Ecole Polytechnique - MAP 361 **Lundi 15 avril 2019**

PC 1 – Probabilités discrètes

Exercice 1 [Evénements indépendants] Soit $\Omega := \{\omega_1, \omega_2, \omega_3, \omega_4\}$ muni de la probabilité uniforme $\mathbb{P}$. On définit les événements $A := \{\omega_1, \omega_2\}$, $B := \{\omega_1, \omega_3\}$ et $C := \{\omega_2, \omega_3\}$. Montrer que A, B et C sont indépendants deux à deux. Comparer $\mathbb{P}(A \cap B \cap C)$ et $\mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$.

On a $\mathbb{P}(\{w_i\}) = 1/|\Omega| = 1/4$ pour $i = 1, \dots, 4$. D'une part, $\mathbb{P}(A) = \mathbb{P}(\{\omega_1\}) + \mathbb{P}(\{\omega_2\}) = 1/2$. De mÃłme, $\mathbb{P}(B) = \mathbb{P}(C) = 1/2$. D'autre part, $A \cap B = \{\omega_1\}$ et donc $\mathbb{P}(A \cap B) = 1/4$. Donc, on a montré que $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$, d'oÃź l'indépendance de A et B. De mÃłme, on montre que A et C sont indépendants et B et C sont indépendants.

Comme $A \cap B \cap C = \emptyset$, on a $\mathbb{P}(A \cap B \cap C) = 0$. En revanche, $\mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C) = 1/8$. Cela implique que A, B et C ne sont pas mutuellement indépendants.

Exercice 2 [Conditionnement] L'exercice suivant est très classique et a de nombreuses variantes. Il illustre l'utilité d'une formulation mathématique rigoureuse pour éviter des pièges et des paradoxes dus à des raisonnements spécieux.
Une famille a deux enfants. On suppose les 4 configurations (ω_1, ω_2) avec ω_i = sexe du $i$ème enfant équiprobables.

1. Quelle est la probabilité pour que les deux enfants soient des filles sachant que le plus jeune enfant est une fille ?
2. Quelle est la probabilité pour que les deux enfants soient des filles sachant que l'enfant plus Ãćgé est une fille ?
3. Quelle est la probabilité pour que les deux enfants soient des filles sachant que l'un des enfants est une fille ?

1. Ici $\Omega = \{(F, F), (G, G), (F, G), (G, F)\}$. L'événement que le plus jeune enfant est une fille correspond Ãă l'ensemble $A := \{(F, F), (G, F)\}$ qui se produit avec probabilité $\mathbb{P}(A) = |A|/|\Omega| = 1/2$. Pour la probabilité que les deux enfants soient des filles sachant que le plus jeune enfant est une fille on obtient alors
$$\mathbb{P}(\{(F, F)\}|A) = \frac{\mathbb{P}(\{(F, F)\} \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(\{(F, F)\})}{\mathbb{P}(A)} = \frac{1/4}{1/2} = \frac{1}{2}.$$
2. Par le mÃłme raisonnement, en notant $B := \{(F, F), (F, G)\}$ l'événement que l'enfant plus Ãćgé est une fille avec $\mathbb{P}(B) = |B|/|\Omega| = 1/2$, on obtient
$$\mathbb{P}(\{(F, F)\}|B) = \frac{\mathbb{P}(\{(F, F)\} \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(\{(F, F)\})}{\mathbb{P}(B)} = \frac{1/4}{1/2} = \frac{1}{2}.$$
3. Notons $C := \{(F, F), (F, G), (G, F)\}$ l'événement que l'un des enfants est une fille. On a $\mathbb{P}(C) = |C|/|\Omega| = 3/4$ et donc
$$\mathbb{P}(\{(F, F)\}|C) = \frac{\mathbb{P}(\{(F, F)\})}{\mathbb{P}(C)} = \frac{1/4}{3/4} = \frac{1}{3}.$$

Exercice 3 [Cartes électroniques défectueuses] On considère une usine fabriquant des cartes électroniques. Lors de la production on sait qu'une carte sur 100 000 est défectueuse. En fin de production, on effectue un test pour savoir si la carte est défectueuse ou non. Lorsque le test donne un résultat positif, ce test déclare la pièce comme défectueuse. Pour les pièces effectivement défectueuses, le test est positif dans 95 % des cas, pour les pièces correctes, le test est positif (donc déclare la pièce comme déféctueuse [1]) dans 1% des cas.

