Le problème des sucres du Grand Khong

Aimé Lachal

15 août 2008

1 Prologue

Le problème des sucres du Grand Khong peut est décrit de manière alternative selon une allégorie mystico-sibylline du ploutocrate sortant d'un pandémonium de sybarites dévoyés où il a croisé quelque échanson dionysiaque, et qui cherche à rejoindre bon an mal an son gynécée. Modélisation :

$$\text{point } A = \text{pandémonium, point } B = \text{gynécée, distance de } A \text{ à } B = 666 \text{ m.}$$

À chaque mètre parcouru, le ploutocrate se pose le problème cornélien suivant : doit-il faire un mètre en avant ou un mètre en arrière ? Le choix est fait au hasard. La question posée est la suivante : en lui imposant la règle de rebrousser chemin dès que l'un des deux points A et B est atteint (il oscille donc entre A et B), quelle est la probabilité que le ploutocrate se trouve à 333 m de A (ou de B) au bout de 123456789 pas (d'un mètre) ? au bout de 1234567890 pas ? (NDLR : la réponse à la deuxième question est triviale ! En revanche, celle de la première l'est nettement moins...)

C'est la célèbre marche aléatoire du dipsomane impénitent (i.e. ivrogne) ou du périssodactyle capricant (i.e. bourrin), ou encore du sphéniscidé claudiquant (i.e. manchot)... Dans le cas de la marche aléatoire classique, les probabilités d'aller en avant ou en arrière sont constantes au cours du temps. Dans notre problème, elles dépendent de la position du ploutocrate au k^{e} pas :

- probabilité de se mouvoir en avant : k/N ;
- probabilité de se mouvoir en arrière : $1 - k/N$.

Il s'agit du modèle d'urnes d'Ehrenfest (utilisé en thermodynamique pour étudier les échanges thermiques), c'est une chaîne de Markov célèbre.

2 Le problème du Grand Khong

Le Grand Khong choisit une boîte au hasard, puis :

- s'il y a un demi-sucre dans cette boîte, il le prend et il y en a un de moins dans la totalité des boîtes ;
- s'il n'y a pas de demi-sucre dans cette boîte, il casse un sucre entier, prend un des deux demi-sucre et laisse l'autre. Il reste un demi-sucre dans la totalité des boîtes.

Ainsi, si le soir du n^{e} jour, il y a k demi-sucres dans la totalité des boîtes, le soir du $(n+1)^{\text{e}}$ jour, il y a soit $k+1$ demi-sucres soit $k-1$ demi-sucres. Il reste donc

- $k-1$ demi-sucres avec la probabilité d'avoir choisi une bonne boîte (il y a k telles boîtes parmi les N), probabilité $= k/N$;

- $k+1$ demi-sucres avec la probabilité d'avoir choisi une mauvaise boîte (il y en a $N-k$ parmi les N), probabilité $= (N-k)/N = 1-k/N$.

Cas particuliers : le raisonnement précédent est valable pour $1 \leqslant k \leqslant N-1$.
- Pour $k=0$, il n'y a pas de demi-sucre, donc quelle que soit la boîte choisie, le Grand Khong doit casser un sucre entier, il reste donc ensuite un demi-sucre dans la totalité des boîtes (et ce avec la probabilité 1).
- Pour $k=N$, toutes les boîtes contiennent un demi-sucre, le Grand Khong n'a pas à casser de sucre, il prend le demi-sucre présent dans la boîte choisie et il reste $N-1$ demi-sucres dans la totalité des boîtes (et ce avec la probabilité 1).

La table 1 ci-dessous résume la discussion précédente.

jour	n^{e}	$(n+1)^{\mathrm{e}}$	probabilité
nombre de demi-sucres	0	1	1
	k $(1\leqslant k\leqslant N-1)$	$\nearrow k-1$ $\searrow k+1$	k/N $1-k/N$
	N	$N-1$	1

TAB. 1 – Partant du n^{e} jour

Notons $p_{k,n} = \text{Prob(le soir du } n^{\mathrm{e}} \text{ jeu, il y a } k \text{ demi-sucres)}$.

