

Statistiques, cours, 2nde

Statistiques, cours, 2nde

F.Gaudon

<http://mathsfg.net.free.fr>

1^{er} octobre 2014

- 1 Vocabulaire et notations
- 2 Fréquences, séries cumulées
- 3 Caractéristiques statistiques
 - Moyenne
 - étendue
 - Médiane
 - Quartiles

- 1 Vocabulaire et notations
- 2 Fréquences, séries cumulées
- 3 Caractéristiques statistiques
 - Moyenne
 - étendue
 - Médiane
 - Quartiles

Définitions et notations :

On considère une série statistique.

- Si la série statistique comporte p valeurs distinctes avec $p \in \mathbb{N}$ (c'est à dire p entier naturel) , les valeurs du caractère étudiées sont notées x_i pour i entier naturel allant de 1 à k . Le nombre d'individus pour la valeur x_i , c'est à dire l'effectif pour la valeur x_i est noté n_i .
- Si la série est regroupée en p classes $[a_i; a_{i+1}[$ pour i allant de 1 à p , où a_i sont des réels tels que $a_i < a_{i+1}$, on prend pour valeurs les centres des classes $c_i = \frac{a_i + a_{i+1}}{2}$ pour i allant de 1 à p .
- L'effectif total N est alors égal à $N = n_1 + n_2 + \dots + n_p$ ce que l'on note aussi $\sum_{i=1}^{i=p} n_i = N$.

Exemple :

On étudie la hauteur des plantes dans un champ de tulipes. La population est l'ensemble des tulipes. Un individu est donc une tulipe. Le caractère étudié est la hauteur des tulipes. C'est un caractère quantitatif. Les valeurs du caractères x_i sont les différentes hauteurs de tulipes relevées. Il peut y avoir pour une hauteur $x_1 = 30$ cm un effectif de $n_1 = 3$ tulipes qui ont cette hauteur. L'effectif total N est le nombre total de tulipes. On peut regrouper les valeurs en classes, par exemple la classe des tulipes qui ont une hauteur dans l'intervalle $[30; 35[$. Le centre de l'intervalle est alors

Exemple :

On étudie la hauteur des plantes dans un champ de tulipes. La population est l'ensemble des tulipes. Un individu est donc une tulipe. Le caractère étudié est la hauteur des tulipes. C'est un caractère quantitatif. Les valeurs du caractères x_i sont les différentes hauteurs de tulipes relevées. Il peut y avoir pour une hauteur $x_1 = 30$ cm un effectif de $n_1 = 3$ tulipes qui ont cette hauteur. L'effectif total N est le nombre total de tulipes. On peut regrouper les valeurs en classes, par exemple la classe des tulipes qui ont une hauteur dans l'intervalle $[30; 35[$. Le centre de l'intervalle est alors $c_1 = \frac{30+35}{2} = 32,5$.

- 1 Vocabulaire et notations
- 2 Fréquences, séries cumulées**
- 3 Caractéristiques statistiques
 - Moyenne
 - étendue
 - Médiane
 - Quartiles

Définition :

On appelle *fréquence* f_i de la série pour la valeur x_i le nombre réel défini par

$$f_i =$$

Définition :

On appelle *fréquence* f_i de la série pour la valeur x_i le nombre réel défini par

$$f_i = \frac{n_i}{N}$$

qui s'exprime aussi en pourcentage en multipliant par 100.

Exemple :

Dans l'exemple des tulipes, s'il y a 3 tulipes de taille 30 cm parmi un effectif total de 65 tulipes, la fréquence de la valeur 3 est $f =$

Exemple :

Dans l'exemple des tulipes, s'il y a 3 tulipes de taille 30 cm parmi un effectif total de 65 tulipes, la fréquence de la valeur 3 est $f = \frac{3}{65} \approx 0,046$ soit 4,6 % environ.

Définition :

On appelle :

- *effectif cumulé croissant (ECC)* pour la valeur x_j , la somme des effectifs des valeurs inférieures ou égales à x_j ;

Définition :

On appelle :

- *effectif cumulé croissant (ECC)* pour la valeur x_i , la somme des effectifs des valeurs inférieures ou égales à x_i ;
- *fréquence cumulée croissante (FCC)* pour la valeur x_i , la somme des fréquences des valeurs inférieures ou égales à x_i .

