Statistiques Exercices

1. Exercices de calcul

1. Tracer l'histogramme de la série ci-dessous et compléter :

classe	[0 ; 100[	[100 ; 140[	[140 ; 160[	[160 ; 200[
centre classe				
effectif	5	4	3	2
fréquence				

effectif total		Q_1	
moyenne		médiane	
variance		Q_3	
écart-type		D_1	

2. Tracer l'histogramme de la série ci-dessous et compléter :

classe	[0 ; 20[	[20 ; 30[	[30 ; 40[	[40 ; 60[
centre classe				
effectif	8	6	6	4
fréquence				

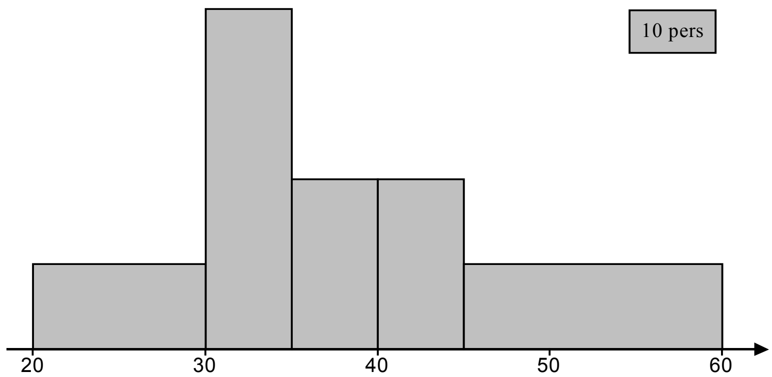
effectif total		Q_1	
moyenne		médiane	
variance		Q_3	
écart-type		D_1	

3. Tracer l'histogramme de la série ci-dessous et compléter :

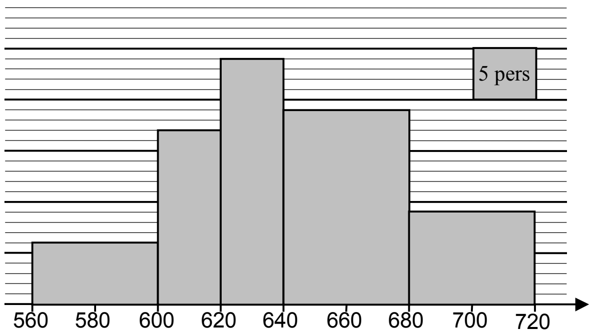
classe	[150 ; 155[	[155 ; 160[	[160 ; 165[	[165 ; 170[	[170 ; 175[	[175 ; 180[
centre classe						
effectif						
fréquence	0,1	0,15	0,25	0,35		0,05

effectif total		Q_1	
moyenne		médiane	
variance		Q_3	
écart-type		D_1	

4. Retrouver le tableau des données à partir de l'histogramme ci-dessous



Retrouver le tableau des données à partir de l'histogramme ci-dessous



2. Exercices de calcul

1. Déterminer la médiane, les 1er et 3ème quartiles, les 1er et 9ème déciles de la série ci-dessous :

Tailles en cm d'un groupe d'enfants de 5 à 7 ans :

104	107	107	107	108	108	109	110	111	111	112	112	112	112	113	113	114	114	114	114	115	115
115	115	115	116	116	117	117	117	118	118	118	119	119	120	120	120	121	121	122	123	123	125

2. L'indice du coût de la construction est donné par le tableau suivant pour la période s'étendant du 2ème trimestre 1997 au 4ème trimestre 1999 :

Année	1997	1997	1997	1998	1998	1998	1998	1999	1999	1999	1999
Trimestre	2	3	4	1	2	3	4	1	2	3	4
Indice	1060	1067	1068	1058	1058	1057	1074	1071	1074	1080	1065

Calculer la médiane, les quartiles, l'écart interquartile et l'étendue de cette série.

3. On a relevé les valeurs de l'indice des prix à la consommation pour l'ensemble des ménages en France (hors tabac) de janvier 1998 à septembre 2000 (base 100 en 1998). On a trouvé :

99,5-99,8-100-100,2-100,2-100,3-100-100-100-99,9-100-99,6-99,9-100,3-100,6-100,6-100,6-100,3-100,5-100,6-100,7-100,7-101,2-101,1-101,2-101,7-101,7-101,9-102,2-102-102,2-102,7.

Déterminer la médiane, les quartiles, le 1er et le 9ème décile de cette série statistique.

4. Nombre de battements du cœur par minutes : Déterminer la médiane, les 1er et 3ème quartiles et les 1er et 9ème déciles de la série ci-dessous puis faire un diagramme en boîtes.

25	66	78	82	86	93	96	99	102	105	110	112	122	129	135	152
28	73	78	83	87	94	97	100	103	106	110	113	122	130	135	154
53	73	78	83	87	95	97	100	103	106	110	114	126	130	136	169
54	74	79	84	89	95	97	101	104	107	111	116	126	130	140	175
55	76	80	85	90	95	97	101	104	107	111	118	126	134	140	176
58	76	81	85	91	96	98	101	104	109	112	119	126	134	140	188
59	78	82	86	92	96	99	102	105	110	112	120	127	135	145	217

5. On a relevé dans 2 tableaux les âges de décès de 90 hommes et 88 femmes

Hommes														
25	65	79	59	54	77	72	33	74	68	66	74	71	64	87
66	77	84	84	79	77	59	60	81	44	65	83	32	66	82
74	78	76	76	54	59	56	78	84	62	48	78	41	60	40
79	70	72	34	37	59	52	68	65	88	54	59	82	76	76
75	61	73	76	52	68	87	62	72	52	71	78	80	80	58
86	75	83	63	63	80	59	84	64	64	78	78	63	75	63

Femmes														
94	82	63	56	90	88	72	76	88	89	94	67	82	82	62
97	88	68	89	91	56	85	77	68	81	88	46	73	97	78
83	75	91	80	95	80	76	63	98	73	66	90	71	73	66
55	53	95	57	69	62	69	80	90	82	73	78	91	88	87
73	80	87	87	84	64	78	77	64	54	83	89	85	86	74
44	84	97	54	76	87	59	80	72	90	90	70	87		

Tracer les diagrammes en boites de ces deux séries. Que peut on en conclure ?

6. On compare les températures maximales moyennes (en °C) de chaque mois de l'année pour deux communes de Haute-Savoie situées à 1000m d'altitude : Chamonix et La Clusaz. (Atlas climatique de la Haute-Savoie, 1991)

Mois	1	2	3	4	5	6	7	8	9	10	11	12
Chamonix	1,5	4	7,5	12	15,5	20	23	22	19	14	6,5	2
La Clusaz	2,5	3,5	6	9,5	14	17	20,5	20,0	17	13	7	3,5

1. Tracer les diagrammes en boîte de ces deux séries (sans les déciles).

2. Interpréter les différences constatés.

7. Deux tireurs s'entraînent au tir à la cible. Ils ont noté leurs résultat en points obtenues au bout de 30 tirs :

Points	50	30	20	10	0
Tireur A	8	9	8	4	1
Tireur B	6	16	3	3	2

1. Calculer la moyenne et l'écart-type pour les séries des résultats des deux tireurs.

2. Comparer ces résultats. Lequel est le plus régulier ?

8. On a mesuré la fluorescence de la chlorophylle α (en millivolts) dans un océan. Calculer la moyenne et l'écart-type de cette série :

Fluorescence	[15;20[	[20;25[	[25;30[	[30;35[	[35;40[	[40;45[
Effectif	8	13	21	20	14	4

9. Un élève a obtenu les notes suivantes à une série de devoirs de français : 8, 12, 10 et 14.

1. Calculer la moyenne et l'écart-type de cette série de note.

2. Son professeur décide d'augmenter toutes les notes d'un point. Que deviennent la moyenne et l'écart-type de la nouvelle série de notes ?

3. Même question s'il décide plutôt d'augmenter toutes les notes de 10%

10. Le tableau ci-dessous donne le coût moyen et le coût médian (en millions de francs) des films de long métrage d'initiative française de 1986 à 1998 :

	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998
Coût moyen	12,5	14,4	18,3	21,0	21,3	23,7	25,9	22,5	26,1	28,1	24,3	31,3	28,6
Coût médian	10,4	12,7	13,5	15,1	15,7	18,5	19,0	17,5	18,0	20,7	17,3	18,6	17,5

1. Comment expliquer le fait que le coût moyen soit toujours supérieur au coût médian ?

2. Interpréter l'accroissement de la différence entre ces 2 coûts à partir de l'année 1994.

11. Étude des températures moyennes au Canada : Le tableau ci-dessous donne les écarts des températures moyennes annuelles par rapport à une température de référence choisie.

Année	1951	1952	1953	1954	1955	1956	1957	1958	1959	1960	1961	1962	1963	1964
Ecart	-0.2	0.9	1	0.2	0.1	-0.5.	0	0.8	-0.3	0.6	-0.1	0.1	0.3	-0.5
Année	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978
Ecart	-0.6	-0.2	-0.3	0.2	0.4	-0.2	0	-1.8	0.7	0.8	-0.2	0.1	0.9	-0.3
Année	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992
Ecart	-0.1	0.4	1.9	-0.9	0.2	0.3	0	0.2	1.4	0.8	-0.2	0	0.5	-0.1
Année	1993	1994	1995	1996	1997	1998	1999	2000						
Ecart	0.5	0.4	0.7	0.1	0.4	2.4	1.8	1						

1. a. Déterminer la médiane, la moyenne et les quartiles de la série. Comparer médiane et moyenne. Commenter.

b. Représenter la série par un diagramme en boite. Est-il symétrique ? Proposer une explication.

c. Déterminer les déciles D_1 et D_9. Représenter la série par un diagramme en boite limité par ces déciles. Est-il symétrique ? Commenter.

2. a. Déterminer la médiane, les quartiles et les déciles D_1 et D_9 de la série la série des 20 dernières années.

b. Représenter sur le même graphique les diagrammes en boites des deux séries (choisir d'utiliser ou non les déciles).

3. Exercice

Une communauté souhaite limiter la longueur des conversations téléphoniques. Elle décide d'envoyer un signal aux 15% des appels les plus longs. On cherche la durée après laquelle un signal doit être envoyé.

Une étude montre que la durée d'appel suit approximativement une loi normale de moyenne $\mu = 8$ min 30 s et d'écart-type $\sigma = 2$ min 15 s.

1. Déterminer les intervalles de centre μ qui contiennent 99%, 95% et 68% des appels.

2. Après quelle durée de conversation faut-il envoyer le signal ?

4. Exercice

Lors de la fabrication d'un lot de fromages de chèvres, on a relevé la masse des fromages fabriqués :

x_i : masse (en g)	[80;85[	[85;90[	[90;95[	[95;100[	[100;105[	[105;110[	[110;115[
n_i : effectifs	5	9	14	18	25	16	7

Dans une production de ce type, tous les fromages ne sont pas commercialisés.

Les fromages dont la masse est comprise dans l'intervalle $\left[\bar{x}-\sigma\,;\bar{x}+\sigma\right]$ sont commercialisés au prix courant.

Les fromages dont la masse est comprise au-delà de l'intervalle $\left[\bar{x}-2\sigma\,;\bar{x}+2\sigma\right]$ ne sont pas commercialisés.

Les autres fromages sont commercialisés au rabais.

1. Peut-on déterminer le nombre de fromages qui seront commercialisés au prix courant ? Peut-on donner un encadrement de ce nombre ?

2. Donner un encadrement du nombre de fromages non commercialisés ?

3. Donner un encadrement du nombre de fromages commercialisés au rabais.

5. Exercice

Analyses sanguines (on admettra que les variables biologiques sont gaussiennes et que les plages de normalité indiquées le sont avec niveau de confiance de 95%)

1. Les triglycérides : Les valeurs normales sont entre 0,5 et 1,5 g/L. Calculer la moyenne μ et l'écart type σ.

2. Le cholestérol : Les valeurs normales sont entre 1,5 et 2,4 g/L. Calculer la moyenne μ et l'écart type σ. Quel intervalle de centre μ contient 99% des cas ?

3. La glycémie : La valeur moyenne est de 0,97 g/L, l'écart type est de 0,09 g/L.

a. Quelle est la plage de normalité ?

b. Pour quel pourcentage de la population la glycémie est-elle supérieure à 1,15 g/L.

6. Exercice

1. Une association de consommateurs contrôle le poids d'un lot de barres chocolatées. Par contrat, le fabricant s'engage à un poids moyen de 50 g, avec 97,5% des poids supérieurs à 48g : quel doit être l'écart type ?

2. Une barre choisie au hasard pèse 49g. Peut-on accuser le fabricant de tricherie ?

3. Un échantillon de 100 barres à un poids de 4,970 kg. On sait que si le fabricant dit vrai, les poids des échantillons ont une moyenne de 5 kg et un écart type de 10 g. Que peut-on dire ?

7. Exercice

L'entreprise « Bien Fondu » fabrique des boîtes de fromage fondu. La masse nette de fromage inscrite sur les boîtes est de 325 grammes. Afin de vérifier que la production est conforme à la déclaration figurant sur les boîtes, le service qualité prélève un échantillon de 20 boîtes produites par la machine. Les valeurs en grammes, relevées sont les suivantes :

Poids	316	321	322	323	324	325	326	327	329
Effectif	1	1	3	3	4	2	2	1	3

1. a. Calculer la moyenne m et l'écart-type σ de cette série statistique.

b. Donner une interprétation concrète de m et σ.

2. La production issue d'une machine est considérée comme conforme si au moins 95% des boîtes de l'échantillon ont une masse appartenant à l'intervalle $\left[m-2\sigma \; ; m+2\sigma \right]$.

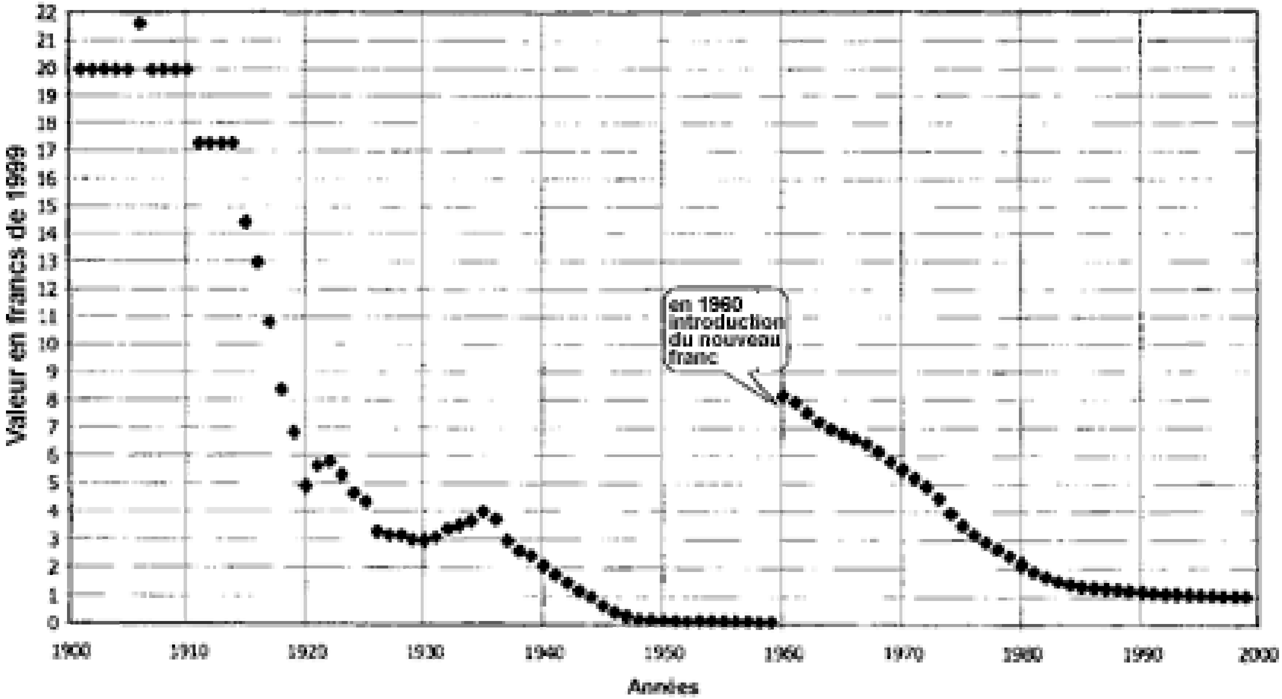
b. La production de la machine de l'entreprise « Bien Fondu » est-elle conforme ? Justifier.

