

Probabilités & statistiques (introduction)

Par Bensalem.H

santé

Analyse des causes des maladies
Mesure de l'efficacité d'un traitement



économie

Analyse des besoins des marchés
Calcul des rendements et déficits



éducation

Étude comparative des approches d'enseignement
Classement des établissements



sociologies

Analyse des comportements sociaux



Extrait de :
<https://www.linternaute.fr/dictionnaire/fr/definition/statistique/>

statistique

statistique , *nom féminin*

Sens 1
Mathématiques Science qui recueille et analyse mathématiquement des données pour une évaluation numérique.

Synonyme : mesure

Traduction anglais : statistic

Sens 2
Ensemble des données numériques concernant l'état ou l'évolution d'un groupe ou d'un phénomène.

Synonymes : description, dénombrement

Traduction anglais : statistic

Effectuer des **calculs** qui **décrivent** (récapitulent) les données relevées lors d'observations concernant un **fait** réel.



Statistique descriptive

Population

échantillon

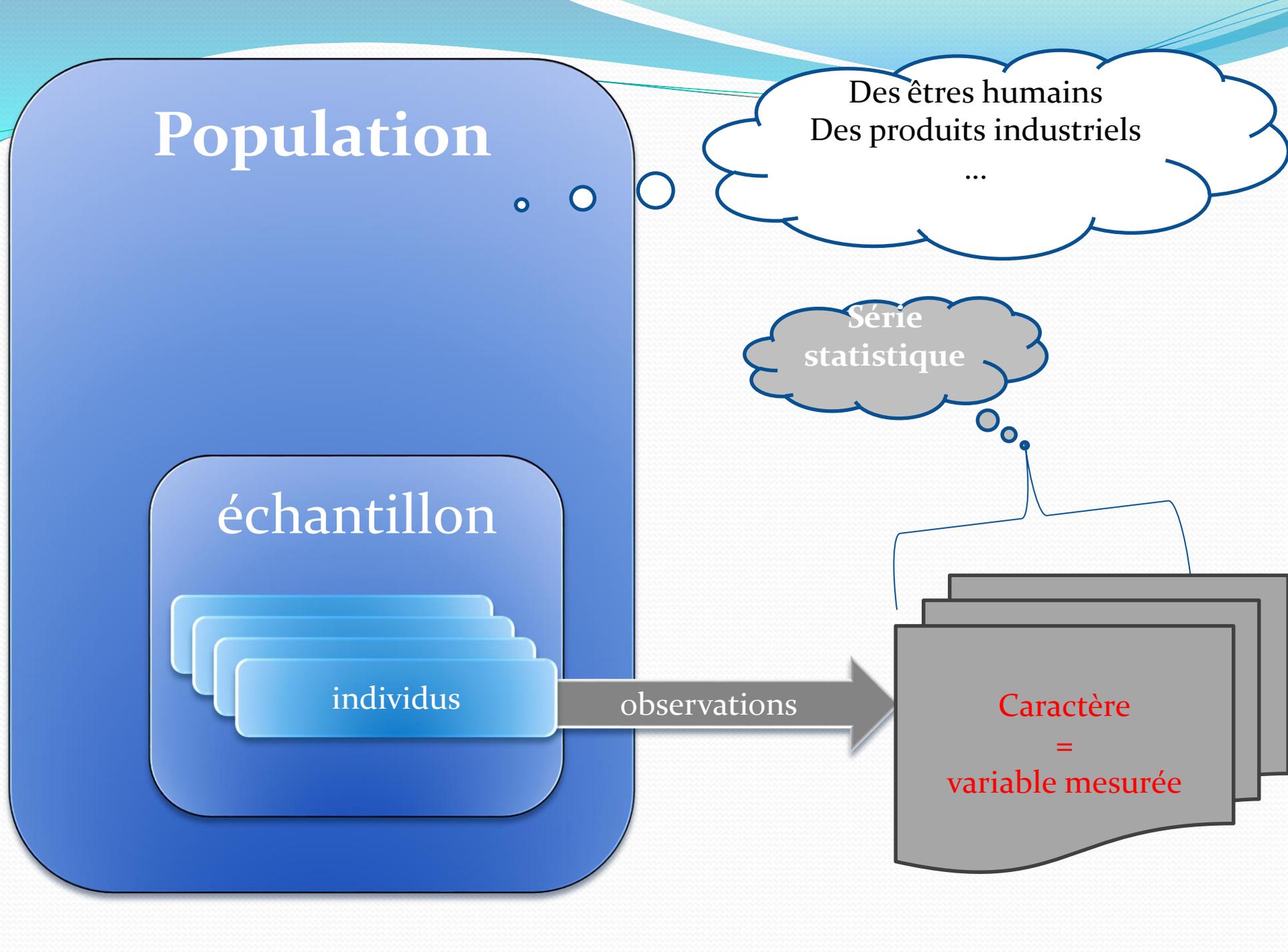
individus

observations

Caractère
=
variable mesurée

Des êtres humains
Des produits industriels
...

Série
statistique



variable

Non
numérique

Ayant des valeurs
numériques:
1,0.25.....

qualitative

quantitative

Catégorielle
avec
hiérarchie:
A,B,C

Sous forme
d'intervalle ou
classe:
Age de [25 ,40ans[

ordinaire

continue

Catégorielle
sans hiérarchie:
blanc, bleu

Exemple: le
nombre d'enfants:
0,1,2,3,4...

nominale

discrete

1	Il m'arrive régulièrement de travailler et d'étudier à la dernière minute. عادتنا اترك الواجبات والدراسة إلى اللحظة الأخيرة							
2	Je comprends très facilement les diverses connaissances théoriques enseignées dans les cours.							

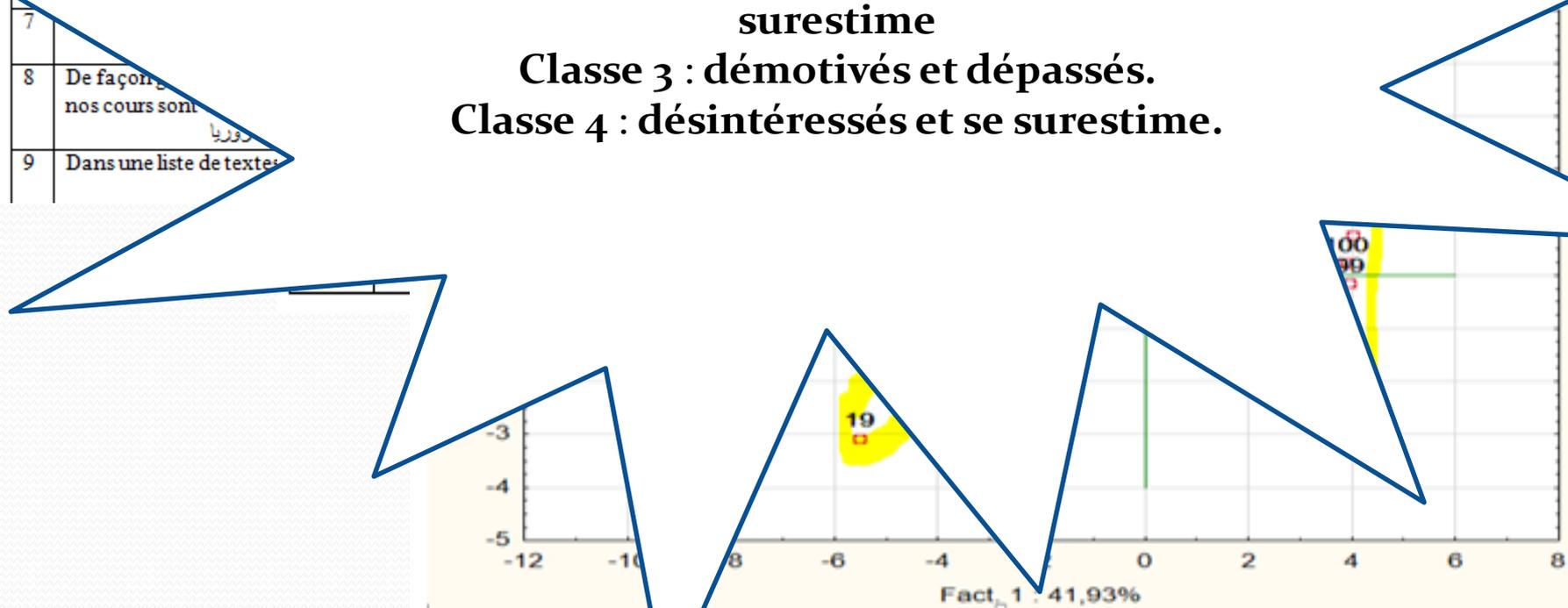
	Indice	Niveau	Potentiel	L'as	Concentration	Communi	de affective	Note	Classe
3	38	5	4	5		5		20	1
4			1	7					
5									
6	99								

