

SÉRIE SIMULÉE OU INVENTÉE : UNE ACTIVITÉ INSTRUCTIVE ET LUDIQUE AU LYCÉE

Jean-François KENTZEL¹

TITLE

Simulated or invented series: informative and fun activity

RÉSUMÉ

Le travail présenté dans le cadre de cet article met en exergue le caractère non intuitif de la notion de hasard. Il est intéressant, par exemple, de travailler avec les élèves ou étudiants sur le fait que, dans une série de deux états équiprobables, le « hasard » n'induit pas un perpétuel changement d'état. Une activité fort instructive peut être proposée dans ce cadre aux élèves de différents niveaux en étudiant la longueur maximale d'une succession de piles ou faces lors d'une série de lancers d'une pièce équilibrée. Des variantes d'une même activité sont présentées afin de permettre de différencier deux séries dont une est obtenue par simulation et l'autre a été inventée par un élève. Cet article met aussi en évidence l'intérêt de l'usage d'un tableur dans la compréhension des phénomènes aléatoires, mais aussi dans l'approche de la difficile notion de test statistique, accompagnée de celle d'erreur associée à la décision.

Mots-clés : variable aléatoire, simulation.

ABSTRACT

The purpose of the work described in this paper is to highlight the non-intuitive character of randomness. For instance, high school or college students find it interesting to discover that in a succession of experiments involving two equally likely outcomes, "randomness" does not imply constant change. The length of runs of heads or tails in independent tosses of a fair coin is particularly instructive in this regard. This point can be illustrated in various ways by asking students at different levels to try and distinguish between a computer-generated random series of coin tosses and another one imagined by a student. This article also shows how a spreadsheet application can help users to understand randomness and grasp the difficult notions of statistical test and risk of error.

Keywords: random variable, simulation.

1 Introduction

Les programmes de mathématiques en lycée ont été modifiés depuis la rentrée 2010, tout d'abord en seconde, puis en première pour la rentrée 2011 et, enfin, les nouveaux programmes de terminales sont diffusés pour la rentrée 2012. Pour les enseignants de lycée non spécialisés en statistique, l'approche de la statistique via l'outil informatique est souvent utile et permet de mieux comprendre ce qui se passe. Ainsi, la notion de fluctuation d'échantillonnage ne devient-elle pas limpide dès lors que les élèves peuvent relancer plusieurs fois la sélection d'échantillon ainsi que le calcul d'estimateurs à partir de la touche F9 d'un tableur ?

¹ Lycée Pardailhan à Auch, jkentzel@ac-toulouse.fr

Série simulée ou inventée : une activité instructive et ludique au lycée

De plus, il est important de réussir à motiver les élèves ou étudiants dans leur apprentissage de la statistique car il est assez évident que l'étude de la statistique leur rendra de grands services dans des domaines bien variés ; pour réaliser cet objectif, l'utilisation de l'informatique est un outil essentiel.

L'activité qui est présentée ci-dessous rentre complètement dans le cadre de la partie du programme d'enseignement spécifique de mathématiques en classe de première, mais sera applicable bien au-delà de ce niveau avec quelques modifications à la marge que nous suggérerons tout au long de ce papier.

Dans l'annexe de l'arrêté du 27-01-2010 modifié du bulletin officiel spécial du 30 septembre 2010 [1], il est dit : « L'étude et la comparaison de séries statistiques menées en classe de seconde se poursuivent avec la mise en place de nouveaux outils dans l'analyse de données. L'objectif est de faire réfléchir les élèves sur des données réelles, riches et variées (issues, par exemple, de fichiers mis à disposition par l'INSEE). La notion de loi de probabilité d'une variable aléatoire permet de modéliser des situations aléatoires, d'en proposer un traitement probabiliste et de justifier certains faits observés expérimentalement en classe de seconde. »

Une activité similaire à celle proposée dans cet article mais n'utilisant pas de simulation réalisée par les élèves avait été proposée en 2009 [2] ; cette activité est accessible à partir du site indiqué en bas de page². Il s'agit ici d'une amélioration d'une activité attribuée à T. Varga et racontée³ par P. Révész à un congrès à Helsinki en 1978.

