Second Midterm - Fall 2016

Instructions

- The use of a mobile phone, or any other means of communication, is forbidden.

- Give all the results with **a two decimals** precision.

- Write answers on a **separate** sheet of paper.

Section are independent and can be treated separately.

1   Exercise I - Warm-Up (1 point)

In a study on the salaries in a firm, the median (M_e), the deciles ($D_1, D_2, \ldots, D_9$) and the quartiles (Q_1, Q_2, Q_3) have been calculated. The salaries ranged between € 1,200 and € 12,000.

Quantile	D_1	Q_1	M_e	D_8	D_9
Salary in €	1,500	2,500	4,000	8,000	10,000

1. Reconstruct the distribution of employees by *salaries bracket*.

2. (Bonus: +1pt) What is the standard deviation of the distribution. (Hint: make the standard assumption of uniformed distribution of wages *within* the brackets).

2   Exercise II - Live from the staffroom (5 points)

A class is divided in two groups of 10 students each for an exam. The results are the following :

Group I	5	14	10	11	14	13	10	9	8	5
Group II	4	11	15	13	6	11	10	12	12	9

1. Give the mode, the mean, the median for each group

2. Give the absolute deviation for **Group 1**

3. Give the standard deviation and the coefficient of variation for **Group 2**

4. Give the lowest quartile, the upper quartile and the interquartile range (IQR) for each group.

A **Tukey BoxPlot** is a standard boxplot with the lower fence of the boxplot equals the lowest quartile minus 1.5 times the IQR whereas the upper fence equals the upper quartile plus 1.5 times the IQR.

Formally :

$$\text{Lower fence} = Q1 - 1.5 * IQR$$
$$\text{Upper fence} = Q3 + 1.5 * IQR$$

5. Draw a Tukey boxplot for each group. Comment.

3 Exercise III - Unions statistics (8 points)

A firm has the following wages :

Wages	Nb of people
10k - 20k	200
20k - 30k	800
30k - 50k	180
50k - 90k	20

1. Determine the following values and explain in a few words what they mean :

   - the mean salary
   - the median salary
   - the standard deviation
   - the medial

2. Draw the Lorenz Curve

3. Determine the Gini coefficient

4 Exercise IV - Hold-up (5 points)

John is borrowing $1,000 at an annual interest rate of 5%.

1. What is the duration of the loan if John pays as many interests as principal ?

2. How many interests should he pay if he makes a 10 years loan ?

3. John makes a 10 years loan. What is the *real* interest rate if he pays the interest at the beginning ?

Samatha invests $5,000 for six months at a 4% annual rate.

4. How many dollars will she earn at the end of the investment ?

5. At the end of the six months, Samantha wants to reinvest the money she earned plus the interests. She has the choice between:

- Invest for one year at a 5% annual rate

- Make a first investment for six months at a 4% annual rate and a second in six months at a 4% annual rate of the principal and interests earned after the first investment.

What is the best choice ?

5   Exercise V - Cop 22 (2 points)

Green is better, an environmental NGO, had measured daily pollution on numerous european cities. They want to estimate the correlation between the averaged dailu pollution and a city's population. Doing some preliminaries statistics, they find the following results :

Statistics	Value
$mean_{pollution}$	5
$stdev_{pollution}$	10
$mean_{population}$	50,000
$stdev_{population}$	1,000
$Cov(pollution, population)$	6.89

Then, they estimate the following regression.

$$pollution = a * population + b + \epsilon_t \quad R_2 = 0.34$$

1. Using the preliminaries statistics, can you find a and b ?

2. What does the R_2 measure (give a <u>precise</u> answer) ?

3. (bonus + 0.5) How could you improve the model ?

***