Probabilités 3 – Vecteurs aléatoires

Correction de l'exercice 14 –

1. (a) Soit X et Y deux variables aléatoires admettant des variances. Alors

$$V(X + Y) = V(X) + V(Y) + 2\text{cov}(X, Y).$$

- (b) On obtient alors facilement par récurrence l'égalité souhaitée.

Soit, pour tout n dans $\mathbb{N}^*$, la propriété $\mathcal{P}(n)$: pour tout n -uplet $(T_1, \dots, T_n)$ de variables aléatoires admettant des variances, on a :

$$V\left(\sum_{k=1}^n T_k\right) = \sum_{k=1}^n V(T_k) + 2 \sum_{1 \leq i < j \leq n} \text{cov}(T_i, T_j).$$

La relation est triviale pour $n = 1$. Ainsi, $\mathcal{P}(1)$ est vrai. On peut aussi remarquer, même si cela n'est formellement pas nécessaire, que $\mathcal{P}(2)$ est la propriété énoncée dans la question 1.

Soit $n \in \mathbb{N}^*$ tel que $\mathcal{P}(n)$ soit vrai, et soit $(T_1, \dots, T_{n+1})$ un $(n+1)$ -uplet de variables aléatoires admettant une variance. Alors, d'après la question 1,

$$V(T_1 + \dots + T_{n+1}) = V(T_1 + \dots + T_n) + V(T_{n+1}) + 2\text{cov}(T_1 + \dots + T_n, T_{n+1}).$$

Ainsi, en utilisant l'hypothèse de récurrence, et en exploitant la bilinéarité de la covariance, on obtient :

$$V(T_1 + \dots + T_{n+1}) = \sum_{k=1}^n V(T_k) + 2 \sum_{1 \leq i < j \leq n} \text{cov}(T_i, T_j) + V(T_{n+1}) + 2 \sum_{i=1}^n \text{cov}(T_i, T_{n+1}).$$

La somme $\sum_{1 \leq i < j \leq n} \text{cov}(T_i, T_j)$ est indexée sur tous les couples $(i, j) \in \llbracket 1, n+1 \rrbracket^2$ tels que $i < j$ et

$j \neq n+1$, tandis que la somme $\sum_{i=1}^n \text{cov}(T_i, T_{n+1})$ regroupe tous les indices $(i, j) \in \llbracket 1, n+1 \rrbracket^2$ tels que $i < j$ et $j = n+1$. Ainsi, en regroupant les deux sommes, on a tous les indices $(i, j) \in \llbracket 1, n+1 \rrbracket^2$ tels que $i < j$. Par conséquent,

$$V(T_1 + \dots + T_{n+1}) = \sum_{k=1}^{n+1} V(T_k) + 2 \sum_{1 \leq i < j \leq n+1} \text{cov}(T_i, T_j).$$

Ainsi, la propriété $\mathcal{P}(n+1)$ est vérifiée.

Par conséquent, $\mathcal{P}(1)$ est vraie, et pour tout n dans $\mathbb{N}^*$, $\mathcal{P}(n)$ entraîne $\mathcal{P}(n+1)$. D'après le principe de récurrence, $\mathcal{P}(n)$ est vraie pour tout n dans $\mathbb{N}^*$.

2. La variable aléatoire $\sum_{k=1}^n X_k$ admet une espérance et une variance, en tant que somme finie de variables admettant une espérance et une variance. De plus, par linéarité de l'espérance, et d'après l'expression des espérances des variables de Bernoulli, on obtient :

$$E\left(\sum_{k=1}^n X_k\right) = \sum_{k=1}^n E(X_k) = \sum_{k=1}^n p^k = p \cdot \frac{1-p^n}{1-p},$$

puisque $p \neq 1$. De plus, les variables X_k , $k \in \llbracket 1, n \rrbracket$, étant mutuellement indépendantes,

$$V\left(\sum_{k=1}^n X_k\right) = \sum_{k=1}^n V(X_k) = \sum_{k=1}^n p^k(1-p^k) = \sum_{k=1}^n p^k - \sum_{k=1}^n p^{2k} = p \cdot \frac{1-p^n}{1-p} - p^2 \cdot \frac{1-p^{2n}}{1-p^2}.$$

Une mise au même dénominateur donne alors :

$$V\left(\sum_{k=1}^n X_k\right) = \frac{p(1+p)(1-p^n) - p^2(1-p^{2n})}{1-p^2} = \frac{p(1-p^n) + p^2(p^{2n} - p^n)}{1-p^2} = \frac{p(1-p^n)(1-p^{n+1})}{1-p^2}.$$

3. (a) Soit $n \in \mathbb{N}^*$. La variable aléatoire Y_n , produit de deux variables de Bernoulli, prend ses valeurs dans $\{0, 1\}$. De plus, l'événement $[Y_n = 1]$ est réalisé si et seulement si les événements $[X_n = 1]$ et $[X_{n+1} = 1]$ sont réalisés. Ainsi,

$$[Y_n = 1] = [X_n = 1] \cap [X_{n+1} = 1],$$

et, les X_i étant mutuellement indépendants, on obtient :

$$P(Y_n = 1) = P(X_n = 1)P(X_{n+1} = 1) = p^n p^{n+1} = p^{2n+1}.$$

Ainsi, Y_n suit une loi de Bernoulli de paramètre p^{2n+1} .