Il s'agira dans une première partie de décrire l'activité proposée, puis d'indiquer les notions soulevées dans le cadre de cette activité et enfin de discuter d'un certain nombre de points.

2 Description de l'activité d'une durée proposée (durée : 1 heure minimum)

L'activité proposée va se décliner en deux actes pourrions-nous dire, ainsi qu'un entracte.

2.1 Premier acte : cacher la suite simulée

Mission : chaque élève doit construire deux séries de valeurs, A et C, dans les colonnes 1 et 3, composées de 200 chiffres prenant les valeurs 0 ou 1 et représentant les résultats de 200 lancers d'une pièce équilibrée (0 pour face, 1 pour pile). Une des deux colonnes (l'élève choisira la première ou la troisième sans l'indiquer) aura été élaborée en utilisant le simulateur de nombres aléatoires du logiciel utilisé (la série sera dite « simulée ») et la seconde sera composée d'une succession de 0 et 1 composée selon le bon vouloir de l'élève (la série sera dite « inventée »).

² <http://pardailhan.entmip.fr/rubrique-des-disciplines/mathematiques/documents-enseignants/differencier-une-suite-aleatoire-d-une-autre-qui-ne-l-est-pas-844.htm>

³ Anecdote trouvée dans « Analytic combinatorics » (pages 52 et 53) de Philippe Flajolet (INRIA) and Robert Sedgewick (Princeton), 2008 ; disponible sur la page web <http://algo.inria.fr/flajolet/Publications/publist.html> .

J.-Fr. Kentzel

Il est donné comme objectif pour l'élève d'empêcher « l'adversaire » de reconnaître quelle est la série inventée⁴ (autrement dit, il doit imiter de son mieux une série simulée). Chaque élève tape « =ent(alea()*2) »⁵ ou « =alea.entre.bornes(0;1) » (suivant qu'on utilise Excel ou Open Office) dans la première cellule de la colonne choisie puis glisse-copie jusqu'à la 200^e ligne de la même colonne. Il s'agit donc de simuler la réalisation de 200 variables aléatoires indépendantes de même loi de Bernoulli de paramètre 1/2.

Dans l'autre colonne, il « invente » une suite de deux cents 0 ou 1 en les tapant un par un. On supprime ensuite les formules alea() en utilisant la notion de recopiage des valeurs uniquement et non des formules (clic droit puis « collage spécial - valeurs »).

2.2 Entracte

Les élèves changent de place afin de s'assurer qu'ils n'ont pas pu voir les colonnes choisies par leur « adversaire ».

Avant de reprendre la manipulation du fichier, l'enseignant définit par N **la longueur de « la » plus longue suite de « piles » dans une série de 200 lancers indépendants d'une pièce équilibrée** ; N est une variable aléatoire⁶.

On montre aux élèves que, pour une série de 200 lancers, $P(N \geq 5) \approx 0,965$ en se contentant la plupart du temps d'une simulation réalisée, par exemple, avec un tableur. On peut trouver cette simulation et la fiche distribuée aux élèves sur le site du Lycée Pardailhan de Auch⁷. Prouver ce résultat est hors de portée des élèves ; dans certaines classes cependant, à titre d'information, nous proposons en annexe de cet article une démonstration par récurrence. On attire leur attention sur le fait que la probabilité que N soit inférieur ou égal à 4 est « faible », de l'ordre de 3,5%, c'est-à-dire inférieure à 5%, seuil utilisé assez fréquemment dans des études statistiques comme caractérisant des événements « peu probables ».

Le professeur peut apprécier, en fonction des capacités de ses élèves et du temps imparti à ce travail, s'il est opportun de faire faire par les élèves eux-mêmes cette simulation (car l'activité est alors infaisable en une heure).

⁴ Cette activité peut être pratiquée, sur 4 colonnes au lieu de 2, même si on a 35 élèves pour 20 ordinateurs. Elle est encore tenable en une heure car les élèves s'aident. On peut d'ailleurs imaginer 4 colonnes par élève, 2 inventées et 2 pseudo-aléatoires (je ne l'ai pas testé).

