Stat 346/446: Theoretical Statistics II: Homework 3 Solutions

Textbook Exercises

7.19 (346 & 446: 3 pts.)Suppose that the random variables $Y_1, \dots, Y_n$ satisfy

$$Y_i = \beta x_i + \epsilon_i, \quad i = 1, \dots, n,$$

where $x_1, \dots, x_n$ are fixed constants, and $\epsilon_1, \dots, \epsilon_n$ are iid $\mathcal{N}(0, \sigma^2)$, σ^2 unknown.

(b) (346 & 446: 2 pts.) Find the MLE of β, and show that it is an unbiased estimator of β.

The likelihood function for $Y_1, \dots, Y_n$ is:

$$L(\beta, \sigma^2 | \mathbf{y}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(Y_i - \beta x_i)^2}{2\sigma^2}\right).$$

Taking the log-likelihood:

$$\log L(\beta, \sigma^2 | \mathbf{y}) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(Y_i - \beta x_i)^2.$$

Expanding the squared term:

$$\sum_{i=1}^{n}(Y_i - \beta x_i)^2 = \sum_{i=1}^{n} Y_i^2 - 2\beta\sum_{i=1}^{n} x_i Y_i + \beta^2 \sum_{i=1}^{n} x_i^2.$$

Thus, the log-likelihood function becomes:

$$\log L(\beta, \sigma^2 | \mathbf{y}) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\left(\sum_{i=1}^{n} Y_i^2 - 2\beta\sum_{i=1}^{n} x_i Y_i + \beta^2\sum_{i=1}^{n} x_i^2\right).$$

To find the MLE, we differentiate $\log L$ with respect to β:

$$\frac{\partial \log L}{\partial \beta} = \frac{1}{\sigma^2}\sum_{i=1}^{n} x_i Y_i - \frac{\beta}{\sigma^2}\sum_{i=1}^{n} x_i^2.$$

Setting this derivative equal to zero:

$$\sum_{i=1}^{n} x_i Y_i - \beta\sum_{i=1}^{n} x_i^2 = 0.$$

Solving for β:

$$\hat{\beta} = \frac{\sum_{i=1}^{n} x_i Y_i}{\sum_{i=1}^{n} x_i^2}.$$

Taking the second derivative:

$$\frac{\partial^2 \log L}{\partial \beta^2} = -\frac{1}{\sigma^2}\sum_{i=1}^{n} x_i^2.$$

Since $\sum x_i^2 > 0$, we have:

$$\frac{\partial^2 \log L}{\partial \beta^2} < 0.$$

Thus, the maximum likelihood estimator (MLE) of β is:

$$\hat{\beta} = \frac{\sum x_i Y_i}{\sum x_i^2}.$$

To check if $\hat{\beta}$ is an unbiased:

$$E[\hat{\beta}] = E\left[\frac{\sum_{i=1}^n x_i Y_i}{\sum_{i=1}^n x_i^2}\right].$$

Since $Y_i = \beta x_i + \epsilon_i$, and $x_1, \ldots, x_n$ are fixed constants:

$$E[Y_i] = \beta x_i.$$

Thus,

$$E[\hat{\beta}] = \frac{\sum_i x_i E[Y_i]}{\sum_i x_i^2} = \frac{\sum_i x_i \cdot \beta x_i}{\sum_i x_i^2} = \beta.$$

The estimator $\hat{\beta}$ is **unbiased**.

(c) (346 & 446: 1 pts.) Find the distribution of the MLE of β.