Exemple :

On a relevé le prix de la baguette de pain dans différentes boulangeries :

Prix en euros (x_i)	0,82	0,83	0,84	0,85	0,86	0,87
Nombre de boulangeries (n_i)	52	40	35	28	23	2
Fréquences (f_i)						

Exemple :

On a relevé le prix de la baguette de pain dans différentes boulangeries :

Prix en euros (x_i)	0,82	0,83	0,84	0,85	0,86	0,87
Nombre de boulangeries (n_i)	52	40	35	28	23	2
Fréquences (f_i)	$\frac{52}{80} \approx 0,289$	0,222	0,194	0,155	0,128	0,011
Effectifs cumulés croissants						

Exemple :

On a relevé le prix de la baguette de pain dans différentes boulangeries :

Prix en euros (x_i)	0,82	0,83	0,84	0,85	0,86	0,87
Nombre de boulangeries (n_i)	52	40	35	28	23	2
Fréquences (f_i)	$\frac{52}{80} \approx 0,289$	0,222	0,194	0,155	0,128	0,011
Effectifs cumulés croissants	52	92	127	155	178	180

- 1 Vocabulaire et notations
- 2 Fréquences, séries cumulées
- 3 Caractéristiques statistiques**
 - Moyenne
 - étendue
 - Médiane
 - Quartiles

Plan

- 1 Vocabulaire et notations
- 2 Fréquences, séries cumulées
- 3 Caractéristiques statistiques**
 - **Moyenne**
 - étendue
 - Médiane
 - Quartiles

Définition :

- Soit x_i les p valeurs distinctes d'une série statistique et n_i les effectifs pour chaque valeur. La *moyenne* noté \bar{x} est donnée par :

Définition :

- Soit x_i les p valeurs distinctes d'une série statistique et n_i les effectifs pour chaque valeur. La *moyenne* noté \bar{x} est donnée par :

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

ce qui s'écrit aussi $\bar{x} = \frac{\sum_{i=1}^{i=p} n_i x_i}{\sum_{i=1}^{i=p} n_i}$

Définition :

- Soit x_i les p valeurs distinctes d'une série statistique et n_i les effectifs pour chaque valeur. La **moyenne** noté \bar{x} est donnée par :

$$\bar{x} = \frac{n_1x_1 + n_2x_2 + \dots + n_px_p}{n_1 + n_2 + \dots + n_p}$$

ce qui s'écrit aussi $\bar{x} = \frac{\sum_{i=1}^{i=p} n_i x_i}{\sum_{i=1}^{i=p} n_i}$

- Dans le cas d'une série où les effectifs des p valeurs x_i sont égaux à 1, la moyenne est donc :

Définition :

- Soit x_i les p valeurs distinctes d'une série statistique et n_i les effectifs pour chaque valeur. La **moyenne** noté \bar{x} est donnée par :

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

ce qui s'écrit aussi $\bar{x} = \frac{\sum_{i=1}^p n_i x_i}{\sum_{i=1}^p n_i}$

- Dans le cas d'une série où les effectifs des p valeurs x_i sont égaux à 1, la moyenne est donc :

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_p}{p}$$

Exemple :

Dans l'exemple du prix du pain, la moyenne est donnée par :

Exemple :

Dans l'exemple du prix du pain, la moyenne est donnée par :

$$\bar{x} = \frac{52 \times 0,82 + 40 \times 0,83 + 35 \times 0,84 + 28 \times 0,85 + 23 \times 0,86 + 2 \times 0,87}{52 + 40 + 35 + 28 + 23 + 2}$$

Exemple :

Dans l'exemple du prix du pain, la moyenne est donnée par :

$$\bar{x} = \frac{52 \times 0,82 + 40 \times 0,83 + 35 \times 0,84 + 28 \times 0,85 + 23 \times 0,86 + 2 \times 0,87}{52 + 40 + 35 + 28 + 23 + 2} \approx 0,84$$

Propriété :

On considère la *distribution des fréquences* d'une série statistique c'est à dire l'ensemble des fréquences des valeurs de la série :

x_1	x_2	\dots	x_p
f_1	f_2	\dots	f_p

Alors, la moyenne de la série statistique est donnée par :

Propriété :

On considère la *distribution des fréquences* d'une série statistique c'est à dire l'ensemble des fréquences des valeurs de la série :

x_1	x_2	\dots	x_p
f_1	f_2	\dots	f_p

Alors, la moyenne de la série statistique est donnée par :

$$\bar{x} = f_1 x_1 + f_2 x_2 + \dots + f_p x_p$$

Preuve :

On a :

Preuve :

On a :

$$\bar{X} = \frac{n_1 X_1}{N} + \frac{n_2 X_2}{N} + \dots + \frac{n_p X_p}{N} =$$

Preuve :

On a :

$$\bar{X} = \frac{n_1 X_1}{N} + \frac{n_2 X_2}{N} + \dots + \frac{n_p X_p}{N} = \frac{n_1}{N} X_1 + \frac{n_2}{N} X_2 + \dots + \frac{n_p}{N} X_p$$

Preuve :

On a :

$$\bar{X} = \frac{n_1 X_1}{N} + \frac{n_2 X_2}{N} + \dots + \frac{n_p X_p}{N} = \frac{n_1}{N} X_1 + \frac{n_2}{N} X_2 + \dots + \frac{n_p}{N} X_p$$

On obtient le résultat en remplaçant $\frac{n_i}{N}$ par f_i pour tout i allant de 1 à p .

