

I] RAPPEL DU VOCABULAIRE USUEL :

- **Population – Echantillon** : ensemble ou partie d'un ensemble d'individus ou d'unités statistiques dont on observe un ou plusieurs **caractères**. Le nombre de ses éléments est sa **taille** ou son **effectif**.
- **Modalité** : valeur prise par le caractère pour un individu de la population.
- **Caractère qualitatif** : caractère dont les modalités sont **repérables** (couleur, forme, marque etc.)
- **Caractère quantitatif** : caractère dont les modalités sont **mesurables**, les mesures étant les valeurs d'une variable statistique (poids, taille, longueur etc.)
- **Variable discrète** : variable qui ne peut prendre qu'un **nombre limité de valeurs**. A chaque valeur correspond un effectif, la série est pondérée.
- **Variable continue** : variable qui peut prendre n'importe quelle valeur d'un **intervalle**.
- **Classe** : **intervalle** partiel appartenant à l'intervalle d'étude.
- La largeur (**amplitude**) de la classe $[a ; b]$ est la différence $b - a$.
- Le **centre de la classe** est le nombre $\frac{a+b}{2}$ (milieu de l'intervalle)
- **Fréquence** d'une valeur ou d'une classe dans une population : c'est le quotient f_i de l'effectif n_i de cette valeur ou de cette classe par l'effectif total N : $f_i = \frac{n_i}{N}$
- Les statistiques que nous étudierons, seront soit à **une variable**, soit à **deux variables**, à valeur discrète ou continue.

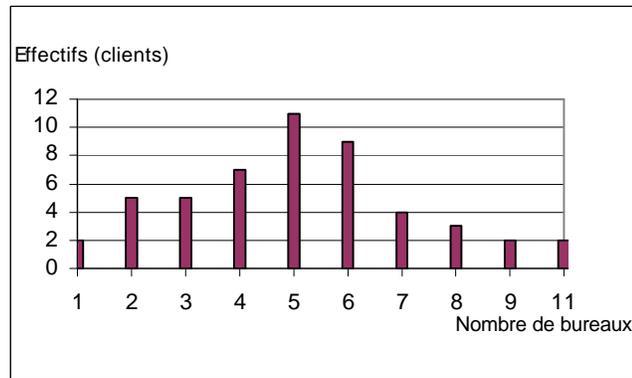
II] SERIES A VALEURS DISCRETES :

1°) Modes de représentation :

On extrait du fichier d'une entreprise d'entretien de bureaux, 50 fiches de clients, ce qui nous constitue une série statistique à variable discrète. Un regroupement par nombre de bureaux fournit alors le **tableau des effectifs et des fréquences** ...

Nombre de bureaux	1	2	3	4	5	6	7	8	9	11
Effectifs (clients)	2	5	5	7	11	9	4	3	2	2
Fréquences	0,04	0,10	0,10	0,14	0,22	0,18	0,08	0,06	0,04	0,04

... permettant de dresser un **histogramme des effectifs**.



2°) Caractéristique de valeur centrale (ou de positions) :

➤ **Mode**

Le mode est la (les) valeur(s) de la variable pour laquelle l'effectif est maximal.

➤ **Médiane**

Les valeurs de la variable étant rangées dans l'ordre croissant, la médiane est alors :

La valeur de rang $q+1$ si l'effectif (impair) de la population est $n = 2q + 1$.

La valeur de rang q si l'effectif (pair) de la population est $n = 2q$.

➤ **Moyenne**

La moyenne de la variable statistique est la moyenne pondérée :

$$\bar{x} = \frac{n_1.x_1+n_2.x_2+\dots+n_p.x_p}{n} = \frac{1}{n} \sum_{i=1}^p x_i.n_i \quad (\text{Avec } n_1+n_2+\dots+n_p = n)$$

Rq : si pour tout i naturel on a : $x_i = a X_i + b$ ($a \neq 0$) on a :

$$\bar{x} = a \bar{X} + b \quad (\text{changement de variable})$$

3°) Caractéristique de dispersion :

Ce sont des valeurs qui caractérisent le plus ou moins grand étalement des termes de la série.

➤ **L'étendue**

L'étendue e d'une série est la différence entre des valeurs extrêmes x_{\min} et x_{\max} :

$$e = x_{\min} - x_{\max}$$

➤ **L'écart moyen**

L'écart moyen e_m d'une série statistique est la moyenne des valeurs absolues des écarts à la moyenne \bar{x} .
$$e_m = \frac{1}{n} \sum_{i=1}^p xi.ni$$

➤ **Variance – Ecart type**

La variance V d'une série statistique est la moyenne des carrés des écarts entre les observations et leur moyenne.

