STAT2006 Basic Concepts in Statistics and Probability II
Tutorial 6 – Midterm Revision: Point Estimation Exercises

Benjamin Chun Ho Chan

January 4, 2019

Abstract

It is an extension of Tutorial 5: Point Estimation. More exercises are provided for students to practice. Note that the concept of method of moments estimator is not covered in tutorials but exercises are provided here. Some are extracted from the textbook *Statistical Inference*, written by Casella and Berger and used in STAT4003.

1 Exercises

Exercise 1. Let $X_1, \ldots, X_n$ be a random sample from the pdf

$$f(x;\theta) = \theta x^{-2}, \;\; 0 < \theta \leq x < \infty.$$

Find the maximum likelihood estimator (MLE) of θ.[1]

Solution. The pdf can be rewritten using indicator function:

$$f(x;\theta) = \theta x^{-2} \mathbb{1}_{\{x \geq \theta\}}.$$

The likelihood function is given by

$$\begin{aligned} L(\theta; x_1, \ldots, x_n) &= \prod_{i=1}^{n} \theta x_i^{-2} \mathbb{1}_{\{x_i \geq \theta\}} \\ &= \theta^n \left(\prod_{i=1}^{n} x_i^{-2} \right) \mathbb{1}_{\{\min_i x_i \geq \theta\}}. \end{aligned}$$

Notice that $\theta^n (\prod_{i=1}^n x_i^{-2})$ is an increasing function in θ. To maximize L, we want to make θ as large as possible. But because of the indicator function, the MLE is given by

$$\hat{\theta}_{\text{MLE}} = \min_i X_i.$$

[1] "Statistical Inference" 2nd ed. (Casella and Berger) Ch7 p.355 Q7.6

Exercise 2. [15-16 Final Q2(a), 6%]
Assume that a random sample of $X_1, \ldots, X_n$ is following a common distribution

$$P(X_i \leq x|\alpha, \beta) = \begin{cases} 0, & \text{if } x < 0; \\ (\frac{x}{\alpha})^\beta, & \text{if } 0 \leq x \leq \alpha; \\ 1, & \text{if } x > \alpha; \end{cases}$$

where the parameters α and β are positive. Assume α and β are both unknown, find the MLEs of α and β.

Solution. First of all, we have to find the pdf from the cdf:

$$\begin{aligned} f(x|\alpha, \beta) &= \frac{d}{dx}\left(\frac{x}{\alpha}\right)^\beta \\ &= \frac{\beta}{\alpha^\beta} x^{\beta-1} \mathbb{1}_{\{0 \leq x \leq \alpha\}}. \end{aligned}$$

The likelihood function is

$$\begin{aligned} L(\alpha, \beta; x_1, \ldots, x_n) &= \prod_{i=1}^n \frac{\beta}{\alpha^\beta} x_i^{\beta-1} \mathbb{1}_{\{0 \leq x_i \leq \alpha\}} \\ &= \frac{\beta^n}{\alpha^{\beta n}} \left(\prod_{i=1}^n x_i\right)^{\beta-1} \mathbb{1}_{\{\min_i x_i \geq 0\}} \mathbb{1}_{\{\max_i x_i \leq \alpha\}}. \end{aligned}$$

For any fixed β, $L = 0$ if $\alpha < \max_i x_i$, and L is a decreasing function of α if $\alpha \geq \max_i x_i$. Thus,

$$\hat{\alpha}_{\text{MLE}} = \max_i X_i.$$

For the MLE of β, calculate

$$\begin{aligned} \frac{\partial}{\partial \beta} \ln L &= \frac{\partial}{\partial \beta}\left[n \ln \beta - \beta n \ln \alpha + (\beta - 1) \ln \prod_{i=1}^n x_i\right] \\ &= \frac{n}{\beta} - n \ln \alpha + \ln \prod_{i=1}^n x_i. \end{aligned}$$

Note that

$$\frac{\partial^2}{\partial \beta^2} \ln L = -\frac{n}{\beta^2} < 0.$$

Set $\frac{\partial}{\partial \beta} \ln L = 0$ and use $\hat{\alpha}_{\text{MLE}} = \max_i X_i$ to obtain

$$\begin{aligned} \hat{\beta}_{\text{MLE}} &= \frac{n}{n \ln \max_i X_i - \ln \prod_{i=1}^n X_i} \\ &= \left[\frac{1}{n} \sum_{i=1}^n (\ln \max_i X_i - \ln X_i)\right]^{-1}. \end{aligned}$$

Exercise 3. Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from a uniform$[\theta, 2\theta]$ distribution, where $\theta > 0$.[2]

(a) Find the method of moments estimator, $\tilde{\theta}$, and find a constant c such that $E(c\tilde{\theta}) = \theta$.

