

Exercice 1 *Bilinéarité de la covariance et indépendance*

On considère un processus bernoullien de probabilité de succès $p \in]0; 1[$ et on pose $q = 1 - p$. Pour chaque n de \mathbb{N}^* , on note S_n et E_n le nombre de succès et d'échecs respectivement au cours des n premières épreuves. On note enfin $\Delta_n = S_n - E_n$.

1. Trouver, sans calculs excessifs, la loi de Δ_n , son espérance et sa variance.
2. Lorsque $\text{Cov}(X, Y)$ existe d'une part, et $\sigma(X)$ et $\sigma(Y)$ existent et sont non nuls, on appelle coefficient de corrélation linéaire de X et Y le nombre

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma(X)\sigma(Y)}.$$

- a) On suppose $m = n$. Que vaut le coefficient de corrélation linéaire $\rho(\Delta_m, \Delta_n)$?
- b) On suppose $m < n$. Justifier l'indépendance de $S_n - S_m$ et S_m .
- c) Calculer, pour tous m et n de \mathbb{N}^* , le coefficient de corrélation linéaire de Δ_m et Δ_n .

Solution (Ex.1 – Bilinéarité de la covariance et indépendance)

1. Comme $S_n + E_n = n$, $\Delta_n = S_n - n + S_n = 2S_n - n$.
Donc $S_n(\Omega) = \llbracket 0; n \rrbracket$ donc $\Delta_n(\Omega) = \{2k - n; k \in \llbracket 0; n \rrbracket\} = \{-n, -n + 2, \dots, n - 2, n\}$
 $\mathbb{E}(\Delta_n) = 2\mathbb{E}(S_n) - n = 2np - n = n(2p - 1) = n(p - q)$, $\mathbb{V}(\Delta_n) = 4\mathbb{V}(S_n) = 4npq$.
2. Rappelons que la covariance est bilinéaire :
 $\text{Cov}(aX + b, cY + d) = ac\text{Cov}(X, Y)$ et $\text{Cov}(X, Y + Z) = \text{Cov}(X, Y) + \text{Cov}(X, Z)$
 - Si $m = n$ alors $\Delta_m = 1 \times \Delta_n + 0$ et $\text{Cov}(\Delta_m, \Delta_n) = 1$. Sinon, on peut revenir à $\text{Cov}(\Delta_m, \Delta_n) = \mathbb{V}(\Delta_n) = 4npq$.
 - Supposons $m < n$.
 $\text{Cov}(\Delta_m, \Delta_n) = \text{Cov}(2S_m - m, 2S_n - n) = 4\text{Cov}(S_m, S_n)$.
Observons que $S_n = S_m + (S_n - S_m)$ avec S_m et $S_n - S_m$ indépendantes car S_m compte le nombre de succès au cours des m premières expériences et $S_n - S_m$ le nombre de succès au cours des $n - m$ suivantes.
 $\text{Cov}(\Delta_m, \Delta_n) = 4\text{Cov}(S_m, S_m + (S_n - S_m)) = 4\text{Cov}(S_m, S_m) + 4\text{Cov}(S_m, S_n - S_m) = 4\mathbb{V}(S_m) + 4 \times 0 = 4mpq$
 - Par symétrie du problème, si $m > n$, alors $\text{Cov}(\Delta_m, \Delta_n) = 4npq$

Bilan : $\text{Cov}(\Delta_m, \Delta_n) = 4\min(m, n)pq$ et du coup $\rho(\Delta_m, \Delta_n) = \frac{\min(m, n)}{\sqrt{mn}} = \min\left(\sqrt{\frac{m}{n}}, \sqrt{\frac{n}{m}}\right)$

Exercice 2 *Marche aléatoire sur \mathbb{Z}*

Soit $p \in]0; 1[$ et $q = 1 - p$. On considère une suite de variables $(X_i)_{i \in \mathbb{N}^*}$ indépendantes toutes de lois définies par

$$\forall i \in \mathbb{N}^*, X_i(\Omega) = \{-1, 1\} \quad \text{et} \quad \mathbb{P}(X_i = 1) = p, \mathbb{P}(X_i = -1) = q.$$

On pose

$$S_0 = 0 \quad \text{et} \quad \forall n \in \mathbb{N}^*, S_n = X_1 + \dots + X_n.$$

On peut se représenter la situation par un mobile se déplaçant sur \mathbb{Z} , considéré comme axe gradué, situé en 0 à l'instant $i = 0$, et se déplaçant à chaque instant $i \in \mathbb{N}^*$ de $X_i = \pm 1$ unité. S_n est alors l'abscisse du mobile à l'issue du n -ième déplacement, $S_0 = 0$ caractérisant l'abscisse nulle au début de l'expérience.

On s'intéresse au premier retour à l'origine du mobile.

Bien que les X_i ne suivent pas exactement une loi de Bernoulli $\mathcal{B}(p)$, elles s'en rapprochent et cette observation permet de trouver sans fatigue la loi des S_n .

1. Donner pour tout $i \in \mathbb{N}^*$, la loi de $Y_i \stackrel{\text{déf.}}{=} \frac{X_i + 1}{2}$.
2. En s'appuyant sur les Y_i , expliciter la loi de S_n .
3. Expliciter en particulier $\mathbb{P}(S_n = 0)$, probabilité d'un retour en 0 à l'issue du n -ième déplacement.

Solution (Ex.2 – Marche aléatoire sur \mathbb{Z})

1. $X_i = -1 \iff Y_i = 0$, et $X_i = 1 \iff Y_i = 1$, tout est dit... : Y_i suit $\mathcal{B}(p)$.
2. Les (Y_i) sont indépendantes et toutes de lois $\mathcal{B}(p)$, donc par stabilité de la loi binomiale, $\Sigma_n \stackrel{\text{déf.}}{=} \sum_{i=1}^n X_i \iff \mathcal{B}(n, p)$.

Donc $\Sigma_n(\Omega) = \llbracket 0; n \rrbracket$ et, $\forall k \in \llbracket 0; n \rrbracket$, $\mathbb{P}(\Sigma_n = k) = \binom{n}{k} p^k q^{n-k}$.

