Probabilités et Statistiques

1.1. Calculs statistiques de base

1.1.1. Election

Les candidats à une élection ont obtenu les voix suivantes :

Candidat	A	B	C	D
Nombre de voix obtenues	51 210	43 821	23 212	8 597

Calculez les fréquences des voix. Faites un « camembert » permettant de visualiser ces résultats.

1.1.2. Prix de vente

On a relevé le prix de vente d'un CD et le nombre de CD vendus chez différents fournisseurs.

Les résultats sont donnés dans le tableau suivant :

Prix de vente en euros, P	12	14	16	18	19
Nombre de CD vendus	97	34	43	20	6

1. Dessinez un histogramme permettant de visualiser P.

2. Quel est le prix moyen d'un CD ?

3. Précisez l'écart-type de cette série (les calculs doivent apparaître).

1.2. Lecture de données, pourcentages

1.2.3. Orientation

Dans le lycée, durant l'année scolaire 2008-2009, l'effectif en seconde était de 450 élèves. Le bilan de l'orientation des élèves de seconde, fin juin 2009, est donné dans le tableau ci-dessous. 45 élèves sont orientés en 1ère STI.

Orientation	Redoublement	1ère L	1ère S	1ère ES	1ère STI	1ère STG	Autres
Pourcentage	10%	12%		28%		4%	6%
Effectifs					45		

1. Compléter le tableau avec le pourcentage des élèves de seconde orientés en 1ère S et en 1ère STI ainsi que les effectifs pour les différentes orientations.

2. Sachant que 80 % des élèves orientés en 1ère STG sont des filles, quel est le nombre de garçons orientés en première STI ?

3. Dans les statistiques du Ministère de l'Education Nationale on a le nombre de reçus au bac (ici 92 % des élèves de Terminale) et le pourcentage d'élèves de Seconde qui iront en Terminale (ici 78 %). Combien parmi les 450 élèves de l'année 2009 auront-ils le bac en 2011 ?

1.2.4. Tri sélectif

Les parties A et B sont indépendantes.

Un organisme collecte et traite les déchets ménagers pour un groupe de communes :

- Un ramassage apporte des déchets qui seront directement enfouis.
- Un autre ramassage apporte des déchets triés par les usagers. Après traitement ces déchets seront recyclés.

Partie A : Étude des quantités traitées en 2007.

En 2007, la quantité totale des déchets recyclés est 1 550 tonnes et la quantité totale de déchets enfouis est 14 500 tonnes.

La population totale concernée par la collecte est de 60 000 habitants.

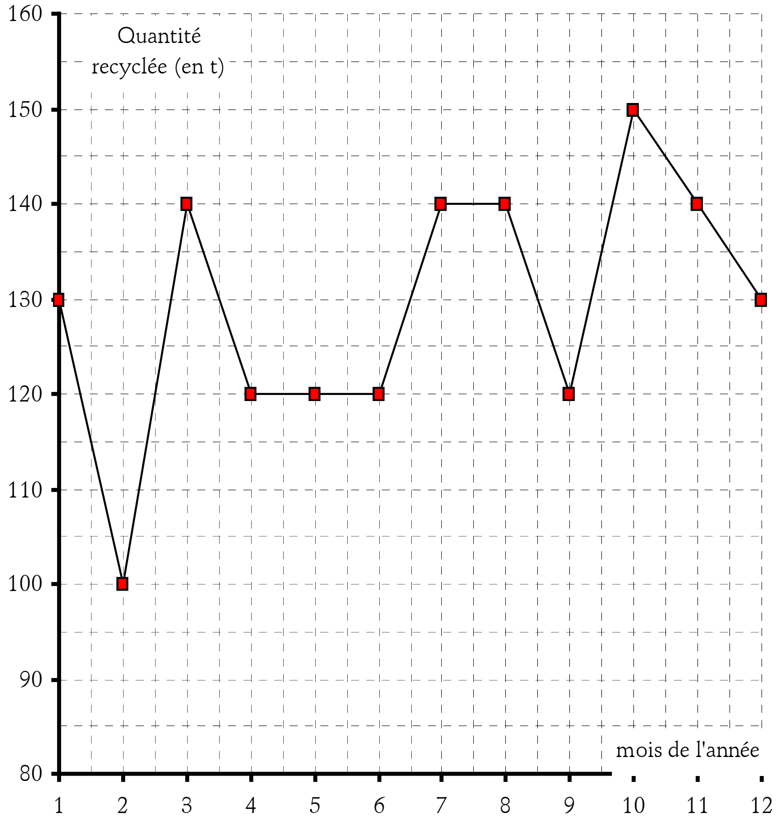
1. Calculer la quantité totale de déchets produite par cette population en 2007 puis la quantité moyenne produite par habitant.

2. Quel pourcentage de la quantité totale représente la quantité recyclée en 2007 ? (arrondir au dixième).

3. La représentation graphique donnée ci-dessous présente, pour 2007, la quantité recyclée (en tonnes) en fonction des mois de l'année. Après l'avoir étudiée, répondre aux questions suivantes :

a. Quel est le mois où la quantité recyclée a été la plus grande ? Quel est le mois où elle a été la plus faible ?

b. Deux campagnes de sensibilisation ont été organisées auprès de la population concernée afin de l'inciter à trier davantage. L'une a été effectuée durant lemois de février et l'autre durant le mois de septembre : ces campagnes de sensibilisation ont-elles eu un impact ? Argumenter.



Partie B : Étude statistique concernant le tri effectué par les habitants.

Afin d'améliorer l'information des usagers sur la façon dont ils doivent trier les déchets ménagers, on réalise un sondage auprès d'un échantillon de 2 000 personnes. À l'aide des questions posées aux usagers, on détermine s'ils savent trier ou non le contenu de leurs poubelles. Voici le résultat de l'étude sous la forme d'un tableau. Tous les pourcentages sont exprimés par rapport à l'effectif total des 2 000 personnes interrogées.

Âge	savent trier	ne savent pas trier
de 18 à 24 ans	8%	7%
de 25 à 34 ans	12%	10%
de 35 à 44 ans	16%	13%
de 45 à 59 ans	8%	5%
de 60 à 69 ans	5%	5%
70 ans et plus	5%	6%

1. Parmi les personnes interrogées, quel est le pourcentage de celles qui ont entre 35 et 44 ans et qui savent trier ?
2. Parmi les personnes interrogées, quel est le pourcentage de celles qui savent trier ?
3. Parmi les personnes interrogées, quel est le pourcentage de celles qui ont entre 60 et 69 ans ?
4. Calculer le nombre de personnes interrogées qui ne savent pas trier.
5. Parmi les personnes interrogées qui ont entre 35 et 44 ans, quel est le pourcentage (à 1 % près) de celles qui savent trier ?

1.3. Moyenne, médiane, boites à moustaches

1.3.5. Etude du temps de déchargement d'une plate-forme

Une société de transport réalise sur une de ses plateformes un contrôle des durées de déchargement de ses camions.

Les durées relevées sont données dans le tableau suivant :

Durées de	Nombre de

déchargement (minutes)	camions
25	2
26	5
27	12
28	36
29	45
30	32
31	8
32	3
Total	143

1. Calculer la moyenne des durées de déchargement en utilisant une méthode au choix. Arrondir le résultat à l'unité.

2. Calculer le pourcentage de camions dont la durée de déchargement est inferieure ou égale à 30 minutes. Arrondir le résultat à 1'unité.

3. La plateforme est performante si les deux conditions suivantes sont respectées :

- la durée moyenne de déchargement doit appartenir à [28 ; 30],
- la durée de déchargement doit être inferieure ou égale à 30 minutes pour plus de 95 % des camions.

Ecrire une phrase justifiant si cette plateforme est performante ou non.

1.3.6. Véhicules automobiles

Le tableau ci-dessous indique la puissance x en chevaux DIN et la cylindrée y en cm^3 de 8 voitures à moteur diesel.

Voiture		A	B	C	D	E	F	G	H
Puissance x		37	55	60	60	65	70	72	76
Cylindrée y		993	1579	1761	1697	1935	1986	1997	2498

Dans le repère orthogonal suivant, on a placé les points M_1 (37 ; 993), M_2 (55 ; 1579),M_3 (60 ; 1761), M_4 (60 ; 1697), M_5 (64 ; 1935), M_6 (70 ; 1986), M_7 (72 ; 1997) et M_8 (76 ; 2498) associés respectivement aux voitures V_1, V_2, V_3, V_4, V_5, V_6, V_7 et V_8.

1. On se propose de tracer une droite d'ajustement de ce nuage de points.

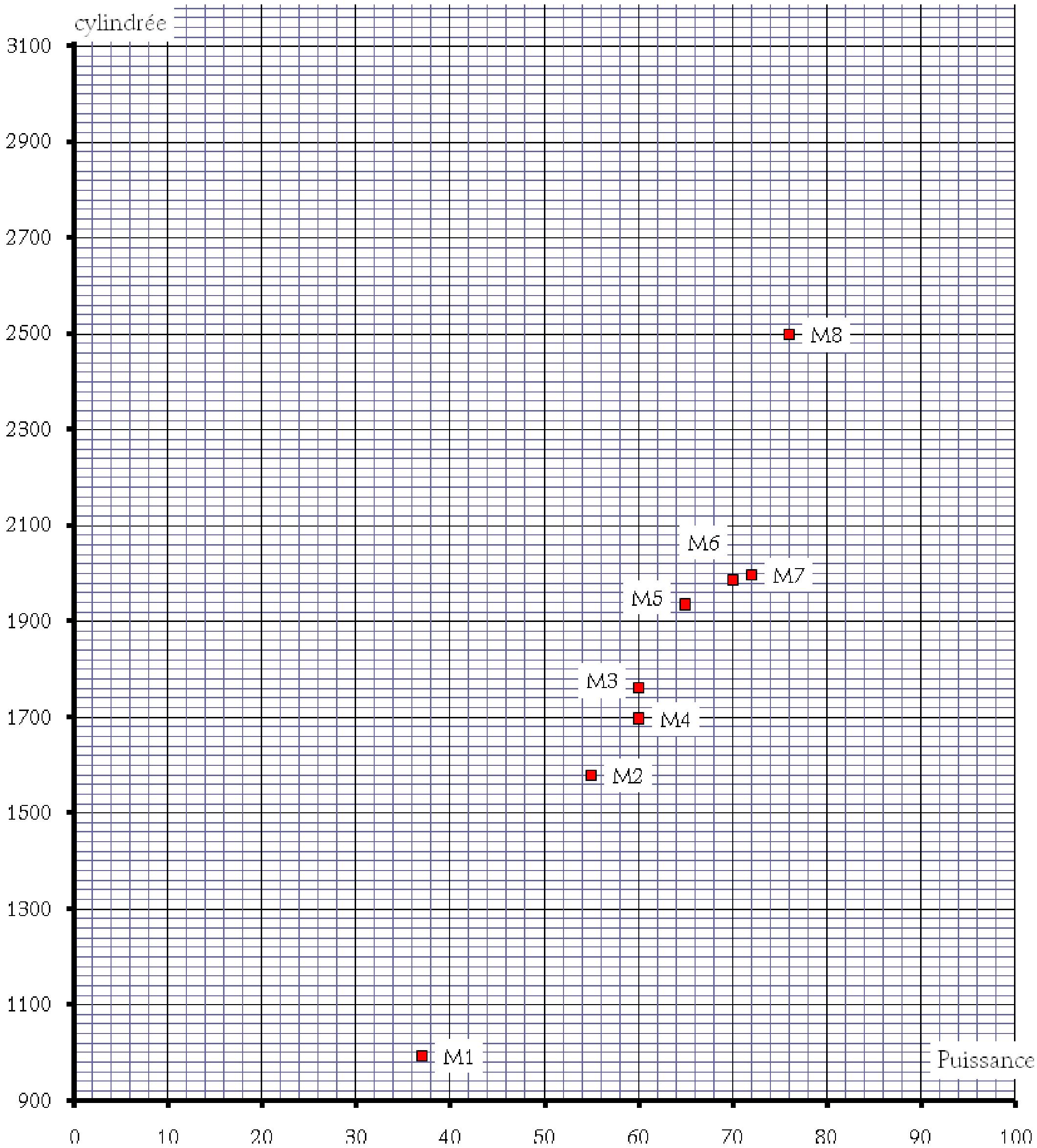
a. Calculer les coordonnées, arrondies à l'unité, du point G de ce nuage de points. On rappelle que son abscisse est la moyenne des abscisses des points, de même pour son ordonnée.

b. On choisit comme droite d'ajustement la droite (AG) où A et G ont pour coordonnées respectives (40 ; 1058) et (62 ; 1806). Tracer la droite d'ajustement (AG).

c. Déterminer une équation de la droite (AG).

2. a. En utilisant la droite d'ajustement (AG), déterminer graphiquement la puissance d'un moteur de cylindrée 2 800 cm^3 en laissant apparents les traits permettant la lecture graphique.

b. En utilisant une équation de la droite (AG), calculer la puissance, arrondie à l'unité, d'un moteur de cylindrée 2 800 cm^3.



1.3.7. Y'a d'la truffe

Un trufficulteur (agriculteur cultivant les truffes, *tuber melanosporum*) décide de tester l'influence de l'arrosage de ses truffières sur la masse des truffes récoltées. Il décide donc de répartir ses récoltes en deux lots de 100 truffes :

- le premier, appelé lot A, provient de truffières ne recevant aucun arrosage ;
- le second, appelé lot B, provient de truffières arrosées.

1. Au moment de la récolte il pèse ses truffes et obtient, pour le lot B, les résultats suivants :

Masse en grammes	15	15,5	16	16,5	17	17,5	18	18,5	19	19,5	20	20,5	21	21,5	22	Total
Nombre de truffes	16	4	20	14	22	4	B	3	2	1	2	0	1	0	3	100

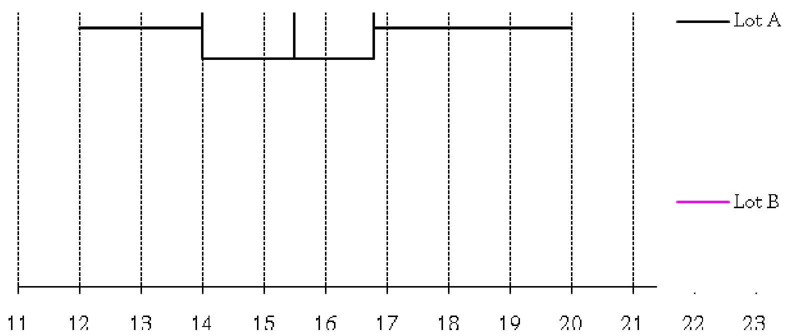
a. Déterminer, pour le lot B, le minimum, le premier quartile Q_1, la médiane M, le troisième quartile Q_3 et le maximum.

b. Construire le diagramme en boite correspondant au lot B.

2. On a représenté ci-dessous le diagramme en boîte correspondant au lot A. Déduire de ce graphique le minimum, le premier quartile Q'_1, la médiane M', le troisième quartile Q'_3 et le maximum du lot A.

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

Phrase 1 : environ la moitié du lot 13 est constitué de truffes d'un poids égal ou supérieur aux trois-quarts des truffes du lot A.
Phrase 2 : en arrosant, on réduit l'écart inter-quartile de masse entre les truffes récoltées.



D'après Bac L Maths-Info France 2008

1.3.8. Test biochimique

On mesure la quantité d'une certaine molécule M dans le sang de plusieurs groupes de personnes :

– un groupe A de 5 000 individus en bonne santé,

– un groupe de 100 individus souffrant d'une même maladie P répartis au hasard en deux groupes :

  – un groupe B de 50 individus qui ne reçoivent pas de traitement,

  – un groupe C de 50 individus qui reçoivent un traitement.

La quantité est mesurée en μg/L (microgramme par litre).

Partie A : Étude du groupe A

La série de données recueillies dans le groupe A (appelée série de référence) correspond à des données gaussiennes. La plage de normalité à 95 % obtenue pour cette série de référence est l'intervalle [120 ; 160].

1. Pour approximativement quel nombre d'individus du groupe A le dosage a-t-il été dans la plage de normalité ?

2. Quelle est la moyenne de la série de référence ?

3. Quel est l'écart-type de la série de référence ?

Partie B : Étude du groupe B

Le tableau ci-dessous donne les résultats obtenus pour le groupe B

Quantité(μg/L)	130	135	140	145	150	155	160	165	170	175	180	185	190
Effectifs	2	3	3	5	3	4	3	7	5	6	3	2	4
Effectifs cumulés croissants													

1. Pour quel pourcentage des individus du groupe B la quantité mesurée est-elle dans la plage de normalité [120 ; 160] ?

2. Calculer les effectifs cumulés croissants de cette série. Compléter le tableau.

3. Déterminer la médiane et les quartiles de cette série statistique.

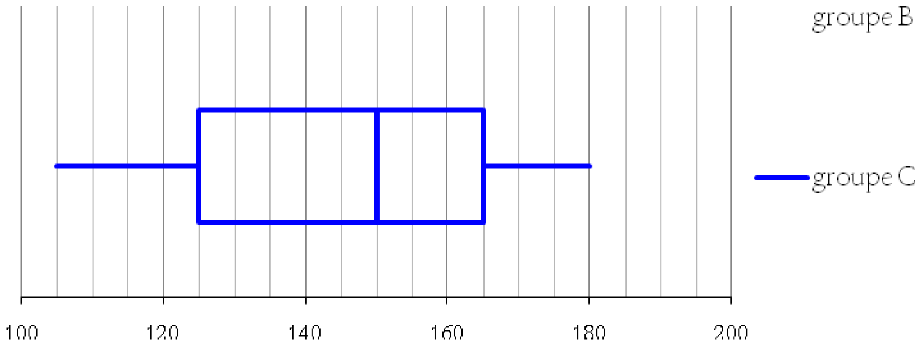
4. Tracer le diagramme en boîtes de cette série ci-dessous. Prendre pour extrémités le minimum et le maximum de la série.

Partie C : Étude du groupe C

Les données recueillies pour le groupe C ont été résumées dans le diagramme en boîtes.

1. Pour approximativement quel pourcentage des individus du groupe C la quantité mesurée est-elle dans la plage de normalité [120 ; 160] ?

2. Quel semble être l'effet du traitement sur les individus du groupe C par comparaison avec ceux du groupe B ?



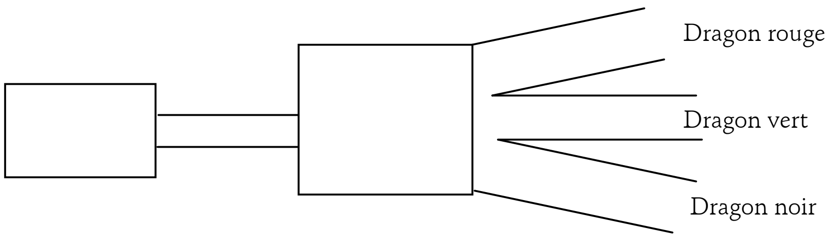
D'après Bac L Maths-Info Nouvelle Calédonie 2008

1.3.9. Dungeons & Dragons

Les trois parties sont indépendantes

Partie A

Kévin joue à un jeu vidéo comportant deux étapes. L'étape 1 consiste à choisir dans un coffre une arme parmi trois proposées, puis à choisir un des trois chemins qui aboutissent chacun à un dragon de force différente.



