

DM 08 : Sujet Oral

Bio Spé

à rendre le lundi 24 novembre

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

- Écrire une fonction Python `NbDiff(L)` prenant en argument une liste L et qui renvoie le nombre d'éléments distincts présents dans cette liste.
 - Écrire une fonction Python `Z(N,k)` qui, prenant en argument les valeurs de N et k , renvoie une simulation de Z_k .
 - Estimer l'espérance de Z_k à l'aide de votre programme, et conjecturer son comportement lorsque :
 - $N = 10$ et $k \rightarrow +\infty$;
 - $k = 10$ et $N \rightarrow +\infty$;
 - $N = k$ et $N \rightarrow +\infty$.
- Déterminer la loi de la variable aléatoire Z_1 et la loi de la variable aléatoire Z_2 . En déduire $E(Z_1)$ et $E(Z_2)$.
- Soit k un entier supérieur ou égal à 1.
 - Déterminer $P(Z_k = 1)$ et déterminer $P(Z_k = k)$.
 - Montrer, pour tout $\ell \in \llbracket 1; N \rrbracket$:

$$P(Z_{k+1} = \ell) = \frac{\ell}{N} P(Z_k = \ell) + \frac{N - \ell + 1}{N} P(Z_k = \ell - 1).$$

- En déduire : $E(Z_{k+1}) = \frac{N-1}{N} E(Z_k) + 1$.
- Montrer alors que pour tout $k \geq 1$: $E(Z_k) = N \left(1 - \left(\frac{N-1}{N} \right)^k \right)$.
 - Déterminer un équivalent de $E(Z_k)$ dans les trois cas suivants, en comparant avec vos résultats numériques de la question 1.c) :
 - lorsque N est fixé et $k \rightarrow +\infty$;
 - lorsque k est fixé et $N \rightarrow +\infty$;
 - lorsque $N = k$ et $N \rightarrow +\infty$.

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 $k \geq 1$, 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 $1 \leq i \leq n$, on note $q_{n,i} = \mathbb{P}(K_n = i)$. 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.

- Montrer que : $q_{n+1,1} = \frac{n!}{(n+\theta)(n-1+\theta) \cdots (1+\theta)}$
 - Pour tous $2 \leq i \leq n$, trouver une relation entre $q_{n+1,i}$, $q_{n,i}$ et $q_{n,i-1}$.
- Soient L_n et P_n les polynômes de degré n suivants

$$P_n(X) = \sum_{i=1}^n q_{n,i} X^i \quad L_n(X) = \prod_{i=0}^{n-1} (X+i)$$

- Donner une relation de récurrence vérifiée par (P_n) .
- En déduire que

$$P_n(X) = \frac{L_n(\theta X)}{L_n(\theta)}$$

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.

- Montrer que $\mathbb{E}(K_n) = P'_n(1)$ et en déduire $\mathbb{E}(K_n)$.
Indication. On pourra prendre le logarithme de P_n .
 - Montrer que $\text{Var}(K_n) = P''_n(1) + P'_n(1) - (P'_n(1))^2$ et calculer $\text{Var}(K_n)$.
 - Dans cette question, on cherche à obtenir directement les résultats de la question précédente.
 - Montrer que $K_n = \sum_{i=1}^n \varepsilon_i$ où les $(\varepsilon_i)_{i=1, \dots, n}$ sont des variables de Bernoulli indépendantes dont on précisera les probabilités de succès respectives.
 - En déduire $\mathbb{E}(K_n)$ et $\text{Var}(K_n)$.
 - Établir la double inégalité
- $$1 + \int_1^n \frac{\theta}{\theta+x} dx \leq \mathbb{E}(K_n) \leq 1 + \int_0^{n-1} \frac{\theta}{\theta+x} dx$$
- Donner un équivalent de $\mathbb{E}(K_n)$ lorsque $n \rightarrow \infty$ (et le justifier).
- Étudier la différence $\text{Var}(K_n) - \mathbb{E}(K_n)$ et en déduire un équivalent de $\text{Var}(K_n)$ lorsque $n \rightarrow \infty$.