Classe 1 : étudiants compétents, engagés, très motivés

Classe 2 : moyennement motivé et se surestime

Classe 3 : démotivés et dépassés.

Classe 4 : désintéressés et se surestime.



Série statistique

L'effectif total N
(taille de l'échantillon)

Modalités x_j
(valeurs distinctes de x)

n_j l'effectif d'une *modalité*
nombre d'apparition de x_j

$f_j = n_j / N$, la fréquence d'une modalité x_j

Tableaux statistiques

Aussi des proportions=
fréquence exprimée sous forme
de pourcentage

Modalité X_i	Effectifs n_i	Effectifs Cumulés N_i	fréquences F_i	Fréquences cumulées $F_i \text{ cum}$
X_1	n_1	$N_1 = n_1$	n_1/N	$\sum_1^i n_i$
X_2	n_2	$N_2 = n_1 + n_2$		
...		$N_j = \sum_1^j n_i$		
Total	N	<hr/>	1	<hr/>

Cas discret:

Cas continue:

Mode

x_i ayant n_i maximum

classe modale ayant n_i maximum

-Idée sur l'ordre de grandeur
-définir le centre des données

X_{ci} est le milieu
d'un intervalle

$$\bar{x} = 1/n \sum_{j=1}^n x_j$$

$$\bar{x} = 1/n \sum_{i=1}^J n_j x_j$$

$$\bar{x} = \sum_{k=1}^n f_i * x_{ci}, \text{ on pose } x_{ci} = (x_{i+1} + x_i)/2$$

Moyenne

Pour une série
ordonnée de X_1 à X_n

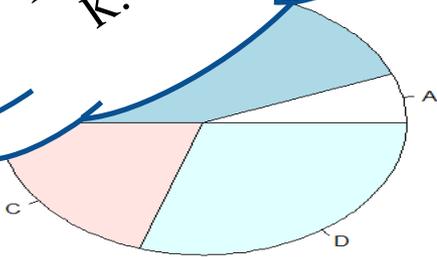
Médiane

Si n est impair : $X_{1/2} = X_{(\frac{n+1}{2})}$
Sinon $X_{1/2} = 1/2(X_{(\frac{n}{2})} + X_{(\frac{n}{2}+1)})$
c'est la valeur ayant la fréquence 0,5
dans la courbe cumulé.