Au départ, le joueur a 100 points. En passant devant le coffre, il doit acheter une arme : il peut acheter une épée pour 80 points, un pistolet pour 50 points ou un arc pour 30 points (sachant que plus l'arme coûte cher, plus elle est efficace).

Le joueur doit alors choisir un des trois chemins pour affronter un dragon à l'étape 1, et pouvoir accéder, après la victoire, à l'étape 2. Vaincre le dragon vert rapporte 150 points, le rouge 180 points et le noir 200 points ; ainsi, si un joueur achète une épée et gagne contre le dragon rouge, il dispose de 200 points à la fin de l'étape 1.

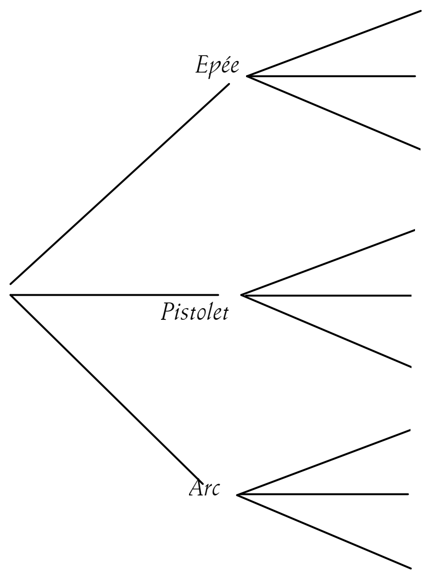
1. Compléter l'arbre en indiquant, pour chacun des neuf choix possibles, le nombre de points restant au joueur à la fin de l'étape 1, en supposant chaque combat victorieux.

2. Sachant qu'il faut au moins 200 points pour accéder à l'étape 2, préciser le seul chemin qui empêche le joueur d'y accéder.

3. En fait Kevin aimerait avoir au moins 220 points pour commencer l'étape 2, car au début de cette étape, il a la possibilité d'acheter une petite fiole qui augmente les pouvoirs du héros.

a. Combien de possibilités Kévin a-t-il ?

b. Kévin veut, de plus, éviter le dragon noir trop difficile à vaincre. Combien lui reste-t-il de choix possibles ? Préciser, pour chacun d'eux, l'arme utilisée et le dragon combattu.



Partie B

Les concepteurs du jeu l'ont fait tester par 80 joueurs confirmés. Les résultats sont donnés dans le tableau 1 ci-dessous.

On a obtenu, à l'aide d'un tableur, le tableau des pourcentages (tableau 2), le tableau des pourcentages en ligne (tableau 3) ainsi que le tableau des pourcentages en colonne (tableau 4). Les résultats ont été

arrondis à l'unité.

1. En utilisant les valeurs du tableau 1, écrire le calcul permettant de trouver le résultat 29 de la cellule B23.

2. À l'aide des tableaux répondre aux questions suivantes :

a. Sur l'ensemble des 80 joueurs, quel est le pourcentage de ceux qui ont combattu le dragon noir ?

b. Parmi les joueurs qui ont vaincu un dragon, quel est le pourcentage de ceux qui ont combattu le dragon vert ?

c. Parmi les joueurs ayant affronté le dragon rouge, quel est le pourcentage de ceux qui ont perdu le combat ?

3. Quelle formule peut-on saisir dans la cellule B10 pour obtenir, par recopie automatique vers le bas et vers la droite, les nombres du tableau 2 ?

4. Quelle formule peut-on saisir dans la cellule B17 pour obtenir, par recopie automatique vers le bas et vers la droite, les nombres du tableau 3 ?

Partie C

Pour savoir si la difficulté du combat contre le dragon noir est acceptable, les concepteurs du jeu ont demandé à 71 joueurs confirmés de combattre ce dragon pendant une heure.

On a relevé, en minutes (valeurs arrondies à la minute), le temps mis par les joueurs pour gagner le combat. Les joueurs ont droit à autant d'essais qu'ils le souhaitent et le choix de l'arme n'est pas imposé. Les résultats obtenus sont consignés dans le tableau 5.

1. Compléter la ligne des effectifs cumulés croissants (ECC) du tableau 5.

2. Pour que la difficulté du combat contre le dragon noir soit jugée acceptable, il faut qu'au moins 75 % des joueurs aient mis au maximum 35 minutes pour gagner le combat.

a. Tracer le diagramme en boîte (on prendra comme extrémités des moustaches les minimum et maximum de la série).

b. Dire, en expliquant, si la difficulté du combat contre le dragon noir est jugée acceptable.

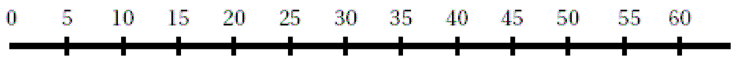
	A	B	C	D
1	tableau 1 : résultats obtenus			
2		combat perdu	combat gagné	total
3	dragon vert	8	30	38
4	dragon rouge	11	16	27
5	dragon noir	9	6	15
6	total	28	52	80
7				
8	tableau 2 : résultats en pourcentage			
9		combat perdu	combat gagné	total
10	dragon vert	10	38	48
11	dragon rouge	14	20	34
12	dragon noir	11	8	19
13	total	35	65	100
14				
15	tableau 3 : pourcentages en ligne			
16		combat perdu	combat gagné	total
17	dragon vert	21	79	100
18	dragon rouge	41	59	100
19	dragon noir	60	40	100
20				
21	tableau 4 : pourcentages en colonne			
22		combat perdu	combat gagné	
23	dragon vert	29	58	
24	dragon rouge	39	31	
25	dragon noir	32	12	
26	total	100	100	

Tableau 5

temps (en min)	5	7	9	11	14	15	17	20	21	22	23	25
effectif	1	2	1	3	4	1	5	4	7	8	6	5
ECC												
temps (en min)	27	30	33	34	37	41	46	49	52	53	60	
effectif	4	3	1	4	1	2	1	2	1	1	4	
ECC												

La colonne « 60 » correspond aux joueurs n'ayant pas réussi à vaincre le dragon noir.

Boîte à moustaches



D'après Bac L Maths-Info Amérique du Sud 2008

1.3.10. Téléchargement (illégal ?)

Le tableau (incomplet), fourni ci-dessous, donne la répartition d'une population de 800 utilisateurs d'internet pour le téléchargement selon leur âge et leur volume de téléchargement mensuel.

Le volume de téléchargement est exprimé en Giga-octets (notés Go) et l'âge en années.

Volume en Go Tranche d'âge	[0 ; 2[	[2 ; 4[	[4 ; 6[	[6 ; 8[	Total
[10 ; 20[	21	51	80	125	217
[20 ; 30[	17			107	223
[30 ; 40[	22	44	50	47	163
[40 ; 50[	30	20	20	12	
[50 ; 60[	42		2	8	
Total	132	158		299	800

Partie A

1. Compléter le tableau. Aucune justification n'est demandée.

2. Les pourcentages demandés dans cette question seront arrondis à l'unité.

a. Parmi ces utilisateurs d'internet, quel pourcentage est dans la tranche d'âge [30 ; 40[?

b. Parmi les utilisateurs d'internet qui téléchargent entre 0 et 2Go par mois, combien représentent, en pourcentage, ceux âgés de 40 ans ou plus ?

Partie B

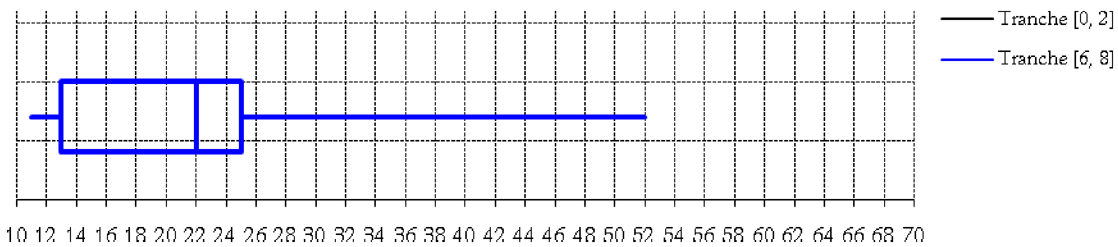
1. Dans la population observée, combien d'utilisateurs d'internet ont moins de 30 ans ? Expliquer alors pourquoi l'âge médian (la médiane) de cette population est nécessairement compris entre 20 et 30 ans..

2. Pour déterminer cet âge médian, on donne, la répartition des âges dans la classe [20 ; 30[. Elle est fournie par le tableau suivant :

Âge	20	21	22	23	24	25	26	27	28	29
Effectif	25	26	30	22	34	21	19	20	14	12

Justifier que l'âge median vaut 24 ans.

3. Les diagrammes en boîte des âges des utilisateurs d'internet qui téléchargent entre 0 et 2 Go et entre 6 et 8 Go sont représentés ci-dessous :



Les propositions suivantes sont-elles vraies ou fausses ? Justifier les réponses.

Proposition a : l'écart interquartile de la série des âges des utilisateurs qui téléchargent entre 0 et 2 Go est plus du double de la série des âges des utilisateurs qui téléchargent entre 0 et 2 Go.

Proposition b : plus de 75 % des utilisateurs qui téléchargent entre 0 et 2 Go ont plus de 26 ans.

Proposition c : plus de la moitié des utilisateurs qui téléchargent entre 6 et 8 Go sont mineurs.

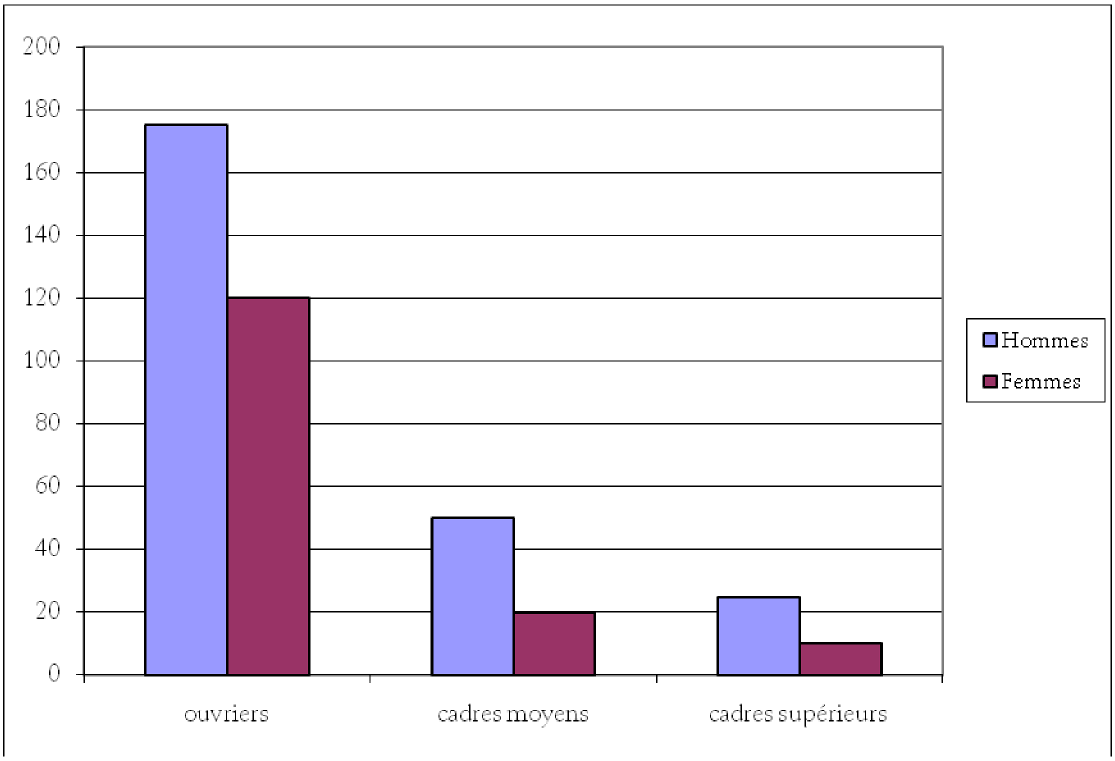
D'après Bac L Maths-Info Polynésie 2008

1.3.11. Smoking ? no Smoking ?

Les parties 1 et 2 sont indépendantes

Partie 1

Une entreprise emploie 400 employés. Le diagramme ci-dessous indique leur répartition selon leur sexe et leur qualification.



Compléter le tableau ci-dessous à partir des données du graphique.

Tableau : Répartition des employés de l'entreprise

	Ouvriers	Cadres moyens	Cadres supérieurs	Total
Hommes				
Femmes				
Total				400

Partie 2

Dans cette entreprise on a dénombré 59 femmes et 130 hommes fumeurs de cigarettes. L'entreprise souhaite proposer à ses employés plusieurs méthodes pour diminuer, voire supprimer, l'usage du tabac. Une enquête est menée parmi les fumeurs, femmes et hommes, pour déterminer la quantité approximative de cigarettes fumées sur une journée.

Elle permet de dresser les deux tableaux suivants : Pour les femmes fumeurs :

Nombre de cigarettes fumées par jour	5	10	15	20	25	30	35	40
Nombres de femmes	10	18	12	8	5	3	2	1

Pour les hommes fumeurs

Nombre de cigarettes fumées par jour	5	10	15	20	25	30	35	40
Nombres d'hommes	15	18	25	35	12	10	10	5

1. Le diagramme en boîte de la série du nombre de cigarettes fumées par jour par les femmes fumeurs est représenté ci-dessous.

Lire la médiane, le premier quartile et le troisième quartiîe de cette série.

2. Déterminer la médiane, le premier quartile et le troisième quartile de la série du nombre de cigarettes fumées par jour par les hommes fumeurs.

Représenter le diagramme en boîte de cette série au-dessus de celui des femmes fumeurs.

3. Calculer le nombre moyen de cigarettes fumées par jour par les femmes fumeurs puis par les hommes rumeurs (arrondir à l'unité).

4. Chacune des phrases suivantes est-elle vraie ou fausse ? Justifier votre réponse.

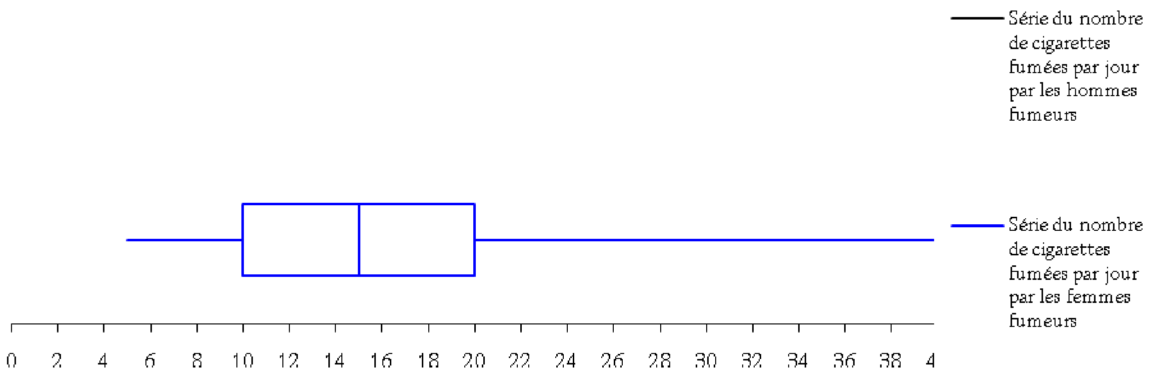
Dans cette entreprise :

a. Parmi les fumeurs, au moins la moitié des hommes fument au plus 20 cigarettes par jour.

b. Parmi les rumeurs, environ la moitié des femmes fument entre 10 et 20 cigarettes par jour.

c. Parmi les fumeurs, les femmes fument en moyenne plus que les hommes.

Diagrammes en boîte :



1.3.12. Densité médicale

Partie A

On considère le tableau suivant, disponible sur le site Internet de l'INSEE. Il donne les effectifs de médecins au 31 décembre pour 1990 et 2002 en France métropolitaine et les perspectives pour 2010, 2015 et 2025. (*Source : ministère de la Santé, de la Jeunesse et des Sports - Drees. Champ : France métropolitaine*)

Perspectives des effectifs de médecins					
	1990	2002	2010	2015	2025
Total	177 410	205 185	202 130	196 737	185 966
dont : médecine générale	93 387	100 541	100 514	99 665	97 119
spécialités médicales	48 033	57 127	56 330	54 453	50 595
spécialités chirurgicales	21 393	24 528	23 788	23 023	21 149
psychiatrie	11 897	13 727	12 291	11 008	8 816
biologie médicale	1 960	3 109	3 037	3 060	3 079
Santé publique et travail	800	6153	6 171	5 528	5 208

Dans cette partie, les pourcentages seront arrondis au dixième.

1. On s'intéresse à l'année 2002. Quel est le pourcentage de médecins en biologie médicale par rapport à la population totale de médecins (on justifiera le résultat) ?

2. On prévoit une augmentation du nombre de psychiatres entre 1990 et 2010. Exprimer cette augmentation en pourcentage.

3. On prévoit une diminution du nombre de médecins de médecine générale entre 2010 et 2025. Exprimer cette diminution en pourcentage.

4. La proportion de chirurgiens dans la population totale de médecins va-t-elle augmenter ou diminuer entre 2010 et 2025 ? Justifier la réponse.

Partie B

Le tableau suivant donne le nombre de médecins pour 100 000 habitants dans 15 pays européens. (Source : Eurostat. Les résultats non disponibles sont indiqués par : . . .)

	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
Allemagne	293	300	307	312	313	319	321	326	331	334	337	339
Autriche	314	302	303	290	277	265	254	313	324	328	340	347
Belgique	359	365	373	381	386	395	405	411	419	449	. . .	444
Danemark	245	248	251	253	259	266	267	269	272	281	285	. . .
Espagne	. . .	265	269	305	309	306	325	349	346	331	329	340
Finlande	264	270	277	286	296	300	306	308	311	313	. . .	. . .
France	313	316	318	320	322	324	325	324	326	329	333	335
Grèce	388	389	389	389	410	426	438	448	439	. . .	. . .	. . .
Irlande	202	199	210	211	214	219	227	223	240	242	259	277
Italie	551	559	566	571	578	583	589	599	603	611	628	636
Luxembourg	215	217	204	213	226	228	233	236	240	239	245	328
Pays-Bas	. . .	. . .	. . .	. . .	. . .	295	311	321	329	339	349	350
Portugal	246	252	255	263	262	259	262	265	264	274	269	. . .
Royaume-Uni	167	168	174	178	184	188	192	195	200	208	218	. . .
Suède	. . .	. . .	286	288	290	297	301	308	318	327	333	. . .

