

EXAM OF STATISTICS**2nd Physiotherapy****Version A****June, 28 2018**

Name:

DNI:

Group:

Duration: 2 hours.

- (2.5 pts.) 1. The time required by a drug A to be effective has been measured in a sample of 150 patients. The table below summarizes 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? Justify the answer.
- (d) Study the kurtosis of the sample.
- (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: $\sum x_i = 3105 \text{ min}$, $\sum x_i^2 = 83650 \text{ min}^2$, $\sum (x_i - \bar{x})^3 = 2277448.69 \text{ min}^3$ y $\sum (x_i - \bar{x})^4 = 82723757.55 \text{ min}^4$.

- (1.5 pts.) 2. A 40% of a population of athletes have a very athletic mother and a 30% a very athletic father. If 50% of athletes in this population have some very athletic progenitor:

- (a) Compute the probability that a random athlete of this population have a very athletic mother if he or she has a very athletic father.
- (b) Compute the probability that a random athlete of this population have a very athletic father if he or she has a non very athletic mother.
- (c) ¿Are the events corresponding to having a very athletic father and having a very athletic mother independent? Justify the answer.

- (2.5 pts.) 3. 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:

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

- (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.

- (1.5 pts.) 4. The average number of injuries in an international tennis tournament is 2.

- (a) Compute the probability that in an international tennis tournament there are more than 2 injuries.
- (b) If a tennis circuit has 6 international tournaments, what is the probability that there are no injuries in some of them?

(2 pts.) 5. A diagnostic test to determine doping of athletes returns a positive outcome when the concentration of a substance in blood is greater than $4 \mu\text{g/ml}$. If the distribution of the substance concentration in doped athletes follows a normal distribution model with mean $4.5 \mu\text{g/ml}$ and standard deviation $0.2 \mu\text{g/ml}$, and in non-doped athletes is normally distributed with mean $3 \mu\text{g/ml}$ and standard deviation $0.3 \mu\text{g/ml}$,

- (a) what is the sensitivity and specificity of the test?
- (b) If there is a 0.1% of doped athletes in a competition, what are the positive and the negative predicted values? Interpret them.