VECTEURS ALÉATOIRES

♦ ***Exercice 1.*** [$\circ$]
On dispose de n boites numérotées de 1 à n. La boite numérotée k contient k boules numérotées de 1 à k. On choisit au hasard une boule dans une boite. On note X le numéro de la boite et Y celui de la boule.

1. Déterminer la loi conjointe de (X, Y).
2. En déduire la loi de Y (on laissera sous forme d'une somme). Déterminer $E(Y)$ puis un équivalent de $E(Y)$ quand $n \longrightarrow +\infty$.

1. On a $X(\Omega) = [\![1; n]\!]$ et $Y(\Omega) = [\![1; n]\!]$. Par ailleurs, on a $Y \leqslant X$. Par conséquent, on a
$$\boxed{(X, Y)(\Omega) = \{(k, \ell) \in [\![1; n]\!]^2 : k \geqslant \ell\}.}$$
Soit $(k, \ell) \in (X, Y)(\Omega)$. On a
$$P(X = k\,,\, Y = \ell) = P(X = k)P(Y = \ell \,|\, X = k).$$
Il est clair que $X \hookrightarrow \mathscr{U}([\![1; n]\!])$ et que, pour tout $k \in [\![1; n]\!]$, on a $Y \,|\, (X = k) \hookrightarrow \mathscr{U}([\![1; k]\!])$, donc
$$P(X = k\,,\, Y = \ell) = \frac{1}{nk}\cdot$$
En conclusion,
$$\boxed{\forall (k, \ell) \in [\![1; n]\!]^2\,, \qquad P(X = k\,,\, Y = \ell) = \begin{cases} \dfrac{1}{nk} & \text{si } k \geqslant \ell, \\ 0 & \text{sinon.} \end{cases}}$$
2. On a
$$\boxed{Y(\Omega) = [\![1; n]\!].}$$
Pour tout $\ell \in Y(\Omega)$, on a
$$P(Y = \ell) = \sum_{k=1}^{n} P(X = k\,,\, Y = \ell) = \sum_{k=\ell}^{n} \frac{1}{nk},$$
donc
$$\boxed{\forall \ell \in [\![1; n]\!]\,, \qquad P(Y = \ell) = \frac{1}{n}\sum_{k=\ell}^{n} \frac{1}{k}\cdot}$$

On a

$$\begin{aligned}
E(Y) &= \sum_{\ell=1}^{n} \ell P(Y=\ell) \\
&= \sum_{\ell=1}^{n} \frac{\ell}{n} \sum_{k=\ell}^{n} \frac{1}{k} \\
&= \frac{1}{n} \sum_{1\leqslant \ell \leqslant k \leqslant n} \frac{\ell}{k} \\
&= \frac{1}{n} \sum_{k=1}^{n} \frac{1}{k} \sum_{\ell=1}^{k} \ell \\
&= \frac{1}{n} \sum_{k=1}^{n} \frac{1}{k} \frac{k(k+1)}{2} \\
&= \frac{1}{n} \sum_{k=1}^{n} \frac{k+1}{2} \\
&= \frac{1}{2n} \Big(\sum_{k=1}^{n} k + \sum_{k=1}^{n} 1 \Big) \\
&= \frac{1}{2n} \Big(\frac{n(n+1)}{2} + n \Big),
\end{aligned}$$

donc

$$\boxed{E(Y) = \frac{n+3}{4} \underset{n\to+\infty}{\sim} \frac{n}{4}.}$$

♦ ***Exercice 2.*** [○]

On lance un dé honnète et on note X le nombre obtenu. On relance ensuite X fois le dé et on note Y le nombre de 1 obtenus parmi ces X tirages.

Déterminer la loi conjointe du couple (X, Y). Les variables X et Y sont-elles indépendantes ?

On a $X(\Omega) = [\![1;6]\!]$ et $Y(\Omega) = [\![0;6]\!]$. Par ailleurs, on a clairement $Y \leqslant X$. Donc

$$\boxed{(X,Y)(\Omega) = \{(k,\ell) \in [\![1;6]\!] \times [\![0;6]\!] : k \geqslant \ell\}.}$$

Soit $(k, \ell) \in (X,Y)(\Omega)$. On a

$$P(X=k\,,\,Y=\ell) = P(X=k)P(Y=\ell \,|\, X=k).$$

Or :

- il est clair que X suit la loi uniforme sur $[\![1;6]\!]$, d'où
$$P(X=k) = \frac{1}{6}\,;$$
- sachant que $X = k$, Y compte le nombre de succès (« obtenir un 1 ») dans une suites de k lancers de dé indépendants, avec une probabilité de succès égale à $1/6$, donc $Y \,|\, (X = k)$ suit la loi $\mathscr{B}(k; 1/6)$, c'est-à-dire
$$P(Y=\ell \,|\, X=k) = \binom{k}{\ell}\Big(\frac{1}{6}\Big)^{\ell}\Big(\frac{5}{6}\Big)^{k-\ell}.$$

Il s'ensuit que

$$P(X=k\,,\,Y=\ell) = \frac{1}{6}\binom{k}{\ell}\Big(\frac{1}{6}\Big)^{\ell}\Big(\frac{5}{6}\Big)^{k-\ell}.$$

En conclusion,

$$\boxed{\forall (k,\ell) \in [\![1;6]\!] \times [\![0;6]\!]\,, \qquad P(X=k\,,\,Y=\ell) = \binom{k}{\ell}\Big(\frac{1}{6}\Big)^{\ell+1}\Big(\frac{5}{6}\Big)^{k-\ell}.}$$

Remarque : C'est le coefficient binomial qui s'occupe d'annuler cette probabilité dès que $\ell > k$.

