2022-2023 MP2I

29. Variables aléatoires, corrigé

Exercice 1. On considère la variable aléatoire X à valeurs dans [1,5] dont la loi est donnée par P(X=k)=k/15 pour $k \in [1,5]$.

1) Il suffit de vérifier que $\sum_{k=1}^{5} P(X=k) = 1$ (les probabilités données étant positives). On a $\frac{5}{5}$

 $\sum_{k=1}^{5} P(X=k) = \frac{5 \times 6}{2 \times 15} = 1. \text{ On a donc bien défini une variable aléatoire. On a}$

$$E(X) = \sum_{k=1}^{5} \frac{k^2}{15} = \frac{5 \times 6 \times 11}{6 \times 15} = \frac{11}{3}.$$

On a $V(X) = E(X^2) - (E(X))^2$ donc :

$$V(X) = \sum_{k=1}^{5} k^2 \times \frac{k}{15} - \frac{121}{9} = \frac{15^2}{15} - \frac{121}{9} = \frac{14}{9}.$$

2) $Y = X^2$ est à valeurs dans $\{1, 4, 9, 16, 25\}$. On a $P(Y = k^2) = P(X = k) = \frac{k}{15}$ pour k variant entre 1 et 5. On a $E(Y) = E(X^2) = 15$. On a $V(Y) = E(Y^2) - (E(Y))^2 = E(X^4) - (E(X^2))^2$. On calcule alors $E(X^4) = \sum_{k=1}^{5} k^4 \times \frac{k}{15} = 295$ d'où V(Y) = 70.

 $Z=\min(X-3,0)$ est à valeurs dans {-2,-1,0}. On a P(Z=-2)=P(X=1)=1/15, P(Z=-1)=P(X=2)=2/15 et P(Z=0)=12/15. On en déduit $E(X)=-2\times 1/15-1\times 2/15+0\times 12/15=-4/15$ et :

$$V(Z) = E(Z^{2}) - (E(Z))^{2} = 4 \times 1/15 + 1 \times 2/15 + 0 \times 12/15 - (4/15)^{2} = \frac{74}{15^{2}}$$

Exercice 2.

- 1) X suit une loi binomiale de paramètres (n,p) si on note p la probabilité d'obtenir face sur un lancer. On a donc E(X) = np.
- 2) On a $Y(\Omega) = [1, n]$. Puisque la famille $(X = j)_{0 \le j \le n}$ est un système complet d'évènements, on a pour tout $k \in [1, n]$:

$$P(Y = k) = \sum_{j=0}^{n} P_{X=j}(Y = k)P(X = j)$$

$$= P(X = 0) \times \frac{1}{n} + 1 \times P(X = k)$$

$$= \frac{(1-p)^{n}}{n} + \binom{n}{k} p^{k} (1-p)^{n-k}.$$

On a alors pour l'espérance :

$$E(Y) = \sum_{k=1}^{n} kP(Y=k)$$

$$= \sum_{k=1}^{n} k \left(\frac{(1-p)^n}{n} + \binom{n}{k} p^k (1-p)^{n-k} \right)$$

$$= \frac{(n+1)(1-p)^n}{2} + \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} - 0$$

$$= \frac{(n+1)(1-p)^n}{2} + np.$$

Exercice 3. On revient à la définition de l'espérance :

$$E(X) = \sum_{k=0}^{N} kP(X = k)$$

$$= 0 + \sum_{k=1}^{N} kP(X = k)$$

$$= 0 + \sum_{k=1}^{N} \sum_{j=1}^{k} P(X = k)$$

$$= \sum_{1 \le j \le k \le N} P(X = k)$$

$$= \sum_{j=1}^{N} \sum_{k=j}^{N} P(X = k)$$

$$= \sum_{j=1}^{N} P(X \ge j)$$

$$= \sum_{j=0}^{N-1} P(X \ge j + 1)$$

$$= \sum_{j=0}^{N-1} P(X > j).$$

On a ici utilisé le fait que les évènements $(X \ge j)$ et $\bigcup_{k=j}^{n} (X = k)$ sont égaux puisque X est à valeurs dans [0, N].