Exemple :

On reprend l'exemple du prix du pain. On a :

$$f_1x_1 + f_2x_2 + \dots + f_px_p$$

Exemple :

On reprend l'exemple du prix du pain. On a :

$$f_1x_1 + f_2x_2 + \dots + f_px_p \approx 0,84$$

Plan

- 1 Vocabulaire et notations
- 2 Fréquences, séries cumulées
- 3 Caractéristiques statistiques**
 - Moyenne
 - étendue**
 - Médiane
 - Quartiles

Définition :

L'étendue d'une série statistique est

Définition :

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

Exemple :

Dans l'exemple du prix du pain, l'étendue est

Définition :

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

Exemple :

Dans l'exemple du prix du pain, l'étendue est $0,87 - 0,82 = 0,05$ soit 5 centimes d'euros.

Plan

- 1 Vocabulaire et notations
- 2 Fréquences, séries cumulées
- 3 Caractéristiques statistiques**
 - Moyenne
 - étendue
 - Médiane**
 - Quartiles

Définition :

La *médiane* est la valeur du caractère

Définition :

La *médiane* est la valeur du caractère qui sépare la série statistique *ordonnée* en deux sous séries de même effectif.

Méthode de détermination :

- Dans le cas d'un caractère discret d'effectif total N impair, la médiane est la valeur de rang

Méthode de détermination :

- Dans le cas d'un caractère discret d'effectif total N impair, la médiane est la valeur de rang $\frac{N+1}{2}$;

Méthode de détermination :

- Dans le cas d'un caractère discret d'effectif total N impair, la médiane est la valeur de rang $\frac{N+1}{2}$;
- dans le cas d'un caractère discret d'effectif total N pair, la médiane est la demi-somme des valeurs de rang

Méthode de détermination :

- Dans le cas d'un caractère discret d'effectif total N impair, la médiane est la valeur de rang $\frac{N+1}{2}$;
- dans le cas d'un caractère discret d'effectif total N pair, la médiane est la demi-somme des valeurs de rang $\frac{N}{2}$ et de la suivante.

Exemples :

Série avec *effectifs égaux à 1* : Soit la série statistique dont les valeurs sont 2 ; 2 ; 3 ; 5 ; 7 ; 7 ; 8 ; 9.

L'effectif total est

Exemples :

Série avec *effectifs égaux à 1* : Soit la série statistique dont les valeurs sont 2 ; 2 ; 3 ; 5 ; 7 ; 7 ; 8 ; 9.

L'effectif total est $N = 8$.

Exemples :

Série avec *effectifs égaux à 1* : Soit la série statistique dont les valeurs sont 2 ; 2 ; 3 ; 5 ; 7 ; 7 ; 8 ; 9.

L'effectif total est $N = 8$. Il est pair donc la médiane est la demi-somme entre les valeurs de rang

Exemples :

Série avec *effectifs égaux à 1* : Soit la série statistique dont les valeurs sont 2 ; 2 ; 3 ; 5 ; 7 ; 7 ; 8 ; 9.

L'effectif total est $N = 8$. Il est pair donc la médiane est la demi-somme entre les valeurs de rang 4 et 5.

Exemples :

Série avec *effectifs égaux à 1* : Soit la série statistique dont les valeurs sont 2 ; 2 ; 3 ; 5 ; 7 ; 7 ; 8 ; 9.

L'effectif total est $N = 8$. Il est pair donc la médiane est la demi-somme entre les valeurs de rang 4 et 5.

La médiane est donc

Exemples :

Série avec *effectifs égaux à 1* : Soit la série statistique dont les valeurs sont 2 ; 2 ; 3 ; 5 ; 7 ; 7 ; 8 ; 9.

L'effectif total est $N = 8$. Il est pair donc la médiane est la demi-somme entre les valeurs de rang 4 et 5.

La médiane est donc $\frac{5+7}{2} = 6$.