From part (b), we get:

$$\hat{\beta}_{MLE} = \frac{\sum x_i Y_i}{\sum x_i^2}.$$

Since $Y_i \sim \mathcal{N}(\beta x_i, \sigma^2)$, $\hat{\beta}$ is a linear combination of normal random variables:

$$\hat{\beta} = \sum_i a_i Y_i, \quad \text{where } a_i = \frac{x_i}{\sum_j x_j^2}.$$

By Corollary 4.6.10, a linear combination of normal variables is also normally distributed. Therefore, $\hat{\beta}$ follows a normal distribution with mean β and variance:

$$\text{Var}(\hat{\beta}) = \sum_i a_i^2 \text{Var}(Y_i).$$

Since $Y_i \sim \mathcal{N}(\beta x_i, \sigma^2)$, the variance of Y_i is σ^2. Thus,

$$\text{Var}(\hat{\beta}) = \sum_i a_i^2 \sigma^2 = \sum_i \left(\frac{x_i}{\sum_j x_j^2}\right)^2 \sigma^2 = \left(\sum_i \frac{x_i^2}{(\sum_j x_j^2)^2}\right) \sigma^2.$$

Since

$$\sum_i x_i^2 = \sum_j x_j^2,$$

Then

$$\text{Var}(\hat{\beta}) = \frac{\sum_i x_i^2}{(\sum_j x_j^2)^2} \sigma^2 = \frac{\sigma^2}{\sum_i x_i^2}.$$

Thus, the MLE $\hat{\beta}$ follows the normal distribution:

$$\hat{\beta} \sim \mathcal{N}\left(\beta, \frac{\sigma^2}{\sum x_i^2}\right).$$

7.20 (346: 4 pts, 446 : 2 pts.) Consider $Y_1, \ldots, Y_n$ as defined in Exercise 7.19.

(a) (346 : 2 pts, 446: 1 pts.) Show that $\sum Y_i / \sum x_i$ is an unbiased estimator of β.

We set

$$\hat{\beta}_1 = \frac{\sum_i Y_i}{\sum_i x_i}.$$

Then

$$E\left[\hat{\beta}_1\right] = E\left[\frac{\sum_i Y_i}{\sum_i x_i}\right].$$

Knows that:

$$E[Y_i] = \beta x_i.$$

Thus,

$$E\left[\sum_i Y_i\right] = \sum_i E[Y_i] = \sum_i (\beta x_i) = \beta \sum_i x_i.$$

Therefore

$$E\left[\hat{\beta}_1\right] = E\left[\frac{\sum_i Y_i}{\sum_i x_i}\right] = \frac{1}{\sum_i x_i} \sum_i E[Y_i] = \frac{1}{\sum_i x_i} \sum_i \beta x_i = \beta.$$

(b) (346 : 2 pts, 446 : 1 pts.) Calculate the exact variance of $\sum Y_i / \sum x_i$ and compare it to the variance of the MLE.

The variance of $\hat{\beta}_1$ is:

$$\operatorname{Var}\left(\frac{\sum_i Y_i}{\sum_i x_i}\right) = \frac{1}{(\sum_i x_i)^2} \sum_i \operatorname{Var}(Y_i).$$

Since $Y_i \sim \mathcal{N}(\beta x_i, \sigma^2)$, we have:

$$\operatorname{Var}(Y_i) = \sigma^2.$$

Thus,

$$\operatorname{Var}\left(\hat{\beta}_1\right) = \frac{1}{(\sum_i x_i)^2} \sum_i \sigma^2 = \frac{\sum_i \sigma^2}{(\sum_i x_i)^2} = \frac{n\sigma^2}{(\sum_i x_i)^2} = \frac{n\sigma^2}{n^2 \bar{x}^2} = \frac{\sigma^2}{n\bar{x}^2}.$$

The variance of the **MLE** $\hat{\beta}$ is:

$$\operatorname{Var}(\hat{\beta}) = \frac{\sigma^2}{\sum_i x_i^2}.$$

Because $\sum_i x_i^2 - n\bar{x}^2 = \sum_i (x_i - \bar{x})^2 \geq 0$, we have $\sum_i x_i^2 \geq n\bar{x}^2$. Hence,

$$\operatorname{Var}(\hat{\beta}) = \frac{\sigma^2}{\sum_i x_i^2} \leq \frac{\sigma^2}{n\bar{x}^2} = \operatorname{Var}\left(\frac{\sum_i Y_i}{\sum_i x_i}\right).$$

Extra Problems

1. (346 & 446 : 2 pts.) Let $X_1, X_2, \ldots, X_n$ be a random sample from the Expo(θ) distribution and let the prior distribution of θ be the Gamma(α, β) distribution. Find the posterior distribution of θ and the Bayes estimator of θ.

