ESD 2016_02 : Probabilités

1. Le sujet

A. L'exercice proposé au candidat

Arnaud et Bernard jouent à un jeu de dés. Le jeu consiste pour chacun d'eux à choisir un nombre compris entre 3 et 18, puis chaque joueur lance trois dés cubiques équilibrés et calcule la somme des nombres portés par les trois faces supérieures. Arnaud choisit le nombre 9, Bernard le nombre 10. Qui a le plus de chances de gagner ?

B. Les réponses proposées par des élèves de seconde

Elève 1

Pour avoir un total égal à 9 on peut avoir $6+2+1$ *ou* $5+3+1$ *ou* $5+2+2$ *ou* $4+4+1$ *ou* $4+3+2$ *ou* $3+3+3$

Pour avoir un total égal à 10 on peut avoir $6+3+1$ *ou* $6+2+2$ *ou* $5+4+1$ *ou* $5+3+2$ *ou* $4+4+2$ *ou* $4+3+2$ *ou* $4+3+3$. *Il y a donc autant de possibilités de faire 9 que de faire 10, je pense que Arnaud et Bernard ont autant de chances de gagner*

Elève 2

J'ai utilisé un tableur pour faire 100 lancers avec la fonction ALEA.ENTRE.BORNES(1;6). Sur cet exemple, j'ai obtenu le 10 plus souvent que le 9 mais en recommençant plusieurs fois 100 lancers, j'ai obtenu parfois le 9 plus souvent que 10, et parfois égalité. Je pense donc qu'Arnaud et Bernard ont autant de chances de gagner.

	A	B	C	D	E	F
1	dé 1	dé 2	dé 3	somme		
2	6	1	5	12		
3	5	3	6	14		
4	4	2	6	12	les 9	13
5	6	4	1	11	les 10	19
6	2	3	5	10		

Elève 3

Avec ma calculatrice j'ai tapé l'algorithme ci-contre : N doit contenir le nombre de 9, D le nombre de 10. Quand j'exécute le programme, il me donne toujours beaucoup plus de 10 que de 9. Je pense que c'est Bernard qui a le plus de chances de gagner, mais je trouve étrange qu'il y ait un tel écart entre les 10 et les 9.

```
N prend la valeur 0
D prend la valeur 0
pour I variant de 1 à 100 faire
|  Choisir un entier R au hasard
|    entre 1 et 6.
|  Affecter à S la valeur 3R
|  si S = 9 alors
|   | N prend la valeur N + 1
|  sinon
|   | D prend la valeur D + 1
|  fin
fin
Afficher N,D
```

C. Le travail à exposer devant le jury

1. Analysez les productions de des élèves en mettant en évidence les compétences acquises et les difficultés rencontrées par chacun d'eux.

2. Présentez une correction de cet exercice telle que vous l'exposeriez devant une classe de seconde, en vous appuyant sur les productions des élèves.

3. Proposez deux exercices sur le thème *probabilités* à des niveaux de classe différents. Vous motiverez vos choix en indiquant les objectifs pédagogiques visés par chacun des exercices.

2. Eléments de correction

Le problème étudié ici est connu sous le nom de « paradoxe du Duc de Toscane ». Le lecteur trouvera facilement de multiples ressources sur son contexte historique et sur son explication.

L'auteur du sujet laisse aux élèves le soin de traduire eux-mêmes la situation en langage mathématique en proposant un énoncé ouvert. L'énoncé utilise d'ailleurs délibérément le mot « chances », sans signification mathématique précise, à la place de « probabilité ».

Cet exercice peut être exploité pour souligner l'opportunité et l'efficacité d'une analyse probabiliste d'une situation dépendant du hasard et, peut-être, pour faire apparaître les limites d'une activité de simulation.

