# Chapter 7: Point Estimation

May 20, 2023

## Exercise 7.1

If we take the product of $f(x|\theta)$ as the likelihood function, all 3 values of $\theta$ attain maximum likelihood of 0.

**Method 1: With perturbation on likelihood function**

We will perturb the likelihood function by $\epsilon$. Now even the probability 0 entries get a probability of $\epsilon$. Then

$$L(\theta = 1|x + \epsilon) = \frac{\epsilon(1/3 + \epsilon)^2(1/6 + \epsilon)^2}{(1 + 5\epsilon)^5}$$
$$L(\theta = 2|x + \epsilon) = \frac{\epsilon(1/4 + \epsilon)^4}{(1 + 5\epsilon)^5}$$
$$L(\theta = 3|x + \epsilon) = \frac{\epsilon^2(1/4 + \epsilon)^2(1/2 + \epsilon)}{(1 + 5\epsilon)^5}$$

We can compare the functions by taking the ratio and letting $\epsilon$ go to 0.

$$\frac{L(\theta = 2|x)}{L(\theta = 1|x)} = \lim_{\epsilon\to 0} \frac{L(\theta = 2|x + \epsilon)}{L(\theta = 1|x + \epsilon)} = \lim_{\epsilon\to 0} \frac{(1/4 + \epsilon)^4}{(1/3 + \epsilon)^2(1/6 + \epsilon)^2} = 1.26$$

Similarly,

$$\frac{L(\theta = 2|x)}{L(\theta = 3|x)} = \lim_{\epsilon\to 0} \frac{(1/4 + \epsilon)^4}{\epsilon(1/4 + \epsilon)^2(1/2 + \epsilon)} = \infty$$

Therefore $\theta = 2$ is the MLE

**Method 2: With +1 smoothing**

we can assume $N$ trials are performed for each $\theta$ and apply laplace smoothing (+1 smoothing).

Then the likelihood functions become

$$L(\theta = 1|x) = \lim_{N\to\infty} \frac{(N/3 + 1)^2(N/6 + 1)^2}{(N + 5)^5}$$
$$L(\theta = 2|x) = \lim_{N\to\infty} \frac{(N/4 + 1)^4}{(N + 5)^5}$$
$$L(\theta = 3|x) = \lim_{N\to\infty} \frac{(N/4 + 1/N)^2(N/2 + 1)}{(N + 5)^5}$$

This is essentially the same as method 1.

# Exercise 7.2

Gamma$(\alpha, \beta)$ has pdf $f(x|\alpha, \beta) = \frac{x^{\alpha-1}e^{-\beta x}\beta^{\alpha}}{\Gamma(\alpha)}$, with $\alpha$ known, the likelihood function is

$$L(\alpha, \beta|x) = \prod_{i=1}^{n} f(x_i|\alpha, \beta) = \prod_{i=1}^{n} \frac{x_i^{\alpha-1}e^{-\beta x_i}\beta^{\alpha}}{\Gamma(\alpha)} = \frac{(\prod_i x_i)^{\alpha-1}}{\Gamma^n(\alpha)} \left[\beta^{n\alpha}e^{-\beta(\sum_i x_i)}\right]$$

Let $\frac{\partial L(\beta|x,\alpha)}{\partial \beta} = 0$, we get the MLE for $\beta$.

$$\hat{\beta} = \frac{\alpha}{\bar{x}}$$

If $\alpha$ is also unknown,

$$\frac{\partial L(\alpha, \beta|x)}{\partial \alpha} = \frac{\partial}{\partial \alpha} \frac{A^{\alpha-1}\beta^{n\alpha}}{\Gamma^n(\alpha)} = 0$$

$$\Rightarrow \sum \ln x_i + n \ln \beta = \psi(\alpha) \Rightarrow \sum \ln x_i + n(\ln \alpha - ln\bar{x}) = \psi(\alpha)$$

where $\psi$ is polygamma function

We can then solve for $\alpha$ numerically.

# Exercise 7.3

$\log$ is monotonic increasing function, when $L$ attains maximum values, $\log(L)$ also attains maximum.

# Exercise 7.6

$f(x|\theta) = \theta x^{-2}$ with $0 < \theta \le x < \infty$. We need to extend the range of $x$ in order to get rid of the $\theta$ dependency with indicator function. So $f(x|\theta) = \mathbb{I}_{\theta \le x}\theta x^{-2}$ and $x \in \mathbb{R}$.

(a) For a sample $X_1, \ldots, X_n$, by factorization theorem, $f(x|\theta) = g(T(x)|\theta)h(x)$. So $h(x) = x^-2$ and $g(T(x)|\theta) = \mathbb{I}_{\theta \le x}\theta = \mathbb{I}_{\theta \le x_i, \forall i}\theta = \mathbb{I}_{\theta \le \min x_i}$. Therefore $T(x) = \min x_i$ is a sufficient statistics.

(b) The likelihood function $L(\theta|x) = \mathbb{I}_{\theta \le x_i}\theta^n(\prod_i x_i)^{-2}$. $L(\theta|x)$ is non zero only when $\theta \le \min x_i$. Then $L(\theta|x) = \theta^n(\prod_i x_i)^{-2}$ when $\theta \le \min x_i$ decreases monotonically as $\theta$ decreases. So it attains maximum when $\theta = \min x_i$