À propos de l'introduction aux probabilités en Troisième

Auteurs : Brigitte Chaput et Claudine Vergne
Commission Inter-IREM Statistique et Probabilités

Emails : brigitte.chaput@educagri.fr ; claudine.vergne@wanadoo.fr

Version – 18 juin 2010

HASARD	**vient de l'arabe**	**AZ-ZAHR**	**qui signifie**	**JET de DÉ**
ALEA	**vient du latin**	**ALEA**	**qui signifie**	**DÉ, JEU de DÉ, JEU de HASARD**
CHANCE	**vient du latin**	**CADERE**	**qui signifie**	**CHOIR, TOMBER**

L'objectif de cet article est de proposer aux professeurs quelques éclairages et quelques pistes pour l'introduction aux probabilités en Troisième. Les propositions n'ont pas de caractère officiel et n'engagent que les auteurs. Le document s'adresse aux enseignants et n'est pas destiné à être transmis directement aux élèves. Nous nous restreignons au cas d'espaces probabilisés finis.

Programme de statistique et probabilités en classe de Troisième (publié le 28 août 2008)

Connaissances	Capacités	Commentaires
1.3. Statistique Caractéristiques de position.	- Une série statistique étant donnée (sous forme de liste ou de tableau ou par une représentation graphique) : • *déterminer une valeur médiane de cette série et en donner la signification ;*	Le travail est conduit aussi souvent que possible en liaison avec les autres disciplines dans des situations où les données sont exploitables par les élèves. L'utilisation d'un tableur permet d'avoir accès à des situations plus riches que celles qui peuvent être traitées « à la main »
Approche de caractéristiques de dispersion.	• *déterminer des valeurs pour les premier et troisième quartiles et en donner la signification ;* • *déterminer son étendue.* - Exprimer et exploiter les résultats de mesures d'une grandeur.	La notion de dispersion est à relier, sur des exemples, au problème posé par la disparité des mesures d'une grandeur, lors d'une activité expérimentale, en particulier en physique et chimie.
1.4. Notion de probabilité	- Comprendre et utiliser des notions élémentaires de probabilité. - Calculer des probabilités dans des contextes familiers.	La notion de probabilité est abordée à partir d'expérimentations qui permettent d'observer les fréquences des issues dans des situations familières (pièces de monnaie, dés, roues de loteries, urnes,etc.). La notion de probabilité est utilisée pour modéliser des situations simples de la vie courante. Les situations étudiées concernent les expériences aléatoires à une ou *à deux épreuves.*

Dans la version 2007 de ce programme, on pouvait lire dans **1.4 Notion de probabilité**, (colonne Commentaires), la phrase supplémentaire : *Certaines de ces situations permettent de rencontrer des cas pour lesquels les probabilités ne sont pas définies à partir de considérations intuitives de symétrie ou de comparaison mais sont approximativement évaluées par les fréquences observées expérimentalement (approche fréquentiste des probabilités).*

Cette phrase, loin d'être anodine, engageait à favoriser une approche fréquentiste des probabilités et le document ressource mis à disposition des enseignants (Ressources pour les classes de 6e, 5e, 4e et 3e du collège - Probabilités au collège - mars 2008) y fait encore référence (pages 5 à 7 par exemple). Sa disparition dans la version du programme la plus récente (28 août 2008) n'est pas anodine non plus. Elle n'a pas toujours été remarquée ou prise en compte : cela contribue à jeter le trouble et à créer des disparités importantes dans les interprétations et les mises en œuvre du programme.

Soulignons que le programme incite à *observer des fréquences* et que le mot *modéliser* apparaît. Notre approche didactique s'appuie sur ces prescriptions.

La brièveté du libellé du programme sur le thème *Notion de probabilité* ne doit pas cacher que de nombreuses précisions et mises au point sont nécessaires pour faire comprendre les notions en jeu.

1. - Conceptions du hasard

La *notion de probabilité* est indissociable de celle de hasard, car elle a pour objet de quantifier l'attente d'un évènement dont la réalisation est considérée comme dépendante du hasard. À leur arrivée en Troisième, les élèves ont déjà été confrontés à plusieurs conceptions du hasard dans leur vie quotidienne. Il convient, pour cette première rencontre avec les phénomènes aléatoires dans le cadre des mathématiques, de clarifier ceux qui, en mathématiques, permettront d'introduire des probabilités.
Au niveau du programme de Troisième, deux conceptions du hasard sont abordées : celle de *hasard du tirage au sort* dans une population statistique[1], avec équiprobabilité garantie par le générateur de hasard, et celle du *hasard bénin*, cause des fluctuations observées lors de la répétition d'une même expérience aléatoire ; ce dernier s'oppose au *hasard sauvage* de la réalisation fortuite d'un évènement accidentel (la tuile qui tombe du toit au moment où vous passez dessous, selon l'exemple donné par Cournot). L'élève arrive avec des conceptions personnelles relatives au hasard et au tirage au sort, qu'il s'est forgées au quotidien dans son entourage, sous l'influence des médias, etc. Il est important de faire émerger ces conceptions de façon à lever les ambiguïtés, les malentendus qui pourraient faire obstacle à la compréhension de l'approche mathématique de la notion de probabilité. Il s'agit de passer d'un hasard *subi* (dont on subit les effets) à un hasard *construit* auquel on peut rationnellement associer une quantification .

C'est au professeur de mathématiques qu'il revient :

- d'assurer transitions et ruptures entre les conceptions personnelles des élèves et l'approche scientifique ;
- de faire observer des phénomènes aléatoires de manière rationnelle par le biais de protocoles expérimentaux ;
- d'installer un vocabulaire qui permette de décrire ces phénomènes et de formuler des observations et des propriétés.

