

Chapitre 7

Statistiques

Statistiques descriptives

7.1 Dans chaque situation exposée ci-dessous,

- a) décrire la population étudiée ;
- b) décrire l'échantillon ;
- c) nommer la variable étudiée ;
- d) décrire l'ensemble des catégories ou des valeurs de la variable ;
- e) donner le type de variable étudiée.

Situation 1

On effectue un sondage auprès de 500 habitants de la ville de Lausanne pour connaître leur chaîne de télévision favorite.

Situation 2

Dans une étude portant sur l'évolution de la situation économique en Suisse de 2000 à 2010, on s'intéresse au taux de chômage annuel de cette décennie.

Situation 3

Afin de déterminer le profil socioéconomique des ménages de la ville de Genève, on a noté le nombre d'enfants par ménage pour un échantillon de 380 ménages.

Situation 4

Selon les données du recensement helvétique de l'an 2000, à la question « Quelle est la langue dans laquelle vous pensez et que vous savez le mieux ? »,

- 63.7% de la population a répondu « l'allemand » ;
- 20.4% de la population a répondu « le français » ;
- 6.5% de la population a répondu « l'italien » ;
- 0.5% de la population a répondu « le romanche » ;
- et 9.0% de la population a cité une langue non nationale ;

7.2 Donner le type de chacune des variables suivantes :

- a) La superficie des lacs de Suisse.
- b) Le pays d'origine des touristes qui visitent la Suisse.
- c) Le nombre d'étudiants dans les gymnases vaudois.
- d) La longueur d'un crayon
- e) Prenez-vous le train chaque semaine ?
 - (a) Oui
 - (b) Non
- f) Ressentez-vous du stress avant de prendre l'avion ?
 - (a) Toujours
 - (b) Souvent
 - (c) Parfois
 - (d) Rarement
 - (e) Jamais

7.3 Pour chaque question, indiquer le type de variable et l'échelle de mesure.

- a) Avez-vous au moins une note insuffisante dans votre bulletin semestriel ?
 1. Non
 2. Oui
- b) Combien de notes insuffisantes avez-vous dans votre bulletin semestriel ?
 1. 0
 2. 1
 3. 2 ou 3
 4. 4 et plus
- c) Combien de note insuffisante avez-vous dans votre bulletin semestriel ?
- d) Quel est votre taux d'échec au semestre ?

$$\left(\text{taux d'échec} = \frac{\text{Nombre de notes insuffisantes}}{\text{Nombre total de notes}} \right)$$

1. 0%
 2. De 1% à 15.9%
 3. De 16% à 49.9%
 4. 50% et plus
- e) Quel est votre taux d'échec au semestre ?
 - f) Dans quelle mesure êtes-vous d'accord avec l'affirmation « Les élèves ayant plus de quatre notes insuffisantes en fin de première année devraient arrêter leurs études gymnasiales ».
 1. Fortement en désaccord
 2. En désaccord
 3. Plutôt d'accord
 4. Totalemment d'accord
 - g) Quelle est votre année de naissance ?

7.4 Dans le bulletin météo du journal local, on trouve notamment pour chaque jour de la semaine l'heure du lever du soleil et la vitesse des vents. Indiquer le type de chacune des deux variables et l'échelle de mesure qui lui est associée.

7.5 D'après l'office fédéral de la statistique, les blessés légers victimes d'accidents de la route en 2013 se répartissent par moyen de locomotion de la façon suivante :

Moyen de locomotion	Blessés légers
Voiture de tourisme	9570
Motocycle	2479
Bicyclette	2435
Piétons	1570
Autres	1196

- Représenter cette situation à l'aide d'un diagramme circulaire.
- Faire de même à l'aide d'un diagramme en barres.
- Peut-on déduire de ces chiffres qu'il est moins dangereux de se déplacer en moto plutôt qu'en voiture ?

7.6 Le nombre de véhicules à moteur mis en circulation en Suisse en 2011 est donné par catégorie dans le tableau suivant :

Catégorie	Nombre
Voitures de tourisme	327'955
Véhicules de transport de personnes	3'950
Véhicules de transport de choses	33'119
Véhicules agricoles	3'714
Véhicules industriels	4'006
Motocycles	48'133
Total des véhicules	420'875

Source : Office fédéral de la statistique, site web Statistique suisse 2012

Représenter ces données graphiquement par un diagramme à rectangles horizontaux et par un diagramme circulaire. Laquelle de ces deux représentations est-elle la plus appropriée ?

7.7 Lors d'un sondage, on a demandé à 820 citoyens suisses leur opinion sur les accords bilatéraux Suisse-UE. Les réponses se répartissent comme suit.

Répartition de selon

Utilité	Effectifs	Pourcentage
Très utiles	95	
Utiles	342	
Nuisibles	210	
Très nuisibles	46	
Sans opinion	127	
Total		

- a) Décrire la population étudiée, nommer la variable considérée et le type d'échelle de mesure.
- b) Compléter le tableau de distribution ci-dessus ainsi que son titre.
- c) Représenter graphiquement la distribution par un diagramme approprié au type de variable.
- d) Calculer le taux de confiance en ces accords, soit le pourcentage de personnes qui estiment les accords bilatéraux utiles ou très utiles.

7.8 Donner la première des classes qui permettraient de grouper une série de 36 données, précises au centième, sachant que la plus petite valeur est 2,65 et la plus grande 18,45.

7.9 On désire grouper en classes les revenus de 80 stagiaires. Le plus petit revenu est de 252 francs et le plus grand de 937 francs. Donner la première classe de la distribution des revenus.

7.10 On a récolté les données suivantes :

314	473	500	812	566	212	606	935	247	474	993	432
262	1080	972	383	975	978	366	322	638	570	1094	270
813	227	950	1030	776	503	398	398	755	650	1008	711
563	930	1054	836	631	519	1019	299	1032	500	918	979
570	592	1023	859	759	990	964	598	1097	803	998	337

- a) Grouper les données en 6 classes de largeur 150 ($[200;350[$, $[350;500[$, etc.) et dresser un tableau de distribution.
- b) Construire l'histogramme et le polygone des fréquences.

7.11 Sur une route où la vitesse est limitée à 80 km/h, on a mesuré la vitesse de 50 véhicules.

84	81	76	71	80	81	83	84	80	83
74	75	92	76	80	82	94	73	83	83
75	81	79	97	78	82	76	78	82	82
78	81	91	68	82	73	82	79	75	77
83	80	77	81	69	78	81	83	87	87

- a) Grouper les données en classes **fermées à droites** et dresser un tableau de distribution.
- b) Construire l'histogramme et le polygone des fréquences.
- c) Compléter l'analyse suivante : « Une des véhicules respectent la limitation de vitesse de 80 km/h et % roulent entre 80 et 85 km/h. En

tenant compte d'une marge de tolérance de 5 km/h, % des véhicules sont amendables. »

d) Pourquoi a-t-on fermé les classes à droite dans ce contexte ?

7.12 Le tableau ci-dessous met en parallèle la distribution de l'âge des Suisses en 1860, lors du premier recensement, et en 2009.

Répartition de la population suisse en 1860 et 2009 selon l'âge.

Âge	1860		2009	
	Effectif	Pourcentage	Effectif	Pourcentage
] 0 ; 10]	518'538	20.6%	763'546	9.8%
] 10 ; 20]	476'347	18.9%	872'579	11.2%
] 20 ; 30]	429'507	17.1%	978'050	12.6%
] 30 ; 40]	362'978	14.4%	1'096'126	14.1%
] 40 ; 50]	287'564	11.4%	1'277'392	16.4%
] 50 ; 60]	230'276	9.2%	1'031'892	13.3%
] 60 ; 70]	138'932	5.5%	840'583	10.8%
] 70 ; 80]	59'549	2.4%	554'034	7.1%
] 80 ; 90]	11'095	0.4%	311'195	4.0%
90 et plus	610	0.0%	60'409	0.8%
Total	2'515'396	99.9% *	7'785'806	100.1% *

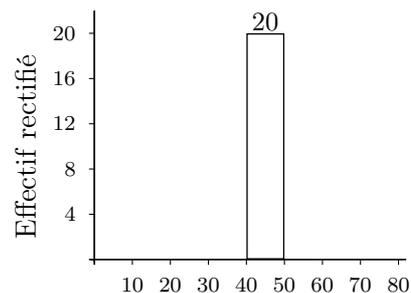
Source : Office fédéral de la statistique, site web Statistique suisse 2012.

