Corrigé Médian - SY01/A17, 2 heures, 2 pages, 5 exercices.

Exercice 1. Dans cet exercice les questions sont indépendantes.

1. Considérons la fonction $F : \mathbb{R} \to [0, 1]$ définie par

$$F(t) = \begin{cases} 0, \text{ si } t < 1, \\ 0.25, \text{ si } 1 \leq t < 5, \\ 1, \text{ si } \geq 5. \end{cases}$$

Donner la loi de la variable aléatoire X dont F est la fonction de répartition.
Corrigé La fonction F étant constante sur $]-\infty, 1[$, $[1, 5[$ et $[5, +\infty[$, la loi de X est $P(X = 1) = F(1) - F(1^-) = \frac{1}{4}$ et $P(X = 5) = F(5) - F(5^-) = \frac{3}{4}$.

2. Soit $n \in \mathbb{N}$. Pour quelle valeur de c, $P(X = n) = \frac{c3^{n-1}}{3^{n+1}-2^{n+1}} \sum_{k=0}^{n} \left(\frac{2}{3}\right)^k$ est-ce une probabilité ?
Corrigé. Nous avons, c doit être positive strictement et

$$P(X = n) = \frac{c3^{n-1}}{3^{n+1} - 2^{n+1}} \sum_{k=0}^{n} \left(\frac{2}{3}\right)^k = \frac{c3^{n-1}}{3^{n+1} - 2^{n+1}} \frac{1 - \left(\frac{2}{3}\right)^{n+1}}{1 - \frac{2}{3}} = \frac{c3^{n-1}}{3^{n+1} - 2^{n+1}} \frac{3^{n+1} - 2^{n+1}}{3^{n+1}(1 - \frac{2}{3})} = \frac{c}{3}.$$

Donc $P(X = n)$ est une probabilité si et seulement si $0 < c \leq 3$.

Exercice 2. Une urne contient N_b boules blanches et N_n boules noires. Posons $N = N_b + N_n$. On tire r boules avec remise dans une urne, il y a alors N^r tirages possibles.
Soit A_k l'événement "on a tiré exactement k" boules blanches.

1. Calculer $Card(A_k)$.
Corrigé : L'événement A_k est réalisé lorsque l'issue est constituée de k boules blanches et $r-k$ boules noires. Il y a C_r^k façons de choisir la position des boules blanches, la position des boules noires est ensuite fixée. Ainsi : $Card(A_k) = C_r^k N_b^k N_n^{r-k}$

2. Déterminer la probabilité de A_k.
Corrigé : $P(A_k) = \frac{C_r^k N_b^k N_n^{r-k}}{N^r} = C_r^k \left(\frac{N_b}{N}\right)^k \left(\frac{N_n}{N}\right)^{r-k}$.

3. Soit X la variable aléatoire égale au nombre de boules blanches tirées. Déterminer la loi de X.
Corrigé : La variable aléatoire X a pour valeurs $\{0, 1, \ldots, r\}$ et pour tout $k \in \{0, 1, \ldots, r\}$, l'événement $\{X = k\} = A_k$, ainsi la loi de X est donnée par $P(X = k) = P(A_k) = C_r^k \left(\frac{N_b}{N}\right)^k \left(\frac{N_n}{N}\right)^{r-k}$.

4. Reconnaître la loi de X.
Corrigé : Il s'agit de la loi Binomiale des paramètres r et $\frac{N_b}{N}$.

Exercice 3. Un joueur tourne une roue. A chaque essai, il a une probabilité $1 - p$ de tomber sur une case lui rapportant 1 euro et p de tomber sur la case « banqueroute » qui lui fera tout perdre et l'éliminera du jeu. Il débute avec une somme initiale égale à 0 euro et souhaite atteindre un gain G_0, après quoi, il quittera le jeu, s'il n'a pas déjà été contraint auparavant. Les résultats successifs sont supposés indépendants. Ceci sera formalisé par un espace probabilisé $(\Omega, \mathcal{F}, P)$ et une suite d'expériences de Bernoulli indépendantes de probabilité p. Le gain final est noté G.

1. Sur un petit graphique, représenter rapidement les deux situations typiques pouvant se produire, et les valeurs possibles de G.
Corrigé. Les deux situations typiques sont l'atteinte du gain avant la banqueroute et la banqueroute avant l'atteinte du gain. Ceci peut se visualiser facilement (voir le cours).

