

I. Population statistique/ Caractère:

1) Etude d'un exemple:

Regardons les deux exemples suivants:

Exemple :1

L'étude suivante donne une répartition de 20 élèves suivant le nombre de villes visitées par chacun d'eux :
5-4-3-4-5-3-1-2-1-4-1-3-4-2-4-3-1-4-1-5

Exemple :2

Les vitesses de 150 voitures ont été détectées sur l'autoroute entre Rabat et Casa, on a obtenu le tableau suivant :

vitesse	$[50,70[$	$[70,90[$	$[90,110[$	$[110,130[$	$[130,150[$
effectif	10	40	60	25	15

2) Définitions:

Population statistique :

La population statistique est l'ensemble qui fait l'objet de l'étude . et cha que élément de cet ensemble est appelé « **individu** » ou « **unité statistique** »

Dans le premier exemple la population statistique est l'ensemble des **élèves**.

Dans le premier exemple la population statistique est l'ensemble des **voitures**.

Caractère : la propriété qu'on veut étudier " chez une population statistique s'appelle « **le caractère** » ou « **la variable statistique** ». " le caractère peut être **quantitatif** ou **qualitatif** .

II. Types de caractères:

1) Caractère quantitatif :

Le caractère quantitatif est un caractère qui peut s'exprimer par des nombres, on distingue le caractère quantitatif **discret** et le caractère quantitatif **continu** .

Caractère discret :

le caractère quantitatif **discret** est celui qui prend des valeurs **isolées** , comme le numéro du mois de naissance d'un élève par exemple.

Caractère continu :

le caractère quantitatif continu est celui qui prend des valeurs très proches, dans ce cas les valeurs du caractères sont rassemblées dans des intervalles qu'on appelle aussi des « **classes** » , comme les hauteurs des élèves par exemple.

2) Caractère qualitatif :

Le caractère qualitatif est un caractère qui ne peut s'exprimer par des nombres, comme la couleur du cheveux des élèves ou leur groupe sanguin ou leur sexe.

III. Effectif - fréquence:

1) En général:

Pour étudier une série statistique:

On classe les valeurs du caractère d'une manière croissante et on note le nombre d'apparition n_i de la valeur x_i du caractère ; n_i est alors le nombre d'unités de la population caractéristique qui prennent la valeur x_i , on l'appelle l'effectif relatif à la valeur x_i du caractère.

Le nombre d'unités de la population dont la valeur du caractère est inférieure à x_i est N_i tel que

$N_i = n_1 + n_2 + n_3 + \dots + n_i$, N_i est l'effectif cumulé relatif à la valeur x_i du caractère. Le nombre N représente l'effectif total.

Dans l'exemple 1 $n_4 = 6$ est l'effectif de la valeur $x_4 = 4$ du caractère.

Le nombre : $f_i = \frac{n_i}{N}$ est appelé « fréquence » relative à la valeur x_i du caractère.

Le nombre : $P_i = 100 \times f_i$ est appelé le pourcentage relatif à la valeur x_i du caractère.

2) Application à un exemple:

Compléter le tableau suivant à partir des données de l'exemple 1:

Caractère x_i : nombre de villes	1	2	3	$x_4 = 4$	5
effectif n_i : nombres d'élèves	5	2	4	$n_4 = 6$	3
effectif cumulé : N_i	5	7	11	$n_4 = 17$	20
fréquence : $f_i = \frac{n_i}{N}$	$\frac{5}{20}$	$\frac{2}{20}$	$\frac{4}{20}$	$\frac{6}{20}$	$\frac{3}{20}$
pourcentage : $P_i = 100 \times f_i$	25	10	20	30	15
fréquence cumulée $F_i = \frac{N_i}{N}$:	$\frac{5}{20}$	$\frac{7}{20}$	$\frac{11}{20}$	$\frac{17}{20}$	1

IV. Paramètres de position : Mode (classe modale) , moyenne , médiane (classe médiane)

1) le mode:

c'est la valeur du caractère ou la classe correspondant au plus fort effectif.

Remarque : pour déterminer le mode , il faut d'abord dresser le tableau des effectifs.

2) La moyenne :

x_i est la valeur de la variable discrète ou le centre dans le cas de variable continue

La moyenne qu'on peut noter \bar{x} est déterminée par la relation suivante: $\bar{x} = \frac{n_1x_1 + n_2x_2 + \dots + n_kx_k}{N}$

En fonction des fréquences, on obtient : $\bar{x} = f_1x_1 + f_2x_2 + \dots + f_kx_k$

Remarque : pour faire le calcul de la moyenne , on complète le tableau statistique pour la ligne $f_i x_i$.