*Les pourcentage totaux ne sont pas exactement égaux à 100% à cause des arrondis.

- a) Quelle représentation graphique mettrait le mieux en évidence les différences de distribution des deux années étudiées ? Justifier la réponse.
- b) Représenter sur un même graphique le polygone des fréquences relatives de ces deux années.
- c) Compléter le texte suivant :
- « La population suisse a plus que entre 1860 et 2009 en passant de à presque d'habitants. En 1860, la population était très jeune : l'aire sous le polygone est plus grande avant ans qu'après. A cette époque, seulement% des habitants avaient plus de 70 ans, contre% actuellement, soit une proportion fois plus élevée. En 1860, le groupe des moins de 20 ans représentait près de% de la population contre% aujourd'hui, soit une proportion réduite de En 1860, la classe la plus représentée est celle des , avec% des habitants, alors qu'en 2009, c'est la classe des avec% des habitants. »

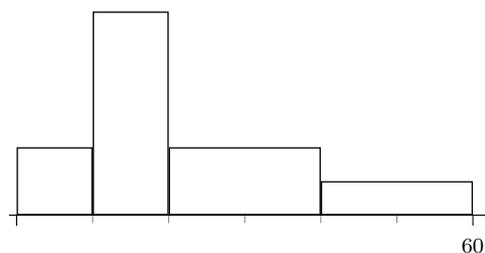
7.13 a) Compléter l'histogramme de la distribution suivante :

Amplitude	Classe	Effectif	Effectif rectifié
	[10; 40 [12	
	[40; 50 [20	
	[50; 60 [18	
	[60; 80 [10	
	Total	60	

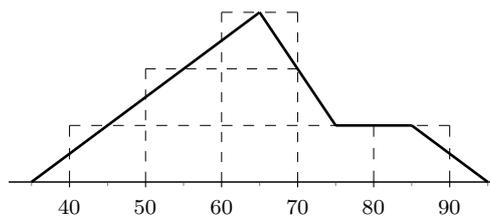


b) Compléter le tableau de distribution en utilisant l'information donnée par l'histogramme.

	Pourcentage
[0 ; [
[; [
[; [
[; 60 [
Total	100%



c) Le polygone de fréquences ci-contre représente une distribution. Quel est le pourcentage des données ayant une valeur comprise entre 50 et 60 ?



7.14 Dans une usine, lors d'un contrôle qualité, le diamètre, en mm, de 50 boulons tirés au hasard dans la production a été mesuré. Les résultats suivants ont été obtenus.

Répartition de selon

Diamètre [mm]	Effectifs
[21.5; 21.8[4
[21.8; 21.9[6
[21.9; 22.0[6
[22.0; 22.1[13
[22.1; 22.2[8
[22.2; 22.3[7
[22.3; 22.5[6
Total	50

- a) Décrire la population étudiée, nommer la variable considérée, le type de la variable et l'échelle de mesure. Compléter le titre du tableau de distribution.
- b) Représenter l'histogramme de ces données.
- c) Représenter le polygone des fréquences cumulées.
- d) Si la valeur nominale du diamètre des boulons est de 22 mm, calculer le pourcentage de boulons qui s'en écartent de plus de 0.3 mm ? Vérifier la cohérence du résultat sur le polygone des fréquences cumulées.

7.15 Déterminer la moyenne, la médiane et le mode ou la classe modale de chaque jeu de données.

a)

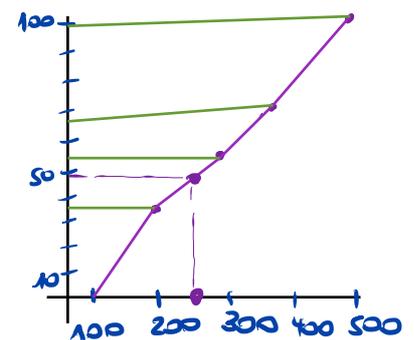
valeur	1	2	3	4	5	6
effectif	1	3	5	5	7	4

- b) -0.5 0.1 0.9 0.3 0.2 -0.6 0 -1.0 0.7 -0.1

c)

classe	[100; 200[[200; 300[[300; 400[[400; 500[
fréquence	32%	20%	12%	36%

F_i 32% 52% 64% 100%



d)

classe	[0; 2[[2; 3[[3; 4[[4; 5[[5; 10[
effectif	12	48	94	42	4

Moyenne : $\frac{150}{0,32} + \frac{250}{0,2} + \frac{350}{0,12} + \frac{450}{0,36}$

Classe modale : [400; 500[

Pour atteindre 50%, il faut ajouter 18% à 32%.

20% ↔ 100 = l
 18% ↔ x
 $x = \frac{100 \cdot 18}{20} = 90$
 $\tilde{x} = 290$

7.16 Un professeur de mathématiques recueille toutes les notes qu'il a mises dans une classe donnée et obtient le tableau de distribution suivant :

Note	1	1.5	2	2.5	3	3.5	4	4.5	5	5.5	6
Effectif	0	1	4	9	9	21	28	33	34	30	21

44 72 105
 Nombre d'élèves : 190 , Médiane : 95^e nde

→ (45)

a) Décrire la variable statistique étudiée et donner son type.

(b) Tracer l'histogramme de cette distribution.)

c) Calculer le pourcentage associé à chacune des valeurs de ce tableau de distribution.

d) Donner le mode, la médiane et la moyenne.

7.17 Lors d'une journée de recrutement de l'armée, on a mesuré la taille en centimètres de 50 hommes âgés de 20 ans et reporté les mesures ci-dessous :

171.5 171.5 172.0 177.0 171.0 169.5 176.0 174.5 170.5 175.0 173.5
 172.5 172.0 173.0 175.5 176.5 173.0 173.5 171.0 169.5 173.5 171.0
 174.0 166.0 173.5 168.0 177.0 170.0 175.0 167.5 176.5 172.5 177.0
 172.5 179.5 168.0 175.0 174.0 178.5 167.0 170.5 176.0 172.0 177.0
 174.0 171.0 179.0 176.0 170.0 170.0

a) Décrire la variable statistique étudiée et donner son type.

b) Grouper les données en 7 classes de deux centimètres de large, comprises entre 166 cm et 180 cm, calculer l'effectif, la fréquence et la fréquence cumulée pour chaque classe.

c) Tracer l'histogramme de cette distribution et tracer le polygone des fréquences cumulées croissantes.

d) Donner la classe modale, la médiane et la moyenne.

7.18 La taille moyenne de 41 250 000 adultes d'un pays est de 1,67 m. Si l'on sait de plus que, dans ce pays, la taille moyenne des femmes est de 1,61 m et celle des hommes de 1,74 m, de combien le nombre de femmes dépasse-t-il le nombre d'hommes ?

7.19 Le prof de maths m'a dit : « Finalement, vous avez 4.5 de moyenne sur les cinq notes de l'année ». Sachant que mes quatre premières notes étaient 5.2, 3.1, 4.4 et 4.2, calculer la cinquième note.

7.20 En utilisant le tableau de distribution de l'exercice 7.12, page 75,

- Calculer l'âge moyen de la population suisse en 1860 et en 2009 et représenter chaque moyenne par un triangle sous l'axe des âges des polygones de fréquences construits au point b) de l'exercice 7.12.
- Calculer l'âge médian de la population suisse en 1860 et en 2009 et marquer chaque médiane par une barre verticale sur le graphique précédent.
- Déterminer la classe modale de l'âge de la population suisse en 1860 et en 2009. Cette notion est-elle représentative dans le cas étudié? Justifier la réponse
- Pourquoi l'âge moyen et l'âge médian de l'année 1860 sont-ils différents? Pourquoi l'âge moyen et l'âge médian de l'année 2009 sont-ils presque égaux?
- Que peut-on conclure en comparant les âges moyens et médians des années 1860 et 2009?

7.21 En utilisant le tableau de distribution de l'exercice 7.14, page 77,

- Calculer le diamètre moyen des boulons et représenter la moyenne par un triangle sous l'axe horizontal de l'histogramme construit au point b) de l'exercice 7.14.
- Estimer la valeur de la médiane à l'aide du polygone des fréquences cumulées construit au point b) de l'exercice 7.14. Calculer de diamètre médian et vérifier sa proximité avec la valeur estimée. Marquer cette valeur par une barre verticale sur l'histogramme.
- Déterminer la classe modale. Cette notion est-elle représentative ici? Justifier la réponse.
- Que peut-on conclure en comparant la moyenne, la médiane et la classe modale sur la forme de la distribution des diamètres des boulons?