2. - Définir la notion de *probabilité* en Troisième ?

Il s'agit d'expliciter en restant à un niveau élémentaire la *notion* de probabilité installée progressivement, et non pas d'envisager le *concept* de probabilité, ce qui nécessiterait de définir un nouvel objet mathématique.
Dans certains manuels scolaires, les auteurs se sont essayés à donner des *définitions* qui se révèlent des plus fantaisistes.

- Certains ont fait un mélange incompréhensible entre deux approches de la notion : *La probabilité d'un évènement A est la proportion probable, parmi tous les cas possibles, des cas où A sera réalisé si on répète un grand nombre de fois l'expérience* (Dimathème, p. 184).

[1] Cf. Lahanier-Reuter, D.

- D’autres ont escamoté la difficulté : *La probabilité d’un évènement A représente les chances que l’évènement se réalise lors d’une expérience aléatoire* (Diabolo, p. 168).
- D'autres ont privilégié une approche fréquentiste, d'autres encore ont choisi de s’en tenir à la conception laplacienne restrictive : rapport du nombre de *cas favorables* au nombre de *cas possibles*.

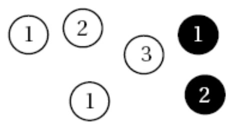
Avec les chercheurs contemporains en didactique des probabilités, nous pensons qu’il vaut mieux s'abstenir de donner une définition[2] plutôt que de tomber dans de telles dérives[3]. Au niveau qui nous intéresse, c'est-à-dire dans le cas discret fini où les issues d’une expérience aléatoire sont énumérables, la seule manière de définir formellement la probabilité serait axiomatique et consisterait à donner comme probabilités des issues une famille de nombres de l'intervalle [0 ; 1] dont la somme est égale à 1.
S'abstenir de donner une définition formelle de la probabilité n'affranchit pas d'être précis dans le choix des termes utilisés : ceux-ci ne doivent pas être en contradiction avec les définitions qui seront données dans les classes ultérieures. Nous proposons plus loin un ensemble de formulations qui pourraient être adoptées en Troisième.
Notons aussi que l'on rencontre dans certains ouvrages des énoncés tels que *deux évènements sont incompatibles s'ils ne peuvent se produire en même temps* ou *l'évènement contraire de A est celui qui se réalise quand A n'a pas lieu* qui ne peuvent pas être pris comme définitions : cette façon de s'exprimer introduit une notion de temporalité qui risque de brouiller la perception des notions abordées.
Quels que soient les choix du professeur, l'approche doit être expérimentale, l’institutionnalisation de quelques notions, du vocabulaire ou de propriétés peut se faire en cours d'étude ou en fin de chapitre, mais pas en préalable.

3. - Expérience aléatoire

Précisons d'abord que dans le cadre des exercices de probabilités proposés aux élèves, il est fondamental de décrire l'expérience aléatoire associée. Cette remarque n'est pas inutile puisqu'on rencontre encore certains exercices ou même sujets d'examens où rien n'indique le caractère aléatoire de la situation, c'est le cas par exemple de l'énoncé de l'exercice de probabilité du Brevet 2009 (centre de Pondichéry).

EXERCICE 4
Cet exercice est un questionnaire à choix multiples. Aucune justification n’est demandée. Pour chacune des questions, trois réponses sont proposées. Une seule est exacte. Chaque réponse exacte rapporte 1 point. Une réponse fausse ou l’absence de réponse n’enlève aucun point.
Pour chacune des trois questions, indiquer sur la copie le numéro de la question et recopier la réponse exacte.

Énoncé :
Un sac contient six boules : quatre blanches et deux noires. Ces boules sont numérotées :
Les boules blanches portent les numéros 1 ; 1 ; 2 et 3 et les noires portent les numéros 1 et 2.

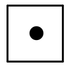
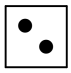


Numéro	Question	Réponse A	Réponse B
1	Quelle est la probabilité de tirer une boule blanche ?	$\frac{2}{3}$	$\frac{6}{4}$
2	Quelle est la probabilité de tirer une boule portant le numéro 2 ?	$\frac{1}{4}$	$\frac{1}{6}$
3	Quelle est la probabilité de tirer une boule blanche numérotée 1 ?	$\frac{1}{3}$	$\frac{2}{4}$

[2] En mathématiques, une définition est la délimitation précise d'un concept dans un cadre théorique en utilisant d'autres concepts et des propriétés caractéristiques.
[3] Cf. aussi l’article de Jean Claude GIRARD, *Quelle définition pour la probabilité au collège ?*

Dans la description d'une expérience aléatoire, d'un point de vue conceptuel et didactique, on peut distinguer trois niveaux.

- Le premier niveau est celui de l'*expérience réelle*. Par exemple, on lance une pièce de 1 € en l'air et on observe son comportement : la pièce peut tomber sur le côté Pile ou sur le côté Face ou sur la tranche ou rouler et se coincer sous un meuble (auquel cas le jeu s'arrête), etc. ; on peut ainsi envisager de nombreux *résultats* de l'expérience, certains farfelus mais pas impossibles. Un autre exemple consiste à lancer en l'air un dé cubique rouge et à observer son comportement : le dé peut retomber sur une des six faces ou cogner un obstacle et rester en équilibre sur une arête (dé cassé) ou se briser[4], etc. ; là-aussi on peut envisager de nombreux résultats possibles.
- Le deuxième niveau, celui de l'*expérience pseudo-concrète*, est une première étape dans la simplification et la modélisation de la réalité. Des objets qui produisent du hasard, on ne retient que certaines propriétés. De la multitude de résultats envisageables, on ne retient que ceux qui sont considérés comme pertinents d'un point de vue probabiliste et qui vont être objets d'étude, et on les interprète en termes d'*issues*.