⁵ Sous l'hypothèse de travailler sous Excel.

⁶ Il est conseillé de ne pas introduire ce vocabulaire dans certaines classes.

⁷ <http://pardailhan.entmip.fr/rubrique-des-disciplines/mathematiques/premiere/reconnaitre-alea-20179.htm>

Série simulée ou inventée : une activité instructive et ludique au lycée

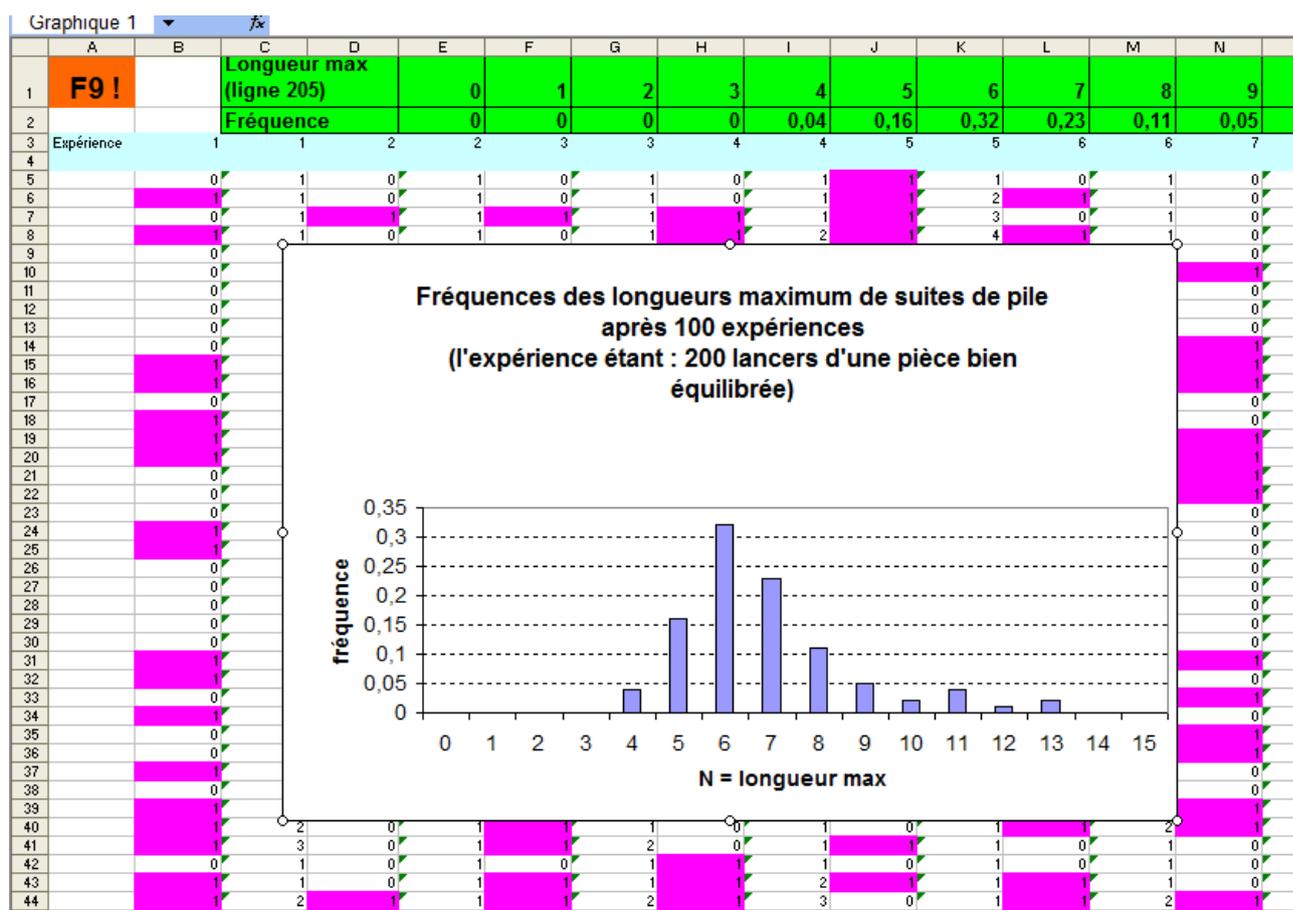


FIGURE 1 – Feuille de simulation de 100 suites de 200 lancers de pièces et distribution observée de la variable N

2.3 Acte 2 : retrouver la série inventée

Le nouvel objectif des élèves est de repérer la série inventée, la A ou la C, de la feuille de leur camarade, en utilisant la valeur prise par N pour les séries A et C.

