Question 845

Vincent Devinck

Question. Soient $X_1, \ldots, X_n$ des variables indépendantes de loi uniforme sur $[0,1]$. On note

$$X^{(1)} \leqslant X^{(2)} \leqslant \cdots \leqslant X^{(n)}$$

les valeurs ordonnées de $\{X_1, \ldots, X_n\}$. Déterminer l'espérance de

$$\sum_{k=1}^{n-1} |\ln X^{(k+1)} - \ln X^{(k)}|$$

Franck Taieb

Solution. Comme $(X_k)_{1 \leqslant k \leqslant n}$ est la suite ordonnée de $\{X_1, \ldots, X_n\}$, on a (par croissance du logarithme sur $\mathbb{R}_+^*$) :

$$\sum_{k=1}^{n-1} |\ln X^{(k+1)} - \ln X^{(k)}| = \sum_{k=1}^{n-1} (\ln X^{(k+1)} - \ln X^{(k)}) = \ln X^{(n)} - \ln X^{(1)} \tag{1}$$

On sait que

$$X^{(1)} = \min(X_1, \ldots, X_n) \qquad \text{et} \qquad X^{(n)} = \max(X_1, \ldots, X_n)$$

Il est aisé de calculer la fonction de répartition d'une variable aléatoire correspondant au maximum (ou au minimum) d'une suite finie de variables aléatoires indépendantes suivant une loi donnée. La proposition suivante fait le lien entre la fonction de répartition et l'espérance d'une variable aléatoire.

Proposition 1. *Soit X une variable aléatoire réelle continue d'univers image $X(\Omega) =]-\infty, 0]$. On suppose que l'intégrale $\displaystyle\int_{-\infty}^{0} P(X \leqslant t)\,\mathrm{d}t$ est convergente. Alors X admet une espérance et, dans ce cas, on a l'égalité*

$$E(X) = -\int_{-\infty}^{0} P(X \leqslant t)\,\mathrm{d}t$$

Démonstration. Pour tout $t \in]-\infty, 0]$, on sait que $P(X \leqslant t) = \displaystyle\int_{-\infty}^{t} \mathrm{d}P_X(x)$ où P_X désigne la loi de probabilité de X. Par conséquent, pour tout $A < 0$, on a

$$\begin{aligned}
\int_A^0 |x|\,\mathrm{d}P_X(x) = \int_A^0 (-x)\,\mathrm{d}P_X(x) &= \int_A^0 \left(\int_x^0 \mathrm{d}t \right) \mathrm{d}P_X(x) \\
&= \int_A^0 \left(\int_A^t \mathrm{d}P_X(x) \right) \mathrm{d}t \\
&= \int_A^0 P(A < X \leqslant t)\,\mathrm{d}t
\end{aligned}$$

En écrivant maintenant $P(A < X \leqslant t) = P(X \leqslant t) - P(X \leqslant A)$ pour tout $t \in]A, 0]$, on obtient

$$\int_A^0 |x| \,\mathrm{d}P_X(x) = \int_A^0 P(X \leqslant t) \,\mathrm{d}t + AP(X \leqslant A)$$

On suppose que l'intégrale $\displaystyle\int_{-\infty}^0 P(X \leqslant t) \,\mathrm{d}t$ est convergente. Montrons alors que

$$\lim_{A \to -\infty} AP(X \leqslant A) = 0 \tag{2}$$

ce qui démontrera la proposition. Par croissance de la probabilité P, on sait que pour tout élément t de l'intervalle $[A, A/2]$, on a l'inégalité

$$P(X \leqslant A) \leqslant P(X \leqslant t)$$

Par croissance de l'intégrale, il vient

$$0 \leqslant -\frac{AP(X \leqslant A)}{2} \leqslant \int_A^{A/2} P(X \leqslant t) \,\mathrm{d}t$$

et comme par hypothèse l'intégrale $\displaystyle\int_{-\infty}^0 P(X \leqslant t) \,\mathrm{d}t$ est convergente, on a

$$\lim_{A \to -\infty} \int_A^{A/2} P(X \leqslant t) \,\mathrm{d}t = 0$$

et la limite annoncée en (2). $\square$

Le proposition suivante contient essentiellement la réponse à la question posée.

