Statistique Descriptive II

Pr. Mohamed El Merouani

1

Compléments du Chapitre 6

Ajustement et Corrélation linéaire

2

Méthode des points médians :

Exemple :

- Un phénomène économique a fait l'objet d'une mesure mensuelle au cours de 20 mois consécutifs.
- Représentons graphiquement la série statistique en question et ajustons par **la méthode des points médians** les points obtenus.

3

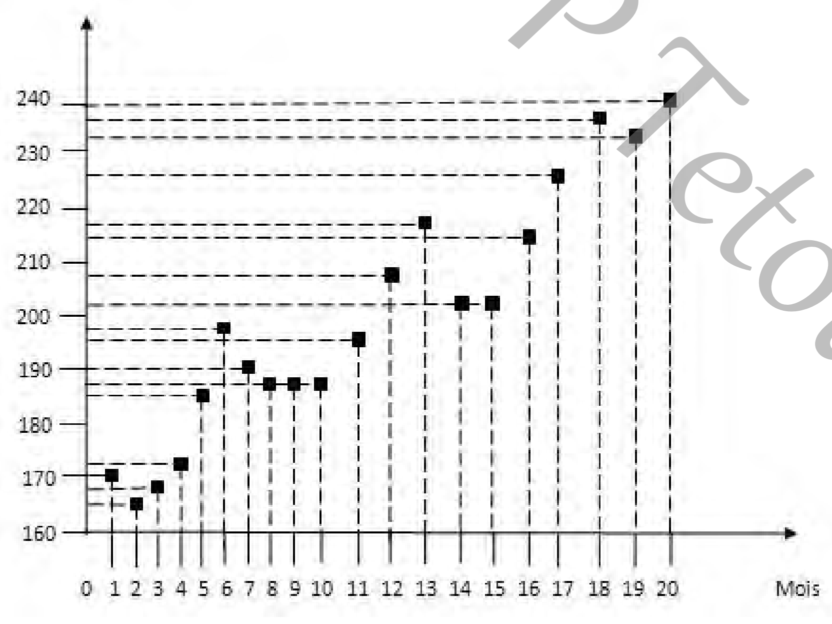
Mois	Mesures	Mois	Mesures	Mois	Mesures	Mois	Mesures
1	170	6	198	11	196	16	216
2	164	7	190	12	208	17	226
3	168	8	188	13	218	18	236
4	172	9	188	14	202	19	232
5	186	10	188	15	202	20	240

4

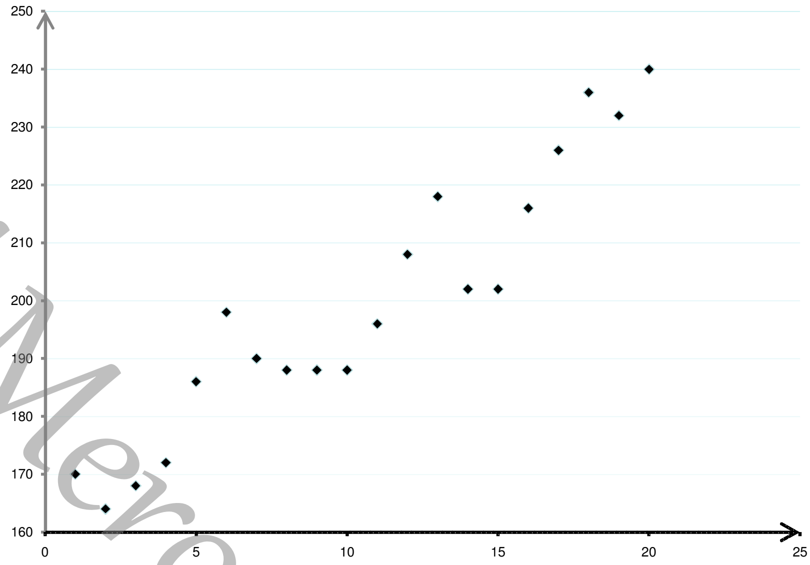
- On représente d'abord les 20 points de coordonnées respectives :

 (1, 170) ; (2 , 164) ; … ;(20 , 240).