Chaque élève doit définir la stratégie à développer en fonction de ce qui a été vu durant l'entracte.

Ces valeurs de N peuvent être déterminées visuellement ou informatiquement : par exemple, pour la série A qui est sur la plage « A1:A200 », il suffit de :

- taper dans la cellule B1 : « =A1 » ;
- taper dans la cellule B2 : « =SI(A2=0;0;B1+1) » ;
- glisser-copier jusqu'en B200 ;
- taper en B201 par exemple : « =MAX(B1:B200) ».

A la fin de l'acte 2 une évaluation rapide est réalisée auprès des élèves ou étudiants pour connaître la proportion ayant reconnu la série inventée. Il s'agit aussi de recueillir les différents arguments avancés les ayant conduits à leur choix.

Un bilan est alors proposé par le professeur, mettant en évidence le rôle de la propriété vue pendant l'entracte. En reprenant globalement les séries imaginées, le constat est fait que la

J.-Fr. Kentzel

propriété selon laquelle la probabilité que N soit supérieur ou égal à 5 n'est « que » de 3,5 % est suffisamment surprenante pour que, dans leur imagination, les élèves n'aient pas « osé » réaliser des suites de 1 aussi longues. Le professeur approuvera donc les élèves qui, face à deux séries pour lesquelles N serait supérieur ou égal à 5 dans l'une et inférieur ou égal à 4 dans l'autre, auront conclu que la première est simulée.

3 La mise en œuvre de l'activité

3.1 Quel objectif ?

L'objectif de cette activité est une modeste sensibilisation aux méthodes de la statistique inférentielle.

Il s'agit d'étudier deux séries de 200 informations binaires successives. L'une est construite comme succession de 200 lancers indépendants d'une pièce équilibrée (on notera 1 pour pile et 0 pour face), ce qui peut être représenté comme une succession de 200 résultats de variables aléatoires indépendantes suivant toutes une loi de Bernoulli de paramètre $1/2$. L'autre série est inventée avec comme objectif de reproduire ce que « le hasard » (au sens défini précédemment) aurait pu permettre de construire.

Dans le cadre de cette activité, la notion de test est donc introduite sans que cela soit vraiment dit, car il s'agit de déterminer la probabilité associée à la longueur maximale de 1 obtenue dans chacune des séries statistiques sous l'hypothèse que la série est une série obtenue par simulation de la loi de Bernoulli de paramètre $1/2$. Il s'agit donc de lire la valeur de la probabilité dans la loi de probabilité de la variable N construite durant la partie de l'activité correspondant à l'extraite.

Cela revient à tester l'hypothèse : A (ou C) est la réalisation d'une succession de 200 variables aléatoires indépendantes identiquement distribuées suivant la loi de Bernoulli de paramètre $1/2$.

3.2 L'implantation de l'activité

Pour que cette activité puisse être réalisée en une heure, l'acte 1 doit impérativement être très court, il est préférable que la notion de fonction `alea()`, par exemple, ait été introduite par ailleurs.

Il est évidemment nécessaire de bien préciser ce que l'élève va devoir faire, c'est-à-dire faire un choix à partir d'un argumentaire.

L'élève a donc deux étapes à réaliser très différentes : dans un premier temps, il doit imaginer une série de 0 et de 1 en s'imaginant maître « du hasard » puis, dans un second temps, il doit être en mesure de reconnaître entre deux séries celle qui est probablement due « au hasard ».

3.3 Pour aller plus loin

Plusieurs variantes de cette activité peuvent être proposées et correspondre à des niveaux d'enseignement différents.

