

3C Théorie de mathématiques

Claude Chevalley

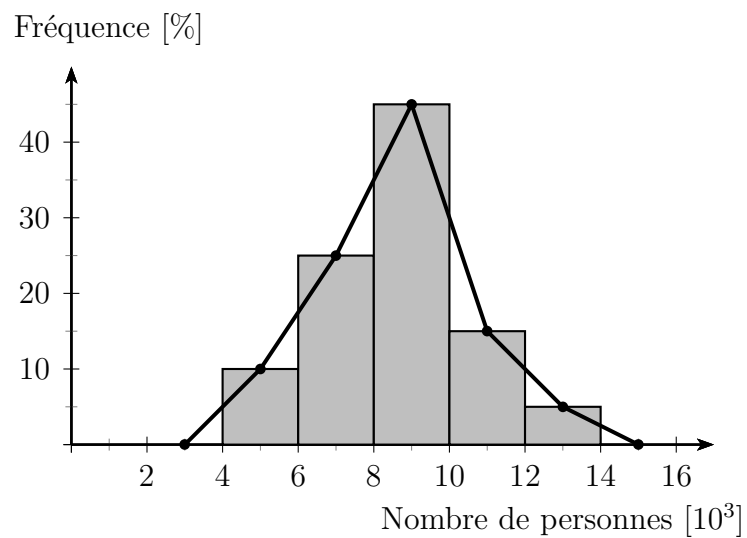


Table des matières

1. Géométrie dans l'espace.....	p.1
2. Graphiques	p.3
3. Statistiques descriptives.....	p.5
4. Probabilités.....	p.22
5. Inférence statistique	p.25

1 Géométrie dans l'espace

1.1 Tracés de solides et calculs de longueurs, d'aires et de volumes

Un **polyèdre** est un solide de l'espace délimité par des polygones appelés du polyèdre. Les côtés des polygones sont appelés les du polyèdre et les sommets des polygones sont appelés les du polyèdre.

Modèle 1. Résoudre l'exercice 1.1.

- a) Esquisse :
- b) Longueur arête = ...
- c) Aire d'une face triangulaire = ... ; Aire d'une face carrée = ...
Aire totale = ...
- d) Volume du prisme droit = ...

Modèle 2. Résoudre l'exercice 1.7.

- a) Esquisse :
- b) Nombre de sommets $s = \dots$; nombre d'arêtes $a = \dots$; nombre de faces $f = \dots$
- c) Aire totale = ...
- d) Volume de cette pyramide régulière à base carrée = ...

Remarque :

Nom du solide	s	f	a	$s + f - a$
Tétraèdre (ex.1.6)				
Prisme droit à base pentagonale (ex.1.8)				
Pyramide à base carrée (ex.1.7)				
Pyramide à base hexagonale (ex.1.9)				
Cube				
Hexaèdre				
Octaèdre				

Relation d'Euler :

$$s + f - a = \dots$$

1.2 Solides de révolution

Modèle 3. Résoudre l'exercice 1.10.

a) Esquisse :

Le solide de révolution est un

Son volume $V = \dots$

b) Esquisse :

Le solide de révolution est un

Son volume $V = \dots$

c) Esquisse :

Le solide de révolution est un

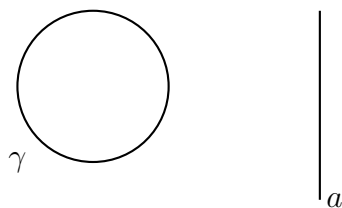
Son volume $V = \dots$

Remarque sur l'exercice 1.11 :

Soit un disque limité par le cercle γ .

On fait tourner cette figure autour d'un axe a .

Cette figure va ainsi générer un solide de révolution, ici un tore (une chambre à air).



2 Graphiques

2.1 Fonction polynomiale

2.1.1 Lecture et interprétation du graphe d'une fonction polynomiale

Modèle 4. Résoudre l'exercice 2.3.

Méthode par graphique :

- 1) Déterminer l'**ordonnée à l'origine** h de f avec $f(0)$: c'est l'intersection du graphe de f et l'axe Oy ;
- 2) Déterminer les **zéros** de f s'ils existent en résolvant $f(x) = 0$: c'est l'intersection du graphe de f et l'axe Ox ;
- 3) Déterminer les coordonnées du **maximum absolu** $M(x_1; y_1)$ de f en résolvant $f(x_1) = y_1 \mid f(x) < y_1$ si $x \neq x_1$;
- 4) Déterminer les coordonnées du **minimum absolu** $m(x_2; y_2)$ de f en résolvant $f(x_2) = y_2 \mid f(x) > y_2$ si $x \neq x_2$;

• Lecture :

- a) L'ordonnée à l'origine : $h = f(0) \cong \dots \Rightarrow H(0; \dots)$
- b) $f(2) \cong \dots$
- c) L'ensemble des zéros de f : $Z_f = \{ \dots ; \dots ; \dots ; \dots \}$
- d) $M(\dots ; \dots)$
- e) $m(\dots ; \dots)$

• Interprétation :

- f) ...
- g) ...
- h) ...
- i) ...

