

Vecteurs aléatoires

La difficulté des exercices est inquée par le nombre d'étoile(s) : applicatif (\star), classique ($\star \star$) et exploratoire ($\star \star \star$).

Cette indication n'a rien d'absolu : ne vous découragez pas pour quelques petites étoiles !

Pour toute question ou remarque, n'hésitez pas à m'écrire à l'adresse mhacini.pro@gmail.com

Rappelons-nous qu'autrefois, et à une époque qui n'est pas encore bien reculée, une pluie ou une sécheresse extrême, une comète traînant après elle une queue fort étendue, les éclipses, les aurores boréales et généralement tous les phénomènes extraordinaires étaient regardés comme autant de signes de la colère céleste. On invoquait le ciel pour détourner leur funeste influence. [En réalité,] la courbe décrite par une simple molécule d'air ou de vapeurs, est réglée d'une manière aussi certaine, que les orbites planétaires : il n'y a de différences entre elles, que celle qu'y met notre ignorance. La probabilité est relative en partie à cette ignorance, en partie à nos connaissances.

– Pierre Simon de Laplace

Mauvaise file ($\star \star$)

Trois personnes entrent simultanément dans une poste ne comportant que deux guichets. On note X_i le temps passé par la i -ème personne au guichet. Les variables X_i sont supposées mutuellement indépendantes et suivant toutes trois la loi uniforme sur $[0, 1]$.

1. Déterminer la fonction de répartition de la variable correspondant au temps que doit attendre la troisième personne pour qu'un guichet se libère, notée Y .
2. En déduire que Y admet une densité, et déterminer l'une de ses densités.
3. Déterminer une densité de la variable aléatoire correspondant au temps total que la troisième personne passe dans le bureau de poste, notée Z .
4. En déduire $E(Z)$ et $V(Z)$.

Correction :

1. Pour tout $x \in \mathbb{R}$, on a $[Y > x] = [X_1 > x] \cap [X_2 > x]$. Et donc, par indépendance de X_1 et X_2 , il vient

$$\begin{aligned} F_Y(x) &= P(Y \leq x) \\ &= 1 - P(Y > x) = 1 - P([X_1 > x] \cap [X_2 > x]) \\ &= 1 - P(X_1 > x)P(X_2 > x) \\ &= 1 - (1 - F_{X_1}(x))(1 - F_{X_2}(x)) = \begin{cases} 0 & \text{si } x \leq 0 \\ 1 - (1-x)^2 & \text{si } 0 < x < 1 \\ 1 & \text{si } x \geq 1 \end{cases} \end{aligned}$$

2. Puisque F_{X_1} et F_{X_2} sont continues sur \mathbb{R} et de classe \mathcal{C}^1 sauf éventuellement en un nombre fini de points, il en est de même de F_Y . Une densité de Y est alors toute fonction coïncidant avec F'_Y là où F_Y est dérivable. Par exemple, on peut prendre

$$f_Y(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ 2(1-x) & \text{si } 0 < x < 1 \\ 0 & \text{si } x \geq 1 \end{cases}$$

3. Par le lemme des coalitions, $Y = \min(X_1, X_2)$ et X_3 sont indépendantes. Puisque la densité de X_3 est bornée, par produit de convolution, une densité de $Y + X_3$ est alors donnée par

$$h(x) = \int_{-\infty}^{+\infty} f_Y(t) f_{X_3}(x-t) dt.$$

Mais on a $f_Y(t) \neq 0 \Leftrightarrow 0 \leq t \leq 1$ et $f_{X_3}(x-t) \neq 0 \Leftrightarrow 0 \leq x-t \leq 1 \Leftrightarrow x-1 \leq t \leq x$.