Série simulée ou inventée : une activité instructive et ludique au lycée

3.3.1 Variante 1

On s'intéresse dans cette variante à la variable aléatoire $L =$ **la longueur de « la » plus longue suite de résultats consécutifs identiques de 0 ou de 1**.

La condition $P(N \geq 5) \geq 0,96$ devient $P(L \geq 6) \geq 0,96$, à peine plus difficile à prouver (voir l'annexe). Cette variante d'activité ainsi que la simulation associée sont accessibles à partir de [ce lien](#)⁸.

Ces variables aléatoires N et L présentent une proximité importante. On précise les notations en posant, pour tout entier $n \geq 1$:

- $N_n =$ la longueur de « la » plus longue suite de « piles » obtenue à l'issue du n -ème lancer ;
- $L_n =$ la longueur de « la » plus longue suite de résultats consécutifs identiques obtenue à l'issue du n -ème lancer.

On peut prouver assez facilement que, pour tous n et k , $P(N_n = k) = P(L_{n+1} = k+1)$. **Les suites définissant les lois de N_n et L_n sont identiques « à un double décalage près ».**

Cette égalité est compréhensible avec les preuves données dans l'annexe qui suit. Si on pose, pour k entier fixé, $p_n = P(N_n = k)$ et $q_n = P(L_n = k+1)$, on montre que les deux suites (p_n) et (q_n) sont définies par la même relation de récurrence, cependant qu'elles ne diffèrent que d'un décalage, d'un rang, pour leurs premières valeurs :

Suite N_n :

- pour $n \in \{0; k-1\}$: $p_n = 0$
- pour $n = k$: $p_n = \frac{1}{2^k}$
- pour $n = k+1$: $p_n = \frac{3}{2^{k+1}}$

Suite L_n :

- pour $n \in \{0; k\}$: $q_n = 0$
- pour $n = k+1$: $q_n = \frac{1}{2^k}$
- pour $n = k+2$: $q_n = \frac{3}{2^{k+1}}$

3.3.2 Variante 2, plus simple

Si on trouve que les variables aléatoires N ou L sont trop compliquées, ce qui serait dommage car le résultat $P(N \geq 5) \approx 0,965$ ou $P(L \geq 6) \approx 0,96$ étonne les élèves (il a même étonné certains probabilistes), on étudie la variable aléatoire $I =$ **le nombre d' « inversions » pile-face ou face-pile au cours des 200 lancers**. Par exemple, I vaut 4 pour la série 1 1 0 1 1 0 0 0 1.

⁸

<http://pardailhan.entmip.fr/rubrique-des-disciplines/mathematiques/premiere/reconnaitre-alea-en-comptant-des-repetitions-21170.htm>

J.-Fr. Kentzel

En classe de seconde on réalise sans peine une simulation et en classe de première on prouve que I suit la loi binomiale $B(199; 1/2)$. Dans les deux cas on obtient facilement que la probabilité que I soit supérieur ou égal à 112 est de l'ordre de 5%.

Comme L et I ont tendance à varier en sens inverse (plus on fait d'inversions, moins on fabrique des suites homogènes longues), on obtient « souvent » une valeur de I plus grande pour la série inventée.

Le professeur approuvera donc les élèves qui auront trouvé justifié, face à deux séries pour lesquelles I serait inférieur ou égal à 111 dans l'une et supérieur ou égal à 112 dans l'autre, de conclure que la première est simulée.

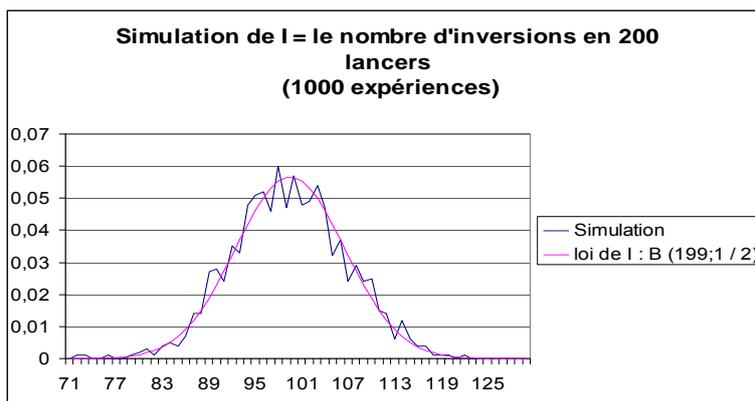


FIGURE 2 – Simulation de 1000 expériences de la valeur de I (voir une simulation et la fiche distribuée aux élèves à partir de ce [lien](#)⁹)

