Université d'Angers
L3 Math & Math appli. Année universitaire 2018-2019

Corrigé du

Devoir maison de Probabilités

À rendre la semaine du 25 février

Exercice 1. (1) Soient $X_1, X_2, \dots, X_n$ des v.a. indépendantes où chaque X_k suit une loi de Bernoulli de paramètre p_k. On pose $S = X_1 + X_2 + \cdots + X_n$.

(a) Exprimer $\mathbb{E}(S)$ et $\mathbb{V}(S)$ en fonction des p_k.

(b) On note $p = (p_1 + p_2 + \cdots + p_n)/n$ la moyenne des probabilités de succès. Vérifier par un calcul direct que $\mathbb{V}(S) + \sum_{k=1}^{n}(p_k - p)^2 = np(1-p)$, puis compléter l'assertion suivante : « À p fixé, l'incertitude sur S est maximale lorsque les p_k sont ...... , puisque ...... ».

(2) Soit X une v.a. de loi de Poisson $\mathcal{P}(\lambda)$, $\lambda > 0$, c'est-à-dire telle que pour tout $n \geq 0$, $\mathbb{P}(X = n) = e^{-\lambda}\lambda^n/n!$. Montrer que

$$\mathbb{E}\left(\frac{1}{1+X}\right) = \frac{1-e^{-\lambda}}{\lambda}.$$

Solution. (1) (a) Par linéarité de l'espérance, on a

$$\mathbb{E}(S) = \sum_{k=1}^{n} \mathbb{E}(X_k) = \sum_{k=1}^{n} p_k.$$

Comme les variables sont indépendantes, on a aussi

$$\mathbb{V}(S) = \sum_{k=1}^{n} \mathbb{V}(X_k) = \sum_{k=1}^{n} p_k(1-p_k).$$

(b) On calcule en tenant compte du fait que $\sum_{k=1}^{n} p_k = np$:

$$\begin{aligned}
\mathbb{V}(S) + \sum_{k=1}^{n}(p_k - p)^2 &= \sum_{k=1}^{n} p_k(1-p_k) + \sum_{k=1}^{n}(p_k - p)^2 \\
&= \sum_{k=1}^{n} p_k - \sum_{k=1}^{n} p_k^2 + \sum_{k=1}^{n} p_k^2 - 2p\sum_{k=1}^{n} p_k + \sum_{k=1}^{n} p^2 \\
&= np - 2np^2 + np^2 \\
&= np(1-p).
\end{aligned}$$

Cette égalité montre que la variance de S est maximale lorsque $\sum_{k=1}^{n}(p_k - p)^2$ est minimal, c'est-à-dire lorsque tous les p_k sont égaux à p. Donc :
« À p fixé, l'incertitude sur S est maximale lorsque les p_k sont égaux à p, puisque la variance est alors maximale ».

(2) Par transfert, on a

$$\begin{aligned}\mathbb{E}\left(\frac{1}{1+X}\right) &= \sum_{n=0}^{\infty} \frac{1}{n+1}\mathbb{P}(X=n)\\ &= e^{-\lambda}\sum_{n=0}^{\infty}\frac{\lambda^n}{(n+1)!}\\ &= e^{-\lambda}\sum_{j=1}^{\infty}\frac{\lambda^{j-1}}{j!}\\ &= \frac{e^{-\lambda}}{\lambda}\left(e^{\lambda}-1\right)\\ &= \frac{1-e^{-\lambda}}{\lambda}.\end{aligned}$$

Exercice 2 (Formule du crible)**.** Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé.

(1) Démontrer que pour tous $A, B, C \in \mathcal{F}$, on a

$$\mathbb{P}(A\cup B\cup C) = \mathbb{P}(A)+\mathbb{P}(B)+\mathbb{P}(C)-\mathbb{P}(A\cap B)-\mathbb{P}(B\cap C)-\mathbb{P}(A\cap C)+\mathbb{P}(A\cap B\cap C).$$

(2) Plus généralement, démontrer par récurrence sur n que pour tous $A_1, A_2, \dots, A_n \in \mathcal{F}$, on a

$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{k=1}^{n}(-1)^{k+1}\sum_{I\in\mathcal{P}_k(n)}\mathbb{P}\left(\bigcap_{i\in I} A_i\right),$$

où $\mathcal{P}_k(n)$ désigne l'ensemble des parties de $\{1, 2, \dots, n\}$ à k éléments.

