RMS 134 [1536]

On considère une famille $(X_k)_{1\leqslant k\leqslant n}$ de variables aléatoires définies sur un même espace probabilisé, indépendantes et suivant toutes la loi géométrique $\mathscr{G}(p)$ avec $0 < p < 1$.
On considère que la variable aléatoire X_k indique l'année de la première floraison de la tulipe k. En considérant que, si une tulipe a fleuri une année, alors elle fleurit toutes les années suivantes, on note X l'année à partir de laquelle toutes les tulipes fleurissent.
1❧ *Exprimer X en fonction des variables aléatoires X_k.*
2❧ *Calculer $\mathbf{P}(X > k)$ pour $k \in \mathbb{N}$. En déduire que X est une variable aléatoire d'espérance finie et calculer cette espérance.*

1❧ Par définition, la i-ième tulipe fleurit à partir de l'année X_i. Donc toutes les tulipes fleurissent l'année k si, et seulement si,

$$\forall\, 1 \leqslant i \leqslant n, \quad k \geqslant X_i,$$

c'est-à-dire (passage au maximum) si

$$k \geqslant \max_{1\leqslant i\leqslant n} X_i.$$

On en déduit que l'année X de la première floraison générale est donnée par

$$X = \max_{1\leqslant i\leqslant n} X_i.$$

2❧ Comme on l'a vu plus haut, pour tout $k \in \mathbb{N}$,

$$[X \leqslant k] = \bigcap_{1\leqslant i\leqslant n} [X_i \leqslant k].$$

Comme les variables aléatoires X_i sont indépendantes, on en déduit que

$$\forall\, k \in \mathbb{N}, \quad \mathbf{P}(X \leqslant k) = \prod_{i=1}^{n} \mathbf{P}(X_i \leqslant k)$$

et comme elles suivent toutes la loi géométrique $\mathscr{G}(p)$,

$$\forall\, k \in \mathbb{N}^*, \quad \mathbf{P}(X \leqslant k) = \mathbf{P}(X_1 \leqslant k)^n = (1 - q^k)^n$$

avec $q = 1 - p \in]0, 1[$ (comme d'habitude).
❧ On en déduit, par passage au complémentaire et par la formule du binôme, que

$$\begin{aligned}\forall\, k \in \mathbb{N}, \quad \mathbf{P}(X > k) &= 1 - (1 - q^k)^n \\ &= \sum_{\ell=1}^{n} (-1)^{\ell+1} \binom{\ell}{n} q^{k\ell}.\end{aligned}$$

Le terme général de la série $\sum \mathbf{P}(X > k)$ est donc une combinaison linéaire de séries géométriques convergentes (de raisons respectives $q, q^2, \dots, q^n$ avec $0 < q < 1$). Par conséquent, la série $\sum \mathbf{P}(X > k)$ est convergente, donc X est une variable aléatoire d'espérance finie et

$$\begin{aligned}\mathbf{E}(X) = \sum_{k=0}^{+\infty} \mathbf{P}(X > k) &= \sum_{\ell=1}^{n} (-1)^{\ell+1} \binom{\ell}{n} \sum_{k=0}^{+\infty} (q^\ell)^k \\ &= \sum_{\ell=1}^{n} \binom{\ell}{n} \frac{(-1)^{\ell+1}}{1 - q^\ell}.\end{aligned}$$

✍ *Il s'agit ici d'une combinaison linéaire d'un nombre fini de séries convergentes, donc la linéarité de la somme suffit pour justifier l'interversion des deux symboles $\sum$. Le théorème de Fubini est hors sujet !*