Couples de variables aléatoires 1

Exercices axés sur le calcul

Exercice 1 ⋆ *Loi d'un couple : une variable fonction d'une autre variable*

Soit X une variable aléatoire suivant la loi uniforme sur $X(\Omega) = \{-1,0,1\}$. Notons $Y = X^2$.

1) Déterminer la loi du couple (X,Y).
2) En déduire la loi de Y.
3) Les variables X et Y sont-elles indépendantes ?
4) Comparer $\mathbb{E}(X \cdot Y)$ et $\mathbb{E}(X) \cdot \mathbb{E}(Y)$. Commenter.

Exercice 2 ⋆⋆ *Loi d'un couple : loi conjointe donnée par un tableau*

La loi conjointe du couple (X,Y) est donnée par le tableau suivant :

$X \setminus Y$	0	1	2
0	1/20	1/4	0
1	17/60	1/4	1/6

1) Déterminer la loi de X et celle de Y.
2) Les variables X et Y sont-elles indépendantes ?
3) Calculer $\mathbb{E}(X)$, $\mathbb{E}(Y)$ et $\mathbb{E}(X \cdot Y)$. Commenter.

Exercice 3 ⋆ *Une boîte puis une boule*

On considère n boîtes numérotées de 1 à n. La boîte numéro k contient k boules numérotées de 1 à k. On choisit au hasard une boîte, puis une boule dans cette boîte. Soient X et Y les numéros de la boîte et de la boule obtenues.

1) Déterminer la loi du couple (X,Y).
2) Calculer $\mathbb{P}(X = Y)$.
3) Déterminer la loi de Y ainsi que son espérance.

Exercice 4 ⋆⋆ *Temps d'attente*

Une urne contient $N - 2$ boules vertes, 1 boule blanche et 1 boule rouge. On tire les boules de l'urne, une à une et sans remise.

1) Soient X_1 le rang d'apparition de la boule blanche et X_2 le rang d'apparition de la boule rouge. Déterminer la loi de X_1, la loi de X_2 et la loi du couple (X_1,X_2).

2) Soit X le rang où on obtient pour la première fois soit la boule blanche, soit la boule rouge. Soit Y le rang où on a obtenu pour la première fois les deux boules blanche et rouge. Déterminer la loi de X et la loi de Y.

Exercices axés sur le raisonnement

Exercice 5 ⋆⋆ *Au plus deux tentatives*

Un standardiste effectue n appels téléphoniques vers n correspondants distincts ($n \geqslant 2$). Pour chaque appel, la probabilité d'obtenir le correspondant demandé est p appartenant à $]0,1[$ et la probabilité de ne pas l'obtenir est q, avec $q = 1 - p$.

1) Soit X le nombre de correspondants obtenus lors de ces n appels. Quelle est la loi de X ? Donner son espérance.

2) Après ces n recherches, le standardiste appelle une deuxième fois chacun des $n - X$ correspondants qu'il n'a pas obtenus la première fois. Soit Y le nombre de correspondants obtenus dans la deuxième série d'appels, et soit $Z = X + Y$ le nombre total de correspondants obtenus.

   a) Quelles sont les valeurs prises par Z ?

   b) Calculer $p_0 = \mathbb{P}(Z = 0)$ et vérifier que $p_1 = npq^{2n-2}(1 + q)$.

   c) Calculer la probabilité conditionnelle $\mathbb{P}_{[X=k]}(Y = \ell)$, pour k appartenant à $[\![0,n]\!]$ et ℓ à $[\![0,n - k]\!]$.

   d) Établir que $\mathbb{P}(Z = s) = \sum_{k=0}^{s} \mathbb{P}\big((X = k) \cap (Y = s - k)\big)$.

   e) Calculer $p_s = \mathbb{P}(Z = s)$, et montrer que Z suit une loi binomiale de paramètres n et $p(1 + q)$.
   Indication. *On pourra vérifier que* $\binom{n}{k}\binom{n-k}{s-k} = \binom{n}{s}\binom{s}{k}$.

Exercice 6 ⋆⋆ *Minimum et maximum de deux géométriques indépendantes*

Soient X et X′ deux variables aléatoires indépendantes suivant respectivement $\mathcal{G}(p)$ et $\mathcal{G}(p')$, où p et p' sont deux éléments de $]0,1[$ fixés. On note également $q = 1 - p$ et $q' = 1 - p'$. Pour toute variable aléatoire T, on notera $\mathbb{F}_T$ sa fonction de répartition.

1) Rappeler ou retrouver $\mathbb{F}_X(k)$ pour tout entier positif k.

2) Soit $Y = \max(X, X')$.

   a) Calculer $\mathbb{F}_Y(k)$ pour tout entier positif k.

   b) La variable Y suit-elle une loi géométrique ?

3) Soit $Z = \min(X, X')$.
   a) Calculer $\mathbb{F}_Z(k)$ pour tout entier positif k.
   b) En déduire que Z suit une loi géométrique dont on précisera le paramètre.
   c) Donner une interprétation de ce résultat en termes d'expériences aléatoires.

Exercice 7 ⋆⋆ *Maximum d'uniformes*

On fixe deux entiers strictement positifs d et n, et on considère d variables aléatoires $X_1, \dots, X_d$ mutuellement indépendantes, suivant chacune la loi uniforme sur l'ensemble $[\![1,n]\!]$. On note alors

$$M(d,n) = \max(X_1, \dots, X_d).$$

1) *Échauffement.* Dans un premier temps, on suppose que d vaut 2, et on considère donc la variable aléatoire $M(2,n) = \max(X_1, X_2)$, que l'on note ici M pour abréger.
   a) Pour chaque entier $k \in [\![1,n]\!]$, déterminer la valeur de $\mathbb{F}_M(k)$.
   b) Justifier successivement que

$$\forall\, k \in [\![1,n]\!], \qquad \mathbb{P}(M = k) = \mathbb{F}_M(k) - \mathbb{F}_M(k-1),$$

que

$$\mathbb{E}(M) = n\,\mathbb{F}_M(n) - \sum_{k=0}^{n-1} \mathbb{F}_M(k)$$

puis que

$$\mathbb{E}(M) = \sum_{k=0}^{n} \big(1 - \mathbb{F}_M(k)\big).$$

   Indication. *On pourra utiliser la linéarité de la sommation et introduire un changement d'indice bien choisi.*
   c) En déduire la valeur de $\mathbb{E}\big(M(2,n)\big)$.

