

Statistiques à 2 variables

Définition. L'ensemble des points M_i de coordonnées $(x_i; y_i)$, est appelé **le nuage de points** associé à la série statistique $(x_1; y_1), \dots, (x_n; y_n)$ à deux variables.

Définition. Le point de coordonnées $(\bar{x}; \bar{y})$, où $\bar{x} = \frac{x_1 + \dots + x_n}{n}$ et $\bar{y} = \frac{y_1 + \dots + y_n}{n}$, sont les moyennes respectives des x_i et des y_i , est appelé le **point moyen** du nuage de points associé à la série.

Définition. Lorsque les points d'un nuage sont sensiblement alignés, on peut construire une droite, appelé **droite d'ajustement** (ou **droite de régression**), passant « au plus près » de ces points.

Remarque. Il existe plusieurs méthodes permettant d'obtenir la droite d'ajustement :

La **méthode au jaugé**, qui consiste à tracer une telle droite sans méthode particulière. Cette méthode est rapide mais pas assez précise.

La **méthode des points moyens** qui consiste à faire passer la droite par 2 points moyens (on divise en deux le nuage de points, on calcule pour chaque groupe le point moyen associé, on place les deux points moyens et on trace la droite ainsi formée)

La **méthode des moindres carrés** qui est celle que l'on utilise le plus souvent car plus précise. Cette méthode s'effectue à la calculatrice.

Avec TI :

- Appuyer sur « **STATS** » puis « **Edite** » et saisir les valeurs de X_i dans L1 et les valeurs de y_i dans L2.
- Appuyer à nouveau sur « **STATS** » puis « **CALC** » et « **RegLin(ax+b)** ».
- Saisir **L1,L2**

Avec CASIO :

- Aller dans le menu « **STAT** ».
- Saisir les valeurs de X_i dans *List1* et les valeurs de y_i dans *List2*.
- Sélectionner « **CALC** » puis « **SET** ».
- Choisir *List1* pour 2Var *XList* et *List2* pour 2Var *YList* puis « **EXE** ».
- Sélectionner « **REG** » puis « **X** » et « **aX+b** ».

Avec Numworks :

- Un menu est dédié aux régressions, il suffit de se remplir les données et ensuite avec le curseur se déplacer sur l'onglet « stat »

Vocabulaire. A partir de l'ajustement obtenu, on peut faire des **estimations** :

On **extrapolera** pour des valeurs à l'extérieur de l'intervalle des valeurs connues.

On **interpola** pour des valeurs à l'intérieur de l'intervalle des valeurs connues.