Paramètres de position

Représentation graphique

Nombre de classe (intervalle)
 K: Formules de Sturge ou
 Yale



Cas continue

Diagramme Circulaire, en secteurs ou en camembert

Histogramme de fréquences

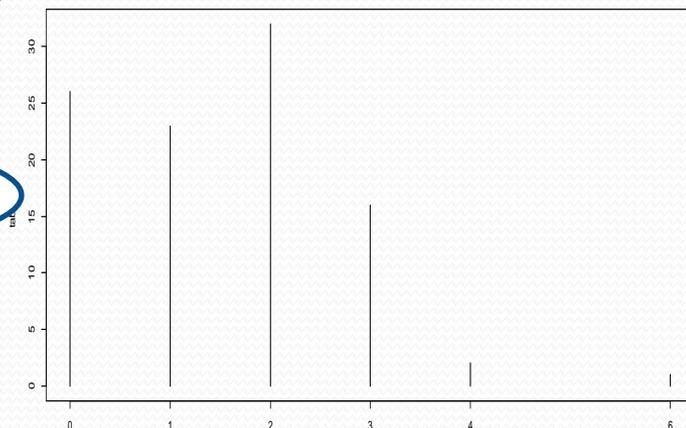
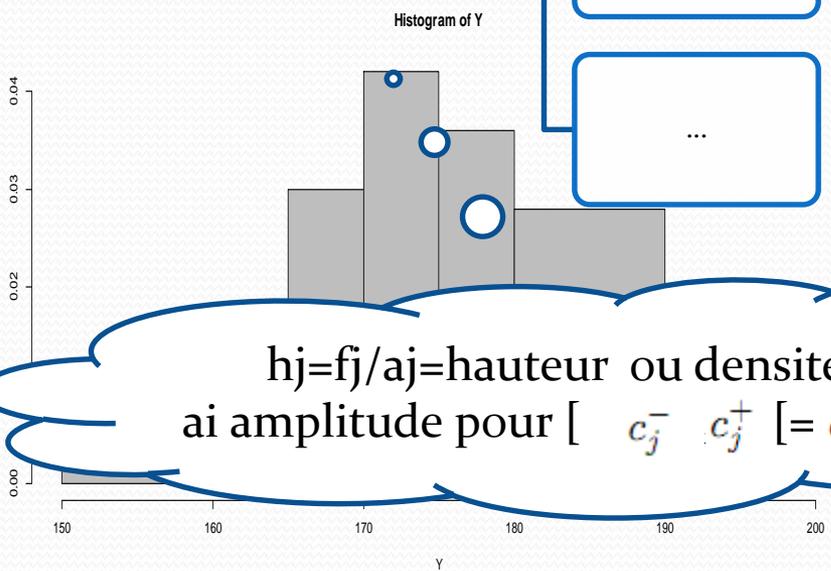
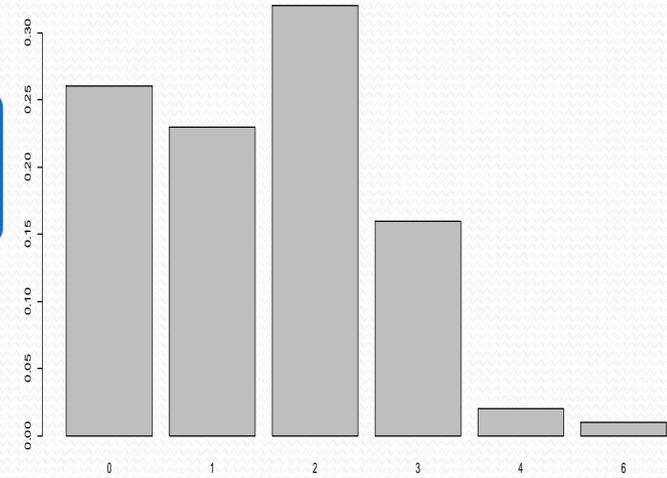
...

Cas discret

Diagramme en barre (ou polygone de fréquences)

Diagramme en bâtonnets (des effectifs)

...



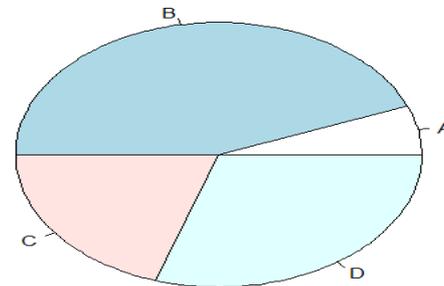
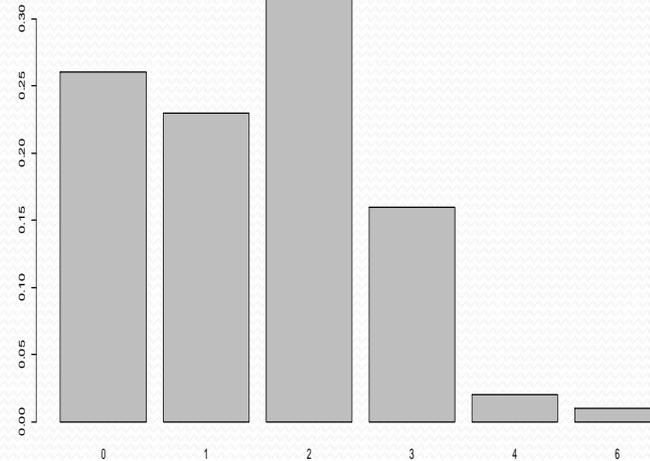
$h_j = f_j / a_j = \text{hauteur ou densité}$
 a_j amplitude pour $[c_j^-, c_j^+]$ $[= c_j^+ - c_j^-]$