Solution. For the uniform$[\theta, 2\theta]$ distribution, we have

$$E(X) = \frac{\theta + 2\theta}{2} = \frac{3}{2}\theta.$$

So we solve

$$\frac{3}{2}\theta = \bar{X}$$

for θ to obtain the method of moments estimator

$$\tilde{\theta}_{\text{MME}} = \frac{2}{3}\bar{X}.$$

First of all, note that

$$E(\tilde{\theta}) = \frac{2}{3}E(\bar{X}) = \frac{2}{3}E(X) = \frac{2}{3} \cdot \frac{3}{2}\theta = \theta,$$

which is unbiased. So the constant c is 1.

(b) Find the MLE, $\hat{\theta}$, and find a constant k such that $E(k\hat{\theta}) = \theta$.

Solution. First of all, write down the pdf as

$$f(x;\theta) = \frac{1}{\theta}\mathbb{1}_{\{\theta \le x \le 2\theta\}}.$$

Then, the likelihood function is

$$\begin{aligned} L(\theta; x_1, \ldots, x_n) &= \prod_{i=1}^{n} \frac{1}{\theta}\mathbb{1}_{\{\theta \le x_i \le 2\theta\}} \\ &= \frac{1}{\theta^n}\mathbb{1}_{\{\theta \le \min_i x_i\}}\mathbb{1}_{\{\max_i x_i \le 2\theta\}} \\ &= \frac{1}{\theta^n}\mathbb{1}_{\{\frac{1}{2}\max_i x_i \le \theta \le \min_i x_i\}}. \end{aligned}$$

Because $\frac{1}{\theta^n}$ is a decreasing function of θ, L is maximized at $\theta = \frac{1}{2}\max_i x_i$. So the MLE is

$$\hat{\theta}_{\text{MLE}} = \frac{1}{2}\max_i X_i.$$

To get $E(\hat{\theta})$, one approach is to first find the distribution of $X_{(n)} \triangleq \max_i X_i$. Its cdf is

$$\begin{aligned} P(X_{(n)} \le s) &= P(X_1 \le s) \times \cdots \times P(X_n \le s) \\ &= \left(\int_\theta^s \frac{1}{\theta}dx\right)^n = \left(\frac{s-\theta}{\theta}\right)^n. \end{aligned}$$

The pdf of $X_{(n)}$ is given by

$$f_{X_{(n)}}(s;\theta) = \frac{d}{ds}\left(\frac{s-\theta}{\theta}\right)^n = \frac{n(s-\theta)^{n-1}}{\theta^n}.$$

[2] "Statistical Inference" 2nd ed. (Casella and Berger) Ch7 p.364 Q7.46

Thus,

$$
\begin{aligned}
E(X_{(n)}) &= \int_{\theta}^{2\theta} s\frac{n(s-\theta)^{n-1}}{\theta^n}ds \\
&= \int_{\theta}^{2\theta} \frac{s}{\theta^n}d(s-\theta)^n \\
&= \left[\frac{s}{\theta^n}(s-\theta)^n\right]_{\theta}^{2\theta} - \int_{\theta}^{2\theta} \frac{(s-\theta)^n}{\theta^n}ds \qquad \text{(integration by parts)} \\
&= \frac{2\theta}{\theta^n}\theta^n - \int_{\theta}^{2\theta} \frac{(s-\theta)^n}{\theta^n}d(s-\theta) \\
&= 2\theta - \left[\frac{(s-\theta)^{n+1}}{\theta^n(n+1)}\right]_{\theta}^{2\theta} \\
&= 2\theta - \frac{\theta^{n+1}}{\theta^n(n+1)} \\
&= \frac{2\theta(n+1)}{n+1} - \frac{\theta}{n+1} \\
&= \frac{2n+1}{n+1}\theta.
\end{aligned}
$$

Therefore,

$$E(\hat{\theta}) = \frac{1}{2}E(X_{(n)}) = \frac{2n+1}{2n+2}\theta.$$

The constant is

$$k = \frac{2n+2}{2n+1}$$

since

$$E(k\hat{\theta}) = \frac{2n+2}{2n+1}E(\hat{\theta}) = \frac{2n+2}{2n+1}\cdot\frac{2n+1}{2n+2}\theta = \theta.$$

(Optional) Exercise 4. [Related to Simple Linear Regression]
Suppose that the random variables $Y_1, \ldots, Y_n$ satisfy