Or : $S_n = \sum_{i=1}^n X_i = \sum_{i=1}^n (2Y_i - 1) = 2 \sum_{i=1}^n Y_i - n = 2\Sigma_n - n$.

Donc : $S_n(\Omega) = \{2k - n | k \in \Sigma_n(\Omega)\} = \{2k - n | k \in \llbracket 1; n \rrbracket\}$

Et : $\forall k \in \llbracket 1; n \rrbracket$, $\mathbb{P}(S_n = 2k - n) = \mathbb{P}(\Sigma_n = k) = \binom{n}{k} p^k q^{n-k}$.

3. $2k - n = 0 \iff k = n/2$, impossible si n est impair, et sinon on applique la formule précédente :

$$\mathbb{P}(S_n = 0) = \begin{cases} 0 & \text{si } n \text{ impair,} \\ \binom{n}{n/2} (pq)^{n/2} & \text{si } n \text{ pair.} \end{cases}$$

Exercice 3 Passerelle

Quels liens entre les deux exercices précédents ?

Solution (Ex.3 – Passerelle) Si, dans le premier exercice on décide qu'un succès est un déplacement d'une unité à droite et un échec une unité à gauche, Δ_n donne l'abscisse. Ainsi $\Delta_n = S_n$ de l'exercice 2...

Exercice 4 Somme aléatoire

Une urne contient six jetons indiscernables au toucher et numérotés de 1 à 6. On lance un dé équilibré et l'on pioche dans l'urne un nombre de jetons égal à la valeur du dé. L'objectif est de déterminer l'espérance et la variance de la variable S donnant la somme totale des numéros des jetons piochés.

Pour i dans $\llbracket 1; 6 \rrbracket$, on note X_i la variable aléatoire indicatrice de l'événement « le jeton numéro i a été pioché ». On note aussi D la valeur donnée par le dé.

1. En exploitant l'espérance de D, déterminer la loi des X_i .
2. Dédire de la variance de D la valeur de la covariance de X_i et X_j pour tous $i \neq j$.
3. Déterminer finalement l'espérance et la variance de S.

Solution (Ex.4 – Somme aléatoire)

1. Par symétrie, les X_i suivent toutes la même loi de type Bernoulli. Notons p le paramètre commun de ces lois.

D suit la loi uniforme de paramètre $\frac{1}{6}$ donc $\mathbb{E}(D) = \frac{7}{2}$.

Or $D = X_1 + \dots + X_6$ donc par linéarité $\mathbb{E}(D) = \sum_{i=1}^6 \mathbb{E}(X_i) = 6p$, d'où

$$\forall i \in \llbracket 1; 6 \rrbracket, \quad X_i \sim \mathcal{B}\left(\frac{7}{12}\right)$$

2. Comme $D \sim \mathcal{U}(\llbracket 1; 6 \rrbracket)$, $\mathbb{V}(D) = \frac{6^2 - 1}{12} = \frac{35}{12}$.

Or $\mathbb{V}(D) = \mathbb{V}\left(\sum_{i=1}^6 X_i\right) = \sum_{i=1}^6 \mathbb{V}(X_i) + \sum_{i \neq j} \text{Cov}(X_i, X_j)$ où la somme double contient exactement $6 \times (6 - 1) = 30$ termes identiques.

Donc $6p(1 - p) + 30\text{Cov}(X_i, X_j) = \frac{35}{12}$, $30\text{Cov}(X_i, X_j) = \frac{35}{12} - \frac{35}{24} = \frac{35}{24}$ donc

$$\forall i \neq j, \quad \text{Cov}(X_i, X_j) = \frac{35}{24 \times 30} = \frac{7}{144}.$$

3. $S = \sum_{i=1}^6 iX_i$ donc par linéarité de l'espérance

$$\mathbb{E}(S) = \sum_{i=1}^6 ip = p \frac{6(6+1)}{2} = \frac{49}{4}.$$

$$\text{Ensuite } \mathbb{V}(S) = \sum_{i=1}^6 \mathbb{V}(iX_i) + \sum_{i \neq j} \text{Cov}(iX_i, jX_j) = \frac{35}{24} \sum_{i=1}^6 i^2 + \frac{7}{144} \sum_{i \neq j} ij$$

$$\text{Or : } \sum_{i \neq j} ij = \left(\sum_i i\right)^2 - \sum_i i^2 = \frac{n^2(n+1)^2}{4} - \frac{n(n+1)(2n+1)}{6} =$$

$$\frac{n(n+1)}{12} (3n(n+1) - 2(2n+1)) = \frac{n(n+1)(3n^2 - n + 1)}{12} \stackrel{n=6}{=} \frac{721}{2}.$$

$$\text{Numériquement (sans trop d'intérêt) : } \mathbb{V}(S) = \frac{5635}{144}.$$

Exercice 5 Péage

À un péage autoroutier n voitures franchissent au hasard et indépendamment l'une des trois barrières de péage mises à leur disposition. On note X_1, X_2 et X_3 les variables aléatoires dénombrant les voitures ayant franchi chacune de ces barrières.

1. Déterminer les lois de X_1 et $X_1 + X_2$.
2. En déduire la covariance de X_1 et X_2 .

Solution (Ex.5 – Péage)

1. $X_1 \sim \mathcal{B}(n, 1/3)$ et $X_1 + X_2 \sim \mathcal{B}(n, 2/3)$.

$$2. \text{Cov}(X_1, X_2) = \frac{1}{2} (\mathbb{V}(X_1 + X_2) - \mathbb{V}(X_1) - \mathbb{V}(X_2)) = \frac{1}{2} \left(\frac{2}{9}n - 2 \frac{2}{9}n \right) = -\frac{n}{9}$$

Exercice 6 Un coefficient de corrélation

Soient λ un réel strictement positif et α un réel de $]0; 1[$.