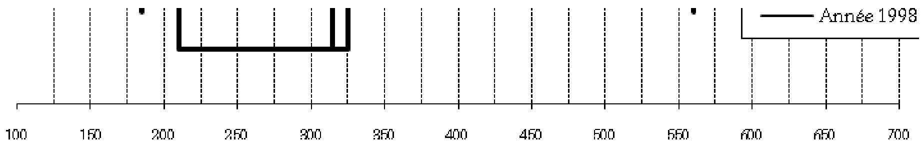
On s'intéresse aux valeurs de l'année 2000.

1. Déterminer la médiane M du nombre de médecins pour 100 000 habitants dans ces 15 pays.

2. Déterminer les quartiles de la série statistique étudiée.

3. Sur le graphique suivant, on a construit le diagramme en boîte correspondant aux données du tableau pour l'année 1998. Les extrémités du diagramme correspondent aux valeurs minimale et maximale de la série. Construire sur le même graphique le diagramme en boîte correspondant aux données du tableau pour l'année 2000.

4. Peut-on dire que les situations des différents pays d'Europe, au regard du nombre de médecins par habitant, se sont rapprochées entre 1998 et 2000 ? Justifier la réponse.



D'après Bac L Maths-Info La Réunion 2008

1.3.13. Mariages

Le tableau ci-dessous est une feuille automatisée de calcul.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1													
2	TABLEAU 1												
3		1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	Total	Fréquence en%
4	Janvier	5 556	6 182	6578	8 152	7 185	6 362	6749	7243	7 157	6 500	67 664	
5	Février	S 342	10 139	9721	9 159	9444	8 726	8 623	9 775	8 399	8 200	90 528	
6	Mars	9845	8 470	8 939	8 947	11 334	10 852	11 036	8 702	9 146	9700	96 971	3,5%
7	Avril	15 978	16 025	18 020	20 721	17430	15 012	16 172	18 013	17812	16 900	17 2083	6,2%
8	Mai	26 516	27 886	25 098	23 371	23 276	26 248	28 335	27 101	24 761	25000	257 592	9,2%
9	Juin	52 923	45 249	47 824	52100	61 737	54 575	50 986	47041	49 044	47 500	50 8979	18,2 %
10	Juillet	50633	47 532	57 541	58 932	42 536	42 763	43241	52 121	55 169	51 400	501 868	17,9 %
11	Août	45 028	42 188	38 847	38 936	39781	45 934	43614	35 079	36 675	34 700	400 782	14,3 %
12	Septembre	31 882	32 556	34887	40 191	40366	31 401	31 248	31 639	32 664	35900	342 734	12,2 %
13	Octobre	16 139	16 452	17 544	15 420	14441	15 811	15894	16 279	15 859	12 900	156 739	5,6%
14	Novembre	9 703	8 215	9522	9 388	8 897	10 011	8 860	8465	8 870	8400	90331	3,2%
15	Décembre	11 439	10 467	11 670	12 605	11 828	11 392	11 205	10 140	10 747	10 200	111 693	4,0%
16	TOTAL	283 984	271 361	286 191	297 922	288 255	279 087	275 963	271 598	276 303	267 300	2 797 964	100,0 %
-	--------------------												
22	TABLEAU 2												
23													
24							Moyenne	Ecart- type					
25					Janvier			675					
26					Février		9053	653					
27					Mars		9697	988					
28					Avril		17208	1515					

29					Mai	25 759	1 663					
30					Juin	50898	4558					
31					Juillet	50187	5742					
32					Août	40078	3797					
33					Septembre	34273	3328					
34					Octobre	15674	1 186					
35					Novembre	9033	569					
36					Décembre	11 169	744					

On a recensé entre 1907 et 2000 le nombre mensuel de mariages en France métropolitaine. Les résultats de l'enquête sont regroupés dans le tableau 1 donné ci-dessous.

1. a. Compléter les cellules M4 et M5 de l'annexe, les résultats seront arrondis à 0,1 %. On ne demande pas le détail des calculs.

b. Donner une formule qui, placée dans la cellule M4 puis recopiée vers le bas jusqu'en M16, permet d'obtenir ces fréquences.

c. Donner une formule qui, placée dans la cellule B16 puis recopiée vers la droite jusqu'en L16, permet d'obtenir les totaux par colonne.

2. Dans cette question, les résultats seront arrondis à 0,1 %.

a. Quel est le pourcentage d'évolution du nombre total de mariages de 1997 à 2000 ? Préciser s'il s'agit d'une augmentation ou d'une diminution.

b. Quel est le pourcentage d'évolution du nombre total de mariages de 2000 à 2006 ? Préciser s'il s'agit d'une augmentation ou d'une diminution.

3. Le tableau 2, calculé pour chaque mois de l'année, le nombre moyen de mariages entre les années 1907 et 2006, ainsi que l'écart-type correspondant. Les nombres sont arrondis à l'unité.

a. Compléter le contenu de la cellule G25 dans le tableau 2. Arrondir à l'unité.

b. Donner une formule qui, placée dans la cellule G25 puis recopiée vers le bas jusqu'en G36, permet d'obtenir ces moyennes.

c. Les nombres moyens de mariages en juin et juillet sont sensiblement les mêmes environ - 50 000 mariages - alors que les écarts-types sont très différents. Interpréter cette différence.

D'après Bac L Maths-Info France 2008

1.3.14. Promos

Partie A

On donne ci-dessous le diagramme en boîte des montants en euros des achats effectués par les clients d'un magasin lors d'une journée de promotion. Les extrémités du diagramme correspondent au montant minimal et au montant maximal des achats effectués par les clients.

Quels sont les cinq renseignements sur les achats effectués dans le magasin lors de la journée de promotion que l'on peut lire, sur ce diagramme ?

Partie B

Le tableau ci-dessous donne les montants en euros, arrondis à l'unité, des achats effectués par les 80 clients du magasin pendant une journée ordinaire.

2	10	14	25	33	39	40	45
3	10	20	26	35	39	40	45
5	10	20	30	36	39	42	45
5	10	20	30	38	40	42	45
5	10	20	30	38	40	42	45
8	10	20	30	38	40	43	46
8	11	20	30	38	40	43	46
8	13	21	30	38	40	43	47
8	14	24	31	39	40	44	55
10	14	24	33	39	40	44	60

1. a. Déterminer le pourcentage de clients ayant effectué des achats pour un montant compris, au sens large, entre 30 et 40 euros.

b. Déterminer le pourcentage de clients ayant effectué des achats pour un montant ne dépassant pas 25 euros.

2. a. Déterminer la médiane de la série des montants d'achats donnée par le tableau ci-dessus.

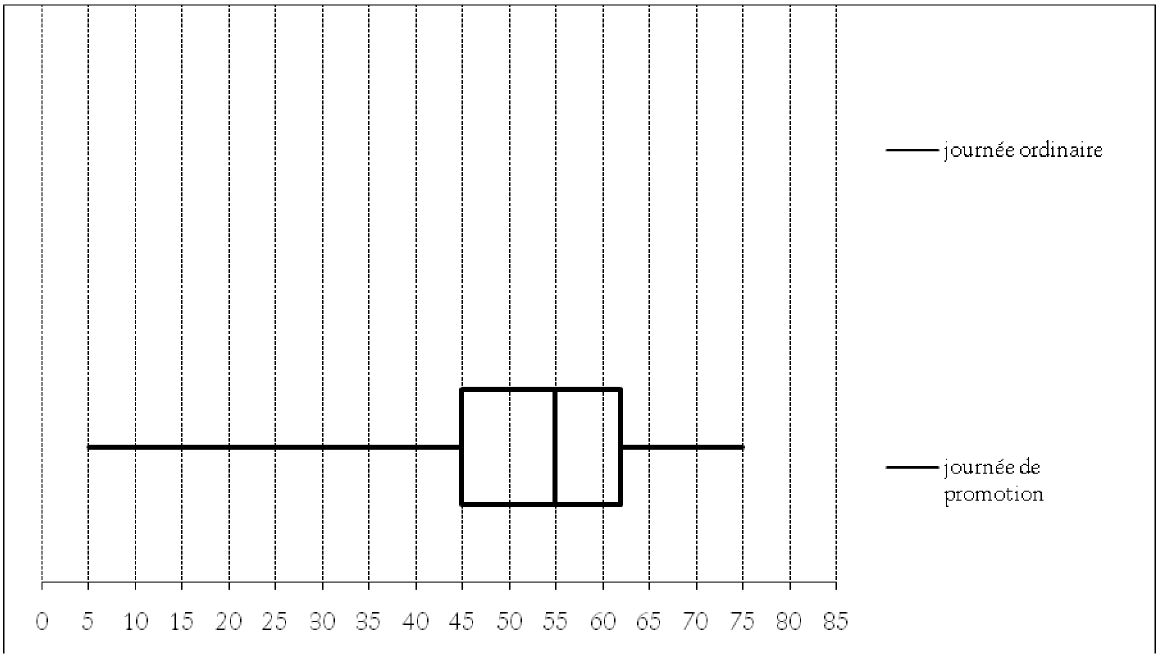
b. Déterminer le premier quartile Q_1 et le troisième quartile Q_3 de cette série.

c. Construire le diagramme en boîte de cette série au dessus du diagramme en boîte donné. On prendra pour extrémités le minimum et le maximum de la série.

3. Le magasin a annoncé sa journée de promotion par une distribution de tracts sur lesquels était indiqué :

« *Grande journée de promotion '.Des prix, des affaires, l'occasion de dépenser moins '.* »

Au vu des deux diagrammes en boîtes quelle analyse peut-on faire de ce message publicitaire ?



D'après Bac L Maths-Info Centres étrangers 2008

1.3.15. Machines

On se propose de comparer la fabrication de moquette sur deux machines A et B.

1. Sur la machine A, on prélève quotidiennement pendant 120 jours un échantillon de 1 m^2 afin d'en contrôler sa masse. On obtient les résultats ci-dessous.

Résultats machine A

Masse (en gramme)	Effectif n_i
[285 ; 290[	11
[290 ; 295[	18
[295 ; 300[	35
[300 ; 305[	20
[305 ; 310[	10
[310 ; 315[	6
[315 ; 320[	8
[320 ; 325[	12

1.1. Tracer l'histogramme des effectifs.

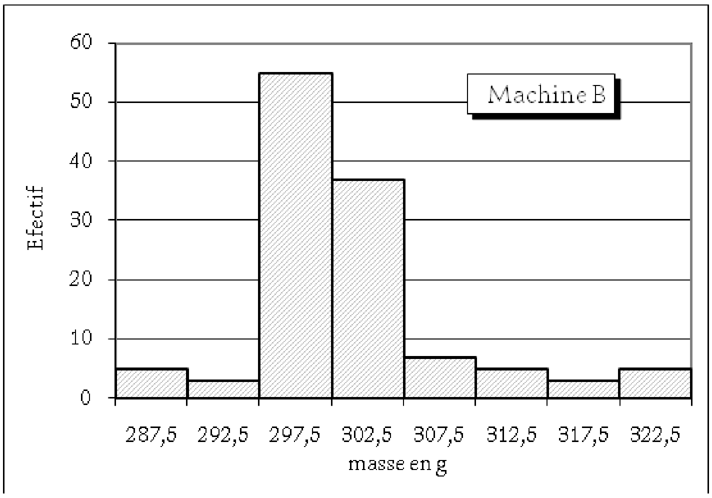
1.2. Calculer la moyenne de cette série à l'unité près.

1.3. Calculer l'écart type de cette série à l'unité près.

2. Sur la machine B, on effectue les mêmes prélèvements et on obtient les résultats suivants :

- la moyenne : $\overline{x} = 301$
- l'écart type : $\sigma = 7$
- l'histogramme des effectifs ci-dessous :

Résultats machine B



En comparant les résultats de ces deux machines, laquelle vous paraît la plus fiable ?

1.3.16. Population francaise en 2030

Le tableau ci-dessous donne la répartition en fonction de l'âge de la population française en 2010 et une prévision pour 2030.

Age	Effectif en 2010	Effectif prévu en 2030
[0;10[	7,6	6,5
[10;20[	7,7	6,8
[20;30[	8,1	7,3
[30;40[	8,3	7,4
[40;50[	8,6	7,9
[50;60[	8,2	8,3
[60;70[	6,3	8,2
[70;80[	4,6	6,4
[80;90[	2,9	3,5
[90;100[	0,6	0,9

L'objectif de cette activité est de calculer, à l'aide d'un ordinateur, différents indicateurs de tendance centrale et de dispersion pour comparer ces deux répartitions.

1. Étude de l'année 2010

Entrée des données dans le tableur

a. Pour entrer les libellés des colonnes, saisir « a » dans la cellule B2, « b » en C2, « n_i » en D2 et « x_i » en E2.

Remarque : pour les indices (ou exposants) cliquer sur Format dans la barre d'outil, sélectionner Cellule puis cliquer sur l'onglet Police et sélectionner Indice (ou exposant)

b. Pour entrer les classes d'âge : saisir verticalement à partir de B3 les bornes inférieures des intervalles de classes ; saisir verticalement à partir de C3 les bornes supérieures des intervalles de classes.

c. Pour entrer les effectifs de l'année 2010 : saisir verticalement les effectifs à partir de D3 ; sélectionner les effectifs et cliquer sur le bouton Σ pour effectuer le total de la colonne ; l'effectif total apparaît en D13.

d. Pour entrer les centres de classes : saisir « =(B3+C3)/2 » en E3 ; cliquer-glisser à partir du coin inférieur droit de E3 jusqu'en E12.

Remarque : pour les calculs qui suivent, on admet que la totalité de l'effectif de chaque classe est affecté en son centre x_i.

Estimation de la moyenne : $\overline{x} = \dfrac{\sum n_i x_i}{N}$

a. Pour entrer le libellé de la colonne de calcul, saisir « $n_i\, x_i$ » en F2.

b. Pour calculer une estimation de la moyenne : saisir « = D3*E3 » en F3 puis faire un cliquer-glisser jusqu'en F12 ; effectuez le total de la colonne avec le bouton Σ ; saisir « Moyenne : » en B16 ; saisir « = F13/D13 » en C16 ; valider, l'estimation de la moyenne apparaît.

Remarque : pour choisir le nombre de chiffres après la virgule, cliquer sur le bouton .

Calcul de l'écart type : $\sigma = \sqrt{\dfrac{\sum n_i x_i^2}{N} - \overline{x}^2}$

a. Pour entrer le libellé de la colonne de calcul, saisir « $n_i\, x_i^2$ » en G2.

b. Pour calculer l'écart-type : saisissez « = D3*E3*E3 » en G3, puis faire un cliquer-glisser jusqu'en G12 ; effectuez le total de la colonne avec le bouton Σ ; saisissez « Ecart type : » en E16 ; saisissez « = RACINE(G13/D13-C16*C16) » en G16 ; validez, la valeur de l'écart type apparaît.

2. Étude de l'année 2030

En procédant comme à la question précédente avec l'aide du tableur, calculer une estimation de la moyenne et l'écart-type de la population française prévue en 2030.

3. Analyse des résultats

Comparer les indicateurs statistiques moyenne et écart type en 2010 avec ceux en 2030. En déduire un commentaire sur l'évolution de la population française prévue sur cette période.

Éléments de réponse

M16 fx =RACINE(N13/K13-J16*J16)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1															
2		a	b	ni	xi	ni xi	ni xi2		a	b	ni	xi	ni xi	ni xi2	
3		0	10	7,6	5	38	190		0	10	6,5	5	32,5	162,5	
4		10	20	7,7	15	115,5	1732,5		10	20	6,8	15	102	1530	
5		20	30	8,1	25	202,5	5062,5		20	30	7,3	25	182,5	4562,5	
6		30	40	8,3	35	290,5	10167,5		30	40	7,4	35	259	9065	
7		40	50	8,6	45	387	17415		40	50	7,9	45	355,5	15997,5	
8		50	60	8,2	55	451	24805		50	60	8,3	55	456,5	25107,5	
9		60	70	6,3	65	409,5	26617,5		60	70	8,2	65	533	34645	
10		70	80	4,6	75	345	25875		70	80	6,4	75	480	36000	
11		80	90	2,9	85	246,5	20952,5		80	90	3,5	85	297,5	25287,5	
12		90	100	0,6	95	57	5415		90	100	0,9	95	85,5	8122,5	
13				62,9		2542,5	138232,5				63,2		2784	160480	
14															
15															
16		Moyenne :	40,42		Ecart type :	23,74			Moyenne :	44,05		Ecart type :	24,47		
17															
18															
19		Population française en 2010							Population française en 2030						
20															
21															

L'écart type en 2010 est peu différent de celui en 2030 (moins d'un an d'écart). La moyenne en 2030 est supérieure de 3,6 ans à celle en 2010.

La prévision prévoie donc que la population française va peu augmenter sur la période étudiée. En revanche elle sera plus âgée en moyenne de 3,5 ans, avec une dispersion par rapport à la moyenne variant peu.

Commentaire

L'élève ne doit pas être en totale autonomie dans la phase d'apprentissage informatique. La présence du professeur est indispensable pour apporter l'aide nécessaire lorsqu'il rencontre des difficultés liées aux différentes procédures à mettre en œuvre.

Cette activité permet, dans un premier temps, d'introduire les outils du tableur lors de l'étude de l'année 2010. Dans un deuxième temps, lors de l'étude de l'année 2030, plus d'autonomie est laissée à l'élève pour consolider l'acquisition de ces outils en lui demandant de reproduire les mêmes procédures.

Comparer deux séries statistiques à l'aide d'indicateurs de tendance centrale et de dispersion donne du sens à l'activité.

L'analyse des résumés numériques des séries pourrait être prolongée en demandant à l'élève de comparer des représentations graphiques obtenues à l'aide du tableur-grapheur.

1.3.17. Qui a les meilleurs réflexes ?

Nous allons déterminer quel est l'élève de la classe qui a les meilleurs réflexes. Pour cela, nous allons réaliser avec chaque élève une expérience simple de mesure indirecte du temps de réaction.

1. Travaux pratiques par groupes de trois élèves

Un des élèves tient une règle, graduée en cm, en la laissant pendre. Un autre place le pouce et l'index au niveau du zéro en bas de la règle, sans la tenir. Lorsque la règle est lâchée sans l'avertir, il essaye de l'attraper au vol. Le troisième élève enregistre, dans le tableau d'un tableur-grapheur, la hauteur h correspondant à la distance entre le bas de la règle et le point attrapé. Ce test est effectué 40 fois par chacun des trois élèves.

2. Traitement informatique des données

a. Comparer vos performances avec ceux des autres élèves de la classe à partir d'indicateurs statistiques calculés à l'aide du tableur.

b. Quel est celui qui a les meilleurs réflexes ? Justifier.