  Par exemple, on lance en l'air une pièce (peu importe sa valeur, sa matière) et on observe si elle tombe sur Pile ou sur Face ; le résultat possible "*la pièce tombe sur la tranche*" n'est pas interprété en termes d'issues. Ou bien on lance un dé cubique (peu importe sa matière, sa couleur) et on observe le nombre de points de la face supérieure. Le résultat  est interprété par l'issue 1, le résultat  est interprété par l'issue 2, etc.
- Le troisième niveau est celui du *modèle mathématique* ou *modèle probabiliste*. Les objets, générateurs de hasard, sont idéalisés : *la* pièce est dite *équilibrée* (on s'intéressera plus loin au sens que l'on peut donner à cela), *le* dé est *homogène* et *régulier*. On introduit un ensemble d'*éventualités*, chacune étant affectée d'un nombre compris entre 0 et 1 (le rôle et les propriétés de ces nombres seront précisés plus loin). Cet ensemble d'éventualités représente en général les issues considérées dans le modèle pseudo-concret. L'ensemble des éventualités est appelé *univers* ou *ensemble fondamental* ou encore *référentiel*, il est noté Ω ou E.

  Par exemple, on peut modéliser les issues Pile et Face du lancer d'une pièce bien équilibrée par le choix des éventualités 0 et 1 et attribuer à chacune la probabilité $\frac{1}{2}$. Pour modéliser le lancer d'un dé cubique régulier dont les 6 faces sont les issues considérées, on peut choisir comme éventualités 1, 2, 3, 4, 5 et 6, chacune étant associée à la probabilité $\frac{1}{6}$.

Une *expérience aléatoire* est un processus
- où le hasard intervient pour produire un effet parmi d'autres possibles ;
- qui est susceptible d'être décrit par un *protocole expérimental,* lequel en permet, au moins par la pensée, la reproductibilité dans les mêmes conditions.

Le *protocole expérimental* est l'ensemble des instructions à suivre pour réaliser l'expérience aléatoire. Il est donc indissociable de celle-ci. Il doit :
- décrire clairement et avec précision les conditions de réalisation de l'expérience de façon à la caractériser et à pouvoir la reproduire dans les mêmes conditions ;
- décrire les observations qui peuvent être attendues d'un point de vue expérimental et présenter la liste des issues (on disait aussi les *cas possibles*) qui peuvent se présenter.

[4] Cf. Ekeland, I.

Le respect du protocole garantit que l'issue d'une expérience ne peut être ni prévue, ni calculée, ni influencée : à notre échelle, elle dépend du hasard. C'est la reproductibilité du protocole qui permet de reproduire l'expérience dans les mêmes conditions ; c'est la non-prédictibilité de l'issue qui confère à l'expérience son caractère aléatoire[5].

Par exemple, pour le jet d'un dé, le protocole peut être le suivant :

- le dé est placé dans un cornet que l'on secoue avant de le lancer ;
- le dé doit choir et rebondir sur un plateau à bord haut ;
- le dé ne doit pas être touché avant son immobilisation ;
- une fois le dé immobilisé, on regarde la face supérieure ;
- les six issues retenues sont notées 1, 2, 3, 4, 5, 6.

Le mot *évènement*[6] au sens probabiliste (pas au sens médiatique !) recouvre plusieurs acceptions suivant le degré d'abstraction, tout comme en géométrie, le mot *cube* désigne à la fois un objet réel référent, un dessin en perspective sur un papier (le signifiant), un solide de la géométrie aux propriétés idéales (le signifié).

- Dans la description de l'expérience aléatoire réelle, un évènement est décrit par une assertion. Par exemple, pour le jet d'un dé, on peut s'intéresser à l'évènement "*la face supérieure du dé montre un nombre pair*". Lorsqu'on réalise concrètement l'expérience, on obtient un résultat que l'on interprète en terme d'issue. Alors l'assertion se révèlera vraie, auquel cas on dit que l'évènement est réalisé, ou fausse.
- Dans le cadre d'un modèle pseudo-concret, un évènement est caractérisé par la liste des issues qui le réalisent (on disait aussi *cas favorables*). Par exemple, pour le jet d'un dé, l'évènement "*obtenir un nombre pair*" est réalisé par les issues 2, 4 et 6.
- Dans le modèle probabiliste, le mot évènement désigne un sous-ensemble de Ω ; par exemple, pour le jet d'un dé, l'évènement "*obtenir un nombre pair*" est $A = \{2, 4, 6\}$. Les singletons de Ω sont appelés *évènements élémentaires* (on distingue ainsi en toute rigueur, l'éventualité e_i de l'évènement élémentaire $\{e_i\}$ qui lui est canoniquement associé)[7].

4. - Observation des fréquences

Le programme précise que *la notion de probabilité est abordée à partir d'expérimentations qui permettent d'observer les fréquences des issues dans des situations familières*. De fait, il convient de ne pas dissocier la description statistique et les probabilités : les outils de la statistique[8] permettent de décrire et de résumer des résultats d'expériences aléatoires, les probabilités permettent de modéliser les phénomènes aléatoires.
Que peut signifier *observer les fréquences* et en quoi cette observation peut-elle déboucher sur la notion de probabilité ?

a) - Distribution de fréquences

On considère une expérience aléatoire donnant lieu à r issues possibles e_1, e_2,… e_r. En répétant n fois cette expérience aléatoire et en faisant la liste des n issues obtenues, on obtient un échantillon de taille n. Si, dans cet échantillon, l'issue e_i apparaît k fois, on dira que la fréquence de e_i dans l'échantillon est $f(e_i) = \frac{k}{n}$. Après avoir constitué un tel échantillon de taille n, on détermine la

[5] Cf. *Autour de la modélisation en probabilités*.