7.22 Déterminer la variance et l'écart-type de chaque jeu de données.

a)

valeur	1	2	3	4	5	6
effectif	1	3	5	5	7	4

b) -0.5 0.1 0.9 0.3 0.2 -0.6 0 -1.0 0.7 -0.1

c)

classe	[100; 200[[200; 300[[300; 400[[400; 500[
fréquence	32%	20%	12%	36%

d)

classe	[0; 2[[2; 3[[3; 4[[4; 5[[5; 10[
effectif	12	48	94	42	4

7.23 On a mesuré la vitesse de 50 véhicules :

Vitesse	Effectif
[65; 70[2
[70; 75[7
[75; 80[15
[80; 85[20
[85; 90[2
[90; 95[3
[95; 100[1

- Décrire la variable statistique étudiée et donner son type.
- Trouver la classe modale.
- Calculer le pourcentage associé à chacune des valeurs de ce tableau de distribution.
- Tracer l'histogramme de cette distribution.
- Calculer les fréquences cumulées.
- Tracer le polygone des fréquences cumulées.
- Calculer la moyenne et l'écart-type de cette distribution.

7.24

- Une série A représente l'âge des cinq membres d'une famille et une série B celui des élèves d'une classe de gymnase. Laquelle des deux séries aura le plus grand écart-type ?
- Un professeur de mathématiques fait passer un travail dans deux classes. Les deux groupes obtiennent la même moyenne, mais l'écart-type de la première classe est plus grand que celui de la seconde. Dans quelle classe peut-on dire que les élèves ont à peu près tous le même niveau sur ce sujet ?
- Dans une région aride du globe, on enregistre les précipitations quotidiennes, en mm, durant 60 jours consécutifs. La moyenne des 60 données est de 0. Quelle est la valeur de l'écart-type ?
- Dans une classe de première année de gymnase, la moyenne d'âge est de 16,16 ans, avec un écart-type de 0,76 an. Si les élèves de cette classes restent les mêmes, que vaudront la moyenne \bar{x} et l'écart-type s en troisième année ?
- Vrai ou faux ? Toutes les données d'une distribution dont la moyenne est 70 et l'écart-type 10 sont comprises entre 60 et 80.

7.25 Un maître rend un test dans une classe de 22 élèves en disant : « La moyenne de la classe est de 4.20 avec un écart-type de 0.83 ». Donner une interprétation de ces informations.

7.26 Au laboratoire de physique, une série de mesures de l'accélération de la pesanteur terrestre a donné les résultats suivants : 9.95 9.85 10.13 9.69 9.47 9.98 9.87 9.46 10.00.

Calculer la moyenne et l'écart-type de ces résultats et interpréter.

7.27 En reprenant les données du tableau de distribution de l'exercice 7.23, page 80,

- Calculer une approximation de la vitesse moyenne et de l'écart-type et interpréter.
- Les données sont-elles homogènes ?

7.28 Deux enseignants, l'un travaillant en Suisse où les tests sont notés de 1 à 6 et l'autre travaillant en France où les tests sont notés de 0 à 20, discutent de leur classe. L'enseignant suisse constate que sa classe a une moyenne de 4.1 avec un écart type de 1.2. L'enseignant français constate que sa classe a une moyenne de 12.5 avec un écart type de 5.3.

- Si x est une note attribuée dans le système français et y une note attribuée dans le système suisse, déterminer la relation entre x et y qui permet de transposer les notes d'un système à l'autre.
- Si on compare les moyennes des ces deux classes, laquelle est la meilleure ?
- Pourquoi le coefficient de variation $\frac{\sigma}{\bar{x}}$ ne permet-il pas de mesurer l'homogénéité des résultats de ces classes ?
- Quelle est la classe la plus homogène ? Justifier la réponse par un calcul adéquat à définir.

7.29 Déterminer la médiane \tilde{x} et les quartiles Q_1 et Q_3 de chaque jeu de données, puis représenter les données sous la forme d'un boxplot.

- | | | | | | | |
|---|---|---|---|---|----|---|
| 1 | 3 | 6 | 6 | 3 | 4 | 1 |
| 2 | 1 | 2 | 5 | 2 | 10 | 0 |
| 6 | 3 | 2 | 3 | 7 | 4 | 2 |

- | | | | | | |
|----------|----|----|----|----|---|
| valeur | -2 | -1 | 0 | 1 | 2 |
| effectif | 18 | 10 | 15 | 12 | 5 |

- | | | | | | |
|-----------|------------|------------|------------|------------|------------|
| classe | [140; 150[| [150; 160[| [160; 165[| [165; 170[| [170; 180[|
| fréquence | 10% | 15% | 40% | 20% | 15% |

7.30 On reprend le tableau de distribution de l'exercice 7.14, page 77,

- a) Calculer la variance et l'écart-type.
- b) Calculer le premier quartile, la médiane et le troisième quartile.
- c) Tracer la boîte à moustaches correspondante.

7.31 On reprend le tableau de distribution de l'exercice 7.17, page 78,

- a) Calculer la variance et l'écart-type.
- b) Calculer le premier quartile, la médiane et le troisième quartile.
- c) Tracer la boîte à moustaches correspondante.

7.32 On reprend le tableau de distribution de l'exercice 7.23, page 80,

- a) Calculer le premier quartile, la médiane et le troisième quartile.
- b) Tracer la boîte à moustaches correspondante.

7.33 Voici les âges de 20 personnes qui se présentent au permis de conduire :

18 19 19 23 36 21 57 23 22 19
18 18 20 21 19 26 32 19 21 20

- a) Donner le type de la variable étudiée.
- b) Calculer la moyenne, la médiane et le mode de cette série statistique. Quelle est le paramètre de position le plus approprié ?
- c) Quel est le pourcentage de personnes de 25 ans au plus qui se présentent à l'examen ?
- d) Quel est la cote z du candidat le plus âgé ? Interpréter le résultat.
- e) Quel âge aurait un candidat avec un score $z = -1$? Est-ce possible ?

7.34 Trois élèves se disputent le prix du meilleur financement de la semaine spéciale dans une école :

- Edgar a vendu 85 tablettes de chocolat, alors que la moyenne de vente est de 52 tablettes par élève avec un écart-type de 13 tablettes.
- Faustine a vendu 25 arrangements de fleurs, avec une moyenne de 12 arrangements et un écart-type de 6.
- Georges a vendu 75 abonnements au journal de l'école, avec une moyenne de 47 abonnements et un écart-type de 10.

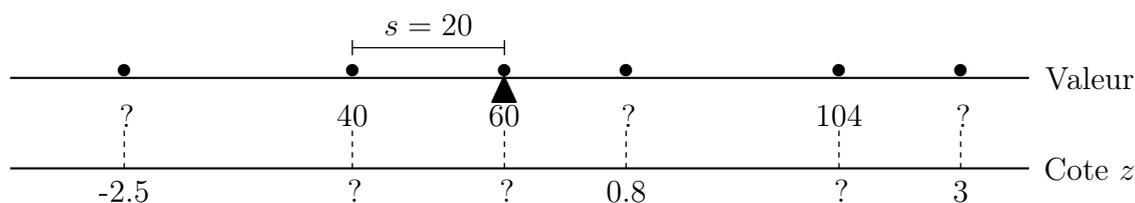
Qui mérite le prix ?

