BAYESIAN STATISTICS

Chapter 7

Instructor: Seokho Lee

Hankuk University of Foreign Studies

7. Estimating Mean of Poisson Distribution

7.1. Count response variable

Examples of the count variables are:

- the number of traffic accidents in a year
- the number of tumors in a cancer patient
- the number of children in a household

The Poisson distribution is frequently used to analyze the count variables.

7.2. Conjugate prior: gamma distribution

Gamma distribution is the conjugate prior for the mean of Poisson distribution.

The probability density function of Gamma(α, β) is

$$f(x|\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-x/\beta}, \quad x \geq 0,\ \alpha > 0,\ \beta > 0$$

- If α is small, then most of mass is near 0 so that the distribution is skewed to the right. If α is large, then the distribution becomes symmetric and eventually similar to the normal distribution. α is called the **shape parameter**.
- The larger β is, the more the distribution spread out. β is called the **scale parameter**.
- $\Gamma(\alpha)$ is the **gamma function**, having the following property:

$$\Gamma(z+1) = z\Gamma(z)$$

- $E(X) = \alpha\beta$ and $var(X) = \alpha\beta^2$

Gamma distribution has two special cases, which are important in statistics: exponential distribution and chi-square distribution.

Definition (Exponential distribution)

When $\alpha = 1$ and $\beta = 1/\lambda$, the gamma distribution is called the **exponential distribution**. Its probability density function is

$$f(x|\lambda) = \lambda e^{-\lambda x}, \quad x \geq 0, \ \lambda > 0.$$

And

$$E(X) = \tfrac{1}{\lambda}, \quad var(X) = \tfrac{1}{\lambda^2}$$

Definition (Chi-square distribution)

When $\alpha = \nu/2$ and $\beta = 2$, the gamma distribution is called the **chi-square distribution**. Its probability density function is

$$f(x|\nu) = \tfrac{1}{\Gamma(\nu/2)2^{\nu/2}} x^{\nu/2-1} e^{-x/2}, \quad x \geq 0, \ \nu > 0.$$

And

$$E(X) = \nu, \quad var(X) = 2\nu.$$

The parameter ν is called the **degree of freedom**.

7.3. Bayesian estimation

Suppose $X \sim \text{Poisson}(\lambda)$ and the prior for λ is $\lambda \sim \text{Gamma}(\alpha, \beta)$. Then the posterior is following:

- prior $\propto \lambda^{\alpha-1} e^{-\lambda/\beta}$
- likelihood $\propto \lambda^{x} e^{-\lambda}$
- posterior
  $= \text{likelihood} \times \text{prior} \propto \lambda^{\alpha+x-1} e^{-\lambda(1/\beta+1)} \sim \text{Gamma}(\alpha+x, (1/\beta+1)^{-1})$

Note that gamma distribution is the conjugate prior for the mean of Poisson distribution.

Bayes estimator for λ is $\hat{\lambda}_B = \frac{\alpha+x}{1/\beta+1}$.

Example (7-1)

Let denote by X the number of cars that a salesperson sells during a year. X is supposed to follow Poisson distribution with mean λ. From his performance in the past, his annual sales average λ follows gamma distribution, $\lambda \sim Gamma(10, .5)$. Suppose he sells 12 cars in this year. Provide the posterior for λ and compare the expected values of λ from the prior and the posterior.

(solution) The posterior is $\lambda \sim \text{Gamma}(22, 1/3)$. Bayes estimator is $22/3 = 7.33$ from the posterior. This becomes larger than the expectation from the prior which is 5 because he sells more cars than the previous years.

Now, consider the case where there are n independent random variables: $X_1, \cdots, X_n \sim$ Poisson(λ) independently. Suppose the prior is $\lambda \sim$ Gamma(α, β). Then

- prior $\propto \lambda^{\alpha-1} e^{-\lambda/\beta}$
- likelihood $\propto \prod_{i=1}^{n} \lambda^{x_i} e^{-\lambda} = \lambda^{\sum_{i=1}^{n} x_i} e^{-n\lambda}$
- posterior $=$ likelihood $\times$ prior $\propto \lambda^{\alpha+\sum_{i=1}^{n} x_i - 1} e^{-\lambda(1/\beta+n)} \sim$ Gamma$(\alpha + \sum_{i=1}^{n} x_i, (1/\beta + n)^{-1})$

The corresponding Bayes estimator is $\hat{\lambda}_B = \frac{\alpha+\sum_{i=1}^{n} x_i}{1/\beta+n}$, which can be rewritten as

$$\hat{\lambda}_B = w_n(\alpha\beta) + (1 - w_n)\bar{x}$$

where $w_n = \frac{1/\beta}{1/\beta+n}$. This is the weighted average between the prior mean ($\alpha\beta$) and the maximum likelihood estimator ($\bar{x}$) with the weight w_n. Since $w_n \to 0$ as $n \to \infty$, the Bayes estimator becomes close to the maximum likelihood estimator when the sample size gets large. However, the Bayes estimator may be heavily affected by the prior when n is small.

Example (7-2)

To investigate the number of children per a household, 100 families were selected randomly. Total numbers of children from 100 households was 235. It is known that the number of children per a family is supposed to follow Poisson distribution with λ. From the past surveys, the prior for λ can be set as Gamma$(10, 1/5)$. Find the posterior for λ.

(solution) The posterior becomes $\text{Gamma}(10 + 235, 1/(5 + 100)) = \text{Gamma}(245, 1/105)$. Thus, the Bayes estimator for λ is $245/105 = 2.33$, which is close to the sample mean $235/100 = 2.35$.

Example (7-2 continue)

Find an 95% EPD credible interval for λ. Is it reasonable to assume that the average number of children per a family is 2?

(solution) The posterior is known to be Gamma$(245, 1/105)$. Thus, 95% EPD credible interval for λ is $(2.0503, 2.6344)$. Since 2 is outside this interval, 2 is not reasonable for λ.

```
> qgamma(0.025,shape=245,scale=1/105)
[1] 2.050282
> qgamma(0.975,shape=245,scale=1/105)
[1] 2.634421
```

When there is no prior information on λ, the vague prior distribution

$$\pi(\lambda) = \frac{1}{\lambda}, \qquad \lambda > 0$$

is popularly used. This prior is improper, but can be viewed as Gamma(α, β) with $\alpha \to 0$ and $\beta \to \infty$. With this prior,

- prior $= \frac{1}{\lambda}$
- likelihood $\propto \lambda^{\sum_{i=1}^n x_i} e^{-n\lambda}$
- posterior prior $\times$ likelihood $\propto \lambda^{\sum_{i=1}^n x_i - 1} e^{-n\lambda} \sim$ Gamma$(\sum_{i=1}^n x_i, 1/n)$.

Thus, the Bayes estimator of λ becomes $\frac{1}{n}\sum_{i=1}^n x_i$, which is the maximum likelihood estimator of λ.

Example (7-2 continue)

Suppose there is no information available on λ. Using the vague prior distribution, find the posterior distribution, Bayes estimator, and 95% EPD credible interval for λ.

(solution) The posterior distribution for λ becomes Gamma$(235, 1/100)$. The Bayes estimator is $235/100 = 2.35$ which is the same as the sample mean. 95% EPD credible interval is $(2.06, 2.66)$.

```
> qgamma(0.025,shape=235,scale=1/100)
[1] 2.059124
> qgamma(0.975,shape=235,scale=1/100)
[1] 2.659813
```