(3) Une application théorique : Soit $(A_n)_{n\geq 1}$ une suite d'éléments de $\mathcal{F}$ tels que $\mathbb{P}(A_i\cap A_j) = 0$ dès que $i \neq j$. Démontrer (en justifiant soigneusement les étapes de votre argument) que

$$\mathbb{P}(\cup_{n\geq 1} A_n) = \sum_{n=0}^{\infty}\mathbb{P}(A_n).$$

(4) Une application pratique : *Ce matin, le facteur doit distribuer* n *lettres adressées à* n *destinataires différents. Au hasard, il en glisse une dans chaque boîte aux lettres. On cherche la probabilité* p_n *qu'au moins l'un des destinataires reçoive la lettre qui lui était adressée.* L'expérience peut être modélisée à l'aide de l'espace probabilisé fini $(\Omega, \mathbb{P})$ où Ω désigne l'ensemble des permutations σ de $\{1, 2, \dots, n\}$ et $\mathbb{P}$ la probabilité uniforme sur Ω.

(a) Pour $i = 1, \dots, n$, calculer la probabilité de l'événement

$$A_i = \{\sigma \in \Omega | \sigma(i) = i\}.$$

(b) Pour $I \subset \{1, 2, \dots, n\}$, déterminer la probabilité de $\cap_{i\in I} A_i$ en fonction du cardinal de I.

(c) La probabilité p_n est celle d'un certain événement A. Exprimer A à l'aide des événements A_i.

(d) Montrer que

$$p_n = \sum_{k=1}^{n} \frac{(-1)^{k+1}}{k!}.$$

(e) Déterminer la limite de p_n lorsque n tend vers l'infini.

Solution. (1) On sait que pour tous événements E et F, on a

$$\mathbb{P}(E \cup F) = \mathbb{P}(E) + \mathbb{P}(F) - \mathbb{P}(E \cap F).$$

En appliquant cette formule avec $E = A \cup B$ et $F = C$ d'abord, puis avec $E = A$ et $F = B$ ensuite, on obtient successivement

$$\begin{aligned}\mathbb{P}(A \cup B \cup C) &= \mathbb{P}(A \cup B) + \mathbb{P}(C) - \mathbb{P}((A \cup B) \cap C)\\ &= \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}((A \cup B) \cap C)\end{aligned}$$

Puisque $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$, le dernier terme se calcule ainsi :

$$\mathbb{P}((A \cup B) \cap C) = \mathbb{P}(A \cap C) + \mathbb{P}(B \cap C) - \mathbb{P}(A \cap B \cap C),$$

et l'on trouve la relation demandée.

(2) Pour $n = 1$, la formule s'écrit $\mathbb{P}(A_1) = \mathbb{P}(A_1)$ donc il n'y a rien à démontrer. Pour $n = 2$, la formule s'écrit

$$\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2).$$

C'est une formule établie dans le cours (qu'on ne va donc pas redémontrer !).

Supposons que la formule soit vérifiée au rang $n \geq 1$ pour tous $A_1, A_2, \ldots, A_n \in \mathcal{F}$, et donnons nous $n+1$ événements $A_1, A_2, \ldots, A_n, A_{n+1} \in \mathcal{F}$. Posons $A = \cup_{k=1}^{n+1} A_k$. En écrivant A comme la réunion de $\cup_{k=1}^{n} A_k$ avec A_{n+1} et en appliquant la formule connue pour deux événements, il vient

$$\mathbb{P}(A) = \mathbb{P}(\cup_{k=1}^{n} A_k) + \mathbb{P}(A_{n+1}) - \mathbb{P}(\cup_{k=1}^{n}(A_k \cap A_{n+1})).$$

Par hypothèse de récurrence, on a d'une part

$$\mathbb{P}(\cup_{k=1}^{n} A_k) = \sum_{k=1}^{n} (-1)^{k+1} \sum_{I \in \mathcal{P}_k(n)} \mathbb{P}\left(\bigcap_{i \in I} A_i\right),$$

et comme une partie I de $\{1, 2, \ldots, n\}$ est une partie de $\{1, 2, \ldots, n, n+1\}$ qui ne contient pas $n+1$, on peut toujours écrire

$$\mathbb{P}(\cup_{k=1}^{n} A_k) = \sum_{k=1}^{n} (-1)^{k+1} \sum_{I \in E_k} \mathbb{P}\left(\bigcap_{i \in I} A_i\right),$$

où E_k désigne l'ensemble des $I \in \mathcal{P}_k(n+1)$ qui ne contiennent pas $n+1$. D'autre part (toujours d'après l'hypothèse de récurrence), on a

$$\mathbb{P}(\cup_{k=1}^{n}(A_k \cap A_{n+1})) = \sum_{k=1}^{n} (-1)^{k+1} \sum_{I \in \mathcal{P}_k(n)} \mathbb{P}\left(\bigcap_{i \in I}(A_i \cap A_{n+1})\right).$$

