

Méthodes statistiques TD n°3

O. Dadoun dadoun@in2p3.fr semaine du 28/09/201

Histogramme

Graphique permettant de représenter la répartition d'une variable.

La forme de l'histogramme indique la distribution statistiques d'une variable.

La dispersion statistique

Tendance des valeurs de la distribution d'un caractère

- à s'étaler de part et d'autre d'une valeur centrale
- et/ou à s'éloigner les unes des autres

L'étendue d'une distribution

Est la différence entre la plus grande et la plus petite valeur d'une distribution : $x_{max} - x_{min}$

Exercice 1

184 étudiants de psycho. devaient décrire leur personnalité. Plusieurs traits de personnalité leur étaient proposés et, pour chacun d'eux, ils devaient se positionner sur une échelle allant de 0 (je ne suis pas du tout comme ça) à 9 (je suis tout à fait comme ça). L'une de ces échelles concernait le trait "audacieux", une autre le trait "ambitieux".

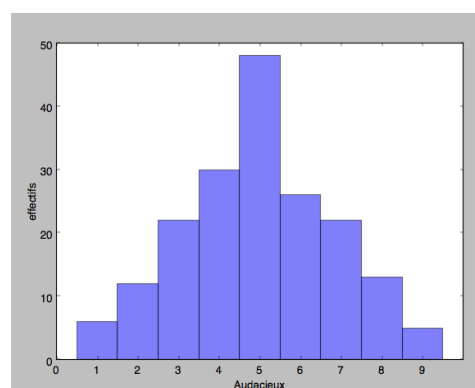
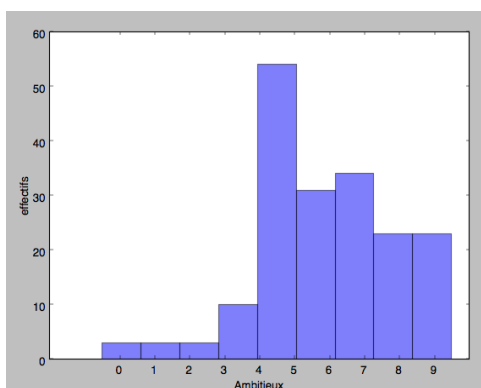
Les résultats sont reportés ci-dessous :

| Modalité (x) | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | |
|----------------------------|---|---|----|----|----|----|----|----|----|----|-------|
| Effectifs pour "audacieux" | 0 | 6 | 12 | 22 | 30 | 48 | 26 | 22 | 13 | 5 | N=184 |
| Effectifs pour "ambicieux" | 3 | 3 | 3 | 10 | 18 | 36 | 31 | 34 | 23 | 23 | N=184 |

1. Type de la variable et son échelle

Variable quantitative discrète échelle d'intervalle

2. Faites une représentation graphique sous forme d'histo. de chacune de ces distributions et comparez ce représentations



On constate de les notes que les sujets se sont attribuées en "audace" sont réparties en miroir autour d'un axe de symétrie situé sur la modalité 5. Plus on s'éloigne de cet axe (dans un sens ou dans l'autre) moins on trouve de sujet. La distribution des notes ambition ne présente pas un tel axe de symétrie et les notes sont tassées vers les valeurs fortes (bcp de sujet se sont attribués bcp d'ambition)

Définitions

La moyenne

Somme de toutes les valeurs prises par une variable divisée par le nombre d'observations

Défaut de la moyenne : elle est très sensible aux valeurs extrêmes.

La médiane

Les valeurs du caractère d'une série statistique étant rangées par ordre croissant, on définit la médiane. C'est un nombre tel qu'il y a autant de valeurs de la série qui lui sont inférieures que de valeurs qui lui sont supérieures.

La valeur de la position $(n+1)/2$ n étant le nb de valeurs dans un ensemble données

Moyenne et médiane sont des valeurs autour desquelles les observations sont réparties ; ce sont deux façons complémentaires de caractériser la tendance centrale.

Exemple

On a demandé à 50 personnes prenant l'autobus, le nombre de fois où chacune de ces personnes a utilisé ce type de transport pendant la semaine écoulée.

Voici les résultats:

| Nb de voyages en autobus | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|--------------------------|---|----|----|----|----|----|----|----|----|-----|
| Effectif | 3 | 3 | 5 | 7 | 6 | 9 | 5 | 4 | 5 | 3 |
| Effectif cumulé | 3 | 6 | 11 | 18 | 24 | 33 | 38 | 42 | 47 | 50 |
| Fréquence(%) | 6 | 6 | 10 | 14 | 12 | 18 | 10 | 8 | 10 | 6 |
| Fréquence cumulé | 6 | 12 | 22 | 36 | 48 | 66 | 76 | 84 | 94 | 100 |

- Nb moyen de voyages: **5.07**
- La médiane se trouve à la position: $(50+1)/2 = 25.5$
c.a.d entre la 25eme et la 26eme position qui à pour valeur 6

Si on présente ce tableau sous forme de liste on a:

| | | | | | | | | | | | | | | | | | | | | |
|----------------------|---|---|---|---|---|---|---|---|---|----|----|-----|----|----|----|-----|----|----|----|----|
| Nb de voyages en bus | 1 | 1 | 1 | 2 | 2 | 2 | 3 | 3 | 3 | 3 | 3 | ... | 5 | 6 | 6 | ... | 9 | 10 | 10 | 10 |
| Effectif cumulé | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | | 24 | 25 | 26 | | 47 | 48 | 49 | 50 |

Exercice

Sur un groupe de 20 personnes on demande combien ils ont d'argent sur eux.

| | | | | | | | |
|-----------------|---|---|----|----|----|----|-----|
| argent (€) | 2 | 5 | 10 | 15 | 20 | 30 | 500 |
| Nb de personnes | 3 | 3 | 3 | 2 | 5 | 3 | 1 |
| Nb cumulé | 3 | 6 | 9 | 11 | 16 | 19 | 20 |

- Calculer la moyenne **38.55 €**
- Calculer la médiane **position de la médiane 10.5 entre 10ème et 11ème position, la médiane vaut 15€**

La moitié des personnes ont 15€ et plus en poche et l'autre moitié 15€ et moins.

Donc si on a un produit qui vaut plus que 15€ la moitié des personnes ne pourront pas l'acheter...

La moyenne vaut 38.55€ si on avait fixé le prix par rapport à la moyenne seul 1 personne aurait pu acheter le produit ...

Cette différence est due à la distribution de la population observée avec un individu pesant très lourd.

La dernière personne (avec les 500€) se fait voler son argent.

La moyenne sur 19 personnes est de 14 € et la médiane de 15€ .

La moyenne est très sensible aux valeurs extrêmes pas la médiane.

Exercice

Calculer la médiane de cette série

5 – 10 – 5 – 8 – 4 – 11 – 6 – 4 – 8 – 7 – 9 – 6 – 13 – 15 – 8 – 7 – 5 – 10

rep. 7.5