Onze fiches de statistique

L'objectif du chapitre de statistique, en classe de seconde, est de sensibiliser les élèves à l'aléatoire et plus particulièrement de leur faire prendre conscience de l'existence de la fluctuation d'échantillonnage dans des cas simples : cette fluctuation est ici celle de distributions des fréquences entre des séries obtenues par répétition d'expériences identiques. La moyenne, la médiane, le maximum, le minimum se calculant à partir de cette distribution des fréquences sont aussi soumis à la fluctuation d'échantillonnage.

Les élèves devront se constituer un bagage d'expériences de référence : lancers de pièces, de dés, tirages de boules dans une urne, obtention de listes de nombres au hasard à l'aide d'une calculatrice ; ces expériences seront des éléments fondateurs de leur perception de l'aléatoire.

On simulera les expériences de référence et on pourra ainsi étudier de longues séries ; on simulera aussi de nouvelles situations à l'aide de listes de chiffres au hasard. Les outils de simulation sont puissants, mais pour les élèves une première étape importante est d'établir un lien entre l'expérience et une simulation de cette expérience : quand les thèmes s'y prêtent (et c'est le plus souvent le cas), la classe pourra être partagée en deux groupes : dans un groupe, les élèves feront effectivement des expériences, dans l'autre les élèves simuleront l'expérience.

Le langage utilisé est celui de la vie courante : égalité des chances, pièce ou dé équilibré, choix au hasard dans un ensemble fini, etc. Une série statistique dont les termes sont les résultats issus de la réplication d'une même expérience sera appelée un échantillon de cette expérience (ou plus simplement échantillon si cela ne prête pas à confusion) ; par exemple : les résultats de n lancers d'un dé, les couleurs de n boules tirées au hasard dans une urne, l'opinion de n personnes choisies au hasard dans une population bien définie (l'expérience que l'on répète est ici le recueil de l'opinion d'une personne choisie au hasard), la taille de n personnes choisies au hasard dans une population bien définie (l'expérience que l'on répète est ici la mesure de la taille d'une personne choisie au hasard), etc.

Les situations traitées mènent à l'observation de résultats qui appellent une explication ; par exemple, si on lance deux pièces équilibrées et qu'on regarde le nombre de *pile* obtenus, la distribution des fréquences de 0,1,2 semble voisine de (1/4,1/2,1/4) ; « pourquoi en est-il ainsi ? » est une question à adresser aux mathématiques, qui justifie de formaliser le langage utilisé et d'introduire à la théorie des probabilités. Le langage probabiliste permettra de se dégager de la polysémie des termes employés dans la vie courante (tels ceux de *chance* ou de *hasard*) ; des éléments simples de théorie des probabilités (introduits en classe de première) rendront intelligibles de nombreuses observations liées aux phénomènes aléatoires.

Le choix pédagogique de ce programme est d'aller de l'observation vers la conceptualisation et non d'introduire d'abord le langage probabiliste pour constater ensuite que tout se passe comme le prévoit cette théorie ; pour des enfants ou des adolescents, le langage des probabilités ne préexiste pas à la perception de l'aléatoire.

En seconde, les listes de chiffres au hasard produites par les ordinateurs sont présentées comme toute autre liste relevant de la répétition d'expériences identiques dont les issues ont des chances égales d'apparaître. Pour les élèves, lancer des pièces, des dés ou obtenir des listes de nombres au hasard avec une calculatrice sont des expériences aléatoires de référence. On simulera aussi quelques expériences simples : il est important que l'élève conçoive lui-même de telles simulations (sans recherche systématique de subtilité algorithmique). En première, on précisera que pour simuler une expérience, il faut d'abord en établir un modèle puis simuler ce modèle.

Les fiches qui suivent couvrent la partie commune du programme et l'ensemble des thèmes proposés : il n'est donc pas question de toutes les traiter ! Elles indiquent, au-delà du contenu du programme, l'esprit dans lequel on peut le travailler. Ce ne sont pas des « fiches élèves », néanmoins quelques exemples de questions les concernant sont proposées (elles apparaissent en bleu). En en-tête de chaque fiche, on indique quelques-uns des éléments du programme abordés à l'occasion du thème considéré. Des aperçus théoriques à l'usage des enseignants sont donnés pour certains thèmes. Des résultats de simulation sont systématiquement fournis, que les enseignants pourront étudier avec leurs élèves, après que ceux-ci aient réalisé ou simulé quelques expériences. Ce sera l'occasion de manipuler des distributions de fréquences (c'est à dire des vecteurs ayant un nombre éventuellement assez grand de composantes) et de comprendre que pour regrouper des distributions de fréquences, il convient de faire des moyennes pondérées.

