Temps d'attente

On dispose de *K* boules portant des numéros de 1 à *K*. On effectue des tirages au hasard, de façon répétée, avec remise de la boule après chaque tirage, jusqu'à ce que tous les numéros soient sortis au moins une fois chacun. Par exemple pour *K*=4 on peut avoir les 10 tirages successifs 2332342441. On appelle *X* le nombre des tirages effectués. Déterminer la loi de *X*, c'est-à-dire la probabilité *p*(*X*=*n*) d'effectuer *n* tirages, *n* étant un nombre supérieur ou égal à *K*, ainsi que la valeur moyenne de *X*.

Nous allons ici donner une méthode de résolution utilisant la méthode du crible. On trouvera dans le livre « Combien ? tome 2 : probabilités » *une autre méthode, celle utilisant les graphes de transition.*

Considérons tous les évènements correspondant à n tirages successifs et comptons-les. Remarquons que le dernier numéro obtenu est un numéro que l'on n'a jamais eu auparavant. Classons les évènements selon le dernier numéro obtenu, ce qui les range en K catégories, ayant toutes le même nombre d'éléments. Prenons par exemple une catégorie, celle se terminant par le nombre K. Les $n-1$ tirages qui précèdent sont formés avec les $K-1$ autres nombres, qui doivent tous être présents au moins une fois. On se place dans le contexte de la formule du crible, en définissant $K-1$ propriétés : la propriété j (avec j entre 1 et $K-1$) signifie que le numéro j n'est pas présent dans le mot concerné. Le nombre des mots est $S_0 - S_1 + S_2 - \ldots + (-1)^{K-1} S_{K-1}$, avec :

- $S_0 = (K-1)^{n-1}$, nombre de façons de construire les mots de longueur $n-1$ à base de $K-1$ numéros.
- S_1 est la somme des nombres A_j, avec A_j = nombre de mots ayant la propriété j, soit $(K-2)^{n-1}$ mots de longueur $n-1$ à base de $K-2$ lettres, d'où
  $S_1 = (K-1)(K-2)^{n-1} = \binom{K-1}{1}(K-2)^{n-1}$
- S_2 est la somme des nombres A_{ij}, avec A_{ij} = nombre de mots ayant les propriétés i et j, soit $(K-3)^{n-1}$ mots de longueur $n-1$ à base de $K-3$ lettres, d'où
  $S_2 = \binom{K-1}{2}(K-3)^{n-1}$
- On continue ainsi jusqu'à S_{K-1}, ou plutôt S_{K-2} puisque $S_{K-1} = 0$. [1]

Finalement, avec les K numéros possibles à la fin, le nombre d'évènements de longueur n est :

[1] Prenons l'exemple de $K = 4$, avec les mots de longueur n se terminant par 4. Comment s'en sortir même si l'on a oublié la formule du crible ? On prend tous les mots de longueur $n-1$ à base de trois numéros 1, 2, 3, soit 3^{n-1} cas, mais il faut enlever tous les mots ne contenant pas 3, et contenant seulement des 1 ou des 2, soit 2^{n-1} mots, ainsi que ceux ne contenant pas 2 et ceux ne contenant pas 1, soit au total $3.\ 2^{n-1}$ mots. Mais ce faisant on a compté deux fois les mots 111…1, 222…2, 333…3 : il convient d'enlever 3 au résultat négatif précédent. Les mots se terminant par 4 sont au nombre de $3^{n-1} - (3.\ 2^{n-1} - 3)$. On retrouve ce que donne la formule du crible. Le nombre d'évènements est donc $4(3^{n-1} - 3.\ 2^{n-1} + 3)$.

$$K\sum_{q=0}^{K-2}(-1)^q\binom{K-1}{q}(K-1-q)^{n-1}$$

Chacun de ces évènements a comme probabilité $(1 / K)^n$, d'où

$$p(X = n) = \left(\frac{1}{K}\right)^{n-1}\sum_{q=0}^{K-2}(-1)^q\binom{K-1}{q}(K-1-q)^{n-1}$$

Calcul de l'espérance *E*(*X*)

Considérons un des évènements élémentaires constitutif de $X = n$, et découpons-le en parties correspondant au temps d'apparition de chaque nouvelle lettre. Notons X_i la longueur de chacune de ces parties, au nombre de K. Par exemple pour $K = 4$, avec le mot <u>3</u> <u>331</u> <u>3134</u> <u>11344</u>2 on a $X = 14$, avec $X_1 = 1$, $X_2 = 3$, $X_3 = 4$, $X_4 = 6$, et $X = X_1 + X_2 + X_3 + X_4$. On a toujours $X_1 = 1$, d'où $E(X_1) = 1$. Chacune des variables X_i suit une loi géométrique : X_2 suit une loi de paramètre $p_2 = (K - 1) / K$, X_3 suit une loi de paramètre $p_3 = (K - 2) / K$, …, et X_K une loi de paramètre $1 / K$.

Avec $E(X)=\sum_{i=1}^{K}E(X_i)$, et $E(X_i)=\frac{1}{p_i}=\frac{K}{K+1-i}$,

$$E(X)=K\left(1+\frac{1}{2}+\frac{1}{3}+...+\frac{1}{K}\right)$$