La simulation accessible sur l'adresse indiquée ci-dessus ne porte que sur 100 expériences (fichier trop lourd pour 1000 expériences). L'intérêt d'une simulation sur 1000 expériences pour I vient du fait que, pour cette variable I , les simulations sur 100 expériences semblent plus chaotiques que celles réalisées pour N (voir les deux graphiques de la figure 3 et une comparaison avec la commande F9 sur ce [lien](#) précédemment cité¹⁰).

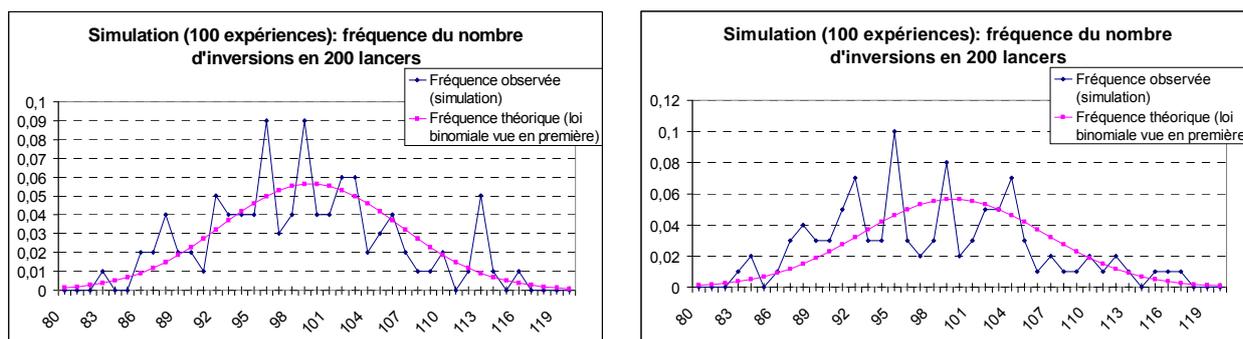


FIGURE 3 – Simulation de 100 expériences de la valeur de I

⁹ <http://pardailhan.entmip.fr/rubrique-des-disciplines/mathematiques/premiere/reconnaitre-alea-en-comptant-des-inversions-21247.htm>

¹⁰ L'approximation $P(N \geq k) \approx \exp(-1/2^{k+1})$ utilisée dans ce fichier est hors de portée pour les élèves de lycée, voir la deuxième référence de la remarque 7.

Série simulée ou inventée : une activité instructive et ludique au lycée

Ceci peut inciter à faire réaliser par les élèves une simulation sur 1000 expériences, chacune des séries devant alors être écrite sur une ligne. Pour mettre en place cette simulation sous Excel, il suffit de :

1. écrire en A2 et B2 : « =ent(alea()*2) » ;
2. mettre en A1 : la valeur 0 ;
3. mettre en A3 : « =\$A\$1 » (détail qui sera utile pour éviter tout problème¹¹ en recopiant la ligne 3 pour une simulation) ;
4. mettre en B3 : « =SI(B2=A2;0;1) » ;
5. recopier B2 et B3 jusqu'à la colonne GR ;
6. taper en GS3 : « =SOMME(A3:GR3) » ;
7. recopier la plage B2:GS3 jusqu'à la ligne 2001.

4 Discussion

4.1 L'évaluation de l'activité initiale

Les élèves déterminent la série inventée dans environ 80% des cas.

