Espérances

Exercice 1 Un problème pratique

Un grand nombre n de gens doivent se soumettre à un test de dépistage. Cette opération peut se gérer de deux façons :

1. Chaque échantillon est testé séparément, auquel cas n tests sont nécessaires.
2. Les échantillons sont mélangés par paquets de k puis le mélange est testé. Si le test est négatif, ce test suffit. Si le test est positif, chacun des k échantillons doit être retesté séparément, ce qui donne un nombre final de tests de $k+1$.

Soit p la probabilité qu'un échantillon soit positif, on suppose les échantillons indépendants et identiquement distribués.

1. Quelle est la probabilité qu'un mélange de k échantillons soit positif ?
2. Quelle est l'espérance de la variable aléatoire X du nombre de tests nécessaires en utilisant la seconde méthode (on supposera n divisible par k) ?
3. Pour p petit, quelle est la valeur de k qui minimise le nombre de tests ?

Solution

1. Dans un paquet de k échantillons, la probabilité que tous les tests soient négatifs est $(1-p)^k$. Ainsi, la probabilité qu'il y en ait au moins un est $1-(1-p)^k$
2. Pour chacun des $\frac{N}{k}$ paquets de k échantillons, qui sont indépendants, les deux valeurs possibles du nombre de tests sont :
   - $k+1$ (on a du refaire k tests après que le premier se soit révélé positif) avec une probabilité $1-(1-p)^k$;
   - 1 (le premier test est négatif) avec une probabilité $(1-p)^k$.

   Au final, on a $\mathbb{E}(X) = \frac{N}{k}\left(\left(1-(1-p)^k\right)(k+1) + (1-p)^k\right) = \frac{N}{k}\left(k+1-k(1-p)^k\right)$.
3. Pour p petit : $(1-p)^k \simeq 1-kp$, donc $\mathbb{E}(X) \simeq (k^2p+1)\frac{N}{k} = N(kp+\frac{1}{k})$. Une étude de variations montre que le minimum est atteint en $\frac{1}{\sqrt{p}}$.

Exercice 2 Un problème discret

Soit $t_1, t_2, \ldots, t_r$ le résultat d'une sélection (avec remise) de r éléments pris dans l'ensemble $(\frac{1}{n}, \frac{2}{n}, \ldots, \frac{n}{n})$. Soit X la variable aléatoire qui donne le plus petit r satisfaisant

$$t_1 + t_2 + \cdots + t_r > 1$$

On souhaite montrer que $\mathbb{E}(X) = (1+1/n)^n$. Pour cela, on peut considérer une variante équivalente de l'expérience, qui consiste à choisir avec remise un ensemble d'éléments $t_1, t_2, \ldots, t_r$ dans l'ensemble $(1, 2, \ldots, n)$. X est alors la variable aléatoire du plus petit r qui satisfait

$$t_1 + t_2 + \cdots + t_r > n$$

1. Montrer que $\mathbb{P}(X \geqslant j+1) = \binom{n}{j}(\frac{1}{n})^j$
2. Montrer que $\mathbb{E}(X) = \sum_{j=0}^{n} \mathbb{P}(X \geqslant j+1)$.
3. En déduire une expression de $\mathbb{E}(X)$ et sa limite quand $n \to \infty$.

Solution

1. On commence par remarquer que : $\mathbb{P}(X \geqslant j+1) = \mathbb{P}((t_1+t_2+\cdots+t_j) \leqslant n)$. Considérons les j nombres $a_1, \ldots, a_j$, définis par :

$$\begin{gathered} a_1 = t_1 \\ a_2 = t_1 + t_2 \\ a_3 = t_1 + t_2 + t_3 \\ \cdots \\ a_j = t_1 + t_2 + \cdots + t_j \end{gathered}$$

   La séquence des a_i est une séquence strictement croissante avec des valeurs distinctes prises dans l'ensemble des entiers compris entre 1 et n. Il y a $\binom{n}{j}$ façons de prendre j nombres distints entre 1 et n. De plus, il y a une bijection entre l'ensemble des telles séquences $(a_1, ..., a_j)$ et l'ensemble des séquences $(t_1, ..., t_j)$ vérifiant $\sum t_k \leqslant n$.
   Chacune de séquences $(t_1, ..., t_j)$ apparaît avec une probabilité $\frac{1}{n^j}$, cqfd.

2.

