cteur : Gestion Commerce

résentation

M 110: Statistique

re Année

lière :

Gestion des
ntreprises
Tronc comun)

statistique

Plan

- **Statistique**
- ***Généralités***
- ***Qualification d’une variable qualitative et variable quantitative***
- ***Représentations graphiques***
- ***les Caractéristiques de distribution***
- ***Régression et corrélation***

Généralités

Définitions :

statistique : méthode scientifique qui vise à observer, collecter, analyser des données quantitatives.

⇩

La statistique descriptive est la partie de la statistique qui sert à décrire un phénomène

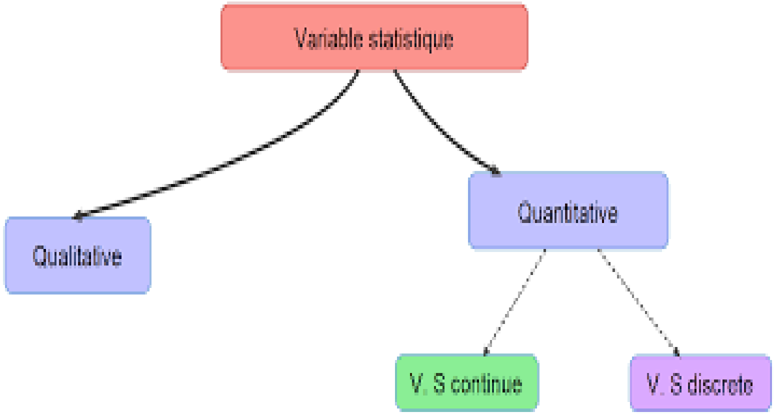
Apport de la statistique aux économistes

Apport de la statistique aux économistes

La statistique est utile aux théoriciens

La statistique est utile aux praticiens de l'économie

Qualification d'une variable qualitative et variable quantitative



Pour une variable qualitative, les modalités ne sont pas mesurables.

Pour une variable quantitative, les modalités sont mesurables. Ce sont :

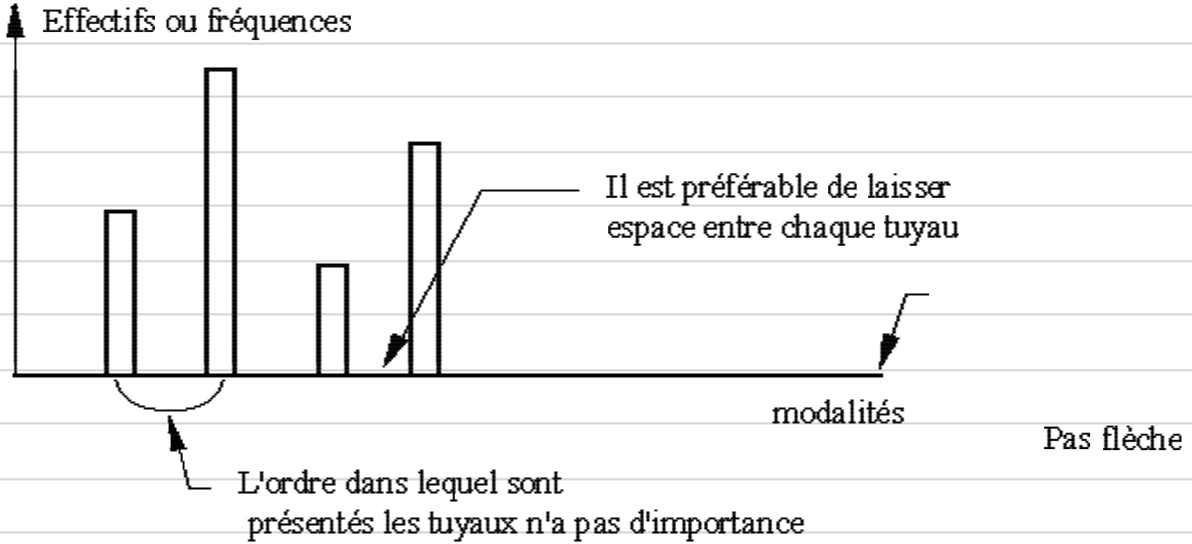
- des valeurs numériques ponctuelles lorsque la variable est **discrète**

- des intervalles lorsque la variable est **continue** ou lorsque la variable est discrète et qu'elle comporte beaucoup de modalités

Représentations graphiques

Variable qualitative :

