COMMENT CHOISIR PARMI N OBJETS AVEC UN SUPER-DÉ À K FACES

par *Denis Labelle*
Université de Montréal

On veut choisir équitablement un objet parmi n au moyen d'un super-dé (ou d'une toupie) à k faces équiprobables. Il y a évidemment une infinité d'algorithmes qui permettent d'effectuer ce choix. Par exemple, si on veut choisir parmi 4 objets avec un dé ordinaire à 6 faces, il y a, entre autres, les trois algorithmes suivants:

ALGORITHME 1

Premier lancer	Second lancer	Objet choisi
1,2,3	1,2,3	1
	4,5,6	2
4,5,6	1,2,3	3
	4,5,6	4

ALGORITHME 2

Premier lancer	Second lancer	Objet choisi
1	...	1
2	...	2
3	...	3
4	...	4
5	R.A.D.	
6	R.A.D.	

ALGORITHME 3

Premier lancer	Second lancer	Objet choisi
1	...	1
2	...	2
3	...	3
4	...	4
5	1,2,3	1
	4,5,6	2
6	1,2,3	3
	4,5,6	4

(note: R.A.D. signifie "retour au départ")

Un algorithme de choix équitable étant donné, on note N le nombre de lancers du super-dé nécessaires pour exécuter cet algorithme. N est une variable aléatoire dont on calcule aisément l'espérance mathématique E(N).

Dans l'exemple précédent où $n = 4$ et $k = 6$ on obtient que, pour les trois algorithmes considérés, $E(N)$ vaut respectivement 2, 3/2 et 4/3. Le troisième algorithme est donc meilleur (plus économique) que les deux autres puisque son exécution demande, en moyenne, moins de lancers du dé.

Théorème 1 :

Pour tout algorithme de choix équitable et $\forall i \geqslant 0$,

$$\mathrm{Prob}(N > i) \geqslant \frac{k^i \pmod n}{k^i}$$

(la notation $k^i \pmod n$ désigne le reste de la division de k^i par n).

Démonstration: Nous donnerons une preuve par contradiction. Supposons qu'il existe un $n > 0$, un $k > 0$ et un $i \geqslant 0$ tels que

$$\mathrm{Prob}(N > i) < \frac{k^i \pmod n}{k^i} .$$

On constate alors que i lancers donnent k^i résultats (ou chemins) possibles. Si on écrit k^i sous la forme $k^i = nq + r$ (où q et r sont respectivement le quotient entier et le reste de la division de k^i par n) on constate que notre supposition initiale implique que plus de $k^i - r = nq$ chemins possibles (sur le total de k^i) ont déjà été affectés à la désignation des objets. Au moins un de ces n objets devra donc avoir été désigné par au moins $q + 1$ chemins.

La probabilité de cet objet sera donc d'au moins

$$\frac{q+1}{k^i} = \frac{1}{n}\left(\frac{nq+n}{k^i}\right) > \frac{1}{n}\left(\frac{nq+r}{k^i}\right) = \frac{1}{n}$$

La probabilité que cet objet soit choisi est donc supérieure à $\frac{1}{n}$ et l'algorithme utilisé n'est donc pas équitable, ce qui contredit la propriété qu'on exige des algorithmes de choix.

Théorème 2 :

$\forall n \geqslant 2$ et $\forall k \geqslant 2$,
il existe un algorithme de choix équitable pour lequel

$$\text{Prob}(N > i) = \frac{k^i \pmod{n}}{k^i} \qquad \forall i \geqslant 0 .$$

<u>*Démonstration:*</u> On obtiendra cet algorithme par construction récurrente.
$\forall i \geqslant 0$, on peut écrire k^i sous la forme $k^i = nq_i + r_i$ où $r_i = k^i \pmod{n}$. Puisque $n \geqslant 2$, un premier lancer est nécessaire. Donc $\text{Prob}(N > 0) = 1$ et l'égalité est satisfaite pour le cas où $i = 0$.