Quelle est la probabilité que la pièce soit effectivement défectueuse lorsque le test est positif ?

Soit D l'évenement 'la pièce est défectueuse'. On a $\mathbb{P}(D) = 1/100000$. On note P l'évenement 'le test est positif'. Traduction de l'énoncé :

$$\mathbb{P}(P|D) = 0.95, \quad \mathbb{P}(P|\overline{D}) = 0.01.$$

On cherche $\mathbb{P}(D|P)$. Formule de Bayes :

$$\mathbb{P}(D|P) = \frac{\mathbb{P}(P|D) \times \mathbb{P}(D)}{\mathbb{P}(P)}.$$

On remarque par les proba totales

$$\mathbb{P}(P) = \mathbb{P}(P|D)\mathbb{P}(D) + \mathbb{P}(P|\overline{D})\mathbb{P}(\overline{D}) = \mathbb{P}(P|D)\mathbb{P}(D) + \mathbb{P}(P|\overline{D})(1 - \mathbb{P}(D)).$$

Enfin, on obtient

$$\mathbb{P}(D|P) = \frac{0.95 * 10^{-5}}{0.95 * 10^{-5} + 0.01 * (1 - 10^{-5})} \approx \frac{10^{-5}}{0 + 10^{-2}} = 10^{-3}.$$

(Avec une calculatrice on obtient la valeur 0.000949.)

Exercice 4 [Limite supérieure d'ensembles] Soit $(A_n)_{n\geq 1}$ une suite de sous-ensembles de Ω.

1. Si $\Omega = \mathbb{R}$, donner $\limsup_{n\to\infty} A_n$ dans les trois cas suivants :
   (a) $A_n = [-1/n, 3 + 1/n]$
   (b) $A_n = [-2 - (-1)^n, 2 + (-1)^n[$
   (c) $A_n = p_n\mathbb{N}$, où $(p_n)_{n\geq 1}$ est la suite ordonnée des nombres premiers et $p_n\mathbb{N}$ est l'ensemble des multiples de p_n,
2. Comparer $\limsup_{n\to\infty}(A_n \cup B_n)$ et $\limsup_{n\to\infty}(A_n \cap B_n)$ avec $\limsup_{n\to\infty} A_n$ et $\limsup_{n\to\infty} B_n$.

1. Rappel de la définition : $\limsup_{n\to\infty} A_n = \cap_{k\geq 1} \cup_{n\geq k} A_n$. Il s'agit de l'événement où :

$$\omega \in \limsup_{n\to\infty} A_n \iff \text{il existe une infinité de } n \text{ tels que } \omega \in A_n.$$

   (a) La suite $(A_n)_{n\geq 1}$ étant monotone décroissante ($A_n \supset A_{n+1}$ pour tout $n \geq 1$), on a pour tout $k \geq 1$, $\cup_{n\geq k} A_n = [-1/k, 3 + 1/k]$. D'une part, on voit que $[0, 3] \subset A_k \subset \cup_{n\geq k} A_n$ pour tout k. D'autre part, pour tout $s < 0$ et pour tout $t > 3$ il existe k tel que $s < -1/k$ et $t > 3 + 1/k$. Donc, $\limsup_{n\to\infty} A_n = [0, 3]$.
   (b) Pour tout n pair, on a $A_n = [-3, 3[$. Pour tout n impair, on a $A_n = [-1, 1[$. Donc, pour tout $k \geq 1$, $\cup_{n\geq k} A_n = [-3, 3[$, et $\limsup_{n\to\infty} A_n = [-3, 3[$.

1. On parle alors de "faux-positifs".

(c) On a $A_1 = 2\mathbb{N} = \{0, 2, 4, 6, ...\}$, $A_2 = 3\mathbb{N} = \{0, 3, 6, 9, ...\}$, ..., $A_n = p_n\mathbb{N} = \{0, p_n, 2p_n, 3p_n, ...\}$. Clairement, $A_n \subset \mathbb{N}$ pour tout n, donc $\limsup_{n\to\infty} A_n \subset \mathbb{N}$. On voit que $0 \in A_n$ pour tout n, donc $0 \in \limsup_{n\to\infty} A_n$. Enfin, pour tout $z \in \mathbb{N}^*$ et pour tout k tel que $z < p_k$, on a $z \notin A_n$ pour tout $n \geq k$. Donc $\limsup_{n\to\infty} A_n = \{0\}$.