[6] L'accentuation *évènement* est conforme à la dernière réforme de l'orthographe ; l'écriture plus ancienne *événement* est encore acceptée.

[7] Dans un exposé élémentaire, le modèle pseudo-concret est confondu avec le modèle probabiliste abstrait, les mots *issue* et *éventualité* deviennent alors synonymes.

[8] Au sens donné à ce terme dans le programme de Troisième, il s'agit en fait de la statistique descriptive, qui est une branche de la science *Statistique*.

fréquence observée de chacune des issue e_1, e_2,… e_r, puis on établit la distribution de ces fréquences :

Issues	e_1	e_2	...	e_r	
Fréquences	f_1	f_2	...	f_r	1

Distribution des fréquences des issues observées sur un échantillon de taille *n*

On voit facilement que la somme des fréquences est égale à 1.

Si l'on produit 3 échantillons de taille 50 du lancer d'un même dé cubique, les distributions de fréquences peuvent être par exemple :

Distribution 1

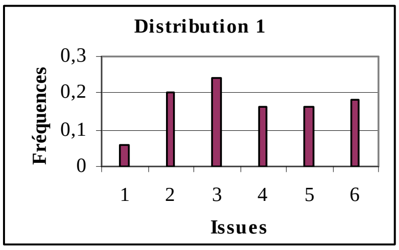
Issues	1	2	3	4	5	6	
Fréquences	$\frac{3}{50}$	$\frac{10}{50}$	$\frac{12}{50}$	$\frac{8}{50}$	$\frac{8}{50}$	$\frac{9}{50}$	1

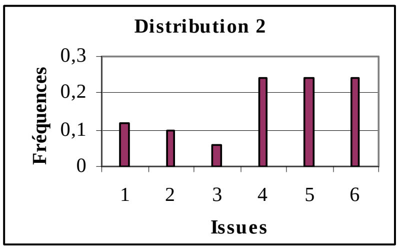
Distribution 2

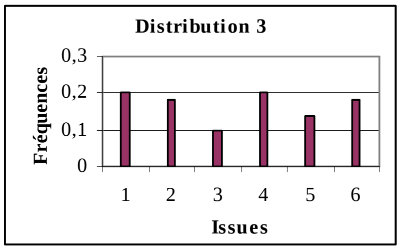
Issues	1	2	3	4	5	6	
Fréquences	$\frac{6}{50}$	$\frac{5}{50}$	$\frac{3}{50}$	$\frac{12}{50}$	$\frac{12}{50}$	$\frac{12}{50}$	1

Distribution 3

Issues	1	2	3	4	5	6	
Fréquences	$\frac{10}{50}$	$\frac{9}{50}$	$\frac{5}{50}$	$\frac{10}{50}$	$\frac{7}{50}$	$\frac{9}{50}$	1







Lancer d'un dé : distribution des fréquences des issues observées sur trois échantillons de taille 50

La première observation fondamentale est celle de la *variabilité* des fréquences d'un échantillon à l'autre : elle est inhérente au caractère aléatoire de l'expérience.

Le constat de l'inévitable variabilité ne va pas de soi et ne doit pas être occulté, il constitue une rupture épistémologique et didactique dans la classe de mathématiques où, habituellement, il est considéré que chacun doit donner la même réponse à une question posée. La prise de conscience de la variabilité est nécessaire pour saisir la nature des questions et des réponses qui sont en jeu dans le calcul de probabilités ainsi que les formulations que l'on adopte pour rendre compte des résultats des calculs.

Dans l'exemple ci-dessus, y a-t-il une distribution *meilleure* que les autres ? L'obtention de la distribution 3 rend-elle caduque celle de la distribution 1 produite antérieurement ? Sur quelle distribution s'appuyer pour prendre une décision, par exemple, relative à la qualité du dé utilisé ?

Comment gérer la diversité et la volatilité des données ? Cette problématique peut motiver la recherche d'un *modèle* de l'expérience aléatoire qui à la fois rende compte du phénomène aléatoire et s'affranchisse de la variabilité. Par exemple pour le *lancer d'un dé cubique*, si l'on considère qu'il est régulier et homogène, il n'y a pas plus de raison d'obtenir le 1 que le 2 ou que tout autre nombre. Les nombres positifs affectés à 1, à 2, ..., à 6 doivent être égaux, de somme 1. Pour ces raisons, le modèle probabiliste attaché peut, sans surprise, être décrit par la distribution de probabilité :

Éventualités	1	2	3	4	5	6	
Probabilités	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	1

Dans la culture courante, ce modèle existe, au moins de manière implicite. On peut être tenté de le poser *a priori*. Toutefois, on enrichit considérablement son sens et sa portée si on le propose après avoir travaillé sur la variabilité des fréquences pour des échantillons de taille donnée et après avoir remarqué que, pour certaines tailles d'échantillons, le nombre $\frac{1}{6}$, admis par tous, ne peut *jamais* être une fréquence observée (par exemple, pour des échantillons de taille 50, les fréquences observées sont toutes de la forme $\frac{k}{50}$, k entier entre 0 et 50).

