

Exercices de TD de statistiques

Rappels

Les exercices 1 à 4 sont des rappels portant sur la prise en main de la calculatrice. Ils sont corrigés en ligne

Exercice 1 : Fractions

Si le résultat d'un calcul est un nombre à virgule, vous pouvez soit demander à la calculatrice de l'afficher sous forme décimale (nombre à virgule) soit sous forme de fraction (avec la touche $\boxed{F \leftrightarrow D}$ pour les casios, ou la fonction $\blacktriangleright \text{Frac}$ du menu $\boxed{\text{MATH}}$ des TI).

- Grâce à cette fonctionnalité de la calculatrice, simplifiez chacune des fractions suivantes : $\frac{-143}{-104}$, $\frac{-10 \times 3}{-10}$, $\frac{-13-9}{4 \times 2}$, $\frac{14 \times 2}{67-4}$, $\frac{-7-48}{126-49}$, $\frac{80-35}{-27-15}$.
- Classez les dans l'ordre croissant ("du plus petit au plus grand").

Exercice 2 : Utilisation de la calculatrice

- La capture d'écran suivante indique un calcul effectué sur une calculatrice.

A calculator screen showing the input -1^2 and the result -1 .

Peut-on en conclure

- que le carré de -1 est -1 ?
- que le carré de 1 est -1 ?
- que -1 est l'opposé du carré de 1 ?

De même, pour chacune des questions ci-dessous, une capture d'écran montre un calcul effectué sur la calculatrice, et trois réponses sont proposées qui interprètent (de manière correcte ou erronée) ce calcul. À chaque fois, une seule des trois réponses proposées est correcte, à vous de déterminer laquelle.

- $\frac{1}{2020 \times 2021} \simeq 0,000000245$
 - $\frac{1}{2020 \times 2021} \simeq 24495274,86$
 - $\frac{1}{2020 \times 2021} \simeq 2,45$
- $-8 \times 44^2 = 123904$
 - $-352^2 = 123904$
 - $(-8 \times 44)^2 = 123904$
- $16 - 25 \times 2021 = -18189$
 - $9 \times 2021 = -18189$
 - $(16 - 25) \times 2021 = -18189$
- $1 + (4 + 2 - 1)^2 = 34$
 - $1 + 4^2 + 2^2 - 1^2 = 34$
 - $(1 + 4)^2 + (1 + 2)^2 + (1 - 1)^2 = 34$

Exercice 3 : Puissances de dix

Calculer (de préférence "de tête") chacune des expressions suivantes :

$47 \div 100$	$0,47 \times 10^2$	$18,885\,392\,4 \times 10^{-5}$
$1,048\,4 \times 10^6$	$8 \times 100\,000\,000$	$111 \times 10\,000\,000$
$19\,820 \div 100\,000$	$11,283\,133 \times 10^{-6}$	8×100

Exercice 4 : Règles de calcul

- On considère les expressions suivantes :

• $3 - 5 + 4 - 2$	• $4 - 5 + 3 - 2$
• $(2 - 5) - (4 - 3)$	• $(3 - (5 + 4 - 2))$

Calculer la valeur de chacune de ces expressions, et constater que certaines sont égales.

- De même, pour chacune des expressions suivantes, calculer les lorsque les variables prennent les valeurs indiquées, et remarquer que certaines d'entre elles sont égales.

(a) Les expressions suivantes

$$\begin{aligned}
 &\bullet 5 \frac{2}{7 \times x} & \bullet \frac{2 \div 5}{7 \times x} \\
 &\bullet \frac{1}{7} \times \frac{2}{5 \times x} & \bullet \frac{5 \times 2}{7 \times x} \\
 &\bullet 2 \div x \times 5 \div 7 & \bullet \frac{2}{7 \times x \div 5}
 \end{aligned}$$

lorsque $x = 5$.

(b) Les expressions

$$\begin{aligned}
 &\bullet \frac{3 \times a}{2} - \frac{3 \times 4}{b} & \bullet 3 \left(\frac{a}{2} - \frac{4}{b} \right) \\
 &\bullet (3 \times a - 4 \times 3) \div (2 - b) & \bullet 3 \frac{(a-4)}{(2-b)}
 \end{aligned}$$

où $a = -8$ et $b = 9$.

(c) Lorsque $x = 9$, $y = 8$ et $z = 6$:

$$\begin{aligned}
 &\bullet 4x + z - 8x & \bullet -8 - 4x + z & \bullet z - y \times \frac{8-4}{y}x \\
 &\bullet (z - 8) - 4x & \bullet z - (8 - 4)x
 \end{aligned}$$

(d) Lorsque $x_1 = -9$ et $x_2 = -6$:

$$\begin{aligned}
 &\bullet 9 + x_1 - x_2 - x_1 + 2 \times 8 + 2 \times 6 \\
 &\bullet 9 - x_2 + 2 - 8 - 6 - (2 - 8 - 6) \\
 &\bullet 9 - x_2 - 0(2 - 8 - 6) \\
 &\bullet 9 - x_2 + 2 \times 8 + 2 \times 6 \\
 &\bullet 9 - x_2 - 1(2 - 8 - 6) \\
 &\bullet 9 - x_2 - (2 - 8 - 6) \\
 &\bullet 9 - x_2 - 2(-8 - 6)
 \end{aligned}$$

(e) Lorsque $a_1 = 5$, $a_2 = -8$, $x_1 = -8$ et $x_2 = -2$:

$$\begin{aligned}
 &\bullet x_2 + 5 \times a_2 + 6^2 \times x_1 + a_1 \\
 &\bullet x_2 \times a_2 + 5 \times a_2 + 6^2(x_1 + a_1) \\
 &\bullet (x_2 + 5) \times a_2 + (x_1 + a_1) \times 6^2 \\
 &\bullet (x_2 + (5 \times a_2)) + 6^2(x_1 + a_1) \\
 &\bullet x_2 + 5 \times a_2 + (6^2 \times x_1 + (a_1)) \\
 &\bullet 5 \times a_2 + (x_2 + (x_1 + a_1) \times 6 \times 6)
 \end{aligned}$$

Compléments

Exercice 5 : Précision des calculs

Maeva, Clara, Jeanne, Erwan et Hedi étudient en première année de psychologie à l'université de Bourgogne.

On leur demande de calculer l'écart-type de la taille de 3 femmes, et au vu des formules du formulaire, cela revient à calculer des nombres notés " $m(X)$ ", " $m(X^2)$ ", " Var " et " s ", donnés par les formules suivantes :

$$\begin{aligned}
 m(X) &= \frac{1,59+1,71+1,63}{3} \\
 m(X^2) &= \frac{1,59^2+1,71^2+1,63^2}{3} \\
 Var &= m(X^2) - (m(X))^2 \\
 s &= \sqrt{Var}
 \end{aligned}$$

Le nombre " s " issu de ce calcul est l'écart-type, exprimé (dans le cas présent) en mètre. On leur demandait de calculer cet écart type à 1 cm près.

Ci-dessous se trouve ce que chacun.e a écrit sur sa copie ainsi que les calculs effectués sur sa calculatrice.

