Université Abdelmalek Essaâdi
Faculté Polydisciplinaire
Tétouan

Année Universitaire : 2014/2015
LEF en Sc. éco. et Gestion
Sections C et D

Rattrapage de Statistique Descriptive
Durée : 1 heure

Problème n°1 : (10 points)
Un tableau statistique se présente comme suit :

xi	−5	−1	3	10	13
yi	33	25	17	3	−3

1. Donner la représentation graphique de y en fonction de x, et s'assurer qu'un ajustement linéaire de y par x semble légitime.
2. Calculer les moyennes arithmétiques $\bar{x}$ et $\bar{y}$ de x et y.
3. Effectuer ensuite l'ajustement par la droite des moindres carrés.
4. Calculer le coefficient de corrélation linéaire entre les variables x et y et conclure.

Problème n°2: (10 points)
Les salaires annuels (en 103 DH) de 100 ouvriers d'une entreprise sont distribués de la manière suivante :

Salaires (en 103 DH) compris entre	Nombres des ouvriers
25 et 26	13
26 et 28	15
28 et 30	21
30 et 32	19
32 et 34	15
34 et 35	10
35 et 36	7

1. Calculer les fréquences relatives et les effectifs cumulés croissants.
2. Déterminer le nombre des ouvriers qui ont un salaire annuel inférieur à 32000 DH.
3. Calculer le salaire annuel le plus fréquent.
4. Calculer la médiane de cette distribution et interpréter le résultat.

Bon courage !

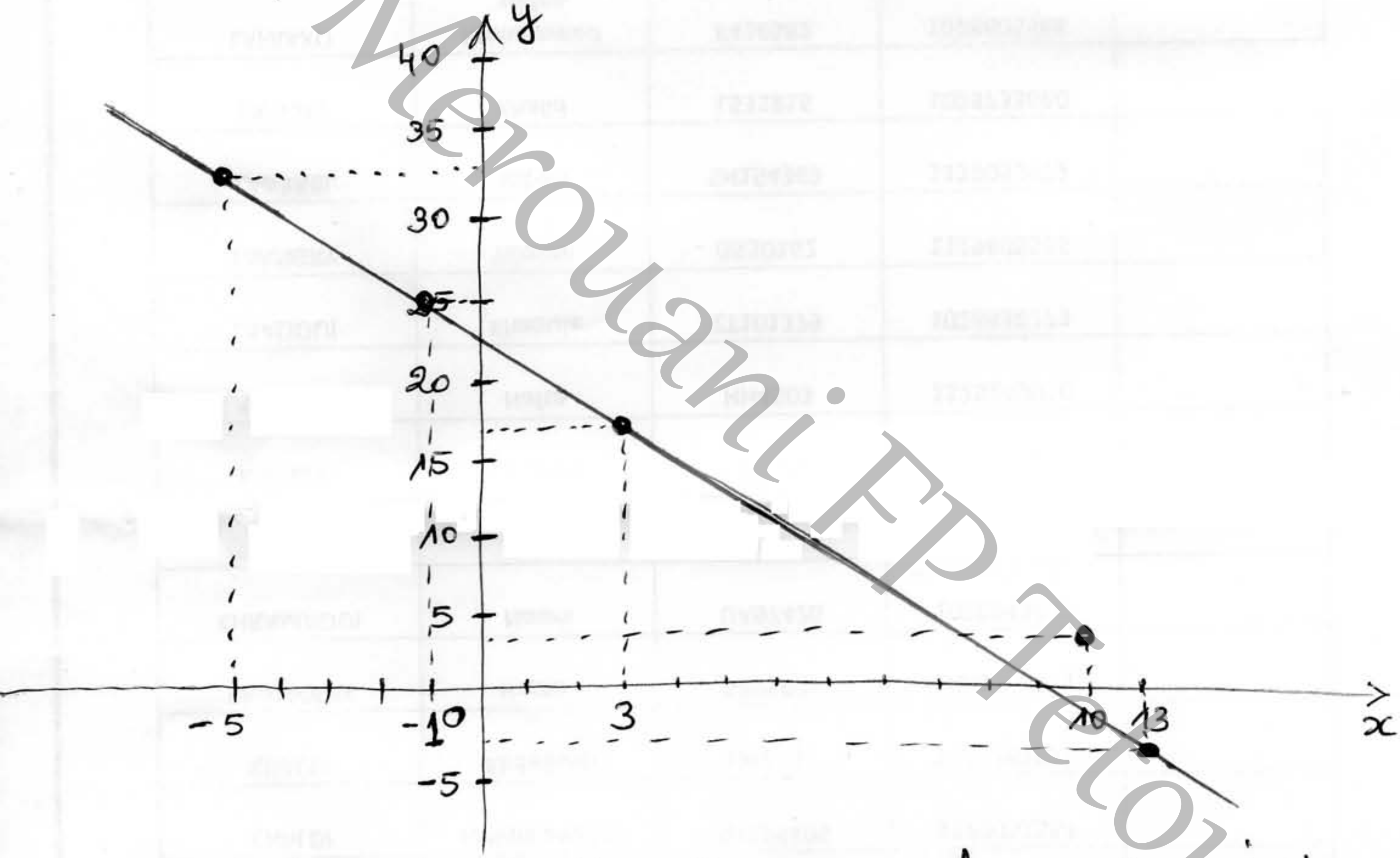
Université Abdelmalek Essaâdi
Faculté Polydisciplinaire
de Tétouan