Le premier lancer donne $k = nq_1 + r_1$ résultats (ou chemins) possibles. On affecte tout de suite q_1 chemins à chacun des n objets. On trouve alors que seulement r_1 chemins parmi k n'ont pas encore désigné d'objet.

Donc $\text{Prob}(N > 1) = \dfrac{r_1}{k} = \dfrac{k \pmod{n}}{k}$ et l'égalité est encore satisfaite pour le cas où $i = 1$.

Supposons l'algorithme ainsi construit jusqu'au i-ième lancer et possédant la propriété que $\text{Prob}(N > i) = \dfrac{k^i \pmod{n}}{k^i} = \dfrac{r_i}{k^i}$. Exactement r_i chemins n'ont donc pas encore été affectés à la désignation des objets. Le $(i+1)$-ième lancer développe donc ces r_i chemins en kr_i nouveaux chemins. En posant alors $kr_i = nq + r$, on affecte tout de suite q chemins à chacun des n objets. On obtient donc que seulement r nouveaux chemins n'ont pas encore été affectés aux objets (après $i+1$ lancers). On trouve alors que

$$\text{Prob}(N > i+1) = \frac{r}{k^{i+1}} .$$

Or $r = kr_i \pmod{n} = k^2 r_{i-1} \pmod{n} = \dots = k^{i+1}\, r^0 \pmod{n} = k^{i+1} \pmod{n}$ et l'égalité est encore conservée en passant de i à $i+1$, ce qui complète la démonstration.

Notation

On notera $F(n,k)$ l'espérance mathématique du nombre de lancers du meilleur algorithme (non nécessairement unique) permettant de choisir parmi n objets avec un super-dé à k faces.

Théorème 3 :

$$\forall n \geqslant 2 \quad \text{et} \quad \forall k \geqslant 2, \; F(n, k) = \sum_{i=0}^{\infty} \frac{k^i (\text{mod } n)}{k^i}$$

Démonstration: Utilisant la formule générale $E(N) = \sum_{i=0}^{\infty} \text{Prob}(N > i)$ (valable pour toute loi de probabilité sur les entiers positifs) on obtient directement du théorème 1 que quel que soit l'algorithme de choix équitable utilisé, $E(N) \geqslant \sum_{i=0}^{\infty} \frac{k^i (\text{mod } n)}{k^i}$. Or le théorème 2 nous permet de construire un algorithme de choix équitable pour lequel cette expression devient uné égalité. Cet algorithme est donc optimal et, par définition, le nombre moyen de lancers nécessaires à son exécution est $F(n,k)$, ce qui complète la démonstration du théorème 3.

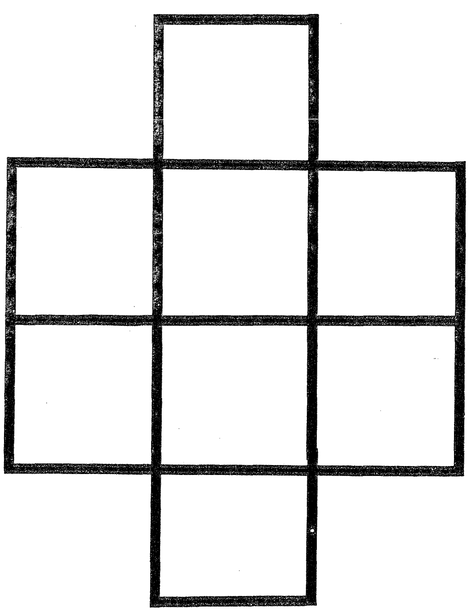
REMARQUE: La formule du théorème 3 s'applique encore même si $n = 1$ ou $k = 1$. En effet, si $n = 1$, aucun lancer du super-dé n'est nécessaire pour choisir cet unique objet. La formule donne alors $F(1,n) = 0$. De même, si $k = 1$ (et $n > 1$), il sera impossible de choisir un objet avec ce dé dégénéré qui ne donne aucune information. La série du théorème 3 est alors divergente et on trouve $F(n,1) = \infty$ (pour $n > 1$).

EXEMPLE: Pour choisir parmi 5 objets avec un super-dé à 3 faces, l'algorithme suivant est optimal.

