Feuille Proba 3
Variables aléatoires discrètes - Corrigé [1]

Les feuilles d'exercices sont découpées en trois types d'exercice :
- Les *indispensables* : à savoir faire en autonomie.
- Les *exercices d'application* : pour mieux maîtriser et comprendre le cours.
- *Pour aller plus loin* : exercices présentant des développements mathématiques ou des études de modélisations de phénomènes issues d'autres disciplines.

Indispensables

Exercice 1. Soit X et Y deux variables aléatoires indépendantes de loi de Bernoulli de paramètre $0 \leqslant p \leqslant 1$
a) Loi de $U = 2X$.
b) Loi de $V = X + Y$.
c) Loi de $W = \max(X, Y)$.
d) Loi de $Z = \min(X, Y)$.

Solution : a) U prend les valeurs 0 et 2. On a $\mathrm{P(U = 2) = P(2X = 2) = P(X = 1) = p}$. Et $\mathrm{P(U = 0) = 1 - P(U = 2) = 1 - p}$.
b) V prend les valeurs 0,1 et 2. On a $\mathrm{P(V = 2) = P(X = 1, Y = 1) = P(X = 1)P(Y = 1)}$ (par indépendance) et donc $\mathrm{P(V = 2) = p^2}$. $\mathrm{P(V = 0) = P(X = 0, Y = 0) = P(X = 0)P(Y = 0) = (1 - p)^2}$. Et enfin $\mathrm{P(V = 1) = 1 - P(V = 0) - P(V = 2) = 1 - p^2 - (1 - p)^2 = 1 - p^2 - 1 + 2p - p^2 = 2p - 2p^2 = 2p(1 - p)}$.
c) W prend les valeurs 0 et 1. On a $\mathrm{P(W = 0) = P(X = 0, Y = 0) = P(X = 0)P(Y = 0) = (1 - p)^2}$. Et donc $\mathrm{P(W = 1) = 1 - (1 - p)^2 = 2p - p^2 = P(2 - p)}$.
d) Z prend les valeurs 0 et 1. On a $\mathrm{P(Z = 1) = P(X = 1, Y = 1) = P(X = 1)P(Y = 1) = p^2}$. Et donc $\mathrm{P(Z = 0) = 1 - p^2}$.

Exercice 2. Soit A un événement. On note $\mathbb{1}_A$ la variable aléatoire qui vaut 1 sur A et 0 ailleurs (fonction indicatrice de A). Montrer que $E(\mathbb{1}_A) = \mathrm{P(A)}$.

Solution :
Par définition de $\mathbb{1}_A$ qui vaut 1 si et seulement si $\omega \in A$, on peut écrire

$$E(\mathbb{1}_A) = 1.\mathrm{P}(\mathbb{1}_\mathrm{A} = 1) + 0.\mathrm{P}(\mathbb{1}_\mathrm{A} = 0) = \mathrm{P}(\mathbb{1}_\mathrm{A} = 1) = \mathrm{P(A)}.$$

Exercice 3. Un joueur joue à la roulette à 37 cases 10 euros sur le 19 et 100 euros sur "pair" : si la bille tombe sur 19 il touchera 36 fois sa mise (soit 360 euros) et si elle tombe sur "pair"(0 exclu), il touchera 2 fois sa mise ; dans tous les autres cas, sa mise va à la banque. Quelle est l'espérance de son gain ?

Solution :
La situation est donc la suivante : si la bille tombe sur 19, ce qui arrive avec une probabilité 1/37, le joueur réalise un gain de $360 - 110 = -250$ euros . Si la bille tombe sur un nombre pair (probabilité 18/37) le gain est de 200-110=90 euros . Enfin, sur toutes les autres cases (probabilité 18/37), le gain est négatif et vaut -110 euros .
L'espérance du gain est donc $250 \times 1/37 + 90 \times 18/37 - 110 \times 18/37 = -110/37 \approx -2,97$ euros .

Exercice 4. Soient X_1 et X_2 deux variables indépendantes, de lois $\mathcal{B}(n_1, p)$ et $\mathcal{B}(n_2, p)$. Quelle est la loi de $X_1 + X_2$?

1. Version du 31 mars 2020

Solution : Une manière intelligente de procéder est la suivante.
Considérons n_1+n_2 variables aléatoires de loi de Bernoulli de paramètre p indépendantes, $Y_1, \ldots, Y_{n_1}, Z_1, \ldots, Z_{n_2}$.
Puisqu'on ne s'intéresse qu'aux lois, on peut remplacer X_1 par $Y_1+\cdots+Y_{n_1}$ et X_2 par $Z_1+\cdots+Z_{n_2}$.
Comme toutes ces variables sont indépendantes, X_1+X_2 a alors même loi que $Y_1+\cdots+Y_{n_1}+Z_1+\cdots+Z_{n_2}$, c'est à dire une loi binômiale de paramètres n_1+n_2 et p.

Exercice 5. Soient X et Y deux variables aléatoires indépendantes telles que X suit une loi de Poisson de paramètre λ et Y une loi de Poisson de paramètre μ. Déterminer la loi de $X+Y$.