Proposition 2. *Les variables aléatoires* $\ln X^{(1)}$ *et* $\ln X^{(n)}$ *admettent une espérance qui valent respectivement*

$$E(\ln X^{(1)}) = -\sum_{k=1}^{n} \frac{1}{k} \qquad \textit{et} \qquad E(\ln X^{(n)}) = -\frac{1}{n}$$

Démonstration. L'univers image des variables aléatoires réelles $\ln X^{(1)}$ et $\ln X^{(n)}$ est $] - \infty, 0]$. Soit $A < 0$. Alors

$$\begin{aligned}\int_A^0 P(\ln X^{(n)} \leqslant t) \,\mathrm{d}t &= \int_A^0 P(\max(X_1, \ldots, X_n) \leqslant e^t) \,\mathrm{d}t \\ &= \int_A^0 P\big((X_1 \leqslant e^t) \cap \cdots \cap (X_n \leqslant e^t)\big) \,\mathrm{d}t\end{aligned}$$

et comme les variables aléatoires $X_1, \ldots, X_n$ sont indépendantes et suivent la loi uniforme sur l'intervalle $[0, 1]$, il vient

$$\int_A^0 P(\ln X^{(n)} \leqslant t) \,\mathrm{d}t = \int_A^0 P(X_1 \leqslant e^t) \ldots P(X_n \leqslant e^t) \,\mathrm{d}t = \int_A^0 e^{nt} \,\mathrm{d}t = \frac{1 - e^{nA}}{n}$$

On déduit de la proposition 1 que la variable aléatoire $\ln X^{(n)}$ admet une espérance qui vaut

$$E(\ln X^{(n)}) = \lim_{A \to -\infty} -\int_A^0 P(\ln X^{(n)} \leqslant t) \,\mathrm{d}t = -\frac{1}{n}$$

D'autre part,

$$\begin{aligned}\int_A^0 P(\ln X^{(1)} \leqslant t) \,\mathrm{d}t &= \int_A^0 P(\min(X_1, \ldots, X_n) \leqslant e^t) \,\mathrm{d}t \\ &= \int_A^0 \big(1 - P(\min(X_1, \ldots, X_n) > e^t)\big) \,\mathrm{d}t \\ &= \int_A^0 \big[1 - P\big((X_1 > e^t) \cap \cdots \cap (X_n > e^t)\big)\big] \,\mathrm{d}t\end{aligned}$$

En utilisant à nouveau l'indépendance des variables aléatoires $X_1, \dots, X_n$ et le fait qu'elles suivent la loi uniforme sur $[0, 1]$, on obtient

$$\begin{aligned}\int_A^0 P(\ln X^{(1)} \leqslant t)\,\mathrm{d}t &= \int_A^0 \left[1 - P(X_1 > e^t) \dots P(X_n > e^t)\right] \mathrm{d}t \\ &= \int_A^0 \left[1 - (1 - e^t)^n\right] \mathrm{d}t\end{aligned}$$

Le changement de variable $u = 1 - e^t$ fournit

$$\int_A^0 P(\ln X^{(1)} \leqslant t)\,\mathrm{d}t = \int_{1-e^A}^0 \frac{1-u^n}{u-1}\,\mathrm{d}u = \int_0^{1-e^A} \frac{1-u^n}{1-u}\,\mathrm{d}u$$

Avec la formule donnant la somme des termes d'une suite géométrique, on obtient

$$\begin{aligned}\int_A^0 P(\ln X^{(1)} \leqslant t)\,\mathrm{d}t &= \sum_{k=0}^{n-1} \int_0^{1-e^A} u^k\,\mathrm{d}t \\ &= \sum_{k=1}^{n} \frac{(1-e^A)^k}{k}\end{aligned}$$

Comme $\lim\limits_{A \to -\infty} (1 - e^A)^k = 1$, la variable aléatoire $\ln X^{(1)}$ admet une espérance (d'après la proposition 1) qui vaut

$$E(\ln X^{(1)}) = \lim_{A \to -\infty} - \int_A^0 P(\ln X^{(1)} \leqslant t)\,\mathrm{d}t = -\sum_{k=1}^{n} \frac{1}{k}$$

$\square$

En utilisant la proposition 2, l'égalité (1) et la linéarité de l'espérance, on conclut finalement que

$$E\left(\sum_{k=1}^{n-1} |\ln X^{(k+1)} - \ln X^{(k)}|\right) = E(\ln X^{(n)}) - E(\ln X^{(1)}) = \sum_{k=1}^{n-1} \frac{1}{k}$$