On a

$$P(X=1\,,\,Y=6) = 0 \qquad \text{et} \qquad P(X=1)P(Y=6) \neq 0,$$

donc

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

♦ ***Exercice 3.*** [$\star$]

Soient X, Y et Z trois variables aléatoires mutuellement indépendantes suivant toutes les trois la loi uniforme sur $[\![1;2n]\!]$. Calculer $P(X=Y)$, $P(X \geqslant 2Y)$ et $P(X+Y \leqslant Z)$.

On a

$$\begin{aligned}
P(X=Y) &= \sum_{k=1}^{2n} P(X=Y\,,\,Y=k) \qquad \text{par filtration}\\
&= \sum_{k=1}^{2n} P(X=k\,,\,Y=k)\\
&= \sum_{k=1}^{2n} P(X=k)P(Y=k) \qquad \text{par indépendance}\\
&= \sum_{k=1}^{2n} \frac{1}{(2n)^2}\\
&= \frac{1}{2n}
\end{aligned}$$

et

$$\begin{aligned}
P(X \geqslant 2Y) &= \sum_{k=1}^{2n} P(X \geqslant 2X\,,\,Y=k) \qquad \text{par filtration}\\
&= \sum_{k=1}^{2n} P(X \geqslant 2k\,,\,Y=k)\\
&= \sum_{k=1}^{2n} P(X \geqslant 2k)P(Y=k) \qquad \text{par indépendance}\\
&= \sum_{k=1}^{n} P(X \geqslant 2k)P(Y=k)\\
&= \sum_{k=1}^{n} \frac{2n-2k+1}{2n}\frac{1}{2n}\\
&= \frac{1}{4n^2}\Big(\sum_{k=1}^{n} 2n - 2\sum_{k=1}^{n} k + \sum_{k=1}^{n} 1\Big)\\
&= \frac{1}{4n^2}\Big(2n^2 - n(n+1) + n\Big)\\
&= \frac{1}{4}
\end{aligned}$$

et

$$\begin{aligned}
P(X+Y \leqslant Z) &= \sum_{k=1}^{2n}\sum_{\ell=1}^{2n} P(X+Y \leqslant Z\,,\, Y=\ell\,,\, Z=k) \qquad \text{par filtration}\\
&= \sum_{k=1}^{2n}\sum_{\ell=1}^{2n} P(X+\ell \leqslant k\,,\, Y=\ell\,,\, Z=k)\\
&= \sum_{k=1}^{2n}\sum_{\ell=1}^{2n} P(X+\ell \leqslant k)P(Y=\ell)P(Z=k) \qquad \text{par indépendance}\\
&= \sum_{k=1}^{2n}\sum_{\ell=1}^{k-1} \frac{k-\ell}{2n}\frac{1}{2n}\frac{1}{2n}\\
&= \frac{1}{8n^3}\sum_{k=1}^{2n}\frac{k(k-1)}{2}\\
&= \frac{1}{16n^3}\Big(\sum_{k=1}^{2n}k^2-\sum_{k=1}^{2n}k\Big)\\
&= \frac{(2n-1)n(2n+1)}{24n^3}\\
&= \frac{4n^2-1}{24n^2}\cdot
\end{aligned}$$

♦ ***Exercice 4.*** [⋆]

Tic et Tac conservent leurs noisettes au creux d'un arbre (leur garde-manger). On y trouve autant de noisettes véreuses que de non-véreuses. Les premières sont indiscernables des autres.

1. Tic choisit n noisettes puis c'est au tour de Tac de faire de même. On note I (respectivement A) le nombre de noisettes véreuses choisies par Tic (respectivement par Tac).
   a) Soit N le nombre total de noisettes du garde-manger. Préciser la loi de I et, pour tout $k \in I(\Omega)$, donner la loi conditionnelle de A sachant $(I=k)$.
   On suppose dorénavant que N est très grand devant n. Donner, dans ce cas, les lois de I et A et préciser alors une propriété du couple (I, A).
   b) Trouver la probabilité que Tic et Tac aient le même nombre de noisettes véreuses.
2. Pour éviter les noisettes véreuses, Tic et Tac s'adressent à Donald qui sait, lui, reconnaître les bonnes noisettes des mauvaises. Tic demande à Donald de lui choisir n noisettes pour son goûter et Tac, plus gourmand, lui en demande le double. Donald, un peu facétieux, glisse une (et une seule) noisette véreuse dans chacun des deux goûters. Tic et Tac les mangent ensuite une par une, à la même vitesse. On note $\widehat{I}$ et $\widehat{A}$ les variables aléatoires égales aux rang d'apparition de la noisette véreuse respectivement pour Tic et Tac.
   a) Déterminer les lois de $\widehat{I}$ et $\widehat{A}$.
   b) Déterminer la probabilité que Tic découvre sa noisette véreuse avant Tac. Vers quelle limite tend cette probabilité lorsque n tend vers $+\infty$?

1. a) La variable I compte le nombre de succès (« choisir une noisette véreuse ») dans une suite de n choix sans remise d'une noisette parmi N avec, au début, une probabilité de succès égale à $1/2$. Donc

$$\boxed{I \hookrightarrow \mathscr{H}(N, n; 1/2).}$$

Soit $k \in [\![0; n]\!]$. Sachant que $I = k$, la variable A compte le nombre de succès (« choisir une noisette véreuse ») dans une suite de n choix sans remise d'une noisette parmi $N-n$ avec, au début, une probabilité de succès égale à $\dfrac{N/2-k}{N-n}$. Donc

$$\boxed{A \,|\, (I=k) \hookrightarrow \mathscr{H}\Big(N-n, n; \frac{N/2-k}{N-n}\Big)\cdot}$$

Lorsque N est très grand devant n, l'approximation de la loi hypergéométrique par la loi binomiale nous dit que tout se passe comme si les choix se faisaient avec remise. Dans ce cas, on obtient

$$\boxed{I \hookrightarrow \mathscr{B}(n;1/2) \qquad \text{et} \qquad A \hookrightarrow \mathscr{B}(n;1/2).}$$