Solution : Tout d'abord, on remarque clairement que $X+Y$ est à valeurs dans $\mathbb{N}$. Soit donc $n \geqslant 0$. Dire que $X+Y=n$ revient à dire que $X=k$ et $Y=k'$ avec $k+k'=n$. De ce fait, en décomposant l'événement $X+Y=n$, on peut écrire

$$\mathrm{P(X+Y=n)} = \sum_{\mathrm{k=0}}^{\mathrm{n}} \mathrm{P(X=k, Y=n-k)}$$

En utilisant l'indépendance de X et Y, cette expression devient

$$\begin{aligned}
&\sum_{k=0}^{n} \mathrm{P(X=k)P(Y=n-k)} \\
&= \sum_{k=0}^{n} e^{-\lambda}\frac{\lambda^k}{k!}e^{-\mu}\frac{\mu^{n-k}}{(n-k)!} \\
&= \frac{e^{-(\lambda+\mu)}}{n!}\sum_{k=0}^{n}\binom{n}{k}\lambda^k\mu^{n-k} \\
&= \frac{e^{-(\lambda+\mu)}}{n!}(\lambda+\mu)^n
\end{aligned}$$

en utilisant le développement binomial. On en déduit que $X+Y$ suit une loi de Poisson de paramètre $\lambda+\mu$.

Exercice 6. Décrire un espace de probabilité $(\Omega, \mathcal{A}, P)$ de l'expérience aléatoire qui consiste à répartir au hasard n boules dans N cases.
On note X le nombre de boules tombant dans la première case.
a) Expliciter les $p_k = \mathrm{P(X=k)}, 0 \leqslant \mathrm{k} \leqslant \mathrm{n}$, $E(X)$ et $\mathrm{Var}(X)$
b) Donner la limite de p_k quand k étant fixé, n et N tendent vers l'infini de telle sorte que $\dfrac{n}{N} \to \lambda, 0 < \lambda < \infty$

Solution : Si l'on suppose les boules numérotées de 1 à n, on peut identifier leur ensemble avec $\{1, 2, \ldots, n\}$. De la même façon, l'ensemble des cases s'identifie à $\{1, 2, \ldots, N\}$. Un résultat de l'expérience aléatoire est donc représenté par le fait d'associer à chaque boule (i.e. à chaque élément de $\{1, 2, \ldots, n\}$) une case (i.e. un élément de $\{1, 2, \ldots, N\}$). Il est donc naturel de prendre pour Ω l'ensemble des applications de $\{1, 2, \ldots, n\}$ dans $\{1, 2, \ldots, N\}$. Ω est bien entendu muni de la probabilité uniforme définie par

$$\forall A \subset \Omega, \mathrm{P(A)} = \frac{\mathrm{CardA}}{\mathrm{N^n}}$$

a) Dire que $X=k$, c'est dire que les $n-k$ autres boules se répartissent dans les $N-1$ cases restantes ; on peut alors dénombrer les éléments de $(X=k)$ ainsi : on choisit k boules parmi les n, qu'on met dans la première case, et on répartit indifférement les $n-k$ autres dans les $N-1$ autres cases.

De ce fait, $\mathrm{Card}(X=k) = \binom{n}{k}(N-1)^{n-k}$ et donc $p_k = \mathrm{P(X=k)} = \dfrac{\mathrm{Card(X=k)}}{\mathrm{N^n}} = \binom{\mathrm{n}}{\mathrm{k}}(1-\dfrac{1}{\mathrm{N}})^{\mathrm{n-k}}(\dfrac{1}{\mathrm{N}})^{\mathrm{k}}$.
X suit une loi binômiale de paramètres n et $\frac{1}{N}$.
On a donc $E(X) = \dfrac{n}{N}$ et $\mathrm{Var}(X) = n.\dfrac{1}{N}(1-\dfrac{1}{N}) = n.\dfrac{N-1}{N^2}$.
b) On a $p_k = \dfrac{n!}{k!(n-k)!}(1-\dfrac{1}{N})^{n-k}(\dfrac{1}{N})^k = \dfrac{1}{k!}\dfrac{n!}{(n-k)!}(1-\dfrac{1}{N})^{n-k}(\dfrac{1}{N})^k$. D'où

$$p_k = \frac{1}{k!}\frac{n}{N}\frac{n-1}{N}\frac{n-2}{N}\cdots\frac{n-k+1}{N}(1-\frac{1}{N})^{n-k} \sim \frac{1}{k!}(\frac{n}{N})^k(1-\frac{1}{N})^{n-k} \sim \frac{\lambda^k}{k!}(1-\frac{1}{N})^{n-k}.$$

Or,

$$(1-\frac{1}{N})^{n-k} = \exp\{(n-k)\ln(1-\frac{1}{N})\} = \exp\{(n-k)(-\frac{1}{N}+o(\frac{1}{N}))\} \to \mathbf{e}^{-\lambda}$$

donc $p_k \sim e^{-\lambda}\dfrac{\lambda^k}{k!}$ et comme ce terme ne dépend pas de n et de N c'est la limite cherchée.