Le programme de statistique comporte dans son libellé deux parties : une première partie purement descriptive et la seconde partie qui parle de fluctuation d'échantillonnage ; comme on pourra le voir avec ces fiches, ces deux parties n'ont pas à être dissociées, les outils de description étant naturellement utilisés pour les séries statistiques, qu'elles soient réelles ou simulées.

Quelques adresses utiles :

Un logiciel pour la formation (articles, cours, simulations dynamique):
http://www.inrialpes.fr/sel/

Des simulations dynamiques des thèmes statistiques de seconde :
http://perso.wanadoo.fr/jpq

Des simulations dynamiques de quelques concepts de base en statistique :
http://www.kuleuven.ac.be/ucs/java/index.htm

Des simulations des thèmes statistiques de seconde avec le logiciel GEOPLANW2 et GEOSPACW:
http://www.irem.univ-mrs.fr

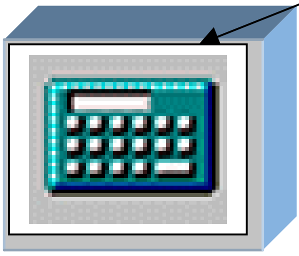
Liste de chiffres au hasard

- Liste de chiffres au hasard
- Utilisation pour des simulations

1) Chiffres au hasard

Les calculatrices de poche et les ordinateurs possèdent une instruction dont le nom est *random*, *RDM*, *rand* ou *alea*. L'appel de cette touche, disons la touche *random*, fournit un nombre décimal compris entre 0 et 1 ; la partie décimale est une liste de k chiffres au hasard ; en faisant n appels de la touche *random*, on obtient une liste de kn chiffres au hasard. Une telle liste pourrait tout aussi bien résulter de kn tirages au hasard et avec remise dans une urne contenant en quantités égales des boules marquées 0 ou 1 ou 2,…, ou 9 ; au niveau de chaque choix, les chiffres ont des chances égales d'apparaître et les choix sont indépendants. Pour les élèves de seconde, la touche *random* fournit des chiffres au hasard de même que des lancers d'un dé équilibré fournissent des chiffres au hasard parmi les chiffres 1,2,3,4,5,6, ou que des lancers d'une pièce équilibrée produisent des listes de chiffres 0 ou 1 au hasard.

Certaines calculatrices affichant k chiffres après la virgule n'affichent cependant pas le dernier ou les derniers chiffres si ceux-ci valent 0 : il convient en ce cas de toujours compléter l'écriture par des 0 jusqu'à avoir k chiffres après la virgule.



RAND

0.95948630
0.74391960
0.98399939
0.55345500
0.85701411
0.43616244
0.86793361
0.55703248
0.03204353
0.35857880
0.82833289
0.40918336
0.18254121

959486307439196098399939553455008570141143616244867933615570324803204353358578808283328940918336182541213142052857217421943549095959068 0 …

2) Simulation de lancers de pièces ou de dés.

Pour simuler des lancers d'une pièce équilibrée avec une touche *random*, on codera les chiffres ; voici deux exemples de codages permettant de transformer une liste de chiffres au hasard en liste de lettres P, F choisies au hasard.

0	1	2	3	4	5	6	7	8	9
F	P	F	P	F	P	F	P	F	P

0	1	2	3	4	5	6	7	8	9
P	P	P	P	P	F	F	F	F	F

Si on voulait utiliser une *urne à chiffres*, (c'est à dire une urne contenant en quantités égales des boules marquées 0 ou 1 ou 2….ou 9) pour simuler des lancers d'un dé équilibré, on pourrait retirer de l'urne les boules marquées 0,7,8,9. S'il est impossible de retirer ces boules, on fera des tirages avec remise sans tenir compte des boules 0,7,8,9. Et on fait de même avec la touche *random* : on ne tient pas compte des chiffres qui n'existent pas dans le lancer d'un dé. Ainsi, a partir d'une liste de chiffres au hasard, on simule une série plus petite de lancers de dés :

9	5	6	4	8	6	3	7	4	3	9	1	9	2	0	8	2	5	4	1
	5	6	4		6	3		4	3		1		2			2	5	4	1

Comparer le temps mis pour obtenir 100 résultats de lancer de dés et pour obtenir une simulation de même taille.

