베이지안통계학 과제4 solution

Jieun Shin

Spring 2022

문제 5-1. (10점)

- $Beta(18, 6)$의 분위수 값 (10점, 약간의 오차도 정답으로 인정)

통신사의 만족률θ는 $\theta \sim Beta(1, 1)$을 따른다. 이용 고객 20명의 만족여부를 $X_i, i = 1, \ldots, 22$라 하면 $X = \sum_{i=1}^{22}$는 $X \sim B(22, \frac{17}{22})$를 따른다.

이로부터 θ의 사후분포는 $\theta|X \sim Beta(17 + 1 = 18, 22 - 17 + 1 = 6)$를 따르고, θ의 95% 최대사후구간은 사후분포를 알고 있으므로 $Beta(18, 6)$의 2.5%, 97.5% 분위수 값으로 구할 수 있다 (또는 경험적으로 구하여도 됨).

```
qbeta(c(0.025, 0.975), 18, 6)
```

```
## [1] 0.5629693 0.8977139
```

문제 5-3. (15점)

(1)

- $\alpha = 4.4, \beta = 6.6$ (5점, 반올림한 $\alpha = 4, \beta = 6$도 정답으로 인정)

$Beta(\alpha, \beta)$로부터 $\frac{\alpha}{\alpha+\beta} = 0.4$와, $Beta(\alpha + 1, \beta)$로부터 $\frac{\alpha+1}{\alpha+1+\beta} = 0.45$의 두 식을 세울 수 있고, 연립방정식을 풀어 $\alpha = 4.4, \beta = 6.6$의 값을 구할 수 있다.

(2)

- $\theta|X \sim Beta(7.4, 13.6)$ (5점, 그림도 정답으로 인정)

θ의 사전분포는 $\theta \sim Beta(4.4, 6.6)$를 따르고, 조사 대상자 10명의 의견을 각 $X_i, i = 1, \ldots, 10$라 하면 $X = \sum_{i=1}^{10} X_i$은 $X \sim B(10, \frac{3}{10})$을 따른다.

이로부터 θ의 사후분포는 $\theta|X \sim Beta(3 + 4.4 = 7.4, 10 - 3 + 6.6 = 13.6)$을 따른다.

(3)

- $f(z|x_1, \ldots, x_{10}) = \binom{10}{z} \frac{Be(7.4+z, 23.6-z)}{Be(7.4, 13.6)}$ (5점, 그림도 정답으로 인정)

새로 추출한 표본을 $Z = X_{11} + \cdots + X_{20}$ 라 하자. 새로운 표본에서 낙채에 판성하는 사람들의 수에 대한 예측분포는

$$\begin{aligned}
f(z|x_1, \ldots, x_{10}) &= \int f(z|\theta, x_1, \ldots, x_{10}) \pi(\theta|x_1, \ldots, x_{10}) d\theta \\
&= \int \binom{10}{z} \theta^z (1-\theta)^{1-z} \frac{\Gamma(4.4+6.6+10)}{\Gamma(4.4+10)\Gamma(6.6+10-3)} \theta^{4.4+3-1} (1-\theta)^{6.6+10-3-1} d\theta \\
&= \binom{10}{z} \frac{\Gamma(4.4+6.6+10)}{\Gamma(4.4+10)\Gamma(6.6+10-3)} \frac{\Gamma(4.4+3+z)\Gamma(6.6+10-3+10-z)}{\Gamma(4.4+6.6+10+10)} \\
&= \binom{10}{z} \frac{Be(4.4+3+z, 6.6+10-3+10-z)}{Be(4.4+3, 6.6+10-3)} \\
&= \binom{10}{z} \frac{Be(7.4+z, 23.6-z)}{Be(7.4, 13.6)}
\end{aligned}$$

을 따른다. 여기서 $Be(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ 이다.