

Chapitre 17

Statistique descriptive

17.1 Fréquences

17.1.1 Effectifs, fréquences et pourcentages

Définition 17.1.

1. *L'effectif total N d'une population P est le nombre d'individus de cette population.*
2. *L'effectif n d'une sous-population S est le nombre d'individus de cette sous-population.*
3. *La fréquence f d'une sous-population est donnée par la formule :*

$$f = \frac{n}{N}.$$

Remarques :

- Une fréquence est un rapport toujours compris entre 0 et 1.
- Le nombre f est aussi appelé *proportion* de S dans P , ou encore *taux* de S par rapport à P .

Exemple : On considère l'équipage de Luffy de One Piece ; parmi les 10 membres de l'équipage, 4 ont des fruits du démon. La fréquence de détenteur de fruit du démon au sein de l'équipage est donc :

$$f = \frac{4}{10} = 0,4.$$

17.1.2 Effectifs et fréquences cumulés

On considère une population répartie en sous-population (ou valeur) x_i d'effectif n_i selon le tableau suivant :

Sous-population	x_1	x_2	...	x_p
Effectif	n_1	n_2	...	n_p
Fréquence	f_1	f_2	...	f_p

Définition 17.2.

1. L'**effectif cumulé croissant** E_{cc} de la valeur x_i est la somme des effectifs de toutes les valeurs inférieures ou égales à x_i :

$$E_{cc} = n_1 + \dots + n_i.$$

2. La **fréquence cumulée croissante** F_{cc} de la valeur x_i est la somme des fréquences de toutes les valeurs inférieures ou égales à x_i :

$$F_{cc} = f_1 + \dots + f_i.$$

Exemple : On considère les résultats des 20 jets de dés 8 successifs donnés dans le tableau ci-dessous.

Face obtenue	1	2	3	4	5	6	7	8
Effectif	2	3	1	2	4	3	1	4
Effectif cumulé croissant	2	5	6	8	12	15	16	20
Fréquence	0,1	0,15	0,05	0,1	0,2	0,15	0,05	0,2
Fréquence cumulée croissante	0,1	0,25	0,3	0,4	0,6	0,75	0,8	1

17.2 Caractéristiques d'une série statistique

17.2.1 Moyenne

Définition 17.3. La **moyenne** – notée \bar{x} – est le nombre

$$\bar{x} = \frac{n_1x_1 + \dots + n_px_p}{N}.$$

★ Vidéo.

Exemple On reprend l'exemple des lancés de dés, la moyenne des lancés est

$$\bar{x} = \frac{1 \times 2 + 2 \times 3 + 3 \times 1 + 4 \times 2 + 5 \times 4 + 6 \times 3 + 7 \times 1 + 8 \times 4}{20} = 4,8.$$

Propriété 17.1. En notant f_i la fréquence de la sous-population x_i , on a

$$\bar{x} = f_1x_1 + \dots + f_px_p.$$

17.2.2 Dispersion d'une série statistique

Définition 17.4. On considère une série statistique dont les valeurs sont **ordonnées par ordre croissant**.

- Si N est **impair** : $N = 2n + 1$, alors la **médiane** est la valeur du terme de rang $n + 1$ de la série.
- Si N est **pair** : $N = 2n$, alors la **médiane** est la moyenne des termes de rang n et $n + 1$ de la série.

Exemple :

1. On considère la suite de valeurs **ordonnées** suivantes :

$$1; 2; 3; 3; 5; 7; 8; 8; 9; 10; 16.$$

Il y a 11 termes dans cette série. La médiane est donc le 6ème terme de la série : 7.

2. On considère la même suite de valeurs **ordonnées** que ci-dessus à laquelle on a ajouté le terme 17 :

$$1; 2; 3; 3; 5; 7; 8; 8; 9; 10; 16, 17.$$

Il y a 12 termes dans cette série. La médiane est donc la moyenne des 6ème et 7ème termes de la série :

$$m = \frac{7 + 8}{2} = 7,5.$$

Définition 17.5. Soit une série statistique rangée par ordre croissant.

— Le **premier quartile** – noté Q_1 – est la plus petite valeur de la série telle que au moins 25% des valeurs soient inférieures ou égales à Q_1 .

— Le **troisième quartile** – noté Q_3 – est la plus petite valeur de la série telle que au moins 75% des valeurs soient inférieures ou égales à Q_3 .

Remarques : Le rang du **premier** – respectivement **troisième** – **quartile** d’une série statistique d’effectif total N est le plus petit entier supérieur ou égal à $\frac{N}{4}$ – respectivement $\frac{3N}{4}$.

Exemple : On reprend l’exemple 1. précédent. Le premier quartile est $Q_1 = 3$ et le troisième est $Q_3 = 9$.

Définition 17.6.

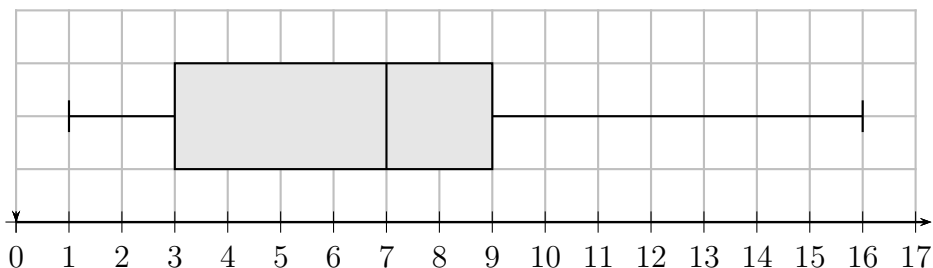
— **L’écart interquartile** e_I d’une série statistique est la différence entre le troisième quartile et le premier quartile de cette série :

$$e_I = Q_3 - Q_1.$$

— **L’étendue** e d’une série statistique est la différence entre la plus grande valeur x_p et la plus petite valeur x_1 de cette série :

$$e = x_p - x_1.$$

Graphiquement, cette répartition est représentée par un diagramme dit « en boîte à moustaches » dont on peut voir un exemple ci-dessus : il s’agit du diagramme en boîtes associé la série 1. de l’exemple précédent.



