Faculté de médecine
1ère année médecine

Corrigé-type de TD N° 1 de Biostatistique-informatique

Exercice 01 :

Variable	Type	Nature	Représentation graphique
Les notes sur 20 des étudiants inscrits en 1ère année médecine	Quantitative	Discrète	Diagramme en bâtons
Cholestérol	Quantitative	Continue	Histogramme
Statut vaccinal	Qualitative	Binaire	Diagramme en bandes
Caractère génétique	Qualitative	Binaire	Diagramme en bandes
Parité	Quantitative	Discrète	Diagramme en bâtons
Groupe sanguin	Qualitative	Nominale	Diagramme en bandes
Stade du cancer de l'estomac	Qualitative	Ordinale	Camembert

Exercice 02 :

1) **Population** : 20 étudiants

   Caractère : Degré de la pratique de lecture.

   Nature : qualitatif **Type :** ordinal

2) **Représentation adéquate :**

Diagramme circulaire

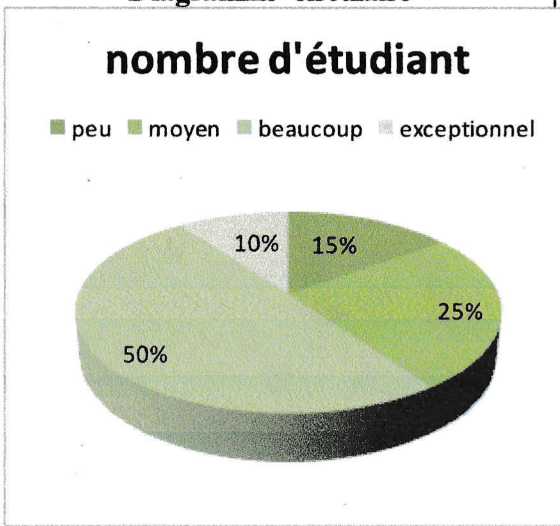
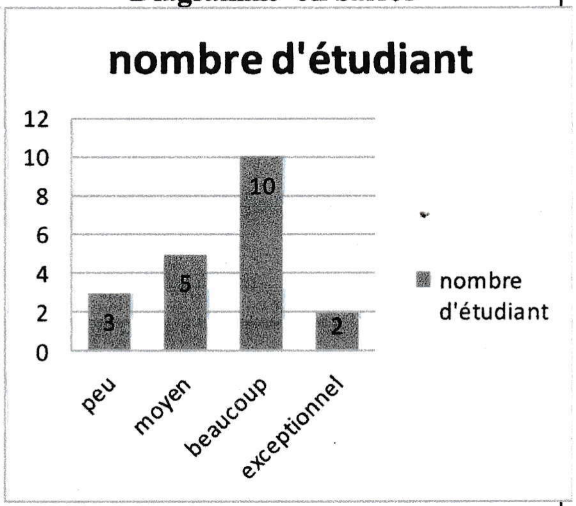


Diagramme en barres



3) Si le caractère présente le nombre de livres lus, dons il s'agit dans ce cas d'un caractère quantitatif.
   - ✓ Soit par des nombres isolés : **Exemple 2, 4, 6,10…**

   Caractère quantitatif discret.
   - ✓ Soit par des classes : **Exemple moins de 2 livres, entre 2 et 5, entre 5 et 10, plus de 10…**

     Caractère quantitatif continu.

Exercice 05 :

Population : 200 lots

Variable : Nombre de comprimés défectueux, quantitative discrète

Représentation : diagramme en bâtons.

Caractéristiques de position :

La moyenne : $\bar{x} = \frac{\sum n_i x_i}{N} = 1.205$

La médiane :

Tout d'abord on calcule les effectifs cumulés :

Nombre de comprimés défectueux par lots	0	1	2	3	4	5
Nombre de lots ni	75	53	39	23	9	1
Effectifs cumulés $n_i \uparrow$	75	128	167	190	199	200

Série paire

$$N\ pair$$

$$M_0 = \frac{\left(\frac{N}{2}\right)^{ième}\ obsevation + \left(\frac{N}{2}+1\right)^{ième}\ obsevation}{2}$$

=1

2ème quartile = médiane=1

Le mode :

L'effectif le plus élevé c'est 75, alors le mode = 0.

