

Information chiffrée

I Proportions

1- Définition

Propriété :

La probabilité d'un évènement A est $P(A) = \frac{\text{nombre d'issues favorables à A}}{\text{nombre total d'issues}} = \frac{\text{Effectif réalisant A}}{\text{Effectif total}} = \frac{n_A}{n_{\text{total}}}$

Une probabilité est un nombre réel compris entre 0 et 1.

2- Réunion et intersection

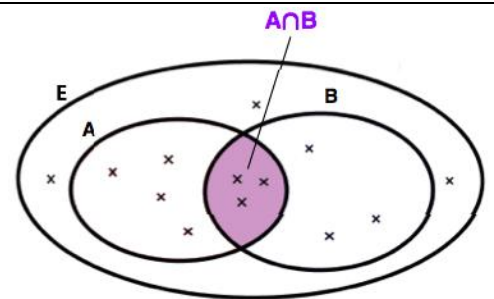
On considère deux sous-populations A et B d'une même population E.

L'**intersection** de A et de B est la sous-population des individus qui appartiennent à la sous-population A **ET** à la sous-population B.

Elle est notée $A \cap B$ et se lit A inter B.

La proportion de cette intersection est égale à :

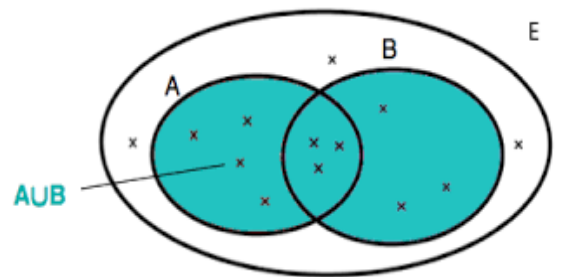
$$P(A \cap B) = \frac{\text{Effectif de } A \cap B}{\text{Effectif total}} = \frac{n_{A \cap B}}{n_E}$$



La **réunion** de A et de B est la sous-population des individus qui appartiennent à la sous-population A **OU** à la sous-population B. Elle est notée $A \cup B$ et se lit A union B.

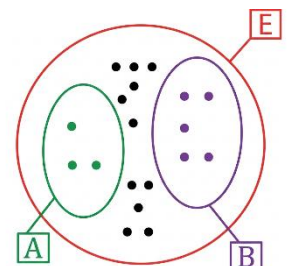
La proportion de cette réunion est égale à :

$$P_{A \cup B} = P_A + P_B - P_{A \cap B}$$



3- Sous-populations disjointes

Deux sous-populations A et B d'une même population E, sont dites **disjointes** lorsqu'elles n'ont aucun individu en commun, c'est-à-dire quand $A \cap B = \emptyset$ et par conséquent quand $P_{A \cap B} = 0$



Application : Exercice 1

II Tableaux croisés

1- Cardinal

Définition :

Soient A et B deux variables étudiées sur une même population. On peut croiser ces deux variables à l'aide d'un **tableau croisé**.

Dans chaque case, on retrouve le **cardinal** du caractère concerné, c'est-à-dire le **nombre d'éléments** de ce caractère.

Exemple : On étudie les employés d'une entreprise selon leur sexe et leur fonction

	Hommes	Femmes	Total
Cadres	56	48	104
Ouvriers	102	94	196
Total	158	142	300

2- Fréquences marginales

Définition :

Les **fréquences (ou probabilités) marginales** de la variable A sont les fréquences de chaque caractère. Elles sont égales à l'effectif de ce caractère sur l'effectif total.

Elle se note $f(A)$ ou $P(A)$

Exemple : On cherche la fréquence des hommes dans l'entreprise

La fréquence marginale des hommes et donc de $\frac{158}{300} \approx 52,7 \%$

	Hommes	Femmes	Total
Cadres	56	48	104
Ouvriers	102	94	196
Total	158	142	300

Elle correspond à la **probabilité** d'être un homme dans cette société.

