UNIVERSITE DE BATNA_2 **FACULTE DE MEDECINE**

SERIE DE TD N° 1 *en BIOSTATISTIQUE 2020/2021*

(1ère ANNEE, 2020_2021)

Rappel de cours :

Définir les notions suivantes en donnant des exemples : Population statistique, individu statistique, caractère étudié, modalité, variable discrète et continue.

- Population statistique : toute étude statistique porte sur un ensemble appelé **population**. Les éléments de cette population sont les **individus** ou unités statistiques.
- Caractère ou variable : au cours d'une étude statistique, on s'intéresse à un aspect des individus de la population appelé variable ou **caractère**. La variable prend différentes valeurs qu'on appelle **modalités**.
- Variable discrète ou continue : si on peut mesurer cet aspect et y faire des opérations et calculs, la variable est dite de nature quantitative. Si cette variable quantitative ne prend que des valeurs dites isolées, elle est **discrète**. Si elle prend n'importe valeur d'un intervalle, elle est **continue**.

Exercice n° 01 :

Le staff médical d'une grande entreprise fait ses petites statistiques sur la pratique du sport par mois de ses employés ; des observations sur 88 employés tirés au sort sont les suivantes :

x_i = nombre de séances par mois	8	12	16	20	24	28
n_i = effectifs partiels	7	20	23	19	14	5

1) Donner la population, le caractère, la nature du caractère et son type.
2) Représenter graphiquement la série statistique.
3) Calculer le mode (M_o), la médiane (Me) et l'écart interquartile (I_Q).
4) Calculer la moyenne, la variance, l'écart type et le coefficient de variation.

Solution

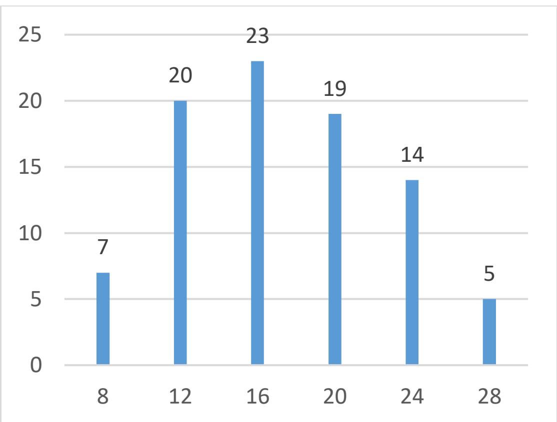
1°) La population : Les employés d'une grande entreprise.

Le caractère : Le nombre de séances de sport pratiquées par mois.

La nature du caractère : une variable statistique quantitative.

Le type du caractère : une quantité discrète.

2°) Représentation graphique



3°) Calcul du mode (M_o), de la médiane (Me) et de l'écart interquartile (I_Q) :

Le mode Mo : 16 (la plus fréquente variable).

La médiane Me : 16

Le premier quartile Q_1 : 12

Le troisième quartile Q_3 : 20

L'écart interquartile I_Q : 8

4°) Calcul de la moyenne, de la variance, de l'écart type et du coefficient de variation.

x_i	8	12	16	20	24	28	**Total**
n_i	7	20	23	19	14	5	**88**
$n_i.x_i$	56	240	368	380	336	140	**1520**
$n_i.x_i^2$	448	2880	5888	7600	8064	3920	**28800**
$n_{icum}\uparrow$	7	27	50	69	83	88	

La moyenne : $\bar{x} = \frac{\sum n_i * x_i}{\sum n_i} = \frac{1520}{88} = 17.2727273 \approx 17.273$

La variance : $\sigma^2 = \frac{\sum n_i . x_i^2}{\sum n_i} - \bar{x}^2 = \frac{28800}{88} - \left(\frac{1520}{88}\right)^2 = 28.92561983 \approx 28.926$

Ecart-type : $\sigma = \sqrt{28.9256} = 5.37825 \approx 5.378$

Le coefficient de variation : $CV = \frac{\sigma}{\bar{x}} = 0.31137 \approx 0.311$

Exercice n° 02 :

La répartition d'un groupe de 20 étudiants classés par degré de lecture est la suivante :

x_i = degré de lecture	Peu	Moyen	Beaucoup	Exceptionnel
n_i = nombre d'étudiants	3	5	10	2

1) Quelle est la population étudiée ; le caractère étudié ainsi que sa nature ?
2) Représenter cette série par un graphe adéquat.
3) Si le caractère est mesuré par le nombre de livres lus, comment se présenterait le tableau statistique.