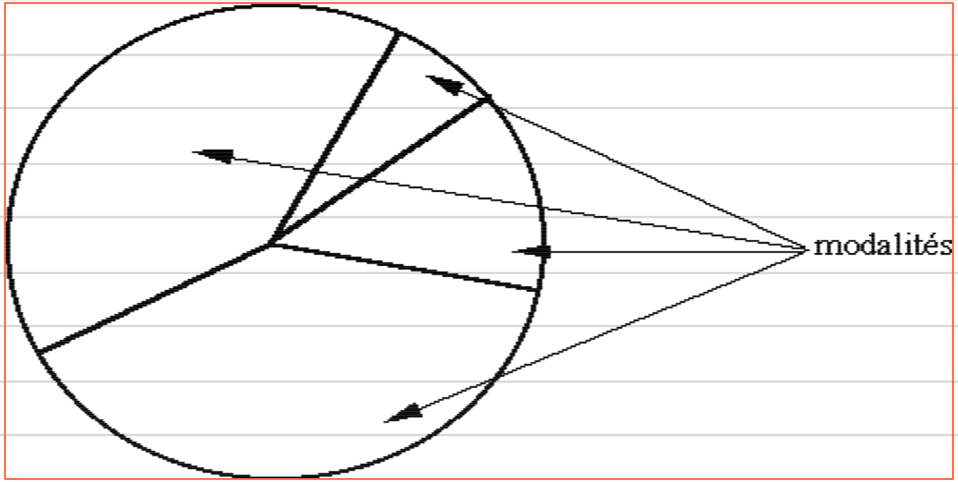
Les tuyaux d'orgue (ou diagramme en barre ou diagramme à bandes



les diagrammes à secteurs (ou camemberts)

-L'effectif total est représenté par un disque.

-Chaque modalité est représentée par un secteur circulaire dont la surface (pratiquement : l'angle au centre) est proportionnelle à l'effectif correspondant



Exemple

Application :

La répartition des candidats convoqués pour participer au Test d'Admissibilité à la Formation en Management (TAFEM 1998) pour l'accession à L'Ecole Nationale de Commerce et de Gestion d'Agadir , selon la série du baccalauréat se présente comme suit :

Série du Bac xi	Nombre de candidats ni
Sciences économiques	250
Sciences mathématiques	200
Sciences expérimentales	400
T.G.A	50
T.G.C	100
Total	1000

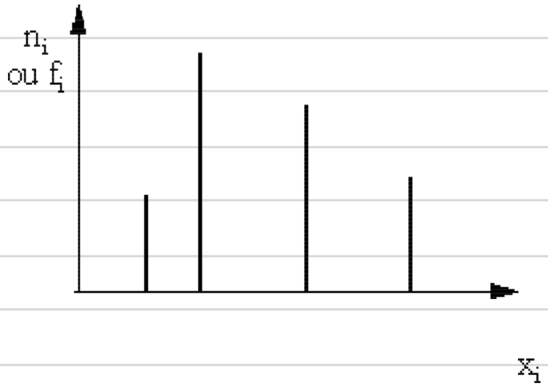
TAF: représentez cette distribution en Tuyaux d'orgues et Diagramme circulaire.

Variable quantitative

Variable discrète

Diagramme différentiel : le diagramme en bâtons

Les valeurs discrètes xi prises par les variables sont placées sur l'axe des abscisses, et les effectifs (ou les fréquences) sur l'axe des ordonnées. La hauteur du bâton est proportionnelle à l'effectif.

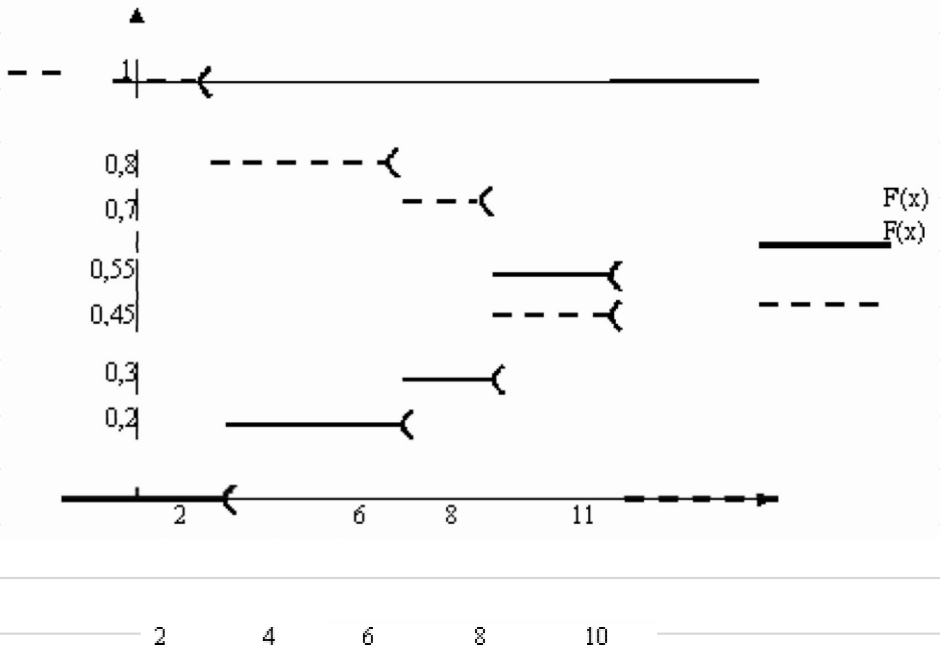


Variable quantitative

Variable discrète

Diagrammes cumulatifs : ils permettent de visualiser l'évolution des effectifs (fréquences) cumulés croissants ou décroissants.