Pour aller plus loin et applications

Exercice 7. Lorsqu'on chauffe un filament de tungstène, on obtient une cathode qui émet des électrons. En plaçant ce dispositif à une extrémité d'un tube cathodique (un tube en verre sous vide), les électrons se propagent alors en ligne droite et vont frapper l'autre extrémité du tube (ce qui produit de la lumière lorsque cette extrémité est recouverte de phosphore ou plus généralement d'un matériau phosphorescent). On suppose que le temps entre deux émissions d'électrons est tel que sur une période donnée, le nombre total Z d'électrons qui ont été émis suit une loi de Poisson avec un certain paramètre $\lambda > 0$.
On suppose de plus qu'au milieu du tube cathodique se trouve une membrane, que chaque électron émis a une probabilité p de franchir. Le fait qu'un électron franchisse ou non la membrane est indépendant de ce que font les autres électrons. On note Y le nombre d'électrons émis qui ont franchi la membrane. Quelle est la loi de Y ?

Solution : Dans ce modèle, on suppose que la cathode est susceptible d'émettre un nombre d'éléctrons quelconque, autrement dit qu'il y en a une infinité à disposition.
Modélisons donc ces électrons en les numérotant par ordre d'émission éventuelle :
l'électron n est celui qui sera émis le n-ième, si toutefois $Z \geqslant n$.
A l' électron n (qu'il soit émis ou pas) est associée une marque X_n qui vaut 1 si l'électron peut franchir la barrière et 0 sinon.
Avec ces notations, on peut donc écrire la variable Y sous la forme

$$Y = \sum_{n=1}^{Z} X_n$$

et il s'agit de trouver la loi de Y. Notons tout d'abord que Y est à valeurs dans $\mathbb{N}$. Soit donc $m \geqslant 0$. En décomposant sur les valeurs de Z (nécessairement plus grandes que m, le nombre d'électrons émis devant être supérieur à celui des électrons émis qui franchissent la barrière), on peut écrire

$$\mathrm{P(Y=m)} = \sum_{\mathrm{k=m}}^{\infty} \mathrm{P}(\sum_{\mathrm{n=1}}^{\mathrm{Z}} \mathrm{X_n = m, Z = k})$$

$$= \sum_{k=m}^{\infty} \mathrm{P}(\sum_{\mathrm{n=1}}^{\mathrm{k}} \mathrm{X_n} = \mathrm{m}, \mathrm{Z} = \mathrm{k}).$$

Par l'indépendance des X_i et de Z, on obtient

$$\mathrm{P(Y = m)} = \sum_{\mathrm{k=m}}^{\infty} \mathrm{P}(\sum_{\mathrm{n=1}}^{\mathrm{k}} \mathrm{X_n} = \mathrm{m})\mathrm{P(Z = k)}.$$

Utilisant alors la loi de Z, et le fait que $\sum_{n=1}^{k} X_n$ suit une loi binomiale de paramètres k et p en tant que somme de variables de Bernoulli indépendantes, on obtient

$$\mathrm{P(Y = m)} = \sum_{\mathrm{k=m}}^{\infty} \mathrm{e}^{-\lambda} \frac{\lambda^{\mathrm{k}}}{\mathrm{k!}} \binom{\mathrm{k}}{\mathrm{m}} \mathrm{p^m (1-p)^{k-m}}$$

$$= \frac{e^{-\lambda}}{m!} p^m \lambda^m \sum_{k=m}^{\infty} \frac{\lambda^{k-m}(1-p)^{k-m}}{(k-m)!} = \frac{e^{-\lambda}}{m!} p^m \lambda^m \sum_{j=0}^{\infty} \frac{\lambda^j (1-p)^j}{j!} = e^{-p\lambda} \frac{(p\lambda)^m}{m!}.$$

On en déduit donc que Y suit une loi de Poisson de paramètre $p\lambda$.

Exercice 8. A) Une usine produit des pièces détachées et cherche à mettre en place une procédure pour estimer la qualité de la production. On met à la disposition du service qualité un (grand) nombre N de pièces, numérotées de 1 à N, et le laboratoire fait un test en en sélectionnant n au hasard sans replacement ($n \leqslant N$).
a) Décrire l'espace de probabilités (Ω_1, P_1) (resp. (Ω_2, P_2)) associé à cette expérience aléatoire quand on regarde la suite (resp. l'ensemble) des numéros obtenus.
b) On suppose qu'une proportion p, $0 < p < 1$, des N pièces sont défectueuses, avec $pN > n$. On note X le nombre de pièces défectueuses figurant parmi les n pièces choisies.
(i) En considérant X définie sur Ω_2 expliciter les $p_k = \mathrm{P_2(X = k)}$.
(ii) En considérant X définie sur Ω_1 montrer que X peut s'écrire comme la somme de n variables aléatoires à valeurs dans $\{0, 1\}$

$$X = \sum_{k=1}^{n} Z_k$$

Calculer les $E(Z_k), \mathrm{cov}(Z_k, Z_l)$. En déduire
$E(X) = np$ et $\mathrm{Var}(X) = np(1-p)(1 - \frac{n-1}{N-1})$
B) On considère les mêmes pièces, mais on en choisit n avec replacement (n peut être plus grand que N). On note Y le nombre de pièces défectueuses observées.
a)
Décrire l'espace de probabilité (Ω, P) associé à cette expérience aléatoire
b) Expliciter les $q_k = \mathrm{P(Y = k)}, 0 \leqslant \mathrm{k} \leqslant \mathrm{n}$ et montrer que, pour k et n fixés, $p_k \to q_k$ quand $N \to \infty$. Commenter.