L'écart type s est la racine carrée positive de la variance.

$$v = \frac{\Sigma(ni \times xi^2)}{N} - \bar{x}^2$$
$$\sigma = \sqrt{v}$$

Signification de l'écart type :

L'expérience montre que dans le cas d'une distribution uni modale et symétrique :

- L'intervalle $[\bar{x} - s ; \bar{x} + s]$ contient environ 68 % des valeurs.
- L'intervalle $[\bar{x} - 2s ; \bar{x} + 2s]$ contient environ 95 % des valeurs.

Voir chapitre sur les lois de distribution.

Effet d'un changement de variable affine sur la variance et l'écart type :

Supposons que l'on ait pour tout i : $xi = a Xi + b$ ($a \neq 0$), on a alors :

$$V_x = a^2 V_X \text{ et } s_x = a s_X$$

4°) Calcul pratique de la moyenne de la variance et de l'écart type :

Ces calculs ne sont généralement pas demandés, la plupart des calculatrices permettant de les effectuer.

La moyenne est donnée par les touches : \bar{x} ou mean, l'écart type par la touche s ou s_n . Dans le cas où les calculs seraient demandés, voir exercice résolu 1.

III] SERIES A VARIABLE CONTINUE :

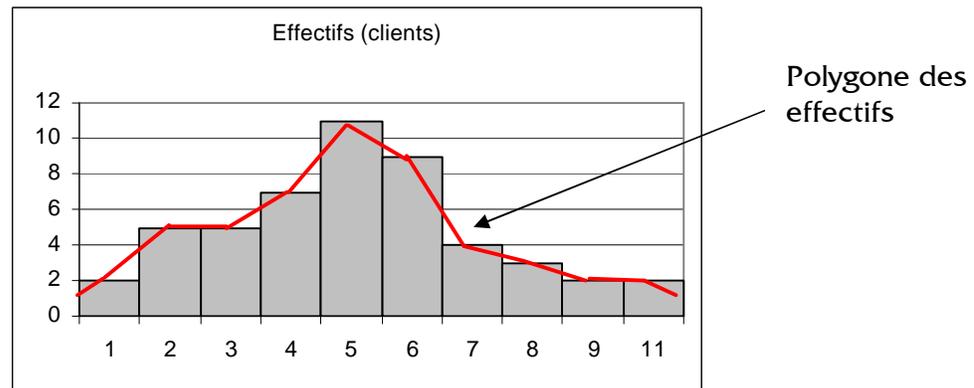
1°) Regroupement en classes :

Une variable continue peut prendre ses n valeurs parmi un effectif total N dans un intervalle de R . On regroupe généralement les valeurs dans des classes adjacentes, en constituant un tableau.

Classes	Centre de classe	Effectif	Fréquence
$[X_0, X_1[$	$x_1 = \frac{X_0+X_1}{2}$	n1	$\frac{n_1}{N}$
$[X_1, X_2[$	$x_2 = \frac{X_1+X_2}{2}$	n2	$\frac{n_2}{N}$
etc.	etc.	etc.	etc.
$[X_{p-1}, X_p[$	$x_p = \frac{X_{p-1}+X_p}{2}$	np	$\frac{np}{N}$

2°) Représentation :

Histogramme - Polygone des effectifs (ou des fréquences)



3°) Effectifs et fréquences cumulés :

Les Effectifs Cumulés Croissants (**ECC**) et les Effectifs Cumulés Décroissants (**ECD**) permettent de déterminer graphiquement la médiane d'une série statistique.

Voir Exercice résolu

4°) Valeurs caractéristiques :

Mode : Il existe une (ou plusieurs) classe modale d'effectif maximal, on prend comme mode le centre de cette classe.

Médiane : il existe une classe médiane, celle qui contient la médiane, valeur de la variable correspondant à 50 % des effectifs cumulés (croissants ou décroissants).

Graphiquement la médiane correspond à l'abscisse du point d'intersection des polynômes des ECC et des ECD.

Quartiles – Déciles :

- Le premier quartile Q_1 est la valeur qui correspond à 25 % des ECC.
- Le deuxième quartile Q_2 est la valeur qui correspond à 50 % des ECC, c'est donc la médiane.

- Le troisième quartile Q_3 est la valeur qui correspond à 75 % des ECC.

Moyenne – Variance – Ecart type :

On se ramène au cas d'une série à variable discrète en prenant pour valeurs les centres de classe (x_i) et en les pondérant par les effectifs de ces classes (produits $x_i.n_i$).