Eléments de réponse

Exemples d'indicateurs statistiques obtenus expérimentalement

	PRÉNOM	Indicateurs de tendance centrale			Indicateurs de dispersion		Classement
		Mode Mo	Médiane Me	Moyenne $\overline{x}$	Étendue e	Écart type σ	
1	Claire	28	25	28,5	37	11,9	
2	Nathan	17	10	10,1	24	6,6	2nd ?
3	Zoé	12	14	15,5	35	5,4	
4	Camille	18	12	12,4	25	5,6	2nd ?
5	Jamal	25	25	26,9	30	7,2	
6	Hubert	25	30	30,4	39	11,7	
7	Romain	15	15	15,2	12	4,2	2nd ?
8	Marie	18	16	15,1	49	10,2	
9	Flora	15	15	15,2	12	4,2	2nd ?
10	Lucie	15	15,5	17,2	29	6,7	
11	Frédéric	12	9	9,5	10	3,1	1er
12	Anaïs	14	17	17,2	27	6,5	
13	Halima	19	14	16	39	11,8	

Commentaires

Cette activité met en œuvre des travaux expérimentaux de deux sortes : des travaux pratiques permettant de construire des séries statistiques individualisées, et des calculs d'indicateurs statistiques à l'aide des TIC.

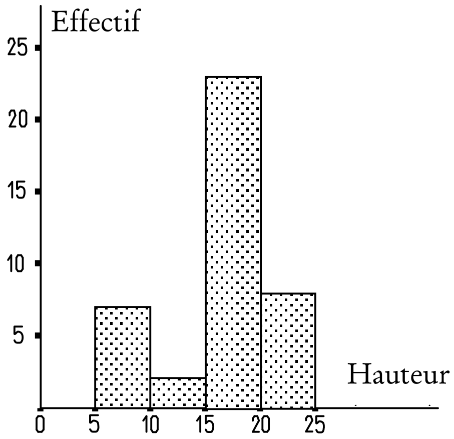
Le calcul « à la main » des indicateurs statistiques ne présente aucun intérêt dans la mesure où les TIC permettent de les obtenir très facilement.

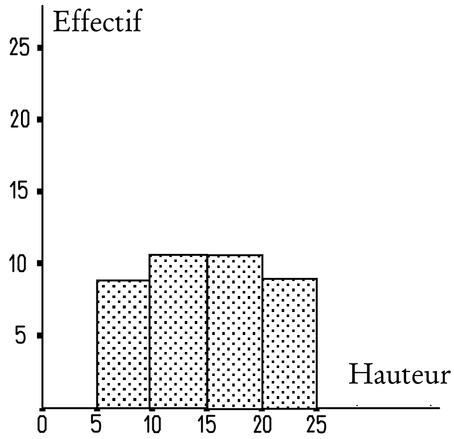
Identifier l'élève qui a les meilleurs réflexes est aisé si l'un d'entre eux se détache des autres en ayant les meilleurs indicateurs (ceux de l'élève Frédéric par exemple). Dans le cas contraire, il est nécessaire d'identifier le ou les indicateurs pertinents. Les élèves doivent rechercher ce ou ces indicateurs, en justifiant leur choix en donnant du sens à ces nombres. Le fait qu'ils soient obtenus à partir de données individualisées facilite leur interprétation et permet de comparer plus facilement des séries deux à deux.

Dans l'activité présentée, si l'objectif est de repérer dans la classe le meilleur chronométreur pour une compétition d'athlétisme, celui ayant la plus faible étendue est le mieux placé. Si l'objectif est de trouver le plus régulier, il faut rechercher les plus petites valeurs du couple « moyenne, écart type »

Il est intéressant de faire remarquer que deux séries de même écart type (ou très peu différent) peuvent, si elles sont éloignées d'une distribution normale, avoir une distribution très différente. Un graphique peut alors montrer davantage qu'un simple résumé numérique.

Les élèves Romain et Flora ont des indicateurs statistiques identiques mais les distributions sont très différentes. Elles montrent que Flora réussit des hauteurs inférieures à 15 cm plus souvent que Romain.





Information

La distance indiquée sur la règle permet d'accéder au temps de réaction. En effet, la hauteur h d'une chute de durée t étant donnée par $h=\frac{1}{2}gt^2$, le temps de réaction est : $t=\sqrt{\frac{2h}{g}}$.

1.3.18. Contrôle d'épaisseur

Le contrôle d'une production de tablettes en bois est effectué toutes les heures par prélèvement d'un échantillon de 20 éléments. On effectue un relevé de l'épaisseur de chaque élément.

L'épaisseur normale est de 18 mm. Une pièce sans défaut est une pièce dont l'épaisseur ne s'écarte pas de plus de 0,5 mm de l'épaisseur normale.

Les pièces ne correspondant pas à cette exigence sont classées en deux catégories :

Défaut mineur	L'épaisseur s'écarte de 0,5 à 1 mm de l'épaisseur normale.	Défaut qui ne nécessite pas d'opération de reprise.
Défaut majeur	L'épaisseur s'écarte de 1 à 2 mm de l'épaisseur normale.	Défaut qui permet d'atteindre le stade suivant sous réserve d'opérations particulières sur cet élément ou un autre.

Lors d'un contrôle on a relevé les épaisseurs suivantes (en mm) :

Epaisseur en mm	16,6	16,8	17,2	17,3	17,6	17,7	17,8	17,9	18	18,2	18,3	18,4	18,8
Effectif	1	1	1	1	3	1	2	2	1	2	1	3	1

1. a. Quel est le nombre de pièces présentant un défaut mineur ?

b. Quel est le nombre de pièces présentant un défaut majeur ?

c. Quel pourcentage de l'échantillon est constitué par des pièces présentant un défaut ?

2. a. Calculer l'épaisseur moyenne $\overline{x}$ de l'échantillon ; arrondir au centième de mm.

b. Calculer l'écart-type σ de l'échantillon.

c. Quel est le pourcentage de pièces dont l'épaisseur est dans l'intervalle .. ?

3. Si l'épaisseur moyenne s'écarte de plus de 0,25 mm de l'épaisseur normale, un réglage de la machine doit être effectué. Est-ce le cas ?

1.4.19. Insécurité routière

Cette activité (pas vraiment du niveau de Seconde mais instructive) consiste en l'illustration de données statistiques à l'aide d'un tableur, puis à leur interprétation, dans un contexte « citoyen ».

Le tableau ci-dessous fournit, pour la France, la vitesse moyenne des véhicules légers, ainsi que le nombre de morts sur les routes, de 1998 à 2006.

Année	Vitesse moyenne des véhicules légers (km/h)	Nombre de morts
1998	88,7	8 437
1999	88,6	8 029
2000	90,1	7 643
2001	89,4	7 720
2002	89,2	7 242
2003	86,8	5 731
2004	84,5	5 593
2005	82,9	5 318
2006	82	4 703

(Source www.securiteroutiere.gouv.fr).

1. a. Représenter, à l'aide d'un tableur, l'évolution de la vitesse moyenne en fonction des années (choisir un « nuage de points reliés par une courbe »).

b. Représenter de même l'évolution du nombre de morts en fonction des années.

c. Comparer les deux graphiques.

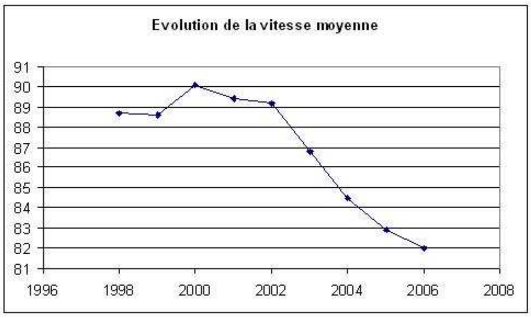
2. a. Représenter, à l'aide d'un tableur, le nuage de points (non reliés) correspondant à la série statistique à deux variables, vitesse et nombre de morts, en plaçant la vitesse en abscisses et le nombre de morts en ordonnées.

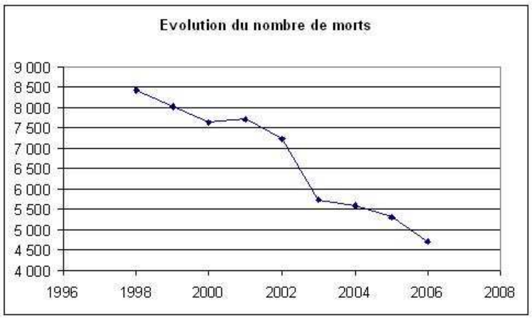
b. Effectuer, à l'aide du tableur, un ajustement affine du nuage précédent.

c. Interpréter le graphique obtenu.

Éléments de réponse

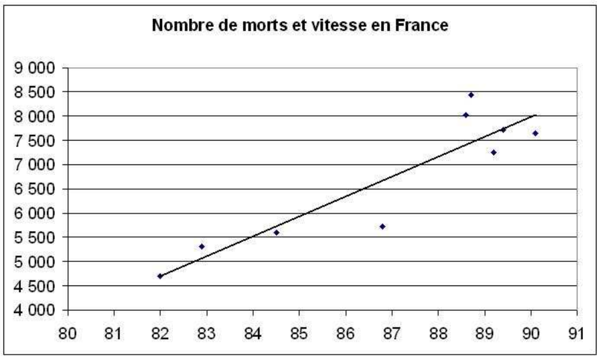
1. a. et b.





c. Les deux graphiques sont très semblables, avec une tendance générale à la baisse (et une petite « bosse » autour de 2001). Ceci conduit à l'idée d'une corrélation entre la vitesse et le nombre de morts, que l'on étudie à la question suivante.

2. a. et b.



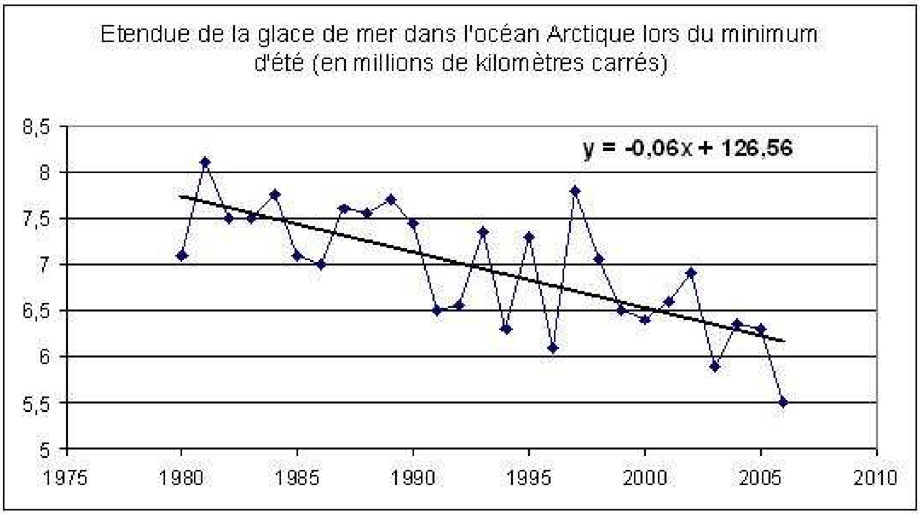
c. La droite indique la « tendance » du nuage : lorsque la vitesse augmente, le nombre de morts à tendance à augmenter.

1.4.20. Glace en Arctique et température globale de la Terre

Source : Nations Unies – GIEC (Groupe d'experts intergouvernemental sur l'évolution du climat – IPCC en anglais) rapport 2007 – http://www.ipcc.ch/

1. Étendue de la glace de mer dans l'océan Arctique

Le graphique suivant donne l'étendue minimale, en millions de km^2, de la glace de mer dans l'océan Arctique, mesurée chaque été de 1980 à 2006.



L'ajustement affine des observations permet de matérialiser la tendance et « d'extrapoler ».

La droite obtenue à l'aide du tableur-grapheur a pour équation $y = -0{,}06\,x + 126{,}56$ où x est l'année et y l'étendue de glace de mer en millions de km^2 calculée avec cette valeur de x.

Les questions suivantes peuvent se poser :

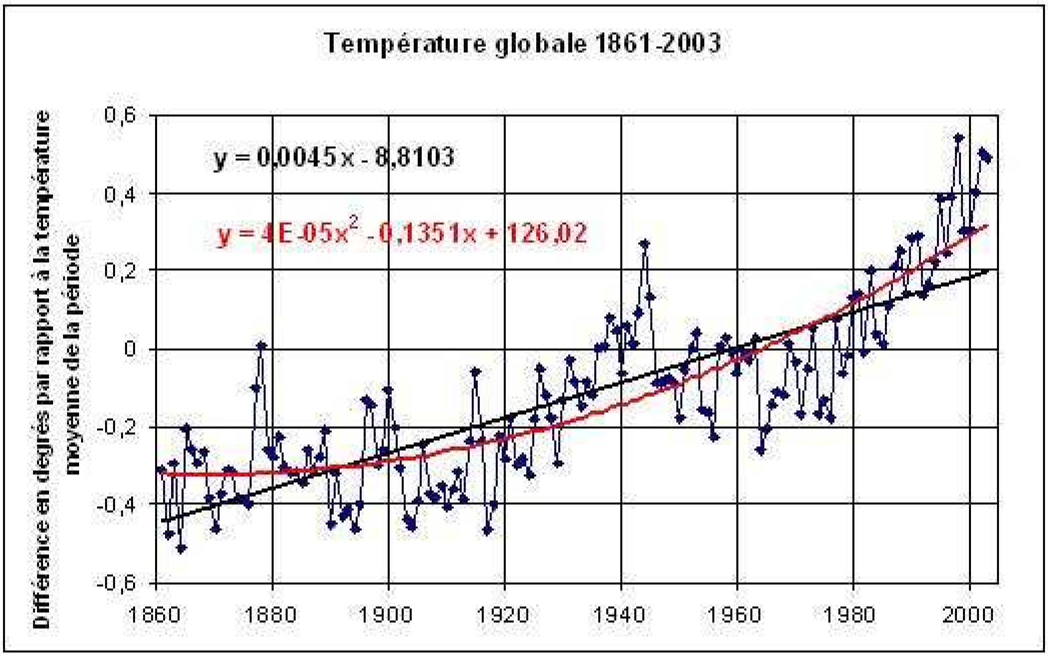
– est-il vrai que « la tendance est à une perte de 60 000 km^2 par an » ?

– si la tendance observée se maintient, quelle serait l'étendue minimale de la glace de mer en Arctique durant l'été 2050 ?

– si la tendance observée se maintient, en quelle année n'y aurait-il plus de glace de mer en Arctique en été ?

2. Température globale à la surface de la Terre

Les données ci-dessous fournissent pour la période 1861-2003 les écarts à la moyenne, de la température globale à la surface de la Terre. La valeur 0 correspond à la moyenne sur la période 1861-2003. Par exemple, en 1900, la température moyenne à la surface de la Terre est de – 0,1067° C en-dessous de la moyenne de la période.



Avec un tableur-grapheur, ces données peuvent être ajustée avec deux modèles mathématiques différents :

– Modèle 1 : la droite d'équation $y = 0{,}0045\,x - 8{,}8103$;

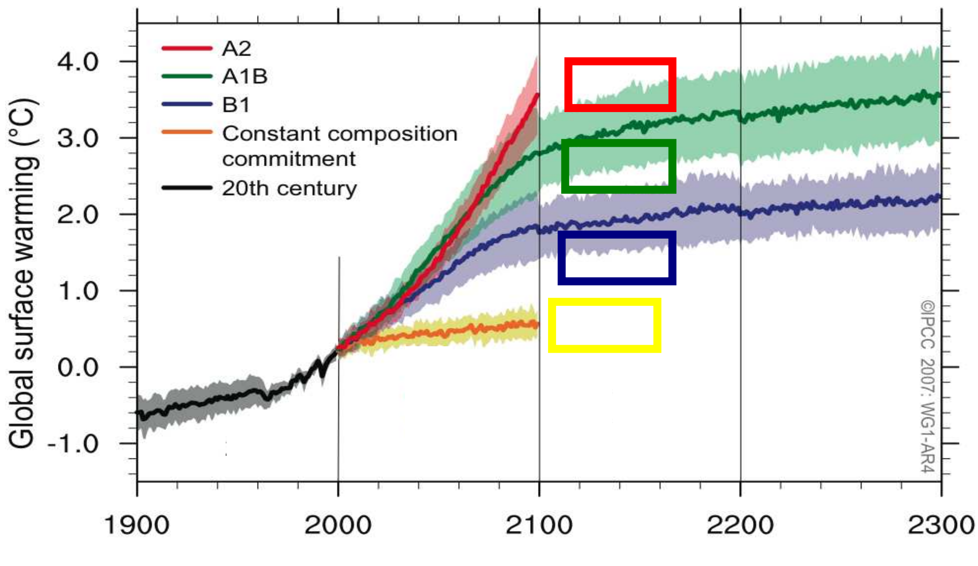
– Modèle 2 : la parabole d'équation $y = 4.10^{-5}\,x^2 - 0{,}1351\,x + 126{,}02$.

Chacun de ces modèles peut être utilisé pour estimer, en 2100, l'écart en degrés de la température de la Terre par rapport à la moyenne de 1861 à 2003 si la tendance observée se poursuit.

– Modèle 1 : + 0,639 7°.

– Modèle 2 : + 18,71°.

Il convient cependant de « relativiser » le caractère prédictif de ces simples calculs, en effectuant par exemple une recherche sur les modélisations utilisées par les chercheurs du GIEC. Les modèles précédents sont plutôt rudimentaires (et le modèle 2 assez inquiétant), le GIEC a mis au point plusieurs modèles, tenant compte, en particulier, des évolutions possibles des concentrations de gaz à effet de serre dans l'atmosphère.



Ces modèles prévoient que le réchauffement continuera si les concentrations de gaz à effet de serre augmentent.

Si les concentrations étaient maintenues au niveau actuel, un réchauffement inexorable de 0,6°C se produirait d'ici à 2100. Un réchauffement plus large se produirait pour les concentrations plus élevées (autres modèles).

Commentaire

Cette étude permet de montrer des ajustements autres qu'affine et ne doit pas faire l'objet de développements théoriques pour d'autres modèles d'ajustement.

1.4.21. Mettre des gants

Travail préparatoire

Quelles sont la température minimale et la vitesse de vent maximale enregistrées par le service météorologique de votre ville au cours des 100 dernières années ?

Informations

Durant les expéditions polaires il est très important de se protéger du froid pour éviter les gelures. Des lésions irréversibles peuvent apparaître aux extrémités (mains, pieds, nez et oreilles) si elles sont soumises plus de 30 minutes à des températures inférieures à −25°C.

Le tableau ci-dessous donne des valeurs de l'indice de refroidissement éolien (IRE) utilisé lors de ces expéditions. Il prend en compte la vitesse du vent pour donner la température réellement ressentie.

