Second Midterm - Spring 2017 - Results

1 Exercise I - Means of distribution (4 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_2	Q_1	M_e	Q_3	D_8	D_9
Salary in €	2,000	3,000	4,000	7,000	9,000	12,000

1. To reconstruct the distribution of employees by salaries bracket, we use the definition of decile, quantiles and median. We know that:
   - 20% of the employees earn less that D_2. Since the minimum wage in the firm is 1200, 20% have a wage between D_2 and 1200.
   - 25% earn less that Q_1, and 5% have a wage between D_2 and Q_1.
   - 50% earn less that M_e and 25% have a wage between Q_1 and M_e.
   - 75% earn less that Q_3 and 25% have a wage between M_e and Q_3.
   - 80% earn less that D_8 and 5% have a wage between Q_3 and D_8.
   - 90% earn less that D_9 and 10% have a wage between D_8 and D_9.
   - 10% earn more than D_9. Since the maximum wage in the firm is 12000, 10% of the employees have a wage of 12000.

   We are then able to construct the following table:

Wages	Frequencies	Cumulative Frequencies
1200 - 2000	20	20
2000 - 3000	5	25
3000 - 4000	25	50
4000 - 7000	25	75
7000 - 9000	5	80
9000 - 12000	10	90
12000 - 12000	10	100

2. The modal classes are 3000 - 4000 and 4000 - 7000 with 25% of the employees each.
   To compute the **mean**, we assume that in each wages' bracket, employees earn the midpoint of the bracket. This assumptions leads us to assume that the average wage of the 20% least paid employees is 1600€ (the middle of the first wages' bracket). The mean is then the average of the middle each wages' bracket:

$$\begin{aligned} Mean &= \frac{20}{100} * 1600 + \frac{5}{100} * 2500 + ... + \frac{10}{100} * 10500 + \frac{10}{100} * 12000 \\ &= 5345€ \end{aligned}$$

To compute the **standard deviation**, we make the same assumption and compute the square root of the weighted square deviation from the mean:

$$St.Dev = \sqrt{\frac{20}{100} * (1600 - 5345)^2 + \frac{5}{100} * (2500 - 5345)^2 + ... + \frac{10}{100} * (12000 - 5345)^2}$$
$$= 3391.97€$$

Hint: Do this calculus on Excel, this is faster !

2 Exercise II - Gini and Freddie (6 points)

Freddie Mac, a mortgage firm, has the following wages' distribution :

Wages	Nb of people	Cumulative frequencies	Middle of the bracket	Wages	Cumulative Wages
10k - 20k	150	150	15k	2250k	2250k
30k - 40k	800	950	35k	28000k	30250k
50k - 70k	50	1000	60k	3000k	33250k

1. First, we add four columns: cumulative people, middle of each bracket, sum of wages given in a bracket (equals the product of the number of people and the middle of each bin) and cumulative sum of wages.
   The **median** wage is the wage that splits the distribution of employees in two. There is 150 + 800 + 50 = 1000 employees in the firm. Thus, the median salary is the average of the wage of the 500^{th} employee and the 501^{th} employee. Both are in the second wages bracket: 30k - 40k.

$$median = 30 + 10 * \frac{500.5 - 150}{800}$$
$$= 34.38k$$

   The **standard deviation** is a measure of how spread out or dispersed a distribution is. The computation is the same as in Exercise I (first you need to compute the mean).

$$Mean = 33250$$
$$St.Dev = 9390.820$$

   The **medial** splits the distribution of *wages* in two equal parts. The total amount of wages is 33,250€ (see the Table). Half of it is 16,625€, which ''belongs' to the second bracket. The medial is thus:

$$Medial = 30 + 10 * \frac{16625 - 2250}{28000}$$
$$= 35.13k$$

2. To draw the **Lorenz Curve**, we need to compute *relative* cumulative frequencies and wages. This is done in the following table:

Wages	Relative cumulative frequencies	Relative cumulative wages
10k - 20k	15%	6.77%
30k - 40k	95%	90.98%
50k - 70k	100%	100%