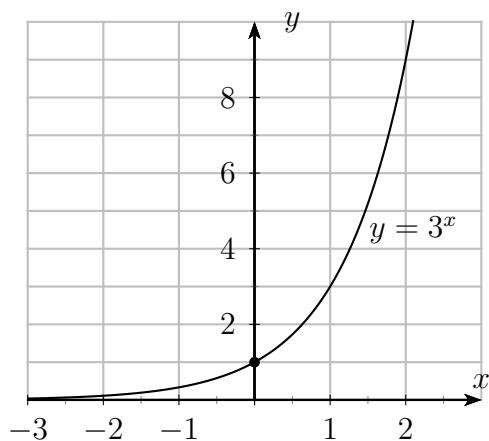
2.2 Fonction exponentielle

2.2.1 Lecture et interprétation du graphe d'une fonction exponentielle

Modèle 5. On donne deux fonctions f et g définies par :

$$f(x) = 3^x \quad ; \quad g(x) = 3^{-x}$$

Graphe de f :



• Lecture :

a) L'ordonnée à l'origine : $h = f(0) = \dots \Rightarrow H(0; \dots)$

b) $f(-1) = \dots$; $f(1) = \dots$; $f(2) = \dots$

c) Zéros de f : $Z_f = \dots$

• Représentation du graphe de g :

d) L'ordonnée à l'origine : $h = g(0) = \dots \Rightarrow H(0; \dots)$

e) $g(-2) = \dots$; $g(-1) = \dots$; $g(1) = \dots$

f) Zéros de g : $Z_g = \dots$

Remarque : ...

3 Statistiques descriptives

La statistique est la science qui traite des éléments suivants :

- Récolte de données ;
- Classification des données ;
- Représentation et analyse des données (**statistiques descriptives**) ;
- Interprétations, conclusions et prévisions pouvant être tirées de l'analyse des données (**inférence statistique** au chapitre 5).

Aujourd'hui la statistique a pris une grande place grâce aux nouvelles techniques et à la puissance des ordinateurs. Géographie, médecine, sciences humaines, sciences économiques, biologie, politique, etc... : aucun domaine n'est épargné.

3.1 Caractéristiques d'un ensemble d'individus

Vocabulaire statistique :

- **Population** : ensemble de toutes les personnes, de tous les objets ou de tous les faits sur lesquels porte l'étude.
- **Individu** : élément de la population.
- **Sondage** : étude réalisée sur un **échantillon** de la population, soit un sous-ensemble de la population (lorsque la population est trop grande pour être étudiée dans son ensemble).
- **Variable statistique** (v.s.) : caractéristique étudiée dans la population.
- **Modalités ou valeurs** : différents états ou valeurs prises par la variable statistique.

Notation : on emploie une lettre majuscule (par exemple X) pour désigner une v.s. et une lettre minuscule (x_i) pour désigner une de ses valeurs.

Si l'échantillon est choisi au hasard, que sa taille est suffisamment grande et qu'il peut être considéré comme représentatif, on peut généraliser certains résultats obtenus à l'ensemble de la population.

Une variable statistique est dite **quantitative** si les valeurs qu'elle peut prendre sont des nombres. Dans le cas contraire, elle est **qualitative**, et ses valeurs s'appellent des **modalités** (ou des **catégories**).

Si les valeurs que peut prendre une variable statistique quantitative sont isolées les unes des autres, on dit que la variable est **discrète**.

Si les valeurs constituent des intervalles de nombres, on dit qu'elle est **continue**.

Modèle 6. Dans chaque situation exposée ci-dessous :

- 1) Décrire la population étudiée;
- 2) Décrire l'échantillon;
- 3) Nommer la variable étudiée;
- 4) Donner le type de variable étudiée.
- 5) Décrire l'ensemble des modalités ou des valeurs de la variable.

Exemples :

- a) On demande à 200 passagers pris au hasard dans un aéroport de donner leur nationalité.
 - 1) Population : ...
 - 2) Echantillon : ...
 - 3) Variable statistique : ...
 - 4) Type de variable : ...
 - 5) Modalités ou valeurs : ...
- b) Durant le mois d'avril, on mesure la température maximale de la journée à La tour-de-Peilz.
 - 1) Population : ...
 - 2) Echantillon : ...
 - 3) Variable statistique : ...
 - 4) Type de variable : ...
 - 5) Modalités ou valeurs : ...
- c) On demande à 5'000 familles résidant en Suisse le nombre d'enfants de leur foyer.
 - 1) Population : ...
 - 2) Echantillon : ...
 - 3) Variable statistique : ...
 - 4) Type de variable : ...
 - 5) Modalités ou valeurs : ...

3.2 Variables qualitative et quantitative discrète

Notation : on désigne la taille de l'échantillon par n et les valeurs prises par la v.s. X dans l'échantillon par x_1, x_2, \dots, x_n .

3.2.1 Tableau de distribution d'une variable qualitative

Une fois les données récoltées, on les regroupe par modalité dans un tableau de distribution.

Modèle 7. On a demandé à tous les élèves d'une classe quelle était leur matière préférée parmi les matières suivantes : français, anglais, maths et musique.

musique	français	français	anglais	français
anglais	musique	maths	musique	musique
musique	musique	musique	français	français
français	anglais	musique	anglais	maths

Tableau de distribution :

Répartition des selon

Matière	Effectif d'élèves = n_i	Fréquence = f_i
Total		

Remarques :

1. La somme des effectifs est toujours égale au nombre N d'individus de la population :

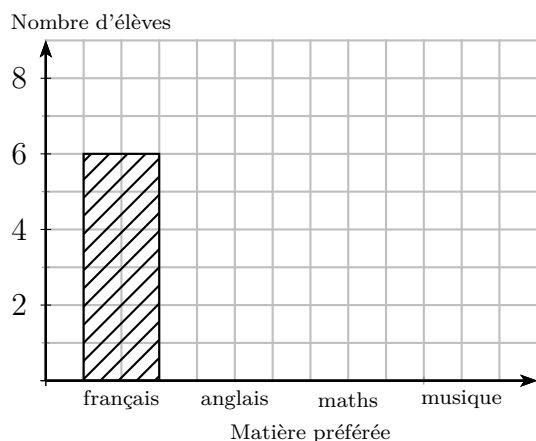
$$n_1 + n_2 + \dots + n_k = N$$

2. La somme des fréquences ($f_i = \frac{n_i}{N}$) est toujours égale à $1 = 100\%$.

3.2.2 Représentation graphique d'une variable qualitative

Modèle 8. Répartition des 20 élèves d'une classe selon leur matière préférée.