3) La médiane:

La médiane peut être notée Me

La médiane est la plus petite valeur du caractère dont l'effectif cumulé est plus grand ou égal à la moitié de l'effectif total.

Remarque : pour déterminer la médiane , il faut d'abord dresser le tableau des effectifs cumulés.

V. Paramètres de dispersion: Etendue, Ecart-moyen, Variance et Ecart-type

1) Etendue:

C'est la différence entre les valeurs extrêmes .

Dans l'exemple 1, la valeur minimale est 1 et la valeur maximale est 5, donc l'étendue est égale à $5-1=4$

Remarque : l'étendue est un enregistrement utile pour constater la dispersion de la série.

2) Ecart-moyen:

C'est la moyenne des écarts à la moyenne \bar{x} .

Il rend bien compte de l'étalement mais n'est pas couramment utilisé.

$$\bar{e} = \frac{n_1|x_1 - \bar{x}| + \dots + n_k|x_k - \bar{x}|}{N}$$

Remarque : pour faire le calcul de l'écart-type , on ajoute les lignes $|x_i - \bar{x}|$ et $n_i|x_i - \bar{x}|$ au tableau statistique .

3) Variance et Ecart-type:

La variance est la moyenne des carrés des écarts à la moyenne \bar{x} .

La variance est noté **Varx** ou **V(x)** ou $\sigma^2(x)$ ou σ^2 (σ étant la notation de l'écart-type)

$$\text{Varx} = \frac{n_1(x_1 - \bar{x})^2 + \dots + n_k(x_k - \bar{x})^2}{N}$$

Remarque : En développant le numérateur et en simplifiant, on trouve : $\text{Varx} = \frac{n_1(x_1)^2 + \dots + n_k(x_k)^2}{N} - (\bar{x})^2$

l'écart-type est la racine carrée de la variance : $\sigma = \sqrt{V}$.

VI. Représentation graphique des données statistiques :

1) Variable discrète:

Si l'on souhaite utiliser représenter les effectifs correspondant à une variable discrète, on obtient un diagramme en bâtons en portant, dans un repère orthogonal, les valeurs x_i en abscisses, les effectifs en ordonnées .

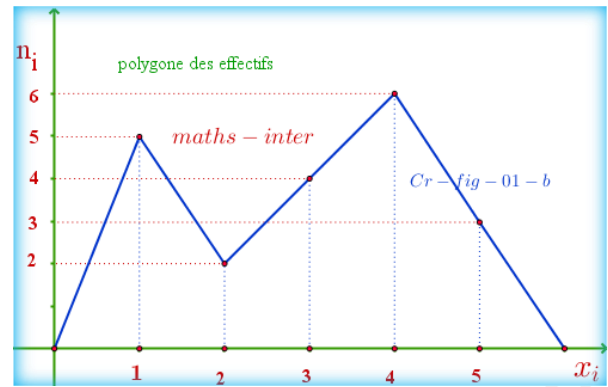
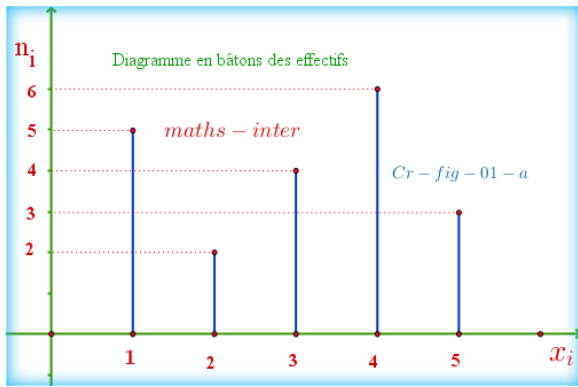
Par un changement d'échelle sur l'axe des ordonnées, on obtient le diagramme en bâtons des fréquences.

Le polygone des effectifs est la ligne polygonale obtenue en joignant les extrémités des bâtons (à la règle). On ferme cette ligne polygonale en considérons des valeurs extrêmes que peut prendre la variable discrète.

Le polygone permet de visualiser plus facilement la variable discrète.

On peut de la même façon représenter le diagramme en bâtons des effectifs cumulés croissants ou décroissant (ou des fréquences), ainsi que les polygones correspondants.

