Corrigé du Médian - SY01/A16

Exercice 1. (Dans cet exercice les questions sont indépendantes)

1. Soient $X_1, X_2, \ldots, X_n$ des variables aléatoires réelles discrètes. On suppose que pour tout $1 \leq i \leq n$, $P(X_i = i^2) = \frac{1}{i+2}$, $P(X_i = -i^2) = \frac{1}{i+2}$, $P(X_i = 0) = \frac{i}{i+2}$. On pose $S_n = X_1 + \ldots + X_n$. Déterminer la valeur de $E[S_n]$.
   Corrigé : $E[X_i] = i^2 P(X_i = i^2) - i^2 P(X_i = -i^2) + 0P(X_i = 0) = 0$. Ainsi $E[S_n] = \sum_{i=1}^n E[X_i] = 0$.
2. Une variable aléatoire à valeurs dans $\mathbb{N}$ peut-elle avoir une loi uniforme (c.a.d. $p_X(k) = P(X = k) = p$ où $p \in [0,1]$, $\forall k \in \mathbb{N}$) ? Justifiez soigneusement votre réponse.
   Corrigé : Si X est à valeurs dans $\mathbb{N}$ et de loi uniforme alors pour tout $k \in \mathbb{N}$ on a : $p_X(k) = P(X = k) = p$ où $p \in [0,1]$. Or p_X est une loi de probabilité si :
   $$1 = \sum_{k=0}^{+\infty} p_X(k) = \sum_{k=0}^{+\infty} p = \begin{cases} 0 & \text{si} \quad p = 0, \\ +\infty & \text{si} \quad p > 0. \end{cases}$$
   Ce qui est donc impossible !
3. Soient $X_1, X_2, \ldots, X_n$ des variables aléatoires indépendantes suivant la même loi :
   $$X_i(\Omega) = \{-1, 1\},\ i \in \{1, \ldots, n\} \text{ et } P(X_i = 1) = \theta,\ P(X_i = -1) = 1 - \theta,\ 0 < \theta < 1.$$
   Pour $k = 1, \ldots, n$, on définit la v.a. $Y_k = \Pi_{i=1}^k X_i$, de loi $P(Y_k = 1) = p_k$ et $P(Y_k = -1) = 1 - p_k$.
   Le but de cet exercice est de déterminer la valeur de p_k en fonction de θ.
   (a) Pour $k = 1, \ldots, n$, déterminer $E[X_k]$ en fonction de θ et $E[Y_k]$ en fonction de p_k.
   Corrigé : $E[X_k] = 2\theta - 1$ et $E[Y_k] = p_k - (1 - p_k) = 2p_k - 1$.
   (b) Pour $k = 1, \ldots, n$, déterminer $E[Y_k]$ en fonction de $E[X_1]$ et de k.
   Corrigé : $E[Y_k] = E\left[\Pi_{i=1}^k X_i\right] = \Pi_{i=1}^k E[X_i] = (E[X_1])^k$.
   (c) A l'aide des deux questions précédentes, déterminer la valeur de p_k en fonction de θ et de k.
   Corrigé : On doit avoir $2p_k - 1 = (2\theta - 1)^k$, d'où $p_k = \frac{1}{2}\left(1 + (2\theta - 1)^k\right)$.

Exercice 2. Pour améliorer la sûreté de fonctionnement d'un serveur informatique, on envisage d'introduire de la redondance, c'est-à-dire d'avoir plusieurs exemplaires des composants importants. On peut réaliser les opérations suivantes :

- on utilise trois alimentations de 300 Watts chacune : le serveur peut continuer à fonctionner avec une alimentation en panne car il consomme au maximum 500 Watts.
- on place les quatre disques durs en configuration RAID 5 : le serveur peut continuer à fonctionner avec un disque dur en panne.

On suppose que la probabilité de panne d'une alimentation est p et que celle d'une panne de disque dur est q. On suppose en outre que tous les composants sont indépendants.