2. D'une part,

$$\begin{aligned}
\omega \in \limsup_{n\to\infty}(A_n \cup B_n) &\iff \text{il existe une infinité de } n \text{ tels que } \omega \in A_n \cup B_n \\
&\iff \text{il existe une infinité de } n \text{ tels que } \omega \in A_n \textbf{ ou } \omega \in B_n \\
&\iff \omega \in \limsup_{n\to\infty} A_n \textbf{ ou } \omega \in \limsup_{n\to\infty} B_n \\
&\iff \omega \in \left(\limsup_{n\to\infty} A_n\right) \cup \left(\limsup_{n\to\infty} B_n\right).
\end{aligned}$$

D'autre part,

$$\begin{aligned}
\omega \in \limsup_{n\to\infty}(A_n \cap B_n) &\iff \text{il existe une infinité de } n \text{ tels que } \omega \in A_n \cap B_n \\
&\iff \text{il existe une infinité de } n \text{ tels que } \omega \in A_n \textbf{ et } \omega \in B_n \\
&\implies \omega \in \left(\limsup_{n\to\infty} A_n\right) \cap \left(\limsup_{n\to\infty} B_n\right).
\end{aligned}$$

On a montré que

$$\begin{aligned}
\limsup_{n\to\infty}(A_n \cup B_n) &= \left(\limsup_{n\to\infty} A_n\right) \cup \left(\limsup_{n\to\infty} B_n\right) \\
\limsup_{n\to\infty}(A_n \cap B_n) &\subset \left(\limsup_{n\to\infty} A_n\right) \cap \left(\limsup_{n\to\infty} B_n\right).
\end{aligned}$$

La dernière inclusion peut Ãłtre stricte. Par exemple, pour $A_{2l} = B_{2l+1} = \{0\}$ et $A_{2l+1} = B_{2l} = \{1\}$ pour tout l, on a $\limsup_{n\to\infty} A_n = \limsup_{n\to\infty} B_n = \{0, 1\}$, alors que $\limsup_{n\to\infty}(A_n \cap B_n) = \emptyset$.

Exercice 5 [Décimation] Proba totale+Borel-Cantelli-va discrete 1. facile, 2. intensif Le nombre d'individus dans une colonie de bactéries est modélisé par une variable aléatoire N suivant une loi de Poisson de paramètre $\lambda > 0$. En présence d'un antibiotique, chaque individu de la population a une probabilité $p \in]0, 1[$ de survivre (indépendamment des autres).

1. Déterminer la loi du nombre N_1 de bactéries survivantes. On adoptera la modélisation suivante :
$$N_1 = \sum_{i=1}^{N} \varepsilon_i,$$
où $(\varepsilon_i)_{i\geq 1}$ est une suite de variables aléatoires indépendantes de loi de Bernoulli de paramètre p qu'on supposera indépendante de N.

2. On expose de manière répétée la colonie de bactéries à ce traitement antibiotique et on note N_k, $k \geq 1$, le nombre de bactéries survivantes après la $k$ième exposition. Montrer que presque sûrement la colonie finit par s'éteindre.