Exercice 4. On utilise le théorème de transfert. On a :

$$E(Y) = \sum_{k=0}^{n} \frac{1}{k+1} \binom{n}{k} p^{k} (1-p)^{n-k}.$$

On simplifie alors cette expression en remplaçant le coefficient binomial par son expression avec les factorielles :

$$\frac{1}{k+1} \binom{n}{k} = \frac{n!}{(k+1)!(n-k)!} = \frac{1}{n+1} \times \frac{(n+1)!}{(k+1)!(n+1-(k+1))!} = \frac{1}{n+1} \binom{n+1}{k+1}.$$

On fait alors apparaître un binôme de Newton (car $p \neq 0$):

$$E(Y) = \frac{1}{n+1} \sum_{k=0}^{n} \binom{n+1}{k+1} p^k (1-p)^{n-k}$$

$$= \frac{1}{n+1} \sum_{k=1}^{n+1} \binom{n+1}{k} p^{k-1} (1-p)^{n+1-k}$$

$$= \frac{1}{p(n+1)} \left(\sum_{k=0}^{n+1} \binom{n+1}{k} p^k (1-p)^{n+1-k} - (1-p)^{n+1} \right)$$

$$= \frac{1}{p(n+1)} \left(1 - (1-p)^{n+1} \right).$$

Exercice 5. Soient X et Y deux variables aléatoires. On suppose $V(X) \neq 0$. Remarquons tout d'abord que V(Y - aX - b) = V(Y - aX) (translater par une constante ne change pas la variance). Posons alors P(a) = V(Y - aX). On a :

$$P(a) = V(Y) - 2a\operatorname{Cov}(X, Y) + a^{2}V(X).$$

Puisque $V(X) \neq 0$, on a que P est un polynôme de degré 2 à coefficients réels. Puisque V(X) > 0, ce polynôme tend vers $+\infty$ en $\pm\infty$. Une étude de dérivée montre que ce polynôme est décroissant puis croissant. Il admet donc un unique minimum en $a = \frac{\operatorname{Cov}(X,Y)}{V(X)}$ (l'unique point d'annulation de la dérivée).

Une fois la valeur de a trouvée (et unique), il existe une unique valeur de b tel que Y - aX - b soit centrée (il s'agit de b = E(Y) + aE(x), qui s'obtient à partir de E(Y - aX - b) = 0 en utilisant la linéarité de l'espérance.

Exercice 7.

1) On a $X(\Omega) = [n, 2n]$. Si on suppose que l'on vide l'urne en entier, le nombre de tirages possibles est $\binom{2n}{n}$ (il faut placer les n boules blanches parmi les 2n tirages et les noires sont automatiquement placées). On a par hypothèse la probabilité uniforme sur l'ensemble des tirages possibles.

Soit $k \in [0, n]$. On veut calculer P(X = n + k). Pour cela, il faut que dans les n + k premiers tirages, il y ait toutes les boules blanches **ET** que la dernière boule tirée soit blanche (car on veut s'arrêter exactement au n + k-ième tirage). On doit donc placer une boule blanche en dernière position et on a alors autant de tirages possibles que le nompbre de façons de placer les k boules noires parmi les n + k - 1 premières positions, ce qui nous fait $\binom{n + k - 1}{k}$ tirages possibles.

On en déduit que $P(X = n + k) = \frac{\binom{n+k-1}{k}}{\binom{2n}{n}}$.

2) Par définition,
$$E(X) = \sum_{k=0}^{n} (n+k) \times \frac{\binom{n+k-1}{k}}{\binom{2n}{n}}$$
. Or, on a :
$$(n+k)\binom{n+k-1}{k} = n\binom{n+k}{k}.$$

Ceci entraine que $E(X) = \frac{n}{\binom{2n}{n}} \sum_{k=0}^{n} \binom{n+k}{k}$. Or, on peut démontrer que $\sum_{k=0}^{n} \binom{n+k}{k} = \binom{2n+1}{n}$

(cela se visualise très bien sur le triangle de Pascal en utilisant le fait que $\binom{n}{0} = \binom{n+1}{0}$ et par utilisation répétée de la formule de Pascal; on peut aussi faire une preuve par récurrence). On en déduit que :

$$E(X) = \frac{(2n+1)n}{n+1}.$$

Exercice 8.