5



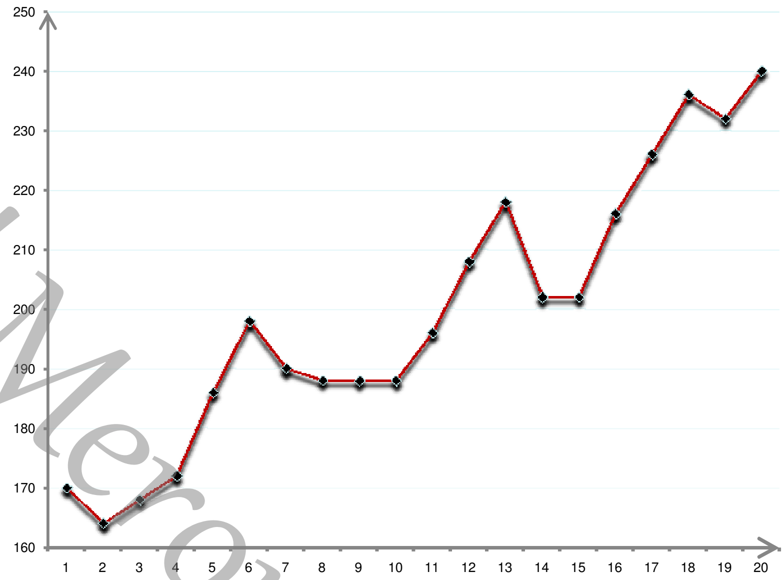
6



7

- On relie ensuite ces points par des segments de droite, obtenant ainsi la représentation graphique de la série statistique étudiée.

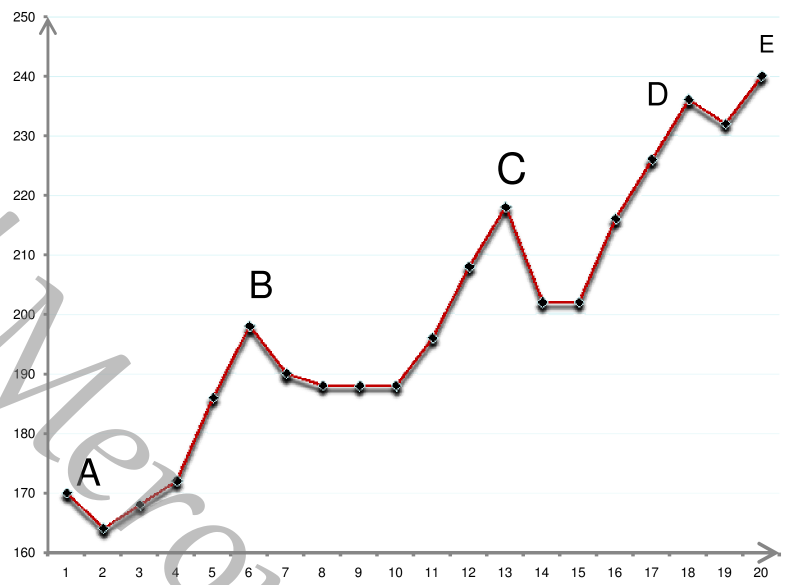
8



9

- Parmi les points représentés certains correspondent à des périodes de pointe (relative), leur ordonnée sont supérieure aux ordonnées des points immédiatement voisins.
- Tels sont les points,
- A(1, 170) – B(6, 198) – C(13, 218) – D(18, 236), E(20, 240)