3. a. Pour cet échantillon, préciser la médiane, le premier quartile et le troisième quartile.

b. Représenter le diagramme en boîte associé à cet échantillon, sur lequel figureront au moins la médiane et les premier et troisième quartiles. (Unité graphique :1 centimètre par gramme.)

8. Exercice

A partir des données publiées par l'INSEE, on a représenté graphiquement l'évolution du pouvoir d'achat du franc de 1901 à 1999, c'est-à-dire sa valeur exprimée en francs de 1999 pour chacune de ces années. Le graphique obtenu figure ci-dessous.



Chacun des points de ce graphique a pour abscisse une année n et pour ordonnée la valeur du franc de l'année n, exprimée en francs de l'année 1999 (ou « francs de 1999 »).

Par exemple : un franc de 1901 valait environ 20 francs de 1999 ; un franc de 1920 valait environ 5 francs de 1999.

Ainsi, une somme de 10 francs de 1901 équivaut environ à une somme de 200 francs de 1999.

1. a. Lire graphiquement la valeur (exprimée en francs de 1999) du franc de 1930, puis du franc de 1940.

b. En utilisant le graphique, expliquer pourquoi une somme de 1 000 francs de 1975 équivaut environ à 3 500 francs de 1999.

2. On veut comparer le prix du pain en 1930, 1940 et 1950.

Selon l'INSEE, la valeur du franc de 1950 est environ 0,144 francs de 1999. Compléter le tableau suivant directement sur le sujet.

	en 1930	en 1940	en 1950
1 kilo de pain coûtait	2,15 francs	3,10 francs	35,10 francs
Valeur en francs de 1999			

3. Marie a acheté un appartement en 1970 pour une somme de 180 000 francs.

a. À quelle somme exprimée en francs de 1999, puis en francs de 1980 correspond son investissement ?

b. En 1980, elle a revendu son appartement 520 000 francs. A-t-elle réalisé un gain ? Expliquer.

9. Exercice

Une machine automatique remplit des paquets de bonbons dont la masse théorique doit être de 250 g.

Le fabricant souhaite vérifier la fiabilité de sa machine. Pour cela, il prend 100 paquets de bonbons au hasard, à la sortie de la machine et les pèse. Il obtient les résultats ci-dessous :

Masse (en g)	Nombre de paquets
[215 ; 225[	7
[225 ; 235[	11
[235 ; 245[	19
[245 ; 255[	26
[255 ; 265[	18

[265 ; 275[	13
[275 ; 285[	6

1. a. Tracer l'histogramme de cette série de données sur la feuille jointe.

b. Peut-on dire que l'on a affaire à des données gaussiennes ? Pourquoi ?

2. En détaillant les calculs,

a Calculer la moyenne $\bar{x}$ de cette série.

b. Calculer l'écart type σ de cette série. En donner une valeur arrondie à un chiffre après la virgule.

3. a. Déterminer le nombre de paquets dont le poids est compris dans l'intervalle $\left[\bar{x}-2\sigma\ ;\bar{x}+2\sigma\right]$.

b. En déduire le pourcentage des paquets dont le poids est compris dans l'intervalle $\left[\bar{x}-2\sigma\ ;\bar{x}+2\sigma\right]$.

c. Pourquoi est-on maintenant sûr que ces données sont gaussiennes ?

10. Exercice

Le tableau suivant donne le nombre d'entrées en milliers dans 2 salles de cinéma entre 1983 et 1990 :

Année	Cinéma A	Cinéma B
1983	4137	4404
1984	4478	4368
1985	4433	4280
1986	3632	3679
1987	3826	4315
1988	3941	5175
1989	4287	4527
1990	4762	4519

1. Déterminer la médiane, et les quartiles de chacune des séries.
2. Dessiner les diagrammes en boîte de chacune de ces deux séries sur deux graphiques différents mais en utilisant la même échelle.
3. Commenter les différences observées.

11. Exercice

Voici les notes de 2 élèves :

Emmanuelle : 3, 17, 19, 5 Franck : 9, 12, 11, 12.

Voici les appréciations de l'enseignant :

A : Elève fantaisiste mais capable. Doit progresser si il/elle fournit un travail régulier.

B : Elève moyen qui s'en sort grâce à un travail régulier.

1. Attribuer à chaque élève son appréciation.

2. a. Calculer pour chaque élève sa moyenne.

b. Sans avoir les notes, ce renseignement permet-il d'attribuer à chaque élève son appréciation ? Pourquoi ?

3. a. Pour chaque élève, calculer l'écart type des notes (à donner avec un chiffre après la virgule).

b. Que signifie l'écart type ?

c. Peut-on grâce à ce dernier renseignement, sans avoir les notes, attribuer à chaque élève son appréciation ? Pourquoi ?

12. Exercice

Le tableau suivant donne le nombre d'entrées en milliers dans 2 salles de cinéma entre 1983 et 1990 :

Année	Cinéma A	Cinéma B
1983	4137	4404
1984	4478	4368
1985	4433	4280

1986	3632	3679
1987	3826	4315
1988	3941	5175
1989	4287	4527
1990	4762	4519

1. Déterminer la médiane et les quartiles de chacune des séries. (Expliquer les calculs.)

2. Dessiner les diagrammes en boîte de chacune de ces deux séries sur deux graphiques différents mais en utilisant la même échelle.

3. Commenter les différences observées.

13. Exercice

Une troupe donne un spectacle dans différentes villes. Le gestionnaire financier de la troupe souhaite voir comment a marché le spectacle en particulier dans 2 villes A et B.

La pièce a été jouée 8 fois dans chacune de ces villes et le tableau ci-dessous donne le nombre d'entrées pour chaque séance.

Séance	ville A	ville B
1	451	413
2	452	447
3	517	443
4	431	363
5	367	382
6	428	394
7	436	428
8	440	476

1. Déterminer la médiane et les quartiles de chacune de ces séries. (Détailler les calculs)

2. Dessiner les diagrammes en boîte de chacune de ces deux séries sur deux graphiques différents mais en utilisant la même échelle.

3. Commenter les différences observées.

14. Exercice

Une machine automatique fabrique des CD audio, dont le diamètre théorique doit être de 12 cm soit 120 mm.

Le fabricant souhaite vérifier la fiabilité de sa machine. Pour cela, il prend un échantillon de 100 CD au hasard à la sortie de la machine et les mesure. Il obtient les résultats ci-dessous :

Taille (en mm)	Nombre de CD
[117 ; 118[	3
[118 ; 119[	27
[119 ; 120[	42
[120 ; 121[	23
[121 ; 122[	5

1. a. Tracer l'histogramme de cette série de données.

b. Peut-on dire que l'on a affaire à des données gaussiennes ? Pourquoi ?

2. En détaillant les calculs,

a. Calculer la moyenne $\bar{x}$ de cette série.

b. Calculer l'écart type σ de cette série. En donner une valeur arrondie à un chiffre après la virgule.

3. a. Déterminer le nombre de CD dont le diamètre est compris dans l'intervalle $\left[\bar{x}-2\sigma\,;\bar{x}+2\sigma\right]$.

b. En déduire le pourcentage de CD dont le diamètre est compris dans l'intervalle $\left[\bar{x}-2\sigma\,;\bar{x}+2\sigma\right]$.

c. Pourquoi est-on maintenant sûr que ces données sont gaussiennes ?

15. Pondichéry, avril 2001

Voici un tableau de données créé sous tableur donnant la mesure des masses à la naissance des filles et des garçons dans une maternité durant l'année 2000.

	A	B	C	D	E
1	masse M (en g)	garçons	filles	totaux	Pourcentages
2	$M < 1500$	3	4		0,3
3	$1500 \le M < 2000$	9	9		
4	$2000 \le M < 2500$	34	46		3,3
5	$2500 \le M < 3000$	156	256		16,8
6	$3000 \le M < 3500$	504	536		42,4
7	$3500 \le M < 4000$	402	288		28,2
8	$M \ge 4000$	142	61		
9	Total	1250			100
10	Moyenne	3424,8			
11	Ecart Type	509,7		506,1	

On prendra 1200 g comme masse moyenne des enfants de masse inférieure à 1500 g et 4300 g comme masse moyenne des enfants de masse supérieure à 4000 g.

1. a. En utilisant la calculatrice, donner la valeur affichée par celle-ci de l'effectif total des filles ainsi que les arrondis à 0,1 près de la moyenne et de l'écart type des masses des filles à la naissance.

b. Reporter ces valeurs dans le tableau.

Pour les garçons l'effectif total est de 1250. D'autre part, la masse moyenne, à 0,1 près, des garçons à la naissance est de 3424,8 g et l'écart type vaut 509,7.

On note m la moyenne des masses à la naissance de tous les enfants (filles et garçons réunis),

a. En utilisant. les valeurs des cellules B9, C9, B10 et C10 quelle formule placeriez-vous dans la cellule D10 pour calculer cette moyenne m ?

b. Effectuer alors ce calcul.

3. La colonne E donne la répartition en pourcentage de chaque classe d'enfants (filles et garçons réunis). Ces pourcentages sont donnés à 0,1 près.

Quelle est la formule qui permet de donner le résultat de la cellule E2 ?

Calculer les résultats manquants de la colonne E.

16. Amérique du Sud, novembre 2002, 8 points

L'entreprise « Bon Fondu » fabrique des boîtes de fromage fondu, sur un même site. Elle utilise trois machines différentes A, B, C. La fabrication du fromage fondu et le conditionnement sont automatisés. Le service qualité est chargé du suivi statistique de la production afin de garantir au mieux le respect des règles prévues par la législation en vigueur.

Partie A

La fabrication d'une journée est de 10 000 tonnes avec la répartition précisée dans le tableau suivant :

Tableau N° 1 : les masses sont exprimées en tonnes

Machine	A	B	C	Totaux
Boîtes sans défaut	1800	4500	2500	M
Boîtes avec défauts de fabrication	180	400	200	780

Boîtes avec défauts de conditionnement	20	100	300	420
Totaux	X	5000	3000	10000

1. Calculer, en justifiant vos calculs, les valeurs de X et de M figurant dans les marges du tableau N° 1 précédent.

Dans les questions suivantes, les résultats demandés seront arrondis à 10^{-1} près.

2. a. Compléter le tableau N° 2, figurant sur la feuille annexe 2, donnant les pourcentages de chaque production par rapport à la production totale.

b. Compléter le tableau N° 3, figurant sur la feuille annexe 2, donnant, par colonnes, les pourcentages par rapport à la production de la colonne pour chacune des machines A, B et C.

c. Compléter le tableau N° 4 donnant, sur chaque ligne, les pourcentages produits par chaque machine par rapport à la production de la ligne (production sans défaut, avec défauts de fabrication ou, enfin, avec défauts de conditionnement).

3. a. Pour la machine A, quel est le pourcentage des boîtes présentant un défaut de fabrication ?

b. Pour la machine B, quel est le pourcentage des boîtes sans défaut ?

c. Parmi les boîtes sans défaut, quel est le pourcentage des boîtes fabriquées par la machine B ?

Partie B

La masse nette de fromage inscrite sur les boîtes est de 320 grammes. Afin de vérifier que la production est conforme à la déclaration figurant sur les boîtes, le service qualité prélève un échantillon de 20 boîtes produites par la machine B. Les valeurs en grammes, ordonnées, sont les suivantes :

315,5	315,5	316	321	322	323	323,5	323,5	324	324
324	325	325,5	326	326	327	328,5	329	329	329

La moyenne m de cette série statistique est 323,85 et son écart-type σ est 4,22.

1. La production issue d'une machine est considérée comme conforme si au moins 95 % des boîtes de l'échantillon ont une masse appartenant à l'intervalle $[m-2\sigma, m+2\sigma]$, où m est la moyenne de l'échantillon et σ son écart-type. La production de la machine B est-elle conforme ? Justifier.

2. a. Pour cet échantillon, préciser la médiane, le premier quartile et le troisième quartile.

b. Représenter le diagramme en boîte associé à cet échantillon, sur lequel figureront au moins la médiane et les premier et troisième quartiles. Unité graphique : 1 centimètre par gramme.

Feuille annexe 2 : À rendre avec la copie

Tableau N° 2 des pourcentages par rapport à l'effectif total

Machine	A	B	C	Totaux
Boîtes sans défaut				
Boîtes avec défauts de fabrication				
Boîtes avec défauts de conditionnement				
Totaux				100 %

Tableau N° 3 des pourcentages par colonne

Machine	A	B	C
Boîtes sans défaut			83,3
Boîtes avec défauts de fabrication			6,7
Boîtes avec défauts de conditionnement			10
Totaux	100 %	100 %	100 %

Tableau N° 4 des pourcentages par lignes

Machine	A	B	C	Totaux
Boîtes sans défaut	20,5	51,1	28,4	100 %
Boîtes avec défauts de fabrication	23,1			100 %
Boîtes avec défauts de conditionnement	4,8			100 %

17. Nouvelle Calédonie, novembre 2002, 12 points

On considère les quatre lettres A, T, C, G. Dans cet exercice, on s'intéresse aux mots de trois lettres (mots ayant un sens ou non) que l'on peut former avec ces lettres.

Ainsi, les mots CAT, TTG et GAG conviennent.

1. a. Déterminer tous les mots de trois lettres distinctes que l'on peut constituer en commençant par la lettre T.

b. Combien de mots de trois lettres distinctes peut-on constituer ? Justifier.

2. Montrer que l'on peut former 64 mots de trois lettres.

3. On veut simuler des tirages demots de trois lettres.

a. Expliquer comment, en utilisant la table de chiffres au hasard donnée ci-dessous, on peut effectuer une telle simulation. L'illustrer par une suite d'exemples.

b. Effectuer cette simulation pour vingt tirages de mots. Donner les vingt mots obtenus ; combien d'entre eux sont formés de trois lettres différentes ?