Diagramme à rectangles



3.3 Variables quantitative continue et discrète à grand nombre de valeurs

3.3.1 Tableau de distribution d'une variable quantitative continue

Modèle 11. On a demandé à ces élèves leur taille en centimètre.

172 157 162 156 167 179 173 173 178 160
 168 171 165 166 184 170 165 164 160 175

Tableau de distribution :

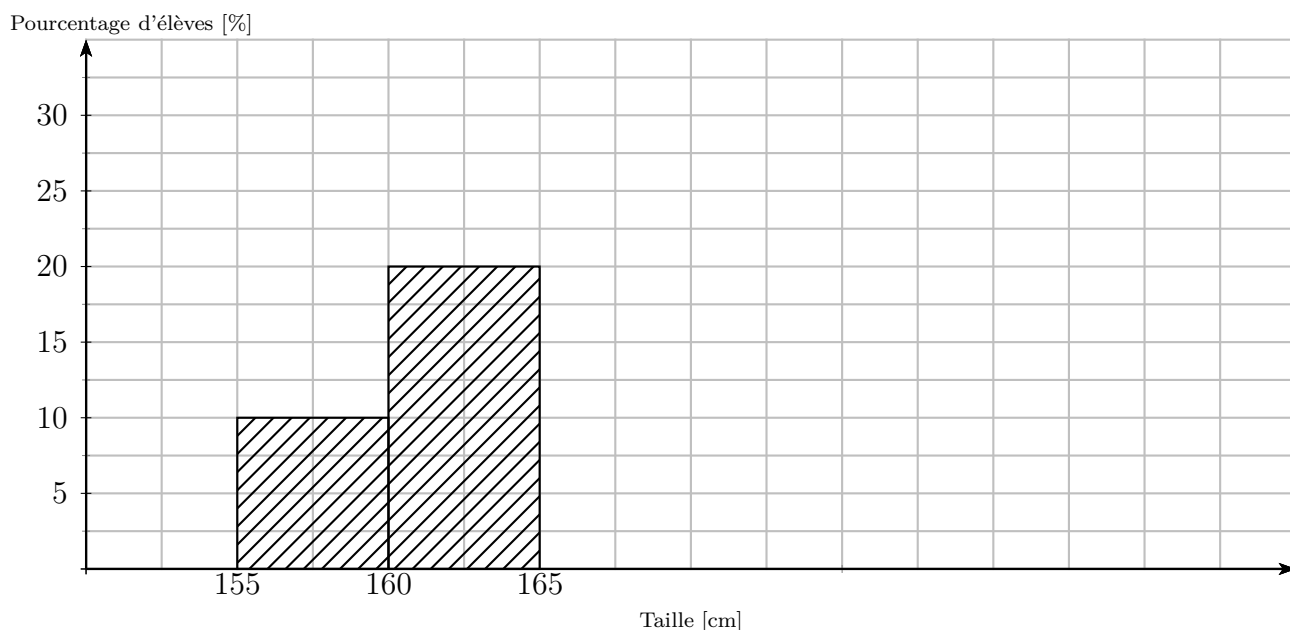
Répartition des selon

Taille	Effectif = n_i	Fréquence = f_i
Total		

3.3.2 Représentation graphique d'une variable quantitative continue

Modèle 12. Répartition des 20 élèves d'une classe selon leur taille.

Histogramme et polygone des fréquences



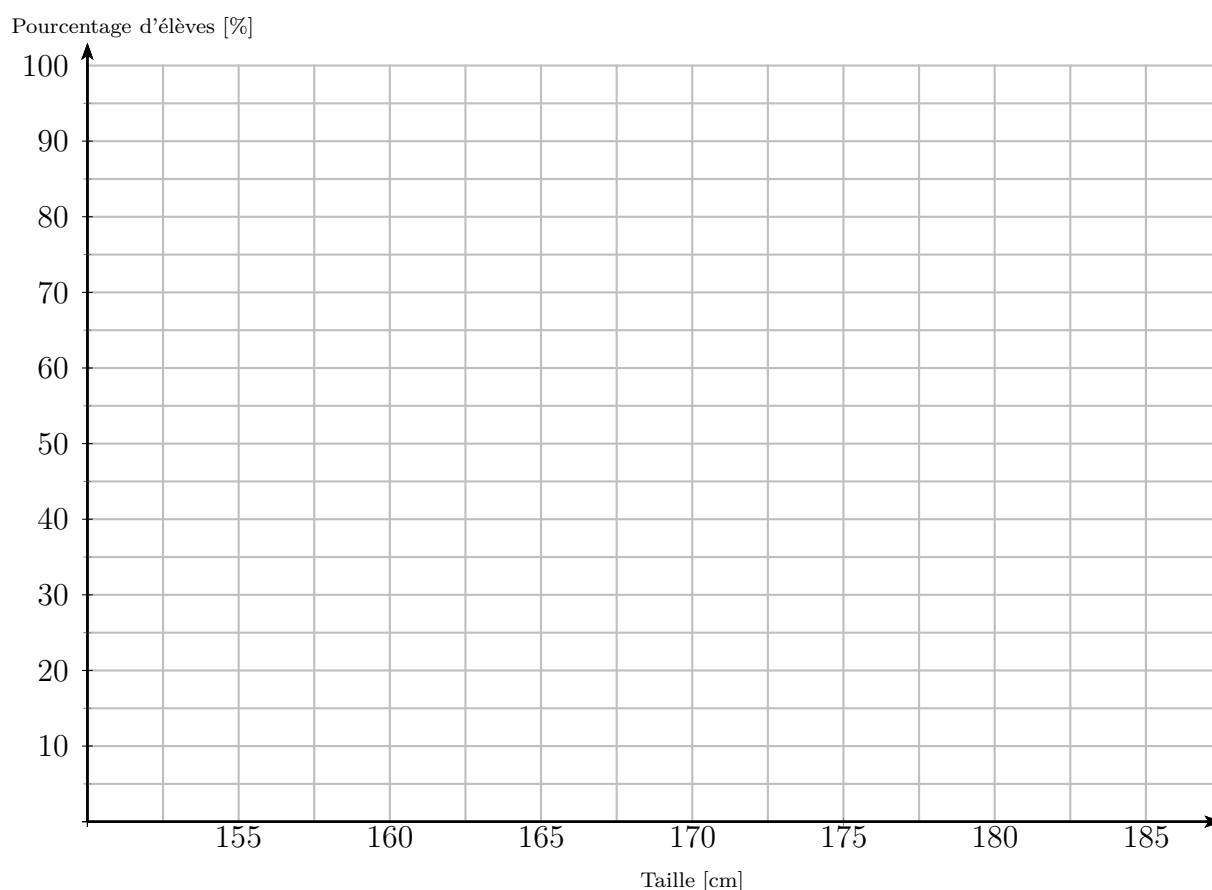
Le **polygone des fréquences** est la ligne polygonale obtenue en joignant les milieux respectifs des côtés supérieurs des rectangles de l'histogramme. On commence et on termine le polygone des fréquences en ajoutant une classe de fréquence nulle avant la première classe et une autre après la dernière classe.