Solution

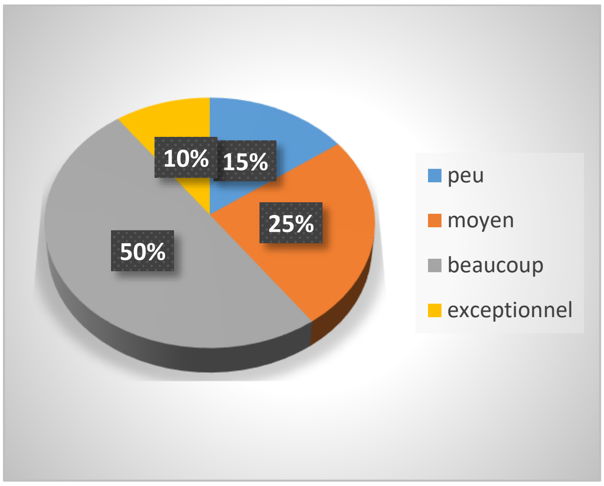
1°) La population étudiée : Le groupe des vingt étudiants.

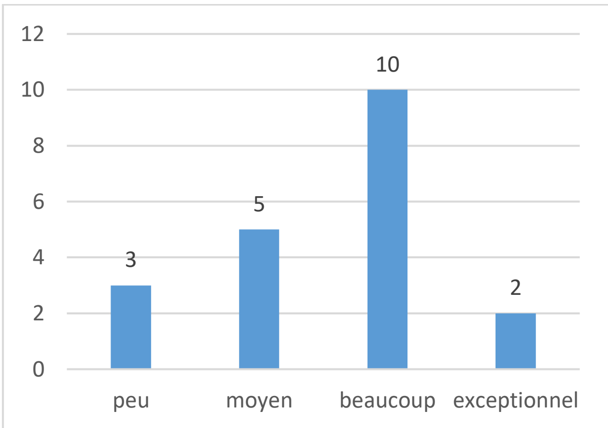
Le caractère : Le degré de la pratique de lecture.

La nature du caractère : une variable qualitative.

Le type du caractère : une qualité ordinale.

2°) Représentation de cette série par un graphe adéquat (circulaire ou en barres) :





3°) Si le caractère est mesuré par le nombre de livres lus, les variables statistiques Xi seraient quatre intervalles qui sont par exemple : moins de 4 livres ; entre 4 et 10 livres ; entre 10 et 20 livres ; plus de 20 livres. La variable statistique x_i serait alors une quantité continue dans ce cas.

Si le nombre de livres lus est donné par des modalités x_i par exemple : 2 livres, 7 livres, 12 livres et 20 livres on dirait alors dans ce cas que la variable statistique x_i est une quantité discrète.

Exercice n° 3 :

Voici les résultats R obtenues dans un exercice par des étudiants lors d'un examen :

14	10.1	17.3	14.8	16	9	12.3	7.9	7	15	6	19	6.3
10.7	5	8.4	7	12	9.6	2.4	13	10.6	17	15	8	3.1
10.5	11	18	3.5	12	9.4	3.4	13.2	11	14	14	5	6
11	11	12	16	8	4							

