SY01 / A21 - TEST - CORRIGÉ

(Durée : 45min - fiche recto-verso A4 autorisée)

Les réponses doivent être justifiées soigneusement.

Exercice I *(8 points)*

Soit une pièce de monnaie telle qu'à chaque lancer la probabilité d'obtenir *pile* est égale à p.
On réalise une suite de parties de *pile* ou *face* indépendantes jusqu'à la première apparition d'un *pile* suivi d'un *face*, ou d'un *face* suivi d'un *pile*.

1. Décrire un ensemble fondamental associé à cette expérience.

2. Calculer la probabilité que le nombre total de lancers soit égal à n, pour $n \in \mathbb{N}, n \geq 2$.

3. Calculer la probabilité que le nombre de lancers soit pair.

Ensuite on réalise une deuxième expérience aléatoire où l'on lance la même pièce jusqu'à la prémière apparition de *pile* deux fois consécutives.

4. Calculer la probabilité de devoir lancer la pièce 4 fois.

Solution:

1. On peut considérer

$$\begin{aligned}\Omega &= \{PF, FP, PPF, FFP, PPPF, FFFP, ...\}\\ &= \bigcup_{n\geq 2} \{\omega = (\omega_1, \omega_2, \ldots, \omega_n) : \omega_i = F, i = 1, \ldots, n-1, \omega_n = P \\ &\qquad \text{ou } \omega_i = P, i = 1, \ldots, n-1, \omega_n = F\}\end{aligned}$$

2. Soit $n \in \mathbb{N}, n \geq 2$, l'évenement "n lancers " se produit s'il y a $n-1$ premiers lancers qui donnent tous pile et le n-ième est face ou inversement si les $n-1$ premiers lancers sont tous face et le n-ième est pile. Or la probabilité d'obtenir $n-1$ pile puis un face est par indépendance des lancers $p^{n-1}(1-p)$, et d'obtenir $n-1$ face puis un pile est $(1-p)^{n-1}p$, alors

$$\mathbb{P}(n \text{ lancers}) = p^{n-1}(1-p) + (1-p)^{n-1}p$$

On peut vérifier aisément que $\sum_{n\geq 2} \mathbb{P}(n \text{ lancers}) = 1$.

3. Le nombre de lancers jusqu'à obtenir pf ou fp peut être pair ou impair, alors nous avons

$$
\begin{aligned}
\mathbb{P}(\text{nombre de lancers est pair}) &= \mathbb{P}\left(\bigcup_{n\in\mathbb{N},n\geq 1}\{2n \text{ lancers}\}\right) \\
&= \sum_{n\in\mathbb{N},n\geq 1} \mathbb{P}\left(\{2n \text{ lancers}\}\right) \\
&= \sum_{n\in\mathbb{N},n\geq 1} p^{2n-1}(1-p) + (1-p)^{2n-1}p \\
&= p(1-p) \sum_{n\in\mathbb{N},n\geq 0} p^{2n} + (1-p)^{2n} \\
&= p(1-p)\left[\frac{1}{1-p^2} + \frac{1}{1-(1-p)^2}\right] \\
&= \frac{p}{1+p} + \frac{1-p}{2-p}.
\end{aligned}
$$

Par example pour $p = 1/2$ cette probabilité vaut $2/3$, et on peut vérifier que de manière générale elle est toujours supérieure à $1/2$. Il n'y a pas de symétrie entre pairs et impairs dans ce modèle.

4. Maintenant $\Omega = \{pp, fpp, ffpp, pfpp, fffpp, fpfpp, pffpp, ...\}$, alors $\mathbb{P}(4 \text{ lancers}) = \mathbb{P}(\omega \in \{ffpp, pfpp\}) = (1-p)^2p^2 + (1-p)p^3 = (1-p)p^2$

Exercice II *(6 points)*

Les résultats d'un test de détection d'une maladie doivent être interprétés en prenant en compte ses performances . Notons I^+ (resp. I^-) le fait qu'un individu soit infecté (resp. non infecté), et T^+ (resp. T^-) le fait qu'un individu soit positif au test (resp. négatif). Les performances intrinsèques du test sont

- La **sensibilité** $S_e = \mathbb{P}\left(T^+ \mid I^+\right)$: la probabilité que le test soit positif sachant que la personne testée est infectée ;
- La **spécificité** $S_p = \mathbb{P}\left(T^- \mid I^-\right)$: la probabilité que le test soit négatif sachant que la personne testée n'est pas infectée.

1. La **valeur prédictive positive** (VPP) d'un test est la probabilité qu'un individu choisit au hasard dans la population soit vraiment infecté sachant son test est positif. Calculer la VPP d'un test pour une population avec une proportion d'individus infectés égale à r, en fonction de S_e, S_p et r.

2. Une étude réalisée dans la ville A montrait que seulement **2%** de la population était infecté. Pour un test avec une **sensibilité de 98%** et une **spécificité de 98%**, calculer la probabilité qu'un individu choisit au hasard dans la ville A soit vraiment infecté lorsque il reçoit un résultat positif de ce test. Ce test remplit-il son rôle dans cette situation dans la ville A ?