7.35 Voici la durée d'hospitalisation en jours de 40 bébés nés à terme :

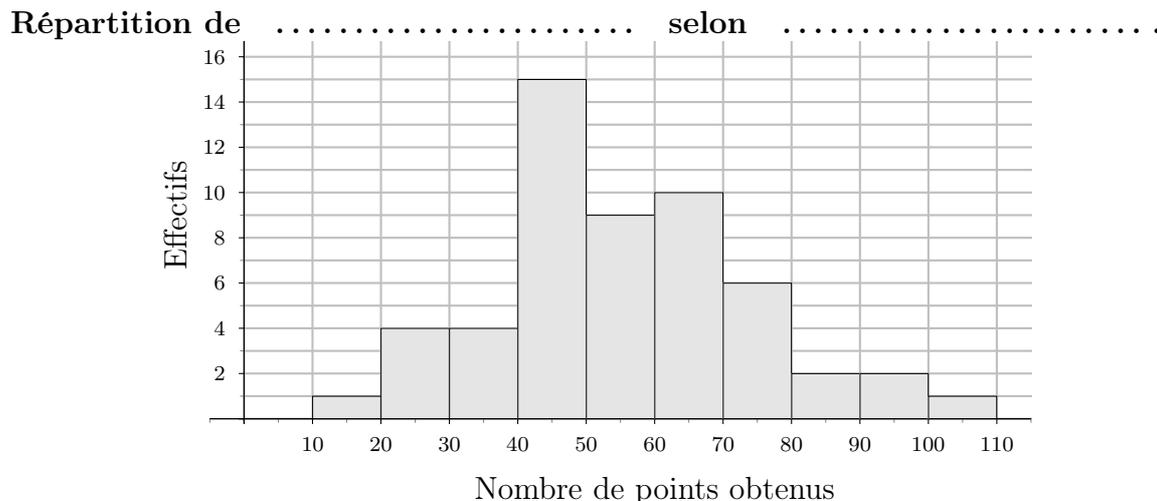
2	1	7	1	33	2	2	3	4	3
4	3	3	10	9	2	5	4	3	3
20	6	2	4	5	2	1	3	3	4
4	2	3	4	3	2	3	4	2	3

- a) Représenter cette distribution sous la forme d'une boîte à moustaches.
- b) Quel est la cote z du bébé qui est resté 20 jours à l'hôpital ?
- c) Calculer le pourcentage des bébés dont l'écart à la moyenne n'excède pas un écart-type. Quelle a été la durée de leur hospitalisation ?

7.36 À l'aide de l'information donnée pour chacun des points du pictogramme ci-dessous, déterminer, selon les cas, la valeur ou la cote z de chaque point du graphique.



7.37 Le nombre de points obtenus par les écoles de Suisse au concours de *Mathématiques sans Frontières* est représenté dans l'histogramme suivant :



- Nommer précisément la variable étudiée, donner son type et le type d'échelle de mesure. Compléter le titre du graphique.
- Calculer la moyenne \bar{x} et l'écart-type σ de ces résultats et interpréter ces mesures. Marquer ces résultats sur le graphique de façon appropriée.
- Quelle est la cote z d'une école ayant obtenu 110 points ?
Quelle est le nombre de points obtenus par une école qui présente une cote z égale à -2 ?
- Les données sont-elles homogènes ? Justifier la réponse.

7.38 Le nombre d'heures de fonctionnement de 50 piles à combustible a été mesuré.

15	238	164	222	764	501	2	43	140	104
492	158	85	311	432	130	308	954	489	491
335	60	209	104	286	229	22	347	326	332
20	225	89	125	61	34	3	287	125	318
91	305	192	491	209	168	869	183	541	552

- Regrouper ces données dans un tableau de distribution en formant des classes d'amplitude égale à 100 heures, avec une dernière classe ouverte « ≥ 600 » et représenter le polygone des fréquences correspondant.
- A l'aide du tableau, estimer par calculs la moyenne et l'écart-type. Représenter sur le graphique la moyenne par un triangle et l'écart-type par un intervalle et interpréter.
- En utilisant la moyenne et l'écart-type obtenus sous b), calculer la cote z des deux valeurs extrêmes. Interpréter et critiquer l'interprétation.
- Représenter le polygone des fréquences cumulées.
- Déterminer graphiquement les quartiles et interpréter.
- La compagnie qui fabrique ces piles garantit leur durée de vie. Ainsi, si une pile achetée dure moins de a heures, la compagnie s'engage à la remplacer gratuitement. D'après cet échantillon, quelle doit être la valeur de a si le fabricant ne veut pas remplacer plus de 3% des piles vendues ?

Solutions des exercices

7.1

Situation 1

- a) *Population* : tous les habitants de la ville de Lausanne.
- b) *Echantillon* : les 500 habitants choisis parmi la population totale.
- c) *Variable étudiée* : la chaîne de télévision préférée d'une personne.
- d) *Ensemble des catégories* : les noms des chaînes que peuvent recevoir les habitants de la ville de Lausanne, pour autant qu'on les retienne pour le sondage.
- e) *Type de variable* : qualitative nominale.

Situation 2

- a) *Population* : la situation économique de la Suisse durant les années comprises entre 2000 et 2010.
- b) *Echantillon* : toute la population est étudiée ici, il n'y a pas d'échantillon.
- c) *Variable étudiée* : le taux de chômage.
- d) *Ensemble des catégories* : tous les pourcentages compris entre 0% et 100%.
- e) *Type de variable* : quantitative continue.

Situation 3

- a) *Population* : les ménages de la ville de Genève.
- b) *Echantillon* : les 380 ménages sélectionnés.
- c) *Variable étudiée* : le nombre d'enfants par ménage.
- d) *Ensemble des catégories* : l'ensemble des nombres entiers inférieurs à 20, en tous cas !
- e) *Type de variable* : quantitative discrète.

Situation 4

- a) *Population* : la population suisse.
- b) *Echantillon* : la quasi-totalité de la population suisse.
- c) *Variable étudiée* : la première langue d'une personne.

- d) *Liste des catégories* : « l'allemand », « le français », « l'italien », « le romanche », « autre langue ».
- e) *Type de variable* : qualitative nominale.

7.2

- a) C'est une variable quantitative continue.
- b) C'est une variable qualitative nominale.
- c) C'est une variable quantitative discrète.
- d) C'est une variable quantitative continue.
- e) C'est une variable qualitative nominale.
- f) C'est une variable qualitative ordinale (les valeurs peuvent être classées).

7.3

- a) Variable qualitative nominale ; échelle nominale.
- b) Variable quantitative discrète ; échelle ordinale.
- c) Variable quantitative discrète ; échelle de rapports.
- d) Variable quantitative continue ; échelle ordinale.
- e) Variable quantitative continue ; échelle de rapports.
- f) Variable qualitative ordinale ; échelle ordinale.
- g) Variable quantitative discrète ; échelle d'intervalle.

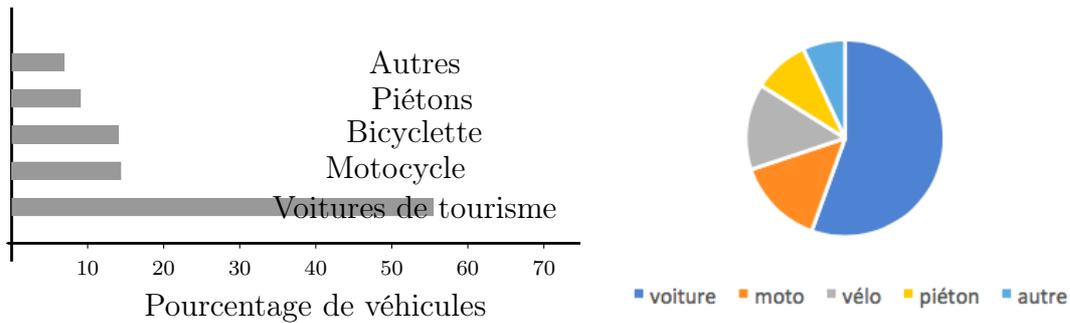
7.4

Heure du lever du soleil : variable quantitative continue ; échelle d'intervalle. La valeur 0h n'indique pas une absence de temps et diviser une heure par une autre n'a pas de sens.

Vitesse des vents : variable quantitative continue ; échelle de rapports. Toutes les opérations mathématiques peuvent être effectuées sur les données.