Voir exercice résolu.

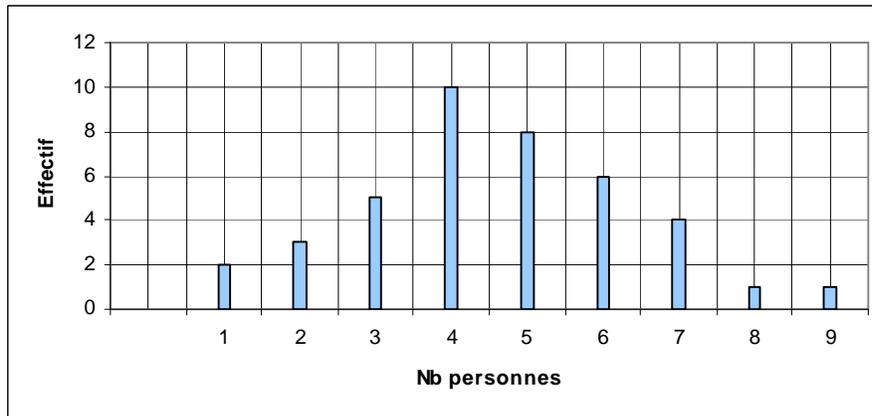
EXERCICE RESOLU – SERIE A VARIABLE DISCRETE

On donne le tableau suivant correspondant au nombre d'employés (n_i) dans 40 entreprises (x_i) de moins de 10 personnes.

x_i	1	2	3	4	5	6	7	8	9
n_i	2	3	5	10	8	6	4	1	1

1. En donner une représentation sous forme d'un diagramme en bâtons, d'un histogramme, d'un diagramme circulaire et d'un diagramme polaire.

Diagramme en bâtons



Histogramme des effectifs

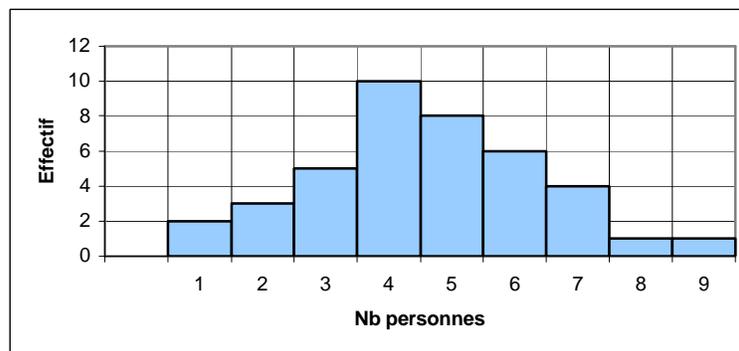


Diagramme circulaire

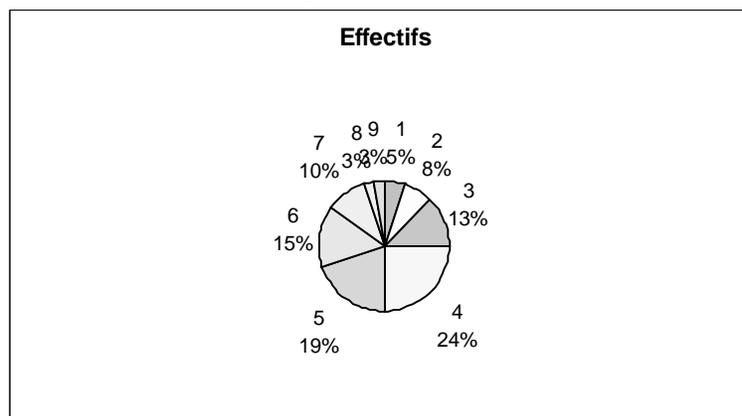
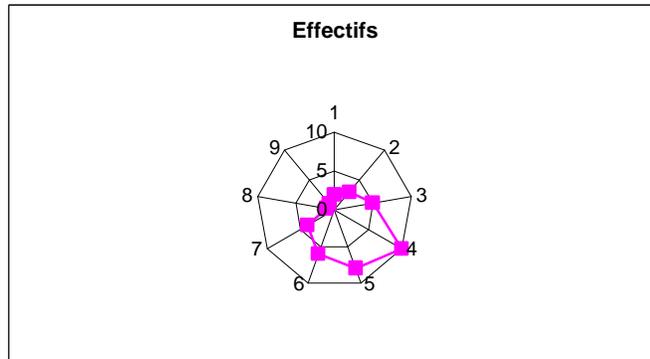
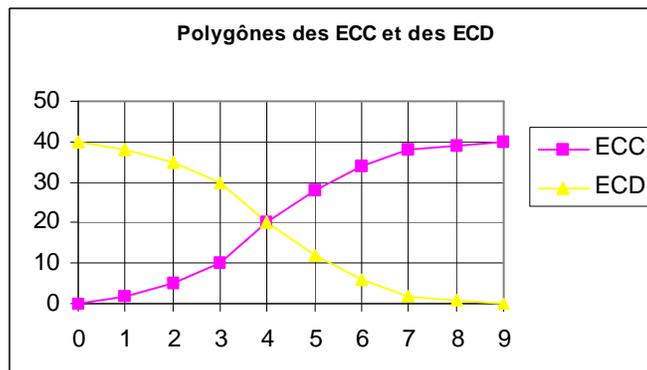