On considère un couple de variables aléatoires discrètes à valeurs dans $\mathbb{N} \times \mathbb{N}$, (X, Y) , dont la loi est donnée par :

$$\forall (i, j) \in \mathbb{N}^2, \quad p_{i,j} = \mathbb{P}((X = i) \cap (Y = j)) = \begin{cases} \frac{\lambda^i e^{-\lambda} \alpha^j (1-\alpha)^{i-j}}{j!(i-j)!} & \text{si } 0 \leq j \leq i \\ 0 & \text{sinon} \end{cases}.$$

1. Vérifier que les nombres $(p_{i,j})_{(i,j) \in \mathbb{N} \times \mathbb{N}}$ définissent la loi d'un couple de variables aléatoires.
2. Démontrer que X et Y suivent chacune une loi de Poisson en précisant leur paramètre.
3. Justifier que X et Y ne sont pas indépendantes.

4. Établir que XY admet une espérance et que $\mathbb{E}(XY) = \alpha\lambda(\lambda + 1)$.

5. Calculer $\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma(X)\sigma(Y)}$.

Solution (Ex.6 – Un coefficient de corrélation)

1. Tous les termes sont positifs.

$$\sum_{i,j} p_{i,j} = \sum_{i=0}^{+\infty} \sum_{j=0}^i p_{i,j} = \sum_{i=0}^{+\infty} \lambda^i e^{-\lambda} \sum_{j=0}^i \frac{1}{i!} \binom{i}{j} \alpha^j (1-\alpha)^{i-j} = e^{-\lambda} \sum_{i=0}^{+\infty} \frac{\lambda^i 1^i}{i!} = 1, \text{ la famille est sommable de somme 1, tout va bien.}$$

2. Avec le SCE ($[Y = j]$) $_{j \in \mathbb{N}}$, la FPT donne

$$\mathbb{P}(X = i) = \sum_{j=0}^{+\infty} p_{i,j} = \sum_{j=0}^i p_{i,j} = e^{-\lambda} \frac{\lambda^i}{i!} \text{ donc } X \sim \mathcal{P}(\lambda).$$

Avec le SCE ($[Y = i]$) $_{i \in \mathbb{N}}$, la FPT donne

$$\mathbb{P}(Y = j) = \sum_{i=0}^{+\infty} p_{i,j} = \sum_{i=j}^{+\infty} p_{i,j} = \frac{e^{-\lambda} \alpha^j \lambda^j}{j!} \sum_{i=j}^{+\infty} \frac{\lambda^{i-j} \alpha^{i-j}}{(i-j)!}$$

$$\mathbb{P}(Y = j) = \frac{e^{-\lambda} \alpha^j \lambda^j}{j!} e^{(1-\alpha)\lambda} = \frac{e^{-\alpha\lambda} (\alpha\lambda)^j}{j!} \text{ donc } Y \sim \mathcal{P}(\alpha\lambda).$$

3. $\mathbb{P}(X = 0, Y = 1) = 0 \neq \mathbb{P}(X = 0)\mathbb{P}(Y = 1)$ donc X et Y ne sont pas indépendantes.

$$4. \mathbb{E}(XY) = \sum_{i,j} ij p_{i,j} = \sum_{i=0}^{+\infty} i \sum_{j=0}^i p_{i,j} = \sum_{i=0}^{+\infty} \frac{\lambda^i e^{-\lambda}}{i!} \sum_{j=0}^i j \binom{i}{j} \alpha^j (1-\alpha)^{i-j}$$

En reconnaissant dans cette dernière somme l'espérance d'une loi $\mathcal{B}(i, \alpha)$

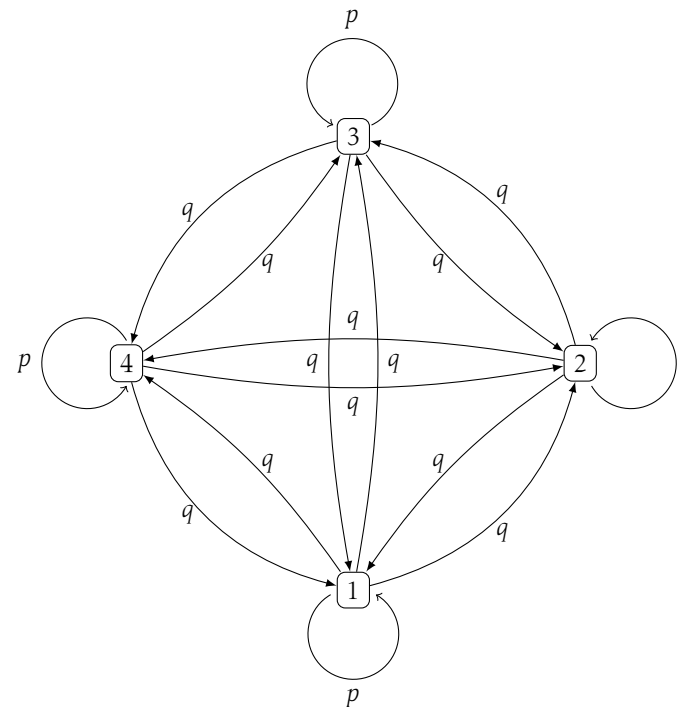
$$\mathbb{E}(XY) = e^{-\lambda} \sum_{i=0}^{+\infty} i \frac{\lambda^i}{i!} \alpha \stackrel{\text{transfert}}{=} \alpha \mathbb{E}(X^2) = \alpha (\mathbb{V}(X) + \mathbb{E}(X)^2) = \alpha(\lambda + \lambda^2), \text{ la famille étant finalement sommable, tout va bien : } \mathbb{E}(XY) = \alpha\lambda(\lambda + 1).$$

5. $\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \alpha\lambda(\lambda + 1) - \lambda\alpha\lambda = \alpha\lambda$, puis

$$\rho(X, Y) = \frac{\alpha\lambda}{\sqrt{\lambda}\sqrt{\alpha\lambda}} = \sqrt{\alpha}.$$

Exercice 7 Exemple de chaîne de Markov à 4 états

On considère quatre points dans le plan numérotés de 1 à 4. Une particule se déplace chaque seconde sur l'ensemble de ces points de la façon suivante : si elle se trouve au point i , elle y reste avec une probabilité égale à $p \in]0; 1[$ ou passe en un point $j \neq i$ de façon équiprobable, donc avec une probabilité $q = \frac{1-p}{3}$.