De plus, les choix s'effectuant alors avec remise, on peut affirmer que

$$\boxed{I \text{ et } A \text{ sont indépendantes.}}$$

b) On cherche la probabilité de l'événement $(I = A)$. Pour cela, on écrit

$$\begin{aligned}
P(I=A) &= \sum_{k=0}^{n} P(I=k\ ,\ I=A) && \text{par filtration}\\
&= \sum_{k=0}^{n} P(I=k\ ,\ A=k)\\
&= \sum_{k=0}^{n} P(I=k)P(A=k) && \text{par indépendance}\\
&= \sum_{k=0}^{n} \binom{n}{k}^2 \left(\frac{1}{2}\right)^{2n}\\
&= \frac{1}{4^n}\sum_{k=0}^{n} \binom{n}{k}^2\\
&= \frac{1}{4^n}\binom{2n}{n} && \text{d'après la formule de Vandermonde,}
\end{aligned}$$

donc

$$\boxed{\text{la probabilité que Tic et Tac choisissent le même nombre de noisettes véreuses vaut } \frac{1}{4^n}\binom{2n}{n}\cdot}$$

2. a) Les rangs d'apparition de la noisette véreuse dans chacun des deux goûters étant tous équiprobables, on a

$$\boxed{\widehat{I} \hookrightarrow \mathscr{U}([\![1;n]\!]) \qquad \text{et} \qquad \widehat{A} \hookrightarrow \mathscr{U}([\![1;2n]\!]).}$$

b) On cherche la probabilité de l'événement $(\widehat{I} < \widehat{A})$. Pour cela, on écrit

$$\begin{aligned}
P(\widehat{I}<\widehat{A}) &= \sum_{k=1}^{n} P(\widehat{I}=k\,,\,\widehat{I}<\widehat{A}) && \text{par filtration}\\
&= \sum_{k=1}^{n} P(\widehat{I}=k\,,\,\widehat{A}>k)\\
&= \sum_{k=1}^{n} P(\widehat{I}=k)P(\widehat{A}>k) && \text{par indépendance}\\
&= \sum_{k=1}^{n} \frac{1}{n}\,\frac{2n-k}{2n}\\
&= \frac{1}{n}\sum_{k=1}^{n} 1 - \frac{1}{2n^2}\sum_{k=1}^{n} k\Big)\\
&= 1 - \frac{n+1}{4n},
\end{aligned}$$

donc

$$\boxed{\text{la probabilité que Tic découvre sa noisette véreuse avant Tac vaut } \frac{3n-1}{4n}\cdot}$$

On voit donc que

$$\boxed{\text{la probabilité que Tic découvre sa noisette véreuse avant Tac tend vers } \frac{3}{4}\cdot}$$

◆ ***Exercice 5.*** [$\circ$]

Guillaume Tell tire n fois sur une pomme posée sur la tête de son fils, avec la probabilité $p \in]0;1[$ d'atteindre la pomme à chaque tir. À l'issue de ces n tirs, son fils (s'il est toujours vivant) comptabilise le nombre de fois où la pomme a été atteinte mais, comme il est distrait, le petit Tell (si petit qu'on le disait mini) donne le bon résultat une fois sur deux et le bon résultat plus 1 le reste du temps. On note X ce résultat.

1. Déterminer l'espérance et la variance de X.
2. Déterminer la loi de X et retrouver ainsi son espérance.

1. Soit Y la variable aléatoire égale au nombre de tirs réussis de Guillaume Tell (sans l'erreur de son fils). La variable Y compte le nombre de succès (« toucher la pomme ») dans une suite de n tirs indépendants avec, à chaque fois, une probabilité de succès égale à p. Donc Y suit la loi $\mathscr{B}(n;p)$.
   Considérons Z la variable qui vaut 0 si le petit Tell donne le bon résultat et 1 sinon. La variable Z suit la loi $\mathscr{B}(1/2)$.
   Les variables Y et Z sont indépendantes puisque le petit Tell ajoute 1 (ou non) sans prendre en compte le résultat de son père.
   On a $X = Y + Z$.
   Donc
   $$E(X) = E(Y) + E(Z) = np + \frac{1}{2} \qquad \text{et} \qquad V(X) = V(Y) + V(Z) = npq + \frac{1}{4},$$
   c'est-à-dire
   $$\boxed{E(X) = np + \frac{1}{2} \qquad \text{et} \qquad V(X) = npq + \frac{1}{4}.}$$
2. On a
   $$\boxed{X(\Omega) = [\![0; n+1]\!].}$$
   Pour tout $k \in [\![0; n+1]\!]$, la formule des probas totales à travers le système complet d'événements $(Y = j)_{j \in [\![0;n]\!]}$, on a
   $$\begin{aligned} P(X = k) &= P(Y = k-1)P(X = k \mid Y = k-1) + P(Y = k)P(X = k \mid Y = k) \\ &= \binom{n}{k-1} p^{k-1} q^{n-k+1} \times \frac{1}{2} + \binom{n}{k} p^k q^{n-k} \times \frac{1}{2}, \end{aligned}$$
   donc
   $$\boxed{\forall k \in [\![0; n+1]\!], \qquad P(X = k) = \frac{1}{2}\binom{n}{k-1} p^{k-1} q^{n-k+1} + \frac{1}{2}\binom{n}{k} p^k q^{n-k}.}$$
   On a
   $$\begin{aligned} E(X) &= \sum_{k=0}^{n+1} kP(X = k) \\ &= \frac{1}{2}\sum_{k=0}^{n+1} k\binom{n}{k-1} p^{k-1} q^{n-k+1} + \frac{1}{2}\sum_{k=0}^{n+1} k\binom{n}{k} p^k q^{n-k} \\ &= \frac{1}{2}\sum_{k=1}^{n+1} k\binom{n}{k-1} p^{k-1} q^{n-k+1} + \frac{1}{2}\sum_{k=0}^{n} k\binom{n}{k} p^k q^{n-k} \\ &= \frac{1}{2}\sum_{\ell=0}^{n} (\ell+1)\binom{n}{\ell} p^{\ell} q^{n-\ell} + \frac{1}{2}\sum_{k=0}^{n} k\binom{n}{k} p^k q^{n-k} \\ &= \frac{1}{2}\sum_{\ell=0}^{n} \ell\binom{n}{\ell} p^{\ell} q^{n-\ell} + \frac{1}{2}\sum_{\ell=0}^{n} \binom{n}{\ell} p^{\ell} q^{n-\ell} + \frac{1}{2}\sum_{k=0}^{n} k\binom{n}{k} p^k q^{n-k} \\ &= \underbrace{\sum_{\ell=0}^{n} \ell\binom{n}{\ell} p^{\ell} q^{n-\ell}}_{\text{espérance de } \mathscr{B}(n;p)} + \frac{1}{2}\underbrace{\sum_{\ell=0}^{n} \binom{n}{\ell} p^{\ell} q^{n-\ell}}_{=(p+q)^n=1} \\ &= np + \frac{1}{2}, \end{aligned}$$
   donc on retrouve que
   $$\boxed{E(X) = np + \frac{1}{2}.}$$