Copie de Maeva

$$\begin{aligned}
 m(X) &= \frac{1,59+1,71+1,63}{3} \simeq 1,64 \\
 m(X^2) &= \frac{1,59^2+1,71^2+1,63^2}{3} \simeq 2,70 \\
 Var &\simeq 2,7 - 1,64^2 \simeq 0,01 \\
 s &\simeq \sqrt{0,01} \simeq 0,10 \text{ mètre.} \\
 \text{Donc l'écart-type est d'environ 10 cm.}
 \end{aligned}$$

Calculatrice de Maeva

$$\begin{aligned}
 &(1.59+1.71+1.63) \div 3 \\
 &\quad 1.643333333 \\
 &(1.59^2+1.71^2+1.63^2) \div 3 \\
 &\quad 2.703033333 \\
 &2.7-1.64^2 \\
 &\quad 0.0104 \\
 &\sqrt{(0.01)} \\
 &\quad 0.1
 \end{aligned}$$

Copie de Clara

$$\begin{aligned}
 m(X) &= \frac{1,59+1,71+1,63}{3} = \frac{4,93}{3} \\
 m(X^2) &= \frac{1,59^2+1,71^2+1,63^2}{3} = \frac{8,1091}{3} \\
 Var &\simeq \frac{8,1091}{3} - \left(\frac{4,93}{3} \right)^2 \simeq 0,0025 \\
 s &= \sqrt{Var} \simeq 0,05 \text{ mètre.} \\
 \text{Donc l'écart-type est d'environ 5 cm.}
 \end{aligned}$$

Calculatrice de Clara

$$\begin{aligned}
 &1.59+1.71+1.63 \\
 &\quad 4.93 \\
 &1.59^2+1.71^2+1.63^2 \\
 &\quad 8.1091 \\
 &8.1091 \div 3 - (4.93 \div 3)^2 \\
 &\quad 0.002488888889 \\
 &\sqrt{(Ans)} \\
 &\quad 0.04988876516
 \end{aligned}$$

Copie de Jeanne

$$m(X) = \frac{1,59+1,71+1,63}{3} \simeq 1,64$$

$$m(X^2) = \frac{1,59^2+1,71^2+1,63^2}{3} \simeq 2,70$$

$$Var = m(X^2) - (m(X))^2 \simeq 0,0025$$

$$s = \sqrt{Var} \simeq 0,05 \text{ mètre.}$$

Donc l'écart-type est d'environ 5 cm.

Calculatrice de Jeanne

```
(1.59+1.71+1.63)/3=M
1.643333333
(1.59^2+1.71^2+1.63^2)/3=N
2.703033333
N-M^2=U
0.002488888889
√(U)
0.04988876516
```

Copie d'Erwan

$$m(X) = \frac{1,59+1,71+1,63}{3} \simeq 1,64$$

$$m(X^2) = \frac{1,59^2+1,71^2+1,63^2}{3} \simeq 2,70$$

$$Var = m(X^2) - (m(X))^2 \simeq 0,0025$$

$$s = \sqrt{Var} \simeq 0,05 \text{ mètre.}$$

Donc l'écart-type est d'environ 5 cm.

Calculatrice d'Erwan

```
(1.59,1.71,1.63)→List 1
(1.59,1.71,1.63)
Sum (List 1)/3=M
1.643333333
Sum ((List 1)^2)/3=N
2.703033333
N-M^2=U
0.002488888889
√(U)
0.04988876516
```

Copie de Hedi

$$s = \sqrt{\frac{1,59^2+1,71^2+1,63^2}{3} - \left(\frac{1,59+1,71+1,63}{3}\right)^2} \simeq 0,05 \text{ mètre.}$$

Donc l'écart-type est d'environ 5 cm.

Calculatrice de Hedi

```
√((1.59^2+1.71^2+1.63^2)/3
-((1.59+1.71+1.63)/3)^2)
0.04988876516
```

(vous pourrez reprendre la rédaction d'un·e des étudiant·e·s ayant trouvé la bonne réponse).

Exercice 6 : Utilisation du symbole $\sum$

1. Calculer, sans utiliser la calculatrice, chacune des sommes suivantes : $\sum_{n=1}^5 n$, $\sum_{n=-5}^5 n$, $\sum_{i=-3}^3 i^2$.

2. Écrivez avec le symbole $\sum$ la somme suivante $\frac{6^2}{6+9} + \frac{7^2}{7+9} + \frac{8^2}{8+9} + \frac{9^2}{9+9} + \frac{10^2}{10+9} + \frac{11^2}{11+9} + \frac{12^2}{12+9} + \frac{13^2}{13+9}$.

3. On pose $x_1 = 6$, $x_2 = 7$, $x_3 = 8$, $x_4 = 9$, $x_5 = 10$, $n_1 = 6$, $n_2 = 5$, $n_3 = 5$, $n_4 = 1$ et $n_5 = 9$.

(a) Calculer $\sum_{i=1}^5 n_i (x_i)^2$ sans utiliser de liste sur la calculatrice.

(b) Calculer à nouveau $\sum_{i=1}^5 n_i (x_i)^2$, en utilisant cette fois-ci des listes sur la calculatrice.

Chapitre 1 : Statistiques descriptives univariées**Exercice 7 : Type de variables**

1. Mme Giraud recueille des données sur

- le nombre de rendez-vous médicaux que différents patients ont pris au cours de l'année 2019
- la popularité des différents footballeurs de l'équipe de France
- les noms des rues de Dijon
- le taux de testostérone parmi les patients souffrant de troubles cognitifs

Dans chacun de ces cas, indiquer quelle est la population, quelle est la variable étudiée, et quelle est la nature de cette variable.

1. Parmi ces différentes réponses, laquelle/lesquelles vous semble(nt) correcte(s) ?

Expliquez pourquoi l'un·e des étudiant·e·s a obtenu(e) une réponse erronée ?

2. On demande ensuite de calculer l'écart type de la taille d'un échantillon de 5 hommes. Comme l'échantillon n'est plus le même, il faut cette fois-ci calculer

$$m(X) = \frac{1,74+1,82+1,89+1,73+1,88}{5}$$

$$m(X^2) = \frac{1,74^2+1,82^2+1,89^2+1,73^2+1,88^2}{5}$$

$$Var = m(X^2) - (m(X))^2$$

$$s = \sqrt{Var}$$

Rédigez ce calcul d'une façon qui empêche des erreurs d'arrondis

2. De même quel est le type de variable lorsqu'on étudie :

- (a) la couleur des voitures stationnées dans le campus
- (b) la pluie tombée à Dijon au cours des différents mois de l'année
- (c) le degré de satisfaction des usagers des transports en commun
- (d) les prénoms des enfants nés en 2019
- (e) le temps hebdomadaire passé devant la télévision par des enfants

Exercice 8 : Taille des ménages

On appelle T la variable statistique indiquant le nombre de personnes d'un ménage. Au sein d'une certaine ville, on obtient les données suivantes :

Nombre de personnes du ménage	1	2	3	4	5	6
Effectifs : nombre de ménages	1853	1122	1925	979	324	123

- Quelles sont les modalités de la variable T ?
- Calculer les fréquences et les fréquences cumulées.
- Calculer la proportion de ménages d'au moins 3 personnes.
- Déterminer $\mathbb{P}_r[T < 4]$.

Exercice 9 : Moyens de transport

1. Utilisation des transports en commun

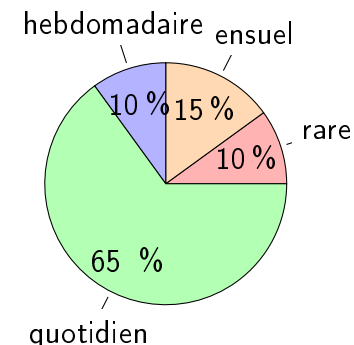
- (a) Un étudiant en sociologie interroge 20 personnes de son entourage sur leur utilisation des transports en commun. Il réalise le graphique ci-dessous après les avoir regroupés en quatre groupes :

rare : ceux qui utilisent les transports en communs moins d'une fois par mois

mensuel : au moins une fois par mois, mais moins d'une fois par semaine

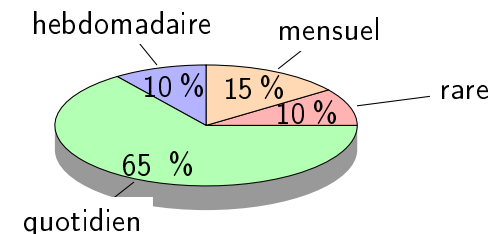
hebdomadaire : entre une et cinq fois par semaine

quotidien : au moins 5 fois par semaine

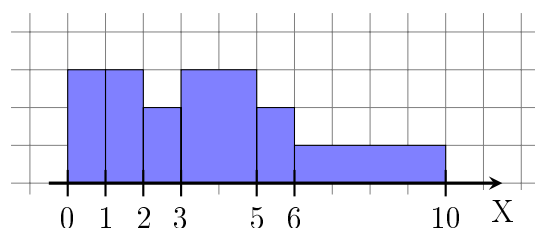


- Déterminer les effectifs des différentes modalités
- Quelle est, au sein de cet échantillon, la proportion d'individus qui utilisent les transports en commun moins d'une fois par semaine ?