-

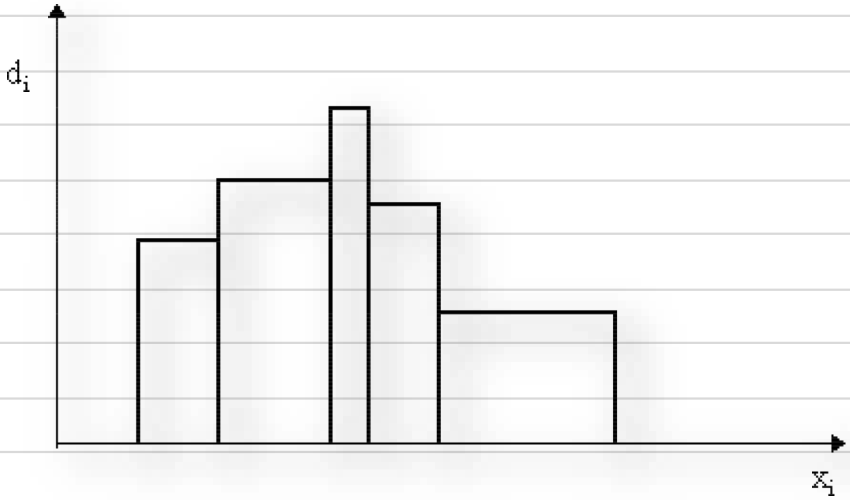


Variable quantitative

Variable classée

Diagramme différentiel : l'histogramme

C'est un ensemble de rectangles contigus, chaque rectangle associé à chaque classe ayant une surface proportionnelle à l'effectif (fréquence) de cette classe.

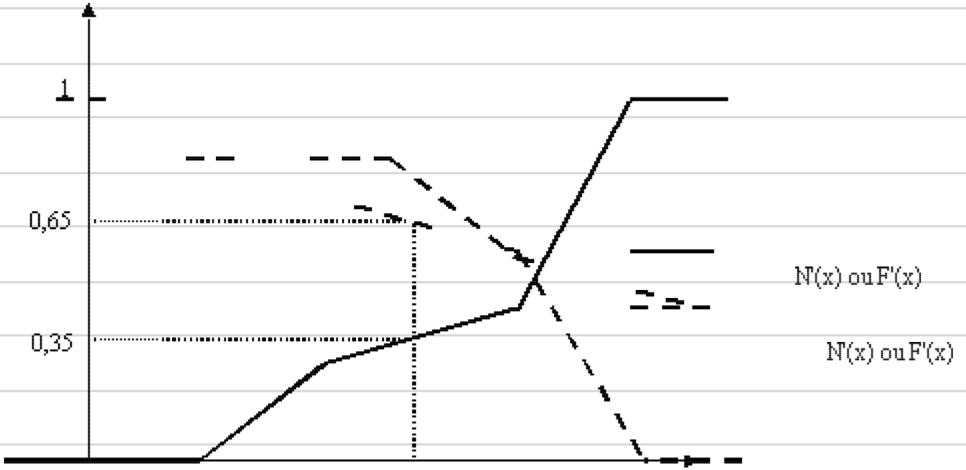


Variable quantitative

Variable classée

Diagrammes cumulatifs

C'est un ensemble de rectangles contigus, chaque rectangle associé à chaque classe ayant une surface proportionnelle à l'effectif (fréquence) de cette classe.



Exemple

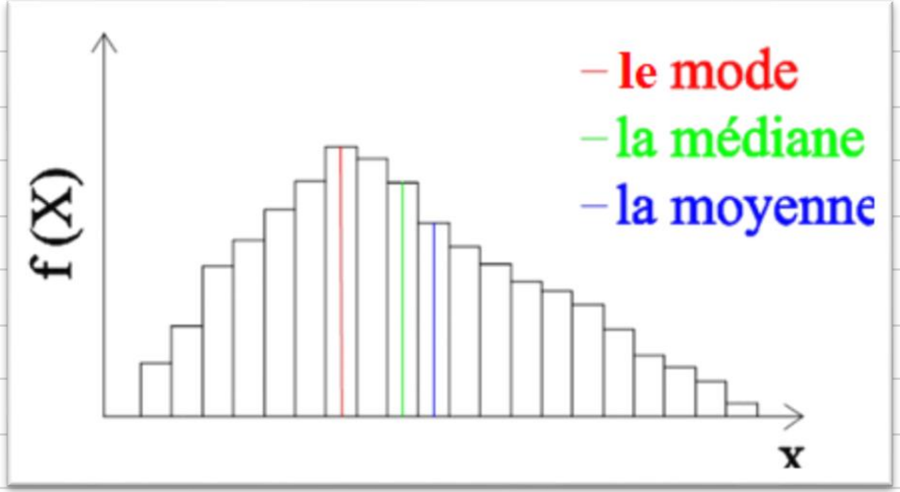
Application :

