

Synthèse Statistique bivariée

Définition : Une série double de taille n portant sur deux caractères quantitatifs x et y , est un ensemble S de n couples de réels.

Définition : Le plan P étant muni d'un repère orthogonal, on peut associer pour chaque couple (x_i, y_i) de la série double un point M_i de coordonnées (x_i, y_i) . L'ensemble des points M_i obtenus s'appelle un nuage de points.

Définition : On appelle point moyen du nuage le point G de coordonnées (\bar{x}, \bar{y}) .

Définition : La covariance d'une série statistique double (X, Y) , notée $s_{X,Y}$ est définie par :

$$s_{X,Y} = \frac{1}{N} \sum_{i=1}^n \sum_{j=1}^n n_{i,j} (x_i - \bar{x})(y_j - \bar{y})$$

Propriété :

1. $s_{XY} = \overline{XY} - \bar{X}\bar{Y}$ (formule de la covariance)
2. $s_{XX} = s^2(X)$

Définition : Le coefficient de corrélation linéaire de la série double (X, Y) , noté r_{XY} , est défini, si X et Y ont des variances non nulles, par :

$$r_{XY} = \frac{s_{XY}}{s(X)s(Y)}$$

Propriété : $r_{XY} \in [-1; 1]$.

Définition : Soit (X, Y) une série statistique double. La droite d'ajustement affine D de y en x , obtenue par la méthode des moindres carrés, est la droite définie par l'équation $y = ax + b$ avec $a = \frac{s_{XY}}{s_X^2}$ et $b = \bar{y} - a\bar{x}$.