Comme on l'a déjà indiqué, le facteur psychologique qui fait réussir assez souvent la reconnaissance de la série inventée est le suivant : pour « faire plus aléatoire », l'élève n'ose pas trop écrire de longues suites de 0 ou de 1, il dissimule le mieux possible la série inventée en lui donnant des inversions de 0 et de 1 assez fréquentes (cette remarque est facile à tester auprès de tout public). Alors qu'en réalité **on a environ 95% de chances d'obtenir au moins 5 succès (pile) consécutifs quand on lance 200 fois une pièce équilibrée**. On obtient donc « souvent » une valeur de N plus faible pour la série inventée.

Le résultat du calcul des probabilités établi durant l'entracte permet seulement d'affirmer que la probabilité de se tromper en rejetant l'hypothèse que la série est simulée alors qu'elle l'est, est inférieure à 5%. L'activité présentée est plus complexe puisque l'on sait que l'une des séries est simulée et l'autre imaginée et que la décision porte sur la distinction entre les deux.

Si les valeurs de N des deux séries sont égales, ou toutes deux supérieures à 5, ou toutes deux inférieures à 5, l'élève se trouve en difficulté pour prendre une décision à partir des observations de N .

Comme cela a été détaillé précédemment, l'activité correspond à une situation de test statistique, pour laquelle prendre une décision est toujours relié à un risque d'erreur. Mais cela est une notion bien difficile à concevoir pour les élèves à qui on a toujours dit que les mathématiques étaient une science exacte. Pour cette raison, il faut éviter de pratiquer cette activité en classe juste avant un contrôle car elle peut être déstabilisante.

4.2 De nombreux développements

Cette activité peut avoir de nombreuses variantes ; certaines d'entre elles ont été présentées. On peut définir de nombreuses variables aléatoires à partir de cette situation et simuler leur distribution.

¹¹ Recopier une plage du type 0/1 avec un tableur peut donner de mauvaises surprises !

J.-Fr. Kentzel

Une généralisation en remplaçant $1/2$ par p quelconque entre 0 et 1 est possible avec la variable $N = N_n$, connue sous le nom de « longest head run ». On obtiendrait alors la formule $P(N_n \geq k) = P(N_{n-1} \geq k) + (1-p)p^k(1 - P(N_{n-k-1} \geq k))$.

La coïncidence décrite au point 3.3.1 ne subsiste pas quand $p \neq 1/2$.

Cette variable aléatoire N_n est étonnante car sa loi se concentre très fortement sur quelques valeurs : on trouve $Var(N_n) = \pi^2 / 6 Ln^2(1/p) + 1/12 + r(n) + \eta(n)$ où $r(n) \leq 0,00006$ et $\eta(n) \rightarrow 0$ quand $n \rightarrow +\infty$ (Gordon, 1986) ; par exemple, l'écart-type de N_{200} vaut environ 1,85 (pour $p = 1/2$).

5 Conclusion

Cette activité peut être déclinée sous de nombreux aspects et peut être la base de différents travaux. Elle permet d'introduire la notion d'aléatoire et discuter la perception peut-être biaisée que les élèves ont de la méthode de repérer le « hasard ». De plus, elle est une source très importante de définition de variables aléatoires, dont la recherche des lois peut être réalisée par simulation.

Mes élèves de classe de première (de S à STG) effectuent cette activité depuis quatre ans. Je ne m'en lasse pas car ils y trouvent un grand intérêt, à cause des trois moments très différents de l'activité et du résultat qui les surprend. Je la fais toujours en une heure mais c'est assez serré ; une heure trente ou deux fois trois quarts d'heure serait plus raisonnable, notamment en STG où la lecture des graphiques est moins aisée. Si on procède ainsi, la partie à développer est probablement l'entracte.

Cette activité est plus en adéquation avec les nouveaux programmes qu'avec les anciens. On rencontre en effet davantage, dans les commentaires des nouveaux programmes, des expressions comme « différence significative » ou « prise de décision », cependant que l'accent est maintenant mis très lourdement sur la loi binomiale en classe de première.

Un dernier détail : faute de temps, je n'insiste pas du tout sur le fait que la série engendrée par l'ordinateur est « pseudo-aléatoire », c'est-à-dire engendrée de façon déterministe à partir d'un algorithme.