Solution:

1. La probabilité d'être infecté lorsque le test est positif (VPP) peut être calculée en fonction de S_e, S_p et p en utilisant le théorème de Bayes et la loi de probabilité totale comme suit

$$\begin{aligned}
\mathbb{P}\left(I^+ \mid T^+\right) = \frac{\mathbb{P}(T^+ \mid I^+)\,\mathbb{P}(I^+)}{\mathbb{P}(T^+)} &= \frac{\mathbb{P}(T^+ \mid I^+)\,\mathbb{P}(I^+)}{\mathbb{P}(T^+ \cap I^+) + \mathbb{P}(T^+ \cap I^-)} \\
&= \frac{\mathbb{P}(T^+ \mid I^+)\,\mathbb{P}(I^+)}{\mathbb{P}(T^+ \mid I^+)\,\mathbb{P}(I^+) + \mathbb{P}(T^+ \mid I^-)\,\mathbb{P}(I^-)} \\
&= \frac{\mathbb{P}(T^+ \mid I^+)\,\mathbb{P}(I^+)}{\mathbb{P}(T^+ \mid I^+)\,\mathbb{P}(I^+) + (1 - \mathbb{P}(T^- \mid I^-))\,\mathbb{P}(I^-)}
\end{aligned}$$

et avec les notations de l'enoncé,

$$= \frac{S_e\, r}{S_e\, r + (1 - S_p)(1 - r)}.$$

2. En remplaçant les valeurs données dans la formule précédente, nous obtenons que dans la ville A

$$\mathbb{P}\left(I^+ \mid T^+\right) = \frac{0.98 \cdot 0.02}{0.98 \cdot 0.02 + 0.02 \cdot 0.98} = 0.5.$$

Cela veut dire que sachant qu'un individu reçoit un test positif, l'individu a autant de chances d'être infecté que de ne pas l'être, donc le test détecte trop d'individus positifs et ne remplit pas son rôle en ce sens. Malgré une sensibilité et une spécificité qui semleraient acceptables, quand la proportion d'individus infectés est trop faible, le taux de faux positifs peut être assez élevé.

Exercice III *(6 points)*

1. On considère toutes les permutations des 26 lettres de l'alphabet. Si l'on choisit une de ces permutations au hasard, quelle est la probabilité qu'elle ne contienne aucun des trois mots suivants : *fish*, *rat*, *bird*.

Solution:

1. Ici l'univers Ω est l'ensemble contenant toutes les permutations des 26 lettres de l'alphabet, donc le total de cas possibles est $|\Omega| = 26!$. Nous sommes dans un modèle d'équiprobabilité, et l'ensemble $\mathcal{F} = \mathcal{P}(\Omega)$, c'est à dire que toute permutation de l'alphabet a la même probabilité $1/|\Omega|$ d'être choisie.

Il faut compter les cas favorables, c'est à dire le nombre de permutations ne contennant aucun des trois mots. Pour cela il est plus adapté de passer par le complémentaire et compter le nombre de permutations contenant au moins l'un des trois mots. On considère les évènements suivants dans Ω :

- $A = \{$permutations ne contenant aucun des 3 mots$\}$
- $F = \{$permutations contenant le mot *fish*$\}$

- $R = \{$permutations contenant le mot *rat*$\}$
- $B = \{$permutations contenant le mot *bird*$\}$

On peut remarquer que $\overline{A} = F \cup R \cup B$.

1. On va compter le nombre de cas favorables dans chacun de ces évènements,
   - Le nombre de permutations avec le mot *fish* peut être compté en considérant ces 4 lettres comme une seule, et donc il nous resteront 26 - 4 autres lettres à placer, donc $|F| = (26 - 4 + 1)! = 23!$ permutations contenant ce mot,
   - Pour le mot *rat* nous avons : $|R| = (26 - 3 + 1)! = 24!$ cas favorables
   - Pour le mot *bird* nous avons : $|B| = (26 - 4 + 1)! = 23!$ cas favorables
2. Maintenant on s'intéresse aux permutations contenant 2 parmi ces trois mots
   - *fish* et *rat* : $|F \cap R| = (26 - 4 - 3 + 2)! = 21!$
   - *fish* et *bird* : $F \cap B = \emptyset$, alors $|F \cap B| = 0$
   - *rat* et *bird* : $R \cap B = \emptyset$, alors $|F \cap B| = 0$
3. Finalement, il n'y a pas des permutations contenant les 3 mots, donc $|F \cap R \cap B| = 0$

En appliquant le principe de Poincaré, nous avons

$$\begin{aligned}
\mathbb{P}(A) &= 1 - \mathbb{P}(\overline{A}) = 1 - \mathbb{P}(F \cup R \cup B) \\
&= 1 - [\mathbb{P}(F) + \mathbb{P}(R) + \mathbb{P}(B) - \mathbb{P}(F \cap R) - \mathbb{P}(F \cap B) - \mathbb{P}(R \cap B) + \mathbb{P}(F \cap R \cap B)] \\
&= 1 - \left[\frac{23! + 24! + 23! - 21!}{26!}\right].
\end{aligned}$$