Application : Exercice 2

III Série statistique à deux variables

1- Définitions

Définition :

Une **série statistique à deux variables** $(x_i ; y_i)$ est constituée d'une liste de n couples de valeurs $(x_1 ; y_1) ; (x_2 ; y_2) ; \dots ; (x_n ; y_n)$.

Le **nuage de points** de cette série est l'ensemble des points du plan de coordonnées $(x_1 ; y_1) ; (x_2 ; y_2) ; \dots ; (x_n ; y_n)$.

Remarque : Le nuage de points permet de visualiser un lien possible entre deux variables ou de repérer des incohérences dans une série statistique à deux variables.

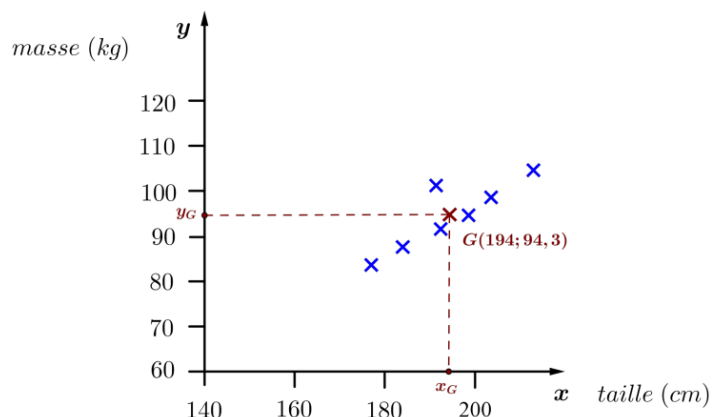
Exemple : Le tableau suivant regroupe la taille et la masse de sept joueurs

Taille (cm)	183	192	177	213	191	199	203
Masse (kg)	88	91	83	105	101	95	97

La représentation graphique est la suivante :

Les **coordonnées du point moyen G** sont $(\bar{x} ; \bar{y})$,
moyenne des x_i , et moyenne des y_i .

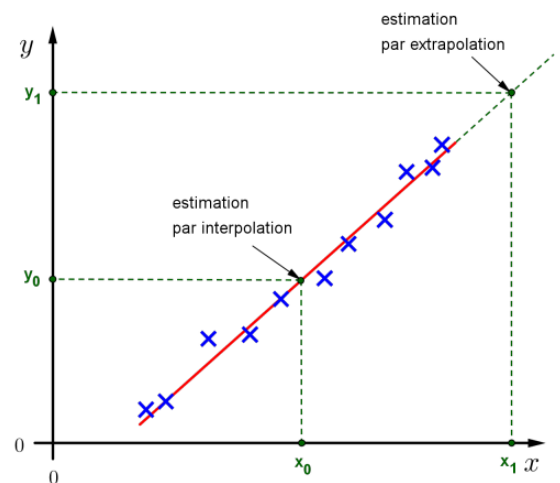
Le nuage de points permet de constater qu'un
des joueurs semble être en surcharge pondérale
par rapport aux autres.



Application : Exercice 3

Remarque : Pour estimer une valeur inconnue dans une série
statistique, on peut utiliser deux procédés :

- La méthode par interpolation quand le calcul est
réalisé dans l'intervalle des valeurs prises par la
série statistique.
- La méthode par extrapolation quand le calcul est
réalisé en dehors des valeurs prises par la série
statistique.



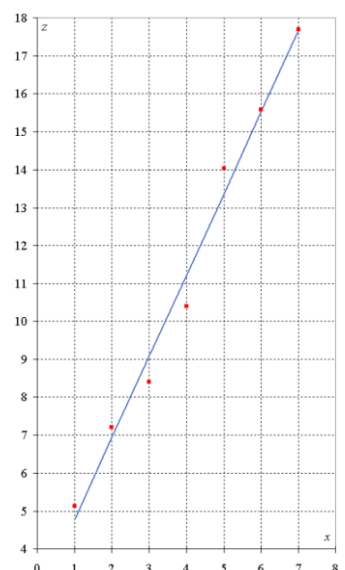
2- Droite d'ajustement

a- Méthode « au jugé »

