# Présenter et traiter des résultats numériques

Pour présenter des valeurs numériques obtenues et pour faciliter leur interprétation, on les ordonne sous forme de tableaux et on construit des représentations graphiques. Différents logiciels tels que Excel ou OpenOffice facilitent beaucoup le traitement des données.

## I. Étude d'une seule variable

On effectue une série de *N* observations et on s'intéresse à une variable numérique *x*. Par exemple on s'intéresse à la taille des 29 élèves d'une même classe.

#### 1. Tableau chronologique

Dans un premier temps on note les valeurs  $x_1, ..., x_n$  prises par la variable x simplement dans l'ordre chronologique de leur obtention. On obtient un tableau chronologique.

On indique l'unité de mesure utilisée.

Ces valeurs doivent être données avec un certain nombre de chiffres significatifs : il faut tenir compte de la précision des appareils de mesure.

#### 2. Tableau des effectifs et diagramme en bâtons

Si une valeur  $x_k$  est obtenue plusieurs fois, on appelle **effectif** le nombre  $n_k$  de fois où x prend la valeur  $x_k$ .

On appelle **fréquence** de la valeur  $x_k$  le rapport  $f_k = \frac{n_k}{n}$ .

On peut ranger les valeurs prises par la variable *x* dans l'ordre croissant:

 $x_1 < x_2 < ... < x_p$  et les placer dans la première colonne d'un tableau. Dans la deuxième colonne on note en face de  $x_k$  son effectif  $N_k$  et dans la troisième la fréquence  $f_k$  de la valeur  $x_k$ .

La somme des effectifs  $n_k$  est égale à la taille n de l'échantillon observé, la somme des fréquences est égale à 1.

On peut représenter ce tableau par un graphique.

On reporte en abscisses les valeurs  $x_1, x_2, ..., x_p$  prises par x et en ordonnées des longueurs proportionnelles aux effectifs correspondants  $n_1, ..., n_p$  (ou aux fréquences  $f_1, ..., f_p$ ). On obtient un **diagramme en bâtons**.

## 3. Tableau récapitulatif et histogramme

Lorsque les valeurs prises par x sont nombreuses, on peut simplifier le tableau précédent.

Si *x* prend ses valeurs dans un intervalle *I*, on fait une partition de *I* en intervalles  $[a_1, a_2[, ...[a_q, a_{q+1}[$ . On place ces intervalles dans la première colonne d'un tableau. Dans la deuxième on note en face de  $[a_k, a_{k+1}[$  son effectif  $N_k$ . Dans une troisième colonne on peut noter la fréquence correspondante  $f_k = \frac{n_k}{n}$ .

Dans notre exemple on a choisi les intervalles:

[160, 165[, [165, 170[, [170, 175[, [175, 180[, [180, 185[, ]185, 190[, [190, 195[.

On peut représenter ce tableau par un histogramme.

On porte sur l'axe des abscisses les nombres  $a_1, a_2, ..., a_{q+1}$ . Puis on trace des rectangles dont l'aire est proportionnelle aux effectifs  $n_k$  (ou aux fréquences  $f_k$ ) et de bases les segments  $[a_k, a_{k+1}]$ .

## 4. Exemple utilisant EXCEL

La variable *x* est la taille des 29 élèves d'une classe.

On saisit les valeurs de *x* dans la plage B2:B30 d'une feuille de calcul.

Dans la plage C2:C7 on porte les nombres 164, 169, 174, 179, 184, 189 définissant les intervalles.

Dans la plage E2:E8 on calcule les effectifs correspondant aux intervalles. Par exemple dans E3 on calcule l'effectif correspondant à l'intervalle [165, 170].

On sélectionne la plage E2:E8 avec la souris. On clique sur l'assistant graphique. On choisit le type histogramme. On appuie sur suivant, puis série. On sélectionne dans Etiquette des abscisses (X) la plage D2:D8 où on a calculé les milieux des intervalles  $[a_k, a_{k+1}]$ . On obtient l'histogramme des effectifs.



**Remarque** : Le graphique appelé par EXCEL histogramme est en fait un diagramme en bâtons : les bases bases des rectangles  $[a_k, a_{k+1}]$  ne sont pas proportionnelles à leurs longueurs.