Diagramme polaire



2. Faire un tableau des ECC et des ECD. Déterminer graphiquement la médiane, puis retrouver le résultat par le calcul.

xi	ni	ECC	ECD	xi.ni	ni.xi ²
1	2	2	40	2	-3,6
2	3	5	38	6	-2,6
3	5	10	35	15	-1,6
4	10	20	30	40	-0,6
5	8	28	20	40	0,4
6	6	34	12	36	1,4
7	4	38	6	28	2,4
8	1	39	2	8	3,4
9	1	40	1	9	4,4
Total	40			184	976



Graphiquement la médiane est de 4

Calcul de la médiane:

- Recherche du rang :

$$\text{Effectif total} / 2 = 40 / 2 = \mathbf{20}$$

- Recherche de la position :

La 20^{ème} valeur (à chercher dans la colonne des ECC) est atteinte pour xi = 4.

- Interpolation linéaire :

Pour $x_i = 3$ on a $n_i = 10$.

Pour $x_i = 4$ on a $n_i = 20$.

On a $20 - 10 = 10$ valeurs sur cette plage pour $4 - 3 = 1$.

La médiane est de $3 + 1 \times \frac{10}{10}$

La médiane est de : **4**

3. Calculer la moyenne, la variance et l'écart type.

Moyenne :

$$\bar{x} = \frac{\sum_{i=1}^9 x_i \cdot n_i}{N} = \frac{184}{40}$$

Soit $\bar{x} = 4,6$

Variance :

$$v = \frac{\sum (n_i \times x_i^2)}{N} - \bar{x}^2 = \frac{976}{40} - 4,6^2$$

$$\sigma = \sqrt{v}$$

Soit $V = 3,24$

Ecart type :

$$s = \sqrt{V}$$

Soit $s = 1,8$



Entraînez vous ... et retrouver ces résultats en utilisant votre calculatrice



TD 1 – APPLICATION DU COURS STATISTIQUES A UNE VARIABLE



Exercice 1

Une usine fabrique des moteurs Diesel. On appelle X le dépassement (exprimé en mm) du dessus du piston par rapport au plan de joint du bloc cylindre. Un étudiant en stage dans cette usine est chargé de tester 50 moteurs. Il obtient les résultats suivants regroupés dans des classes de largeur 2 mm.

Centre de classe X_i	Effectifs n_i	F (%)	ECC	ECD	$x_i.n_i$	$n_i.x_i^2$
0,75	1					
0,77	2					
0,79	2					
0,81	6					
0,83	12					
0,85	16					
0,87	7					
0,89	3					
0,91	1					
Total						

1. Compléter ce tableau.
2. Déterminer la moyenne, la variance, la médiane et l'écart type.
3. Représenter cette série par un histogramme.
4. Représenter le polygone des ECC et des ECD puis retrouver graphiquement la valeur de la médiane

Exercice 2

Une machine produit en grand nombre des axes de moteurs électriques. Une étude portant sur la longueur (en mm) de 100 axes a donné les résultats suivants :

Classes	[89.7; 89.8[[89.8;89.9[[89.9;90[[90 ;90.1[[90.1 ;90.2[[90.2 ;90.3[
Effectifs	3	14	36	33	13	1

1. Tracer l'histogramme et le polygone des effectifs.
2. Calculer la médiane et les deux quartiles Q_1 et Q_3 .
3. Calculer la moyenne et l'écart type.
4. Calculer le pourcentage des axes dont la longueur est comprise dans l'intervalle $[\bar{x} - s ; \bar{x} + s]$.

Eléments de réponses :

2. Médiane : 89,992
3. $\bar{x} = 89,992$ et $s = 0,101$
4. 67,95 %