Remarque :
- au démarrage du processus (soir du $n = 0^{\mathrm{e}}$ jour), il n'y a pas de demi-sucre ;
- au soir du $n = 1^{\mathrm{er}}$ jour, il y a un demi-sucre ;
- au soir du $n = 2^{\mathrm{e}}$ jour, il y a 0 ou 2 demi-sucres ;
- au soir du $n = 3^{\mathrm{e}}$ jour, il y a 1 ou 3 demi-sucres ;
- au soir du $n = 4^{\mathrm{e}}$ jour, il y a 0, 2 ou 4 demi-sucres, etc.

On voit de proche en proche que le numéro n du jour et le nombre k de demi-sucres ont la même parité. Ainsi, lorsque n et k n'ont pas la même parité, $p_{k,n} = 0$. On peut ainsi répondre d'ores et déjà à la question correspondant au cas $n = 1234567890$: $k = 333$ et n n'ont pas la même parité, donc $p_{333,n} = 0$, la probabilité demandée est nulle !

Examinons la situation le soir du $(n+1)^{\mathrm{e}}$ jour : il y a k demi-sucres. À l'aide de la table 1, on voit qu'il y avait donc la veille $k-1$ demi-sucres (avec probabilité $1-(k-1)/N$) ou $k+1$ demi-sucres (avec probabilité $(k+1)/N$). Voir table 2.

jour	n^{e}	$(n+1)^{\mathrm{e}}$	probabilité
nombre de demi-sucres	1	0	1/N
	$k-1$ $k+1$	$\searrow$ $\nearrow$ k	$(k+1)/N$ $1-(k-1)/N$
	$N-1$	N	1/N

TAB. 2 – Partant du $(n+1)^{\mathrm{e}}$ jour

Cette discussion nous permet d'exprimer $p_{k,n+1}$ en fonction de $p_{k-1,n}$ et $p_{k+1,n}$:

$$\begin{aligned}
p_{k,n+1} &= \text{Prob(le soir du } (n+1)^{\text{e}} \text{ jour, il y a } k \text{ demi-sucres)}\\
&= \text{Prob(le soir du } n^{\text{e}} \text{ jour, il y a } k+1 \text{ demi-sucres et Khong a choisi une bonne boîte)}\\
&\quad + \text{Prob(le soir du } n^{\text{e}} \text{ jour, il y a } k-1 \text{ demi-sucres et Khong a choisi une mauvaise boîte)}\\
&= \frac{k+1}{N}\,\text{Prob(le soir du } n^{\text{e}} \text{ jour, il y a } k+1 \text{ demi-sucres)}\\
&\quad + \left(1-\frac{k-1}{N}\right)\text{Prob(le soir du } n^{\text{e}} \text{ jour, il y a } k-1 \text{ demi-sucres)}
\end{aligned}$$

d'où la récurrence, pour $1 \leqslant k \leqslant N-1$:

$$p_{k,n+1} = \frac{k+1}{N}\,p_{k+1,n} + \left(1-\frac{k-1}{N}\right)p_{k-1,n}.$$

Pour $k=0$ et $k=N$, on a les relations

$$p_{0,n+1} = \frac{1}{N}\,p_{1,N} \quad \text{et} \quad p_{N,n+1} = \frac{1}{N}\,p_{N-1,N}.$$

3 Résolution du problème

Les données du problème sont les suivantes : $N=666$, $k=333$ et $n=123456789$. Comme n est grand, on pourrait considérer que $p_{k,n}$ et $p_{k,n+1}$ sont sensiblement identiques : grave erreur ! car nous avons remarqué que les parités de k et de n jouaient un rôle important. En fait, on peut considérer que $p_{k,n}$ et $p_{k,n+2}$ sont sensiblement identiques. Formalisons ceci en distinguant les n pairs et les n impairs : on pose

$$q_k = \lim_{n\to\infty,\, n \text{ pair}} p_{k,n} \quad \text{et} \quad r_k = \lim_{n\to\infty,\, n \text{ impair}} p_{k,n}.$$