♦ ***Exercice 6.*** [$\star$]

On note X le nombre d'enfants d'une famille, F le nombre de filles et G le nombre de garçons de sorte que $X = F + G$. On suppose qu'il y a autant de filles que de garçons à la naissance et que les naissances sont indépendantes (du point de vue du sexe des enfants).

1. On suppose que X est constante. À l'aide de $V(F+G)$, calculer $\operatorname{Cov}(F, G)$. Interpréter.
2. On suppose que X suit la loi binomiale de paramètres $n \in \mathbb{N}$ et $p \in]0; 1[$. Déterminer les lois de F et G. Calculer $\operatorname{Cov}(F, G)$ et donner son signe. *Donnée :* $\binom{n}{i}\binom{i}{k} = \binom{n}{k}\binom{n-k}{i-k}$.

1. Supposons que $X = n$ où $n \in \mathbb{N}$.
   On a
   $$V(F+G) = V(F) + V(G) + 2\operatorname{Cov}(F, G).$$
   Comme $F + G = X = n$, on a
   $$V(F+G) = 0.$$
   La variable F compte le nombre de succès (avoir une fille) dans une série de n naissances indépendantes avec une probabilité de succès $1/2$. Donc $F \hookrightarrow \mathscr{B}(n, 1/2)$. De même, on a $G \hookrightarrow \mathscr{B}(n, 1/2)$. Il s'ensuit que
   $$V(F) = V(G) = \frac{n}{4}\cdot$$
   On obtient alors que
   $$0 = \frac{n}{4} + \frac{n}{4} + 2\operatorname{Cov}(F, G),$$
   c'est-à-dire
   $$\boxed{\operatorname{Cov}(F, G) = -\frac{n}{4}\cdot}$$
   On constate que $\operatorname{Cov}(F, G) < 0$, ce qui traduit le fait que, dans une famille où le nombre d'enfants est fixé, plus il y a de filles, moins il y a de garçons et vice-versa.

2. On pose $q = 1 - p$. On a $F(\Omega) = [\![0; n]\!]$. Soit $k \in \mathbb{N}$. La formule des probabilités totales appliquée à l'événement $(F = k)$ à travers le système complet d'événements $(X = i)_{i \geqslant 0}$ donne
   $$\begin{aligned} P(F=k) &= \sum_{i=0}^{n} P(X=i)P(F=k \mid X=i) \\ &= \sum_{i=k}^{n} \binom{n}{i} p^i q^{n-i} \binom{i}{k} \frac{1}{2^i} && \text{car } F \mid (X=i) \hookrightarrow \mathscr{B}\Big(i, \frac{1}{2}\Big) \\ &= \binom{n}{k} \sum_{i=k}^{n} \binom{n-k}{i-k} \Big(\frac{p}{2}\Big)^i q^{n-i} && \text{par la formule de l'énoncé} \\ &= \binom{n}{k}\Big(\frac{p}{2}\Big)^k \sum_{j=0}^{n-k} \binom{n-k}{j} \Big(\frac{p}{2}\Big)^i q^{n-k-j} && \text{en posant } j = i-k \\ &= \binom{n}{k}\Big(\frac{p}{2}\Big)^k \Big(\frac{p}{2} + q\Big)^{n-k} && \text{binôme.} \end{aligned}$$
   On a donc
   $$\boxed{F \hookrightarrow \mathscr{B}\Big(n; \frac{p}{2}\Big)\cdot}$$
   De même, on a
   $$\boxed{G \hookrightarrow \mathscr{B}\Big(n; \frac{p}{2}\Big)\cdot}$$
   On a
   $$V(F+G) = V(X) = npq, \qquad \text{et} \qquad V(F) = V(G) = n\frac{p}{2}\Big(1 - \frac{p}{2}\Big),$$
   donc la formule $V(F+G) = V(F) + V(G) + 2\operatorname{Cov}(F, G)$ donne
   $$2\operatorname{Cov}(F, G) = npq - 2n\frac{p}{2}\Big(1 - \frac{p}{2}\Big),$$
   c'est-à-dire
   $$\boxed{\operatorname{Cov}(F, G) = -\frac{np^2}{4} < 0.}$$

♦ ***Exercice 7.*** [$\star$] (Les lapins de Chantal)

Un groupe de ℓ lapins surgit devant un groupe de c chasseurs. Chaque lapin, armé d'un fusil à carottes, tire sur un chasseur au hasard et le touche avec la probabilité $p \in]0;1[$. Calculer l'espérance du nombre C de chasseurs carottés.