Modèle 13. Répartition des 20 élèves d'une classe selon leur taille.

Polygone des fréquences cumulées (ou courbe de fréquences cumulées)

Taille	Effectif = n_i	Fréquence = f_i	Fréquence cumulée = F_i
[155 ; 160 [2	10 %	
[160 ; 165 [4	20 %	
[165 ; 170 [5	25 %	
[170 ; 175 [5	25 %	
[175 ; 180 [3	15 %	
[180 ; 185 [1	5 %	
Total	20	100 %	

polygone des fréquences cumulées



30 % des élèves mesurent

..... % des élèves mesurent moins de 1.75 mètre.

5 % des élèves mesurent au moins

Remarque

L'abscisse du point correspondant à une fréquence cumulée croissante de 50 % s'appelle la de la v.s. (ici \cong ).

3.4 Mesure de tendance centrale

But : résumer une série statistique par une seule valeur.

3.4.1 Moyenne

Calcul de la moyenne dans le cas discret

Modèle 14. On reprend les dernières notes d'anglais des 20 élèves d'une classe (modèle p.8).

5 4.5 3.5 5 6 3.5 4 2.5 4 4.5
4 4.5 4.5 4.5 3 4 4.5 5 3.5 4

Calcul de la moyenne arithmétique \bar{x} (se lit x barre) :

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n}$$

$\bar{x} = \dots$

Calculatrice :

Calcul de la moyenne à partir du tableau de distribution (formule plus rapide) :

Répartition des 20 élèves d'une classe selon leur note d'anglais

Note (modalité) = c_i	Effectif = n_i	Fréquence = f_i
1 / 1.5 / 2	0	0 %
2.5	1	5 %
3	1	5 %
3.5	3	15 %
4	5	25 %
4.5	6	30 %
5	3	15 %
5.5	0	0 %
6	1	5 %
Total	20	100 %

$$\bar{x} = \frac{n_1 \cdot c_1 + n_2 \cdot c_2 + \dots + n_k \cdot c_k}{n} = f_1 \cdot c_1 + f_2 \cdot c_2 + \dots + f_k \cdot c_k$$

$\bar{x} = \dots$

Calculatrice :

Calcul de la moyenne dans le cas continu

Modèle 15. On reprend les tailles des 20 élèves d'une classe (modèle p.9).

172	157	162	156	167	179	173	173	178	160
168	171	165	166	184	170	165	164	160	175

Calcul direct de la moyenne (à partir des données brutes) :

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n}$$

$\bar{x} = \dots$

Calcul de la moyenne à partir du tableau de distribution (formule plus rapide) :

Répartition des 20 élèves d'une classe selon leur taille

Taille	Valeur centrale = c_i	Effectif = n_i	Fréquence = f_i	Fréquence cumulée = F_i
[155 ; 160 [2	10 %	10 %
[160 ; 165 [4	20 %	30 %
[165 ; 170 [5	25 %	55 %
[170 ; 175 [5	25 %	80 %
[175 ; 180 [3	15 %	95 %
[180 ; 185 [1	5 %	100 %
Total		20	100 %	

$$\bar{x} = \frac{n_1 \cdot c_1 + n_2 \cdot c_2 + \dots + n_k \cdot c_k}{n} = f_1 \cdot c_1 + f_2 \cdot c_2 + \dots + f_k \cdot c_k$$

$\bar{x} = \dots$

Remarque

La valeur de la moyenne n'est pas la même selon la méthode utilisée. Dans le deuxième cas, on ne dispose plus des données brutes, on doit donc estimer la valeur de \bar{x} avec l'information disponible.

Représentation graphique de la moyenne

On place un pivot (▲) sous l'axe horizontal, au point correspondant à la moyenne, celle-ci coïncide avec le centre d'équilibre du diagramme. Les modèles 10 et 12 sont à compléter.

3.4.2 Médiane

Calcul de la médiane dans le cas discret

La médiane partage une série de données triées en deux parties égales.