- (b) Lorsqu'il entre les données dans un tableur, l'ordinateur réalise un graphique ressemblant à celui ci-contre. Que penser d'un tel graphique ?



2. **Utilisation de la voiture** Pour le même échantillon, il désigne par X le temps hebdomadaire passé dans une voiture (exprimé en heures). Il trace un histogramme sur une feuille quadrillée (et constate avec surprise que ses rectangles tombent exactement sur le quadrillage de la feuille) :



(a) Déterminer les fréquences des différentes classes

X	$[0; 1[$	$[1; 2[$	$[2; 3[$	$[3; 5[$	$[5; 6[$	$[6; 10[$
Fréquence						

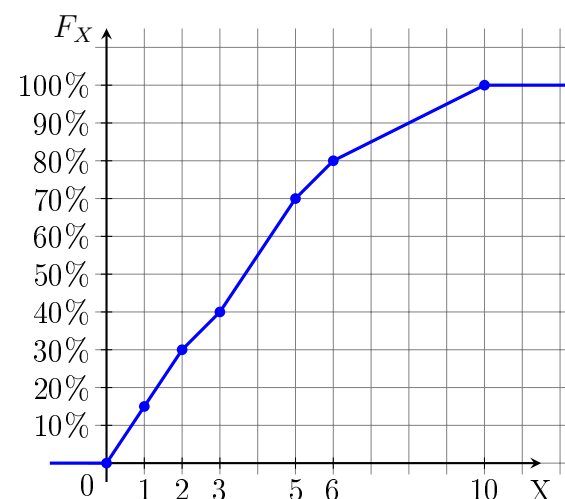
(b) Commentez : comparer les fréquences de la première et de la dernière classe, ainsi que la hauteur des rectangles correspondants.

3. **Polygone des fréquences cumulées** Enfin, il décide de représenter (toujours sur son papier quadrillé) le polygone des fréquences cumulées des données de la question 2. Celui-ci est reproduit à la fin de l'énoncé de cet exercice.

(a) Déterminer graphiquement la médiane et les quartiles approchés de X .

(b) Déterminer graphiquement les proportions $\mathbb{P}_r[X \leq 4]$, $\mathbb{P}_r[X \geq 7]$ et $\mathbb{P}_r[4 \leq X \leq 7]$ (de manière approchée).

(c) On considère l'individu qui passe le moins de temps en voiture parmi les 10% d'individus qui passent le plus de temps en voiture. Combien de temps passe-t-il environ par semaine dans sa voiture ?



Exercice 10 : “Partenariat régional économique global”

On considère dans cet exercice les pays signataires du *Partenariat régional économique global*, un accord de libre échange dans le pacifique. Les pays concernés sont les pays ci-dessous :

Australie	Birmanie	Brunei	Cambodge
Monarchie 25,7 M.hab.	République 56,6 M.hab.	Monarchie 0,5 M.hab.	Monarchie 16,9 M.hab.
Chine	Corée du Sud	Indonésie	Japon
République 1 394 M.hab.	République 51,7 M.hab.	République 267 M.hab.	Monarchie 125,5 M.hab.
Laos	Malaisie	Nouvelle-Zélande	Philippines
République 7,4 M.hab.	Monarchie 32,7 M.hab.	Monarchie 4,9 M.hab.	République 109,2 M.hab.
Singapour	Thaïlande	Viêt Nam	
République 6,2 M.hab.	Monarchie 69 M.hab.	République 98,7 M.hab.	

1. Quelle est la proportion de monarchies parmi ces pays ?
2. Parmi l'ensemble des habitants de ces états, quelle proportion vit dans une république ?

3. Parmi ces pays, quelle est la population moyenne des républiques et celle des monarchies ? Calculer aussi les médianes et écarts type.

Exercice 11 : Développement psychomoteur

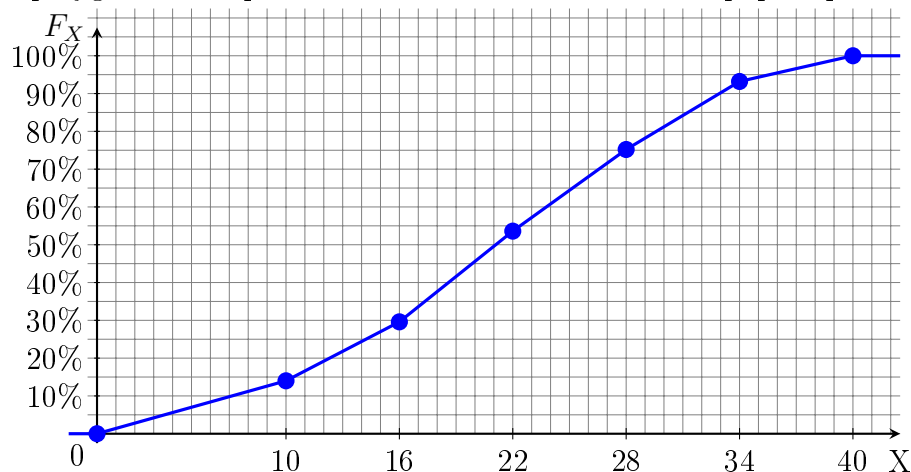
Les données ci-dessous décrivent, pour un ensemble de 160 bébés âgés de deux ans, la variable statistique X représentant le score de développement psychomoteur (SDP) :

SDP	[54; 64[	[64; 74[	[74; 84[	[84; 94[	[94; 104[	[104; 114[	[114; 124[	[124; 134[
Effectif	1	13	26	49	36	23	8	4

- Calculer la médiane et les quartiles de X .
- Calculer la moyenne et l'écart type de X .

Exercice 12 : Méthodes d'apprentissage

On teste une méthode « d'apprentissage par dessins commentés » sur un échantillon de 250 enfants. On appelle X la note obtenue par les enfants à l'issue de l'apprentissage (notée entre 0 et 40). On représente le polygone de fréquences cumulées de ces notes sur un papier quadrillé :



- Déterminer graphiquement la médiane et les quartiles de la variable statistique X .
- (a) Donner une approximation de $\mathbb{P}_r[X < 25]$.

- (b) Quelle est environ la proportion d'enfants dont la note est supérieure à 18 ?

- (c) Combien vaut environ $\mathbb{P}_r[20 \leq X < 30]$?

3. Plus précisément, les notes recueillies sont données ci-dessous :

X	[0; 10[	[10; 16[	[16; 22[	[22; 28[	[28; 34[	[34; 40[
Effectif	35	39	60	54	45	17

- Calculer la médiane et les quartiles approchés de la variable statistique X (cette fois ci sans lecture graphique).
- Calculer la note moyenne m et l'écart type s de l'échantillon.
- Donner une approximation de la proportion des enfants dont la note est comprise entre $m - s$ et $m + s$.

Exercice 13 : Calculer la moyenne et l'écart type des données de l'exercice 8.