7.5

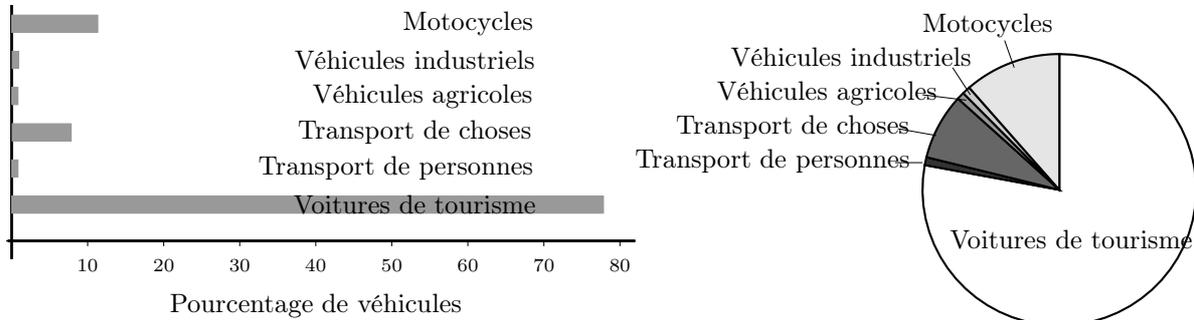
Moyen de locomotion	Blessés légers	Pourcentage	Angle
Voitures de tourisme	9570	55,5%	199,7°
Motocycle	2479	14,4%	51,7°
Bicyclette	2435	14,1%	50,8°
Piétons	1570	9,1%	32,8°
Autres	1196	6,9%	25 °
Total	17250	100%	360°



Non, on ne peut pas en déduire qu'il est moins dangereux de se déplacer en moto qu'en voiture car on ne sait pas le nombre total d'utilisateurs d'une voiture ou d'une moto. Il faudrait comparer les pourcentages relatifs et non absolus.

7.6 Répartition par catégorie des véhicules à moteur mis en circulation en CH en 2011.

Catégorie	Nombre	Pourcentage	Angle
Voitures de tourisme	327'955	77.9%	280.5°
Transport de personnes	3'950	0.9%	3.4°
Transport de choses	33'119	7.9%	28.3°
Véhicules agricoles	3'714	0.9%	3.2°
Véhicules industriels	4'006	1.0%	3.4°
Motocycles	48'133	11.4%	41.2°
Total des véhicules	420'875	100%	360°



Les deux représentations graphiques conviennent car on traite une variable qualitative relevée sur une échelle nominale. On atteint toutefois la limite de visibilité des petites parts sur le diagramme circulaire.

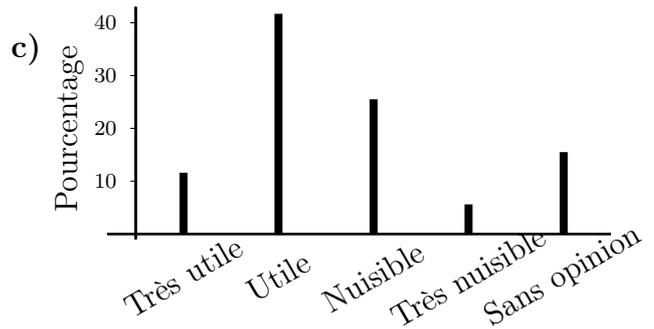
7.7

- a) Population étudiée : les citoyens suisses, variable : opinion sur les accords bilatéraux, Si on excepte la catégorie "sans opinion", échelle ordinale; sinon échelle nominale.

Répartition de 820 répondants selon leur opinion sur les accords bilatéraux.

b)

Utilité	Effectifs	Frq
Très utiles	95	11.6%
Utiles	342	41.7%
Nuisibles	210	25.5%
Très nuisibles	46	5.6%
Sans opinion	127	15.5%
Total	820	99.9%

d) $11.6\% + 41.7\% = 53.3\%$ des sondés sont favorables aux accords bilatéraux.

7.8 $k \approx 6$, $E = 18.45 - 2.65 = 15.8$, amplitude théorique : $\frac{15.8}{6} \approx 2.63$.

Amplitude choisie : 2.5, première classe [2.5 ; 5.0 [et on formera finalement 7 classes.

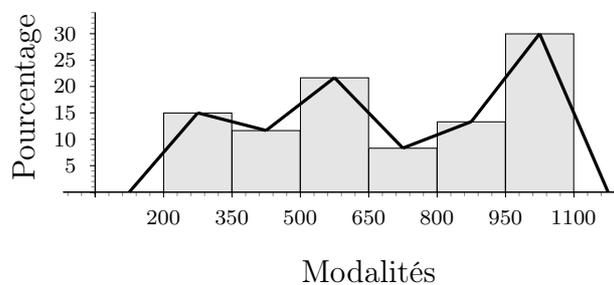
7.9 $k \approx 9$, $E = 937 - 252 = 685$, amplitude théorique : $\frac{685}{9} \approx 76$.

Amplitude choisie : 100, première classe [250 ; 350 [et on formera finalement 7 classes.

7.10

Répartition des données.

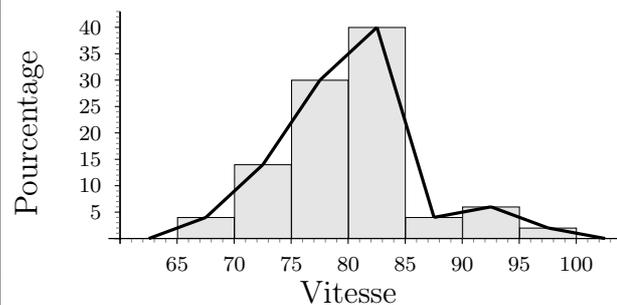
Classes	Effectif	Pourcentage
[200; 350[9	15%
[350; 500[7	11.67%
[500; 650[13	21.67%
[650; 800[5	8.33%
[800; 950[8	13.33%
[950; 1100[18	30%
Total	60	100%



7.11

Répartition de 50 véhicules selon leur vitesse.

Vitesse [km/h]	Effectif	Pourcentage
]65; 70]	2	4%
]70; 75]	7	14 %
]75; 80]	15	30%
]80; 85]	20	40%
]85; 90]	2	4%
]90; 95]	3	6%
]95; 100]	1	2%
Total	50	100%



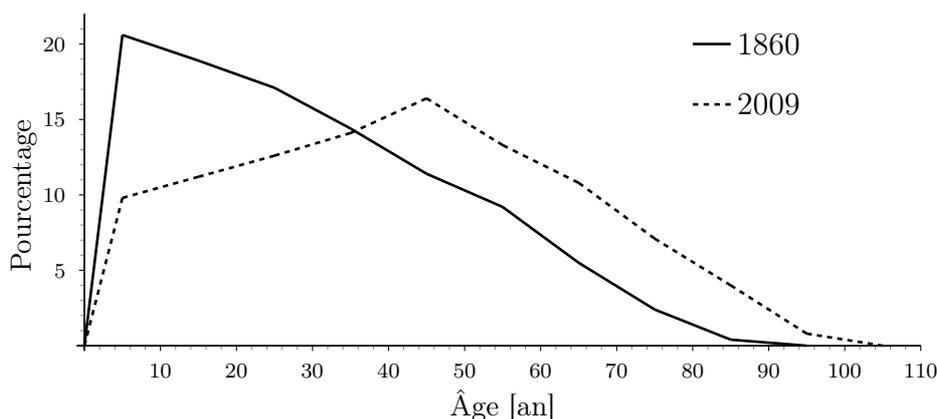
c) Une petite moitié des véhicules respectent la limitation de vitesse de 80 km/h et 40% roulent entre 80 et 85 km/h. En tenant compte d'une marge de tolérance de 5 km/h, 12% des véhicules sont amendables.

d) Parce que les véhicules roulant exactement à 80 km/h respectent la limitation et doivent être groupées avec ceux roulant plus lentement.

7.12

a) Le plus approprié est de représenter sur un même graphique le polygone des fréquences de chacune des deux années. Deux histogrammes superposés produiraient un graphique illisible. On utilise les fréquences relatives car les deux distributions n'ont pas le même effectif total.