b) - Analogie entre distribution de fréquences et loi de probabilité

Un deuxième type d'observations concerne les calculs de fréquences. Ceux-ci correspondent assez bien à l'intuition des élèves et si l'on s'en tient aux programmes en vigueur, ils pourraient être travaillés depuis la classe de Cinquième, à partir de l'étude de populations dans le cadre du chapitre *Gestion de données.*

Par exemple, à partir de la distribution des fréquences des issues d'un échantillon de taille 150 du lancer d'un dé cubique ci-dessous, demander de compléter la distribution, de calculer la fréquence d'un nombre pair, d'en déduire la fréquence d'un nombre impair... permet de mettre en œuvre des propriétés des fréquences telles que : la somme des fréquences est 1 ; la fréquence d'un évènement est la somme des fréquences des issues réalisant cet évènement, le complémentaire à 1 de la fréquence d'un évènement donne la fréquence de l'évènement contraire...

Issues	1	2	3	4	5	6
Fréquences	$\frac{23}{150}$	$\frac{23}{150}$	$\frac{27}{150}$		$\frac{30}{150}$	$\frac{24}{150}$

L'objectif est d'exploiter les analogies entre *Distribution des fréquences des issues observées sur un échantillon de taille donnée* et *Loi de probabilité sur un ensemble fini* de façon à donner du sens à la modélisation. Dans le cas fini qui nous intéresse, la définition mathématique (pour le professeur) du concept de probabilité peut être la suivante :

> Dans un modèle probabiliste fini, une loi de probabilité P sur l'ensemble fini $\Omega = \{e_1, e_2, \dots, e_r\}$ est une application définie sur l'ensemble des parties de Ω, à valeurs dans $[0\,;1]$. La somme des probabilités de tous les évènements élémentaires de Ω est égale à 1. La probabilité d'un évènement est la somme des probabilités des évènements élémentaires qui le constituent.

De fait, P est déterminée dès que l'on connait les éventualités e_i et les probabilités correspondantes $P(\{e_i\})$, c'est-à-dire dès que l'on connaît la distribution de probabilité :

Éventualités	e_1	e_2	...	e_r	
Probabilités	$P(\{e_1\})$	$P(\{e_2\})$	...	$P(\{e_r\})$	1

En pratique, par abus d'écriture, on écrit le plus souvent $P(e_i)$ au lieu de $P(\{e_i\})$ et plus simplement encore p_i.[9]

Les analogies sont résumées dans le tableau suivant, restreint au cas où les éventualités de l'univers correspondent aux issues du modèle pseudo-concret.

<table>
<tr><th>Distribution des fréquences des issues observées sur un échantillon de taille donnée</th><th>Loi de probabilité P sur un ensemble fini Ω</th></tr>
<tr><td>
<table>
<tr><td>Issues</td><td>e_1</td><td>e_2</td><td>...</td><td>e_r</td><td></td></tr>
<tr><td>Fréquences</td><td>f_1</td><td>f_2</td><td>...</td><td>f_r</td><td>1</td></tr>
</table>
Pour tout $i \in [1\,, r]$, $f_i \geq 0$
</td><td>
<table>
<tr><td>Éventualités</td><td>e_1</td><td>e_2</td><td>...</td><td>e_r</td><td></td></tr>
<tr><td>Probabilités</td><td>p_1</td><td>p_2</td><td>...</td><td>p_r</td><td>1</td></tr>
</table>
Pour tout $i \in [1\,, r]$, $p_i \geq 0$
</td></tr>
</table>

[9] On rencontre parfois la notation Prob(A) pour désigner la probabilité d'un évènement A. Cette notation, qui s'apparente à une abréviation, n'est pas opportune pour désigner l'image de l'évènement A par une loi de probabilité donnée, car cette loi variant d'un exercice à l'autre, l'écriture Prob laisserait imaginer qu'elle pourrait revêtir un caractère absolu, indépendamment de l'expérience considérée.

• $\sum_{i=1}^{r} f_i = 1$	• $\sum_{i=1}^{r} p_i = 1$
• Pour un évènement A, la fréquence de A est : $$f(A) = \sum_{e_i \text{ réalisant } A} f_i$$	• Pour un évènement A, la probabilité de A est : $P(A) = \sum_{e_i \in A} p_i$
• non$A$ (ou bien $\bar{A}$) désignant l'évènement contraire de A : $$f(\text{non}A) = 1 - f(A).$$	• $\bar{A}$ désignant l'évènement contraire de A, complémentaire de A dans Ω : $$P(\bar{A}) = 1 - P(A).$$
• Si A et B sont incompatibles (sans issues communes) : $$f(A \text{ ou } B) = f(A) + f(B)$$	• Si A et B sont incompatibles (i.e. $A \cap B = \emptyset$) : $$P(A \cup B) = P(A) + P(B)$$