Vitesse du vent (en km/h) Température de l'air (en °C)	0	10	20	30	40	50
– 10	– 10	– 15	– 18	– 20	– 21	– 23
– 5	– 5	– 9,5	– 12	– 13	– 13	– 14
0	0	– 3	– 5	– 6,5	– 7	– 7,5
5	5	3	1	0	– 0,5	– 1
10	10	9	7,5	7	6	5,5

Exemple : une température de l'air de –10°C avec un vent de 50 km/h correspond à une température ressentie (IRE) de –23°C.

Problème

La température ressentie par le conducteur d'un scooter à 50 km/h est la même que celle donnée dans le tableau des indices de refroidissement éolien (IRE) pour une vitesse du vent de 50 km/h.

Pour un déplacement de 30 minutes à 50 km/h et sans gant, à partir de quelle température de l'air les mains du conducteur vont-elles se geler avec apparition de lésions irréversibles ?

1.4.22. Accidents

La Direction de la Sécurité Routière relève, tous les deux ans, le nombre de personnes tuées dans les accidents de la route. Le tableau ci-dessous indique, à partir de 1983, le rang x de l'année ainsi que le nombre y de personnes décédées dans un accident de la route au cours de cette année.

Année	1983	1985	1987	1989	1991	1993
x	0	1	2	3	4	5
y	11 946	1 0454	9 855	10 528	9 617	9 052
Année	1995	1997	1999	2001	2003	2005
x	6	7	8	9	10	11
y	8 413	7 989	8 029	7 720	5 731	5 318

1. Dans le plan rapporté à un repère orthogonal $(O\,;\vec{i},\vec{j})$, représenter, sur une feuille de papiermillimétré, le nuage des points de coordonnées $(x\,;y)$ associé aux données du tableau.

On prendra pour unités graphiques : − 1 cm pour une unité sur l'axe des abscisses et 1 cm pour 1 000 unités sur l'axe des ordonnées.

2. a. Calculer les coordonnées du point moyen G_1 des six premiers points et celles du point moyen G_2 des six derniers points.

b. Placer ces points sur le graphique et tracer la droite (G_1G_2).

c. Montrer qu'une équation de la droite (G_1G_2) est $y = -507x + 11\,509{,}5$.

3. On considère que la droite (G_1G_2) permet de fournir une bonne approximation du nombre de décès dans les accidents de la route jusqu'en 2010.

a. Utiliser le graphique afin d'estimer le nombre de décès causés par un accident de la route en 2009.On fera apparaître les traits de construction.

b. Déterminer par le calcul en quelle année on peut espérer que le nombre de tués par accident de la route soit inférieur à 4 500.

1.5. Probabilités élémentaires

1.5.23. QCM

1. Le nombre d'allocataires du RMI âgés de plus de 50 ans est passé de 150 000 en 1995 à 262 500 en 2005. Entre 1995 et 2005, ce nombre a augmenté d'environ :

a. 42,9 %	b. 112,5 %	c. 75 %	d. 57,1 %

2. Dans une cage, il y a cinq lapins, deux blancs et trois noirs, et quatre cochons dinde, deux blancs et deux marrons. La probabilité qu'un animal choisi au hasard dans la cage soit blanc est :

a. $\frac{2}{9}$	b. $\frac{1}{3}$	c. $\frac{2}{5}+\frac{2}{4}$	d. $\frac{4}{9}$

4. Dans un club, on a recueilli les lieux de séjour des 120 membres pour les dernières vacances.

Chacun avait choisi un séjour à la mer ou bien à la campagne. Les résultats de l'enquête sont consignés dans le tableau ci-dessous.

	Mer	Campagne
Femmes	60	20
Hommes	10	30

On choisit une personne au hasard dans ce club. La probabilité que ce soit un homme ou une personne ayant passé ses dernières vacances à la mer est :

a. $\frac{1}{12}$	b. $\frac{11}{12}$	c. $\frac{5}{6}$	d. $\frac{3}{4}$

1.5.24. Grippe

Partie A

Dans un lycée de 1 280 élèves, 300 élèves se font vacciner contre la grippe. Pendant l'hiver, il y a une épidémie de grippe et 10 % des élèves contractent la maladie. De plus 3 % des élèves vaccinés ont la grippe.

1. Reproduire et compléter le tableau suivant, sans justifier les réponses :

	Nombre d'élèves ayant eu la grippe	Nombre d'élèves n'ayant pas eu la grippe	Total
Nombre d'élèves vaccinés			
Nombre d'élèves non vaccinés			
Total			1 280

Pour les trois questions suivantes, tous les résultats seront arrondis à 0,001 près.

2. On choisit au hasard l'un des élèves de ce lycée, tous les élèves ayant lamême probabilité d'être choisis. On considère les évènements suivants :

A : « L'élève a été vacciné » ;

B : « L'élève a eu la grippe » ;

C : « L'élève a été vacciné et a eu la grippe ».

a. Calculer la probabilité des évènements A, B et C.

b. Calculer la probabilité de l'évènement $A \cup B$.

3. On choisit au hasard un des élèves vaccinés. Calculer la probabilité de l'événement : « L'élève a eu la grippe ».

4. On choisit au hasard un des élèves non vaccinés. Calculer la probabilité de l'évènement : « L'élève a eu la grippe ».

5. Expliquer pourquoi on peut en déduire que ce vaccin a été efficace pour les élèves de ce lycée.

Partie B

Dès l'apparition des premiers symptômes de l'épidémie, l'infirmière du lycée relève pendant 8 jours le nombre d'élèves malades. Le tableau ci-dessous indique les résultats observés.

Numéro du jour	1	2	3	4	5	6	7	8
Nombre d'élèves grippés	2	5	9	14	17	23	27	31

1. Construire dans un repère orthogonal le nuage de points associé cette série statistique. On prendra les unités suivantes : en abscisse 2 cm pour 1 jour et en ordonnée 1 cm pour 2 élèves.

2. Calculer les coordonnées du point moyen G du nuage et le placer sur le graphique.

3. On appelle G_1 le point moyen des quatre premiers points de ce nuage et G_2 le point moyen des quatre derniers points.

a. Déterminer les coordonnées de G_1 et de G_2.

b. Placer les points G_1 et G_2 sur le graphique puis tracer la droite (G_1G_2).

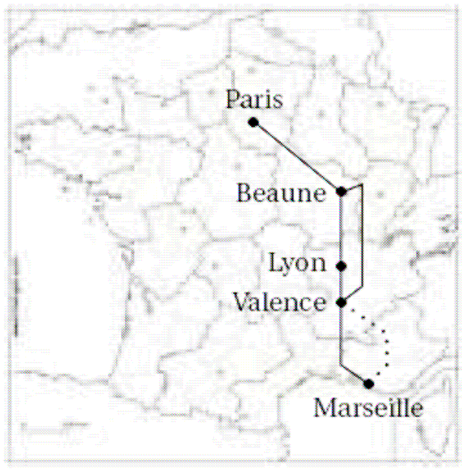
c. Déterminer une équation de la droite (G_1G_2) sous la forme $y = mx + p$. On donnera les valeurs exactes de m et de p.

4. En utilisant l'équation de la droite (G_1G_2), estimer à partir de combien de jours au moins 5 % des élèves du lycée seront atteints par la grippe.

1.5.25. Bison futé

Lors des journées « rouges » selon Bison Futé, l'autoroute qui relie Paris à Marseille est surchargée. Il est donc conseillé de prendre un itinéraire de délestage entre Beaune et Valence (qui ne passe pas par Lyon) afin d'éviter les éventuels « bouchons » autoroutiers.

Entre Valence et Marseille il est également conseillé de prendre la route départementale représentée sur la carte par des pointillés.



Bison Futé a publié les résultats d'une étude portant sur les habitudes des automobilistes sur le trajet entre Paris et Marseille lors de ces journées « rouges ».

Il s'avère que :

• 40 % des automobilistes prennent l'itinéraire de délestage entre Beaune et Valence ;

• parmi les automobilistes ayant suivi l'itinéraire de délestage entre Beaune et Valence, 30 % prennent la route départementale de Valence à Marseille ;

• parmi les automobilistes n'ayant pas suivi l'itinéraire de délestage entre Beaune et Valence, 60 % prennent la route départementale de Valence à Marseille.

On note :

B l'évènement « l'automobiliste prend l'itinéraire de délestage entre Beaune et Valence » et $\overline{B}$ l'évènement contraire ;

V l'évènement« l'automobiliste prend la route départementale entre Valence et Marseille » et $\overline{V}$ l'évènement contraire.

1. a. Représenter la situation à l'aide d'un arbre pondéré.

b. Montrer que la probabilité de l'évènement $\overline{B}\cap\overline{V}$ est $p\left(\overline{B}\cap\overline{V}\right)=0,24$ et interpréter ce résultat.

c. Calculer la probabilité que l'automobiliste ne choisisse pas la route départementale entre Valence etMarseille.

2. On donne les temps de parcours suivants :

Paris – Beaune (par autoroute) : 4 heures ;

Beaune – Valence (par autoroute, en passant par Lyon) : 5 heures ;

Beaune – Valence (par itinéraire de délestage, en ne passant pas par Lyon) : 4 heures ;

Valence – Marseille (par autoroute) : 5 heures ;

Valence – Marseille (par la route départementale) : 3 heures.

a. Calculer les temps de parcours entre Paris et Marseille, selon l'itinéraire choisi.

Compléter le tableau ci-dessous donnant les diverses probabilités de durée du trajet nécessaires pour se rendre de Paris à Marseille selon l'itinéraire choisi.

Temps en heures	11			14
Probabilité				0,24

b. Calculer le temps moyen de durée du trajet en heures et en donner une interprétation (la conversion en heure minute seconde n'est pas utile).

1.5.26. CK ou D&G ?

Les 800 élèves d'un lycée possèdent une montre, soit du type M_1 soit du type M_2.

- Il y a 70% de montres de type M_1.
- La moitié des montres de type M_1 a un bracelet en cuir.
- 16,25% des montres de type M_1 ont un bracelet métallique.
- Parmi les montres de type M_2, il y a trois fois plus de montres à bracelet en tissu que de montres à bracelet métallique.
- Il n'existe pas de montres de type M_2 avec un bracelet en cuir.

1. Reproduire et compléter le tableau suivant :

	Cuir	Métal	Tissu	Total
M_1				
M_2				
Total				800

2. Parmi l'ensemble de toutes les montres quel est le pourcentage des montres de type M_2 à bracelet en tissu ?

Parmi les montres de type M_2, quel est le pourcentage de celles qui ont un bracelet métallique ?

Dans les questions suivantes, les probabilités seront données à 10^{-3} près.

3. On choisit un élève au hasard parmi les 800 élèves du lycée.

Calculer la probabilité de chacun des évènements suivants :

A « la montre de l'élève a un bracelet métallique » ;

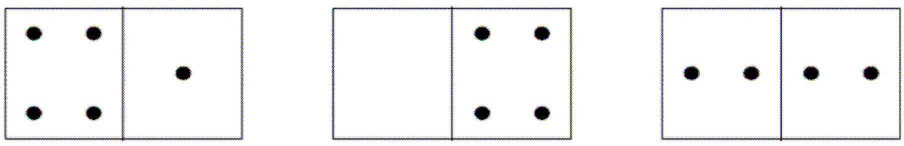
D « la montre de l'élève est de type M_2 ».

4. Définir par une phrase les évènements $A \cap B$ et $A \cup B$ puis calculer leur probabilité.

5. On choisit au hasard un élève ayant une montre de type M_1. Quelle est la probabilité de l'évènement C « la montre de l'élève a un bracelet en tissu » ?

1.5.27. Dominos

Un jeu de dominos est constitué de 28 dominos distincts. On rappelle qu'un domino est partagé en deux parties, chacune portant un nombre de 0 à 6 représenté par des points. Un double est un domino dont les deux parties portent le même nombre. Exemples de dominos :



1. Écrire la liste des 28 dominos distincts.

2. Un joueur tire un domino au hasard.

a. Quelle est la probabilité qu'il obtienne un double ?

b. Quelle est la probabilité d'obtenir un domino dont la somme des nombres situés sur les deux parties soit divisible par 3 ? (On rappelle que 0 est divisible par tout entier non nul.)

c. À chaque domino tiré on associe la différence entre le plus grand et le plus petit nombre. Par exemple, si le domino tiré porte le nombre 1 et le nombre 4, on prend la valeur $4-1=3$.

Quelles sont les valeurs maximum et minimum que l'on peut obtenir ?

Compléter le tableau ci-dessous.

Différence X						
Probabilité associée						

Quelle est la valeur moyenne de X ?

Donner une interprétation de ce résultat.

1.5.28. Boules

On tire au hasard une boule d'une urne contenant deux boules rouges notées R_1 et R_2, une boule verte notée V et deux boules bleues notées B_1 et B_2. On ne remet pas la boule tirée et on effectue un second tirage d'une boule.

On appelle résultat un couple dont le premier élément est la boule obtenue au premier tirage et le second, celle obtenue au second tirage, par exemple (R_1, B_2).

1. Déterminer à l'aide d'un tableau ou d'un arbre l'ensemble des résultats possibles.

2. On compléte la situation précédente par une régle du jeu :

- pour chaque boule rouge tirée, on gagne 1 euro ;
- pour chaque boule verte tirée, on gagne 2 euro ;
- pour chaque boule bleue tirée, on perd 2 euro.

a. Lister les gains possibles (une perte est considérée comme un gain négatif) et déterminer la probabilité de chacun de ces gains.

b. Calculer le gain moyen d'un joueur. Le jeu est-il équitable ?

1.5.29. Biologie

Des étudiants en agronomie procédent au croisement de deux variétés de pois, l'une ayant des graines jaunes et lisses, l'autre des graines vertes et ridées.

En premiére génération, appelée F1, les graines obtenues sont toutes semblables entre elles, elles sont jaunes et lisses.

Les étudiants croisent alors entre eux les individus de la génération F1, pour obtenir la génération F2.

L'observation de 5431 graines issues de la génération F2 montre que :

- 4 069 graines sont jaunes dont 3 057 lisses ;
- 341 graines sont vertes et ridées.

1. Reproduire et compléter le tableau suivant :

	graines jaunes	graines vertes	Total
graines lisses			
graines ridées			
Total			5 341

2. On tire au hasard une graine parmi les 5 431 de cet échantillon, tous les tirages étant équiprobables. Calculer la probabilité des événements suivants : A : « La graine est jaune » ; B : « La graine est lisse ».

3. On considére les événements suivants : $A \cap B$; $A \cup B$; $\overline{A}$ et $\overline{A} \cap \overline{B}$ où $\overline{A}$ et $\overline{B}$ désignent les événements contraires respectifs de A et B.

Définir chacun de ces événements par une phrase, puis calculer leur probabilité.

4. On prend, au hasard, une graine jaune. Quelle est la probabilité de l'événement C « la graine est ridée » ?

1.5.30. Loterie

Une roue de loterie munie d'un index fixe est divisée en secteurs de mêmes dimensions et de différentes couleurs. Le jeu consiste à miser 5 euros, à faire tourner la roue et à noter la couleur du secteur désigné par l'index à l'arrêt de la roue. On admet que chaque secteur a la même probabilité d'apparaître. La roue comporte :

- n secteurs rouges qui font perdre la mise (gain du joueur : -5 €) ;
- 6 bleus où l'on reçoit 5 € (gain du joueur nul) ;
- 3 verts où l'on reçoit 20 €;
- 1 jaune où l'on reçoit 100 €.

1. Dans cette question, la roue comporte 14 secteurs rouges ($n = 14$).

a. Déterminer les gains possibles (positifs ou négatifs) du joueur.

b. Calculer la moyenne des gains et interpréter ce résultat.

2. Dans cette question, la roue comporte n secteurs rouges et son propriétaire désire gagner en moyenne au moins 15% des sommes misées.

a. Montrer que le gain moyen d'un joueur est : $\dfrac{-5n+140}{n+10}$.

b. Déterminer le nombre minimum n de secteurs rouges que doit comporter la roue pour que le propriétaire soit content.

1.5.31. Bibliothèque

Une municipalité décide de regrouper tous les ouvrages de trois petites bibliothèques de quartier en un même lieu et de créer une bibliothèque municipale. On convient de noter b_1, b_2 et b_3 ces trois bibliothèques de quartier.

Le stock de b_1 constituera ainsi 50 % de l'ensemble des ouvrages réunis dans la bibliothèque municipale, celui de b_2 constituera 30 % de cet ensemble et celui de b_3 constituera 20 % de cet ensemble.

Un examen minutieux du stock révèle que :

• 12 % des ouvrages provenant de b_1 sont en mauvais état ;

• 10 % des ouvrages provenant de b_2 sont en mauvais état ;

• 15 % des ouvrages provenant de b_3 sont en mauvais état.

On prélève au hasard un ouvrage dans le stock de la bibliothèque municipale et on note sa provenance et son état.

On appelle les événements suivants :

B_1 l'évènement : « L'ouvrage prélevé provient de la bibliothèque b_1 » ;

B_2 l'évènement : « L'ouvrage prélevé provient de la bibliothèque b_2 » ;

B_3 l'évènement : « L'ouvrage prélevé provient de la bibliothèque b_3 » ;

E l'évènement : « L'ouvrage prélevé est en bon état » et $\overline{E}$ son contraire.

1. a. Donner la valeur de $p(B_1)$, probabilité de l'évènement B .

b. On sait que l'ouvrage choisi vient de b_1, quelle est la probabilité que l'ouvrage soit en bon état ?.

2. Reproduire sur la copie l'arbre de probabilité ci-contre et le compléter par les sept probabilités manquantes.

3. a. Montrer que $p(B_1 \cap E)=0{,}44$.

Calculer $p(B_2 \cap E)$ et $p(B_3 \cap E)$.

b. En déduire que la probabilité qu'un ouvrage prélevé au hasard soit en bon état est égale à 0,88.

4. Caractériser par une phrase l'évènement $B_1 \cup \text{E}$ puis calculer sa probabilité.

5. Je prends à la bibliothèque b_1 un livre toutes les semaines pendant 1 an que je rapporte consciencieusement.

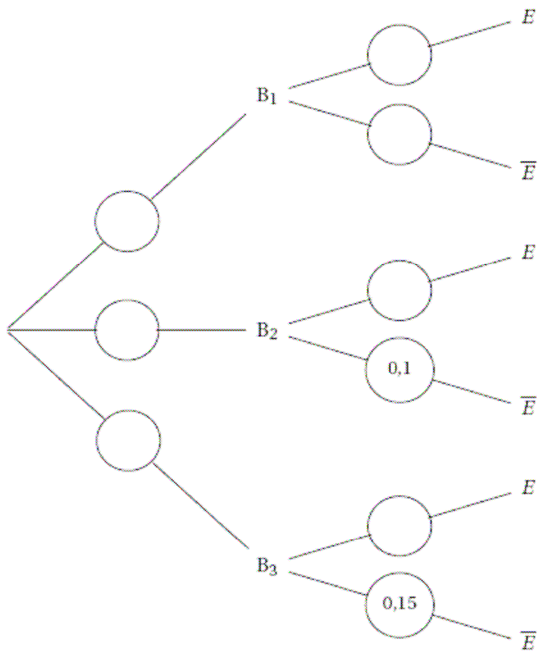
a. Quelle est la probabilité que je n'ai jamais choisi un livre en mauvais état ?

b. Au plus un livre en mauvais état ?

c. Que des livres en mauvais état ?

d. Au moins un livre en mauvais état ?

