Les allumettes de Banach

Résumé

Ce document accompagne la page Web consacrée aux simulations du problème des deux boîtes d'allumettes de Banach.
On pourra comparer les valeurs obtenues lors de ces simulations aux valeurs théoriques calculées ci-dessous.
Le problème modifié est étudié en premier et la solution est trouvée par deux logiciels de calcul symbolique MuPAD et Maxima.

1 Problème modifié (cas 1)

1.1 Énoncé.

Banach,[1] qui était fumeur, avait deux boîtes d'allumettes, l'une dans la poche gauche, l'autre dans la poche droite.
Chaque fois qu'il allumait sa pipe, il puisait au hasard dans l'une des boîtes. Immanquablement, au bout d'un certain temps, il prenait la dernière allumette d'une boîte.
On suppose que les deux boîtes sont changées simultanément et qu'elles contiennent initialement le même nombre n d'allumettes.
Quel est, en moyenne, le nombre d'allumettes restantes, dans l'autre boîte, dès que l'une est vide ?

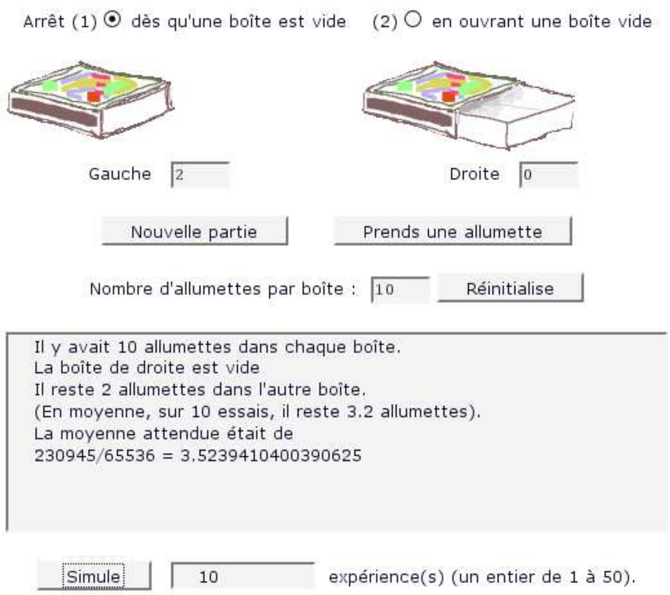

FIGURE 1 – *Simulations*

1.2 Solution

En notant G $= (0, 1)$ ou D$= (1, 0)$ le retrait d'une allumette de la boîte gauche ou droite, les événements sont les mots de la forme DGGDGDDDGD...GDG ou les chemins (Fig : 2) correspondants sur le quadrillage depuis l'origine $(0, 0)$ jusqu'à l'un des bords supérieurs du quadrillage ci-dessus, (n, n)

1. Ce texte est différent de l'énoncé classique. Habituellement l'arrêt n'a lieu qu'au moment où l'on ouvre une boîte vide (cf. cas 2). Les espérances mathématiques des deux problèmes sont liées par la relation $E_2(n) = E(n + 1) - 1$.

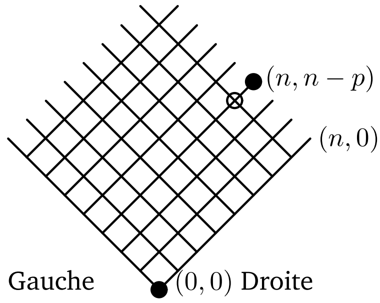

FIGURE 2 – *Marche aléatoire dans un carré de* $\mathbb{N} \times \mathbb{N}$

excepté.
Lorsqu'il reste p allumettes dans l'une des boîtes et aucune dans l'autre, ces suites ont pour longueur $n + (n-p)$.
La lettre de droite du mot (la dernière lettre écrite) est G ou D selon que la boîte vide est celle de gauche ou de droite.
Lorsque la dernière lettre du mot est G (resp. D), les suites contiennent n fois la lettre G (resp. D) et $n-p$ fois l'autre lettre D (resp. G).
Il faut tenir compte du fait que dans chaque cas la lettre finale du mot est imposée. Le nombre de telles suites est dans chacun des deux cas le nombre de combinaisons de $n-1$ lettres parmi $2n-p-1$.
Chacune de ces suites, de longueur $2n-p$ est obtenue avec une probabilité égale à $2^{-(2n-p)}$.
L'espérance mathématique $E(n)$ est donc égale à