3) Simulation de familles de 4 enfants

Faisons les hypothèses suivantes :
- chaque naissance a autant de chances d'être celle d'un garçon ou celle d'une fille ,
- le sexe d'un enfant d'une famille ne dépend pas du sexe des enfants précédents.

Code 1	
F	G
0	5
1	6
2	7
3	8
4	9

```
9594 8630 7439 1960 9839 9939 5534 5500 8570 1411 …
GGGF GGFF GFFG FGGF GGFG GGFG GGFF GGFF GGGF FFFF …
GGGF FFGF GFGG GGFF GFGG GGGG GGGF GGFF FGGF GFGG …
```

Code 2	
F	G
0	1
2	3
4	5
6	7
8	9

4) Listes de nombres au hasard

On veut obtenir des listes de nombres au hasard entre 0 et 99 (resp. 0-999 et 0-9999) ; pour choisir au hasard dans 1-100, on ajoutera 1.

9594863074391960983999395534550085701411436162448679336155703248032043533585788082833289409183361825412131420528572174219435490959590680 …

2

95 94 86 30 74 39 19 60 98 39 99 39 55 34 55 00 85 70 14 11 43 61 62 44 86 79 33 61 55 70 32 48 03 20 43 53 35 85 78 80 82 83 32 89 40 91 83 36 18 25 41 21 31 42 05 28 57 21 74 21 94 35 49 09 59 59 06 80 …

3

959 486 307 439 196 098 399 939 553 455 008 570 141 143 616 244 867 933 615 570 324 803 204 353 358 578 808 283 328 940 918 336 825 412 131 420 528 572 174 219 435 490 959 590 680 …

4

9594 8630 7439 1960 9839 9939 5534 5500 8570 1411 4361 6244 8679 3361 5570 3248 0320 4353 3585 7880 8283 3289 4091 8336 1825 4121 3142 0528 5721 7421 9435 4909 5959 0680 …

5) Utiliser les nombres au hasard pour faire des phrases aléatoires

0 ou 1 : **belle marquise**

2 ou 3 : **vos beaux yeux**

4 ou 5 : **me font**

6 ou 7 : **mourir**

8 ou 9 : **d'amour**

9594863074391960983999395534550085701411436162448679336155703248032043533585788082833289409183361825412131420528572174219435490959590680 …

95948630 | 74391 | 960983999395 | 5345500857 | 01411436162448 | 67933615 | 5703248 | 0320435335857 | 8808283328940918336 | 182541213142052857 | 217 4219 | 4354909595906 | 80 …

95630 74391 96035 53087 04368 67315 57038 03487 80246 18257 21749 43906 80 …

95630

D'amour me font mourir vos beaux yeux belle marquise

ou

Mourir me font vos beaux yeux d'amour belle marquise ?

74391

exercices

Tous les tirages sont «avec remise».
Toutes les simulations seront faites en utilisant des tables de chiffres au hasard produits par une calculatrice ou un tableur. L'exercice consiste à écrire l'algorithme de simulation.

1) Simuler 10 lancers de 5 pièces équilibrées de couleurs différentes.
Simuler 10 lancers de 5 pièces équilibrées indiscernables.

Ces deux simulations consomment 50 chiffres au hasard. Pour la première, un résultat est une liste de 5 lettres P ou F choisies au hasard ; pour la seconde, un résultat est un nombre entre 0 et 5 qui est par exemple le nombre de piles.

2) Dans une urne, il y a 10 boules numérotées; simuler 20 tirages au hasard dans cette urne.
Dans une urne, il y a 94 boules numérotées; simuler 20 tirages au hasard dans cette urne.
Dans une urne, il y a 874 boules numérotées; simuler 20 tirages au hasard dans cette urne.

Pour la première simulation, on code les boules par les dix chiffres. Pour la deuxième, on prend les chiffres deux par deux et on ne tient pas compte de 00 et des entiers supérieurs ou égaux à 95. Pour la troisième, on prend les chiffres trois par trois et on ne tient pas compte de 0 et des entiers supérieurs ou égaux à 875.

3) Dans une urne, il y a 5 boules rouges et 5 boules noires.
Simuler les couleurs obtenues lors de 10 tirages au hasard dans cette urne.