Solution : A)
a) L'ensemble des N pièces est assimilé à $\{1, 2, \ldots, N\}$. La différence entre la situation où on regarde la suite ou l'ensemble des numéros obtenus vient évidemment du fait qu'on tient compte de l'ordre dans le premier cas mais pas dans le second.
De ce fait, Ω_1 est l'ensemble des n-uplets de points distincts de $\{1, 2, \ldots, N\}$. $\mathrm{Card}\Omega_1 = A_N^n = \frac{N!}{(N-n)!}$. P_1, probabilité uniforme est définie par $\mathrm{P_1(A)} = \frac{\mathrm{CardA}}{\mathrm{A_N^n}}$ pour $A \subset \Omega_1$.

Ω_2 est l'ensemble des parties de $\{1, 2, \ldots, N\}$ à n éléments. $\mathrm{Card}\Omega_2 = \binom{N}{n} = \dfrac{N!}{(N-n)!n!}$ et $P_2(A) = \dfrac{\mathrm{Card}A}{\binom{N}{n}}$ pour $A \subset \Omega_2$.

b) L'hypothèse dit donc que $p.N$ pièces sont défectueuses parmi les N pièces.

(i) Pour $0 \leqslant k \leqslant p.N$, dire que $(X = k)$ revient à dire que la partie à n éléments choisie contient k pièces choisies parmi les $p.N$ défectueuses, et $n - k$ parmi les $N - pN = (1-p)N$ restantes.

Donc,

$$\mathrm{Card}(X = k) = \binom{pN}{k}.\binom{(1-p)N}{n-k}$$

et

$$p_k^N = \frac{\binom{pN}{k}.\binom{(1-p)N}{n-k}}{\binom{N}{n}} \tag{1}$$

(ii) Dans le cas où l'on considère la suite des résultats, on doit avoir un indicateur pour chaque pièce choisie du fait qu'elle est ou non défectueuse.

Pour $1 \leqslant k \leqslant n$, $Z_k = 1$ si la k-ième pièce choisie est défectueuse, $Z_k = 0$ sinon.

Clairement $X = \sum_{k=1}^{n} Z_k$.

Notons que $Z_k = 1$ signifie que la k-ième pièce est choisie défectueuse. Bien entendu, le fait de dire qu'il s'agit de la k-ième pièce relève de l'arbitraire de la numérotation que nous donnons au modèle et donc par symétrie nous pouvons affirmer que tous les Z_k ont même loi. D'après les hypothèses, la probabilité de tirer une pièce défectueuse est p. On a donc $E(Z_k) = p$.

Regardons maintenant $\mathrm{Cov}(Z_k, Z_l) = E(Z_k Z_l) - E(Z_k)E(Z_l)$.

Si $k = l$, il s'agit de $E(Z_k^2) - E(Z_k) = 0$.

Supposons $k \neq l$.

On a $E(Z_k Z_l) = \mathrm{P(Z_k = 1, Z_l = 1)}$. Le même raisonnement de symétrie que précédemment nous permet d'affirmer qu'on a $E(Z_k Z_l) = E(Z_1 Z_2)$. Par ailleurs, raisonnons : si $Z_1 = 1$, pour choisir la deuxième pièce, il restera $pN - 1$ pièces défectueuses parmi les $N - 1$ pièces restantes et les chances de choisir la deuxième défectueuse sont donc $\dfrac{pN-1}{N-1}$.

Donc $E(Z_k Z_l) = \dfrac{pN}{N}\dfrac{pN-1}{N-1}$ et $\mathrm{Cov}(Z_k, Z_l) = p.\dfrac{pN-1}{N-1} - p^2 = \mathrm{P}(\dfrac{\mathrm{pN-1}}{\mathrm{N-1}} - \mathrm{p}) = \mathrm{p}.\dfrac{\mathrm{p-1}}{\mathrm{N-1}}$

On déduit

$$E(X) = nE(Z_1) = np$$

$$\begin{aligned}
\mathrm{Var}(X) = \mathrm{Var}(\sum_{k=1}^{n} Z_k) &= \sum_{k=1}^{n} \mathrm{Var}(Z_k) + 2 \sum_{1 \leqslant k < l \leqslant n} \mathrm{Cov}(Z_k, Z_l) \\
&= n\mathrm{P(1-p) + (n^2 - n)p\frac{p-1}{N-1}} \\
&= n\mathrm{P(1-p)[1 - \frac{n-1}{N-1}]}
\end{aligned}$$

B)

a) Cette fois ci, puisqu'on remplace les pièces, l'espace Ω est simplement $\{1, 2, \ldots, N\}^n$. La probabilité uniforme P est définie par $\mathrm{P(A) = \dfrac{CardA}{N^n}}$ pour $A \subset \Omega$.

b) Un tirage tel que $Y = k$ peut être décrit de la façon suivante : on choisit les k places où se trouvent les pièces défectueuses puis on remplit ces places avec une pièce défectueuse choisie parmi les pN (éventuellement plusieurs fois la même) et les autres places avec $n - k$ pièces parmi les $N - pN = (1-p)N$ restantes.