Soit $i \in [\![1;c]\!]$. On note X_i l'indicatrice de l'événément C_i : « le i-ème chasseur est touché » de sorte que $X_i \hookrightarrow \mathscr{B}\big(P(C_i)\big)$. On pose $q = 1-p$ et on note V_i la variable égale au nombre de lapins qui ont visés le i-ème chasseur. D'après la formule des probabilités totales à travers le s.c.e. $(V_i = j)_{j\in[\![0;\ell]\!]}$, on a

$$P(C_i) = \sum_{j=0}^{\ell} P(V_i = j)P(C_i \,|\, V_i = j).$$

La variable V_i compte le nombre de succès (le lapin vise le i-ème chasseur) dans une série de ℓ lapins indépendants avec une proba de succès $1/c$, donc

$$V_i \hookrightarrow \mathscr{B}(\ell, 1/c).$$

Par ailleurs, on a

$$P(C_i \,|\, V_i = j) = 1 - P(\overline{C_i} \,|\, V_i = j) = 1 - q^j.$$

Donc

$$\begin{aligned}
P(C_i) &= \sum_{j=0}^{\ell} \binom{\ell}{j}\Big(\frac{1}{c}\Big)^j\Big(1-\frac{1}{c}\Big)^{\ell-j}(1-q^j) \\
&= \sum_{j=0}^{\ell} \binom{\ell}{j}\Big(\frac{1}{c}\Big)^j\Big(1-\frac{1}{c}\Big)^{\ell-j} - \sum_{j=0}^{\ell} \binom{\ell}{j}\Big(\frac{q}{c}\Big)^j\Big(1-\frac{1}{c}\Big)^{\ell-j} \\
&= 1 - \Big(\frac{q}{c} + 1 - \frac{1}{c}\Big)^{\ell} \\
&= 1 - \Big(\frac{c-p}{c}\Big)^{\ell}.
\end{aligned}$$

Ainsi,

$$X_i \hookrightarrow \mathscr{B}\Big(1 - \Big(\frac{c-p}{c}\Big)^{\ell}\Big)$$

Dès lors, on a

$$E(C) = E(X_1 + \cdots + X_c) = E(X_1) + \cdots + E(X_c)$$

ce qui donne

$$\boxed{E(C) = c\Big(1 - \Big(\frac{c-p}{c}\Big)^{\ell}\Big).}$$

♦ ***Exercice 8.*** [$\star$] (Toute hypergéométrique a la même loi qu'une somme de Bernoullis)

On considère une urne contenant N boules avec une proportion p de boules blanches et une proportion $q = 1-p$ de boules noires. On effectue dans cette urne une succession de n tirages sans remise (avec $n \in \mathbb{N}^*$). Pour tout $k \in [\![1;n]\!]$, on note X_k la variable aléatoire qui vaut 1 lorsque le k-ème tirage donne une boule blanche et 0 sinon. Enfin, pour tout $k \in [\![1;n]\!]$, on pose $S_k = X_1 + \cdots + X_k$, de sorte que S_n suit la loi $\mathscr{H}(N;n,p)$.

On veut démontrer que, pour tout $k \in [\![1;n]\!]$, X_k suit la loi $\mathscr{B}(p)$ et retrouver l'espérance et la variance de la loi $\mathscr{H}(N;n,p)$. On suppose donc que ces valeurs ne sont pas connues.

1. On démontre, de deux manières, que les X_k sont identiquement distribuées sur la loi $\mathscr{B}(p)$.
   a) Démontrer que $\forall k \in [\![1;n-1]\!]\,,\ P(X_{k+1} = 1) = \big(Np - E(S_k)\big)/(N-k)$ et conclure.
   b) À l'aide d'un dénombrement, démontrer que $\forall k \in [\![1;n]\!]\,,\ P(X_k = 1) = p$ et conclure.
2. Démontrer que $\forall i, j \in [\![1;n]\!]\,,\ (i \neq j) \Rightarrow \big(P(X_i = 1, X_j = 1) = p(Np-1)/(N-1)\big)$.
3. Retrouver l'espérance et la variance de la loi $\mathscr{H}(N;n,p)$.

Tout d'abord, on peut noter que S_n compte le nombre de succès (« obtenir une boule blanche ») dans une succession de n tirages sans remise parmi N boules, avec au départ une probabilité de succès p. On reconnaît donc un schéma hypergéométrique, ce qui justifie l'affirmation de l'énoncé :

$$\boxed{S_n \hookrightarrow \mathscr{H}(N,n;p).}$$

1. a) La formule des probabilités totales à travers le système complet d'événements $(S_k = \ell)_{\ell \in [\![0;k]\!]}$ donne

$$P(X_{k+1} = 1) = \sum_{\ell=0}^{k} P(S_k = \ell) P(X_{k+1} = 1 \mid S_k = \ell).$$

Or, pour tout $\ell \in [\![0;k]\!]$, lorsque $(S_k = \ell)$ s'est produit, on sait qu'il reste dans l'urne $N-k$ boules dont $Np - \ell$ sont blanches, donc

$$P(X_{k+1} = 1 \mid S_k = \ell) = \frac{Np - \ell}{N-k}\cdot$$

Il s'ensuit que

$$\begin{aligned} P(X_{k+1} = 1) &= \sum_{\ell=0}^{k} P(S_k = \ell) \frac{Np - k}{N-k} \\ &= \frac{Np}{N-k} \sum_{\ell=0}^{k} P(S_k = \ell) - \frac{1}{N-k} \sum_{\ell=0}^{k} k P(S_k = \ell) \\ &= \frac{Np}{N-k} \times 1 - \frac{1}{N-k} \times E(S_k), \end{aligned}$$

donc

$$\boxed{P(X_{k+1} = 1) = \frac{Np - E(S_k)}{N-k}\cdot}$$

On procède par réccurence forte sur l'assertion : $\mathscr{Q}(k)$: $X_k \hookrightarrow \mathscr{B}(p)$.

Initialisation : Il est clair que $X_1 \hookrightarrow \mathscr{B}(p)$ donc $\mathscr{Q}(1)$ est vraie.