Exemple de chaîne de Markov à 4 états

On note X_0 une variable aléatoire de loi P_0 donnant la position X_0 en l'instant $n = 0$,

$$X_n \text{ la position du point à l'instant } n \text{ et } P_n = \begin{pmatrix} \mathbb{P}(X_n = 1) \\ \vdots \\ \mathbb{P}(X_n = 4) \end{pmatrix} \text{ la loi de } X_n.$$

1. Montrer **RIGOREUSEMENT** qu'il existe une matrice Q, que l'on déterminera, telle que

$$\forall n \in \mathbb{N}, \quad P_n = Q^n P_0.$$

La matrice Q s'appelle « matrice de transition » de cette chaîne de Markov.

2. a) Justifier que Q est diagonalisable.

b) Déterminer **RAPIDEMENT** les valeurs propres et sous-espaces propres de Q.

c) Montrer que la suite $(Q^n)_{n \in \mathbb{N}}$ converge. On note L sa limite.

3. a) Justifier l'existence d'une matrice orthogonale $U \in \mathcal{O}_4(\mathbb{R})$ telle que tUQU soit diagonale.

- b) En déduire que L est la matrice du projecteur orthogonal de $\mathcal{M}_{4,1}(\mathbb{R})$ sur $\text{Vect}(V)$ où V désigne le vecteur de $\mathcal{M}_{4,1}(\mathbb{R})$ dont tous les coefficients valent 1.
- c) Déterminer finalement la limite de P_n lorsque n tend vers $+\infty$. Sa limite dépend-elle de la position initiale de la particule ?
- d) Quelle est la loi limite de la suite de variables (X_n) ?

Généralités sur les matrices stochastiques

On dit qu'une matrice M de $\mathcal{M}_n(\mathbb{R})$ est *stochastique* si

$$\textcircled{1} \quad \forall (i, j) \in \llbracket 1; n \rrbracket^2, \quad m_{i,j} \geq 0,$$

$$\textcircled{2} \quad \forall i \in \llbracket 1; n \rrbracket, \quad \sum_{j=1}^n m_{i,j} = 1.$$

On dit de plus que M est *strictement stochastique* si, dans $\textcircled{1}$, toutes les inégalités sont strictes.

Soit U le vecteur de $\mathcal{M}_{n,1}(\mathbb{R})$ dont tous les coefficients valent 1

- 4. a) Soit $M \in \mathcal{M}_n(\mathbb{R})$.
Montrer que M est stochastique (respectivement strictement stochastique) si, et seulement si, $MU = U$ et tous les coefficients de M sont positifs (respectivement strictement positifs).
- b) En déduire que le produit de deux matrices stochastiques (respectivement strictement stochastiques) est une matrice stochastique (respectivement strictement stochastique).
- c) Si M est stochastique et si la suite (M^n) converge, que peut-on dire de $\lim_{n \rightarrow +\infty} M^n$?

- 5. Soit M une matrice stochastique et λ une valeur propre de M.
Soit $V = (v_i)_{1 \leq i \leq n} \in \mathcal{M}_{n,1}(\mathbb{K})$ un vecteur propre de M associé à la valeur propre λ .
On note $k \in \llbracket 1; n \rrbracket$ un indice tel que $v_k = \max_{1 \leq i \leq n} |v_i|$.

- a) Justifier que $|\lambda| |v_k| \leq |v_k|$.
- b) Que peut-on en déduire pour λ ?
- c) Montrer que $|\lambda - m_{k,k}| \leq 1 - m_{k,k}$.

On suppose de plus, jusque la fin de cette question, que M est strictement stochastique.

- d) Montrer que, si $\lambda \neq 1$, alors $|\lambda| < 1$.
- e) Soit X un vecteur propre réel de M associé à 1 et x_k son coefficient de plus grande valeur absolue. Quitte à remplacer X par $-X$, on suppose $x_k > 0$.

Justifier que $x_k = \sum_{j=1}^n m_{k,j} x_j$ et en déduire que X est proportionnel à U.

- 6. Montrer que si M est strictement stochastique et diagonalisable, alors la suite $(M^m)_{m \in \mathbb{N}}$ converge vers une matrice de projecteur de rang 1.

Solution (Ex.7 – Exemple de chaîne de Markov à 4 états)

- 1. • $([X_n = k]_{1 \leq k \leq 4})$ étant un système complet d'événements, en appliquant la formule des probabilités totales :

$$\begin{aligned} \forall j \in \llbracket 1; 4 \rrbracket, \\ \mathbb{P}(X_{n+1} = j) &= \mathbb{P}(X_n = 1) \mathbb{P}_{[X_n=1]}(X_{n+1} = j) \\ &+ \mathbb{P}(X_n = 2) \mathbb{P}_{[X_n=2]}(X_{n+1} = j) \\ &+ \mathbb{P}(X_n = 3) \mathbb{P}_{[X_n=3]}(X_{n+1} = j) \\ &+ \mathbb{P}(X_n = 4) \mathbb{P}_{[X_n=4]}(X_{n+1} = j) \end{aligned}$$

$$\text{Or } \mathbb{P}_{[X_n=k]}(X_{n+1} = j) = \begin{cases} p & \text{si } j = k \\ q & \text{si } j \neq k \end{cases}$$

Donc les $\mathbb{P}(X_{n+1} = j)$ sont des combinaisons linéaires des $\mathbb{P}(X_n = j)$ avec

$$P_{n+1} = QP_n \quad \text{où} \quad Q = \begin{pmatrix} p & q & q & q \\ q & p & q & q \\ q & q & p & q \\ q & q & q & p \end{pmatrix}.$$

- Comme Q est indépendante de n, une récurrence immédiate montre que

$$\forall n \in \mathbb{N}, \quad P_n = Q^n P_0.$$

- 2. a) $Q \in \mathcal{S}_n(\mathbb{R})$ donc Q est diagonalisable.
- b) $\text{rg}(Q - (p - q)I_r) = 1$ donc $p - q \in \text{Sp}(Q)$ et $\dim(\text{SEP}(Q, p - q)) = 3$.
Il reste une valeur propre λ de multiplicité 1 à trouver puisque Q est diagonalisable.