On a donc $\mathrm{Card}(Y=k) = \binom{n}{k}(pN)^k(N-pN)^{n-k}$ d'où $q_k = \dfrac{\mathrm{Card}(Y=k)}{N^n} = \binom{n}{k}p^k(1-p)^{n-k}$ (loi binômiale de paramètres n et p).
On a alors :

$$p_k^N = \frac{\binom{pN}{k}\binom{(1-p)N}{n-k}}{\binom{N}{n}} = \frac{\frac{(pN)!}{k!(pN-k)!}\,\frac{((1-p)N)!}{(n-k)![(1-p)N-n+k]!}}{\frac{N!}{n!(N-n)!}}.$$

Donc, en "nettoyant" l'expression, on obtient

$$p_k^N \sim \frac{\frac{1}{k!}(pN)^k\frac{1}{(n-k)!}[(1-p)N]^{n-k}}{\frac{1}{n!}N^n} = \binom{n}{k}p^k(1-p)^{n-k} = q_k$$

L'interprétation est alors la suivante : si N est très grand, le fait de prendre les pièces avec ou sans remplacement n'est pas très différent puisque la proportion de pièces défectueuses restantes ne change presque pas. Il est donc logique que la probabilité soit asymptotiquement celle obtenue avec replacement. □

Exercice 9. Construire un exemple d'espace de probabilités et de variable aléatoire discrète non constante pour laquelle $\mathrm{Var}(X) = 0$. Formuler une condition nécessaire et suffisante pour qu'une variable aléatoire discrète ait une variance nulle.

Solution :
Par exemple, on peut construire le modèle suivant *ad hoc*. Soit $\Omega = \{0,1\}$ muni de la probabilité qui met toute la masse sur le point 1 : $\mathrm{P}(\{1\}) = 1$ (et donc nécessairement $\mathrm{P}(\{0\}) = 0$.
On considère la variable X définie sur Ω par $X(0) = 1$ et $X(1) = 0$. Elle n'est donc pas constante, au sens des applications. Pourtant, on a, utilisant le fait que $X^2 = X$,

$$E(X^2) = E(X) = 1.\mathrm{P(X=1)} + 0.\mathrm{P(X=0)} =$$
$$= \mathrm{P(X=1)} = \mathrm{P}(\{0\}) = 0$$

et donc $\mathrm{Var}(X) = 0$.
La subtilité qui peut paraître spécieuse sur cet exemple est en fait d'importance : elle montre qu'il convient de bien distinguer entre un événement vide et un événement de probabilité nulle. Cette dernière catégorie joue un rôle très important en calcul des probabilités.

Exercice 10. Soit X une variables aléatoire à valeurs entières. On pose $p_k = \mathrm{P(X=k)}$ et $q_k = \displaystyle\sum_{j\geqslant k+1} p_j$

Montrer que $E(X) = \displaystyle\sum_{j\geqslant 0} q_j$

Solution : Par définition, on a $E(X) = \displaystyle\sum_{k=0}^{\infty} kp_k$. Or,

$$\sum_{k\geqslant 0} q_k = \sum_{k\geqslant 0}\sum_{j\geqslant k+1} p_j = \sum_{j\geqslant 1}\sum_{k=0}^{j-1} p_j$$

où l'interversion des sommes est justifiée car tous les termes sont positifs. Comme $\displaystyle\sum_{k=0}^{j-1} p_j = jp_j$, on en déduit que

$$\sum_{k\geqslant 0} q_k = \sum_{j\geqslant 1} jp_j = \sum_{j\geqslant 0} jp_j = E(X)$$

ce qu'on voulait démontrer.

Encore plus exigeant (et plus excitant)

Deux applications fondamentales de la théorie des probabilités.

Exercice 11. (Ruine du joueur) Un joueur va au casino avec une fortune $a \in \mathbb{N}$. A chaque partie, il peut gagner 1 euro avec une probabilité p et perdre 1 euro avec une probabilité $q = 1 - p$.
Son but est de jouer jusqu'à l'obtention de la fortune $c \geqslant a, c \in \mathbb{N}$ mais il doit s'arrêter s'il est ruiné.
On note $s_c(a)$ sa probabilité de succès (atteindre c avant la ruine).
a) Calculer $s_c(0)$ et $s_c(c)$
b) Montrer, pour $a > 0$, en raisonnant sur ce qui s'est passé au premier coup, la relation

$$s_c(a) = ps_c(a+1) + qs_c(a-1)$$

c) Déduire la valeur de $s_c(a)$
d) Application numérique :
Calculer la valeur précédente avec $a = 900, c = 1000$; $a = 100, c = 20000$ dans les cas $p = 0,5$ et $p = \dfrac{18}{38}$.