   Indication. *On pourra admettre ou redémontrer que* $\sum_{k=1}^{n} k^2 = \frac{1}{6}\, n(n+1)(2n+1)$.
   d) Déterminer un équivalent de $\mathbb{E}\big(M(2,n)\big)$ lorsque n tend vers l'infini.

2) On revient au cas général, et d est donc désormais un entier strictement positif quelconque. Reprendre les questions précédentes, et montrer que

$$\mathbb{E}\big(M(d,n)\big) \underset{n\to\infty}{\sim} \frac{n\,d}{d+1}.$$

3) Dans cette dernière partie, on vérifie numériquement le résultat qui vient d'être obtenu.
   a) Écrire en Python une fonction d'en-tête `simulation_M(d, n, N)` qui renvoie N simulations aléatoires indépendantes de la variable $M(d,n)$. *On pourra importer le module* `numpy` *sous l'alias* `np`, *et le sous-module* `numpy.random` *sous l'alias* `rd`.
   b) Écrire en Python une fonction d'en-tête `moyenne_empirique(d, n, N)` qui fasse appel à la fonction précédente et qui renvoie la moyenne des valeurs des N simulations.
   c) Examiner la valeur de `moyenne_empirique(d, n, N) / n` pour différentes valeurs de d, n et N.

Exercice 8 ⋆⋆⋆ *Fonctions génératrices et stabilité*

On établit ici le résultat de stabilité des lois binomiales sous l'addition indépendante en se servant de quelques propriétés simples des *fonctions génératrices*.
Si X est une variable aléatoire finie, à valeurs entières positives, la fonction génératrice de X est la fonction

$$g_X : \begin{cases} [0,1[& \to & \mathbb{R} \\ t & \mapsto & \mathbb{E}(t^X). \end{cases}$$

1) Déterminer une expression simple de $g_X(t)$ pour $t \in [0,1[$ si X suit une loi de Bernoulli de paramètre p, puis si X suit une loi uniforme sur $[\![1,n]\!]$, où n est un entier positif fixé.

2) On suppose dans cette question que $n \in \mathbb{N}$ et que X est une variable aléatoire à valeurs dans $[\![0,n]\!]$.

   a) Justifier que
   $$\forall t \in [0,1[, \quad g_X(t) = \sum_{k=0}^{n} \mathbb{P}(X = k)\, t^k.$$

   b) Justifier que si deux variables aléatoires à valeurs dans $[\![0,n]\!]$ ont la même loi, alors elles ont la même fonction génératrice.

   c) Montrer que
   $$\forall k \in [\![0,n]\!], \quad \mathbb{P}(X = k) = \tfrac{1}{k!}\, g_X^{(k)}(0),$$
   où $g^{(k)}$ désigne la dérivée k-ième de g_X.

   d) En déduire que si deux variables aléatoires à valeurs dans $[\![0,n]\!]$ ont la même fonction génératrice, alors elles ont la même loi.

3) Soient X et Y deux variables aléatoires finies. Montrer que si X et Y sont indépendantes, alors
$$g_{X+Y} = g_X \cdot g_Y$$
en tant que fonctions de $[0,1[$ dans $\mathbb{R}$.

4) Dans cette dernière question, on note γ la fonction génératrice associée à une variable de loi $\mathcal{B}(p)$, telle que calculée lors de la première question.

   a) Montrer que si X est une variable de loi $\mathcal{B}(n,p)$, alors $g_X = \gamma^n$.

   b) Montrer que si X_1 et X_2 sont deux variables indépendantes, de loi respective $\mathcal{B}(n_1,p)$ et $\mathcal{B}(n_2,p)$, alors $g_{X_1+X_2} = \gamma^{n_1+n_2}$.
   Retrouver le fait que $X_1 + X_2$ suit la loi $\mathcal{B}(n_1 + n_2,p)$.

Corrections

Exercices axés sur le calcul

Exercice 1 *Loi d'un couple : une variable fonction d'une autre variable*

1) La variable Y est à valeurs dans $\{0,1\}$. De plus, si $X = 0$, alors $Y = 0$, et si $X = \pm 1$, alors $Y = 1$. De ce fait, la loi de (X,Y) peut être décrite par le tableau suivant :

Y \ X	−1	0	1
0	0	1/3	0
1	1/3	0	1/3

2) On en déduit que $\mathbb{P}(Y=0) = 1/3$ et que $\mathbb{P}(Y=1) = 2/3$. Autrement dit, $\boxed{Y \hookrightarrow \mathcal{B}(2/3).}$

3) Intuitivement, puisque Y est fonction de X, il faut s'attendre à ce que X et Y ne soient pas indépendantes. Plus précisément, la valeur de X détermine entièrement celle de Y.
Formalisons cette idée. D'une part

$$\begin{aligned} \mathbb{P}([X=0]\cap[Y=0]) &= \mathbb{P}(X=0) && \text{(puisque si } X=0\text{, alors } Y=0) \\ &= 1/3, \end{aligned}$$

et d'autre part

$$\begin{aligned} \mathbb{P}(X=0)\cdot\mathbb{P}(Y=0) &= 1/3\cdot 1/3 \\ &= 1/9. \end{aligned}$$

On constate donc que $\mathbb{P}([X=0]\cap[Y=0]) \neq \mathbb{P}(X=0)\cdot\mathbb{P}(Y=0)$, ce qui prouve que

$$\boxed{\text{X et Y ne sont pas indépendantes.}}$$

4) Pour X et Y, on revient à la définition de l'espérance : ces deux variables étant finies, elles admettent automatiquement une espérance, et

$$\mathbb{E}(X) = \sum_{k\in X(\Omega)} k\,\mathbb{P}(X=k) = (-1)\cdot 1/3 + 0\cdot 1/3 + 1\cdot 1/3 = \boxed{0.}$$