1. Pour tout $k \geq 0$, on a en appliquant la formule des probabilités totales
$$\mathbb{P}(N_1 = k) = \sum_{n=0}^{+\infty} \mathbb{P}(N_1 = k | N = n)\mathbb{P}(N = n).$$
On remarque que $\mathbb{P}(N_1 = k|N = n) = 0$ si $k > n$. Par ailleurs, pour $n \geq k$, on a
$$\mathbb{P}(N_1 = k|N = n) = \mathbb{P}\left(\sum_{i=1}^{N} \varepsilon_i = k|N = n\right) = \mathbb{P}\left(\sum_{i=1}^{n} \varepsilon_i = k|N = n\right) = \mathbb{P}\left(\sum_{i=1}^{n} \varepsilon_i = k\right),$$
où la dernière égalité vient de l'indépendance de N et des ε_i. Pour conclure, on sait que $\sum_{i=1}^{n} \varepsilon_i$ suit une loi binomiale de paramètres n et p et donc
$$\mathbb{P}\left(\sum_{i=1}^{n} \varepsilon_i = k\right) = \binom{n}{k} p^k (1-p)^{n-k}.$$
On en déduit que
$$\begin{aligned}\mathbb{P}(N_1 = k) &= e^{-\lambda} \sum_{n=k}^{+\infty} \binom{n}{k} p^k (1-p)^{n-k} \frac{\lambda^n}{n!} = e^{-\lambda} \sum_{n=k}^{+\infty} \frac{(\lambda p)^k (\lambda(1-p))^{n-k}}{k!(n-k)!} \\ &= e^{-\lambda} \frac{(\lambda p)^k}{k!} e^{\lambda(1-p)} = \frac{(\lambda p)^k}{k!} e^{-p\lambda}.\end{aligned}$$
Conclusion : N_1 suit une loi de Poisson de paramètre $p\lambda$.
2. Soit $(\varepsilon_n^k)_{n\geq 1, k\geq 1}$ un tableau de Bernoulli de paramètre p indépendantes. On définit de proche en proche N_k par la formule
$$N_k = \sum_{i=1}^{N_{k-1}} \varepsilon_i^{(k)},$$
avec la convention $N_0 = N$ (indépendante des $\varepsilon_n^{(k)}$). D'après la question précédente, N_k suit une loi de Poisson de paramètre $\lambda_k > 0$ vérifiant $\lambda_k = p\lambda_{k-1}$. On en déduit immédiatement que $\lambda_k = p^k \lambda$. La population n'est pas éteinte à l'instant k si et seulement si $N_k \geq 1$, événement dont la probabilité vaut
$$\mathbb{P}(N_k \geq 1) = 1 - \mathbb{P}(N_k = 0) = 1 - e^{-p^k \lambda} \sim p^k \lambda$$
lorsque $k \to \infty$.
1 ère méthode : Comme $p \in]0, 1[$, on conclut que $\sum_{k\geq 1} \mathbb{P}(N_k \geq 1) < +\infty$. Le lemme de Borel-Cantelli entraîne donc que $\mathbb{P}(\limsup\{N_k \geq 1\}) = 0$ et donc, en passant au complémentaire, $\mathbb{P}(\liminf\{N_k = 0\}) = 1$. On conclut en remarquant que l'événement $\liminf\{N_k = 0\}$ est l'événement "la population finit par s'éteindre."
2ème méthode : On peut également remarquer que
$$\{N_{k+1} \geq 1\} \subset \{N_k \geq 1\}$$
Les événements $A_k = \{N_k \geq 1\}$ forment donc une suite décroissante pour l'inclusion. On sait alors qu'en posant $A = \cap_{k\geq 1} A_k$, on a
$$\mathbb{P}(A) = \lim_{k\to+\infty} \mathbb{P}(A_k) = \lim_{k\to+\infty} 1 - e^{-p^k \lambda} = 0.$$

L'ensemble $A = \cap_{k\geq 1}\{N_k \geq 1\}$ représentant l'événement "la population ne s'éteint jamais", on conclut qu'avec probabilité 1 la population finit par s'éteindre.
Pour faire le lien avec la première méthode, on peut noter que comme les événements sont décroissants,

$$\limsup A_n = \cap_{k\geq 1} \cup_{l\geq k} A_l = \cap_{k\geq 1} A_k = A.$$

Remarque : le processus $(N_k)_{k\geq 0}$ n'est autre qu'un arbre de Galton-Watson avec loi de reproduction de Bernoulli de paramètre p et condition initiale de loi de Poisson.

Exercice 6 [Loi gÃľomÃľtrique] On modélise le jeu de pile ou face par une suite $(X_n)_{n\geq 1}$ de variables aléatoires indépendantes de loi de Bernoulli de paramètre $p \in [0,1]$, en codant 1 pour succès (pile) et 0 pour échec (face) : $\mathbb{P}(X_1 = 1) = 1 - \mathbb{P}(X_1 = 0) = p$.

1. On pose $T_1 = \inf\{n \geq 1 : X_n = 1\}$. Que reprÃľsente la variable aléatoire T_1, qu'elle est sa loi, sa moyenne, sa variance ?
2. Soit $k \geq 2$; on s'intéresse à la variable T_k définie par $T_k = \inf\{n \geq 1 : \sum_{i=1}^n X_i = k\}$ représentant l'instant où le joueur réalise son $k$ème succès. Déterminer la loi de T_k.
3. Posons $T_0 = 0$ et $\Delta_k = T_k - T_{k-1}$ pour $k \geq 1$. Montrer que les variables aléatoires Δ_k sont indépendantes et de même loi.