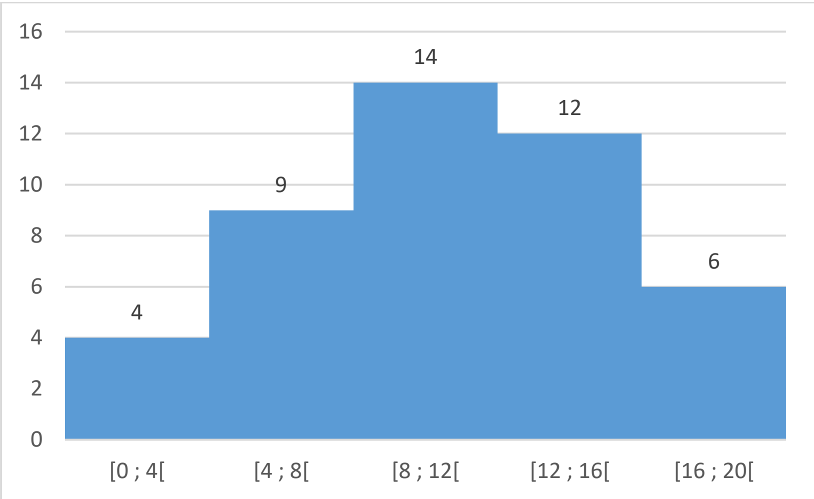
Compléter le tableau suivant :

x_i = classe de notes	[0 ; 4[	[4 ; 8[	[8 ; 12[	[12 ; 16[	[16 ; 20]	Total
n_i = effectifs partiels	4	9	14	12	6	45

1) Représenter graphiquement la série statistique.
2) Calculer le mode (Mo) et la médiane (Me) par interpolation linéaire.
3) Calculer la moyenne, la variance, l'écart type et le coefficient de variation.

Solution

1°) Représentation graphique de cette série



2°) Calcul du mode (Mo) par interpolation linéaire : la classe modale est [8 ; 12[

Le Mode : Mo = $x_{i-1} + k \frac{n_i - n_{i-1}}{2n_i - (n_{i+1} + n_{i-1})} = 8 + 4 \frac{14-9}{2*14-(12+9)} = 10.857$

Calcul de la médiane (Me) par interpolation linéaire : la classe médiane est [8 ; 12[

La médiane : Me $= x_{i-1} + k \frac{\frac{N}{2} - n_{(i-1)cum}}{n_{icum} - n_{(i-1)cum}} = 8 + 4 \frac{\frac{45}{2} - 13}{27 - 13} = 10.714$

<u>Interpolation de Q_1 et Q_3</u> : $Q_1 = x_{i-1} + k.\frac{\frac{N}{4} - n_{(i-1)cum}}{n_{icum} - n_{(i-1)cum}}$ et $Q_3 = x_{i-1} + k.\frac{\frac{3N}{4} - n_{(i-1)cum}}{n_{icum} - n_{(i-1)cum}}$

3°) Calculer la moyenne, la variance, l'écart type et le coefficient de variation.

Classe de R	$0 \leq r < 4$	$4 \leq r < 8$	$8 \leq r < 12$	$12 \leq r < 16$	$16 \leq r < 20$	Σ
Centre c_i	2	6	10	14	18	
n_i	4	9	14	12	6	45
$n_i c_i$	8	54	140	168	108	478
$n_i c_i^2$	16	324	1400	2352	1944	6036
n_{icum} ↑	4	13	27	39	45	
f_i	0.089	0.2	0.311	0.267	0.133	1
$f_i c_i$	0.178	1.2	3.11	3.738	2.394	$\bar{x}$ = 10.62

La moyenne $\bar{x} = \frac{\sum n_i * c_i}{\sum n_i} = \frac{478}{45} = 10.622\ldots$ $\quad \sigma^2 = \frac{\sum n_i.c_i^2}{\sum n_i} - \bar{x}^2 = \frac{6036}{45} - \left(\frac{478}{45}\right)^2 = 21.3017284$

$\sigma = \sqrt{21.30173} = 4.615$ Le coefficient de variation $CV = \frac{\sigma}{\bar{x}} = \frac{4.61537955}{10.6222222} = 0.434502259$

Exercice n° 4 :

Dans un centre de renseignements téléphoniques, une enquête est effectuée sur un échantillon de 320 clients, afin de diminuer le temps d'attente subi par la clientèle. Les résultats sont consignés dans le tableau suivant :