Dans une urne, il y a 50 boules rouges et 50 boules noires.
Simuler les couleurs obtenues lors de 10 tirages au hasard dans cette urne.

Dans une urne, il y a 3645 boules rouges et 3645 boules noires.
Simuler les couleurs obtenues lors de 10 tirages au hasard dans cette urne.

Les trois simulations sont identiques puisqu'il y a autant de chances de tirer une boule rouge ou une boule blanche.

4) Dans une urne, il y a 73 boules rouges et 24 boules noires.
Simuler les couleurs obtenues lors de 10 tirages au hasard dans cette urne.

Dans une urne, il y a 7300 boules rouges et 2400 boules noires.
Simuler les couleurs obtenues lors de 10 tirages au hasard dans cette urne.

Les deux simulations sont identiques : on prendra les chiffres d'une liste de chiffres au hasard deux par deux, on ne tiendra pas compte du 00, 98,99. Les nombres de 1 à 73 seront codés rouge, les nombres de 74 à 97 seront codés noir.

5) Simuler 10 lancers d'une pièce déséquilibrée.

On pourrait simuler ce lancer si le fabriquant disait par exemple : à chaque lancer, on a trois fois plus de chances de tomber sur pile que sur face ; dans ce cas, sous réserve que ce que dit le fabriquant soit valide, on pourrait prendre les chiffres d'une liste au hasard deux par deux et coder par face les nombres de 00 à 74, par pile les autres.

6) A partir d'une liste de nombres choisis au hasard entre 0 et 99, comment construire une liste de nombres choisis au hasard parmi les nombres pairs compris entre 0 et 99 ?

Les élèves pour qui le résultat n'est pas intuitif pourront se reporter mentalement au tirage dans une urne à chiffres.

7) Si dans une liste de chiffres choisis au hasard, on élimine un chiffre sur deux, obtient-on encore une liste de chiffres au hasard ?

Les élèves pour qui le résultat n'est pas intuitif pourront se reporter mentalement au tirage dans une urne à chiffres.

8) Dans une liste de chiffres au hasard, on supprime un chiffre s'il est strictement inférieur au précédent ; obtient-on encore une liste de chiffres au hasard ?

Réfléchir sur la monotonie de la série obtenue.

De plus en plus de lancers d'un dé

- Echantillon de taille n d'une expérience.
- Notion de distribution de fréquences et de fluctuation d'échantillonnage.
- La distribution des fréquences de lancers d'un dé ***équilibré***, calculée sur une série de taille n varie d'une série à l'autre ; on observe que cette fluctuation est d'autant plus faible que n est grand ; quand n augmente, la distribution des fréquences se rapproche de (1/6, 1/6, 1/6, 1/6, 1/6, 1/6).
- La distribution des fréquences de lancers d'un dé ***non équilibré***, calculée sur une série de taille n, varie d'une série à l'autre ; on observe que cette fluctuation est d'autant plus faible que n est grand.

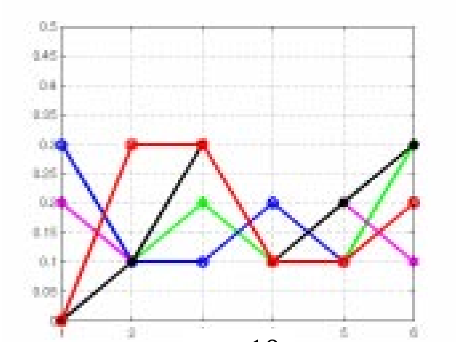
Dé équilibré

Une expérience consiste à lancer *n* fois un dé équilibré **:** les résultats constituent un échantillon de taille n de lancers d'un dé équilibré**.** La distribution de fréquences associée à cet échantillon est la liste des fréquences d'apparition de chacune des six faces. **A l'aide de listes de chiffres au hasard produites par un ordinateur, on simule 5 expériences pour chacune des valeurs de *n* suivantes : 10, 100, 1000 et 10000.**

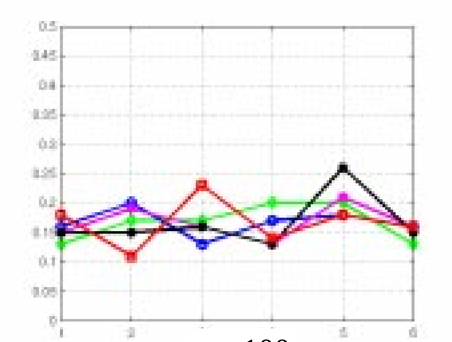
Les graphes ci-dessous représentent, pour chaque valeur de *n*, les 5 distributions de fréquences, avec en abscisse le numéro de face du dé et en ordonnée la fréquence associée.

