

Série statistiques à deux variables

I. Présentation

1. **Définition**

Déf : On considère une population sur laquelle on étudie simultanément deux caractères.

Ex : Dans la classe de Term.STG1, on observe les notes de Français écrit et Français oral obtenues aux épreuves anticipées du bac.

Rmq : On cherchera s'il existe une corrélation entre les deux caractères.

2. **Nuage de points**

Déf : Soit (x_i) et (y_i) l'ensemble des valeurs prises par les caractères X et Y d'une série double. Alors, l'ensemble des points $M(x_i; y_i)$ dans un repère $(O; \vec{i}, \vec{j})$ est appelé nuage de points.

3. **Point moyen**

Déf : On appelle point moyen d'un nuage de points le point $G(\bar{x}; \bar{y})$ où \bar{x} est la moyenne de la série X et \bar{y} la moyenne de la série Y.

II. Ajustement affine

Principe : déterminer un éventuel lien entre les deux caractères. Pour cela, on observe le nuage de points, si ce dernier semble plutôt « allongé », alors on définit une droite d'ajustement, c'est-à-dire que $Y = aX + b$.

Il existe plusieurs méthodes pour déterminer l'équation de cette droite :

1. **Au jugé**

On trace approximativement une droite la plus près possible de chacun des points.

2. **Méthode de Mayer**

On divise la série en deux (si possible de même effectif) on détermine le point moyen G_1 de la première moitié puis G_2 de la deuxième moitié. La droite de Mayer est la droite (G_1G_2) .

3. **Méthode des moindres carrés**

Son équation est déterminée par la calculatrice.

Propriété : elle passe par le point moyen.