Les caractéristiques de dispersion :

La variance =1.473. Ecart-type =1.214. Etendue = 5 - 0 =5

1er quartile

N/4=50, donc le 1er quartile = $(50)^{\text{ième}}$ $obsevation = 0$

3ème quartile

3N/4=150, donc le 3ème quartile = $(150)^{\text{ième}}$ $obsevation = 2$

Coefficient de variation :$CV = \frac{\sigma}{\bar{x}}$

Exercice 03 :

Voici les 72 résultats d'un examen de Bio-statistique :

Données brutes

12 8 15 11 4 7 13 2 9 10 17 13 14 3 6 8 12 9 16 16 12 9 4 15

0 3 13 2 18 5 6 11 10 14 6 8 17 10 14 11 16 10 8 10 9 11 10 14

7 13 19 14 10 15 12 13 6 12 11 9 13 16 15 13 5 10 7 16 10 8 16 11

- **Variable quantitative discrète :**

Notes	0	2	3	4	5	6	7	8	9	10
Effectifs	1	2	2	2	2	4	3	5	5	9
Effectifs cumulés	1	3	5	7	9	13	16	21	26	35
Notes	11	12	13	14	15	16	17	18	19	
Effectifs	6	5	7	5	4	6	2	1	1	
Effectifs cumulés	41	46	53	58	62	68	70	71	72	

- **Variable quantitative continue :**

Classes	Centres	Effectifs	Effectif s cumulés
[0 ; 4[	2	5	5
[4 ; 8[	6	11	16
[8 ; 12[	10	25	41
[12 ;16[	14	21	62
[16 ; 20[	18	10	72

1ère partie :

La moyenne

=10.58 La médiane

En 1er lieu on calcule les effectifs cumulés, on

obtient N=72 Série paire

$$M_0 = \frac{(N/2)^{i\grave{e}me} \ obsevation + (\frac{N}{2}+1)^{i\grave{e}me} obsevation}{2}$$

$$M_0 = \frac{(36)^{i\grave{e}me} \ obsevation + (37)^{i\grave{e}me} obsevation}{2}$$

$$\frac{11+11}{2} = 11$$

Le mode = 10 (l'effectif le plus élevé est 9)

Les quartiles

1er quartile

N/4=18, donc le 1er quartile $= (18)^{i\grave{e}me}\ obsevation = 8$

3ème quartile

3N/4=54, donc le 3ème quartile $= (54)^{i\grave{e}me}\ obsevation = 14$

L'interquartile= 14-8 =6. **La variance** = 17.60.

L'écart-type = 4.195

Coefficient de variation : $CV \ \frac{\sigma}{\bar{x}} = \frac{4.195}{10.58} = 0.3965 \ = (40\%) =$
=

Le mode :

L'effectif le plus élevé c'est 75, alors le mode = 0.

Les caractéristiques de dispersion :

La variance =1.473. Ecart-type =1.214. Etendue = 5 - 0 =5

1er quartile

N/4=50, donc le 1er quartile $= (50)^{\text{ième}}\ obsevation = 0$

3ème quartile

3N/4=150, donc le 3ème quartile $= (150)^{\text{ième}}\ obsevation = 2$

Coefficient de variation : $CV = \frac{\sigma}{\bar{x}}$

Exercice 03 :

Voici les 72 résultats d'un examen de Bio-statistique :

Données brutes

12 8 15 11 4 7 13 2 9 10 17 13 14 3 6 8 12 9 16 16 12 9 4 15

0 3 13 2 18 5 6 11 10 14 6 8 17 10 14 11 16 10 8 10 9 11 10 14

7 13 19 14 10 15 12 13 6 12 11 9 13 16 15 13 5 10 7 16 10 8 16 11

- **Variable quantitative discrète :**

Notes	0	2	3	4	5	6	7	8	9	10
Effectifs	1	2	2	2	2	4	3	5	5	9
Effectifs cumulés	1	3	5	7	9	13	16	21	26	35
Notes	11	12	13	14	15	16	17	18	19	
Effectifs	6	5	7	5	4	6	2	1	1	
Effectifs cumulés	41	46	53	58	62	68	70	71	72	

- **Variable quantitative continue :**

Suite de corrigé-type de TD 1

Exercice 04 :

A partie des centres de classes, vous identifiez les bornes de classes :

a. $L'amplitude : h = C_{i+1} - C_i = 30 - 20 = 10$

b. $x_i = C_i + \frac{h}{2}$ et $x_{i-1} = C_i - \frac{h}{2}$

On obtient :