1. La variable aléatoire T_1 correspond à la date du premier succès. Déterminer la loi de T_1 consiste à donner $\mathbb{P}(T_1 = k)$ pour tout $k \geq 1$. On remarque que

$$\{T_1 = k\} = \{X_1 = 0, \ldots, X_{k-1} = 0, X_k = 1\} = \left(\cap_{i=1}^{k-1}\{X_i = 0\}\right) \cap \{X_k = 1\}$$

et donc par indépendance et équidistribution des variables X_i on trouve

$$\mathbb{P}(T_1 = k) = (1-p)^{k-1}p.$$

Autrement dit, la variable T_1 suit la loi géométrique de paramètre p. Pour l'espÃľrance de T_1, on trouve

$$\mathbb{E}[T_1] = \sum_{k=1}^{+\infty} k\mathbb{P}(T_1 = k) = \sum_{k=1}^{+\infty} k(1-p)^{k-1}p$$

donc en notant $f(x) = \sum_{k=0}^{+\infty}(1-x)^k = \frac{1}{x}$, $x \in]0,1]$, on voit que

$$\mathbb{E}[T_1] = -pf'(p) = \frac{1}{p}.$$

Le même raisonnement peut être adapté pour montrer que la variance est donnÃľe par

$$\mathrm{Var}(T_1) = \frac{1-p}{p^2}.$$

2. Déterminons la loi de T_k. Remarquons tout d'abord que $\mathbb{P}(T_k = n) = 0$ si $n < k$. Pour $n \geq k$, on a

$$\{T_k = n\} = \left\{\sum_{i=1}^{n-1} X_i = k-1, X_n = 1\right\}.$$

Par indépendance, on a donc

$$\mathbb{P}(T_k = n) = \mathbb{P}\left(\sum_{i=1}^{n-1} X_i = k-1\right)\mathbb{P}(X_n = 1) = p\mathbb{P}\left(\sum_{i=1}^{n-1} X_i = k-1\right).$$

Par ailleurs, $\sum_{i=1}^{n-1} X_i$ suit une loi binomiale de paramètres $n-1$ et p :

$$\mathbb{P}\left(\sum_{i=1}^{n-1} X_i = j\right) = \binom{n-1}{j} p^j (1-p)^{n-1-j}, \qquad \forall j \in \{0, 1, \ldots, n-1\}.$$

D'où la formule

$$\mathbb{P}(T_k = n) = \binom{n-1}{k-1} p^k (1-p)^{n-k}, \qquad \forall n \geq k.$$

3. Afin de montrer l'indépendance des variables Δ_k, remarquons que pour toute suite d'entiers $n_1, \ldots, n_{k+1} \geq 1$, on a

$$\begin{aligned}
&\{\Delta_1 = n_1, \Delta_2 = n_2, \ldots, \Delta_{k+1} = n_{k+1}\} \\
=&\{T_1 - T_0 = n_1, T_2 - T_1 = n_2, \ldots, T_{k+1} - T_k = n_{k+1}\} \\
=&\{T_1 - T_0 = n_1, T_2 - T_1 = n_2, \ldots, T_k - T_{k-1} = n_k\} \cap \bigcap_{j=1}^{n_{k+1}-1} \{X_{n_1+\cdots+n_k+j} = 0\} \cap \{X_{n_1+\cdots+n_k+n_{k+1}} = 1\}.
\end{aligned}$$

Comme l'événement $\{T_1 - T_0 = n_1, T_2 - T_1 = n_2, \ldots, T_k - T_{k-1} = n_k\}$ s'exprime en fonction des variables $X_1, \ldots, X_{n_1+\cdots+n_k}$ uniquement, on en déduit que

$$\begin{aligned}
&\mathbb{P}(\Delta_1 = n_1, \Delta_2 = n_2, \ldots, \Delta_{k+1} = n_{k+1}) \\
&\quad = \mathbb{P}(\Delta_1 = n_1, \Delta_2 = n_2, \ldots, \Delta_k = n_k)(1-p)^{n_{k+1}-1} p
\end{aligned}$$

et par une récurrence immédiate que

$$\mathbb{P}(\Delta_1 = n_1, \Delta_2 = n_2, \ldots, \Delta_{k+1} = n_{k+1}) = \prod_{j=1}^{k+1} (1-p)^{n_j - 1} p.$$

On a montrÃľ que les variables Δ_k sont indÃľpendantes et de loi gÃľomÃľtrique de paramÃĺtre p.