Année : 2014 / 2015
LEF en sc. éco. & Gestion
Semestre 1

Corrigés du rattrapage de Statistique Descriptive

Pb. n° 1 :

1°) Représentation graphique de y en fonction de x :



On voit clairement, d'après ce graphe, qu'un ajustement linéaire de y par x semble légitime.

2°) Calcul des moyennes arithmétiques de x et de y :

Ces deux variables statistiques sont unitaires et ona pour chacune 5 valeurs. Donc $\bar{x} = \frac{1}{5}\sum x_i$ et $\bar{y} = \frac{1}{5}\sum y_i$

D'où $\bar{x} = \frac{20}{5} = 4$ et $\bar{y} = \frac{75}{5} = 15$.

3°) Ajustement de y par x en utilisant la méthode des moindres carrées :

On cherche la droite $y = ax + b$

avec $a = \dfrac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sum_i (x_i - \bar{x})^2}$ et $b = \bar{y} - a\bar{x}$

x_i	y_i	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})(y_i - \bar{y})$	$(y_i - \bar{y})^2$
−5	33	−9	18	81	−162	324
−1	25	−5	10	25	−50	100
3	17	−1	2	1	−2	4
10	3	6	−12	36	−72	144
13	−3	9	−18	81	−162	324
20	75	0	0	224	−448	896

Alors

$a = \dfrac{-448}{224} = -2$ et $b = 15 + 2 \times 4 = 23$

donc $\boxed{y = -2x + 23}$

4°) Coefficient de corrélation linéaire entre x et y :

$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \cdot \sum_i (y_i - \bar{y})^2}}$$

$$r = \frac{-448}{\sqrt{224 \times 896}} = \frac{-448}{\sqrt{200704}} = \frac{-448}{448}$$

$\boxed{r = -1}$

x et y sont parfaitement (fortement) corrélées, mais négativement !

Pb. n° 2 :

$[e_{i-1}, e_i[$	n_i	f_i	n_{icc}	a_i	h_i
[25, 26[	13	0,13	13	1	13
[26, 28[	15	0,15	28	2	7,5
[28, 30[	21	0,21	49	2	10,5
$\boxed{[30, 32[}$	19	0,19	68	2	9,5
[32, 34[	15	0,15	83	2	7,5
[34, 35[	10	0,1	93	1	10
[35, 36[	7	0,07	100	1	7
	100				

1°) voir la 3ème et le 4ème colonne du tableau ci-dessus.

2°) Pour déterminer le nombre des ouvriers qui ont un salaire annuel inférieur à 32000 DH, on cherche parmi les nicc, la valeur en face de la classe qui a comme borne sup. la valeur 32. D'après le tableau en haut, cette valeur est : 68 ouvriers.

3°) Le salaire annuel le plus fréquent correspond au mode Mo de cette série statistique. Après avoir calculé les amplitudes, on voit que l'on est dans le cas des amplitudes différentes. Donc, on calcul les hi (les fréquences moyennes par unité d'amplitude : $h_i = \frac{n_i}{a_i}$). Alors, le classe modale correspond à hi la plus grande. c'est-à-dire, la classe [25, 26[. On applique la formule $M_o = e_{i-1} + \frac{h_{i+1}}{h_{i-1} + h_{i+1}} a_i$

$$M_0 = 25 + \frac{7,5}{0+7,5} \times 1 = 26$$

Donc, le salaire annuel le plus fréquent est 26000 DH.

4°) La médiane "Mé" :

$N = 100 \Rightarrow \frac{N}{2} = 50$, cette valeur n'existe pas parmi les $n_i cc$. Alors, on cherche la valeur qui la dépasse pour la 1ère fois, c'est 68. D'où, la classe médiane est $[30, 32[$, on applique la formule

$$Mé = e_{i-1} + \frac{\frac{N}{2} - n_{i-1}cc}{n_i} \times a_i$$

$$= 30 + \frac{50-49}{19} \times 2 = 30,1$$

Interprétation: Mé = 30100 DH/an

Il y a 50 ouvriers qui gagnent moins que 30100 DH/an et autres 50 ouvriers qui gagnent plus que 30100 DH/an.