Quelle est alors la fréquence d'apparition des mots de cette nature ?

EXERCICE 2 : Table de chiffres au hasard.

72432	77372	46210	25703	18412
50237	64312	80814	75120	33549
58061	02571	58258	34743	92043
45152	71434	30278	96654	10783
23670	42367	04950	15824	38193
35710	49301	02047	88463	01415
26715	53714	39182	76434	97502
21040	82379	91768	42893	34271

18. Nouvelle Calédonie, novembre 2002 8 points

On étudie grâce à un tableur et à une calculatrice les communications téléphoniques d'une famille durant la période du 16 juin au 15 août 2000.

I. On s'intéresse d'abord à la durée des communications téléphoniques vers les téléphones mobiles pendant la période du 16 juin au 15 août.

Sur la feuille annexe, figure une copie de l'écran d'une calculatrice où est tracé un diagramme en boîte représentant la série relative à la durée de ces communications.

Sur ce diagramme sont entre autres indiqués :

• le minimun (10 secondes),

• le premier quartile (50 secondes),

• le troisième quartile (2 minutes 50 secondes),

• et le maximum (5 minutes 20 secondes).

Le pas de la graduation est de 10 secondes.

1. Quelle information a-t-on sur le pourcentage des communications téléphoniques qui ont duré moins de 50 secondes et sur celui des communications qui ont duré plus de 2min 50 s ?

2. a. Lire graphiquement la médiane et donner le résultat en minutes, secondes.

b. Peut-on dire qu'au moins la moitié des communications ont duré moins de 2 minutes ?

II. On s'intéresse ensuite à la durée des communications téléphoniques locales, toujours pendant la période du 16 juin au 15 août.

On étudie plus particulièrement les communications téléphoniques des quinze derniers jours du mois de juin. Les données figurent dans le cadre 1 de la feuille annexe.

On lit, par exemple, que le 16 juin il y a eu une communication téléphonique d'une durée de 8 minutes et 8 secondes ce qui est noté 0 : 08 : 08.

1. Pour cette période, quel est le jour où il y a eu le plus grand nombre de communications téléphoniques locales ?

2. Pour ce jour-là, calculer la durée moyenne d'une communication téléphonique locale.

III. On considère maintenant l'ensemble des communications téléphoniques locales durant la période du 16 juin au 15 août et on s'intéresse à la série constituée par la durée de ces appels.

Dans le cadre 2 de la feuille annexe figure un tableau regroupant les appels en fonction de leur durée. En utilisant les données de ce cadre :

1. a. Déterminer le pourcentage des appels qui ont duré moins de 3 minutes.

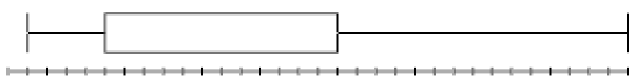
b. Justifier que la médiane de la série est comprise entre 1 minute et 2 minutes.

2. À l'aide d'un tableur on a obtenu les résultats figurant dans le cadre 3 de la feuille annexe.

En utilisant des données pertinentes de ce cadre, construire un diagramme en boîte correspondant à cette série (on prendra comme échelle 1 cm pour 1 minute).

Annexe de l'exercice 1

Diagramme en boîte



Cadre 1			
Dates	Durée des communications	Dates	Durée des communications
16 juin	0:08:08	28 juin	0:01:07
16 juin	0:11:07	29 juin	0:00:20
16 juin	0:01:00	24 juin	0:02:03
16 juin	0:12:22	24 juin	0:01:56
16 juin	0:12:48	24 juin	0:01:35
16 juin	0:07:29	24 juin	0:00:17
16 juin	0:11:36	24 juin	0:00:17
16 juin	0:09:28	24 juin	0:03:32
18 juin	0:02:30	24 juin	0:00:30
18 juin	0:02:54	24 juin	0:00:05
19 juin	0:00:10	24 juin	0:02:57
19 juin	0:05:29	24 juin	0:01:18
19 juin	0:01:05	25 juin	0:05:06
19 juin	0:01:21	25 juin	0:00:04
19 juin	0:00:18	25 juin	0:00:56

Cadre 2	
Durée des communications	Nombre de communications
$0 \leq d < 30$ s	49
30 s $\leq d < 1$ min	31
1min $\leq d < 2$ min	47
2 min $\leq d < 3$ min	21
3min $\leq d < 5$ min	34
5min $\leq d < 10$ min	28
10 min $\leq d < 20$ min	19
nombre total d'appels	229

19 juin	0:13:58	25 juin	0:13:21
20 juin	0:01:08	26 juin	0:01:22
20 juin	0:07:59	26 juin	0:02:54
20 juin	0:04:31	26 juin	0:04:36
20 juin	0:04:53	26 juin	0:00:35
21 juin	0:00:01	26 juin	0:03:00
21 juin	0:01:53	26 juin	0:00:16
21 juin	0:01:28	26 juin	0:01:15
21 juin	0:01:18	26 juin	0:03:47
21 juin	0:01:10	26 juin	0:00:30
21 juin	0:00:34	27 juin	0:07:28
22 juin	0:00:08	27 juin	0:11:29
23 juin	0:01:05	27 juin	0:01:27
28 juin	0:03:39	27 juin	0:01:00
28 juin	0:03:43	27 juin	0:00:56

Cadre 3	
moyenne =	0:03:11
médiane =	0:01:36
premier quartile =	0:00:36
troisième quartile =	0:04:21
minimum =	0:00:01
maximum =	0:14:01
premier décile =	0:00:12
neuvième décile =	0:09:28

19. France, juin 2003, 8 points (c)

Dans tout l'exercice les tailles sont exprimées en centimètre.

1. L'équipe de soins de la maternité « Beaux jours » a relevé la taille des nouveaux-nés. Pendant la troisième semaine du mois de janvier 2003, il y a eu 9 naissances. Les tailles sont données dans le tableau ci-dessous :

48	50,5	51,5	50	52,5	50	49	53	50

a. Calculer la moyenne des tailles de ces 9 nouveaux-nés.

b. Déterminer la médiane des tailles de ces 9 nouveaux-nés.

2. Sur la totalité du mois de janvier 2003, il y a eu 57 naissances à la maternité « Beaux jours ». Les 57 tailles sont données dans le tableau ci-dessous :

Taille en cm	46	47,5	48	48,5	49	49,5	50	50,5	51	51,5	52	52,5	53
Effectif	1	2	3	5	5	7	9	8	7	5	2	2	1

a. Calculer la moyenne des tailles de ces 57 nouveau-nés.

b. Déterminer la médiane des tailles de ces 57 nouveaux-nés en précisant la démarche.

c. Calculer le pourcentage de nouveaux-nés ayant une taille inférieure ou égale à 49 cm. Donner la réponse arrondie à 0,1 %.

d. Parmi toutes ces tailles, déterminer la plus petite taille t telle qu'au moins les trois quarts des nouveaux-nés aient une taille inférieure ou égale à t centimètres. Quel paramètre de la série des tailles a-t-on ainsi trouvé ?

e. Tracer le diagramme en boîte correspondant à ces tailles sur l'axe D_1 de l'annexe 1 (à remettre avec la copie).

3. L'étude statistique de la taille, en centimètre, des 64 nouveaux-nés durant le même mois de janvier 2003 à la maternité « Bon accueil » a donné les résultats suivants :

Minimum	Maximum	Moyenne	Médiane	Premier quartile	Troisième quartile
46	53	49,3	49	48	50,5

a. Tracer le diagramme en boîte correspondant à ces tailles sur l'axe D_2 de l'annexe 1.

b. Parmi les deux maternités « Beaux jours » et « Bon accueil », une seule possède un service pour les naissances prématurées. En utilisant les deux diagrammes en boîte tracés précédemment, peut-on trouver laquelle ? Justifier votre réponse.

c. Les deux maternités « Beaux jours » et « Bon accueil » sont les seules maternités de la même ville. Calculer la moyenne des tailles des nouveaux-nés en janvier 2003 dans les maternités de cette ville.

Les données de l'énoncé permettent-elles de déterminer la médiane des tailles de ces nouveaux-nés ? Si oui, la déterminer; sinon expliquer pourquoi.

ANNEXE 1 : à rendre avec la copie

Exercice 1

clinique « Beaux jours »

clinique « Bon accueil »

20. Liban, juin 2003, 9 points

Le tableau ci-dessous donne :

- la répartition par classes d'âge d'un échantillon de 1000 personnes représentatif de la population française en 2000 ;
- la répartition par classes d'âges d'un échantillon de 1 000 personnes, telle qu'elle est prévue pour l'année 2025.

classe d'âge / année	]0 ; 20]	]20 ; 60]	]60 ; 66]	]66 ; 76]	]76 ; 86]	]86 ; 100]
2000	198	442	162	126	56	16
2025	136	379	212	166	821	25

(Sources : « *Tableaux de l'économie française* », d'après des données de l'INSEE)

Une telle prévision est utile pour planifier les investissements dans les domaines du logement, des maisons de retraite, des écoles, des hôpitaux, des transports... On suppose que la répartition dans chaque classe est uniforme et on remplacera chaque classe par son centre.

1. À l'aide de la calculatrice, donner les résultats arrondis à 10^{-1} près, de la moyenne, de l'écart-type, de la médiane, du premier quartile et du troisième quartile pour la série concernant l'année 2025.

2. On réalise le même type de prévision pour l'année 2050. On souhaite alors comparer les indicateurs des années 2000 et 2050. Pour cela, on dispose du tableau ci-dessous où les résultats sont arrondis à 10^{-1} près.

indicateur / année	moyenne	écart-type	médiane	1^{er} quartile	$3^{ème}$ quartile
2000	44,8	22,4	40	40	63
2050	56,4	22,2	63	40	71

Pour les séries des années 2000 et 2050, réaliser les boîtes à moustaches non élaguées par rapport au même axe, en les construisant l'une en dessous de l'autre et en prenant 1 cm pour 10 ans. On rappelle que les boîtes à moustaches sont aussi appelées diagrammes en boîtes, diagrammes en boîtes et moustaches, diagrammes de Tuckey ou boîtes à pattes.

3. Les phrases suivantes sont-elles vraies ou fausses ? Justifier.

a. En 2050, on prévoit que plus d'une personne sur deux aura au moins 60 ans.

b. En 2000, il y au moins 75% des personnes âgées de 63 ans ou moins.

c. La dispersion par rapport à la moyenne des âges est supérieure en 2000 à celle prévue en 2050.

d. On prévoit qu'au moins trois personnes sur quatre aura 71 ans ou moins en 2050.

e. En 2050, on prévoit que la moitié de la population aura moins que l'âge moyen.

f. On prévoit que le pourcentage de la population dont l'âge est compris entre 40 et 63 ans baissera environ de moitié entre 2000 et 2050.

21. Amérique du Sud, novembre 2003, 10 points

On a demandé aux 28 élèves d'une classe de Première L de prendre leur pouls au repos et de compter le nombre de battements cardiaques pendant une minute. On obtient ainsi une série statistique à partir des résultats obtenus, rassemblés dans un tableau :

Nombre de battements par minute	44	59	62	63	65	67	68	70	72	73
Effectifs	1	1	1	1	1	1	1	2	2	2
Nombre de battements par minute	74	75	76	77	79	80	82	83	90	100
Effectifs	2	1	2	1	1	2	3	1	1	1

1. a. Quels sont les nombres maximal et minimal de battements par minute des élèves de la classe?

b. Déterminer la médiane de cette série. À l'aide d'une phrase, donner une interprétation de ce résultat.

c. Déterminer l'écart interquartile de cette série.

2. Représenter la série statistique par un diagramme en boîte sur lequel figureront les valeurs extrêmes, le premier et le troisième quartile ainsi que la médiane (unité graphique : 1 cmpour 5 battements parminute).

3. À l'aide de la calculatrice, calculer le nombre moyen de battements x (le résultat sera arrondi au dixième).

4. a. On admet que l'écart type σ de cette série vaut environ 10,2. Calculer le pourcentage d'élèves qui se trouvent dans l'intervalle $[x-\sigma\ ; x+\sigma\ ]$.

b. Peut-on dire qu'un quart des élèves ont un nombre de battements en dehors de cet intervalle ?

5. Pour tous les élèves du lycée, la même expérience est menée. On obtient une série de données que l'on suppose gaussiennes.

a. La plage de normalité à 95% est l'intervalle [53 ; 94] . À l'aide d'une phrase utilisant le nombre de battements, interpréter ce renseignement.

b. Calculer alors le nombre moyen de battements par minute, puis l'écart type de cette série.

22. Nouvelle Calédonie, novembre 2003, 8 points

Dans un hôpital universitaire, des chercheurs étudient un aspect de la croissance des jeunes dont l'âge est compris entre 1 an et 20 ans sur un échantillon de 165 personnes.

Étude de la répartition de l'échantillon selon l'âge et le sexe

Cette répartition est donnée par le tableau suivant :

Tranche d'âge	1 à 5 ans	6 à 10 ans	11 à 15 ans	16 à 20 ans	Total
Filles	11	36	32	10	89
Garçons	15	29	24	8	76
Total	26	65	56	18	165
Taille moyenne en cm pour les deuxsexes	93,9	119,3	146,9	161,2	

1. Les pourcentages demandés seront arrondis à 10^{-2}.

a. Quelle est, en pourcentage, la part de l'échantillon représentée par les filles ?

b. Quelle est, en pourcentage la part de l'échantillon représentée par les enfants de 1 à 5 ans ?

c. Parmi les filles, quelle est en pourcentage, la part de celles qui ont de 1 à 5 ans ?

2. Calculer, au centimètre prés, la taille moyenne des jeunes de l'échantillon.

3. a. Le tableau, ci-dessous, donne la série des âges des 76 garçons de l'échantillon.

2	2	3	3	3	3	4	4	4	4	4	4	5	5	5	6	6	6	6
6	6	7	7	7	7	8	8	8	8	8	8	8	9	9	9	9	10	10
10	10	10	10	10	10	11	11	11	11	11	12	12	12	12	13	13	13	13
13	13	13	13	14	14	15	15	15	15	15	16	16	17	17	18	18	18	20

Déterminer la médiane, puis les premier et troisième quartiles de cette série.

b. On admet que le minimum, le maximum, la médiane, les premier et troisième quartiles de la série relative aux âges des filles sont respectivement 2, 20, 11, 8 et 13.

Sur un même graphique représenter les deux séries précédentes par des diagrammes en boites sur lesquels figureront au moins la médiane, les premier et troisième quartiles (unité : 1 cm pour 2 ans).

23. Polynésie, septembre 2003, 8 points

1. On a demandé aux 35 élèves d'une classe de première, la première L_1, le temps consacré à la lecture pendant une semaine. Les résultats sont consignés dans le diagramme en boîte numéro 1 de la feuille annexe à rendre avec la copie.

a. Donner les valeurs du premier quartile Q_1 et du troisième quartile Q_3.

b. Pour cette classe, le temps moyen de lecture est de 4 heures et le temps médian de lecture est de 3 heures.