2. Quelle est la probabilité d'obtenir le gain G_0 ?
Corrigé. La variable aléatoire G ne peut prendre que deux valeurs : $\{0, G_0\}$. Pour que le gain soit égal à G_0, il faut gagner les G_0 premiers essais. Ainsi, si on note $A_i = \{X_i = 1\}$, on a

$$P(G = G_0) = P(\bar{A}_1 \cap \ldots \cap \bar{A}_{G_0}) = P(\bar{A}_1) \times \ldots \times P(\bar{A}_{G_0}) = (1 - p)^{G_0}.$$

3. En déduire l'espérance du gain final $E[G]$. Interpréter le résultat lorsque $1 > p > 0$ et G_0 tend vers l'infini.
   Corrigé. L'espérance de G est bien définie puisque c'est une variable aléatoire bornée, et on a $E\left[G\right] = 0 \times P(G = 0) + G_0 \times P(G = G_0) = G_0 \times (1-p)^{G_0}$. On constate que si $1 > p > 0$, l'espérance tend vers 0 quand le gain G_0 tend vers l'infini.

Exercice 4. Dans un mot binaire, c'est-à-dire une suite de 0 ou 1, on appelle k-séquence toute suite de k 1 consécutifs n'étant ni précédée ni suivie de 1. Par exemple, le mot (110011010001110) possède une 1-séquence, deux 2-séquences et une 3-séquence.
On considère à présent un mot binaire aléatoire de longueur n : $(X_1, X_2, \ldots, X_n)$, les X_i étant des variables indépendantes et de même loi de Bernoulli $B(p)$ où $p \in [0,1]$ est fixé. On note alors $A_k = \{$le mot possède au moins une k-séquence$\}$.

1. Que vaut $P(A_n)$?
   Corrigé : Si un mot de longueur n contient une n-séquence, c'est que ce mot ne contient que des 1. Donc par indépendance on a $P(A_n) = P(X_1 = 1, X_2 = 1, \ldots, X_n = 1) = p^n$.
2. On considère le cas $n = 4$ et $k = 2$. Calculer $P(A_2)$ en fonction de p.
   Indication : Combien y a-t-il de 2-séquences possibles dans un mot de longueur 4 ?
   Corrigé : On ne peut avoir qu'une seule 2-séquence dans un mot de longueur 4. En effet si le mot comporte une 2-séquence, elle est
   – soit placée au début, et dans ce cas le mot est (1100) ou (1101),
   – soit placée au milieu et dans ce cas le mot est (0110),
   – soit placée à la fin et dans ce cas le mot est (0011) ou (1011).
   Donc dans tous les cas on voit qu'il n'y a pas d'autre 2-séquence.
   Nous avons donc $A_2 = S_1 \cup S_2 \cup S_3$; où $S_i = \{$la 2-séquence commence en position $i\}$. Ces trois événements sont incompatibles, donc $P(A_2) = P(S_1) + P(S_2) + P(S_3) = P(X_1 = 1, X_2 = 1, X_3 = 0) + P(X_1 = 0, X_2 = 1, X_3 = 1, X_4 = 0) + P(X_2 = 0, X_3 = 1, X_4 = 1)$.
   Les X_i étant indépendantes, on obtient $P(A_2) = (1-p)(3-p)p^2$.
3. On suppose n quelconque et $k \geq n/2$. Écrire une formule généralisant celle de la question précédente et en déduire la valeur de $P(A_k)$ en fonction de p, n et k. La formule est-elle aussi vraie lorsque $k < n/2$?
   Corrigé : Lorsque $k \geq n/2$, comme dans l'exemple précédent, on ne peut avoir qu'une seule k-séquence dans le mot, car pour qu'il y ait deux k-séquences ou plus, il faudrait $2k$ 1 plus un 0 pour les séparer, ce qui est impossible car $2k + 1 > n$.
   Pour calculer $P(A_k)$, on considère toutes les positions possibles pour la k-séquence dans le mot. On a ainsi, en définissant $S_i =$"la k-séquence commence en position i", $A_k = S_1 \cup S_2 \cup \ldots \cup S_{n-k+1}$, et ces événements étant incompatibles, on a $P(A_k) = P(S_1) + \ldots + P(S_{n-k+1})$. Pour le calcul des $P(S_i)$ on regarde d'abord le cas où la séquence est en début ou en fin ($i = 1$ ou $i = n - k + 1$), ainsi $P(S_1) = P(X_1 = 1, X_2 = 1, \ldots, X_k = 1, X_{k+1} = 0) = P(X_1 = 1)P(X_2 = 1) \ldots P(X_k = 1)P(X_{k+1} = 0) = p^k(1-p)$. De même on aura $P(S_{n-k+1}) = p^k(1-p)$ (probabilité que le mot se termine par un 0 suivi de k 1). Dans tous les autres cas ($2 \leq i \leq n-k$) on a
   $$P(S_i) = P(X_{i-1} = 0, X_i = 1, \ldots, X_{i+k-1} = 1, X_{i+k} = 0) = (1-p)^2 p^k.$$
   Ainsi $P(A_k) = 2p^k(1-p) + (n-k-1)(1-p)^2 p^k = p^k(1-p)[2 + (n-k-1)(1-p)]$. Dans le cas $k < n/2$, le calcul n'est plus valable car il peut y avoir plusieurs k-séquences dans le mot, et les événements S_i ne sont plus incompatibles.