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

where $x_1, \ldots, x_n$ are fixed constants, and $\epsilon_1, \ldots, \epsilon_n$ are iid $N(0, \sigma^2)$, σ^2 unknown. Find the MLE of β, and show that it is an unbiased estimator of β. Also, find the distribution of the MLE of β.[3]

Solution. Note that $\epsilon_i = Y_i - \beta x_i \overset{\text{i.i.d.}}{\sim} N(0, \sigma^2)$. The likelihood function is

$$
\begin{aligned}
L(\beta, \sigma^2; y_1, \ldots, y_n) &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}}\exp\bigg(-\frac{1}{2\sigma^2}(y_i - \beta x_i)^2\bigg) \\
&= (2\pi\sigma^2)^{-n/2}\exp\bigg(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i^2 - 2\beta x_i y_i + \beta^2 x_i^2)\bigg) \\
&= (2\pi\sigma^2)^{-n/2}\exp\bigg(-\frac{1}{2\sigma^2}\sum_{i=1}^{n} y_i^2 + \frac{\beta}{\sigma^2}\sum_{i=1}^{n} x_i y_i - \frac{\beta^2}{2\sigma^2}\sum_{i=1}^{n} x_i^2\bigg).
\end{aligned}
$$

[3] "Statistical Inference" 2nd ed. (Casella and Berger) Ch7 p.358 Q7.19

The log-likelihood function is then

$$\ln L(\beta, \sigma^2; y_1, \ldots, y_n) = -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln \sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^n y_i^2 + \frac{\beta}{\sigma^2}\sum_{i=1}^n x_i y_i - \frac{\beta^2}{2\sigma^2}\sum_{i=1}^n x_i^2.$$

For a fixed value of σ^2,

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

Setting $\frac{\partial \ln L}{\partial \beta} = 0$, we get

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

Note that

$$\frac{\partial^2 \ln L}{\partial \beta^2} = -\frac{1}{\sigma^2}\sum_{i=1}^n x_i^2 < 0,$$

so it is a maximum. Because $\hat{\beta}$ does not depend on σ^2, the MLE of β is

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

Check that

$$\begin{aligned} E(\hat{\beta}_{\text{MLE}}) &= \frac{\sum_{i=1}^n x_i E(Y_i)}{\sum_{i=1}^n x_i^2} \\ &= \frac{\sum_{i=1}^n x_i \cdot \beta x_i}{\sum_{i=1}^n x_i^2} \\ &= \beta. \end{aligned}$$

So $\hat{\beta}_{\text{MLE}}$ is an unbiased estimator of β. We can write $\hat{\beta}_{\text{MLE}}$ as

$$\hat{\beta}_{\text{MLE}} = \sum_{i=1}^n a_i Y_i, \quad \text{where } a_i = \frac{x_i}{\sum_{j=1}^n x_j^2} \text{ are constants.}$$

Note that any linear combination of normal random variables are also normally distributed. Hence $\hat{\beta}_{\text{MLE}}$ is normally distributed with mean β, and

$$\begin{aligned} \text{Var}(\hat{\beta}_{\text{MLE}}) &= \sum_{i=1}^n a_i^2 \text{Var}(Y_i) \\ &= \sum_{i=1}^n \left(\frac{x_i}{\sum_{j=1}^n x_j^2}\right)^2 \sigma^2 \\ &= \frac{\sum_{i=1}^n x_i^2}{(\sum_{j=1}^n x_j^2)^2}\sigma^2 \\ &= \frac{\sigma^2}{\sum_{i=1}^n x_i^2}. \end{aligned}$$

In other words,

$$\hat{\beta}_{\text{MLE}} \sim N\left(\beta, \frac{\sigma^2}{\sum_{i=1}^n x_i^2}\right).$$

(Optional) Remark: For multiple linear regression, the model is $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, where $\boldsymbol{X}$ is the design matrix and all others are vectors. Under multivariate normal assumption, i.e. $\boldsymbol{\epsilon} \sim N(\boldsymbol{0}, \sigma^2\boldsymbol{I})$, then $\hat{\boldsymbol{\beta}}_{\text{MLE}} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{Y}$ and $\hat{\boldsymbol{\beta}}_{\text{MLE}} \sim N(\boldsymbol{\beta}, \sigma^2(\boldsymbol{X}^T\boldsymbol{X})^{-1})$. It is a famous and well-known result. Interested students should be familiar with linear algebra and consider to take STAT3008 Applied Regression Analysis in the future.