1st Semester, 2021-2022

3rd Year, Math & CompSci

Mathematical Statistics

Seminar Exercises: Week 7

Recap. *Throughout this class, $X_1, X_2, \dots, X_n, \dots$ will be i.i.d. random variables that follow the distribution of a given characteristic X with finite mean and variance.*

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

- $E(X+Y) = E(X) + E(Y)$;
- $V(X+Y) = V(X) + V(Y)$, *if X and Y are independent;*
- *If $E(X) = \mu$ and $V(X) = \sigma^2$, then $E(\overline{X}) = \mu$ and $V(\overline{X}) = \frac{\sigma^2}{\mu}$;*
- ***The Strong Law of Large Numbers (SLLN):***

  If $(X_n)_{n\in\mathbb{N}}$ is a sequence of i.i.d. random variables with $X_n \sim X$, then

$$\overline{X}_n := \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{a.s.} E(X)$$

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

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

- *If $X \sim \mathcal{N}(\mu, \sigma)$, then:*

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \ x \in \mathbb{R}$$

*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**: $B := E(\overline{\theta}) - \theta$);*
- ***absolutely correct estimator**: $E(\overline{\theta}) = \theta$, $\lim\limits_{n\to\infty} V(\overline{\theta}) = 0$;*
- ***consistent estimator**: $\overline{\theta} \xrightarrow{p} \theta$;*
- *The **likelihood function** of the sample $X_1, X_2, \ldots, X_n$:*

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

- ***Fisher's (quantity of) information** relative to θ:*

  $$I_n(\theta) = E\left(\left(\frac{\partial \ln L(X_1, X_2, \ldots, X_n; \theta)}{\partial \theta}\right)^2\right)$$

  If the range of X does not depend on θ:

  $$I_n(\theta) = -E\left(\frac{\partial^2 \ln L(X_1, X_2, \ldots, X_n; \theta)}{\partial \theta^2}\right)$$

  or

  $$I_n(\theta) = nI_1(\theta)$$

- *The **efficiency** of an absolutely correct estimator $\overline{\theta}$ is*

  $$e(\overline{\theta}) = \frac{1}{I_n(\theta)V(\overline{\theta})}$$

  *$\overline{\theta}$ is an **efficient estimator** for θ if $e(\overline{\theta}) = 1$*

Exercise 1. Let $X \begin{pmatrix} -1 & 1 \\ \frac{1-\theta}{2} & \frac{1+\theta}{2} \end{pmatrix}$, where $\theta \in (0,1)$ is a parameter. Prove that the sample mean $\overline{X} = \frac{1}{n}\sum_{j=1}^{n} X_j$, $n \in \mathbb{N}$, is an absolutely correct estimator of θ. Is this estimator efficient?

Exercise 2. Let $X \sim Unif([0,\theta])$, where $\theta > 0$ is a parameter. Consider the estimator $\bar{\theta} = c_n \cdot \max\{X_1, X_2, \ldots, X_n\}$, where $c_n \in \mathbb{R}$ depends only on $n \in \mathbb{N}$. Find c_n such that $\bar{\theta}$ is unbiased. Is $\bar{\theta}$ absolutely correct?

Exercise 3. Let $X \sim Unid(\theta)$, where $\theta \in \mathbb{N}^*$ is a parameter, *i.e.* $X \begin{pmatrix} k \\ \frac{1}{\theta} \end{pmatrix}_{k=\overline{1,\theta}}$. Consider the estimator $\overline{\theta} = \max\{X_1, \ldots, X_n\}$, $n \in \mathbb{N}$. Prove that $\overline{\theta}$ is biased, but $E(\overline{\theta}) \to \theta$, as $n \to \infty$.

Exercise 4. Let $X \sim N(\mu, \sigma)$. For a random sample $X_1, X_2, \ldots, X_n$ we consider the estimator $\overline{s} = \frac{1}{n}\sqrt{\frac{\pi}{2}}\sum_{i=1}^{n}\left|X_i - \mu\right|$. Show that it is an absolutely correct estimator for σ and find its efficiency.

Exercise 5. Prove that the sample moment of order 2:

$$\overline{\mu}_2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$$

is a consistent estimator of the variance $V(X)$. Deduce that the sample standard deviation is a consistent estimator of the standard deviation of $\sigma = \sqrt{V(X)}$.

Hint: For a sequence $(X_n)_{n\in\mathbb{N}}$ of random variables, almost sure convergence implies convergence in probability:

$$X_n \overset{a.s.}{\to} X \Longrightarrow X_n \overset{p}{\to} X$$