Comme $Q^T = 4p = 3 \times (p - q) + \lambda$, $\lambda = p + 3q = 1$.

On peut aussi observer que la somme des coefficients sur chaque ligne vaut $p +$

$$3q = 1 \text{ donc } Q \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = 1 \times \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \text{ donc } 1 \in \text{Sp}(Q).$$

$$\text{SEP}(Q, p - q) = \text{Vect} \left(\begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix} \right) \text{ et } \text{SEP}(Q, 0) = \text{Vect} \left(\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \right).$$

c) Il existe une matrice inversible M telle que

$$M^{-1}QM = D = \text{diag}(1, p - q, p - q, p - q)$$

Alors $Q^n = MD^nM^{-1}$. Or $|p - q| = \left| p - \frac{1-p}{3} \right| = \left| \frac{4p-1}{3} \right| \in]-1/3; 1[$ donc $(p - q)^n \xrightarrow{n \rightarrow +\infty} 0$, et $D^n \xrightarrow{n \rightarrow +\infty} \text{diag}(1, 0, 0, 0)$.

Comme M et M^{-1} ne dépendent pas de n,

$$Q^n \xrightarrow{n \rightarrow +\infty} M \text{diag}(1, 0, 0, 0) M^{-1}.$$

3. a) Comme $Q \in \mathcal{S}_4(\mathbb{R})$, le théorème spectral assure l'existence d'une matrice orthogonale $U \in \mathcal{O}_4(\mathbb{R})$ telle que U^TQU soit égale à D diagonale.

b) On a : $L = \lim_{n \rightarrow +\infty} UD^nU^T = U \lim_{n \rightarrow +\infty} D^nU^T = U\Delta U^T$ où $\Delta = \text{diag}(1, 0, 0, 0)$.

$L^2 = U\Delta^2U^T = U\Delta U^T = L$ et $L^T = U^T\Delta^T U^T = U\Delta U^T = L$ donc L est la matrice du projecteur orthogonal de $\mathcal{M}_{4,1}(\mathbb{R})$ sur $\text{Im}(L) = \text{SEP}(L, 1)$.

Enfin $\text{rg}(L) = \text{rg}(\Delta) = 1$ et $LV = V$. En effet, $QV = V$ car $V \in \text{SEP}(Q, 1)$ donc : $\forall n \in \mathbb{N}, Q^nV = V$, donc $\lim_{n \rightarrow +\infty} Q^nV = V$ donc $LV = V$. Donc $\text{Im}(V) = \text{Vect}(V)$, c'est gagné!

c) Déterminer finalement la limite de P_n lorsque n tend vers $+\infty$. $\lim_{n \rightarrow +\infty} P_n =$

$$\lim_{n \rightarrow +\infty} Q^n P_0 = LP_0 = p_{\text{Vect}(V)}(P_0) = (P_0 | V')V' \text{ où } V' = \frac{1}{\|V\|}V = \frac{1}{2}V \text{ est un vecteur directeur unitaire de } \text{Vect}(V).$$

$$\text{En notant } P_0 = \begin{pmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \\ \pi_4 \end{pmatrix}, \lim_{n \rightarrow +\infty} P_n = \frac{1}{4}(\pi_1 + \pi_2 + \pi_3 + \pi_4)V = \frac{1}{4}V.$$

$$\lim_{n \rightarrow +\infty} P_n = \begin{pmatrix} 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{pmatrix}.$$

Sa limite ne dépend pas de la position initiale de la particule caractérisée par les valeurs des π_i .

d) La loi limite de la suite de variables (X_n) est la loi uniforme sur $[[1; 4]]$.

Généralités sur les matrices stochastiques

4. a) $\forall i \in [[1; n]], (MU)_i = \sum_{j=1}^n m_{i,j}u_j = \sum_{j=1}^n m_{i,j}$, donc M vérifie ② si, et seulement si,

$MU = U$. Et évidemment positivité (stricte éventuellement).

b) Soit M et N stochastiques.

$\forall (i, j) \in [[1; n]]^2, (MN)_{i,j} = \sum_{k=1}^n m_{i,k}n_{k,j}$ donc le produit matriciel conserve la (stricte) positivité des coefficients.

$MNU = M(NU) = MU = U$ donc MN vérifie (Σ).

c) $\forall k \in \mathbb{N}^*, M^kU = U$ par récurrence immédiate. Et si (M^k) converge vers L, alors $\lim_{k \rightarrow +\infty} M^kU = LU$ d'une part et $\lim_{k \rightarrow +\infty} M^kU = U$ d'autre part. Donc si M est stochastique et la suite de ses puissances (M^k) converge, alors $\lim_{k \rightarrow +\infty} M^k$ est stochastique, et positive par convergence par coordonnées toujours positives. On perd éventuellement la STRICTE positivité, comme le montre l'exemple introductif.

5. a) Soit $\lambda \in \text{Sp}(M)$ et $V = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \neq 0$ tel que $MV = \lambda V$. Soit $k \in [[1; n]]$ tel que $|v_k| =$

$\max_{1 \leq i \leq n} |v_i| > 0$. La k-ème ligne de $MV = \lambda V$ fournit

$$\sum_{j=1}^n m_{k,j} v_j = \lambda v_k \quad (\mathcal{P})$$

Et par l'inégalité triangulaire

$$|\lambda| |v_k| \leq \sum_{j=1}^n m_{k,j} |v_j| \leq \sum_{j=1}^n m_{k,j} |v_k| \leq 1 \times |v_k|$$

b) ...donc $|\lambda| \leq 1$ puisque $|v_k| > 0$.

c) Je continue avec les mêmes notations, en supposant de plus $|\lambda| = 1$. (\mathcal{P}) fournit