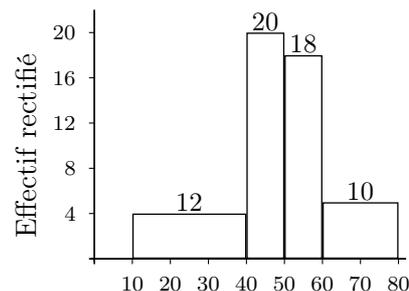
b) **Répartition de la population suisse en 1860 et 2009 selon l'âge.**



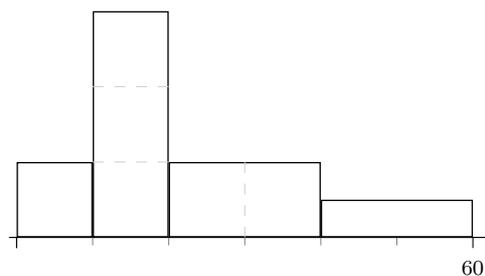
c) « La population suisse a plus que triplé entre 1860 et 2009 en passant de 2,5 millions à presque 7,8 millions d'habitants. En 1860, la population était très jeune : l'aire sous le polygone est plus grande avant 30 ans qu'après. A cette époque, seulement 3% des habitants avaient plus de 70 ans, contre 12% actuellement, soit une proportion quatre fois plus élevée. En 1860, le groupe des moins de 20 ans représentait près de 40% de la population contre 21% aujourd'hui, soit une proportion réduite de moitié. En 1860, la classe la plus représentée est celle des 0 à 10 ans, avec 20.6% des habitants, alors qu'en 2009, c'est la classe des 40 à 50 ans avec 16.4% des habitants. »

7.13

	Amplitude	Classe	Effectif	Effectif rectifié
a)	30	[10; 40[12	4
	10	[40; 50[20	20
	10	[50; 60[18	18
	20	[60; 80[10	5
		Total	60	



	Pourcentage
b)	[0; 10[14.3% (1/7)
	[10; 20[42.9% (3/7)
	[20; 40[28.6% (2/7)
	[40; 60[14.3% (1/7)
Total	100%

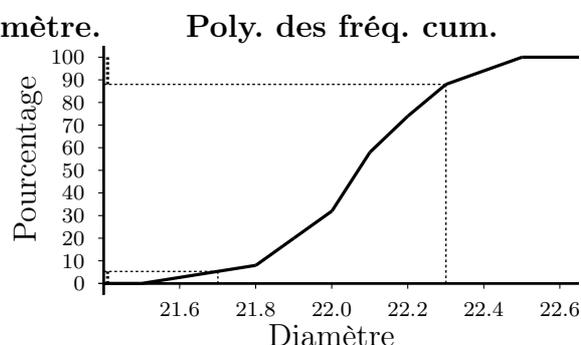
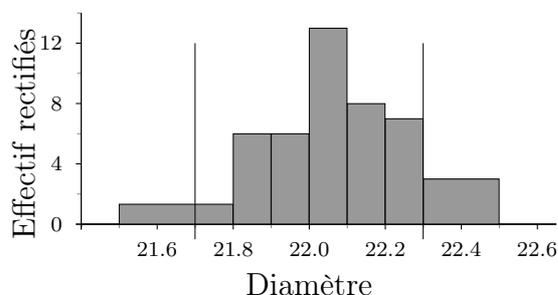


c)
$$\frac{\text{Aire de la portion comprise entre les abscisses 50 et 60}}{\text{Aire totale du polygone}} = \frac{2}{8} = 25\%.$$

7.14

a) Population étudiée : Tous les boulons de la production, variable : diamètre des boulons, variable quantitative continue, échelle de rapport.

Répartition de 50 boulons selon leur diamètre.



d) $\frac{4}{50} \cdot \frac{2}{3} = 5.3\%$ des boulons ont un diamètre < 21.7 mm et $\frac{6}{50} = 12\%$ ont un diamètre > 22.3 mm. Ainsi, 17.3% des boulons ont un diamètre qui s'écarte de plus de 0.3 mm de la valeur nominale.

7.15

a) moyenne : $\bar{x} = 4.04$ médiane : $\tilde{x} = 4$ mode : $M = 5$

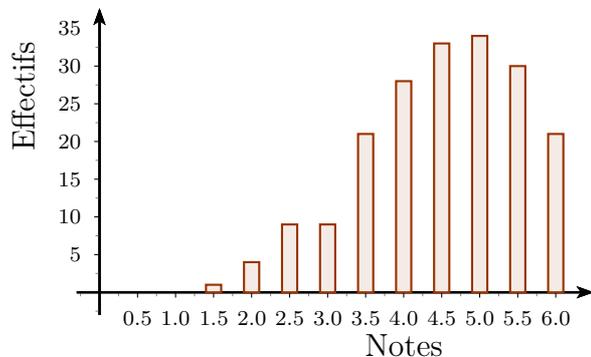
b) moyenne : $\bar{x} = 0$ médiane : $\tilde{x} = 0.05$ le mode n'existe pas

c) moyenne : $\bar{x} = 302$ médiane : $\tilde{x} = 290$ classe modale : [400; 500[

d) moyenne : $\bar{x} = 3.4$ médiane : $\tilde{x} \cong 3.43$ classe modale : [3; 4[

7.16 a) On étudie l'ensemble des travaux écrits passés par les élèves de cette classe. La variable est la note qui figure sur le travail après correction. Il s'agit d'une variable quantitative discrète qui peut prendre 11 valeurs.

b)



c)

Note	Effectif	Fréquence [%]
1.5	1	0.5
2	4	2.1
2.5	9	4.7
3	9	4.7
3.5	21	11.1
4	28	14.7
4.5	33	17.4
5	34	17.9
5.5	30	15.8
6	21	11.1
Total	190	100

d) le mode vaut 5, la médiane $\tilde{x} = 4.5$ et la moyenne $\bar{x} \cong 4.49$

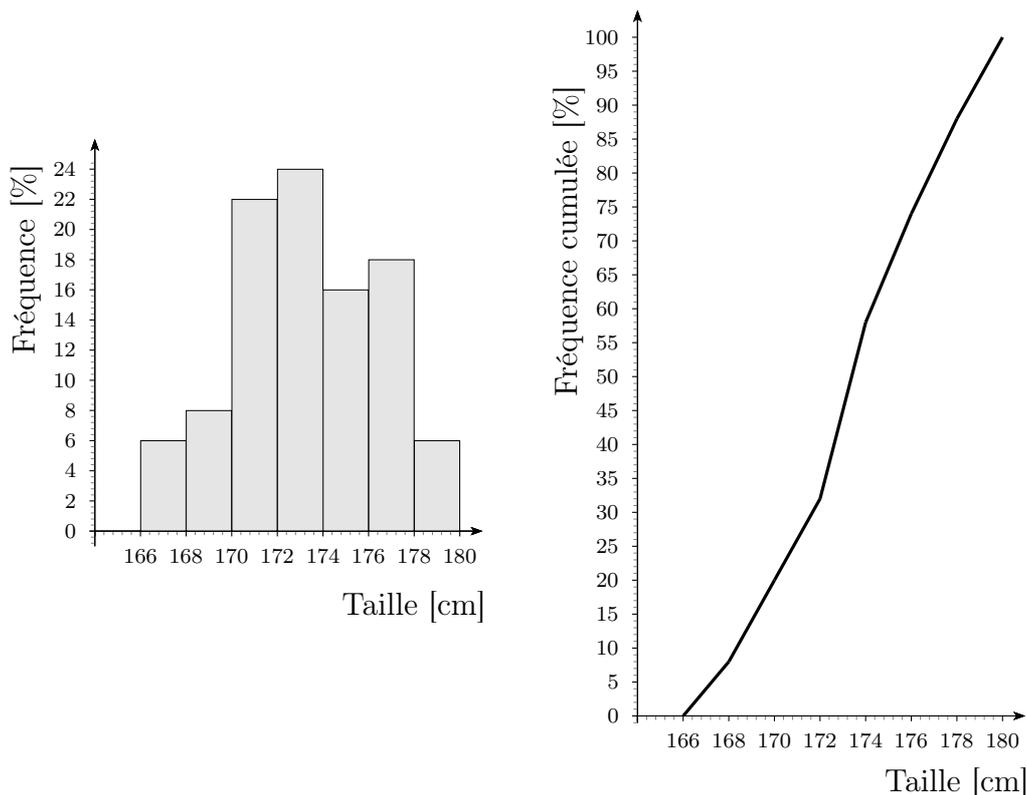
7.17

a) La variable étudiée est la taille en centimètres, variable quantitative continue.

b)

Taille [cm]	Effectif	Fréq.	Fréq. cumulées
[166; 168[3	6%	6%
[168; 170[4	8%	14%
[170; 172[11	22%	36%
[172; 174[12	24%	60%
[174; 176[8	16%	76%
[176; 178[9	18%	94%
[178; 180[3	6%	100 %
Total	50	100%	-

c) Répartition de 50 hommes selon leur taille en cm.



d) La classe modale : $[172; 174[$, la médiane $\tilde{x} \cong 173.17$, la moyenne $\bar{x} \cong 173.28$

7.18 Soit x le nombre de femmes. On peut écrire

$$x \cdot 1.61 + (41\,250\,000 - x) \cdot 1.74 = 41\,250\,000 \cdot 1.67$$

Et donc, $x \simeq 21\,211\,538.46 \simeq 21\,211\,538$, vu que l'on ne considère pas des fractions de personnes.