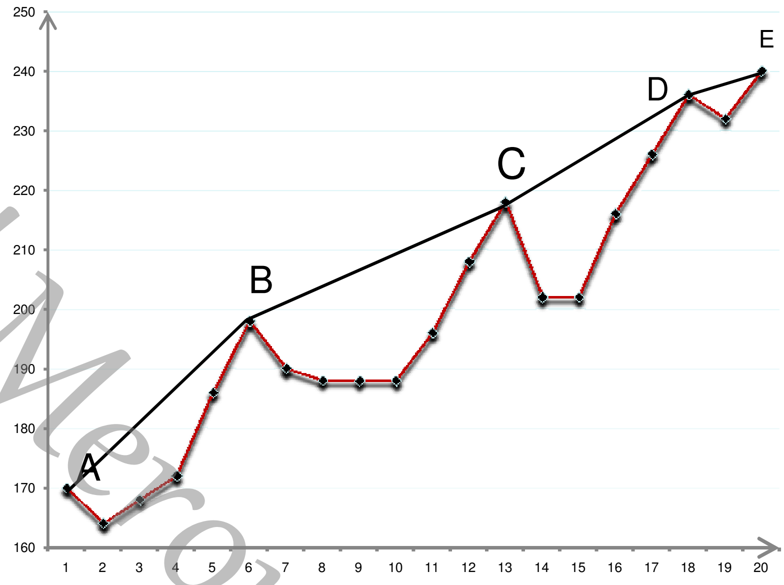
10



11

- Construisons les segments AB, BC, CD, DE.
- Tous les points de la représentation se situent sur ces segments, ou au dessous.

12



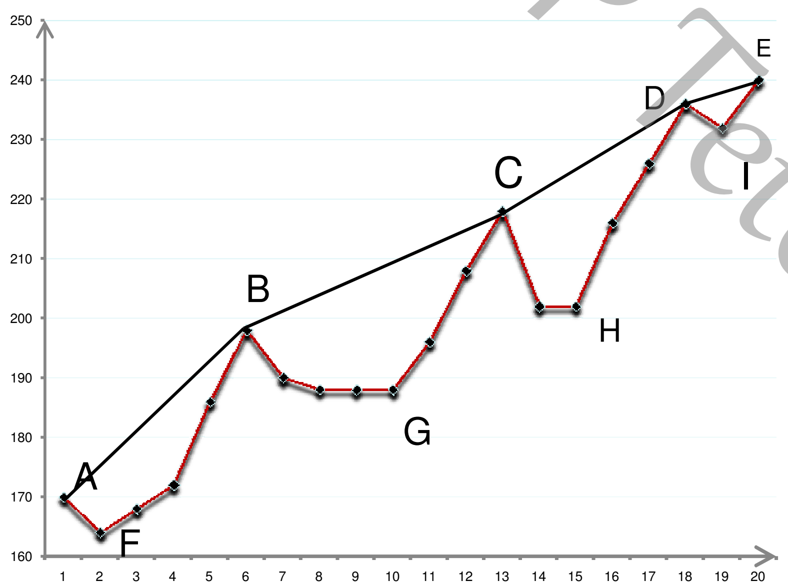
13

- Nous avons ainsi construit une sorte de toit à notre représentation graphique.
- D'autres points, parmi les vingt points représentés, correspondent à des périodes de creux (relatif) leur ordonnée étant inférieure aux ordonnées des points immédiatement voisins.

14

- Tels sont les points F(2, 164) – G(10, 188) – H(15, 202) – I(19, 232).