Représentation graphique

Variable qualitative

Diagramme en barre (de fréquences ou effective)

Diagramme Circulaire, en secteurs ou en camembert



Fonction de répartition

- Appelée aussi fonction cumulative:
- $F : \mathbb{R} \rightarrow [0,1]$
- F est croissante
- $\lim_{x \rightarrow -\infty} F(x) = 0$
- $\lim_{x \rightarrow +\infty} F(x) = 1$

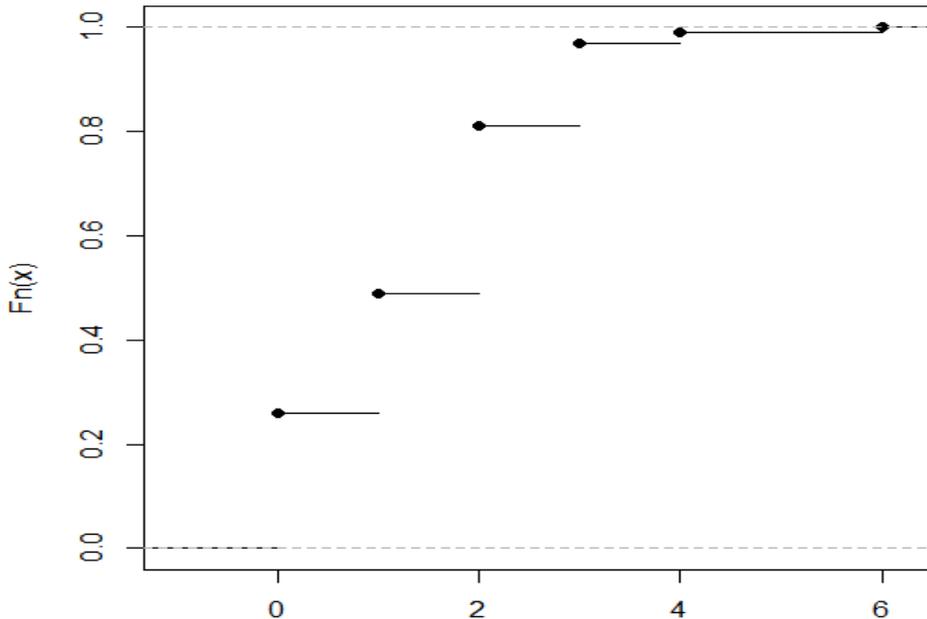
Variable quantitative

F est constante par
morceau, une fonction en
escalier

F continue et affine
(linéaire) par morceau.