Si \tilde{x} est la médiane d'une série statistique, il y a donc 50 % des données qui sont plus petites ou égales à \tilde{x} et 50 % qui sont plus grandes ou égales à \tilde{x} .

Modèle 16. Dans l'exemple des notes d'anglais des 20 élèves d'une classe (modèle p.8).

Calcul de la médiane \tilde{x} (se lit x tilde) :

$$\tilde{x} = \begin{cases} x_{\frac{n+1}{2}} & \text{si } n \text{ est impair} \\ \frac{1}{2} (x_{\frac{n}{2}} + x_{\frac{n}{2}+1}) & \text{si } n \text{ est pair} \end{cases}$$

Il y a deux formules différentes, car si n est impair et les données sont triées, il y a une donnée au milieu de la série. Si n est pair, par contre, il y a deux données qui sont au milieu, et on utilise donc la moyenne arithmétique de ces deux valeurs.

Remarque importante :

Cette mesure de tendance centrale est plus robuste que la moyenne, elle est moins affectée par les valeurs extrêmes.

Exemple :

Dans une entreprise de 35 employés, supposons que le patron gagne 40'000 francs par mois, alors que les 34 employés gagnent 3'000 francs par mois.

Calcul du revenu mensuel moyen : $\bar{x} = \dots$

Cette moyenne ne reflète en rien la réalité des travailleurs de cette entreprise. La valeur extrême du salaire du patron a un impact trop grand sur la moyenne.

Calcul de la médiane : $\tilde{x} = \dots$

Il est correct de dire que le salaire moyen dans cette entreprise est de mais il vaudrait mieux dire que le salaire **médian** est de par mois, donc qu'au moins la moitié des employés gagnent

Calcul de la médiane dans le cas continu

Modèle 17. Dans l'exemple des tailles des 20 élèves d'une classe (modèle p.9).

Calcul direct de la médiane (à partir des données brutes) :

$$\tilde{x} = \frac{1}{2} \left(x_{\frac{n}{2}} + x_{\frac{n}{2}+1} \right)$$

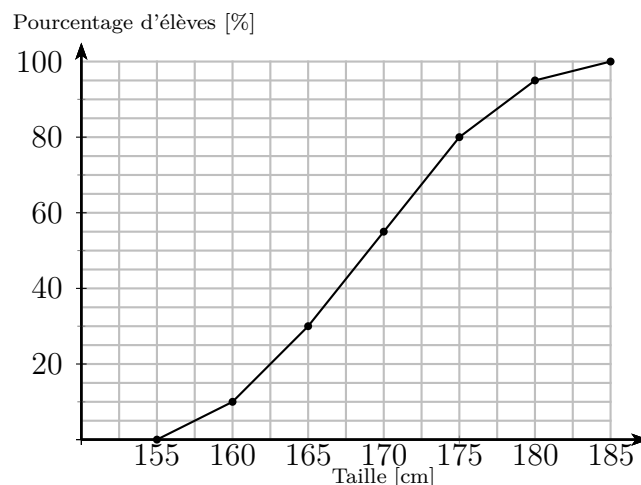
$\tilde{x} = \dots$

Si on ne dispose que du tableau de distribution, on doit estimer la médiane autrement.

Répartition des 20 élèves d'une classe selon leur taille

Polygone de fréquences cumulées

Taille	Centre	Effectif	Fré.	Fré.cum.
[155 ; 160 [157.5	2	10 %	10 %
[160 ; 165 [162.5	4	20 %	30 %
[165 ; 170 [167.5	5	25 %	55 %
[170 ; 175 [172.5	5	25 %	80 %
[175 ; 180 [177.5	3	15 %	95 %
[180 ; 185 [182.5	1	5 %	100 %
Total		20	100 %	



Classe médiane :

Calcul de la médiane par proportionnalité entre les fréquences et les valeurs :

Remarque

La valeur de la médiane n'est pas la même selon la méthode utilisée. Dans le deuxième cas, on ne dispose plus des données brutes, on doit donc estimer la valeur de \tilde{x} avec l'information disponible.

3.4.3 Mode et classe modale

Le mode est la valeur (ou la catégorie dans le cas qualitatif) qui revient le plus souvent dans une série statistique.

La classe modale est la classe qui regroupe le plus de données dans le cas d'une variable continue.

Remarques

1. Le mode ou la classe modale ne sont significatifs que si leur effectif est largement plus grand que celui des autres modalités ou des autres classes.
2. Le mode est la seule mesure de tendance centrale qui peut être utilisée pour une variable qualitative.

Modèle 18. Déterminer le mode pour chacune des parties :

a) Dans l'exemple des matières préférées, le mode est

Interprétation :

b) Dans l'exemple des notes d'anglais, le mode est

Interprétation :

c) Dans l'exemple des tailles, la classe modale est

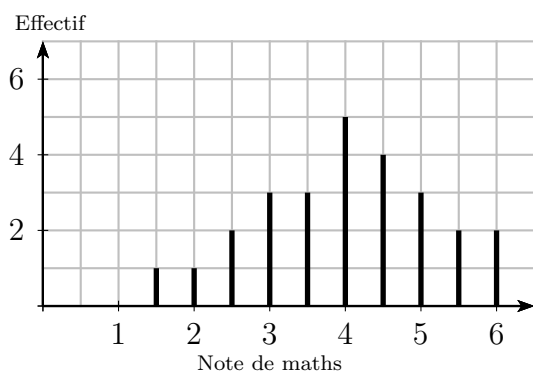
Interprétation :

3.5 Mesures de dispersion

Lorsqu'on résume une série statistique par une mesure de tendance centrale (souvent la moyenne), on ne donne aucune information sur la manière dont les données se répartissent autour de cette valeur : sont-elles toutes assez proches de la moyenne, ou trouve-t-on des valeurs très dispersées autour de celle-ci ? Cette question nécessite de donner une valeur supplémentaire, appelée mesure de dispersion.

