Chapitre 17

Statistiques à une variable

17.1 Fréquences

17.1.1 Effectifs et fréquences

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 :

— Le nombre f est aussi appelé *proportion* de S dans P, ou encore taux de S par rapport à P.

— Une fréquence étant une proportion, elle est toujours comprise entre 0 et 1.

Exemple : Parmi les 150 premiers pokémons, 12 sont de types feu. La proportion de pokémons feu parmi les 150 premiers est donc

$$f = \frac{\text{effectif pokémons feu}}{\text{effectif total pokémons}} = \frac{12}{150} = 0,08 = 8\%.$$

17.1.2 Effectifs et fréquences cumulés

On considère une population repartie 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 *:*

$$E_{cc}(x_i) = n_1 + ... + n_i.$$

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 *:*

$$F_{cc}(x_i) = f_1 + ... + f_i.$$

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

Exercices : 17.1 ; 17.10.

17.2 Moyennes

17.2.1 Moyenne simple

Définition 17.3. *On considère une série de valeurs* x_i *d'effectifs respectifs* n_i *et d'effectif total* N*. La* ***moyenne*** *de cette série – notée* $\bar{x}$ *– est le nombre*

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

Valeur	x_1	x_2	...	x_p
Effectif	n_1	n_2	...	n_p

Remarque : Cette moyenne est en réalité la moyenne arithmétique. Il en existe d'autres comme les moyennes :

Géométrique : $\bar{x} = (x_1^{n_1} \dots x_p^{n_p})^{\frac{1}{N}}$;

Harmonique : $\dfrac{N}{\bar{x}} = \dfrac{n_1}{x_1} + \cdots + \dfrac{n_p}{x_p}$.

Le choix d'une moyenne par rapport à une autre dépend de la situation. Par exemple, la moyenne géométrique convient très bien aux situations d'évolutions successives et permet de calculer des coefficients et taux d'évolutions moyens.

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

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

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

$$\bar{x} = f_1 x_1 + ... + f_p x_p.$$

17.2.2 Moyenne pondérée

Définition 17.4. *On considère une série de valeurs x_i avec les poids (ou coefficients) respectifs α_i. La* ***moyenne pondérée*** *de cette série – notée $\bar{x}$ – est le nombre*

$$\bar{x} = \frac{\alpha_1 x_1 + ... + \alpha_p x_p}{\alpha_1 + \cdots + \alpha_p}.$$

Valeur	x_1	x_2	...	x_p
Poids	α_1	α_2	...	α_p

Remarques :

- La moyenne exprimée à l'aide des effectifs est un cas particulier de la moyenne pondérée. En effet, il s'agit du cas où le poids est égal à l'effectif : $\alpha_i = n_i$.
- À nouveau, il s'agit d'une moyenne pondérée arithmétique ; il existe également des moyennes pondérées géométrique harmonique, etc.

Exemple : Hermione vient de recevoir ses notes de potions avec les coefficients associés, elles sont données dans le tableau ci-contre.

Note / 20	20	19	18	20
Coefficient	1	2	0,5	3

Sa moyenne pondérée est donc

$$\frac{1 \times 20 + 2 \times 19 + 0,5 \times 18 + 3 \times 20}{1 + 2 + 0,5 + 3} = \frac{127}{6,5} \simeq 19,5.$$

17.2.3 Moyenne d'une série avec valeurs regroupées en classe

Définition 17.5. *On considère une série de valeurs réparties en classes $[x_i ; x_{i+1}[$ d'effectifs respectifs n_i et d'effectif total N.*

Classe	$[x_1 ; x_2[$	$[x_2 ; x_3[$	...	$[x_p ; x_{p+1}]$
Effectif	n_1	n_2	...	n_p

- *On appelle* ***centre*** *de la classe $[x_i ; x_{i+1}[$ le réel $c_i = \dfrac{x_i + x_{i+1}}{2}$.*
- *La* ***moyenne*** *de cette série – notée $\bar{x}$ – est le nombre*

$$\bar{x} = \frac{n_1 c_1 + ... + n_p c_p}{N}.$$

Exemple : Au sein du cours de métamorphose, les temps d'apprentissage d'une nouvelle technique par les élèves sont variés. Le tableau ci-dessous donnent les effectifs par intervalles de temps mis pour maîtriser la technique.