Représentez graphiquement la distribution de 50 étudiants en fonction de leur taille suivante :

Taille en cm xi	Nombre d'étudiants
150-160	16
160-165	6
165-170	12
170-175	14
175-180	2
Total	50

les Caractéristiques de distribution

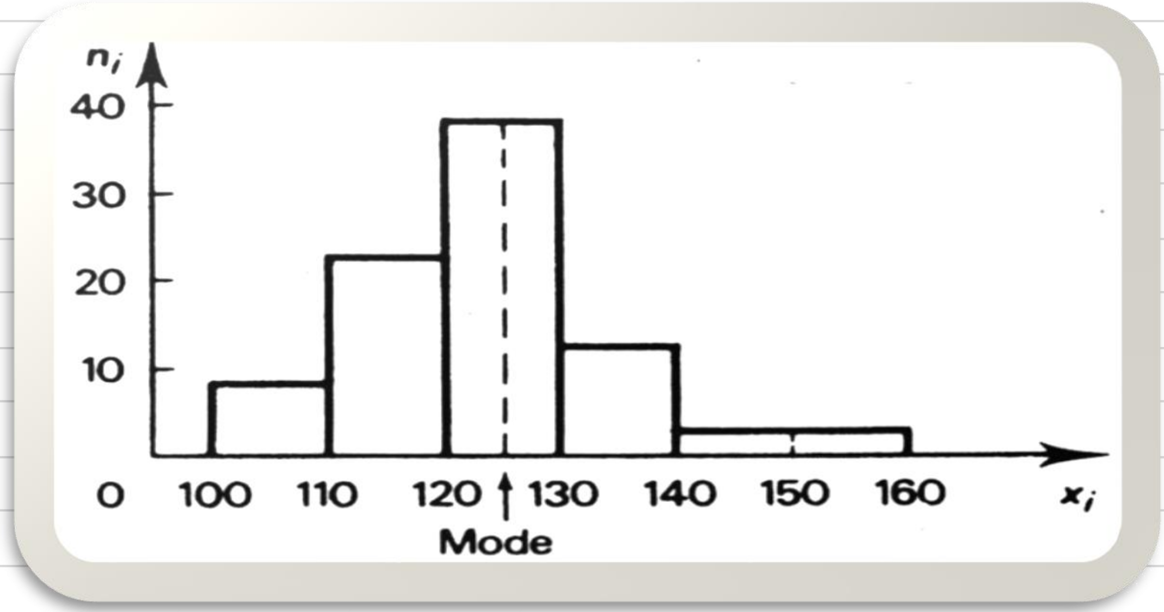
Caractéristiques de tendance centrale et de position



Les caractéristiques de tendance centrale essayent de donner la valeur la plus représentative d'un ensemble de valeurs numériques.

Le Mode

- **C'est la valeur observée d'effectif maximum.**



Exemple

Application :

Déterminez la valeur modale de la distribution suivante, de 50 étudiants selon leur taille :

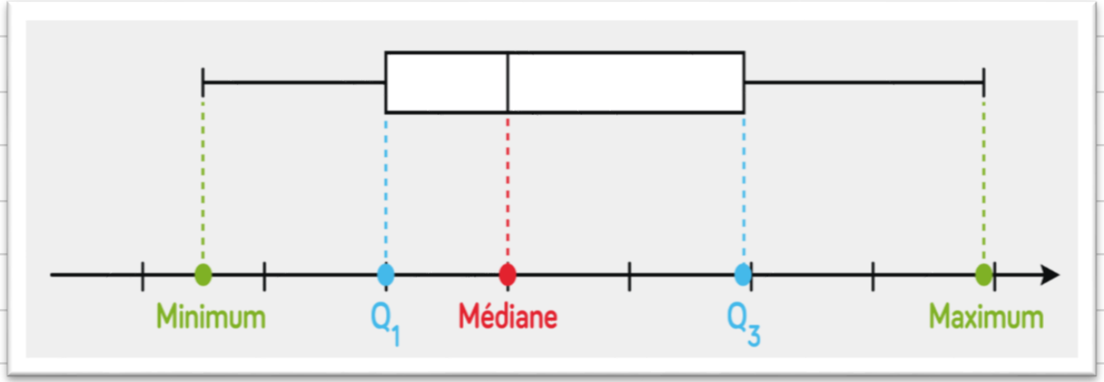
Taille en cm : xi	Nombre d'étudiants : ni
150-160	15
160-170	6
170-175	10
175-180	16
185-200	3
Total	50

Eléments de réponse :

Mo = 173.77 cm

Médiane

- **Les valeurs étant rangées par ordre croissant, c'est la valeur de la variable qui sépare les observations en deux groupes d'effectifs égaux.**



Exemple

Application :

Soit la série statistique suivante :