Pour illustrer ces mesures de dispersion, nous allons nous baser sur les notes de maths de trois classes parallèles, données par des diagrammes en bâtons.

Classe A

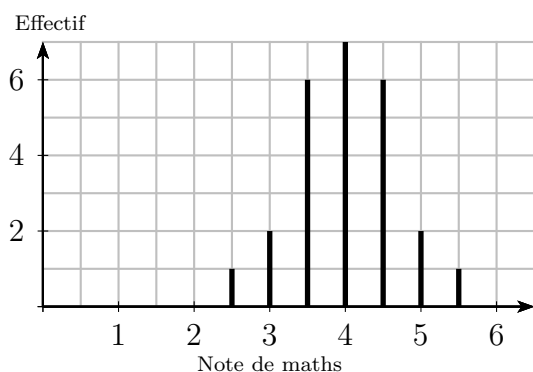


Moyenne :

Médiane :

Mode :

Classe B

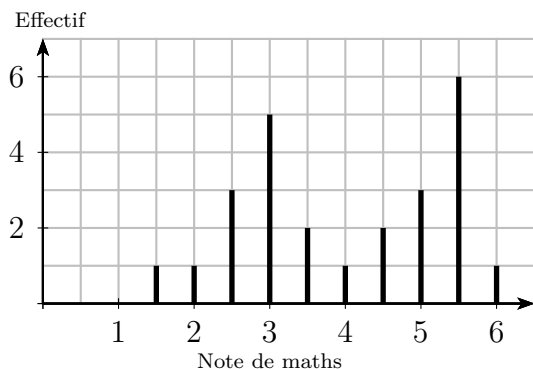


Moyenne :

Médiane :

Mode :

Classe C



Moyenne :

Médiane :

Mode :

Remarque

Ces mesures de tendance centrale ne suffisent pas à décrire les différences entre ces trois classes.

3.5.1 Etendue

Modèle 19. L'étendue est la "distance" entre la plus petite et la plus grande valeur.

Classe	A	B	C
Etendue			

Cette mesure permet de différencier les situations des classes ... et ... , mais pas des classes ... et ...

3.5.2 Variance et écart-type

La variance s^2 est l'écart quadratique moyen à la moyenne. Elle se calcule par la formule suivante :

$$s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n}$$

Dans le cas de données regroupées par modalités ou par classes, la formule devient :

$$s^2 = \frac{n_1(c_1 - \bar{x})^2 + n_2(c_2 - \bar{x})^2 + \dots + n_k(c_k - \bar{x})^2}{n} = f_1(c_1 - \bar{x})^2 + f_2(c_2 - \bar{x})^2 + \dots + f_n(c_k - \bar{x})^2$$

Comme la variance est calculée à partir de grandeurs au carré, on définit l'écart-type, noté s , comme la racine de la variance. On obtient ainsi une mesure de la dispersion dans la même unité que les mesures initiales.

Calculatrice :

Modèle 20. Calculer la variance et l'écart-type des classes A , B et C à l'aide des diagrammes en bâtons de la page 16.

Classe	A	B	C
Variance			
Ecart-type			

Grâce à ces nouvelles mesures, on peut maintenant affirmer que les notes de la classe ... sont les moins dispersées autour de la moyenne, et que celles de la classe ... sont les plus dispersées.

3.6 Mesure de position

3.6.1 Quantiles

La médiane est une valeur qui partage les données de l'échantillon en deux groupes de taille égale : 50% des données sont inférieures ou égales à la médiane, et 50% des données lui sont supérieures ou égales.

Cette idée se généralise pour n'importe quel pourcentage. Par exemple, quelle est la valeur qui sépare les 25% les plus petits des 75% les plus grands ?

Un quantile à $p\%$ est une valeur qui est supérieure ou égale aux $p\%$ des données les plus petites, et inférieure ou égale au reste des données. On le note $q_p\%$.

Cas particuliers

- Les quartiles (Q_1, Q_2, Q_3) sont les quantiles à 25%, 50% et 75%. Ils partagent les données en quatre parties égales. Le deuxième quartile (Q_2) est égal à la médiane.
- Les quintiles (V_1, V_2, V_3, V_4) sont les quantiles à 20%, 40%, 60% et 80%. Ils partagent les données en cinq parties égales.
- Les déciles (D_1, D_2, \dots, D_9) sont les quantiles à 10%, 20%, ..., 90%. Ils partagent les données en 10 parties égales.
- Les centiles (C_1, C_2, \dots, C_{99}) sont les quantiles à 1%, 2%, ..., 99%. Ils partagent les données en cent parties égales.

Les quantiles se déterminent en utilisant le même principe que pour la médiane.

Remarque

Pour que les quantiles aient du sens, il faut que l'échantillon soit suffisamment grand. On ne calculera jamais le premier décile d'une distribution composée d'une dizaine de valeurs !

3.6.2 Boxplot

Un boxplot (ou boîte à moustache) est une manière de représenter graphiquement la distribution d'une variable statistique en faisant apparaître la médiane, les quartiles et les deux valeurs extrêmes (la plus petite et la plus grande).