Exercice 14 : Type de variables

- Mr Riviere recueille des données sur
 - la sensibilité aux questions écologiques des habitants d'un même quartier
 - le temps mis par des rats pour sortir d'un labyrinthe
 - les langues officielles des différents pays européens
 - le nombre de personnes des ménages Français
 Dans chacun de ces cas, indiquer quelle est la population, quelle est la variable étudiée, et quelle est la nature de cette variable.
- De même quel est le type de variable lorsqu'on étudie :
 - les marques des téléphones vendus en 2019
 - le nombre de crises d'épilepsie de différents patients au cours du mois d'octobre 2019
 - le nombre moyen d'enfants par femme dans chaque pays européen

- (d) le nombre d'enfants de chaque femme dijonnaise
- (e) le poids des appareils photos vendus en 2019
- (f) le nombre de pages des livres d'une bibliothèque
- (g) les numéros de téléphones figurant dans les pages blanches dijonnaises
- (h) la douleur ressentie par différents patients atteints de la même pathologie
- (i) les numéros de sécurité sociale des enseignants de l'université de Bourgogne
- (j) le nombre de pièces des appartements en location à Dijon

Exercice 15 : Questionnaire

Un chercheur établit un questionnaire visant à évaluer les compétences logiques de jeunes adolescents. Il fait remplir ce questionnaire à 13 sujets, et pour chacun d'eux il note dans un tableau le symbole « ✓ » si la réponse donnée est correcte, « 0 » si la réponse donnée est fausse, et a laissé la case vide si le sujet n'a pas répondu. Il obtient le tableau suivant :

Individu	A	B	C	D	E	F	G	H	I	J	K	L	M
Question 1	✓	0	0			✓	✓			0			✓
Question 2	0	0	✓	0	✓	✓	0	✓	✓	0		✓	
Question 3	✓	0	0		✓		✓	✓	0			✓	0
Question 4		0	✓	0			✓	✓	✓	✓	✓		✓
Question 5	0	✓		0	0	✓	0	✓	✓	0	0		✓
Question 6			✓	✓	✓		✓	✓	✓	0	0	✓	0
Question 7	✓	0			✓		✓	✓			0		✓
Question 8	✓	✓			✓	0		0	✓	✓		✓	✓
Question 9		✓	✓	0	✓	✓	✓	✓	✓		0		
Question 10	0	0					✓		0	✓	✓	0	0

- Quelle a été le taux de réponse à la question 7 ?
- Quelle proportion de sujets a répondu correctement à la question 3 ?

- Parmi les individus ayant répondu à la question 8, quelle proportion a donné la bonne réponse ?
- Le chercheur décide de résumer par une note les réponses données au questionnaire. Il attribue un point par bonne réponse (et zéro point par réponse erronée ou manquante). Calculer les notes (sur 10) attribuée à chaque jeune interrogé.
- Calculer la moyenne, l'écart type et la médiane de ces notes.

Exercice 16 : Regroupement en classes

Voici les notes obtenues par un groupe d'élèves lors d'un contrôle sur 10 points :

9,5 ; 7 ; 6 ; 7 ; 7 ; 3,5 ; 4 ; 6 ; 7,5 ; 1 ; 8 ; 6 ; 4 ; 9,5 ; 8 ; 4 ; 10 ; 3 ; 3 ; 5 ; 6,5 ; 7,5 ; 4,5 ; 3 ; 6 ; 5 ; 2 ; 3 ; 8 et 7

- Faire un tri des données en indiquant l'effectif de chaque note.
- Quelle est la taille de l'échantillon ?
- Donner la moyenne, l'écart type et la médiane.
- Transformer les données en les rangeant en classes d'amplitude 1,5. Calculer les fréquences et les fréquences cumulées des classes.
- Calculer la moyenne, l'écart type et la médiane des notes regroupées en classes. Que remarque-t-on ?

Exercice 17 : Âge et performances mémorielles

On considère un échantillon de 16 personnes de 35 à 80 ans, auxquels on attribue une note indiquant leurs performances mémorielles. On note leur âge X et leur performances mémorielles Y .

Les données mesurées sont les suivantes :

X	80	56	67	72	67	51	80	42	56	47	53	37	48	46	44	35
Y	22	38	33	36	22	47	30	45	48	55	38	81	46	33	51	60

Calculer la moyenne et l'écart-type de l'âge X et des performances mémorielles Y au sein de l'échantillon.

- Suite à une perte d'échantillon, on ne connaît plus le niveau de stress d'un rat, dont on a mesuré qu'il met 686ms à réagir aux stimuli visuels. Déterminer environ son niveau de stress.
- Tracer le nuage de point des variables X et Y .
- Commenter vos réponses aux questions précédentes en vous appuyant sur le nuage de points.

Exercice 20 : Revenus et espérance de vie

On étudie le revenu par habitant X (en milliers de “dollars internationaux” par an) et l'espérance de vie Y dans 12 pays différents :

pays	revenu/hab. X	espérance de vie Y
Afghanistan	1,9	60,5
Burundi	0,8	59,6
Chili	24,6	80,5
Djibouti	3,6	63,5
Espagne	38,2	82,8
Finlande	44	81,1
Grenade	14,8	73,6
Honduras	5,5	74,6
Iran	20	75,5
Jordanie	12,5	74,1
Koweït	69,7	74,7
Laos	7,4	65,7

- Déterminer le coefficient de corrélation linéaire.
- Déterminer le coefficient de corrélation des rangs (de Spearman).
- Commenter (on pourra tracer le nuage de points).

Exercice 21 : Extraversion

140 sujets ont rempli un questionnaire de personnalité d'Eysenck comprenant, entre autres, une échelle de sociabilité (X) et une échelle d'impulsivité (Y). On a les résultats suivants.

$$\sum_{i=1}^{120} x_i = 845, \quad \sum_{i=1}^{120} x_i^2 = 6636, \quad \sum_{i=1}^{120} y_i = 448$$

$$\sum_{i=1}^{120} y_i^2 = 2076, \quad \sum_{i=1}^{120} x_i y_i = 3254.$$

On souhaite savoir si ces deux échelles mesurent plusieurs manifestations d'un même caractère : l'extraversion. Si c'est le cas, on pourra additionner les scores obtenus aux deux échelles pour former une échelle unique d'extraversion.

- Calculer les moyennes, les écarts types et le coefficient de corrélation des deux échelles X et Y .
- Serait-il pertinent de résumer ces deux échelles par une échelle d'extraversion ?

Exercice 22 : Reconnaissance de forme

Pour étudier l'entraînement d'enfants passant un test de reconnaissance de forme, on mesure les performances d'enfants qui ont déjà effectué un certain nombre de fois un exercice similaire.

Pour un échantillon de 14 enfants, on mesure le temps de réponse Y des enfants, tandis que le nombre d'exercices similaires qu'ils ont déjà effectués auparavant est noté X . On obtient les résultats suivants :

Entraînement X	0	1	1	2	2	4	5	7	7	8	8	8	10	10
Temps de réponse Y	3,7	1,9	3,7	3	4	2,6	2,8	1,9	2	1,5	1,7	2,7	1,1	1,4

- Calculer le coefficient de corrélation des rangs de Spearman. Interpréter la valeur de ce coefficient.
- Calculer les moyennes et écarts type des variables X et Y , puis leur coefficient de corrélation linéaire. Interpréter la valeur de ce coefficient.

- Tracer le nuage de points correspondant à ces données. Vous conforte-t-il dans les interprétations données aux questions précédentes ?
- Si un enfant a déjà fait 6 exercices de reconnaissance de forme, alors à combien estimeriez-vous son temps de réponse si on lui fait à nouveau passer ce test ?
- Un enfant a mis 1,6 secondes pour répondre au test. Combien de fois estimeriez-vous qu'il avait déjà fait un test similaire pour s'entraîner ?

Exercice 23 : Résolution de problèmes mathématiques

On a mis au point deux méthodes d'apprentissage pour la résolution de problèmes mathématiques. La première méthode ($\mathcal{MV}$) est uniquement verbale, la seconde ($\mathcal{ME}$) est écrite. Ces méthodes sont testées sur des groupes d'enfants. Le tableau suivant représente les notes X et Y obtenues par ces 18 élèves à deux épreuves relatives à ces apprentissages.

$X : \mathcal{MV}$	7	0,5	9	5,5	9,5	3	8,5	5,5	5,5	8	9	4,5	5,5	5	9,5	3	9,5	5,5
$Y : \mathcal{ME}$	6	1	9	5,5	10	5	10	7	6	10	9	6,5	5	5	10	5	9	5

- Dessiner le nuage statistique de ces variables.
- Calculer le coefficient de corrélation des rangs de Spearman des deux variables X et Y .
- Calculer les moyennes, les écarts types et le coefficient de corrélation linéaire des deux variables X et Y .
- On souhaite désormais estimer à quelle note s'attendre avec la méthode écrite pour un enfant ayant obtenu une note $x = 3,5$ avec la méthode verbale :
 - Quelle droite de régression peut-on utiliser pour répondre à cette question ? Donner son équation.
 - Si un enfant a obtenu une note $x = 3,5$, donner une estimation de sa note y .