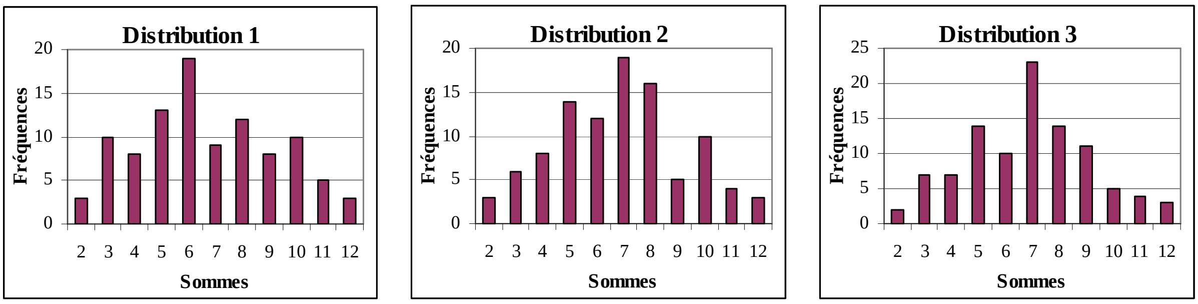
Rq : La notation $\bar{A}$ *pour l'évènement contraire de A est recommandée dans le programme.*

c) - Validité d'un modèle

Dans certains cas, l'observation des fréquences peut jouer le rôle de *contrôle* de la validité d'un modèle.
Considérons, par exemple, l'expérience qui consiste à lancer en même temps deux dés cubiques réguliers et à observer la somme des points des faces supérieures. On se persuade vite que les issues que l'on peut raisonnablement attacher à cette expérience aléatoire sont 2, 3, …, 12. Un néophyte en calcul de probabilités est souvent tenté alors de modéliser cette expérience aléatoire par la distribution de probabilité (biais psychologique d'équiprobabilité) :

Éventualités	2	3	4	5	6	7	8	9	10	11	12
Probabilités	$\frac{1}{11}$	$\frac{1}{11}$	$\frac{1}{11}$	$\frac{1}{11}$	$\frac{1}{11}$	$\frac{1}{11}$	$\frac{1}{11}$	$\frac{1}{11}$	$\frac{1}{11}$	$\frac{1}{11}$	$\frac{1}{11}$

Or la production de quelques échantillons de taille donnée (pas trop petite) va conduire à mettre en doute sa validité. Il faut alors rechercher un nouveau modèle, qui sera plus en accord avec les résultats de l'expérimentation et, si possible, explicatif.



Lancer de deux dés : distribution des fréquences des sommes observées sur trois échantillons de taille 100

Des problèmes de ce type, soulevant un désaccord entre un modèle (implicite parfois) et les résultats de l'expérimentation ont joué historiquement un rôle important dans l'avancée des recherches et la formalisation de la théorie des probabilités[10]. Ils illustrent clairement la dialectique

[10] On peut penser au problème du Grand Duc de Toscane. Celui-ci avait observé que lorsqu'on lance trois dés et que l'on fait la somme des faces apparues, on obtient 10 un peu plus souvent que 9. Or, se disait-il, 10 se décompose de six façons en somme de trois entiers compris entre 1 et 6 (6+3+1, 6+2+2, 5+4+1, 5+3+2, 4+4+2, 4+3+3) et il en est de

entre Statistique et Probabilité, qui s'alimentent l'une de l'autre. À l’origine, ces problèmes ont été résolus en considérant des modèles d'*équiprobabilité*.

5. - Situations d'équiprobabilité

Pour les situations familières invoquées dans le programme, la modélisation pourra s'appuyer sur des raisons de symétries (de la pièce, du dé...) et conduire au choix du modèle d’équiprobabilité en considérant qu'il n'y a pas plus de raison d'obtenir une issue plutôt qu'une autre : c'est le *principe de raison insuffisante* invoqué par Laplace à propos d’issues *également possibles*[11]. Dans le cas où l’on peut supposer que les issues d’une expérience aléatoire sont *d’égale possibilité*, on est amené à attribuer à chaque éventualité la probabilité égale à $\frac{1}{\text{cardinal}(\Omega)}$. La probabilité d'un évènement A est alors obtenue en ajoutant ce nombre autant de fois qu'il y a d'éventualités dans A. C'est ainsi que l'on retrouve la formule que Laplace avait donnée comme premier principe du calcul des probabilités $P(A)=\frac{\text{cardinal}(A)}{\text{cardinal}(\Omega)}$ et qu’il énonçait sous la forme $\frac{\textit{nombre de cas favorables}}{\textit{nombre de cas possibles}}$.

Ainsi, pour le problème de la somme de 2 dés cubiques, on peut considérer que les 2 dés sont distincts, par exemple l'un est rouge, l'autre vert, et choisir comme univers l'ensemble des couples $(x\,,y)$, x désignant le nombre obtenu sur le dé rouge, y celui obtenu sur le dé vert[12].

Dans ce cas, chaque couple est *également possible* et on attribue à chacun la probabilité $\frac{1}{36}$. L'évènement "*la somme est 7*" étant égal à $\{(1\,,6)\,;(2\,,5)\,;(3\,,4)\,;(4\,,3)\,;(5\,,2)\,;(6\,,1)\}$, sa probabilité est $\frac{6}{36}$. On obtient facilement les probabilités ci-dessous, plus en accord avec les distributions de fréquences représentées plus haut.

Sommes	2	3	4	5	6	7	8	9	10	11	12
Probabilités	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Dans la plupart des exercices au niveau de la classe de Troisième, l'énumération des éventualités est très facile et il serait regrettable que la mention *Calculs de probabilités dans des contextes familiers* soit interprétée de telle sorte qu'elle conduise les professeurs à tomber dans l’ornière de calculs complexes de dénombrements. Il faut s’en tenir à des calculs très simples, soit directs, soit mettant en jeu la règle d’addition : *la probabilité que se réalise l’un au moins de deux évènements incompatibles (ou disjoints) est la somme de leurs probabilités*.