Références

- [1] Bulletin officiel spécial n°9 du 30 septembre 2010, Programme d'enseignement spécifique de mathématiques en classe de première de la série scientifique, arrêté du 21-7-2010 - J.O. du 28-8-2010 (NOR MENE1019634A).
- [2] Flajolet, P. and R. Sedgewick (2009), *Analytic combinatorics*, Cambridge University Press, London, 52-53.
- [3] Gordon, L., M. F. Schilling, and M. S. Waterman (1986), An Extreme Value Theory for Long Head Runs, *Probab. Th. Rel. Fields*, **72**, 279-287.
- [4] Kentzel, J.-Fr. (2009), Différencier une suite aléatoire d'une autre qui ne l'est pas, *Bulletin APMEP*, **481**, 181-187.

Série simulée ou inventée : une activité instructive et ludique au lycée

- [5] Revesz, P. (1978), Strong theorems on coin tossing. In Proceedings of the International Congress of Mathematicians, *Acad. Sci. Fennica*, 749–754.
- [6] Mathématikos, Site Internet de Jean Paul Quelen, <http://jpb.pagesperso-orange.fr/> .
- [7] Delahaye, J.-P. (2006), Les dés pipés du cerveau, *Belin*.
- [8] <http://pardailhan.entmip.fr/rubrique-des-disciplines/mathematiques/documents-enseignants/doc-k-3359.htm>

Annexe : preuves des résultats $P(N \geq 5) \geq 0,96$ et $P(L \geq 6) \geq 0,96$

On donne une preuve de ces résultats valable en classe de première car n'utilisant que des dénombrements. Ces deux preuves sont quasiment identiques à cause de la proximité des lois de N et L .

1) On lance 200 fois une pièce bien équilibrée.

Pour tout entier n entre 1 et 200 (ou plus), on désigne par A_n l'événement :

$$\{N_n \geq 5\} = \{\text{il existe (au moins) une suite de 5 piles consécutifs à l'issue du } n^{\text{e}} \text{ lancer}\}.$$

En notant p_n pour $P(A_n)$, on prouve facilement que la suite $(p_n)_{n \geq 1}$ est définie par :

$$\text{si } 1 \leq n \leq 4, p_n = 0 ; p_5 = \frac{1}{2^5} ; p_6 = \frac{3}{2^6} ; \text{si } n \geq 7, p_n = p_{n-1} + \frac{1}{2^6}(1 - p_{n-6}). \quad (*)$$

((*) est vraie pour $n = 6$ en posant $p_0 = 0$, ce qui n'a pas de sens concret)

Pour démontrer cette égalité il suffit d'écrire que

$$P(A_n) = P(A_{n-1} \cup (\bar{A}_{n-1} \cap B_{n-1}))$$

où B_{n-1} est l'événement « les 4 derniers lancers sont des piles ».

On peut représenter $(\bar{A}_{n-1} \cap B_{n-1})$ comme ci-dessous :

Lancer	1	2	3		n-6	n-5	n-4	n-3	n-2	n-1	n
Résultat	Série quelconque sans succession de plus de quatre 1					0	1	1	1	1	1

$$P(A_n) = P(A_{n-1} \cup (\bar{A}_{n-1} \cap B_{n-1})) = P(A_{n-1}) + P(\bar{A}_{n-1} \cap B_{n-1}) = p_{n-1} + (1 - p_{n-6}) \cdot \frac{1}{2^6}$$

2) Pour tout entier n entre 1 et 200 (ou plus), on désigne par C_n l'événement :

$$\{L_n \geq 6\} = \{\text{il existe (au moins) une suite de 6 résultats consécutifs égaux à l'issue du } n^{\text{e}} \text{ lancer}\}.$$

En notant q_n pour $P(C_n)$, on prouve facilement que la suite $(q_n)_{n \geq 1}$ est définie par :

$$\text{si } 1 \leq n \leq 5, q_n = 0 ; q_6 = \frac{1}{2^5} ; \text{si } n \geq 7, q_n = q_{n-1} + \frac{1}{2^6}(1 - q_{n-6}). \quad (**)$$

On peut voir une preuve plus formalisée et utilisant l'indépendance des lancers sur les [pages](#) de l'excellent site de Jean Paul Quelen.