Exemples :

Série avec *effectifs non tous égaux à 1* : on fait une étude sur le prix de la baguette de pain dans différentes boulangeries. On obtient le tableau *d'effectifs* suivant :

Prix en euros (valeurs x_i)	0,82	0,83	0,84	0,85	0,86	0,87
Nombre de boulangeries (effectifs n_i)	52	40	35	28	23	2
Effectif cumulés croissants	52	92	127	155	178	180

Exemples :

Série avec *effectifs non tous égaux à 1* : on fait une étude sur le prix de la baguette de pain dans différentes boulangeries. On obtient le tableau *d'effectifs* suivant :

Prix en euros (valeurs x_i)	0,82	0,83	0,84	0,85	0,86	0,87
Nombre de boulangeries (effectifs n_i)	52	40	35	28	23	2
Effectif cumulés croissants	52	92	127	155	178	180

Il y a 6 *valeurs distinctes* mais

Exemples :

Série avec *effectifs non tous égaux à 1* : on fait une étude sur le prix de la baguette de pain dans différentes boulangeries. On obtient le tableau *d'effectifs* suivant :

Prix en euros (valeurs x_i)	0,82	0,83	0,84	0,85	0,86	0,87
Nombre de boulangeries (effectifs n_i)	52	40	35	28	23	2
Effectif cumulés croissants	52	92	127	155	178	180

Il y a 6 *valeurs distinctes* mais 180 *valeurs* au total. L'effectif total est

Exemples :

Série avec *effectifs non tous égaux à 1* : on fait une étude sur le prix de la baguette de pain dans différentes boulangeries. On obtient le tableau *d'effectifs* suivant :

Prix en euros (valeurs x_i)	0,82	0,83	0,84	0,85	0,86	0,87
Nombre de boulangeries (effectifs n_i)	52	40	35	28	23	2
Effectif cumulés croissants	52	92	127	155	178	180

Il y a 6 *valeurs distinctes* mais 180 *valeurs* au total. L'effectif total est $N = 180$, il est donc pair.

Exemples :

Série avec *effectifs non tous égaux à 1* : on fait une étude sur le prix de la baguette de pain dans différentes boulangeries. On obtient le tableau *d'effectifs* suivant :

Prix en euros (valeurs x_i)	0,82	0,83	0,84	0,85	0,86	0,87
Nombre de boulangeries (effectifs n_i)	52	40	35	28	23	2
Effectif cumulés croissants	52	92	127	155	178	180

Il y a 6 *valeurs distinctes* mais 180 *valeurs* au total. L'effectif total est $N = 180$, il est donc pair. La médiane est donc la demi-somme entre les valeurs de rang

Exemples :

Série avec *effectifs non tous égaux à 1* : on fait une étude sur le prix de la baguette de pain dans différentes boulangeries. On obtient le tableau *d'effectifs* suivant :

Prix en euros (valeurs x_i)	0,82	0,83	0,84	0,85	0,86	0,87
Nombre de boulangeries (effectifs n_i)	52	40	35	28	23	2
Effectif cumulés croissants	52	92	127	155	178	180

Il y a 6 *valeurs distinctes* mais 180 *valeurs* au total. L'effectif total est $N = 180$, il est donc pair. La médiane est donc la demi-somme entre les valeurs de rang 90 et 91.

Exemples :

Série avec *effectifs non tous égaux à 1* : on fait une étude sur le prix de la baguette de pain dans différentes boulangeries. On obtient le tableau *d'effectifs* suivant :

Prix en euros (valeurs x_i)	0,82	0,83	0,84	0,85	0,86	0,87
Nombre de boulangeries (effectifs n_i)	52	40	35	28	23	2
Effectif cumulés croissants	52	92	127	155	178	180

Il y a 6 *valeurs distinctes* mais 180 *valeurs* au total. L'effectif total est $N = 180$, il est donc pair. La médiane est donc la demi-somme entre les valeurs de rang 90 et 91.

D'après le tableau des effectifs cumulés, les valeurs de rang 90 et 91 sont

Exemples :

Série avec *effectifs non tous égaux à 1* : on fait une étude sur le prix de la baguette de pain dans différentes boulangeries. On obtient le tableau *d'effectifs* suivant :

Prix en euros (valeurs x_i)	0,82	0,83	0,84	0,85	0,86	0,87
Nombre de boulangeries (effectifs n_i)	52	40	35	28	23	2
Effectif cumulés croissants	52	92	127	155	178	180

Il y a 6 *valeurs distinctes* mais 180 *valeurs* au total. L'effectif total est $N = 180$, il est donc pair. La médiane est donc la demi-somme entre les valeurs de rang 90 et 91.