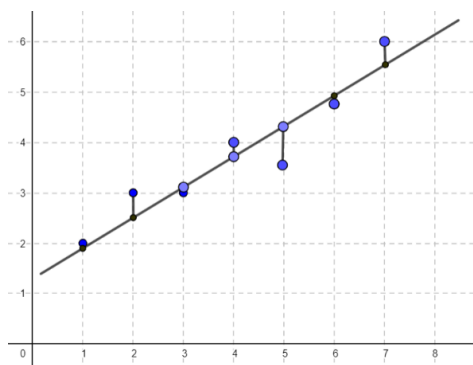
La méthode « au jugé » consiste à tracer une droite passant au plus près
des différents points du nuage.

Après avoir tracé cette droite, on peut en déterminer son équation
graphiquement ou par le calcul en déterminant les coordonnées de deux
points appartenant à la droite (Méthodes précédentes).

Application : Exercice 4



b- Méthode des moindres carrés



L'équation de la droite d'ajustement peut être aussi déterminée à l'aide d'une calculatrice ou d'un tableur, par la **méthode des moindres carrés**.

[Vidéo explicative et tutos Casio/TI](#)

[Tuto Numworks](#)

c- Méthode des points moyens (ou droite de Mayer)

Cette méthode consiste à déterminer deux points moyens du nuage de points et de déterminer par le calcul l'équation de la droite passant par ces deux points moyens.

Exemple : On considère la série statistique suivante

Superficie X	2	2	3	4	5	6	7	7,6
Nombre d'exploitations Y	14	26	31	29	44	40	54	50

Les valeurs du caractère X sont 8 au total. On va donc partager la série en deux séries de 4 valeurs :

Série 1

X	2	2	3	4
Y	14	26	31	29

Point moyen $G_1 (\bar{X}_1 ; \bar{Y}_1)$:

$$\bar{X}_1 = \frac{2 + 2 + 3 + 4}{4} = 2,75$$

$$\bar{Y}_1 = \frac{14 + 26 + 31 + 29}{4} = 25$$

Série 2

X	5	6	7	7,6
Y	44	40	54	50

Point moyen $G_2 (\bar{X}_2 ; \bar{Y}_2)$:

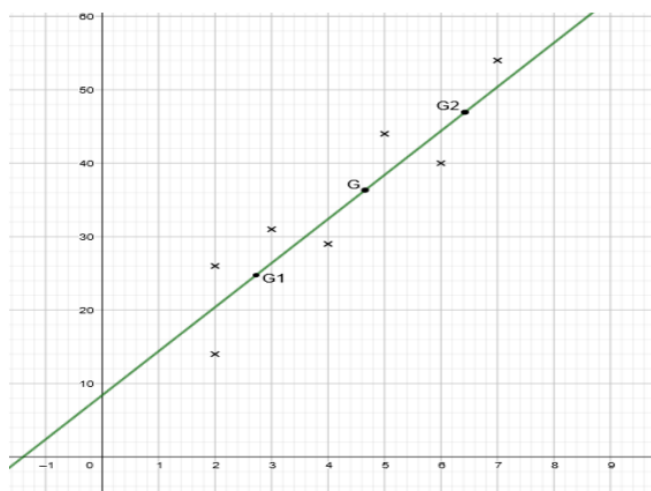
$$\bar{X}_2 = \frac{5 + 6 + 7 + 7,6}{4} = 6,4$$

$$\bar{Y}_2 = \frac{44 + 40 + 54 + 50}{4} = 47$$

On peut ensuite déterminer l'équation de la droite ($G_1 G_2$) grâce à la méthode vue précédemment pour déterminer l'équation réduite d'une droite à l'aide des coordonnées des deux points G_1 et G_2 :

$$a = \frac{47 - 25}{6,4 - 2,75} \approx 6 \quad b = 25 - 6 \times 2,75 = 8,4$$

Donc : l'équation de la droite d'ajustement ($G_1 G_2$) est $y = 6x + 8,4$



Application : Exercice 5