Devoir maison n° 8

À rendre le 24 novembre 2025

Sujet d'oral Agro-véto 2022

Soit N un entier supérieur ou égal à 2. On considère une urne contenant N boules indiscernables au toucher, numérotées de 1 à N.

On procède à des tirages successifs d'une boule avec remise de la boule dans l'urne avant le tirage suivant.

On note pour tout $k \geqslant 1$, X_k le numéro obtenu au k-ième tirage, et Z_k le nombre de numéros distincts obtenus au cours des k premiers tirages.

1. a) Écrire une fonction Python `NbDiff(L)` prenant en en argument une liste `L` et qui renvoie le nombre d'éléments distincts présents dans cette liste.

   b) Écrire une fonction Python `Z(N,k)` qui, prenant en argument les valeurs de N et k, renvoie une simulation de Z_k.

   c) Estimer l'espérance de Z_k à l'aide de votre programme, et conjecturer son comportement lorsque :

      (i) $N = 10$ et $k \to +\infty$;

      (ii) $k = 10$ et $N \to +\infty$;

      (iii) $N = k$ et $N \to +\infty$.

2. Déterminer la loi de la variable aléatoire Z_1 et la loi de la variable aléatoire Z_2. En déduire $E(Z_1)$ et $E(Z_2)$.

3. Soit k un entier supérieur ou égal à 1.

   a) Déterminer $P(Z_k = 1)$ et déterminer $P(Z_k = k)$.

   b) Montrer, pour tout $\ell \in [\![1; N]\!]$: $P(Z_{k+1} = \ell) = \dfrac{\ell}{N}P(Z_k = \ell) + \dfrac{N-\ell+1}{N}P(Z_k = \ell - 1)$.

   c) En déduire : $E(Z_{k+1}) = \dfrac{N-1}{N}E(Z_k) + 1$.

4. Montrer alors que pour tout $k \geqslant 1$: $E(Z_k) = N\left(1 - \left(\dfrac{N-1}{N}\right)^k\right)$.

5. Déterminer un équivalent de $E(Z_k)$ dans les trois cas suivants, en comparant avec vos résultats numériques de la question 1.c) :

   a) lorsque N est fixé et $k \to +\infty$;

   b) lorsque k est fixé et $N \to +\infty$;

   c) lorsque $N = k$ et $N \to +\infty$.

Facultatif : la métaphore de la cantine (ENS).

Dans ce problème il s'agit de caractériser la loi du nombre d'espèces représentées dans un échantillon de n individus, et leurs abondances respectives, à l'aide d'un unique paramètre.

Soit θ un réel strictement positif. Des individus numérotés $1, 2, \dots, n$, arrivent successivement dans une salle de restaurant contenant une infinité de tables infiniment longues. Le premier individu s'assied à une table au hasard. Pour tout entier $k \geqslant 1$, lorsque l'individu $k+1$ arrive, il choisit au hasard un des k convives déjà attablés avec la probabilité $1/(k+\theta)$, et s'assied à la même table, ou occupe une nouvelle table avec la probabilité $\theta/(k+\theta)$.

L'entier K_n désigne le nombre de tables occupées lorsque n convives se sont installés et pour $1 \leqslant i \leqslant n$, on note $q_{n,i} = \mathbb{P}(\mathrm{K}_n = i)$. La répartition de ces n convives en K_n tables est une métaphore pour la répartition d'un échantillon de n individus vivants en K_n espèces.

1. a) Montrer que : $q_{n+1,1} = \dfrac{n!}{(n+\theta)(n-1+\theta)\cdots(1+\theta)}$

   b) Pour tous $2 \leqslant i \leqslant n$, trouver une relation entre $q_{n+1,i}, q_{n,i}$ et $q_{n,i-1}$.

2. Soient L_n et P_n les polynômes de degré n suivants

$$\mathrm{P}_n = \sum_{i=1}^{n} q_{n,i}\mathrm{X}^i \qquad \mathrm{L}_n = \prod_{i=0}^{n-1}(\mathrm{X}+i)$$

   a) Donner une relation de récurrence vérifiée par (P_n).

   b) En déduire que

$$\mathrm{P}_n = \frac{\mathrm{L}_n(\theta\mathrm{X})}{\mathrm{L}_n(\theta)}$$

On admettra que cette équation caractérise la loi de K_n, mais dans la question suivante, on se concentre sur son espérance et sa variance.

3. a) Montrer que $\mathbb{E}(\mathrm{K}_n) = \mathrm{P}'_n(1)$ et en déduire $\mathbb{E}(\mathrm{K}_n)$.
   Indication : On pourra prendre le logarithme de P_n.

   b) Montrer que $\mathbb{V}(\mathrm{K}_n) = \mathrm{P}''_n(1) + \mathrm{P}'_n(1) - \big(\mathrm{P}'_n(1)\big)^2$ et calculer $\mathbb{V}(\mathrm{K}_n)$.

4. Dans cette question, on cherche à obtenir directement les résultats de la question précédente.

   a) Montrer que $\mathrm{K}_n = \displaystyle\sum_{i=1}^{n} \varepsilon_i$ où les $(\varepsilon_i)_{i=1,\dots,n}$ sont des variables de Bernoulli indépendantes dont on précisera les probabilités de succès respectives.

   b) En déduire $\mathbb{E}(\mathrm{K}_n)$ et $\mathbb{V}(\mathrm{K}_n)$.

5. a) Établir la double inégalité

$$1 + \int_1^n \frac{\theta}{\theta+x}\,\mathrm{d}x \leqslant \mathbb{E}(\mathrm{K}_n) \leqslant 1 + \int_0^{n-1} \frac{\theta}{\theta+x}\,\mathrm{d}x$$

   b) Donner un équivalent de $\mathbb{E}(\mathrm{K}_n)$ lorsque $n \to \infty$ (et le justifier).

6. Étudier la différence $\mathbb{V}(\mathrm{K}_n) - \mathbb{E}(\mathrm{K}_n)$ et en déduire un équivalent de $\mathbb{V}(\mathrm{K}_n)$ lorsque $n \to \infty$.