We are given that $X_1, X_2, \ldots, X_n$ is a random sample from an Expo(θ) distribution:

$$X_i \sim \text{Expo}(\theta)$$

which has the PDF:

$$f(x_i \mid \theta) = \theta e^{-\theta x_i}, \quad x_i > 0, \quad \theta > 0.$$

Since $X_1, \dots, X_n$ are independent:

$$f(\mathbf{x} \mid \theta) = \prod_{i=1}^{n} \theta e^{-\theta x_i} = \theta^n e^{-\theta \sum_{i=1}^{n} x_i}.$$

The prior distribution for θ follows a Gamma(α, β) distribution:

$$\theta \sim \text{Gamma}(\alpha, \beta)$$

$$p(\theta \mid \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta}, \quad \theta > 0, \quad \alpha > 0, \quad \beta > 0.$$

Thus by Bayes Theorem, the posterior distribution is proportional to

$$\begin{aligned} p(\theta \mid \mathbf{x}) &\propto f(\mathbf{x} \mid \theta) p(\theta) \\ &\propto \theta^n e^{-\theta \sum_{i=1}^{n} x_i} \times \theta^{\alpha-1} e^{-\beta\theta} \\ &= \theta^{(n+\alpha)-1} e^{-(\beta + \sum_{i=1}^{n} x_i)\theta} \end{aligned}$$

Which is the kernel of a Gamma distribution:Gamma$(\alpha + n, \beta + \sum_{i=1}^{n} x_i)$. Thus, the posterior distribution of θ is Gamma$(\alpha + n, \beta + \sum_{i=1}^{n} x_i)$. The Bayes estimator of θ is given by the posterior mean, which for a Gamma distribution Gamma(α', β') is:

$$E[\theta \mid \mathbf{x}] = \frac{\alpha'}{\beta'}.$$

From the posterior distribution:

$$E[\theta \mid \mathbf{x}] = \frac{\alpha + n}{\beta + \sum_{i=1}^{n} x_i}.$$

Thus, the Bayes estimator of θ is:

$$\hat{\theta}_{\text{Bayes}} = \frac{\alpha + n}{\beta + \sum_{i=1}^{n} x_i}.$$

2. (346 & 446 : 1 pts.) Suppose that the time in minutes required to serve a customer at a certain facility has an exponential distribution for which the value of the parameter θ is unknown and that the prior distribution of θ is a gamma distribution for which the mean is 0.2 and the standard deviation is 1. If the average time required to serve a random sample of 20 customers is observed to be 3.8 minutes, what is the posterior distribution of θ?

The time required to serve a customer follows an Exponential distribution with rate parameter θ:

$$X_i \sim \text{Expo}(\theta)$$

which has the PDF:

$$f(x_i \mid \theta) = \theta e^{-\theta x_i}, \quad x_i > 0, \quad \theta > 0.$$

The prior distribution of θ follows a Gamma distribution:

$$\theta \sim \text{Gamma}(\alpha, \beta).$$

We are given that the prior mean is 0.2 and the standard deviation is 1. Using the mean and variance formulas for the Gamma distribution:

$$E[\theta] = \frac{\alpha}{\beta} = 0.2, \quad \text{Var}(\theta) = \frac{\alpha}{\beta^2} = 1.$$