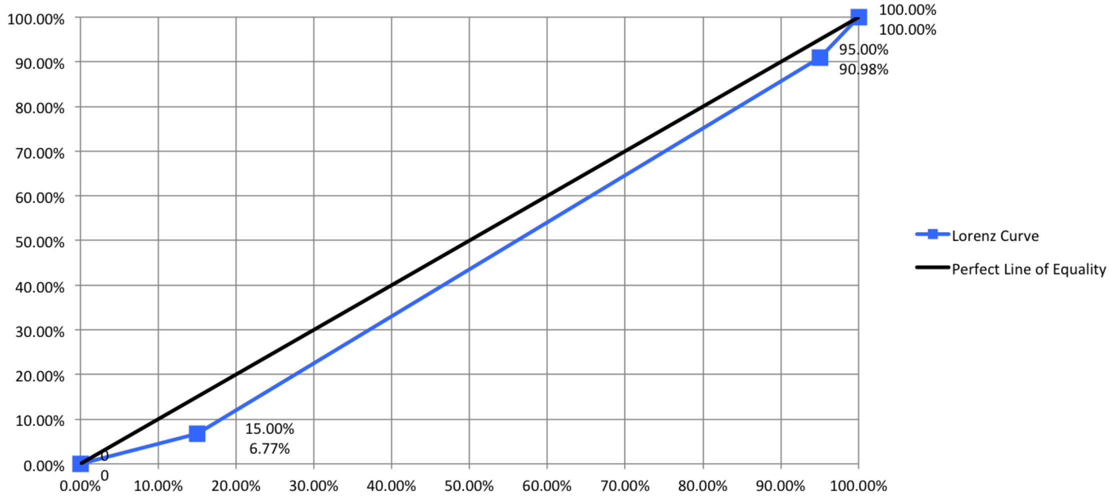

Figure 1: Lorenz Curve

3. To compute the **Gini coefficient**, we divide the B area into three areas, a triange B1, and two trapezoide B2 and B3.

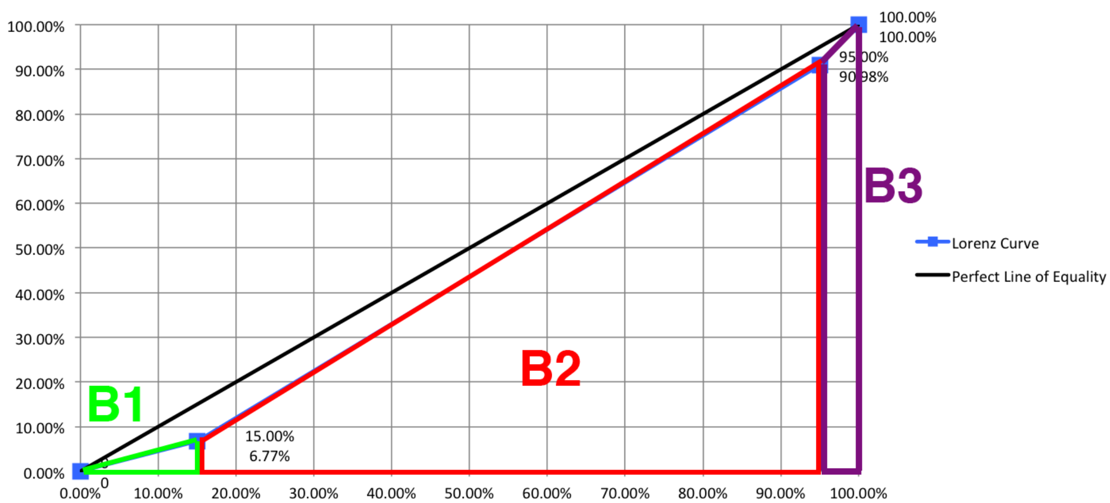

Figure 2: Lorenz Curve

B1 is a triangle and B2 and B3 are rectangular trapezoid :

$$
\begin{aligned}
B1 &= \frac{15}{100} * \frac{6.77}{100} * \frac{1}{2} \\
&= 0.005 \\
B2 &= \frac{95-15}{100} * \frac{6.77+90.98}{100} * \frac{1}{2} \\
&= 0.391 \\
B3 &= \frac{5}{100} * \frac{90.98+100}{100} * \frac{1}{2} \\
&= 0.048
\end{aligned}
$$

We compute $B = B_1 + B_2 + B_3 = 0.44$.
Gini coefficient:

$$G = 1 - 2 * B$$
$$= 0.11$$

3 Exercise III - Mothers and prisoners (4 points)

DataIsFun Inc., a consulting firm specialised in statistics, has gathered data on the number of prisoners in France (in thousands) each year since 1980 and the average age of the mother at child birth. The following table reports their findings:

Variable	Statistics	Value
Age of the mother	mean	28.96
Nb of prisoners	mean	53.45
Age of the mother	standard deviation	1.18
Nb of prisoners	standard deviation	10.89
Covariance		11.29