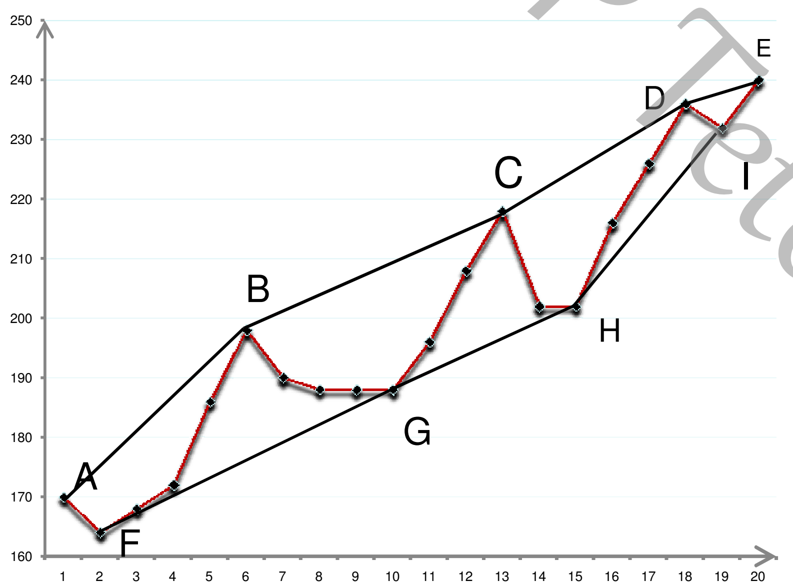
15



16

- Construisons les segments FG, GH, HI.

17



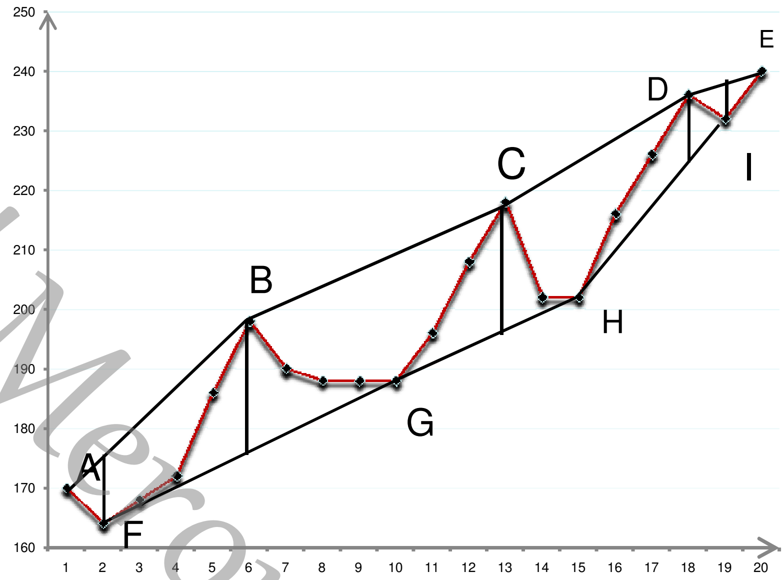
18

- Tous les points de la représentation se situent sur ces segments ou au-dessous.
- Nous avons ainsi construit une sorte de plancher à la représentation.
- Toit et plancher obtenus déterminent une sorte de couloir à l'intérieur duquel doit se trouver la fonction d'ajustement cherchée.

19

- Pour déterminer cette fonction, à partir de chacun des points A, B, …, H, I, on mène une parallèle à l'axe des ordonnées, parallèle partant du point en question, vers le haut quand elle part d'un point de creux, vers le bas quand elle part d'un point de pointe, mais limitée à la paroi du couloir.