De même, $\boxed{\mathbb{E}(Y) = 2/3.}$

(on peut par exemple se servir du fait que Y suit une loi de Bernoulli). De plus, le *théorème de transfert* (dans le cas de sommes finies, donc automatiquement absolument convergentes) donne

$$\mathbb{E}(X\cdot Y) = \sum_{\substack{k\in X(\Omega)\\ \ell\in Y(\Omega)}} k\,\ell\,\mathbb{P}([X=k]\cap[Y=\ell]) = (-1)\cdot 1\cdot 1/3 + 1\cdot 1\cdot 1/3 = 0.$$

En particulier, $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$, *et ce sans que* X *et* Y *soient pour autant indépendantes.*

Remarque

On peut généraliser de la façon suivante :
Soit X une variable aléatoire de loi uniforme sur $[\![-n,n]\!]$ et $Y = X^2$.
Alors X et $-X$ ont même loi. De même X^3 et $-X^3$. Donc $\mathbb{E}(X) = 0$ et $\mathbb{E}(XY) = \mathbb{E}(X^3) = 0$.
Ainsi $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$, sans que X et Y soient indépendantes.

Exercice 2 *Loi d'un couple : loi conjointe donnée par un tableau*

1) On procède par sommation des lignes et des colonnes, en commençant par les sommes les plus simples et en concluant en utilisant le fait que la somme de toutes les probabilités vaut 1. On trouve d'une part

$$\mathbb{P}(X = 0) = 6/20 = 3/10, \quad \text{puis}, \quad \mathbb{P}(X = 1) = 14/20 = 7/10,$$

et d'autre part

$$\mathbb{P}(Y = 2) = 2/12 = 1/6, \quad \mathbb{P}(Y = 1) = 6/12 = 1/2, \quad \text{et donc} \quad \mathbb{P}(Y = 0) = 1/3.$$

2) La présence d'une case nulle entraîne qu'une certaine valeur de X empêche une certaine autre valeur de Y : ici, si $X = 0$, alors $Y \neq 2$. Transformons cette idée en une preuve rigoureuse du fait que X et Y ne sont pas indépendantes : d'une part

$$\mathbb{P}([X = 0] \cap [Y = 2]) = 0$$

(événement impossible) ; d'autre part

$$\mathbb{P}(X = 0) \cdot \mathbb{P}(Y = 2) = 3/10 \cdot 1/6 \neq 0,$$

ce qui montre que $\boxed{\text{X et Y ne sont pas indépendantes.}}$

3) Pour X et Y, on revient à la définition de l'espérance : ces deux variables étant finies, elles admettent automatiquement une espérance, et

$$\mathbb{E}(X) = \sum_{k \in X(\Omega)} k\,\mathbb{P}(X = k) = 1 \cdot 7/10 = 7/10.$$

(on peut d'ailleurs utiliser l'expression connue pour l'espérance d'une variable de Bernoulli). De même, $\mathbb{E}(Y) = 5/6$.
De plus, *le théorème de transfert* (dans le cas de sommes finies, donc automatiquement absolument convergentes) donne

$$\begin{aligned}\mathbb{E}(X \cdot Y) &= \sum_{\substack{k \in X(\Omega) \\ \ell \in Y(\Omega)}} k\,\ell\,\mathbb{P}([X = k] \cap [Y = \ell]) \\ &= 1 \cdot 1 \cdot 1/4 + 1 \cdot 2 \cdot 1/6 \qquad \text{(les 4 autres termes étant nuls)} \\ &= 7/12.\end{aligned}$$

Or on remarque également que $\quad \mathbb{E}(X) \cdot \mathbb{E}(Y) = 7/10 \cdot 5/6 = 7/12.$

et en particulier, $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$, *et ce sans que* X *et* Y *soient pour autant indépendantes.*

Exercice 3 *Une boîte puis une boule*

1) La variable X est à valeurs dans $[\![1,n]\!]$, et la variable Y est également à valeurs dans $[\![1,n]\!]$ puisque si l'on choisit la boîte numéro n, tous les numéros de boule entre 1 et n sont accessibles. De plus, soit $(k,\ell) \in [\![1,n]\!]^2$. La *formule des probabilités composées* donne

$$\begin{aligned}\mathbb{P}([X = k] \cap [Y = \ell]) &= \mathbb{P}(X = k) \cdot \mathbb{P}_{[X=k]}(Y = \ell) \\ &= \frac{1}{n} \cdot \begin{cases} 0 & \text{si } \ell > k \\ 1/k & \text{si } \ell \in [\![1,k]\!]. \end{cases}\end{aligned}$$

On a donc entièrement déterminé la loi du couple (X,Y).

2) Décomposons cet événement à l'aide de la *formule des probabilités totales*, en se servant du système complet d'événements $([X = k])_{k\in[\![1,n]\!]}$:

$$\begin{aligned}\mathbb{P}(X = Y) &= \sum_{k=1}^{n} \mathbb{P}(X = k) \cdot \mathbb{P}_{[X=k]}(X = Y) \\ &= \sum_{k=1}^{n} \frac{1}{n} \cdot \mathbb{P}_{[X=k]}(Y = k) \\ &= \sum_{k=1}^{n} \frac{1}{n} \cdot \frac{1}{k} = \frac{1}{n} \cdot \sum_{k=1}^{n} \frac{1}{k} = \boxed{\frac{1}{n} H_n,}\end{aligned}$$

où H_n désigne le n-ième nombre harmonique.

3) De la même manière, si $\ell \in [\![1,n]\!]$, décomposons l'événement $Y = \ell$ à l'aide de *la formule des probabilités totales*. On écrit

$$\begin{aligned}\mathbb{P}(Y = \ell) &= \sum_{k=1}^{n} \mathbb{P}([X = k] \cap [Y = \ell]) \\ &= \sum_{k=1}^{\ell-1} \underbrace{\mathbb{P}([X = k] \cap [Y = \ell])}_{=0 (\ell > k)} + \sum_{k=\ell}^{n} \mathbb{P}([X = k] \cap [Y = \ell]) \\ &= 0 + \sum_{k=\ell}^{n} \mathbb{P}([X = k] \cap [Y = \ell]) \\ &= \sum_{k=\ell}^{n} \frac{1}{n} \cdot \frac{1}{k} = \frac{1}{n} \sum_{k=\ell}^{n} \frac{1}{k} = \frac{1}{n} (H_n - H_{\ell-1})\end{aligned}$$

relation de Chasles

avec la même notation que dans la question précédente.