Hérédité : Fixons $k \in [\![1;n-1]\!]$ telle que $\mathscr{Q}(1), \ldots, Q(k)$ sont vraies et démontrons $\mathscr{Q}(k+1)$. On a

$$\begin{aligned} P(X_{k+1} = 1) &= \frac{Np - E(S_k)}{N-k} \\ &= \frac{Np - kp}{N-k} \qquad \text{car } E(S_k) = \sum_{\ell=1}^{k} E(X_\ell) = \sum_{\ell=1}^{k} p = kp \\ &= p, \end{aligned}$$

donc, puisque $X_{k+1}(\Omega) = \{0;1\}$, on a bien $X_{k+1} \hookrightarrow \mathscr{B}(p)$. Cela démontre $\mathscr{Q}(k+1)$.

Conclusion : Le principe de récurrence nous dit que $\mathscr{Q}(k)$ est vraie pour tout $k \in [\![1;n]\!]$, c'est-à-dire que les X_i sont identiquement distribuées sur la loi $\mathscr{B}(p)$.

Remarque : Cette façon de démontrer que les X_i suivent la loi $\mathscr{B}(p)$ n'est pas la plus simple. Cependant, elle peut être généralisée à d'autres types de tirages que les tirages sans remise. On pourra, dans ce sens, se renseigner sur le modèle des urnes de Pólya.

b) On numérote les boules. Les boules numérotées de 1 à Np sont blanches, les autres (numérotées de $Np+1$ à N) sont noires.

On travaille sur l'univers des n-uplets sans répétition d'éléments de $[\![1;N]\!]$. Le cardinal de cet univers est $N(N-1)\cdots(N-n+1)$.

Soit $k \in [\![1;n]\!]$. Pour dénombrer l'événement $(X_k = 1)$, on choisit une des Np boules blanches pour la placer en k-ème position dans notre n-uplet : Np choix, puis on choisit successivement $n-1$ entiers parmi les $N-1$ restants : $(N-1)(N-2)\cdots(N-n+1)$ choix. Cela donne $\operatorname{card}(X_k = 1) = Np(N-1)(N-2)\cdots(N-n+1)$.

La formule de loi uniforme donne

$$P(X_k = 1) = \frac{Np(N-1)(N-2)\cdots(N-n+1)}{N(N-1)\cdots(N-n+1)} = p.$$

Comme $X_k(\Omega) = \{0;1\}$, on en déduit $X_k \hookrightarrow \mathscr{B}(p)$.

En conclusion,

$$\boxed{\text{les variables } X_i \text{ sont identiquement distribuées sur la loi } \mathscr{B}(p).}$$

On a ainsi démontré (de deux manières) que

$$\boxed{\text{pour toute variable } Z \text{ suivant } \mathscr{H}(N;n;p)\text{, il existe des variables } X_1,\dots,X_n \text{ identiquement distribuées sur la loi } \mathscr{B}(p) \text{ (non indépendantes) telles que } Z \text{ a la même loi que } X_1+\cdots+X_n.}$$

Remarque : On notera que l'on a pas démontré que Z est la somme de n bernoullis mais seulement que Z a la même loi que S_n qui, elle, est la somme de n bernoullis.

2. On reprend le raisonnement de dénombrement de la question précédente. On numérote les boules : les boules numérotées de 1 à Np sont blanches, les autres (numérotées de $Np+1$ à N) sont noires.
   a) On travaille sur l'univers des n-uplets sans répétition d'éléments de $[\![1;N]\!]$. Le cardinal de cet univers est $N(N-1)\cdots(N-n+1)$.

   Soient $i,j\in[\![1;n]\!]$ avec $i\neq j$. Pour dénombrer l'événement $(X_i=1,X_j=1)$, on choisit successivement deux des Np boules blanches pour les placer en i-ème et j-ème position dans notre n-uplet : $Np(Np-1)$ choix, puis on choisit successivement $n-2$ entiers parmi les $N-2$ restants : $(N-2)(N-3)\cdots(N-n+1)$ choix. Cela donne $\mathsf{card}(X_i=1,X_j=1)=Np(Np-1)(N-2)\cdots(N-n+1)$.

   La formule de loi uniforme donne

$$P(X_i=1,X_j=1)=\frac{Np(Np-1)(N-2)\cdots(N-n+1)}{N(N-1)\cdots(N-n+1)}=\frac{p(Np-1)}{N-1}.$$

   En conclusion,

$$\boxed{\forall i,j\in[\![1;n]\!]\,,\qquad (i\neq j)\Rightarrow\Big(P(X_i=1,X_j=1)=\frac{p(Np-1)}{N-1}\Big).}$$

3. La linéarité de l'espérance nous dit immédiatement que

$$E(S_n)=E(X_1)+\cdots+E(X_n)=np.$$

Par ailleurs, la formule sur la variance d'une somme nous dit que

$$V(S_n)=V(X_1+\cdots+X_n)=V(X_1)+\cdots+V(X_n)+2\sum_{1\leqslant i<j\leqslant n}\mathrm{Cov}(X_i,X_j).$$

Or

$$E(X_iX_j)=1\times P(X_i=1,X_j=1)+0=\frac{p(Np-1)}{N-1},$$

donc

$$\mathrm{Cov}(X_i,X_j)=E(X_iX_j)-E(X_i)E(X_j)=\frac{p(Np-1)}{N-1}-p^2.$$

Cela donne

$$V(S_n)=np(1-p)+\frac{n(n-1)}{2}\Big(\frac{p(Np-1)}{N-1}-p^2\Big)=np(1-p)\frac{N-n}{N-1}.$$

En conclusion,

$$\boxed{\text{l'espérance de la loi } \mathscr{H}(N;n;p) \text{ vaut } np \text{ et sa variance est } np(1-p)\frac{N-n}{N-1}.}$$

♦ ***Exercice 9.*** [⋆]
Soit $n\geqslant 2$. Soient X et Y deux variables aléatoires indépendantes suivant toutes les deux la loi uniforme sur $[\![1;n]\!]$. On pose $U=X-Y$ et $V=\min(X,Y)$. Déterminer les lois de U et V. Sont-ce des variables indépendantes ?