Intervalle de temps	[0 ; 5[	[5 ; 10[	[10 ; 15[	[15 ; 20[
Effectif	1	3	10	1

Les centres respectifs des classes ci-dessus sont : $2,5\,;\ 7,5\,;\ 12,5\,;\ 17,5$. La moyenne de cette série est

$$\frac{1\times 2,5+3\times 7,5+10\times 12,5+1\times 17,5}{1+3+10+1}=\frac{167,5}{15}\simeq 11.$$

Le temps moyen d'apprentissage est donc de 11 minutes.

17.2.4 Propriétés de la moyenne

Proposition 17.2. ***[Linéarité de la moyenne]*** *Soient une série statistique de moyenne $\bar{x}$, a et b deux réels.*

1. *Si l'on multiple chaque valeur de la série par a, alors on obtient une nouvelle série de moyenne $a\times\bar{x}$.*
2. *Si l'on ajoute b à chaque valeur de la série, alors on obtient une nouvelle série de moyenne $b+\bar{x}$.*

Exemple : Fred et George vendent leurs farces et attrapes avec un prix moyen hors taxes (HT) de 3 gallions. Suite à une phase d'inflation, ils décident de monter les prix tous leurs articles de 1 gallion. D'après la propriété précédente, le nouveau prix moyen hors taxes est de $3+1=4$ gallions. Par ailleurs, le ministère de la Magie impose une taxe de 20% sur tous les articles de farces et attrapes. Le nouveau prix moyen, toutes taxes comprises, d'un article de Fred et George est donc de $1,2\times 4=4,8$ gallions.

Remarque : La moyenne est très sensible aux valeurs extrêmes, à leur ajout ou leur retrait. Elle peut donc fournir un résultat peu représentatif de la série considérée ; elle est un mauvais indicateur de dispersion.

Exemple : Harry et Ron eu respectivement 10 et 8 sur 20 à leur dernière évaluation de potions. À eux deux, ils ont une moyenne de $\dfrac{10+8}{2}=9$. Ajoutons à présent le note de Hermione, cette dernière a eu, comme à son habitude, 20 sur 20. Avec elle, la moyenne est de $\dfrac{10+8+20}{3}\simeq 12,7$. Cette moyenne n'est ni représentative du sous-groupe Harry - Ron, ni de Hermione.

Exercices : 17.2 à 17.5 ; 17.11 et 17.12.

17.3 Indicateurs de dispersion d'une série statistique

17.3.1 Médianes et quartiles

Définition 17.6. *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 6^e^ terme de la série : 7.

2. On considère la même suite de valeurs **ordonnées** que ci-dessus à laquelle on a ajouté le terme 100 :

$$1\,;\,2\,;\,3\,;\,3\,;\,5\,;\,7\,;\,8\,;\,8\,;\,9\,;\,10\,;\,16\,;\,100.$$

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

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

Remarque : On remarque que contrairement à la moyenne, l'ajout d'une valeur extrême n'a que très peu d'influence sur la médiane.

Définition 17.7. *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 17.8.

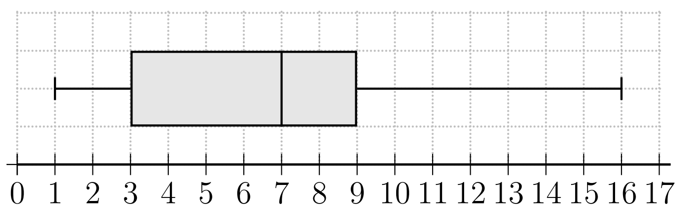
— ***L'écart interquartile*** *e_I d'une série statistique est la différence entre le troisième quartile et le premier quartile de cette série :*

$$e_I = Q_3 - Q_1.$$

— ***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.