Ensuite, puisque la variable Y est finie, elle admet automatiquement une espérance, et

$$\begin{aligned}
\mathbb{E}(Y) = \sum_{\ell=1}^{n} \ell\, \mathbb{P}(Y=\ell) &= \sum_{\ell=1}^{n} \ell \cdot \frac{1}{n} \sum_{k=\ell}^{n} \frac{1}{k} = \frac{1}{n} \sum_{1 \leqslant \ell \leqslant k \leqslant n} \ell \cdot \frac{1}{k} \\
&= \frac{1}{n} \sum_{k=1}^{n} \frac{1}{k} \sum_{\ell=1}^{k} \ell && \text{somme arithmétique} \\
&= \frac{1}{n} \sum_{k=1}^{n} \frac{1}{k} \cdot \frac{k(k+1)}{2} && \text{par factorisation et linéarité} \\
&= \frac{1}{2n} \left(\sum_{k=1}^{n} k + \sum_{k=1}^{n} 1 \right) && \text{somme arithmétique} \\
&= \frac{1}{2n} \left(\frac{n(n+1)}{2} + n \right) \\
&= \boxed{\frac{n+3}{4}.}
\end{aligned}$$

Exercice 4 *Temps d'attente*

1) Puisque l'urne contient une boule blanche et une boule rouge, ces deux boules jouent un rôle équivalent, et donc X_1 et X_2 ont la même loi.

On peut tirer au maximum N boules, puisque l'expérience se fait sans remise, et chaque boule peut être tirée à n'importe quel moment, donc $X_1(\Omega) = [\![1,N]\!]$.

Fixons un entier $k \in [\![1,N]\!]$; en tenant compte de l'évolution de la composition de l'urne au fil des tirages et en appliquant *la formule des probabilités composées*, on obtient

$$\begin{aligned}
\mathbb{P}(X_1 = k) &= \overbrace{\frac{\cancel{N-1}}{N} \times \frac{\cancel{N-2}}{\cancel{N-1}} \times \cdots \times \frac{\cancel{N-(k-1)}}{\cancel{N-(k-1)+1}}}^{k-1 \text{ fois une boule verte ou la rouge}} \times \overbrace{\frac{1}{\cancel{N-(k-1)}}}^{\text{boule blanche}} && \text{simplification} \\
&= \frac{1}{N}
\end{aligned}$$

d'où l'on déduit que $\boxed{X_1 \hookrightarrow \mathcal{U}([\![1,N]\!]).}$

et de même pour X_2.

• Fixons désormais deux entiers k et ℓ dans $[\![1,N]\!]$ et calculons $\mathbb{P}([X_1 = k] \cap [X_2 = \ell])$. Les deux boules ne pouvant pas être tirées en même temps, cette probabilité est nulle lorsque $k = \ell$; supposons à présent que $k \neq \ell$. Par symétrie des rôles, on peut d'ailleurs supposer que $k < \ell$, c'est-à-dire que l'on va seulement examiner les situations où la boule blanche est tirée avant la boule rouge.

D'après *la formule des probabilités composées,*

$$\mathbb{P}([X_1 = k] \cap [X_2 = \ell]) = \overbrace{\frac{N-2}{N} \times \frac{N-3}{N-1} \times \cdots \times \frac{N-(k-1)-1}{N-(k-1)+1}}^{k-1 \text{ boules vertes}} \times \underbrace{\overbrace{\frac{1}{N-(k-1)}}^{\text{boule blanche}}}_{\text{tirage } k}$$

$$\times \underbrace{\overbrace{\frac{N-k-1}{N-k} \times \frac{N-k-2}{N-k-1} \times \cdots \times \frac{N-(\ell-1)}{N-(\ell-1)+1}}^{\ell-k-1 \text{ boules vertes}}}_{\text{tirages } k+1 \text{ à } \ell-1} \times \underbrace{\overbrace{\frac{1}{N-(\ell-1)}}^{\text{boule rouge}}}_{\text{tirage } \ell}.$$

Les dénominateurs correspondant au nombre de boules dans l'urne à chaque étape, et les numérateurs au nombre de boules favorables à l'événement considéré. Par télescopage de produit,

$$\mathbb{P}([X_1 = k] \cap [X_2 = \ell]) = \frac{1}{N(N-1)}.$$

et de même lorsque $k > \ell$. En résumé,

$$\boxed{\forall (k,\ell) \in [\![1,N]\!]^2, \quad \mathbb{P}([X_1 = k] \cap [X_2 = \ell]) = \begin{cases} 0 & \text{si } k = \ell \\ \frac{1}{N(N-1)} & \text{si } k \neq \ell. \end{cases}}$$

Remarque

On peut aussi obtenir ces résultats de la façon suivante.

Pour avoir équiprobabilité des façons de vider l'urne, on distingue toutes les boules en les numérotant. L'univers est l'ensemble des permutations : N! tirages possibles.

Pour la boule rouge en position k : 1 possibilité pour la boule rouge, $(N-1)!$ pour les boules restantes. Donc

$$\mathbb{P}(X_1 = k) = \frac{\operatorname{card}(X_1 = k)}{\operatorname{card}(\Omega)} = \frac{(N-1)!}{N!} = \frac{1}{N}.$$

Pour la boule rouge en position k et la boule verte en position ℓ $(k \neq \ell)$: 1 possibilité pour la boule rouge et la boule verte $(N-2)!$ pour les boules restantes.

