Statistics I Final Exam, 23 June 2015.
Degrees in ADE, DER-ADE, ADE-INF, ADE-INT, FICO, ECO, ECO-DER, TUR.

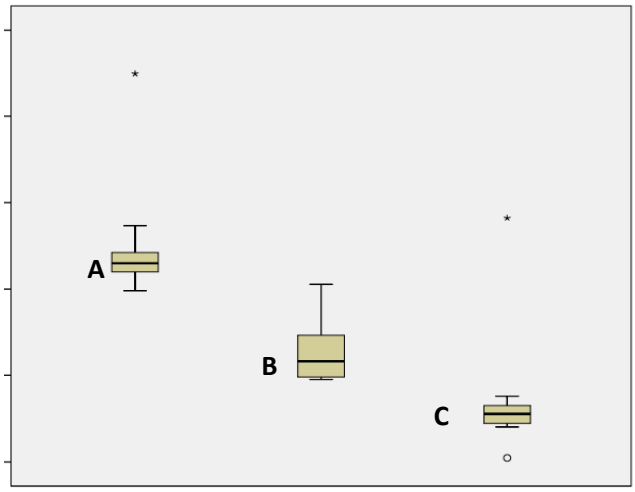
EXAM RULES: 1) Use separate booklets for each problem. 2) Perform the calculations with at least two significant decimal places. 3) You may not leave the exam during the first 30 minutes. 4) You are not allowed to leave the classroom without handing in the exam.

1. The following tables contain information about 10 companies of the Dow Jones. In particular: X_1 ="highest-paid CEO (in million dollars)" and X_2 ="share price (in dollars)". Source: El País, 8th May 2016.

Tabla 1 / Table 1

X1	X2
44,91	106,11
27,29	87,81
24,2	53,95
23,79	113,69
23,37	29,97
22,58	158,31
22,03	100,31
21,98	64,21
20,01	110,68
19,82	147,6

Figura 1 / Figure 1



   (a) (1 point) Draw the box-plot for X_2 and identify the outliers (if any). Justify your answer.
   (b) (0.5 points) Determine if X_1 and X_2 have the same type of asymmetry. Justify your answer.
   (c) (0.5 points) Determine which box-plot (A, B, C) corresponds to X_1. Justify your answer.
   (d) (0.5 points) If 1 euro = 1.14 dollars, calculate the average salary for those CEOs (in million euros) and the variance.

2. The weekly amount of newspapers and magazines (in tens) that are sold in shopping center of a given city can be represented by a random variable with the following probability density function:

$$f(x) = \begin{cases} k\,(x-1)(3-x), & 1 \le x \le 3, \\ 0, & \text{otherwise}, \end{cases}$$

   where k is an appropriate constant.

   (a) (0.5 points) Find the value of k so that $f(x)$ can be considered a probability density function.
   (b) (0.5 points) Obtain the cumulative probability function.
   (c) (1 point) Calculate the probability that in a given week, the amount of sold newspapers and magazines is greater than the mean.
   (d) (0.5 points) Calculate the variance of the weekly amount of sold newspapers and magazines.

3. In a city of 3.5 millon inhabitants there are three urban transport systems: metro, bus and tram. In general, in a working day, the amount of travellers are 1.500.000 for the metro, 750.000 for the bus and 450.000 for the tram. Moreover, it is known that, 30% of metro travellers also use the bus, 10% of metro travellers also use the tram and 5% of metro travellers also use both bus and tram. Finally, 15% of bus travellers also use the tram. (Hint: An inhabitant can take or not the urban transport).

   (a) (0.75 points) Calculate the probability that, in a working day, an inhabitant uses only one of the three transport systems.

   (b) (0.5 points) Calculate the probability that, in a working day, an inhabitant uses at least one transport system.

   (c) (0.75 points) When only a transport system is used, there is a 2% probability of having a delay of more that 5 minutes in a working day. However, the probability of having such a delay rises to 7% when combining more than one transport system in a working day. Calculate the probability that an inhabitant suffers a delay of more that 5 minutes in a working day.

   (d) (0.5 points) With the same information as in part (c) and given that a traveller suffered a delay of more than 5 minutes, calculate the probability that this traveller took more than one transport system.

4. Let X be a r.v. taking values on the set $S = \{2, 3, 5, 7\}$, with the following probability function:

$$P(X = x) = \begin{cases} 0.10 & \text{if } x = 2, \\ 0.25 & \text{if } x = 3, \\ 0.35 & \text{if } x = 5, \\ 0.30 & \text{if } x = 7, \\ 0 & \text{otherwise.} \end{cases}$$

   (a) (1 point) Consider a simple random sample of size $n = 150$ with the same distribution of X. Calculate the probability that the sample mean takes values between 4.6 and 5.0.

   (b) (0.5 points) Compare the result obtained in part (a) with the lower bound that would be obtained using Chebyschev's inequality. Is this result contradictory? Explain when it is appropriate to use the approximation given by Chebyschev's inequality.

   (c) (0.5 points) Let Y be a new r.v. being equal to 1 if $X < 5$, and 0 otherwise. Determine the probability model of Y, give its expectation and variance.

   (d) (0.5 points) In a sample of size $n = 250$, it happened that 75 times X took values smaller than 5. Obtain the 98% confidence interval for the mean of Y.