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{x=1}^{n} x\mathbb{P}(X=x) \\
&= \sum_{x=1}^{n}\sum_{i=1}^{x} \mathbb{P}(X=x) \\
&= \sum_{1\leqslant i\leqslant x\leqslant n} \mathbb{P}(X=x) \\
&= \sum_{i=1}^{n} \mathbb{P}(X\geqslant i)
\end{aligned}$$

3. $\mathbb{E}(X) = \sum_{j=1}^{n} \binom{n}{j}\frac{1}{n^j} = (1+\frac{1}{n})^n = e^{n\ln(1+\frac{1}{n})} = e^{1+o(\frac{1}{n})} \to e$

Exercice 3 Un problème théorique (et pratique aussi, finalement)

Soit $X_1, X_2, \ldots, X_n$ une collection de variables aléatoires discrètes ayant des espérances finies.

1. Montrer que, pour toute variable aléatoire Y,

$$\mathbb{E}\left(\sum_{i=1}^{n} X_i \mid Y=y\right) = \sum_{i=1}^{n} \mathbb{E}(X_i \mid Y=y)$$

2. On lance 2 dés. Soit X la variable aléatoire de la somme des deux lancés. Soit X_1 la variable aléatoire du résultat du premier lancer. Calculer $\mathbb{E}(X \mid X_1)$ en fonction de X_1.
3. $\mathbb{E}(X \mid X_1)$ est donc une variable aléatoire fonction de X_1. Plus généralement, démontrer le théorème suivant :

$$\mathbb{E}(X) = \mathbb{E}(\mathbb{E}(X \mid Y))$$

4. Considérons maintenant un programme dont la procédure principale inclut un appel à la procédure $\mathcal{S}$. Chaque appel à $\mathcal{S}$ lance récursivement des copies de $\mathcal{S}$, où le nombre de copies suit une loi aléatoire binomiale de paramètres n et p. Ces variables aléatoires sont indépendantes. On cherche l'espérance du nombre de copies de $\mathcal{S}$.
Raisonnons en terme de *génération*. Le lancement initial de $\mathcal{S}$ est dans la génération 0. Sinon, un processus de génération i a été lancé par un processus de génération $i-1$. Notons Y_i le nombre de copies de $\mathcal{S}$ de génération i. On voit facilement que $\mathbb{E}(Y_0) = 1$ et $\mathbb{E}(Y_1) = np$.
Supposons que l'on connaisse y_{i-1}, la valeur de Y_{i-1}. Comment peut-on calculer l'espérance de Y_i ?
5. Donner l'espérance du nombre total de copies au bout de i générations.

Solution

1. Vu en cours.
2. $\mathbb{E}(X|X_1) = \sum x\mathbb{P}(X|X_1) = \sum_{x=X_1+1}^{X_1+6} x\frac{1}{6} = X_1 + \frac{7}{2}$
3. $\mathbb{E}(\mathbb{E}(X|Y)) = \sum_y \mathbb{E}(X|Y=y)\mathbb{P}(Y=y) = \mathbb{E}(X)$
4. Soit $Z_1, ..., Z_{y_{i-1}}$ des v.a. i.i.d. suivant une loi binômiale de paramètres (n,p).

$$\begin{aligned}
\mathbb{E}\left(Y_i \,|\, Y_{i-1}=y_{i-1}\right) &= \mathbb{E}\left(\sum_{k=1}^{y_{i-1}} Z_k \,\middle|\, Y_{i-1}=y_{i-1}\right) \\
&= \sum_j j\mathbb{P}\left(\sum_{k=1}^{y_{i-1}} Z_k = j \,\middle|\, Y_{i-1}=y_{i-1}\right) \\
&= \sum_j j\mathbb{P}\left(\sum_{k=1}^{y_{i-1}} Z_k = j\right) \\
&= \mathbb{E}\left(\sum Z_k\right) \\
&= \sum \mathbb{E}\left(Z_k\right) \\
&= y_{i-1}np
\end{aligned}$$

5. $\mathbb{E}(Y_i) = \mathbb{E}(\mathbb{E}(Y_i \mid Y_{i-1})) = \mathbb{E}(()\, Y_{i-1} np) = (np)^i$. L'espérance du nombre de total de copies est donc $\sum_{k=0}^{i} (np)^k = \frac{1-(np)^{i+1}}{1-np}$.

Exercice 4 Un problème de boules

Vous avez une urne contenant deux boules noires et deux boules blanches, et 80 dollars. Chance ! Vous allez pouvoir jouer à un super jeu. Sortez les boules une par une sans remise. A chaque fois, pariez la moitié de votre fortune que c'est une blanche qui apparait.

Quelle est la fortune finale que vous pouvez espérer ?

Solution Quel que soit l'ordre de mes gains et pertes, je vais perdre 2 fois et gagner deux fois. Perdre revient à diviser l'argent qui reste par 2, gagner revient à multiplier l'argent qu'il nous reste par $3/2$. Par commutativité de la multiplication, on obtient toujours 45.