Modèle 21. On suppose que le nombre de périodes d'absences par année des élèves d'un gymnase se répartit de la manière suivante :

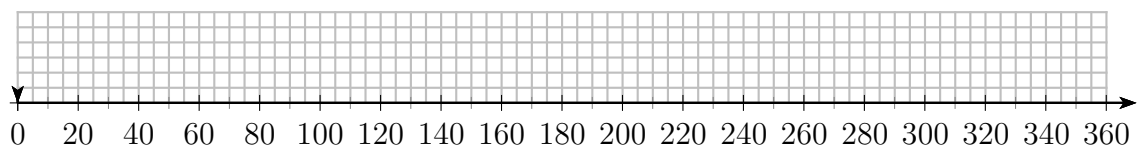
Périodes d'absence	$[0 ; 10 [$	$[10 ; 30 [$	$[30 ; 60 [$	$[60 ; 90 [$	$[90 ; 120 [$	$[120 ; 360 [$	Total
Fréquence	14 %	60 %	15 %	7 %	2 %	2 %	100 %

a) Calcul de la médiane (Q_2) :

b) Calcul du premier quartile (Q_1) :

c) Calcul du troisième quartile (Q_3) :

d) Box-Plot :



3.6.3 Cote Z

Modèle 22. Un gymnase souhaite engager un ancien étudiant pour donner des cours d'appui de mathématiques. Les quatre candidats ont suivi leur troisième année dans quatre gymnases différents, mais on souhaite tout de même déterminer le meilleur étudiant en fonction de ses résultats à l'examen de maturité.

Candidat	Nom de l'élève	Note moyenne de son gymnase	Ecart-type de son gymnase
Loïc	4.5	3.7	1.1
Muriel	5	4.1	0.6
Antonin	5.5	5.1	0.4
Eloïse	5	4.0	0.9

Si le gymnase ne se fie qu'à la note de l'élève, il devrait engager ...

S'il tient aussi compte de la moyenne du gymnase, il engagera plutôt ...

Enfin, en tenant compte de l'écart-type du gymnase, il choisira alors ...

Pour décrire la position d'une donnée par rapport à une distribution, on utilise la cote Z .

$$\text{Cote } Z \text{ de } x_i : z_{x_i} = \frac{x_i - \bar{x}}{s}$$

La cote Z mesure la distance d'une valeur à la moyenne, mesurée en nombre d'écart-type.

Cote Z de Loïc :

Cote Z de Muriel :

Cote Z d'Antonin :

Cote Z d'Eloïse :

Interprétation de la cote Z

Une cote Z positive signifie que la valeur est supérieure à la moyenne, alors qu'une cote Z négative indique qu'elle est en dessous de la moyenne.

Une cote Z de 3 ou plus, ou de -3 ou moins indique une valeur très rare. La cote Z permet donc d'identifier des situations exceptionnelles ou peu plausibles.

Modèle 23. Un cinéma accueille en moyenne 120 spectateurs les soirs de semaine, avec un écart-type de 14 spectateurs. Il décide de proposer une offre spéciale le mardi soir, avec des places à tarif réduit. Le mardi suivant, 172 spectateurs assistent à la projection. Peut-on déduire que l'offre spéciale a eu de l'effet ?

Un lundi soir, une exposition a lieu tout près du cinéma. Ce même soir, le cinéma vend 104 billets. Le gérant se plaint de l'effet négatif de l'exposition, qui lui aurait "volé" des clients. Est-ce justifié ?

4 Probabilités

4.1 Définition de la notion de probabilité

- **Univers** : ensemble U de tous les cas (ou issues) possibles.
- **Événement** : un sous-ensemble de cas possibles, donc un sous-ensemble de U .
- U est l'événement **certain**.
- \emptyset est l'événement **impossible**.

Notation : $|U|$ désigne le nombre d'issues contenues dans U .

La probabilité d'un événement A , notée $P(A)$ est définie par :

$$P(A) = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}} = \frac{|A|}{|U|}$$

C'est un **nombre** sans unité compris entre 0 et 1 (100%) exprimant les « chances » que l'événement se réalise.

On se trouve dans une situation d'équiprobabilité si l'on peut décrire l'expérience aléatoire à l'aide d'un univers U dont tous les cas ont les mêmes « chances » de se réaliser. Dans une situation d'équiprobabilité, la probabilité d'un événement A vérifie :

1)

$$0 \leq P(A) \leq 1$$

En particulier, $P(U) = 1$, où U est l'univers des cas possibles

2) Probabilité d'un événement contraire :

Soit $B = U \setminus A$ l'événement comportant les cas non compris dans l'événement A .

$$P(B) = 1 - P(A)$$

4.2 Probabilités en utilisant la combinatoire

Modèle 24. Résoudre l'exercice 4.3.

$|U| = \dots$

a) $|A| = \dots$

b) $|B| = \dots$

c) $|C| = \dots$

4.3 Probabilités en utilisant un diagramme de Venn

Modèle 25. Résoudre l'exercice 4.8.

$$|U| = \dots$$

Diagramme de Venn :

a) $|A| = \dots$

b) $|B| = \dots$

4.4 Probabilité conditionnelle

Probabilité conditionnelle de « A sachant B »

Soient A et B deux événements avec $p(B) > 0$.

La probabilité conditionnelle de A sachant que B est réalisé (plus simplement, probabilité de A sachant B), notée $P(A|B)$, est définie par :

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Modèle 26. Résoudre l'exercice 4.15.

$$|U| = \dots$$

$$P(A) = \dots$$

$$P(A|B) = \dots$$

$$P(A|C) = \dots$$

$$P(A|\bar{B}) = \dots$$

$$P(A|\bar{C}) = \dots$$

4.5 Probabilités avec diagramme en arbre

Modèle 27. Résoudre l'exercice 4.23.

$$|U| = \dots$$

Diagramme en arbre :

a) ...

b) ...