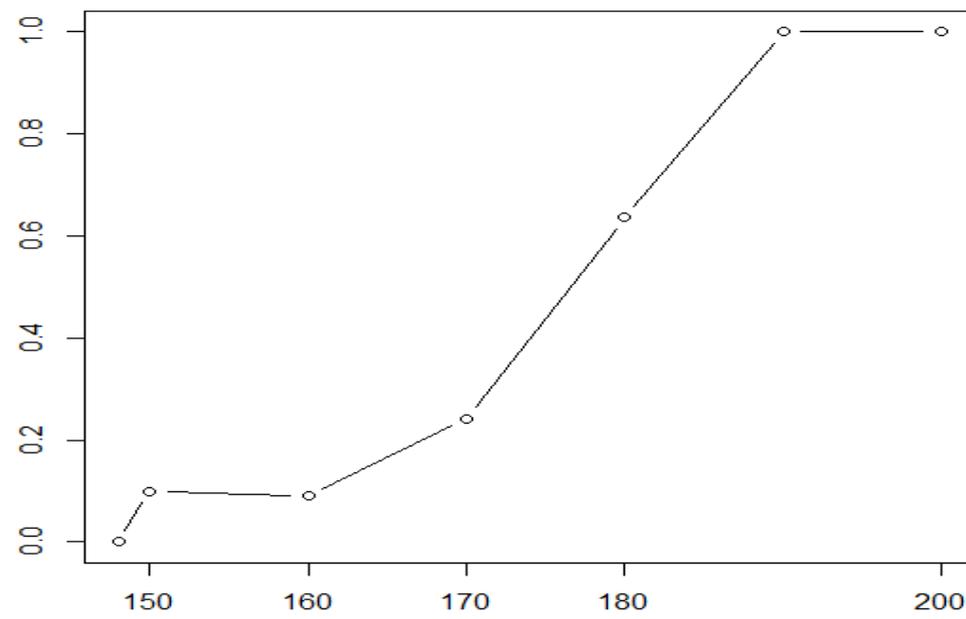
Cas discret

$$F(x) = \begin{cases} 0 & x < x_1 \\ F_j & x_j \leq x < x_{j+1} \\ 1 & x_J \leq x. \end{cases}$$



Cas continu

$$F(x) = \begin{cases} 0 & x < c_1^- \\ F_{j-1} + \frac{f_j}{c_j^+ - c_j^-} (x - c_j^-) & c_j^- \leq x < c_j^+ \\ 1 & c_J^+ \leq x \end{cases}$$



Indicateur de dispersion des données

Paramètres de dispersion

Étendue

$$= X_{max} - X_{min}$$

Quantiles

Pour tout $0 \leq p \leq 1$, un quantile d'ordre p est la fraction de l'échantillon x_p

Une médiane est un quantile d'ordre $p = 0,5 = 1/2$.

Le quantile q_1 d'ordre $p = 0,25 = 1/4$

et q_3 d'ordre $0,75 = 3/4$, Premier et troisième quartile.

Écart inter-quantile

$$= q_3 - q_1$$

théorème de KOENIG

Variance

$$\text{Var}(x) = \delta_x^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2$$

$$\text{var}(x) = \delta_x^2 = \frac{1}{n} \sum_{j=1}^J n_j (x_j - \bar{x})^2 = \frac{1}{n} \sum_{j=1}^J n_j x_j^2 - \bar{x}^2$$