Exercice 7 [Absence de mÃľmoire]

1. Montrer que pour toute variable aléatoire X à valeurs dans $\{1, 2, \ldots\}$, les propriétés suivantes sont équivalentes :
   (i) X suit une loi géométrique ;
   (ii) La loi de $X - n$ sachant $\{X > n\}$ est identique à la loi de X pour tout entier $n \geq 0$.
2. Utiliser cette propriÃľtÃľ pour calculer la variance de la loi géométrique.

1. Supposons que X suit une loi géométrique de paramètre p. Alors

$$\mathbb{P}(X > n) = \sum_{k=n+1}^{+\infty} (1-p)^{k-1} p = p(1-p)^n \frac{1}{1-(1-p)} = (1-p)^n.$$

Par conséquent, pour tout $k \geq 1$,

$$\mathbb{P}(X - n = k | X > n) = \frac{\mathbb{P}(X = n+k)}{\mathbb{P}(X > n)} = \frac{(1-p)^{n+k-1} p}{(1-p)^n} = (1-p)^{k-1} p = \mathbb{P}(X = k).$$

Autrement dit, la loi géométrique est sans mémoire : la probabilité d'attendre un pile durant k unités de temps sachant qu'on a déjà attendu n unités de temps ne dépend pas de n.
Réciproquement, supposons que X soit une variable aléatoire à valeurs dans $\mathbb{N}^*$ vérifiant cette propriété d'absence de mémoire et montrons que la loi de X est géométrique. Pour tout $n \geq 0$, on a donc

$$\mathbb{P}(X = k+n) = \mathbb{P}(X > n)\mathbb{P}(X = k), \qquad \forall k \geq 1.$$

En prenant $n = 1$ et en posant $q = \mathbb{P}(X > 1) = 1 - \mathbb{P}(X = 1)$ on a donc

$$\mathbb{P}(X = k+1) = q\mathbb{P}(X = k), \qquad \forall k \geq 1,$$

d'où l'on tire $\mathbb{P}(X = k) = \mathbb{P}(X = 1)q^{k-1} = (1-q)q^{k-1}$. Le cas $q = 1$ est exclu car sinon on aurait $\mathbb{P}(X = k) = 0$ pour tout $k \geq 1$ ce qui contredit le fait que $\mathbb{P}(X \geq 1) = 1$. Lorsque $q = 0$, on a $\mathbb{P}(X = 1) = 1$ et donc la loi de X est la masse de Dirac en 1. On en conclut que, dans tous les cas, X suit une loi géométrique de paramètre $p = 1 - q \in]0, 1]$.

2. On peut utiliser cette propriété d'absence de mémoire pour calculer les moments d'ordre 1 et 2 de X. Illustrons ceci pour le moment d'ordre 2. Par l'absence de mémoire, on a

$$\mathbb{E}[X^2] = \mathbb{E}[(X-1)^2 | X > 1] = \frac{\mathbb{E}[(X-1)^2 \mathbf{1}_{X>1}]}{\mathbb{P}(X > 1)} = \frac{\mathbb{E}[(X-1)^2]}{\mathbb{P}(X > 1)},$$

et donc

$$(1-p)\mathbb{E}[X^2] = \mathbb{E}[X^2] - 2\mathbb{E}[X] + 1$$

Comme on sait que $\mathbb{E}[X] = 1/p$, on trouve $\mathbb{E}[X^2] = \frac{2}{p^2} - \frac{1}{p}$. On en déduit la variance de X : $\mathrm{Var}(X) = (1-p)/p^2$.

Exercice 8 Soit X_1, X_2 des variables aléatoires indépendantes de loi de Poisson de paramètres respectifs $\theta_1 > 0$ et $\theta_2 > 0$.

1. Calculer la loi de $X_1 + X_2$.
2. Calculer $\mathbb{E}(z^{X_1} | X_1 + X_2)$.
3. Calculer $\mathbb{E}(X_1 | X_1 + X_2)$.

