# Exam Statistics $29^{th}$ October 2021 (A)

Maura Mezzetti

## NAME:

## ID:

**This is a closed book exam. Answer all the following questions and solve all the following exercises. You have two hours to complete the exam.**

## Exercise 1

1. The test statistic

$$\frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}}$$

   for testing the mean of a normal population follows t-distribution when

   (a) $\sigma$ is known and n is small

   (b) $\sigma$ is unknown and n is small

   (c) $\sigma$ is known and n is large

   (d) $\sigma$ is unknown and n is large

2. Best critical region (UMP) to test a simple hypothesis $H_0$ against a simple $H_1$ is given by

   (a) Rao-Blackwell Theorem

   (b) Factorization Theorem

   (c) Cramer-Rao lower bound

   (d) Neymann-Pearson Lemma.

3. Which of the following will increase the power of a test?

   (a) Increase $n$

   (b) Increase $\alpha$

   (c) Reduce the amount of variability in the sample

(d) Consider an alternative hypothesis further the null

(e) All of the previous will increase the power of the test

4. Let $T$ be the (MVUE) minimum variance unbiased estimator for $\theta$. Which of the following is true?

   (a) $Var(T) <$ Cramer-Rao lower bound

   (b) $Var(T) =$ Cramer-Rao lower bound

   (c) $Var(T) >$ Cramer-Rao lower bound

   (d) $Var(T) \geq$ Cramer-Rao lower bound

## Exercise 2

Let $(X_1, \ldots, X_n)$ be independent identically distributed random variables with p.d.f.

$$f(x) = \theta(1-x)^{\theta-1} I_{(0,1)}(x) \qquad \theta > 1$$

1. Find a sufficient statistics for $\theta$
2. Find $\hat{\theta}_{MLE}$ maximum likelihood estimator (MLE) for $\theta$ and discuss properties of this estimator.
3. Compute the score function and the Fisher information.
4. Specify asymptotic distribution of $\hat{\theta}_{MLE}$.
5. Suppose form a random sample of 200 random variables you obtain $\sum_{i=1}^{200} log\,(1-x_i) = -50$. Let $\alpha = 0.05$, Complete ALL of the following questions:

   (a) Find the Likelihood Ratio test statistic for testing $H_0 : \theta = 5$ versus $H_1 : \theta \neq 5$, specify the asymptotic distribution and verify the null hypothesis.

   (b) Find the Wald test statistic for testing $H_0 : \theta = 5$ versus $H_1 : \theta \neq 5$, specify the distribution and verify the null hypothesis.

   (c) Find the Score test statistic for testing $H_0 : \theta = 5$ versus $H_1 : \theta \neq 5$, specify the distribution and verify the null hypothesis.

## Exercise 3

Let $(X_1, \ldots, X_n)$ be a random sample of i.i.d. random variables distributed as follows:

$$f(x;\theta) = \frac{2}{\theta^2}(\theta - x), \qquad 0 \leq x \leq \theta$$

1. Apply the method of the moments to find the estimator $\hat{\theta}_{MOM}$ of the parameter $\theta$.
2. Calculate the bias and the mean square error of the estimator $\hat{\theta}_{MOM}$
3. Study the consistency of $\hat{\theta}_{MOM}$
4. What does it happen when you try to apply the maximum likelihood method to find and estimator of the parameter $\theta$?

## Exercise 4

The literacy rate of a nation measures the proportion of people age 15 and over who can read and write. A statistician is interested in the estimation of parameter $\theta$, the literacy rate for women in Afghanistan and she has to choose between random samples from a Bernoulli or a geometric distributions (with the same $\theta$).

i.e. Given $n$ she can choose between the two following experiments:

1. She randomly select $n$ Afghani women and count the number of women who are literate.
2. She run $n$ different experiments, for each experiment, she keeps selecting women untill she finds a literate one; $x_i$ is the number of Afghani women she asks until one says that she is literate.

Which of the two experiment will give the more precise inference on $\theta$ (Justify your answer).

### Reminder of the geometric distribution

Geometric distribution is used for modeling the number of failures until the first success:

$$P(X = k) = (1 - \theta)^k \theta \qquad E(X) = \frac{1 - \theta}{\theta}$$

## Exercise 5

The manager of a gas station wants to estimate the median time necessary to change of oil. The actual time varies from costumer to costumer. However, one can assume that this time will be well represented by an exponential random variable. The following values were recorded from 10 clients randomly selected (the time is in minutes):

$$4,\ 5,\ 5,\ 3,\ 7,\ 8,\ 8,\ 5,\ 5,\ 10$$

Propose $MLE$ estimator and compute an estimate of the median time necessary to change oil.

## Exercise 6

Provide correct statement for Neyman Pearson Lemma and provide an example of application.