

1 Introduction-vocabulaire

Les premières études statistiques étaient des recensements démographiques : on en a conservé le vocabulaire.

- L'ensemble sur lequel porte l'étude statistique s'appelle la population.
- Un élément de cet ensemble est un individu.
- La population est étudiée selon un ou plusieurs caractères (ou variables).
- Le caractère est :
soit quantitatif (on peut le mesurer) ; on lui donne souvent le nom de variable statistique.
soit Qualitatif (on ne peut pas le mesurer)
- L'effectif d'une classe statistique est le nombre d'éléments de la population observés dans cette classe.

Pour la suite on notera x_i (i entier naturel) les valeurs du caractère et n_i les effectifs correspondants.

$N = n_1 + n_2 + \dots$ est l'effectif total.

- Le mode (ou classe modale) est la valeur que la variable statistique prend le plus souvent. C'est à dire la valeur du caractère ou de la classe qui a le plus grand effectif.

2 Effectifs cumulés-fréquences

Définition : Fréquences

La fréquence d'une classe statistique est le rapport de l'effectif de cette classe à l'effectif total de la population. (la fréquence peut être exprimée en pourcentage)

La fréquence d'une classe statistique est le rapport de l'effectif de cette classe à l'effectif total de la population : $f_i = \frac{n_i}{N}$

Définition : Effectifs et fréquences cumulées

L'effectif (la fréquence) cumulé(e) croissant(e) d'une classe statistique x_i est la somme des effectifs (fréquences) des valeurs inférieures ou égales à x_i .

3 Etude de deux exemples

Pour la suite, on utilisera ces deux exemples pour appliquer les définitions...

Cas 1

Dans un service de maintenance, on a répertorié le nombre d'interventions par jour sur un mois. On a obtenu la distribution suivante :

Nombre d'interventions x_i	2	3	5	6	7	8	9	Total
Nombre de jours n_i	2	5	4	6	7	8	9	30
Fréquences f_i	$\frac{2}{30}$	$\frac{5}{30}$	$\frac{4}{30}$	$\frac{6}{30}$	$\frac{7}{30}$	$\frac{8}{30}$	$\frac{9}{30}$	$\frac{30}{30}$
Effectifs cumulés croissants	2	7	11	17	24	32	41	

Cas 2

Dans une succursale de banque, on a noté le montant des 2000 versements effectués au guichet pendant la journée.

Montant (en euros)	0, 500	500, 750	750, 1000	1000, 1500	1500, 3000
effectif	440	320	400	480	360
Fréquences	0,22 ou 22%	0,16 ou 16%	0,20 ou 20%	0,24 ou 24%	0,18 ou 18%
Effectifs cumulés	440	720	1120	1600	1960
Effectifs cumulés décroissants	2000	1560	1240	860	480

Exemple 1

Compléter ces deux tableaux après avoir précisé la population, le caractère et le mode de chacune de ces deux séries de données.

4 représentation graphique d'une série statistique

4.1 nuage de points

On place les points de coordonnées (x_i, n_i) dans un repère orthogonal.

Exemple 2

Construire ce graphique pour le cas 1

4.2 histogramme

L'aire de chacun des rectangles est proportionnelle à l'effectif.

Exemple 3

Construire ce graphique pour le cas 2 en prenant 1 cm (ou un carreau) pour 250 euros et 1 cm² (ou l'aire de un carreau) pour 50 versements.

4.3 diagramme circulaire

La mesure de l'angle du secteur angulaire est proportionnelle à l'effectif correspondant sachant que l'effectif total correspond 360°.

4.4 courbe des effectifs (ou fréquences) cumulés(ées)

On place les points dont les abscisses sont données par les valeurs de la variable et les ordonnées par le effectifs (ou fréquences) cumulés (ées) dans un repère.

S'il s'agit d'une série de données à variable continue, on suppose dans ce cas que la répartition est uniforme dans chaque classe et l'abscisse correspond à la valeur maximale de chacune des classes pour les effectifs cumulés croissants et à la valeur minimale de chaque classe s'il s'agit des effectifs cumulés décroissants (voir cas 2).

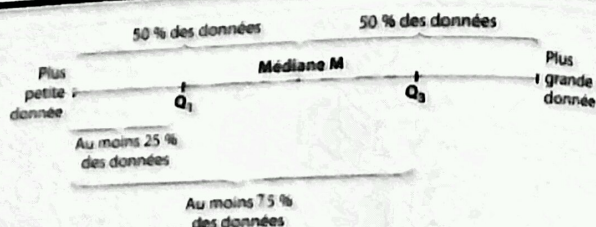
Exemple 4

Construire le graphique des effectifs cumulés croissants pour le cas 1 puis pour le cas 2

5.3 Quartiles

Définition : Quartiles

Le premier quartile est plus petite valeur Q_1 telle que au moins un quart (25%) des données soient inférieures ou égales à Q_1 .
Le troisième quartile est plus petite valeur Q_3 telle que au moins un trois quarts (75%) des données soient inférieures ou égales à Q_3 .