Exercices : 17.6 à 17.9 ; 17.13 et 17.14.

17.3.2 Écart-type

Définition 17.9. *L'**écart-type** σ d'une série de valeurs x_i d'effectifs respectifs n_i et d'effectif total N est la distance moyenne entre les valeurs x_i et la moyenne $\bar{x}$ de la série :*

$$\sigma = \sqrt{\frac{n_1(x_1 - \bar{x})^2 + \cdots + n_p(x_p - \bar{x})^2}{N}}.$$

Remarque : Plus l'écart-type est grand, plus la série est dispersée ; plus l'écart-type est petit, plus la série est resserrée.

Exemple : Une série avec une moyenne de 100 et un écart-type de 20 est plus dispersée qu'une série avec une moyenne de 100 et un écart-type de 10 .

17.4 Capacités attendues

— Calculer une moyenne pondérée à l'aide d'effectifs ou de fréquences.

— Déterminer les indicateurs de dispersion d'une série statistique à l'aide de calculs, tableaux ou graphiques.

— Décrire verbalement les différences entre deux séries statistiques, en s'appuyant sur des indicateurs ou sur des représentations graphiques données.

17.5 Exercices

17.5.1 Progresser

Effectifs et fréquences

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

1. 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 ?

Moyennes

Exercice 17.2. Calculer la moyenne de la série statistique de l'exercice 17.1.

Exercice 17.3. Harry vient de recevoir ses notes de métamorphose avec les coefficients associés, elles sont données dans le tableau ci-contre. Calculer sa moyenne.

Note / 20	15	12	10	14	17
Coefficient	1	0,5	2	0,5	4

Exercice 17.4. Construire une série statistique de moyenne 100 et dont 90% des valeurs sont nulles. Qu'en déduisez-vous ?

Exercice 17.5. Le tableau ci-dessous donne la répartition des consommations annuelles d'électricité en kWh d'appartements au sein d'un immeuble.

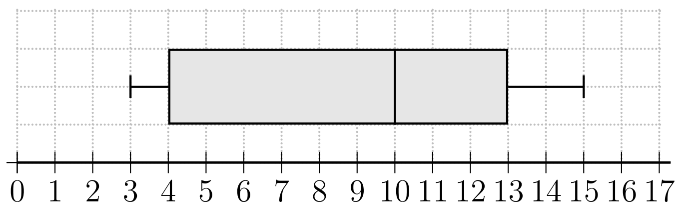
Consommation	[3 000 ; 4 000[	[4 000 ; 5 000[	[5 000 ; 6 000[	[6 000 ; 7 000[	[7 000 ; 8 000[
Appartements	5	7	10	8	5

1. Calculer la consommation moyenne d'un appartement.
2. La pose d'un nouvel appareil au sein de chaque appartement de l'immeuble baisse la consommation de chacun d'entre eux de 10% en contrepartie d'une hausse de 100 kWh pour alimenter l'appareil. Quelle est la nouvelle consommation moyenne des appartements de l'immeuble ?

Médiane et quartiles

Exercice 17.6. Calculer la médiane et les quartiles de la série statistique puis construire le diagramme en boîte de l'exercice 17.1.

Exercice 17.7. Lire sur le diagramme en boîte ci-dessous la médiane, les quartiles, le minimum et le maximum de la série statistique associée.



Exercice 17.8. [Économie] La fréquence des salaires (nets) par intervalle (IS) au sein de la Société Nationale du Tchoutchou est donnée par le tableau suivant :

IS (€)	[1 500 ; 2 000[	[2 000 ; 2 500[	[2 500 ; 3 500[	[3 500 ; 5 000[	[5 000 ; 100 000]
Fréquence	0,36	0,41	0,2	0,02	0,01
FCC					