Donc $\mathbb{P}([X_1 = k] \cap [X_2 = \ell]) = \dfrac{\operatorname{card}([X_1 = k] \cap [X_2 = \ell])}{\operatorname{card}(\Omega)} = \dfrac{(N-2)!}{N!} = \dfrac{1}{N(N-1)}.$

2) La variable X ne peut prendre la valeur N, car cela signifierait que l'on a tiré $N-1$ boules vertes, or l'urne n'en contient initialement que $N-2$. De plus, si $k \in [\![1,N-1]\!]$, alors

$$\begin{aligned}\mathbb{P}(X = k) &= \frac{N-2}{N}\frac{N-3}{N-1} \cdots \frac{N-(k-1)-1}{N-(k-1)+1} \cdot \frac{2}{N-(k-1)} \\ &= \frac{1}{N}\frac{1}{N-1}\frac{N-(k-1)}{1}\frac{N-(k-1)-1}{1} \cdot \frac{2}{N-(k-1)} \\ &= \frac{2(N-k)}{N(N-1)}.\end{aligned}$$

De plus, au vu de la loi du couple (X_1,X_2), on peut déduire la loi de Y de celle de X : en imaginant que l'on renverse l'ordre dans lequel les tirages sont effectués, la probabilité que la dernière des deux boules blanche et rouge soit tirée à l'instant k, et suivie de $N-k$ boules vertes, est la même que la probabilité que la première des deux boules blanche et rouge soit tirée à l'instant $N+1-k$, et précédée de $N-k$ boules vertes. Autrement dit, $Y(\Omega) = [\![2,N]\!]$ et, pour tout entier $k \in [\![2,N]\!]$,

$$\mathbb{P}(Y = k) = \mathbb{P}(X = N+1-k) = \frac{2\left(N-(N+1-k)\right)}{N\,(N-1)} = \boxed{\frac{2\,(k-1)}{N\,(N-1)}.}$$

Exercices axés sur le raisonnement

Exercice 5 *Au plus deux tentatives*

1) On est en présence d'une répétition de n expériences de Bernoulli, que l'on peut supposer indépendantes et de même probabilité de succès p, et X compte le nombre de succès.

$$\boxed{X \hookrightarrow \mathcal{B}(n,p).}$$

et en particulier $\mathbb{E}(X) = np$.

2) a) Le nombre total de correspondants obtenus est compris entre 0 (dans le pire des cas) et n (car on ne peut pas obtenir une même personne deux fois de suite).

b) • Dire que $Z = 0$, c'est dire qu'aucun correspondant n'a été obtenu ni pendant la première, ni pendant la deuxième phase. La probabilité de cet événement est donc

$$p_0 = q^n \cdot q^n = q^{2n}.$$

• Dire que $Z = 1$, c'est dire que $X = 0$ et que $Y = 1$, ou que $X = 1$ et que $Y = 0$, ces deux événements étant incompatibles. De ce fait,

$$p_1 = \underbrace{q^n}_{\mathbb{P}(X=0)} \times \underbrace{\binom{n}{1} q^{n-1}\,p}_{\mathbb{P}_{[X=0]}(Y=1)} + \underbrace{\binom{n}{1} q^{n-1}\,p}_{\mathbb{P}(X=1)} \times \underbrace{q^{n-1}}_{\mathbb{P}_{[X=1]}(Y=0)}$$
$$= n\,q^{n-1}\,p\,(q^{n-1} + q^n) = n\,q^{2n-2}\,p\,(1+q).$$

c) Si l'on a réussi à obtenir k correspondants lors de la première phase, il en reste $n-k$ à appeler.
La probabilité que l'on en obtienne ℓ est donnée par la loi binomiale $\mathcal{B}(n-k,p)$:

$$\forall\, \ell \in [\![0,n-k]\!], \quad \mathbb{P}_{[X=k]}(Y = \ell) = \binom{n-k}{\ell} p^\ell q^{n-k-\ell}.$$

d) Il ne s'agit de rien d'autre que de *la formule des probabilités totales* appliquée au système complet d'événements $(X = k)_{k \in [\![0,n]\!]}$, en restreignant la somme aux indices inférieurs à s puisque Y ne peut être négatif.

e) Écrivons : *la formule des probabilités composées* donne

$$\begin{aligned} p_s &= \sum_{k=0}^{s} \mathbb{P}(\mathrm{X}=k)\cdot \mathbb{P}_{[\mathrm{X}=k]}(\mathrm{Y}=\ell) \\ &= \sum_{k=0}^{s} \binom{n}{k} p^k q^{n-k} \cdot \binom{n-k}{s-k} p^{s-k} q^{n-s} = p^s q^{n-s} \sum_{k=0}^{s} \binom{n}{k}\binom{n-k}{s-k} q^{n-k}. \end{aligned}$$

Or $\quad \displaystyle\binom{n}{k}\binom{n-k}{s-k} = \frac{n!}{k!\,(n-k!)} \cdot \frac{(n-k)!}{(s-k)!\,(n-s)!} = \frac{n!}{k!} \cdot \frac{1}{(s-k)!\,(n-s)!}.$

Et, $\quad \displaystyle\binom{n}{s}\binom{s}{k} = \frac{n!}{s!\,(n-s)!} \cdot \frac{s!}{k!\,(s-k)!} = \frac{n!}{(n-s)!} \cdot \frac{1}{k!\,(s-k)!}.$

Les deux expressions sont égales. Revenons à présent à notre calcul de probabilité :

$$\begin{aligned} p_s &= p^s q^{n-s} \sum_{k=0}^{s} \binom{n}{s}\binom{s}{k} q^{n-k} \\ &= p^s q^{n-s} \binom{n}{s} \sum_{k=0}^{s} \binom{s}{k} q^{n-k} \\ &= p^s q^{n-s} \binom{n}{s} \sum_{k=0}^{s} \binom{s}{k} q^{n-s+s-k} \\ &= p^s q^{2n-2s} \binom{n}{s} \sum_{k=0}^{s} \binom{s}{k} q^{s-k} \\ &= p^s q^{2n-2s} \binom{n}{s} (1+q)^s \\ &= \binom{n}{s} \big(p(1+q)\big)^s (q^2)^{n-s}. \end{aligned}$$

binôme de Newton

On est bien en présence d'une loi binomiale de paramètres n et $p(1+q)$.
On vérifie que $1-q^2 = (1-q)(1+q) = p(1+q)$ et donc que $1-p(1+q) = q^2$.