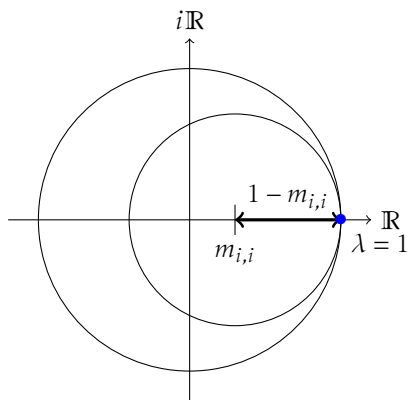
$$(\lambda - m_{k,k}) x_k = \sum_{j \neq k} m_{k,j} x_j \text{ donc } \lambda - m_{k,k} = \sum_{j \neq k} m_{k,j} \frac{x_j}{x_k}$$

L'inégalité triangulaire permet d'écrire

$$1 - m_{k,k} = |\lambda - m_{k,k}| \leq |\lambda - m_{k,k}| \leq \sum_{j \neq k} m_{k,j} \left| \frac{x_j}{x_k} \right| \leq \sum_{j \neq k} m_{k,j} = 1 - m_{k,k}.$$

Puisque les membres extrêmes sont égaux, ces inégalités sont toutes des égalités et en particulier $1 - m_{k,k} = \left| \lambda - m_{k,k} \right|$.

Donc λ est sur le cercle de centre $m_{k,k}$ et de rayon $1 - m_{k,k}$, et comme $|\lambda| = 1$, λ est aussi sur le cercle unité, et il n'y a qu'une possibilité car ces deux cercles sont tangents en 1 :



Donc $\lambda = 1$.

d) On a $|\lambda| \leq 1$ et $|\lambda| = 1 \implies \lambda = 1$, donc $\lambda \neq 1 \implies |\lambda| < 1$.

e) Donc $\text{SEP}(M, 1) = \text{Vect}(U)$.

f) Si on suppose de plus M diagonalisable, alors M est semblable à $D = \text{diag}(1, \lambda_2, \dots, \lambda_n)$ avec $\forall i \in \llbracket 2; n \rrbracket, |\lambda_i| < 1$ d'après ③.

Du coup, $D^k \xrightarrow{k \rightarrow +\infty} \Delta = \text{diag}(1, 0, \dots, 0)$.

Écrivons $M = PDP^{-1}$. Alors $M^k \xrightarrow{k \rightarrow +\infty} L = P\Delta P^{-1}$.

Ce qui prouve déjà que la suite (M^k) converge.

De plus L est semblable à Δ , de rang 1, donc L est de rang 1.

Et $L^2 = P\Delta^2 P^{-1} = P\Delta P^{-1} = L$ car $\Delta^2 = \Delta$, donc L est une matrice de projecteur de rang 1.

Et à quoi elle ressemble ?

Donc toutes ses lignes sont proportionnelles, mais L est stochastique car limite d'une suite de matrices stochastiques. Donc ses lignes sont toutes égales entre elles, car leur somme vaut toujours 1. Donc

$$M^k \xrightarrow{k \rightarrow +\infty} L = \begin{pmatrix} p_1 & p_2 & \dots & p_n \\ p_1 & p_2 & \dots & p_n \\ \vdots & \vdots & & \vdots \\ p_1 & p_2 & \dots & p_n \end{pmatrix} \text{ où } p_1 + p_2 + \dots + p_n = 1.$$

Exercice 8 Loi multinomiale et matrice de covariance

Une urne contient des boules de r couleurs différentes C_1, \dots, C_r , en proportion respective p_1, \dots, p_r de sorte que : $p_1 + \dots + p_r = 1$. On tire au hasard et avec remise n boules dans cette urne et on note, pour chaque i entre 1 et r , v_i le nombre de boules de couleur C_i obtenues au cours des n tirages, de sorte que : $v_1 + \dots + v_r = n$ (étonnant non ?).

1. On note X le vecteur aléatoire (v_1, \dots, v_r) .
 - a) Décrire l'ensemble $X(\Omega)$ des valeurs possibles de X .
 - b) Donner la loi de X .
 - c) Qu'obtient-on lorsque $r = 2$?
2. a) Donner, pour chaque i entre 1 et r , la loi, l'espérance et la variance de v_i .
 - b) Donner, pour chaque i et j distincts entre 1 et r , la loi de $v_i + v_j$. En déduire la covariance du couple (v_i, v_j) .
3. Pour chaque i entre 1 et r , on pose $y_i = \frac{v_i - np_i}{\sqrt{np_i}}$.

Les v.a.r. y_i sont-elles centrées, c'est-à-dire d'espérance nulle ? Sont-elles réduites, c'est-à-dire de variance égale à 1 ?

4. a) On note C la matrice de variance-covariance des $(y_i)_{1 \leq i \leq n}$:

$$C \stackrel{\text{déf.}}{=} (\text{Cov}(y_i, y_j))_{1 \leq i, j \leq r}$$

Montrer que $C = I_r - N$ où $N = L^T \cdot L$ avec $L = (\sqrt{p_1}, \dots, \sqrt{p_r}) \in \mathcal{M}_{1,r}(\mathbb{R})$.

- b) Montrer que : $N^2 = N$ et que : $\text{rg}(N) = 1$.
- c) Que dire de $\text{rg}(C)$?

5. Justifier que C est symétrique positive mais non définie.

Solution (Ex.8 – Loi multinomiale et matrice de covariance)

1. a) $X(\Omega) = \{(k_1, \dots, k_r) \in \mathbb{N}^r, \sum_{i=1}^r k_i = n\}$.

b) Soit $\mathcal{T} = \{(c_1, \dots, c_n) \in \llbracket 1; r \rrbracket^n\} = \llbracket 1; r \rrbracket^n$ l'ensemble des tirages possibles. Dénombrons les tirages favorables à $X = (k_1, \dots, k_r) \in X(\Omega)$. Pour construire un tirage favorable, il faut choisir les rangs de k_1 tirages amenant la couleur C_1 et il y a $\binom{n}{k_1}$ possibilités.

Puis il faut choisir les rangs de k_2 tirages amenant la couleur C_2 et il y a $\binom{n-k_1}{k_2}$ possibilités.

