**COURS 2nde STATISTIQUES : DEFINITIONS ET EXEMPLES Fiche 1/2**

On considère les résultats d’une évaluation de mathématiques d'une classe de 2nde de 25 élèves, que l'ont range dans un tableau:

11, 13, 8, 16, 11, 7, 13, 12, 11, 12, 12, 8, 11, 8, 12, 11, 11, 7, 16, 8, 11, 8, 12, 8, 16

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Notes | 7 | 8 | **11** | 12 | 13 | 16 | total |
| Effectifs | 2 | 6 | **7** | 5 | 2 | 3 | 25 |
| Eff. cumulés. croissants. | 2 | 8 | **15** | 20 | 22 | 25 |  |
| Fréquences | 0,08 | 0,24 | 0,28 | 0,2 | 0,08 | 0,12 | 1 |
| Fréquences en % | 8 | 24 | 28 | 20 | 8 | 12 | 100 |

E. C. C.

F.C.C. Fréquences cumulées croissantes % 8 32 60 80 88 100 X

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Définitions | | Dans notre exemple | | |
| On appelle **population** un ensemble soumis à une étude statistique | | La population est | | |
| On appelle **individu** un élément d'une population. | | Un individu est | | |
| On appelle **caractère** le critère suivant lequel on étudie la population. Suivant les cas le caractère peut être   * **discret** (âge, pointure, couleur des yeux, ...) * ou **continu** (la taille, les salaires,... ). Dans ce dernier cas les valeurs sont données par classe ([160, 165[ pour la taille)(*dans ce cas on demande souvent de calculer la moyenne de la classe ici : (160 + 165) : 2 = 325 :2=162,5, et on revient à un caractère discret) c’est le centre de la classe.* | | Le caractère est LA NOTE obtenue à une évaluation.  C'est un caractère DISCRET. | | |
| On appelle **effectif total** le nombre total d'individus dans la population. | | L'effectif total est N = 25 | | |
| On appelle **effectif** le nombre d'individus qui correspondent au même caractère. | | 7 est l'effectif de la note 11 : Il y a 7 élèves qui ont eu la note 11. | | |
| On appelle **effectif cumulé croissant** le nombre d'individus qui correspondent au même caractère et aux caractères précédents. | | Il y a élèves qui ont eu une note inférieur ou égale à 11 ( élèves ont eu 11 ou moins de 11) | | |
| On appelle **fréquence** le rapport entre l'effectif d'une valeur et l'effectif total.  Pour l’avoir en pourcentage : on multiplie par 100 ce rapport.  On peut calculer les **fréquences cumulées croissantes** | | • Fréquence de la note 7/20 : 2/25 = 0,08  f en % : 0,08 x 100 = 8%  • Il y a 8 élèves qui ont eu 7 et 8 sur 20. (≤ 8)  f en % : 32% | | |
| On appelle **moyenne** ou **moyenne arithmétique** le rapport entre la somme des valeurs et le nombre de valeurs. | | Moyenne arithmétique : (7+8+11+12+13+16) / 6 = 11,17 (ne sert pas aux élèves.) | | |
| On appelle **moyenne pondérée**, une moyenne dans laquelle chaque valeur possède un coefficient.  x : Valeur n : son coefficient  Capture d’écran 2015-05-12 à 15.25.00.png | | Moyenne pondérée :   = 10,92 | | |
| On appelle **médiane** la valeur située au milieu de la liste quand celle-ci est classée dans l'ordre (dé)croissant.  Si le nombre de valeurs est pair, la médiane est la moyenne des deux valeurs situées au milieu. | | 7, 7, 8, 8, 8, 8,Q1 8, 8, 11, 11, 11, 11, 11, 11, 11, 12, 12, 12, 12, 12, 13, 13, 16, 16, 16   donc la médiane est la 13ième valeur. | | |
| On appelle **étendue** d'une série statistique, la différence entre la valeur la plus élevée et la valeur la moins élevée.  Quand cette valeur est "grande" par rapport à l'ensemble des valeurs possibles, on dit que la série est **dispersée.** | | La valeur la plus élevée : 16 La valeur la moins élevée : 7 Étendue = 16-7 = 9 Les notes "s'étalent" donc sur 9 points , la série est **dispersée.** (comparer aux 20 points possibles de différence) | | |
| Représentations graphiques : | | | | |