6. - Stabilisation des fréquences

Malgré la restriction apportée dans la version du programme du 28 août 2008, *l'observation des fréquences des issues* peut bien sûr amener à constater que, pour des échantillons de très grande taille, la variabilité des fréquences observées d'un échantillon à l'autre est moindre : il y a une relative stabilisation des fréquences. La comparaison des distributions des fréquences pour les trois

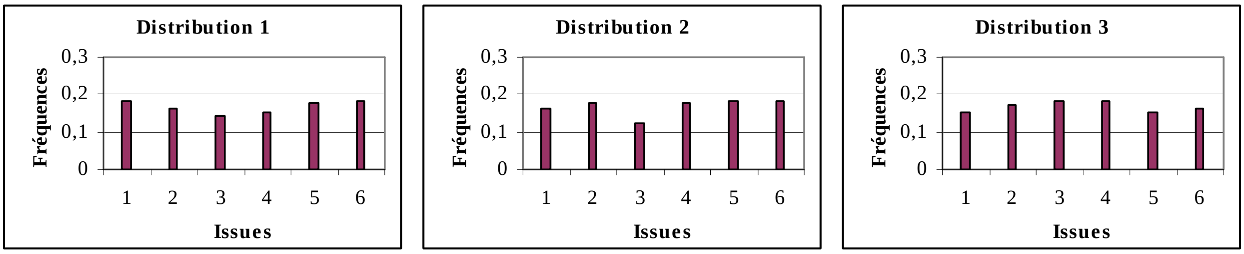
---

même pour 9 (6+2+1, 5+3+1, 5+2+2, 4+4+1, 4+3+2, 3+3+3). Le Duc de Toscane avait exposé ce paradoxe à Galilée qui rédigea vers 1620 un mémoire sur les jeux de dés pour lui répondre (ce mémoire n'a été publié qu'en 1718). Un poème attribué à Richard de Fournival, *De Vetula*, décrit dès 1260, les 216 *manières de tomber* équipossibles, correspondant aux 56 combinaisons observables sur 3 dés réalisant les 16 totaux possibles.

[11] Début de l’Essai philosophique sur les probabilités de 1814.

[12] Il est intéressant de remarquer que l'attribut *couleur du dé*, considéré comme non pertinent pour la modélisation de l'expérience aléatoire *lancer d'un dé cubique équilibré* peut aider à la modélisation de l'expérience aléatoire *lancer de 2 dés cubiques équilibrés et somme des faces*. Mais d’autres distinctions peuvent être proposées : tailles différentes, jets successifs… Cette distinction de couleurs n’est là que pour aider la pensée, car les probabilités des évènements considérés dans cette expérience ne dépendent évidemment pas de la couleur des dés.

échantillons de taille 500 ci-dessous avec celles obtenues pour le même dé pour des échantillons de taille 50 (**4- a)**) illustre ce phénomène.

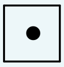
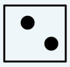
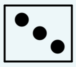


Lancer d'un dé : distribution des fréquences
des issues observées sur trois échantillons de taille 500

On constate ainsi un effet du *hasard bénin*, cela sera précisé en classe de Seconde à l'aide de la notion d'*intervalle de fluctuation*. Mais en adoptant cette démarche, on évite de *définir* la probabilité d'une issue par des notions vagues de *limite de fréquences* qui posent des problèmes épistémologiques non pertinents à ce niveau et qui dépassent les enjeux de cette classe ; une approche, même très vulgarisée de la loi des grands nombres n'est pas un objectif du programme de Troisième.

7. - Que dire aux élèves ?

De tout ce qui précède, que peut-on communiquer à des élèves de Troisième ? Le vocabulaire lié à la notion de probabilité peut être introduit progressivement sur des exemples. Voici une proposition de formulations que l'on pourrait adopter en Troisième pour l'étude du lancer d'un dé équilibré.

- Chaque lancer d'un dé cubique produit un *résultat* comme , , , …, dé cassé, dé perdu...

  Dans l'étude expérimentale du lancer d'un dé cubique, on ne s'intéresse qu'à certains résultats ; on retient comme *issues* les résultats des faces supérieures que l'on note 1, 2, ... 6.

- Si l'on imagine le lancer d'un dé parfaitement régulier, les 6 issues sont des éventualités également possibles. On peut modéliser le lancer d'un dé équilibré par la donnée des éventualités et de probabilités qui leur sont attachées :

Éventualités	1	2	3	4	5	6	
Probabilités	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	1

  On dit : la probabilité d'obtenir 1 est $\frac{1}{6}$, la probabilité d'obtenir 2 est $\frac{1}{6}$, etc.

- Un évènement est un ensemble d'éventualités.

  Par exemple, l'évènement "*le nombre obtenu est impair*" est noté si nécessaire B = {1 , 3 , 5}.
  On dit que 1 *réalise B*, de même 3 *réalise B*...On peut dire aussi *B est réalisé par* 1...
  On dit aussi 4 *ne réalise pas B.*

- Pour calculer la probabilité d'un évènement, on effectue la somme des probabilités des éventualités qui le réalisent.

  Par exemple : $P(B) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2} = 0{,}5$.

Avec les élèves, en pratique, il n'y a pas lieu de se focaliser sur la distinction entre *issue* et *éventualité*, un mot est souvent employé à la place de l'autre. La nuance est qu'une éventualité est ce qu'on peut attendre comme issue quand on effectue l'expérience.

Nous évitons ainsi, en Troisième, de parler d'évènement élémentaire et d'énoncer la phrase connue de tous, explicitement au programme de Seconde : *la probabilité d'un évènement est la somme des probabilités des évènements élémentaires qui le constituent*. Notre choix est guidé par le souci de ne pas alourdir le vocabulaire en Troisième et justifié par l'abus d'écriture précité entre $P(\{e_i\})$ et $P(e_i)$.