Exercice 24 : Groupe d'étudiants

On demande à un groupe d'étudiants leur nombre de frères et sœurs, leur humeur, leur taille, et la note qu'ils ont eue au baccalauréat. On obtient les données suivantes :

Anais
de très bonne humeur 1 frère/sœur 1m69 bac : 14,6/20

Nicolas
de très bonne humeur 4 frères/sœurs 1m82 bac : 14,26/20

Catherine
de très bonne humeur 0 frère/sœur 1m65 bac : 11,07/20

Olivier
de relativement bonne humeur 2 frères/sœurs 1m74 bac : 15,97/20

Enzo
de très bonne humeur 3 frères/sœurs 1m91 bac : 12,66/20

Loïc
de très bonne humeur 2 frères/sœurs 1m74 bac : 11,14/20

Isabelle
de très bonne humeur 1 frère/sœur 1m62 bac : 15,44/20

Bernard
de mauvaise humeur 0 frère/sœur 1m78 bac : 11,89/20

Rémi
de mauvaise humeur 1 frère/sœur 1m81 bac : 13,23/20

Mélanie
de bonne humeur 4 frères/sœurs 1m60 bac : 13,07/20

Corinne
de très bonne humeur 1 frère/sœur 1m69 bac : 11,02/20

Calculer le coefficient de corrélation linéaire entre la taille et la note obtenue au bac, pour les étudiants de cet échantillon.

Exercice 25 : Thérapie pour réduire l'angoisse

Pour mettre en évidence l'efficacité d'une thérapie visant à réduire l'angoisse de personnes victimes d'agressions, nous avons observé 16 sujets avant et après la thérapie en affectant à chaque sujet un score (plus le score est élevé, plus fort est le niveau d'angoisse). Les données sont les suivantes.

Avant	33	31	39	31	23	24	37	28	23	14	31	37	23	38	19	32
Après	21	14	29	28	13	16	28	21	25	8	24	19	11	38	16	27

- Calculer les coefficients de corrélation linéaire et de Spearman entre les deux phases.
- Parmi les deux droites de régression, y en a-t-il une qui a plus d'intérêt ? Déterminer son équation.

Chapitre 3 : probabilités

Exercice 26 : Emploi

Un petit immeuble dijonnais compte 5 habitants, dont la situation d'emploi est la suivante :

Nom	Mme Bernard	Mr Martin	Mme Allard	Mr Roche	Mme Lemoine
Situation	chômage	retraite	retraite	emploi	emploi

On choisit au hasard 3 habitants parmi ces 5 habitants de l'immeuble.

- Lister tous les choix possibles de trois habitants, et conclure que chaque cas a une probabilité de 10%.
- Quelle est la probabilité d'avoir
 - exactement 2 femmes dont une est au chômage ?
 - moins de 2 hommes ?
 - deux personnes en situation d'emploi ?
- On note X le nombre de femmes parmi les trois personnes choisies au hasard.
Déterminer la loi de X (c'est à dire, calculer chaque probabilité $\mathbb{P}[X = \dots]$).

Exercice 27 : Factorielle, coefficients binomiaux

- Simplifier les quantités $\frac{17!}{15!}$, $\frac{18!}{16! \times 2!}$.
- Écrire à l'aide de deux factorielles l'expression ci-contre : $6 \times 7 \times 8$.
- Calculer les nombres suivants : $\binom{18}{2}$, $\binom{12}{3}$, $\binom{12}{9}$, $\binom{9}{0}$, $\binom{11}{1}$.

Exercice 28 : Croyance religieuse

1. Cyril et Valérie sont deux enfants nés cette année à Dijon. Compte tenu de l'environnement familial dans lequel ils vont grandir, on considère que chacun d'eux a 60% de chances de devenir athée (et donc 40% de chances de développer au contraire une croyance religieuse). Quelle est la probabilité
 - (a) Qu'ils développent tous deux une croyance religieuse.
 - (b) Qu'un seul d'entre eux devienne athée.
 - (c) Qu'ils deviennent tous les deux athées.

Argumenter que ces probabilités correspondent à une loi binomiale, dont vous donnerez les paramètres.

2. Au sein d'une famille de 5 personnes, on constate que 2 personnes ont une croyance religieuse (les 3 autres personnes sont donc athées). On choisit au hasard 2 personnes au sein de cette famille, et on note X le nombre de personnes athées parmi ces deux personnes choisies au hasard.
Calculer la loi de X .

3. Au sein d'une ville de 2000 personnes, on constate que 40% (c'est à dire 800 personnes) ont une croyance religieuse. On choisit au hasard 2 personnes au sein de cette ville, et on note X le nombre de personnes athées parmi ces deux personnes choisies au hasard.
Calculer la loi de X .

4. Comparer entre elles les lois obtenues dans ces trois situations.

Exercice 29 : Loi binomiale

On considère une loi binomiale $X \sim \mathcal{B}(n, p)$ de moyenne $m(X) = 4,8$ et de variance $\text{Var}(X) = 2,88$.

- 1) Calculer n et p .
- 2) Calculer les probabilités $\mathbb{P}[4 \leq X \leq 6]$ et $\mathbb{P}[X > 9]$.
- 3) Trouver toutes les valeurs de k telles que $\mathbb{P}[X \geq k] \leq 0,05$.

Exercice 30 : Troubles de l'humeur et emploi

On considère trois personnes souffrant de troubles de l'humeur : Alice, Bernard et Cécile. Parmi eux, Alice et Bernard exercent un emploi alors que Cécile est sans emploi.

1. On choisit au hasard le nom d'une de ces trois personnes, puis à nouveau le nom d'une de ces trois personnes au hasard (ce peut être la même personne – ou pas).
 - (a) Lister les neuf possibilités pour ces deux noms choisis au hasard.
 - (b) Parmi ces possibilités, combien comptent uniquement des personnes en situation d'emploi ?
 - (c) Quelle est la proportion, parmi ces possibilités, qui comptent une personne sans emploi, et une qui exerce un emploi ?
 - (d) Pour un tel choix aléatoire de deux noms, on note X le nombre de noms choisis qui correspondent à une personne exerçant un emploi. Quelle est la loi de variable X ?
 - (e) Retrouver le résultat de la question 2, en utilisant cette loi.
 - (f) Déterminer de même $\mathbb{P}[X = 0]$ et $\mathbb{P}[X = 2]$.
2. On décide d'une autre façon de choisir deux noms : on choisit un premier nom au hasard, puis on choisit le second nom en demandant qu'ils soit différent du premier. Reprendre, avec cette nouvelle façon de choisir les deux noms, les questions 1a, 1b et 1c. On prendra garde qu'en question 1a le nombre de cas ne sera plus égal à neuf.

Exercice 31 : Collection de dessins

Mme Chauvet est pédiatre et utilise fréquemment de petites illustrations pour stimuler les enfants. Elle a acheté à cet effet 5 cartes illustrées qu'elle stocke dans un tiroir.

1. Elle mélange les cartes, et constate dans quelle ordre elles se retrouvent à l'issue du mélange. Quelle était la probabilité qu'à l'issue du mélange, elles se retrouvent précisément dans cet ordre ?
2. Si elle remélange les cartes, quelle est la probabilité qu'après le mélange
 - (a) la carte qui était en haut du tas soit à nouveau en haut du tas ?
 - (b) les deux cartes qui étaient en haut du tas soient à nouveau en haut du tas, et dans le même ordre ?
 - (c) les deux cartes qui étaient en haut du tas soient à nouveau en haut du tas, mais dans l'ordre inverse ?
 - (d) les deux cartes qui étaient en haut du tas soient à nouveau en haut du tas, dans n'importe quel ordre ?
 - (e) l'ensemble des cartes soient exactement dans le même ordre qu'avant ?
3. Répondre aux mêmes questions en supposant qu'il y ait cette fois-ci 35 cartes.
4. Peut-on conclure des probabilités calculées que certains événements ont particulièrement peu de chances de se produire ?