Ainsi, pour $x \leq 0$, on a $f_Z(x) = 0$. De même, pour $x-1 \geq 1 \Leftrightarrow x \geq 2$, on a $f_Z(x) = 0$. Si $x \in [0, 1]$, il vient

$$f_Z(x) = \int_0^x 2(1-t) dt = \left[-(1-t)^2 \right]_0^x = 1 - (1-x)^2 = 2x - x^2.$$

Enfin, pour $1 \leq x \leq 2$, on a

$$f_Z(x) = \int_{x-1}^1 2(1-t) dt = (2-x)^2.$$

Une densité de Z est donc

$$f_Z(x) = \begin{cases} 2x - x^2 & \text{si } 0 \leq x \leq 1 \\ 4 - 4x + x^2 & \text{si } 1 < x \leq 2 \\ 0 & \text{sinon} \end{cases}$$

4. Notons que $0 \leq Z \leq 2$, de sorte que l'existence de l'espérance de Z est garantie. On a alors

$$E(Z) = \int_{-\infty}^{+\infty} t f_Z(t) dt = \int_0^1 t(2t-t^2) dt + \int_1^2 t(2-t)^2 dt = \frac{5}{6}.$$

De même, le moment d'ordre 2 de Z est donné par

$$E(Z^2) = \int_{-\infty}^{+\infty} t^2 f_Z(t) dt = \int_0^1 t^2(2t-t^2) dt + \int_1^2 t^2(2-t)^2 dt = \frac{5}{6}.$$

Et donc, par la formule de Huygens, la variance de Z vaut

$$V(Z) = E(Z^2) - E(Z)^2 = \frac{5}{6} - \frac{25}{36} = \frac{5}{36}.$$

Minimum (★★)

Un sac contient N boules numérotées de 1 à N . On effectue dans ce sac n tirages uniformes d'une boule avec remise et on note Z_n le plus petit numéro obtenu. Déterminer la loi de Z_n .

Correction :

Commençons par remarquer que $Z_n(\Omega) = \llbracket 1, N \rrbracket$. Notons X_i le numéro de la boule obtenue lors du i -ème tirage, de sorte que $X_i \hookrightarrow \mathcal{U}(\llbracket 1, N \rrbracket)$. De plus, les tirages ayant lieu avec remise, X_1, \dots, X_n sont mutuellement indépendantes. On a alors $Z_n = \min(X_1, \dots, X_n)$.

Pour tout $k \in \llbracket 1, N \rrbracket$, on a donc $[Z_n \geq k] = \bigcap_{i=1}^n [X_i \geq k]$, et par indépendance des X_i ,

$$P(Z_n \geq k) = \prod_{i=1}^n P(X_i \geq k) = P(X_1 \geq k)^n.$$

Or

$$P(X_1 \geq k) = \sum_{\ell=k}^N P(X_1 = \ell) = \sum_{\ell=k}^N \frac{1}{N} = \frac{N-k+1}{N}.$$

Ainsi, il vient $P(Z_n \geq k) = \left(\frac{N-k+1}{N}\right)^n$.

Enfin, puisque Z_n est à valeurs entières, on a $[Z_n \geq k] = [Z_n = k] \cup [Z_n \geq k+1]$, de sorte que

$$P(Z_n \geq k) = P(Z_n = k) + P(Z_n \geq k+1) \Leftrightarrow P(Z_n = k) = P(Z_n \geq k) - P(Z_n \geq k+1).$$

Et donc, pour tout $k \in \llbracket 1, N \rrbracket$,

$$P(Z_n = k) = \left(\frac{N-k+1}{N}\right)^n - \left(\frac{N-k}{N}\right)^n.$$

Notons que pour $k = N$, on obtient $P(Z_n = N) = \frac{1}{N^n}$, ce qui était prévisible, puisque le seul moyen que la plus petite des boules porte le numéro N est d'avoir tiré n fois la boule N .

Marche aléatoire (★★)

Soit $p \in]0, 1[$. On considère une suite $(X_n)_{n \in \mathbb{N}^*}$ de variables aléatoires mutuellement indépendantes, de même loi, et telles que $P(X_1 = -1) = 1 - p$ et $P(X_1 = 1) = p$. On pose $S_n = \sum_{k=1}^n X_k$.

1. Déterminer deux réels a et b tels que $Y_n = aX_n + b$ suive une loi de Bernoulli de paramètre p .
2. En utilisant la question précédente, déterminer la loi de S_n et calculer son espérance.

Application :

3. Un ivrogne sort du bar, et chaque seconde avance d'un mètre vers la droite avec probabilité p , ou d'un mètre vers la gauche avec probabilité $1 - p$.
En moyenne, où se trouve-t-il après 10 secondes ?
4. Proposer un programme Python simulant la marche d'un tel ivrogne pendant 1000 secondes, avec $p = \frac{1}{2}$.