On pourra noter $P(1) = \frac{1}{6}$, $P(2) = \frac{1}{6}$, etc.

8. - Chance ? Chances ?

Dans le chapitre *Probabilités* des manuels scolaires, on voit fleurir dans les activités d'introduction, et parfois dans les définitions, de nombreuses expressions utilisant le mot *chance*. Ce mot du langage courant est susceptible d'être connoté subjectivement en fonction de la personne qui l'emploie. En l'utilisant, quelle part d'affectivité risque-t-on de laisser passer dans les représentations qui vont s'installer chez les élèves ? Par ailleurs, dans les expressions qui l'utilisent couramment, le mot prend des sens différents.

Considérons par exemple l'expérience aléatoire qui consiste à choisir au hasard dans une urne un jeton parmi 10, indiscernables au toucher et marqués chacun d'un numéro de 1 à 10. On gagne si le numéro tiré est 9 ou 10, on perd dans les autres cas. Concernant les *chances* de gagner, on peut rencontrer plusieurs formulations.

- "Il y a 2 *chances* sur 10 de gagner". Dans ce cas, le mot *chances* semble désigner les différentes issues réalisant le gain, les cas favorables. Mais si l'on considère que 2 sur 10 exprime la proportion de numéros gagnants sur le nombre total de numéros, 2 *chances* sur 10 *chances* au total renvoie aux 10 issues ou cas possibles.
- Mathématiquement parlant, il est tout aussi acceptable de dire "Il y a une *chance* sur 5 de gagner". Mais alors que représente le mot *chance* ?
- Si l'on énonce "Il y a 20 % de *chance* de gagner", le mot *chance* semble désigner la grandeur dont on prend 20 %, c'est-à-dire ici 1, la probabilité de l'univers.
- "Il y a 4 fois plus de *chance* de perdre que de gagner" exprime une comparaison entre la probabilité de perdre et celle de gagner.
- Sans compter qu'un élève, amené hors des mathématiques par sa perception du mot *chance* peut s'exclamer "Oh moi, je n'ai aucune *chance* de gagner, je ne gagne jamais !".

Ce simple exemple montre la polysémie attachée au mot *chance* dans des expressions courantes. Même en mathématiques, la polysémie existe (pensons au mot *cube* déjà évoqué), mais elle doit être évitée autant que possible. Selon nous, le mot *chance* doit être écarté du langage mathématique. Un des objectifs de l'enseignement des probabilités au niveau du collège est précisément de faire passer les élèves du langage familier des chances au langage mathématique rigoureux des probabilités. L'opportunité en est donnée au moment de la modélisation d'une expérience aléatoire. Une fois qu'un modèle est adopté, on effectue des calculs de probabilités (et non pas de *chances*) et on exprime les résultats des calculs en termes de probabilités. Par exemple, dans le tirage d'un jeton de l'urne, on dira que la probabilité de *tirer 9* est $\frac{1}{10}$ ou 0,1 ; que la probabilité de gagner est $\frac{2}{10}$ ou $\frac{1}{5}$ ou 0,2. Ce passage d'un langage en termes de *chance* à un langage en termes de probabilité nous paraît être un enjeu important de la classe de Troisième.

9. - Simulation ?

Le programme ne parle pas de simulation. La simulation d'échantillons est certes, pour le professeur, un outil efficient pour produire des données, illustrer des phénomènes... Peut-on, doit-on faire travailler les élèves sur la simulation instrumentée par une calculatrice ou un ordinateur dès la Troisième ? Les avis divergent sur la question. Quoiqu'il en soit, avant toute simulation d'une expérience aléatoire, il paraît indispensable de faire réaliser l'expérience concrètement plusieurs fois par les élèves pour qu'ils saisissent précisément où le hasard intervient.

Au-delà de l'introduction du calcul des probabilités, comme le soulignait en 2002 le rapport au Ministre de la Commission de Réflexion sur l'Enseignement des Mathématiques, il est important de créer une culture de l'aléatoire chez les élèves, sans l'imposer de façon dogmatique.

Références bibliographiques

BELLHOUSE, D. R. (2000). De Vetula: a medieval manuscript containing probability calculations, *International Statistical Review*, 68 (2), 123-136.

Commission Inter-IREM Statistique et Probabilités (2001). *Autour de la modélisation en probabilités.* M. Henry (éd.), Presses Universitaires Franc-Comtoises, col. Didactiques, Besançon.

Commission de Réflexion sur l'enseignement des mathématiques – Rapport d'étape – Statistique et probabilités (avril 2001)

- publié dans :
  KAHANE, J.-P. (2002). *L'enseignement des sciences mathématiques : Rapport au ministre. - Commission de réflexion sur l'enseignement des mathématiques*, sous la direction de Jean-Pierre Kahane. Odile Jacob.
- en ligne sur le site http://smf4.emath.fr/Enseignement/CommissionKahane/RapportStatistiqueProba/RapportStatistiqueProba.pdf

EKELAND, I. (1991). *Au hasard. La chance, la science et le monde*. Seuil

LAHANIER-REUTER, D. (1999). *Conceptions du hasard et enseignement des probabilités et statistiques.* P.U.F.

LAPLACE, P. S. (1814). *Essai philosophique sur les probabilités* (5ème édition, 1825), préface de René THOM, postface de B. BRU, Editions Bourgois, 1986.

Ressources pour les classes de 6e, 5e, 4e et 3e du collège - Probabilités au collège (mars 2008) :
http://media.eduscol.education.fr/file/Programmes/17/6/doc_acc_clg_probabilites_109176.pdf