

Statistique à deux variables

1 - Série statistique

a - Série statistique à deux variables

Une série statistique à deux variables (ou série statistique doubles) est une série statistiques où deux caractères sont étudiés simultanément.

On n'étudiera que des séries statistiques doubles dont les deux caractères étudiés sont quantitatives. Si, pour chacun des n individus de la population, on note x_i et y_i les valeurs prises par les deux caractères, on peut alors présenter la série statistique sous la forme d'un tableau :

Caractère x	x_1	x_2	$\dots$	x_n
Caractère y	y_1	y_2	$\dots$	y_n

b - Nuage de points

Dans un repère orthonormal, l'ensemble des points M_i de coordonnées $(x_i ; y_i)$ constitue le nuage de points associé à la série statistique à deux variables.

c - Point moyen

Le points moyen d'un nuage de n points M_i de coordonnées $(x_i ; y_i)$, noté G , est le point de coordonnées $(\bar{x} ; \bar{y})$ où :

- le réel $\bar{x}$ est la moyenne des x_i et $\bar{x} = \frac{x_1 + x_2 + \dots + x_p}{n}$;
- le réel $\bar{y}$ est la moyenne des y_i et $\bar{y} = \frac{y_1 + y_2 + \dots + y_p}{n}$.

2 - Ajustement affine d'une série statistique à deux variables

a - Définitions :

Effectuer un ajustement d'un nuage de points consiste à trouver une fonction dont la courbe représentative "approche" le nuage, c'est-à-dire dont la courbe passe au plus près des points du nuage.

Quand le nuage présente une forme "rectiligne", la courbe cherchée est une droite d'équation $y = mx + p$. Dans ce cas, on parle d'ajustement affine.

Remarques :

1. Tous les nuages de points ne peuvent pas être approchés par un ajustement affine.
2. Même si le nuage peut être approché par un ajustement affine, il n'y a pas unicité de la droite d'ajustement.

b - Propriété :

On admettra que, pour que l'ajustement affine soit le meilleur possible, il faut que la droite d'ajustement passe par le point moyen G du nuage de points.

c - Méthode des moindres carrés :

→ Utilisation de la calculatrice.