

Statistiques à deux variables

1. Rappels

Pour une population E d'effectif total N,

Modalités	x_1	x_2	...	x_i	...	x_p
Effectifs	n_1	n_2	...	n_i	...	n_p

- L'effectif total est $N = n_1 + n_2 + \dots + n_p = \sum_{i=1}^n n_i$
- La fréquence est : $f_i = \frac{n_i}{N}$ pour i entier compris entre 1 et p.
- La somme des fréquences est $\sum_{i=1}^n f_i = 1$
- La moyenne arithmétique est : $\bar{x} = \frac{\sum_{i=1}^n n_i x_i}{N}$
- La variance V est le réel : $V(x) = \frac{\sum_{i=1}^n n_i (x_i - \bar{x})^2}{N}$

On montre que $V(x) = \frac{\sum_{i=1}^n n_i x_i^2}{N} - \bar{x}^2$

- L'écart type est la racine carrée de la variance $\sigma = \sqrt{V(x)}$

2. Séries statistiques à deux variables

Pour une population donnée, on peut étudier conjointement deux caractères X et Y.

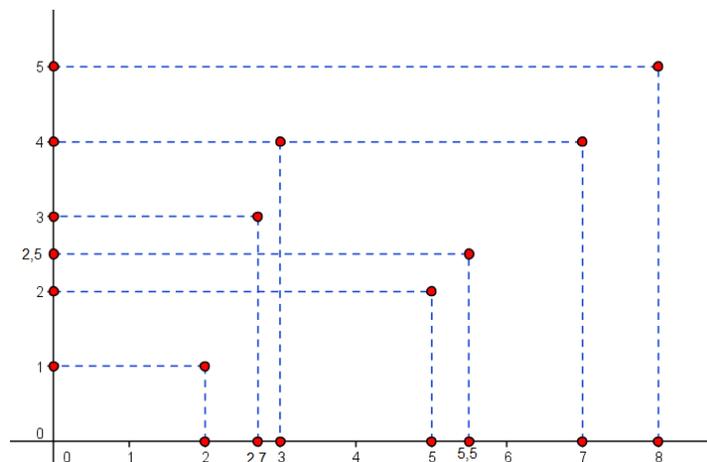
X_i	x_1	x_2	...	x_n
Y_i	y_1	y_2	...	y_n

2.1 Nuage de points

Le nuage de points associé à la série statistique double (X, Y) est l'ensemble des n points $M_i(x_i, y_i)$.

Exemple
Pour la série

x_i	2	2,7	3	5	5,5	7	8
y_i	1	3	4	2	2,5	4	5



2.2 Point moyen

Le point moyen G a est le point qui a pour coordonnées (\bar{x}, \bar{y}) .

2.3 Ajustement affine par la méthode de Mayer

Un ajustement affine est la détermination d'une droite qui se rapproche au mieux un nuage de points représentant une série statistique à deux variable

La méthode de Mayer consiste à partager la série en deux sous séries de même effectif. On obtient alors deux points moyens : $G_1(\bar{x}_1, \bar{y}_1)$ et $G_2(\bar{x}_2, \bar{y}_2)$

La droite (G_1G_2) représente la droite de Mayer.

La forme de son équation est : $y = ax + b$ avec $a = \frac{\bar{y}_2 - \bar{y}_1}{\bar{x}_2 - \bar{x}_1}$ et $b = \bar{y} - a \bar{x}$

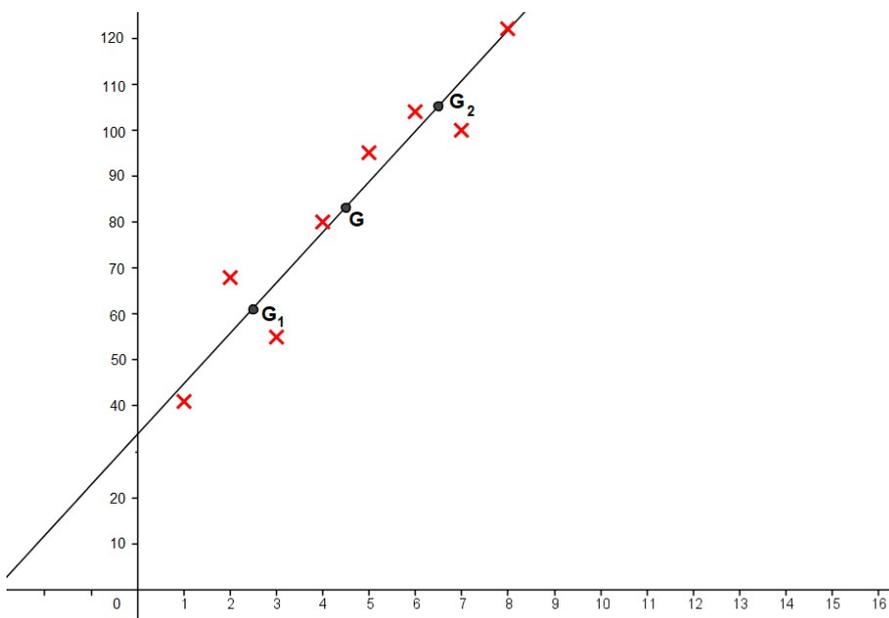
2.4 Exercice résolu

Le tableau suivant donne le chiffre d'affaire d'une entreprise, exprimé en million d'Ariary pendant huit années consécutives.

Nombre d'années (x_i)	1	2	3	4	5	6	7	8
Chiffre d'affaires (y_i)	41	68	55	80	95	104	100	122

- 1) Représenter graphiquement le nuage de points associé à cette série. Sur l'axe des x , 1 cm représente une année, et sur l'axe des y 1 cm représente 20 million. Prendre comme unité respectivement 0 et 30.
- 2) Calculer les coordonnées du point moyen G .
- 3) On partage cette série en deux sous séries de même effectif.
 - a) Calculer les coordonnées des deux points moyens G_1 et G_2 .
 - b) Par la méthode de Mayer faire un ajustement linéaire
 - c) Estimer le chiffre d'affaires en 12 ans.

Réponses



2) $\bar{x} = 4,5$ et $\bar{y} = 10,5$

$G(4,5 ; 10,5)$

3) a) $\bar{x}_1 = 2,5$ et $\bar{y}_1 = 6,1$

$G_1(2,5 ; 6,1)$

$\bar{x}_2 = 6,5$ et $\bar{y}_2 = 10,5$

$G_2(6,5 ; 10,5)$

b) $y = 1,1x + 3,4$

c) Pour $x = 12$, $y = 16,6$, donc le chiffre d'affaires dans 12 ans est 16,6 millions