★ Vidéo (médiane); vidéo (quartile 1); vidéo (quartile 2); vidéo (étendue).

17.3 Attendus et savoir-faire

- Calculer une fréquence. Remplir un tableau d'effectifs et de fréquences.
- Calculer une moyenne à l'aide d'effectifs ou de fréquences.
- Déterminer médiane et quartiles d'une série statistique.
- Lire et faire un diagramme en boîte.

17.4 Exercices

17.4.1 Démarrage

Exercice 17.1. On considère 20 lancers de dé 8 successifs.

1. Reproduire et compléter le tableau ci-dessous.

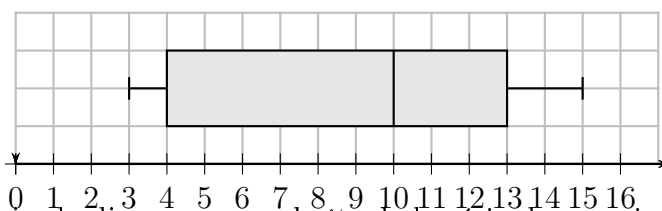
Face obtenue	1	2	3	4	5	6	7	8
Effectif	4	2	4	3	1	2	1	3
Effectif cumulé croissant								
Fréquence								
Fréquence cumulée croissante								

2. Quelle est la fréquence de lancers dont le résultat est strictement inférieur à 3 ? Supérieur ou égal à 5 ?
3. Quelle est la fréquence de lancers dont le résultat est compris entre 3 et 6 ?

Exercice 17.2. Calculer la moyenne de la série statistique de l'exercice précédent.

Exercice 17.3. Calculer la médiane et les quartiles de la série statistique du premier exercice.

Exercice 17.4. Lire sur le diagramme en boîte ci-dessous la médiane, les quartiles, le minimum et le maximum de la série statistique associée.



Exercice 17.5. Construire le diagramme en boîte de la série du premier exercice.

17.4.2 Approfondissement

Exercice 17.6. Lors d'une évaluation sur 10, la classe 1 a les résultats suivants :

Notes	0	1	2	3	4	5	6	7	8	9	10
Effectifs	0	2	4	3	5	8	6	2	3	1	0

- Reproduire le tableau ci-dessous et y ajouter une ligne donnant les effectifs cumulés croissants.
- Donner la médiane et les quartiles de cette série.
- Quelles sont l'étendue et l'intervalle interquartile de cette série ?
- Calculer la moyenne de cette série.
- On compare la classe 1 à la classe 2 dont on a les données suivantes :

— Moyenne : 6,5.	— 1 ^{er} quartile : 2.	— min : 0.
— Médiane : 5.	— 3 ^e quartile : 7.	— max : 10.

 - Réaliser les diagrammes en boîte de ces séries.
 - Comparer les deux classes.

Exercice 17.7. [Économie] La fréquence des salaires (nets) au sein de l'entreprise nationale Tchoutchou est donnée par le tableau suivant :

Intervalle de Salaire (en euro)	[1500 ; 2000[[2000 ; 2500[[2500 ; 3500[[3500 ; 5000[[5000 ; 100000]
Fréquence	0,36	0,41	0,2	0,02	0,01

- Reproduire le tableau ci-dessous et y ajouter une ligne donnant les fréquences cumulées croissantes.
- Donner la médiane et les quartiles de cette série (on pourra prendre le milieu des intervalles). Que signifient ces valeurs ?
- Réaliser le diagramme en boîte de cette série.
- Calculer la moyenne de cette série (on pourra prendre à nouveau le milieu des intervalles). Que signifie cette valeur ?
- Une personne de votre entourage qualifie les salariés de l'entreprise Tchoutchou de « privilégiés » car leur salaire moyen est supérieur à 2500€. Qu'en pensez-vous ?
- Critiquer la catégorie de salaire [5000 ; 100000[.

Exercice 17.8. Construire une série statistique de moyenne 100 et dont 90% des valeurs sont nulles. Qu'en déduisez-vous ?

Exercice 17.9. [Agriculture et Santé] Un producteur de fruits et légumes – afin d’obtenir un label – fait réaliser par un laboratoire une étude sur ses produits afin de déterminer les quantités de pesticides présents sur ceux-ci. Le dosage est le rapport de la masse de pesticide détecté (en μg) sur la masse du fruit ou légume (en kg). Le label est attribué aux conditions suivantes :

- le dosage moyen est inférieur à $0,8\mu\text{g} \cdot \text{kg}^{-1}$;
- le dosage médian est inférieur à $1\mu\text{g} \cdot \text{kg}^{-1}$;
- le dosage est inférieur à $1,3\mu\text{g} \cdot \text{kg}^{-1}$ pour au moins 75% des fruits et légumes ;
- le dosage dépasse $1,9\mu\text{g} \cdot \text{kg}^{-1}$ pour au maximum 10% des fruits et légumes.

Les résultats du laboratoires sont donnés dans le tableau ci-dessous.

Dosage détecté (en $\mu\text{g} \cdot \text{kg}^{-1}$)	0,1	0,4	0,7	1	1,3	1,6	1,9	2,2
Nombre de produits ayant au moins cette dose	300	230	145	99	63	40	23	10

Quelle sera la décision du laboratoire ?