1st Semester, 2021-2022

3rd Year, Math & CompSci

Mathematical Statistics

Seminar Exercises: Week 8

Recap. *Throughout this class, $X_1, X_2, \ldots, X_n, \ldots$ will be i.i.d. random variables that follow the distribution of a given characteristic X.*

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

- *If $X \sim Unif[a, b]$, then:*

$$f_X(x) = \begin{cases} \frac{1}{b-a}, \text{ if } x \in [a, b] \\ 0, \text{ otherwise} \end{cases}$$

- *If $X \sim Unid(a)$, then:*

$$X \begin{pmatrix} k \\ \frac{1}{a} \end{pmatrix}_{k=\overline{1,a}}$$

*A **point estimator** for the target parameter θ is a statistic:*

$$\overline{\theta} = \theta(X_1, X_2, \ldots, X_n)$$

We have the following notions:

- ***unbiased estimator**: $E(\overline{\theta}) = \theta$ (the **bias**: $B := E(\overline{\theta}) - \theta$)*

- ***minimum-variance unbiased estimator (MVUE)*** *if it has lower variance than any other unbiased estimator for* θ*:*

$$V(\overline{\theta}) \leq V(\hat{\theta}),\ \forall \hat{\theta} \text{ with } E(\hat{\theta}) = \theta$$

- *The* ***likelihood function*** *of the sample* $X_1, X_2, \ldots, X_n$*:*

$$L(X_1, X_2, \ldots, X_n; \theta) = \prod_{i=1}^{n} f(X_i; \theta);$$

The statistic

$$S = S(X_1, \ldots, X_n)$$

is called:

- ***sufficient*** *for (the estimation of) the parameter* θ*, if:*

$$f(x_1, \ldots, x_n; \theta \mid S = s) = f(x_1, \ldots, x_n \mid S = s)$$

- ***complete*** *for the family of probability distributions* $(f(x;\theta))_{\theta \in A}$ *if for every measurable function* ϕ *we have the implication:*

$$E(\phi(S)) = 0, \forall \theta \in A \Rightarrow P(\phi(S) = 0) = 1, \forall \theta \in A$$

Theorem (Fisher's Factorization Criterion)**.** *A statistic*

$$S = S(X_1, X_2, \ldots, X_n)$$

is sufficient for θ*, if and only if the likelihood function*

$$L(X_1, X_2, \ldots, X_n; \theta) = \prod_{i=1}^{n} f(X_i; \theta)$$

can be factored into two nonnegative functions

$$L(x_1, x_2, \ldots, x_n; \theta) = g(x_1, x_2, \ldots, x_n) \cdot h(s; \theta),$$

where $s = S(x_1, x_2, \ldots, x_n)$*.*

Theorem (Lehmann-Scheffé). *Let* $\hat{\theta} = \hat{\theta}(X_1, X_2, \ldots, X_n)$ *be an unbiased estimator for* θ *and* $S = S(X_1, X_2, \ldots, X_n)$ *be a sufficient and complete statistic for* θ*. Then the estimator*

$$\overline{\theta} = \overline{\theta}(X_1, X_2, \ldots, X_n) = E(\hat{\theta} \mid S)$$

is a MVUE.

Exercise 1. Let $f(x;\theta) = e^{a(x)\alpha(\theta)+b(x)+\beta(\theta)}$, for x in the range of X, where θ is a parameter of X and a, α, b, β are measurable functions, be a probability density function of the (discrete or continuous) characteristic X. Prove that the statistic

$$S = S(X_1, \ldots, X_n) = \sum_{i=1}^{n} a(X_i)$$

is sufficient for θ.

Exercise 2. Let $X \sim Bern(p)$, where $p \in (0,1)$.

(a) Prove that

$$S = X_1 + \ldots + X_n$$

is a sufficient and complete statistic for p.

(b) Show that $\overline{X}$ is an unbiased estimator for p.

(c) Find the MVUE of p.

Exercise 3. Let $X \sim Unif[0,\theta]$, where $\theta > 0$ is a parameter.

(a) Prove that

$$S = \max\{X_1, \ldots, X_n\}$$

is a sufficient and complete statistic for θ.

(b) Show that

$$\overline{\theta} = \frac{n+1}{n} \max(X_1 \ldots, X_n)$$

is an unbiased estimator for θ.

(c) Find the MVUE of θ.

Hint: If $X \sim Unif[0, \theta]$, then the pdf of $S = \max(X_1, X_2 \ldots, X_n)$ is:

$$f_S(x) = 1_{[0,\theta]} \cdot \frac{nx^{n-1}}{\theta^n}$$

Exercise 4. Let $X \sim Unid(\theta)$, where $\theta \in \mathbb{N}^*$ is a parameter.

1. Prove that
$$S = \max\{X_1, \ldots, X_n\}$$
is a sufficient and complete statistic for θ.

2. Show that
$$\overline{\theta} = \frac{S^{n+1} - (S-1)^{n+1}}{S^n - (S-1)^n}$$
is an unbiased estimator for θ.

3. Find the MVUE of θ.

Hint: If $X \sim Unid(\theta)$, then the pdf of $S = \max(X_1, X_2 \ldots, X_n)$ is:

$$f_S(x) = 1_{\{1,\ldots,\theta\}} \cdot \left(\left(\frac{x}{\theta}\right)^n - \left(\frac{x-1}{\theta}\right)^n\right)$$