Temps minutes	[0 ;5[	[5 ;10[	[10 ;15[	[15 ;20[	[20 ;25[	[25 ;30[	[30 ;35]	Total
Nb clients n_i	32	56	74	78	36	30	14	320
$n_{i\ cum}$ croissant	32	88	162	240	276	306	320	
Centre de classe c_i	2.5	7.5	12.5	17.5	22.5	27.5	32.5	
$n_{i\ cum}$ décroissant	320	288	232	158	80	44	14	
$n_i * c_i$	80	420	925	1365	810	825	455	4880

1) Quelle est la population étudiée ?
2) Quel est le caractère étudié ?
3) Compléter la ligne des effectifs cumulés croissants du tableau précédent.
4) Déterminer la classe modale de cette série ainsi que le mode par interpolation linéaire.
5) Quelle est l´étendue de cette série ?
6) Calculer le temps d'attente moyen.
7) Construire l'histogramme de cette série et calculer la valeur de la variance.
8) Quel est le pourcentage de clients qui attendent au moins 20 secondes ?
9) Quel est le pourcentage de clients qui attendent moins de 10 secondes ?

Solution

1°) La population : les clients du centre de renseignements téléphoniques.

2°) Le caractère c'est les différents temps d'attente en secondes dans ce centre ; la nature de ce caractère est une variable statistique quantitative continue.

3°) Voir tableau

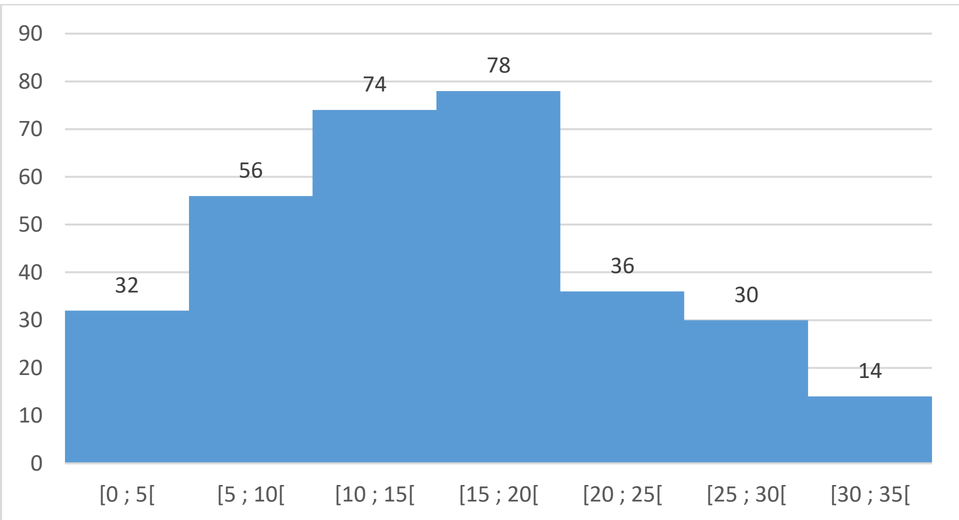
4°) La classe modale = [15 ; 20[: classe dont l'effectif est le plus grand.

Le mode : $Mo = x_{i-1} + k \frac{n_i - n_{i-1}}{2n_i - (n_{i+1} + n_{i-1})} = 15 + 5 \frac{78 - 74}{2 * 78 - (74 + 36)} = 15.434782608 \approx 15.435$

5°) L'étendue de la série : $E = x_{max} - x_{min} = 35 - 0 = 35$

6°) Le temps d'attente moyen : $\bar{x} = \frac{\sum n_i * c_i}{\sum n_i} = \frac{4880}{320} = 15.25$

7°) L'histogramme est le suivant :



8°) Le pourcentage des clients attendant au moins 20 secondes (ç-à-d ≥ 20) est : $\frac{80}{320} = 25\ \%$

9°) Le pourcentage des clients attendant moins de 10 secondes (ç-à-d < 10) est : $\frac{88}{320} = 27.5\ \%$