$$E(n) = 2 \times \sum_{p=1}^{p=n} p \times \frac{\binom{2n-p-1}{n-1}}{2^{2n-p}} = \sum_{p=1}^{p=n} p \times \frac{\binom{2n-p-1}{n-1}}{2^{2n-p-1}}$$

1.3 Calculs

Quelques calculs effectués par le système de calcul formel très répandu PARI_{GP} sont indiqués à la figure (Fig : 3). Des résultats de simulations sont donnés à la figure (Fig : 4)

```
? E(n)=sum(p=0,n,p*binomial(2*n-p-1,n-1)/2^(2*n-p-1));
? E(5)
%1 = 315/128
```

FIGURE 3 – *Calculs à l'aide de PARI/GP*

n	$E(n)$ (valeurs attendues)	moyennes de 50 simulations
5	2.4609375	2.34
10	3.52394104	3.34
20	5.014827505	4.38
30	6.154690381	5.42
40	7.114230302	6.9
50	7.958923739	7.98

FIGURE 4 – *Valeurs obtenues lors des simulations*

1.4 Expression sous la forme d'un produit de fractions

On peut montrer que $\dfrac{E(n+1)}{E(n)} = \dfrac{2n+1}{2n} = 1 + \dfrac{1}{2n}$
Les algorithmes permettant d'obenir une formule close pour l'identité (1.2) sont décrits dans le livre A=B de Marko Petkovsek, Herbert Wilf et Doron Zeilberger.
On obtient alors très facilement l'espérance mathématique ci-après (5).

Les numérateurs et dénominateurs successifs sont les suites A001803 et A046161 de l'Encyclopédie des suites entières de Sloane. La suite A002457 correspond à $4^{n-1}E(n)$ qui est aussi égal à $\dfrac{(2n-1)!}{(n-1)!^2}$.

n	1	2	3	4	5	6
$E(n)=\frac{1.3.5.7...(2n-1)}{1.2.4.6...(2n-2)}$	1	$\frac{3}{2}$	$\frac{15}{8}$	$\frac{35}{16}$	$\frac{315}{128}$	$\frac{693}{256}$

FIGURE 5 – *Espérance mathématique*

```
*----*    MuPAD 2.0.0 -- The Open Computer Algebra System
...
>> ?sum
sum -- definite and indefinite summation
...
Background
 o The function sum implements Abramov's algorithm for rational functions,
   Gosper's algorithm for hypergeometric functions, and Zeilberger's
   algorithm for the definite summation of holonomic functions.
...
>> a(n):=sum(p*binomial(2*n-p-1,n-1)/2^(2*n-p-1),p=0..n)
                                   n
(4 n binomial(2 n - 1, n - 1) + 4 2  binomial(n - 2, n - 1) -
     n                           2 n
4 n 2  binomial(n - 2, n - 1)) / 2
>> expand((1-n)*binomial(n - 2, n - 1))
                                    0
```

FIGURE 6 – *Calcul de l'espérance par MuPAD 2.0.0*

(On trouvera quelques références sur le site, en particulier les marches aléatoires sont évoquées).
À partir de la somme $E(n)=\sum_{p=0}^{p=n} p\times\frac{\binom{2n-p-1}{n-1}}{2^{2n-p-1}}$, les logiciels de calcul symboliques MuPAD (Fig : 6) et Maxima (Fig : 7) calculent l'expression $E(n)=\frac{4n\binom{2n-1}{n-1}}{2^{2n}}=4^{-(n-1)}\frac{(2n-1)!}{(n-1)!^2}$.