Exercice 5. Un message doit être transmis d'un point à un autre à travers N canaux successifs. Ce message peut prendre deux valeurs, 0 ou 1. Durant le passage par un canal, le message a la probabilité $p \in]0,1[$ d'être bruité, c'est-à-dire d'être transformé en son contraire, et $(1-p)$ d'être transmis fidèlement. Les canaux se comportent indépendamment les uns des autres.

1. Notons I_n l'événement : "en sortie du n-ème canal, le message est le même que celui transmis initialement." Exprimer $P(I_{n+1})$ en fonction de $P(I_n)$ et de p.
   Corrigé. Pour que l'événement I_{n+1} ait lieu, de deux choses l'une : ou bien I_n était réalisé et le message a été bien transmis dans le $(n+1)$-ème canal, ou bien $\bar{I}_n$ était réalisé et le message a été mal transmis dans le $(n+1)$-ème canal. Par la formule des probabilités totales on a $P(I_{n+1}) = P(I_{n+1}|I_n)P(I_n) + P(I_{n+1}|\bar{I}_n)P(\bar{I}_n) = P(I_{n+1}) = (1-p)P(I_n) + p(1 - P(I_n))$.
2. En notant $p_n = P(I_n)$, donner une relation de récurrence entre p_{n+1} et p_n. Que vaut p_1 ?
   Corrigé. On a donc la relation de récurrence :$p_{n+1} = (1-p)p_n + p(1-p_n) = (1-2p)p_n + p$, avec $p_1 = 1-p$, I.e. probabilité que le message n'ait pas été bruité dans le premier canal.
3. On considère une suite $(u_n)_{n\geq 1}$ vérifiant la relation de récurrence : $u_{n+1} = (1-2p)u_n + p$. Vérifier que la suite $(v_n)_{n\geq 1}$, définie par $v_n = u_n - 1/2$ pour tout $n \geq 1$, est géométrique. En déduire v_n en fonction de p et v_1.
   Corrigé. On écrit : $v_{n+1} = u_{n+1} - 1/2 = (1-2p)u_n + p - 1/2$ et en remplaçant u_n par $v_n + 1/2$, il vient $v_{n+1} = (1-2p)v_n$, donc la suite $(v_n)_{n\geq 1}$ est géométrique de raison $(1-2p)$. On en déduit : $\forall n \in \{1, \ldots, N\} \quad v_n = (1-2p)^{n-1}v_1$.
4. En déduire p_n en fonction de p pour tout $n \in \{1, \ldots, N\}$.
   Corrigé. On a la même relation pour p_n que pour $u_n = v_n + 1/2$ et puisque $p_1 = (1-p)$, on en déduit que : $\forall n \in \{1, \ldots, N\} \quad p_n = \frac{1}{2} + (1-p)\,(1-2p)^{n-1}$.
5. Que vaut $\lim_{N\to+\infty} p_N$?
   Indication : distinguer suivant les valeurs de p.
   Corrigé. Si : $0 < p < 1$ on est dans le cas d'un bruitage aléatoire. On a $\lim_{N\to+\infty} p_N = \frac{1}{2}$, c.a.d. dès que le nombre de canaux devient grand, on est incapable de retrouver le message initial de façon fiable. On peut aussi remarquer que : Si $p = 0$: la transmission est fiable et on retrouve bien sûr $p_N = 1$ pour tout N. Si $p = 1$: chaque passage dans un canal change de façon certaine le message, donc p_n dépend de la parité du nombre de canaux : $p_{2N} = 1$ et $p_{2N+1} = 0$.