Solution :
Le but de cet exercice est l'étude d'un des plus fameux problèmes du calcul des probabilités, celui de la ruine du joueur.
Commençons par formaliser la situation : à chaque instant n, le joueur joue et obtient un résultat aléatoire X_n valant 1 ou -1 suivant qu'il gagne ou perd. Par hypothèse, les variables (X_n) sont supposées indépendantes de même loi donnée par

$$\mathrm{P(X_n = 1) = p = 1 - P(X_n = 0) = 1 - q}$$

Dans ce jeu, le joueur s'arrête de jouer à deux conditions : soit il est ruiné, soit sa fortune a atteint un objectif fixé à l'avance c.
a) S'il part avec une fortune nulle, il ne peut commencer à jouer. De ce fait, on a $s_0(0) = 1$ et $s_c(0) = 0$ si $c > 0$.
Si, au contraire, il part avec une fortune c, il ne commence pas non plus à jouer et $s_c(s) = 1$.
b)
Proposons d'abord une démonstration intuitive. Au premier coup le joueur a une probabilité p (resp. q) d'avoir sa fortune qui passe à $a+1$ (resp. $a-1$). Comme ensuite tout recommence indépendamment du premier coup (d'où la multiplication) il repart avec $a+1$ (resp. $a-1$) comme nouvelle fortune initiale et se retrouve dans la situation d'un joueur qui entre au casino avec la fortune en question. On est dans une situation dite *markovienne* .
On peut se contenter dans un premier temps de cette approche intuitive ; néanmoins la démonstration complète (qui suit) nécessite plus de technique.
Notons, pour $n \geqslant 0$, A_n^a l'événement : "le joueur réalise son but au n-ième coup en partant de la fortune a", c'est à dire qu'au n-ième coup pour la première fois sa fortune atteint c sans qu'il n'ait jamais été ruiné auparavant. L'événement "le joueur a atteint son but à partir de la fortune a" est donc $\bigcup_{n \in \mathbb{N}} A_n^a$.

On peut écrire, pour $a < c$, notant $S_k = \sum_{r=1}^{k} X_r$,

$$\begin{aligned} A_n^a = (0 < a + S_1 < c, 0 < a + S_2 < c, \dots \\ \dots, 0 < a + S_{n-1} < c, a + S_n = c) \end{aligned} \qquad (2)$$

Si on pose $S'_k = \sum_{r=2}^{k+1} X_r$ (on a donc $S_{k+1} = S'_k + X_1$), décomposant A_n^a suivant les valeurs de X_1, on a pour $n \geqslant 1$,

$$A_n^a = A_n^a \cap (X_1 = 1) \cup A_n^a \cup (X_1 = -1)$$

$= (X_1 = 1)$

$$\cap(a+1 < c, 0 < (a+1) + S'_1 < c, \ldots$$
$$\ldots, 0 < (a+1) + S'_{n-2} < c, (a+1) + S'_{n-1} = c)$$

$\bigcup (X_1 = -1) \cap$

$$(0 < a-1, 0 < (a-1) + S'_1 < c, \ldots$$
$$\ldots, 0 < (a-1) + S'_{n-2} < c, (a-1) + S'_{n-1} = c) \tag{3}$$

Réécrivons (**??**) sous la forme

$$A_n^a = (X_1 = 1) \cap \widetilde{A}_{n-1}^{(a+1)} \cup (X_1 = -1) \cap \widetilde{A}_{n-1}^{(a-1)} \tag{4}$$

Puisque les (X_i) sont indépendantes, on constate que X_1 est indépendante de $\widetilde{A}_{n-1}^{(a+1)}$ et $\widetilde{A}_{n-1}^{(a-1)}$ et que, puisque S'_k a même loi que S_k, $\widetilde{A}_{n-1}^{(a+1)}$ (resp. $\widetilde{A}_{n-1}^{(a-1)}$) ont mêmes probabilités que A_{n-1}^{a+1} (resp. A_{n-1}^{a-1}). En passant aux probabilités dans (**??**), on a pour $n \geqslant 1$,

$$\mathrm{P(A_n^a) = pP(A_{n-1}^{a+1}) + qP(A_{n-1}^{a-1})} \tag{5}$$

Remarquant alors que les ensembles A_n^a (resp. A_{n-1}^{a+1}, resp. A_{n-1}^{a-1}) sont disjoints, on a en sommant (**??**) en n

$$\mathrm{P(\bigcup_{n\geqslant 1} A_n^a) = pP(\bigcup_{n\geqslant 0} A_n^{a+1}) + qP(\bigcup_{n\geqslant 0} A_n^{a-1})} \tag{6}$$

Notons enfin que comme $a > 0$ et $a < c$, le joueur ne réussit pas au 0-ième coup et donc $\mathrm{P(A_0^a) = 0}$. On tire donc de (**??**) l'égalité cherchée

$$s_c(a) = ps_c(a+1) + qs_c(a-1) \tag{7}$$

c) Pour c fixé, on a donc affaire à une suite $(s_c(a))_{0\leqslant a\leqslant c}$ vérifiant la relation de récurrence (**??**) et les conditions aux bornes du a).

Les résultats classiques sur les suites récurrentes disent que la solution va s'écrire en fonction des solutions de l'équation $pr^2 - r + q = 0$. Un calcul rapide montre que $\dfrac{1+ \mid 1-2p \mid}{2p}$ et $\dfrac{1- \mid 1-2p \mid}{2p}$ sont solutions et deux cas se présentent donc ;

<u>1er cas</u> : $p \neq q$

Alors les deux solutions sont $r_1 = \dfrac{q}{p}$ et $r_2 = 1$.