On a

$$\boxed{U(\Omega)=[\![1-n;n-1]\!].}$$

Pour tout $k\in[\![1-n;0]\!]$, on a

$$P(U=k)=\sum_{j=1-k}^{n}P(X=k+j\,,\,Y=j)\underset{\text{par }\perp}{=}\sum_{j=1-k}^{n}\frac{1}{n^2}=\frac{n+k}{n}$$

et, pour tout $k \in [\![1;n-1]\!]$, on a

$$P(U=k) = \sum_{j=1}^{n-k} P(X=k+j\,,\,Y=j) \underset{\text{par } \perp}{=} \sum_{j=1}^{n-k} \frac{1}{n^2} = \frac{n-k}{n^2}$$

donc

$$\boxed{\forall k \in [\![1-n;n-1]\!]\,, \qquad P(U=k) = \frac{n-|k|}{n^2}.}$$

On a

$$\boxed{V(\Omega) = [\![1;n]\!].}$$

Pour tout $k \in [\![1;n]\!]$, on a

$$\begin{aligned} P(V \leqslant k) &= P\big(\max(X,Y) \leqslant k\big) \\ &= P(X \leqslant k\,,\,Y \leqslant k) \\ &= P(X \leqslant k)P(Y \leqslant k) \qquad \text{par } \perp\!\!\!\perp \\ &= \frac{k^2}{n}, \end{aligned}$$

donc

$$\begin{aligned} P(V=k) &= P(V \leqslant k) - P(V \leqslant k-1) \\ &= \frac{k^2}{n} - \frac{(k-1)^2}{n} \\ &= \frac{2k-1}{n^2}. \end{aligned}$$

Donc

$$\boxed{\forall k \in [\![1;n]\!]\,, \qquad P(V=k) = \frac{2k-1}{n^2}.}$$

On a

$$P(U=0\,,\,V=1) = P(X=Y=1) = \frac{1}{n^2}$$

alors que

$$P(U=0)P(V=1) = \frac{1}{n}\frac{2n-1}{n} = \frac{2n-1}{n^2},$$

donc

$$\boxed{U \text{ et } V \text{ ne sont pas indépendantes.}}$$

♦ ***Exercice 10.*** [⋆]

On lance deux dés et on note S la somme des points obtenus. On voudrait savoir s'il est possible de truquer les dés pour que S suive une loi uniforme.

1. Soit X une variable aléatoire telle que $X(\Omega) = [\![m;n]\!]$ où $m < n$. Démontrer que l'application $\varphi_X : t \longmapsto E(t^X)/t^m$ est une fonction polynomiale dont on majorera le degré.
2. On suppose que S suit une loi uniforme. Déterminer φ_S et préciser les racines de ce polynôme (dans $\mathbb{C}$). Justifier que l'on ne peut pas trouver deux variables A et B telles que $A(\Omega) = B(\Omega) = [\![1;6]\!]$ et $\varphi_S = \varphi_A\varphi_B$. Conclure.

1. On a, pour tout $t \in \mathbb{R}$,

$$\varphi_X(t) = \frac{1}{t^m}E(t^X) = \sum_{k=m}^{n} p_k t^{k-m} \qquad \text{où} \qquad p_k = P(X=k),$$

donc

$$\boxed{\varphi_X \text{ est bien une fonction polynomiale de degré au plus } n-m.}$$

2. On a $S(\Omega) = [\![2;12]\!]$, donc, pour tout $t \in \mathbb{R}$, on a

$$\varphi_S(t) = \frac{1}{t^2} \sum_{k=2}^{12} \frac{1}{11} t^k,$$

c'est-à-dire

$$\boxed{\forall t \in \mathbb{R}, \qquad \varphi_S(t) = \frac{1 + t + \cdots + t^{10}}{11}.}$$

D'après la règle de sommation d'une somme géométrique de raison différente de 1, on a

$$\forall t \in \mathbb{R} \setminus \{1\}, \qquad \varphi_S(t) = \frac{t^{11} - 1}{11(t-1)},$$

ce qui prouve que

$$\boxed{\text{les racines de } \varphi_S \text{ sont les racines 11-èmes de l'unité distinctes de 1.}}$$

Supposons qu'il existe deux variables aléatoires A et B telles que $A(\Omega) = B(\Omega) = [\![1;6]\!]$ et $\varphi_S = \varphi_A \varphi_B$. Alors, d'après la question 1, φ_S est le produit de deux polynômes réels de au plus degré 5 et comme φ_S est précisément de degré 10, on peut affirmer que φ_A et φ_B sont exactement de degré 5. Or tout polynôme réel de degré impair possède au moins une racine réelle (pour le voir, on applique le TVI en constatant que les limites en $-\infty$ et $+\infty$ sont infinies et de signes contraires) donc φ_S possède des racines réelles. Comme les éléments de $\mathbb{U}_{11} \setminus \{1\}$ ne sont pas réels, on aboutit à une contradiction. Donc

$$\boxed{\begin{array}{l}\text{on ne peut pas trouver deux variables aléatoires } A \text{ et } B \\ \text{telles que } A(\Omega) = B(\Omega) = [\![1;6]\!] \text{ et } \varphi_S = \varphi_A \varphi_B.\end{array}}$$

Notons A et B les variables aléatoires égales aux résultats respectifs du premier et du second dé. Par indépendance, on a

$$\forall t \in \mathbb{R}, \qquad \varphi_S(t) = \frac{1}{t^2} E(t^S) = \frac{1}{t^2} E(t^A t^B) = \frac{1}{t^2} E(t^A) E(t^B) = \varphi_A(t) \varphi_B(t),$$

et l'on vient de voir qu'une telle relation est impossible à mettre en place. Par suite,

$$\boxed{\text{on ne peut pas truquer deux dés indépendants pour que leur somme suive une loi uniforme.}}$$

♦ ***Exercice 11.*** [⋆]

Soit $n \geqslant 2$. On répartit n jetons numérotés de 1 à n dans n boîtes numérotées elles aussi numérotées de 1 à n de telle sorte que chaque boîte contienne un seul jeton.