1) D'après le cours, $\mathbb{V}(X_n) = 4n \times 1/2 \times 1/2 = n$ et on a $\mathbb{E}(X_n) = 4n \times 1/2 = 2n$.

On a $\mathbb{P}(|X_n-2n|\geq n)=\mathbb{P}(|X_n-\mathbb{E}(X_n)|\geq n)$. D'après l'inégalité de Tchebychev appliquée en $\varepsilon=n>0$, on en déduit que :

$$\mathbb{P}(|X_n - \mathbb{E}(X_n)| \ge n) \le \frac{\mathbb{V}(X_n)}{n^2} = \frac{1}{n}.$$

Puisqu'une probabilité est toujours positive, on en déduit par théorème des gendarmes que $\lim_{n\to+\infty} \mathbb{P}(|X_n - \mathbb{E}(X_n)| \ge n) = 0$.

2) On a $\mathbb{P}(|X_n-2n| \leq n) = \mathbb{P}(|X_n-2n| = n \cup (|X_n-2n| < n) = \mathbb{P}(|X_n-2n| = n) + \mathbb{P}(|X_n-2n| < n)$ car l'union est disjointe. On a donc finalement :

$$\mathbb{P}(|X_n - 2n| \le n) = \mathbb{P}(|X_n - 2n| = n) + 1 - \mathbb{P}(|X_n - 2n| \ge n).$$

Il ne reste donc plus qu'à trouver la limite de

$$\mathbb{P}(|X_n - 2n| = n) = \mathbb{P}(X_n = n) + \mathbb{P}(X_n = 3n) = \binom{4n}{n} \frac{1}{2^n} \times \frac{1}{2^{3n}} + \binom{4n}{3n} \frac{1}{2^{3n}} \times \frac{1}{2^n} = 2 \times \binom{4n}{n} \times \frac{1}{2^{4n}}.$$

Or, on a d'après la formule de Stirling:

$$\binom{4n}{n} = \frac{(4n)!}{n!(3n)!} \sim \frac{(4n)^{4n}\sqrt{8\pi n}e^n e^{3n}}{e^{4n}n^n(3n)^{3n}\sqrt{2\pi n}\sqrt{6\pi n}} \sim \left(\frac{4^4}{3^3}\right)^n \times \frac{\sqrt{2}}{\sqrt{3\pi n}}$$

On a donc:

$$\mathbb{P}(|X_n - 2n| = n) \sim \left(\frac{2^4}{3^3}\right)^n \times \frac{2\sqrt{2}}{\sqrt{3\pi n}}.$$

Or, $0 \le \frac{2^4}{3^3} = \frac{16}{27} < 1$. On en déduit donc que $\lim_{n \to +\infty} \mathbb{P}(|X_n - 2n| = n) = 0$. Par somme de limites, on en déduit d'après la question précédente que :

$$\lim_{n \to +\infty} \mathbb{P}(|X_n - 2n| \le n) = 1.$$

Plus simplement, on a que $\mathbb{P}(|X_n-2n|>n)\leq \mathbb{P}(|X_n-2n|\geq n)$ puisque l'évènement $(|X_n-2n|>n)$ entraine (est inclus dans) l'évènement $(|X_n-2n|\geq n)$. Puisqu'une probabilité est positive, on en déduit par théorème des gendarmes que $\lim_{n\to+\infty}\mathbb{P}(|X_n-2n|>n)=0$. En passant au complémentaire, on obtient donc que $\lim_{n\to+\infty}\mathbb{P}(|X_n-2n|\leq n)=1$.

3) On a par définition de la loi binomiale :

$$\mathbb{P}(|X_n - 2n| \le n) = \mathbb{P}(n \le X_n \le 3n) = \sum_{k=n}^{3n} \binom{4n}{k} \frac{1}{2^k} \times \frac{1}{2^{4n-k}} = \frac{1}{2^{4n}} x_n.$$

D'après la question précédente, on a donc $x_n \sim 2^{4n} \sim 16^n$.

Exercice 9. (m) Soit X une variable aléatoire de moyenne μ et de variance σ^2 . On fixe $\alpha > 0$.

1) Montrer que $P(\mu - \alpha \sigma < X < \mu + \alpha \sigma) \ge 1 - \frac{1}{\alpha^2}$.