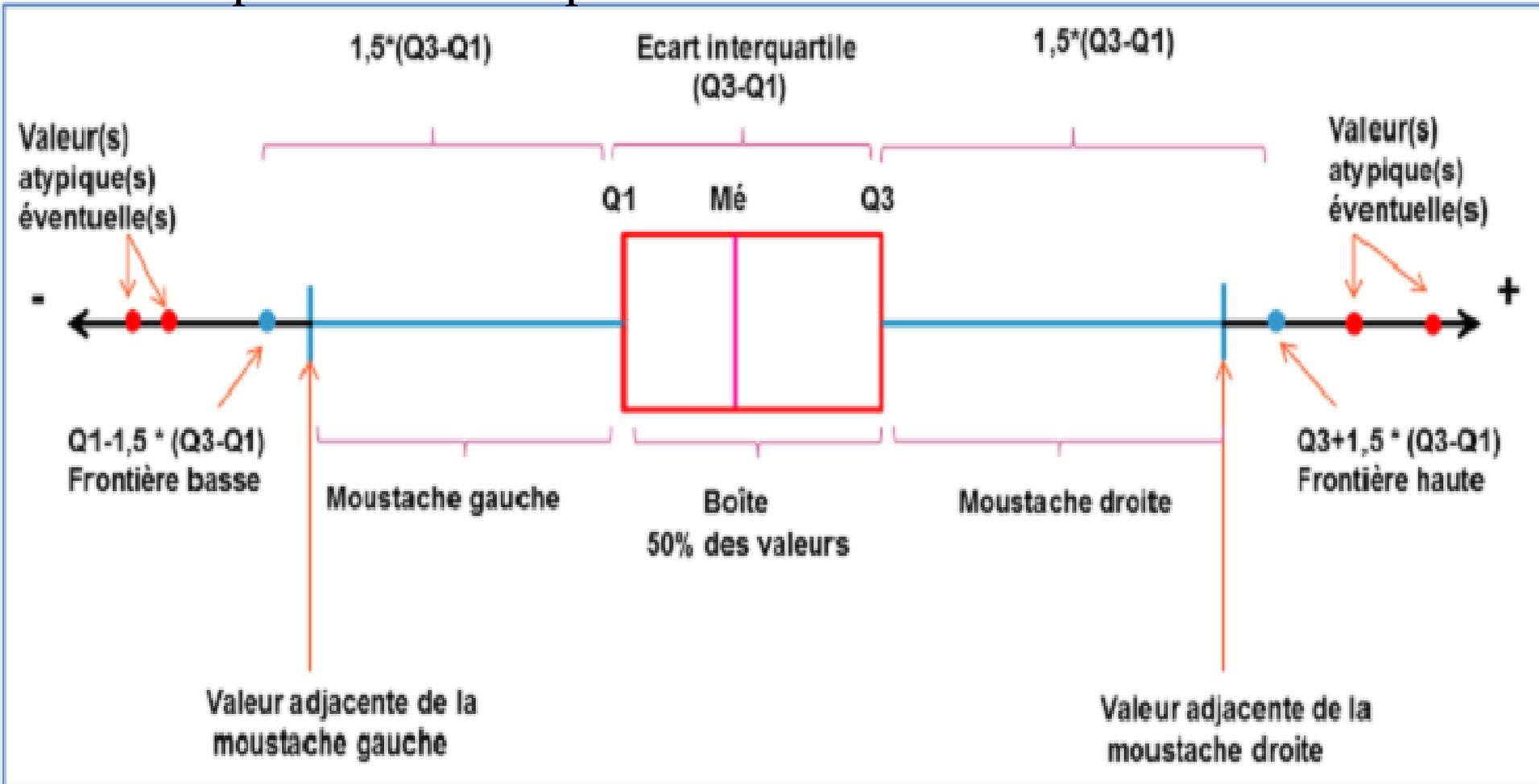
$$\text{var}(X) = \sum_{i=1}^n f_i (x_{ci} - \bar{x})^2$$

Écart type

$$\delta_x = \sqrt{\delta_x^2} = \sqrt{\text{var}(x)}$$

Représentation graphique

-Diagramme en boîte ou boîte à moustache, résume les données sur la série et permet de comparer les séries.



Paramètres bi-variés

Covariance

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

mesure la dépendance
linéaire entre deux
variables

Coefficient
de
corrélation

$$\Gamma_{xy} = \frac{\delta_{xy}}{\sigma_x \sigma_y}$$
$$-1 \leq \Gamma_{xy} \leq 1$$