| stat01 **Diagramme en bâtons** | stat02**Histogramme** | | stat03  **Diagramme circulaire** angle = fréquence × 360 | stat04  **Diagramme semi-circulaire** angle = fréquence × 180 |
| Définitions | | | Dans notre exemple fiche 2/2 | |
| **LES QUARTILES :**  On appelle **premier quartile noté Q1** le plus petit nombre tel qu’au moins 25 % des valeurs de la série soient inférieures ou égales à ce nombre.  Q2 = Médiane. 50% des valeurs lui sont inférieures.  On appelle **troisième quartile Q3** le plus petit nombre tel qu’au moins 75 % des valeurs de la série soient inférieures ou égales à ce nombre.  Remarques :  . Q1 et Q3 sont obligatoirement des valeurs de la série étudiée. . Q1, Me, Q3 partagent la série en 4 sous séries. . Q1 et Q3 sont les médianes des sous séries obtenues.  **L’intervalle interquartile** d’une série statistique est l’intervalle  [ Q1 ; Q3]  **L’écart interquartile** d’une série statistique est le nombre Q3 – Q1  Remarque  L’écart interquartile mesure la dispersion des valeurs autour de la médiane;  Plus l’écart est petit, plus les valeurs de la série appartenant à l’intervalle interquartile sont concentrées autour de la médiane.  Contrairement à **l’étendue** (notée e) qui mesure l’écart entre la plus grande et la plus petite valeur, l’écart interquartile élimine les valeurs extrêmes qui peuvent être douteuses, cependant il ne tient compte que de 50% de l’effectif. | | | **Déterminons Q1:**  **N est impair.**  **N x 25/100= N/4 = 25/4 = 6,25 : Q1 est donc la moyenne entre la 6ième et la 7ième valeur : (8 +8) / 2 = 8**  **Q1 est la médiane de la sous série qui comporte 12 notes.  Q1 = 8 (/20).**  **Déterminons Q3 :**  **Nx3/4 = 25x3/4 = 18,75 : Q3 est donc la moyenne entre la 18ième et la 19ième valeur : ( 12+12) / 2 = 12.**  **Q3 est la médiane de la sous série ( supérieure à Médiane) .**  **Intervalle interquartile :**  **Dans l’intervalle [ 8 ; 12], on retrouve 50% de la population.**  **Ecart interquartile :**  Q3- Q1 = 12-8 = 4 (4 points sur 20) | |
| **LES DECILES :**  Les déciles notées D1, D2, …. D9, partage la série en 10 sous groupe de même effectif. Me = D5 .  Ce sont aussi des valeurs de la série ou des moyennes .  **D1 :** c’est la valeur de la série pour laquelle 10% des valeurs lui sont inférieures ou égales.  **D9 :** c’est la valeur de la série pour laquelle 90% des valeurs lui sont inférieures ou égales. | | | **Déterminer D1 :**  **N/10 = 2,5 Donc D1 sera la moyenne entre le 2 et le 3 ième valeur : (7+8) /2 = 7,5**  **Déterminer D9 :**  **Nx9/10 = 25x9/10 = 22,5 Donc D9 est la moyenne entre la 22ième et la 23ième valeur du caractère.(13+16)/2 = 14,5.**  *90% des élèves ont eu une note inférieure à 14,5 et donc 10 % ont eu une note supérieure à 14,5.* | |
| **Les caractéristiques de position :**  moyenne (mais peu d’intérêt dans les comparaisons), MEDIANE, Q1, Q3.  **Les caractéristiques de dispersion :** étendue, intervalle interquartile ou écart interquartile et L’ECART TYPE.  La **médiane** comme paramètre de position et **l'intervalle interquartile** comme paramètre de dispersion fournissent une bonne description d'une série statistique. Ces deux paramètres permettent une bonne comparaison et analyse entre deux séries d’une même étude sur des individus différents. | | | | |
| **COURS 1 S LE DIAGRAMME EN BOÎTE : BOITE A MOUSTACHES**  La série de notes étudiée ici se résume en:  - le minimum Min =  - le 1er quartile Q1 =  - la médiane Me =  - le 3ème quartile Q3 =  - le maximum Max =  Ces 5 données permettent de construire un diagramme en boîte ou boîte à moustaches. En priorité : étudier et tracer une échelle.   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | | | | |

Voir fiche TI ou CASIO pour utilisation de la calculatrice.