Exercice 32 : Ressources humaines

Dans une grande entreprise, la direction demande aux personnes exerçant des responsabilités de noter les employés placés sous leur responsabilité, et leur impose de donner la note A (la meilleure note) à 20% des employés, la note B à 30% des employés et la note C (la moins bonne note) à 50% des employés.

On choisit parmi les employés de l'entreprise, un échantillon de 25 employés au hasard "avec remise".

1. Calculer la probabilité que l'échantillon contienne exactement 3 employés notés A .

2. Quelle est la probabilité que l'échantillon contienne au moins 3 individus notés A ?
3. Quelle probabilité y a-t-il que l'échantillon contienne au moins 12 individus notés A ou B ?
4. Calculer la probabilité que l'échantillon contienne moins que 8 personnes notées C .
5. Quel est le nombre moyen d'individus notés B au sein d'un échantillon aléatoire de 25 employés ?

Exercice 33 : Thérapie contre la dépression

Un psychologue a développé une nouvelle thérapie contre la dépression, et il affirme qu'elle permet la rémission de 70% des patients.

En interrogeant un échantillon de 20 patients, on constate 8 rémissions.

- 1) S'il y avait 70% de rémission parmi l'ensemble des patients, quelle loi suivrait le nombre de rémissions au sein d'un échantillon de 20 patients choisis au hasard avec remise ? Quel serait le nombre moyen de rémissions au sein d'un tel échantillon ?
- 2) Sous cette hypothèse, quelle serait la probabilité d'avoir au maximum 8 rémissions ?
- 3) Conclure : vous semble-t-il vraisemblable qu'il y ait, comme l'affirme ce psychologue, 70% de rémissions parmi l'ensemble des patients ?

Exercice 34 : Nombre de filles et de garçons

Dans une classe, il y a 16 filles et 12 garçons. On choisit au hasard 4 élèves distincts.

1. Quel est le nombre de choix possibles ?
2. Quel est le nombre de choix ne comportant que des garçons ?
3. Quelle est la proportion de choix ne comportant que des filles ?
4. Quelle est la probabilité, en choisissant ainsi un échantillon de 4 élèves au hasard, que cet échantillon comporte 3 filles et 1 garçon ?
5. Quel est le nombre de choix comportant au plus un garçon ?

Chapitre 4 : Loi Normale

Exercice 35 : Loi normale centrée réduite

On désigne par Z la loi normale centrée réduite (de densité $f(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})$). Pour $z > 0$, on pose

$$F(z) = \mathbb{P}[Z \leq z] = \int_{-\infty}^z f(x) dx.$$

- Pour cette question, on pose $z_1 = 1,05$; $z_2 = 1,92$ et $z_3 = 0,588$.
 - Déterminer $F(z_1)$, $F(z_2)$ et $F(z_3)$. Chacune de ces probabilités peut être obtenue à partir de la table du formulaire, ou directement en utilisant les fonctions spécifiques de la calculatrice. Assurez-vous de savoir utiliser chacune de ces deux méthodes.
 - Calculer : $\mathbb{P}[z_1 < Z \leq z_2]$, $\mathbb{P}[-z_2 < Z < -z_1]$, $\mathbb{P}[z_1 \leq Z < 5]$, $\mathbb{P}[-z_2 \leq Z \leq z_1]$, $\mathbb{P}[-z_1 < Z < z_1]$, $\mathbb{P}[Z > z_2]$ et $\mathbb{P}[Z < z_1]$. On peut calculer ces probabilités à partir des valeurs obtenues dans la question a) ou en utilisant directement les fonctions spécifiques de la calculatrice. Assurez-vous de savoir utiliser chacune de ces deux méthodes.
- Déterminer des valeurs approchées de z_4 , z_5 , et z_6 tels que $F(z_4) = 0,7967$; $F(z_5) = 0,0256$ et $F(z_6) = 0,8782$.
 - Déterminer des valeurs approchées de z_7 , z_8 , et z_9 tels que $\mathbb{P}[Z \leq z_7] = 0,99$; $\mathbb{P}[Z > z_8] = 0,9515$; $\mathbb{P}[Z < z_9] = 0,9345$. On peut le faire à partir de la table du formulaire, ou bien obtenir une réponse directe de la calculatrice. Assurez-vous de savoir utiliser chacune de ces deux méthodes.

Exercice 36 : Images mentales

Dans une expérience sur les images mentales, on demande à un échantillon de 337 enfants d'apprendre une liste de 40 mots en leur montrant, pour chaque mot, un dessin de l'objet correspondant au mot. On désigne par X le nombre dont se souvient chaque enfant un jour après. On obtient :

Nb de mots X	[5; 10[	[10; 15[	[15; 20[	[20; 25[	[25; 30[	[30; 35[	[35; 40[
Effectif	7	70	114	115	28	1	2

- Calculer la moyenne m et l'écart-type s de la variable statistique X .
- Quels seraient les effectifs théoriques si X suivait la loi normale $\mathcal{N}(m; s)$?

Exercice 37 : Désirabilité sociale

On étudie les scores (notés X) obtenus par des enfants sur une échelle de désirabilité sociale (D.S.). Une étude a montré que pour cette échelle, les scores ont, sur l'ensemble des enfants de CM2, une moyenne $\mu \simeq 10,42$ et un écart-type $\sigma \simeq 2,18$.

Dans la suite, on effectue des calculs en supposant que X suit la loi normale $\mathcal{N}(10,42; 2,18)$.

- Calculer la proportion théorique des scores supérieurs à 10,97.
- Trouver la valeur de a telle que $\mathbb{P}[X < a] \simeq 0,95$.
- Trouver les quartiles de la variable X .
- Quel est le score minimal des 20 % de ceux qui ont les scores D.S. les plus élevés ?
- Quel est le score maximal des 35 % de ceux qui ont les scores D.S. les moins élevés ?

Exercice 38 :

Lors d'un concours le pourcentage d'échecs est de 55%. On choisit au hasard (avec remise) un échantillon de n étudiants et on désigne par S_n le nombre d'échecs au sein de l'échantillon.

- On prend $n = 16$. Calculer $p_k = \mathbb{P}[S_{16} \leq k]$ pour $k = 0; 1; 2; 3$.

2. Pour $n = 100$, on souhaite calculer $\mathbb{P}[50 \leq S_{100} \leq 60]$:
 - a) Calculer $\mathbb{P}[50 \leq S_{100} \leq 60]$ avec la calculette.
 - b) Argumenter que l'on peut approcher la loi de S_{100} par une loi normale, et calculer $\mathbb{P}[50 \leq S_{100} \leq 60]$ de manière approchée.
On vérifiera que la précision de l'approximation est bien meilleure si on fait la *correction de continuité*.
3. Pour $n = 125$, déterminer de même $\mathbb{P}[59 \leq S_{125} \leq 79]$ (on pourra utiliser l'approximation par une loi normale en faisant une correction de continuité).
4. Pour $n = 125$, trouver une valeur de a telle que $\mathbb{P}[S_{125} < a] \simeq 0,03$ (pour simplifier les calculs, on pourra omettre la correction de continuité dans cette question).

Exercice 39 : Indice de masse corporelle dans la population française

L'indice de masse corporelle mesure la corpulence d'un individu. Dans cet exercice on le notera X . Il s'agit d'un élément de diagnostic de dénutrition ou d'obésité, dont l'interprétation est la suivante :

$X < 16,5$ dénutrition ou famine	$25 \leq X < 30$ surpoids
$16,5 \leq X < 18,5$ maigreur	$30 \leq X < 35$ obésité modérée
$18,5 \leq X < 25$ corpulence normale	$35 \leq X < 40$ obésité sévère
	$X \geq 40$ obésité morbide

Une enquête épidémiologique conduite en 2012 par l'INSERM, KANTAR HEALTH et ROCHE, donne la répartition suivante de l'IMC au sein d'un échantillon de 25 714 personnes, supposé représentatif de la population française :

IMC	$[-\infty; 18,5[$	$[18,5; 25[$	$[25; 30[$	$[30; 35[$	$[35; 40[$	$[40; +\infty[$
fréquence (en %)	3,5	49,2	32,3	10,7	3,1	1,2

1. Déterminer la probabilité (« *fréquence théorique* ») qu'aurait chaque classe si X suivait la loi normale $\mathcal{N}(25; 5)$.