19 17 20 18 17 17 20 19 15 16 20 23 22 14 15 24

TAF : Calculez la médiane de cette série

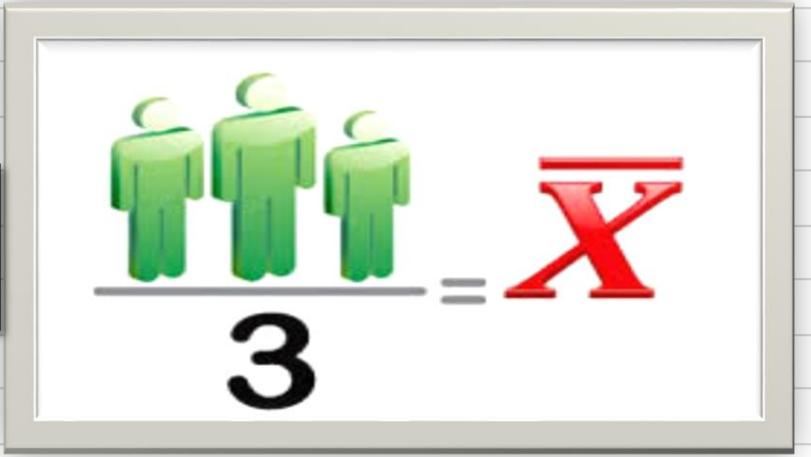
Eléments de réponse :

Me=18.5

Moyenne arithmétique

- Si xi sont les observations d'une variable discrète ou les centres de classe d'une variable

$\bar{x}$ est égale à $\sum_{i=1}^{k} \frac{n_i x_i}{n} = \sum_{i=1}^{k} f_i x_i$



Exemple

Application :

Déterminez la taille moyenne des 50 étudiants dont la distribution par taille se présente commesuit :

Taille en cm xi	Nombre d'étudiants
150-160	16
160-165	6
165-170	12
170-175	14
175-180	2
Total	50

Eléments de réponse : x = 168.3 cm

Moyenne géométrique

Moyenne géométrique

Si x_i sont les observations d'une variable quantitative, la moyenne géométrique est égale

$$G = \sqrt[n]{x_1^{n_1} \times \ldots\ldots \times x_k^{n_k}}$$

à

Moyenne harmonique

$$H = \frac{\sum ni}{\sum ni\, 1/xi}$$

Exemple

Application :

Un cycliste effectue une traversé de 50 kms. Pendant les 20 premiers kms il roulait avec unevitesse constance de km/h, les 15 kms suivants à une vitesse constante de 30 km/h. Du point kilométrique 35 au 55 la vitesse de notre cycliste n'est que de 10 km/h et au-delà du point kilométrique sa vitesse n'est que de 5 km/h.

TAF :

Quelle est la vitesse de ce cycliste sur l'ensemble du parcours ?

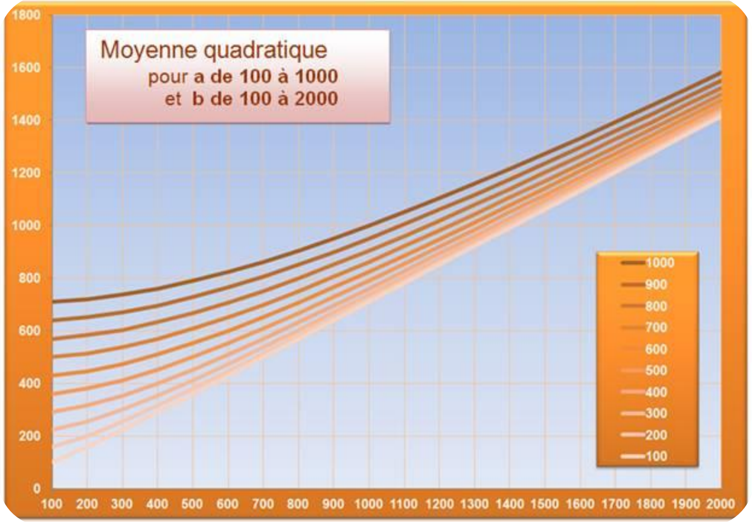
Eléments de réponse :

H = 16.67

Moyenne quadratique

Si xi sont les observations d'une variable quantitative, la moyenne harmonique est égale à

$$Q = \sqrt{\frac{n_1 x_1^2 + \ldots + n_k x_k^2}{n}}$$



Quartiles

Ce sont des caractéristiques de position.

Les quartiles sont obtenus lorsqu'on a cumulé 25, 50, 75% de la population
Les déciles sont obtenus lorsqu'on a cumulé 10, 20,. , 90% de la population
Les centiles sont obtenus lorsqu'on a cumulé 1, 2,. , 99% de la population

Exemple

Application :

Soit la population de 80 salariés classés d'après le niveau de leur salaire journalier.