Il y a donc 3 173 077 femmes de plus que d'hommes.

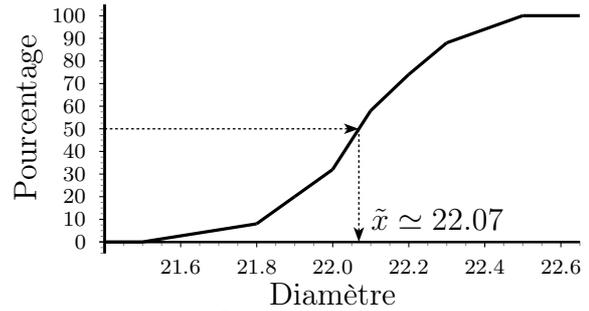
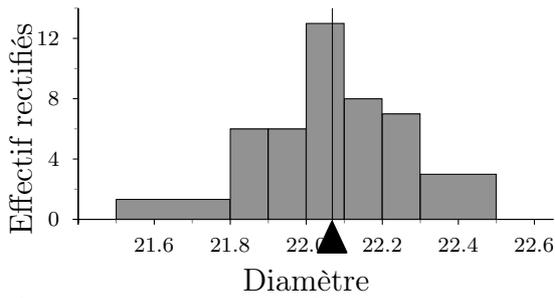
7.19 La cinquième note est 5.6

7.20 a) $\bar{x}_{1860} = 29.1$ ans, $\bar{x}_{2009} = 41.1$ ans **b)** $\tilde{x}_{1860} = 26.1$ ans, $\tilde{x}_{2009} = 41.4$ ans

c) 0 à 10 ans pour 1860 et 40 à 50 ans pour 2009. Ces classes modales sont peu significatives car leurs effectifs ne sont pas beaucoup plus élevés que ceux des autres grandes classes.

d) En 1860, l'âge moyen est plus élevé que l'âge médian, car les quelques personnes très âgées tirent la moyenne vers le haut. En 2009, les âges moyen et médian sont identiques, car la répartition de la population autour de ces mesures est symétrique. **e)** La population est plus vieille en 2009 qu'en 1860.

7.21



a) $\bar{x} = 22.068$ mm b) $\tilde{x} = 22.069$ mm c) La classe modale $[22.0 ; 22.1[$ est significative car son effectif est nettement plus élevé que ceux des autres classes. d) La classe modale contient la moyenne et la médiane qui sont très proches. La distribution est de type normale, en forme de cloche.

7.22

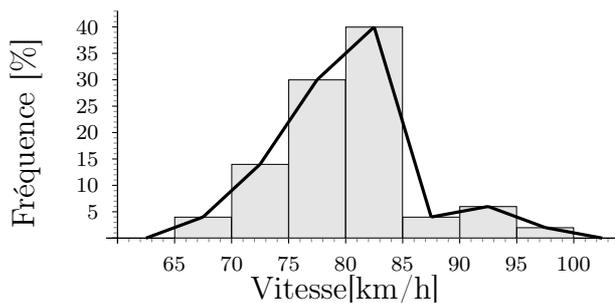
- a) variance : $s^2 = 1.9584$ écart-type : $s \cong 1.40$
- b) variance : $s^2 = 0.306$ écart-type : $s \cong 0.55$
- c) variance : $s^2 = 16096$ écart-type : $s \cong 126.87$
- d) variance : $s^2 = 1.135$ écart-type : $s \cong 1.07$

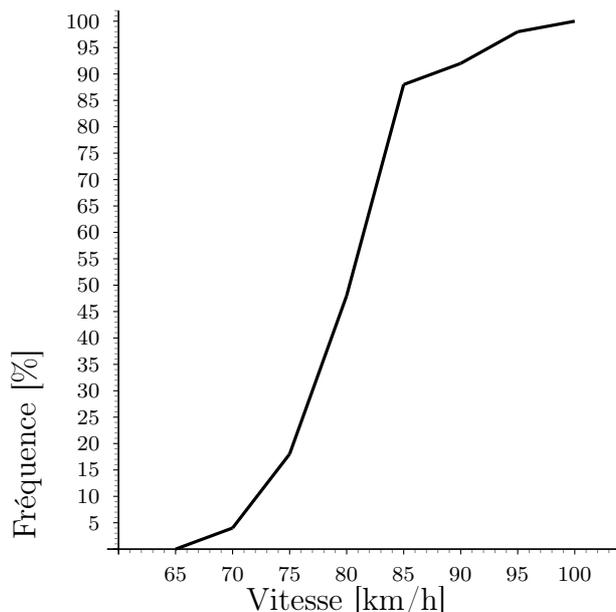
7.23

- a) On étudie la vitesse des véhicules, variable quantitative continue.
- b) Classe modale : $[80 ; 85 [$

	Vitesse [km/h]	Effectif	Frq.	Frq. cum
c)	$[65; 70[$	2	4%	4%
	$[70; 75[$	7	14 %	18%
	$[75; 80[$	15	30%	48%
	$[80; 85[$	20	40%	88%
	$[85; 90[$	2	4%	92%
	$[90; 95[$	3	6%	98%
	$[95; 100[$	1	2%	100%
	Total	50	100%	-

Répartition de 50 véhicules selon leur vitesse.





g) Moyenne $\bar{x} = 80,1 \text{ km/h}$. Ecart-type $s \cong 6,02$.

7.24

- C'est la série A. En effet, il y a plus d'écart à la moyenne dans une famille que dans une classe.
- Plus l'écart-type est faible, plus les résultats sont proches. C'est donc dans la deuxième classe que les élèves ont à peu près le même niveau sur ce sujet.
- L'écart-type et la moyenne sont tous les deux nuls, car aucune des données n'est inférieure à 0. On peut aussi écrire $s = \bar{x} = 0$
- Nouvelle moyenne : $\bar{x}' = 18,16$. Nouvel écart-type : $s' = s = 0,76$.
- C'est faux ; une part importante des données est comprise entre 60 et 80, mais pas toutes.

7.25 Une pluralité de notes sont comprises entre 3.37 et et 5.03, c'est-à-dire entre 3.5 et 5.0 si les notes sont arrondies au demi-point.

7.26 $\bar{x} = 9,822$ et $s = 0,222$.

Une pluralité de mesures donne une accélération comprise entre 9.600 et 10.044 m/s^2 .

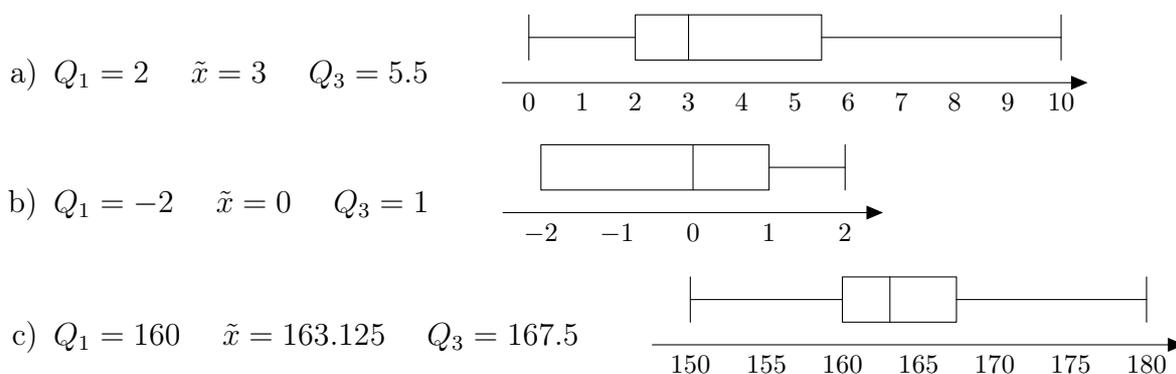
7.27 a) $\bar{x} = 80,1$ $s = 6,0$. Une pluralité de véhicules roulent entre 74.1 km/h et 86.1 km/h .

b) coefficient de variation = $\frac{6,0}{80,1} = 0,075 = 7,5\% < 15\%$. Les données sont homogènes.