4.6 Espérance mathématique

Modèle 28. Résoudre l'exercice 4.39.

a)

x_i	1	2	3	4	5	Total
$P(X = x_i)$						

b) $\mathbb{E}(X) = \dots$

5 Inférence statistique

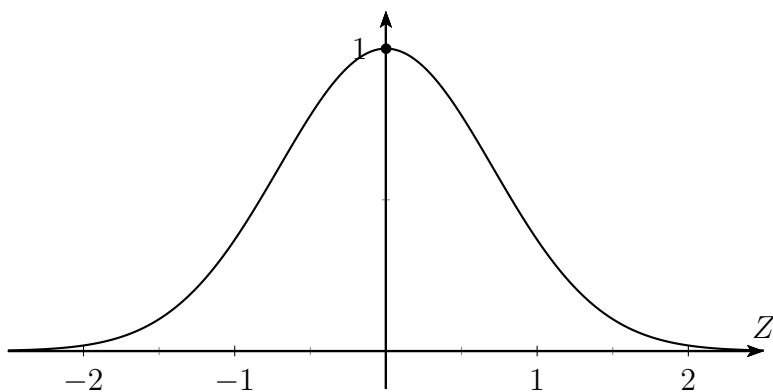
5.1 Loi normale

Modèle 29. Résoudre l'exercice 5.4 (d et h)

Normalisation

Si X est une variable aléatoire suivant une loi normale d'espérance μ et de variance σ^2 , alors

$$X \sim \mathcal{N}(\mu; \sigma^2) \Rightarrow Z = \frac{X - \mu}{\sigma} \sim \mathcal{N}(0; 1)$$



Modèle 30. Résoudre l'exercice 5.6

a) • $\mu = \dots$ et $\sigma = \dots$

- $X \sim \mathcal{N}(50; 16) \Rightarrow \dots$

- $z_{60} = \dots$

- $P(X > 60) = \dots$

f) • $\mu = \dots$ et $\sigma = \dots$

- $X \sim \mathcal{N}(0; 10) \Rightarrow \dots$

- $z_1 = \dots$

- $P(|X| > 1) = \dots$

5.2 Théorème central limite (TCL)

Notations : Population de taille N et échantillon de taille n .

Condition : $n \geq 30$ ou population qui suit une loi normale.

Si X_1, X_2, \dots, X_n sont des variables aléatoires indépendantes et suivent la même loi normale d'espérance μ et de variance σ^2 , alors

$$\frac{X_1 + X_2 + \dots + X_n}{n} = \bar{X} \sim \mathcal{N}\left(\mu; \frac{\sigma^2}{n}\right) \quad \text{si } n \rightarrow +\infty$$

et

$$Z = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}} \sim \mathcal{N}(0; 1) \quad \text{avec} \quad \sigma_{\bar{X}} = \begin{cases} \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}} & \text{si } n < \frac{N}{20} \text{ (petit échantillon)} \\ \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}} & \text{sinon (grand échantillon)} \end{cases}$$

Modèle 31. Résoudre l'exercice 5.14.

$\mu = \dots$ et $\sigma = \dots$

Taille population : $N = \dots$ et taille échantillon : $n = \dots$

5.3 Intervalles de confiance

Condition : population qui suit une loi normale.

Notations :

//////////	Population	Échantillon
Taille	N	n
Moyenne	μ	\bar{x}
Variance	σ^2	s^2

I est un intervalle de confiance au niveau de confiance $1 - \alpha$ (en %) pour μ si

$$P(\mu \in I) = 1 - \alpha$$

α s'appelle le **seuil de l'intervalle de confiance**. Il représente le risque d'erreur.

La marge d'erreur se détermine comme suit :

$$E = q_{1-\alpha/2} \cdot \sigma_{\bar{X}}$$

où $q_{1-\alpha/2} = z \iff \Phi(z) = 1 - \alpha/2$ (voir loi normale $\mathcal{N}(0;1)$ formulaires p.28).

Alors l'intervalle de confiance I pour μ est : $I = [\bar{x} - E; \bar{x} + E]$

Modèle 32. Résoudre l'exercice 5.22.

Taille échantillon : $n = \dots$

5.4 Test d'hypothèses

Tests d'hypothèse pour une moyenne

Condition : population qui suit une loi normale.

Soit H_0 une hypothèse appelée hypothèse nulle : l'échantillon provient d'une population de moyenne $\mu = \mu_0$

Soit H_1 une autre hypothèse appelée hypothèse alternative : l'échantillon provient d'une population de moyenne $\mu \neq \mu_0$ (test bilatéral)

On définit le **seuil de signification** appelé α (compris entre 1% et 5%) tel que :

$$\alpha = \text{probabilité de se tromper en rejetant } H_0$$

Le test d'hypothèse pour une moyenne comporte quatre étapes :

1. Hypothèses et seuil de signification :

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu < \mu_0 : \text{test unilatéral à gauche} \quad (\text{ou } \mu > \mu_0 : \text{test unilatéral à droite})$$

2. Conditions d'application et représentation graphique (voir formulaires p.26) :

Si $n \geq 30$, alors \bar{X} suit approximativement une loi normale de moyenne $\mu_{\bar{x}}$ et d'écart-type $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

3. Point critique :

$$z = q_{1-\alpha} \iff \Phi(z) = 1 - \alpha \text{ (selon p.28)}$$

$$E = z \cdot \sigma_{\bar{x}} \Rightarrow c = \mu - E \text{ (ou } c = \mu + E \text{ si test unilatéral à droite)}$$

4. Règle de décision :

Rejeter H_0 si la moyenne \bar{x} de l'échantillon est inférieure à c (ou supérieure à c si test unilatéral à droite).

Modèle 33. Résoudre l'exercice 5.30.

Population : $\mu = \dots$

Echantillon : $n = \dots$; $\bar{x} = \dots$; $\hat{\sigma} = \dots$