On cherche alors les solutions sous la forme $s_c(a) = \alpha\rho^a + \beta$ où $\rho = \dfrac{q}{p}$. En injectant les conditions aux bords $s_c(0) = 0$ et $s_c(c) = 1$, on obtient $0 = \alpha + \beta$ et $1 = \alpha\rho^c + \beta$, et donc

$$s_c(a) = \frac{\rho^a - 1}{\rho^c - 1}$$

<u>2ème cas</u> : $p = q = \dfrac{1}{2}$

Dans ce cas, on tire de (**??**), $s_c(a+1) - s_c(a) = s_c(a) - s_c(a-1)$ et donc $s_c(a) = a.k$ où $k = s_c(1)$. Comme $s_c(c) = 1$, on en déduit que $s_c(a) = \dfrac{a}{c}$.

d) Pour estimer les applications numériques données, il s'agit donc de distinguer le cas $p = \frac{1}{2}$ (jeu de pile ou face) et $p = \frac{18}{38} \approx 0,47$ (roulette américaine).
1er jeu : $a = 900, c = 1000$
a) pile ou face : $s_c(a) = \frac{a}{c} = \frac{9}{10}$
b) roulette : $\rho = \frac{10}{9}$ et $s_c(a) \approx 3.10^{-5}$.

2ème jeu : $a = 100, c = 20000$
a) pile ou face : $s_c(a) = 5.10^{-3}$
b) roulette : $s_c(a) \approx 3.10^{-911}$! ! ! ! !

Remarque :
Ces résultats spectaculaires, pas du tout intuitifs *a priori*, montrent bien toute la subtilité d'un jeu prolongé. Sur un coup, la différence entre le jeu de pile ou face et la roulette n'est pas flagrante, mais elle devient vertigineuse lors de sa répétition. C'est la tromperie (légale...) utilisée par les casinos pour appâter les joueurs invétérés.

Exercice 12. (Arrêt optimal) Une princesse a n prétendants numérotés par ordre de mérite décroissant 1,2,...,n. Elle doit en choisir un. Le problème est qu'ils défilent un par un au hasard devant elle et qu'elle ne peut revenir sur son choix si elle en a laissé partir un. Elle doit donc adopter une stratégie pour avoir le plus de chance de choisir le meilleur...
Soit Ω l'ensemble des permutations de $\{1, 2, \ldots, n\}$. Ω est muni de la probabilité uniforme.
$\sigma \in \Omega$ représente un tirage du hasard (les prétendants défilent dans l'ordre $\sigma(1), \ldots, \sigma(n)$).
Pour $1 \leqslant k \leqslant n$, on introduit la variable Y_k qui est le rang de $\sigma(k)$ dans l'ensemble $\{\sigma(1), \ldots, \sigma(k)\}$ rangé en ordre décroissant : $Y_k = 1$ signifie que $\sigma(k)$ est le plus grand parmi l'ensemble $\{\sigma(1), \ldots, \sigma(k)\}$, $Y_k = 2$ signifie qu'il y a exactement un élément de $\{\sigma(1), \ldots, \sigma(k)\}$ plus grand que $\sigma(k)$ etc...
Par convention, on pose $Y_{n+1} = 0$.
a) Montrer que

$$F : \sigma \to (Y_1(\sigma), \ldots, Y_n(\sigma))$$

définit une bijection de Ω sur $\Pi = \{1\} \times \cdots \times \{1, 2, \ldots, n\}$
b) En déduire que les variables Y_j sont indépendantes et que Y_k suit la loi uniforme sur $\{1, 2, \ldots, k\}$.
c) Soit $\tau_r = \inf\{k \geqslant r, Y_k = 1\}$ $(= n + 1$ si cet ensemble est vide).
Calculer la probabilité pour qu'au temps τ_r, le prétendant qui se présente soit le meilleur (i.e. le numero 1). Comment choisir r^* pour maximiser la probabilité précédente ? Trouver un équivalent de r^* quand n tend vers l'infini.

Solution : a) Observons que $Y_2 = 1$ signifie que $\sigma(2) > \sigma(1)$ tandis que $Y_2 = 2$ dit que $\sigma(2) < \sigma(1)$. De même, $Y_3 = k_3$ permet de classer $\sigma(3)$ par rapport à $\sigma(1)$ et $\sigma(2)$ et un raisonnement par récurrence immédiat montre que la donnée de $(k_1, k_2, \ldots, k_n)$ revient à ordonner $\sigma(1), \sigma(2), \ldots, \sigma(n)$ sous la forme

$$\sigma(i_1) < \sigma(i_2) < \cdots < \sigma(i_n). \tag{8}$$

Comme σ est une permutation de $\{1, 2, \ldots, n\}$, (**??**) impose

$$\sigma(i_1) = 1, \sigma(i_2) = 2, \ldots, \sigma(i_n) = n$$

ce qui définit σ de manière unique. F est bien une bijection.
b) On déduit alors que pour $(k_1, k_2, \ldots, k_n) \in \Pi$

$$\mathrm{P}(\mathrm{Y_1 = k_1, \ldots, Y_n = k_n}) = \mathrm{P}(\mathrm{F^{-1}(k_1, \ldots, k_n)}) = \frac{1}{\mathrm{n!}}$$