En itérant le procédé, il y a :

$$\binom{n}{k_1} \binom{n-k_1}{k_2} \dots \binom{n-k_1-\dots-k_{r-1}}{k_r} = \frac{n!}{k_1!k_2!\dots k_r!}$$

tirages favorables à $X = (k_1, \dots, k_r)$, et chacun d'eux a une probabilité $p_1^{k_1} p_2^{k_2} \dots p_r^{k_r}$ de se réaliser, par indépendance des expériences.

Les tirages étant deux à deux incompatibles, la loi de X est

$$\forall (k_1, \dots, k_r) \in X(\Omega), \mathbb{P}(X = (k_1, \dots, k_r)) = \frac{n!}{k_1!k_2!\dots k_r!} p_1^{k_1} p_2^{k_2} \dots p_r^{k_r}.$$

c) Pour $r = 2$, $\mathbb{P}(X = (k_1, k_2)) = \frac{n!}{k_1!k_2!} p_1^{k_1} p_2^{k_2} = \binom{n}{k_1} p_1^{k_1} (1-p_1)^{n-k_1}$, on retombe sur la loi binomiale.

2. a) $v_i \hookrightarrow \mathcal{B}(n, p_i)$, $\mathbb{E}(v_i) = np_i$ et $\mathbb{V}(v_i) = np_i(1-p_i)$.

b) $v_i + v_j \hookrightarrow \mathcal{B}(n, p_i + p_j)$.

$$\begin{aligned} \text{Cov}(v_i, v_j) &= \frac{1}{2} (\mathbb{V}(v_i + v_j) - \mathbb{V}(v_i) - \mathbb{V}(v_j)) \\ &= \frac{1}{2} (n(p_i + p_j)(1-p_i-p_j) - np_i(1-p_i) - np_j(1-p_j)) \\ &= \frac{1}{2} (-2np_i p_j) = -np_i p_j \end{aligned}$$

3. • $\mathbb{E}(y_i) = \frac{1}{\sqrt{np_i}} (\mathbb{E}(v_i) - np_i) = 0$: y_i est centrée.

• $\mathbb{V}(y_i) = \frac{1}{np_i} \mathbb{V}(v_i) = 1 - p_i \neq 1$: y_i n'est pas réduite.

4. a) • $\forall i \in \llbracket 1; r \rrbracket, c_{i,i} = \text{Cov}(y_i, y_i) = \mathbb{V}(y_i) = 1 - p_i = 1 - \sqrt{p_i p_i}$

• $\forall (i, j) \in \llbracket 1; r \rrbracket^2$ avec $i \neq j$, comme $\text{Cov}(aX + b, cY + d) = ac \text{Cov}(X, Y)$, on a

$$c_{i,j} = \text{Cov}(y_i, y_j) = \frac{1}{\sqrt{np_i} \sqrt{np_j}} \text{Cov}(v_i, v_j) = -\sqrt{p_i p_j}$$

• On a bien : $C = I_r - N$ où $N = (\sqrt{p_i p_j})_{1 \leq i, j \leq r} = {}^t L \cdot L$ avec $L = (\sqrt{p_1}, \dots, \sqrt{p_r})$.

b) • $N^2 = {}^t L L {}^t L L = {}^t L (L {}^t L) L$ or $L {}^t L = (\sum_{i=1}^r p_i) = (1)$ donc $N^2 = {}^t L L = N$.

• Les colonnes de N sont toutes proportionnelles à sa première colonne qui est non nulles : $C_j = \frac{\sqrt{p_j}}{\sqrt{p_1}} C_1$, donc $\text{rg}(N) = 1$.

c) Remarquons que N est une matrice de projecteur et C est la matrice du projecteur complémentaire. Ce sont même des matrices de projecteurs orthogonaux pour \mathbb{R}^r muni du produit scalaire canonique car elles sont symétriques.

$$\text{En effet : } C^2 = (I_r - N)^2 = I_r - 2N + N^2 = I_r - N = C.$$

Voici deux arguments pour le rang.

• Soit $X \in \mathcal{M}_{r,1}(\mathbb{R})$.

$$X \in \text{Ker}(C) \iff CX = 0 \iff X - NX = 0 \iff NX = X \iff X \in \text{Im}(N)$$

Donc $\text{Ker}(C) = \text{Im}(N)$ et $\text{rg}(C) = r - \dim(\text{Ker}(C)) = r - \text{rg}(N) = r - 1$.

• C est une matrice de projecteur donc diagonalisable avec $\text{Sp}(C) \subset \{0, 1\}$, donc

$$\text{semblable à une matrice } \left(\begin{array}{c|c} I_k & 0 \\ \hline 0 & 0 \end{array} \right).$$

$$\text{Alors } \text{rg}(C) = k = \text{Tr}(C) = \sum_{i=1}^r (1 - p_i) = r - \sum_{i=1}^r p_i = r - 1.$$

5. $C \in \mathcal{S}_r(\mathbb{R})$, $\text{Sp}(C) \subset \{0; 1\}$ donc $C \in \mathcal{S}_r^+(\mathbb{R})$ mais $0 \in \text{Sp}(C)$ donc $C \notin \mathcal{S}_r^{++}(\mathbb{R})$.

Exercice 9 X et Y-X indépendantes.

1. Soit $p \in]0; 1[$ et $q = 1 - p$. Soit Y une variable vérifiant $Y(\Omega) = \mathbb{N}$ et, pour tout n de \mathbb{N} , $\mathbb{P}(Y = n) = q^2(n+1)p^n$.

Vérifier que les conditions précédentes définissent bien une loi de variable aléatoire.

2. Étudier l'existence, et calculer éventuellement, $\mathbb{E}(Y)$.

3. Soit X une variable telle que, pour tout k de \mathbb{N} ,

$$\mathbb{P}_{[Y=n]}(X = k) = \begin{cases} \frac{1}{n+1} & \text{si } 0 \leq k \leq n \\ 0 & \text{sinon} \end{cases}.$$

a) Déterminer la loi de X.

b) Montrer que Y - X suit la loi de X.

c) Montrer que Y - X et X sont indépendantes.

