

Solution de TD N° 1 de Biostatistique-informatique 2021/2022

Exercice 01 :

Pour chacune des variables ci-dessous, indiquer leur nature, préciser les représentations graphiques adaptés.

Variable	Type	Nature	Représentation graphique
Les notes sur 20 des étudiants inscrits en 1 ^{ère} 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 :

Une crème dermatologique est testée sur un échantillon de 150 femmes également réparties en trois groupes en fonction de leur type de peau : mate, normale ou claire. On s'intéresse à l'éventuelle réaction allergique à la crème dermatologique en fonction du type de peau.

Modalités	Mate	Normale	Claire
Effectifs	3	7	13

- 1) Identifier la population le caractère, le type la nature la représentation convenable
- 2) Quelle est la proportion des femmes de cet échantillon ayant développé une allergie ?
- 3) Calculer cette proportion en fonction du type de peau ?
- 4) Donnez une représentation graphique en camembert des réactions allergisantes en fonction du type de peau ?

Solution :

Modalités	Mate	Normale	Claire
Nombre de femmes	50	50	50

En premier lieu

Echantillon : 150 femmes

Caractère : type de la peau

Caractère : qualitatif nominal

Représentation adéquate : camembert ou diagramme en bandes

En deuxième lieu

Type de peau	Mate	Normale	Claire
Nombre d'allergies	3	7	13

Echantillon : 23 femmes ayant développé une allergie.

Caractère : type de la peau

Caractère : qualitatif nominal

Représentation adéquate : Camembert ou diagramme en bandes

Nombre total de femmes :150

Nombre total d'allergies :23

Proportion des femmes ayant développé une allergie = $23/150 = 0.15333$ (15.33%)

Type de peau	Nbre d'allergies	Nbre de non-allergies	Total	Fréquences relatives
Mate	3	47	50	3/50=6%
Normale	7	43	50	7/50=14%
Claire	13	37	50	13/50=26%
Total	23	127	150	

- 6% des femmes ayant la peau mate ont développé une allergie à cette crème.
- 14% des femmes ayant la peau normale ont développé une allergie à cette crème.
- 26% des femmes ayant la peau claire ont développé une allergie à cette crème.

3)

Type de peau	Mate	Normale	Claire
Nombre d'allergies	3	7	13
Fréquences relatives	3/23=0,13	7/23=30%	13/23=57%
Angles	47°	110°	203°

Exercice 03 :

Lors d'un contrôle d'une chaîne de médicaments, on s'intéresse au nombre de comprimés défectueux dans un lot. L'étude de 200 lots a donné les résultats suivants :

Nombre de comprimés défectueux par lots	0	1	2	3	4	5
Nombre de lots	75	53	39	23	9	1

- 1) Quelle est la population, la variable et son type et la représentation adéquate ?
- 2) Calculer les paramètres de tendance centrale ?
- 3) Calculer la variance, l'écart-type, l'étendue, l'intervalle interquartile et le coefficient de variation du nombre de comprimés défectueux pour ces 200 lots.

Solution :

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 n_i	75	53	39	23	9	1
Effectifs cumulés $n_i \uparrow$	75	128	167	190	199	200

Série paire :

N pair

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

$$M_0 = \frac{(100)^{\text{ième}} \text{ observation} + (101)^{\text{ième}} \text{ observation}}{2}$$

$$\frac{1+1}{2} = 1$$

2^{ème} 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

1^{er} quartile

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

3^{ème} quartile

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

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

Exercice 04 :

Voici les 72 résultats d'un examen de Biostatistique :

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 :**

Remplir le tableau suivant :

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 :**

Remplir le tableau suivant :

Classes	Centres	Effectifs	Effectifs 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

Pour chacune des variables ci-dessus, indiquer leur nature, les représentations graphiques adaptés et calculer les paramètres statistiques.

Solution :

1^{ère} partie :

La moyenne = 10.58

La médiane

En 1^{er} lieu on calcule les effectifs cumulés, on obtient

N=72 Série paire

$$M_0 = \frac{(N/2)^{\text{ième observation}} + (\frac{N}{2} + 1)^{\text{ième observation}}}{2}$$

$$M_0 = \frac{(36)^{\text{ième observation}} + (37)^{\text{ième observation}}}{2}$$

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

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

Les quartiles

1^{er} quartile

N/4=18, donc le 1^{er} quartile = (18)^{ième} observation = 8

3^{ème} quartile

3N/4=54, donc le 3^{ème} quartile = (54)^{ième} observation = 14

L'interquartile= 14-8 =6

La variance=17.60

L'écart-type=4.195

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

2^{ème} partie :

Les paramètres de tendance centrale :

La moyenne = 11,11

La classe modale : [8 ; 12[

$$M_0 = x_{i-1} + h \frac{n_i - n_{i-1}}{2n_i - (n_{i+1} + n_{i-1})} = 8 + 4 \frac{25 - 11}{50 - (21 + 11)} = 11.11$$

$N/2=36$, la classe médiane : [8 ; 12[

$$M_e = x_{i-1} + h \frac{N/2 - n_{i-1} \uparrow}{n_{i\uparrow} - n_{i-1} \uparrow} = 8 + 4 \frac{36 - 16}{41 - 16} = 11.2$$

Paramètres de dispersion :

La variance= 19.21

L'écart-type= 4.38

Les quartiles :

$$\frac{N}{4} = 18 \quad Q_1 \in [8; 12[$$

$$Q_1 = x_{i-1} + h \frac{N/4 - n_{i-1} \uparrow}{n_{i\uparrow} - n_{i-1} \uparrow} = 8.32$$

$$\frac{3N}{4} = 54 \quad Q_3 \in [12; 16[$$

$$Q_3 = x_{i-1} + h \frac{3N/4 - n_{i-1} \uparrow}{n_{i\uparrow} - n_{i-1} \uparrow} = 14.48$$

L'interquartile= 14.48 - 8.32= 6.16

$E=20-0=20$

Coefficient de variation : $CV = \frac{\sigma}{\bar{x}} = 0.39$