EXAM OF STATISTICS
(DESCRIPTIVE STATISTICS AND REGRESSION)

2nd Physiotherapy **Version A** **March, 25 2022**

Name: DNI: Group:

Duration: 1 hour.

(5 pts.) 1. The time required by a drug A to be effective has been measured (in minutes) in a sample of 150 patients. The table below summarize the results.

Response time	Patients
$(0, 5]$	5
$(5, 10]$	15
$(10, 15]$	32
$(15, 20]$	36
$(20, 30]$	42
$(30, 60]$	20

(a) Are there outliers in the sample? Justify the answer.

(b) What is the minimum time for the 20% of patients with highest response time?

(c) What is the average response time? Is the mean representative?

(d) Can we assume that the sample comes from a normal population?

(e) If we take another sample of patients with mean 18 min and standard deviation 15 min, in which group is greater a response time of 25 min?

Use the following sums for the computations: $\sum x_i n_i = 3105$ min, $\sum x_i^2 n_i = 83650$ min^2, $\sum (x_i - \bar{x})^3 n_i = 206851.65$ min^3 y $\sum (x_i - \bar{x})^4 n_i = 8140374.96$ min^4.

(5 pts.) 2. We have measured the average number of weekly hours of study X and the score in a clinic entrance test Y of 8 candidates, getting the following results:

$\sum x_i = 15.9$, $\sum \log(x_i) = 3.852$, $\sum y_j = 41.5$, $\sum \log(y_j) = 11.511$,
$\sum x_i^2 = 42.23$, $\sum \log(x_i)^2 = 5.559$, $\sum y_j^2 = 274.25$, $\sum \log(y_j)^2 = 20.742$,
$\sum x_i y_j = 103.3$, $\sum x_i \log(y_j) = 28.047$, $\sum \log(x_i) y_j = 32.616$.

(a) Compute the equations of the logarithmic and exponential regression models of the score as a function of the average number of hours of study.

(b) Use the best of the previous models to predict the score for somebody that study 3.2 hours a week. Is this prediction reliable?