1. Soit un serveur avec alimentations redondantes : calculer la probabilité de panne du serveur en supposant qu'aucun autre composant que les alimentations ne peut tomber en panne.
   Corrigé : L'état de l'alimentation i est représenté par une variable X_i distribuée selon la loi $B(p)$ (loi de Bernoulli) qui vaut 1 si l'alimentation est en panne. Le nombre d'alimentation en panne, est la somme $X_1 + X_2 + X_3$ de loi $B(3, p)$. Le serveur est en panne si deux au moins des alimentations sont en panne. On a donc
   $$P(A) = P(X_1 + X_2 + X_3 = 2) + P(X_1 + X_2 + X_3 = 3) = C_3^2 p^2(1-p) + C_3^3 p^3 = p^2(3 - 2p).$$

2. Soit un serveur avec disques durs RAID 5 : calculer la probabilité de panne du serveur en supposant qu'aucun autre composant que les disques dur ne peut tomber en panne.
   Corrigé : On modélise les disques durs par des variables Y_j de loi de Bernoulli $B(q)$, indépendantes. Le nombre de disques en panne est alors une variable de loi Binomiale $B(4,q)$. Le serveur est en panne si au moins deux des disques sont en panne,

$$P(D) = P\left(\sum_{j=1}^{4} Y_j \geq 2\right) = P\left(\sum_{j=1}^{4} Y_j = 2\right) + P\left(\sum_{j=1}^{4} Y_j = 3\right) + P\left(\sum_{j=1}^{4} Y_j = 4\right)$$

$$= C_4^2 q^2(1-q)^2 + C_4^3 q^3(1-q) + C_4^4 q^4 = q^2(6(1-q)^2 + 4q(1-q) + q^2) = q^2(3q^2 - 8q + 6).$$

3. Si $p = q$, quelle solution de redondance est la plus intéressante ?
   Corrigé : Lorsque $p = q$. On a $P(A) - P(D) = p^2(3-2p) - p^2(3p^2 - 8p + 6) = p^2(-3p^2 + 6p - 3)$. Le signe de cette différence est celui de $f(p) = -3p^2 + 6p - 3$. Or $f'(p) = 6(1-p)$ s'annule en $p = 1$ et $f'(p) > 0$ sur $[0,1]$, donc $f(p)$ est croissante sur cet intervalle. Or, $f(1) = 0$ et $f(0) = -3$. Donc, sur l'intervalle $[0,1]$, $f(p) \leq 0$. La probabilité de défaillance des alimentations est donc toujours inférieure à celle des disques durs.

Exercice 3. Un gardien de nuit doit ouvrir l'une des portes à contrôler pendant sa tournée, dans le noir. Il possède un trousseau de 10 clés d'allures semblables, mais une seule peut ouvrir la porte en question. L'essai des clés se fait au hasard. Le gardien dispose de deux méthodes :

- Méthode A, qui consiste à essayer chaque clé uniformément parmi les clés restantes.
- Méthode B, qui consiste à essayer une clé après avoir agité le trousseau (c'est à dire il fait un choix uniforme entre toutes les clés).

1. On appelle X_A la variable aléatoire qui désigne le nombre des clés essayées (y compris celle qui donne satisfaction) par la Méthode A. Déterminer la loi de probabilité de X_A.
2. De même, on appelle X_B la variable aléatoire analogue que l'on obtient par la méthode B. Déterminer la loi de probabilité de X_B.
3. Quelle est la probabilité d'essayer plus de 8 clés par la méthode A ? par la méthode B ?
4. Le gardien utilise la méthode A lorsqu'il est à jeûn et la méthode B lorsqu'il est ivre. Un cambrioleur caché à l'intérieur sait que le gardien est ivre un jour sur trois. Quelle est la probabilité conditionnelle pour que le gardien soit ivre, sachant que les 8 premiers essais ont échoués. On notera : **A** l'événement "le gardien utilise la méthode A", **B** l'événement "le gardien utilise la méthode B" et **H** l'événement "8 essais ont échoué".

Corrigé :

1. X_A est une variable aléatoire dont les valeurs possibles sont les entiers de 1 à 10. Ces dix valeurs sont équiprobables, $P(X_A = n) = \frac{1}{10}$, pour $n = 1, 2 \ldots, 10$. En effet, $P(X_A = 1) = \frac{1}{10}$ (un cas favorable parmi dix possibles). $P(X_A = 2)$ est la probabilité conditionnelle de succés au deuxième essai sachant que le premier a échoué, $P(X_A = 2) = \frac{1}{9}\frac{9}{10} = \frac{1}{10}$. De même, la probabilité pour que le n-ième essai échoue sachant que les précédents ont échoué est $\frac{10-n}{10-n+1}$. Donc la probabilité pour que le n-ième essai réussise sachant que les précédents ont échoué est $\frac{1}{10-n+1}$. D'où $P(X_A = n) = \frac{9}{10}\frac{8}{9}\ldots\frac{10-n+1}{10-n+2}\frac{1}{10-n+1} = \frac{1}{10}$.
2. X_B est une variable aléatoire dont les valeurs possibles sont les entiers strictement positifs. Quel que soit l'essai, sa probabilité de réussite est $\frac{1}{10}$, et celle d'échec est $\frac{9}{10}$. D'où

$$P(X_B = n) = (\frac{9}{10})^{n-1}\frac{1}{10}.$$

3. La probabilité d'essayer plus de 8 clés est la probabilité d'échouer les 8 premières fois, soit $\frac{2}{10}$ par la méthode A et $\left(\frac{9}{10}\right)^8$ (soit $\sum_{k=9}^{\infty}\left(\frac{9}{10}\right)^{k-1}\frac{1}{10}$) par la méthode B.