1. Pour $k \in \mathbb{N}$ on a

$$\begin{aligned}
\mathbb{P}(X_1 + X_2 = k) &= \sum_{l=0}^{\infty} \mathbb{P}(X_1 + X_2 = k | X_2 = l)\mathbb{P}(X_2 = l) \quad \text{(par la formule des proba totale)} \\
&= \sum_{l=0}^{k} \mathbb{P}(X_1 = k - l | X_2 = l)\mathbb{P}(X_2 = l) \\
&= \sum_{l=0}^{k} \mathbb{P}(X_1 = k - l)\mathbb{P}(X_2 = l) \quad \text{(par indÃľpendance)} \\
&= \sum_{l=0}^{k} \frac{\theta_1^{k-l} e^{-\theta_1}}{(k-l)!} \frac{\theta_2^l e^{-\theta_2}}{l!} \\
&= \frac{e^{-\theta_1 - \theta_2}}{k!} \sum_{l=0}^{k} \binom{k}{l} \theta_1^{k-l} \theta_2^l \\
&= \frac{(\theta_1 + \theta_2)^k e^{-(\theta_1 + \theta_2)}}{k!}.
\end{aligned}$$

Donc, $X_1 + X_2$ suit la loi de Poisson de paramÃĺtre $\theta_1 + \theta_2$.

2. D'aprÃĺs la dÃľfinition de l'espÃľrance conditionnelle on a

$$G(z) := \mathbb{E}[z^{X_1} | X_1 + X_2 = m] = \sum_{k=0}^{\infty} z^k \mathbb{P}(X_1 = k | X_1 + X_2 = m)$$

Bien Ãľvidemment, on $\mathbb{P}(X_1 = k | X_1 + X_2 = m) = 0$ pour tout $k > m$. Pour tout $0 \leq k \leq m$ on trouve

$$\begin{aligned} \mathbb{P}(X_1 = k | X_1 + X_2 = m) &= \frac{\mathbb{P}(X_1 = k, X_1 + X_2 = m)}{\mathbb{P}(X_1 + X_2 = m)} = \frac{\mathbb{P}(X_1 = k, X_2 = m-k)}{\mathbb{P}(X_1 + X_2 = m)} \\ &= \frac{\mathbb{P}(X_1 = k)\mathbb{P}(X_2 = m-k)}{\mathbb{P}(X_1 + X_2 = m)} \quad \text{(par indÃľpendance)} \\ &= \frac{\frac{\theta_1^k e^{-\theta_1}}{k!} \frac{\theta_2^{m-k} e^{-\theta_2}}{(m-k)!}}{\frac{(\theta_1+\theta_2)^m e^{-(\theta_1+\theta_2)}}{m!}} \\ &= \binom{m}{k} \frac{\theta_1^k \theta_2^{m-k}}{(\theta_1 + \theta_2)^m}. \end{aligned}$$

En fait, $X_1 | X_1 + X_2 = m$ suit une loi binomiale de paramÃĺtres m et $\theta_1/(\theta_1 + \theta_2)$. On en dÃľduit que

$$\mathbb{E}[z^{X_1} | X_1 + X_2 = m] = \frac{1}{(\theta_1 + \theta_2)^m} \sum_{k=0}^{m} \binom{m}{k} (z\theta_1)^k \theta_2^{m-k} = \left(\frac{z\theta_1 + \theta_2}{\theta_1 + \theta_2} \right)^m.$$

3. On dÃľduit des propriÃľtÃľ de la fonction gÃľnÃľratrice que

$$\mathbb{E}(X_1 | X_1 + X_2 = m) = G'(1) = \left. \frac{m\theta_1 \left(z\theta_1 + \theta_2\right)^{m-1}}{\left(\theta_1 + \theta_2\right)^m} \right|_{z=1} = \frac{m\theta_1}{\theta_1 + \theta_2}.$$

Exercice 9 [Fonction génératrice] Déterminer une condition nécessaire et suffisante pour que les réels a et k soient tels que la suite (p_n) définie, pour $n \geq 0$, par $p_n = k(\frac{a}{a+1})^n$ soit la loi de probabilité d'une variable aléatoire à valeurs dans $\mathbb{N}$. Donner alors la fonction génératrice d'une telle variable aléatoire.

Il faut (1) $p_n \in [0,1]$ pour tout n et que (2) $\sum_{n=0}^{+\infty} p_n = 1$. (1) si et seulement si $a \geq 0$ et $1 \geq k \geq 0$ (prendre $n = 2$ puis $n = 1$, $n = 0$). On a $\sum p_n = k(a+1)$. Donc condition : $a \geq 0$ et $k = \frac{1}{a+1}$. Fonction génératrice

$$G(z) = \sum_{n \geq 0} p_n z^n = \frac{1}{a + 1 - az}.$$