	Classes en dhs	Ni	ni cumulés
1	90 à 100	5	5
2	100 à 110	9	14
3	110 à 120	16	30
4	120 à 130	25	55
5	130 à 140	13	68
6	140 à 150	7	75
7	150 à 160	3	78
8	160 à 170	2	80
Total		80	

TAF : calculez la médiane et les deux quartiles

Eléments de réponse :

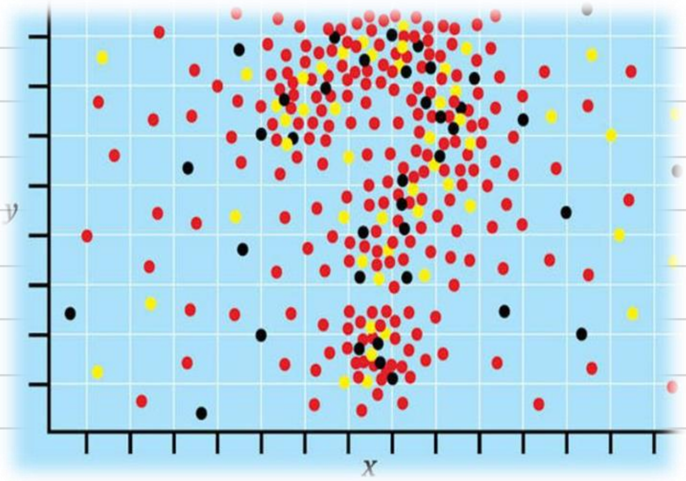
Me = 124

$Q_1 = 110 + (10x6)/16 = 113.7$

$Q_3 = 130 + (10x5)/13 = 133.8$

les Caractéristiques de distribution

Caractéristiques de dispersion



Comme leur nom l'indique, ces caractéristiques essayent de synthétiser par une seule valeur numérique la dispersion de toutes les valeurs observées.

Étendue

C'est la différence entre la plus grande et la plus petite observation

Exemple

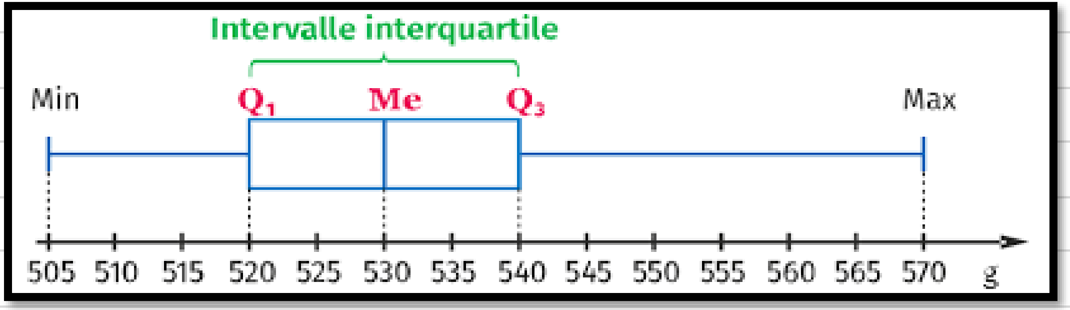
Application :

Quelle est l'étendue de la série statistique suivante :10 390 395 405 410 1000

Eléments de réponse : Etendue = 990

Intervalle inter-quartile

C'est la différence entre le troisième et le premier quartile



Exemple

Application :

Reprenez les données de l'application sur les quartiles et calculez l'intervalle inter-quartile.

Eléments de réponse :

Q3-Q1=20

Variance et écart-type

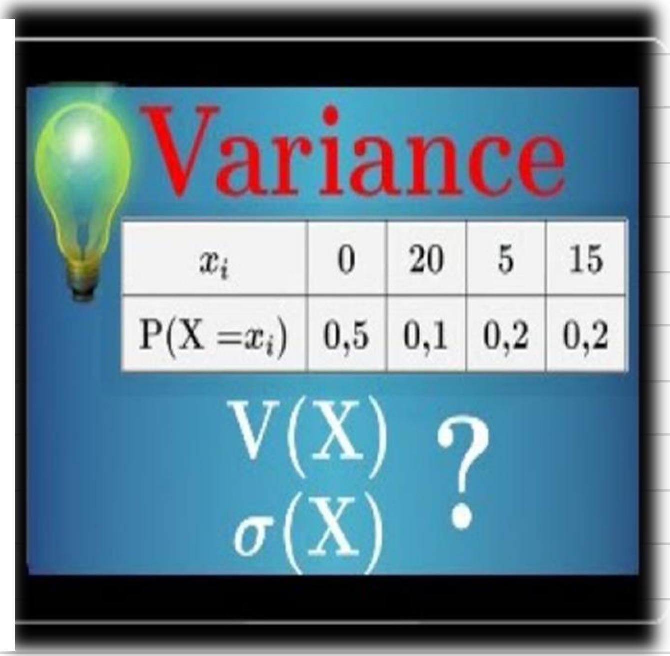
Si xi sont les observations d'une variable discrète ou les centres de classe d'une variable classée, la variance

V est égale à $\sum_{i=1}^{k} \frac{n_i(x_i - \bar{x})^2}{n} = \sum_{i=1}^{k} f_i(x_i - \bar{x})^2$