OK avec l'inégalité de Tchebychev.

2) On pose la variable aléatoire $Y = (\alpha(X - \mu) + \sigma)^2$. Déterminer E(Y) en fonction de α et σ .

OK, on a trouvé $E(Y) = \sigma^2(1 + \alpha^2)$.

3) En déduire que $P(X \ge \mu + \alpha \sigma) \le \frac{1}{1 + \alpha^2}$.

Comme on l'a fait, on utilise l'inégalité de Markov à la variable aléatoire Y (qui est positive). On a donc pour tout a>0:

$$P(Y \ge a) \le \frac{E(Y)}{a}$$
.

On applique ceci en $a = \sigma^2(1+\alpha^2)^2$, ce qui donne puisque les événements $Y \ge a$ et $|\alpha(X-\mu)+\sigma| \ge \sigma(1+\alpha^2)$ sont égaux (par stricte croissante de $u \mapsto \sqrt{u}$ sur \mathbb{R}_+) que :

$$P(|\alpha(X-\mu)+\sigma| \ge \sigma(1+\alpha^2)) \le \frac{1}{1+\alpha^2}.$$

Or, si on considère l'événement $(X \ge \mu + \alpha \sigma)$, ceci est équivalent (puisque $\alpha > 0$) à $\alpha(X - \mu) + \sigma \ge (1 + \alpha^2)\sigma$. On a donc cet événement qui est inclus dans $|\alpha(X - \mu) + \sigma| \ge \sigma(1 + \alpha^2)$ et donc :

$$P(X \ge \mu + \alpha \sigma) \le P(|\alpha(X - \mu) + \sigma| \ge \sigma(1 + \alpha^2)) \le \frac{1}{1 + \alpha^2}.$$

Exercice 11. Une urne contient N boules numérotées de 1 à N.

1) On tire successivement et avec remise n boules de l'urne. On note X et Y les variables aléatoires égales au plus petit et au maximum des numéros obtenus.

Comme les tirages se font avec remise, on peut considérer $\Omega = [1, N]^n$, muni de la probabilité uniforme (puisque tous les évènements élémentaires ont même probabilité).

a) On a $X(\Omega) = [\![1,N]\!]$ puisque le plus petit numéro tiré va de 1 (si on tire la boule 1 une fois) à N (si on ne tire que la boule N lors des tirages). Soit $x \in [\![1,N]\!]$. L'évènement $(X \ge x)$ est réalisé quand au cours des n tirages toutes les boules ont un numéro plus grand que x. Il faut donc tirer nos boules uniquement parmi les numéros $[\![x,N]\!]$, ce qui entraine (N-x+1) tirages possibles. On a donc :

$$P(X \ge x) = \frac{(N - x + 1)^n}{N^n}.$$

Ceci entraine que $P(X = x) = P(X \ge x) - P(X \ge x + 1) = \frac{(N - x + 1)^n - (N - x)^n}{N^n}$.

b) Soit $y \in [1, N]$. On procède de même. $(Y \le y)$ signifie que les boules sont toutes tirées parmi les boules de 1 à y. On a donc $P(Y \le y) = \frac{y^n}{N^n}$ puis :

$$P(Y = y) = P(Y \le y) - P(Y \le y - 1) = \frac{y^n - (y - 1)^n}{N^n}.$$

On a également Y à valeurs dans [1, N].

c) Soit $(x,y) \in [\![1,N]\!]^2$. Remarquons que si x>y, alors $P((X\geq x)\cap (Y\leq y))=0$. Si x=y, cela signifie que l'on a tiré n fois la boule numéro x et donc $P((X=x)\cap (Y=x))=\frac{1}{N^n}$. Il reste alors le cas x< y. On a alors que l'évènement $(X\geq x)\cap (Y\leq y)$ correspond à tirer toutes les boules comprises entre x et y. On a donc, de même qu'avant :

$$P((X \ge x) \cap (Y \le y)) = \frac{(y - x + 1)^n}{N^n}.$$

On peut alors écrire:

$$\begin{array}{lcl} P((X=x)\cap (Y=y)) & = & P((X=x)\cap (Y\leq y)) - P((X=x)\cap (Y\leq y-1)) \\ & = & P((X\geq x)\cap (Y\leq y)) - P((X\geq x+1)\cap (Y\leq y)) - (P((X\geq x)\cap (Y\leq y-1)) - P((X\geq x)\cap (Y\leq y)) - P((X$$

On en déduit que si
$$x < y$$
, $P((X = x) \cap (Y = y)) = \frac{(y - x + 1)^n + (y - x - 1)^n - 2(y - x)^n}{N^n}$.