On remarque maintenant que

$$\bigcap_{i\in I}(A_i \cap A_{n+1}) = \left(\bigcap_{i\in I} A_i\right) \cap A_{n+1} = \bigcap_{j\in J} A_j,$$

où l'on a posé $J = I \cup \{n+1\}$. Lorsque I décrit l'ensemble $\mathcal{P}_k(n)$ des parties à k éléments de $\{1, 2, \ldots, n\}$, J décrit l'ensemble des parties à $k+1$ éléments de $\{1, 2, \ldots, n, n+1\}$ qui contiennent $n+1$, c'est-à-dire l'ensemble

$$F_{k+1} = \mathcal{P}_{k+1}(n+1) \setminus E_{k+1}.$$

On peut donc écrire

$$\mathbb{P}(\cup_{k=1}^n (A_k \cap A_{n+1})) = \sum_{k=1}^{n} (-1)^{k+1} \sum_{J\in F_{k+1}} \mathbb{P}\left(\bigcap_{j\in J} A_j\right) = \sum_{k=2}^{n+1} (-1)^k \sum_{J\in F_k} \mathbb{P}\left(\bigcap_{j\in J} A_j\right).$$

Donc

$$\mathbb{P}(A) = \sum_{k=1}^{n} (-1)^{k+1} \sum_{J\in E_k} \mathbb{P}\left(\bigcap_{j\in J} A_j\right) + \mathbb{P}(A_{n+1}) + \sum_{k=2}^{n+1} (-1)^{k+1} \sum_{J\in F_k} \mathbb{P}\left(\bigcap_{j\in J} A_j\right).$$

Analysons maintenant les différents termes de cette expression. L'ensemble E_1 est constitué des parties $\{1\}, \{2\}, \ldots, \{n\}$, donc le terme d'indice $k = 1$ de la première somme est simplement $\mathbb{P}(A_1) + \mathbb{P}(A_2) + \cdots + \mathbb{P}(A_n)$. Son regroupement avec $\mathbb{P}(A_{n+1})$ se récrit sous forme condensée

$$\sum_{J\in\mathcal{P}_1(n+1)} \mathbb{P}\left(\bigcap_{j\in J} A_j\right).$$

Pour $k = 2, \ldots, n$, le regroupement du terme d'indice k de la première somme avec celui de la seconde donne

$$(-1)^{k+1}\left(\sum_{J\in E_k} + \sum_{J\in F_k}\right) \mathbb{P}\left(\bigcap_{j\in J} A_j\right) = (-1)^{k+1} \sum_{J\in\mathcal{P}_k(n+1)} \mathbb{P}\left(\bigcap_{j\in J} A_j\right)$$

puisque E_k et F_k forment une partition de $\mathcal{P}_k(n+1)$. Enfin, comme E_{n+1} est vide, F_{n+1} n'est autre que $\mathcal{P}_{n+1}(n+1)$, si bien que le terme d'indice $k = n+1$ de la seconde somme dans l'expression obtenue pour $\mathbb{P}(A)$ s'écrit aussi

$$(-1)^{k+1} \sum_{J\in\mathcal{P}_k(n+1)} \mathbb{P}\left(\bigcap_{j\in J} A_j\right),$$

avec $k = n+1$. D'où, finalement,

$$\mathbb{P}(A) = \sum_{k=1}^{n+1} (-1)^{k+1} \sum_{J\in\mathcal{P}_k(n+1)} \mathbb{P}\left(\bigcap_{j\in J} A_j\right).$$