1. a) On fixe $k \in [\![1;n]\!]$ et l'on considère $1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n$. Déterminer la probabilité p_k de l'événement « pour tout $\ell \in [\![1;k]\!]$, le jeton i_ℓ est dans la boîte i_ℓ » ?
   b) Déterminer la probabilité π_n de l'événement « aucun jeton n'est dans la boîte portant son numéro ». Quelle est la limite de π_n lorsque n tend vers $+\infty$?
2. Pour tout $i \in [\![1;n]\!]$, on note X_i la variable aléatoire qui vaut 1 si le jeton i est dans la boîte i et 0 sinon. On pose $X = X_1 + \cdots + X_n$ le nombre de jetons qui sont dans une boîte portant leur numéro.
   a) Pour $i \in [\![1;n]\!]$, déterminer la loi de X_i. En déduire l'espérance de X.
   b) Pour $i, j \in [\![1;n]\!]$ distincts, calculer $P(X_i = X_j = 1)$. En déduire la variance de X.

1. Considérons l'univers Ω constitué des permutations de $[\![1;n]\!]$ muni de la loi uniforme. On a donc $\operatorname{card} \Omega = n!$
   a) L'événement « pour tout $\ell \in [\![1;k]\!]$, le jeton i_ℓ est dans la boîte i_ℓ » est le sous-ensemble de Ω constitué des permutations qui se contentent de permuter les éléments de $[\![1;n]\!] \setminus \{i_1, i_2, \ldots, i_k\}$. Le cardinal de cet événement est donc égal à $(n-k)!$ On en déduit que

$$\boxed{p_k = \frac{(n-k)!}{n!}.}$$

b) Notons A_i l'événement « le jeton i est dans la booîte i ». On veut calculer $\pi_n = P(\overline{A_1} \cap \cdots \cap \overline{A_n}) = P(\overline{A_1 \cup \cdots \cup A_n})$. On a

$$\begin{array}{rcll} \pi_n & = & 1 - P(A_1 \cup \cdots \cup A_n) & \\ & = & 1 - \displaystyle\sum_{k=1}^{n} (-1)^{k-1} \sum_{1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n} P(A_{i_1} \cap \cdots \cap A_{i_k}) & \text{(crible)} \\ & = & 1 + \displaystyle\sum_{k=1}^{n} (-1)^k \sum_{1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n} \frac{(n-k)!}{n!} & \text{d'après b)} \\ & = & 1 + \displaystyle\sum_{k=1}^{n} (-1)^k \binom{n}{k} \frac{(n-k)!}{n!} & \text{car il y a } \binom{n}{k} \text{ termes dans la somme } \sum_{1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n} \\ & = & 1 + \displaystyle\sum_{k=1}^{n} \frac{(-1)^k}{k!}, & \end{array}$$

donc

$$\boxed{\pi_n = \sum_{k=0}^{n} \frac{(-1)^k}{k!}\cdot}$$

Comme $\sum_{k \geqslant 0} (-1)^k / k!$ est une série exponentielle convergente de somme e^{-1}, on en déduit que

$$\boxed{\lim \pi_n = \frac{1}{\mathrm{e}} \approx 0,37.}$$

2. a) On a $X_i(\Omega) = \{0; 1\}$ donc X_i suit une loi de Bernoulli (comme toutes les fonctions indicatrices d'ailleurs). Or, avec les notations de la question 1, on a $P(X_i = 1) = p_1 = 1/n$. Donc

$$\boxed{X_i \hookrightarrow \mathscr{B}\Big(\frac{1}{n}\Big)\cdot}$$

Alors $E(X) = E(X_1 + \cdots + X_n) = E(X_1) + \cdots + E(X_n) = n \times 1/n$, donc

$$\boxed{E(X) = 1.}$$

Remarque : Une permutation de n objets possède donc en moyenne un point fixe, ceci quel que soit n !

b) Toujours avec les notations de la question 1, on a

$$P(X_i = 1, X_j = 1) = p_2 = \frac{1}{n(n-1)}\cdot$$

On en déduit que

$$\begin{array}{rcl} E(X_i X_j) & = & 1 \times 1 \times P(X_i = 1, X_j = 1) + 1 \times 0 \times P(X_i = 1, X_j = 0) \\ & & +0 \times 1 \times P(X_i = 0, X_j = 1) + 0 \times 0 \times P(X_i = 0, X_j = 0) \\ & = & \dfrac{1}{n(n-1)}, \end{array}$$

donc

$$\mathrm{Cov}(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j) = \frac{1}{n(n-1)} - \frac{1}{n}\frac{1}{n} = \frac{1}{n^2(n-1)}\cdot$$

Par conséquent,

$$\begin{array}{rcl} V(X) & = & V(X_1 + \cdots + X_n) \\ & = & \displaystyle\sum_{i=1}^{n} V(X_i) + 2 \sum_{1 \leqslant i < j \leqslant n} \mathrm{Cov}(X_i, X_j) \\ & = & \displaystyle\sum_{i=1}^{n} \frac{1}{n}\Big(1 - \frac{1}{n}\Big) + 2 \sum_{1 \leqslant i < j \leqslant n} \frac{1}{n^2(n-1)} \\ & = & 1 - \dfrac{1}{n} + 2\dbinom{n}{2} \dfrac{1}{n^2(n-1)} \\ & = & 1 - \dfrac{1}{n} + \dfrac{1}{n}, \end{array}$$

donc

$$\boxed{V(X) = 1.}$$