2) Cette fois, on tire sans remise et simultanément, ce qui revient à tirer n boules parmi N. Une réalisation est donc une partie à n éléments de [1, N] et toutes les possibilités sont équiprobables.

On garde alors exactement la même méthode. Ceci nous donne que $P(X \ge x) = \frac{\binom{N-x+1}{n}}{\binom{N}{n}}$ (cela revient à choisir les n boules tirées entre x et N donc parmi un ensemble à N-x+1 boules). On en déduit que :

$$P(X=x) = \frac{\binom{N-x+1}{n} - \binom{N-x}{n}}{\binom{N}{n}} = \frac{\binom{N-x}{n-1}}{\binom{N}{n}} \quad \text{d'après la formule de Pascal.}$$

On obtient ensuite $P(Y \leq y) = \frac{\binom{y}{n}}{\binom{N}{n}}$ puis $P(Y = y) = \frac{\binom{y-1}{n-1}}{\binom{N}{n}}$, toujours à l'aide de la formule de Pascal. Enfin, si $x, y \in [\![1,N]\!]$, on obtient :

$$P((X=x)\cap (Y=y)) = \frac{\binom{y-x}{n-1} - \binom{y-x-1}{n-1}}{N^n} = \frac{\binom{y-x-1}{n-2}}{\binom{N}{n}}.$$

Exercice 13. Puisque X et Y sont à valeurs dans $\{0,1\}$, alors Z aussi. On a donc que Z suit une loi de Bernoulli de paramètre $r = \mathbb{P}(Z = 1)$. Or, on a :

$$\mathbb{P}(Z=1) = \mathbb{P}(X=1 | JY=1) = \mathbb{P}(X=1) + \mathbb{P}(Y=1) - \mathbb{P}((X=1) \cap (Y=1)).$$

En effet, puisque Z est le maximum des deux, il suffit que X ou Y soit égal à 1 pour que Z soit égal à 1 (et Z vaut 0 si et seulement si X et Y sont tous les deux nuls). On a donc par indépendance de X et Y:

$$\mathbb{P}(Z=1) = p + q - \mathbb{P}(X=1)\mathbb{P}(Y=1) = p + q - pq.$$

On a donc Z qui suit une loi de Bernoulli de paramètre r = p + q - pq.

On pose X_i la variable aléatoire valant 1 si le premier archer touche la cible i et Y_i la variable aléatoire valant 1 si le second archer touche la cible i. X_i et Y_i suivent des lois de Bernoulli de paramètre p et q. On note alors $Z_i = \max(X_i, Y_i)$ qui suit donc une loi de Bernoulli. Le nombre de cibles touchées au total est donc $\sum_{i=1}^n Z_i$. Or, puisque les variables aléatoires $X_1, \ldots, X_n, Y_1, \ldots, Y_n$ sont mutuellement indépendantes, le lemme des coalitations assure que les Z_1, \ldots, Z_n sont mutuellement indépendantes. On en déduit que $\sum_{i=1}^n Z_i$ suit une loi binomiale de paramètre (n, p+q-pq).

Pour les cibles non touchées, on en a exactement $n-\sum_{i=1}^n Z_i=\sum_{i=1}^n (1-Z_i)$. Or, $1-Z_i$ suit une loi de Bernoulli de paramètre 1-r=1-(p+q-pq)=(1-p)(1-q). Par le même argument que ci-dessus, on en déduit que le nombre de cibles non touchées suit une loi binomiale de paramètre (n,(1-p)(1-q)).

Exercice 15. Puisque X + Y et X - Y sont indépendantes, on a V((X + Y) + (X - Y)) = V(X + Y) + V(X - Y). Ceci entraine que :

$$V(2X) = V(X) + 2\operatorname{Cov}(X, Y) + V(Y) + V(X) - 2\operatorname{Cov}(X, Y) + V(Y) = 2V(X) + 2V(Y).$$
 Or, on a $V(2X) = 4V(X)$. On a donc l'égalité voulue.