1. Analyse des travaux d'élèves.

Elève 1.
Réponse incorrecte.
Cet élève est peut-être un descendant authentique du Duc de Toscane. Sa production pourrait en effet avoir été inspirée par son hypothétique ancêtre lui-même …

Selon lui, comme il y a autant de façons de décomposer le nombre 9 que le nombre 10 en somme de trois entiers compris entre 1 et 6, les deux joueurs ont « autant de chances de gagner ».
Cet élève a su comprendre la situation et s'engager dans une démarche (dénombrer les décompositions de 9 et de 10 en sommes de trois entiers compris entre 1 et 6).

La difficulté rencontrée par cet élève est son incapacité à traduire cette situation concrète en langage mathématique (langage des probabilités). Cet élève s'exprime exclusivement en langage naturel (comme l'atteste l'emploi des termes « possibilités » et « chances »).

Il faudrait amener cet élève à se rendre compte de deux choses :

- Que certaines de ses « possibilités » se produisent plus souvent que d'autres (par exemple, trois nombres différents donnés sont obtenus plus souvent qu'un nombre répété et un autre nombre).
- Que son langage naturel est limité et ne peut décrire correctement la situation. En particulier l'emploi d'un tel langage n'est pas en mesure de pondérer les « possibilités » de quoi que ce soit qui ressemble à une probabilité.

Bougnègue.
Aujourd'hui Bougnègue propose une simulation sur tableur de l'expérience. Il a réservé trois colonnes, chacune simulant le résultat du lancer d'un dé. Une quatrième colonne fait la somme des trois nombres. La cellule F4 dénombre les sommes égales à 9, la cellule F5 les sommes égales à 10 (à l'aide de la fonction =SOMME.SI).
Il a sans doute écrit les formules adéquates en ligne 1 puis tiré vers le bas sur les quatre premières colonnes.
Sa simulation est correcte mais inadaptée au but recherché (qui n'a pas été identifié par cet élève).

Cet élève sait « élaborer une simulation numérique prenant appui sur la modélisation et utilisant un logiciel ».

Il faudrait faire préciser à cet élève qu'est-ce qui lui laisse penser que les deux joueurs ont « autant de chances » et reformuler avec lui les conclusions de son étude, à savoir que l'on ne peut pas écarter l'hypothèse d'égalité des chances (ce qui n'est pas du tout la même chose qu'affirmer qu'il y a bien égalité des chances). Mais on ne peut pas écarter non plus l'hypothèse d'inégalité…

Cet élève se heurte à deux difficultés :

- Il ne fait aucune analyse mathématique préalable de la situation permettant d'assigner à la simulation un objectif précis. C'est la difficulté majeure.

- Il n’est donc pas en mesure de paramétrer sa simulation en fonction de ce qu’il cherche à tester. Sa simulation n’est pas assez fine pour mettre en évidence la différence de fréquence entre une somme égale à 9 et une somme égale à 10.

Chouquerouste.
Démarche et conclusion incorrectes.
Chouquerouste propose une simulation de l’expérience à l’aide d’un algorithme.
Son algorithme est certes cohérent et démontre certaines connaissances de la notion de boucle (boucle For … EndFor) et de la notion de test alternatif (instruction If … Then … Else … EndIf).

On ne peut parler pour autant de « compétence » en matière de mise en œuvre d’algorithmes car son algorithme ne décrit pas la situation proposée :

- Chaque pas de la boucle simule le lancer d’un seul lancer dont on multiplie le résultat par 3.
- L’algorithme compte en réalité le nombre N de lancers où le résultat est 3 et le nombre D où le résultat est différent de 3.

Il obtient de ce fait que la fréquence de ce qu’il croit être une somme égale à 10 est environ cinq fois supérieure à celle de ce qu’il croit être une somme égale à 9. Même s’il conclue que le 10 sort beaucoup plus souvent que le 9, on ne peut en aucun cas parler de « compétence » en matière de prise de décision (il est amené à enfoncer une porte grande ouverte …)

Chouquerouste fait preuve d’une certaine velléité d’attitude critique par rapport à ses résultats, puisqu’il met en doute lui-même sa conclusion, mais on ne peut parler dans ce cas non plus de « compétence » car ce doute ne l’amène pas à revoir l’écriture de son algorithme.