$$E(n)=\frac{(2n-1)!}{4^{n-1}\times(n-1)!^2}=\frac{(2n-1)!}{(2^{n-1}\times(n-1)!)^2}=\frac{1.3.5.7\ldots(2n-1)\times 2.4.6\ldots(2n-2)}{(2.4.6\ldots(2n-2))^2}=\frac{1.3.5.7\ldots(2n-1)}{1.2.4.6\ldots(2n-2)}.$$

1.5 Fonction génératrice

Comme $\frac{1}{\sqrt{1-x^2}}=1+\frac{1}{2}x^2+\frac{1.3}{2.4}x^4+\frac{1.3.5}{2.4.6}x^6+\ldots+\frac{1.3.5\ldots(2n-1)}{2.4.6\ldots 2n}x^{2n}+\ldots$

on trouve aisément à l'aide d'une dérivation et une multiplication par x^2 que
$x^2\,(1-x^2)^{\frac{-3}{2}}=1\,x^2+\frac{3}{2}x^4+\frac{15}{8}x^6+\frac{35}{16}x^8+\frac{315}{128}x^{10}+\frac{693}{256}x^{12}+\ldots+\frac{3.5.7\ldots(2n-1)}{2.4.6\ldots(2n-2)}x^{2n}+\ldots,$
$x\,(1-x)^{\frac{-3}{2}}=1\,x+\frac{3}{2}x^2+\frac{15}{8}x^3+\frac{35}{16}x^4+\frac{315}{128}x^5+\frac{693}{256}x^6+\ldots+\frac{3.5.7\ldots(2n-1)}{2.4.6\ldots(2n-2)}x^n+\ldots$
dont le coefficient de x^n est $E(n)$.
$x\,(1-x)^{\frac{-3}{2}}$ est donc la fonction génératrice de l'espérance $E(n)$.

2 Problème classique (cas 2)

2.1 Espérance mathématique déduite du cas précédent

Si on ajoute une allumette (fictive) dans chaque boîte on passe du problème classique au problème modifié, il suffit de retrancher 1 de l'espérance $E(n+1)$ du problème modifié à $n+1$ allumettes par boîte, pour obtenir celle $E_2(n)=E(n+1)-1$ du problème classique à n allumettes. L'allumette fictive étant par exemple celle que l'on (ne) prend (pas) à la dernière ouverture de la boîte.

$$E_2(n)=E(n+1)-1=-1+\sum_{p=1}^{p=n+1} p\times\frac{\binom{2n-p+1}{n}}{2^{2n-p+1}}=\sum_{p=1}^{p=n+1}(p-1)\times\frac{\binom{2n-p+1}{n}}{2^{2n-p+1}}=\sum_{k=0}^{k=n}k\binom{2n-k}{n}2^{-(2n-k)}$$

(La somme des probabilités est 1).
On obtient aussi $E_2(n)=E(n+1)-1=\frac{1.3.5.7...(2n+1)}{1.2.4.6...(2n)}-1=\frac{2n+1}{2^{2n}}\binom{2n}{n}-1$ qui est bien plus simple que la relation de récurrence $E_2(n)=\frac{2n+1}{2n}\,E_2(n-1)+\frac{1}{2n}$.

```
Maxima 5.9.0 http://maxima.sourceforge.net
...
(C1) describe(nusum);
Info from file /usr/local/info/maxima.info:
 - Function: NUSUM (exp,var,low,high)
     performs indefinite summation of exp with respect to var using a
     decision procedure due to R.W. Gosper.  exp and the potential
     answer must be expressible as products of nth powers, factorials,
     binomials, and rational functions.  The terms "definite" and
     "indefinite summation" are used analogously to "definite" and
     "indefinite integration".  To sum indefinitely means to give a
     closed form for the sum over intervals of variable length, not
     just e.g. 0 to inf.  Thus, since there is no formula for the
     general partial sum of the binomial series, NUSUM can't do it.
...
(C2) nusum(p*binomial(2*n-p-1,n-1)/2^(2*n-p-1),p,0,n);

Dependent equations eliminated:  (1)
                          4 n BINOMIAL(2 n - 1, n - 1)
(D2)                      ----------------------------
                                       2 n
                                      2
```

