

Estimation

Instructions:

Chaque dia contient une question. Réfléchissez à la réponse qui vous semble juste, ensuite cliquez pour que la réponse apparaisse.

Rappel: l'examen se fera sous forme de questions à choix multiple (QCM)

On demande à 25 étudiants de 1^{ère} année de médecine s'ils sont satisfaits du cours de statistique. Les réponses possibles à la question étaient: « Pas satisfait », « Moyennement satisfait » et « Satisfait ». 4 étudiants ont répondu « Pas satisfait », 12 « Moyennement satisfait » et 9 « Satisfait ».

- 1) Quel est le type de cette variable ?
- 2) Calculez le tableau de fréquences de cette variable
- 3) Peut-on calculer le mode pour une variable de ce type ?
- 4) Peut-on calculer la moyenne de ces observations ?

1) Type de la variable: qualitative ordinaire. Cette variable ne mesure pas une quantité chiffrée donc elle est qualitative. De plus, il y a un ordre naturel dans les modalités.

2) Fréquences:

- « Pas satisfait »: $n=4$ (16%)
- « Moyennement satisfait »: $n=12$ (48%)
- « Satisfait »: $n=9$ (36%)

3) Le mode est la modalité la plus fréquente. Ici, le mode est « Moyennement satisfait ».

4) Non. La moyenne ne se calcule que pour des variables quantitatives

Dans un échantillon de 6 individus, on a collecté les observations suivantes: 1, 2, 1, 2, 1, 2

- 1) Calculez la moyenne, la variance et l'écart-type de ces observations
- 2) Calculez la médiane de ces observations

Moyenne: $(1+2+1+2+1+2) / 6 = 9 / 6 = 1.5$

Variance: $[(1-1.5)^2+(2-1.5)^2+(1-1.5)^2+(2-1.5)^2+(1-1.5)^2+(2-1.5)^2] / (6-1)$
 $= 6 * 0.5^2 / (6-1) = 6 * 0.25 / 5 = 1.5 / 5 = 0.3$

Ecart-type: $\sqrt{0.3} = 0.55$

Pour calculer la médiane, ordonnez les observations: 1, 1, 1, 2, 2, 2
Comme le nombre d'observations est pair (n=6 observations), la médiane est le « milieu » entre les 3^{ème} et 4^{ème} observations ordonnées, donc entre 1 et 2. La médiane est $(1+2)/2=1.5$.

Ici, la médiane est égale à la moyenne. Ce n'est pas toujours le cas.

Donnez le mode et la médiane des 24 observations suivantes:

Valeur (modalité)	Fréquence (nombre de réponses égales à la valeur)
0	5
1	2
2	6
3	7
4	3
5	1

La moyenne des observations est-elle plus grande ou plus petite que la médiane ?

Les 24 observations de l'échantillon sont: 0, 0, 0, 0, 0, 1, 1, 2, 2, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 4, 4, 4, 5

Mode – valeur la plus fréquente: 3 (valeur observée 7 fois)

Médiane: 2

Le nombre d'observations est pair (24), donc la médiane est à mi-chemin entre la 12^{ème} observation la plus grande (qui est 2) et la 13^{ème} (qui est 2).

La moyenne est égale à : $(5*0+2*1+6*2+7*3+3*4+1*5)/24=52/24=2.2$

Pour chacune des situations suivantes, citez un type de représentation graphique permettant de décrire les résultats:

- 1) Comparaison du taux de cholestérol HLD entre des patients diabétiques et non diabétiques.
- 2) Etude de l'association entre l'âge et le taux de cholestérol HDL dans la population générale.
- 3) Etude du nombre de personnes par canton avec un taux de cholestérol LDL supérieur à 4.14 mmol/l.

- 1) Les box plots permettent de décrire la distribution du taux de cholestérol (variable quantitative continue). Une représentation appropriée à la question serait de tracer un box plot chez les patients diabétiques et un autre chez les patients non diabétiques (cf slide 26 du cours 2). Des histogrammes (slide 25 du cours 2) seraient aussi appropriés.
- 2) La relation entre 2 variables continues mesurées chez les mêmes sujets peut être visuellement examinée à l'aide d'un nuage de points (scatter plot, slide 27 du cours 2).
- 3) Un diagramme en bâton (bar chart) peut être utilisé: chaque barre représente le nombre de personnes **dans un canton** avec un taux de cholestérol LDL supérieur à 4.14 mmol/l. Un diagramme circulaire serait aussi approprié.