

TD 1 : Statistique Descriptive Univariée et Généralités

L2 Info HLMA303 : Statistique descriptive et probabilités

Exercice 1 : Pour chacune des variables suivantes, spécifier son type (i.e. si elle est quantitative, qualitative, continue, discrète, nominale, ordinale). Proposer en ensemble de modalités adéquates pour chacune d'elle.

1. Nombre d'enfants d'une famille
2. Couleur des yeux
3. Catégorie socio-professionnelle
4. Département de naissance
5. Niveau de scolarité
6. Revenu
7. Poids
8. Sexe
9. Age
10. Langue maternelle
11. Type de voiture
12. Taille
13. Nombre de grains de beauté
14. Taille de grains de beauté
15. Type de tumeur

Exercice 2 : On a mesuré sur un échantillon de $n = 100$ jours, la vitesse du courant dans un canal (en cm/s). Les observations sont les suivantes :

4 11 16 17 24 25 25 27 27 28 29 29 31 31 32 32 33 34 34 34 34 35 35
35 36 38 42 42 43 43 46 46 47 48 48 49 49 49 50 50 50 50 51 51 51 52
53 54 54 54 54 55 55 55 56 56 56 57 57 57 57 57 57 58 58 58 59 59 60
62 62 62 62 63 63 63 64 66 66 69 70 72 73 74 76 77 77 78 82 82 83 84
84 84 86 87 88 92 94 99

1. Décrire la population, les individus, la variable quantitative d'intérêt.
2. Construire le tableau de la Distribution Observée (DO) des effectifs pour 10 classes équiréparties : $[0, 10[$; $[10, 20[$; ... ; $[90, 100[$.
3. Représenter la DO des fréquences à l'aide d'un histogramme. Quelle est l'aire de l'histogramme si l'on choisit de représenter les fréquences en ordonnées ? Quelle modification faut-il apporter au graphique pour obtenir un histogramme d'aire égale à 1 ? Cela change-t-il l'information qualitative apportée par ce graphique ?
4. Représenter les DO des fréquences cumulées croissantes et décroissantes sur le même graphique.
5. Calculer la vitesse moyenne m_e à partir des observations brutes, puis à partir des observations groupées en classes, commenter.

6. Calculer la vitesse médiane observée Med à partir des observations brutes, puis à partir des observations groupées en classes, comparer les deux méthodes de calcul.
7. Donner la classe modale et la valeur du mode observé. Comparer les valeurs observées de la moyenne m_e , de la médiane Med et du mode de la vitesse du courant. Ce résultat était-il prévisible à l'examen de l'histogramme?
8. Calculer les 3 premiers quartiles en utilisant la DO des fréquences cumulées. Construire le boxplot ou boîte à moustaches de cette DO.
9. Calculer l'étendue, l'écart-type (en utilisant les données regroupées en classes), et l'écart inter-quartile.

Exercice 3 : L'étude du taux de cholestérol sur un échantillon de 100 personnes a donné les résultats suivants :

Taux de cholestérol (gr/l)	Effectifs n_k
[1,1.4[6
[1.4,1.6[13
[1.6,1.8[16
[1.8,2[22
[2,2.2[18
[2.2,2.4[10
[2.4,2.6[6
[2.6,2.8[4
[2.8,3[3
[3,3.4[2

1. Tracer l'histogramme des fréquences
2. Tracer la courbe des fréquences cumulées croissantes et décroissantes
3. Déterminer le mode et la moyenne de cette distribution
4. Déterminer la médiane, le 1er et le 3ème quartile graphiquement et par le calcul.

Exercice 4 : Les pesées de 50 nouveaux-nés d'une maternité ont permis d'établir le tableau suivant :

Classe (kg)	Effectifs n_k
[2,2.4[6
[2.4,2.8[10
[2.8,3.4[18
[3.4,3.8[12
[3.8,4.2[4

On suppose que dans chaque classe, les poids des nouveaux-nés sont uniformément répartis.

1. Déterminer le mode de cette distribution
2. Calculer le 1er, 2ème et 3ème quartile
3. Retrouver ces résultats à partir d'un digramme des fréquences cumulées
4. Déterminer le quatrième centile.

Exercice 5 : Des notes obtenues par 90 candidats lors d'un test d'effort sont regroupées dans le tableau suivant où n_i désigne le nombre de candidats :

Note	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
n_i	2	0	0	4	0	3	4	10	6	5	8	13	10	7	8	5	2	0	3

1. Tracer le diagramme en bâtons des effectifs et celui des fréquences.
2. Tracer la courbe des fréquences cumulées
3. Déterminer le mode et la moyenne de cette distribution.