4. Soit A :"le gardien utilise la méthode A", B :"le gardien utilise la méthode B" et H :"8 essais ont échoué" . Nous avons $P(A) = \frac{2}{3}$; $P(B) = \frac{1}{3}$; $P(H|A) = 0.2$ et $P(H|B) = (\frac{9}{10})^8$. On cherche $P(B|H)$. Par la formule de Bayes on a donc

$$P(B|H) = \frac{\frac{1}{3}(\frac{9}{10})^8}{\frac{2}{3}\frac{2}{10} + (\frac{9}{10})^8\frac{1}{3}}$$

Exercice 4. Soit $p \in]0,1[$ un nombre réel. Soient X et Y deux variables aléatoires indépendantes qui suivent toutes deux la loi géométrique de paramètre $1-p$.

1. Calculer la loi de $X+Y$, en précisant l'ensemble des valeurs possibles de la variable aléatoire $X+Y$.
   Corrigé : Puisque X et Y sont à valeurs dans $\mathbb{N}^*$, la variable aléatoire $X+Y$ est à valeurs dans $\{2,3,\ldots\}$. Pour $n \in \{2,3,\ldots\}$, on a

$$P(X+Y=n) = \sum_{k=1}^{n-1} P(X=k, Y=n-k) = \sum_{k=1}^{n-1} P(X=k)P(Y=n-k)$$

   où la dernière égalité est due à l'indépendance. D'où

$$P(X+Y=n) = \sum_{k=1}^{n-1}(1-p)p^{k-1}(1-p)p^{n-k-1} = (1-p)^2\sum_{k=1}^{n-1} p^{n-2} = (1-p)^2(n-1)p^{n-2}.$$

   ainsi, $\forall n \in \{2,3,\ldots,\}$ $P(X+Y=n) = (1-p)^2(n-1)p^{n-2}$.
2. Déterminer l'espérance de $X+Y$.
   Corrigé : Puisque X et Y admettent un moment d'ordre 1, il en est de même de $X+Y$ et, en utilisant l'égalité $E[X] = E[Y] = \frac{1}{1-p}$, on trouve $E[X+Y] = E[X] + E[Y] = \frac{2}{1-p}$.
3. Décrire l'événement $\{X=Y\}$ à l'aide des événements $\{X=n, Y=n\}$ pour $n \geq 1$, puis calculer $P(X=Y)$.
   Corrigé : L'événement $\{X=Y\} = \cup_{n\geq 1}\{X=n, Y=n\}$. On a donc, en utilisant la σ-additivité de P puis l'indépendance de X et de Y,

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

4. Expliquer pourquoi $P(X>Y) = P(Y>X)$. En déduire la valeur de $P(X>Y)$ à l'aide de la question 3.
   Corrigé : Puisque X et Y sont indépendantes et de même loi, on a l'égalité $P(X>Y) = P(Y>X)$. Par ailleurs, $1 = P(X=Y) + P(X>Y) + P(Y>X) = P(X=Y) + 2P(X>Y)$. Il s'ensuit que $P(X>Y) = \frac{1}{2}\left(1 - \frac{1-p}{1+p}\right) = \frac{p}{1+p}$.
5. On définit la variable aléatoire Z par $Z = 1_{\{\text{ impair }\}}(X)$ (fonction indicatrice de l'ensemble des nombres impairs). Déterminer la loi de Z.
   Corrigé : La v.a. Z prend ses valeurs dans $\{0,1\}$. L'événement $\{Z=0\}$ est l'événement "X prends les valeurs paires" et l'événement $\{Z=1\}$ est l'événement "X prends les valeurs impaires". On a donc

$$P(Z=1) = \sum_{n=1}^{\infty} P(X=2n-1) = (1-p)\sum_{n=1}^{\infty} p^{2n-2} = \frac{1-p}{1-p^2} = \frac{1}{1+p}.$$

   et donc $P(Z=0) = \frac{p}{1+p}$. La variable aléatoire Z suit donc la loi de Bernoulli de paramètre $\frac{1}{1+p}$.