**Moyenne et écart type :** Les fonctions MOYENNE et ECARTYPEP d'EXCEL permettent de calculer la moyenne et l'écart type des variables étudiés.

## II. Étude de deux variables

Au cours d'une série de *N* observations on peut s'intéresser simultanément à deux variables numériques *x* et *y*. On peut par exemple s'intéresser à la taille et au poids des 29 élèves de la classe. On dispose des valeurs  $x_1, ..., x_n$  de la variable *x* et  $y_1, ..., y_n$  de la variable *y*. On cherche à déterminer si il existe une fonction *f* telle que f(x) soit une approximation de *y*.

## 1. Représentation graphique

Dans le plan xOy on représente les valeurs prises par les deux variables par les points  $M_1(x_1, y_1), ..., M_n(x_n, y_n)$ . On obtient un **nuage de points**.

Le problème d'ajustement consiste à chercher, parmi les fonctions d'un certain type, la fonction f telle que les approximations  $y_1 = f(x_1), ..., y_n = f(x_n)$  soient les meilleures possibles.

## 2. Droite de régression

Lorsque le nuage de points a une forme presque rectiligne, on cherche à trouver la "meilleure droite" représentant l'ajustement : c'est la droite de régression.

## Tableau de la taille et du poids de 29 élèves d'une classe :

Taille en cm	Poids en kg							
178	76							
187	76							
172	62							
161	58							
168	58	90						
168	53						_	
174	50	f(x) = 0.9	97x - 107,53		• _	9		-
180	60	70	5					
172	58		_	_				
167	50	60		•				
168	52	50		_				
165	57		-	-				
168	54	40						
163	56							
193	81	30 - 155 160	165 170	175	180	185	190	195
187	70	100 100	100 110		100	100	100	100
173	71							
177	63							
166	54							
160	46							
173	58							
182	74							
179	75							
160	46							
172	55							
180	73							
169	56							
167	56							
190	80							
173,06	60,9655172	covariance	74,6575505					
8,75699143	9,86978237	coeff.cor.lin	0.86379609					

On peut utiliser EXCEL pour traiter les valeurs observées :

On saisit les valeurs des variables x et y sur deux colonnes juxtaposées, on sélectionne avec la souris les valeurs de x et y. On clique sur l'assistant graphique et on choisit le type nuage de points. Un graphique apparaît sur l'écran. On place la souris sur un point et on clique sur le bouton droit de la souris. En sélectionnant Ajouter une courbe de tendance, puis en choisissant le type linéaire, on obtient le tracé de la droite de régression.

**Corrélation :** Les fonctions COVARIANCE et COEFFICIENT.CORRELATION d'Excel permettent de calculer ces quantités.

Parfois les points du nuage des points sont obtenus en faisant la moyenne de plusieurs valeurs. EXCEL offre la possibilité de faire apparaître sur le graphique l'écart type.

# Évolution de la taille des hommes et des femmes en France de 1970 à 2001.



**Avec Excel** : On sélectionne les trois premières colonnes (étiquettes comprises) et on clique sur l'assistant graphique. On choisit "nuage de points".

On place la souris sur l'un des points de la courbe et on clique avec le côté droit sur "Format de la série des données", puis "Barre d'erreur Y", "afficher les deux", "Personnalisée". On sélectionne ensuite la plage D2:D5. L'écart type apparaît. On peut aussi sélectionner "Ajouter une courbe de tendance" pour obtenir la droite de régression.

On peut aussi cliquer sur Axe des X ou Y, sélectionner "Format de l'axe", "échelle" pour modifier les valeurs max et min des axes.

**Avec OpenOffice** : Dans "type de diagramme" on choisit "XY (dispersion)" qui est l'équivalent de "nuage de points" d'Excel.

Dans "plage des données" on choisit "Première ligne comme étiquette".

Dans "éléments du diagramme" on peut choisir des titres pour l'axe des X, des Y etc...

On place la souris sur l'un des points de la courbe et on clique avec le côté droit sur "Barre d'erreurs des Y". puis "Plage des cellules" et "valeur identique pour les deux". On sélectionne ensuite la plage D2:D5. L'écart type apparaît. On peut aussi sélectionner "Ajouter une courbe de tendance" pour obtenir la droite de régression.

Dans Format de la barre des outils on peut sélectionner "Axe" pour modifier les axes de X ou Y.