FIGURE 7 – *Calcul de l'espérance par Maxima*

```
? taylor((1-x)^(-3/2)-(1-x)^(-1),x)
%1 = 1/2*x + 7/8*x^2 + 19/16*x^3 + 187/128*x^4 + 437/256*x^5 +
1979/1024*x^6 + 4387/2048*x^7 + 76627/32768*x^8 + 165409/65536*x^9 +
707825/262144*x^10 + 1503829/524288*x^11 + 12706671/4194304*x^12 +
26713417/8388608*x^13 + 111868243/33554432*x^14 +
233431331/67108864*x^15 + O(x^16)
```

FIGURE 8 – *Fonction génératrice*

2.2 Solution directe

En considérant les chemins du quadrillage (Voir 2), on montre d'abord que $E_2(n) = \sum_{k=0}^{k=n} k\binom{2n-k}{n}2^{-(2n-k)}$

On considère la série $\sum_{i\geq 0} z^i = \frac{1}{1-z}$.

En dérivant n fois on obtient $\sum_{i\geq 0} \binom{n+i}{n} z^i = \frac{1}{(1-z)^{n+1}}$.

Puis en faisant $i = n-k$ et $z = \frac{1}{2}$ on a $\sum_{k\geq 0}^{n} \binom{2n-k}{n}2^{-(n-k)} = 2^{n+1}$.

2.3 Fonction génératrice

La fonction génératrice est $(1-x)^{\frac{-3}{2}} - (1-x)^{-1} = \frac{1}{(1-x)\sqrt{1-x}} - \frac{1}{1-x} = \frac{1}{1-x}\left(\frac{1}{\sqrt{1-x}} - 1\right)$.
Les espérances $E_2(n)$ sont les coefficients des x^n dans le développement en série $\sum_n E_2(n)x^n$ de cette fonction génératrice.

Un calcul de PARI$_{\text{GP}}$ (Fig : 8) donne bien les valeurs attendues.

2.4 Fonction génératrice d'une loi de probabilité

Au chapitre 8 sur les Probabilités discrètes du livre Mathématiques concrètes de R.L. Graham , Donald E. Knuth , O. Patashnik le problème est résolu à l'exercice 48 en considérant a) une fonction génératrice double $P(w,z) = \sum_{m,n} p_{m,n} w^m z^n$ où $p_{m,n}$ est la probabilité que les deux boîtes soient vides lorsque la première choisie est vide (les deux boîtes contenant initialement m et n allumettes), b) la probabilité $P_{k,m,n}$ que la seconde boîte contienne exactement k allumettes lorsque la première choisie est vide.

La question c) permet d'obtenir la formule close $\sum_k k\ P_{k,m,n} = \frac{2n+1}{2^{2n}}\binom{2n}{n} - 1$ allumettes dans la seconde boîte en utilisant la propriété $\sum_{k\geq m}\binom{m+k}{k}2^{-k} = 2^m$.

Les auteurs signalent que l'espérance vaut $2\sqrt{n/\pi} - 1 + \mathrm{O}(n^{-1/2})$ (Fig : 9).

n	$E_2(n)$	$2\sqrt{n/\pi} - 1$
10	2.70014	2.56825
50	7.03851	6.97885
100	10.326	10.2838
200	14.9876	14.9577
300	18.5685	18.5441
400	21.5887	21.5676
500	24.2502	24.2313
600	26.6568	26.6395
700	28.8701	28.8541
800	30.9303	30.9154
900	32.8655	32.8514
1000	34.6959	34.6825

FIGURE 9 – *Valeurs approchées*

Table des matières

Table des figures