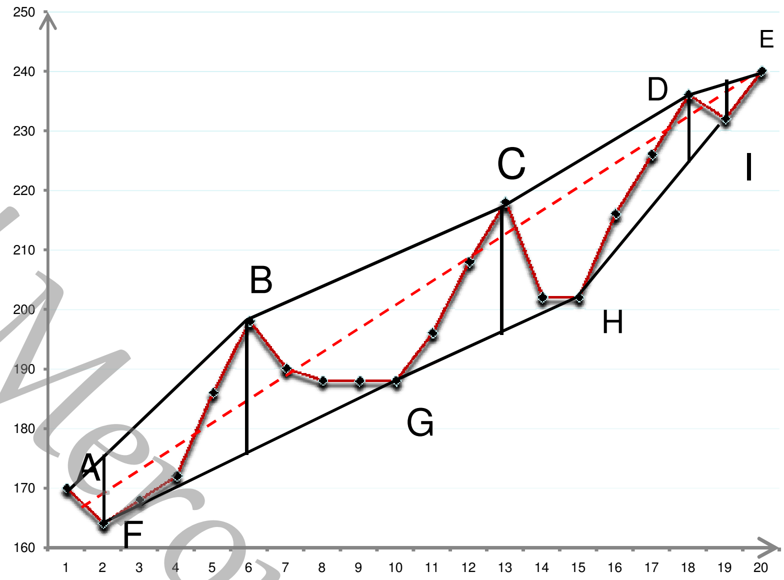
20



21

- On détermine sans difficulté le milieu, le point médian, de chacun de ces segments.
- Les points médians ainsi obtenus sont joints, dans l'ordre, et donnent la fonction d'ajustement cherchée, en pointillé sur la représentation.

22



23

Méthode de Mayer

- L'ensemble des points à ajuster est partagé 2 sous-ensembles de même effectif, dans l'ordre où les points se présentent.
- Chacun de ces 2 sous-ensembles est alors remplacé par le point dont les coordonnées sont les suivants:

24

- En abscisse, la moyenne arithmétique des abscisses des points du sous-ensemble.
- En ordonnée, la moyenne arithmétique des ordonnées des points du sous-ensemble.

- On dispose alors de 2 points dont les coordonnées sont connues, par lesquels passe la droite dite « Droite de Mayer », droite d'ajustement d'équation facile à déterminer.

25

Exemple:

- (Déjà utilisé) qui portait 6 points (donc de 2 sous-ensembles de 3 points)

x_i	n_i
2	7
4	10
6	13
8	15
9	20
13	28

26

- Les coordonnées des 3 premiers points du tableau conduiront à:
  - abscisse du point représentant le premier sous-ensemble: $$\frac{2+4+6}{3}=4$$
  - Ordonnée de ce point: $$\frac{7+10+13}{3}=10$$

27

- Les coordonnées des 3 derniers points du tableau conduisent à:
  - abscisse du point représentant le 2ème sous-ensemble: $$\frac{8+9+13}{3}=10$$
  - Ordonnée de ce point: $$\frac{15+20+28}{3}=21$$

28

- La droite de Mayer, d'équation y=ax+b, passe par les deux points de coordonnées (4, 10) et (10, 21).
- Donc $\begin{cases} 10 = 4a + b \\ 21 = 10a + b \end{cases} \Rightarrow \begin{cases} a = 1{,}83 \\ b = 2{,}67 \end{cases}$
- D'où, la droite de Mayer cherchée est:

$$y = 1{,}83\,x + 2{,}67$$

29

Exemple de calcul du coefficient de corrélation linéaire entre deux variables X et Y :

30

x_i	y_i	$x_i - \bar{X}$	$y_i - \bar{Y}$	$(x_i - \bar{X})^2$	$(y_i - \bar{Y})^2$	$(x_i - \bar{X})(y_i - \bar{Y})$
16	20	-10.1	-10.4	102.01	108.16	+105.04
18	24	-8.1	-6.4	65.61	49.96	+51.84
23	28	-3.1	-2.4	9.61	5.76	7.44
24	22	-2.1	-8.4	4.41	70.56	17.64
28	32	+1.9	+1.6	3.61	2.56	3.04
29	28	+2.9	-2.4	8.41	5.76	-6.96
26	32	-0.1	+1.6	0.01	2.56	-0.16
31	36	+4.9	+5.6	24.01	31.36	+27.44
32	41	+5.9	+10.6	34.81	112.36	+62.54
34	41	+7.9	+10.6	62.41	112.36	83.74
261	304	0	0	314.90	492.40	+351.60