Remarque On peut définir de même les déciles. Le premier décile est la plus petite valeur D_1 telle que au moins 10% des valeurs soient inférieures ou égales à D_1 .

Exemple 8

Déterminer les premiers et troisième quartiles pour chacun des deux cas.

6 caractéristiques de dispersion

Définition : Etendue et écart interquartile

L'étendue d'une série statistique est la différence entre la plus grande et la plus petite valeur du caractère.

L'écart interquartile est l'écart entre le premier et le troisième quartile $Q_3 - Q_1$.

7 linéarité de la moyenne

Propriété : Linéarité de la moyenne

Soit $\bar{x}$ la moyenne de la série de données (x_i, n_i) ($1 \leq i \leq q$).

Si on ajoute α (α réel) à chacun des termes de la série de données alors la moyenne de cette série de données (x_i, n_i) est alors $\bar{x} + \alpha$.

Si on multiplie chacun des termes de cette série de données par k ($k \in \mathbb{R}^*$) alors la moyenne de cette série de données (kx_i, n_i) est $k\bar{x}$.

Preuve :

$$\frac{x_1 + \alpha n_1 + x_2 + \alpha n_2 + \dots + x_q + \alpha n_q}{N} = \frac{x_1 n_1 + x_2 n_2 + \dots + x_q n_q}{N} + \frac{\alpha(n_1 + n_2 + \dots + n_q)}{N} = \bar{x} + \alpha$$

$$\frac{kx_1 + n_1 + kx_2 + n_2 + \dots + kx_q + n_q}{N} = k \frac{x_1 n_1 + x_2 n_2 + \dots + x_q n_q}{N} = k\bar{x}$$

Exemple 9

5 caractéristiques de position et de dispersion

5.1 moyenne

Définition : Moyenne

La moyenne d'une série statistique est notée $\bar{x} = \frac{n_1 x_1 + n_2 x_2 + \dots + n_q x_q}{N}$.

Dans le cas d'une série regroupée en classe, on utilise le centre des classes pour faire le calcul de la moyenne.

Remarque On peut aussi calculer la moyenne à partir des fréquences.

$$\bar{x} = f_1 x_1 + f_2 x_2 + \dots + f_p x_p = \frac{n_1}{N} x_1 + \frac{n_2}{N} x_2 + \dots + \frac{n_p}{N} x_p = \frac{n_1 x_1 + n_2 x_2 + \dots + n_p x_p}{N}$$

Utilisation de la calculatrice graphique.

[Menu] puis [Stats]

Entrer les données dans la liste 1 (pour les valeurs) et dans la liste 2 pour les effectifs.

Si des données sont déjà présentes dans les listes, les effacer en faisant défiler le bandeau en bas de l'écran [F6] puis Del-A [F4].

Après avoir entré les données.

Calc [F2] puis SET [F6] pour paramétrer la calculatrice correctement.

On utilise les statistiques à une variable.

IVAR XLstat doit correspondre aux valeurs du caractère et IVAR Freq aux effectifs (choisir la liste correspondante avec les touches F1, F2).

Ensuite, revenir à l'écran précédent [Exit] ou [Quit] puis IVAR [F1].

Les résultats s'affichent alors sur l'écran.

Exemple 5

Calculer la moyenne pour les deux cas manuellement puis avec le menu statistiques.

Exemple 6

Dans le lycée, lors d'un devoir commun à trois classes de seconde, on a obtenu les moyennes suivantes :

seconde 1 : 26 élèves et une moyenne de 11.

seconde 2 : 28 élèves et une moyenne de 10.

seconde 3 : 31 élèves et une moyenne de 12.

Calculer la moyenne générale de ce contrôle pour les trois classes réunies.

5.2 Médiane

Définition : Médiane

La médiane d'une série statistique est une valeur de la variable telle qu'il y ait autant d'observations ayant une valeur supérieure à la médiane que d'observations ayant une valeur inférieure à la médiane.

Si la série est regroupée en classe, la classe à laquelle appartient la médiane est appelée classe médiane.

Remarque On peut utiliser les effectifs cumulés croissants pour déterminer la médiane.

Exemple 7

Déterminer la médiane pour les deux cas et en déduire la signification.

2. LA MOYENNE : UN AUTRE EXEMPLE DE VALEUR CENTRALE

Définition 3. La moyenne $\bar{x}$ des caractères x_i d'une série statistique est :

$$\bar{x} = \frac{x_1 + \dots + x_N}{N} = \frac{1}{N} \sum_{i=1}^N x_i = \frac{n_1 a_1 + \dots + n_k a_k}{N} = \frac{1}{N} \sum_{j=1}^k n_j a_j$$

Exemple 2. Dans l'exemple 1, $\bar{x} =$

L'avantage de la moyenne réside dans le fait qu'elle est linéaire : on peut calculer la moyenne de deux groupes dont on connaît les effectifs et les moyennes.