Compléter le diagramme en boîte numéro 1, en plaçant le temps moyen (le marquer par une croix **x**) et le temps médian (le marquer par une barre verticale dans la boîte).

c. Pourquoi peut-on affirmer qu'au moins 26 élèves de ce groupe lisent 5 heures par semaine ou moins ? Justifier la réponse.

2. On pose à la classe de Première L_2, composée de 25 élèves, la même question. Les résultats individuels sont consignés dans le tableau ci-dessous :

Temps de lecture (heures)				
3	6	3	5	3
3	4	6	4	2
4	5	8	2	5
7	2	7	4	5
5	4	3	6	9

On considère la série statistique formée des 25 temps de lecture.

a. Déterminer pour cette série statistique le minimum, le maximum, la médiane, la moyenne arithmétique. Déterminer le premier quartile Q_1 et le troisième quartile Q_3.

b. Construire le diagramme en boîte numéro 2 correspondant à cette deuxième classe, en complétant la feuille annexe.

3. Quel est le temps moyen de lecture de l'ensemble des 60 élèves formé par les deux classes ?

Annexe à rendre avec la copie

Exercice 1
Diagramme numéro 1 Classe de première L1

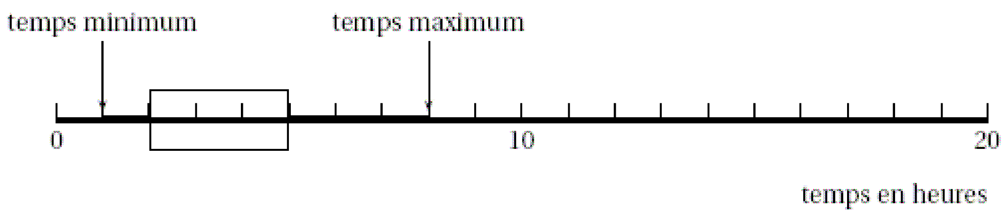
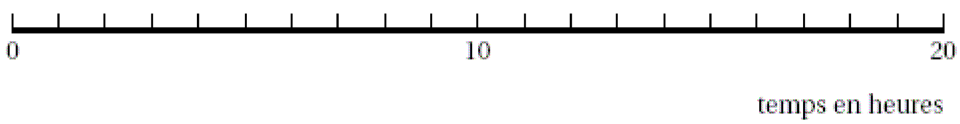


Diagramme numéro 2 Classe de première L2

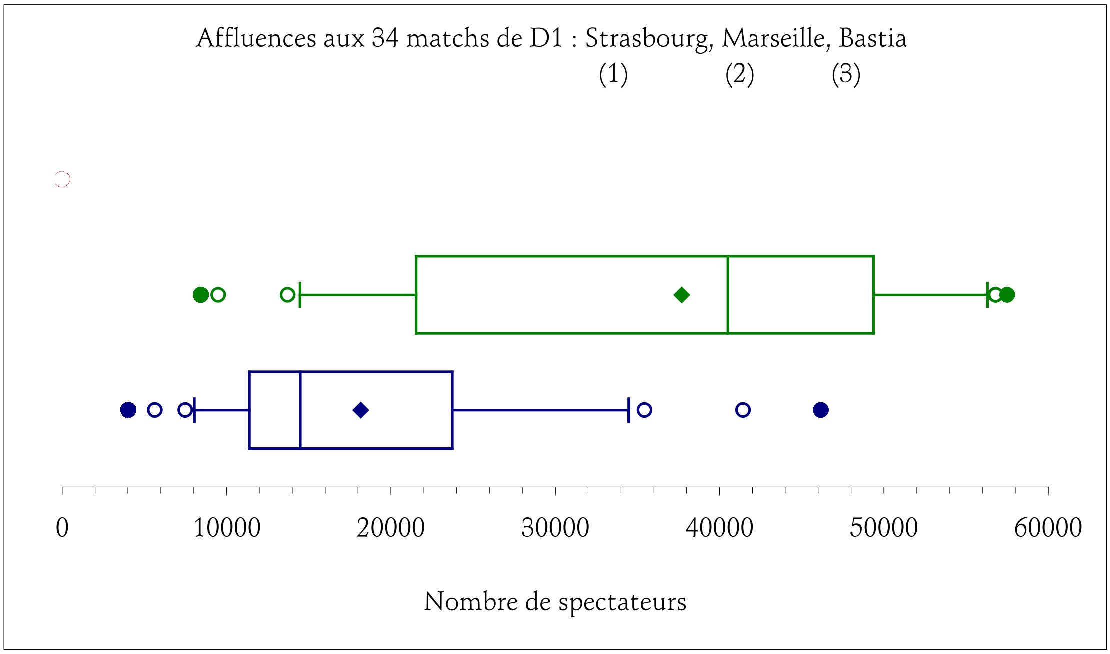


24. Football : Championnat de France de D1 - Saison 1998-1999

Nombre de spectateurs lors des 34 matchs de championnat joués par Strasbourg, Marseille et Bastia.

Pour chaque club de 1ère division, on a relevé le nombre de spectateurs à chacun des 34 matchs de la saison 1998 - 1999. Voici, ci-dessous, les diagrammes en boîtes représentant les deux séries ainsi obtenues pour Strasbourg (1) et Marseille (2).

Outre les marqueurs habituels (déciles, quartiles et médianes) on y lit aussi les moyennes (losanges).



1. Pour Strasbourg et Marseille, lire la moyenne et la médiane (à 1000 unités près).
2. Faire une phrase d'interprétation concrète pour chacune de ces quatre valeurs.
3. Dire si les affirmations suivantes sont possibles ou impossibles en donnant un argument pour chaque réponse.

a. Strasbourg a joué 20 matchs devant plus de 20 000 spectateurs.

b. Marseille n'a joué que 8 matchs devant moins de 20 000 spectateurs.

c. Marseille a joué 10 matchs devant plus de 50 000 spectateurs.

d. Strasbourg a joué 26 matchs devant moins de 24 000 spectateurs.

4. Comparer les affluences aux matchs de Strasbourg et Marseille en 98-99. Dire, en particulier, pour quel club l'écart-type de la série est le plus élevé.

5. Pour le club Bastia, la série obtenue est : (résultats arrondis à 1 000 unités près).

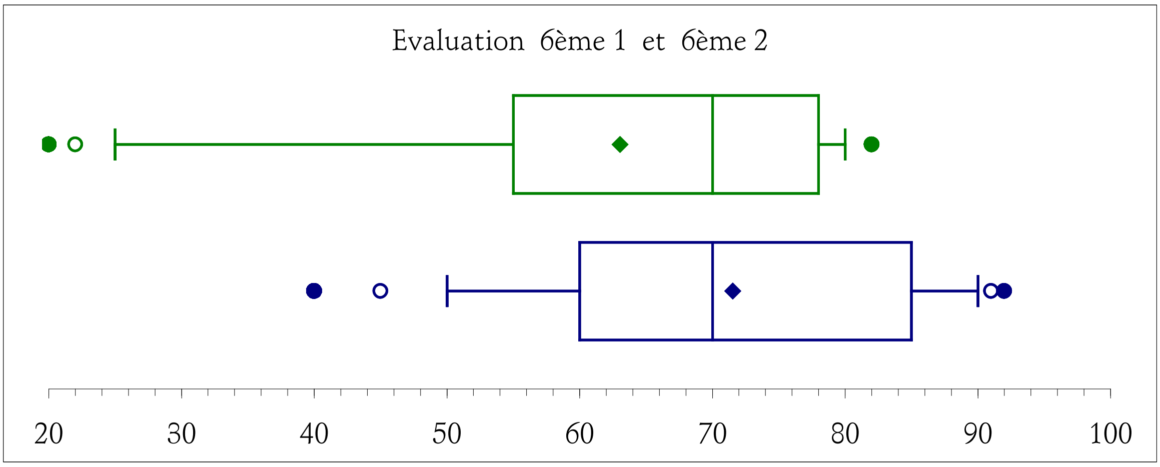
Minimum	3000	Quartile 3	13000
Décile 1	4000	Décile 9	26000
Quartile 1	5000	Maximum	48000
Médiane	7000	moyenne	11000

Représenter cette série par un diagramme en boîte sur le graphique ci-dessus.

25. Evaluation en 6ème

Voici, sous forme de diagrammes en boites, les résultats de deux classes de 6ème de deux collèges de la région parisienne à l'évaluation de la rentrée 1999.

Les diagrammes résument les résultats en % d'items réussis dans les classes de 6ème 1 et 6ème 2. Outre les éléments habituels (max, min, déciles, quartiles, médianes), on peut encore lire le pourcentage moyen de réussite (losanges).



1. Donner le pourcentage moyen de réussite dans chaque classe.
2. Donner la médiane commune aux deux classes. Comment interpréter ce nombre ?
3. Comment expliquer le fait que, en 6ème 2, le pourcentage moyen soit inférieur à la médiane ?
4. Il y avait 25 élèves en 6ème 1. Combien d'élèves de 6ème 1, au moins, ont obtenu un score:
   - inférieur ou égal à 60%,
   - inférieur ou égal à 70%.
5. Comparer les profils des deux classes en ce qui concerne les scores de réussites à l'évaluation ?

26. Amérique du Nord, juin 2004, 12 points

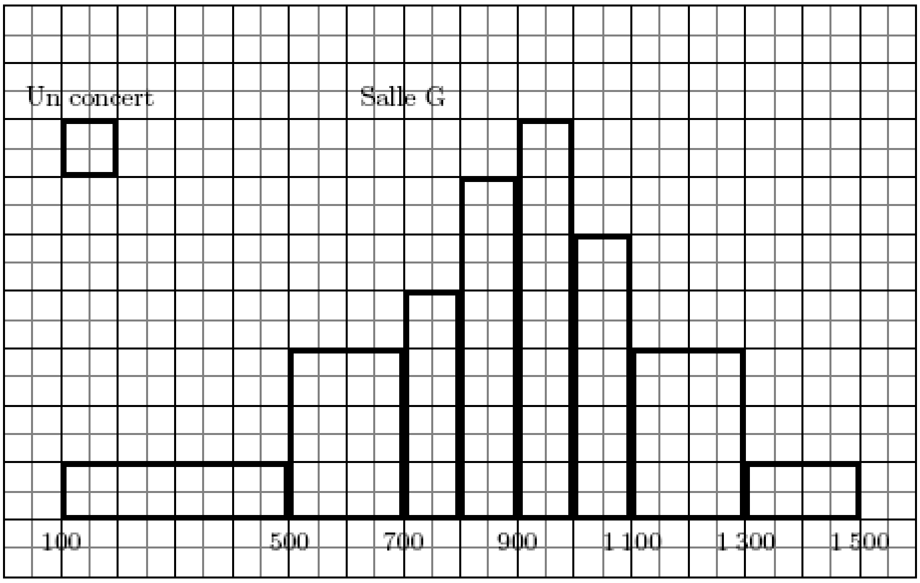
Dans une ville existent deux salles de spectacles ayant programmé chacune 40 concerts durant la saison 2004/2005. La salle G est spécialisée dans la musique classique et la salle J dans le jazz.

Pour la salle G, les résultats en nombre de spectateurs sont indiqués par l'histogramme ci-dessous.

1. a. A la vue de cet histogramme, peut-on penser que ces données sont gaussiennes ? Justifier.

b. Remplir le tableau ci-dessous à l'aide de l'histogramme, directement sur le sujet.

Nombre de spectateurs	[100 ; 500[	[500 ; 700[	[700 ; 800[	[800 ; 900[	[900 ; 1000[	[1000 ; 1100[	[1100 ; 1300[	[1300 ; 1500[
Nombre de concerts		6						



2. a. Calculer, en utilisant les milieux de classes, la moyenne $\bar{x}$ de cette série statistique puis sa variance V et son écart-type σ (on arrondira les résultats à l'unité prés et on laissera les calculs sur la copie). Que signifie $\bar{x}$?

b. Pour cette question, on suppose que $\bar{x}=870$ et $\sigma=284$. Calculer le nombre de concerts pour lesquels le nombre de spectateurs est dans $\left[\bar{x}-2\sigma\,;\bar{x}+2\sigma\right]$.

A quel pourcentage cela correspond-il ? Cela confirme-t-il ou non la réponse à la question 1.a.? Justifier.

3. Les statistiques concernant la salle J sont données sur la feuille de calcul ci-dessous, réalisée à l'aide d'un tableur.

Les cellules A5 à A11 contiennent les classes de nombres de spectateurs, toutes d'amplitude 200.

Les cellules B5 à B11 contiennent les milieux de classes.

Les cellules C5 à C11 contiennent les nombres de concerts correspondant aux classes de la colonne A.

	A	B	C	D
1				
2	Classes	Milieux de	Nombre de	Spectateurs
3		Classes	Concerts	2004/2005
4				
5	[0 ; 200[	100	4	400
6	[200 ; 400[	300	8	
7	[400 ; 600[	500	4	
8	[600 ; 800[	700	2	
9	[800 ; 1000[	900	6	
10	[1000 ; 1200[	1100	10	
11	[1200 ; 1400[	1300	6	
12				
13		Somme :	40	30400
14				

15			Moyenne :	
16				

a. Le gérant veut obtenir, en utilisant le tableur, le nombre moyen de spectateurs par concert pour la saison 2004/2005.

Dans la cellule D5 figure 400 qui représente le nombre de spectateurs susceptibles d'avoir assisté aux quatre concerts relatifs à la première classe.

Quelle formule le gérant a-t-il saisi dans D5, sachant qu'elle doit être recopiée jusqu'à D11, pour obtenir les nombres concernant les autres classes ?

Inscrire les résultats des cellules D6 à D11 directement dans le tableau ci-dessus.

Quelle formule le gérant a-t-il saisi dans D13 ?

Quelle formule doit-il saisir dans D15 pour avoir le nombre moyen de spectateurs par concert dans la salle J ? Inscrire ce nombre dans la cellule D15 ci-dessus.

b. Trouver, pour la série concernant la salle J, les classes respectives contenant la médiane et les quartiles du nombre de spectacles.

27. Centres étrangers juin 2004, 10 points

À la fin des délibérations d'un examen comportant trois épreuves, un professeur relève les résultats de ses 30 élèves aux épreuves n° 1, n° 2 et n° 3. Ces notes sont regroupées dans le tableau suivant :

	Effectifs		
Notes sur 20	Épreuve n°1	Épreuve n°2	Épreuve n°3
5	0	3	0
6	6	0	0
7	5	5	2
8	8	0	1
9	1	8	6
10	3	0	3
11	0	3	5
12	2	4	0
13	0	0	2
14	1	1	6
15	2	4	3
16	2	2	2

1. Dans cette question, on s'intéresse à la série statistique E_1 formée des notes à l'épreuve n° 1.

a. Déterminer, pour cette série statistique, le minimum et le maximum.

b. Déterminer la médiane. Justifier.

c. Déterminer les 1^{er} et $3^{ème}$ quartiles. Justifier.

d. Tracer le diagramme en boîte correspondant à cette série E_1, sur la feuille fournie en annexe, avec le minimum et le maximum pour valeurs extrêmes.

2. On s'intéresse maintenant à la série statistique E_2 formée des notes à l'épreuve n° 2.

a. Dresser le diagramme en boîte correspondant à cette série, sur la feuille annexe, avec le minimum et le maximum pour valeurs extrêmes. On précisera les valeurs utilisées.

b. Calculer la moyenne arithmétique de la série E_2.

c. Donner la valeur de l'écart-type de la série E_2.