On observe que :

- pour toute valeur de *n*, les distributions de fréquences des échantillons de taille n varient d'une expérience à l'autre : nous dirons qu'il y a **fluctuation d'échantillonnage**.
- lorsque *n* augmente, les distributions de fréquences « s'aplatissent » ; chaque fréquence semble se rapprocher de 1/6 (le 1/6 traduisant ici l'hypothèse que le dé est équilibré).

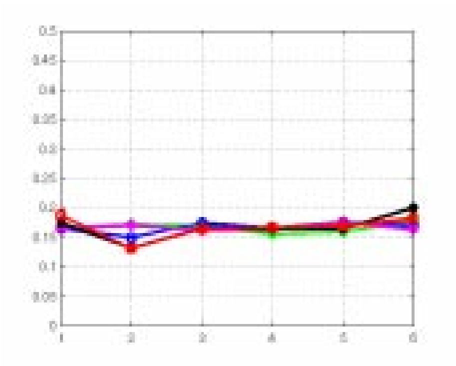


n = 10

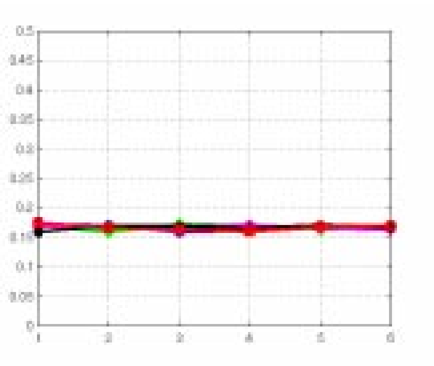


n = 100

Pour n=10, des représentations graphiques se chevauchent ; reconstituer les cinq distributions des fréquences.



n = 1000



n = 10000

Dé truqué

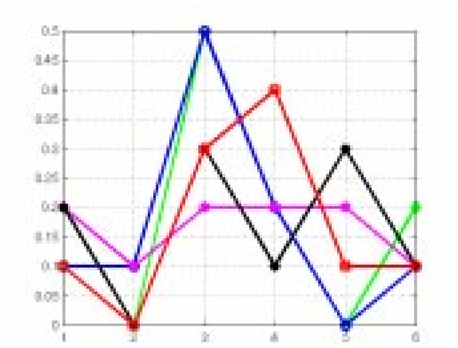
Une expérience consiste à lancer *n* fois un dé vendu comme dé truqué (i.e. à chaque lancer, toutes les faces n'ont pas les mêmes chances d'être observées) ; les résultats constituent un échantillon de taille n du lancer de ce dé**.** La distribution de fréquences associée à cet échantillon est la liste des fréquences d'apparition de chacune des six faces.

On fait 5 expériences dans lesquelles *n* prend successivement les valeurs 10, 100.

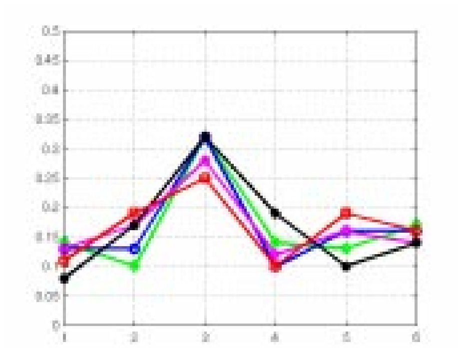
Les graphes ci-dessous représentent, pour différentes valeurs de *n*, les 5 distributions des fréquences, avec en abscisse le numéro de face du dé et en ordonnée la fréquence associée.