On a aussi $V = \sum_{i=1}^{k} \frac{n_i x_i^2}{n} - \bar{x}^2$

c.à.d moyenne des carrés - carré de la moyenne

On utilise plus couramment l'écart type qui est la racine carrée de la variance et qui a l'avantage d'être un nombre de même dimension que les données (contrairement à la variance qui en est le carré)



Coefficient de variation

$$CV = \frac{\sigma}{x}$$

Exemple

Applications :

App.1- Les séries suivantes représentent la mesure d'un caractère auprès des individus d'unepopulation :

a. 6 1 8 10 5 4 11 3 2 9 7 12 13

b. 19 17 7 1 4 24 15 22 10 13

c. 15 12 17 15 20 15 20 15 15 9 7

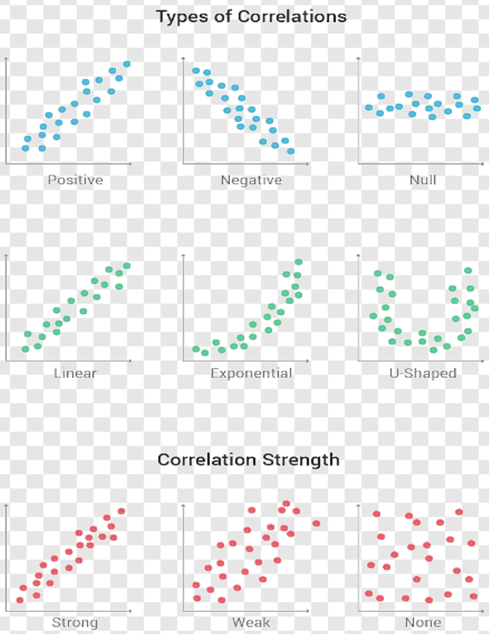
d. 21 25 34 10 20 27 14 20 34

Dans chacun de ces cas calculez : la moyenne, la médiane, le mode ,la variance, l'écart type et lecoefficient de variation.

Eléments de réponse :

a. $\bar{x}=7$, Me=7, pas de mode, $\sigma^2=14$, $\sigma=3.74$, V=53.4%
b. $\bar{x}=13.2$, Me=14, pas de mode, $\sigma^2=52.76$, $\sigma=7.26$, V=55%
c. $\bar{x}=14.5$, Me=15, Mo=15, $\sigma^2=14.61$, $\sigma=3.82$, V=26.3%
d. $\bar{x}=22.8$, Me=21, deux modes : 20 et 34, $\sigma^2=59.28$, $\sigma=7.70$, V=33.8%

Régression et corrélation

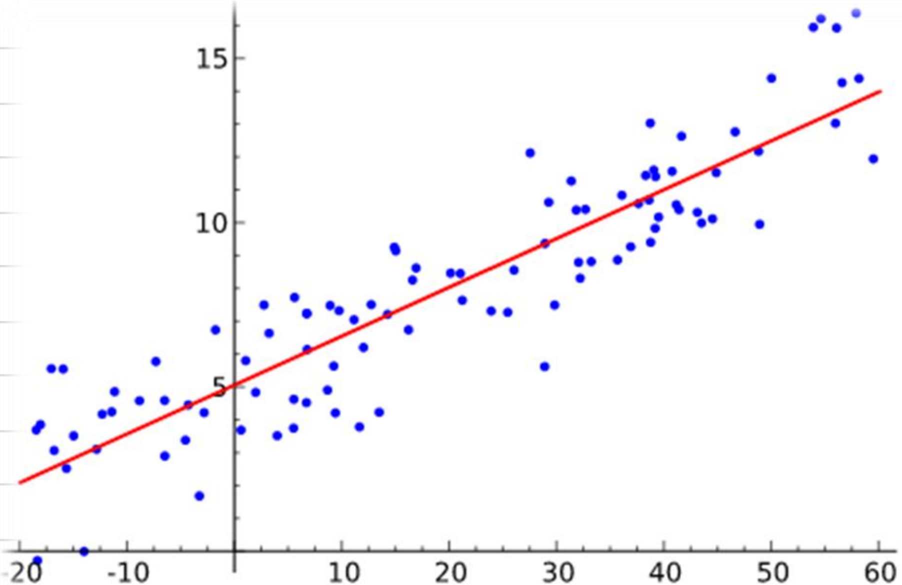


Lorsqu'on observe deux variables quantitatives sur les mêmes individus, on peut s'intéresser à une liaison éventuelle entre ces deux variables.

La régression fournit une expression de cette liaison sous la forme d'une fonction mathématique.

La corrélation renseigne sur l'intensité de cette liaison.