Correction :

1. $Y_n = \frac{1}{2}(X_n + 1)$ ne prend que les valeurs 0 et 1, avec $P(Y_n = 1) = P(X_n = 1) = p$, donc suit une loi de Bernoulli $\mathcal{B}(p)$.
2. On a

$$S_n = \sum_{k=1}^n X_k = \sum_{k=1}^n (2Y_k - 1) = 2 \sum_{k=1}^n Y_k - n$$

et $\sum_{k=1}^n Y_k$ suit une loi binomiale $\mathcal{B}(n, p)$ car les Y_i sont indépendantes. De plus, S_n prend ses valeurs dans $\llbracket -n, n \rrbracket$, et donc pour tout $k \in \llbracket -n, n \rrbracket$,

$$\begin{aligned} P(S_n = k) &= P\left(2 \sum_{k=1}^n Y_k = k + n\right) \\ &= P\left(\sum_{k=1}^n Y_k = \frac{k+n}{2}\right) = \begin{cases} \binom{n}{\frac{k+n}{2}} p^{\frac{k+n}{2}} (1-p)^{\frac{2-k-n}{2}} & \text{si } k+n \text{ est impair} \\ 0 & \text{si } k+n \text{ est pair} \end{cases} \end{aligned}$$

On a alors, par linéarité de l'espérance,

$$E(S_n) = \sum_{i=1}^n E(X_i) = \sum_{i=1}^n ((-1)(1-p) + p) = n(2p - 1).$$

Application :

3. Si on note X_i la variable aléatoire qui vaut 1 si l'ivrogne se déplace vers la gauche et -1 s'il se déplace d'un mètre sur la droite, alors S_n désigne sa position à l'instant n . En particulier, on a

$$E(S_{10}) = 2 \times 10 \times p - 10 = 20p - 10.$$

Remarque : Si $p = 0.5$, alors en moyenne il est revenu à son point de départ, que si $p > 0.5$, alors il se sera en moyenne déplacé vers la gauche, et que si $p < 0.5$, alors il se sera en moyenne déplacé vers la droite.

4. `import random`



```
def simulation_ivrogne(n_secondes, p):
    position = 0

    for _ in range(n_secondes):
        tirage = random.random()
        if tirage < p:
            position = position + 1
        else:
            position = position - 1
    return position

duree = 1000
p = 0.5

resultat = simulation_ivrogne(duree, proba_p)
print(f"Après {duree} secondes, l'ivrogne est à la position : {resultat}")
```

Poissons géométriques (★★★)

Toutes les variables aléatoires de l'exercice sont définies sur un espace probabilisé (Ω, \mathcal{A}, P) .

On considère une suite $(X_n)_{n \in \mathbb{N}^*}$ de variables aléatoires indépendantes, suivant toutes la loi géométrique de paramètre p , où p est un réel de $]0, 1[$. On pose de plus $q = 1 - p$.

Pour tout entier naturel n non nul, on pose : $S_n = \sum_{i=1}^n X_i$.

On considère également une variable aléatoire N , à valeur dans \mathbb{N}^* , possédant une espérance et indépendante des variables aléatoires X_n .

Pour tout ω de Ω , on pose $S(\omega) = \sum_{i=1}^{N(\omega)} X_i(\omega)$ et on admet que $S = \sum_{i=1}^N X_i$ est une variable aléatoire.

Enfin, on rappelle la formule suivante, que l'on pourra utiliser sans la justifier. Pour tous entiers naturels r et s tels que $r \leq s$, on a :

$$\sum_{j=r}^s \binom{j}{r} = \binom{s+1}{r+1}$$

1. Déterminer $S_n(\Omega)$. Montrer que l'on a :

$$\forall k \in S_n(\Omega), \quad P([S_n = k]) = \binom{k-1}{n-1} p^n q^{k-n}$$

2. a. Vérifier pour tout $n \in \mathbb{N}^*$, l'existence de l'espérance conditionnelle $E(S | [N = n])$ et donner sa valeur.
b. En déduire $E(S)$.
3. On suppose dans cette question que N suit la loi géométrique de paramètre p . Déterminer la loi de S .