(on créera un petit programme donnant ces probabilités pour un nombre de semaines N et on l'utilisera avec N=52).



1.5.32. Médicament

Un laboratoire cherche à tester l'apparition d'éventuels effets secondaires liés à la prise d'un médicament. Pour cela, il sélectionne un échantillon de personnes en bonne santé parmi lesquelles 25 % ont entre 18 et 24 ans, 50 % ont entre 25 et 49 ans et 25 % ont 50 ans et plus. Suite à la prise de ce médicament, 9 % des personnes ayant entre 18 et 24 ans, 7 % des personnes ayant entre 25 et 49 ans et 12 % des 50 ans et plus ont vu apparaître des effets secondaires.

On choisit au hasard une personne ayant participé à ce test. On note :

A l'évènement « la personne a entre 18 et 24 ans » ;

B l'évènement « la personne a entre 25 et 49 ans » ;

C l'évènement « la personne a 50 ans ou plus » ;

S l'évènement « la personne a vu apparaître des effets secondaires suite à la prise dumédicament ».

$\overline{\text{S}}$ est l'évènement contraire de S.

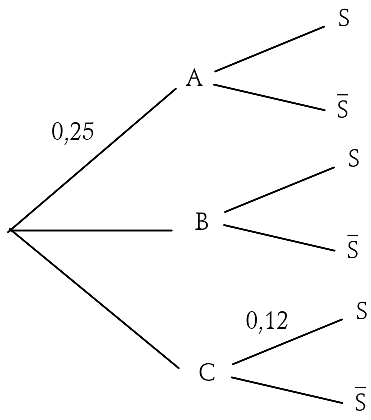
1. Reproduire et compléter l'arbre de probabilités suivant.

2. Calculer la probabilité de l'évènement $\text{A} \cap \text{S}$.

3. Montrer que la probabilité de choisir une personne ayant vu apparaître des effets secondaires est égale à 0,0875.

4. On choisit une personne n'ayant pas vu d'effets secondaires liés à la prise de ce médicament.

Quelle est la probabilité qu'elle ait entre 18 et 24 ans ? On arrondira la réponse à 10^{-4} près.



1.6.33. TP EXCEL Nous voulons une fille

TP EXCEL Nous voulons une fille !

Remarque technique : on utilise ici la fonction **=ALEA()** d'Excel. Cette fonction renvoie un nombre au hasard (aléatoire) compris entre 0 et 1 (1 non compris).

A chaque saisie d'une nouvelle cellule Excel recalcule toute la feuille, ce qui n'est pas très pratique parfois. Pour éviter cet inconvénient, faire « Outils », « Options », « Calcul » puis mettre « Recalcul » à « sur ordre » et décocher « recalcul avant enregistrement ».

Dorénavant vous devrez appuyer sur la touche « F9 » pour recalculer votre feuille.

Dans une petite ville (imaginaire bien sûr !), il y a 2000 couples. Ils ont tous au moins un enfant et tous veulent une fille. Cependant ils ne veulent pas plus de quatre enfants.

Ils appliquent donc tous la stratégie suivante : ils feront un autre enfant tant qu'ils n'ont pas une fille ou quatre garçons.

Nous admettrons que chacun de ces couples peut avoir autant d'enfants qu'il le désire, qu'à chaque naissance, ils ont autant de chance d'avoir un garçon qu'une fille et qu'ils n'ont pas de jumeaux.

1. a. Donner toutes les compositions de familles possibles dans cette ville :

b. Pensez-vous que le nombre de filles dans cette ville sera supérieur, inférieur ou à peu près égal à celui des garçons (expliquez…) ?

2. Nous allons simuler ce qui se passe pour 2000 couples et chercher des réponses aux questions suivantes :

- Quelle sera la proportion de garçons et de filles dans cette ville ?
- Quel sera le nombre moyen d'enfants par famille dans cette ville ?

a. Dans la cellule A1 d'une feuille Excel, mettre la formule : = SI(ALEA()<0,5;"G";"F")

Que simule-t-on ainsi ?

b. Recopier cette formule jusqu'en D1. Qu'a-t-on simulé dans les cellules A1 à D1 ?

c. Dans la cellule F1, mettre la formule : =A1. Dans la cellule G1, mettre la formule : =SI(F1="G";B1;"")

Que fait la formule précédente ?

d. Recopier la formule de G1 en H1 et I1. Qu'a-t-on simulé dans les cellules F1 à I1 ?

e. Recopier la ligne 1 jusqu'à la ligne 2000.

f. Dans les cellules A2002 à A2007 mettre les titres suivants :

Garçons	
Filles	
Total	
Prop Garçons	(pour proportion de garçons)
Prop Filles	(pour proportion de filles)
Enf / famille	(pour nombre moyen d'enfants par famille)

g. Dans la cellule B2002, mettre la formule : =NB.SI(F1:I2000;"G")
Dans la cellule B2003, mettre la formule donnant le nombre total de filles dans les 2000 familles.
Dans la cellule B2004, mettre la formule donnant le nombre total d'enfants
Dans la cellule B2005, mettre la formule donnant la proportion de garçons
Dans la cellule B2006, mettre la formule donnant la proportion de filles dans la ville.
Observer les résultats en faisant plusieurs tests avec la touche « F9 ».
Les proportions de garçons et de filles semblent-elles égales ?

Dans la cellule B2007, mettre la formule donnant le nombre moyen d'enfants par couple.

Etude théorique

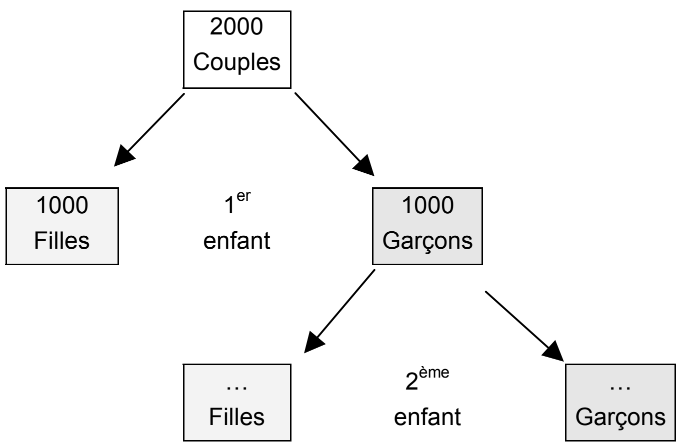
1. a. 2000 couples ont un premier enfant. A quel nombre théorique de naissances de filles peut-on s'attendre pour ce premier enfant ?

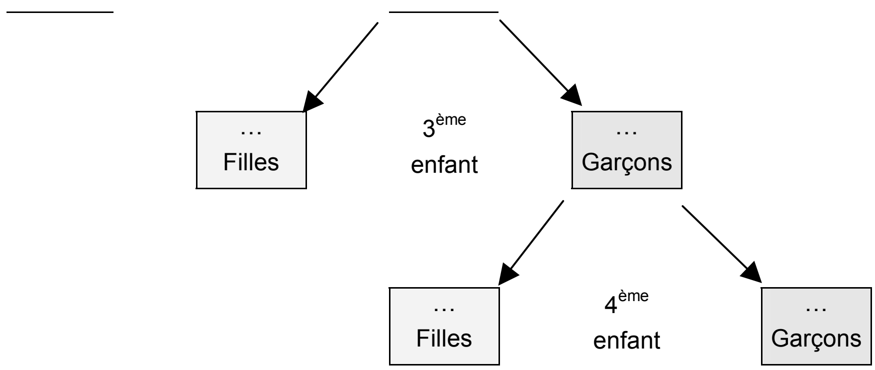
b. A quel nombre de couples avec un enfant peut-on s'attendre dans notre ville ?

c. 1000 couples ayant déjà un garçon ont un deuxième enfant. A quel nombre théorique de naissances de filles peut-on s'attendre pour ce 2ème enfant ?

d. quel nombre de couples ayant deux enfants peut-on s'attendre dans notre ville ?

2.a. Compléter le schéma suivant, indiquant la répartition théorique des naissances pour les 2000 couples de notre ville :





b. Pour 2000 couples, donner la répartition théorique attendue :

Enfants	F	GF	GGF	GGGF	GGGG
Effectifs					

c. Déterminer :

- le nombre théorique de filles pour les 2000 couples ;
- le nombre théorique de garçons pour les 2000 couples ;
- le nombre théorique moyen d'enfants pour les 2000 couples.

Comparer la simulation faite au 2. à ces résultats théoriques.

4. Refaire la simulation et l'étude théorique en supposant maintenant que les couples de la ville ont un autre enfant tant qu'ils n'ont pas une fille ou cinq garçons.

(Pour la simulation, dans le menu « Edition ; Déplacer ou copier une feuille », on peut cocher « Créer une copie » puis modifier la copie obtenue.)

1.6.34. Probabilité d'avoir le même anniversaire

1. Lister toutes les dates anniversaires des élèves de la classe. Y'en a-t-il deux identiques ? Pensez-vous que cela soit fréquent ?

2. Comme c'est difficile à calculer avec 30 élèves (ou plus) et 365 jours on va commencer avec plus simple : 4 élèves et 12 mois.

a. Relevez les effectifs des mois de naissance dans votre classe.

mois	1	2	3	4	5	6	7	8	9	10	11	12
effectif												

b. Prenez deux personnes au hasard, quelle est la probabilité qu'elles soient né le même mois ? D'abord avec le tableau précédent puis en général.

c. Quelle est la probabilité que 3 personnes distinctes soient néees le même mois ? Quelle est la probabilité qu'elles ne soient pas nées le même mois ? Mêmes questions avec 4.

4. Quelle est la probabilité que deux élèves ne soient pas nés le même jour dans une classe de 4 élèves ? Soient nés le même jour ?

5. Quelle est la probabilité que deux élèves ne soient pas nés le même jour dans une classe de 30 élèves ? Soient nés le même jour ?

6. Voici deux propositions d'algorithme pour modéliser cette situation.

a. Expliquez les significations de chaque ligne puis tester de manière à répondre à la question précédente.

Algorithme 1	Algorithme 2
Variables : n, i entiers ; q réel. Entrée : Saisir n. Initialisation : q prend la valeur 1. Traitement : Pour i variant de 1 à $n-1$, q prend la valeur $q\times\frac{365-i}{365}$. Sortie : Afficher $1-q$.	Variables : i entier ; p, q réel. Entrée : Saisir p. Initialisation : q prend la valeur 1, i la valeur 0. Traitement : Tant que $q > 1-p$, q prend la valeur $q\times\frac{365-i}{365}$, i prend la vameur $i+1$. Sortie : Afficher i.

b. Combien faut-il d'élèves dans une classe pour être sûr à 99% que deux élèves au moins ont même date anniversaire ?

1.6.35. Dans le grenier

En fouillant dans un vieux cahier d'écolier datant de 2018 j'ai retrouvé cet algorithme… Aussitôt, me précipitant sur mon Apple Iclad-5G++ je rentre le programme par synthèse vocale et je vois se dessiner les positions des points de coordonnées (i, x).

Algorithme 1	
Variables :	x, n, i, entiers ; q réel.
Entrée : Saisir n.	
Initialisation :	x et q prennent la valeur 0.
Traitement :	Pour i variant de 1 à n, Si alea(0…1)<0,5 alors x prend la valeur $x+1$, sinon x prend la valeur $x-1$. Fin si. Si $x=0$ alors q prend la valeur $q+1$. Fin si. Fin pour.
Sortie :	Afficher x. Afficher $1-q/n$.

1. Expliquer ce que fait ce programme et à quoi correspondent les variables x, n et q.

2. Compléter le tableau suivant :

n	100	1 000	10 000	100 000
$1-q/n$				

Quelle conclusion pouvez-vous tirer de ces résultats ?

3. a. Modifier l'algorithme pour garder une trace du nombre de fois où x prend une valeur donnée entre $-n$ et n.
b. Relever vos données et faire un graphique pour $n=50$: histogramme avec x en abscisse et n en ordonnée.
4. Un trader se dit que cet algorithme pourrait lui servir de modèle pour le cours d'une action en Bourse. Quelle est alors la signification pratique des diverses variables pour le trader ?
5. *Pour aller plus loin* : que se passe-t-il si on passe dans le plan puis dans des espaces de dimension supérieure ? On pourra consulter Wikipedia, article Marche aléatoire.

1.7. Fluctuations d'une fréquence

1.7.36. Taux anormal de cas de leucémie

Une petite ville des États-Unis a connu 9 cas de leucémie chez de jeunes garçons en l'espace de 10 années. Doit-on, comme l'ont alors affirmé les autorités, en accuser le hasard ?

Cet exemple montre les enjeux de la méthode statistique.

Woburn est une petite ville industrielle du Massachusetts, au Nord-Est des États-Unis. Du milieu à la fin des années 1970, la communauté locale s'émeut d'un grand nombre de leucémies infantiles survenant en particulier chez les garçons dans certains quartiers de la ville. Les familles se lancent alors dans l'exploration des causes et constatent la présence de décharges et de friches industrielles ainsi que l'existence de polluants. Dans un premier temps, les experts gouvernementaux concluent qu'il n'y a rien d'étrange. Mais les familles s'obstinent et saisissent leurs propres experts. Une étude statistique montre qu'il se passe sans doute quelque chose « d'étrange ».

Le tableau suivant résume les données statistiques concernant les garçons de moins de 15 ans, pour la période 1969-1979 (Source : *Massachusetts Department of Public Health*).

Population des garçons de moins de 15 ans à Woburn selon le recensement de 1970 : n	Nombre de cas de leucémie infantile observés chez les garçons à Woburn entre 1969 et 1979	Fréquence des leucémies aux Etats-Unis (garçons) : p
5 969	9	0,000 52

La question statistique qui se pose est de savoir si le hasard seul peut raisonnablement expliquer le nombre de leucémies observées chez les jeunes garçons de Woburn, considérés comme résultant d'un échantillon prélevé dans la population américaine.

La population des États-Unis étant très grande par rapport à celle de Woburn, on peut considérer que l'échantillon résulte d'un tirage avec remise et simuler des tirages de taille n avec le tableur.

Il est aisé de simuler sur le tableur 100 échantillons de taille $n = 5\,969$ prélevés au hasard dans une population de garçons où la probabilité de leucémie est $p = 0{,}00052$ (cas « normal ») en utilisant l'instruction :

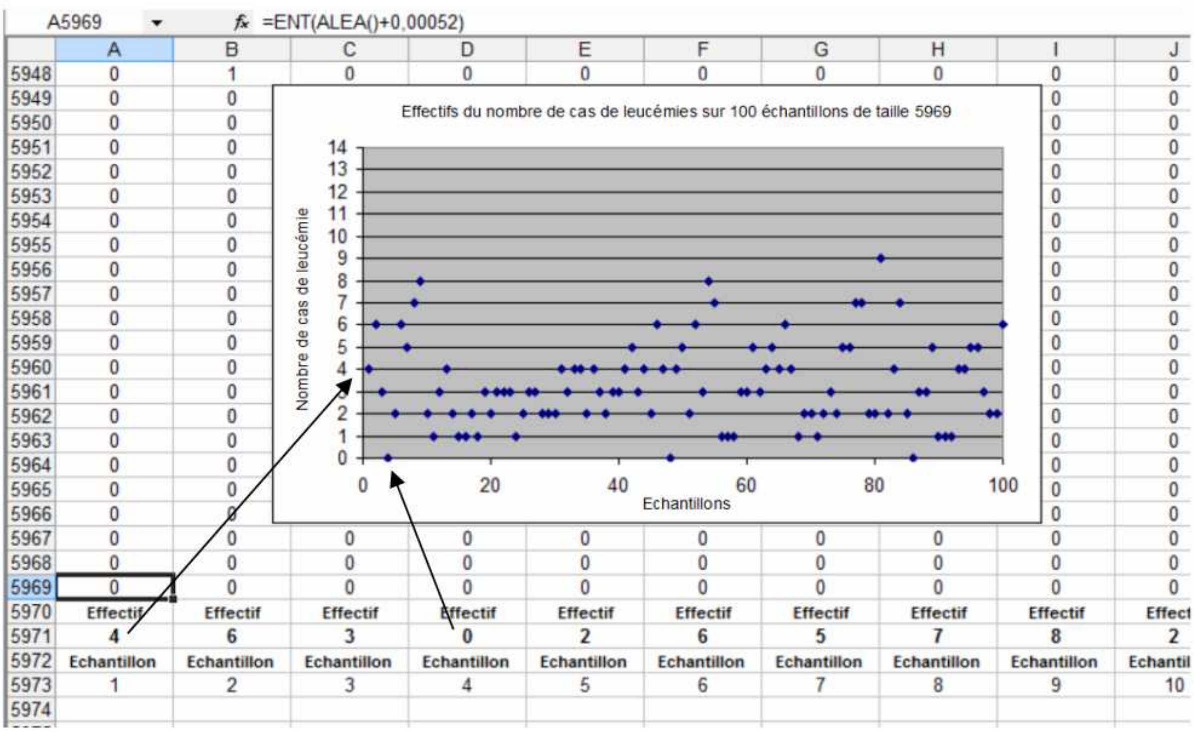
=ENT(ALEA()+0,00052).

(l'instruction =ALEA() génère un nombre aléatoire dans l'intervalle [0, 1[.)

L'instruction =ALEA()+0,00052 génère donc un nombre aléatoire dans l'intervalle [0,000 52 ; 1,000 52[. Ainsi, =ENT(ALEA()+0,000 52), où ENT désigne la partie entière, vaut la plupart du temps 0 (non malade) et vaut 1 (malade) avec la probabilité 0,000 52.

Sur chaque échantillon, en faisant la somme, on obtient le nombre de cas observés, sous l'hypothèse d'une probabilité « normale ».

Il est possible de représenter sur un graphique les 100 résultats observés sur les échantillons ainsi simulés.



Les simulations montrent que le nombre de cas observés à Woburn (9 cas) est extrêmement rare (de l'ordre de 1 % des simulations sur un grand nombre d'essais), sous l'hypothèse d'une probabilité « normale ».

Il est donc raisonnable de penser que le niveau très « significativement » élevé des leucémies infantiles observées chez les garçons de Woburn n'est pas dû au hasard.

Ce taux anormalement élevé de leucémies est officiellement confirmé par le Département de Santé Publique du Massachusetts en avril 1980. Les soupçons se portent alors sur la qualité de l'eau de la nappe phréatique qui, par des forages, alimente la ville. On découvre ainsi le syndrome du trichloréthylène.

1.7.37. Défauts de peinture

Dans une usine automobile, on contrôle les défauts de peinture de type « grains ponctuels sur le capot ».