2. En particulier, si X suit la loi normale $\mathcal{N}(25; 5)$, quelle est la proportion (*théorique*) d'obésité sévère ou morbide (c'est à dire $\mathbb{P}[X \geq 35]$) ?
3. Au contraire, quelle est dans l'échantillon de cette étude épidémiologique, la proportion d'obésité sévère ou morbide (c'est à dire $\mathbb{P}_r[X \geq 35]$) ?
4. On cherche désormais à comparer les résultats des questions 2 et 3 : on souhaite déterminer si la proportion $\mathbb{P}_r[X \geq 35]$ est compatible avec la probabilité $\mathbb{P}[X \geq 35]$ obtenue sous l'hypothèse d'une loi normale.
 - (a) Si on suppose qu'une personne choisie au hasard a une probabilité de 2,3% d'être en obésité sévère ou morbide, quelle est la loi du nombre de personnes en obésité sévère ou morbide dans un échantillon de 25 714 français ? Le fait que l'échantillon soit *avec* ou *sans remise* a-t-il la moindre importance ?
 - (b) Cette loi peut-elle être approximée par une loi normale, et si oui laquelle ?
 - (c) En utilisant cette approximation par une loi normale, calculer la probabilité qu'un échantillon aléatoire de 25 714 personnes contienne au moins 4,3% de personnes en obésité sévère ou morbide.
 - (d) Comparer avec l'échantillon de l'étude épidémiologique mentionnée en début d'exercice.
Conclure qu'en 2012, l'IMC des français ne suivait pas la loi de probabilité $\mathcal{N}(25; 5)$.

Exercice 40 : Loi normale avec moyenne et écart type

1. Pour une loi normale centrée réduite, quelle est la probabilité qu'une observation soit :
 - (a) supérieure ou égale à 1,5 ?
 - (b) comprise entre -1 et 1 ?
 - (c) comprise entre -2 et 2 ?
 - (d) comprise entre -3 et 3 ?

2. Pour une loi normale de moyenne $\mu = 123$ et d'écart-type $\sigma = 9$, (c'est-à-dire $X \rightsquigarrow \mathcal{N}(123; 9)$), quelle est la probabilité qu'une observation soit :

- (a) supérieure ou égale à $\mu + 1,5\sigma$?
- (b) comprise entre $\mu - \sigma$ et $\mu + \sigma$?
- (c) comprise entre $\mu - 2\sigma$ et $\mu + 2\sigma$?
- (d) comprise entre $\mu - 3\sigma$ et $\mu + 3\sigma$?

3. Reprendre les questions précédentes avec la loi normale $\mathcal{N}(12,53; 1,34)$. Que remarque-t-on ?

Exercice 41 :

On évalue les niveaux de dépression au moyen du test Inventaire Multiphasique de la Personnalité du Minnesota (MMPI). Au vu du nombre important de questions du MMPI, on considère que quand on fait passer ce test à un individu choisi au hasard, son score suit une loi normale. De plus le MMPI est normalisé pour avoir une moyenne $\mu = 50$ et un écart-type $\sigma = 10$.

Si l'on considère qu'une note supérieure à 70 traduit un état pathologique, combien s'attend-on à trouver de personnes pathologiquement dépressives sur un ensemble de 50 000 personnes ?

Exercice 42 :

La proportion des enfants qui réussissent parfaitement une épreuve graphique d'organisation perceptive est de 60%. On choisit au hasard (avec remise) un échantillon de n individus et on note X la variable aléatoire égale au nombre d'enfants dans l'échantillon qui réussissent l'épreuve.

- 1. On prend $n = 28$. Préciser la loi de probabilité de X et calculer la probabilité : $\mathbb{P}[15 \leq X \leq 18]$.
- 2. On prend $n = 280$.

- (a) On peut approcher la loi de X par une loi normale. Justifier et préciser laquelle. Dans la suite, si vous utilisez cette approximation, il faudra faire une correction de continuité.
- (b) Calculer les probabilités $\mathbb{P}[X \geq 180]$ et $\mathbb{P}[150 \leq X \leq 185]$.

Exercice 43 : Pile ou face Vous jouez à « pile » ou « face » avec une pièce de monnaie équilibrée. Avec quelle probabilité la fréquence de « pile » est-elle comprise entre 0,48 et 0,52 si vous lancez la pièce :

- i) 100 fois.
- ii) 1000 fois.
- iii) 2000 fois.
- iv) 10000 fois.

Exercice 44 :

On considère que le nombre de potentiels d'action (PA) émis en une minute par un neurone au repos suit une loi normale de moyenne 100 et d'écart-type 10.

- 1. Quelle est la probabilité pour qu'un neurone émette plus de 110 PA en 1 mn s'il n'est pas activé ?
- 2. Quelle est la probabilité pour qu'un neurone émette moins de 80 PA en 1 mn s'il n'est pas activé ?
- 3. Quelle est la probabilité pour qu'un neurone émette entre 90 et 110 PA en 1 mn s'il n'est pas activé ?
- 4. Un neurone émet plus de 150 PA en une minute. Selon vous, est-il plus vraisemblable que ce neurone était activé ou au repos lors de l'enregistrement ? Justifiez votre réponse.

Chapitre 5 : Estimation

Exercice 45 : Pile ou face

Vous jouez à « pile » ou « face » avec une pièce de monnaie. Vous la lancez 800 fois et obtenez 443 fois pile.

1. Estimez, avec la confiance 99%, la probabilité qu'a cette pièce de tomber sur pile quand on la lance.
2. Conclure : pouvez-vous affirmer (avec la confiance 99%) que cette pièce soit biaisée ?

Exercice 46 : Efficacité d'un traitement

Un fabricant de médicaments affirme qu'en 10 jours, 40% des malades qui sont traités par l'un de ses produits guérissent de leur maladie.

Pour vérifier cette affirmation du fabricant, on l'administre à 800 malades, et on constate que parmi cet échantillon de 800 patient traités, il y en a 447 qui guérissent en 10 jours.

1. Donner une estimation de la vraie proportion p de guérisons (pour l'ensemble des patients que l'on traiterait avec ce médicament) avec une confiance de 95%.
2. Que peut-on conclure par rapport à l'affirmation du fabricant ?
3. On veut réduire la marge de l'estimation à 1% tout en atteignant une confiance de 98%. Quelle doit alors être la taille minimale de l'échantillon à considérer ?

Exercice 47 : Satisfaction de clients

Afin de répondre au mieux au désir de sa clientèle, une entreprise réalise un sondage pour connaître l'avis des consommateurs sur un produit qu'elle fabrique. Sur 1218 personnes interrogées, 618 se sont déclarées satisfaites. L'entreprise essaye alors d'améliorer la qualité de son produit, puis elle réalise un deuxième sondage un an après : sur 905 personnes interrogées, il y en a alors 482 qui sont satisfaites.

1. On désigne par p_1 et p_2 les proportions de satisfaits pour l'ensemble des clients avant et après l'amélioration apportée. Donner une estimation de p_1 et p_2 par intervalle de confiance en prenant comme confiance $c = 0,9$. Peut-on dire que la proportion de satisfaits pour l'ensemble de la clientèle a augmenté ?
2. Lors du deuxième sondage, quelle aurait du être la taille de l'échantillon pour avoir une estimation de p_2 à 1% près avec une confiance de 0,99 ?

