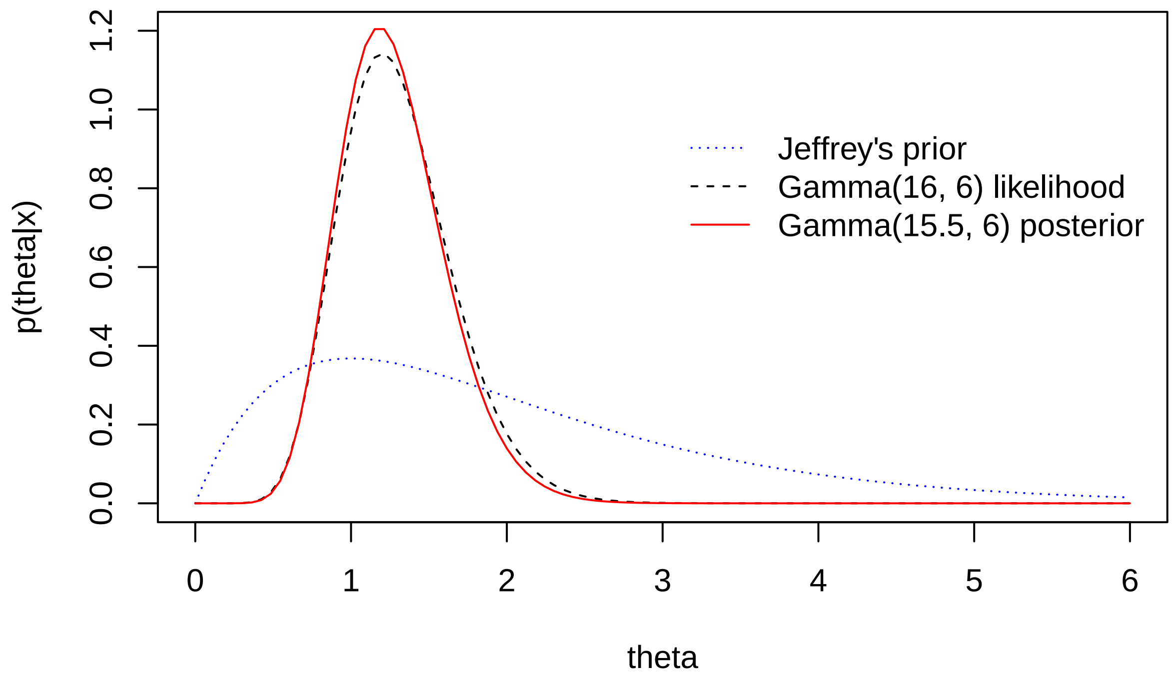
이로부터 θ의 사후분포 $\pi(\theta; \sum_{i=1}^{10} x_i)$를 다음과 같이 구할 수 있다.

$$\begin{aligned}
\pi(\theta; \sum_{i=1}^{10} x_i) &= f(x_1, \ldots, x_{10}) \cdot \pi(\theta) \\
&\propto \prod_{i=1}^{10} e^{-\theta}\theta^{x_i} \cdot e^{-\theta}\theta^{2-1} \\
&= e^{-10\theta}\theta^{\sum_{i=1}^{10} x_i} \cdot e^{-\theta}\theta^{1} \\
&= e^{-11\theta} \cdot \theta^{(\sum_{i=1}^{10} x_i + 2) - 1} \\
&= e^{-11\theta} \cdot \theta^{14-1}
\end{aligned}$$

이는 $Gamma(14, 11)$분포이며, 밀도함수는 다음과 같이 그려진다.

```r
theta = seq(0, 6, length = 100)
plot(theta, dgamma(theta, 2, 1), lty = 3, type = "l", xlab = "theta", ylab = "p(theta|x)",
     ylim = c(0, 1.2), col = "blue") # 사전분포
lines(theta, dgamma(theta, 13, 10), lty = 2, col = "black") # 가능도함수
lines(theta, dgamma(theta, 14, 11), col = 'red') # 사후분포

legend(3, 1, legend = c("Jeffrey's prior", "Gamma(16, 6) likelihood", "Gamma(15.5, 6) posterior"),
       lty = c(3, 2, 1), bty = "n", col = c("blue", "black", "red"))
```



(2)