3. Quels commentaires pouvez-vous faire en comparant les deux diagrammes en boîte correspondant aux séries E_1 et E_2.

4. On note E_3 la série statistique formée des notes à l'épreuve n° 3. On admet que l'écart-type de la série E_3 est 2,7.

a. Calculer la moyenne arithmétique de la série E_3.

b. Calculer le pourcentage d'élèves ayant une note inférieure ou égale à 9 dans l'épreuve n° 3.

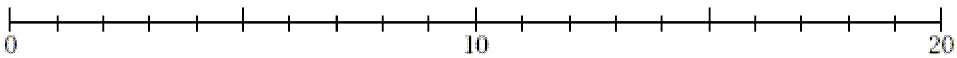
c. Quels commentaires pouvez-vous faire en comparant les résultats de l'épreuve n° 2 avec ceux de l'épreuve n° 3 ?

5. Sachant que la moyenne arithmétique à l'épreuve n° 1 est 9,13 et que cette épreuve n° 1 est affectée du coefficient 3 et les épreuves n° 2 et n° 3 du coefficient 1, quelle est la moyenne arithmétique, sur 20, des notes des 30 élèves à cet examen ?

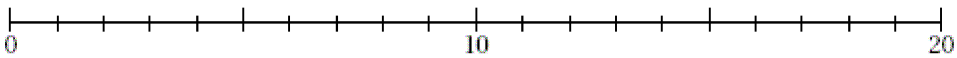
Annexe (à rendre avec la copie)

Exercice 2

Série statistique E_1 - Diagramme en boîte.



Série statistique E_2 - Diagramme en boîte.



28. France, juin 2004, 9 points

Le tableau (incomplet) ci-dessous donne la répartition des 800 chefs d'exploitation agricole d'une région selon leur âge et l'aire de la Surface AgricoleUtile (S.A.U.) de leur exploitation. L'aire est exprimée en hectares (ha) et l'âge en années.

S.A.U / tranche d'âge	[0 ; 10[	[10 ; 30[	[30 ; 50[	[50 ; 100[	TOTAL
[15 ; 25[	2	1	5	3	
[25 ; 35[	21		16	28	84
[35 ; 45[	40	33		59	148
[45 ;55[	17		53	123	
[55 ; 65[	110	60	70	57	297
TOTAL	190	180		270	800

Partie A

1. Compléter le tableau (on recopiera sur la copie les colonnes complétées correspondant à une Surface Agricole Utile de [10 ; 30[et de [30 ; 50[).

2. Les pourcentages demandés dans cette question seront arrondis à 0,1%.

a. Parmi les chefs d'exploitation agricole, quel est le pourcentage de ceux dans la tranche d'âge [25 ; 35[?

b. Parmi les chefs d'exploitation agricole, quel est le pourcentage de ceux âgés de strictement moins de 45 ans et possédant au moins 30 ha de Surface Agricole Utile ?

c. Parmi les chefs d'exploitation agricole de 55 ans ou plus, quel est le pourcentage de ceux qui ont une Surface Agricole Utile de moins de 10 ha ?

d. Parmi les chefs d'exploitation agricole de Surface AgricoleUtile de moins de 10 ha, quel est le pourcentage de ceux âgés de 55 ans ou plus ?

Partie B

1. a. Combien de chefs d'exploitation agricole ont strictement moins de 45 ans ?

Strictement moins de 55 ans ?

b. Expliquer pourquoi l'âge médian des chefs d'exploitation agricole est nécessairement entre 45 et 55 ans.

Pour déterminer l'âge médian, la répartition des âges dans la classe [45 ; 55[est donnée par le tableau suivant :

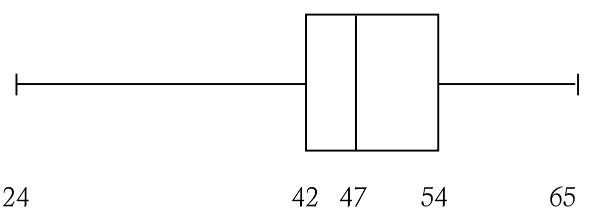
ÂGE	45	46	47	48	49	50	51	52	53	54
EFFECTIF	18	21	24	31	30	31	30	27	28	20

Combien de chefs d'exploitation ont 45 ans ou moins ? Justifier que l'âge médian est de 51 ans.

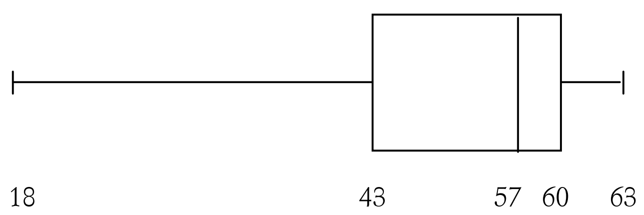
c. Les premier et troisième quartiles de la série des âges sont 42 et 58. Construire le diagramme en boîte de cette série en prenant en compte les valeurs extrêmes 18 et 65. On choisira comme échelle 2mm pour une année.

2. Le diagramme en boîte des âges des chefs d'exploitation de 50 ha à 100 ha et celui des chefs d'exploitation demoins de 10 ha sont représentés ci-dessous.

Exploitations de 50 ha à 100 ha



Exploitations de moins de 10 ha



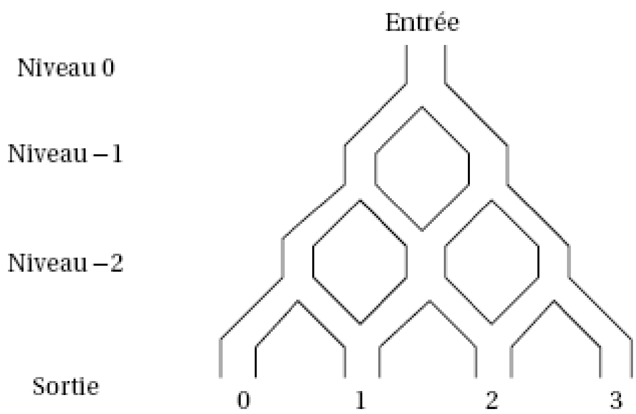
Un journaliste a écrit : « Dans leur ensemble les chefs d'exploitation de 50 à 100 ha sont plus jeunes que les chefs d'exploitation de moins de 10 ha. »

Commenter cette affirmation en utilisant ces diagrammes en boîtes.

29. Liban, juin 2004, 8 points

Une souris descend dans une canalisation (schématisée par la figure ci-dessous) aboutissant aux sorties 0, 1, 2, 3. On suppose quelle progresse vers l'arrivée en se dirigeant au hasard à chaque niveau vers la droite ou vers la gauche pour accéder au niveau inférieur.

Un parcours possible peut donc se coder GGD, où G signifie « aller vers la gauche » et D « aller vers la droite », à chacun des trois niveaux. On s'intéresse alors au numéro de la sortie de la souris.



Partie A Étude théorique

Trouver tous les chemins possibles (éventuellement l'aide d'un arbre) et compléter alors le tableau des fréquences théoriques (tableau 1 de l'annexe de l'exercice 1).

Partie B Simulation à l'aide d'un tableur

À l'aide d'un tableur, on effectue une simulation de 100 progressions de la souris dans la canalisation : on obtient ainsi les fréquences correspondant à chacune des sorties possibles de la souris. On note alors la fréquence correspondant à la sortie n°1 obtenue.

En effectuant 50 simulations, on obtient 50 fréquences correspondant à la sortie n°1 (ces fréquences sont relevées dans le tableau 2 de l'annexe de l'exercice 1).

1. On admet que la série des 50 fréquences a pour moyenne $m = 0{,}364$ et pour écart-type $s = 0{,}051$, résultats donnés avec trois chiffres après la virgule. Calculer le pourcentage de valeurs de la série situées dans l'intervalle $[m-2s\ ;\ m+2s]$.

Ce résultat correspond-il à ce que l'on peut attendre d'une série gaussienne ou normale ? Justifier.

2. On effectue ensuite deux séries de 50 simulations, l'une correspondant à 500 progressions de la souris, l'autre à 1000 progressions et on obtient 50 fréquences de la sortie n°1 pour chaque série.

Le graphique de l'annexe de l'exercice 1 représente les diagrammes en boîte (ou boîtes à moustaches) de ces deux séries.

Dessiner, sur le même graphique, le diagramme en boîte qui correspond à la série des 50 simulations effectuées dans la question 1. en calculant tous les éléments nécessaires pour construire ce type de boîte.

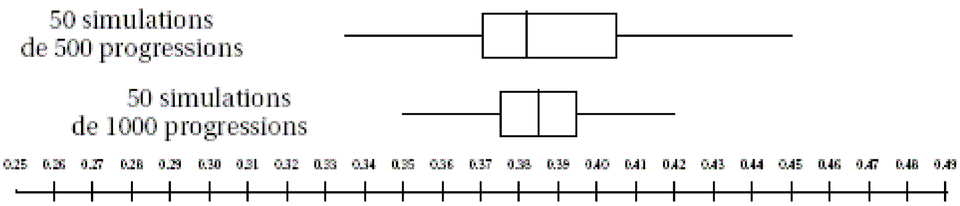
3. a. À l'aide des trois diagrammes, déterminer la série qui semble donner les fréquences les plus proches de la fréquence théorique.

b. Que faudrait-il faire pour s'en approcher encore davantage ?

Annexe 1

Sortie n°	0	1	2	3
Nombre de chemins possibles				
Fréquences théoriques en %				

0,250	0,260	0,290	0,290	0,300	0,300	0,310	0,310	0,320	0,320
0,320	0,320	0,330	0,330	0,330	0,340	0,340	0,340	0,340	0,350
0,350	0,350	0,350	0,350	0,360	0,360	0,370	0,370	0,370	0,370
0,380	0,380	0,380	0,390	0,390	0,390	0,390	0,400	0,400	0,410
0,410	0,420	0,420	0,420	0,430	0,430	0,450	0,460	0,470	0,470



30. Polynésie, juin 2004, 9 points

Les parties A et B sont indépendantes

L'infirmière d'un lycée décide de mener une enquéte sur la qualité des repas servis à la cantine scolaire de son établissement.

Partie A

Elle réalise cette enquête lors du repas de midi du 26 septembre 2003 auprès des 150 élèves des classes de Première. Elle dispose des renseignements suivants :

- 105 mangent au lycée ce jour-là ;
- 131 ne sont pas allergiques au tait
- 3 sont allergiques au lait et ne mangent pas au lycée ce jour-là.

1. Compléter le tableau donné en annexe et donner le nombre d'élèves de première qui mangent au lycée ce 26 septembre et ne sont pas allergiques au lait.

2. L'infirmière fait des propositions de repas aux élèves participant à l'enquête en précisant que tout menu doit comporter obligatoirement une entrée, un plat principal, un accompagnement, un fromage et un dessert. Ces propositions sont données ci-dessous :

Entrée	Oeuf
Plat principal	Viande (portion de 120 g) Poisson (portion de 120 g) Frites (portion moyenne)
Accompagnement	Légumes verts (portion moyenne) Pâtes (portion moyenne
Fromage	Fromage blanc (portion de 100 g) Gruyère (portion de 30 g) Bleu (portion de 30 g)
Dessert	1 petit suisse Fruit (150 g)

Chaque élève compose son menu.

Quel est le nombre de menus différents que l'on peut obtenir à partir des propositions faites par l'infirmière ?

Partie B

Dans une circulaire du Ministre de l'Éducation Nationale relative à la restauration scolaire, il est écrit :

« L'apport minimal de calcium (...) est rarement assuré (...). Le repas de midi doit donc apporter pour les adolescents 300 à 400 mg de calcium ». (*B.O. Spécial n° 9 du 28 juin 2001*)

L'infirmière décide de détecter les éventuelles carences en calcium. Elle mène une autre enquête dans laquelle elle interroge les élèves de première qui ont mangé au lycée le 26 septembre 2003 à midi et qui ne sont pas allergiques au lait. À partir du menu choisi par chacun d'eux, elle calcule l'apport en calcium correspondant. Elle note les résultats de cette enquête dans le tableau donné ci-dessous. On appellera

S_1 la série statistique ainsi formée.

Apport en calcium (en mg)	106	173	190	192	198	231	259	315	317	341	407	409
Nombre d'élèves	1	2	7	6	11	16	21	7	3	6	6	3

1. Parmi les élèves participant à cette enquête, quel est le pourcentage de ceux pour lesquels l'apport en calcium lors de ce repas est conforme à la recommandation ministérielle ?

2. Donner la moyenne et l' écart-type de la série statistique S_1.

3. Jugeant la moyenne de l'apport en calcium trop faible, l'infirmière décide de distribuer à chaque élève un verre de lait. Elle suppose que tous les élèves boivent ce verre de lait et ajoute l'apport en calcium correspondant, soit 120 mg, aux résultats précédents, elle obtient une nouvelle série S_2 d'apports en calcium. Comment déduire des valeurs calculées pour la série S_1 la moyenne et l'écart-type de la série S_2 ?

4. En fait, après leur départ, l'infirmière s'aperçoit que 15 élèves n'ont pas bu leur verre de lait, mais elle ne sait pas lesquels. Déterminer l'apport moyen en calcium. Peut-on donner l'écart-type ?

ANNEXE

À rendre avec la copie

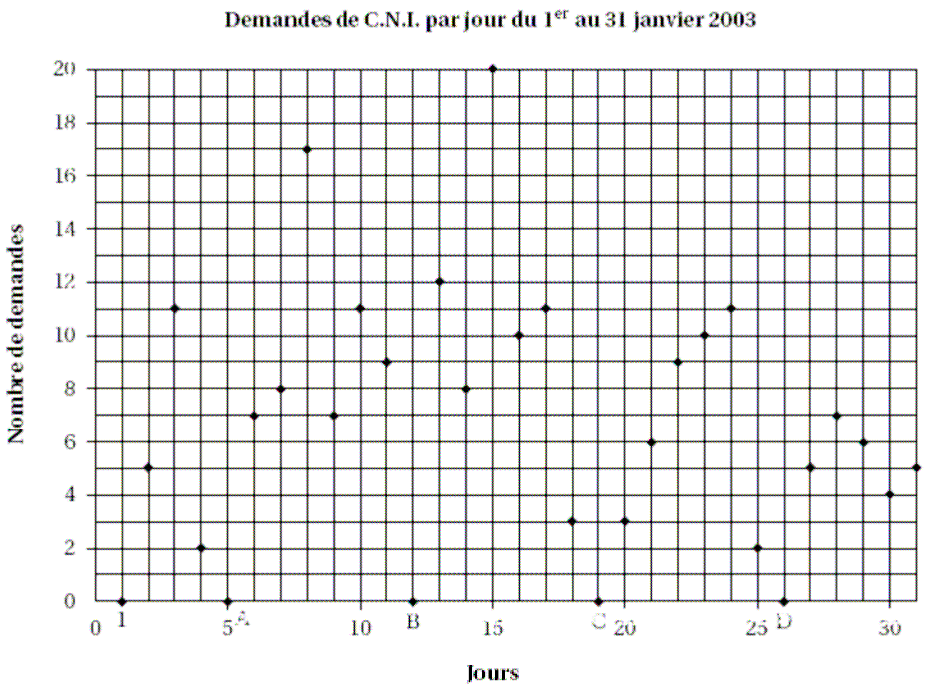
EXERCICE 1

	Élèves mangeant au lycée	Élèves ne mangeant pas au lycée	Total
Élèves allergiques au lait			
Élèves non allergiques au lait			
Total			

31. Amérique du Sud, novembre 2004, 8 points

Les données chiffrées de cet exercice proviennent du service « Formalités administratives » d'une commune de 51 137 habitants de l'Est de la France. Ce service est ouvert du lundi matin au samedi douze heures et reçoit, entre autres, les demandes de cartes nationales d'identité (C.N.I.).