Exercice 16. D'après l'énoncé, X_1 suit une loi binomiale de paramètre $(n, \frac{1}{3})$. On en déduit que $\mathbb{V}(X_1) = \frac{2n}{9}$. Par symétrie, X_2 et X_3 suivent la même loi que X_1 donc $\mathbb{V}(X_2) = \frac{2n}{9}$. Puisque le nombre de voitures qui franchit le péage est égal à n, on a $X_1 + X_2 + X_3 = n$, ce qui entraine que :

$$\mathbb{V}(X_1 + X_2) = \mathbb{V}(n - X_3) = (-1)^2 \mathbb{V}(X_3) = \frac{2n}{9}$$

Or, on a aussi $\mathbb{V}(X_1 + X_2) = \mathbb{V}(X_1) + 2\operatorname{Cov}(X_1, X_2) + \mathbb{V}(X_2)$, ce qui entraine que :

$$Cov(X_1, X_2) = -\frac{n}{9}.$$

En particulier, les variables aléatoires X_1 et X_2 ne sont pas indépendantes!

Exercice 17.

1) La trace est linéaire et l'espérance également donc :

$$E(\operatorname{tr}(M)) = E(\sum_{i=1}^{n} X_{i,i})$$
$$= \sum_{i=1}^{n} E(X_{i,i})$$
$$= 0.$$

Pour le déterminant, on utilise d'abord la linéarité de l'espérance et le fait que pour des variables aléatoires indépendantes, l'espérance du produit est le produit des espérances. On a donc :

$$E(\det(M)) = E(\sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^n X_{\sigma(i),i})$$
$$= \sum_{\sigma \in S_n} \varepsilon(\sigma) E(\prod_{i=1}^n X_{\sigma(i),i}).$$

Puisque tous les $X_{\sigma(i),i}$ sont pour des variables différentes (et donc indépendantes), on a alors :

$$E(\prod_{i=1}^{n} X_{\sigma(i),i}) = \prod_{i=1}^{n} E(X_{\sigma(i),i}) = 0.$$

On a donc $E(\det(M)) = 0$.

2) On a $V(\det(M)) = E(\det(M)^2) - (E(\det(M)))^2 = E(\det(M) \times \det(M))$. On a alors par linéarité de l'espérance :

$$V(\det(M)) = E\left(\sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^n X_{\sigma(i),i} \times \sum_{\sigma' \in S_n} \varepsilon(\sigma') \prod_{j=1}^n X_{\sigma(j),j}\right)$$
$$= \sum_{\sigma \in S_n} \sum_{\sigma' \in S_n} \varepsilon(\sigma) \varepsilon(\sigma') E\left(\prod_{i=1}^n X_{\sigma(i),i} \prod_{j=1}^n X_{\sigma(j),j}\right).$$

Encore une fois, si il y a un seul X_{i_0,j_0} tout seul, en utilisant l'indépendance, cette variable aléatoire sera indépendante des autres et on aura un terme en $E(X_{i_0,j_0}) = 0$ donc le produit sera nul. Pour que l'espérance soit non nulle, il faut donc que $\sigma = \sigma'$. On a donc finalement :

$$V(\det(M)) = \sum_{\substack{\sigma \in S_n \\ \sigma \in S_n}} \varepsilon(\sigma)^2 E(\prod_{i=1}^n X_{\sigma(i),i})^2)$$
$$= \sum_{\substack{\sigma \in S_n \\ n!.}} 1 \times E(1)$$

3) Une fois la première colonne de M fixée (on a 2^n possibilités), chaque colonne doit être proportionnelle à celle-ci. Or, il n'y a par exemple pour la colonne C_2 que 2 colonnes colinéaires à C_1 (soit les mêmes coefficients, soit opposés). De même pour les autres colonnes. On a donc en tout $2^n \times 2 \times 2 \times \ldots \times 2 = 2^{2n-1}$ matrices de rang 1. Puisque au total on a 2^{n^2} matrices, on a donc la probabilité d'avoir une matrice de rang égale à $\frac{1}{2^{n^2-2n+1}} = \frac{1}{2^{(n-1)^2}}$.