Centres C_i	Classes $x_{i-1} ; x_i$	Effectifs	Effectifs cumulés Croissants $n_{i\uparrow}$	Effectifs cumulée décroissants $n_{i\downarrow}$
20	15-25	10	10	200
30	25-35	15	25	190
40	35-45	35	60	175
50	45-55	50	110	140
60	55-65	50	160	90
70	65-75	40	200	40

En utilisant la calculatrice, on obtient :

- ✓ La moyenne : $m = \frac{\sum_{i=1}^{6} x_i n_i}{200} = 51.75$
- ✓ La variance : Var=194.44
- ✓ L'étendue : $E = x_p - x_0 = 75 - 15 = 60$
- ✓ Cv=0.27

L'effectif le plus élevé est 50, donc deux classes modales c'est-à-dire série bimodale

Mais :

1^er mode : *La classe modale :* $M_{01} \in [45 - 55[$

$$M_{01} = 45 + 10\frac{50 - 35}{2 \times 50 - (35 + 40)} = 55$$

2^ème mode : *La classe modale :* $M_{02} \in [55 - 65[$

$$M_{o2} = 55 + 10\frac{50 - 50}{2 \times 50 - (50 + 40)} = 55$$

$$M_{01} = M_{o2}$$

Pourquoi ? Parce que les deux classes modales sont ***adjacentes***

La médiane : *N =200 paire, la classe médiane :* $M_e \in [45 - 55[$

$$M_e = x_{i-1} + h\frac{\frac{N}{2} - n_{i-1\uparrow}}{n_{i\uparrow} - n_{i-1\uparrow}} 45 + 10\frac{100-60}{110-60} = 53$$

Le 1^er^ quartile : *N/4 =50* $Q_1 \in [35 - 45[$

$$Q_1 = x_{i-1} + h\frac{\frac{N}{4} - n_{i-1\uparrow}}{n_{i\uparrow} - n_{i-1\uparrow}} = 35 + 10\frac{50-25}{60-25} = 42.14$$

Le 3^ème^ quartile : *3N/4 =150* $Q_3 \in [55 - 65[$

$$Q_3 = x_{i-1} + h\frac{\frac{3N}{4} - n_{i-1\uparrow}}{n_{i\uparrow} - n_{i-1\uparrow}} = 55 + 10\frac{150-110}{160-110} = 63$$

6) une 2^ème^ série statistique ayant un coefficient de variation égale à 0.72, oui c'est possible d'utiliser ce coefficient pour comparer la dispersion par ce qu'il s'exprime sans unité de mesure.

Exercice 05 :

Modalités	***0***	***1***	***2***	***3***	***4***	***5***
Effectifs	***75***	***53***	***39***	***23***	***9***	***1***
Effectifs cumulés	***75***	***128***	***167***	***190***	***199***	***200***

1) **Population** : 200 lots c'est-à-dire 200 unités

Caractère : Nombre de comprimés défectueux

Nature et type: quantité discrète.

2) **Représentation graphique** : Diagramme en bâtons.
3) *Le mode est la modalité la plus fréquente* $M_o=0$

Le premier quartile Q_1 = 50^ième^ observation =0

La médiane Me = 100^ième^ observation = 1
Le troisième quartile Q_3 = 150^ième^ observation =2
La moyenne = 1.205
L'écart-type = 1.106
L'étendue : $E = x_{max} - x_{min} = 5 - 0 = 5$
Le coefficient de variation : CV =0.92

Exercice 06 :

1) *Non pas toujours*

Exemple :

Modalités	***1***	***2***	***3***	***4***	***5***
Effectifs	***7***	***8***	***10***	***14***	***11***
Effectifs cumulés	***7***	***15***	***25***	***39***	***50***

N paire

$$M_e = \frac{observation\ d'ordre\frac{N}{2} + observation\ d'ordre\ (N+1)/2}{2} = \frac{3+4}{2} = 3.5$$

Mais, pour le deuxième exemple : dans le cas pair la médiane est parmi les observations (parmi les modalités)

Changement dans les effectifs:

Modalités	***1***	***2***	***3***	***4***	***5***
Effectifs	***7***	***8***	***11***	***13***	***11***
Effectifs cumulés	***7***	***15***	***26***	***39***	***50***

$$M_e = \frac{observation\ d'ordre\frac{N}{2} + observation\ d'ordre\ (N+1)/2}{2} = \frac{3+3}{2} = 3$$

2) *Oui*

3) *Oui*

4) *Non, seulement le coefficient de variation*