STATISTICS – Exercises SESSION N° 1

About Chapter :

- Chapter 1 : DESCRIPTIVE STATISTICS

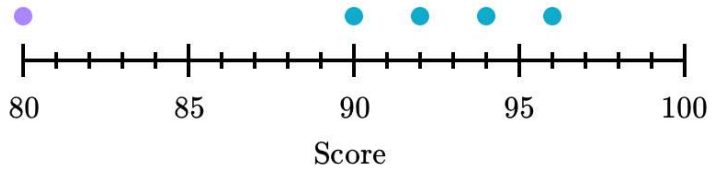
TD1 – 1 :

1. What is the median of the following statistical series? 1, 2, 4, 6, 4
2. The following table gives the distribution of the number of points scored by each player during a Quidditch season.

Player	Fred	George	Olivier	Angelina	Harry
Number of points	11	12	7	3	?

If the team scored on average 8 points, how many points did Harry score?

3. Anna played 5 rounds of golf this year. During the final round she obtained the lowest score of 80 points.



What is the impact of this last game on the mean and on the median of the scores of the previous games? Interpret the results.

TD1 – 2 :

Calculate the mean, median and mode for the following samples:

1) (3 ; 5 ; 2 ; 6 ; 5 ; 9 ; 5 ; 2 ; 8 ; 6)

2) (1.28 ; 2.16 ; 0.75 ; 1.44 ; 2.05 ; 0.65 ; 1.26 ; 1.73 ; 1.81 ; 0.92)

Last update on October 27th, 2022

TD1 – 3 :

The table below represents the distribution of the intelligence quotient (IQ) of 100 students. This distribution is grouped into 9 classes of width 10.

Class mark (mid-value)	59.5	69.5	79.5	89.5	99.5	109.5	119.5	129.5	139.5
frequency	1	2	9	22	33	22	8	2	1

1) Calculate the mean and standard deviation of this distribution

2) Determine the values of the 1st and 3rd quartile of the distribution. Calculate the Semi-interquartile range.

3) Calculate the coefficient of symmetry $\gamma_1 = \frac{\mu_3}{\sigma^3}$ and the coefficient of kurtosis $\gamma_2 = \frac{\mu_4}{\sigma^4} - 3$. Compare these values to the standard normal distribution.

STATISTICS – Exercises SESSION N° 2

About chapters :

- Chapter 2 : Statistical theory of estimation

TD2 – 1 :

Let $x_1,\ldots,x_n$, be a sample of size n, results of independent observations of a random variable X whose probability law depends on a parameter. Determine the estimators of this parameter by the Maximum Likelihood method and calculate their precision for the following situations:

1) X follows an exponential law $f_X(x,\lambda) = \lambda e^{-\lambda x}$ for $x \geq 0$ and 0 otherwise.
2) Let us define $a = 1/\lambda$. Calculate the minimum variance (Cramer - Rao bound) for an estimator of the parameter a. Draw a conclusion.

2) X follows a Bernoulli law with parameter p

Remark : Here you must conclude by indicating whether the estimator is efficient and with minimum variance for the two situations.

TD2 – 2 :

An urn contains an unknown proportion of red balls and white balls. A sampling with replacement of size 60 gave a percentage of 70% for the red balls.

Calculate the 95% and 99.73% confidence limits of the proportion of red balls in the urn.

TD2 – 3 :

We assume that the random variable X that represents the weight of a given object follows a Gaussian distribution of mean *m* and standard deviation σ .

The following table shows the results of 10 weighings of the same object (in grams):

72.20	72.24	72.26	72.30	72.36	72.39	72.42	72.48	72.50	72.54

1) Calculate the point estimates of the mean *m* and of the standard deviation σ of the variable X.

2) The standard deviation is unknown but you can consider the estimated value in question 1), calculate a confidence interval at the threshold of 5% of the mean *m* .

Last update on October 27th, 2022

Now, consider the variance is known and equal to the value calculated in 1), how does the interval change? Calculate and compare the 2 intervals.

3) The standard deviation of the balance, measured from many previous studies, is indeed, for an object of about this mass, 0.08. In this question, we consider then $\sigma = 0.08$. Determine the level of significance α so that the resulting confidence interval is $[72.31; 72{,}43]$.

TD2 – 4 :
Consider the following sample:

Sample : (1.28 ; 2.16 ; 0.75 ; 1.44 ; 2.05 ; 0.65 ; 1.26 ; 1.73 ; 1.81 ; 0.92)

Calculate the sample mean and sample variance. Deduce the confidence intervals for the mean, variance and standard deviation of the population. We will assume that the random variable follows a Gaussian law.

STATISTICS – Exercises SESSION N° 3

About Chapters :

- Chapter 3 : Hypothesis testing

TD3 – 1 :

During an extrasensory perception experiment, a person is isolated in a room and tasked to guess the color (red or blue) of a card chosen randomly from a 50-card deck. The cards were previously well shuffled. The person does not know the total number of blue or red cards in the deck.

Suppose the person correctly guessed the color of 32 cards. Determine if it is possible to conclude that the person possesses extrasensory perception (sixth sense) at the 0.05 and 0.01 significance level.

TD3 – 2 :

A manufacturer delivers a batch of bulbs with an average lifetime specification of 1600 hours. After inspection on a sample of 100 bulbs, the average lifespan is 1,578 hours with a standard deviation of 120 hours.

1) Can the delivery be considered to meet specification (that is, the average lifetime is 1600 hours) at the 5% significance level?

2) In fact, what the buyer is interested in is whether the lifetime is equal to or greater than the specification. Can the delivery be considered to meet this condition at the 5% significance level?

TD3 – 3:

You are hired by TOYOTA for a project to design a low-consumption car model. First of all, it is necessary to ensure that fuel consumption is the primary factor behind the purchase of this type of car. After performing a survey of 120 owners of this type of car, 40 considered that the criterion "consumption" was the decisive criterion of purchase.

1) Let p *be* the proportion of owners who favor the “consumption” criterion. You are tasked to test the following hypotheses at a given level of significance α :

$H_0 : p = 0.4$

$H_1 : p \neq 0.4$

Suppose that the proportion of owners who favored the criterion " consumption" within the total population is exactly 40%. What is the probability of concluding that H_0 should be rejected? Test these hypotheses by considering the level of significance $\alpha = 0.05$.

2) Build a confidence interval of 95% for the proportion of owners whose choice was motivated mainly by the fuel consumption.

3) The company claims that, for this model of cars, the average fuel consumption μ is equal to 7 liters per 100 km on urban roads (50 km/h). On a random sample of 5 cars (the population is assumed to be normal), the consumption was (in liters per 100 km): 8.1; 8.5; 8.9; 9.7; 9.8

At the threshold $\alpha = 0.05$, do you confirm the consumption announced by the manufacturer?

TD3 – 4 *:

Two groups A and B are each composed of 100 people with the same disease. A new drug was given to group A and placebo treatment to group B (referred to as a control group). It was observed that 75 patients from group A and 65 from group B were cured.

1) Test the hypothesis that the new drug is effective in the cure of the disease by considering a significance level of 1%.

2) Answer the previous question by considering groups of 300 people, and the fact that there were 225 people healed in group A and 195 in group B. Any Comments? Compare both results.