Exemple : représentons les données de l'exemple1



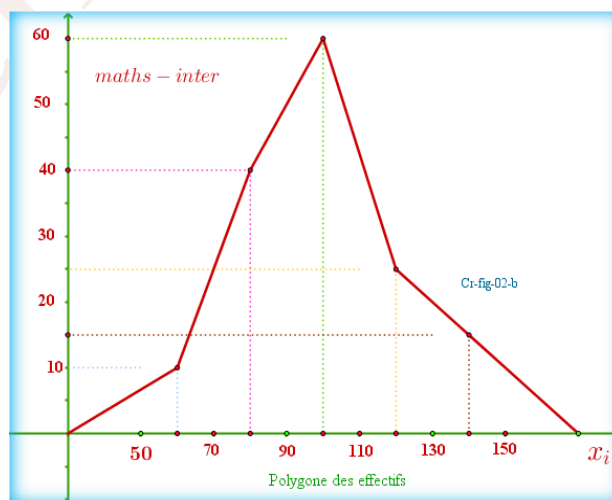
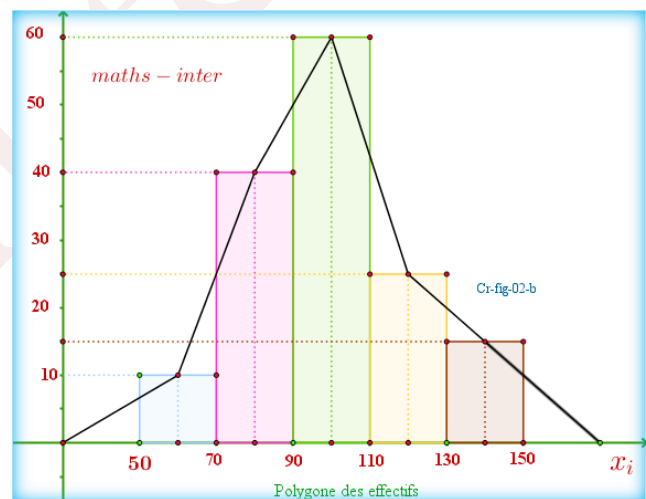
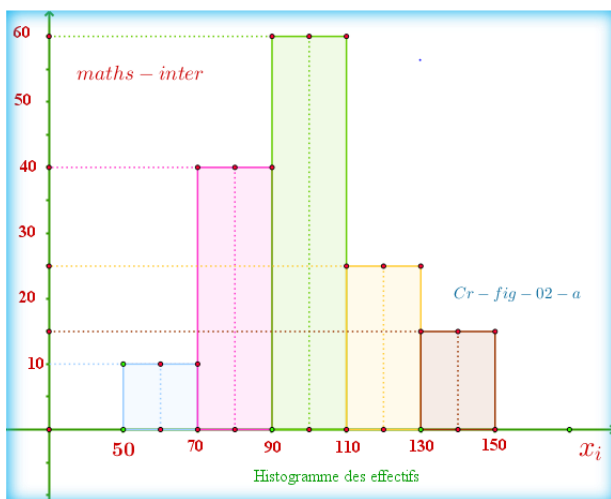
2) **Variable Continue: Représentation des effectifs**

On trace principalement des histogrammes, c'est à dire une suite de rectangles dont l'aire est proportionnelle à l'effectif.

Dans la pratique, on porte en abscisses les classes, et en ordonnées les effectifs, si les classes ont même amplitudes.

On trace alors des rectangles dont les bases représentent les amplitudes de la classe, et la hauteur l'effectif correspondant. La méthode est la même lorsqu'on représente les fréquences les effectifs cumulés croissant ou décroissant (et les fréquences du même genre)

Exemple : représentons les données de l'exemple2



VII. Détermination analytique de la médiane dans le cas des classes :

1) **Méthode analytique:**

On a besoin de construire le polygone des effectifs cumulés .

On lit la médiane sur l'axe des abscisses (voir la figur en bas), comme étant l'abscisse du point qui a pour ordonnée la moitié de l'effectif total.

2) Méthode analytique:

Du polygone des effectifs cumulés on déduit une méthode géométrique qui amène à la détermination de la médiane (voir la figur en bas):

$$\frac{N_2 - N_1}{X_2 - X_1} = \tan \hat{CDA} = \tan \hat{BSA} = \frac{N_2 - \frac{N}{2}}{X_2 - M}$$