1. Compléter le tableau ci-dessus en donnant les fréquences cumulées croissantes (FCC).
2. Donner la médiane et les quartiles de cette série (on pourra prendre le milieu des intervalles). Que signifient ces valeurs ?
3. Réaliser le diagramme en boîte de cette série.
4. Calculer la moyenne de cette série. Que signifie cette valeur ?
5. Une personne de votre entourage qualifie les salariés de la Société Nationale du Tchoutchou de « privilégiés » car leur salaire moyen est supérieur à 2 500€. Qu'en pensez-vous ?
6. Critiquer la catégorie de salaire [5 000 ; 100 000]. Calculer le salaire moyen sans cette catégorie, 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 $0,8\mu\mathrm{g}\cdot\mathrm{kg}^{-1}$;

— le dosage médian est inférieur à $1\mu\mathrm{g}\cdot\mathrm{kg}^{-1}$;

— le dosage est inférieur à $1,3\mu\mathrm{g}\cdot\mathrm{kg}^{-1}$ pour au moins 75% des fruits et légumes ;

— le dosage dépasse $1,9\mu\mathrm{g}\cdot\mathrm{kg}^{-1}$ pour au maximum 10% des fruits et légumes.

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

Dosage détecté (en $\mu\mathrm{g}\cdot\mathrm{kg}^{-1}$)	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 ?

17.5.2 S'entraîner

Effectifs et fréquences

Exercice 17.10. Lors d'une évaluation sur 10, la classe 1 a les résultats donnés dans le tableau ci-dessous.

Note	0	1	2	3	4	5	6	7	8	9	10
Effectif	0	2	4	3	5	8	6	2	3	1	0
Effectif cumulé croissant											
Fréquence											
Fréquence cumulée croissante											

1. Compléter le tableau ci-dessus.
2. Quelle est la fréquence de notes strictement inférieures à 4 ? Supérieures ou égales à 7 ?
3. Quelle est la fréquence de notes comprises entre 3 et 7 ?

Moyennes

Exercice 17.11. Ron vient de recevoir ses notes de sortilège avec les coefficients associés, elles sont données dans le tableau ci-contre. Calculer sa moyenne.

Note / 20	10	13	8	9	10
Coefficient	0,5	1	1	0,5	2

Exercice 17.12. Le tableau ci-dessous donne la répartition des âges des membres d'une assemblée citoyenne qui vient d'être tirée au sort.

Âge	$[20\,;30[$	$[30\,;40[$	$[40\,;50[$	$[50\,;60[$	$[60\,;70[$	$[70\,;80[$
Membres	4	6	8	7	10	5

1. Calculer l'âge moyen de l'assemblée.
2. L'assemblée est renouvelée tous les trois ans, quel sera l'âge moyen de l'assemblée au moment du renouvellement ?

Médiane et quartiles

Exercice 17.13. On reprend la série de notes à une évaluation de l'exercice 17.10.

1. Donner la médiane et les duquartiles de la série de notes.
2. Quelles sont l'étendue et l'intervalle interquartile de cette série ?
3. Calculer la moyenne de cette série.
4. On compare la classe 1 à la classe 2 dont on a les données suivantes :

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

   (a) Réaliser les diagrammes en boîte de ces séries.
   (b) Comparer les deux classes.

5. Le professeur décide de mettre les notes sur 20, quelle sera la moyenne des groupes ?

Exercice 17.14. Dans chaque cas, déterminer la ou les bonnes réponses.

1. Je suis un indicateur très peu sensible aux valeurs extrêmes de la série. Je suis...

   (a) la médiane   (b) la moyenne   (c) l'étendue   (d) le minimum

2. Je suis un indicateur qui rend compte de la dispersion de la série. Je suis...

   (a) la médiane   (b) la moyenne   (c) l'étendue   (d) l'écart-type

3. Deux classes ont effectué un devoir commun. Dans la première, comptant 35 élèves, la moyenne est 11,5. Dans la seconde, comptant 24 élèves, la moyenne est 9,5. La moyenne globale est :

   (a) 10,5   (b) 10,7   (c) On ne peut pas savoir

4. En France, le salaire médian est inférieur au salaire moyen.

   (a) 50% des salariés français gagnent plus que la moyenne.
   (b) 50% des salariés français gagnent moins que la moyenne.