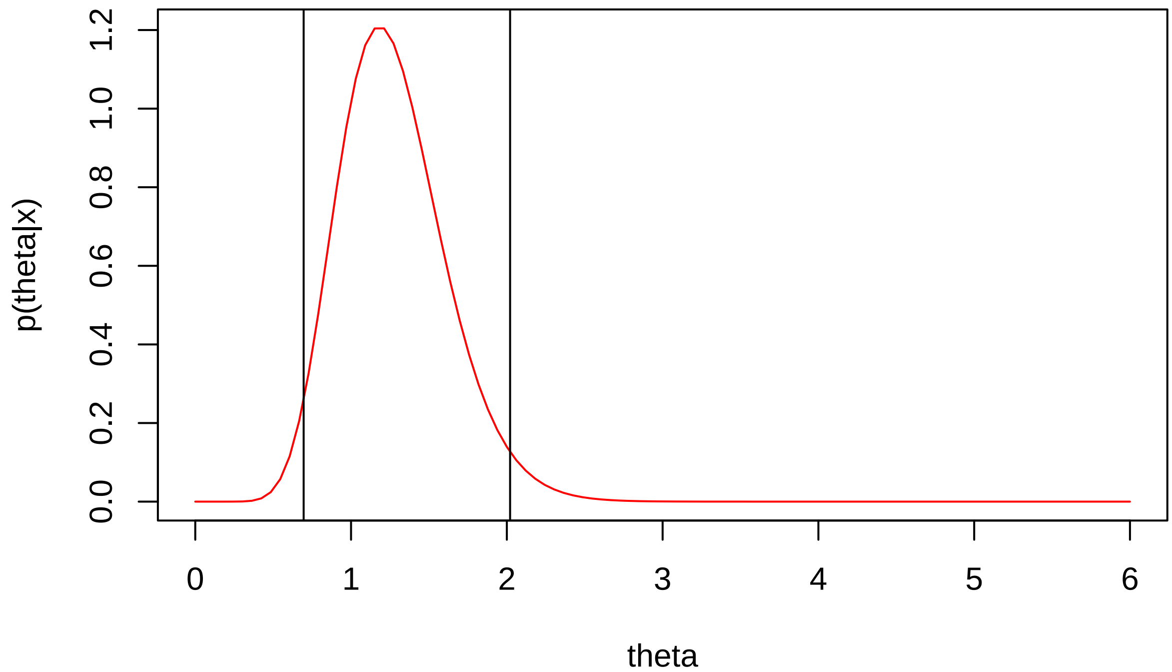
- 95% 최대 사후구간 구하기 (5점, 경험적 구간도 정답으로 인정)

사후분포를 알고 있으므로 이론적인 분위수 값으로 95% 최대 사후구간 (0.696, 2.021)을 구할 수 있다.

```r
qgamma(c(0.025, 0.975), 14, 11)
```

```
## [1] 0.6958118 2.0209451
```

```r
plot(theta, dgamma(theta, 14, 11), col = 'red', type = 'l', xlab = "theta",
     ylab = "p(theta|x)") # 사후분포
abline(v = qgamma(c(0.025, 0.975), 14, 11))
```

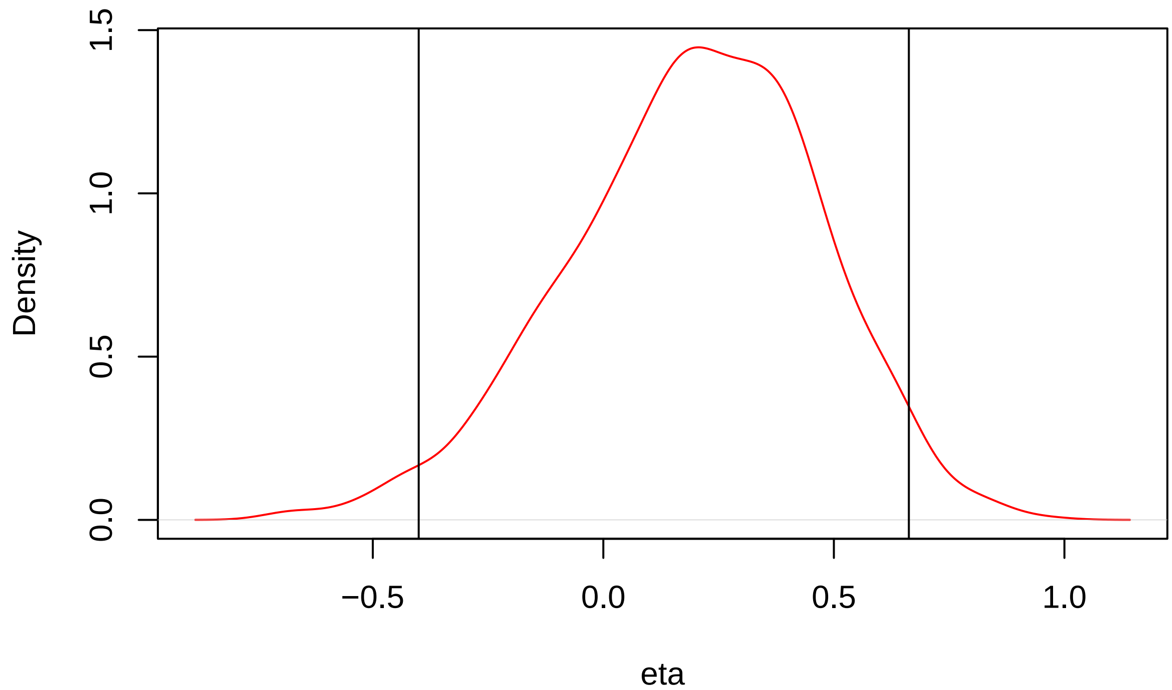


(3)

- $\log \theta$ 의 사후밀도 함수 (2점)
- $\log \theta$ 의 95% 사후구간 (3점)

1. 사후분포로부터 n 개의 $\theta_i, i = 1, \ldots, n$ 를 생성한다.
2. $\eta_i = \log \theta_i$ 인 η_i 를 만든다.
3. η_i 들로부터 2.5%, 97.5% 분위수에 해당하는 값으로 사후구간을 구한다.

```
n = 1000
theta = rgamma(n, 14, 11)
eta = log(theta)
plot(density(eta), col = 'red', main = '', xlab = 'eta')
abline(v = quantile(eta, c(0.025, 0.975)))
```




```
quantile(eta, c(0.025, 0.975))
```

```
##       2.5%      97.5%
## -0.4004954  0.6626553
```