Sujet 12 – Le collectionneur de vignettes

Vous faites la collection de vignettes présentes dans vos paquets de céréales. La collection entière comprend N vignettes différentes et vous achetez un paquet de céréales par semaine, combien de temps vous faudra-t-il pour compléter votre collection? Après avoir travaillé sur les résultats théoriques, nous pourrons estimer ce temps dans certains cas plus généraux : vous vous alliez avec votre voisin qui fait la même collection que vous et partagez vos doublons, le fabricant de céréales vous permet d'obtenir la vignette de votre choix en échange de 10 autres vignettes, ...

Collecte de Vignettes et Loi Géométrique

- Lorsque l'on suppose que toutes les vignettes sont équiprobables, c'est-à-dire que chaque type de vignette a une probabilité de **1/n** d'être obtenu lors de l'achat d'une boîte, on peut étudier la durée nécessaire pour compléter une collection de vignettes.

- Soit **T** la variable aléatoire qui représente le nombre de paquets à acheter (et par extension, le nombre de semaines nécessaires) pour compléter la collection de vignettes.
- L'obtention d'une nouvelle vignette dans un paquet **k** suit une loi géométrique de paramètre **p = n-(k−1)/n** , où **n** est le nombre total de vignettes à collecter. Ainsi, $T_{n,k}$, une variable aléatoire représentant le nombre de paquets à acheter pour obtenir une nouvelle vignette lorsque vous en avez déjà **k** différentes, suit également une loi géométrique de paramètre **p**.

L'espérance de $T_{n,k}$ est donnée par : $E(T_{n,k}) = 1/p = n/n-(k-1)$

Par conséquent, la variable aléatoire **T** qui représente le temps nécessaire pour compléter la collection entière est la somme de toutes les $T_{n,k}$ où **k** varie de **1** à **n**,

$T = \sum_{k=1,n} T_{n,k}$

en utilisant la linéarité de l'espérance, l'espérance de T est :

$$
\begin{aligned}
E(T_n) &= \sum_{k=1,n} E(T_{n,k}) \\
&= \sum_{k=1,n} n/n-(k-1) \\
&= n/n + n/(n-1) + n/(n-2) + \ldots + n/2 + n/1 \\
&= n \cdot (1/n + 1/(n-1) + 1/(n-2) + \ldots + 1/2 + 1/1) \\
&= n \cdot (1 + 1/2 + \ldots + 1/(n-2) + 1/(n-1) + 1/n) \\
&= n \cdot \sum_{k=1,n} 1/k
\end{aligned}
$$

D'où

$E(T_n) = n \cdot Hn$

En utilisant le développement asymptotique de **Hn**, on obtient :

$E(T_n) = n \cdot Hn = n \cdot \ln(n) + \gamma \cdot n + 1/2 + o(1)$

Où $\gamma \approx$ **0.5772156649** est la constante d'Euler-Mascheroni.

REF :https://fr.wikipedia.org/wiki/Problème du collectionneur de vignettes