Correction :

1. Puisque chaque X_i suit une loi géométrique, on a $X_i(\Omega) = \mathbb{N}^*$. Ainsi, la somme des n termes vaut au minimum n . On a donc $S_n(\Omega) = [\![n, +\infty]\!]$.

Montrons par récurrence la propriété

$$\mathcal{P}(n) : \forall k \geq n, P(S_n = k) = \binom{k-1}{n-1} p^n q^{k-n}.$$

Initialisation ($n = 1$) :

$$P(S_1 = k) = P(X_1 = k) = pq^{k-1}. \text{ La formule donne } \binom{k-1}{0} p^1 q^{k-1} = 1 \cdot pq^{k-1}.$$

Héritéité :

Supposons $\mathcal{P}(n)$ vraie. On a $S_{n+1} = S_n + X_{n+1}$. Par la formule des probabilités totales (convolution discrète):

$$P(S_{n+1} = k) = \sum_{j=n}^{k-1} P(S_n = j)P(X_{n+1} = k-j).$$

En utilisant l'hypothèse de récurrence :

$$\begin{aligned} P(S_{n+1} = k) &= \sum_{j=n}^{k-1} \binom{j-1}{n-1} p^n q^{j-n} \times p q^{k-j-1} \\ &= p^{n+1} q^{k-(n+1)} \sum_{j=n}^{k-1} \binom{j-1}{n-1}. \end{aligned}$$

On pose le changement d'indice $i = j - 1$. La somme devient :

$$\sum_{i=n-1}^{k-2} \binom{i}{n-1}$$

D'après la formule rappelée dans l'énoncé avec $r = n - 1$ et $s = k - 2$:

$$\sum_{i=n-1}^{k-2} \binom{i}{n-1} = \binom{(k-2)+1}{(n-1)+1} = \binom{k-1}{n}$$

Finalement :

$$P(S_{n+1} = k) = \binom{k-1}{n} p^{n+1} q^{k-(n+1)}$$

Ce qui achève la récurrence.

2. a. Conditionnellement à $[N = n]$, S est la somme déterministe de n variables X_i indépendantes. Par linéarité de l'espérance :

$$E(S \mid [N = n]) = E(S_n) = \sum_{i=1}^n E(X_i) = nE(X_1)$$

Comme $X_1 \in \mathcal{G}(p)$, $E(X_1) = \frac{1}{p}$. Donc :

$$E(S \mid [N = n]) = \frac{n}{p}$$

b. D'après la formule de l'espérance totale:

$$E(S) = \sum_{n=1}^{+\infty} E(S \mid [N = n])P(N = n) = \sum_{n=1}^{+\infty} \frac{n}{p} P(N = n) = \frac{1}{p} \sum_{n=1}^{+\infty} nP(N = n) = \frac{1}{p} E(N)$$

On sait que N possède une espérance, donc $E(S)$ existe et vaut $\frac{E(N)}{p}$.

3.

On applique la formule des probabilités totales :

$$P(S = k) = \sum_{n=1}^{+\infty} P(S = k \mid N = n)P(N = n)$$

On sait que $P(S = k \mid N = n) = P(S_n = k)$, qui est nul si $n > k$. La somme s'arrête donc à k .

$$\begin{aligned} P(S = k) &= \sum_{n=1}^k \binom{k-1}{n-1} p^n q^{k-n} \times p q^{n-1} \\ &= \sum_{n=1}^k \binom{k-1}{n-1} p^{n+1} q^{k-1} \\ &= p^2 q^{k-1} \sum_{n=1}^k \binom{k-1}{n-1} p^{n-1} \end{aligned}$$

On pose $j = n - 1$. La somme va de 0 à $k - 1$:

$$P(S = k) = p^2 q^{k-1} \underbrace{\sum_{j=0}^{k-1} \binom{k-1}{j} p^j 1^{(k-1)-j}}_{(1+p)^{k-1}}$$

On obtient donc :

$$P(S = k) = p^2 q^{k-1} (1 + p)^{k-1} = p^2 [q(1 + p)]^{k-1}$$

Or $q(1 + p) = (1 - p)(1 + p) = 1 - p^2$. Ainsi :

$$\forall k \in \mathbb{N}^*, \quad P(S = k) = p^2 (1 - p^2)^{k-1}$$

S suit une loi géométrique de paramètre p^2 .