Remarque
On peut donner une interprétation plus simple à ce résultat : dans l'ensemble de l'expérience, chaque correspondant peut se retrouver dans une et une seule de trois situations possibles; soit il répond au premier appel (probabilité p), soit il répond au second appel mais pas au premier (probabilité qp), soit il ne répond pas du tout (probabilité q^2). Dès lors, la probabilité qu'un correspondant individuel réponde à au moins un des appels est $p + qp = p(1+q)$.

Exercice 6 *Minimum et maximum de deux géométriques indépendantes*

1) Deux principaux points de vue possibles :

- On utilise, en les additionnant, les probabilités (connues) $\mathbb{P}(X = k)$. On obtient ainsi

$$\begin{aligned}\mathbb{F}_X(k) &= \sum_{\substack{i \in X(\Omega) \\ i \leqslant k}} \mathbb{P}(X = i) = \sum_{i=1}^{k} q^{i-1} p \\ &= p \cdot \frac{1-q^k}{1-q} = 1 - q^k\end{aligned}$$

somme géométrique de raison différente de 1

puisque $p + q = 1$.

- On interprète X comme le temps d'attente avant le premier succès dans une répétition d'expériences de Bernoulli indépendantes et de même paramètre p.
Écrivons $\mathbb{F}_X(k) = \mathbb{P}(X \leqslant k) = 1 - \mathbb{P}(X > k)$. La probabilité $\mathbb{P}(X > k)$ est la probabilité que le premier succès survienne strictement après l'instant k, ce qui revient à dire que les k premières tentatives se sont toutes soldées par un échec. Donc par indépendance

$$\mathbb{P}(X > k) = q^k$$

et l'on retrouve le résultat.

2) a) Écrivons : $\mathbb{F}_Y(k) = \mathbb{P}(\max(X,X') \leqslant k)$. Or,

$$\begin{aligned}[\max(X,X') \leqslant k] \iff [X \leqslant k] \cap [X' \leqslant k]; \\ \mathbb{F}_Y(k) &= \mathbb{P}([X \leqslant k] \cap [X' \leqslant k]) \\ &= \mathbb{P}(X \leqslant k) \cdot \mathbb{P}(Y \leqslant k) \\ &= (1-q^k)\cdot(1-q'^k) = 1 - q^k - q'^k + (qq')^k.\end{aligned}$$

Passage aux probabilités

indépendance

b) La fonction de répartition de Y n'est donc pas de la forme de celle associée à une loi géométrique, donc on peut s'attendre à ce que la loi de Y ne soit pas géométrique. Justifions-le rigoureusement : pour que Y suive une loi géométrique, les probabilités $\mathbb{P}(Y = k)$ devraient être en progression géométrique. En particulier, on devrait avoir

$$\frac{\mathbb{P}(Y=3)}{\mathbb{P}(Y=1)} = \left(\frac{\mathbb{P}(Y=2)}{\mathbb{P}(Y=1)}\right)^2$$

Calculons ces deux quotients : on peut écrire

$$\begin{aligned}\mathbb{P}(Y=1) &= (1-q)(1-q'), \\ \mathbb{P}(Y=2) &= (1-q^2)(1-q'^2) - (1-q)(1-q') \\ &= (1-q)(1-q')((1+q)(1+q')-1)\end{aligned}$$

et

$$\begin{aligned}\mathbb{P}(Y=3) &= (1-q^3)(1-q'^3) - (1-q)(1-q') \\ &= (1-q)(1-q')((1+q+q^2)(1+q'+q'^2)-1)\end{aligned}$$

d'où

$$\begin{aligned}\frac{\mathbb{P}(Y=2)}{\mathbb{P}(Y=1)} &= (1+q)(1+q')-1\\ &= q+q'-qq'\end{aligned}$$

et donc

$$\left(\frac{\mathbb{P}(Y=2)}{\mathbb{P}(Y=1)}\right)^2 = q^2+q'^2+(qq')^2+2qq'\,(1-q-q')$$

tandis que

$$\begin{aligned}\frac{\mathbb{P}(Y=3)}{\mathbb{P}(Y=1)} &= (1+q+q^2)(1+q'+q'^2)-1\\ &= q+q'+q^2+q'^2+qq'+q\,q'^2+q^2\,q'+(qq')^2\\ &= q+q'+q^2+q'^2+(qq')^2+qq'(1+q+q')\end{aligned}$$

d'où il résulte que

$$\begin{aligned}\frac{\mathbb{P}(Y=3)}{\mathbb{P}(Y=1)}-\left(\frac{\mathbb{P}(Y=2)}{\mathbb{P}(Y=1)}\right)^2 &= q+q'+qq'(-1+3q+3q')\\ &= (q+q')(1+3qq')-qq'\end{aligned}$$

ce qui est en général strictement positif, parce que q et q' sont tous deux strictement supérieurs à qq', et que $1+3qq'$ est supérieur à 1. Bref,

$$\boxed{\text{Y n'est pas géométrique.}}$$

3) a) On procède de la même manière, mais cette fois-ci il est judicieux de passer au complémentaire pour se ramener à la probabilité d'une intersection :

$$\mathbb{F}_Z(k) = \mathbb{P}(\min(X,X') \leqslant k) = 1-\mathbb{P}(\min(X,X') > k).$$

Or $$[\min(X,X') \leqslant k] \iff [X>k]\cap[X'>k]\,;$$

donc

$$\begin{aligned}\mathbb{F}_Z(k) &= 1-\mathbb{P}([X>k]\cap[X'>k])\\ &= 1-\mathbb{P}(X>k)\cdot\mathbb{P}(Y>k)\\ &= 1-q^k\cdot q'^k = 1-(qq')^k.\end{aligned}$$

par indépendance

b) On aboutit donc à une fonction de répartition identique à celle associée à une loi géométrique de probabilité d'échec qq', c'est-à-dire à la loi $\boxed{\mathcal{G}(1-qq').}$

c) Voyons X et X′ comme les temps d'attente avant d'obtenir le premier succès dans deux répétitions d'expériences de Bernoulli indépendantes et de paramètre respectif p et p', et supposons que ces deux répétitions ont lieu *en parallèle*. Alors Z correspond au temps d'attente avant d'obtenir un premier « succès collectif », c'est-à-dire au premier instant où l'expérience de Bernoulli associée à X, *ou* celle associée à X′, connaît son premier succès. Dans cette interprétation, un « échec collectif » a une probabilité qq' de se produire, et on retrouve ainsi la loi de Z ainsi que son paramètre.