Partie A



1. Combien de demandes ont été déposées le 3 janvier, le 12 janvier ?

2. À quel jour de la semaine correspondent les points A, B, C et D situés sur l'axe des abscisses. Justifier votre réponse.

3. Un agent de ce service affirme que lemercredi est un jour d'affluence particulière. Qu'en pensez-vous ?

Partie B

On a extrait du graphique précédent les nombres de demandes de C.N.I. traitées par jour, pour chacun des jours où le service est ouvert le matin et l'après-midi (les lundis, mardis, mercredis, jeudis et vendredis) au cours du mois de janvier 2003 :

5 ; 11 ; 7 ; 8 ; 17 ; 6 ; 11 ; 12 ; 8 ; 20 ; 10 ; 11 ; 3 ; 6 ; 9 ; 10 ; 11 ; 5 ; 7 ; 6 ; 4 ; 5.

1. Calculer le nombre moyen de demandes de C.N.I. traitées par jour de cette série (le résultat sera arrondi à l'entier le plus proche).

2. Déterminer la médiane m, le premier quartile Q_1, le troisième quartile Q_3 de cette série.

3. Construire le diagramme en boîte de cette série sur la feuille annexe.

4. On estime que l'organisation du service est efficace si pendant au moins la moitié des jours où le service est ouvert le matin et l'après-midi, le nombre de demandes traitées journellement est dans l'intervalle [6 ; 11].

L'organisation est-elle satisfaisante ? Justifier votre réponse.

32. Nouvelle Calédonie, novembre 2004, 10 points

On a étudié les fréquences cardiaques d'un groupe de 60 sportifs amateurs hommes et femmes (appelé groupe I), pratiquant leur sport de 2 à 4 fois par semaine.

La fréquence cardiaque est le nombre de pulsations du coeur parminute.

Pour chacun de ces sportifs du groupe I, on mesure **la fréquence cardiaque au repos (FCR),** c'est-à-dire la fréquence cardiaque la plus faible rencontrée chez cette personne, mesurée après plusieurs essais après une longue période de calme et de repos.

Les résultats de cette étude sont récapitulés dans te tableau ci-dessous où les fréquences cardiaques au repos (FCR) des 60 sportifs du groupe I sont classées par ordre croissant.

Age	FCR
42	42
41	43
61	45
51	45
41	46
27	46
33	46
40	48
55	48
31	48
32	48
35	48
44	49
40	50
36	50

Age	FCR
50	50
35	50
24	50
23	50
52	50
36	51
31	51
35	51
60	51
29	52
30	52
49	52
32	52
40	52
47	52

Age	FCR
37	52
42	52
21	52
40	53
34	53
35	53
28	53
55	53
49	53
31	53
35	53
38	54
53	54
42	54
54	54

Age	FCR
41	54
31	55
50	55
32	55
22	55
42	55
52	55
18	57
51	59
22	59
23	59
53	59
50	59
28	59
47	61

1. a. Déterminer la médiane ainsi que les premier et troisième quartiles de la série des FCR.

b. Construire sur l'axe D_1 de l'annexe 2 à rendre avec la copie, un diagramme en boîte pour cette série.

2. a. Compléter le tableau de l'annexe 2 et tracer, sur la copie, une représentation graphique de la série des FCR des 60 sportifs du groupe I.

b. Calculer la moyenne x de cette série.

3. a. On suppose que les FCR des sportifs du groupe I sont des données gaussiennes dont l'écart-type σ est égal à 4,06. Déterminer l'intervalle $[52-2\sigma\ ;\ 52+2\sigma]$. Comment nomme-t-on cet intervalle ?

b. Calculer le pourcentage de sportifs dont la FCR est située dans cet intervalle. Était-il possible de prévoir ce résultat ? Expliquer.

4. On souhaite comparer les FCR des sportifs du groupe I aux FCR d'un groupe de 60 personnes pratiquant peu d'activité physique (appelé groupe II).

L'étude des FCR des personnes du groupe II a donné les résultats suivants :

- Moyenne : 59,8
- Écart-type : 6,23
- Médiane : 60
- Premier quartile : 57
- Troisième quartile : 63
- Valeur minimale : 45
- Valeur maximale : 70

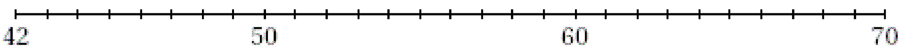
a. Sur l'axe D_2 de l'annexe 2 à rendre avec la copie, tracer un diagramme en boîte pour les FCR des personnes du groupe II.

b. Quelle incidence semble avoir la pratique régulière d'activités sportives sur la FCR d'un individu ?

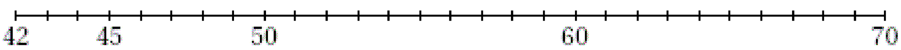
Annexe à rendre avec la copie

Annexe 2

Axe D_1



Axe D_2



Tableau

FCR	42	43	45	46	48	49	50	51	52	53	54	55	57	59	61
Nombre d'individus															

33. Antilles, septembre 2004, 12 points

Écriture de mots

La langue française comporte 26 lettres de l'alphabet plus les lettres avec accents ou tréma, soit 36 caractères qui permettent d'écrire les mots.

Un mot est une liste de caractères distincts ou non ayant un sens ou non, par exemple « cab » et « eta » sont deux mots.

Un mot simple est un mot dont les caractères sont tous distincts. Par exemple « cab» est un mot simple mais « cca » n'est pas un mot simple.

La longueur d'un mot est le nombre de caractères qui le composent : par exemple, le mot « littéraire » a pour longueur 10.

Partie A - Nombre de mots possibles de longueur donnée

On souhaite calculer :

- le nombre N de mots possibles de longueur inférieure ou égale à 5.
- le nombre S de mots simples possibles ayant une longueur donnée inférieure ou égale à 5.

On décide d'utiliser un tableur. La feuille de calcul correspondant à ce travail est donnée ci-dessous. Compléter ce tableau au fur et à mesure.

	A	B	C
1	Longueur du mot	Nombre de mots possibles	Nombre de mots simples possibles
2	1	36	36
3	2	1 296	1 260
4	3		
5	4		
6	5		
7	Total		

1. Calcul de N

a. Justifier les résultats des cellules B2 et B3.

b. On admet que les résultats de la colonne B sont les premiers termes d'une suite géométrique. Montrer que la raison de cette suite est égale à 36. Donner le premier terme.

c. Quel type de croissance cette suite traduit-elle ?

d. Quelle formule doit-on saisir dans la cellule B3 pour que par recopie on obtienne les termes de la suite jusqu'à la cellule B6 ?

e. Compléter la colonne B jusqu'en cellule B6.

f. Quelle formule doit-on saisir dans la cellule B7 pour obtenir N ? Calculer N.

2. Calcul de S

a. Justifier les résultats des cellules C2 et C3.

b. Justifier que l'on peut saisir dans la cellule C3 la formule suivante **= C2*(36 – A2)** pour que par recopie jusqu'en la cellule C6 on obtienne les nombres demandés.

c. Compléter la colonne C.

d. Quelle formule doit-on saisir dans la cellule C7 pour obtenir le nombre demandé S ? Calculer S.

Partie B

Un texte de Charles Perrault est écrit en quatre langues :

Les amours de la règle et du compas

En français	*En anglais*
Les amours de la règle et du compas	**A love story between a ruler and a compass**
Toutefois nos amours, répliqua le compas	Hovewer, our love, replied the compass
Produiront des enfants qui vaincront le trépas	Will produce children who will overcome death
De nous deux sortira la belle architecture	Fromus both a beautiful architecture will come out
Et mille nobles arts pour polir la nature, [. . .]	And a thousand noble arts to enhance nature
Le compas aussitôt sur un pied se dressa ,	Immediately, the compass stood on his foot
Et de l'autre, en tournant un grand cercle traça.	While he drew a great circle with the other one
La règle en fut ravie et soudain se vint mettre	The ruler was delighted and suddenly came to lie
Dans le milieu du cercle, et fit le diamètre.	In the center of the circle and draw a diameter
Son amant l'embrassa, l'ayant à sa merci	Her lover kissed her, having her at his mercy
Tantôt s'élargissant et tantôt raccourci	Either widening or shortening

Et l'on vit naître de leurs doctes postures Triangles et carrés et mille autres figures	And came to birth, from their learned posture Triangles and squares and a thousand other figures
En italien **Gli amori della riga del compasso** Tuttavia, i nostri amori, replicó il compasso Produrranno figli che vinceranno iltra passo, Da noi due uscira la bell'arcittetura, Emille nobili arti per raffinare la natura. Subito el compasso su in piede si raddrizz'o, E dell'altro, girando, un gran cerchio disegn'o. La riga ne fu meravigliata, e ad une tratto venne a collocarsi Nelmezzo del cerchio, e fece il diametro. Siccome era in babia dell'amante, questo bacio, Ora allargandosi, ora accoorciato, E dalle loro dotte posture, si video nascere Triangoli e quadrati emille altre figure.	En allemand **Die Liebschaften des Lineals und des Kompass** Immerhin wird unsere Liebe Kinder erzeugen, Erwidere der Kompass, die den Tod überwinden werden. Aus uns beidenwerden schöne Architktur und tausende vornehme Künste entstehen, um die Natur zu verfeinern. Sogleich erhob sich der Kompass auf einen Fuß Und mit dem anderen entwarf er einen groußen Kreiss. Das Lineal war entzückt und bildete den Durchmesser. Sien Liebhaber umarmte es, es war ihmausgeliefert. Bald dehnte er sich aus, bald zog er sich zusamla men. Aus ihren gelehten Haltungen entwickelten sich Quadrate und Dreiecke und tausende andere Gestalten.

Le tableau donne le nombre de mots d'une longueur donnée dans chacune des langues.

Longueur du mot	1	2	3	4	5	6	7	8	9	10	11	12	Total
Nombre de mots en français	6	32	9	7	14	19	4	6	4	1	1	1	104
Nombre de mots en anglais	8	10	24	16	13	8	12	7	3	1	1	1	104
Nombre de mots en italien	10	19	11	9	15	10	8	7	3	3	1	3	99
Nombre de mots en allemand	0	7	29	8	7	11	7	11	6	2	2	3	93

Construire les diagrammes en boîte des quatre séries statistiques correspondant aux quatre langues.

34. France, septembre 2004, 8 points

On s'intéresse au jeu « Keno » de la Française Des Jeux.

L'une des façons de jouer est la suivante : dans une grille contenant une fois chacun les nombres de 1 à 70, on choisit 10 numéros. Un tirage au sort de 20 numéros a lieu : une grille est gagnante dans l'un des deux cas suivants :

- soit aucun des numéros sortis n'a été trouvé ;
- soit au moins cinq numéros sortis ont été trouvés.

Dans l'annexe 1 on trouve un extrait tiré des règles figurant au dos des bulletins.

Sur 10000 bulletins, on a obtenu les résultats suivants :

nombre de numéros trouvés	0	1	2	3	4	5	6	7	8	9	10
effectif	254	1253	2521	2922	1962	822	220	41	5	0	0

Par exemple, le nombre de bulletins où on a trouvé exactement deux bons numéros est de 2521.

1. a. Combien y a-t-il de bulletins gagnants ?

b. Quel pourcentage cela représente-t-il ?

c. Ce pourcentage est-il proche du « 1 sur 7,4 » annoncé dans le tableau de l'annexe ?

2. Sur l'échantillon observé, combien un bulletin contient-il de bons numéros en moyenne ?

3. Déterminer, en expliquant votre démarche, la médiane ainsi que le premier et le troisième quartile de la série résumée par le tableau.

4. Construire le diagramme en boîte correspondant.

5. Les affirmations suivantes sont-elles vraies ou fausses ? Justifier la réponse en utilisant uniquement les indicateurs de la série.

a. Au moins lamoitié des bulletins comporte au plus 2 bons numéros.

b. 25 % au plus des bulletins comportent 4 bons numéros ou davantage.

c. Au moins 50 % des bulletins comportent de 2 à 4 bons numéros.

6. Les 10 000 joueurs ont misé 3 € chacun : ils ont donc dépensé 30 000 €. Calculer le total des gains redistribués.

<u>Annexe</u>

Numéros joués par grille	Vos chances totales de gagner	Numéros trouvés par grille	Vos chances totales de gagner	Gain x fois la mise	Gain pour une mise de 1,5 €	Gain pour une mise de 3 €
10 numéros	1 sur 7,4	10	1 sur 2147181	x 200 000	300 000 €	600 000 €
		9	1 sur 47238	x 2 500	3750 €	7 500 €
		8	1 sur 2571	x 100	150 €	300 €
		7	1 sur 261	x 10	15 €	30 €
		6	1 sur 44	x 5	7,5 €	15 €
		5	1 sur 12	x 2	3 €	6 €
		0	1 sur 39	x 2	3 €	6 €

35. Pondicherry, avril 2005, 8 points

Un fabricant de barres chocolatées a fait imprimer, en grande quantité, le même nombre d'images de trois chanteuses Mlle Pinson, Mlle Rossignol et Mlle Décibel.

L'image de Mlle Pinson porte le n° 1 celle de Mlle Rossignol le n° 2, et celle de Mlle Décibel le n° 3. Une machine insère au hasard une image dans chaque barre chocolatée fabriquée.

Il y a autant de barres chocolatées contenant l'image de chaque chanteuse.

Chaque jour, Aline achète une barre chocolatée. Elle voudrait obtenir la collection complète des trois chanteuses et se demande au bout de combien de jours elle l'obtiendra.

<u>Partie 1</u>

Aline a répertorié à l'aide d'un arbre les différentes images qu'il est possible d'obtenir sur trois jours. Cet arbre, partiellement complété, se trouve dans l'annexe 2.

Par exemple, la 3e possibilité 1 1 3 signifie que le premier jour, la barre chocolatée contient l'image de Mlle Pinson, le deuxième jour, elle contient celle de Mlle Pinson, et le troisième jour celle de Mlle Décibel.

1. Parmi ces 27 possibilités, combien en compte-t-on qui permettent d'obtenir une collection complète ?

2. Y a-t-il plus de 25% des cas dans lesquels on obtient une collection complète ? Justifiez.

Aline veut obtenir la collection complète. Son argent de poche étant limité, elle aimerait estimer le nombre de jours au bout desquels elle peut espérer obtenir la collection complète. Elle va pour cela effectuer des simulations.