(3) (Application théorique) Quel que soit $n \geq 1$, d'après la formule du crible, on a
$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{k=1}^{n} (-1)^{k+1} \sum_{I \in \mathcal{P}_k(n)} \mathbb{P}\left(\bigcap_{i\in I} A_i\right).$$
Si I est une partie de $\{1, 2, \ldots, n\}$ ayant au moins deux éléments, disons $k \neq \ell$, alors $\cap_{i\in I} A_i$ est inclus dans $A_k \cap A_\ell$, donc $\mathbb{P}(\cap_{i\in I} A_i) \leq \mathbb{P}(A_k \cap A_\ell) = 0$ par hypothèse. Ainsi, tous les termes d'indice $k \geq 2$ dans la formule du crible sont nuls. Il en résulte que
$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = (-1)^{1+1} \sum_{I \in \mathcal{P}_1(n)} \mathbb{P}\left(\bigcap_{i\in I} A_i\right) = \sum_{i=1}^{n} \mathbb{P}(A_i),$$
puisque $\mathcal{P}_1(n) = \{\{1\}, \{2\}, \ldots, \{n\}\}$. Maintenant, on observe que
$$\bigcup_{n\geq 1} A_n = \bigcup_{n\geq 1} B_n,$$
avec $B_n = \cup_{i=1}^{n} A_i$. Comme la suite d'événements (B_n) est croissante, on sait que la suite de terme général $\mathbb{P}(B_n)$ converge et vérifie
$$\lim_{n\to\infty} \mathbb{P}(B_n) = \mathbb{P}\left(\bigcup_{n\geq 1} B_n\right) = \mathbb{P}\left(\bigcup_{n\geq 1} A_n\right).$$
Or, on vient de voir que $\mathbb{P}(B_n) = \sum_{i=1}^{n} \mathbb{P}(A_i)$ est la n-ème somme partielle de la série de terme général $\mathbb{P}(A_i)$. Donc
$$\sum_{i=1}^{\infty} \mathbb{P}(A_i) = \mathbb{P}\left(\bigcup_{n\geq 1} A_n\right).$$

(4) (Application pratique)

(a) Pour fixer les idées, on commence par dénombrer $A_1 = \{\sigma \in \Omega | \sigma(1) = 1\}$. Soit donc $\sigma \in A_1$. Pour la valeur de $\sigma(1)$, on n'a pas de liberté (c'est égal à 1) ; pour $\sigma(2)$ on peut choisir n'importe quel entier entre 2 et n, donc $n-1$ possibilités. Pour $\sigma(3)$ on peut choisir n'importe quel entier entre 2 et n, mais différent de $\sigma(2)$ (qu'on a déjà fixé), ce qui fait $(n-2)$ possibilités, et ainsi de suite. Donc le cardinal de A_1 est $(n-1)!$. Pour les autres A_i, le résultat est le même (pas de liberté pour le choix de $\sigma(i)$ et ensuite on choisit une par une et sans répétition les images $\sigma(j), j \neq i$ parmi les entiers $\{1, 2, \ldots, n\} \setminus \{i\}$). Donc
$$\mathbb{P}(A_i) = \frac{|A_i|}{|\Omega|} = \frac{(n-1)!}{n!} = \frac{1}{n}.$$

(b) Soit I une partie non vide de $\{1, 2, \ldots, n\}$. On a
$$\cap_{i\in I} A_i = \{\sigma \in \Omega | \forall i \in I, \sigma(i) = i\},$$
c'est-à-dire que les éléments de $\cap_{i\in I} A_i$ sont les permutations qui fixent les éléments de I. Appelons $i_1, i_2, \ldots, i_k$ les éléments de I et $j_1, j_2, \ldots, j_\ell$, avec

$\ell = n-k$, les éléments de son complémentaire J dans $\{1, 2, \ldots, n\}$. On cherche le nombre de permutations σ qui vérifient

$$\sigma(i_1) = i_1, \sigma(i_2) = i_2, \ldots, \sigma(i_k) = i_k.$$

Pour les images $\sigma(i_1), \sigma(i_2), \ldots, \sigma(i_k)$ il n'y a donc pas de liberté ; ensuite $\sigma(j_1)$ peut être n'importe quel élément de J, ce qui fait ℓ possibilités ; pour $\sigma(j_2)$, on peut choisir n'importe quel élément de J différent de $\sigma(j_1)$, ce qui fait $\ell - 1$ possibilités ; et ainsi de suite. Donc la cardinal de $\cap_{i\in I} A_i$ est $\ell! = (n-k)!$ et

$$\mathbb{P}(\cap_{i\in I} A_i) = \frac{(n-k)!}{n!},$$

où k est le cardinal de I.

(c) Le résultat de l'expérience aléatoire décrite dans l'énoncé peut être modélisé par une permutation σ, où $\sigma(i) = j$ s'interprète par « le destinataire i reçoit a lettre adressée à j ». L'événement A dont on cherche la probabilité p_n est l'ensemble des permutations σ pour lesquelles il existe au moins un i tel que $\sigma(i) = i$. C'est donc simplement la réunion des événements A_i.