Exercice 7 *Maximum d'uniformes*

1) a) Par définition,

$$\mathbb{F}_{\mathrm{M}}(k) = \mathbb{P}(\mathrm{M} \leqslant k) = \mathbb{P}\big(\max(\mathrm{X}_1,\mathrm{X}_2) \leqslant k\big) = \mathbb{P}\big([\mathrm{X}_1 \leqslant k] \cap [\mathrm{X}_2 \leqslant k]\big),$$

car le plus grand de deux nombres est inférieur à k si et seulement si les deux nombres sont inférieurs à k ; donc, par indépendance

$$\mathbb{F}_{\mathrm{M}}(k) = \mathbb{P}(\mathrm{X}_1 \leqslant k) \cdot \mathbb{P}(\mathrm{X}_2 \leqslant k) = \mathbb{P}(\mathrm{X}_1 \leqslant k)^2 = \left(\frac{k}{n}\right)^2 \cdot$$

puisque X_1 suit une loi uniforme sur $[\![1,n]\!]$ et qu'elle a donc k chances sur n d'appartenir à $[\![1,k]\!]$.

b) Comme M est à valeurs entières,

$$\begin{aligned}[\mathrm{M} = k] &\iff [\mathrm{M} \leqslant k] \setminus [\mathrm{M} < k] \\ &\iff [\mathrm{M} \leqslant k] \setminus [\mathrm{M} \leqslant k-1]\end{aligned}$$

donc

$$\begin{aligned}\mathbb{P}(\mathrm{M} = k) &= \mathbb{P}(m \leqslant k) - \mathbb{P}(\mathrm{M} \leqslant k-1) \\ &= \mathbb{F}_{\mathrm{M}}(k) - \mathbb{F}_{\mathrm{M}}(k-1).\end{aligned}$$

Ensuite, la variable M est finie, donc admet une espérance, et

$$\begin{aligned}\mathbb{E}(\mathrm{M}) &= \sum_{k=1}^{n} k\, \mathbb{P}(\mathrm{M} = k) \\ &= \sum_{k=1}^{n} k \left(\mathbb{F}_{\mathrm{M}}(k) - \mathbb{F}_{\mathrm{M}}(k-1)\right) && \text{linéarité de la sommation} \\ &= \sum_{k=1}^{n} k\, \mathbb{F}_{\mathrm{M}}(k) - \sum_{k=1}^{n} k\, \mathbb{F}_{\mathrm{M}}(k-1)) && \text{changement d'indice } k \leftarrow k-1 \\ &= \sum_{k=1}^{n} k\, \mathbb{F}_{\mathrm{M}}(k) - \sum_{k=0}^{n-1} (k+1)\, \mathbb{F}_{\mathrm{M}}(k) \\ &= n\, \mathbb{F}_{\mathrm{M}}(n) + \sum_{k=0}^{n-1} \big(k - (k+1)\big)\, \mathbb{F}_{\mathrm{M}}(k).\end{aligned}$$

Or puisque M est à valeurs dans $[\![1,n]\!]$, sa fonction de répartition vaut 1 en n, donc

$$\mathbb{E}(\mathrm{M}) = n + \sum_{k=0}^{n-1} (-1)\, \mathbb{F}_{\mathrm{M}}(k) = \sum_{k=0}^{n-1} (1 - \mathbb{F}_{\mathrm{M}}(k)) = \sum_{k=0}^{n} (1 - \mathbb{F}_{\mathrm{M}}(k)).$$

c) De ce fait,

$$\begin{aligned}\mathbb{E}\big(\mathrm{M}(2,n)\big) &= \sum_{k=0}^{n}\left(1-\left(\frac{k}{n}\right)^2\right) = n+1-\frac{1}{n^2}\sum_{k=0}^{n}k^2 \\ &= n+1-\frac{1}{n^2}\cdot\frac{n(n+1)(2n+1)}{6} = (n+1)\left(1-\frac{2n+1}{6n}\right) \\ &= \boxed{(n+1)\,\frac{4n-1}{6n}.}\end{aligned}$$

d) En simplifiant chaque facteur en ne conservant que son terme prépondérant, on obtient

$$\boxed{\mathbb{E}\big(\mathrm{M}(2,n)\big) \underset{n\to\infty}{\sim} \frac{2}{3}\,n.}$$

2) En adaptant les premières questions, on trouve

$$\mathbb{E}\big(\mathrm{M}(d,n)\big) = \sum_{k=0}^{n}\left(1-\left(\frac{k}{n}\right)^d\right) = n\cdot\frac{1}{n}\sum_{k=0}^{n} f\left(\frac{k}{n}\right)$$

où l'on a introduit la fonction $f : \left\{\begin{array}{ccc}[0,1] & \to & \mathbb{R}\\ t & \mapsto & 1-t^d.\end{array}\right.$

On reconnaît *une somme de Riemann* et, la fonction f étant continue, on en déduit que

$$\mathbb{E}\big(\mathrm{M}(d,n)\big) \sim n\cdot\int_0^1 (1-t^d)\,\mathrm{d}t = n\cdot\left(1-\frac{1}{d+1}\right)$$

ce qui conduit au résultat annoncé.

3) a)

```
import numpy as np
import numpy.random as rd

def simulation_M(d, n, N):
    res = []
    for _ in range(N):
        X = rd.randint(1, n+1, d)
        res.append(max(X))
    return res


def moyenne_empirique(d, n, N):
    M = simulation_M(d, n, N)
    return np.mean(M)
```

b) On peut par exemple exécuter des scripts de la forme suivante :

```
d, N = 5, 100
print(d/(d+1),
[moyenne_empirique(d, n, N) / n for n in range(100, 110)])
```

c) On doit observer des valeurs d'autant plus proches de $d/(d+1)$ que n est plus grand !