Exercice 48 : Résistance à la persuasion

On pense que la résistance à la persuasion passe par une réaction active des sujets qui développent intérieurement des contre-arguments. Pour tester cette hypothèse on considère l'opinion qu'ont les étudiants sur la coopération entre étudiants et enseignants pour établir les programmes. Cette opinion est mesurée par un questionnaire pour lequel, sur l'ensemble des étudiants, la moyenne est 13,7 et l'écart type est 2,4.

1. On soumet dans un premier temps un échantillon de 50 étudiants à un argumentaire persuasif contre cette coopération, puis on mesure leur opinion à l'aide du même questionnaire. On obtient les notes suivantes :

Classe	[7 ; 9[	[9 ; 11[	[11 ; 13[	[13 ; 15[	[15 ; 17[	[17 ; 19[
Effectif	5	6	15	17	4	3

- a) Calculer la moyenne expérimentale $m_e(X)$ et l'écart type expérimental $s_e(X)$ de l'échantillon.
- b) Estimer pour l'ensemble des étudiants la note moyenne après avoir entendu un tel argumentaire (on l'appelle $\mu(X)$). On déterminera un intervalle de confiance, avec la confiance $c = 0,95$.
- c) Peut-on dire avec un risque d'erreur de 5% qu'en moyenne, la note (traduisant l'opinion en faveur d'une coopération entre étudiants et enseignants) a diminué après l'argumentaire ? Justifiez votre réponse.

- d) Quelle devrait être la taille minimale de l'échantillon à prélever pour estimer la moyenne $\mu(X)$ à 0,5 points près avec une confiance de 0,99 ?
2. On considère désormais un échantillon de 26 étudiants à qui on demande, pendant qu'ils écoutent l'argumentaire, de réaliser des « opérations » arithmétiques (réciter mentalement la table de multiplication par 8). La moyenne et l'écart type expérimentaux sont alors $m_e = 9$ et $s_e = 4,2$.
- a) Avec une confiance de 0,95 déterminer la note moyenne de tous les étudiants s'ils écoutaient l'argumentaire en récitant des tables de multiplication. (*On pourra supposer que les notes suivent une loi normale.*)
- b) Comparer cette moyenne à celles qu'ils auraient sans réciter les tables de multiplication (en étant donc en mesure de réfléchir à des contre-arguments). Ces résultats sont-ils compatibles avec l'hypothèse théorique que l'on souhaitait tester ?
- c) Donner une estimation de l'écart type des notes, dans le deuxième cas (en récitant les tables de multiplication) avec un risque d'erreur de 5%.

Exercice 49 : Motivation lors d'un test

Un expérimentateur qui étudie la prise de décisions a demandé à 30 enfants de résoudre le plus grand nombre possible de problèmes en 30 minutes. Il a expliqué à un groupe de 16 enfants qu'il voulait tester leur aptitude innée à résoudre des problèmes, et à un autre groupe (14 enfants) qu'il ne s'agissait que d'une tâche destinée à les occuper. Les résultats obtenus sont les suivants :

Aptitude innée	33	18	27	21	27	22	26	25	33	19	29	30	14	20	32	27
Occupation	12	15	24	9	21	12	17	14	24	9	20	15	18	16		

On suppose que les résultats varient normalement.

1. Calculer les moyennes et les écarts types des deux échantillons.

2. Donner une estimation par intervalle de confiance des moyennes et des écarts types associés aux deux conditions d'expérience. *On fera cette estimation avec une confiance de 95%.*
3. Que peut-on conclure au vu de ces résultats ?

Exercice 50 : Taux de cholestérol

On étudie les taux de cholestérol dans le sang chez les hommes et les femmes de plus de 50 ans. On suppose qu'ils suivent des lois normales.

Dans un échantillon de 18 hommes de plus de 50 ans on a mesuré un taux moyen de 192,4mg/dL avec un écart type de 45,2mg/dL.

Dans un échantillon de 22 femmes de plus de 50 ans on a mesuré un taux moyen de 185,7mg/dL avec un écart type de 36,4mg/dL.

1. Donner des estimations des écarts types des deux populations.
2. En utilisant les estimations par intervalle de confiance à 95%, peut-on conclure que les femmes de plus de 50 ans ont en moyenne moins de cholestérol que les hommes de plus de 50 ans ?

Exercice 51 : Épreuve de dictée

Le tableau suivant représente les résultats obtenus par un échantillon de 15 enfants de CE2 dans deux épreuves de dictée préparée : une liste de 20 mots (épreuve $\mathcal{E}_1$) et un texte (épreuve $\mathcal{E}_2$).

Élève	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Épreuve 1	19	17,5	18	15	15,5	17,5	16,5	16	17	12	14,5	18,5	15	13	14
Épreuve 2	14	13	13,5	11,5	12	13	12,5	12	13	9	11	14	11,5	10	11

1. Peut-on affirmer, avec un risque d'erreur de 5%, qu'une des deux épreuves donne, pour des enfants de CE2, une note moyenne plus élevée que l'autre épreuve ?
2. Donner des estimations des écarts types.

Exercice 52 :

Un psychologue étudie le taux de chômage de patients atteint de boulimie, afin de déterminer s'il diffère du taux de chômage de l'ensemble de la population.

1. En France, le taux de chômage est de 10% (selon les données de l'INSEE).

On choisit au hasard un échantillon de 117 français. Quelle est la probabilité d'avoir, au sein de cet échantillon, entre 9% et 11% de chômeurs ?

Justifiez bien les éventuelles approximations que vous seriez amené à faire

2. Sur un échantillon de 117 patients atteints de boulimie, le psychologue a constaté que 5,98% d'entre eux étaient au chômage.

- (a) Avec la confiance $c = 90\%$, déterminer un intervalle de confiance pour le taux de chômage des personnes boulimiques.

Peut on en déduire (avec la confiance 90%) que le taux de chômage des personnes boulimiques est différent de celui du reste de la population ?

- (b) Avec la confiance $c = 98\%$, estimer la proportion de chômeurs au sein de l'ensemble des personnes atteintes de boulimie.

Peut on en déduire (avec le risque d'erreur $\alpha = 2\%$) que le taux de chômage des personnes boulimiques est différent de celui du reste de la population ?

- (c) Pour pouvoir estimer à 1% près le taux de chômage des personnes boulimiques avec une confiance 98%, quelle taille d'échantillon faudrait il considérer ?

Exercice 53 : Une chercheuse, Mme Lefebvre, se demande si la perception des couleurs diffère en fonction du sexe des individus.

1. On étudie tout d'abord la prévalence du Daltonisme.

On considère d'une part un échantillon de 1000 femmes, parmi lesquelles 7 sont daltoniennes, et d'autre part un échantillon de 1000 hommes, parmi lesquels 94 sont daltoniens.

Avec la confiance $c = 99\%$, peut on déduire que la prévalence du daltonisme soit plus importante chez l'homme que chez la femme ?

2. La chercheuse pense que, bien au delà de la question du daltonisme, les filles sont bien plus douées que les garçons pour détecter d'infimes nuances de couleurs. Elle teste cette hypothèse sur des enfants de 7 ans, auxquels elle fait passer un test standardisé de perception des couleurs. Sur un échantillon de 15 filles, elle obtient les données suivantes :

Score	7	6	8	0	6	6	5	2	7	7	6	4	7	6	7
-------	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Estimer, avec la confiance 98%, le score moyen des filles à ce test.

3. Sur un échantillon de 40 garçon auxquels la chercheuse fait passer le test, le score moyen est de 4,27 avec un écart-type de 1,73.

Estimer, avec la confiance 98%, le score moyen des garçons à ce test.

Peut on affirmer, avec la confiance 98%, que – comme s'y attendait la chercheuse – les filles ont en moyenne une meilleure perception des couleurs que les garçons ?

4. Estimez de même les écart-types des garçons et des filles, avec la confiance 98%. Peut on conclure qu'ils diffèrent selon le sexe ?
5. Quelle taille d'échantillon aurait-il fallu considérer pour estimer le score moyen des filles à 0,25 point près, avec une confiance 98% ?

Remarque : Une grande partie des données de ces exercices sont fictives, elles visent simplement à illustrer les outils mathématiques introduits en cours.