31

- Moyenne arithmétique de *X* : $\bar{X} = \frac{261}{10} = 26{,}1$
- Moyenne arithmétique de *Y* : $\bar{Y} = \frac{304}{10} = 30{,}4$
- Coefficient de corrélation linéaire :

$$r = \frac{\sum_i (x_i - \bar{X})(y_i - \bar{Y})}{\sqrt{\sum_i (x_i - \bar{X})^2 (y_i - \bar{Y})^2}} = \frac{351{,}60}{\sqrt{314{,}90 \times 492{,}40}}$$

32

$$= \frac{351,60}{\sqrt{155056,76}} = \frac{351,60}{393,77} = +0,89$$

- Donc on a une corrélation positive, comme l'indique déjà la représentation graphique et assez serrée, le coefficient r ayant une valeur absolue voisine de 1.

33

Droite de régression linéaire: Méthode de moindres carrées:

34

Exemple:

- Cherchons l'équation de la droite d'ajustement ou de régression linéaire par la méthode des moindres carrées, pour l'exemple de six points déjà vu en Méthode de Mayer:
- On a le tableau des calculs suivant:

35

x_i	y_i	$x_i y_i$	x_i^2	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i-\bar{x})(y_i-\bar{y})$	$(x_i-\bar{x})^2$
2	7	14	4	-5	-8,5	42,5	25
4	10	40	16	-3	-5,5	16,5	9
6	13	78	36	-1	-2,5	2,5	1
8	15	120	64	+1	-0,5	-0,5	1
9	20	180	81	+2	+4,5	9	4
13	28	364	169	+6	+12,5	75	36
42	93	796	370	0	0	145	76

36

$$\bar{x} = \frac{42}{6} = 7 \qquad \bar{y} = \frac{93}{6} = 15{,}5$$

Calcul des paramètres a et b en utilisant les formules:

$$a = \frac{\sum_{i=1}^{k} x_i y_i - \bar{y}\sum_{i=1}^{k} x_i}{\sum_{i=1}^{k} x_i^2 - \bar{x}\sum_{i=1}^{k} x_i} \quad et \quad b = \bar{y} - a\bar{x}$$

37

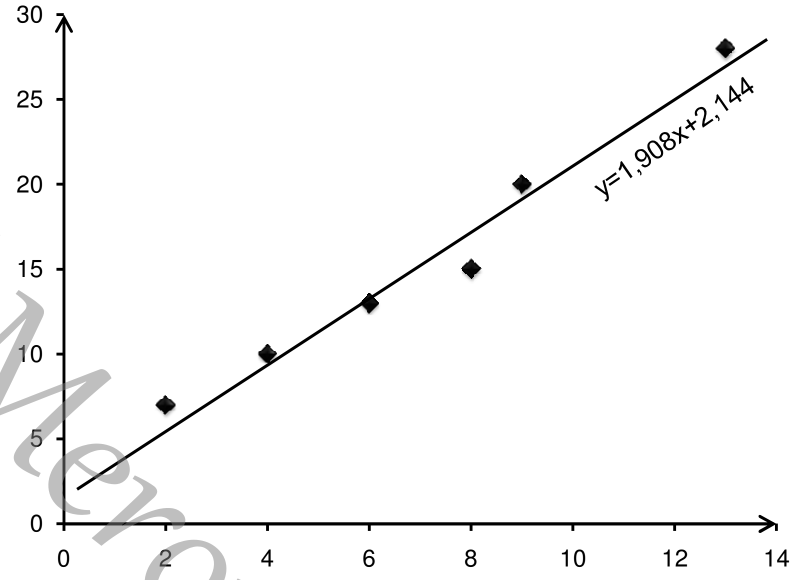
$$a = \frac{796 - 15{,}5 \times 42}{370 - 7 \times 42} = \frac{796 - 651}{370 - 294} = \frac{145}{76} \approx 1{,}908$$

$$b = 15{,}5 - 1{,}908 \times 7 = 15{,}5 - 13{,}356 = 2{,}144$$

Equation de la droite d'ajustement:

$$y = 1{,}908\ x + 2{,}144$$

38



39