En sommant sur tous les n-uples $(k_1, \dots, k_{i-1}, k_{i+1}, \dots, k_n)$, avec

$$k_1 = 1, k_2 \in \{1, 2\}, \dots, k_{i-1} \in \{1, 2, \dots, i-1\}, k_{i+1} \in \{1, 2, \dots, i+1\}, \dots k_n \in \{1, 2, \dots, n\}$$

on obtient

$$\mathrm{P}(\mathrm{Y_i = k_i}) = \frac{1 \cdot 2 \dots (\mathrm{i}-1) \cdot (\mathrm{i}+1) \dots \mathrm{n}}{\mathrm{n}!} = \frac{1}{\mathrm{i}}$$

et Y_i suit donc une loi uniforme sur $\{1, 2, \dots, i\}$. De plus

$$\mathrm{P}(\mathrm{Y_1 = k_1, \dots, Y_n = k_n}) = \prod_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{P}(\mathrm{Y_i = k_i})$$

d'où l'indépendance des Y_i.

c) Il s'agit de calculer la probabilité de l'événement $\{Y_{\tau_r} = 1, Y_{\tau_r+1} > 1, \dots, Y_n > 1\}$. Décomposant sur les valeurs de τ_r, on peut écrire

$$\begin{aligned}
&\mathrm{P}(\mathrm{Y}_{\tau_\mathrm{r}} = 1, \mathrm{Y}_{\tau_\mathrm{r}+1} > 1, \dots, \mathrm{Y_n} > 1) \\
= &\sum_{k=r+1}^{n} \mathrm{P}(\tau_\mathrm{r} = \mathrm{k}, \mathrm{Y_k} = 1, \mathrm{Y_{k+1}} > 1, \dots, \mathrm{Y_n} > 1) \\
= &\sum_{k=r+1}^{n} \mathrm{P}(\mathrm{Y_r} > 1, \mathrm{Y_{r+1}} > 1, \dots, \mathrm{Y_{k-1}} > 1, \mathrm{Y_k} = 1, \mathrm{Y_{k+1}} > 1, \dots, \mathrm{Y_n} > 1) \\
= &\sum_{k=r+1}^{n} \frac{r-1}{r} \frac{r}{r+1} \cdots \frac{k-2}{k-1} \cdot \frac{1}{k} \cdot \frac{k}{k+1} \cdots \frac{n-1}{n} \\
= &\frac{r-1}{n} \sum_{k=r+1}^{n} \frac{1}{k-1} \\
= &\frac{r-1}{n} \sum_{k=r}^{n-1} \frac{1}{k}.
\end{aligned}$$

On doit choisir r pour maximiser cette probabilité. Posons

$$s_r = (r-1) \frac{1}{n} \sum_{k=r}^{n-1} \frac{1}{k}. \tag{9}$$

On a

$$\begin{aligned}
s_{r+1} - s_r = \; & r \frac{1}{n} \sum_{k=r+1}^{n-1} \frac{1}{k} - (r-1) \frac{1}{n} \sum_{k=r}^{n-1} \frac{1}{k} \\
= \; & \frac{1}{n} \sum_{k=r+1}^{n-1} \frac{1}{k} - \big(1 - \frac{1}{r}\big) \frac{1}{n} \\
= \; & \frac{1}{n} \Big[\sum_{k=r+1}^{n-1} \frac{1}{k} - \frac{r-1}{r} \Big]
\end{aligned} \tag{10}$$

et comme $\displaystyle\sum_{k=r+1}^{n-1} \frac{1}{k} - \frac{r-1}{r}$ est décroissant en r, et $s_2 - s_1 > 0$, $s_n - s_{n-1} < 0$, il existe un unique r^* tel que $s_{r^*+1} - s_{r^*} < 0$ et $s_{r^*} - s_{r^*-1} > 0$: c'est le r^* qui maximise (**??**). D'après (**??**), on a

$$\frac{r^*-1}{r^*} > \sum_{k=r^*+1}^{n-1} \frac{1}{k} \quad \text{et} \quad \frac{r^*-2}{r^*-1} < \sum_{k=r^*}^{n-1} \frac{1}{k}.$$

Posons $m_{r^*} = \sum_{k=r^*}^{n-1} \frac{1}{k}$. On a

$$1 - \frac{1}{r^*} + \frac{1}{r^*} > m_{r^*} > 1 - \frac{1}{r^* - 1}$$

soit

$$1 > m_{r^*} > 1 - \frac{1}{r^* - 1}. \tag{11}$$

Il est classique que $\sum_{k=1}^{n-1} \frac{1}{k} - \log(n-1) = \gamma + o(1)$ (où γ est la constante d'Euler). On a donc

$$m_{r^*} = \log \frac{n-1}{r^* - 1} + o(1).$$

On déduit alors de (**??**) que $\log \frac{n}{r^*} \sim 1$ soit $r^* \sim \frac{n}{\mathrm{e}}$.