- (b) Par conséquent, pour tout $n \in \mathbb{N}^*$, $E(Y_n) = p^{2n+1}$, et ainsi, par linéarité de l'espérance :

$$E(Z_n) = \sum_{k=1}^n E(Y_k) = \sum_{k=1}^n p^{2k+1} = p^3 \sum_{k=0}^{n-1} p^{2k} = p^3 \cdot \frac{1 - p^{2n}}{1 - p^2}.$$

4. (a) Soit $n \in \mathbb{N}^*$. De la même manière, $Y_n Y_{n+1}$ prend ses valeurs dans $\{0, 1\}$, donc il va s'agir d'une variable de Bernoulli. Les variables Y_n et Y_{n+1} n'étant pas indépendantes, on ne peut pas raisonner de même ; on repasse donc aux variables X_i . On a :

$$Y_n Y_{n+1} = X_n X_{n+1}^2 X_{n+2},$$

ainsi, $Y_n Y_{n+1}$ prend la valeur 1 si et seulement si les trois variables X_n , X_{n+1} et X_{n+2} prennent la valeur 1 (sinon l'une au moins s'annule, et $Y_n Y_{n+1}$ prend la valeur 0). Ainsi

$$[Y_n Y_{n+1} = 1] = [X_n = 1] \cap [X_{n+1} = 1] \cap [X_{n+2} = 1],$$

et l'indépendance des X_i amène :

$$P(Y_n Y_{n+1} = 1) = P(X_n = 1)P(X_{n+1} = 1)P(X_{n+2} = 1) = p^n p^{n+1} p^{n+2} = p^{3n+3}.$$

Ainsi, $Y_n Y_{n+1}$ suit une loi de Bernoulli de paramètre p^{3n+3} . Son espérance est donc :

$$E(Y_n Y_{n+1}) = p^{3n+3}.$$

- (b) Puisque $\text{cov}(Y_k, Y_\ell) = \text{cov}(Y_\ell, Y_k)$, on peut se contenter des couples (k, ℓ) tels que $k \leq \ell$. Soit $(k, \ell) \in (\mathbb{N}^*)^2$ tel que $k \leq \ell$. Trois cas se produisent :

- si $k = \ell$, alors $\text{cov}(Y_k, Y_\ell) = \text{cov}(Y_k, Y_k) = V(Y_k) = p^{2n+1}(1 - p^{2n+1})$;
- si $k < \ell - 1$, alors Y_k ne dépendant que de X_k et X_{k+1} , et Y_ℓ ne dépendant que de X_ℓ et $X_{\ell+1}$, puisque les indices $k, k+1, \ell$ et $\ell+1$ sont deux à deux distincts et que les variables X_i sont mutuellement indépendantes, on en déduit que Y_k et Y_ℓ sont indépendantes, et par conséquent, $\text{cov}(Y_k, Y_\ell) = 0$;
- si $k = \ell - 1$, alors

$$\text{cov}(Y_k, Y_\ell) = \text{cov}(Y_k, Y_{k+1}) = E(Y_k Y_{k+1}) - E(Y_k)E(Y_{k+1}) = p^{3k+3} - p^{2k+1} p^{2k+3} = p^{3k+3}(1 - p^{k+1}).$$

- (c) D'après la question 1 et les covariances calculées dans la question précédente, on obtient donc :

$$\begin{aligned} V(Z_n) &= \sum_{k=1}^n V(Y_k) + 2 \sum_{1 \leq i < j \leq n} \text{cov}(Y_i, Y_j) \\ &= \sum_{k=1}^n p^{2k+1}(1 - p^{2k+1}) + 2 \sum_{i=1}^{n-1} \text{cov}(Y_i, Y_{i+1}) \\ &= p^3 \cdot \frac{1 - p^{2n}}{1 - p^2} - p^6 \cdot \frac{1 - p^{4n}}{1 - p^4} + 2 \sum_{i=1}^{n-1} (p^{3k+3} - p^{4k+4}) \\ &= p^3 \cdot \frac{1 - p^{2n}}{1 - p^2} - p^6 \cdot \frac{1 - p^{4n}}{1 - p^4} + 2p^6 \cdot \frac{1 - p^{3n-3}}{1 - p^3} - 2p^8 \cdot \frac{1 - p^{4n-4}}{1 - p^4}. \end{aligned}$$

On ne peut pas simplifier substantiellement cette expression.