On observe que :

- pour toute valeur de *n*, les distributions de fréquences varient d'une expérience à l'autre : nous dirons qu'il y a **fluctuation d'échantillonnage**.
- on observe que pour n=10, les distributions des fréquences pour ce dé et pour un dé équilibré (cf. figure correspondante pour n=10) ne sont pas sensiblement différentes.
- Pour n=100, le déséquilibre commence à se voir.



n = 10

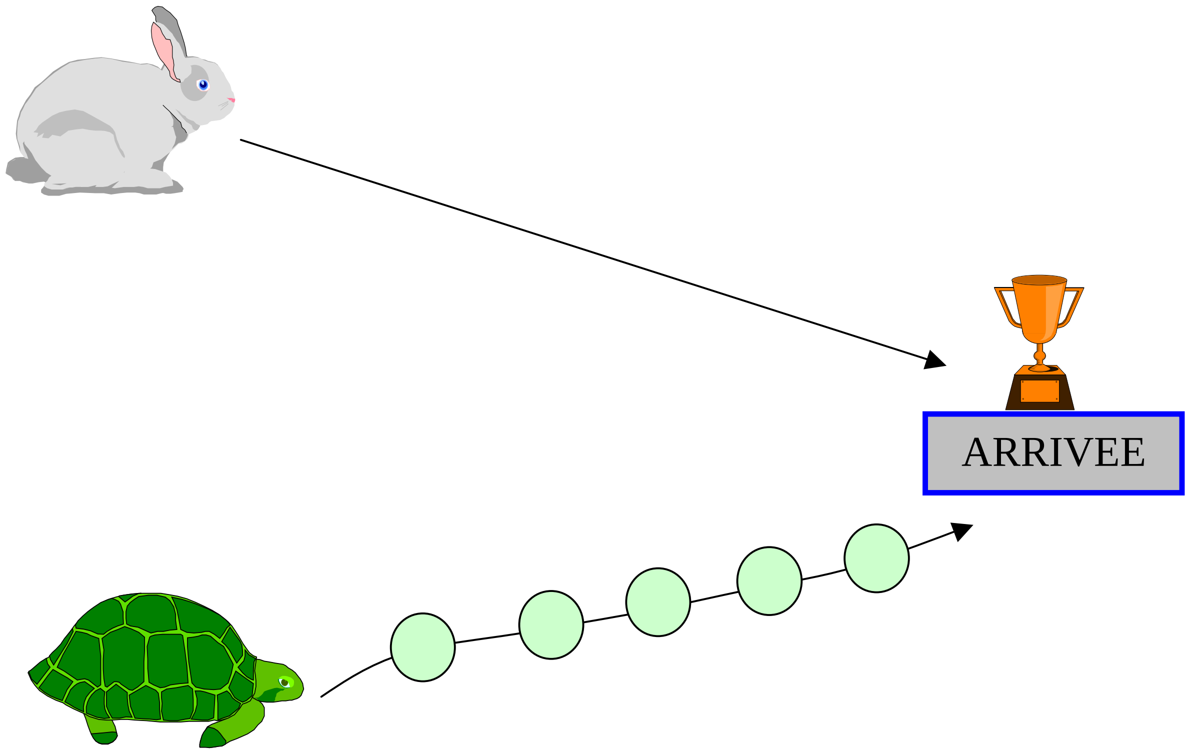


n = 100

Pour n=10, des représentations graphiques se chevauchent ; reconstituer les cinq distributions de fréquences.

Le lièvre et la tortue

- Expérimentation et simulation.
- Opérations sur les effectifs et sur les fréquences.
- Un exemple de jeu non équitable.
- La fluctuation d'échantillonnage diminue avec la taille des séries observées.



Une partie du jeu du lièvre et de la tortue se déroule ainsi :

1. **On lance un dé.**

   - ***Si le dé tombe sur 1, 2, 3, 4 ou 5 la tortue avance d'une case.***
     Elle a 5 cases à franchir avant d'atteindre l'arrivée.
     La partie est alors terminée, la tortue a gagné.

   - ***Si le dé tombe sur 6 le lièvre atteint directement l'arrivée.***
     La partie est alors terminée, le lièvre a gagné.

2. **La partie continue jusqu'à ce qu'il y ait un gagnant.**

Quelle est la situation la plus enviable : celle du lièvre ou celle de la tortue ?

1) Un jeu est composé de 10 parties

Pour un jeu, on recueille des données sous forme d'effectifs et on calcule des distributions de fréquences.

	(le lièvre, la tortue)	Les lancers de dés
Effectifs	(n, 10 - n)	(n_1, n_2, n_3, n_4, n_5, n_6)
Distribution des fréquences	(f, 1 - f)	(f_1, f_2, f_3, f_4, f_5, f_6)

Voici les résultats obtenus par dix élèves qui ont simulé un jeu.