Ajustement d'un nuage de points à une fonction mathématique



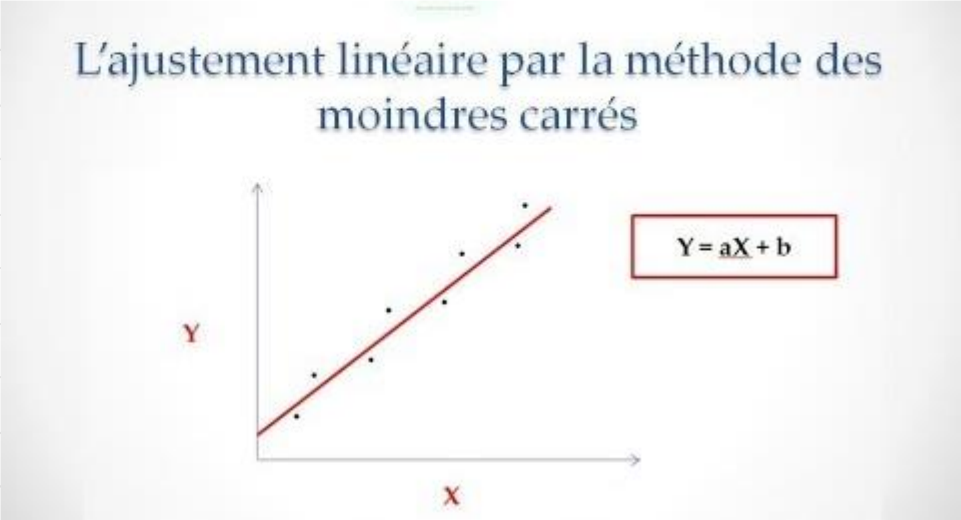
Ajustement linéaire par la méthode des moindres carrés

- Lorsque le nuage de points (xi , yi) est à peu près rectiligne, on peut envisager d'exprimer la liaison entre x et y sous forme de fonction affine y = ax + b

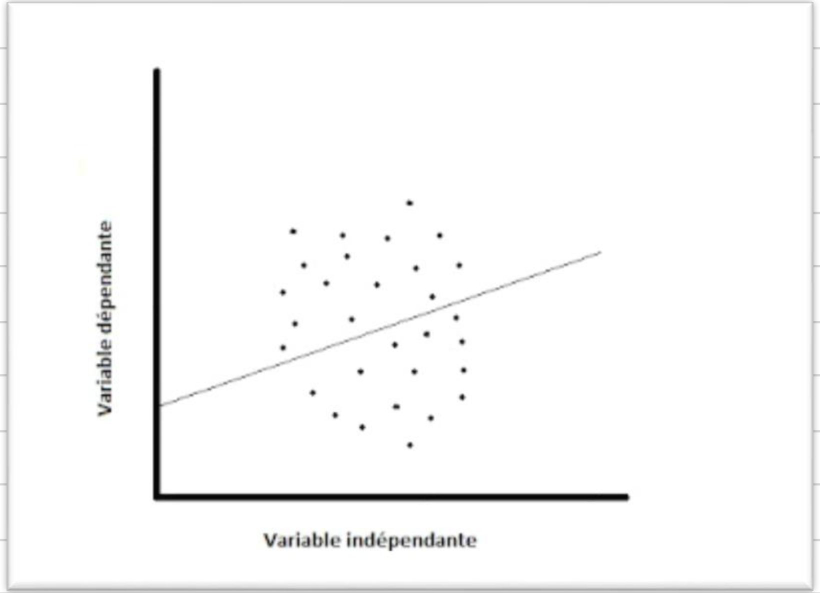
$$\sum_i e_i^2 = \sum_i (y_i - ax_i - b)^2$$

$$b = \bar{y} - a\bar{x}$$

$$a = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$



Mesure de l'intensité de la relation linéaire entre deux variables



Covariance

$$\mathrm{Cov}(x,y) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$$

$\mathrm{Cov}(x,y) > 0 \Leftrightarrow$ x et y varient dans le même sensx

$\mathrm{Cov}(x,y) < 0 \Leftrightarrow$ et y varient en sens contraire

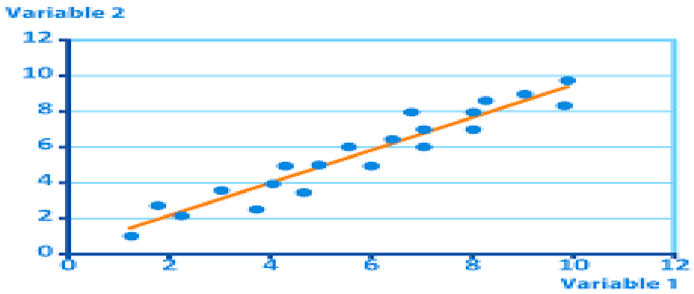
$\mathrm{Cov}(x,y) = \mathrm{Cov}(y,x)$

$\mathrm{Cov}(x,x) = V(x)$

$$Cov(X,Y) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

population count minus one
(Bessel's Correction)

Droites de régression



$$a = \frac{\operatorname{cov}(x,y)}{V(x)} \quad \text{et} \quad b = \bar{y} - a\bar{x}$$

Dy/x : $y = ax + b$ avec

$$a' = \frac{\operatorname{cov}(x,y)}{V(y)} \quad \text{et} \quad b' = \bar{x} - a'\bar{y}$$

Dx/y : $x = a'y + b'$ avec

Merci de votre attention