Solving for α and β:

$$\frac{\alpha}{\beta} = 0.2, \quad \frac{\alpha}{\beta^2} = 1.$$

From the first equation: $\alpha = 0.2\beta$. Substituting into the second equation:

$$\frac{0.2\beta}{\beta^2} = 1 \quad \Rightarrow \quad 0.2 = \beta.$$

Using $\alpha = 0.2\beta = 0.2(0.2) = 0.04$, we obtain:

$$\alpha = 0.04, \quad \beta = 0.2.$$

A random sample of 20 customers is observed, and the average service time is 3.8 minutes, so:

$$\sum_{i=1}^{20} x_i = 20 \times 3.8 = 76.$$

Given that $X_1, X_2, \ldots, X_n \sim \text{Expo}(\theta)$ and $\theta \sim \text{Gamma}(\alpha, \beta)$, as we concluded from Extra Problem 1, the posterior distribution is:

$$\theta \mid X_1, \ldots, X_n \sim \text{Gamma}(\alpha + n, \beta + \sum_{i=1}^{n} x_i),$$

where:

- $n = 20$ (number of observations),
- $\sum_{i=1}^{20} x_i = 76$,
- Prior parameters: $\alpha = 0.04$, $\beta = 0.2$.

Thus, the posterior parameters are:

$$\alpha' = \alpha + n = 0.04 + 20 = 20.04,$$

$$\beta' = \beta + S = 0.2 + 76 = 76.2.$$

The posterior distribution of θ is:

$$\theta \mid X_1, \ldots, X_n \sim \text{Gamma}(20.04, 76.2).$$

3. (446 : 2 pts.) For a distribution with mean $\mu \neq 0$ and standard deviation $\sigma > 0$, the *coefficient of variation* of the distribution is defined as $\sigma/|\mu|$. Consider again the problem described in extra problem 2. Suppose that the coefficient of variation of the prior gamma distribution of θ is 2. What is the smallest number of customers that must be observed in order to reduce the coefficient of variation of the posterior distribution to 0.1?

The coefficient of variation (CV) for a random variable X is defined as:

$$CV(X) = \frac{\sigma}{|\mu|}$$

For a Gamma-distributed random variable $X \sim \text{Gamma}(\alpha, \beta)$, the mean, variance, and coefficient of variation are:

$$E[X] = \frac{\alpha}{\beta}, \quad V(X) = \frac{\alpha}{\beta^2}, \quad CV(X) = \frac{1}{\sqrt{\alpha}}.$$

The prior distribution of θ is $\text{Gamma}(\alpha, \beta)$ with a prior CV of 2:

$$CV(\theta) = \frac{1}{\sqrt{\alpha}} = 2.$$

Solving for α:

$$\frac{1}{\sqrt{\alpha}} = 2 \quad \Rightarrow \quad \alpha = \frac{1}{4} = 0.25.$$

We want to find the smallest n such that the posterior CV is at most 0.1 Given that $X_1, X_2, \ldots, X_n \sim \text{Expo}(\theta)$ and $\theta \sim \text{Gamma}(\alpha, \beta)$, as we concluded from Extra Problem 1, the posterior distribution is:

$$\theta \mid X_1, \ldots, X_n \sim \text{Gamma}(\alpha + n, \beta + \sum_{i=1}^{n} x_i),$$

From this, the posterior shape parameter is:

$$\alpha' = \alpha + n = 0.25 + n.$$

The posterior coefficient of variation is:

$$CV(\theta_{\text{post}}) = \frac{1}{\sqrt{\alpha'}} = \frac{1}{\sqrt{0.25 + n}}.$$

We want:

$$\frac{1}{\sqrt{0.25 + n}} \leq 0.1.$$

$$\Rightarrow \sqrt{0.25 + n} \geq 10.$$

$$\Rightarrow 0.25 + n \geq 100.$$

Solving for n:

$$n \geq 99.75.$$

Since n (the number of observations of the customers) must be a whole number, the smallest valid n is:

$$n = 100.$$

The smallest number of customers that must be observed to reduce the coefficient of variation of the posterior distribution to 0.1 is 100.