Nombre de fois où le lièvre a gagné

Nombre de fois où la tortue a gagné

Nombre de fois où le dé est tombé sur i

	n	10 - n	n_1	n_2	n_3	n_4	n_5	n_6
Elève 1	7	3	5	6	7	3	6	7
Elève 2	6	4	2	7	10	9	11	6
Elève 3	6	4	4	8	6	9	6	6
Elève 4	6	4	8	7	8	8	6	6
Elève 5	5	5	8	10	10	5	7	5
Elève 6	8	2	7	3	9	2	4	8
Elève 7	7	3	3	5	11	13	9	7
Elève 8	7	3	5	2	5	8	9	7
Elève 9	5	5	9	5	9	9	9	5
Elève 10	4	6	4	10	13	10	4	4
Total	**61**	**39**	**55**	**63**	**88**	**76**	**71**	**61**

Fréquence de l’événement : “ le lièvre a gagné ” (f)

Fréquence de l’événement : “ la tortue a gagné ” ($1 - f$)

Distribution des fréquences (f_1 à f_6)

	f	$1 - f$	f_1	f_2	f_3	f_4	f_5	f_6
Elève 1	0.7	0.3	0.1471	0.1765	0.2059	0.0882	0.1765	0.2059
Elève 2	0.6	0.4	0.0444	0.1556	0.2222	0.2000	0.2444	0.1333
Elève 3	0.6	0.4	0.1026	0.2051	0.1538	0.2308	0.1538	0.1538
Elève 4	0.6	0.4	0.1860	0.1628	0.1860	0.1860	0.1395	0.1395
Elève 5	0.5	0.5	0.1778	0.2222	0.2222	0.1111	0.1556	0.1111
Elève 6	0.8	0.2	0.2121	0.0909	0.2727	0.0606	0.1212	0.2424
Elève 7	0.7	0.3	0.0625	0.1042	0.2292	0.2708	0.1875	0.1458
Elève 8	0.7	0.3	0.1389	0.0556	0.1389	0.2222	0.2500	0.1944
Elève 9	0.5	0.5	0.1957	0.1087	0.1957	0.1957	0.1957	0.1087
Elève 10	0.4	0.6	0.0889	0.2222	0.2889	0.2222	0.0889	0.0889
100 parties	**0.61**	**0.39**	**0.1356**	**0.1504**	**0.2116**	**0.1788**	**0.1713**	**0.1524**

Représenter graphiquement les distributions de fréquences des élèves 1 et 2 et celle qui correspond aux 100 parties. On explicitera les calculs relatifs à la dernière ligne du tableau.

Mais le tableau ci-dessus concerne 100 parties et on peut se demander si avec cent autres parties, les résultats ne seraient pas nettement différents. Nous allons donc faire plus de simulations.

2) Un jeu est une suite de 1000 parties et sera donc simulée.

Nombre de fois où le lièvre a gagné

Nombre de fois où la tortue a gagné

Nombre de fois où le dé est tombé sur *i*

	n	$1000\text{-}n$	n_1	n_2	n_3	n_4	n_5	n_6
Jeu 1	672	328	674	645	646	667	682	672
Jeu 2	659	341	661	646	633	682	705	659
Jeu 3	666	334	671	643	631	657	672	666
Jeu 4	674	326	627	668	656	689	648	674
Jeu 5	671	329	670	671	660	634	661	671
Jeu 6	682	318	658	646	639	615	647	682
Jeu 7	667	333	688	657	643	637	729	667
Jeu 8	683	317	628	732	630	671	651	683
Jeu 9	656	344	678	639	703	651	650	656
Jeu 10	671	329	637	649	647	670	685	671
Total	**6701**	**3299**	**6592**	**6596**	**6488**	**6573**	**6730**	**6701**

Fréquence de l'événement : “ le lièvre a gagné ”

Fréquence de l'événement : “ la tortue a gagné ”