D'après le tableau des effectifs cumulés, les valeurs de rang 90 et 91 sont 0,83 et 0,83

Donc la médiane est 0,83.

Plan

- 1 Vocabulaire et notations
- 2 Fréquences, séries cumulées
- 3 Caractéristiques statistiques**
 - Moyenne
 - étendue
 - Médiane
 - Quartiles**

Définition :

- Le premier quartile noté Q_1 de la série statistique est

Définition :

- Le premier quartile noté Q_1 de la série statistique est la plus petite valeur telle qu'au moins 25 % des valeurs de la série lui soient inférieures ou égales ;

Définition :

- Le premier quartile noté Q_1 de la série statistique est la plus petite valeur telle qu'au moins 25 % des valeurs de la série lui soient inférieures ou égales ;
- le troisième quartile noté Q_3 de la série statistique est

Définition :

- Le premier quartile noté Q_1 de la série statistique est la plus petite valeur telle qu'au moins 25 % des valeurs de la série lui soient inférieures ou égales ;
- le troisième quartile noté Q_3 de la série statistique est la plus petite valeur telle qu'au moins 75 % des valeurs de la série lui soient inférieures ou égales.

Définition :

- Le premier quartile noté Q_1 de la série statistique est la plus petite valeur telle qu'au moins 25 % des valeurs de la série lui soient inférieures ou égales ;
- le troisième quartile noté Q_3 de la série statistique est la plus petite valeur telle qu'au moins 75 % des valeurs de la série lui soient inférieures ou égales.

Détermination pratique :

On suppose la série ordonnée dans l'ordre croissant des valeurs du caractère. Soit N l'effectif total.

- Si $\frac{N}{4}$ est un entier alors Q_1 est la valeur de rang $\frac{N}{4}$ et Q_3 est la valeur de rang $\frac{3N}{4}$;

Détermination pratique :

On suppose la série ordonnée dans l'ordre croissant des valeurs du caractère. Soit N l'effectif total.

- Si $\frac{N}{4}$ est un entier alors Q_1 est la valeur de rang $\frac{N}{4}$ et Q_3 est la valeur de rang $\frac{3N}{4}$;
- si $\frac{N}{4}$ n'est pas un entier, alors Q_1 est la valeur dont le rang suit le rang $\frac{N}{4}$ et Q_3 est la valeur dont le rang suit le rang $\frac{3N}{4}$.

Exemple :

Prix en euros (valeurs x_i)	0,82	0,83	0,84	0,85	0,86	0,87
Nombre de boulangeries (effectifs n_i)	52	40	35	28	23	2
Effectif cumulés croissants	52	92	127	155	178	180

L'effectif total N est 180.

Exemple :

Prix en euros (valeurs x_i)	0,82	0,83	0,84	0,85	0,86	0,87
Nombre de boulangeries (effectifs n_i)	52	40	35	28	23	2
Effectif cumulés croissants	52	92	127	155	178	180

L'effectif total N est 180.

$\frac{N}{4} = 45$ donc le premier quartile est la valeur de rang 45 soit

Exemple :

Prix en euros (valeurs x_i)	0,82	0,83	0,84	0,85	0,86	0,87
Nombre de boulangeries (effectifs n_i)	52	40	35	28	23	2
Effectif cumulés croissants	52	92	127	155	178	180

L'effectif total N est 180.

$\frac{N}{4} = 45$ donc le premier quartile est la valeur de rang 45 soit
 $Q_1 = 0,82$

Exemple :

Prix en euros (valeurs x_i)	0,82	0,83	0,84	0,85	0,86	0,87
Nombre de boulangeries (effectifs n_i)	52	40	35	28	23	2
Effectif cumulés croissants	52	92	127	155	178	180

L'effectif total N est 180.

$\frac{N}{4} = 45$ donc le premier quartile est la valeur de rang 45 soit

$Q_1 = 0,82$.

$\frac{3N}{4} = 135$ donc le troisième quartile est la valeur de rang 135
soit

Exemple :

Prix en euros (valeurs x_i)	0,82	0,83	0,84	0,85	0,86	0,87
Nombre de boulangeries (effectifs n_i)	52	40	35	28	23	2
Effectif cumulés croissants	52	92	127	155	178	180

L'effectif total N est 180.

$\frac{N}{4} = 45$ donc le premier quartile est la valeur de rang 45 soit
 $Q_1 = 0,82$.

$\frac{3N}{4} = 135$ donc le troisième quartile est la valeur de rang 135
 soit $Q_3 = 0,85$.