Exercice 8 *Fonctions génératrices et stabilité*

1) • Supposons que $\mathrm{X} \hookrightarrow \mathcal{B}(p)$, et fixons une valeur de $t \in [0,1[$. D'après *le théorème de transfert*,

$$\begin{aligned} g_{\mathrm{X}}(t) &= \mathbb{E}(t^{\mathrm{X}}) \\ &= \sum_{k \in \mathrm{X}(\Omega)} t^k\, \mathbb{P}(\mathrm{X} = k) \\ &= t^0\, \mathbb{P}(\mathrm{X} = 0) + t^1\, \mathbb{P}(\mathrm{X} = 1) \\ &= 1 \cdot q + t \cdot p \qquad\qquad \text{(en notant } q = 1 - p) \\ &= q + p\, t. \end{aligned}$$

• Supposons ensuite que $\mathrm{X} \hookrightarrow \mathcal{U}([\![1,n]\!])$. Toujours d'après *le théorème de transfert*,

$$\begin{aligned} g_{\mathrm{X}}(t) &= \sum_{k=1}^{n} t^k\, \mathbb{P}(\mathrm{X} = k) \\ &= \sum_{k=1}^{n} t^k \cdot \frac{1}{n} = \frac{1}{n} \sum_{k=1}^{n} t^k = \frac{1}{n} t \cdot \frac{1 - t^n}{1 - t} \end{aligned}$$

puisqu'il s'agit d'une *somme géométrique* de raison $t \neq 1$, de premier terme t et contenant n termes.

2) a) Il s'agit là encore d'appliquer *le théorème de transfert*, ce qui est licite puisque la variable X est finie.

b) Si X et Y ont la même loi, alors $\mathbb{P}(\mathrm{X} = k) = \mathbb{P}(\mathrm{Y} = k)$ pour toute valeur de k, et donc l'expression précédente entraîne que $g_{\mathrm{X}}(t) = g_{\mathrm{Y}}(t)$ pour toute valeur de t, ce qui signifie que X et Y ont la même fonction génératrice.

c) Pour éviter les confusions, remplaçons par i l'indice de la somme définissant g_{X}. Par *linéarité de la dérivation*, si k est un entier positif, alors pour toute valeur de t

$$\begin{aligned} g_{\mathrm{X}}^{(k)}(t) &= \sum_{i=0}^{n} \mathbb{P}(\mathrm{X} = k) \cdot k(k-1)(k-2)\cdots(k-i+1)\, t^{k-i} \\ &= \sum_{i=0}^{n} \mathbb{P}(\mathrm{X} = k)\, \frac{k!}{(k-i)!}\, t^{k-i} \\ &= \sum_{i=0}^{n} \mathbb{P}(\mathrm{X} = k)\, i! \binom{k}{i} t^{k-i}. \end{aligned}$$

Or, lorsque $t = 0$, l'expression $\binom{k}{i} t^{k-i}$ est nulle pour $i > k$ (car le coefficient binomial est nul), pour $i < k$ (car l'exposant de $t = 0$ est strictement positif), et vaut 1 pour $i = k$ (car le coefficient binomial et la puissance sont tous deux égaux à 1). Dès lors

$$g_{\mathrm{X}}^{(k)}(0) = \mathbb{P}(\mathrm{X} = k)\, k!$$

ce qui équivaut au résultat demandé.

d) Si X et Y ont la même fonction génératrice, alors, pour tout $k \in \mathbb{N}$, les expressions $g_{\mathrm{X}}^{(k)}(0)$ et $g_{\mathrm{Y}}^{(k)}(0)$ sont égales, donc $\mathbb{P}(\mathrm{X} = k) = \mathbb{P}(\mathrm{Y} = k)$, ce qui revient à dire que X et Y ont la même loi.

3) Fixons une valeur de $t \in [0{,}1[$. On écrit

$$\begin{aligned} g_{\mathrm{X+Y}}(t) &= \mathbb{E}(t^{\mathrm{X+Y}}) \\ &= \mathbb{E}(t^{\mathrm{X}} \cdot t^{\mathrm{Y}}). \end{aligned}$$

Or, les variables X et Y étant indépendantes, c'est aussi le cas de t^{X} et t^{Y} (on peut par exemple invoquer un cas particulier du lemme des coalitions). Donc

$$\begin{aligned} g_{\mathrm{X+Y}}(t) &= \mathbb{E}(t^{\mathrm{X}}) \cdot \mathbb{E}(t^{\mathrm{Y}}) \\ &= g_{\mathrm{X}}(t) \cdot g_{\mathrm{Y}}(t). \end{aligned}$$

4) a) Introduisons n variables aléatoires $\mathrm{X}_1, \dots, \mathrm{X}_n$ de loi $\mathcal{B}(p)$ mutuellement indépendantes. Alors X et $\mathrm{X}_1 + \cdots + \mathrm{X}_n$ ont la même loi ($\mathcal{B}(n{,}p)$), donc la même fonction génératrice ; or, par indépendance mutuelle, $g_{\mathrm{X}_1+\cdots+\mathrm{X}_n} = g_{\mathrm{X}_1} \cdots g_{\mathrm{X}_n} = \gamma \cdots \gamma = \gamma^n$, ce qu'il fallait démontrer.

b) Puisque $g_{\mathrm{X}_1} = \gamma^{n_1}$, que $g_{\mathrm{X}_2} = \gamma^{n_2}$ et que X_1 et X_2 sont indépendantes, les résultats précédents entraînent que $g_{\mathrm{X}_1+\mathrm{X}_2} = \gamma^{n_1+n_2}$. Donc la fonction génératrice de $\mathrm{X}_1 + \mathrm{X}_2$ est celle d'une variable aléatoire de loi $\mathcal{B}(n_1 + n_2{,}p)$. Or on a vu que si deux variables aléatoires (ici à valeurs dans $[\![0{,}n_1+n_2]\!]$) ont la même fonction génératrice, alors elles ont la même loi, donc $\mathrm{X}_1 + \mathrm{X}_2 \hookrightarrow \mathcal{B}(n_1 + n_2{,}p)$.