Lorsque le processus est sous contrôle, on a 20 % de ce type de défauts.

Lors du contrôle aléatoire de 50 véhicules, on observe 26 % de défauts (13 sur 50). Faut-il s'inquiéter ?

Éléments de réponse

En supposant que la situation est sous contrôle, c'est-à-dire que la proportion de capots présentant ce défaut (minime) de peinture dans la production totale est $p = 0{,}20$, un échantillon aléatoire de 50 capots présentera une proportion de défauts comprise, dans plus de 95 % des cas, entre $0{,}20 - \frac{1}{\sqrt{50}}$, soit environ 6 %, et $0{,}20 + \frac{1}{\sqrt{50}}$, soit environ 34 %. Il n'y a donc pas lieu de considérer une observation de 26 % de défauts sur un échantillon de taille 50 comme « anormale ».

Ce type de contrôle de qualité a effectivement été pratiqué par un constructeur d'automobiles français. Il s'agissait de détecter une amélioration significative du procédé de peinture grâce à cet indicateur de défaut, quasiment invisible pour le client.

1.7.38. Naissances à pile ou face

On propose ici une séquence en trois étapes.

ÉTAPE 1 : énoncé

Les données statistiques suivantes ont été relevées : en 2000, dans le village de Xicun, en Chine, il est né 20 enfants, parmi lesquels 16 garçons ; dans la réserve indienne d'Aamjiwnaag, située au Canada à proximité d'industries chimiques, il est né entre 1999 et 2003, 132 enfants dont 46 garçons.

Ces observations sont-elles le fruit du hasard ?

Éléments de réponse

Un temps est laissé aux élèves pour se « débrouiller » avec ces chiffres puis saisir les pistes issues d'une « tempête de cerveaux ». Il faut s'attendre à des calculs de statistique descriptive (pourcentage, fréquence), à une comparaison avec une proportion de 50 %, et voir émerger la problématique du hasard.

La proportion de garçons à la naissance est habituellement estimée à 50 %. En réalité elle est, de façon étonnamment stable, de 51,2 %, mais on peut ici prendre sans inconvénient le modèle « d'une chance sur deux » pour analyser les données proposées.

Une première idée, pour se situer par rapport à un a priori de 50 %, consiste à calculer le pourcentage de garçons (ou de filles) dans chaque cas :

	Xicun	Aamjiwnaag
Pourcentage de garçons	80 %	34,8 %

Les valeurs trouvées sont loin du résultat « attendu ». Mais peut-être est-ce le hasard ? Pour préciser la réponse à la question posée, il faut étudier les propriétés du hasard en supposant, pour simplifier, qu'à la naissance on a une chance sur deux d'avoir une fille ou un garçon.

ÉTAPE 2 : expérimentation avec des pièces de monnaie

Lancer 20 fois une pièce de monnaie et noter le nombre de « pile ».

Recommencer l'expérience une dizaine de fois ou regrouper les résultats obtenus dans la classe.

Comment peut-on utiliser ces expériences pour commenter les statistiques de Xicun ?

Pourquoi l'expérimentation avec des pièces ne permet-elle pas de répondre complètement au problème posé ?

Éléments de réponse

On peut décider que « pile » correspond à « garçon » et « face » à « fille » (ou le contraire). On constate que pratiquement aucune expérience ne donne 16 piles (ou plus), mais que cela peut se produire.

Le résultat chinois est très étonnant et laisse plutôt penser que « la pièce est truquée ».

Remarque pour le professeur

Le nombre de « pile » peut être considéré comme correspondant à la réalisation d'une variable aléatoire de loi binomiale de paramètres $n = 20$ et $p = 0,5$. La probabilité, sur 20 lancers, d'avoir un nombre de « pile » inférieur ou égal à 15 est environ 0,994.

Ce résultat peut s'obtenir sur un tableur par l'instruction : =LOI.BINOMIALE(15;20;0,5;VRAI). C'est-à-dire que l'observation d'un résultat comparable à celui du village chinois a 0,6 % de chances de se produire (quand on suppose le hasard « une chance sur deux » à l'œuvre).

Pour Aamjiwnaag, il faudrait lancer la pièce un trop grand nombre de fois.

Faire constater qu'un plus grand nombre d'expériences permettrait de s'assurer des observations faites pour 20 lancers, permet alors de proposer un TP de simulation sur tableur-grapheur (ou faire une vidéo projection devant la classe ou encore distribuer le fichier sur des ordinateurs).

Cette première expérimentation informatique permet d'étudier ensuite des échantillons de 132 lancers en utilisant le même protocole de simulation.

ÉTAPE 3 : travaux pratiques sur tableur

Un énoncé possible

Simulation d'un lancer de pile ou face.

a. Entrer en cellule A1 la formule =ALEA() (avec des parenthèses vides) puis appuyer de nombreuses fois sur la touche F9.

Compléter la phrase : « La formule =ALEA() affiche un nombre décimal tiré au hasard entre .... et .... ».

b. Compléter la phrase : « La formule =ALEA()+0,5 affiche un nombre décimal tiré au hasard entre .... et .... ».

Vérifiez votre réponse en entrant la formule =ALEA()+0,5 en cellule A1.

c. En complétant la phrase suivante, trouver une règle permettant de simuler un tirage à pile ou face à l'aide de la fonction ALEA :

« Si en utilisant la formule =ALEA()+0,5 le tableur affiche un nombre de l'intervalle [0 ,5 ; ....[, on considère que la pièce est tombée sur face, si au contraire le tableur affiche un nombre de l'intervalle [........ , ...[, on considère que l'on a obtenu pile. ».

d. On souhaite améliorer la simulation en utilisant la fonction ENT qui affiche la partie devant la virgule d'un nombre positif (partie entière).

Entrer en A1 la formule =ENT(ALEA()+0,5). Faire plusieurs fois F9.

Quels résultats obtient-on ? Quelle est leur signification dans le jeu de pile ou face ?

Simulation de 20 lancers de pile ou face.

Recopier le contenu de la cellule A1 jusqu'en A20 (pointeur de la souris en forme de croix noire).

En A21 entrer la formule =SOMME(A1:A20)/20 . Que calcule cette formule ?

Simulation de 100 expériences.

Sélectionner les cellules de A1 à A21 puis recopier vers la droite jusqu'en colonne CV.

Sélectionner la ligne 21 (en cliquant sur la tête de ligne) puis cliquer sur l'icône de l'assistant graphique et demander un « nuage de points ».

Faire de nombreuses fois F9.

D'après vos observations, sur 20 lancers, une fréquence de « pile » égale ou supérieure à 0,8 :

- ☐ ne se produit jamais ;
- ☐ se produit environ 20 fois sur 100 ;
- ☐ se produit environ 10 fois sur 100 ;
- ☐ se produit environ 1 fois sur 100.

Que pouvez-vous déduire de ces simulations à propos des naissances à Xicun en 2000 ?

Cas de 132 naissances.

a. Sur une autre feuille de calcul, simuler 100 expériences de 132 lancers de pile ou face et représenter les fréquences de « pile » comme précédemment.

D'après le graphique, donner un intervalle autour de 0,5 dans lequel se trouve la « grande majorité » des points.

b. Que pouvez-vous déduire de ces simulations à propos des naissances à Aamjiwnaag ?

Éléments de réponse

1. a. « La formule =ALEA() affiche un nombre décimal tiré au hasard entre 0 et 1 ».

b. « La formule =ALEA()+0.5 affiche un nombre décimal tiré au hasard entre 0,5 et 1,5 ».

c. « Si en utilisant la formule =ALEA()+0.5 le tableur affiche un nombre de l'intervalle [0,5 ; 1[on considère que la pièce est tombée sur face, si au contraire le tableur affiche un nombre de l'intervalle [1 ; 1,5[on considère que l'on a obtenu pile. »

d. On obtient comme résultats 0 ou 1. Le résultat 0 correspond à face et 1 correspond à pile.

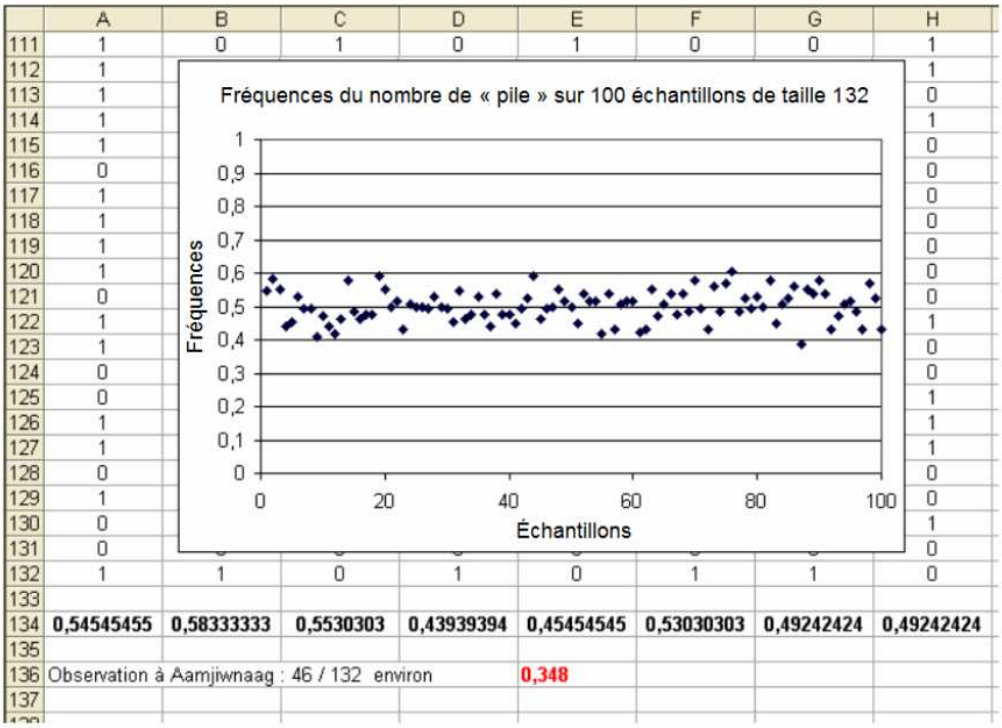
2. La formule calcule la fréquence de « pile » sur les 20 lancers.

3. D'après les observations, sur 20 lancers, une fréquence de « pile » égale ou supérieur à 0,8

☑ se produit environ 1 fois sur 100.

4. On peut en déduire que les statistiques de Xicun sont très étonnantes. Il est peu probable qu'elles s'expliquent par le seul hasard. C'est une « alerte » qui doit inciter à rechercher des causes extérieures.

5. a. La « grande majorité » des points sont dans l'intervalle [0,4 ; 0,6].



b. Le résultat 0,348 observé à Aamjiwnaag ne se trouve pas dans l'intervalle [0,4 ; 0,6].

Il est peu probable, qu'il s'explique par le seul hasard. C'est une « alerte » qui doit inciter à rechercher des causes extérieures.

Commentaires

Il est important de préciser que les « réponses » apportées par l'activité précédente ne sont que des réponses « statistiques ». Les résultats observés sur les naissances à Xicun et Aamjiwnaag sont « bizarres » (et préoccupants). Rien de plus ne peut être dit quant aux causes, mais ces résultats doivent inciter à enquêter. C'est une obligation morale, car on a établi une « preuve statistique », rationnelle, qu'il se passe sans doute quelque chose d'inhabituel.

Pour le cas de Xicun, la cause probable est l'acquisition dans ce village (en 1999) d'une machine à ultra-sons bon marché, permettant aux médecins de déterminer le sexe du fœtus.

(Source : *Washington Post* du 29 mai 2001.)

Dans le cas d'Aamjiwnaag, une enquête sanitaire est menée. En effet, depuis Seveso, le rôle de certains polluants sur les déséquilibres du sex-ratio est connu.

(Sources : *Science et Vie* février 2006 – *Environmenthal Health Perspectives* octobre 2005, article en anglais en ligne.)

1.7.39. Segment aléatoire

Cette activité illustre l'approche « fréquentiste » d'une probabilité, c'est-à-dire l'observation de la « stabilisation relative des fréquences quand n augmente » vers la probabilité de l'événement.

Deux points A et B sont pris « au hasard » sur un segment de longueur 1.

Ceci peut être réalisé grâce à la fonction « random » de la calculatrice qui donne « au hasard » un nombre compris entre 0 et 1, qui sera l'abscisse du point.

Quelle est la probabilité de l'événement : « la longueur AB est supérieure à 0,5 » ?

Éléments de réponse

– Expérimentation avec une calculatrice :

La différence rand – rand correspond, au signe près, à la distance *AB*.

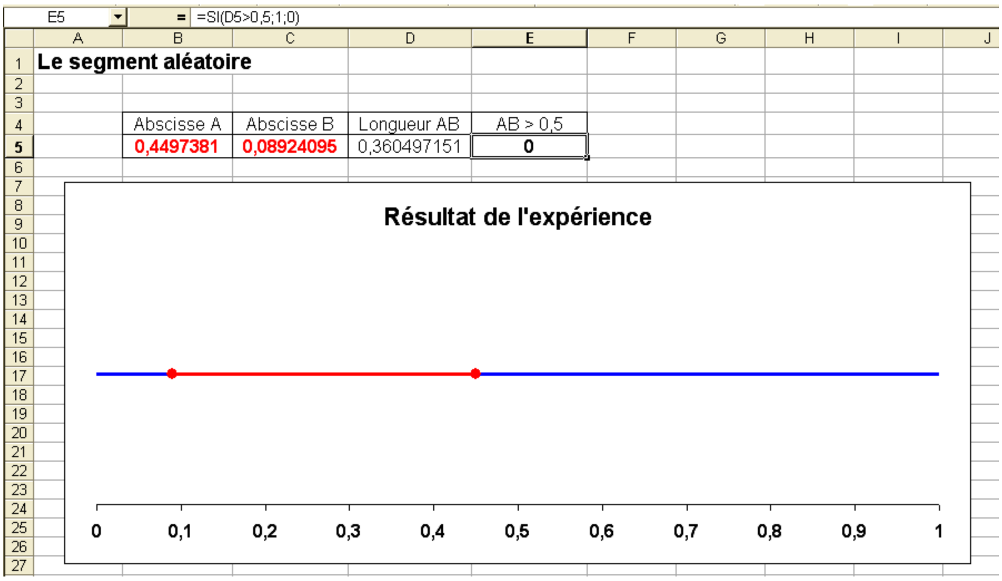
Il suffit d'appuyer 10 fois sur Enter pour simuler 10 expériences.

Les résultats de la classe peuvent ensuite être mutualisés.

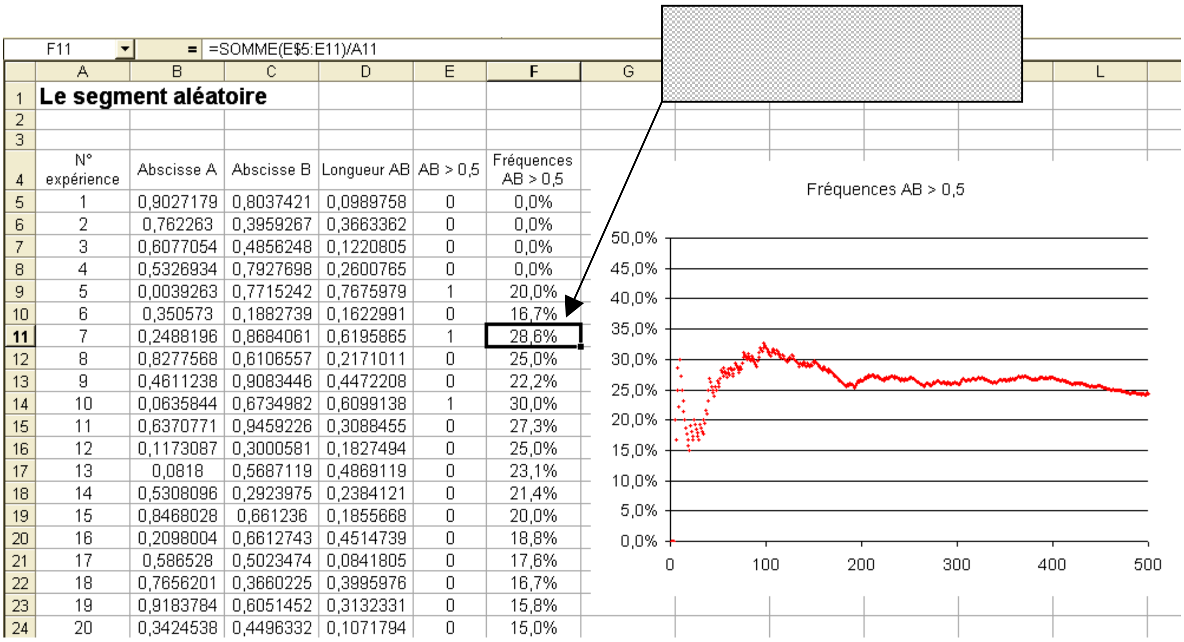
– Visualisation sur un tableur :

La simulation de l'expérience est simple à mettre en place sur un tableur : sur l'image d'écran ci-dessous, est entrée en B5 et en C5 la formule =ALEA(). En D5 la formule =ABS(B5−C5) fournit la distance AB.

En entrant en E5 la formule =SI(D5>0,5;1;0) la valeur 1 est obtenue lorsque l'événement « $AB > 0,5$ » est réalisé et la valeur 0 est obtenue lorsqu'il ne l'est pas.



Par recopie, l'expérience simulée peut être répétée et constituer un échantillon (de taille par exemple 500) des résultats de cette expérience. En faisant F9, on multiplie les observations.



N° expérience	Abscisse A	Abscisse B	Longueur AB	AB > 0,5	Fréquences AB > 0,5
1	0,9027179	0,8037421	0,0989758	0	0,0%
2	0,762263	0,3959267	0,3663362	0	0,0%
3	0,6077054	0,4856248	0,1220805	0	0,0%
4	0,5326934	0,7927698	0,2600765	0	0,0%
5	0,0039263	0,7715242	0,7675979	1	20,0%
6	0,350573	0,1882739	0,1622991	0	16,7%
7	0,2488196	0,8684061	0,6195865	1	28,6%
8	0,8277568	0,6106557	0,2171011	0	25,0%
9	0,4611238	0,9083446	0,4472208	0	22,2%
10	0,0635844	0,6734982	0,6099138	1	30,0%
11	0,6370771	0,9459226	0,3088455	0	27,3%
12	0,1173087	0,3000581	0,1827494	0	25,0%
13	0,0818	0,5687119	0,4869119	0	23,1%
14	0,5308096	0,2923975	0,2384121	0	21,4%
15	0,8468028	0,661236	0,1855668	0	20,0%
16	0,2098004	0,6612743	0,4514739	0	18,8%
17	0,586528	0,5023474	0,0841805	0	17,6%
18	0,7656201	0,3660225	0,3995976	0	16,7%
19	0,9183784	0,6051452	0,3132331	0	15,8%
20	0,3424538	0,4496332	0,1071794	0	15,0%

La fréquence l'événement « $AB > 0,5$ » se stabilise autour de 25 %.