Premier lancer	Second lancer	Troisième lancer	Quatrième lancer	Objet choisi
1	1			1
	2			2
	3			3
2	1			4
	2			5
	3	1		1
		2		2
		3		3
3	1	1		4
		2		5
		3		1
	2	1		2
		2		3
		3		4
	3	1		5
		2	1	1
			2	2
			3	3
		3	1	4
			2	5
			3	(retour au départ)

Le théorème 3 permet alors de calculer que $F(5,3) = \frac{51}{20} = 2.55$

La grille



Placer les nombres: 1 à 8 dans les cases en respectant les règles suivantes:

1) deux nombres consécutifs ne peuvent occuper des cases adjacentes;
2) deux nombres consécutifs ne peuvent occuper des cases placées en diagonale.

Journal A.M.Q

région de l'Estrie

Table partielle des valeurs de la fonction F(n,k)

	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10	k=11	k=12
n=2	1.000	1.500	1.000	1.250	1.000	1.167	1.000	1.125	1.000	1.100	1.000
n=3	2.667	1.000	1.333	1.458	1.000	1.167	1.270	1.000	1.111	1.192	1.000
n=4	2.000	2.250	1.000	1.250	1.333	1.458	1.000	1.125	1.200	1.283	1.000
n=5	3.600	2.550	2.133	1.000	1.200	1.377	1.442	1.462	1.000	1.100	1.196
n=6	3.667	2.500	2.333	2.083	1.000	1.167	1.317	1.375	1.444	1.467	1.000
n=7	3.429	2.518	2.159	2.212	2.057	1.000	1.143	1.273	1.326	1.381	1.448
n=8	3.000	2.250	2.000	2.083	2.111	2.042	1.000	1.125	1.240	1.283	1.333
n=9	4.667	2.000	2.476	2.351	2.000	2.088	2.032	1.000	1.111	1.221	1.250
n=10	4.600	3.300	2.467	2.250	2.200	2.193	2.068	2.025	1.000	1.100	1.199
n=11	4.848	3.248	2.467	2.167	2.124	2.110	2.154	2.054	2.020	1.000	1.091
n=12	4.667	3.250	2.333	2.083	2.000	2.042	2.079	2.125	2.044	2.017	1.000
n=13	4.708	3.115	2.421	2.548	2.322	2.223	2.198	2.040	2.102	2.037	2.014
n=14	4.429	3.643	2.270	2.560	2.257	2.167	2.143	2.139	2.027	2.076	2.031
n=15	4.267	3.550	2.133	2.458	2.200	2.120	2.067	2.087	2.111	2.017	2.065
n=16	4.000	3.450	2.000	2.468	2.148	2.042	2.000	2.025	2.048	2.077	2.000
n=17	5.765	3.580	3.216	2.394	2.225	2.317	2.211	2.184	2.164	2.021	2.063
n=18	5.667	3.350	3.190	2.440	2.000	2.272	2.175	2.125	2.111	2.121	2.000
n=19	5.723	3.432	3.164	2.359	2.509	2.231	2.148	2.073	2.063	2.059	2.087
n=20	5.600	3.300	3.133	2.250	2.533	2.193	2.090	2.025	2.000	2.017	2.033

La fonction F(n,k) possède plusieurs propriétés intéressantes. Je laisse au lecteur curieux le soin de démontrer celles que voici: (Il est chaque fois sous-entendu que n, m et k sont des entiers $\geqslant 2$).

1. Si $n = mk$ alors $F(n,k) = F(m,k) + 1$.

2. Pour tout n, m et k, $F(n,k) \leqslant F(n,m)F(m,k)$.

3. Les fonctions $D(n,k) = \log(F(n,k)) + \log(F(k,n))$ et
   $d(n,k) = F(n,k) + F(k,n) - 2$
   satisfont les quatre axiomes d'une distance sur $\{2,3,4,\ldots\}$.

4. $n < k \Longrightarrow F(n,k) \leqslant 3/2$, l'égalité n'étant atteinte que pour le couple $(n,k) = (2,3)$.