

Probabilités, Statistiques, Combinatoire : Devoir maison

À rendre pour le dimanche 12 avril

Les différents exercices sont indépendants. Il est important de justifier soigneusement toutes vos réponses ; chaque fois que vous calculez une probabilité ou une probabilité conditionnelle, ou que vous interprétez l'énoncé d'un exercice en termes de probabilités ou de probabilités conditionnelles, faites-le le plus explicitement possible.

Rappel de notation : $\mathbb{P}_B(A)$ désigne la probabilité conditionnelle parfois notée $\mathbb{P}(A|B)$.

Exercice 1

Une société de voyages doit faire vérifier tous les bagages abandonnés afin de s'assurer qu'ils ne sont pas piégés, mais cherche à diminuer le coût de ces vérifications. Elle a développé un système de test automatique qui semble extrêmement prometteur : à partir d'une simple photo prise par les caméras de surveillance, celui-ci prend une décision qui est soit "colis suspect" soit "colis non suspect". Les tests du système semblent indiquer que celui-ci déclare suspects 90% des colis piégés, et seulement 10% des colis non piégés (fausses alertes). Par ailleurs, on estime que seulement 0.1% des colis abandonnés sont effectivement piégés.

On modélise la situation par les deux événements suivants :

- P : "Le bagage abandonné est piégé"
- S : "Le système déclare le bagage suspect"

1. Interprétez les données de l'énoncé comme des probabilités ou probabilités conditionnelles faisant intervenir les événements S et P et leurs complémentaires.
2. Que valent $\mathbb{P}_S(P)$ et $\mathbb{P}_{\bar{S}}(P)$? Pensez-vous qu'un tel système devrait être utilisé, et pourquoi ?

Exercice 2

Un fabricant de lampes torches essaie de rendre ses produits plus robustes en doublant le nombre d'ampoules par lampe. Chaque lampe torche est alors composée d'une pile et de deux ampoules. La lampe tombe en panne si la pile est défaillante, ou si les deux ampoules sont défaillantes.

Les ampoules sont d'assez bonne qualité, chacune a seulement 1% de chances d'être défaillante. En revanche la pile est assez fragile, elle a 5% de chances d'être défaillante.

On modélise alors le problème avec les événements suivants :

- L_1 : "La première ampoule est défaillante"
- L_2 : "La deuxième ampoule est défaillante"
- P : "La pile est est défaillante"
- T : "La lampe torche est en panne"

1. Soient A et B deux événements, rappelez (sans preuve) comment calculer $\mathbb{P}(A \cup B)$ si on ne suppose pas les deux événements disjoints.
2. Quelles sont les probabilités décrites dans l'énoncé ?

3. D'après l'énoncé, comment s'exprime l'événement T en fonction de L_1 , L_2 et P ?
4. En supposant que les événements L_1 , L_2 et B sont globalement indépendants, que vaut alors $\mathbb{P}(T)$?
5. En fait, le constructeur constate que souvent, lorsque la première ampoule est défectueuse, la deuxième l'est aussi : lorsque la première ampoule est défectueuse, la deuxième a 90% de chances de l'être aussi. Il constate aussi que lorsque les deux ampoules sont défectueuses la pile a autant de chance qu'avant d'être défectueuse.

On note \mathbb{P}' les probabilités dans ce nouveau modèle. Calculer $\mathbb{P}'(L_1 \cap L_2)$, $\mathbb{P}'(L_1 \cap L_2 \cap P)$ et $\mathbb{P}'(T)$.

Exercice 3

Note : cet exercice est assez long et comporte beaucoup de questions. Vous pouvez parfaitement *utiliser* le résultat d'une question précédente même si vous n'avez pas réussi à la traiter.

Soit m un mot sur l'alphabet $\{a, b\}$:

$$m = m[0]m[1] \cdots m[\ell - 1]$$

où ℓ est un entier ≥ 1 , ou bien $m = \varepsilon$ (le mot vide) et $\ell = 0$. L'entier ℓ est la longueur du mot.

Le mot m est dit "dominé par les a " si toute position de ce mot est étiquetée par la lettre a ou bien a une voisine étiquetée par la lettre a . Plus formellement :

$$\forall i \in [0, \ell - 1], \exists j \in [0, \ell - 1], |i - j| \leq 1 \text{ et } m[j] = a.$$

Par exemple, si

$$m_0 = \varepsilon, m_1 = aaabba, m_2 = ababaab, m_3 = baaabba, m_4 = aaabbbba, m_5 = bbaa, m_6 = b$$

m_0, m_1, m_2, m_3 sont dominés par les a mais m_4, m_5, m_6 ne sont pas dominés par les a . On note L l'ensemble de tous les mots de $\{a, b\}^*$ qui sont dominés par les a .