Cet élève se heurte à deux difficultés :

- Il ne fait aucune analyse mathématique préalable de la situation permettant d’assigner à la simulation un objectif précis. C’est la difficulté majeure, exactement comme pour l’élève précédent.
- Ses connaissances en algorithmique sont insuffisantes. Le test « If … Then … ElseIf ... (Else) … EndIf » ou équivalent, d’ailleurs non exigible au niveau de la seconde, est nécessaire dans son algorithme pour trier les somme égales à 9 et celles égales à 10 des autres sommes.

Il faudrait commencer par lui proposer de remplacer 9 par 10 dans l’écriture de son algorithme. Il devrait s’attendre à obtenir à peu près les mêmes résultats, mais inversés, et se rendre compte que, bizarrement, il obtient toujours $N=0$; $D=100$ (ceci pour l’amener à corriger l’instruction de saisie des résultats par des lancers de trois dés *différents*). Lui demander ensuite si les sommes 9 et 10 sont les seules à pouvoir être obtenues (pour qu’il se rende compte que le test « If … Then … Else … EndIf » n’est pas suffisant pour trier les deux sommes 9 et 10).

D’une manière ou d’une autre, et quelle que soit la remise en question proposée à propos de leurs simulations, les élèves 2 et 3 devront tous deux être amenés tôt ou tard à modéliser mathématiquement l’expérience (par un espace probabilisé).

2. Correction de l’exercice.
Le lecteur trouvera facilement différents scénarios de correction de cet exercice parmi lesquels il pourra choisir.

Les trois élèves apportent des éléments intéressants sur lesquels il est possible de s’appuyer.

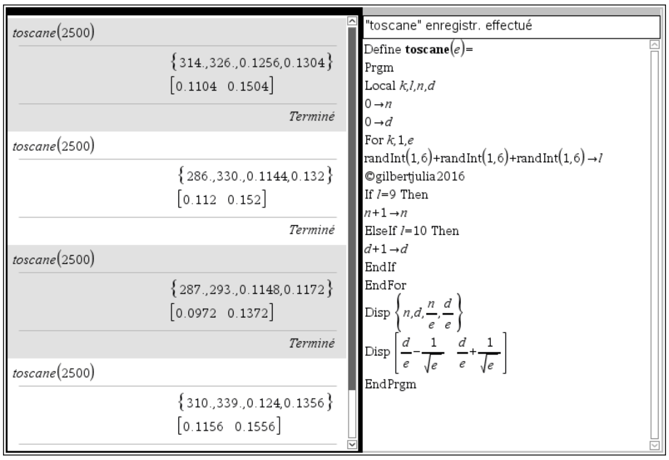
- L’élève 1 a trouvé les six décompositions possibles de 9 et 10 en sommes de trois nombres.
- L’élève 2 a clairement différencié les trois dés : « dé 1 », « dé 2 », « dé 3 ». (On peut éventuellement leur attribuer des couleurs).
- L’élève 3 a eu l’idée d’un algorithme qui, une fois corrigé, permet une simulation portant sur un grand nombre d’essais.

Un scénario parmi d’autres.

Deux élèves pensent que les deux joueurs ont les mêmes « chances » de gagner, le troisième n’a pas les moyens de se prononcer. Ont-ils raison ?

On propose d’effectuer une simulation plus fine de l’expérience, en exploitant un algorithme semblable à celui de l’élève 3 (corrigé bien entendu …). Le nombre de 2500 essais devrait apparaître « suffisant » pour se faire une idée assez précise.

On fait afficher par le programme **toscane(e)** les fréquences des sommes 9 et 10 ainsi qu’un intervalle de confiance au seuil 95% de la probabilité d’obtenir une somme égale à 10.



L’écran ci-contre nous laisse une impression mitigée.

D’une part on obtient dans tous les cas un peu moins souvent la somme 9 que la somme 10 mais d’autre part la fréquence de la somme 9 n’est jamais en dehors de l’intervalle de confiance qui s’affiche. Aucune idée précise comme on l’espérait $_{gj}$ …

Une étude mathématique apparaît nécessaire, comme l’a ébauchée l’élève 1 en décomposant de diverses façons les nombres 9 et 10 en sommes de trois entiers.