**COURS 1S STATISTIQUES : DEFINITIONS ET EXEMPLES *ACoco***

**Ajustement affine de la MOYENNE:**

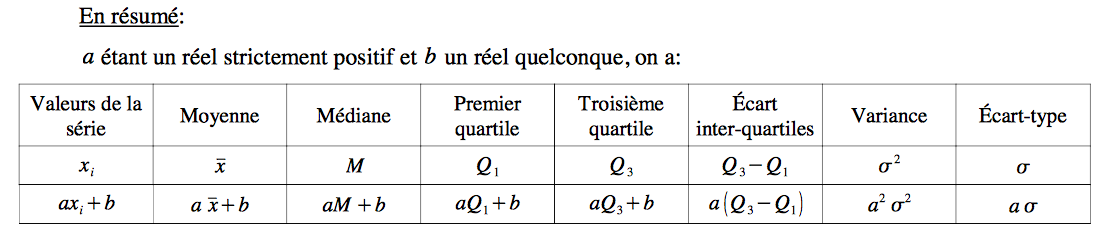
Ajouter un réel b aux notes, les multiplier par un réel non nul a, revient à ajouter b à la moyenne et à la multiplier par a aussi.

Exemple :

Imaginons qu’on multiplie les notes par 1,5 et qu’on ajoute 0,5 point à la série donnée ci dessus, ALORS la moyenne sera  :

1,5 x 10,92 + 0,5 = 16,38 + 0,5 = 16,88.

**AUTRES CARACTERES après ajustement affine :**



**LA VARIANCE et L’ECART TYPE d’UNE SERIE STATISTIQUE :**

Pendant la semaine du 13 au 19 septembre 2010, on a relevé les températures minimales et maximales à Brest :

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Température minimale en °C | 8,8 | 12,2 | 13,5 | 12,7 | 8,5 | 7,7 | 5,2 |  |
| Température maximale en °C | 19,5 | 19,9 | 18,6 | 17,8 | 18 | 17,3 | 18,1 |  |

Les températures maximales semblent plus « régulières » que les températures minimales.

Le but de cette activité est d’introduire une nouvelle caractéristique d’une série statistique pour mesurer sa **dispersion autour de la moyenne.** On pourra alors **comparer la « régularité »** de deux séries.

Dans les quatre premières questions, on considère seulement les températures minimales.

1) Calculer la température minimale moyenne x.

2) Calculer les écarts à la moyenne : ( x i  - x). Compléter le tableau.

3) Qu’observe-t-on quand on calcule la moyenne de ces différences ? ON TROUVE 0…

4) Ce qui précède amène à ne considérer que des quantités positives. Pour cela, on peut calculer les valeurs absolues ou les carrés. Les carrés, moins naturels, ont cependant été choisis car les propriétés mathématiques sont ensuite beaucoup plus intéressantes. Compléter le tableau.

5) Calculer la moyenne des carrés des écarts à la moyenne 

Le nombre obtenu s’appelle **la variance** de la série statistique, on le note V.  =8,02

LA VARIANCE d’UNE SERIE STATISTIQUE est la moyenne des écarts au carré des valeurs à la moyenne.

1. Pour compenser l’utilisation des carrés et se ramener à une quantité de même nature (même unité) que les termes de la série statistique, on calcule maintenant la racine carrée de la variance V.

Ce nouveau nombre s’appelle **L’ECART TYPE** de la série statistique, on le note  = 2,83.

MOYENNE

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Température minimale en °C | 8,8 | 12,2 | 13,5 | 12,7 | 8,5 | 7,7 | 5,2 | 9,8 |
| Ecart à la moyenne  x i  - x | 8,8-9,8  =-1 | 12,2-9,8  = 2,4 | 13,5-9,8  =3,7 | 12,7-9,8  =2,9 | 8,5-9,8  =-1,3 | 7,7-9,8  =-2,1 | 5,2-9,8  =-4,6 | 0 |
| Carré de l’écart à la moyenne  (x i  - x)2 | 1 | 5,76 | 13,69 | 8,41 | 1,69 | 4,41 | 21,16 | Variance :  8,02 |

1. Calculer sur le même principe l’écart type de la série des températures maximales : Ecart type = 0, 87
2. Finalement, quelle série paraît la plus « régulière » :

**COURS 2nde STATISTIQUES : DEFINITIONS ET EXEMPLES Fiche PROF**

On considère les résultats d'un devoir d'une classe de 25 élèves que l'on range dans un tableau.:

11, 13, 8, 16, 11, 7, 13, 12, 11, 12, 12, 8, 11, 8, 12, 11, 11, 7, 16, 8, 11, 8, 12, 8, 16

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Notes | 7 | 8 | **11** | 12 | 13 | 16 | total |
| Effectifs | 2 | 6 | **7** | 5 | 2 | 3 | 25 |
| Eff. cumulés. croissants. | 2 | 8 | **15** | 20 | 22 | 25 |  |
| Fréquences | 0,08 | 0,24 | **0,28** | 0,20 | 0,08 | 0,12 | 1 |
| Fréquences en % | 8 | 24 | **28** | 20 | 8 | 12 | 100 |

|  |  |  |  |
| --- | --- | --- | --- |
| Définitions | | Dans notre exemple | |
| On appelle **population** un ensemble soumis à une étude statistique | | La population est la classe de 25 élèves. | |
| On appelle **individu** un élément d'une population. | | Un individu est un élève. | |
| On appelle **caractère** le critère suivant lequel on étudie la population. Suivant les cas le caractères peuvent être   * **discret** (âge, pointure, couleur des yeux, ...) * ou **continu** (la taille, les salaires,... ). Dans ce dernier cas les valeurs sont données par **classe** ([160, 165[ pour la taille)(*dans ce cas on demande souvent de calculer la moyenne de la classe ici : (160 + 165) : 2 = 325 :2=162,5, et on revient à un caractère discret). C’est le centre de la classe.* | | Le caractère est la note obtenue par chaque élève.  C'est un caractère discret. | |
| On appelle **effectif total** le nombre total d'individus dans la population. | | L'effectif total est 25. | |
| On appelle **effectif** le nombre d'individus qui correspondent au même caractère. | | 7 est l'effectif de la note 11 : Il y a 7 élèves qui ont eu la note 11. | |
| On appelle **effectif cumulé croissant** le nombre d'individus qui correspondent au même caractère et aux caractères précédents. | | Il y a 15 élèves qui ont eu une note inférieur ou égale à 11 (15 élèves ont eu 11 ou moins de 11) | |
| On appelle **fréquence** le rapport entre l'effectif d'une valeur et l'effectif total.  Pour l(avoir en pourcentage : on multiplie par 100 ce rapport  On peut calculer les **fréquences cumulées croissantes** | | f = 7 : 25 = 0,28 =  f en % : 0,28 x 100 = 28% (28% des élèves ont eu 11)  il y a 6+2 = 8 élèves qui ont eu 7 et 8 sur 20.  f en % : 8 : 25 x 100 = 0,32 x100 = 32% des élèves de la classe !! | |
| On appelle **moyenne** (ou **moyenne arithmétique**) le rapport entre la somme des valeurs et le nombre de valeurs. | | Moyenne des valeurs : ( 7 + 8 + 11 + 12 + 13 + 16) : 6 = 11,17 NE SERT PAS VRAIMENT…. | |
| On appelle **moyenne pondérée**, une moyenne dans laquelle chaque valeur possède un coefficient. | | Moyenne pondérée = stat05 La moyenne de la classe pour ce devoir est 10,92. | |
| On appelle **médiane** la valeur située au milieu de la liste quand celle-ci est classée dans l'ordre (dé)croissant.  Si l’EFFECTIF TOTAL est pair, la médiane est la moyenne des deux valeurs situées au milieu. | | 7, 7, 8, 8, 8, 8, 8, 8, 11, 11, 11, 11, **11**, 11, 11, 12, 12, 12, 12, 12, 13, 13, 16, 16, 16 La médiane est 11. Elle sépare la population en deux sous groupe d’effectif 50% de N=25, soit Deux ss gr de 12 individus. ( le 13ième est la médiane.) | |
| On appelle **étendue** d'une série statistique, la différence entre la valeur la plus élevée et la valeur la moins élevée.  Quand cette valeur est "grande" par rapport à l'ensemble des valeurs possibles, on dit que la série est **dispersée.** | | La valeur la plus élevée : 16 La valeur la moins élevée : 7 Étendue = 16 – 7 = 9 points sur 20. Les notes "s'étalent" donc de 7 à 16, la série est dispersée. | |
| Représentations graphiques : | | | |
| stat01 **Diagramme en bâtons** | stat02**Histogramme** | stat03  **Diagramme circulaire** angle = fréquence × 360 | stat04  **Diagramme semi-circulaire** angle = fréquence × 180 |