Block 1. Descriptive statistics

PROPOSED PROBLEMS

Problem 1.1. The statistics grades for a group of students were recorded. The data were then organised, with the following result:

Grades	0	1	2	3	4	5	6	7	8	9	10
Number of students	4	2	6	15	5	16	20	6	4	2	2

a. Determine the appropriate frequency distribution for the grades.
b. Find the percentage of students that passed the subject (those who got a grade of 5 or higher).
c. Find the percentage of students that got grades higher than 7.
d. If you had to grant 8 students a place in the Erasmus programme, by how much should you increase the passing grade?
e. Graphically represent the appropriate distributions in this problem.

Problem 1.2. The values relative to the number of students and engineering faculties in a given country are:

Students	0-100	100-200	200-300	300-400	400-500	500-600	600-700	700-800	800-900
Number of faculties	18	37	11	19	20	12	5	3	2

a. Construct the table of frequencies appropriate for the data.
b. Find the number of faculties with more than 400 students.
c. Find the percentage of faculties with more than 200 and less than 500 students.

Problem 1.3. In a laboratory, a magnitude was measured and measurements were made for several values of that magnitude, as reported in the following table:

x_i	1	3	4	6	10
n_i	5	12	20	8	5

a. Study the symmetry of this distribution both analytically and graphically.
b. Quantify the degree of pointedness for this distribution.
c. Study the normality of this distribution.

Notice that this is an integral problem. You must therefore perform a complete analysis.

Problem 1.4. The distribution of people's income is described using the Pareto principle, . Here, and represent the percentage of people with an income equal to or lower than and the minimum income, respectively. On the other hand, is a structural parameter. Adjust the Pareto principle to the following data and assess the result.

$\ln x$	0,8	1,1	1,4	1,7	2,1	2,6	3,0	3,4	4,0
y	0,02	0,100	0,190	0,350	0,500	0,745	0,800	0,90	1

Problem 1.5. The Cobb-Douglas production function, , can relate an investment to the gross domestic product. and are constants that represent the output elasticities of labour and capital. Adjust the function to the data:

y_i	2,5	2,8	3,2	4,0	5,0	5,9	7,1	9,0	11,4	13,3	14,9	17,0	20,0
K_i	0,5	0,6	0,7	1,1	1,4	1,4	1,7	1,8	2,3	2,6	2,8	3,6	3,9

Determine the best fit that allows us to obtain a regression curve. Assess the result.

a. Determine the mean, median, standard deviation, variance, and coefficients of skewness and kurtosis for the three variables.
b. Construct the correlation matrix and find the degree of dependency between the variables.
c. Construct the covariance matrix and interpret the results.