7.28 a) $y = \frac{x}{4} + 1$ et $x = 4 \cdot (y - 1)$ b) La classe française. c) Les notes suisses se mesurent sur une échelle d'intervalle et non de rapport. De plus, dans les deux cas, les notes sont bornées supérieurement.

d) $\frac{1.2}{6-1} = 0.24 < \frac{5.3}{20-0} = 0.263$. La classe suisse est plus homogène que la française.

7.29



7.30 a) $s^2 = 1.074$ écart-type : $s \cong 1.04$

b) Q_1 , 48^e valeur $\Rightarrow Q_1 = 4$, médiane $\tilde{x} = \frac{4.5 + 4.5}{2} = 4.5$
et Q_3 , 143^e valeur $\Rightarrow Q_3 = 5,5$,

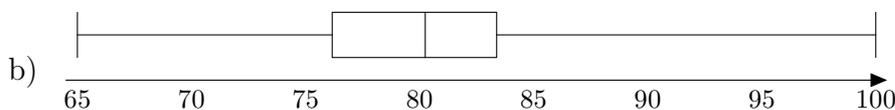


7.31 a) $s^2 \cong 1.074$ écart-type : $s \cong 1.04$

b) $Q_1 = 171$ cm $\tilde{x} \cong 173,17$ $Q_3 \cong 175,88$ cm



7.32 a) $Q_1 \cong 76.17$ $\tilde{x} = 80.25$ $Q_3 \cong 83.38$



7.33

a) Variable quantitative continue (traitée comme discrète).

b) La médiane vaut 20.5 et la moyenne $\bar{x} = 23.55$ et le mode vaut 19. La médiane est la mesure la plus appropriée, car elle n'est pas influencée par les très grandes valeurs contrairement à la moyenne. Le mode ne présente pas une fréquence suffisante par rapport aux autres valeurs.

- c) Il y a 80% des candidats qui ont moins de 25 ans.
 d) On calcule la cote z comme suit :

$$z = \frac{57 - 23.55}{8.93} \simeq 3.75$$

L'âge du candidat est très éloigné de la moyenne. Cela constitue une exception.

- e) Si la cote z vaut -1 , cela implique que $x \simeq 14.62 < 18$. Cette situation ne peut pas se produire, vu que l'âge minimal pour se présenter à l'examen du permis de conduire est de 18 ans !

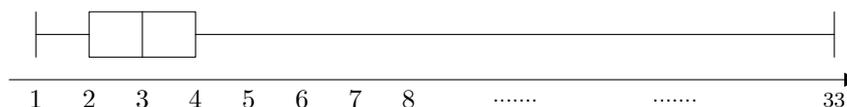
7.34 On a calculé la cote z du nombre d'objets vendus pour chacun des élèves concernés :

$$z_E \simeq 2.54 \quad z_F \simeq 2.17 \quad z_G \simeq 2.8$$

C'est Georges qui obtiendra le prix, vu que la cote z de ses ventes est la plus élevée.

7.35

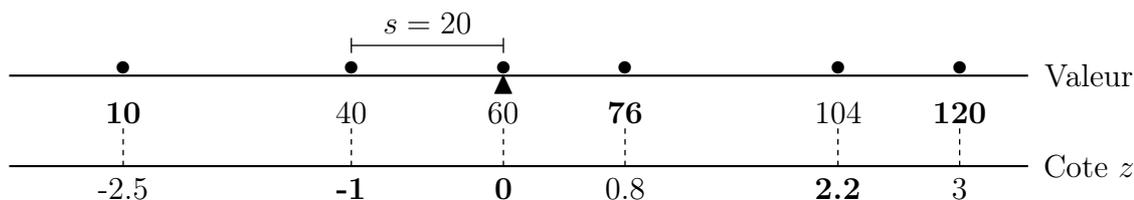
a) $Q_1 = 2 \quad \tilde{x} = 3 \quad Q_3 = 4$



b) $z \simeq (20 - 4.6)/5.55 \simeq 2.77$

- c) Si $z = -1$, alors $x = -0.95$ et donc $x = 1$. Si $z = 1$, alors $x = 10.15$. L'ensemble des valeurs de cette variable comprises entre 1 et 10 représentent le 95% de toutes les valeurs. On peut donc affirmer que 95% des bébés sont restés entre 1 et 10 jours.

7.36



7.37 a) Le nombre de points obtenus est une variable quantitative discrète, mesurée sur une échelle de rapport. Titre du graphique : Répartition de 54 écoles suisses selon le nombre de points obtenus au concours.

b) $\bar{x} = 55.37$; $\sigma = 18.85$. Une pluralité d'écoles ont obtenus entre 37 et 74 points.

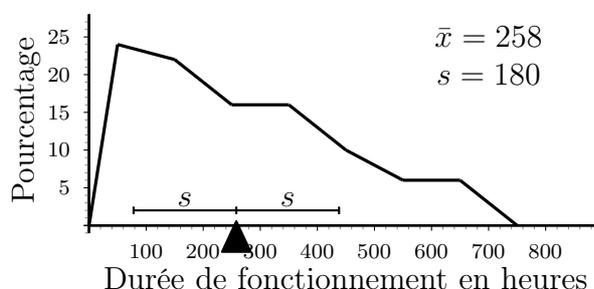
c) $z_{110} = 2.9$ et $x_{-2} \simeq 18$ points. **d)** $CV = 34\% > 15\%$. Les résultats ne sont pas du tout homogènes.

7.38

a) Répartition de 50 piles à combustible selon leur durée de fonctionnement.

Durée [h]	Effectif	Pourcentage
[0; 100[12	24%
[100; 200[11	22%
[200; 300[8	16%
[300; 400[8	16%
[400; 500[5	10%
[500; 600[3	6%
≥ 600	3	6%
Total	50	100%

b) Polygone des fréquences.

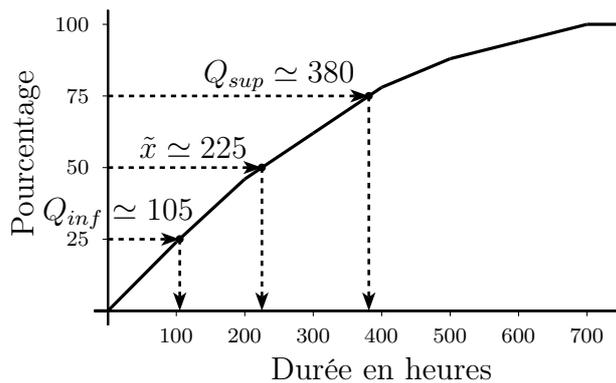


Une pluralité de piles ont une durée de fonctionnement comprise entre 78 et 438 heures.

c) $z_{min} = \frac{2 - 258}{180} = -1.42$ et $z_{max} = \frac{954 - 258}{180} = 3.87$.

D'après les cotes z , une durée de fonctionnement de 954 heures est exceptionnelle alors qu'une durée de fonctionnement de 2 heures ne constitue pas un cas particulièrement rare. Cette dernière interprétation n'est toutefois pas valide pour ces données dont la plus petite cote z possible est $\frac{0 - 258}{180} = -1.4\bar{3}$.

d) Polygone des fréquences cumulées.



e) 50% des piles fonctionnent entre 105 et 380 heures, 25% des piles fonctionnent moins de 105 heures et 25% des piles plus de 380 heures.

A partir des données brutes, on obtient les quartiles suivants :

$$Q_0 = x_{min} = 2, \quad Q_1 = Q_{inf} = 104,$$

$$Q_2 = \bar{x} = 215.5, \quad Q_3 = Q_{sup} = 335,$$

$$Q_4 = x_{max} = 954.$$

Par calcul sur les données regroupées en classe, on obtient les valeurs suivantes :

$$Q_0 = 0, \quad Q_1 = 104.5, \quad Q_2 = 225,$$

$$Q_3 = 381.25, \quad Q_4 = 700$$

f) 3% des piles de l'échantillon ont duré moins de $a = \frac{3\%}{24\%} \cdot 100 = 12.5$ heures.

Ainsi, les piles ayant duré moins de 12.5 heures devraient être remplacées gratuitement.