Solution (Ex.9 – X et Y-X indépendantes.)

1. $\forall n \in \mathbb{N}, \mathbb{P}(Y = n) \geq 0$, et

$$\sum_{n=0}^{+\infty} \mathbb{P}(Y = n) = q^2 \sum_{n=0}^{+\infty} (n+1)p^n = q^2 \frac{d}{dp} \left(\frac{1}{1-p} \right) = \frac{q^2}{(1-p)^2} = 1 \dots \text{gagné!}$$

$$2. \sum_{n=0}^{+\infty} n\mathbb{P}(Y = n) = q^2 \sum_{n=0}^{+\infty} n(n+1)p^n = q^2 p \sum_{n=1}^{+\infty} n(n+1)p^{n-1} = q^2 p \frac{d^2}{dp^2} \left(\frac{1}{1-p} \right)$$

$$\sum_{n=0}^{+\infty} \mathbb{P}(Y = n) = q^2 p \frac{2}{(1-p)^3} = \frac{p}{q}.$$

3. a) À l'aide du SCE $([Y = n])_{n \in \mathbb{N}}$, la FPT fournit

$$\mathbb{P}(X = k) = \sum_{n=0}^{+\infty} \mathbb{P}(Y = n)\mathbb{P}(X = k|Y = n) = \sum_{n=k}^{+\infty} q^2(n+1)p^n \frac{1}{n+1} = q^2 p^k \sum_{n=k}^{+\infty} p^{n-k} =$$

$$qp^k.$$

On pourra remarquer que X + 1 suit la loi $\mathcal{G}(q)$.

$$b) \mathbb{P}(Y - X = k) = \sum_{n=0}^{+\infty} \mathbb{P}(Y = n)\mathbb{P}(Y - X = k|Y = n)$$

$$= \sum_{n=0}^{+\infty} \mathbb{P}(Y = n)\mathbb{P}(X = n - k|Y = n) = \sum_{n=k}^{+\infty} q^2(n+1)p^n \frac{1}{n+1} \text{ car } \mathbb{P}(X = n - k|Y = n) = 0$$

si $n - k < 0$

On retrouve la même somme que pour la loi de X.

4. Soit $(i, j) \in \mathbb{N}^2$.

$$\mathbb{P}(Y - X = i, X = j) = \mathbb{P}(Y = i + j, X = j) = \mathbb{P}(Y = i + j)\mathbb{P}_{[Y=i+j]}(X = j)$$

$$\mathbb{P}(Y - X = i, X = j) = q^2(i + j + 1)p^{i+j} \frac{1}{i + j + 1} = (qp^i)(qp^j)$$

$$\mathbb{P}(Y - X = i, X = j) = \mathbb{P}(Y - X = i)\mathbb{P}(X = j).$$

Donc Y - X et X sont indépendantes.

Exercice 10 Bienaymé-Tchebychev et Markov

1. Soit $\lambda > 0$ et X une variable aléatoire suivant la loi de Poisson de paramètre λ .

Justifier que $\mathbb{P}(X \geq 2\lambda) \leq \frac{1}{\lambda}$.

2. a) Soit Y une variable d'espérance μ et d'écart-type $\sigma > 0$.

Montrer à l'aide de la variable $Z = (\alpha(Y - \mu) + \sigma)^2$ que, pour tout $\alpha > 0$,

$$\mathbb{P}(Y \geq \mu + \alpha\sigma) \leq \frac{1}{1 + \alpha^2}.$$

b) En déduire une nouvelle majoration de $\mathbb{P}(X \geq 2\lambda)$.

Solution (Ex.10 – Bienaymé-Tchebychev et Markov)

1. Par l'inégalité de Bienaymé-Tchebychev appliquée à X avec $\varepsilon = \lambda > 0$:

$$\mathbb{P}(|X - \lambda| \geq \lambda) \leq \frac{\lambda}{\lambda^2}.$$

Or $[|X - \lambda| \geq \lambda] = [X \geq 2\lambda] \cup [X \leq 0]$ donc $[X \geq 2\lambda] \subset [|X - \lambda| \geq \lambda]$, et

$$\mathbb{P}(X \geq 2\lambda) \leq \frac{1}{\lambda}.$$

2. a) Par l'inégalité de Markov appliquée à $Z \geq 0$ avec $\varepsilon > 0$, $\mathbb{P}(Z \geq \varepsilon) \leq \frac{\mathbb{E}(Z)}{\varepsilon}$

$$\text{Or } \mathbb{E}(Z) = \alpha^2 \mathbb{E}((Y - \mu)^2) + 2\alpha\sigma \mathbb{E}(Y - \mu) + \sigma^2 = \alpha^2 \sigma^2 + 0 + \sigma^2.$$

$$\text{Donc } \mathbb{P}(Z \geq \varepsilon) \leq \frac{\sigma^2(\alpha^2 + 1)}{\varepsilon}.$$

$$\text{Avec } \varepsilon = \sigma^2(\alpha^2 + 1)^2, \mathbb{P}(Z \geq \varepsilon) \leq \frac{1}{1 + \alpha^2}.$$

$$\text{Enfin : } [Y \geq \mu + \alpha\sigma] \subset [Y - \mu \geq \alpha\sigma] \subset [\alpha(Y - \mu) + \sigma \geq (\alpha^2 + 1)\sigma]$$

$$\text{Donc : } [Y \geq \mu + \alpha\sigma] \subset [Z \geq (\alpha^2 + 1)^2 \sigma^2] \subset [Z \geq \varepsilon], \text{ d'où}$$

$$\mathbb{P}(Y \geq \mu + \alpha\sigma) \leq \frac{1}{\alpha^2 + 1}.$$

b) Avec $\alpha = 1$, $\mathbb{P}(X \geq 2\lambda) \leq \frac{1}{2}$.

Remarque : cette deuxième majoration est plus précise lorsque $\frac{1}{2} < \frac{1}{\lambda}$ c'est-à-dire $\lambda < 2$.