Commentaire

L'avantage de cette situation est que la probabilité de l'événement considéré, qui est égale à 0,25, n'est pas « intuitive ». Il est d'ailleurs intéressant de demander aux élèves une évaluation « a priori » de la réponse, avant d'expérimenter

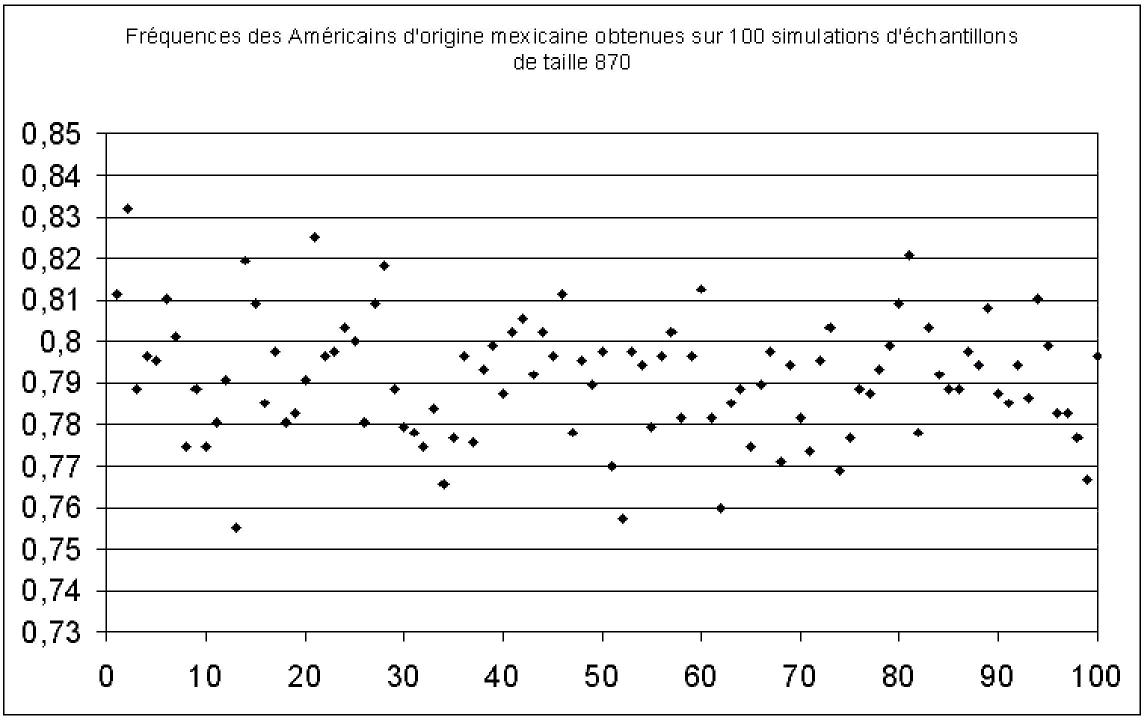
1.7.40. Contester un jugement

(Source : *Prove it with figures* – H. Zeisel et D Kaye.)

En Novembre 1976 dans un comté du sud du Texas, Rodrigo Partida était condamné à huit ans de prison. Il attaqua ce jugement au motif que la désignation des jurés de ce comté était discriminante à l'égard des Américains d'origine mexicaine. Alors que 79,1% de la population de ce comté était d'origine mexicaine, sur les 870 personnes convoqués pour être jurés lors d'une certaine période de référence, il n'y eut que 339 personnes d'origine mexicaine.

1. Quelle est la fréquence des jurés d'origine mexicaine observée dans ce comté du Texas ?

2. La simulation sur un tableur du prélèvement d'échantillons aléatoires de taille n = 870 dans une population où la fréquence des habitants d'origine mexicaine est p = 0,791. Les fréquences des habitants d'origine mexicaine observées sur 100 échantillons simulés sont représentées ci-dessous.



a. Calculer les bornes de l'intervalle $[\,p-\frac{1}{\sqrt{n}}\ ;\ p+\frac{1}{\sqrt{n}}\,]$. (Arrondir à 10^{-2}).

b. Quel est le pourcentage des simulations fournissant une fréquence en dehors de l'intervalle précédent ?

3. Sur les simulations, est-il arrivé au hasard de fournir une fréquence d'habitants d'origine mexicaine comparable à celle des jurés d'origine mexicaine observée dans ce comté du Texas ?

4. Comment expliquez-vous cette situation ?

Éléments de réponse

1. La fréquence observée des jurés d'origine mexicaine est $f=\frac{339}{870}$ soit environ 0,39.

2. a. On a $[\,0{,}791-\frac{1}{\sqrt{870}}\ ;\ 0{,}791+\frac{1}{\sqrt{870}}\,]$ c'est-à-dire environ [0,76 ; 0,82].

b. Sur le graphique des 100 simulations, 4 points sont en dehors de l'intervalle précédent, soit 4 % des cas.

3. Non. La fréquence observée 0,39 est très loin des valeurs obtenues sur les simulations.

4. La constitution des jurys n'est sans doute pas totalement aléatoire.

Commentaires

Les données étudiées constituent une « preuve statistique » du fait que la constitution de ces jurys n'est pas totalement aléatoire, c'est-à-dire que ceux ci ne sont pas « représentatifs » de la population, du point de vue du caractère hispanique. Les calculs précédents montrent qu'il n'est pas possible de considérer que les jurys résultent d'un tirage au sort où chaque élément de la population a les mêmes chances d'être choisi. Mais c'est tout ce que l'on peut dire et, en particulier, il n'est pas possible de se prononcer sur les causes et porter des accusations de discrimination raciale. L'étude statistique précédente doit inciter à enquêter sur les conditions de constitution des jurys. Le constat pourra alors être fait que pour être juré on doit maîtriser la langue anglaise (écrite et parlée), ce qui n'est pas le cas de la majorité de la population d'origine hispanique, que pendant les 11 années correspondant à l'étude, la proportion des hispaniques dans la population a évolué, et que la proportion d'hispaniques dans les jurys a également évolué au cours de ces 11 années.

1.7.41. Pile ou face et contrôle de qualité

Lors de certains contrôles de qualité en cours de fabrication dans l'industrie (diamètre d'une pièce par exemple), des cartes de contrôle reposent sur la procédure suivante : la moyenne de la cote surveillée (le diamètre de la pièce par exemple) est calculée sur des échantillons aléatoires prélevés régulièrement en fin de fabrication. Ces moyennes sont reportées sur une carte de contrôle (comme ci-dessous). Si une série de sept points consécutifs se trouve du même côté de la « moyenne attendue » (la norme visée), le processus doit être surveillé pour déceler une éventuelle « dérive » dans le processus de fabrication.L'explication du choix du nombre 7 se trouve dans la résolution du problème de probabilités suivant : une pièce de monnaie équilibrée est lancée 7 fois, quelle est la probabilité de l'événement A : « la pièce est tombée 7 fois sur pile » ?

1. Estimer la valeur de cette probabilité à l'aide de simulations sur un tableur.

2. Lancer 3 fois de suite une pièce de monnaie équilibrée.

a. Dessiner un arbre figurant tous les résultats possibles de l'expérience.

b. À l'aide de l'arbre précédent, calculer la probabilité de l'événement « la pièce est tombée trois fois sur pile ».

3. Par analogie avec le cas de 3 lancers, donner la probabilité de l'événement A et comparer avec l'estimation de la question 1.

Éléments de réponse

1. Une feuille de calcul telle que celle montrée ci-après peut être constituée. Selon les circonstances, cette feuille de calcul peut-être totalement ou partiellement fabriquée par les élèves, en salle informatique ou en utilisant un vidéo projecteur.

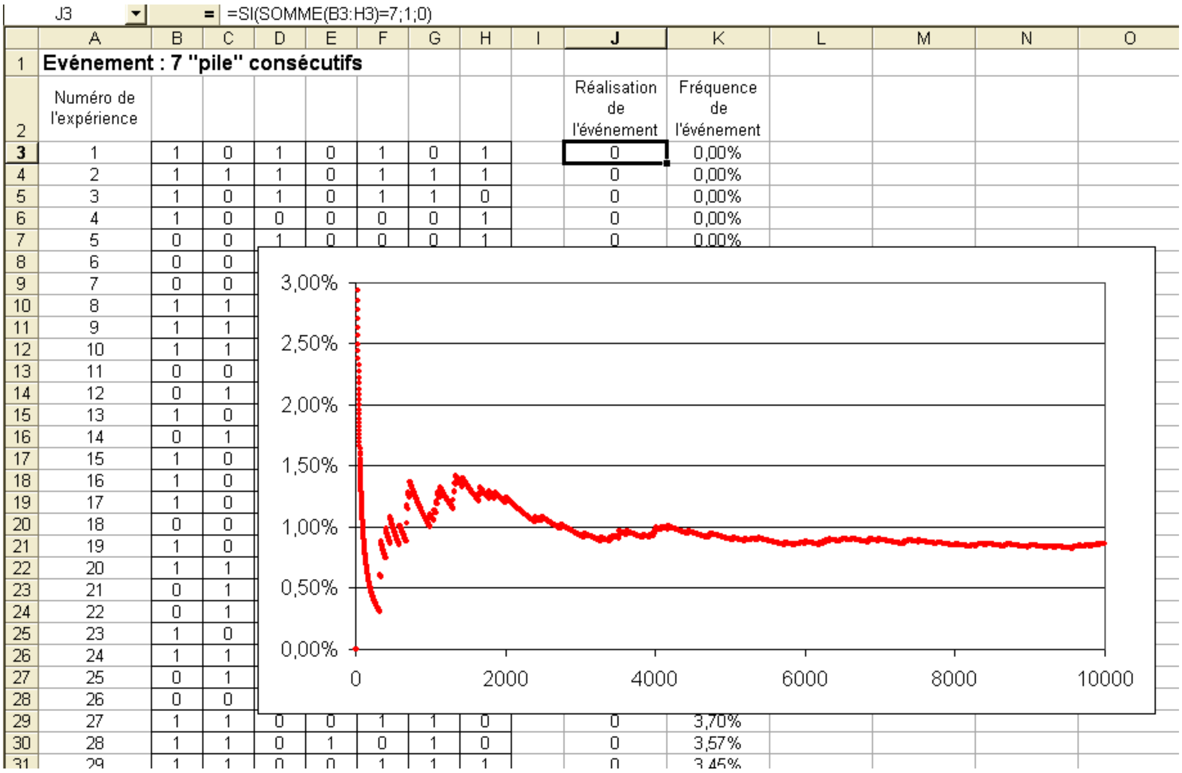
Pour constituer cette feuille de calcul, 7 lancers à pile ou face sont simulés en introduisant en cellule B3 la formule =ENT(ALEA()+0,5) qui a été recopiée vers la droite.

La réalisation ou non de l'événement A est codée par 1 ou 0 avec l'introduction en cellule J3 de la formule =SI(SOMME(B3:H3)=7;1;0) .

Le calcul des fréquences cumulées de l'événement A est obtenu en introduisant en K3 la formule =J3/A3 puis en K4 la formule =SOMME(J$3:J4)/A4 .

Il faut sélectionner la ligne 4 puis recopier vers le bas, par exemple jusqu'à la ligne 10 002 pour obtenir la fréquence de l'événement A sur 10 000 expériences.

La touche F9 permet de répéter les 10 000 expériences et de faire constater que la probabilité de l'événement A est un peu inférieure à 1 %.



2. b. La probabilité de l'événement PPP est $\left(\frac{1}{2}\right)^3 = 0{,}125$.

3. On a $P(A) = \left(\frac{1}{2}\right)^7$, soit environ 0,008 c'est-à-dire 0,8 %.

Ce résultat est conforme aux observations des simulations.

1.7.42. Premier tour des présidentielles 2002

Voici un extrait d'article, publié dans le journal « Le Monde » par le statisticien Michel Lejeune, après le premier tour de l'élection présidentielle de 2002.

« Pour les rares scientifiques qui savent comment sont produites les estimations, il était clair que l'écart des intentions de vote entre les candidats Le Pen et Jospin rendait tout à fait plausible le scénario qui s'est réalisé. En effet, certains des derniers sondages indiquaient 18 % pour Jospin et 14 % pour Le Pen. Si l'on se réfère à un sondage qui serait effectué dans des conditions idéales [...], on obtient sur de tels pourcentages une incertitude de plus ou moins 3 % étant donné la taille de l'échantillon [...]. »

1. Si l'on tient compte de l'incertitude liée au sondage, entre quels pourcentages pourraient se situer réellement (à 95 % de confiance) les deux candidats lorsque le sondage donne 18 % pour l'un et 14 % pour l'autre ?

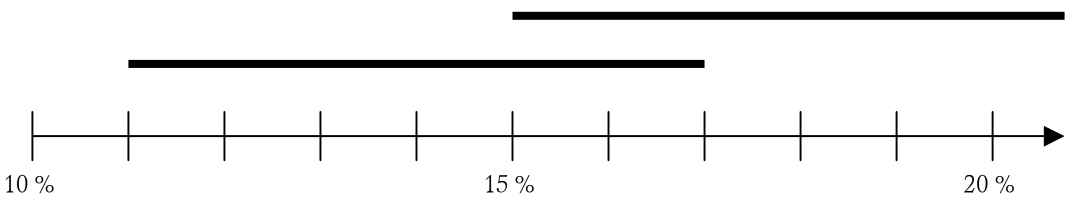
2. Représenter sur un même graphique les deux « fourchettes » calculées à la question précédente. Peut-on prévoir l'ordre des candidats ?

3. Au premier tour de l'élection présidentielle de 2002, L. Jospin a obtenu 16,18 % des voix et J.-M. Le Pen 16,86 %.

Expliquer la phrase « l'écart des intentions de vote entre les candidats Le Pen et Jospin rendait tout à fait plausible le scénario qui s'est réalisé ».

Éléments de réponse

1. Pour L. Jospin, entre 15 % et 21 %. Pour J.-M. Le Pen, entre 11 % et 17 %.
2. Un dessin possible.



Si on utilise ces fourchettes, on ne peut pas prévoir l'ordre des candidats car elles ont une partie commune.

3. La phrase correspond au fait que les pourcentages obtenus à l'élection sont situés dans les fourchettes du sondage.

Commentaires

Voici un exemple où quelques notions mathématiques de base (pourcentage, représentation des nombres, intervalle, intersection) sont nécessaires à la bonne compréhension d'un article de presse de la rubrique société ou politique.

La lecture du texte est une des difficultés de l'exercice, mais montre aussi l'intérêt de la situation.

L'aspect fluctuation d'échantillonnage (plus ou moins 3 % sur un échantillon de taille mille, c'est-à-dire $\frac{1}{\sqrt{1000}} \approx 0{,}03$) n'est pas pris en compte dans cet exercice, qui se veut élémentaire. La « confiance » est donnée au statisticien. Il est possible d'expérimenter la pertinence de ce 3 % (environ 95 % de confiance) par simulation.

1.7.43. Surréservation

Deux énoncés sont proposés. Dans le premier, l'élève est guidé dans la démarche de résolution et dans le second il peut faire preuve de son autonomie et de sa prise d'initiative dans la résolution d'un problème.

Énoncé 1

Une compagnie aérienne dispose d'un avion de 100 places et vend 107 réservations.

L'objectif est d'évaluer la probabilité de surréservation de cette compagnie, autrement dit le risque que plus de 100 passagers se présentent à l'embarquement.

1. On suppose que toute personne réservant une place d'avion a une chance sur 10 de ne pas se présenter à l'embarquement. Réaliser une simulation du nombre de personnes se présentant à l'embarquement d'un vol de 100 places pour 107 réservations, sur un échantillon aléatoire obtenu à l'aide d'un tableur.

Pour cela, dans une feuille de calcul du tableur :

* saisir « =ENT(ALEA()+0,9) » dans la cellule A1 et recopier cette formule vers la droite jusqu'en DC1 pour obtenir 107 réalisations ,

* saisir « =SOMME(A1:DC1) » dans la cellule DD1.

2. Réaliser une simulation du nombre de personnes se présentant à l'embarquement de 1 000 vols de 100 places pour 107 réservations à chaque vol.

3. Déterminer, pour cette simulation de 1 000 vols, la proportion des cas où l'effectif des passagers se présentant à l'embarquement est supérieur à 100. Pour cela :

* dans une cellule de votre choix, utiliser la formule « =NB.SI(DD1:DD1000;">100") »,

* dans une cellule de votre choix, en déduire la fréquence demandée.

4. a. En utilisant la touche F9, réaliser plusieurs simulations, puis évaluer la probabilité que plus de 100 personnes se présentent à l'embarquement.

b. Évaluer, en pourcentage, le risque de surréservation pour la compagnie aérienne.

Éléments de réponse

1. Tout type de tableur convient, par exemple Excel ou OpenOffice Calc.

Il suffit d'inscrire la formule « =ENT(ALEA()+0,9) » dans la cellule A1, de la recopier horizontalement pour qu'elle soit calculée 107 fois, puis d'effectuer la somme.

L'élève doit comprendre que lorsque la formule affiche 1, le passager se présente à l'embarquement et lorsqu'elle affiche 0, le passager ne se présente pas.

2. Il s'agit de sélectionner les cellules de la simulation de la question 1. puis de recopier vers le bas.

3. L'instruction NB.SI fournit un effectif. On s'attache à la différence qui est faite entre effectif et fréquence.

4. L'élève doit comprendre que la probabilité de surréservation est la valeur autour de laquelle fluctuent les fréquences lorsqu'on appuie sur la touche F9.

On accepte toute évaluation comprise entre 0,06 et 0,10.

Signalons pour le professeur que le calcul de cette probabilité peut s'effectuer par la formule « =1-LOI.BINOMIALE(100;107;0,9;VRAI) » qui donne comme réponse environ 0,08 (ou 8 %).

Énoncé 2

On suppose qu'une personne réservant une place d'avion a une chance sur 10 de ne pas se présenter à l'embarquement. Une compagnie dispose d'un avion de 100 places et vend 107 réservations. Le but de l'exercice est d'évaluer la probabilité de surréservation.

1. Sur un tableur, réaliser une simulation du nombre de personnes se présentant à l'embarquement lorsqu'il y a 107 réservations.

On peut utiliser la formule =ENT(ALEA()+0,9).

2. Sur un tableur, réaliser une simulation sur un échantillon de taille 1 000 de l'expérience aléatoire précédente et déterminer, pour cette simulation, la fréquence des cas où plus de 100 personnes se présentent à l'embarquement.

On peut utiliser la formule =NB.SI(plage ;" > 100").

3. À l'aide des simulations réalisées, est-il possible d'évaluer le risque de surréservation que prend la compagnie ? On peut utiliser la touche F9.