Est-il exact que ces décompositions se produisent toutes aussi souvent les unes que les autres (autrement dit, sont-elles *équiprobables* ?)

On propose aux élèves de s’intéresser de plus près aux décompositions $5+3+1_{gjulia2016}$, $5+2+2$ et $3+3+3$ (par exemple) en demandant (par exemple) comment elles apparaissent dans les colonnes du tableur de l’élève 2. L’objectif est de mettre en évidence qu’il y a six combinaisons pour obtenir $5+3+1$, trois pour $5+2+2_{gjulia2016}$ et une seule pour $3+3+3$.

Sans nécessairement introduire la totalité du vocabulaire probabiliste utile ici, il serait intéressant de parler à ce moment des 216 *éventualités* liées à l’expérience : les triplets (x, y, z) où x, y, z sont des entiers de l’ensemble $\{1,2,3,4,5,6\}$, toutes *équiprobables*.

Il reste à faire l’inventaire des types de décompositions :

	Trois nombres différents	Un nombre répété deux fois	Trois fois le même nombre	Probabilité
Somme 9	3	2	1	$3\times\frac{6}{216}+2\times\frac{3}{216}+\frac{1}{216}=\frac{25}{216}$
Somme 10	3	3	0	gilbertjulia2016 $3\times\frac{6}{216}+3\times\frac{3}{216}=\frac{27}{216}=\frac{1}{8}$

La probabilité que Bernard gagne est supérieure à la probabilité que Arnaud gagne. (On dénonce au passage le terme « chances » du langage courant et on lui substitue la « probabilité » d'un évènement).

En synthèse on met en évidence le fait qu'une simulation peut donner un verdict contraire à la réalité. Dans les circonstances ci-dessus, les simulations n'écartent pas l'hypothèse de même probabilité mais on souligne que *elle ne la démontre pas*, en aucun cas.

On insiste sur l'opportunité de définir un cadre probabiliste structurant l'expérience : déterminer un ensemble d'éventualités affectées d'une probabilité.

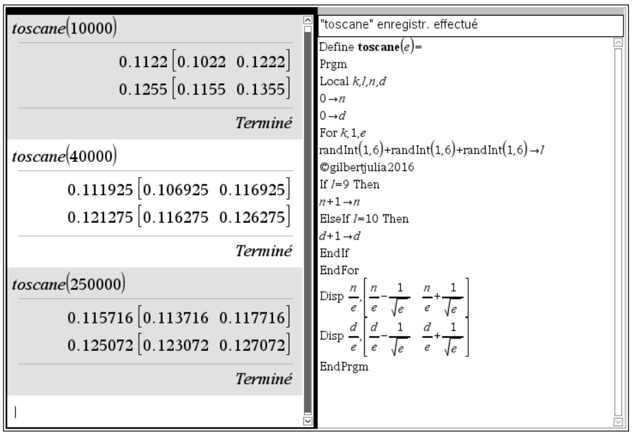
3. Pour aller plus loin

On peut faire remarquer que l'écart entre les deux probabilités est minime (moins de 0,01).

Le Duc de Toscane a dû observer un nombre de parties colossal pour se rendre compte[1] de l'existence de cet écart.

Le programme **toscane(e)** un peu modifié calcule ci-contre la fréquence sur *e* essais de la somme 9 et celle de la somme 10 et détermine aussi des intervalles de confiance au seuil 95 % des probabilités (supposées inconnues) de « obtenir 9 » et « obtenir 10 ».

On observe que dans les trois cas ci-contre, la fréquence de l'un n'est pas dans l'intervalle de confiance de l'autre. Cependant, ce n'est que dans le troisième cas que l'on obtient deux intervalles de confiance disjoints.



---

[1] Il est d'ailleurs étrange que ce perspicace Duc de Toscane ne se soit pas rendu compte que 5+3+2 était plus fréquent que 4+4+2 … Le paradoxe n'est peut-être qu'une légende.