<u>Partie 2</u>

Elle effectue une simulation en faisant afficher à sa calculatrice une liste aléatoire de nombres, de telle manière que chacun des nombres 1, 2 et 3 ait la même chance d'apparition.

Voici la liste qu'elle obtient

$$1 - 1 - 1 - 1 - 1 - 3 - 1 - 2 - 1 - 3$$

Selon cette simulation, les cinq premiers jours, Aline découvre dans sa barre chocolatée l'image de Mlle Pinson, le 6^e jour, celle de Mlle Décibel, le 7^e jour celle de Mlle Pinson, le 8^e jour celle de Mlle Rossignol, le 9^e jour celle de Mlle Pinson, et le 10^e celle de Mlle Décibel.

Aline est donc en possession de la collection complète au 8^e jour.

Imaginez, sur le modèle précédent, une liste de 9 nombres conduisant à la collection complète obtenue au 5^e jour.

<u>Partie 3</u>

Pour se faire une idée plus précise, Aline effectue 1000 simulations à l'aide de sa calculatrice. Les résultats obtenus figurent dans le tableau suivant :

nombre de jours nécessaires à l'obtention de la collection complète	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Effectifs	227	203	179	126	99	56	40	25	18	12	4	3	3	2	1	2

Cela signifie par exemple que parmi les 1 000 simulations, 203 sont des situations pour lesquelles la collection complète des images est obtenue au 4^{e} jour.

1. Déterminez la médiane, le premier et le troisième quartile de cette série statistique.

2. Aline formule deux remarques en observant ces résultats simulés.

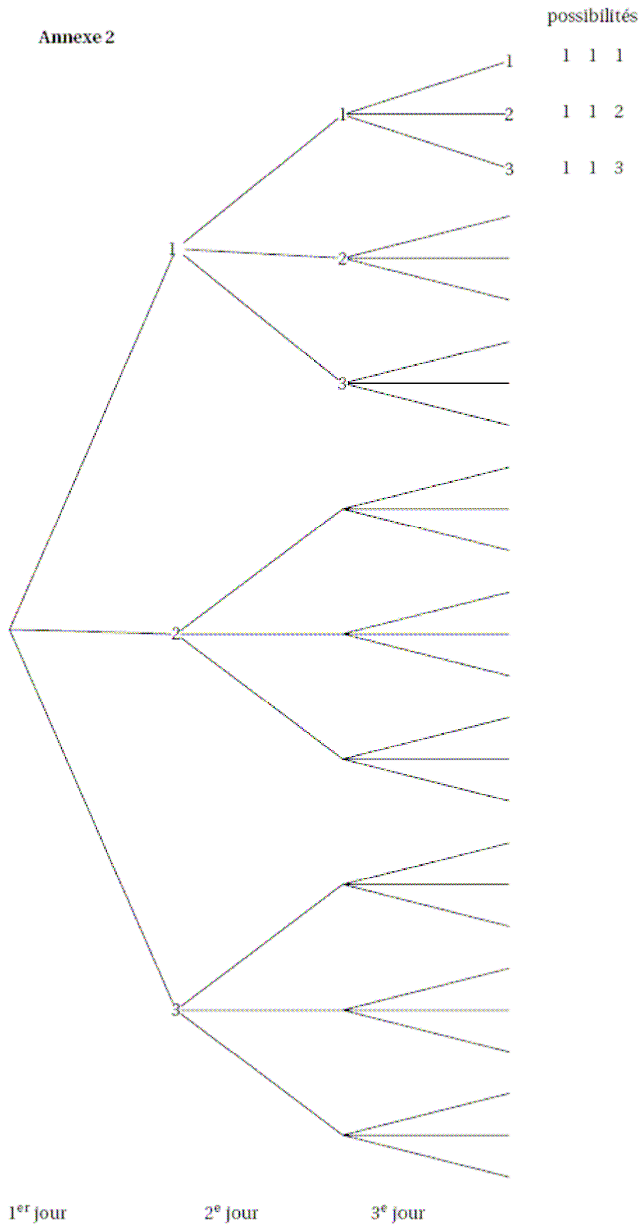
Remarque 1 : Dans au moins 50% des situations simulées la collection complète est obtenue au plus tard le ...... jour.

Remarque 2 : Dans . . . . . . % des situations simulées, la collection complète des images est obtenue au plus tard le 7^{e} jour.

Recopiez ces remarques sur votre copie et complétez-les.

3. Aline affirme : « Au bout de 18 jours, je suis sûre d'obtenir la collection complète ». Que pensez-vous de cette affirmation ?

Annexe 2



36. Amérique du Nord, juin 2005, 12 points

Les jeux olympiques de 2004 se sont déroulés en Grèce, à Athènes. Le tableau donné en annexe a été obtenu à l'aide d'un tableur. Il indique, pour chacun des 75 pays participants et dans l'ordre des colonnes de A à J :

- le rang du pays dans le classement officiel (classement effectué suivant le nombre de médailles d'or gagnées puis, en cas d'égalité, suivant le nombre de médailles d'argent et, en cas de nouvelle égalité, suivant le nombre de médailles de bronze),
- le nom du pays,
- les nombres de médailles d'or, d'argent et de bronze obtenues par pays,
- le nombre total de médailles par pays,
- le pourcentage de médailles gagnées par pays, arrondi à 10^{-1} près,
- la population du pays en millions d'habitants, arrondie à 10^{-1} près,
- le nombre total de médailles d'or par million d'habitants du pays, arrondi à 10^{-2} près,
- le rang du pays dans le classement du nombre de médailles d'or par million d'habitants.

Partie A

Pour le tableau de l'annexe 1, on utilise la convention suivante : la notation F8, par exemple, est l'adresse de la cellule située à l'intersection de la colonne F et de la ligne 8. Le contenu de trois cellules a été volontairement masqué. Le format des cellules contenant des valeurs numériques de ce tableau est celui des nombres lus.

1. Parmi les cellules B79, E13, F33, H8 et J1 de ce tableau, indiquer celles qui contiennent :

a. du texte,

b. une variable,

c. une formule.

2. On a obtenu les résultats de la colonne F du tableau à l'aide d'une formule saisie dans la cellule F2.

a. Quelle est cette formule ?

b. Quelle est l'action la plus rapide donnant les formules permettant d'obtenir les résultats lus dans les cellules F3 à F76 ?

c. Donner alors la formule contenue dans la cellule F8.

3. a. Indiquer la formule contenue dans la cellule I8.

b. Quel est le résultat affiché dans la cellule I8 ?

4. a. Donner la formule contenue dans la cellule C79 qui permet de calculer le nombre total de médailles d'or gagnées.

b. Donner la formule contenue dans la cellule C80 qui permet de calculer la moyenne du nombre de médailles d'or gagnées, par pays.

c. Donner la formule contenue dans la cellule C81 qui permet de calculer le plus grand nombre de médailles d'or gagnées.

5. a. Calculer le pourcentage de médailles gagnées par les Etats-Unis par rapport au nombre total de médailles distribuées (arrondi à 10^{-1} près).

b. Indiquer la formule permettant d'obtenir le résultat de la cellule G2, sachant que cette formule doit utiliser la cellule F79 et doit être recopiable pour obtenir les résultats des cellules G3 à G76.

6. On compare les classements des colonnes A et J du tableau. Donner deux exemples illustrant l'intérêt de réaliser ces deux classements.

Partie B

1. Dans cette question, on s'intéresse à la série concernant les médailles de bronze. Déterminer, en expliquant votre démarche, la médiane de cette série.

Pour cela, utiliser et compléter si besoin le tableau de l'annexe 2 qui sera à joindre à la copie.

2. Maintenant, on s'intéresse aux deux séries concernant les médailles d'or et les médailles d'argent. On a obtenu les résultats suivants :

Indicateur / médailles	moyenne	écart-type	1^{er} quartile	médiane	3^{e} quartile

D'or	4,0	6,75	1	2	4
D'argent	4,0	6,31	0	2	5

Utiliser ce tableau pour répondre aux questions suivantes :

a. La dispersion par rapport à la moyenne du nombre de médailles d'or gagnées est-elle supérieure à celle des médailles d'argent ? Justifier.

b. Interpréter sous forme d'une phrase le fait que le 1er quartile de la série des médailles d'argent soit nul.

c. Que peut-on dire des phrases suivantes ? Justifier.

Phrase 1 : au moins 25 % des pays participant aux jeux olympiques n'ont pas gagné de médaille d'argent.

Phrase 2 : plus de 50 % des pays participant aux jeux olympiques ont gagné au moins trois médailles d'or.

ANNEXE

	A	B	C	D	E	F	G	H	I	J
1	Rang officiel	Pays	Or	Argent	Bronze	Total	Pourcentage de médailles gagnées	Population (en millions) d'habitants	Médailles d'or par million d'habitants	Rang par rapport au nombre de médailles d'or par million d'habitants
2	1	Etats Unis	35	39	29	103		291,5	0,12	33
3	2	Chine	32	17	14	63	6,8%	1288,7	0,03	52
4	3	Russie	27	27	38	92	9,9%	145,5	0,19	24
5	4	Australie	17	6	16	49	5,3%	19,9	0,85	3
6	5	Japon	16	9	12	37	4,0%	127,5	0,13	31
7	6	Allemagne	14	16	18	48	5,2%	82,6	0,17	28
8	7	France	11	9	13	33	3,6%	6,6		57
9	8	Italie	10	11	11	32	3,4%	57,2	0,18	25
10	9	Corée du Sud	9	12	9	30	3,2%	47,9	0,19	22
11	10	Grande-Bretagne	9	9	12	30	3,2%	59,2	0,15	27
12	11	Cuba	9	7	11	27	2,9%	11,3	0,80	4
13	12	Ukraine	9	5	9	23	2,5%	47,8	0,19	22
14	13	Hongrie	8	6	3	7	1,8%	10,1	0,79	5
15	14	Roumanie	8	5	6	19	2,0%	21,6	0,37	11
16	15	Grèce	6	6	4	16	1,7%	11,0	0,55	8
17	16	Norvège	5	0	1	6	0,6%	4,6	1,09	2
18	17	Pays-Bas	4	9	9	22	2,4%	16,2	0,25	18
19	18	Brésil	4	3	3	10	1,1%	176,5	0,02	53
20	19	Suède	4	1	2	7	0,8%	9,0	0,44	9
21	20	Espagne	3	11	5	19	2,0%	41,3	0,07	42
22	21	Canada	3	6	3	12	1,3%	31,6	0,10	37
23	22	Turquie	3	3	4	10	1,1%	71,2	0,04	48
24	23	Pologne	3	2	5	10	1,1%	38,6	0,08	40
25	24	Nouvelle-Zélande	3	2	0	5	0,5%	4,0	0,75	7
26	25	Thaïlande	3	1	4	8	0,9%	63,1	0,05	47
27	26	Belarus	2	6	7	15	1,6%	9,9	0,20	21
28	27	Autriche	2	4	1	7	0,8%	8,2	0,24	19
29	28	Ethiopie	2	3	2	7	0,8%	70,7	0,03	51
30	29	Slovaquie	2	2	2	6	0 ,6%	5,4	0,37	11
31	29	Iran	2	2	2	6	0 ,6%	66,6	0,03	50
32	31	Taïwan	2	2	1	5	0,5%	22,6	0,09	38
33	32	Géorgie	2	2	0	4	0,4%	4,7	0,43	10
34	33	Bulgarie	2	1	9	12	1,3%	7,5	0,27	15
35	34	Ouzbékistan	2	1	2	5	0,5%	25,7	0,08	40

36	34	Jamaïque	2	1	2	5	0,5%	2,6	0,77	6
37	36	Maroc	2	1	0	3	0,3%	30,4	0,07	41
38	37	Danemark	2	0	6	8	0,8%	5,4	0,37	11
39	38	Argentine	2	0	4	6	0,6%	36,9	0,05	46
40	39	Chili	2	0	1	3	0,3%	15,8	0,13	30
41	40	Kazakhstan	1	4	3	8	0,9%	14,8	0,07	43
42	41	Kenya	1	4	2	7	0,8%	31,6	0,03	49
43	42	République tchèque	1	3	4	8	0,9%	10,2	0,10	35
44	43	Afrique du Sud	1	3	2	6	0,6%	44,0	0,02	53
45	44	Croatie	1	2	2	5	0,5%	4,3	0,23	20
46	45	Lithuanie	1	2	0	5	0,3%	3,5	0,29	14
47	46	Suisse	1	1	3	5	0,5%	7,3	0,14	29
48	46	Egypte	1	1	3	5	0,5%	72,1	0,01	55
49	48	Indonésie	1	1	2	4	0,4%	220,5	0,01	56
50	49	Zimbabwe	1	1	1	3	0,3%	12,6	0,08	39
51	50	Azerbaïdjan	1	0	4	5	0,5%	8,2	0,12	32
52	51	Belgique	1	0	2	3	0,3%	10,4	0,10	36
53	52	Israël	1	0	1	2	0,2%	6,7	0,15	28
54	52	Bahamas	1	0	1	2	0,2%	0,3	3,23	1
55	54	Rép. Dominicaine	1	0	0	1	0,1%	8,7	0,12	34
56	54	Emirats	1	0	0	1	0,1%	3,9	0,26	16
57	54	Eire	1	0	0	1	0,1%	4,0	0,25	17
58	54	Cameroun	1	0	0	1	0,1%	15,7	0,06	45
59	58	Corée du Nord	0	4	1	5	0,5%	22,7	0,00	57
60	59	Lettonie	0	4	0	4	0,4%	2,3	0,00	57
61	60	Mexique	0	3	1	4	0,4%	104,9	0,00	57
62	61	Portugal	0	2	1	3	0,3%	10,4	0,00	57
63	62	Serbie-Monténégro	0	2	0	2	0,2%	10,7	0,00	57
64	62	Finlande	0	2	0	2	0,2%	5,2	0,00	57
65	64	Slovénie	0	1	3	4	0,4%	104,9	0,00	57
66	65	Estonie	0	1	2	3	0,3%	1,4	0,00	57
67	66	Paraguay	0	1	0	1	0,1%	6,2	0,00	57
68	66	Inde	0	1	0	1	0,1%	1068,6	0,00	57
69	66	Hong-Kong	0	1	0	1	0,1%	6,8	0,00	57
70	69	Vénézuela	0	0	2	2	0,1%	25,7	0,00	57
71	69	Nigéria	0	0	2	2	0,1%	133,9	0,00	57
72	71	Trinidad-et-Tobago	0	0	1	1	0,1%	1,3	0,00	57
73	71	Syrie	0	0	1	1	0,1%	17,5	0,00	57
74	71	Mongolie	0	0	1	1	0,1%	2,5	0,00	57
75	71	Erythrée	0	0	1	1	0,1%	4,4	0,00	57
76	71	Colombie	0	0	1	1	0,1%	44,2	0,00	57
77										
78										
79		Total des Médailles gagnées	301	301	327	929				
80		Nombre moyen de médailles gagnées	4,0	4,0		12,4				
81		Plus grand Nombre de médailles gagnées	35	39	38	103				
82		(selon les résultats officiels des J.O et l'institut national d'études démographiques)								

Partie B, question 1.