1. Expliquer pourquoi $m_4 \notin L, m_5 \notin L, m_6 \notin L$.

On définit

$$\begin{aligned} L_a &= \{m \in \{a, b\}^* \mid a \cdot m \in L\} \\ L_b &= \{m \in \{a, b\}^* \mid b \cdot m \in L\} \end{aligned}$$

(un mot est dans L_a si, lorsqu'on ajoute un a devant, on obtient un mot de L ; un mot est dans L_b si, lorsqu'on ajoute un b devant, on obtient un mot de L)

2. Montrer que $L_b \subseteq L$.
3. Vérifier que $bba \in L_a$. Est-ce-que $bba \in L$? Est-il vrai que $L_a \subseteq L$?
4. Montrer que

$$\begin{aligned} L &= \{\varepsilon\} \cup a \cdot L_a \cup b \cdot L_b \\ L_a &= \{\varepsilon\} \cup b \cdot L \cup a \cdot L_a \\ L_b &= a \cdot L_a \end{aligned}$$

Notations : ε est le mot "vide" ; pour tout mot m et tout ensemble de mots M , $m \cdot M$ est l'ensemble de tous les mots obtenus en concaténant le mot m avec un mot (n'importe lequel) de M . Par exemple, $ab \cdot \{a, ab\} = \{aba, abab\}$.

5. Montrer que, dans les deux premières formules données à la question 4, les unions sont *disjointes*.

On cherche à dénombrer les mots dominés par les a . On note u_n (respectivement, v_n, w_n) le nombre de mots de longueur n de L (respectivement, L_a, L_b).

6. Montrer que

$$u_0 = 1, \quad v_0 = 1, \quad w_0 = 0$$

et pour tout $n \geq 0$

$$u_{n+1} = v_n + w_n, \quad v_{n+1} = u_n + v_n, \quad w_{n+1} = v_n$$

Indication : on s'appuyera sur les égalités de la question 4.

7. Calculer u_n pour $n \in [0, 5]$.

8. Montrer que u_n vérifie la récurrence : $\forall n \geq 0$

$$u_{n+3} = u_{n+2} + u_{n+1} + u_n.$$

Calculer u_{10} .

9. On tire aléatoirement un mot m de longueur 10 sur l'alphabet $\{a, b\}$ selon le procédé suivant :

— on joue 10 fois de suite à pile ou face, avec une pièce de monnaie : on obtient les résultats r_0, r_1, \dots, r_9 où $r_i \in \{a, b\}$; à chaque tirage, on traduit "Pile" par a et "Face" par b .

— on pose $m := r_0 r_1 \cdots r_9$.

On obtient ainsi la loi uniforme sur l'ensemble des mots de longueur 10 sur $\{a, b\}$.

Quelle est la probabilité de tirer un mot m qui soit dominé par les a ?

Exercice 4

On réunit 365 personnes, toutes nées une année non bisextile. On fait l'hypothèse que chaque personne est née un jour de l'année tiré uniformément parmi les 365 jours de l'année, et que les jours de naissance des personnes sont indépendants entre eux.

On notera D_i , pour chaque entier $1 \leq i \leq 365$, la variable aléatoire qui désigne le jour de naissance de la i -ème personne (vu comme un entier entre 1 et 365 : 1 pour le jour de l'An, 15 pour le 15 janvier, 32 pour le 1er février, 365 pour le 31 décembre).

1. D'après l'énoncé, quelle est la loi de D_1 ? Quelle relation a-t-on entre les variables aléatoires D_i pour $1 \leq i \leq 365$?
2. On note N_1 la variable aléatoire qui compte le nombre de personnes nées le 1er janvier. Donner une expression de N_1 en fonction des variables D_i .
3. Quelle est **exactement** la loi de N_1 ? Par quelle loi peut-on l'approcher ?
4. Donner une formule exacte, pour la probabilité qu'aucune des 365 personnes ne soit née le 1er janvier. Comparer à ce que donne la loi approchée décrite à la question précédente.
5. On note M_1 la variable aléatoire qui compte le nombre de personnes nées en janvier, et M_2 la variable aléatoire qui compte le nombre de personnes nées en février. Quelle est la loi exacte de M_1 ? De M_2 ? De $M_1 + M_2$?
6. Les variables aléatoires M_1 et M_2 sont-elles indépendantes ?
7. Donner une formule exacte pour la probabilité que les 365 personnes soient toutes nées un jour différent de l'année.