La remarque faite sur la parité montre que $q_k = 0$ pour les k impairs et que $r_k = 0$ pour les k pairs. D'autre part, comme les $p_{k,n}, 0 \leqslant k \leqslant N$, constituent une loi de probabilité, on a la relation $\sum_{k=0}^{N} p_{k,n} = 1$ et donc

$$\sum_{k=0}^{N} q_k = 1 \quad \text{et} \quad \sum_{k=0}^{N} r_k = 1.$$

La relation de récurrence de la section précédente donne, à la limite, le système

$$\left\{\begin{aligned}
r_k &= \frac{k+1}{N}\,q_{k+1} + \left(1-\frac{k-1}{N}\right)q_{k-1}\\
q_k &= \frac{k+1}{N}\,r_{k+1} + \left(1-\frac{k-1}{N}\right)r_{k-1}
\end{aligned}\right.$$

avec les conditions

$$r_0 = \frac{1}{N}\,q_1 \quad \text{et} \quad q_0 = \frac{1}{N}\,r_1.$$

Introduisons ensuite $u_k = q_k + r_k$. On a

$$u_k = \left\{\begin{array}{ll} q_k & \text{si } k \text{ est pair,}\\ r_k & \text{si } k \text{ est impair.}\end{array}\right.$$

En ajoutant les deux équations du système précédent, on obtient

$$u_k = \frac{k+1}{N}\, u_{k+1} + \left(1 - \frac{k-1}{N}\right) u_{k-1}$$

avec $u_0 = \frac{1}{N} u_1$ et $\sum_{k=0}^{N} u_k = 2$. Invoquons à présent une petite astuce d'initié : on écrit le coefficient 1 devant u_k sous la forme $1 = \frac{k}{N} + (1 - \frac{k}{N})$, ce qui conduit à

$$\frac{k+1}{N}\, u_{k+1} - \left(1 - \frac{k}{N}\right) u_k = \frac{k}{N}\, u_k - \left(1 - \frac{k-1}{N}\right) u_{k-1}.$$

On voit ainsi que la suite de terme général $u'_k = \frac{k+1}{N}\, u_{k+1} - (1 - \frac{k}{N}) u_k$ est constante puisque $u'_k = u'_{k-1}$ pour tout k, constante de valeur $u'_1 = \frac{1}{N}\, u_1 - u_0 = 0$. Ainsi, pour tout k, $u'_k = 0$, puis $\frac{k+1}{N}\, u_{k+1} - (1 - \frac{k}{N}) u_k = 0$, soit

$$u_k = \frac{N-k+1}{k}\, u_{k-1}$$

qui donne, en extrapolant,

$$\begin{aligned} u_k &= \frac{N+1-k}{k}\, u_{k-1} = \frac{(N-k+2)(N-k+1)}{k(k-1)}\, u_{k-2} \\ &= \cdots = \frac{N(N-1)\cdots(N-k+2)(N-k+1)}{k(k-1)\cdots 2.1}\, u_0 = \binom{N}{k} u_0. \end{aligned}$$

Enfin, la condition $\sum_{k=0}^{N} u_k = 2$ et l'identité $\sum_{k=0}^{N} \binom{N}{k} = 2^N$ donnent $u_0 = 2/2^N$ puis

$$u_k = \frac{\binom{N}{k}}{2^{N-1}}.$$

Ainsi, puisque $u_k = q_k$ lorsque k pair et $u_k = r_k$ lorsque k impair, on obtient l'approximation pour n grand de même parité que k :

$$\boxed{p_{k,n} \approx \frac{\binom{N}{k}}{2^{N-1}}.}$$

Lorsque k et n n'ont pas la même parité, on a tout simplement $p_{k,n} = 0$.

Réponse au problème posé : pour $k = 333$, $N = 666$ et $n = 123456789$, k et n ont la même parité et on peut considérer que n est grand (!), donc

$$p_{k,n} \approx \frac{\binom{666}{333}}{2^{665}} = \frac{666!}{333!^2 \times 2^{665}} \approx 0,06181159056 \approx 61812 \times 10^{-6}.$$

Pour $n = 1234567890$, comme nous l'avions anoncé, la réponse est triviale : $p_{k,n} = 0$!

Les probabilités recherchées sont donc respectivement 061812 millionièmes et 0.