Distribution de fréquences

	f	$1 - f$	f_1	f_2	f_3	f_4	f_5	f_6
Jeu 1	0.672	0.328	0.1691	0.1618	0.1621	0.1673	0.1711	0.1686
Jeu 2	0.659	0.341	0.1658	0.1621	0.1588	0.1711	0.1769	0.1653
Jeu 3	0.666	0.334	0.1703	0.1632	0.1602	0.1668	0.1706	0.1690
Jeu 4	0.674	0.326	0.1583	0.1686	0.1656	0.1739	0.1636	0.1701
Jeu 5	0.671	0.329	0.1689	0.1691	0.1664	0.1598	0.1666	0.1691
Jeu 6	0.682	0.318	0.1693	0.1662	0.1644	0.1582	0.1665	0.1755
Jeu 7	0.667	0.333	0.1711	0.1634	0.1599	0.1584	0.1813	0.1659
Jeu 8	0.683	0.317	0.1572	0.1832	0.1577	0.1680	0.1630	0.1710
Jeu 9	0.656	0.344	0.1705	0.1607	0.1768	0.1637	0.1634	0.1649
Jeu 10	0.671	0.329	0.1609	0.1639	0.1634	0.1692	0.1730	0.1695
Moyenne	**0.6701**	**0.3299**	**0.1661**	**0.1662**	**0.1635**	**0.1656**	**0.1696**	**0.1689**

- **On observe que les distributions varient moins d'une ligne à l'autre pour ce tableau que dans le cas où un jeu comporte 10 parties.**
- **Le tableau ci-dessus totalise 10 000 parties : c'est déjà beaucoup. On constate que le lièvre gagne plus souvent, et que sur les 10 000 parties simulées, il gagne dans 67% des cas.**
- **D'où sort ce 67% ? On conçoit que ce pourcentage ne dépend que du fait que le dé est équilibré, donc calculable à partir de cette hypothèse, mais comment faire ce calcul ?**

Expliquer le phénomène observé, c'est démontrer dans un cadre théorique que le lièvre gagne et calculer ses chances théoriques de gagner : mais pour cela, il faut donner une définition mathématique des hypothèses faites, à savoir que le dé est équilibré et les lancers indépendants les uns des autres.
La théorie qui permet de démontrer tout cela est la théorie des probabilités. On montre que la « chance théorique » du lièvre est 1 - $(5/6)^6$, soit à peu près 0,665.
Voici de nouvelles simulations ; on pourra observer que quand n augmente, les fréquences se rapprochent des valeurs théoriques.

NOMBRE DE PARTIES

VALEURS THEORIQUES

Pour f **:** $1 - (5/6)^6 \approx$ **0.6651**
Pour 1- f : $(5/6)^6 \approx$ **0.3349**
Pour les f_i : $1/6 \approx$ **0.1667**

	f	$1 - f$	f_1	f_2	f_3	f_4	f_5	f_6
10	**0.6100**	**0.3900**	**0.1356**	**0.1504**	**0.2116**	**0.1788**	**0.1713**	**0.1524**
1000	**0.6701**	**0.3299**	**0.1661**	**0.1662**	**0.1635**	**0.1656**	**0.1696**	**0.1689**
10000	**0.6667**	**0.3333**	**0.1667**	**0.1671**	**0.1660**	**0.1659**	**0.1668**	**0.1675**
100000	**0.6651**	**0.3349**	**0.1666**	**0.1667**	**0.1667**	**0.1669**	**0.1664**	**0.1671**

- On peut s'interroger sur la durée moyenne d'une partie (nombre de lancers de dés utilisés pour cette partie).

On pourra faire quelques remarques diverses ; par exemple :
- Si la tortue gagne, la partie dure 6 coups.
- Il y a une chance sur 6 que la partie dure 1 coup.
- Parmi les parties de 6 coups, le lièvre en gagne environ une sur six.

On n'a pas recueilli les durées des parties : on ne peut pas calculer la distribution des fréquences de cette durée ; en statistique, on a toujours intérêt à réfléchir à tout ce qui peut être intéressant avant de faire des expériences, pour ne pas risquer d'avoir à tout recommencer. On peut tout de même calculer des durées moyennes : pour les 100 parties du premier tableau, la durée moyenne est 4,14 et pour les 10 000 parties simulées, la durée moyenne est 3,958. Là aussi, la seule hypothèse faite concerne les lancers de dés et il devrait y avoir une formule qui permette de calculer une valeur théorique μ pour cette moyenne. Cette formule existe bien et se démontre par le calcul des probabilités :

$$\mu = \sum_{i=1}^{i=5} i \times \frac{5^{i-1}}{6^i} + 6 \times \frac{5^5}{6^5} \approx 3{,}99$$

- On peut modifier le jeu, c'est-à-dire changer le nombre de cases pour la tortue et faire franchir des cases au lièvre, en nombre inférieur, supérieur ou identique à celles de la tortue et étudier les distributions de fréquences (*f*, 1 - *f*) correspondantes.