1st Semester, 2021-2022

3rd Year, Math & CompSci

Mathematical Statistics

Seminar Exercises: Week 13

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.*

Let X be a random variable. The probability density functions for the laws that we will use in this problem set are:

- $X \sim N(\mu, \sigma)$: $f_X(x) = \dfrac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$, $x \in \mathbb{R}$;
- $X \sim Bern(p)$: $f_X(x) = p^x(1-p)^{1-x}$, $x \in \{0, 1\}$;
- $X \sim Exp(\lambda)$: $f_X(x) = \lambda e^{-\lambda x}$, $x > 0$;
- $X \sim \Gamma(a, b)$: $f_X(x) = \dfrac{1}{\Gamma(a) b^a} \cdot x^{a-1} e^{\frac{-x}{b}}$, $x > 0$;
- $X \sim Poisson(\lambda)$: $f_X(x) = e^{-\lambda} \dfrac{\lambda^x}{k!}, x \in \mathbb{Z}_{\geq 0}$

Proposition. Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed random variables, that follow the distribution of the variable X.

- If $X \sim N(\mu, \sigma)$, then $X_1 + \ldots + X_n \sim N(n\mu, \sigma\sqrt{n})$;
- If $X \sim Exp(\lambda)$, then $X_1 + \ldots + X_n \sim Gamma(n, \frac{1}{\lambda})$;

- **Central Limit Theorem:** If n is big enough, then:

$$\frac{\frac{X_1+\ldots+X_n}{n} - E(X)}{\frac{\sigma_X}{\sqrt{n}}} \sim N(0,1)$$

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

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

Type I error rate (Significance level):

$$\alpha = P(\textit{type I error}) = P(\textit{reject } H_0 \mid H_0) = P(TS \in RR \mid \theta = \theta_0)$$

Type II error rate:

$$\beta = P(\textit{type II error}) = P(\textit{not reject } H_0 \mid H_1) = P(TS \notin RR \mid H_1)$$

Power of a Test on a parameter θ*:*

$$\pi(\theta^*) = P(\textit{reject } H_0 \mid \theta = \theta^*) = P(TS \in RR \mid \theta = \theta^*) = 1 - \beta(\theta^*)$$

Theorem (Neyman-Pearson)**.** *Let X be a characteristic with pdf $f(x;\theta)$, with $\theta \in A \subseteq \mathbb{R}$ unknown. Suppose we test on θ the simple hypotheses:*

$$H_0 : \theta = \theta_0$$

$$H_1 : \theta = \theta_1$$

based on a random sample $X_1, X_2, \ldots, X_n$. Then for a fixed $\alpha \in (0,1)$, a ***most powerful test*** *is the test with rejection region given by*

$$RR = \left\{ \frac{L(\theta_1)}{L(\theta_0)} \geq k_\alpha \right\}$$

where the constant $k_\alpha > 0$ depends only on α and the sample variables.

<u>Remark:</u> *Finding such a most powerful test amounts to finding k_α so that $\alpha = P((X_1, \ldots, X_n) \in RR \mid H_0)$*

Exercise 1. Let X be a characteristic that has a normal distribution $N(m, 1)$, where $m \in \mathbb{R}$ is unknown.

(a) For a random sample of size 9 for X, find a most powerful test with the significance level 5% for the hypothesis $H_0 : m = 0$ against $H_1 : m = 1$. Find the power of this test.

(b) For the following sample data of X:

$$-1, 0.5, -0.25, -0.75, 0.75, 1.25, 0.5, 1, 0.25$$

accept or reject the hypothesis $H_0 : m = 0$ against $H_1 : m = 1$, using the obtained test.

Exercise 2. Let X be a characteristic that has a Bernoulli distribution with parameter p, where $p \in (0, 1)$ is unknown.

(a) For a random sample of (large) size n for X, find a most powerful test with the significance level 5% for the hypothesis $H_0 : p = 0.5$ against $H_1 : p = 0.45$.

(b) Let p be the probability of heads for a coin. Assume we toss the coin 900 times and we get 441 heads. Accept or reject the hypothesis $H_0 : p = 0.5$ against $H_1 : p = 0.45$, using the obtained test.

Exercise 3. Let X be a characteristic that has an $Exp(\lambda)$ distribution, where $\lambda > 0$ is unknown.

(a) For a random sample of size 10 for X, find a most powerful test with the significance level 5% for the hypothesis $H_0 : \lambda = 1$ against $H_1 : \lambda = 3$. Find the power of this test.

(b) Assume that X is the time (in minutes) spent by a client to buy a bus ticket from a vending machine and that the total time spent by a random sample of 10 clients is 4.9 minutes. Accept or reject the hypothesis $H_0 : \lambda = 1$ against $H_1 : \lambda = 3$, using the obtained test.

Exercise 4. Let X be a characteristic that has a Poisson distribution with parameter λ, where $\lambda > 0$ is unknown.

(a) For a random sample of size 100 for X, find a most powerful test with the significance level 5% for the hypothesis $H_0 : \lambda = 10$ against $H_1 : \lambda = 11$. Find the power of this test.

(b) Assume that X is the number of received calls during an hour in a call center and that for a random sample of 100 hours we have o total of 1062 calls. Accept or reject the hypothesis $H_0 : \lambda = 10$ against $H_1 : \lambda = 11$, using the obtained test.