They want to estimate the following model

$$nb\,of\ prisoners = a * age\ of\ the\ mother + b$$

1. To compute we use the formulas in the lecture's notes, with y is the number of prisoners and x the age of the mother.

$$\begin{aligned} a &= \frac{covariance(y,x)}{var(x)} \\ &= \frac{cov(x,y)}{\sigma_x^2} \\ &= \frac{11.29}{1.18^2} \\ &= 8.11 \\ b &= mean_y - a * mean_x \\ &= 10.89 - 8.11 * 28.96 \\ &= -181.37 \end{aligned}$$

2. We compute the correlation coefficient rho and the coefficient of determination R_2:

$$\begin{aligned} \rho &= \frac{cov(x,y)}{\sigma_x * \sigma_y} \\ &= \frac{11.29}{1.18 * 10.89} \\ &= 0.88 \\ R_2 &= \rho^2 \\ &= 0.77 \end{aligned}$$

3. Our findings imply a positive correlationship between the age of the mother at child birth and the number of prisoners. The correlation is pretty good (R2=0.77) but it seems that age of the mother has nothing to do with number of prisoners. One of the main driver of the number of prisoners seems to be the increase of the total population in France. The model can be thus corrected by adding evolution of the population.

4 Exercise IV - Out of tiny acorns mighty oaks grow (4 points)

Josh has the choice between two investments of 1000$, each of them for 10 years.

- Investment A: annual interest rate: 5% (**simple** interests)
- Investment B: annual compound interest : 3% (**compound** interests).

1. We compute profit with investment A and profit with investment B.

$$\begin{aligned} Profit_A &= 1000 * (1 + \frac{5}{100} * 10) \\ &= 1500 \\ Profit_B &= 1000 * (1 + \frac{3}{100})^{10} \\ &= 1343.92 \end{aligned}$$

   Jos should choose **Investment A**.

2. We want $Profit_A$ to equal $Profit_B = 1343.92$

$$\begin{aligned} 1000 * (1 + \frac{i}{100} * 10) &= 1343.92 \\ 1000 + 1000 * \frac{i}{100} * 10 &= 1343.92 \\ \frac{i}{100} &= \frac{343.92}{100 * 10} \\ i &= 3.4392 \end{aligned}$$

   The interest rate should be **3.44%**.

3. We use the same formula except that n, the number of years, is the unknown while the interest rate, i, equals 5%. We have:

$$\begin{aligned} 1000 * (1 + \frac{5}{100} * n) &= 1343.92 \\ n &= 343.92 * \frac{100}{5} * \frac{1}{1000} \\ &= 6.88\, years \end{aligned}$$

   For profit with investment A to equals profit with investment B, the duration of investment A should be **6.88 years**.

   Hint: Since profit with investment A is higher than profit with investment B when the duration is the same for both investments (question 1), we should find a *lower* duration for investment A *or* a *lower* interest rate for profit with investment A to be equal to profit with investment B.

4. There are two ways to answer that question:

   - A good which costs 1000€ now will costs $1000 * (1 + \frac{3}{100})^{10} = 1343.92$ in ten years because of inflation. The *real* profit is thus $1500 - 1343.92 = 156.08$€.
   - One can also compute the present value of receiving 1500€ in ten years. It is:

$$\frac{1500}{(1 + \frac{3}{100})^{10}} - 1000 = 116.14$$

5 Exercise V - The Art of the deal (3 points)

1. Let's compute the interests first:

$$\begin{aligned} interests &= 1000 * \frac{5}{100} * 1 \\ &= 50 \end{aligned}$$

The capital which is indeed borrowed by Eric is thus $1000 - 50 = 950$€. The effective interest rate is thus:

$$\begin{aligned} i^* &= \frac{50}{950} \\ &= 5.26\% \end{aligned}$$

2. The average maturity is the average of durations weighted by the part of due capital

$$\begin{aligned} d^* &= 30 * \frac{1000}{1000 + 3000} + 90 * \frac{3000}{1000 + 3000} \\ &= 75\ days. \end{aligned}$$

3. The commercial present value of repaying the loan in 180 days or in 360 days should be the same:

$$\begin{aligned} X * (1 - \frac{7}{100}\frac{360}{360}) &= 100,000 * (1 - \frac{7}{100} * \frac{180}{360}) \\ X &= \frac{100,000 * (1 - \frac{7}{100} * \frac{180}{360})}{(1 - \frac{7}{100} * \frac{360}{360})} \\ X &= 116,140\text{ €} \end{aligned}$$

***