Dans le tableau suivant, les données de la deuxième ligne indiquent que 14 pays n'ont obtenu aucune médaille de bronze.

Nombre de médailles de bronze	Effectif	
0	14	
1	15	
2	14	
3	7	
4	6	
5	2	
6	2	
7	1	
9	4	
11	2	
12	2	
13	1	
14	1	
16	1	
18	1	
29	1	
38	1	

37. Antilles, juin 2005, 8 points

Le tableau 1 donné dans l'annexe donne la répartition, selon le prix, de 1 000 billets « séjour » vendus par une agence de voyage pour la période de mai à septembre.

L'annexe est à rendre avec la copie.

Les pourcentages demandés seront arrondis au dixième. Les autres résultats seront arrondis au centième.

1. Calculer, en pourcentage, les fréquences de la répartition, selon le prix, de ces 1 000 billets, puis calculer les fréquences cumulées croissantes. Compléter le tableau 1.

2. Construire dans le repère de l'annexe 1 le diagramme en bâtons des fréquences de la série des prix de l'ensemble de ces séjours.

3. a. Déterminer la médiane ainsi que les premier et troisième quartiles de la série des prix des billets.

b. Construire sur l'axe donné en annexe le diagramme en boîte de cette série de prix.

c. En utilisant ce diagramme en boîte, recopier et compléter la phrase suivante :

« Au moins 75 % des billets ont un prix inférieur ou égal à . . . . . . ».

4. Calculer la moyenne x des prix des billets.

5. L'agence réalise un bénéfice sur chaque billet qui s'élève à 12% du prix de vente du billet. Quel est le bénéfice moyen par billet ?

Annexe (à rendre avec la copie)

Tableau 1

Destination du séjour	Malte	Baléares	Corse	Tunisie	Turquie	Grèce	Crète	République dominicaine	Égypte
Prix du séjour en euros	550	600	650	700	750	775	800	875	900

Nombre de billets « séjour » vendus	30	75	150	130	210	175	150	60	20
Fréquencesen%	3								
Fréquences cumulées croissantes en %	3								100

Diagramme en bâtons

Diagramme en boîte

38. France, juin 2005, 10 points

Un site de vente aux enchères sur Internet désire réaliser une étude statistique de sa clientèle.

Les responsables de l'étude utilisent un échantillon de 3000 clients, parmi les plus réguliers du site.

Partie A

La première question concerne l'âge des clients considérés. Les résultats sont donnés par l'histogramme ci-dessous.

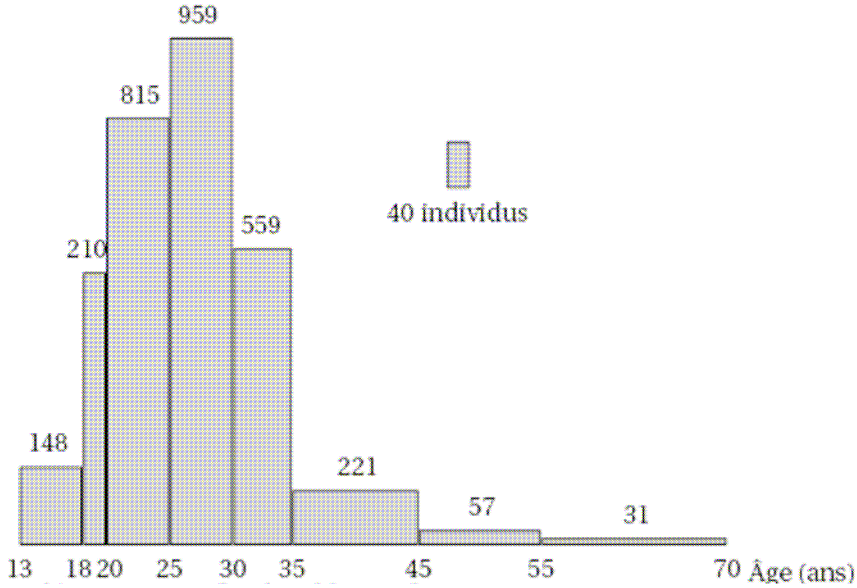
1. Compléter, sans justifier, le **tableau 3** figurant en annexe.

2. À l'aide de la calculatrice, déterminer sans justifier (on arrondira les résultats au dixième) :

a. L'âge moyen m des 3000 clients du site de vente aux enchères.

b. L'écart type σ de la série des âges des clients.

3. Peut-on estimer que le pourcentage des individus qui ont un âge appartenant à la plage $[m-\sigma\ ;\ m+\sigma]$ est supérieur ou égal à 75 % ?



Partie B

La seconde question posée aux 3000 clients porte sur la durée moyenne de connexion en minute durant une période d'une semaine.

1. L'étude a montré que la série des durées moyennes de connexion suit une loi de Gauss de moyenne $\mu \approx 83,5$ et d'écart type $s \approx 26,6$.

a. Déterminer la plage de normalité à 95% de cette série.

b. À combien peut-on estimer le nombre de clients dont la durée moyenne de connexion par semaine est située en dehors de cette plage ?

2. Pour cette série, le premier quartile Q_1 est 65, la médiane Me est 85 et le troisième quartile Q_3 est 100.

a. Quel est le nombre minimum de clients dont la durée moyenne de connexion par semaine sur le site est inférieure ou égale à 65 minutes ?

b. Les responsables du site espéraient qu'au moins 1000 personnes se connecteraient en moyenne 1 heure et 40 minutes ou plus par semaine. Cet objectif est-il atteint ?

Annexe

Tableau 3

classe	centre de la classe	effectif	fréquence (en %)
[13 ; 18[			
[18 ; 20[			
[20 ; 25[	22,5		
[25 ; 30[	27,5		32
[30 ; 35[	32,5		18,6
[5 ; 45[	40		7,4
[45 ; 55[			1,9
[55 ; 70[			1
Total	×	3000	100

39. La Réunion, juin 2005, 10 points

Les trois parties sont indépendantes

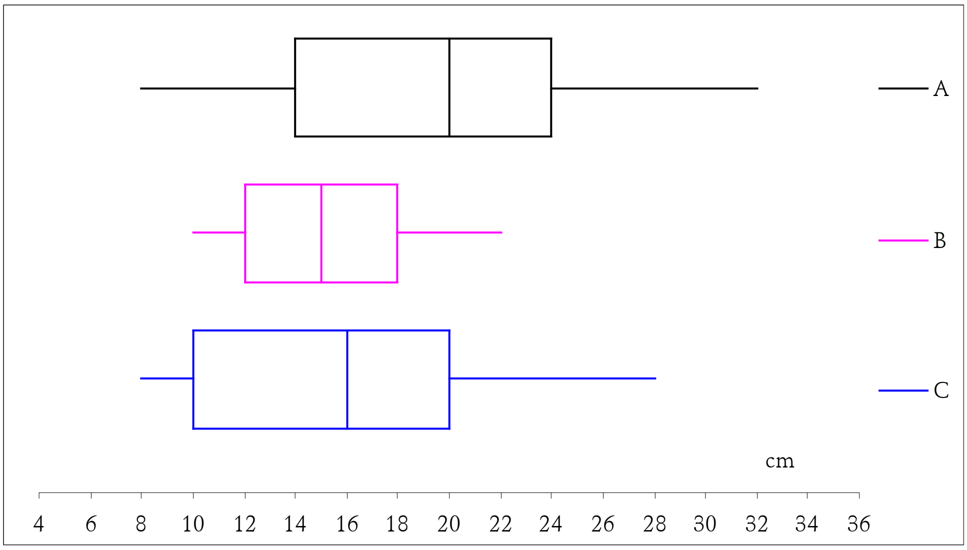
Partie A

Une fabrique de boules de pétanque conçoit des boules de compétition de différentes masses et de différents diamètres. Les trois masses proposées sont 700 g, 720 g et 745 g et pour chacune de ces masses trois diamètres sont proposés : 71 mm, 75 mm, 79 mm.

1. Combien y a-t-il de types de boules fabriquées dans cette entreprise ?

2. Un champion régional décide d'acheter des boules de 720 g, mais il hésite sur le diamètre. Pour faire son choix, il place un cochonnet à 9 mètres, pointe 200 fois avec chacune des boules de différents diamètres et mesure la distance au cochonnet. Voici les diagrammes en boîte élagués aux déciles représentant ce test.

Les extrémités du diagramme sont respectivement le premier et le neuvième décile.



Voici quelques sensations du joueur après le test :

« Avec la boule de 79 mm j'ai réussi de très bons lancers mais également de très mauvais. Avec la boule de 71 mm, j'ai eu de très bonnes sensations, la moitié de mes lancers était à moins de 16 cm du cochonnet et j'en ai réussi de très beaux. Mais ma préférence va à la boule de 75 mm avec laquelle je suis plus régulier. »

Associez à chaque type de boule le diagramme en boîte correspondant. Justifiez votre réponse.

Partie B : Vente annuelle en 2004

Les tableaux 1 et 2, extraits d'une feuille automatisée de calcul, donnent pour l'année 2004 la répartition des ventes de boules de pétanque suivant le diamètre et la masse.

	A	B	C	D	E
1	Tableau 1 : effectif				
2	Diamètre (mm) Masse (en g)	71	75	79	Total
3	700	2157	3123	1803	7083
4	720	3003	4122	2310	9435
5	745	2124	2982	1923	7029
6	Total	7284	10227	6036	23547
7	Tableau 2 : En pourcentage par rapport à l'effectif total				
8	Diamètre (mm) Masse (en g)	71	75	79	Total
9	700				

10	720				
11	745				
12	Total				100 %

1. Calculez le pourcentage, arrondi à 0,1%, de boules de 720 grammes et de diamètre 75mm vendues en 2004.

2. Calculez le pourcentage, arrondi à 0,1%, de boules de 700 grammes et de diamètre inférieur ou égal à 75 mm vendues en 2004.

3. Parmi les boules de 700 g vendues en 2004, quel est le pourcentage, arrondi à 0,1%, de boules de diamètre 79 mm ?

4. Le tableau des pourcentages est au format pourcentage. On propose ci-dessous des formules à écrire dans la cellule B10 et à recopier dans le reste du tableau. Citez la (ou les bonne(s) formule(s).

A : =B3/E6   B : =B3/E6   C : = B$3/E6   D : B3/23547

Il n'est pas demandé de compléter le tableau.

Partie C : Boules rejetées

Le tableau ci-dessous donne la répartition desmasses réelles de 2 500 boules de compétition de 720 grammes fabriquées :

Masse (En g)	719,5	719,6	719,7	719,8	719,9	720	720,1	720,2	720,3	720,4	720,5
Effectif	10	53	178	385	441	524	478	201	165	51	14

1. Calculez la moyenne μ des masses de ces 2 500 boules. Vous arrondirez à 3 chiffres après la virgule.

2. On suppose que les données suivent une répartition gaussienne. L'écart-type s de cette série est égal à 0,185. Déterminez la plage de normalité à 95% de cette série.

3. L'entreprise rejette en tant que boules de compétition les boules qui ne sont pas dans la plage de normalité. Calculez le pourcentage, arrondi à 0,1%, de boules rejetées.

40. Liban, juin 2005, 8 points

1 000 élèves de différents lycées ont mesuré la masse volumique du laiton par la méthode du flacon. Les résultats arrondis au dixième ont été regroupés dans le tableau suivant :

Masse volumique en g/cm^3	8	8,1	8,2	8,3	8,4	8,5	8,6	8,7	8,8	8,9	9	9,1
Effectif	3	19	42	100	200	250	190	113	50	20	7	6

1. Tracer le diagramme en bâtons de cette série (unités graphiques : 1 cm pour 0,1 g/cm^3 en abscisse en graduant à partir de 7,9 g/cm^3 et 1 cm pour 20 élèves en ordonnée).

2. a. Déterminer, en précisant votre méthode, le premier quartile q_1, la médiane$m$ et le troisième quartile q_3 de cette série.

b. Tracer le diagramme en boîte de cette série en y faisant figurer q_1, m, q_3 et les valeurs extrêmes de la série (unité : 1 cm pour 0,1 g/cm^3).

c. On note l la longueur de l'intervalle interquartile. Calculer le pourcentage des élèves ayant mesuré une masse volumique comprise dans l'intervalle $[m-l\ ;\ m+l]$.

3. a. Déterminer la valeur exacte de la moyenne μ de cette série.

b. Déterminer la valeur approchée à 10^{-3} par défaut de l'écart type σ de cette série.

c. Calculer le pourcentage des élèves ayant mesuré une masse volumique comprise dans l'intervalle $[\mu-2\sigma\ ;\ \mu+2\sigma]$, puis dans l'intervalle $[\mu-3\sigma\ ;\ \mu+3\sigma]$.

41. Polynésie, juin 2005, 8 points

La série suivante donne le nombre de jours de neige par année, à Paris, de 1900 à 1948.

Les 49 valeurs de cette série ne sont pas classées par ordre chronologique mais par ordre croissant.

1– 5 – 6 – 6 – 6 – 7 – 7– 7 – 8 – 8

9 –10 –10 –11 –11 – 11 – 12 – 12 – 12 – 12

13 – 13 – 13 – 14 – 14 – 14 – 14 – 15 – 16 – 17

17 – 17 – 18 – 18 – 18 – 18 – 19 – 19 – 20 – 20

20 – 23 – 26 – 29 – 29 – 31 – 32 – 32 – 34

1. a. Calculer le nombre moyen x de jours de neige par année, à Paris, sur la période 1900-1948 (le résultat sera arrondi au dixième).

b. Déterminer la médiane, *med*, ainsi que le premier et le troisième quartile, Q_1 et Q_3, de cette série. Justifier chaque réponse.

2. Les nombres de jours de neige par an, à Paris, ont également été relevés de 1949 à 1997. On fournit ci-dessous les caractéristiques de cette nouvelle série statistique.

Minimum	Premier quartile Q_1	Médiane *med*	Troisième quartile Q_3	Maximum	Moyenne
1	7	12	18	36	13,3

a. Donner l'écart interquartile de chacune des deux séries statistiques donnant le nombre de jours de neige par an, à Paris, sur la période 1900-1948, puis sur la période 1949-1997.

b. Construire sur l'annexe le diagramme en boîtes de chacune de deux séries étudiées.

c. Comparer ces deux diagrammes en boîtes.

3. La série des nombres de jours de neige par an, à Paris, sur la période 1900-1948, a pour écart-type s et celle des nombres de jours de neige par an, à Paris, sur la période 1949-1997 a pour écart-type s'. On admet que $s \approx 7{,}82$ et que $s' \approx 8{,}01$.

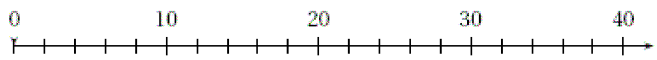
Que signifie le fait que l'écart-type soit plus élevé pour la deuxième période que pour la première ?

4. À propos de ces relevés météorologiques, divers commentaires ont été relevés dans la presse, dont celui-ci : « Des hivers demoins en moins neigeux au cours du siècle ».

Que peut-on penser de ce commentaire ?

ANNEXE
À rendre avec la copie

Période 1900–1948



Période 1949–1997