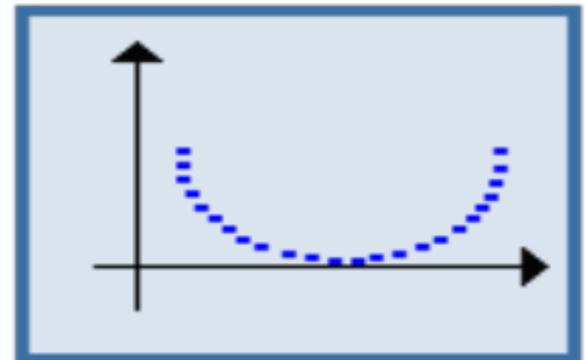
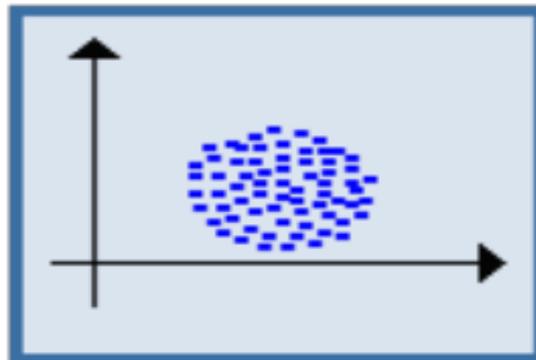
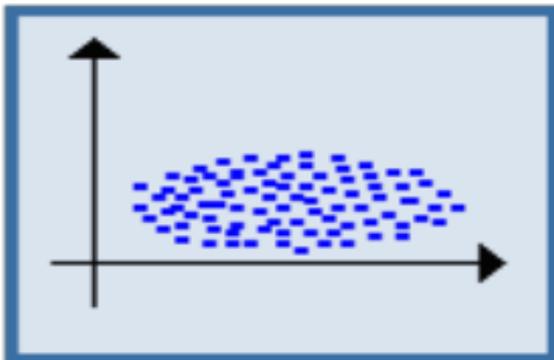
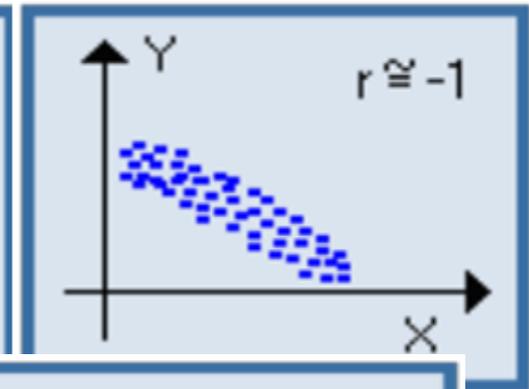
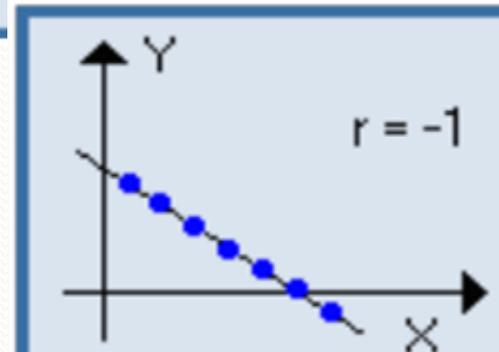
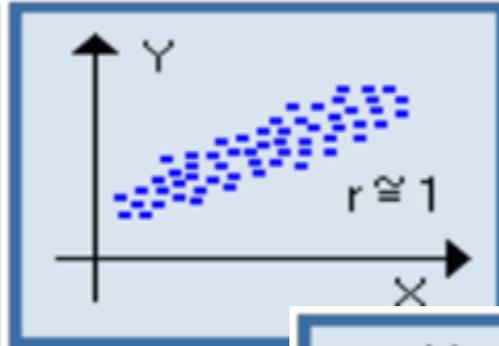
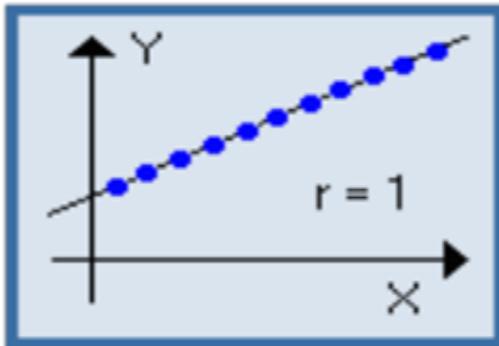
Tableaux de contingence (tableau croisé)

où $n(i,j)$ est l'effectif des observations présentant à la fois la caractéristique x_i et la caractéristique v_j .

	v_1	v_2	v_j		
x_1	$n_{1,1}$	$n_{2,1}$...		
x_2	$n_{2,1}$	N_{22}	...		
x_i	$n_{i,j}$		

REPRESENTATION GRAPHIQUE

- Nuage de points



web graphie

- Sur R: http://cerim.univ-lille2.fr/fileadmin/user_upload/statistiques/michael_genin/Cours/Formation_R_univLille2.pdf
- <http://samuserensemble.canalblog.com/archives/2017/05/24/35307276.html>
- Introduction:
<https://www.youtube.com/watch?v=Mw5yv3SYkMk>
- Boite à moustache:
<https://www.youtube.com/watch?v=oRMzUlhoy6E>
- Exemple:
<https://www.youtube.com/watch?v=xLxA9wCQAxQ>