Propriété 1

La moyenne des images des valeurs du caractère par une fonction affine du type $f(x) = ax + b$ est l'image de la moyenne $f(\bar{x}) = a\bar{x} + b$.

Si la série z est composée de deux séries x et y d'effectifs respectifs p et q , on a : $\bar{z} = \frac{p\bar{x} + q\bar{y}}{p+q}$

Preuve. On a : $\frac{(ax_1 + b) + \dots + (ax_N + b)}{N} = \frac{a(x_1 + \dots + x_N) + Nb}{N} = a\bar{x} + b$

$$\bar{z} = \frac{n_1 + \dots + n_N}{p+q} = \frac{p\bar{x} + q\bar{y}}{p+q} = \frac{p\bar{x} + q\bar{y}}{p+q}$$

Exemple 3. On suppose qu'au même contrôle, un groupe de 20 élèves a obtenu une moyenne de $\bar{x} = 11,4$. Calculer la moyenne des 30 élèves (les 20 plus les 10 de l'exemple 1).

$$\bar{z} = \frac{10 \times 10,9 + 20 \times 11,4}{20 + 10} = 11,23$$

Exemple 4. Une entreprise compte 99 salariés gagnant 1 200 € par mois et un dirigeant gagnant 81 200 €. Quel est le salaire moyen ? Est-il représentatif de la situation ?

$$\bar{s} = \frac{1200 \times 99 + 1 \times 81200}{99 + 1} = 2000$$

3. LA MÉDIANE : UN TROISIÈME EXEMPLE DE VALEUR CENTRALE

Définition 4. On suppose les x_i rangés dans l'ordre croissant.

• Une médiane M est telle que 50% des caractères sont au dessus de M et 50% en dessous.

Si N est impair, la médiane est la valeur du milieu.

Si N est pair, une médiane est la moyenne des deux valeurs du milieu.

• Le premier quartile Q_1 est une médiane de la série des x_i où $i > \frac{N+1}{4}$.

• Le troisième quartile Q_3 est une médiane de la série des x_i où $i < \frac{3N+1}{4}$.

• L'intervalle interquartile est $[Q_1, Q_3]$.

• L'écart interquartile est $Q_3 - Q_1$. C'est le paramètre de dispersion associé à la médiane.

Exemple 5. Dans l'exemple 1 :

Min : 7 ; 1 ; 9 ; 9 ; 11 ; 11 ; 14 ; 18

Max : 18 ; 14 ; 11 ; 9 ; 9 ; 1 ; 7

Méd : 11 ; 11 ; 11 ; 11 ; 11 ; 11 ; 11

$Q_1 = 9$; $Q_3 = 14$; $Q_3 - Q_1 = 5$

L'inconvénient de la médiane est qu'elle n'est pas linéaire. Son intérêt est d'être peu sensible aux valeurs extrêmes.

Exemple 6. Dans l'exemple 4, le salaire médian (la médiane des salaires) est :

4. UTILISER LA CALCULATRICE

Exemple 7. Durant quatre cours de mathématiques, un professeur a fait tracer à chacun de ses élèves un carré de 10 cm de côté et mesurer la diagonale. Ces mesures sont consignées ici :

Mesure	13,7	13,8	13,9	14	14,1	14,2	14,3	14,4	14,5	14,6	N	Méd	Min	Max
1	1	1	2	2	0	4	4	1	1	0	19	14,5	13,7	14,6
2	2	2	2	2	4	0	1	1	1	0	18	14,2	13,8	14,5
3	0	2	3	0	2	4	3	0	2	2	18	14,2	13,8	14,5
4	0	0	3	1	3	2	4	3	0	0	16	14,2	13,8	14,5

2nd F2
13.7
13.8
13.9
14
14.1
14.2
14.3
14.4
14.5
14.6

STAT (edit et 1) (Editor)

1 13.7
2 13.8
3 13.9
4 14
5 14.1
6 14.2
7 14.3
8 14.4
9 14.5
10 14.6

modalités → L1, effectifs → L2

EDIT NEW TESTS
1 Stats 1-Var
2 Stats 2-Var
3 Reg-Mod
4 Reg-Inv/Ansb
5 Reg-Quad
6 Reg-Log
7 Reg-Exp
8 Reg-Pow

STAT (edit et 1) (Stats 1-Var)

Stats 1-Var L1:L
n=19
Σx=373.9
Σx²=5045.89
σx=14.23
σx²=202.5

L1 2nd (edit et 1)

Stats 1-Var
n=19
Σx=373.9
Σx²=5045.89
σx=14.23
σx²=202.5

Ecran 1

Stats 1-Var
n=19
Σx=373.9
Σx²=5045.89
σx=14.23
σx²=202.5

Ecran 2

Exemple 8. Remplir le tableau de l'exemple 7 à la calculatrice.

La moyenne des mesures des quatre classes est :

Représenter le diagramme en boîte des secondes 2,3 et 4.

