Probability and Statistics

Y-DATA School of Data Science

P&P 6

Due: 14.12.2022

- This final assignment contains one question from each topic.
- The work is individual, NOT in pairs.
- The weight of the assignment is 1/6 (i.e. no extra weight).

PROBLEM 1. During military training, three aircraft independently attacked a target, each exactly once. The first aircraft usually hits the target with a 0.6 probability, the second with a 0.4 probability, and the third with a 0.3 probability. After the training, it turned out that only one aircraft had hit the target. What is the probability that the 3rd one hit it?

PROBLEM 2. A fair die is rolled until the total sum of points becomes 3 or more. Let X denote the number of rolls until the above happens. Find the PMF and the CDF of X. Calculate $E(X)$ and $Var(X)$.

PROBLEM 3. Let X be a continuous random variable with PDF

$$f_X(x) = x/4, \text{ if } 1 < x \leq 3 \text{ and 0 otherwise}$$

(1) Verify that $f_X(x)$ is indeed a PDF.
(2) Find $E(X)$.
(3) Let $Y = X^2$. Find $E(Y)$ and $Var(Y)$.
(4) Let $A = \{X \geq 2\}$. Find $P(A)$, $f_{X|A}(x)$ and $E(X|A)$.

PROBLEM 4. Let $X_1, ..., X_n$ be an i.i.d. sample from the distribution with density

$$f_\theta(x) = \frac{x}{\theta} e^{-\frac{x^2}{2\theta}}, x \geq 0, \theta \geq 0$$

(1) Find the maximum likelihood estimator for θ.
(2) Evaluate the MLE for the sample $(0.5, 0.5, 1)$

PROBLEM 5. It is known that an existing drug improves the health situation of 40% of the patients. Let p be the probability of a positive effect of a **new** drug. After making sure that the side effects of the new drug are not different than those of the existing drug, the ministry of health wants to test whether the new drug has a bigger rate of success. To this end, a random sample of size n is drawn and a hypothesis test is planned.

(1) Formulate the hypotheses H_0 and H_1.
(2) We reject the null if $\{\hat{p} > c\}$. Find c using the normal approximation for significance level $\alpha = 0.05$.
(3) Compute the power of the test at the point $p = 0.55$ (you should get some function of n).
(4) Following the previous part, find the minimal sample size n for which the power at the point $p = 0.55$ is at least 0.8.