(d) D'après la formule du crible, on a donc

$$\begin{aligned} p_n &= \sum_{k=1}^{n} (-1)^{k+1} \sum_{I\in\mathcal{P}_k(n)} \mathbb{P}\left(\bigcap_{i\in I} A_i\right) \\ &= \sum_{k=1}^{n} (-1)^{k+1} \sum_{I\in\mathcal{P}_k(n)} \frac{(n-k)!}{n!} \\ &= \sum_{k=1}^{n} (-1)^{k+1} \binom{n}{k} \frac{(n-k)!}{n!} \\ &= \sum_{k=1}^{n} \frac{(-1)^{k+1}}{k!}. \end{aligned}$$

(e) On a

$$p_n = -\sum_{k=1}^{n} \frac{(-1)^k}{k!} = 1 - \sum_{k=0}^{n} \frac{(-1)^k}{k!},$$

donc

$$\lim_{n\to\infty} p_n = 1 - e^{-1} \approx 0.63.$$

Exercice 3. On tire un entier aléatoire N selon une loi de Poisson $\mathcal{P}(\lambda)$. Sachant $N = n$, on effectue n répétitions indépendantes d'une même épreuve de Bernoulli de probabilité de succès p. On note S le nombre de succès et E le nombre d'échecs.

(1) Sachant $N = n$, quelle est la loi de S ? Et celle de E ?

(2) Pour tous $i, j \geq 0$, exprimer l'événement $\{S = i, E = j\}$ à l'aide des variables S et N. En déduire que

$$\mathbb{P}(S = i, E = j) = e^{-\lambda} \lambda^{i+j} \frac{p^i (1-p)^j}{i!j!}.$$

(3) Déterminer les lois de S et E. Ces deux variables sont-elles indépendantes ?

Solution. (1) Sachant $N = n$, le nombre de succès S est la somme de n variables indépendantes de même loi $\mathcal{B}(p)$, donc S suit une loi binomiale $\mathcal{B}(n, p)$. Ainsi, pour tout $0 \leq k \leq n$, on a

$$\mathbb{P}(S = k|N = n) = \binom{n}{k} p^k q^{n-k},$$

où l'on a posé $q = 1 - p$. Toujours sachant $N = n$, il est assez clair que E suit une loi $\mathcal{B}(n, q)$. On peut le vérifier ainsi : puisque $N = S + E$, on a

$$\mathbb{P}(E = j, N = n) = \mathbb{P}(S = n - j, N = n),$$

donc

$$\mathbb{P}(E = j|N = n) = \mathbb{P}(S = n - j|N = n) = \binom{n}{n-j} p^{n-j} q^{n-(n-j)} = \binom{n}{j} q^j p^{n-j}.$$

(2) Puisque $S + E = N$, on a

$$\{S = i, E = j\} = \{S = i, N = i + j\},$$

donc

$$\begin{aligned}
\mathbb{P}(S = i, E = j) &= \mathbb{P}(S = i, N = i + j) \\
&= \mathbb{P}(S = i|N = i + j)\mathbb{P}(N = i + j) \\
&= \binom{i+j}{i} p^i q^j e^{-\lambda} \frac{\lambda^{i+j}}{(i + j)!} \\
&= e^{-\lambda} \lambda^{i+j} \frac{p^i q^j}{i! j!}.
\end{aligned}$$

(3) La probabilité précédente peut être mise sous une forme produit $g(i)h(j)$. Par exemple, on peut écrire

$$\mathbb{P}(S = i, E = j) = \left(e^{-\lambda p} \frac{(\lambda p)^i}{i!}\right) \cdot \left(e^{-\lambda q} \frac{(\lambda q)^j}{j!}\right)$$

On en déduit que les variables S et E sont indépendantes, et qu'il existe $c > 0$ tel que

$$\mathbb{P}(S = i) = c e^{-\lambda p} \frac{(\lambda p)^i}{i!},$$

pour tout $i \geq 0$, et

$$\mathbb{P}(E = j) = c^{-1} e^{-\lambda q} \frac{(\lambda q)^j}{j!},$$

pour tout $j \geq 0$. Puisqu'en l'absence des constantes c et c^{-1}, les termes de droite sont déjà des germes de probabilité (ceux des lois de Poisson $\mathcal{P}(\lambda p)$ et $\mathcal{P}(\lambda q)$ respectivement), on a nécessairement $c = c^{-1} = 1$. Donc S suit la loi $\mathcal{P}(\lambda p)$ et E suit la loi $\mathcal{P}(\lambda q)$.